IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 1 November 2023, accepted 22 January 2024, date of publication 24 January 2024, date of current version 16 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3358182

RESEARCH ARTICLE

Short-Term Load Forecasting in Smart Grids Using Hybrid Deep Learning

MASHAEL M. ASIRI [1], GHADAH ALDEHIM [2], FAIZ ABDULLAH ALOTAIBI [3], MRIM M. ALNFIAI [4], MOHAMMED ASSIRI [5], AND AHMED MAHMUD [6]

[1]Department of Computer Science, College of Science and Art at Mahayil, King Khalid University, Abha 61421, Saudi Arabia
[2]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[3]Department of Information Science, College of Arts, King Saud University, Riyadh 11437, Saudi Arabia
[4]Department of Information Technology, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia
[5]Department of Computer Science, College of Sciences and Humanities Aflaj, Prince Sattam bin Abdulaziz University, Aflaj 16273, Saudi Arabia
[6]Research Center, Future University in Egypt, New Cairo 11835, Egypt

Corresponding author: Mrim M. Alnfiai (m.alnofiee@tu.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/ 65 /44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R387), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2024R838), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

ABSTRACT Load forecasting in Smart Grids (SG) is a major module of current energy management systems, that play a vital role in optimizing resource allocation, improving grid stability, and assisting the combination of renewable energy sources (RES). It contains the predictive of electricity consumption forms over certain time intervals. Load Forecasting remains a stimulating task as load data has exhibited changing patterns because of factors such as weather change and shifts in energy usage behaviour. The beginning of advanced data analytics and machine learning (ML) approaches; particularly deep learning (DL) has mostly enhanced load forecasting accuracy. Deep neural networks (DNNs) namely Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) have achieved popularity for their capability to capture difficult temporal dependencies in load data. This study designs a Short-Load Forecasting scheme using a Hybrid Deep Learning and Beluga Whale Optimization (LFS-HDLBWO) approach. The major intention of the LFS-HDLBWO technique is to predict the load in the SG environment. To accomplish this, the LFS-HDLBWO technique initially uses a Z-score normalization approach for scaling the input dataset. Besides, the LFS-HDLBWO technique makes use of convolutional bidirectional long short-term memory with an autoencoder (CBLSTM-AE) model for load prediction purposes. Finally, the BWO algorithm could be used for optimal hyperparameter selection of the CBLSTM-AE algorithm, which helps to enhance the overall prediction results. A wide-ranging experimental analysis was made to illustrate the better predictive results of the LFS-HDLBWO method. The obtained value demonstrates the outstanding performance of the LFS-HDLBWO system over other existing DL algorithms with a minimum average error rate of 3.43 and 2.26 under FE and Dayton grid datasets, respectively.

INDEX TERMS Energy management, short-term prediction, artificial intelligence, hyperparameter optimization, bio-inspired algorithm.

I. INTRODUCTION

With the incorporation of digital control and information and communication (ICT) technologies, Smart grid (SG), an intelligent technology has attracted wide attention [1], [2], [3]. The combination of these current technologies can have a high effect on managing and making decisions on user energy consumption [4]. On the other hand, equalizing the supply-demand profile and the optimum decision appears hard with no consideration of the actual load-demand profiles

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

with higher performance [5]. Hence, particular focus is previously provided for carrying electricity load forecasting (ELF) [6], [7]. While these approaches and techniques give satisfactory outcomes. Still, the incorporation of unstable energy resources because of the dynamic load demand patterns [8]. Market variations, it is quite a chance to enhance the performance of these methods with respect to error-probability, accuracy, complexity, and so on [9]. Besides, other aspects namely unstable climate, humidity, temperature, social standards, calendar signs, and occupancy patterns also affected the prediction accuracy [10], [11].

Accurate load forecast refers to both usability and provider to increase their electricity rate savings because of spot rate launch—most causes that usability indicate developing attention towards SG application [12]. The involved usability predictions of the future rate or load signal can depend on the previous activity of the energy consumption of the users [13]. With respect to the estimated rates or load signals, the user changes their energy consumption plans subjected to reduced electrical energy rate and comfort levels [14]. Load prediction depends on time to be categorized such as short-term, medium-term, and long-term predictions. The prediction time is usually over months in medium-term forecasting [15]. These methods are employed by usability for maintenance scheduling, hydro reservoir management, and fuel preparation. Further, short-term load prediction is divided into very short-term, and short-term forecasts [16]. In very short-term prediction, a forecast time from minutes or seconds to hours and model application in flow control. In short-term predicting, it forecasts periods from hours to weeks and models applications in modifying demand and formation, thus, utilized to establish provides for the electricity market [17]. The short-term prediction methods have been important in daily activities, estimation of net exchange, unit commitments and planning operation, and analysis of system security. In long-term prediction, the forecast period is for years. Usability employs these categories of methods for devising the capability of maintenance and grid planning [18]. With machine learning (ML) technologies exploited widely in industries, data-driven techniques are progressively utilized for predicting and analyzing load data like support vector machine (SVM), deep learning (DL), relevance vector machine (RVM) and random forest (RF) methods [19]. Meanwhile, accurate load prediction is required by usability to correctly devise the current grid functions for effective controlling of their resources, This study targets to accurate load forecasting approach [20].

This study designs a Short-Load Forecasting scheme using a Hybrid Deep Learning and Beluga Whale Optimization (LFS-HDLBWO) approach. The main aim of the LFS-HDLBWO system is to predict the load in the SG environment. To accomplish this, the LFS-HDLBWO technique initially uses a Z-score normalization method for scaling the input dataset. Besides, the LFS-HDLBWO methodology makes use of convolutional bidirectional long short-term memory with an autoencoder (CBLSTM-AE) model for load prediction purposes. Finally, the BWO algorithm is used for optimal hyperparameter selection of the CBLSTM-AE algorithm, which helps to enhance the overall prediction results. A wide-ranging experimental analysis was made to show the improved predictive outcomes of the LFS-HDLBWO technique. In short, the key contributions are summarized as follows.

- An automated LFS-HDLBWO technique comprising Z-score normalization, CBLSTM-AE-based prediction, and BWO-based hyperparameter tuning has been developed for load prediction in the SG environment. This unique fusion controls the strengths of either DL or optimizer approaches for enhancing load forecasting.
- By employing the CBLSTM-AE approach, the LFS-HDLBWO method utilizes a sophisticated DL structure that captures complex temporal dependencies and designs from load data, enhancing the accuracy of load forecasts.
- The BWO methodology was combined for optimum hyperparameter selection of the CBLSTM-AE approach. This optimizer step fine-tunes the DL approach, contributing to higher load prediction outcomes.

The rest of the paper is organized as follows. Section II provides the related works and section III offers the proposed model. Then, section IV gives the result analysis and section V concludes the paper.

II. RELATED WORKS

Motwakel et al. [21] introduced an innovative wild horse optimizer technique with a DL-based STLF system (WHODL-STLFS) for SGs. This developed algorithm is primarily employed to develop the WHO method for the optimum choosing features. Moreover, attention-based long short-term memory (ALSTM) is deployed for learning the activities of electricity load demand. Lastly, an artificial algae optimizer (AAO) method is implemented as the hyperparameter optimization of the ALSTM technique. Alrasheedi and Almalaq [22] designed hybrid DL algorithms to improve the effectiveness of Saudi SG load forecasting. A benchmark approach by various standard DL techniques comprising LSTM, conventional neural networks (CNNs), artificial neural network (ANN), gated recurrent unit (GRU), recurrent neural network (RNN), and various actual databases are utilized to confirm these developed techniques. In [23], a novel hybrid clustering-based DL algorithm was presented for STLF at the distributing transformer's level with higher scalability. A k-Medoid-based method could be exploited for clustering while the predicting methods have been produced for several clusters of load profiles. The clustering of the distributed transformer is dependent upon the resemblance in power utilization profiles. It contains 6-layers and utilizes Adam optimizer through the TensorFlow model.

Chen et al. [24] suggested a hybrid framework depends on ResNet and LSTM techniques. Primarily, the information with the number of feature parameters has been recreated and input to ResNet to extract features. Secondarily, the feature extraction vectors are employed as an input of LSTM for short-term load prediction. Finally, a real-time model has been employed for comparing this approach with other methods. In [25], a hybrid architecture was designed by integrating Feature Engineering and SAMF with the WNN technique. Feature Engineering extracts the unrelated data and superficial features to ensure higher computation efficiency. Alternatively, the SAMF incorporates the wavelet transform and time-frequency field characteristics as well as modifies the WNN framework's related features. In conclusion, the SAMF was employed for tuning the control measurements of WNN. Zarei and Ghaffarzadeh [26] introduced a multiobjective optimizer of the AC optimum power flow (AC-OPF) issue corresponding to demand-response (DR). The DR-based OPF method includes reducing the system rate via the concurrent involvement of reactive and active power in DR and improving the computation accuracy by demand forecast dependent upon earlier data utilizing the DL algorithm. Lastly, the better DR values were resolved by utilizing the TOPSIS approach.

Chen et al. [27] designed a new ResNet-based technique for the load forecasting of the next 24 hr. This developed approach consists of a backbone network, ensemble model, feature extraction method, and ResNet. The multiscale features could be extracted from raw data to provide them into one snapshot paradigm that has been modelled with a backbone network and a ResNet. In [28], the authors developed a multi-output Gaussian processes (MOGP) regression method for forecasting 24 load values of the following days according to aspects like temperature, dew, and load point values of earlier days. The effectiveness of this presented MOGP technique was studied and compared to the determination and multi-linear regression techniques.

Ali et al. [29] present a control technique that concentrates on a sophisticated FL method. Advanced fuzzy control takes overloading and difference from demand profile as input that mitigates these disturbances by integrating optimum power dispatch of renewable energy resources (RERs). Hong et al. [30] examine a short-term residential load forecasting structure that creates usage of the spatio-temporal correlation present in appliances' load data by the DL method. In addition, this technique dependent upon DNN and iterative ResBlock has been presented for learning the correlation among distinct power consumption behaviors for STLF. In [31], it has been decided that the load spatial-temporal distribution and the load can be applicable in the SG, and the influence of load spatial-temporal distribution was assumed in the forecasting model. It sifted multi-variate load series extremely appropriate to forecast points with correlation analysis and modeled the load sequences from the perspective of time with LSTM.

III. THE PROPOSED MODEL

In this study, we have designed an automated load prediction using the LFS-HDLBWO approach on the SG environment. The main aim of the LFS-HDLBWO technique is to predict the load in the SG platform. To accomplish this, the LFS-HDLBWO method includes Z-score normalization, CBLSTM-AE model, and BWO-assisted hyperparameter tuning. Fig. 1 demonstrates the total flow of the LFS-HDLBWO system.

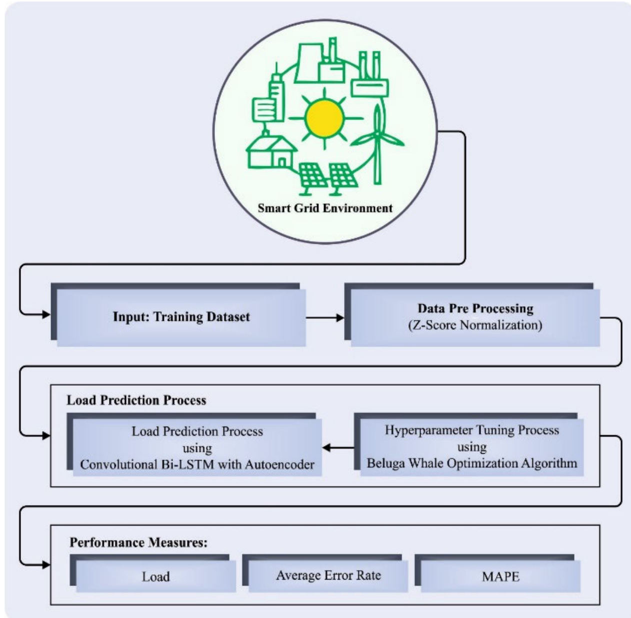

FIGURE 1. Overall flow of LFS-HDLBWO algorithm.

A. DATA PRE-PROCESSING

To pre-process the input, Z-score normalization is utilized. Z-score normalization, otherwise called standardization, is a statistical approach used to standardize and transform arithmetical data by modifying it to a combined scale. This procedure includes computing the Z-score for all the data points that characterize how many standard deviations a data point is additionally the mean of the database. The equation to calculate the Z-score of data point x with standard deviation σ and mean μ is shown below:

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

B. LOAD PREDICTION USING THE CBLSTM-AE MODEL

For the prediction of the load in the SG environment, the CBLSTM-AE model is used. The proposed model comprises numerous frameworks involving 2 CNN layers and an AE model encompassing of LSTM layer as a decoder and BLSTM as an encoder [32]. Fig. 2 illustrates the infrastructure of the CBLSTM-AE approach.

1) CNN

CNN can store varied irregular trends and is mainly skilled at extracting complex features. These features decrease the parameters required to make predictions, thereby decreasing the network computation. The CNN uses the concept of weight distribution in non-linear problems including ECP. The CNN has a hidden layer comprising an activation function, a pooling layer, and a convolution layer. The convolution layer converts input data into feature mapping. Next, the pooling layer samples the mapping features for extracting convolution features, thus decreasing the size of feature mapping. The downsampling process and feature extraction of CNN decreases the computing time, which makes it best-fitted for the model.

2) BLSTM-AE

The output layer of CNN can be provided as input to the BLSTM layer that acts as an input to the AE layer. However, the CNN extract relevant features and the BLSTM-AE layers are applied for sequence prediction and data analysis. BLSTM combines the data sequence in predicting backwards and forward directions. At the same time, *AE* is especially adapted for learning representation, to realize unsupervised input in the feature vectors. It includes an encoding and decoding model for encoding the input series previously decoded with internal representation. Therefore, the Bi-LSTM-AE is used to learn the temporal dependency of the data, positively affecting the prediction outputs.

3) LSTM-AE

As opposed to the BLSTM of the encoder, an LSTM-AE has been utilized as a decoder to decrease the complexity of this developed architecture. Also, the single LSTM can able to learn from the temporal dependency of the data. Before we proceed to two FC layers, the data encoded in the output of Bi-LSTM-AE could be decoded through a single layer of LSTM-AE for the last prediction outputs. Consider the input vector $x_i^m = \{x_1, x_2, \cdots, x_n\}$, where x^m signifies the varied input vectors, and provides the input vectors x_{i^m} into CNN layers, the resultant output can be shown in Eq. (2).

$$y_{ij}^m = \sigma\left(b_j^m + \sum_{m=1}^{M} w_{m,j}^1 x_{i+m-1,j}^0\right) \tag{2}$$

where b_j^m refers to the bias for j^{th} feature maps and σ denotes the activation function. y_{ij}^m shows the output vector x_{ij}^m of the prior layer. m implies the index value of the filter, w denotes the kernel weight, and Eq. (3) shows the output vector for the k^{th} convolution layer.

$$y_{ij}^{m(k)} = \sigma\left(b_j^{m(k)} + \sum_{m=1}^{M} w_{m,j}^{m(k)} x_{i+m-1,j}^0\right) \tag{3}$$

The pooling and convolution layers downsample the activation from mapping features to lessen the network computation costs and several parameters. The max-pooling layer shown in (4) exploits the maximal value from the prior layer for the down-sampling that helps to adjust the model over-fitting.

$$P_{ij}^{m(k)} = \max_{r \in R} y_{i \times T + rj}^{k-1} \tag{4}$$

In Eq. (4), T indicates the stride defining the length of the input dataset and y denotes the pooling size. The output from the max-pooling layer can be provided to the input of Bi-LSTM layers via the gating unit. BLSTM encompasses forget, input, and output gates in both directions, and all the gates are activated once the memory cell updates the state using the following expressions:

$$i_t = \sigma\left(W_{piPt} + W_{hi}h_{t-1} + W_{ci} \cdot c_{t-1} + b_i\right) \tag{5}$$
$$f_t = \sigma\left(W_{pfPt} + W_{hf}h_{t-1} + W_{cf} \cdot c_{t-1} + b_f\right) \tag{6}$$
$$o_t = \sigma\left(W_{poPt} + W_{ho}h_{t-1} + W_{Co} \cdot c_t + b_o\right) \tag{7}$$

Here P_t represents the output of the maxpooling layer at t time. σ represents the activation function, c_t shows the cell state, and b denotes to the bias. h_t denotes the hidden layer of BLSTM cells viz., updated at t steps in both directions. i_t, f_t, and o_t correspondingly show the input, forget, and output gates. The hidden and cell states are defined by the gating units of the BLSTM as follows:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \sigma\left(W_{pc}p_t + W_{hc}h_{t-1} + b_c\right) \tag{8}$$
$$h_t = o_t \cdot \sigma\left(c_t\right) \tag{9}$$

The output layer of Bi-LSTM is concatenated to these two directions as follows:

$$\bar{y} = \sigma\left(\overleftrightarrow{W_y h_t} + b_y\right) \tag{10}$$

where the resulting output $\hat{y} = \sigma(W_y h_t + b_y)$ shows the input to a 2 FC dense layers formulated in (11) for the last estimated output. The output of Bi-LSTM $\bar{y}$ is considered as an input of decoded LSTM.

$$d_i^k = \sum_j w_{ji}^k - 1\left(\sigma\left(\hat{y}_i^{k-1}\right) + b_i^{k-1}\right) \tag{11}$$

The Bi-LSTM-AE accept the feature from the CNN, and the CNN layer extracts spatial features in the input dataset for learning temporal dependency.

C. HYPERPARAMETER TUNING USING THE BWO ALGORITHM

Finally, the BWO method can be utilized for the optimum hyperparameter selection of the CBLSTM-AE method. BWO is a new swarm-based metaheuristic approach based on the behaviours of beluga whales in nature [33]. Like other metaheuristic techniques, BWO comprise of exploration stage and the exploitation stage. The steps of basic BWO are summarized below.

Initialization stage: The beluga whale is utilized as a searching agent owing to the population-assisted BWO. All the whales have been considered as a possible solution that can be updated. The position of the searching agent can be shown in the matrix form as follows:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \tag{12}$$

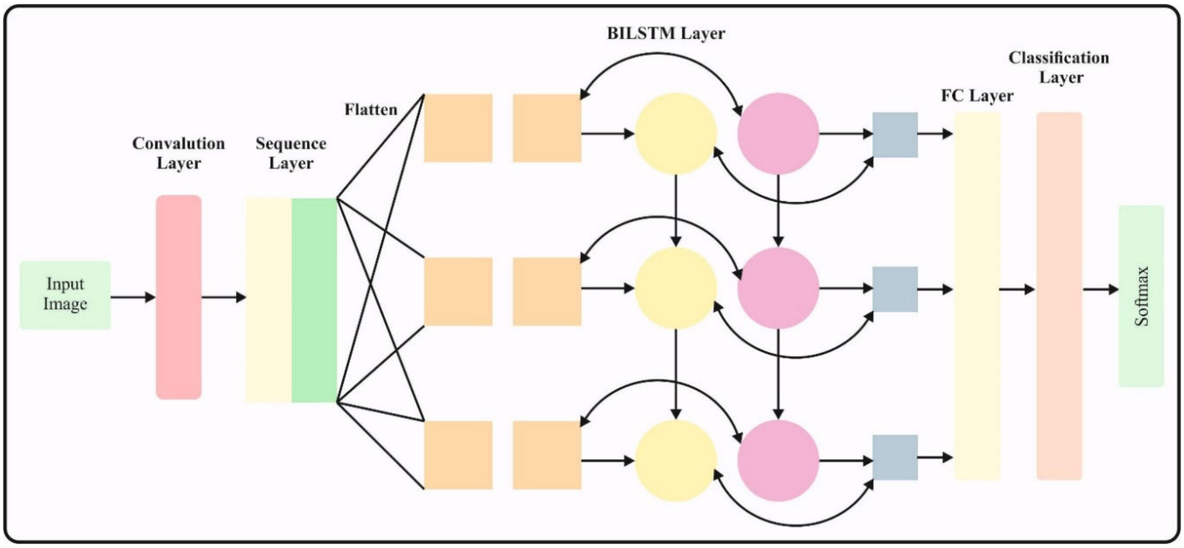

FIGURE 2. Architecture of CBLSTM-AE model.

In Eq. (12), n denotes the population size of searching agents and d refers to the dimensionality of the problem. The fitness value for beluga whales can be stored using the given matrix:

$$F_x = \begin{bmatrix} f\left(x_{1,1}, x_{1,2}, \ldots, x_{1,d}\right) \\ f\left(x_{2,1}, x_{2,2}, \ldots, x_{2,d}\right) \\ f\left(x_{n,1}, x_{n,2}, \ldots, x_{n,d}\right) \end{bmatrix} \tag{13}$$

The balance factor B_f is represented as:

$$B_f = B_0\left(1 - \frac{T}{2T_{\max}}\right) = \begin{cases} > 0.5, & \text{exploration stage} \\ \leq 0.5, & 0 \text{ exploitation stage} \end{cases} \tag{14}$$

In Eq. (14), $T_{\max}$ shows the existing iterations , (T) is the maximal amount of iterations. In all the iterations, B_0 refers to randomly generated integers within [0, 1]. The variation range of B_f is reduced at (0,1) to (0,0.5) as T rounds increase, which shows the significant modification in possibilities for the exploration and exploitation stages, whereas the probabilities of the exploitation stage increase as T rounds increase.

Exploration phase: The exploration phase includes an examination of the swimming behaviors of the beluga whale. The location of the whale can be defined by the pair swim, and this position is changed by the following expression:

$$\begin{cases} X_{i,j}^{T+1} = X_{i,p_j}^{T} + \left(X_{r,p_1}^{T} - X_{i,p_j}^{T}\right)(1 + r_1)\sin(2\pi r_2), \\ \qquad j = even \\ X_{i,j}^{T+1} = X_{i,p_j}^{T} + \left(X_{r,p_1}^{T} - X_{i,p_j}^{T}\right)(1 + r_1)\cos(2\pi r_2), \\ \qquad j = odd \end{cases} \tag{15}$$

where T refers to the latest iteration, $X_{i,j}^{T+1}$ signifies the new location of i^{th} beluga whale on the j^{th} dimension, and p_j indicates the randomly generated value derived from the j^{th} set. The existing location for r^{th} and i^{th} are signified as $X_{r,P1}^{T}$ and $X_{i,Pj}^{T}$, correspondingly, r denotes the randomly chosen beluga whale, r_1 and r_2 are the randomly generated numbers within [0, 1], and cos $(2\pi r_2)$ and $\sin(2\pi r_2)$ shows that the fin of the mirrored face of the upward beluga whales.

Exploitation phase: This phase draws inspiration from the hunting behaviors of beluga whales. They forage and migrate together based on their closeness towards other whales. Therefore, beluga whale hunts by sharing data regarding the presented position for one another and selecting the best candidate amongst them. Levy flight (LF) method is added to the exploitation stage for improving convergence as follows:

$$X_i^{T+1} = r_3 X_{best}^{T} - r_4 X_i^{T} + C_1.L_f.\left(X_r^{T} - X_i^{T}\right) \tag{16}$$

In Eq. (16), X_i^T and X_r^T denote the existing location of the i^{th} beluga whales and a random beluga whale, correspondingly; X_i^{T+1} and X_{best}^{T} show the new location of i^{th} beluga whales and the optimal location for the whole population; r_3 and r_4 is a randomly produced value within [0, 1]; and $C_1 = 2r_4(1-T/T_{\max})$ shows the random jump strength.

The L_f, or LF function, is calculated by the following expression:

$$L_f = 0.05 \times \frac{u \times \sigma}{|v|^{1/\beta}} \tag{17}$$

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{(1+\beta)}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}}\right)^{1/\beta} \tag{18}$$

Now, β represents a constant as 1.5 and u and v indicate the random value with uniform distribution.

Whale falls: Polar bears, killer whales, and people pose a threat to beluga whales through foraging and migration. Mostly, whales are very smart and avoid the dangers of sharing data. Nonetheless, a few whales didn't thrive and died in the Deep Ocean. This condition is named 'whale falls,' and provides food for the other creatures. The BWO replicate the process of whale falls by choosing the possibility of whale falls from the individual population. According to the possibility of whale fall viz., selected to deal with the changes in the group, it is assumed that the beluga whale either fell into the deep sea or migrated. The step size and the position of beluga whales where they fall are used to define the updated location for keeping the population size constant. The mathematical formula can be given as:

$$X_i^{T+1} = r_5 X_i^T - r_6 X_r^T + r_7 X_{step} \tag{19}$$

In Eq. (19), r_5, r_6 and r_7 represent the random number lies in [0, 1], and X_{step} refers to the whale fall's step size and is formulated by Eq. (20):

$$X_{step} = (u_b - l_b) \exp\left(-\frac{C_2 T}{T_{\max}}\right) \tag{20}$$

$$C_2 = 2W_f \times n \tag{21}$$

Now C_2 shows the step factor which relates to the population size and probability of whale falls, and u_b and l_b indicate the upper and lower boundaries, correspondingly. The boundaries of the design variable, the maximal amount of iterations, and the iteration, each having an impact on C_2.

In the BWO model, the possibility of a whale falling (W_f) is computed as a linear function:

$$W_f = 0.1 - \frac{0.05T}{T_{\max}} \tag{22}$$

During the optimization process, the probability of whale falls dropped from 0.1 to 0.05, representing the risk of beluga whale reductions as they approach the food source.

The BWO determines the hyperparameter included in the CBLSTM-AE algorithm. The MSE is regarded as an objective function and is shown below:

$$MSE = \frac{1}{T} \sum_{j=1}^{L} \sum_{i=1}^{M} \left(y_j^i - d_j^i\right)^2 \tag{23}$$

whereas M and L correspondingly denote the resulting values of layer and data, y_j^i and d_j^i show the obtained and proper magnitude for $thej^{th}$ unit from the resulting layer at t time.

IV. PERFORMANCE VALIDATION

This section inspects the load predictive results of the LFS-HDLBWO method on two datasets: the FE GRID dataset and the Dayton GRID dataset. For comparative result analysis, the proposed model is compared with existing models [21] such as FCRBM, AFC-ANN, Bi-level, MI-ANN, and LSTM.

The commonly used measures to evaluate the performance of the predictive results are defined as follows:

- Average Error Rate (AER) is a metric utilized for evaluating the outcome of a predictive method or system, classically in the context of forecast or predictive tasks.

TABLE 1. Predictive outcome of LFS-HDLBWO technique on FE Grid dataset.

FE GRID Dataset					
Load (kW)					
Timeslots (hrs)	Actual	LSTM	AFC-ANN	BiLevel	LFS-HDLBWO
1	689	672	665	669	687
2	694	677	683	699	693
3	712	712	702	743	711
4	731	747	752	779	729
5	751	777	793	800	749
6	776	763	811	816	775
7	785	766	794	799	782
8	783	751	762	784	780
9	762	733	743	757	760
10	753	725	733	751	748
11	739	718	719	738	740
12	723	706	727	702	721
13	715	709	716	727	716
14	721	740	747	757	721
15	736	770	768	784	734
16	766	775	789	794	764
17	786	776	799	782	785
18	786	760	794	772	784
19	759	745	783	762	758
20	747	705	747	740	745
21	722	685	709	699	720
22	703	677	686	695	702
23	674	683	675	675	673
24	688	688	648	669	687

The AER has been computed by summing up the separate errors (variances between predictive and actual values) and dividing by the total count of predictions.

- Mean Absolute Percentage Error (MAPE) is a generally utilized accuracy metric to evaluate the outcome of forecasting methods, particularly in the context of time sequence analysis and prediction. MAPE procedures the average percentage difference between the predictive and actual values.
- Execution Time (EXET) for load prediction refers to the count of time it takes for a load-forecasting system to create predictions.

Table 1 and Fig. 3 reveal the predictive outcomes of the LFS-HDLBWO method on the FE Grid dataset [21]. The outcomes show that the LFS-HDLBWO method accomplishes closer predictive values. On 1hr, the LFS-HDLBWO technique predicted the load of 687kW with the actual value of 689kW. Besides, on 10hr, the LFS-HDLBWO technique predicted a load of 748kW with an actual value of 753kW. Along with that, in 24 hours, the LFS-HDLBWO technique predicted a load of 687kW with an actual value of 688kW.

Table 2 and Fig. 4 reveal the predictive analysis of the LFS-HDLBWO system on the DAYTON Grid dataset. The

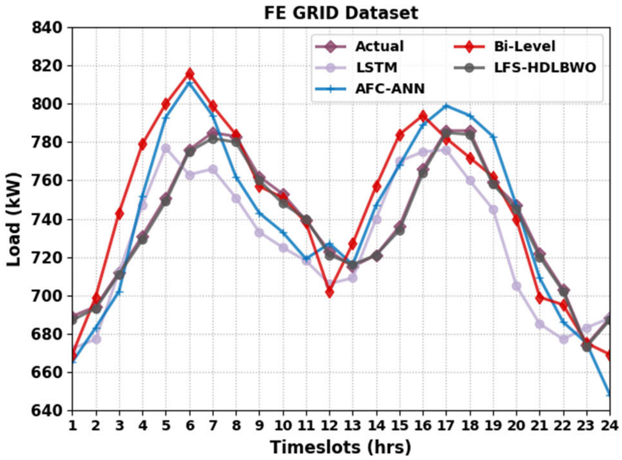

FIGURE 3. Predictive outcome of LFS-HDLBWO technique on FE Grid dataset.

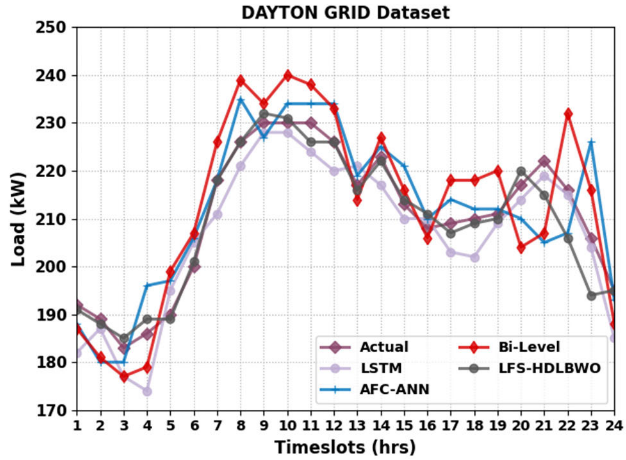

FIGURE 4. Predictive outcome of LFS-HDLBWO technique on DAYTON Grid dataset.

TABLE 2. Predictive outcome of LFS-HDLBWO technique on DAYTON Grid dataset.

DAYTON GRID Dataset					
Load (kW)					
Timeslots (hrs)	Actual	LSTM	AFC-ANN	BiLevel	LFS-HDLBWO
1	192	182	188	187	191
2	189	187	180	181	188
3	183	177	180	177	185
4	186	174	196	179	189
5	190	195	197	199	189
6	200	205	206	207	201
7	218	211	218	226	218
8	226	221	235	239	226
9	230	228	227	234	232
10	230	228	234	240	231
11	230	224	234	238	226
12	226	220	234	233	226
13	217	221	219	214	216
14	223	217	225	227	222
15	213	210	221	216	214
16	208	210	210	206	211
17	209	203	214	218	207
18	210	202	212	218	209
19	211	209	212	220	210
20	217	214	210	204	220
21	222	219	205	207	215
22	216	215	207	232	206
23	206	204	226	216	194
24	195	185	193	188	195

achieved outcome shows that the LFS-HDLBWO system obtains closer predictive values. On 1hr, the LFS-HDLBWO method predicted the load of 191kW with the actual value of 192kW. Besides, on 10hr, the LFS-HDLBWO approach predicted a load of 231kW with an actual value of 230kW. Along with that, on 24hr, the LFS-HDLBWO method predicted the load of 195kW with the actual value of 195kW.

In Table 3 and Fig. 5, a comparison execution time (EXET) result of the LFS-HDLBWO system with other

TABLE 3. EXET outcome of LFS-HDLBWO technique with other methods on two datasets.

Execution Time (min)		
Methods	FE GRID Dataset	DAYTON GRID Dataset
FCRBM	2.72	1.73
AFC-ANN	1.93	1.87
Bilevel	1.72	1.76
MI-ANN	1.28	0.99
LSTM	1.11	1.08
LFS-HDLBWO	0.30	0.19

methodologies can be provided. The obtained outcome indicates that the LFS-HDLBWO system achieves reduced EXET values on two datasets. On the FE grid dataset, the LFS-HDLBWO algorithm offers decreasing EXET of 0.30min whereas the FCRBM, AFC-ANN, Bilevel, MI-ANN, and LSTM techniques provide increasing EXET of 2.72, 1.93, 1.72, 1.28, and 1.1 min respectively. Additionally, on the DAYTON grid dataset, the LFS-HDLBWO technique offers decreasing EXET of 0.19min while the FCRBM, AFC-ANN, Bilevel, MI-ANN, and LSTM methods provide increasing EXET of 1.73, 1.87, 1.76, 0.99, and 1.08min correspondingly.

In Table 4 and Fig. 6, a comparison average error rate (AER) result of the LFS-HDLBWO technique with other models is provided. The outcomes indicate that the LFS-HDLBWO method accomplishes reduced AER values on both datasets. On the FE grid dataset, the LFS-HDLBWO system offers a decreasing AER of 3.43 while the LSTM, AFC-ANN, and Bilevel techniques provide increasing AER of 20.98, 18.71, and 19.33 respectively. Additionally, on the DAYTON grid dataset, the LFS-HDLBWO system offers a decreasing AER of 2.26 whereas the FCRBM, AFC-ANN, Bilevel, MI-ANN, and LSTM methods provide an increasing AER of 5.72, 6.98, and 8.78 correspondingly.

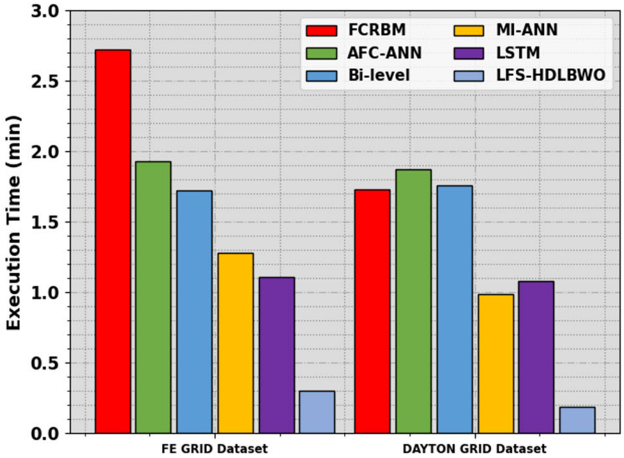

FIGURE 5. EXET outcome of LFS-HDLBWO technique on two datasets.

TABLE 4. AER outcome of LFS-HDLBWO technique with other methods on two datasets.

Avg. Error Rate		
Methods	FE GRID	DAYTON GRID
LSTM	20.98	5.72
AFC-ANN	18.71	6.98
BiLevel	19.33	8.78
LFS-HDLBWO	3.43	2.26

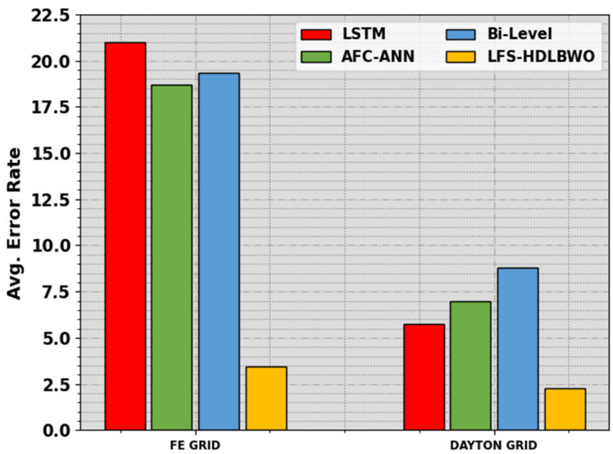

FIGURE 6. EXET outcome of LFS-HDLBWO technique on two datasets.

The MAPE results of the LFS-HDLBWO technique with recent models under varying data sizes are reported in Table 5 and Fig. 7. The attained value shows that the LFS-HDLBWO technique reaches the least MAPE values. On 60 samples, the LFS-HDLBWO technique offers reduced MAPE of 0.342% while the LSTM, AFC-ANN, and Bilevel methods provide increased MAPE of 3.490%, 0.903%, and 3.018% correspondingly. Also, on 420 samples, the LFS-HDLBWO system offers a reduced MAPE of 0.264% but, the LSTM, AFC-ANN, and Bilevel techniques provide maximum MAPE of 2.832%, 0.617%, and 0.589% correspondingly. Moreover, on 720 samples, the LFS-HDLBWO method offers a reduced MAPE of 0.059% however, the LSTM, AFC-ANN, and Bilevel methods provide increased MAPE of 2.485%, 0.458%, and 2.392% correspondingly.

TABLE 5. MAPE outcome of LFS-HDLBWO technique with other methods under varying data size.

MAPE (%)				
Data size (samples)	LSTM	AFC-ANN	BiLevel	LFS-HDLBWO
60	3.490	0.903	3.018	0.342
120	3.479	0.821	2.768	0.273
180	3.391	0.787	2.731	0.258
240	3.192	0.633	2.696	0.251
300	3.152	0.670	2.632	0.228
360	2.948	0.708	2.654	0.230
420	2.832	0.617	2.589	0.264
480	2.810	0.651	2.533	0.266
540	2.734	0.585	2.595	0.216
600	2.709	0.623	2.553	0.254
660	2.796	0.619	2.537	0.242
720	2.485	0.458	2.392	0.059

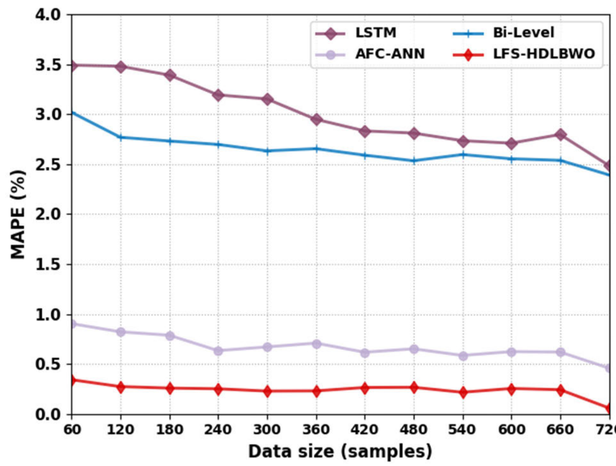

FIGURE 7. MAPE outcome of LFS-HDLBWO technique under varying data size.

The EXET results of the LFS-HDLBWO method with current techniques under varying data sizes are reported in Table 6 and Fig. 8. The experimental outcome shows that the LFS-HDLBWO technique reaches the least EXET values. On 60 samples, the LFS-HDLBWO technique offers reduced EXET of 0.31s whereas the LSTM, AFC-ANN, and Bilevel approaches provide increased EXET of 2.19s, 1.93s, and 3.53s correspondingly. Also, on 420 samples, the LFS-HDLBWO technique offers reduced EXET of 0.33s whereas the LSTM, AFC-ANN, and Bilevel approaches provide increased EXET of 2.34s, 1.97s, and 3.58s respectively. Moreover, on 720 samples, the LFS-HDLBWO technique offers a reduced EXET of 0.49s whereas the LSTM, AFC-ANN, and Bilevel approaches provide maximum EXET of 2.43s, 2.16s, and 3.72s correspondingly.

TABLE 6. EXET outcome of LFS-HDLBWO technique with other methods under different data size.

Execution Time (sec)				
Data size (samples)	LSTM	AFC-ANN	BiLevel	LFS-HDLBWO
60	2.19	1.93	3.53	0.31
120	2.16	1.94	3.53	0.34
180	2.23	1.92	3.58	0.32
240	2.27	1.96	3.58	0.33
300	2.29	1.96	3.60	0.34
360	2.29	1.95	3.58	0.32
420	2.34	1.97	3.65	0.33
480	2.32	2.00	3.63	0.33
540	2.34	2.04	3.73	0.41
600	2.39	2.05	3.72	0.44
660	2.43	2.12	3.72	0.44
720	2.43	2.16	3.72	0.49

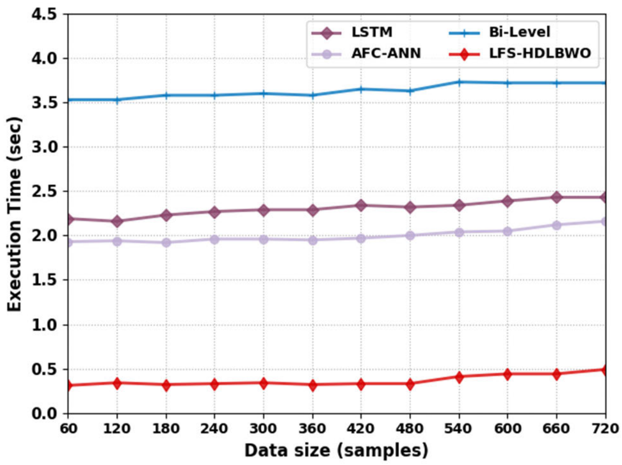

FIGURE 8. EXET outcome of LFS-HDLBWO technique under varying data size.

These outcomes highlight the greater performance of the LFS-HDLBWO technique on the load predictive system.

V. CONCLUSION

In this study, we have designed an automated load prediction using the LFS-HDLBWO approach on the SG environment. The major intention of the LFS-HDLBWO technique is to predict the load in the SG environment. To accomplish this, the LFS-HDLBWO method includes Z-score normalization, CBLSTM-AE model, BWO based hyperparameter tuning approach to scale the input dataset. Finally, the BWO algorithm could be used for optimal hyperparameter selection of the CBLSTM-AE algorithm, which helps to improve the overall prediction results. A wide-ranging experimental analysis was made to illustrate the better predictive outcomes of the LFS-HDLBWO system. The outcome value indicates the outstanding performance of the LFS-HDLBWO model over other existing DL algorithms. The demonstrated potential of this hybrid approach suggests its significance in addressing short-term load forecasting challenges within the smart grid domain, paving the way for improved load management and energy efficiency strategies. Through extensive experimentation, it has demonstrated superior predictive capabilities, outperforming existing deep learning models. Further work can focus on real-time load forecasting, enabling immediate decision-making and dynamic grid management. Leveraging data from the Internet of Things (IoT) devices and sensors can provide additional real-time information for improved load forecasting. Besides, combining LFS-HDLBWO with other forecasting models and techniques, such as ensemble methods, may lead to even more accurate predictions.

Declarations

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/ 65 /44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R387), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2024R838), King Saud University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

REFERENCES

[1] Z. Tong, Y. Zhou, and K. Xu, "An intelligent scheduling control method for smart grid based on deep learning," *Math. Biosci. Eng.*, vol. 20, no. 5, pp. 7679–7695, 2023.

[2] A.-X. Wang and J.-J. Li, "A novel cloud-edge collaboration based short-term load forecasting method for smart grid," *Frontiers Energy Res.*, vol. 10, Aug. 2022, Art. no. 977026.

[3] A. K. Shaikh, A. Nazir, N. Khalique, A. S. Shah, and N. Adhikari, "A new approach to seasonal energy consumption forecasting using temporal convolutional networks," *Results Eng.*, vol. 19, Sep. 2023, Art. no. 101296.

[4] P. K. S., P. Amitasree, G. R. Vamshi, and V. S. K. Devi, "Deep learning based load forecasting for futuristic sustainable smart grid," in *Proc. Int. Conf. Electron. Renew. Syst. (ICEARS)*, Mar. 2022, pp. 35–40.

[5] H. Habbak, M. Mahmoud, K. Metwally, M. M. Fouda, and M. I. Ibrahem, "Load forecasting techniques and their applications in smart grids," *Energies*, vol. 16, no. 3, p. 1480, Feb. 2023.

[6] V. T. Tran, M. H. Le, M. T. Vo, Q. T. Le, V. H. Hoang, N.-T. Tran, V.-T. Nguyen, T.-A.-T. Nguyen, H. N. Nguyen, V. T. T. Nguyen, and T. T. Nguyen, "Optimization design for die-sinking EDM process parameters employing effective intelligent method," *Cogent Eng.*, vol. 10, no. 2, Dec. 2023, Art. no. 2264060.

[7] C.-N. Wang, F.-C. Yang, V. T. T. Nguyen, and N. T. M. Vo, "CFD analysis and optimum design for a centrifugal pump using an effectively artificial intelligent algorithm," *Micromachines*, vol. 13, no. 8, p. 1208, Jul. 2022.

[8] A. Ahmad, N. Javaid, A. Mateen, M. Awais, and Z. A. Khan, "Short-term load forecasting in smart grids: An intelligent modular approach," *Energies*, vol. 12, no. 1, p. 164, Jan. 2019.

[9] S. Khan, I. Hayder, M. Habib, M. Ahmad, S. Mohsin, F. Khan, and K. Mustafa, "Enhanced machine-learning techniques for medium-term and short-term electric-load forecasting in smart grids," *Energies*, vol. 16, no. 1, p. 276, Dec. 2022.

[10] B. Ibrahim, L. Rabelo, E. Gutierrez-Franco, and N. Clavijo-Buritica, "Machine learning for short-term load forecasting in smart grids," *Energies*, vol. 15, no. 21, p. 8079, Oct. 2022.

[11] 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.
[12] S. Rai and M. De, "Analysis of classical and machine learning based short-term and mid-term load forecasting for smart grid," *Int. J. Sustain. Energy*, vol. 40, no. 9, pp. 821–839, Oct. 2021.
[13] Z. Tavassoli-Hojati, S. F. Ghaderi, H. Iranmanesh, P. Hilber, and E. Shayesteh, "A self-partitioning local neuro fuzzy model for short-term load forecasting in smart grids," *Energy*, vol. 199, May 2020, Art. no. 117514.
[14] C. Li, "Designing a short-term load forecasting model in the urban smart grid system," *Appl. Energy*, vol. 266, May 2020, Art. no. 114850.
[15] S. Salehimehr, B. Taheri, and M. Sedighizadeh, "Short-term load forecasting in smart grids using artificial intelligence methods: A survey," *J. Eng.*, vol. 2022, no. 12, pp. 1133–1142, Dec. 2022.
[16] D. Wu, Y. T. Xu, M. Jenkin, J. Wang, H. Li, X. Liu, and G. Dudek, "Short-term load forecasting with deep boosting transfer regression," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 5530–5536.
[17] J. M. Aguiar-Pérez and M. Á. Pérez-Juárez, "An insight of deep learning based demand forecasting in smart grids," *Sensors*, vol. 23, no. 3, p. 1467, Jan. 2023.
[18] N. Jha, D. Prashar, M. Rashid, S. K. Gupta, and R. K. Saket, "Electricity load forecasting and feature extraction in smart grid using neural networks," *Comput. Electr. Eng.*, vol. 96, Dec. 2021, Art. no. 107479.
[19] A. Yasin, V. P. Singh, M. Ahsan, M. J. Awan, and R. Mubashar, "Efficient residential load forecasting using deep learning approach," *Int. J. Comput. Appl. Technol.*, vol. 68, no. 3, p. 205, 2022.
[20] J.-W. Xiao, P. Liu, H. Fang, X.-K. Liu, and Y.-W. Wang, "Short-term residential load forecasting with baseline-refinement profiles and bi-attention mechanism," *IEEE Trans. Smart Grid*, 2023.
[21] A. Motwakel, E. Alabdulkreem, A. Gaddah, R. Marzouk, N. M. Salem, A. S. Zamani, A. A. Abdelmageed, and M. I. Eldesouki, "Wild horse optimization with deep learning-driven short-term load forecasting scheme for smart grids," *Sustainability*, vol. 15, no. 2, p. 1524, Jan. 2023.
[22] A. Alrasheedi and A. Almalaq, "Hybrid deep learning applied on Saudi smart grids for short-term load forecasting," *Mathematics*, vol. 10, no. 15, p. 2666, Jul. 2022.
[23] D. Syed, H. Abu-Rub, A. Ghrayeb, S. S. Refaat, M. Houchati, O. Bouhali, and S. Bañales, "Deep learning-based short-term load forecasting approach in smart grid with clustering and consumption pattern recognition," *IEEE Access*, vol. 9, pp. 54992–55008, 2021.
[24] X. Chen, W. Chen, V. Dinavahi, Y. Liu, and J. Feng, "Short-term load forecasting and associated weather variables prediction using ResNet-LSTM based deep learning," *IEEE Access*, vol. 11, pp. 5393–5405, 2023.
[25] M. ZulfiqAr, M. Kamran, M. B. Rasheed, T. Alquthami, and A. H. Milyani, "A short-term load forecasting model based on self-adaptive momentum factor and wavelet neural network in smart grid," *IEEE Access*, vol. 10, pp. 77587–77602, 2022.
[26] A. Zarei and N. Ghaffarzadeh, "Optimal demand response-based AC OPF over smart grid platform considering solar and wind power plants and ESSs with short-term load forecasts using LSTM," *J. Sol. Energy Res.*, vol. 8, no. 2, pp. 1367–1379, 2023.
[27] W. Chen, G. Han, H. Zhu, L. Liao, and W. Zhao, "Deep ResNet-based ensemble model for short-term load forecasting in protection system of smart grid," *Sustainability*, vol. 14, no. 24, 2022, Art. no. 16894.
[28] A. Ghasempour and M. Martínez-Ramón, "Short-term electric load prediction in smart grid using multi-output Gaussian processes regression," in *Proc. IEEE Kansas Power Energy Conf.*, Apr. 2023, pp. 1–6.
[29] M. Ali, M. Adnan, and M. Tariq, "Optimum control strategies for short term load forecasting in smart grids," *Int. J. Elect. Power Energy Syst.*, vol. 113, pp. 792–806, 2019.
[30] Y. Hong, Y. Zhou, Q. Li, W. Xu, and X. Zheng, "A deep learning method for short-term residential load forecasting in smart grid," *IEEE Access*, vol. 8, pp. 55785–55797, 2020.
[31] G. Yan, T. Han, W. Zhang, and S. Zhao, "Short-term load forecasting of smart grid based on load spatial–temporal distribution," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT Asia)*, May 2019, pp. 781–785.
[32] O. Jogunola, B. Adebisi, K. V. Hoang, Y. Tsado, S. I. Popoola, M. Hammoudeh, and R. Nawaz, "CBLSTM-AE: A hybrid deep learning framework for predicting energy consumption," *Energies*, vol. 15, no. 3, p. 810, Jan. 2022.
[33] Z. Chen and D. Song, "Modeling landslide susceptibility based on convolutional neural network coupling with metaheuristic optimization algorithms," *Int. J. Digit. Earth*, vol. 16, no. 1, pp. 3384–3416, Oct. 2023.