Received 23 June 2022, accepted 11 July 2022, date of publication 19 July 2022, date of current version 28 July 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3192433

RESEARCH ARTICLE

A Short-Term Load Forecasting Model Based on Self-Adaptive Momentum Factor and Wavelet Neural Network in Smart Grid

MUHAMMAD ZULFIQAR[1,2], MUHAMMAD KAMRAN[2], MUHAMMAD BABAR RASHEED[3], (Senior Member, IEEE), THAMER ALQUTHAMI[4], (Senior Member, IEEE), AND AHMAD H. MILYANI[4]
[1]Department of Telecommunication Systems, Bahauddin Zakariya University, Multan 60000, Pakistan
[2]Department of Electrical Engineering, University of Engineering and Technology, Lahore, Lahore 54890, Pakistan
[3]Escuela Politécnica Superior, Universidad de Alcalá, 28805 Alcalá de Henares, Spain
[4]Electrical and Computer Engineering Department, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding authors: Muhammad Babar Rasheed (muhammad.rasheed@uah.es) and Thamer Alquthami (tquthami@kau.edu.sa)

The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia has funded this project, under grant no. (RG-34-135-42). The authors, therefore, acknowledge with thanks DSR for their technical and financial support. This project also received funding from the European Union Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 754382, GOT ENERGY TALENT.

ABSTRACT Short-term load forecasting plays an essential role in the efficient management of electrical systems. Building an optimization model that will enhance forecasting accuracy is a challenging task and a concern for electrical load prediction. Due to Artificial Neural Networks (ANNs), the final result depends on initial random weights and thresholds that affect the stability of the forecast. Although much devotion is being given to improving the forecast accuracy, convergence, complexity, and resilience need to be considered for stable predictive models. To overcome this limitation, this work has jointly considered the Wavelet Neural Network (WNN) and Self-Adaptive Momentum Factor (SAMF) to achieve fast convergence, stability, and high accuracy. The proposed hybrid model is developed by combining the Feature Engineering (FE) and SAMF with the WNN model. The FE removes the irrelevant data and shallow features to ensure high computational performance. In contrast, the SAMF combines the wavelet transform's time and frequency domain properties and adjusts the WNN model's corresponding parameters. This ensures the global optimum solution while returning accurate predictive results. Finally, the SAMF is used to tune the control parameters of WNN by initializing the random weights and thresholds to accelerate the convergence rate and improve the accuracy compared to the Back-Propagation (BP) method. The proposed hybrid model is tested on the real-time datasets taken from the Australian states of (New South Wales (NSW), and Victoria (VIC)). Experimental results show that the developed model outperforms other benchmark models such as WNN-IGA, BPNN, WNN-AMBA, and Enhanced WNN in terms of instability, rate of convergence, and accuracy.

INDEX TERMS Load forecasting, self-adaptive momentum factor, wavelet transform, wavelet neural networks, convergence accuracy.

I. INTRODUCTION

Short-term load forecasting (STLF) is an essential task in the energy management system (EMS) for maintenance scheduling [1], energy generation accumulation, operation and planning of an energy utility system [2], security and market demand assessments, load switching, cost reductions, and ensuring a continuous supply of electricity [3]. The incomplete and insufficient predictive capabilities can cause significant power losses and even power outages. Therefore, developing appropriate forecasting strategies and enhanced predictive capabilities has become crucial [4]. However, the resilience and accuracy of load forecasts (LFs) depend on various variables such as population perspective, weather,

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

geographic limits, time, consumer behavior, power load conditions, energy prices, holidays, law, technological progress, politics, and economy, environmental change, and social activities. These variables make forecasting a problematic task [5].

In recent years, LF approaches have been divided into three main classes: classical statistical techniques, intelligent forecasting processes, and multiple forecasting models, including mixed, composite, fused, and hybrid forecasting models. The classical statistical approaches mainly depend on statistical and mathematical modeling. They have Box-Jenkins models [6], the Regression Analysis (RA) method [7], Auto-Regressive Integrated Moving Average (ARIMA) [8], Exponential Soothing (ES) [9], Kalman Filtering (KF) method [10], and the State-Space (SS) model [11]. However, these approaches are processed by linear analysis and are not sufficient for nonlinear load series [12]. The intelligent forecasting techniques are processed in non-linear time series to enhance the effectiveness and performance of the forecasting model [13]–[18]. The model does not require quantitative correlation or complex mathematical modeling. Intelligent approaches are widely used in STLF because they are adequate to deal with indeterminate processes (like load time series) [19], [20]. For instance, ANN [21]–[23], Support Vector Machines (SVM) [24], Expert Systems (ES) [25] and Fuzzy Logic (FL) [26] have been developed. Furthermore, the two major network management factors, stability and adaptability characterize a network with fast convergence and stable learning [27]. The use of ANN for STLF has received a lot of attention. However, a single predictive model rarely works well in all cases [28], [29]. Each method has its drawbacks, and it is not always possible to achieve the desired level of accuracy. For example, linear regression cannot capture non-linear and seasonal features. The Gray Forecasting (GF) technique can only solve specific problems with exponential biases. The ES depends upon the knowledge base. The final result of ANN relies on the thresholds and initial random weights that influence the forecasting instability. Therefore, the use of multiple predictive models to leverage a single model has been presented [30]–[33] to improve the forecast accuracy. Researchers used several optimization algorithms to optimize the thresholds and initial random weights of the ANN to address the instability of the final result for building up the STLF hybrid model.

ANNs are the most widely implemented methods in forecasting electrical energy consumption. However, there are some advantages and disadvantages needed to be acknowledged when ANNs are used, as listed in Table. 1. Since the complexity of electrical energy system is very high due to several factors, the ability of ANN in performing non-linear analysis is an advantage in executing electrical energy consumption forecasting.

The forecasting combinations initiated by Elliott and Timmermann [34] have long been considered an effective and efficient method to intact the forecast stability. It is an improvement over a single model. Later, Diebold and Pauly [35] and Pesaran & Timmermann [36] developed some enhancements in forecasting combination methodology. In [37], the authors used the adaptability of individual models and combined different types of ANN to anticipate electric load to address these shortcomings,. The fully unified and fusion predictive models are insightful solutions that take full advantage of the desired functionalities of individual models to ensure exceptional efficacy [31], [33]. For ELF, a hybrid model based on improved differential evolution (IDE) and WNN is being devised [38]. To validate the proposed model, valid comparisons with other models such as genetic algorithm ANN (GA-ANN), evolutionary programming ANN (EP-ANN), and particle swarm optimization ANN (PSO-ANN) are performed. In [6], the prediction accuracy is improved by adjusting the SVR parameters in a hybrid model based on DE and SVR. The proposed model outperforms the SVR, BP-ANN, and regression models. A model thata combines SVR and the Fruitfly Algorithm (FFA) has been developed in [39] to solve the parameter selection problem and improve the accuracy of load prediction. In addition, new approaches have been developed that hybridize the firefly optimization algorithm (FFOA) with the SVR model to adjust hyperparameters and optimally guarantee accurate load prediction [4], [40]. The above hybrid model can be regarded as promising and valuable in the field of improving prediction accuracy by properly adjusting SVR hyperparameters. However, the authors of these articles focus on optimizing random weights and bias initialization, or properly adjusting and selecting hyperparameters. Also, none of these models considered accuracy, stability, and rate of convergence at the same time. As a result of numerous analyzes and investigations, one factor (optimization of random weighting and bias initialization, or correct setting and selection of hyperparameters) and one criterion (accuracy, stability, rate of convergence) are not enough. Therefore, a robust hybrid model is needed to overcome the problems of existing models while improving prediction accuracy and stability with fast convergence rates. This paper presents a navel WNN prediction model based on the SAMF to overcome low convergence problems and abruptly fall into the local optimum. The adaptive momentum factor (MF), which utilizes the wavelet theory, is a practical scheme to solve the shortcomings of conventional algorithms [41]–[43]. In article [44], the MF does not accelerate convergence in the initial stages. Still, it oscillates the error curve of the network, making it challenging to adjust the impulse term value. This paper proposes an adaptive MF with a weight update phase in which the network automatically adjusts the MF to solve the problem of dealing with a steep and slow failures on curved surfaces.

A. CONTRIBUTIONS

- This paper proposes a novel robust hybrid framework, which integrates FE and SAMF with WNN in order to to cater to non-linear time-series predictions. The FE overcomes the redundancy and irrelevancy (dimensionality

TABLE 1. Advantages and disadvantages of ANN.

Artificial Neural Networks	
Advantages	Disadvantages
1. A Neural Network can execute the task that a linear program cannot	1. The neural networks need training to operate
2. Neural networks are parallel in nature. When an element of the neural networks fails, it can continue without any problems	2. Needs to be emulated since the architecture of neural network is different from the architecture of microprocessor
3. A Neural Network can learn and no need to be programed	3. For a large neural network, high processing time is required
4. Can be executed in any application	
5. Can be implemented without any problem	

reduction) problem, and SAMF algorithm intelligently selects and tunes hyperparameters of the WNN model to improve the forecast accuracy and stability with a fast convergence rate simultaneously. The integration of FE and SAMF algorithm effectively boosts the performance of the SVR model.

- The WNN model has high computational complexity and is poor in processing uncertain information. In load forecasting, redundant and irrelevant features pose complexity that slows down the training process of WNN and also negatively affects the forecast accuracy.To overcome these problems, FE consists of a hybrid feature selector (HFS) technique by fusing two algorithms: XGboost and decision tree classifiers (DTC) schemes, to monitor and control the feature selection process and feature extractor (FX) based on RFE to resolve the dimensionality reduction issue. In this way, the computational efficiency of the SVR model is enhanced.
- The SAMF acquires and tunes the appropriate parameters of WNN model to avoid trapping into local optimum and returns accurate forecasting results. Besides, most literature studies are immersed in a forecast accuracy refinement. However, the forecasting model's effectiveness and productiveness are determined equally by its stability and convergence rate.
- To evaluate the effectiveness and applicability of the proposed framework, actual half-hourly load data of two states of Australia (New South Wales NSW) and Victoria (VIC)) are employed as a case study.
- Experimental results show that the proposed framework outperforms benchmark frameworks in terms of accuracy, stability, and convergence rate.

B. PAPER ORGANIZATION

The rest of the paper is organized as follows: The developed model is described in section 2. Section 3 describes the performance evaluation measures. Simulation results, performance assessment,comparison of models and discussion are described in Section 4. Finally, we have completed the work by showing possible future directions in section 5.

II. PROPOSED MODEL

This work presents a novel FE-SAMF-WNN model for STLF. The devised model has three main modules: (i) FE module comprising of hybrid feature selector (HFS) and extractor,

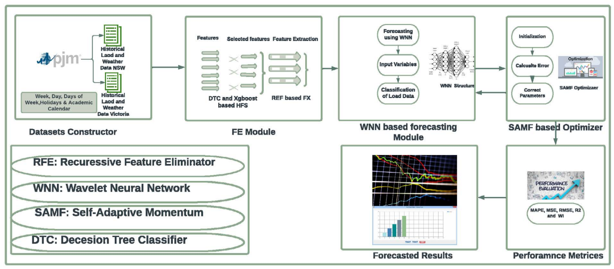


FIGURE 1. Schematic architecture of a new hybrid framework developed based on FE, WNN, SAMF for electrical load prediction.

(ii) forecasting module based on WNN, and (iii) optimization module based on SAMF. Firstly, the real site historical data are retrieved from cloud services or warehouses and given to the pre-processing phase, where cleansing, integration, rectification, outlier removal, normalization, and structuring operations are performed on the datasets to prepare the data for effective forecasting. The clean data is fed to the FE phase, which uses a hybrid feature selector and extractor to remove redundant, irrelevant, and ambiguous features and select the most relevant features. The prepared data is divided into training and testing data samples. The constructed training and testing samples are given to the WNN-based forecaster that forecasts future electric load. The forecaster output is given to the optimizer part of the framework based on the SAMF, which selects and tunes WNN hyperparameters to yield accurate, stable, and fast forecasting results. The pictorial diagram of the proposed framework is depicted in Fig. 1. In this developed hybrid model, SAMF plays an important role in optimizing WNN's thresholds and initial random weights. This enables the developed model to generate predictive significance and evaluate accuracy, stability, and rate of convergence.

A. FE

FE transforms the data into an easily interpretable format. Translucent data efficiently interprets intelligent model prediction problems and improves accuracy. FE picks and removes attributes from the input data so that the ML frameworks can sufficiently capture. Diverse data mining techniques include Principal component analysis (PCA), DTC, Mutual information (MI), Relief-F, Random Forest, and XGboost. These techniques pick and remove attributes from the data. XGboost and DTC among these techniques, are considered for feature selection (FS), and RFE is devised for feature extraction (FX). The FE module of the devised model comprises two phases: (i) FS part and (ii) FX part. The explicit illustration is as follows.

1) FS PHASE

This section describes the model's FS process. We have proposed a hybrid feature selector (HFS) that combines XGboost, DTC, and defined thresholds i.e. μ to control feature selection. HFS consists of two feature evaluators i.e. α and β. These two evaluators calculate the feature importance separately. In the FS process, the features are selected by joining the feature importance generated by the two evaluators. Feature selection is based U^{α} and U^{β}, which can be normalized by:

$$U^{\alpha} = U^{\alpha}/\max\left(U^{\alpha}\right) \tag{1}$$

$$U^{\beta} = U^{\beta}/\max\left(U^{\beta}\right) \tag{2}$$

Then the FS perform as:

$$F_s = \begin{cases} \text{reserve} & U^{\alpha}\left[\tau_j\right] + U^{\alpha}\left[\tau_j\right] > \mu \\ \text{drop,} & U^{\beta}\left[\tau_j\right] + U^{\beta}\left[\tau_j\right] \leq \mu \end{cases} \tag{3}$$

$U^{\alpha}\left[\tau_j\right]$ represents the feature importance calculated by evaluator XGBoost, $U^{\beta}\left[\tau_j\right]$ shows feature importance given by the DT. μ is the threshold controlling the FS. Features have also redundancy among them. To remove further redundancy and dimension reduction, they sent to FX pahse.

2) FX PHASE

The FX process is described in this section. The features selected by HFS are considered to have no irrelevant features, however, it contains redundant features. To reduce dimension and redundancy of features, RFE is applied for removing redundancy. To find a suitable low dimensional embedding, data needs non-linear mapping in electricity load forecasting. Thus RFE is applied to reduce nonlinear dimension.

B. WNN MODULE

Wavelet neural network (WNN) is a combination of wavelet analysis theory and neural network (NN). Wavelet analysis has the characteristics of time-frequency localization, which can analyze the data variously and extract the local information of the signal effectively [45]. The NN has the ability of self-learning and self-adaptation [46]. WNN has both the characteristics of wavelet analysis and NNs. The expansion factor and translation factor of the wavelet are adaptively adjusted during the training process of the network. Therefore, WNN can extract the local features of the training data to the maximum extent, which has a strong nonlinear approximation ability, stability, and faster convergence speed [47]. WNN is a significant improvement over the original NN with multi-layer feed-forward [7]. WNNs convert the original hidden layer activation function to a wavelet function, which has excellent symmetry and local features based on wavelet analysis. Considering $\theta(t) \in L^2(r)$, where $L^2(r)$ denotes

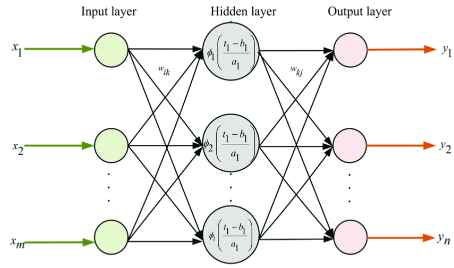


FIGURE 2. Structure of WNN.

the real-number square integral space and that the Fourier function $\theta(\omega)$, if $\theta(\omega)$ meets the following requirements as presented in Eq. 4:

$$\int_r \frac{\left|\theta'(\omega)\right|^2}{\omega} < \infty \tag{4}$$

However, $\theta(t)$ is a fundamental parental wavelet or wavelet feature and a required precondition for a wavelet function [48], [49]. Scalability (a) and shifting transformation (b) can have the following functions as represented in Eq. 5, after scaling and shifting:

$$\theta_{a,b}(t) = \frac{1}{\sqrt{a}}\theta\left(\frac{t-b}{a}\right) \tag{5}$$

where $\theta a, b(t)$ is the spectral analytical function. The following functions can achieved by taking the intermediate component of the signal $x(t)$ [50]:

$$f_x(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(t)\theta\left(\frac{t-b}{a}\right)dt \tag{6}$$

where $a > 0$, If the wavelets is linearized by transfer function:

$$f_x(a,b) = \frac{1}{\sqrt{a}}\Delta t\sum_{k=1}^{N} x(k\Delta t)\phi\left(\frac{k\Delta t - b}{a}\right) \tag{7}$$

In Eq. 7, the lens is pushed towards or away from the target by the scaling factor a, while the lens is parallel to the target by the translational factor b. By altering the wavelet basis function, the local characteristics of the signal can achieve localization of time series and time-frequency. The Morlet function [51] is mostly used, as presented in Eq. 8:

$$\phi(x) = \cos(1.70x)\exp\left(-\frac{x^2}{2}\right) \tag{8}$$

Wavelet transform has the properties of multi-resolution analysis in local features in the time and frequency domains. The network structure depicts in Fig. 2.

In Fig.2, to train WNN features, following three steps involved: (1) Assume that there are three layers of WNN, with M input neurons, I hidden layer neurons, and J output layer neurons. (2) Assume the input layer's m_{th} neuron is x_m, the hidden layer's i_{th} neuron is k_i, and the output layer's j_{th} neuron is y_j. (3) The relative weights of the connections ω_{mi} and ω_{ij}. Assume $f(.)$ is the hidden layer activation function (wavelet function), and $g(.)$ is the output layer scale parameter. Fig.3

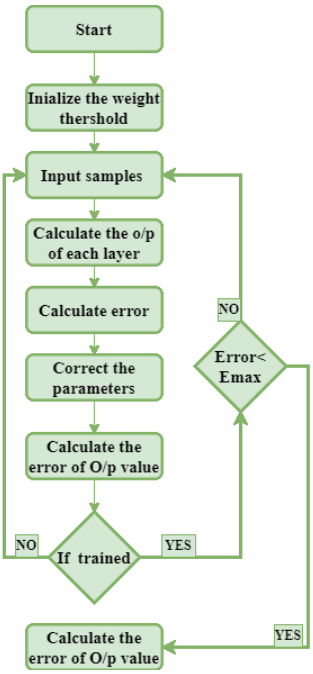


FIGURE 3. Training flow chart of WNN.

indicates the structural flow chart of the wavelet nerve. The output of the input layer is equal to the whole network's input signal, and the i_{th} neuron input is equal to the weighted amount of the input layer $v_M^m(n)$:

$$v_M^m(n) = x(n) \tag{9}$$

$$u_I^i(n) = \sum_{m=1}^{M} \omega_{mi} v_M^m(n) \tag{10}$$

As a result, the output of each neuron's i_{th} hidden layer is:

$$v_I^i(n) = f\left(u_I^i(n)\right) \tag{11}$$

The input of the j_{th} neuron in the output layer is equal to the weighted sum of $v_I^i(n)$:

$$u_J^j(n) = \sum_{i=1}^{I} \omega_{ij} v_I^i(n) \tag{12}$$

The output of the j_{th} neuron is in the output layer:

$$v_J^j = g\left(u_J^j(n)\right) \tag{13}$$

where the output layer's error is:

$$e_j(n) = d_j(n) - v_J^j(n) \tag{14}$$

The network's overall error is:

$$e(n) = \frac{1}{2}\sum_{j=1}^{J} e_j^2(n) \tag{15}$$

When the predicted error value exceeds the target error value set. First, alter the weighing parameter ω_{ij} between the hidden layer and the output layer. Calculate the gradient set $\frac{\partial e(n)}{\partial \omega_{ij}}$ of error $\omega_i j$, Adjust in the other way along this axis after that:

$$\Delta\omega_{ij}(n) = -\varphi\frac{\partial e(n)}{\partial \omega_{ij}} \tag{16}$$

$$\omega_{ij}(n+1) = \Delta\omega_{ij}(n) + \omega_{ij}(n) \tag{17}$$

The partial derivative may be used to calculate the gradient:

$$\frac{\partial e(n)}{\partial \omega_{ij}} = -e_j(n)g^t\left(u_J^j(n)\right)v_I^i(n) \tag{18}$$

where, the weight correction is expressed in Eq. 19 [52]:

$$\Delta\omega_{ij}(n) = \varphi e_j(n)g'\left(u_J^j(n)\right)v_I^i(n) \tag{19}$$

The local gradient is introduced [53]:

$$\begin{aligned}\delta_J^j &= -\frac{\partial e(n)}{\partial u^j(n)} = -\frac{\partial e(n)}{\partial e(n)}\cdot\frac{\partial e_j(n)}{\partial v^j(n)}\cdot\frac{\partial v_J^j(n)}{\partial u^j(n)}\\ &= e_j(n)g'\left(u_J^j(n)\right)\end{aligned} \tag{20}$$

As a result, the weight adjustment can be presented as $\Delta\omega_{ij}(n) = \varphi\delta_J^j v_I^i(n)$. We can compute the weight adaptation from the hidden layer to the input layer as:

$$\Delta\omega_{mi}(n) = \varphi\delta_J^j v_M^m(n) \tag{21}$$

Finally, through use of continual learning, decide if the error fulfills the criteria.

C. SAMF MODULE

The conventional NN is slowly converging and may easily be reduced to the optimal local practice [54], [55]. The method is sensitive to the momentum factor parameter since it has a substantial influence on the convergence rate and steady-state error. This paper offers a SAMF for WNN. Assume there are x, y, and z nodes in the input, hidden, and output layers, respectively. $S = \left(S_{u,v}\right)_{y\times z}$ is the loss function that connects the hidden and output layers, $T = \left(T_{u,v}\right)_{x\times y}$ is the input and hidden layer weight matrix, $h(.)$ is the function for activation and vector function definition [56] is described in Eq. 22:

$$F(x) = \left(h\left(x_1\right), h\left(x_2\right), \ldots, h\left(x_n\right)\right)^T \tag{22}$$

For any x, $x = (x_1, x_2, \ldots x_n)^T \in R^n$. For any input vector $\theta \in R^m$, the output of the network is

$$\gamma = F\left(S\left(F\left(T^\theta\right)\right)\right) \tag{23}$$

Among them $\gamma \in R^j$, let

$$\begin{aligned}w &= \left(S_{11}, S_{12}, \ldots, S_{1y}, S_{21}, \ldots S_{yz},\right.\\ &\qquad \left.T_{11}, \ldots, T_{1x}, T_{21}, \ldots, T_{xy}\right)^T\\ &= \left(w_1^T, w_2^T\right)^T\end{aligned} \tag{24}$$

where

$$w_1 = \left(S_{11}, S_{12}, \ldots, S_{1y}, S_{21}, \ldots S_{yz}\right)^T \tag{25}$$

$$w_2 = \left(T_{11}, \ldots, T_{1x}, T_{21}, \ldots, T_{xy}\right)^T \tag{26}$$

w_1, w_2 are the network proprietary vector values. The offline method updates the weight value after all of the samples are submitted and not according to the criterion if the samples are significant. As a result, the online algorithm uses here [57]. Set the goal output $O \epsilon R_j$ of the input $\theta \in R^m$, and the error presented in Eq. 27:

$$E = \frac{1}{2}\left(O - F\left(G\left(F\left(H^{\theta}\right)\right)\right)\right)^T \cdot \times \left(O - F\left(G\left(F\left(H^{\theta}\right)\right)\right)\right) \quad (27)$$

During the K_{th} training, samples θ^k, O^k are randomly selected from sample set N. Set $\nabla E(k)$ as the gradient of w in $w(k)$ when E is $\theta = \theta^k$, $O = O^k$, $k = \{1, 2, 3, \ldots, n\}$. At this point, when the initial value $w(0)$ and $w(1)$ are given, the weight value is updated in Eq. 28:

$$w(k+1) = w(k) - \varphi \nabla E(k) + \alpha(w(k) - w(k-1)) \quad (28)$$

make $\nabla_{w_1} E(k)$ and $\nabla_{w_2} E(k)$ respectively, when

$$\theta = \theta^k,$$

$O = O^k$, the vector weight at $w = w(k)$ is the gradient E. The online algorithm with adaptive algorithm weight update form at this moment has been established in Eq. 29:

$$w_1(k+1) = w_1(k) - \varphi \nabla_{w_1} E(k) + e^{-\delta - \left\|\nabla_{w_1} E(k)\right\|}\left(w_1(k) - w_1(k)\right) \quad (29)$$

$$w_2(k+1) = w_2(k) - \varphi \nabla_{w_2} E(k) + e^{-\delta - \left\|\nabla_{w_2} E(k)\right\|}\left(w_2(k) - w_2(k)\right) \quad (30)$$

The algorithm weight acquires the ability to withstand vibration and increase the convergence speed by incorporating the momentum component, where *delta* is a constant that governs the magnitude of the momentum factor. When the gradient is greater, the error changes more quickly, and the error hook face is steeper. When the gradient is small, the error change is gradual and the error hook face is rather uniform.

D. WNN MODEL PARAMETERS TUNING WITH SAMF

The hyperparameters settings influence the prediction accuracy and stability of the WNN model. Still, no recommended algorithm exists that can optimally tune the parameters of the WNN model. Thus, in this work, the SAMF is introduced to optimally tune hyperparameters of the WNN model to ensure accurate and stable forecasting performance. The developed SAMF determines optimal values of hyperparameters and takes the error function as an optimization objective to further improve forecast accuracy by reducing error. The error function is expressed in terms of the popular form MAPE as:

$$MAPE = \left(\frac{1}{M}\sum_{\tau-1}^{M}\frac{\left|u_\tau^r - u_\tau^\beta\right|}{\left|u_\tau^p\right|}\right) \times 100 \quad (31)$$

The MAPE is used as an objective function by the SAMF. Furthermore, the SAMF algorithm selects and tunes hyperparameters for optimal training of WNN. By using MAPE as an objective function, we determine the optimal function as per the best values of MAPE. Finally, WNN model employes the SAMF to get more accuarte and stable predictions with stable values of MAPE.

E. PROPOSED MODEL COMPUTATIONAL COMPLEXITY

This work demonstrates that the WNN performance is mainly determined by its parameter setting, and different parameter settings correspond to different data structures and characteristics. For this reason, the parameter setting is most informative in WNN modeling process and can represent the structure and characteristic of a specific training data set. However, the computational complexity of WNN is $O(Z \times M^3)$ (where M is the size of the training dataset, and Z is the evaluation number of the parameter selection process), and how to reduce this computational complexity is an open issue. Hence its application is often limited for large-scale data. More specifically, the WNN modeling process may suffer a computationally expensive inference, thus increase the running time for the WNN. Thus,WNN based on SAMF is an effective method to reduce the computational complexity of WNN learning.

Supposing that the number of training data is denoted by M, the number of model evaluation is denoted by Z, the size of momentum factor is denoted by f, the step size of SAMF-WNN is denoted by u, the dimension of input data is denoted by b. Generally, u is a relatively small number, $f \leq M$ the large sample learning. As mentioned earlier, the model selection complexity of the WNN model is $O(Z \times M^3)$. Thus, the model selection complexity of the proposed SAMF-WNN model is $O(Z \times u \times f^3)$. The complexity of SAMF algorithm in this paper is equivalent to the generation algorithm of non-repetitive random sequence, that is, the complexity of self-adapative momentum process is $O(u)$. In summary, the computational complexity of SAMF-WNN is $O(u) + O(Z \times u \times f^3) = O(Z \times u \times f^3)$. In order to compare the complexity of different non-linear models, the mainstream non-linear models support vector machine (SVM) is considered. The following Table 2 shows the computational complexities of standard WNN, SAMF-WNN and SVM [58]. We can observe from the above Table2 that the SAMF-WNN significantly speeds up the calculation speed of WNN: $O(Z \times u \times f^3) \leq O(Z \times M^3)$ indicates the complexity of SAMF-WNN is lower than the standard WNN; When The momentum factor satisfies $f < M^{1/3}$, the SAMF-WNN will have lower computational complexity than SVM. Furthermore, we also calculated the complexity of benchmark models and observed that proposed model have lower computational complexity as compared to the other benchmark models as depicted in Table 2.

III. PERFORMANCE ASSESSMENT METRICS

This work uses stability, accuracy, and rate of convergence to test the effectiveness of the developed model. Statistical tests and Diebold Mariano (DM) test [59] are used to justify these three goals.

TABLE 2. Computational complexities of proposed and other forecasting models (single and hybrid).

Models	Computational Complexity
WNN	$O(Z \times M^3)$
SAMF-WNN	$O(Z \times u \times f^3)$
SVM	$O(Z \times M^2)$
(SVM-SAMF	$O(Z \times u \times f^{2/3})$
BPNN	$O(Z \times M^{2/3})$
AMBA-WNN	$O(Z \times u \times f^2)$
IGA-WNN	$O(Z \times u \times f^{1.3})$
Proposed FE-SAMF-WNN	$O(Z \times u \times f^{2.7})$

A. STATISTICAL INDICATORS

Five standard statistical indicators are used: Pearson correlation coefficient, MAPE, MSE, RMSE, and Willmott's index (WI) [60].

$$MAPE = \left(\frac{1}{M}\sum_{\tau-1}^{M}\frac{\left|u_\tau^r - u_\tau^\beta\right|}{\left|u_\tau^p\right|}\right) \times 100 \tag{32}$$

$$MSE = \frac{1}{M}\sum_{\tau=1}^{M}\left(u_\tau^\alpha - u_\tau^\beta\right)^2 \tag{33}$$

$$RMSE = \sqrt{\frac{1}{M}\sum_{\tau=1}^{M}\left(u_\tau^\alpha - u_\tau^\beta\right)^2} \tag{34}$$

$$R = \frac{g\sum u_\tau^\beta u_\tau^0 - \sum u_\tau^\beta u_\tau^{\delta^\gamma}}{\sqrt{g\sum u_\tau^{\beta^2} - \left(\sum u_\tau^\beta\right)^2 - \sqrt{g\sum u_\tau^{\beta 2-(\sum u'_\tau)^2}}}} \tag{35}$$

$$WI = 1 - \frac{\frac{1}{M}\sum_{\tau=1}^{M}\left(u_\tau^B - u_\tau^\beta\right)^2}{\sum_{\tau=1}^{M}\left(\left|u_\tau^q - \bar{u}_\tau^\alpha\right| + \mid u_\tau^\tau - \bar{u}_\tau^\tau\right)^2} \tag{36}$$

where M is the number for forecasting steps, u_τ^α and $\bar{u}_\tau^\alpha$ are the targeted and mean targeted values, while u_τ^β and $\bar{u}_\tau^\beta$ are the forecasted and mean forecasted values respectively.

IV. SIMULATION RESULTS AND PERFORMANCE EVALUATION

This section describes the experimental results and performance evaluation of the developed FE-WNN-SAMF prediction model. The simulation is run on MATLAB 2017b to demonstrate the effectiveness of the developed model. The developed model is compared to benchmark models such as WNN-IGA [52], BPNN, WNN-AMBA [61], and Enhanced WNN (without SAMF). These models are considered benchmark models because their architecture is comparable to the proposed framework. To guarantee fair experiments between the proposed and benchmark models, we set the neuron numbers of ANNs according to the most accurate one. In general, there is not a clear theory to help people develop the neuron number of an ANN. This study selected the number of nodes by trial and error [62], [63]. After we did many experiments, we set the neuron number of each model according to the most accurate one (as shown in Table 3) to evaluate the proposed model more efficiently. Then, based on the obtained neuron number by trial and error, each ANN experiment was repeated 200 times to avoid the uncertainty effects. The reliability of the final results and the independence of the initial random weight values of the ANNs are ensured. The control parameters remained constant to compare the proposed and existing benchmark models properly. The developed model is evaluated using the historical data of AEMO. The data comprises New South Wales (NSW), and Victoria (VIC) states of Australia. Different variables (humidity, temperature, dew point, and hour of the day) that are typical for the training forecasting model are used. Historical information covers the four years from 2017 to 2021 at hourly intervals. The data consists of humidity, temperature, and load values. Historical information covers the four years from 2017 to 2021 at hourly intervals. The datasets goes through the FE module of the proposed model and transforms it into the desired format to determine the necessary functionality from the specified dataset. The processed data is divided into training datasets and test datasets. 80% of the information is used for training, and 20% is used for testing. The statistical data is shown in the Table 4.

A. FS USING XGboost AND DTC

In the FE module, XGboost and DTC-based HFS are applied to AEMO (NSW) data, selecting preferred attributes and discarding irrelevant attributes. Each feature sequence is in vector format. Vectors with different timestamps have feature values. To predict the energy load of the data, called load demand, we can extract features that slightly affect the energy load. HFS calculates the relationship between feature and electrical load, as shown in the Fig. 4. From the Fig 4, we can see that the order of most features exceeds the 0.5 value of μ. Five characteristics, such as DACC, RTMLC, RTCC, DAMLC, and RSP, are of lower grade than the selected μ and are therefore discarded during the selection process. Features with a value more excellent than μ are retained, and features with a value less than μ are discarded. Use different thresholds to control the FS process and visualize the importance of FE in the WNN model. For example, updating μ from 0.50 to 0.55, 0.65, and 0.75 will reduce RTLM, RTEC, Rgcp, and RT demand, respectively. The results show that as μ increases, many features are dropped, reducing prediction accuracy while increasing training speed. Important features are sent to the FX phase for dimension reduction.

B. FX USING RFE

Fig. 5 shows the FX through REF and comparison with PCA, accumulative contribution comparison among REF with PCA is shown. When the accumulative contribution rate reaches 95% REF extract most of the component thus we select REF to guarantee forecasting accuracy.

TABLE 3. Experimental parameter values.

Model	Experimental parameter	Default value
IGA-WNN	IGA population scale	100
	IGA maximum number of iteration times	500
	IGA cross rate	0.4
	IGA mutation rate	0.2
	Convergence tolerance	10−5
	WNN maximum number of iteration times	1000
	WNN convergence value	0.00004
	WNN learning rate	0.01
	Neuron number in the input layer	12
	Neuron number in the hidden layer	7
	Neuron number in the output layer	1
BPNN	Neuron number in the input layer	12
	Neuron number in the hidden layer	6
	Neuron number in the output layer	1
	Learning velocity	0.1
	Maximum number of trainings	1000
	Training requirements precision	0.00004
AMBA-WNN	Maximum number of iteration times of CS	500
	Population size of CS	100
	pa of CS	0.25
	Convergence tolerance of CS	10−5
	WNN maximum number of trainings	1000
	WNN learning rate	0.1
	WNN training requirement accuracy	0.00004
	Neuron number in the input layer	8
	Neuron number in the hidden layer	6
	Neuron number in the output layer	1
Enhanced-WNN	Maximum number of trainings	1000
	Learning rate	0.1
	Training accuracy requirement	0.00004
	Neuron number in the input layer	9
	Neuron number in the hidden layer	5
	Neuron number in the output layer	1
Proposed	Maximum number of iteration times	100
	Population size	100
	β_0	1.0/3.5
	γ	0.001
	Convergence tolerance	10−5
	GRNN radial basis function expansion	0.1−2.0
	GRNN maximum number of training	1000
	Training requirement precision	0.00004
	Neuron number in the input layer	7
	Neuron number in the pattern layer	7
	Neuron number in the summation layer	1
	Neuron number in the output layer	1

TABLE 4. Statistical parameters of the historical datasets. (NSW and VIC.)

Stats of Australia	Days/Week	Maximum	Minimum	Mean	Medium	STD
NSW	Monday	7144	5016	6210.95	7656	730.87
	Thursday	11 312	5908	8170.79	8756	1176.98
	Saturday	9354	6191	7581.5	8091	656.79
	Week	10538	6584	8298.9	8234	672.78
VIC	Friday	4407	2795	3542.31	4534	456.27
	Saturday	5757	3254	4763.04	4641	268.39
	Sunday	4782	3214	4069.39	4134	361.88
	Week	4993	3624	3989.95	4341	372.59

C. LEARNING ASSESSMENT

Australia's New South Wales training and testing datasets obtained from the FE module are sent to the SAMF-based WNN model for network training and validation. Fig. 6 shows thec learning behavior over several iterations using the testing and training datasets. The proposed framework is generalized. In addition, it is not affected by overfitting or underfitting. In addition, the results shown in Fig.6 that the proposed framework has a small gap between training and test errors. There is no variance and bais. Therefore, the developed model was trained to predict the electrical load of the day-ahead and week ahead. Investigations are conducted to estimate the effectiveness of the developed model, taking into account the rate of convergence, accuracy, and stability. The detailed test results are as follows:

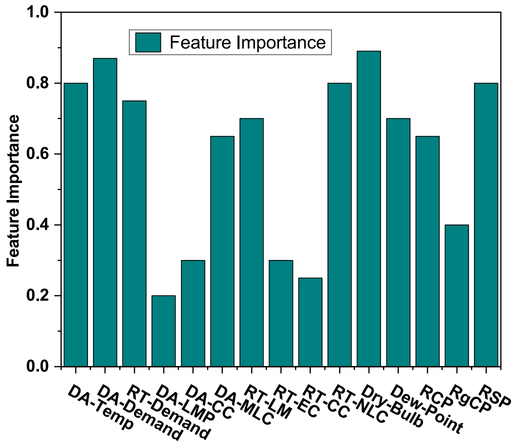


FIGURE 4. Features selected by HFS (XGboost & DTC): (a) NSW.

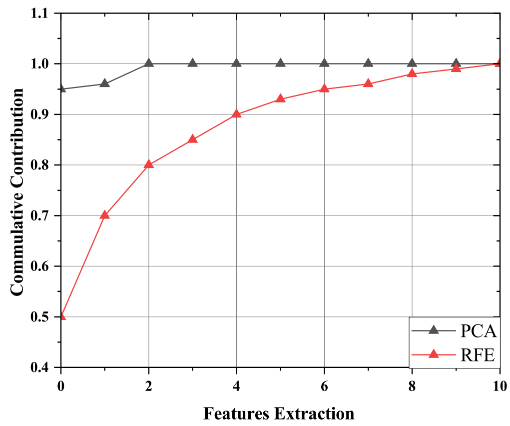


FIGURE 5. FX through REF and comparison with PCA.

D. TEST-I: AUSTRALIAN's STATE (NSW)

In this experiment, we consider actual hourly resolution-based data from the Australian state (NSW) as a case study to validate the developed model for STLF in terms of stability, accuracy, and rate of convergence. The Figs. 7, 8 and Tables 5, 6 show the pictorial and numerical results confirmed during the 200 iteration. The proposed model closely follows the target load curve approximated to the benchmark models. However, the developed model shows optimistic performance proximate to other models.

Results demonstrate the following observations:

1) The statistical analysis for the 200 iterative experiments is depicted in Table 5. The results show stability test. From the Table 5, MAPEs of the devised and benchmark frameworks are 0.3190, 0.3489, 0.5671, 0.4910, and 0.2190 respectively. The proposed model has lower MAPE as compared to the other models. This shows that the devised model is more stable and consistent as compared to the benchmarks models as shown in Tables 5 and 6.

TABLE 5. Stability analysis in terms of DM test & MAPE and forecasting accuracy of proposed and other benchmark models considering the Australian states (NSW, and VIC) hourly resolution load datasets.

Testing methods	Frameworks	Experiment-I (NWS)	Experiment-II (VIC)
DM test for stability	WNN-IGA	4.013	4.1815
	WNN-AMBA	2.2010	2.5255
	BPNN	2.9292	2.9169
	Enhanced WNN	2.4012	2.7151
	Proposed (FE-SAMF-WNN)	**0.5892**	**0.4911**
Stability test (MAPE)	WNN-IGA	0.3190	0.3289
	WNN-AMBA	0.5217	0.5134
	BPNN	0.6728	0.6217
	Enhanced WNN	0.6167	0.6145
	Proposed (FE-SAMF-WNN)	**0.2190**	**0.2219**
Forecasting Accuracy	WNN-IGA	95.6%	96.38%
	WNN-AMBA	93.2%	92.8%
	BPNN	90.1%	91.4%
	Enhanced WNN	87.14%	87.5%
	Proposed (FE-SAMF-WNN)	**98.5%**	**97.5%**

TABLE 6. Performance evolution of proposed and other benchmark frameworks considering hourly resolution load dataset of NSW based on MSE, MAPE, RMSE, R, and WI as well computational time.

Models	Days/Week	MAPE	MSE	RMSE	R	WI	Computational Time (s)
WNN-IGA	Monday	2.967	5467.15	31.1670	0.76	0.68	121
	Week	3.215	5317.17	30.1561	0.79	0.67	122
WNN-AMBA	Monday	2.617	6757.16	27.1781	0.81	0.72	145
	Week	3.167	6987.18	29.1781	0.79	0.71	143
BPNN	Monday	3.124	7689.12	31.1891	0.67	0.81	109
	Week	3.675	7845.11	30.1781	0.69	0.80	110
Enhanced WNN (without SAMF)	Monday	2.878	5689.18	34.1781	0.71	0.83	100
	Week	2.971	5781.1	35.1781	0.71	0.90	101
Proposed framework (FE-SAMF-WNN)	Monday	**1.4942**	**3918.16**	**20.1891**	**0.92**	**0.96**	**97**
	Week	**1.567**	**4017.13**	**22.1781**	**0.91**	**0.95**	**98**

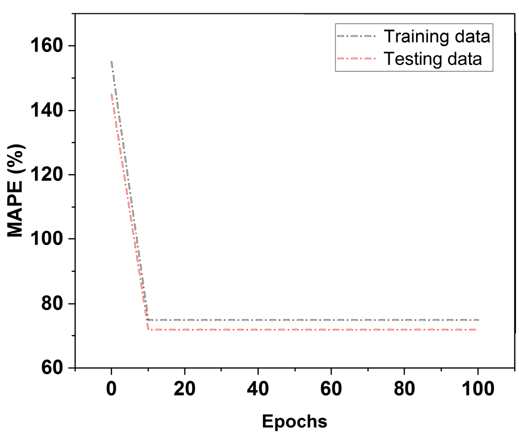
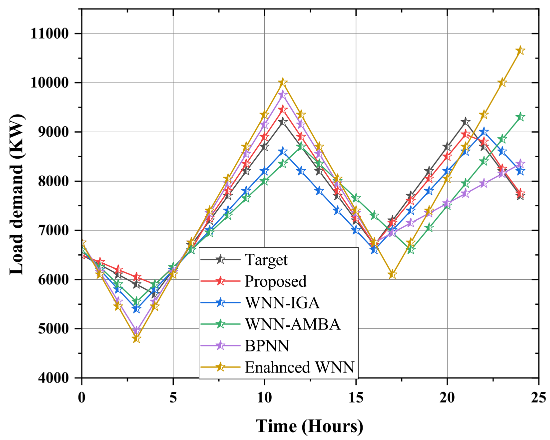


FIGURE 6. The learning assessment of devised model on training and testing datasets from AEMO (a) NSW.

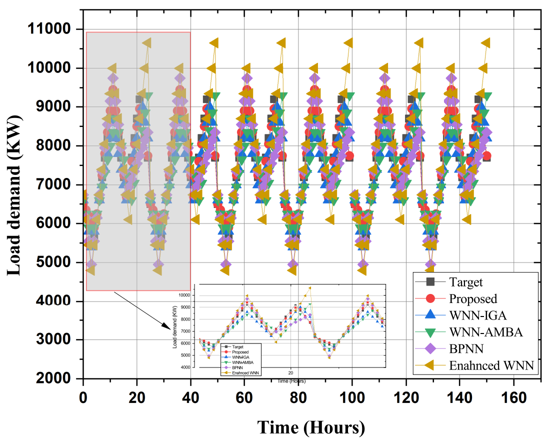
2) The devised model achieves an accuracy of 98.2% with the lowest mean error of 0.52% approximated to other existing models. Forecast accuracy results are shown in Table 5. The experimental investigation demonstrates that the proposed model for evaluating accuracy surpasses the other benchmark models.

3) The convergence rate of the devised model compared to the benchmark models is evaluated with three test functions. The test functions results are illustrated in Tables 6. We observed that the devised model converged after 21 iterations. While other benchmark frameworks such as WNN-IGA, WNN-AMBA, BPNN, and Enhanced WNN (without SAMF) converged after 43, 131, and 129 iterations, respectively. The network training times for the devised and other the benchmark models are 97s, 121s, 145s, and 105s, respectively. The results demonstrate that the proposed model has a high convergence rate with low computational complexity like other existing models.

Remark 1: The proposed framework is considered superior to the benchmark frameworks in all performance metrics based on experimental results. This remarkable performance is because the FE extracts repetitious, irrelevant, and extraneous features from the prepared data and determines the desired features that contribute significantly to forecasting efficiency. However, the SAMF optimizes the WNN model to enhance accuracy, stability, and convergence rate simultaneously. Therefore, integrating FE and SAMF can improve

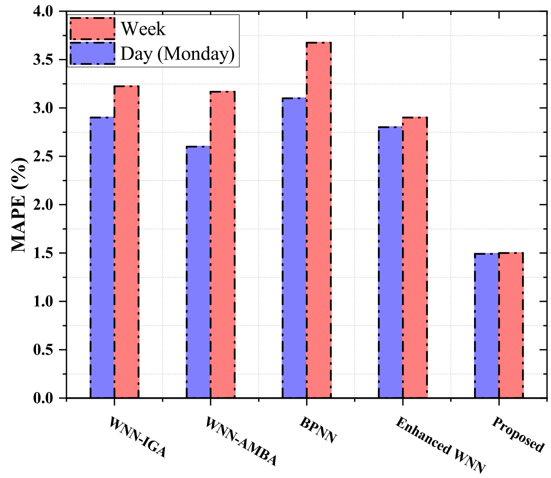


(a) day-ahead forecasting.

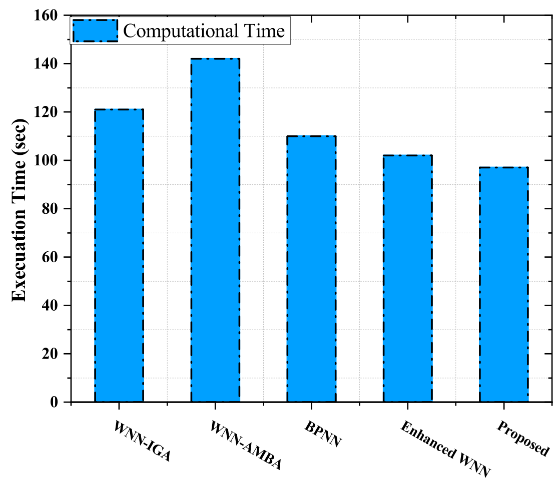


(b) Week ahead forecasting.

FIGURE 7. **Relative analysis of the devised FE-WNN-SAMF model on Australian's state (NSW) hourly load data considering stability, convergence rate, and accuracy: day-ahead forecasting (a) Monday; Week ahead forecasting (b) from 12/04/2017 to 12/04/2017.**



(a) Accuracy



(b) Convergence rate profile.

FIGURE 8. **Relative analysis of the devised FE-WNN-SAMF model on Australian's state (NSW) hourly load data considering stability, convergence rate, and accuracy: (a) Accuracy; (b) Convergence rate profile.**

predictive efficacy and be effectively used with different methods in other areas.

E. TEST-II: AUSTRALIAN's STATE (VIC)

In this experiment, we consider actual hourly resolution-based data from the Australian state (VIC) as a case study to validate the developed model for STLF in terms of stability, accuracy, and rate of convergence. The Tables 5, 8 show the numerical results confirmed during the 200 iteration. The proposed model closely follows the target load curve approximated to the benchmark models. However, the developed model shows optimistic performance proximate to other models.

Results demonstrate the following observations:

1) The statistical analysis for the 200 iterative experiments is depicted in Table 5. The results show stability test. From the Table 5, MAPEs of the devised and benchmark frameworks are 0.3289, 0.5134, 0.6217, 0.6145, and 0.2219 respectively. The proposed model has lower MAPE as compared to the other models. This shows that the devised model is more stable and consistent as compared to the benchmark models as shown in Tables 5.
2) The devised model achieves an accuracy of 97.5% with the lowest mean error of 0.49% approximated to other existing models. Forecast accuracy results are shown in Table 5. The experimental investigation demonstrates that the proposed model for evaluating accuracy surpasses the other benchmark models.
3) The convergence rate of the devised model compared to the benchmark models is evaluated with three test functions. The test functions results are illustrated in Tables 5. We observed that the devised model converged after 21 iterations. While other benchmark frameworks such as WNN-IGA, WNN-AMBA, BPNN, and Enhanced WNN (without SAMF)

TABLE 7. Relative assessment of WNN model with and without SAMF and FE, and WNN with other optimization modules for STLF.

State	Metrics	WNN model with and without FE, SAMF and Other optimization algorithms				
		Without FE & SAMF	With other optimization algorithms		With SAMF	With FE and SAMF
		WNN	WNN-IGA	WNN-AMBA	WNN-SAMF	**FE-SAMF-WNN**
NSW	MAPE	7.101	5.272	4.114	2.4194	1.2082
	MSE	9186.79	7911.75	6718.12	5189.12	4178.17
	RMSE	129.118	75.011	55.150	39.1111	25.9191
	R	0.5315	0.7191	0.7259	0.8188	0.9478
	WI	0.4484	0.7182	0.7712	0.7812	0.9319

TABLE 8. Performance evolution of proposed and other benchmark frameworks considering hourly resolution load dataset of VIC based on MSE, MAPE, RMSE, R, and WI as well computational time.

Models	Days/Week	MAPE	MSE	RMSE	R	WI	Computational Time (s)
WNN-IGA	Monday	2.819	5471.15	33.1670	0.78	0.71	121
	Week	3.198	5287.17	32.1561	0.80	0.69	123
WNN-AMBA	Monday	2.514	6687.16	26.1781	0.81	0.73	144
	Week	2.918	7001.18	30.1781	0.83	0.75	142
BPNN	Monday	3.017	7718.12	30.1891	0.69	0.82	109
	Week	3.571	7981.11	29.1781	0.70	0.83	110
Enhanced WNN (without SAMF)	Monday	2.910	5516.18	35.1761	0.81	0.85	100
	Week	3.101	5617.1	36.2781	0.78	0.91	101
Proposed framework (FE-SAMF-WNN)	Monday	**1.4942**	**3888.16**	**22.1791**	**0.91**	**0.94**	**96**
	Week	**1.567**	**4117.13**	**24.1811**	**0.89**	**0.93**	**97**

TABLE 9. comparisons of the mean difference of MAPE, RMSE and R^2 of all models across various datasets (NSW and VIC).

Model	Input variable		MAPE/%		RMSE		R^2		Run time (second)	
	History load	Temperature	NSW	VIC	NSW	VIC	NSW	VIC	NSW	VIC
WNN	added	Added	4.3652	5.4969	443.3448	532.0077	0.7507	0.7439	1.02	0.95
BPNN	added	Added	4.2673	4.9968	475.3198	513.4513	0.6795	0.6512	1.32	1.15
SVM	added	added	2.5255	3.0476	351.1483	346.4087	0.9064	0.8293	19.46	13.73
FE-SVM-PSO	added	added	2.5026	2.9890	253.5623	335.7369	0.9128	0.8513	72028.92	63530.27
WNN-IGA	added	No	2.4954	2.9465	183.4849	351.1443	0.9249	0.8646	186.23	104.76
WNN-AMBA	added	No	4.2136	4.8424	345.2895	431.6853	0.8695	0.8262	31620.85	21705.13
FE-SAMF-WNN	added	Added	0.7866	1.0783	113.9876	153.7615	0.9542	0.9247	18729.09	16169.29

converged after 44, 132, and 128 iterations, respectively. The network training times for the devised and other the benchmark models are 96s, 123s, 142s, and 103s, respectively. The results demonstrate that the proposed model has a high convergence rate with low computational complexity like other existing models. Furthermore, it is added that complexity of the framework depends upon the time at which it converged.

Remark 2: The proposed FE-SAMF-WNN framework is expected to achieve the lowest MAPE from the above experiments, demonstrating that the proposed model is more accurate than other benchmark models. On the other hand, FE-SAMF-WNN is the most stable framework and has a minor standard deviation of MAPE among the benchmark frameworks. FE-SAMF-WNN includes FE and the optimization process, but its response time is minimal compared to the benchmark frameworks. Therefore, it is clear that the proposed three-purpose framework is the most accurate, stable, and fastest of all the frameworks analyzed. The main reasons for choosing different Australian states such as NSW and VIC for the case study are the significant differences in geographic features, climatic characteristics, geographic location, economic development, industrial structure, and regional size. Suppose the proposed forecasting framework performs better in significantly diverse environments. In that case, it could be reasonably concluded that the developed framework has notable forecasting performance and broad applicability in various areas for the same indicator under different circumstances and is practicable in price, generation, and wind speed forecasting of the power system.

F. LOAD FORECASTING MODEL CONSIDERING SEASONAL FACTORS

The performance is ranked in terms of MAPE, RMSE and R^2 of the 200 times in NSW and VIC. Prediction accuracy measure of different models for the two testing datasets(NSW and VIC) are presented in details in Table 9. Table 9 shows the comparisons of MAPE, RMSE and R^2 of all models across various datasets. In all models, the single BPNN model has the worst performance on R^2. The reason might be the intrinsic limitations of the single BPNN model for load forecasting, such as overfitting on training sets and local optimum. Comparing the classical WNN-IGA model and the WNN-AMBA models, the WNN-AMBA model achieves

TABLE 10. The DM test values between the proposed model and other models in the two real-world datasets.

Model	NSW		VIC	
	DM	P-value	DM	P-value
WNN	7.2125	0.0000∗	7.6345	0.0000∗
BPNN	10.3444	0.0000∗	9.8127	0.0000∗
SVM	4.9834	0.0000 ∗	5.7852	0.0000∗
FE-PSO-SVR	6.1951	0.0000∗	6.2917	0.0000∗
WNN-IGA	5.5726	0.0000∗	4.9281	0.0000∗
FE-SAMF-WNN	6.9731	0.0000∗	7.1458	0.0000∗

lower accuracy because of not considering the other relevant covariates, such as the temperature. However, hybrid models WNN-AMBA and WNN-IGA obtain better performance than single models. Furthermore, FE-PSO-BPNN has good performance on R^2, which is due to the optimization technique of PSO and the feature engineering. Comparing the WNN-IGA model, the proposed FE-SAMF-WNN model is more accurate by using the optimization of SAMF. For the analysis of error measures above, the proposed hybrid model was achieved higher performance than other examined models. Error measures with the best performance are highlighted in bold in Table 9. Furthermore, the FE-PSO-SVR modeling framework proposed in [64], which dealt with univariate short-term load forecasting based on SVR optimized using PSO, is compared with the proposed model of this paper for emphasizing mainly the importance of exogenous variable and the effectiveness of multivariate decomposition. To further verify the performance of prediction, statistical method (DM test [59]) is employed to determine whether a statistically significant difference exist among the examined models for load forecasting. Note that the DM test can effectively eliminate the constraint of random sampling dynamics and provide comprehensive evaluation for accuracy and stability. The DM test results in the testing sets are shown in Table 4. The forecasting errors of single and different hybrid models were tested using the DM test, and the DM test results indicated that the FE-SAMF-WNN test result consistently exceed the 5% significance level upper bound in all cases. According to the comprehensively evaluation of DM test, it can be reasonably concluded that the proposed modeling framework not only has higher forecasting accuracy than other models, but also shows significant differences at a certain significant level, which further verifies the performance robustness of the proposed model in load forecasting.

G. DISCUSSION

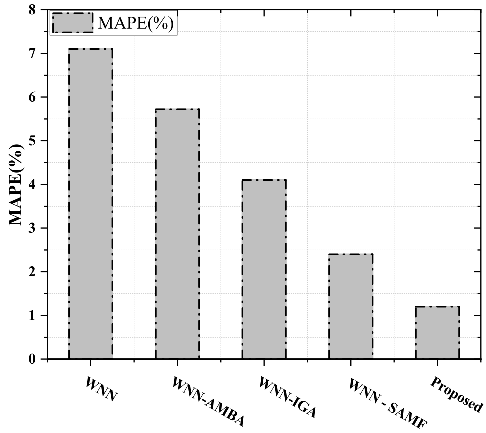
This section describes comparative assessments, including the importance of predicting outcomes and effectiveness with each purpose aspect. The proposed framework has reduced MAPE compared to WNN-IGA, WNN-AMBA, BPNN, and enhanced WNN (without SAMF) models. In addition, the proposed framework shows stable performance compared to WNN-IGA, WNN-AMBA, BPNN, and enhanced WNN (without SAMF) models. Similarly, the accuracy of the proposed framework is 98.2%. In contrast, WNN-IGA has an accuracy of 95.6%, WNN-AMBA has 93.2%, BPNN has 93.2%, and Enhanced WNN (without SAMF) has 87.54%. In addition, comparative experiments are performed using WNNs that do not use SAMF, WNNs that use SAMF, and WNNs that use other optimization algorithms. The usefulness of the devised model over WNN with and without various optimization algorithms is reflected in Fig. 9. The results acquired are shown in Table. 7. The performance of WNN optimized by SAMF is functional compared with WNN optimized by other benchmark algorithms because SAMF oscillates the error curve of the network in which the network automatically adjusts the momentum coefficient according to a steep and slow error on the curved surface. Hence, the devised three-objective FE-WNN-SAMF model is accurate, stable, and converged among the benchmark models as analyzed in Fig. 9 for future ELF.

H. COMPARISON OF WNN MODEL WITH SVM

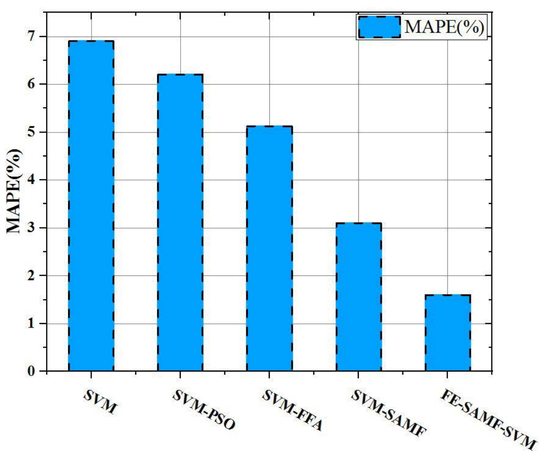
Comparative experiments are conducted using SVM without optimization algorithms, SVM with optimization algorithms, and SVM with proposed SAMF. The purpose is to reflect the comparison of the proposed WNN model with the SVM model. The obtained results are listed in Table 11. It is evident from the results that SVM without optimization algorithms has the highest MAPE (6.901), MSE (9089.89), RMSE (125.181), and lowest R (0.5189) and WI (0.4978). In contrast, the SVM with optimization algorithms (PSO and FFA) has a relatively low error and correlation statistics (R and WI) near 1. Thus, the results illustrate that the SVM model with optimization algorithms produced better performance (lowest MAPE, RMSE, MSE, and largest R and WI) than the SVM model without optimization algorithms. This relatively better behavior is due to the integration of the optimization algorithms with the SVM model. On the other hand, the SVM model with the SAMF has the lowest error statistics (MAPE (1.6082), MSE (4678.17), RMSE (30.9191)), and near one correlation coefficients (0.8878, 0.8919). The minimum error statistics and R and WI near 1.0 means a perfect match between observed and predicted values. This significantly improved the performance of the SVM model with the SAMF compared to the SVM model with other optimization algorithms because the SAMF algorithm uses an adaptive control procedure that improves both local and global searchability. From the comparison mentioned above, we observed that the single/individual SVM model has a relatively low error and correlation statistics (R and WI) near one as compared to single/individual WNN model, as depicted in Tables 7 and 11. In contrast, the hybridization of the WNN (with FE and SAMF) model has a relatively low error and correlation statistics (R and WI) near one as compared to the hybridization of SVM (with FE and SAMF). Based on the analysis, it is evident that the proposed framework (FE-SAMF-WNN) achieves the lowest MAPE values compared to the framework (FE-SAMF-SVM), which endorses that the devised model FE-SAMF-WNN has fast, accurate, and stable performance.

TABLE 11. Comparative assessment of SVM model with and without SAMF and FE, and SVM with other optimization modules for STLF.

State	Metrics	SVM model with and without FE, SAMF and Other optimization algorithms				
		Without FE & SAMF	With other optimization algorithms		With SAMF	With FE and SAMF
		SVM	SVM-PSO [65]	SVM-FFA [66]	SVM-SAMF	FE-SAMF-SVM
NSW	MAPE	6.90	6.278	5.124	3.1194	1.6082
	MSE	9089.89	7412.75	6567.12	5099.12	4678.17
	RMSE	125.181	76.077	65.145	49.110	30.9191
	R	0.5189	0.6917	0.7317	0.8210	0.8878
	WI	0.4978	0.7219	0.7612	0.7678	0.8919



(a) Relative evaluation of proposed WNN model with and without SAMF and FE for ELF.



(b) Relative evaluation of proposed SVM model with and without SAMF and FE for ELF

FIGURE 9. Comparison of Proposed WNN model with SVM model.

I. LIMITATIONS AND FUTURE WORK

Power grids are anticipated to be more complex with the development of emerging renewable energy (RE) technologies. The uncertainties of smart grid systems are increasing as many factors may simulate electricity demand. This paper focuses not on the future load demand from a long-term perspective but the short-term load fluctuation. The forecasting framework devised in this research does not consider other related factors but is only based on the detailed historical short-term load. Many key impacts may be missing, and there are also significant research gaps there. From the perspective of life cycle assessment (LCA), research on the whole system from "cradle to grave" is introduced, which can be employed in forecasting models. Moreover, scientific scenarios can be inducted to combine long-term and short-term forecasting, and more work must be done in related fields. Follow-up studies could be performed in the future work, including but not limited to:

- The forecasting model can consider additional factors or parameters to improve the STLF effectiveness.
- Research on energy systems, particularly the use of RE, needs to be studied so that the distribution and structures of future RE can be well known, which is a critical factor for STLF.
- Paying close attention to developing data cleaning technologies to deal with irregular and unstable short-term load data so that the adverse impacts of noise can be effectively controlled.
- A dynamic model selection strategy could be considered when determining the weights of hybrid or combined models.
- LCA-based modeling and scenario analysis can be introduced into forecasting models.
- More case studies in different smart grid systems could be done to show the scalability of the proposed forecasting model.

J. APPLICATIONS AND POLICIES OF STLF

STLF plays a pivotal role in balanced energy distribution, economics, and power system's secure and reliable operations. Accurate load forecasting reduces power grid collapse, alleviates costs and risks, improves power grid security, and helps policymakers in optimal planning and decision making, thereby making power grid cost-effective and environment-friendly. Moreover, it provides a strategic insight into the present power situation to estimate the amount of energy imported or exported. In the smart grid (SG), STLF is imperative for strategic decisions like operation and planning management, maintenance scheduling, load switching, power generation expansion, security and demand monitoring assessments, and ensuring a reliable provision of electricity. In contrast, electric load forecast overestimation or underestimation can introduce various challenges to the SG strategic decisions.

Electric load forecast overestimation results in the establishment of unnecessary spinning reserves, generation capacity, and insufficient energy distribution, leading to increased operating cost. Contrarily, electric load forecast underestimation poses reliability, power quality, security, and monitoring problems. Therefore, accurate and precise ELF is needed for SG distribution system operators (DSOs) to ensure sustainable, secure, and reliable power system operation. Inadequate and ineffective forecasting abilities of DSOs lead to significant power losses and even blackouts. For example, New York suffered from a severe blackout on 17 July 2006. Due to the blackout, many people of Queens borough of New York were affected, hundreds of businesses were stopped, prison complex at Rikers Island started their standby generators, escalators and elevators in commercial and industrial buildings were paralyzed, traffic signals were non-functional in roads and streets, and various incidents occurred at two terminals of the airport. The electrical power systems were entirely restored until 18 Aug 2006.

Similarly, a large blackout was recorded in Moscow on the morning of 25 May 2005. At the same time, the power system of Tula Oblast and Kaluga Oblast was also disrupted, which is 200 kilometers apart from Moscow. This blackout adversely affected residential activity, commercial business, industrial production, transportation, and communication services [3]. Such blackouts might be avoided if an early warning message is issued based on accurate and reliable forecasts to timely enact response measures. Thus, large-area blackout causes economic loss. Moreover, inaccurate prediction increases the utility cost; for example, a 1% increase in the prediction error will increase the overall utility cost by 10 million [67]. With this motivation, efforts are initiated to ensure fast, accurate, and stable load forecasting in order to guarantee secure, safe, and reliable power grid operation, which alleviates both economic and power losses [4].

V. CONCLUSION

STLF plays a vital role in the economics and safety of energy system operations. Promising forecasts have a significant effect on the economy. Recently, industry and academia are increasingly focusing on STLF. Exemplary predictive efficiency can minimize risks and costs, facilitate the security of energy systems, assist managers in making optimal plans, and improve grid management's financial and social benefits. Robustness, convergence rate, and accuracy are crucial in predictive models. It is desirable to build an STLF method that achieves these three aspects simultaneously. The present research proposes a navel WNN-based hybrid model FE-WNN-SAMF that uses SAMF to optimize the threshold values and initial random weights of the WNN for STLF. Furthermore, the FE module is integrated to improve computational performance and address dimensionality concerns. The goal is to achieve improved accuracy, outstanding stability, and fast convergence simultaneously. The developed model compares with benchmark frameworks such as WNN-IGA, WNN-AMAB, BPNN, and Enhanced WNN (without SAMF), considering stability, accuracy, and convergence rate. Based on many predictions and analyses, We can conclude that the effectiveness of the developed model outweighs the effectiveness of the other models considered.

Accordingly, the proposed hybrid model can become a suitable modeling framework in load forecasting. However, we considered only the historical load and the corresponding temperature as input variables of load forecasting. Some other lagged variables and more exogenous variables might be examined for improving the forecasting performance. Additionally, SAMF and its relevant factors are usually time-consuming to search for the optimal parameters of WNN. Therefore, some heuristic algorithms should be called upon for addressing the computational cost in the further investigation.

REFERENCES

[1] D. Masa-Bote, M. Castillo-Cagigal, E. Matallanas, E. Caamaño-Martín, A. Gutiérrez, F. Monasterio-Huelín, and J. Jiménez-Leube, "Improving photovoltaics grid integration through short time forecasting and self-consumption," *Appl. Energy*, vol. 125, pp. 103–113, Jul. 2014.

[2] F. Javed, N. Arshad, F. Wallin, I. Vassileva, and E. Dahlquist, "Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting," *Appl. Energy*, vol. 96, pp. 150–160, Aug. 2012.

[3] L. Xiao, W. Shao, C. Wang, K. Zhang, and H. Lu, "Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting," *Energy*, vol. 180, pp. 213–233, Oct. 2016.

[4] L. Xiao, W. Shao, T. Liang, and C. Wang, "A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting," *Appl. Energy*, vol. 167, pp. 135–153, Apr. 2016.

[5] T.-Y. Wang and H.-M. Chiang, "Solving multi-label text categorization problem using support vector machine approach with membership function," *Neurocomputing*, vol. 74, no. 17, pp. 3682–3689, Oct. 2011.

[6] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, "Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique," *Energy Buildings*, vol. 126, pp. 94–103, Aug. 2016.

[7] J. Zhang, S. Yan, Y. Liu, W. Zhu, and Z. Zhao, "A novel wavelet neural network load forecasting algorithm with adaptive momentum factor," in *Proc. IEEE 5th Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2021, pp. 1673–1678.

[8] N. Amral, C. S. Ozveren, and D. King, "Short term load forecasting using multiple linear regression," in *Proc. 42nd Int. Universities Power Eng. Conf.*, Sep. 2007, pp. 1192–1198.

[9] L. Wei and Z. Zhen-gang, "Based on time sequence of ARIMA model in the application of short-term electricity load forecasting," in *Proc. Int. Conf. Res. Challenges Comput. Sci.*, Dec. 2009, pp. 11–14.

[10] X. Zhang, J. Wang, and K. Zhang, "Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm," *Electr. Power Syst. Res.*, vol. 146, pp. 270–285, May 2017.

[11] W. Christiaanse, "Short-term load forecasting using general exponential smoothing," *IEEE Trans. Power App. Syst.*, vols. PAS-90, no. 2, pp. 900–911, Mar. 1971.

[12] H. Nie, G. Liu, X. Liu, and Y. Wang, "Hybrid of ARIMA and SVMs for short-term load forecasting," *Energy Proc.*, vol. 16, pp. 1455–1460, Jan. 2012.

[13] W.-M. Lin, H.-J. Gow, and M.-T. Tsai, "An enhanced radial basis function network for short-term electricity price forecasting," *Appl. Energy*, vol. 87, no. 10, pp. 3226–3234, Oct. 2010.

[14] J. Wang, L. Li, D. Niu, and Z. Tan, "An annual load forecasting model based on support vector regression with differential evolution algorithm," *Appl. Energy*, vol. 94, pp. 65–70, Jun. 2012.

[15] R. K. Jain, K. M. Smith, P. J. Culligan, and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy," *Appl. Energy*, vol. 123, pp. 168–178, Jun. 2014.

[16] H. Liu, H.-Q. Tian, D.-F. Pan, and Y.-F. Li, "Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks," *Appl. Energy*, vol. 107, pp. 191–208, Jul. 2013.

[17] C. García-Martos, J. Rodríguez, and M. J. Sánchez, "Modelling and forecasting fossil fuels, CO_2 and electricity prices and their volatilities," *Appl. Energy*, vol. 101, pp. 363–375, Jan. 2013.

[18] T. Alquthami, M. Zulfiqar, M. Kamran, A. H. Milyani, and M. B. Rasheed, "A performance comparison of machine learning algorithms for load forecasting in smart grid," *IEEE Access*, vol. 10, pp. 48419–48433, 2022.

[19] N. D. Uri, "Forecasting peak system load using a combined time series and econometric model," *Appl. Energy*, vol. 4, no. 3, pp. 219–227, Jul. 1978.

[20] K. Metaxiotis, A. Kagiannas, D. Askounis, and J. Psarras, "Artificial intelligence in short term electric load forecasting: A state-of-the-art survey for the researcher," *Energy Convers. Manage.*, vol. 44, no. 9, pp. 1525–1534, Jun. 2003.

[21] J. Song, J. Wang, and H. Lu, "A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting," *Appl. Energy*, vol. 215, pp. 643–658, Apr. 2018.

[22] Y. Xie, J. Liu, T. Huang, J. Li, J. Niu, C. M. Mak, and T.-C. Lee, "Outdoor thermal sensation and logistic regression analysis of comfort range of meteorological parameters in Hong Kong," *Building Environ.*, vol. 155, pp. 175–186, May 2019.

[23] K. Berk, A. Hoffmann, and A. Müller, "Probabilistic forecasting of industrial electricity load with regime switching behavior," *Int. J. Forecasting*, vol. 34, no. 2, pp. 147–162, Apr. 2018.

[24] Y. Li, J. Che, and Y. Yang, "Subsampled support vector regression ensemble for short term electric load forecasting," *Energy*, vol. 164, pp. 160–170, Dec. 2018.

[25] S. Rahman and R. Bhatnagar, "An expert system based algorithm for short term load forecast," *IEEE Trans. Power Syst.*, vol. PS-3, no. 2, pp. 392–399, May 1988.

[26] Y. Bodyanskiy, S. Popov, and T. Rybalchenko, "Multilayer neuro-fuzzy network for short term electric load forecasting," in *Proc. Int. Comput. Sci. Symp.* Russia: Springer, 2008, pp. 339–348.

[27] Y. Cai, J.-Z. Wang, Y. Tang, and Y.-C. Yang, "An efficient approach for electric load forecasting using distributed ART (adaptive resonance theory) & HS-ARTMAP (Hyper-spherical ARTMAP network) neural network," *Energy*, vol. 36, no. 2, pp. 1340–1350, Feb. 2011.

[28] J. Wang, Y. Song, F. Liu, and R. Hou, "Analysis and application of forecasting models in wind power integration: A review of multi-step-ahead wind speed forecasting models," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 960–981, Jul. 2016.

[29] H.-C. Kim, S. Pang, H.-M. Je, D. Kim, and S. Y. Bang, "Constructing support vector machine ensemble," *Pattern Recognit.*, vol. 36, no. 12, pp. 2757–2767, Dec. 2003.

[30] N. Liu, Q. Tang, J. Zhang, W. Fan, and J. Liu, "A hybrid forecasting model with parameter optimization for short-term load forecasting of micro-grids," *Appl. Energy*, vol. 129, pp. 336–345, Sep. 2014.

[31] F. Yu and X. Xu, "A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network," *Appl. Energy*, vol. 134, no. 134, pp. 102–113, Dec. 2014.

[32] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renew. Sustain. Energy Rev.*, vol. 34, pp. 243–254, Jun. 2014.

[33] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renew. Sustain. Energy Rev.*, vol. 50, pp. 1352–1372, Oct. 2015.

[34] G. Elliott and A. Timmermann, "Economic forecasting," *J. Econ. Literature*, vol. 46, no. 1, pp. 3–56, 2008.

[35] F. X. Diebold and P. Pauly, "Structural change and the combination of forecasts," *J. Forecasting*, vol. 6, no. 1, pp. 21–40, 1987.

[36] M. H. Pesaran and A. Timmermann, "Selection of estimation window in the presence of breaks," *J. Econometrics*, vol. 137, no. 1, pp. 134–161, Mar. 2007.

[37] P.-F. Pai and W.-C. Hong, "Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms," *Electr. Power Syst. Res.*, vol. 74, no. 3, pp. 417–425, 2005.

[38] G.-C. Liao, "Hybrid improved differential evolution and wavelet neural network with load forecasting problem of air conditioning," *Int. J. Electr. Power Energy Syst.*, vol. 61, pp. 673–682, Oct. 2014.

[39] G. Cao and L. Wu, "Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting," *Energy*, vol. 115, pp. 734–745, Nov. 2016.

[40] A. Kavousi-Fard, H. Samet, and F. Marzbani, "A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting," *Exp. Syst. Appl.*, vol. 41, no. 13, pp. 6047–6056, 2014.

[41] Q. Zhang, "Using wavelet network in nonparametric estimation," *IEEE Trans. Neural Netw.*, vol. 8, no. 2, pp. 227–236, Mar. 1997.

[42] S. Yao, C. J. Wei, and Z. Y. He, "Evolving wavelet neural networks for function approximation," *Electron. Lett.*, vol. 32, no. 4, p. 360, Feb. 1996.

[43] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox User's Guide*. Natick, MA, USA: The Math Works, 1996.

[44] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox User's Guide*. Natick, MA, USA: The Math Works, 2010, pp. 77–81.

[45] R. Y. Sun, Y. M. Shan, and B. Wang, "Study of pavement identification approach based on wavelet analysis," *Appl. Mech. Mater.*, vol. 44 pp. 3632–3639, Dec. 2010.

[46] H. M. Zhong, J. Liu, Q. F. Chen, and N. Liu, "A method based on wavelet neural network for power system short term load forecasting," *Appl. Mech. Mater.*, vol. 521, pp. 303–306, Feb. 2014.

[47] M. Zhu, W. D. Liu, and W. S. Hu, "Application of wavelet neural network in power load forecasting," *Adv. Mater. Res.*, vols. 181–182, pp. 439–442, Jan. 2011.

[48] C.-C. Yu and B.-D. Liu, "A backpropagation algorithm with adaptive learning rate and momentum coefficient," in *Proc. Int. Joint Conf. Neural Networks. (IJCNN)*, 2002, pp. 1218–1223.

[49] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network-based energy price forecasting in electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1423–1432, Aug. 2008.

[50] S. S. Iyengar and V. V. Phoha, *Foundations of Wavelet Networks and Applications*. London, U.K.: Chapman & Hall/CRC, 2018.

[51] J. Ashmead, "Morlet wavelets in quantum mechanics," 2010, *arXiv:1001.0250*.

[52] H.-J. Yang and X. Hu, "Wavelet neural network with improved genetic algorithm for traffic flow time series prediction," *Optik-Int. J. Light Electron Opt.*, vol. 127, no. 19, pp. 8103–8110, 2016.

[53] C. Bjorland, L. Caffarelli, and A. Figalli, "Non-local gradient dependent operators," *Adv. Math.*, vol. 230, nos. 4-6, pp. 1859–1894, 2012.

[54] Y. Zhongmin, P. Fei, S. Yuhui, W. Jinqiu, and Y. Xiaohong, "Short-term prediction of photovoltaic power generation output based on GA-BP and POS-BP neural network," *Power Syst. Protection Control*, vol. 43, no. 20, pp. 83–89, 2015.

[55] X. Ma, C. Zhou, and I. J. Kemp, "Interpretation of wavelet analysis and its application in partial discharge detection," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 9, no. 3, pp. 446–457, Jun. 2002.

[56] S. N. Chukanov, "Definitions of invariants for N-dimensional traced vector fields of dynamic systems," *Pattern Recognit. Image Anal.*, vol. 19, no. 2, pp. 303–305, Jun. 2009.

[57] K. Kawamura and A. Miyamoto, "Condition state evaluation of existing reinforced concrete bridges using neuro-fuzzy hybrid system," *Comput. Struct.*, vol. 81, nos. 18-19, pp. 1931–1940, Aug. 2003.

[58] A. Nafees, S. Khan, M. F. Javed, R. Alrowais, A. M. Mohamed, A. Mohamed, and N. I. Vatin, "Forecasting the mechanical properties of plastic concrete employing experimental data using machine learning algorithms: DT, MLPNN, SVM, and RF," *Polymers*, vol. 14, no. 8, p. 1583, Apr. 2022.

[59] F. Diebold and R. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 20, no. 1, pp. 134–144, 1995.

[60] P. Jiang, R. Li, N. Liu, and Y. Gao, "A novel composite electricity demand forecasting framework by data processing and optimized support vector machine," *Appl. Energy*, vol. 260, Feb. 2020, Art. no. 114243.

[61] J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, no. 1, pp. 774–781, Sep. 2018.

[62] J. Zhang and A. J. Morris, "A sequential learning approach for single hidden layer neural networks," *Neural Netw.*, vol. 11, no. 1, pp. 65–80, Jan. 1998.

[63] P. J. Gemperline, J. R. Long, and V. G. Gregoriou, "Nonlinear multivariate calibration using principal components regression and artificial neural networks," *Anal. Chem.*, vol. 63, no. 20, pp. 2313–2323, Oct. 1991.

[64] X. Wang and Y. Wang, "A hybrid model of EMD and PSO-SVR for short-term load forecasting in residential quarters," *Math. Problems Eng.*, vol. 2016, pp. 1–10, Dec. 2016.

[65] W. Hu, L. Yan, K. Liu, and H. Wang, "PSO-SVR: A hybrid short-term traffic flow forecasting method," in *Proc. IEEE 21st Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Dec. 2015, pp. 553–561.

[66] M. Nait Amar and N. Zeraibi, "A combined support vector regression with firefly algorithm for prediction of bottom hole pressure," *Social Netw. Appl. Sci.*, vol. 2, no. 1, pp. 1–12, Jan. 2020.

[67] G. Notton and C. Voyant, "Forecasting of intermittent solar energy resource," in *Advances in Renewable Energies and Power Technologies*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 77–114.

MUHAMMAD ZULFIQAR received the B.Sc. and M.S. degrees in electrical engineering from Bahauddin Zakariya University, Multan, Pakistan. He is currently pursuing the Ph.D. degree with the University of Engineering and Technology (UET), Lahore. He is a lifetime Professional Engineer with the Pakistan Engineering Council. He is working as a Lecturer with the Department of Telecommunication Systems, Bahauddin Zakariya University. His research interests include optimization, planning, energy management, and machine learning applications in smart/microgrids.

MUHAMMAD KAMRAN was the Dean of the Faculty of Electrical Engineering. He joined the University of Engineering and Technology (UET), Lahore, in 1994, after attaining industrial experience by serving PEL, SIEMENS, and NESPAK. His responsibilities include monitoring academic activities in electrical engineering, computer engineering, computer science, and biomedical engineering programs in all campuses of UET. Moreover, his responsibilities include teaching graduate and undergraduate courses and the management of faculty. Furthermore, academic issues related to the B.Sc. degree in electrical and biomedical engineering technology are also dealt which include curriculum modifications and NTC accreditation. Since December 2007, he has been a Professor with the Department of Electrical, Electronics and Telecommunication, UET (New Campus) and a HEC Approved Supervisor for the M.Sc. and Ph.D. degrees Program. He actively takes courses at the Center for Energy Research and Development at New Campus. He has been nominated as the Curriculum Committee Member of the Higher Education Commission in Electronics. He is a member of the selection board of various national universities. He has successfully supervised two Ph.D. Students and 33 master's students. He is currently supervising almost ten scholars in postgraduate studies of which three are Ph.D. students. He is looking forward to engaging students in various industrial and research-oriented projects.

MUHAMMAD BABAR RASHEED (Senior Member, IEEE) received the master's and Ph.D. degrees from COMSATS University, Islamabad, in 2013 and 2017, respectively. He is currently working as a GET-COFUND Marie Curie Fellow with the Universidad de Alcalá (UAH), Spain. Previously, he was serving as an Associate and an Assistant Professor for the Department of Electronics and Electrical Systems, The University of Lahore, Pakistan. He has authored over 40 peer-reviewed papers in well-reputed journals and conference proceedings and supervised/supervising more than ten students in their final year projects and theses. His research interests include LP, NLP, heuristic optimizations, machine learning, smart grids, electric vehicles, and demand response. He has obtained postdoctoral fellowships from Durham University, U.K., and King Abdulaziz University (KAU), Saudi Arabia, in 2019 and 2020, respectively. He is an active Reviewer of many esteemed journals and conferences including IEEE TRANSACTIONS, IEEE ACCESS, IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS, *Applied Energy*, and *Energies*.

THAMER ALQUTHAMI (Senior Member, IEEE) received the Doctor of Philosophy (Ph.D.) degree in electrical engineering with a minor in mathematics from the Georgia Institute of Technology. He is an experienced Assistant Professor with a demonstrated history of working in higher education and industry. His research interests include smart grids, renewable energy, power system operation and control, complex system modeling and simulation, energy audit, energy efficiency and savings, and data analytics. He is skilled in PSCAD/EMTDC, Python, PTI/PSSE, RTDs, Building Automation Implementation, R-Statistics, and C++.

AHMAD H. MILYANI received the B.Sc. (Hons.) and M.Sc. degrees in electrical and computer engineering from Purdue University, in 2011 and 2013, respectively, and the Ph.D. degree in electrical engineering from the University of Washington, in 2019. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, King Abdulaziz University, Jeddah, Saudi Arabia. His research interests include power systems operation and optimization, renewable and sustainable energy, power electronics, electric vehicles, and machine learning.