IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 9, 2020, accepted July 18, 2020, date of publication July 21, 2020, date of current version August 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3010782

Efficient Energy Planning With Decomposition-Based Evolutionary Neural Networks

TANVEER AHMAD[1, 2, 3], DONGDONG ZHANG[1], (Member, IEEE), AND WAHAB ALI SHAH[4]

[1]School of Electrical Engineering, Guangxi University, Nanning 530004, China
[2]Energy and Electricity Research Centre, International Energy College, Jinan University, Zhuhai 519070, China
[3]State Key Laboratory of Internet of Things for Smart City, Department of Electrical and Computer Engineering, University of Macau, Taipa 999078, China
[4]Department of Electrical Engineering, Namal Institute, Mianwali 42250, Pakistan

Corresponding author: Dongdong Zhang (dongdongzhang@yeah.net)

This work was supported by the National Natural Science Foundation of China under Grant 61720106009. The authors gratefully acknowledged the support of UM Macao Postdoctoral Fellowship of University of Macao.

ABSTRACT Load forecasting/prediction is a challenging task in the energy markets that require adequate attention in generating stable and reliable load demand to deal with energy management and planning strategies. Accurate load prediction is critical for electrical power systems operations, but nonlinear loads involve high volatility. Forecasting these kinds of complex load characteristics requires highly accurate forecasting devices. It is generally accomplished by constructing models on relative details, including weather and previous data on load demand. This study proposes six decomposition-based evolutionary neural networks for city-scale and building energy forecasting. These evolutionary neural networks have also been trained using historical load data, and weather records are already considered to have a significant effect on energy usage (for example, wind speed, precipitation, dry-dew point temperature, relative humidity, clouds fraction and mean sea level perception). The network structure of the model is built and based on error estimation, trend map, and appropriate method of measuring evolutionary neural networks efficiency. The model's hidden layers and neurons network structure is selected based on recent smart models to enhance its accuracy. Several measures were used to improve performance, such as seasonal smoothing changes, coarse-graining of the humidity ratio, load-oriented day-type classification monitoring, outlier removal, and complex climate information. Results show that used models have a very high fitting accuracy and low error rate for short, medium, and long-term forecasting/planning tasks. These models have the potential to reduce overfitting issues, thereby improving load forecasting efficiency. Further, proposed models may also use for forecasting solar and wind demand. The efficiency of the model compared to the existing two models. Forecast results show superiority and efficiency in short, medium and long-term energy forecasting and offer the ability to manage unbalanced utility loads.

INDEX TERMS Evolutionary neural networks, energy forecasting, energy security, selective combination prediction, energy planning models.

I. INTRODUCTION

Energy systems' distribution infrastructure and their real-time operations are growing increasingly worldwide, with a traditional drift of change in energy quality, carbon pollution reduction, and energy management. A higher distribution of energy operations connecting conventional and renewable energy production would be an integral part of the future grid, offering both thermal and electricity cooperation at the energy consumption level— residential, town, commercial and industrial consumers [1]. Small-scale grids cover the large area of district-level electricity distribution and could be classified as each stand-alone, grid-connected, or use-by-residential, military, commercial, etc. Moreover, potential energy sources, most of them focused on photovoltaic cells, fossil fuel generators and wind turbines [2]–[4]. The performance estimation of stand-alone and hybrid renewable energy networks [2], renewable heating strategies, smart

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

energy systems approach [3], the climate impact on the cost of energy of grid-connected solar concentrator systems and energy yield are briefly investigated [4], it gives possible ways to be used in real-time applications. Energy management preparation will usually be recommended as detailed-

ACRONYMS AND NOMENCLATURE

AI	Artificial intelligence
ANNs	Artificial neural networks
CFNN	Cascade-forward neural networks
C-SL	City-scale level
CV	Coefficient of variation
GDMALB-NN	Gradient descent with momentum and adaptive learning rate backpropagation neural networks
GLRM	The generalized linear regression model
LANN	Linear artificial neural networks
LMANN	Levenberg-Marquardt algorithm neural networks
L-RNN	Layer-recurrent neural networks
MAPE	Mean absolute percentage error
ML	Machine learning
M_{slp}	Mean sea level perception
PB-CBNN	Powell Beale-conjugate gradient backpropagation neural networks
RBNN	Resilient backpropagation neural networks
RGSVM	Regression support vector machine
RLO	The coefficient of correlation lower limit
RUP	The coefficient of correlation upper limit
R	Coefficient of correlation
SL	System load
WSHP	Water source heat pump

NOMENCLATURE

F	Arbitrary error-function
d	Bias vectors
CLD_f	Clouds fraction
q^u	Correlation matrix of input vectors
w_{ko}	Corresponding weight
t	Data samples
M	Data samples of CV
T_{dp}	Dew point temperature (°F)
$D_{A-demand}$	Day-ahead demand
T_{db}	Dry-bulb temperature (°F)
l	Element of the input vector at the *lth* iteration.
e	Explicit derivative
$\hat{Q}_i$	Forecasted amount
eY	The gradient of PB-CBNN model
$JV^{n,m}$	Indices layers set to connect the layer *m* and *n*
M_n^g	Indices layers set directly to connect the layer *n*
y_1, y_n	Input parameters
x	Input parameters of RBNN
o	Input vector with linear neurons
$q_1 \ldots q_s$	Inputs of LANN
eY	The gradient of PB-CBNN model
$JV^{n,m}$	Indices layers set to connect the layer *m* and *n*
M_n^g	Indices layers set directly to connect the layer *n*
y_1, y_n	Input parameters
x	Input parameters of RBNN
o	Input vector with linear neurons
$q_1 \ldots q_s$	Inputs of LANN
eY	The gradient of PB-CBNN model
V	Layer weights of CFNN
s	Learning rate
c	Learning rate of LANN
b	Linear transfer function
$\bar{P}$	Mean of forecasted amount
h	Minimize the error between input variables
f	Model error
y_k	Model output
T	Model output
q	Multiple data set of inputs
Q_i	Net measurement
k	Neuron
n	Number of neurons
z	Output of CFNN
g	Overall model nodes
A	The parameter of the hidden layer of PB-CBNN model
P_{d-load}	Previous day load
P_{w-load}	Previous week load
G	Random variables
RT_{CC}	Real-time congestion component
$R_{t-demand}$	Real time demand
RT_{MLC}	Real-time marginal loss component
RT_{EC}	Real-time energy consumption
H_r	Relative humidity (%)
EM_{nm}	Set of all delays between layer *n* and *m*
Em_{nm}	Set of delays in the tapped delay
S_{load}	System load
R_i	Target value
u	Period
$\mho^-, \mho^+$	Uncertainty factor
W	Weights
X	Weight and bias variables
x_q	The weight of LANN model
ko	Weight of neuron
net_o	Weight sum of the inputs of neurons
W_d	Wind direction (Degree)
W_s	Wind speed (m/s)
P_i	Forecasted amount
iY	Forecasting direction of PB-CBNN
eX	The gradient of GDMALB-NN model

consisting of the quantitative measure inferred from an engineering model [5]. Accordingly, a combination district or city-scale energy management and planning approach as well as supporting technology for a specific strategy is the criteria of the day that must be developed to improve the existing regional/local energy management for districts and cities to be environmentally sustainable. Accordingly, environmentally friendly urban energy management and planning allow a combination of energy conservation and urban planning to be accelerated and encouraged towards integrated local/urban energy management activity. In addition, it establishes the enabling technologies and the systematic implementation process in real time [6]. On the basis of the above aspects, this study proposed six evolutionary neural networks for city-scale and building-level energy management to assist electrical generation firms or utilities, independent power producers, commercial and industrial consumers in making their specific future decisions based on historical past environment and energy consumption data.

The remaining sections of this work are categorized with the following parts-Section-II discusses the application of forecasting methods and Section II presents the proposed models for energy forecasting and statistics for output error estimation. In addition, it displays the performance evaluation indicators to calculate the consistency of the proposed models. Section IV offers an overview and study of the effect of weather variables on energy use at the city level, and data for training and validation. Section V displays the results and description of the forecasts.

A. EXISTING LITERATURE

Smart energy management plays an integral role in energy forecasting and management. The energy prediction is currently a hot area due to the planning of utility energy, load management in the building sector, as well as its resource management and operations. Different authors have conducted a large number of researches to reduce the forecast error and increase the accuracy of energy prediction by applying the machine learning (ML) and data mining models [9]. Li *et al.* [10] proposed the retrofit approach for energy saving for city buildings and to refine the energy forecasting algorithms for urban buildings by applying the (EnergyPlus) program based on different types of data sets for commercial and domestic users. Relevant work on large-scale and data-driven load planning requirement dependent techniques are discussed in comparison [11], and brings greater attention to the introduction of modern energy algorithms and recent development in the field of potential real-time forecasting for building energy strategies.

The new forecasting methods can be classified into four main groups, including: (i) classifications-based physics models; (ii) data-driven algorithms; (iii) hybrid-based techniques; and (iv) large-scale energy forecasting techniques for building projects. The methods are used in various energy systems and offer promising results in tasks of modeling and optimization. Biswas *et al.* [12] proposed the neural network algorithm, which would be useful in supporting the nonlinear characteristics of climate or energy data. The modeling results of previous studies show that the algorithms of the neural network function exquisitely better with the specific performance coefficient that is greater than 0.9. The basic idea of neural networks (NN) modeling was explored about 5-decades earlier, but its applicability has shifted to widespread and more comprehensive over the last 2-decades due to significant technological improvements to address issues with higher computing capacity and rapid processing pace [13], [14]. Kumar *et al.* [13] demonstrate that the energy analysis of buildings using artificial neural networks for heating and cooling purposes is a possible method for load planning. In reference [14], analyzes of different factors such as ambient weather conditions, the structure of the building and its characteristics which directly affect the heating and cooling load of the building. The artificial neural networks (ANNs) have been used extensively in many real-time applications, including robotics, aerospace, energy, economics, medicine, neurology and psychology [13]. Such technologies not only offer substantial efficiency but also raising the real-life human effort.

Rahman *et al.* [15] proposed a recurrent model to achieve medium and long-term prediction, e.g. time period for about one week, energy consumption curve characteristics in the residential and commercial market. Because of their nonlinear properties and characteristics, the ANNs are widely applied in the present scenario in evaluating energy prediction issues. However, the NN also relates to the different forms of optimization tasks, such as evaluating the optimal network variables to resolve the prediction error [16]–[18]. The connection of building energy consumption to the load demand behavior of the occupants in commercial buildings is briefly analyzed in reference [16], and it explains the actual change in load demand by changing occupants' behavior. The accuracy of the decision support model based on ANN demonstrates its dominance in terms of the success of energy planning for non-residential building stocks [17]. SVMs were widely used to address nonlinear regression and time series problems [19]. Perera *et al.* [20], used the statistical and machine learning models for load planning and management. Further, the exponential smoothing model [21], hybrid approach for probabilistic forecasting [22], [23], deep neural networks [24]–[26], chaotic time series model [27], deep reinforcement learning [28], [29], nonlinear autoregressive and random forest approaches [30], [31], convolutional neural network [32], modal decomposition based ensemble learning [33] and artificial intelligence models [34] are largely used for load demand forecasting analysis.

Biswas and Robinson [35] proposed a new method for managing energy consumption and developed an ANN-based approach to generate an automated energy utilization control strategy. Three performance evaluation indices are then applied to figure out the best model network structure being implemented to establish the regulation of energy consumption use. Length of short-term energy forecasting begins from

hours to one week. The medium-term forecasting period is graded from days or one week to specific months. In addition, the long-term forecast indicates the energy requirement for 1-year or longer, which helps to know the energy management of the power grid customer. The consumer load pattern [36], helps to establish potential sources of power generation and grid station infrastructure [37]. Statistical methods and artificial intelligence (AI) can be applied customarily to assess the physical solutions of various energy systems (''Wind speed and wind direction prediction using rapid learning, fully validated neural predictor network,'' 2014)

The AI allows efficient preparation and management of the bulk energy while the effects of the model are much quicker than traditional knowledge discovery and computational neuroscience-based models [38]. Different studies have been performed on demand-side load management and then its effects calculated in real-time predictions [39], microgrid load with deep learning and solar power prediction [40], short-term demand-side planning load prediction [41], and a new hybrid approach commonly used in the prediction of load demand and electricity levels throughout the smart grid infrastructure [42]. The AI is a mixture of neural networks, Fuzzy logic [43], vector supporting computer, genetic algorithms, and ML-based methods [44], used in load planning and control. Particle swarm optimization and genetic simulations are the evolutionary load planning approaches [45]. The use of mathematical models such as Kalman filters [46], autoregressive moving average and multiple regression algorithms provide knowledge of energy demand [47]. Many state-of-the-art methods were commonly used for very short, medium [48] and long-term [49] predictions. In building energy demand forecasting, a vast range of load control and planning methods are applied and can be discussed through detailed literature; Ho and Hsu [50], used ANNs in peak load demand analysis; Ho *et al.* [51] and Dash *et al.* [52] used to estimate peak load requirements to obtain accurate load curves; the ANNs allow for faster training, smaller training sets, and error [53]–[55], and the highest prediction accuracy for 5-day and 1-hour forward prediction [54]. Johannesen *et al.* [56], analyzed several regression tools for urban electricity demand forecasting, including the k-Nearest neighbor regressor, linear regression and random forest regressors. Results demonstrate that the k-Nearest neighbor regressor is reliable for long-term forecasting, and the random forest regressor provides higher precision for short-term energy demand forecasting. ML models, including Gaussian process compact regression, binary decision tree, generalized linear regression, and stepwise Gaussian process regression models are commonly used for short, medium and long-term energy demand forecasts [57]. They suit well, but the time for the computation is high. The ML models offer higher performance, accuracy and least predictive error based on limited data [58].

The literature review summarized that the different approaches to load planning give the idea that the ANNs algorithm could examine the characteristics of nonlinear climate trends and data on energy consumption if the hidden layers of ANNs are best decided. In this study, the six evolutionary-based neural models used for multiscale energy management and planning, including: i) resilient neural backpropagation networks (RBNN); ii) linear neural artificial networks (LANN); iii) cascade-forward neural networks (CFNN); iv) Powell-beale function with conjugate gradient backpropagation neural networks (PB-CBNN); v) Powell beale-conjugate gradient backpropagation neural networks (GDMALB-NN) and vi) layer-recurrent neural networks (L-RNN) offer their specific demand forecasting features. The RBNN render complex mathematical function that accepts the numeric inputs and outputs of various data types. RBNN's accuracy with backpropagation function is highly accurate and sensitive to the number of learning rates from the network. Inherently the LANN can define complex nonlinear relationships between dependent and independent data and environment and energy consumption variables. But LANN's drawbacks are prone to the network's overfitting problems. The training time for the CFNN network is 100 times faster than that of conventional perceptron network. This CFNN characteristic gives a connection between cascade networks that is ideal for a large number of training data sets. The preparation of the PB-CBNN network is very reliable and provides greater precision in predicting results right after adjustment of different kinds of network parameters. There are various kinds of methods used to train the GDMALB-NN, including stochastic gradient descent, stochastic momentum gradient descent, adaptive gradient, the accelerated gradient of Nesterov, Adam and RMSprop. The GDMALB-NN output is very slow to convergence, with that the large training inputs. GDMALB-NN has the advantage of being able to use imbalanced raw data, parallel processing capability and having distributed memory to save the data. The efficiency of L-RNN forecasts is enhanced by iteratively adjusting the weights of the network. The L-RNN is analytically tractable; it is simple to implement and has the ability to handle the different sets of input parameters used in network training.

B. IMPORTANCE OF ENERGY PREDICTION

Demand forecasts is the prediction of electricity needed to meet demand in the short, medium and long term. This analysis allows utilities to run and maintain their customers' supply [59]. Electrical load prediction is an effective method that allows electrical generation and distribution firms to improve their productivity and profits. It allows them to prepare their resources and activities to ensure the requisite energy for all customers. The advantages of load forecasting includes: i) allow the utilities to plan well as understand future consumption or demand for load; ii) minimize the service provider costs. Achieving future long-term load interpretation assists the company in planning and deciding commercially sustainable investments in future generation and transfer; iii) helps to define resources necessary for the construction of power plants like fuels, and other services needed to guarantee consistent and cost-effective energy production and delivery

to customers. For short, medium and long-term planning, this is essential; and iv) the load prediction contributes to forecasting the future for the scale, location and design of potential production facilities.

C. CONTRIBUTION TO THIS STUDY

The aim of this study is to present an analysis on the predictability of energy using the evolutionary models of neural networks. To address the vulnerability issue and improve predictability accuracy, this work proposes models for higher 7-day, 1-month and 1-year energy forecasting for multi-purpose utilities and building heating and cooling demand forecasting by applying six ANN models. ANN models used in this research have been used in serval fields to perform different types of tasks, but in this study, approaches are new to predict energy demand. Through various hidden neuron layers, we proposed a new network structure of these models that would be useful for the implementation of a small and large amount of energy and environmental data collection. Two separate sets of input function variables are selected and implemented as model inputs. The Grubbs method is used to find out the outlier. Error and forecast accuracy compared to three existing ML models. Specifically, it was proposed that characteristic ANNs should simultaneously forecast aggregate energy demand for short, medium- and long-term duration. In addition, a higher predictive accuracy, a simple network reading estimate proposed for predicting confidence interval and length of point estimation at the same time. Via experiments, it is concluded that the models are also capable of forecasting solar and wind energy demand based on the wind and solar energy data available in real time.

II. THE PROPOSED STRUCTURE FOR LOAD PLANNING

This section describes the proposed energy prediction structure for load planning and management.

A. SCHEMATIC STRUCTURE

Fig. 1 shows the forecasting method used to construct energy prediction algorithms. The forecasting structure of the network is based on two key components, including climate variables and real-time system load. System load (SL) represents the aggregated energy consumption of domestic, commercial, industrial, utilities and various private consumers in real time. Smart energy meters are used to measure data on energy consumption in real time at different nodes of power transmission and distribution networks. In the distribution of loads, four major parts are considered, including: (i) high tension and low tension distribution lines; (ii) different types of consumers, including the private and government sectors; (iii) distribution transformers; (iv) grid stations; (v) transmission lines; and (iv) utility control centers.

The distribution lines are the link between the grid station and the distribution transformer. It delivers high volts of power supply from the power grid to the distribution transformer. The distribution transformers receive a high volt's supply from the main grid station and convert it to different

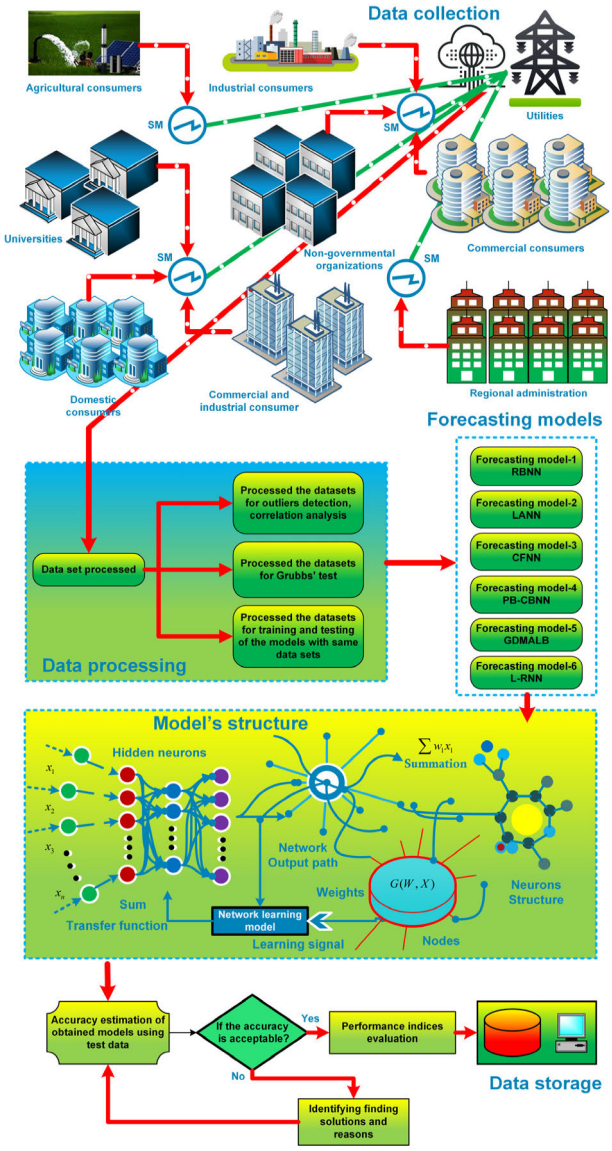

FIGURE 1. The proposed structure of this study.

low volt supply for end-users. The transmission lines are the main link between the power generation and the power grids. It has been used to supply a large amount of energy at a distance with the least amount of energy losses. Utility control centers receive from consumers the net signals of energy supply, energy demand, energy generation and actual load demand and consumption. Smart energy meters are installed at different ends, including a power station, a distribution transformer at the consumer premises to measure real-time energy supply and demand.

1) CLIMATE DATA

The impact of climate change on energy consumption is important and significant [60]. In this study, data are obtained from ISO-New England for the real-time reading of the 1-hour duration of climate and energy consumption. Eight

TABLE 1. Correlation between system load and several input feature parameters.

Paremeters	R	RUP (95.0%)	RLO (95.0%)
T_{db} (°F)	0.142	0.162	0.121
T_{dp} (°F)	0.012	0.246	-0.008
H_r (%)	-0.206	-0.226	-0.104
W_d (Degree)	0.057	0.076	0.0366
W_s (m/s)	0.175	0.255	0.014
CLD_f	-0.047	-0.068	-0.003
M_{slp}	0.015	0.144	-0.005
T (Hours)	0.495	0.514	0.102
RT_{EC} (MWh)	0.416	0.501	0.399
RT_{CC} (MWh)	-0.125	-0.145	-0.014
RT_{mlc} (MWh)	0.101	0.180	0.008

climate parameters were combined to predict energy, including dry bulb temperature (T_{db}), dew point temperature (T_{dp}), relative humidity (H_r), wind speed (W_s), wind direction (W_d), cloud fraction (CLD_f) and mean sea level perception (M_{slp}). The relationship between T_{db} vs SL and T_{dp} is detailed in Fig. 2. The temperature response to the SL gradually increases or decreases with a temperature between 40° F and 60°F as shown in Fig 2(a). Similar patterns are observed in Fig. 2(b). The graphics of T_{db} and T_{dp} vs. SL are taken below the 5th-degree polynomial to determine the behavior of the load pattern. The different colors demonstrate the energy requirement at different times of the day.

T_{db} is the dominant factors that change the requirement for cooling and heating energy consumption. T_{db} shows about 49.70% variation in energy demand for total building cooling and heating load [61]. Human perception research shows that T_{dp} between 40°F and 60°F is good for human comfort. Although, since H_r only affects the building, the latent cooling load demand is typically higher than 12.0°C. However, the increase in T_{dp} leaves a higher impact on building energy consumption [62], [63], and is considered to be one of the key parameters for forecasting load demand. From this study [64], the authors suggest that the increase in W_s directly raises the energy requirement [65]. In addition, the impact of temperature on energy consumption [66]–[68], humidity ratio [69], precipitation [70], [71], wind speed [72], [73] and cloud cover are briefly discussed [74]. These studies show how important climate parameters are to be considered as a key component of load planning and management.

2) ENERGY CONSUMPTION DATA

The actual energy consumption used to test models and energy planning, such as the normalized SL curve, is shown in Fig. 3. Load spikes in both summer and winter show different trends in both summer and winter energy requirements. This dark red curve shows the average load consumption of the system. It helps to draw a benchmark between the actual load, the upper and lower load curve bounds. Fig. 4 shows the average energy consumption from January to December. It has been shown that 8 to 22 hours a day needs higher

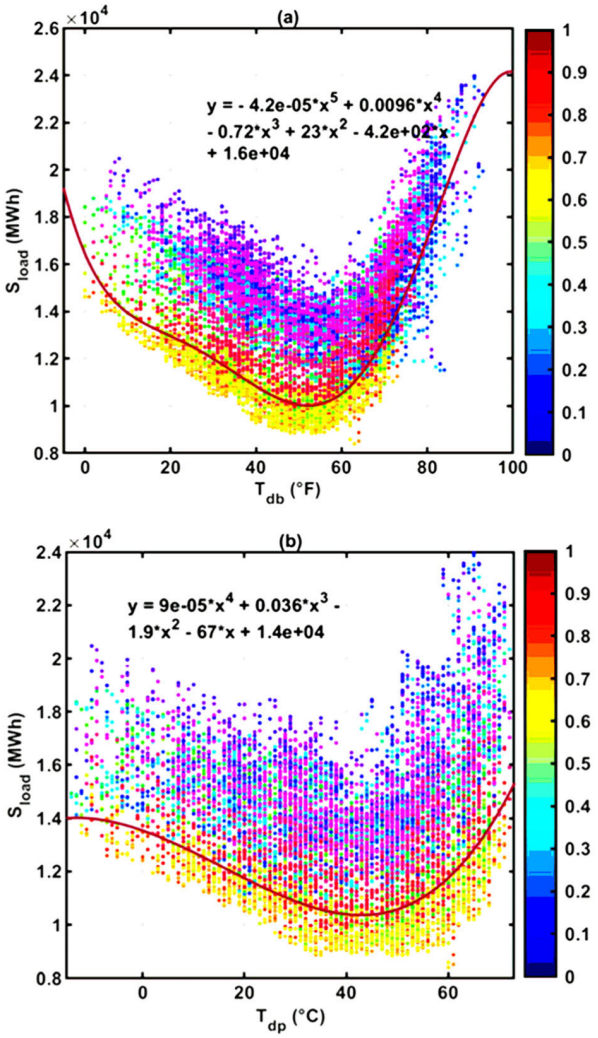

FIGURE 2. Scatter plot of SL and temperature (a) T_{db} vs. SL, (b) T_{dp} vs. SL.

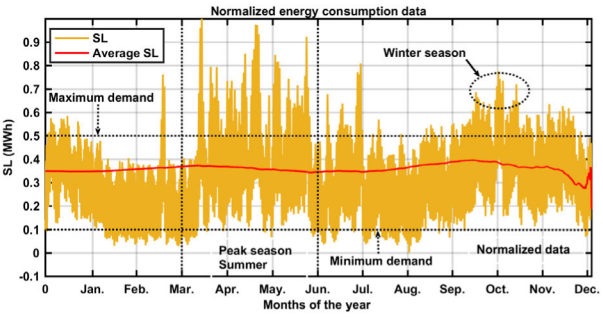

FIGURE 3. Normalized load-duration curve (using 1-hour sample time).

consumption than the rest of the day. The peak of daily energy demand in February and November is 18 to 22 hours in the evening. During August, the pattern has changed considerably; in the morning, the energy requirement is higher than the rest of the day. Data on energy consumption in real time is taken from ISO-New England [75].

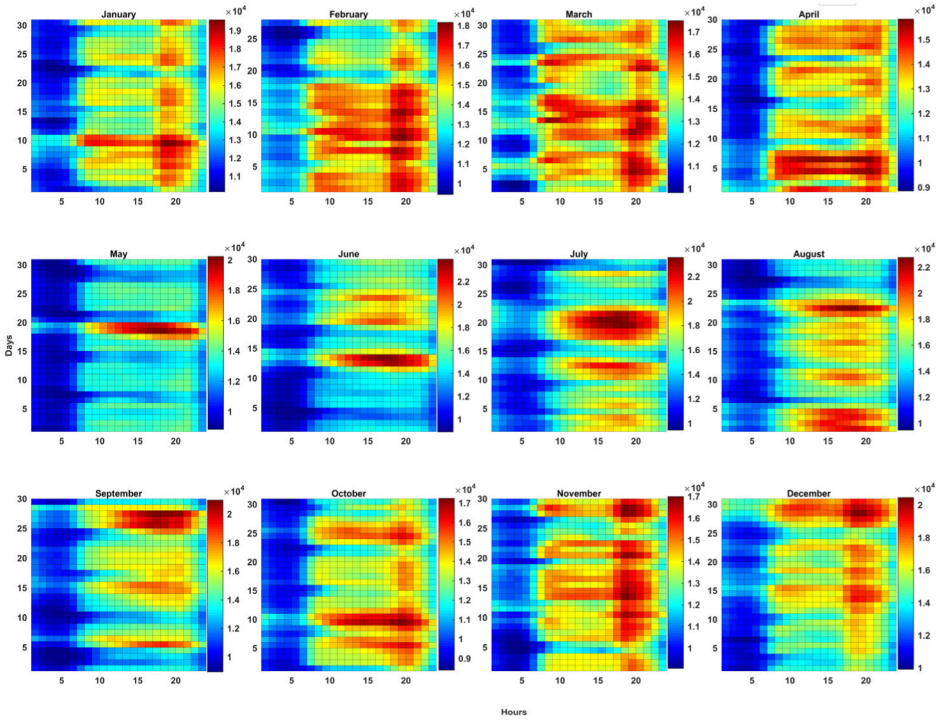

FIGURE 4. Average energy consumption of January to December by the time of day (hours) and days in a month.

3) CALENDAR EFFECTS

Energy use is linked to the time of the day, the day of the week and the month of the year. The availability of 8760 hourly readings for (1-year), 744 hourly readings for (1-month) and 168 hourly readings for (7-day) makes it possible to calculate the time or calendar impacts of the prediction algorithms.

4) DATA TRANSFORMATION AND OUTLIER DETECTION

The Grubbs approach was used to identify outliers in raw data [76]. The Grubbs approach determines the outliers of the dataset that are chosen in the continuous sequence of $y_1 \leq y_2 \leq y_3 \leq \ldots y_o$. This shows Grubbs test importance to analyze the total number of data points and $y_1 \leq y_2 \leq y_3 \leq \ldots y_o$ shows the number of data samples.

a: PEARSON CORRELATION ANALYSIS

Pearson's method is an approach for examining the correlation between two continuous and quantitative parameters. Table 1 shows the correlation between the SL and the different input parameters. The Pearson range of constant data starts from -1 to $+1$. The $+1$ and -1 give a higher correlation in the positive and negative directions, respectively [77].

b: FEATURE SELECTION

A feature selection technique can be considered as the integration of a search method that recommends new feature subsets and an appraisal metric that grades the various feature subsets. The K-Nearest-Neighbour (KNN) nonlinear function is employed as an indicator for determining the feature dependence of the target. In sample x under KNN, the approximate target is defined:

$$\hat{g}_y = \frac{1}{A} \sum_{y \in \acute{O}(y)} g(g') f^{-e(y,y')/\beta} \tag{1}$$

where $O(y)$ explicates the K-Nearest-Neighbour of sample y and $e\left(y, y'\right) = \sum_{j=1}^{o} \left(y_j - y'_j\right)^2 * x_j^2$ is distance between y' and y. Where o represents the total number of input feature variables (weather and climate variables) and x is the vector weight and x_i is the specific weight assigned to i^{th} feature. β shows the Gaussian decay factor and $A = \sum_{y \in \acute{O}(y)} g(g') f^{-e(y,y')/\beta}$ represents the normalization factor. For more details, please see the references [78], [79].

5) LENGTH OF ENERGY CONSUMPTION DATASET AND PREDICTION PERIOD

More accurate network training data set makes prediction algorithms more widely known. In the meantime, the efficiency of the model can be changed if the raw data changes significantly over time. For example, if large energy consumers or domestic users have recently moved from a gas heating system to an electrical heating system, the energy consumption curve will change significantly and raw or previously reported data would lose their relevance, which is why we used three different dataset lengths, ranging from 7 days, one month and one year.

III. THE PROPOSED STRUCTURE OF ANNs MODELS

ANNs models are applications and functions for the analysis, description and use of data algorithm patterns. These types of data could be used for exploratory data analysis, descriptive

statistics, and probabilities in relation to raw data, irregular amounts for hypotheses tests, and Monte Carlo simulations.

A. ENERGY PREDICTION TECHNIQUES

The proposed energy forecasting models are briefly discussed and analyzed in this section.

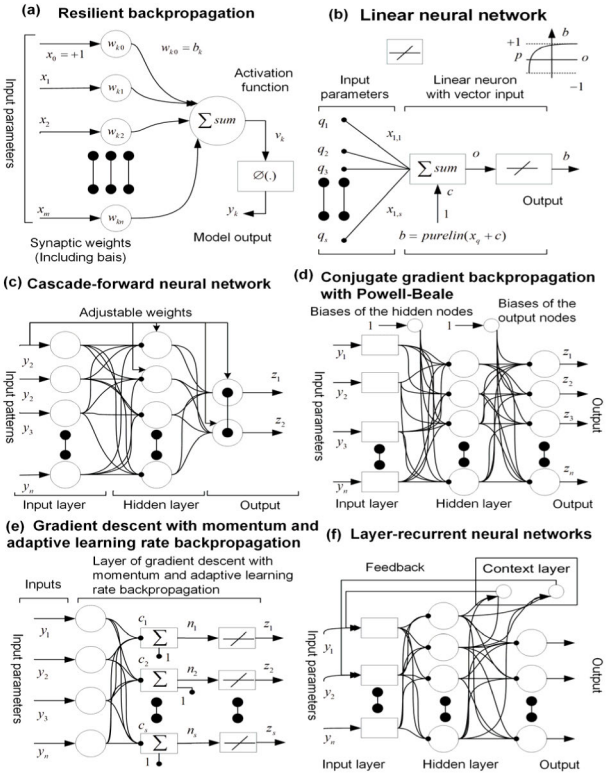

FIGURE 5. Schematic overview of supervised based energy forecasting models.

1) RESILIENT BACKPROPAGATION NEURAL NETWORKS

The schematic overview for resilient backpropagation is seen in Fig. 5(a) [80]. It shows the input parameters like x_1 to x_m, the weight of hidden layers $w_{ko}...w_{kn}$ and the output y_k of the network. The updated numbers are determined for each weight in the above manner, and further, each network weight is adjusted by its own updated number, in the opposite hidden layer direction of the partial derivative of that weight, so it is helpful to reduce the total number of the error function.

$$\frac{\partial F}{\partial w_{ko}} = \frac{\partial F, \partial T_k, \partial (net_{ko})}{\partial T_k, \partial\,(net_o)\,, \partial w_{ko}} \tag{2}$$

$$\Delta_{ko}^{u} = \begin{Bmatrix} \mho^{+} * \Delta_{ko}^{u-1}, & kg \dfrac{\partial F^{u-1}}{\partial w_{ko}} * \dfrac{\partial F^{u}}{\partial w_{ko}} > 0 \\ \mho^{-} * \Delta_{ko}^{u-1}, & kg \dfrac{\partial F^{u-1}}{\partial w_{ko}} * \dfrac{\partial F^{u}}{\partial w_{ko}} < 0 \\ \Delta_{ko}^{u-1}, & else \end{Bmatrix}, where\, 0 < \mho^{-} < 1 \mho^{+} \tag{3}$$

The partial derivative by each time span of the similar weight w_{ko} shifts its symbol that designates the previous update being extremely large and the model skipping across the local lower value, the updated amount Δ_{ko} is reduced from the factor $\mho^{-}$. However, the u determine the different slots of time, x shows the different input variables, g demonstrates the total nodes and y_k is the algorithm output.

2) LINEAR NEURAL NETWORKS

The LANN operates as a multi-layer neural perception network, but it has a linear transfer function. The LANN, such as perception, solves a linear kind of problem [81]. The schematic sketch of the LANN is shown in Fig. 5(b) of this. The LANN has similar characteristics and structure to the perceptron. The basic difference between the LANN and the perceptron structure is that the LANN used a linear transfer function to activate the network.

The matrix weight of x and the output of the network can be described as

$$b = purelin\,(o) = purelin\left(x_q + c\right) = x_q + c \tag{4}$$

Like the multi-layer perceptron, the LANN maintains a boundary that makes decisions that are defined by the vector inputs when the net inputs are on or near zero or zero. For $o = o$, the equation $x_q + c = 0$ designates like a boundary for decision.

$$y\,(l+3) = y\,(l) + 2bf\,(l) q^{u}(l) \tag{5}$$

$$d\,(l+3) = d\,(l) + 2bf\,(l) \tag{6}$$

The LANN network issues, such as bias and network weights, is the sum of the error square and reduces the low network values that are adjusted due to the linear operation of the model with a minimum one-error value. In most cases, the linear function directly minimizes its error for the rendered target vectors and input vectors.

3) CASCADE-FORWARD NEURAL NETWORKS

The CFNN is entirely related to the feed-forward artificial neural networks but gets a combination of each previous layer and input to the subsequent layers. CFNN networks, more than two or two-layer, can determine each finite output-input correlation in a mediately and well-defined manner from the hidden layer and neurons [82], [83]. Figure 5(c) shows the schematic diagram of the CFNN model. The predictor vector variable $y_1 \ldots y_n$ shows the total number of input layers. Output from the hidden neurons are distributed to the output layer of the network. The sigmoid transfer function is applied to the different types of classification issues. The basic equations for simulating the model are listed below to calculate an arbitrary CFNN network. The input layer can be estimated as n

$$o^{n}(u) = \sum \cdot_{m \epsilon M_n^g} \sum \cdot_{e \epsilon EM_{nm}} MV^{n,m}\,(e)\, b^{m}\,(u-e) + \sum \cdot_{m \epsilon m_n^g} \sum \cdot_{e \epsilon Em_{nm}} JV^{n,m}\,(e)\, q^{m}\,(u-e) + c^{n} \tag{7}$$

$$b^n(u) = g^n(o^n)(u)) \tag{8}$$

The CFNN also has relations with the model input for all 3-input layers. Further relationships can improve forecast accuracy when the model determines the covetous correlation of input variables [84].

4) POWELL-BEALE FUNCTION WITH CONJUGATE GRADIENT BACKPROPAGATION NEURAL NETWORKS

Powell Beale has developed a PB-CBNN algorithm and details of this model are presented in the reference [85]. This method is restarted again, when a small orthogonal function remains between the previous and current gradients. The PB-CBNN can be trained with any parameters and net-input, transfer and weight functions that have different derivatives of multiple input functions. Backpropagation is used to determine the different types of performance derivatives (*perf*) including the bias and weight of parameters Y. Each parameter is fixed in the following equation during the term:

$$Y = Y + h^* iY \tag{9}$$

$$eY = -hY + iY_{old} * A \tag{10}$$

PB-CBNN has a higher efficiency when compared to the conjugate gradient backpropagation (Polak-Ribiére) in some cases to solve complex problems. However, the efficiency and performance of any intended problem is difficult to predict. The model training shall terminate the following conditions- (a) the higher number of epochs shall be achieved; (b) the time limit shall be exceeded at its maximum level; (c) the performance of the model shall be fixed at the intended purpose or goal; and (d) and the gradient of the model shall be reduced below (*mingrad*). The schematic overview of the PB-CBNN is shown in Fig 5 (d). For PB-CBNN, it is proposed that hidden node biases and output node biases change the accuracy of the network. For PB-CBNN, the direction of the network search is modified periodically or reset to the negative of the gradient function.

5) THE GRADIENT DESCENT WITH MOMENTUM AND ADAPTIVE LEARNING RATE BACKPROPAGATION NEURAL NETWORKS

In particular, GDMALB-NN has been successful in optimizing artificial neural networks [86]. As a result, various models, such as conjugate gradient, free Hessian, etc., have been developed to address several types of optimization problems, and GDMALB-NN is used in this study. Back-propagation training using an adaptive-based learning-rate, including the (*traingda*) function is used. Backpropagation is used to estimate the performance of derivatives (*dperf*) including the bias and weight of parameters X. Each parameter is modified using the following gradient descent method:

$$eX = \frac{1s * dperf}{eX} \tag{11}$$

From every epoch, if the efficiency of the model reduces objective learning values, the variable *(Lr-INC)* can be improved. If the networkability of the model improves from above except for the variable *(max-perf-inc)*, then the learning-rate is modified from the parameter *(lr-dec)* as well as the shift that improved the efficiency that could not be achieved. Fig.5 (e) demonstrated the basic structure of the model GDMALB-NN. The $c_1 \ldots c_s$ shows that the learning rate decreases or increases on the basis of each iteration that information is obtained from the GDMALB-NN network. Hidden layers with backpropagation and adaptive learning rate improve predictive accuracy and reduce error to a minimum. Parameters $y_1 \ldots y_n$ explain the different number input variables, such as weather and energy consumption data.

6) LAYER-RECURRENT NEURAL NETWORKS

Models with more than one or one feedback are called L-RNN [87]. The two approaches are applied to the feedback construction of the L-RNN model. The first is a single neuron feedback, and the second is a local feedback; however, this type of model is shown in Fig. 5 (f). The L-RNN are very useful for solving time series problems. The characteristics of the L-RNN provide an infinite dynamic response to input data for time series forecasting. The context layers act as a bridge between the hidden layer and the network feedback to the inputs. The feedback system shows the variation between real-time data used as input feature variables and the predictive trends of hidden layers. Based on these results, the weights of the network are adjusted to improve the accuracy of the forecast and to reduce the error.

The network nodes are added to the continuous hidden layers, the input layer, layer to layer, and then the output layer of the model is finished. Network neurons are numbered starting with y, z numbers and actions in the hidden, input and output layers. It's considered the given vector, input y to obtain the desired model output $e_l = e_{l1}, e_{l2}$. The net model output is $z = (z_{l1} z_{l2})$. Appropriate measurements for this task can be calculated from the following equation:

$$F_l = (e_{l1} - z_{l1})^2 + (e_{l1} - z_{l1})^2 \tag{12}$$

and more generally would be if an output

$$F_l = \sum_{j=1}^{n} (e_{lj} - z_{lj})^2 \tag{13}$$

In analyzing the error function of the weights in the LRNN and acquire the nearest expression of the derivatives from the weight function in term of model's weights, the weights define the input given output. The provided weight number o, the error can be calculated from the following function:

$$F = g(x) = g(x_1, x_2, \ldots, x_o) \tag{14}$$

However, F is the error function of model weights that holds the changes in the single weight x, which leave an impact on the error F.

Decomposition-based evolutionary neural networks are well known to be difficult to train computationally. Computational delays and sample effects can significantly influence the efficiency of decomposition-based evolutionary neural

networks. Usually, a network's closed-loop responses are unreliable if these variables are not addressed. For this purpose, we incorporate sampling effects and computational delays in modeling for energy forecasting network, in order to accurately plan and simulate a system's prediction accuracy. We perceived the large constant gradient descent factor needed to achieve an acceptable precise accuracy that helps strengthen the argument. The solution based on decomposition is very effective with a large number of nodes and connection size, but impractical computing capacity is needed. Increasing the number for specific instants considered the minimum number and contact distance needed to complete the path. The result was a two-dimensional matrix consisting of the way to construct the dataset.

B. FORECASTING PERFORMANCE EVALUATION STATISTICS

The mean absolute percentage error (MAPE) shows a scale-independent index, a simple way to measure the dispersion of the models [88].

$$MAPE = \frac{1}{t}\sum_{i=1}^{t}\left|\frac{Q_i - \hat{Q}_i}{Q_i}\right| * 100 \tag{15}$$

where $\hat{Q}_i$ shows the forecasted values, Q_i explicates the actual values and t represents the total number of data samples.

The coefficient of variation (CV) represents the forecasting variation between actual consumption and the demand to be forecasted [89].

$$CV = \frac{\frac{1}{M}\sqrt{\sum_{i=1}^{M}(P_i - R_i-)^2}}{\bar{P}} * 100 \tag{16}$$

where P_i shows the forecasted values, R_i represents the target values, $\bar{P}$ explicates the mean of forecasted values and M presents the total number of data samples used for forecasting analysis.

IV. VARIABLES SELECTION FOR ENERGY PREDICTION

This section explains the variables of the input features and the data for model training, and the validation is briefly analyzed.

A. THE IMPACT OF MAIN FEATURES OF LOAD DEMAND

Time is a key factor in medium and long-term energy forecasting, as its impact on consumer energy needs is more significant [90]. Unforeseen sea breeze and moon downpours influence weather variables that may decrease T_{db} in different scenarios [91]. The impact of T_{db} during the summer period will increase the energy requirement and the reduction of T_{db} will change the energy condition as well as the peak load demand.

B. MODEL TRAINING AND VALIDATION

For the test and training algorithm, 70% of the net energy use and climate data samples are used for network training and 30% of the remaining data samples are used for validation. The forecasting intervals are divided into three sessions: i) 7-day forecast; ii) 1-month forecast; and iii) 1-year forecast. The climate and SL are further divided into two parts: Part-I) with 13 input climate and load characteristics; and Part-II) with 16 input climate and load characteristics. The selection of input features for Part-1 are T_{db}, T_{dp}, $T_{db-min}T_{db-max}$, T_{dp-min}, T_{dp-max}, Months, Hours, H_r, H_{r-min}, H_{r-max}, P_{d-load} and P_{w-load}. The Part-II input feature parameters are T_{db}, $R_{t-demand}$, T_{dp}, $D_{A-demand}$, H_r, W_d, W_s, CLD_f, mean sea level perception (M_{slp}), real-time energy consumption (RT_{EC}), real-time congestion component (RT_{CC}), real-time marginal loss component (RT_{mlc}), Previous day load (P_{d-load}), previous week load (P_{w-load}), months and S_l. The cumulative data samples for seven days in advance, one month in advance and one year in advance are 168, 744 and 8760 data samples, respectively. For one week ahead, the first 118 data samples used for training and the remaining 25 + 25 data samples used for testing and validation. For the next month forecasting, 521 data samples are used for training and the remaining 112 + 111 data samples are used for model testing and validation, respectively. For the next year forecasting, 6132 data samples are used for training and the remaining 1314 + 1314 data samples are used for model testing and validation, respectively. The duration of the training dataset, for the 7-day forecast is started from January 01, 2017 to January 07, 2017, for the 1-month forecast dataset, from January 01, 2017 to January 31, 2017, and for the 1-year forecast dataset, from January 01, 2017 to December 31, 2017.

Testing and validation dataset duration is considered with the following sequence, for example, the 7-day forecast datasets are taken from January 08, 2018, to January 09, 2018, the 1-month forecast datasets are taken from February 01, 2018, to February 10, 2018, and the 1-year forecast datasets are taken from February 01, 2017, to March 31, 2018.

Table 2 represents the development of six evolutionary neural networks based on decomposition. The different parameters for the occurrence model used are similar. But some of them are different, and it's because of enhancing the accuracy of the model. For example, the learning levels of RBNN, LANN, CFNN and PB-CBNN differ. It is acquired by hit and trail method in order to obtain the best model accuracy. Changing the learning rate also depends on the design of the data. For example, for the non-linear trend, the learning rate varies from linear trends to seasonal trends. We choose the best combination of learning rates for the week ahead, one month ahead, and one-year head forecast. Table 2 sets out the same parameters for the three forecast periods chosen, as indicated above. The network layers of the model dimensions of Part-I and Part-II are summarized in Table 3. This shows the full layer and the delay in the network/feedback dimensions. We set these parameters according to our sequence, but the nature of the data and function parameters can be changed.

TABLE 2. The training occurrence to model's training variables.

Dimensions	RBNN	LANN	CFNN	PB-CBNN	GDMALB-NN	L-RNN
Change in input weight	0.07	0.07	0.05	0.01	0.01	0.05
Max. weight change	50	50	50	100	50	50
Learning rate	0.01	0.02	0.05	0.5	0.01	0.05
Max. epochs for training	1000	1000	1000	1000	1000	1000
Gradient performance	1^{e-5}	1^{e-5}	1^{e-6}	1^{e-10}	1^{e-5}	1^{e-6}
Increment to weight change	1.2	1.4	1.2	1.2	1.2	1.2
Epochs between displays	25	25	25	25	25	25

TABLE 3. The used network layer structure of proposed decomposition-based evolutionary neural networks for Part-I and Part-II.

Dimensions	RBNN	LANN	CFNN	PB-CBNN	GDMALB-NN	L-RNN
Num. layers	2	1	2	2	2	2
Num. outputs	1	1	1	1	1	1
Num. layer delays	0	0	0	0	0	1
Feedback delays	0	0	0	0	0	1
weight elements	55	17	197	181	171	209
Sample time	1	1	1	1	1	1
Bias connect	[1; 1]	True	[1; 1]	[0 1; 1]	[1; 1]	[1; 1]
Input connect	[1; 0]	True	[1; 1]	[0 1; 0]	[1; 0]	[1; 0]
Layer connect	[0 0; 1 0]	False	[0 0; 1 0]	[0 0; 1 0]	[0 0; 1 0]	[1 0; 1 0]
Output connect	[0 1]	True	[0 1 0]	[0 1 0]	[0 1]	[0 1]
Hidden layer	3	6	10	10	10	8

V. FORECASTING RESULTS AND DISCUSSION

This part of study shows the modeling results, discussion, the forecasting error and the forecasting performance comparison of the proposed model with the existing models.

A. MODELING RESULTS OF PROPOSED MODELS

In this section, modelling results of proposed models are discussed.

1) SEVEN-DAY FORECASTING

For seven-day ahead forecasting, complete information on the comparison, including other performance assessment indices is provided in Table 4. The model L-RNN makes a higher accurate forecast of 1.507%, 1.962% in Part-I and 0.074%, 0.137% in Part-II, respectively, in terms of MAPE and CV. It also concerns the fact that Part-II is a better predictive performance of RBNN, LANN, CFNN, PB-CBNN, GDMALB-NN and L-RNN compared to Part-I. The two-sample t-test (p-value test) for hypothesis trials uses the estimated probability to evaluate if the null hypothesis is rejected by evidence. The p-value test values of Part-I and Part-II are near or close to 1. For Example, the p-value of RBNN, LANN, CFNN, PB-CBNN, GDMALB-NN, L-RNN in Part-I and Part-II are 0.8525, 0.9987, 0.9730, 0.8055, 0.7048, 0.9450 and 0.9372, 0.9871, 0.9702, 0.9934, 0.8397 0.9864, respectively. The p-values show that at 5% significance level, does not reject the null hypothesis. It means accuracy between actual and forecasted demand higher. Fig. 6 shows the forecast performance of the ANNs. The LANN shows a quiet dispersion in accuracy compared to the others in Part-I. There is a minor change in the performance of GDMALB-NN and RBNN in Part-II. The L-RNN shows that it is superior in forecasting and gives higher accuracy.

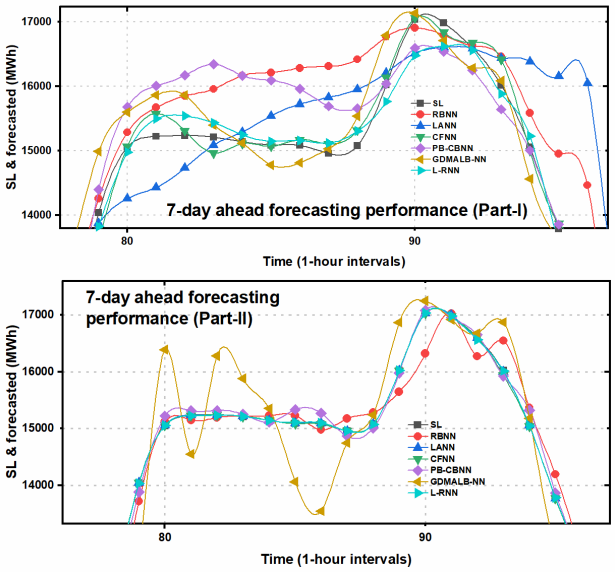

FIGURE 6. Forecasting performance for 7-day ahead prediction.

2) ONE-MONTH AHEAD FORECASTING RESULTS

Fig. 7 tends to be associated the forecast data numbers of the regression method with the SL and shows that the accuracy of the model is adequate. It can be observed that there is a

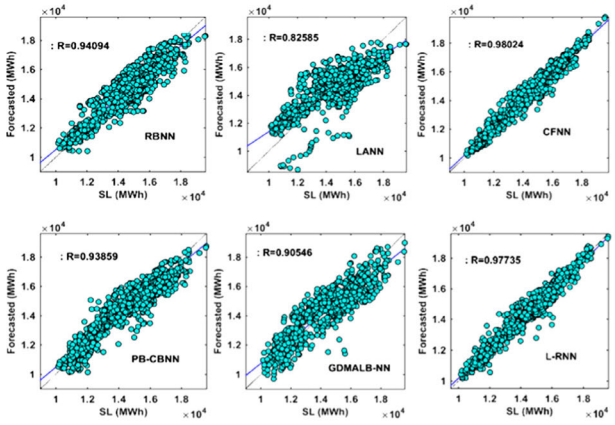

FIGURE 7. The actual vs. forecasted energy requirement for 1-month with limited climate and energy consumption data information.

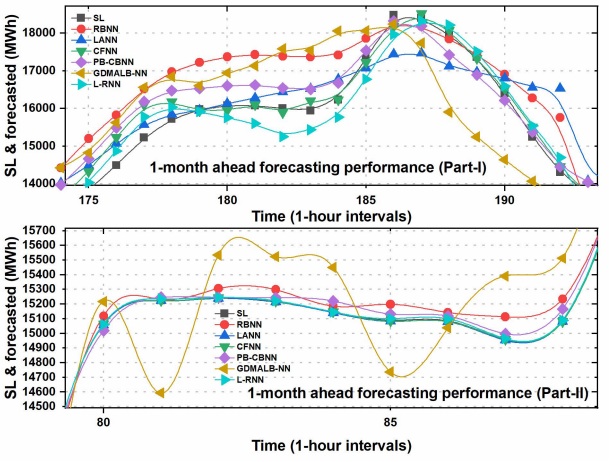

FIGURE 8. Forecasting performance for 1-month ahead prediction.

good relationship between forecast and modeled energy data. In addition, the L-RNN gives a good fit.

There is a small variation in energy forecast that can be notified of the performance of the LANN algorithm. To avoid a large number of regression plots, Fig. 7 shows the forecasting performance of Part-1 and the performance of Part-II is shown in Table 5 for MAPE and CV. The correlation coefficient for RBNN, LANN and CFNN is 0.940, 0.825 and 0.980, respectively. The correlation coefficient of PB-CBNN, GDMALB-NN and L-RNN is 0.938, 0.905 and 0.977, respectively. RBNN and L-RNN

Usually, it is seen that increasing the maximum number of input features at each model node can improve predictive performance. Part-II MAPE of RBNN, LANN, CFNN, PB-CBNN, GDMALB-NN and L-RNN is 0.680%, 0.137%, 0.083%, 0.422%, 1.501% and 0.161%, respectively, and is higher than Part-I. The p-value of GDMALB-NN of Part-I shows the week hypothesis as compared with Part-II. But the p-value of RBNN and LANN in Part-II is higher than Part-I. LRNN, CFNN shows a strong correlation in term of the p-value. For the two prediction cases, CFNN and L-RNN provide the most precise results which can be seen in Fig. 8.

3) ONE-YEAR AHEAD FORECASTING RESULTS

The correlation coefficient in the previous section shows that the forecast performance of Part-II is highly correlated with the SL compared to Part-I. The results of the MAPE and CV are summarized in Table 6 for the long-term forecasting accuracy of the ANNs. The CFNN algorithm, which increases the regression capability in optimization, provides more accurate energy forecasting results. The GDMALB-NN algorithm used in different other areas does not work as well in this study such as the other selected algorithms. In three cases, the GDMALB-NN algorithm performance is relatively low for 7-day, 1-month and 1-year forecast. The other five forecasting models, such as RBNN, LANN, CFNN, PB-CBNN and L-RNN, deliver similar and improved forecasting performance.

The results of the model show that the CFNN can achieve the highest performance with MAPE of 3.435% and 0.161%, respectively, in predicting the 1-year energy consumption for both Part-I and Part-II. The results presented in Tables 6 and Fig.9 are observed and show how a better agreement on the accuracy of the forecast. Generally speaking, improved performance and efficiency algorithms lead to better performance of the part-II features. The values of p-value test also show the higher accuracy between actual and forecasting demand.

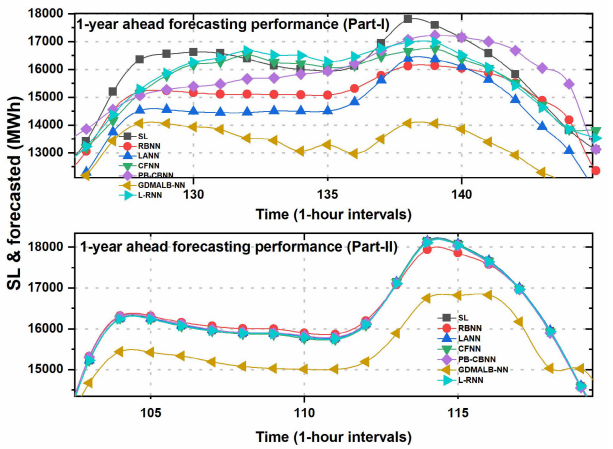

FIGURE 9. Forecasting performance for 1-year ahead prediction.

The regression plot of actual and forecasted 1-year load demand with limited data on climate and energy consumption is shown in Fig. 10. In this scenario, PB-CBNN, CFNN and RBNN have the lowest correlation coefficient. For example, the CFNN correlation coefficient is 0.969, which is similar to PB-CBNN and GDMALB-NN. The LANN displays a quiet low correlation coefficient of 0.869, but is still higher and almost 0.9, which is considered to be higher accuracy. The L-RNN provides a correlation coefficient of 0.896 between actual and predicted load demand.

TABLE 4. Performance indices for 7-day ahead prediction.

Category	Indices	RBNN	LANN	CFNN	PB-CBNN	GDMALB-NN	L-RNN
Part-I	MAPE (%)	3.9153	7.0929	2.0463	3.1775	3.8192	1.5076
	CV (%)	4.9265	8.5004	2.7566	3.9519	4.6626	1.9625
	p-value test	0.8525	0.9987	0.9730	0.8055	0.7048	0.9450
Part-II	MAPE (%)	2.2145	0.0732	0.0878	0.7947	6.5298	0.0747
	CV (%)	2.8498	0.0989	0.1353	0.9826	8.1133	0.1375
	p-value test	0.9372	0.9871	0.9702	0.9934	0.8397	0.9864

TABLE 5. One-month ahead of forecasting performance.

Category	Indices	RBNN	LANN	CFNN	PB-CBNN	GDMALB-NN	L-RNN
Part-I	MAPE (%)	4.166	7.440	2.381	4.324	5.471	2.503
	CV (%)	5.260	8.759	3.073	5.365	6.607	3.289
	p-value test	0.8861	0.9986	0.9929	0.7275	0.5363	0.9126
Part-II	MAPE (%)	0.680	0.137	0.083	0.422	1.501	0.161
	CV (%)	1.123	0.678	0.669	0.815	1.962	0.690
	p-value test	0.9847	0.9989	0.9780	0.9923	0.8793	0.9789

TABLE 6. One-year ahead prediction results.

Category	Indices	RBNN	LANN	CFNN	PB-CBNN	GDMALB-NN	L-RNN
Part-I	MAPE (%)	5.260	6.560	3.435	3.942	6.340	3.468
	CV (%)	6.853	9.198	4.538	5.007	8.238	4.564
	p-value test	0.5206	0.9991	0.8818	0.7460	0.9081	0.9901
Part-II	MAPE (%)	0.547	0.186	0.161	0.323	3.024	0.173
	CV (%)	0.713	0.295	0.267	0.499	3.771	0.274
	p-value test	0.9480	0.9985	0.9989	0.9937	0.4501	0.9775

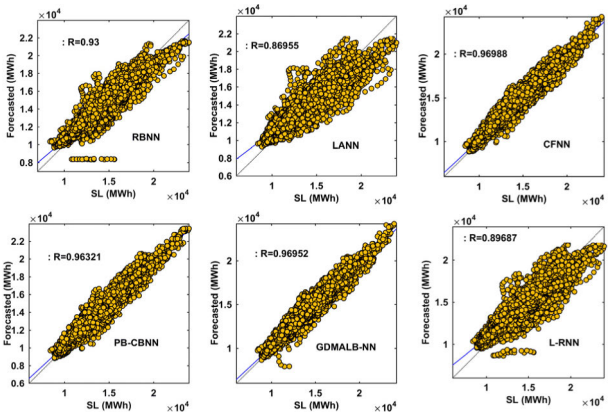

FIGURE 10. The actual vs. forecasted energy requirement for 1-year with limited climate and energy consumption data information.

B. FORECASTING ERROR OF ANNs

Generally, the error plot shows that part-II with a large number of input variables is a more appropriate choice for energy planning using the proposed ANNs. For the 7-day forecast, LANN, CFNN and L-RNN report the least forecast error compared to the MALB-NN and RBNN models. The same trend can be observed for 1-month and 1-year ahead load planning. The CFNN model shows similar accuracy for Part-I for the three intervals used, but there is a small dispersion in LANN performance. Overall, the performance for load planning is satisfactory.

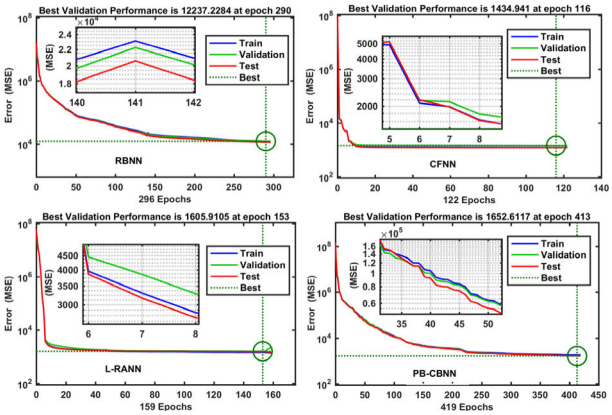

FIGURE 11. Training, validation, and test performance of RBNN, CFNN, PB-CBNN and L-RANN models.

Fig.11 shows the training, validation, and the test performance of the models RBNN, CFNN, PB-CBNN and L-RANN. This test, validation and training ended with the RBNN algorithm when the test error reached 290 iterations. The CFNN, PB-CBNN and L-RANN training, validation and test performance are similar at different iterations. No sub-

stantial over-fitting was observed for 290,116, 413 and 156 iterations of RBNN, CFNN, PB-CBNN and L-RNN algorithms. In addition, the forecasting models do not fit the account of changing scenarios, such as sudden changes in climate parameters. The choice of historical climate data observations may also have an impact on the results. This is due to a sudden change in weather or an irregularity in data when measured from the site.

The validation and test curves are similar. Unless the test curve increased substantially before the validity curve increased, any overfitting may have occurred. But in our case, it does not happen. Nevertheless, too low MSE could lead to over-refining. If the MSE is higher, it provides the following reason for its acceptability.

- The final mean-square error is small.
- The test set error and the validation set error have similar characteristics.
- No significant overfitting has occurred by iteration (where the best validation performance occurs)

The validation and test set has similar characteristics in our case, and there has been no significant overfitting that can be seen in Fig.11. Secondly, we have chosen 1000 epochs for training, testing and validation, and we have only mentioned the first few epochs to demonstrate the best results in validation.

C. PERFORMANCE COMPARISON OF FORECASTING MODELS

To evaluate the efficiency and performance of the model, three existing energy demand forecasting models are used: i) LMANN [92]; ii) Regression Support Vector Machine (RGSVM) [93]; and Generalized Linear Regression Model (GLRM) [55]. In reference [92], the authors proposed six energy, predictive models. The input parameters of the algorithm included weather data, time, and real-time demand for the water source heat pump (WSHP). The output of the algorithms was the use of WSHP energy. In this study, the modeling period was categorized into three basic sections-seven days, fourteen days and one month in advance from July 8, 2016, to August 7, 2016.

Reference [93] has been used to predict future energy requirements for ST perceptive in the building environment through the Hybrid Support Vector Regression (SVR) algorithm. The GLRM was developed in reference [55], and from this research shows the ML algorithms used to predict the energy use of WSHP for medium and ST perspectives. The best forecasting comparison of current and proposed model performance can be seen in Fig. 12. The error ratio (percent) found that Part-II had a lower forecast error with the exception of GDMALB-NN. In Part-I, the GDMALB-NN percentage of MAPE and CV is almost similar to the other models. It's in Fig. 12, the performance forecasting of Part-I is quite different from that of Part-II. The performance of GDMALB-NN in Part-II is poor compared to the other models but is higher in Part-I. It has been shown that GDMALB-NN renders higher load forearming performance with the least number of input variables.

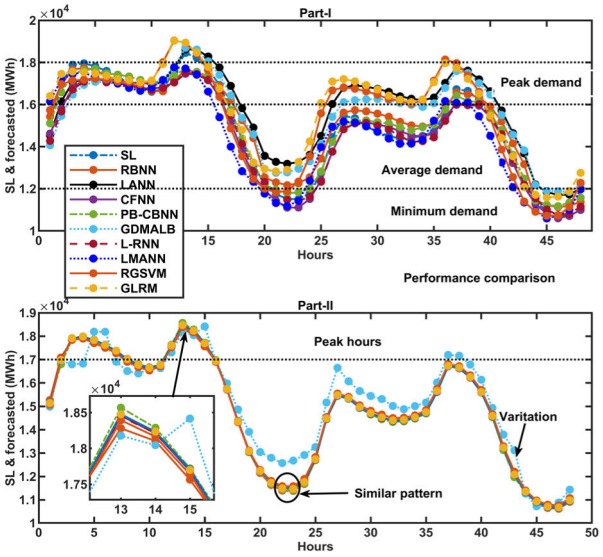

FIGURE 12. System load and forecasted energy requirement comparison with the previous models.

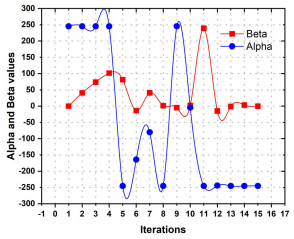

FIGURE 13. Training network of the SVR model.

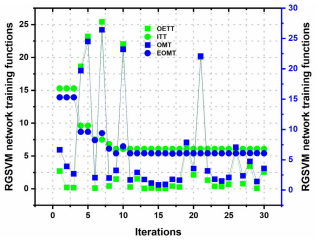

FIGURE 14. Network performance parameters of the SVR model.

Fig.13 and Fig.14 represents the performance of the training network variables of the SVR model. The alpha and beta values have changed by changing the total number of iterations. It can be observed that the Object Evaluation Time Trace (OETT), Iteration Time Trace (ITT), Objective Minimum Trace (OMT) and Estimated Objective Minimum Trace (EOMT) provide a good fit for SVR network training.

Fig.15 shows error contrasts with current LMNN, RGSVM and GLRM models. Horizontal number 1-9 explains the

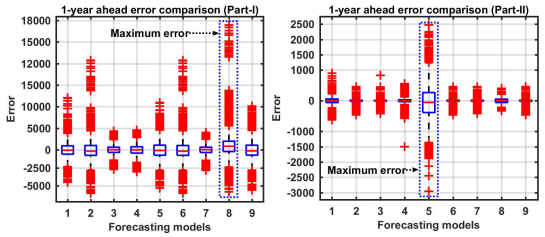

FIGURE 15. Error comparison of proposed models with the existing model.

forecasting model with the following sequences, e.g. RBNN, LANN, CFNN, PB-CBNN, GDMALB-NN, L-RNN, LMANN, RGSVM and GLRM. In Part-I, RGSVM has a higher forecast error than other models. CFNN, PB-CBNN and LMANN have similar efficiency and accuracy. The performance of Part II is promising. In this scenario, the only GDMALB-NN would give a quiet, higher error than the others.

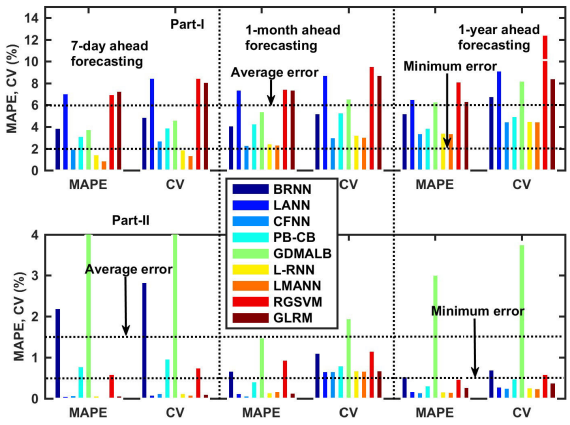

FIGURE 16. The MAPE and CV comparison of proposed and existing forecasting models.

The statistical performance assessment comparison of the proposed models and the existing models is shown in Fig. 16. Our proposed model results are more promising, compared to the GDMALB-NN model, which provides performance dispersion for Part II forecasting. But the results of Part-I do not show the same pattern. The reason is that GDMALB-NN delivers more accurate results with fewer input variables. The performance of the RGSVM is still poor in Part-I and Part-II. The BRBNN performance forecast in Part-II for the 7-day forecast is low but higher for the one month and one year forecast ahead.

In the past decades, many researchers have used ANNs to plan and manage load demand. It has been shown that ANNs, which are used in historical data to load demand planning in dynamic short, medium and long-term perceptions, are used in this study. Modeling results show that forecast accuracy for the short, medium and long term is similar except for quiet dispersion in some models. By means of a comprehensive review in the literature section, it is determined that if the model works well for short-term forecasting, it is not very useful for medium and long periods of time. However, our results have shown that these models are capable of predicting load demand in the short, medium and long term in different sectors with different types of input variables. As the total number of input feature variables increases, the forecasting performance is higher than the limited input feature parameters. Each model has its own unique features, such as LANN, which are associated with easier access to information, saving time and making the task easier for load planning. The L-RNN gives high accuracy in the load forecast that is very close to the actual system load. The RBNN demonstrates its ability to perform with inadequate (unbalanced) data sets or knowledge. The storage requirement for PB-CBNN is much larger than that of the Polak-Ribiére gradient back prorogation. In addition, the accuracy of the forecasting is compared with the three performance assessment indexes and three existing forecasting models. The basic objective of using the three existing forecasting models was to verify the accuracy of the proposed ANNs. In this study, the same input feature data sets used for the training of existing models that were used for the proposed models. The basic objective of comparing the results of evolutionary-based ANNs with LMANN, SVR and GLRM is that similar climate and energy consumption data but different time intervals have been used for these models. This study has shown that evolutionary-based ANNs have the ability to learn and to model complex and nonlinear relationships. It has been shown that the evolutionary-based ANN algorithms used in this research to perform a load planning task have significantly improved performance in terms of higher accuracy and variation over different periods.

VI. CONCLUSION

In this study, the evolutionary-based six ANN models used with a variety of variable features for large-scale utilities and building heating and cooling load planning and management. Three short, medium, and long range forecasting intervals are used for modeling analysis. Furthermore, the variables of the input feature are divided into two parts. Part-I contains thirteen climate and energy parameters with minimum and maximum data limits, and Part-II sets out sixteen input parameters. The LANN and L-RNN determine their prediction perfection when compared to the other model. The LANN, including the features of Part-II, achieves better predictive performance for the 7-day and 1-month forecast. Modeling results conclude that a significant increase in predictive efficiency and performance can be achieved by applying a large number of input feature datasets as presented in Part II. The 7-day, 1-month and 1-year forecast performance of Part-II is better than that of Part-I. The artificial neural network used in this study can provide moderate and accurate predictions in three sessions. The LANN MAPE for Part-I and Part-II is 14% and 3 %, respectively, and provides higher forecasting performance. RBNN and CFNN are more modest than GDMALB-NN in terms of weekly, monthly and yearly performance. Precision

and accuracy of models were compared to three ML and data mining models (LMANN, RGSVN, GLRM) in order to verify their effectiveness. Compared to the existing studies in graphical plots, MAPE and CV, the performance of the proposed algorithm is far superior to the previous studies. The variation between the forecast and the SL could be shown as a sign of anomalies detected in different energy management operations on a different scale.

These models are useful for independent power producers, utilities, commercial and industrial consumers who can structure the load demand profile as well as energy management in real-time applications. Such research can also be applied to energy operators for energy planning purposes. Energy operators can design their resources, generation schedules and power outage planning in accordance with long, medium and short-term energy profiles. In the future, more input features variables will be used to train models and different sets of large data under different climate conditions.

ACKNOWLEDGMENT

The authors gratefully acknowledged the support of UM Macao Postdoctoral Fellowship of University of Macau.

REFERENCES

[1] W. Ma, S. Fang, G. Liu, and R. Zhou, "Modeling of district load forecasting for distributed energy system," *Appl. Energy*, vol. 204, pp. 181–205, Oct. 2017, doi: 10.1016/j.apenergy.2017.07.009.

[2] S. Goel and R. Sharma, "Performance evaluation of stand alone, grid connected and hybrid renewable energy systems for rural application: A comparative review," *Renew. Sustain. Energy Rev.*, vol. 78, pp. 1378–1389, Oct. 2017, doi: 10.1016/j.rser.2017.05.200.

[3] H. Lund, "Renewable heating strategies and their consequences for storage and grid infrastructures comparing a smart grid to a smart energy systems approach," *Energy*, vol. 151, pp. 94–102, May 2018, doi: 10.1016/j.energy.2018.03.010.

[4] E. F. Fernández, D. L. Talavera, F. M. Almonacid, and G. P. Smestad, "Investigating the impact of weather variables on the energy yield and cost of energy of grid-connected solar concentrator systems," *Energy*, vol. 106, pp. 790–801, Jul. 2016, doi: 10.1016/j.energy.2016.03.060.

[5] C. Bhowmik, S. Bhowmik, A. Ray, and K. M. Pandey, "Optimal green energy planning for sustainable development: A review," *Renew. Sustain. Energy Rev.*, vol. 71, pp. 796–813, May 2017, doi: 10.1016/j.rser.2016.12.105.

[6] I.-A. Yeo, S.-H. Yoon, and J.-J. Yee, "Development of an urban energy demand forecasting system to support environmentally friendly urban planning," *Appl. Energy*, vol. 110, pp. 304–317, Oct. 2013, doi: 10.1016/j.apenergy.2013.04.065.

[7] T. Ahmad, H. Chen, and Y. Huang, "Short-term energy prediction for district-level load management using machine learning based approaches," *Energy Procedia*, vol. 158, pp. 3331–3338, Feb. 2019, doi: 10.1016/j.egypro.2019.01.967.

[8] T. Ahmad, H. Chen, J. Shair, and C. Xu, "Deployment of data-mining short and medium-term horizon cooling load forecasting models for building energy optimization and management," *Int. J. Refrig.*, vol. 98, pp. 399–409, Feb. 2019, doi: 10.1016/j.ijrefrig.2018.10.017.

[9] T. Ahmad and H. Chen, "Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment," *Energy*, vol. 160, pp. 1008–1020, Oct. 2018, doi: 10.1016/j.energy.2018.07.084.

[10] W. Li, Y. Zhou, K. Cetin, J. Eom, Y. Wang, G. Chen, and X. Zhang, "Modeling urban building energy use: A review of modeling approaches and procedures," *Energy*, vol. 141, pp. 2445–2457, Dec. 2017, doi: 10.1016/j.energy.2017.11.071.

[11] M. Bourdeau, X. Q. Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," *Sustain. Cities Soc.*, vol. 48, Jul. 2019, Art. no. 101533, doi: 10.1016/j.scs.2019.101533.

[12] M. A. R. Biswas, M. D. Robinson, and N. Fumo, "Prediction of residential building energy consumption: A neural network approach," *Energy*, vol. 117, pp. 84–92, Dec. 2016, doi: 10.1016/j.energy.2016.10.066.

[13] R. Kumar, R. K. Aggarwal, and J. D. Sharma, "Energy analysis of a building using artificial neural network: A review," *Energy Buildings*, vol. 65, pp. 352–358, Oct. 2013, doi: 10.1016/j.enbuild.2013.06.007.

[14] H.-X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renew. Sustain. Energy Rev.*, vol. 16, no. 6, pp. 3586–3592, Aug. 2012, doi: 10.1016/j.rser.2012.02.049.

[15] A. Rahman, V. Srikumar, and A. D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Appl. Energy*, vol. 212, pp. 372–385, Feb. 2018, doi: 10.1016/j.apenergy.2017.12.051.

[16] H. N. Rafsanjani, C. R. Ahn, and J. Chen, "Linking building energy consumption with occupants' energy-consuming behaviors in commercial buildings: Non-intrusive occupant load monitoring (NIOLM)," *Energy Buildings*, vol. 172, pp. 317–327, Aug. 2018, doi: 10.1016/j.enbuild.2018.05.007.

[17] M. Beccali, G. Ciulla, V. Lo Brano, A. Galatioto, and M. Bonomolo, "Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in southern italy," *Energy*, vol. 137, pp. 1201–1218, Oct. 2017, doi: 10.1016/j.energy.2017.05.200.

[18] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2016, pp. 308–313, doi: 10.1109/SmartGridComm.2016.7778779.

[19] P.-F. Pai and W.-C. Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting," *Energy Convers. Manage.*, vol. 46, no. 17, pp. 2669–2688, Oct. 2005, doi: 10.1016/j.enconman.2005.02.004.

[20] S. Perera, S. Dissanayake, D. Fernando, S. De Silva, and W. Rankothge, "Supply and demand planning of electricity power: A comprehensive solution," in *Proc. IEEE Conf. Inf. Commun. Technol.*, Dec. 2019, pp. 1–6, doi: 10.1109/CICT48419.2019.9066184.

[21] L. Di Persio, A. Cecchin, and F. Cordoni, "Novel approaches to the energy load unbalance forecasting in the italian electricity market," *J. Math. Ind.*, vol. 7, no. 1, Dec. 2017, doi: 10.1186/s13362-017-0035-y.

[22] G.-F. Fan, X. Wei, Y.-T. Li, and W.-C. Hong, "Forecasting electricity consumption using a novel hybrid model," *Sustain. Cities Soc.*, vol. 61, Oct. 2020, Art. no. 102320, doi: 10.1016/j.scs.2020.102320.

[23] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A hybrid approach for probabilistic forecasting of electricity price," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463–470, Jan. 2014, doi: 10.1109/TSG.2013.2274465.

[24] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2017, pp. 1483–1488, doi: 10.1109/ISIE.2017.8001465.

[25] R. Porteiro, S. Nesmachnow, and L. Hernández-Callejo, "Short term load forecasting of industrial electricity using machine learning," in *Communications in Computer and Information Science Smart Cities. ICSC-CITIES*, vol. 1152, S. Nesmachnow and L. H. Callejo, Eds. Cham, Switzerland: Springer, 2020.

[26] T. Ahmad and H. Chen, "Deep learning for multi-scale smart energy forecasting," *Energy*, vol. 175, pp. 98–112, May 2019, doi: 10.1016/j.energy.2019.03.080.

[27] D. Niu, Y. Wang, C. Duan, and M. Xing, "A new short-term power load forecasting model based on chaotic time series and SVM," *J. Univers. Comput. Sci.*, vol. 15, no. 13, pp. 2744–2763, 2009, doi: 10.3217/jucs-015-13-2726.

[28] N. M. M. Bendaoud and N. Farah, "Using deep learning for short-term load forecasting," *Neural Comput. Appl.*, Apr. 2020, doi: 10.1007/s00521-020-04856-0.

[29] T. Liu, Z. Tan, C. Xu, H. Chen, and Z. Li, "Study on deep reinforcement learning techniques for building energy consumption forecasting," *Energy Buildings*, vol. 208, Feb. 2020, Art. no. 109675, doi: 10.1016/j.enbuild.2019.109675.

[30] T. Ahmad and H. Chen, "Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems," *Sustain. Cities Soc.*, vol. 45, pp. 460–473, Feb. 2019, doi: 10.1016/j.scs.2018.12.013.

[31] T. Ahmad and D. Zhang, "Novel deep regression and stump tree-based ensemble models for real-time load demand planning and management," *IEEE Access*, vol. 8, pp. 48030–48048, 2020.
[32] Q. Huang, J. Li, and M. Zhu, "An improved convolutional neural network with load range discretization for probabilistic load forecasting," *Energy*, vol. 203, Jul. 2020, Art. no. 117902, doi: 10.1016/j.energy.2020.117902.
[33] C. Xu, H. Chen, W. Xun, Z. Zhou, T. Liu, Y. Zeng, and T. Ahmad, "Modal decomposition based ensemble learning for ground source heat pump systems load forecasting," *Energy Buildings*, vol. 194, pp. 62–74, Jul. 2019, doi: 10.1016/j.enbuild.2019.04.018.
[34] T. Ahmad, H. Chen, and W. A. Shah, "Effective bulk energy consumption control and management for power utilities using artificial intelligence techniques under conventional and renewable energy resources," *Int. J. Electr. Power Energy Syst.*, vol. 109, pp. 242–258, Jul. 2019, doi: 10.1016/j.ijepes.2019.02.023.
[35] M. A. R. Biswas and M. D. Robinson, "Performance estimation of direct methanol fuel cell using artificial neural network," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, 2016, Art. no. V06BT07A022, doi: 10.1115/imece2015-51723.
[36] F. M. Andersen, H. V. Larsen, and R. B. Gaardestrup, "Long term forecasting of hourly electricity consumption in local areas in denmark," *Appl. Energy*, vol. 110, pp. 147–162, Oct. 2013, doi: 10.1016/j.apenergy.2013.04.046.
[37] F. M. Andersen, H. V. Larsen, and T. K. Boomsma, "Long-term forecasting of hourly electricity load: Identification of consumption profiles and segmentation of customers," *Energy Convers. Manage.*, vol. 68, pp. 244–252, Apr. 2013, doi: 10.1016/j.enconman.2013.01.018.
[38] M. A. Mat Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," *Renew. Sustain. Energy Rev.*, vol. 70, pp. 1108–1118, Apr. 2017, doi: 10.1016/j.rser.2016.12.015.
[39] V. Stavrakas and A. Flamos, "A modular high-resolution demand-side management model to quantify benefits of demand-flexibility in the residential sector," *Energy Convers. Manage.*, vol. 205, Feb. 2020, Art. no. 112339, doi: 10.1016/j.enconman.2019.112339.
[40] L. Wen, K. Zhou, S. Yang, and X. Lu, "Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting," *Energy*, vol. 171, pp. 1053–1065, Mar. 2019, doi: 10.1016/j.energy.2019.01.075.
[41] M. Glavan, D. Gradišar, S. Moscariello, D. Juričić, and D. Vrančić, "Demand-side improvement of short-term load forecasting using a proactive load management–A supermarket use case," *Energy Buildings*, vol. 186, pp. 186–194, Mar. 2019, doi: 10.1016/j.enbuild.2019.01.016.
[42] A. Ghasemi, H. Shayeghi, M. Moradzadeh, and M. Nooshyar, "A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management," *Appl. Energy*, vol. 177, pp. 40–59, Sep. 2016, doi: 10.1016/j.apenergy.2016.05.083.
[43] K. V. Zúñiga, I. Castilla, and R. M. Aguilar, "Using fuzzy logic to model the behavior of residential electrical utility customers," *Appl. Energy*, vol. 115, pp. 384–393, Feb. 2014, doi: 10.1016/j.apenergy.2013.11.030.
[44] A. S. Ahmad, M. Y. Hassan, M. P. Abdullah, H. A. Rahman, F. Hussin, H. Abdullah, and R. Saidur, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 102–109, May 2014, doi: 10.1016/j.rser.2014.01.069.
[45] Z. Uykan, "Fast-convergent double-sigmoid hopfield neural network as applied to optimization problems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 6, pp. 990–996, Jun. 2013, doi: 10.1109/TNNLS.2013.2244099.
[46] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output," *Appl. Energy*, vol. 99, pp. 154–166, Nov. 2012, doi: 10.1016/j.apenergy.2012.03.054.
[47] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas, and P. Liatsis, "Electricity demand loads modeling using AutoRegressive moving average (ARMA) models," *Energy*, vol. 33, no. 9, pp. 1353–1360, Sep. 2008, doi: 10.1016/j.energy.2008.05.008.
[48] N. Abu-Shikhah and F. Elkarmi, "Medium-term electric load forecasting using singular value decomposition," *Energy*, vol. 36, no. 7, pp. 4259–4271, Jul. 2011, doi: 10.1016/j.energy.2011.04.017.
[49] L. Ekonomou, "Greek long-term energy consumption prediction using artificial neural networks," *Energy*, vol. 35, no. 2, pp. 512–517, Feb. 2010, doi: 10.1016/j.energy.2009.10.018.
[50] K. Ho and Y. Hsu, "Short term load forecasting using a multilayer neural network with an adaptive learning algorithm," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 141–149, Feb. 1992, doi: 10.1109/59.141697.
[51] K.-L. Ho, Y.-Y. Hsu, C.-F. Chen, T.-E. Lee, C.-C. Liang, T.-S. Lai, and K.-K. Chen, "Short term load forecasting of taiwan power system using a knowledge-based expert system," *IEEE Trans. Power Syst.*, vol. 5, no. 4, pp. 1214–1221, 1990, doi: 10.1109/59.99372.
[52] P. K. Dash, H. P. Satpathy, and A. C. Liew, "A real-time short-term peak and average load forecasting system using a self-organising fuzzy neural network," *Eng. Appl. Artif. Intell.*, vol. 11, no. 2, pp. 307–316, 1998, doi: 10.1016/S0952-1976(97)00069-9.
[53] I. Drezga and S. Rahman, "Input variable selection for ANN-based short-term load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1238–1244, 1998, doi: 10.1109/59.736244.
[54] I. Drezga and S. Rahman, "Short-term load forecasting with local ANN predictors," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 844–850, 1999, doi: 10.1109/59.780894.
[55] T. Ahmad and H. Chen, "Utility companies strategy for short-term energy demand forecasting using machine learning based models," *Sustain. Cities Soc.*, vol. 39, pp. 401–417, May 2018, doi: 10.1016/j.scs.2018.03.002.
[56] N. J. Johannesen, M. Kolhe, and M. Goodwin, "Relative evaluation of regression tools for urban area electrical energy demand forecasting," *J. Cleaner Prod.*, vol. 218, pp. 555–564, May 2019, doi: 10.1016/j.jclepro.2019.01.108.
[57] T. Ahmad, H. Chen, R. Huang, G. Yabin, J. Wang, J. Shair, H. M. Azeem Akram, S. A. Hassnain Mohsan, and M. Kazim, "Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment," *Energy*, vol. 158, pp. 17–32, Sep. 2018, doi: 10.1016/j.energy.2018.05.169.
[58] C. Feng and J. Zhang, "Assessment of aggregation strategies for machine-learning based short-term load forecasting," *Electr. Power Syst. Res.*, vol. 184, Jul. 2020, Art. no. 106304, doi: 10.1016/j.epsr.2020.106304.
[59] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192–1205, Jan. 2018, doi: 10.1016/J.RSER.2017.04.095.
[60] M. Beccali, M. Cellura, V. Lo Brano, and A. Marvuglia, "Short-term prediction of household electricity consumption: Assessing weather sensitivity in a mediterranean area," *Renew. Sustain. Energy Rev.*, vol. 12, no. 8, pp. 2040–2065, Oct. 2008, doi: 10.1016/j.rser.2007.04.010.
[61] M. Li, J. Cao, J. Guo, J. Niu, and M. Xiong, "Response of energy consumption for building heating to climatic change and variability in tianjin city, China," *Meteorol. Appl.*, vol. 23, no. 1, pp. 123–131, Jan. 2016, doi: 10.1002/met.1537.
[62] H. Wang and Q. Chen, "Impact of climate change heating and cooling energy use in buildings in the united states," *Energy Buildings*, vol. 82, pp. 428–436, Oct. 2014, doi: 10.1016/j.enbuild.2014.07.034.
[63] L. Wang, X. Liu, and H. Brown, "Prediction of the impacts of climate change on energy consumption for a medium-size office building with two climate models," *Energy Buildings*, vol. 157, pp. 218–226, Dec. 2017, doi: 10.1016/j.enbuild.2017.01.007.
[64] H. E. Thornton, A. A. Scaife, B. J. Hoskins, and D. J. Brayshaw, "The relationship between wind power, electricity demand and winter weather patterns in great britain," *Environ. Res. Lett.*, vol. 12, no. 6, Jun. 2017, Art. no. 064017, doi: 10.1088/1748-9326/aa69c6.
[65] W. P. Bell, P. Wild, J. Foster, and M. Hewson, "Wind speed and electricity demand correlation analysis in the Australian National Electricity Market: Determining wind turbine generators' ability to meet electricity demand without energy storage," *Econ. Anal. Policy*, vol. 48, pp. 182–191, Dec. 2015, doi: 10.1016/j.eap.2015.11.009.
[66] G. Franco and A. H. Sanstad, "Climate change and electricity demand in california," *Climatic Change*, vol. 87, no. S1, pp. 139–151, Mar. 2008, doi: 10.1007/s10584-007-9364-y.
[67] D. B. Belzer, M. J. Scott, and R. D. Sands, "Climate change impacts on U.S. commercial building energy consumption: An analysis using sample survey data," *Energy Sour.*, vol. 18, no. 2, pp. 177–201, Mar. 1996, doi: 10.1080/00908319608908758.
[68] A. Pardo, V. Meneu, and E. Valor, "Temperature and seasonality influences on Spanish electricity load," *Energy Econ.*, vol. 24, no. 1, pp. 55–70, 2002, doi: 10.1016/S0140-9883(01)00082-2.
[69] S. S. Wyer, *A Treatise on Producer-Gas and Gas-Producers*. Sydney, NSW, Australia: Wentworth Press, 2019, p. 298.

[70] A. Azadeh, M. Saberi, and O. Seraj, "An integrated fuzzy regression algorithm for energy consumption estimation with non-stationary data: A case study of iran," *Energy*, vol. 35, no. 6, pp. 2351–2366, Jun. 2010, doi: 10.1016/j.energy.2009.12.023.

[71] H. Shakouri G. R. Nadimi, and F. Ghaderi, "A hybrid TSK-FR model to study short-term variations of the electricity demand versus the temperature changes," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1765–1772, Mar. 2009, doi: 10.1016/j.eswa.2007.12.058.

[72] R. Zmeureanu and G. Renaud, "Estimation of potential impact of climate change on the heating energy use of existing houses," *Energy Policy*, vol. 36, no. 1, pp. 303–310, Jan. 2008, doi: 10.1016/j.enpol.2007.09.021.

[73] W. E. Woodson, B. Tillman, and P. Tillman, *Human Factors Design Handbook*, 2nd ed. New York, NY, USA: McGraw-Hill, 1992, p. 846.

[74] R. Osczevski and M. Bluestein, "The new wind chill equivalent temperature chart," *Bull. Amer. Meteorol. Soc.*, vol. 86, no. 10, pp. 1453–1458, Oct. 2005, doi: 10.1175/BAMS-86-10-1453.

[75] *ISO-NEW England*. Accessed: Sep. 1, 2018. [Online]. Available: https://www.iso-ne.com/library

[76] C. Xu and H. Chen, "Abnormal energy consumption detection for GSHP system based on ensemble deep learning and statistical modeling method," *Int. J. Refrig.*, vol. 114, pp. 106–117, Jun. 2020, doi: 10.1016/j.ijrefrig.2020.02.035.

[77] P. Sedgwick, "Pearson's correlation coefficient," *BMJ*, vol. 345, no. 1, p. e4483, Jul. 2012, doi: 10.1136/bmj.e4483.

[78] A. Navot, L. Shpigelman, N. Tishby, and E. Vaadia, "Nearest neighbor based feature selection for regression and its application to neural activity," in *Proc. Adv. Neural Inf. Process. Syst.*, 2005, pp. 995–1002.

[79] H.-X. Zhao and F. Magoulès, "Feature selection for predicting building energy consumption based on statistical learning method," *J. Algorithms Comput. Technol.*, vol. 6, no. 1, pp. 59–77, Mar. 2012, doi: 10.1260/1748-3018.6.1.59.

[80] MathWorks. (2019). *Resilient Backpropagation*. MATLAB R2006a, p. 2019. [Online]. Available: http://de.mathworks.com/help/nnet/ref/trainrp.html

[81] (2018). *Linear Neural Networks*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/linear-neural-networks.html

[82] (2018). *Cascade-Forward Neural Network*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/cascadeforwardnet.html

[83] S. A. Abdalkader, "Backpropagation neural network algorithm for forecasting soil temperatures considering many aspects: A comparison of different approaches," in *Proc. 5th Int. Conf. Inf. Technol. (ICIT)*, May 2011, pp. 1–8.

[84] O. De Jesus and M. T. Hagan, "Backpropagation algorithms for a broad class of dynamic networks," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 14–27, Jan. 2007.

[85] M. J. D. Powell, "Restart procedures for the conjugate gradient method," *Math. Program.*, vol. 12, no. 1, pp. 241–254, Dec. 1977, doi: 10.1007/BF01593790.

[86] M. Moreira and E. Fiesler, "Neural networks with adaptive learning rate and momentum terms," *Tech. Rep.*, vol. 95, vol. 4, pp. 1–29, 1995.

[87] W. Xie, A. Noble, and A. Zisserman, "Layer recurrent neural network," in *Proc. 5th Int. Conf. Learn. Represent.*, Toulon, France, 2017, pp. 1–15.

[88] C. Zhang, H. Wei, X. Zhao, T. Liu, and K. Zhang, "A Gaussian process regression based hybrid approach for short-term wind speed prediction," *Energy Convers. Manage.*, vol. 126, pp. 1084–1092, Oct. 2016, doi: 10.1016/j.enconman.2016.08.086.

[89] P. Xue, Y. Jiang, Z. Zhou, X. Chen, X. Fang, and J. Liu, "Multi-step ahead forecasting of heat load in district heating systems using machine learning algorithms," *Energy*, vol. 188, Dec. 2019, Art. no. 116085, doi: 10.1016/j.energy.2019.116085.

[90] F. Amara, K. Agbossou, Y. Dubé, S. Kelouwani, A. Cardenas, and J. Bouchard, "Household electricity demand forecasting using adaptive conditional density estimation," *Energy Buildings*, vol. 156, pp. 271–280, Dec. 2017, doi: 10.1016/j.enbuild.2017.09.082.

[91] K. Papakostas, T. Mavromatis, and N. Kyriakis, "Impact of the ambient temperature rise on the energy consumption for heating and cooling in residential buildings of Greece," *Renew. Energy*, vol. 35, no. 7, pp. 1376–1379, Jul. 2010, doi: 10.1016/j.renene.2009.11.012.

[92] T. Ahmad and H. Chen, "Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches," *Energy Buildings*, vol. 166, pp. 460–476, May 2018, doi: 10.1016/j.enbuild.2018.01.066.

[93] Y. Chen and H. Tan, "Short-term prediction of electric demand in building sector via hybrid support vector regression," *Appl. Energy*, vol. 204, pp. 1363–1374, Oct. 2017, doi: 10.1016/j.apenergy.2017.03.070.

TANVEER AHMAD received the Ph.D. degree from the School of Energy and Power Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China, in July 2019. He is currently a Postdoctoral Fellow with the State Key Laboratory of Internet of Things for Smart City, Department of Electrical and Computer Engineering, University of Macao, Macao, China. His research interests include energy forecasting, city-scale and district-level load demand planning and management in the smart grid environment, and modeling and optimization of multiple smart energy systems.

DONGDONG ZHANG (Member, IEEE) was born in Jining, China, in 1990. He received the B.Eng. degree in electrical engineering and automation from the College of Electrical Engineering, Qilu University of Technology, Shandong, China, in 2013, the M.S. degree in electric power system and automation from North China Electric Power University, in 2016, and the Ph.D. degree in electrical engineering from Xi'an Jiaotong University, in 2019.

He is currently an Assistant Professor with Guangxi University. His research interests include modeling and optimization of multiple energy systems, energy market, and electrical machines and its driving system design.

WAHAB ALI SHAH was born in Pakistan, in 1989. He received the bachelor's degree from the CECOS University of IT Emerging Sciences (CEIT), Peshawar, Pakistan, in 2011, the master's degree from Near East University (NEU), Mersin, Turkey, in 2016, and the Ph.D. degree from the School of Electrical and Electronic Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China. Since 2019, he has been an Assistant Professor with the Department of Electrical Engineering, Namal Institute, Mianwali, Pakistan. His current research interests include observation and simulation of the long air-gap discharge in the laboratory, lightning to tall buildings in nature, insulation design of power systems, and high-voltage engineering.