Volume 18, Issue 2, 182432 (1-13)

Majlesi Journal of Electrical Engineering (MJEE)

https://dx.doi.org/10.57647/j.mjee.2024.1802.32

Hybrid optimization based on deep learning approach for short-term load forecast of electricity demand in buildings

Charan Sekhar*, Ratna Dahiya

Department of Electrical Engineering, NIT Kurukshetra, Kurukshetra, India.

*Corresponding author: charan_61900135@nitkkr.ac.in

Original Research

Received:
25 December 2023
Revised:
3 February 2024
Accepted:
10 February 2024
Published online:
25 June 2024

Abstract:

Due to the growing popularity of microgrids in buildings, the foreseeable electricity demand for a building draws the attention of many researchers. The precise short-term demand forecast efficiently directs building managers and operators for interactions with electrical distribution systems, daily operational decisions, and energy conservation. This research proposes a hybrid optimization-based deep learning (DL) approach to increase the accuracy of short-term forecasts. The present work employs the Bilateral Long Short-Term Memory (BiLSTM) network-based DL technique because the BiLSTM technique has an exceptional ability to manage nonlinear interactions in data and learn the temporal dependencies. The performance of the BiLSTM technique is improved using the optimally determined hyperparameters via a hybrid optimization algorithm that combines particle swarm optimization (PSO) and grey wolf optimization (GWO). The exploration ability of GWO and exploitation ability of PSO are effectively combined in the hybrid optimization GWO-PSO. The performance of the recommended approach is assessed using a case study of an educational building. The performance of the proposed model is compared to existing nonoptimal BiLSTM and single optimization-based BiLSTM for short-term load forecast.

Keywords: Short-term load forecast; Buildings; Bilateral long-short-term memory; GWOPSO; Hyperparameter optimization

1. Introduction

In recent years, the concern for energy conservation has been growing. Buildings worldwide account for one-third of total energy consumption and 55% of total electricity consumption [1, 2]. Buildings are also account for 27% of global operational CO_2 emissions [3]. Accurate electricity demand prediction of buildings helps the operators to take mandatory energy management actions, leading to considerable cost and energy savings. The uncertainties in demand at the building level make demand prediction much more challenging than on a traditional grid demand forecast. Factors like the purpose of the building, weather conditions, people's lifestyle, and technological advancements affect the consumption pattern in buildings, leading to continuous fluctuations in demand [4]. Thus, recent research is becoming more attentive to building demand forecasts. Building categories include residential, educational, office, hospital, hotel, and industrial, and their consumption patterns vary depending on the building's intended use [5]. The demand prediction for educational buildings is the subject of the current effort.

Load forecasting is divided into very short-term load forecast (VSTLF), short-term load forecast (STLF), medium-term load forecast (MTLF), and long-term load forecast (LTLF) based on forecast time horizon. The forecasting period for VSTLF is from a few minutes to a few hours, and for STLF is from one day to one week. An MTLF considers the period from weeks to a few months, while the LTLF period is up to 10 years from a year [6]. STLF is fruitful for demand management, market decisions, unit commitment, and economic scheduling [7]. The researchers developed different forecast technologies for STLF, which are classified into statistical and artificial intelligence. Statistical methods, like auto regression integrated moving average (ARIMA) and exponential smoothing, which are fast, simple, and straightforward, have become apparent and are based on the linear relations between variables [8]. As power consumption patterns have become more volatile, statistical approaches' ability to predict outcomes has declined. Artificial intelligence (AI) techniques are becoming increasingly popular because of their ability to understand nonlinear data, robust computation capabilities, and methods for processing data in forecasting applications.

1.1 Literature review

Over the past thirty years, many Machine Learning (ML) and Deep Learning (DL) techniques have been developed for load forecasting. Various machine learning (ML) models like multiple regression and genetic programming [20], Multiple Linear Regression (MLR), Multilayer Perceptron (MLP) and Support Vector Regression (SVR) [21], artificial neural networks (ANN) [22, 23] and adaptive neuro-fuzzy inference system (ANFIS) [24] have been used to estimate short term electricity demand in various types of buildings. Seven ML techniques have been compared to calculate the next hour's residential energy consumption [25]. A two-step hybrid ML approach using extreme gradient boosting (XGBoost), random forest (RF), CATBoost, Light gradient-boosting machine (LGBM), MLP, and Long Short-Term Memory (LSTM) has been proposed to forecast net zero building load [26]. ML approaches have the significant drawback of dealing with each instant independently, disregarding the relationships between the time instants. The deep learning (DL) techniques have outstanding capabilities to handle nonlinear relations of data and better account for the connection among time instances, leading to a surge in their usage in forecast applications. The work in [27] combined Recurrent Neural Networks (RNN) with wavelets (WT) to develop self-recurrent wavelet-based neural networks for forecasting the energy demand of educational buildings.

Cai et al. proposed two DL models, gated RNN and gated convolution neural networks (CNN). They compared their performance with the Seasonal Auto-Regressive Integrated Moving Average with Exogenous Variables (SARIMAX) model for the case of an academic building, a school building, and a grocery store time horizon of 1-h and 24-h ahead forecast. In contrast, the main disadvantage of the work is the neglected forecast during weekends [28]. After experimenting with various combinations of deep learning models on household electricity consumption, Sajjad et al. proposed a hybrid technique based on the CNN and the gated recurrent unit (GRU) approach [29]. Koukaras et al. proposed a hybrid strategy by the ensemble with weightage to each method to predict an intelligent home's next step energy load forecast [30]. Using appliance data, Kong et al. addressed the residential load forecast problem with Long Short-Term Memory (LSTM) [31]. They also showed that the addition of Feed-Forward Neural Networks (FFNN) to LSTM has improved prediction accuracy by continuing their work for case studies of Smart Grid Smart City (SGSC) data [32]. The work in [33] developed two different structures using LSTM and MLP neural networks to estimate the electricity load in medium- and long-term time horizons with manual neural network hyperparameters optimization. The self-attention-based LSTM network is developed to forecast Baghdad city electricity demand [34].

The aforementioned methods and approaches for predicting electrical consumption in buildings and home loads had chosen hyperparameters manually. It would help if the user or developer were an expert in the subject to improve forecast accuracy using the manual set of hyperparameters. Some researchers have developed automatic parameter tweaking using optimization methods to solve this issue. Different variations of Particle Swarm Optimization (PSO) have been applied to determine the optimal parameters of ANN [9], Back Propagation Neural Networks (BPNN) [10], LSTM [11], and CNN and LSTM [12] techniques to estimate the electricity demand and prices in the grid and buildings and electricity markets. A modified JAYA optimization algorithm has been proposed to optimize the XGBoost parameters, which has improved the prediction performance for cooling and heating loads in a building [13]. J. Wang et al. developed employed Adaptive Particle Swarm Optimization (APSO) to determine the model weights of a hybrid model of Seasonal Auto-Regressive Integrated Moving Average (SARIMA), Exponential smoothing, and weighted support vector machine (SVM) to predict the weekdays' electrical demand of New South Wales [35].

The Grey Wolf Optimization (GWO) was used to determine the optimal ANN structure to predict long-term demand in Iran [14]. The work [15] used grasshopper optimization (GHO) to determine the parameters of the ANN structure to estimate the intelligent grid from Australia. Tayab et al. addressed the short-term load forecast problem using Harris Hawks optimization (HHO) and FFNN and compared its performance with PSO-based FFNN [16]. Panda et al. compared the Bat algorithm (BA) over the Cuckoo search (CSO) to determine the ANN network weights for obtaining a better network to forecast the load on the power plant [17]. A hybrid DL technique based on CNN-BiGRU supported by XGBoost is developed to predict the residential building loads in Mashhad city. The parameters of CNN and GRU networks are optimized using the PSO technique [18]. An optimized ML model is developed for a residential building using different ML techniques using a biogeography optimization algorithm (BBO) [19].

A comparison of related works is presented in Table. 1 The research gaps observed from the study are as follows:

• The DL approaches are more effective at extracting temporal connections. Their use in prediction is currently on the rise. Yet the utilization of bidirectional LSTM (BiLSTM) is not up to mark in use for forecast, which overcomes the overfitting problem of the LSTM technique.

• The researchers have ascertained the optimum values of the forecast network parameters by employing various single metaheuristic optimization methods, including PSO, JAYA, HHO, BA, BBO, and GWO. The single optimization methods may be subject to local optimum. PSO, HHO, and BA methods suffer from premature convergence, requiring more iterations, increasing population, and increasing com-

Table 1. Summary of the related works.

Ref (Year)	Method	Optimal parameter	Load or demand	No. of iterations, population	Findings and gaps
[9](2015)	ANN-PSO-PCA	Weights and threshold values	Energy Prediction Shootout Contest I and a campus building in East China	10, 20	No DL technique, Single optimization method, considered MTLF, may not be suitable for STLF.
[10](2023)	FFNN-PSO	Weights	Ghana grid	-,-	No DL technique, Single optimization method, MTLF, no comparison with other optimization techniques.
[11](2021)	LSTM-PSO	Weights of the LSTM units	Electricity price	-	Single optimization method, no comparison with other optimization techniques.
[12](2019)	CNN-LSTM-PSO-Genetic algorithm	Learn rate, no. of filters, filter size, drop factor, layer size, learn rate, cell type, pool size	Residential	-, 270	Single optimization method, no comparison with other optimization models, higher population.
[13](2022)	XGBoost-JAYA	N-rounds, max. depth, eta, γ, colsample bytree, min. Child weight, sub-sample.	Residential	30, 50	No DL technique, Single optimization method,
[14](2023)	ANN-GWO	ANN parameters	Iraq grid demand	-, -	No DL technique, Single optimization method, Long Term forecast
[15](2020)	ANN-GHO	ANN parameters	Smart grids of Victoria and New South Wales	100, 200	There is no DL technique, Single optimization method, or comparison with other optimization techniques.
[16](2020)	FFNN-HHO	Weights and biases of ANN	Queensland Electrical grid	-, -	No DL technique, Single optimization method, Compared with PSO
[17](2023)	ANN-CSO -BA	Weights of neurons	Power plant load in MW	-	CSO, BA has its specific parameters, No DL technique, and a Single optimization method
[18](2024)	CNN-BiGRU -PSO	No. of BiGRU layers, hidden neurons, learn rate, train epochs, CNN filters, activation function, dropout.	Residential loads of a city, data of 5 years	-, 20	Single optimization technique, grid-level data, more population of optimization technique.
[19](2024)	ANN, MLP, RBF, -BBO	Parameters of ML techniques	Residential load	1000, -	Single optimization, No DL technique.
Proposed	BILSTM-GWOPSO	No. of layers, hidden units in first, second, and third layers, max. Epochs, minibatch size, learn rate, drop rate.	College Building– Commercial load	10, 10	Applied Hybrid Optimization, executed with fewer iterations, made a comparison with other optimization methods.

putation cost. Even though the GWO algorithm is simple and requires no parameter setting, it suffers from slow convergence.
• Despite the increased concern about building electricity forecasts in recent years, most research has concentrated on residential building consumption forecasts, with very little research on college electricity usage.
Major contributions of the present research are to demonstrate the performance of hybrid optimization GWO-PSO based DL technique as a new model as follows:
• A college electricity consumption is considered as a case study with more non-uniformity than residential and indus-

trial consumption profiles to address this problem.

• The DL-based BiLSTM networks have been used rarely, yet the technique has better learning capability than other DL and ML techniques. The BiLSTM technique learns the data in the forward and backward directions, enhancing prediction performance.

• The hybrid optimization technique based on GWO and PSO has been used to determine the optimal hyperparameters of the BiLSTM network. The hybridization of GWO and PSO effectively balances exploitation and exploration. The algorithm is simple to implement, like PSO and GWO, and overcomes the limitations of local optima problem.

• A comparison is made of the performance of the proposed hybrid optimization technique against state-of-the-art techniques such as PSO, GWO, and JAYA optimization.

2. Materials and methods

The two techniques of the proposed methodology are discussed in the present section. They are the BiLSTM technique, which works on the DL concept for forecast, and the hybrid optimization technique based on GWO-PSO to optimize the hyperparameters.

2.1 Bilateral long short-term memory

Although McCulloch and Pitts proposed the concept of deep learning in 1943 [36], the idea roved over decades due to a lack of adequate training algorithms and computing resources. A better understanding of the techniques and substantial growth in the electronic industry has improved computing capabilities. A particular version of RNN, known as the LSTM approach, was proposed by Hochreiter and Schmidhuber. It uses a forget gate and memory cell to solve the vanishing gradient problem [37, 38]. LSTM preserves the internal memory cell state throughout the life cycle to set up temporal connections. The LSTM network comprises four gates: input gate, forget gate, cell candidate, and output gate. The input sequence to the LSTM unit is x_1, x_2,.... x_{t-1}, x_t, x_{t+1},.... x_n. The input weights are W_{fh}, W_{ih}, W_{gh}, W_{oh}, W_{fx}, W_{ix}, W_{gx}, W_{ox} and the bias are b_f, b_i, b_g, b_o.
The sigmoid function is

$$\sum(x) = \frac{1}{1+e^{-x}} \tag{1}$$

The hyperbolic tangent function is

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1} \tag{2}$$

The forget gate is

$$F_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \tag{3}$$

The input gate is

$$I_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \tag{4}$$

$$G_t = \tanh\left(W_{gh}h_{t-1} + W_{gx}x_t + b_g\right) \tag{5}$$

The output gate is

$$O_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \tag{6}$$

$$C_t = C_{t-1} \odot F_t + I_t \odot G_t \tag{7}$$

$$h_t = \tanh(C_t) \odot O_t \tag{8}$$

The LSTM commonly starts with zero initialization, i.e., $h_0 = 0$, $C_0 = 0$. It consists of a sigmoid and tanh function to process the data. The sigmoid function that transforms the input values into a range of 0 to 1, acts as a soft switch and blocks the signal if the gate value is 0.
Otherwise, it will pass. The previous cell state C_{t-1} interacts with the transitional output h_{t-1} and current input x_t. The interaction controls which components should be improved, removed, and kept. The input gate signal controls internal state preservation. Forget gate contains the data that is to be omitted from the previous state C_{t-1} and the output gate decides which cell state C_t should pass to LSTM output.
The proposed research uses BiLSTM, an improved network of LSTM where the information flows in both directions to learn the long-term dependencies in both forward and backward directions, i.e., it learns from past and future values [42]. Yet, both layers take input states the same. The BiLSTM has a double memory cell due to two LSTM cells. The functional diagram of BiLSTM is shown in Fig. 1. The information of the forward LSTM and backward LSTM units is stored as hidden states $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ respectively, at a time, 't'. The final hidden state will be computed by concatenating two hidden states.

$$h_t = \partial(\overrightarrow{h_t}, \overleftarrow{h_t}) \tag{9}$$

Concatenation can be an average function, addition, or multiplication.

2.2 GWO and PSO-based hybrid optimization

GWO serves as the initial phase for the hybrid optimization to get better positions, and the next phase is followed by the PSO algorithm, which serves the function of position updating. A GWO algorithm based on the hunting behaviors of grey wolves covers the gang hunts and how leadership is exercised. It includes wolves tracking down, encircling, and attacking their prey to catch them successfully. In this process, grey wolves use leadership skills at different levels, such as alpha, beta, delta, and omega [43]. The hierarchy of the wolf's leadership is shown in Fig. 2. The initial step of hunting is encircling the prey and mathematically modeling it as at the current time step is

$$D = |C \times Z_p(t) - Z(t)| \tag{10}$$

The positions are updated using the

$$Z(t+1) = Z_p(t) - A \times D \tag{11}$$

The determination of A and C vectors using

$$A = (2 \times a) \bullet r_1 - a; C = 2 \bullet r_2 \tag{12}$$

A & C are coefficient vectors, and the components of vector 'a' are decreased linearly from 2 to 0 to attain optima over the iterations. The Z_P and Z indicates the position vector of the prey and grey wolves. For the mathematical model, the alpha, beta, and delta positions are in better places, and the

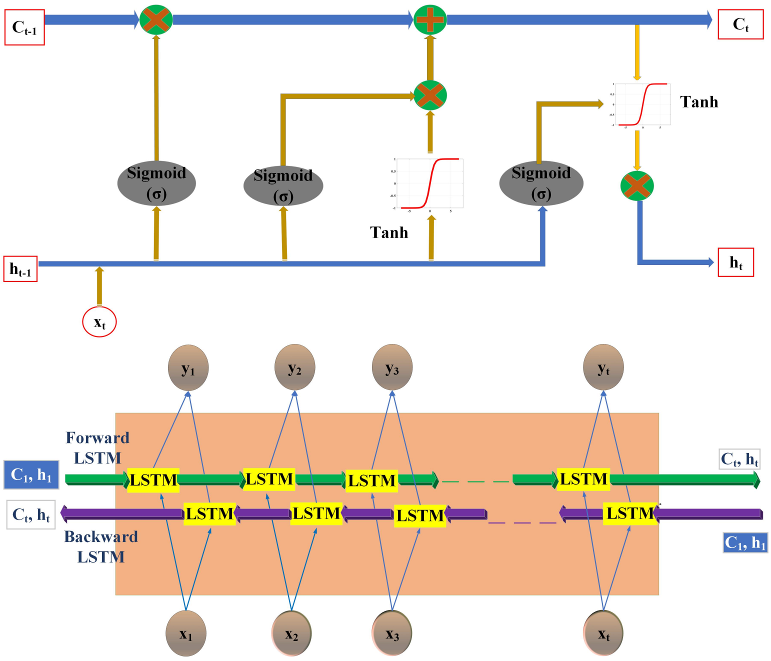

Figure 1. (a) LSTM and (b) BiLSTM architectures [32, 39, 40].

remaining update their positions within the circle, yet the updating concerns the alpha, beta, and delta locations. The leader alpha guides the hunting, and the alpha's location is considered the final best solution.

$$\begin{aligned} D_\alpha &= |C_1 \bullet Z_\alpha - Z|; \\ D_\beta &= |C_2 \bullet Z_\beta - Z|; \\ D_\delta &= |C_3 \bullet Z_\delta - Z| \end{aligned} \tag{13}$$

$$Z_1 = Z_\alpha - A_1 D_\alpha;\ Z_2 = Z_\beta - A_2 D_\beta;\ Z_3 = Z_\delta - A_3 D_\delta \tag{14}$$

The final stage in the hunting of grey wolves is attacking the prey. This stage is mathematically modeled using the variable 'a' by decreasing the value to zero. After determining the positions Z_1, Z_2 and Z_3 in response to alpha, beta, and delta positions, the updated positions are calculated using the PSO method rather than the usual GWO technique, which uses the average of three positions [44].

Kennedy and Eberhart devised the PSO method to optimize linear and nonlinear functions. PSO is based on how birds move, eat, and rest and the underlying concept of information sharing amongst the birds to reach the target [45]. Each position is represented with a particle in the algorithm and considered to move with specific velocities. The associated velocities in N-dimensional space are defined as $v_i = (v_{i1}, v_{i2}, \ldots\ldots, v_{in})$ where (i=1, 2, 3......p) of 'p' number of particles [35]. Velocity update equation:

$$v_{new} = w\Big(v_{old} + C_1 r_1 (Z_1 - x_{ij}) + C_2 r_2 (Z_2 - x_{ij}) + C_3 r_3 (Z_3 - x_{ij})\Big) \tag{15}$$

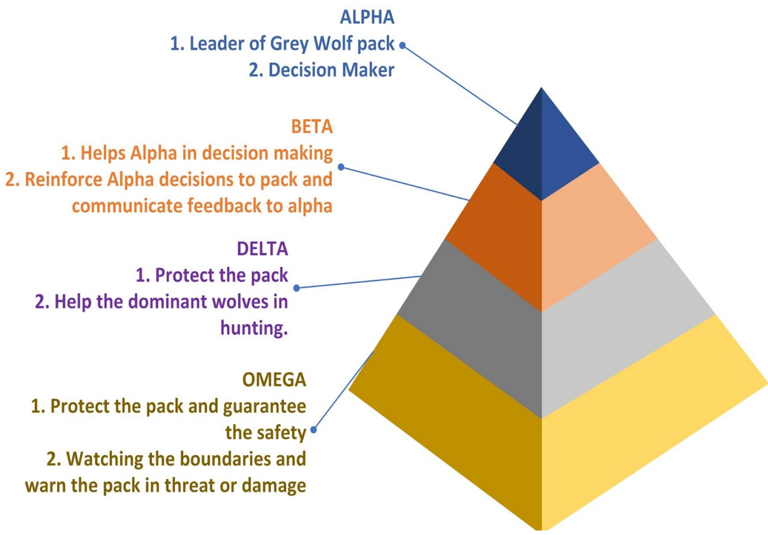

Figure 2. Grey wolves hierarchy[41].

The use of local best P_{best} and global best g_{best} in the velocity update equation is replaced with alpha, beta, and delta in the hybrid optimization. Where r_1, $_2$ and r_3 denotes the random numbers between 0 and 1. The velocities update of the traditional PSO algorithm is improved using weighted velocities to obtain new velocities of particles.

$$w = w_{\max} - (w_{\max} - w_{\min}) \times \frac{it}{it_{\max}} \tag{16}$$

Where W, W_{min} and W_{max} represent inertia weight, minimum, and maximum inertia weight coefficients, respectively. it and it_{max} indicate the current iteration and maximum iteration number, respectively. The new position has been calculated using the following:

$$x_{ij,new} = x_{ij,old} + v_{new} \tag{17}$$

3. Methodology for short-term load forecast

This section describes the case study and the proposed methodology for predicting load. The method includes data processing and follows the optimized BiLSTM structure.

3.1 Case study description

The work considers hourly college building consumption data over a year to examine the performance of the proposed method. The data belongs to one of the six reference buildings of EnergyPlus, which are publicly available. The average campus consumption is 12958.6626 kW, and the standard deviation is 2453.6136 kW [46].

3.2 Methodology

The research proposes a new methodology for an improved BiLSTM network, which is divided into three parts. The first part is the data processing phase, followed by optimizing the BiLSTM network during training, which will be used in the next stage, and executing the optimized BiLSTM network for the final forecast in the third part. The functional flowchart of the proposed methodology indicates the execution of a hybrid optimization GWO-PSO-based BiLSTM network in Fig. 3. The data is processed for better accuracy after importation. The research employs a data standardization mechanism that converts the unprocessed electrical data to Z-score scaling with a zero mean and one standard deviation. This modification reduces the computational time required to handle data quickly during training and lessens the impact of outliers. The formula for data standardization is :

$$x_{i,\text{standard}} = \frac{x_i - \bar{x}}{\sigma_x} \tag{18}$$

where x_i,$\bar{x}$ and σ_x represent the data sequence, the mean, and the standard deviation of data, respectively.
The second phase is the execution of the hybrid optimization algorithm. The objective function of the optimization technique used in the present investigation, mean square error (MSE), was chosen to find the BiLSTM network's optimal parameters. Table. 2. indicates the considered upper and lower boundaries of BiLSTM network parameters for optimization. After the initialization of the developed hybrid optimization algorithm, the fitness values of the parameters are estimated using the BiLSTM network with the training data. A simple BiLSTM network has a sequence of steps through the electricity consumption data flows to predict the demand. It contains four layers; the first is the input sequence layer to create the given input size. The

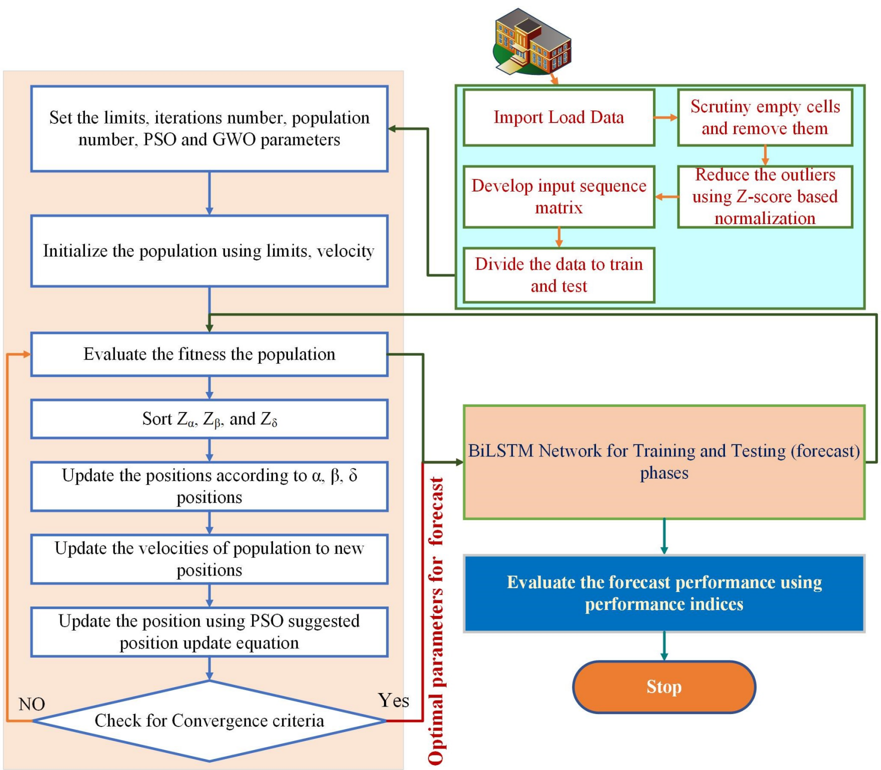

Figure 3. Generalized flow chart for the optimization of BILSTM network parameters.

Table 2. Investigation span of hyperparameter for BiLSTM network optimization.

Parameters	Lower Limit	Upper Limit
Number of hidden Units in the first layer	175	225
Number of hidden units in the second layer	75	125
Number of hidden units in the third layer	25	75
Number of layers	1	3
Number of Epochs	100	200
Initial learn rate	10^{-5}	10^{-3}
Minibatch size	16	256
Learn rate drop factor	0.1	0.4

second and third layers are the BiLSTM layer, and the fully connected layer predicts the demand for assigned hidden units. The regression layer calculates the half mean square error for the regression job [37].
After executing the developed GWO-PSO hybrid optimization, the optimal parameters are given to the BiLSTM network for final training and forecast using test data. The forecasted data is compared with actual data to evaluate the performance in terms of performance metrics. The number of iterations and population size remains equal for all optimization strategies to reasonably contrast the performance of the suggested GWO-PSO-based hybrid optimization BiLSTM techniques with single optimization-based models like PSO-BiLSTM, JAYA-BiLSTM, and GWO-BiLSTM. The number of iterations is 10, and the population size is 10. The JAYA and GWO algorithms do not have algorithm-specific parameters like PSO.
The main influencing parameters of RNN techniques are the numbers of layers and hidden units in the layers. The remaining parameters are the number of epochs, minibatch size, learning rate, and drop rate, and the corresponding values for the simulation are 300, 64, 0.005, and 0.2, respectively, for the traditional BiLSTM network.

3.3 Evaluation metrics

Five performance evaluation measures are used to show the suggested technique performance. Such as Root Mean Square Error (RMSE), Normalized Root Mean Square Error (NRMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and MSE. The aforementioned performance metrics' formulas are [39]

$$MAE = \frac{\Sigma^{n}{}_{i=1}|\,Y_{real} - Y_{pred}|}{n} \tag{19}$$

$$MSE = \frac{\Sigma^{n}{}_{i=1}(Y_{real} - Y_{pred})^2}{n} \tag{20}$$

$$MAPE = \frac{1}{n}\Sigma^{n}{}_{i=1}|\,\frac{Y_{real} - Y_{pred}}{Y_{real}}| \tag{21}$$

$$RMSE = \sqrt{\frac{\Sigma^{n}{}_{i=1}(Y_{real} - Y_{pred})^2}{n}} \tag{22}$$

$$NRMSE = \frac{\frac{\Sigma^{n}{}_{i=1}(Y_{real} - Y_{pred})^2}{n}}{\frac{\Sigma^{n}{}_{i=1}Y_{real}}{n}} \tag{23}$$

Where 'n' indicates forecasted time horizon lead steps.

4. Results and discussions

The observations and analysis of the proposed model are done in the form of evaluation metrics and errors per sample. The proposed hybrid optimization model's results are compared with developed models such as the BiLSTM technique without optimization and single optimization-based BiLSTM techniques such as PSO-BiLSTM, JAYA-BiLSTM, and GWO-BiLSTM. The comparison is done for

Table 3. One-day performance metrics.

Technique/ metrics	BiLSTM	PSO BiLSTM	JAYA BiLSTM	GWO BiLSTM	GWOPSO BiLSTM
MAE (kW)	6.24959	5.94820	5.43241	3.87060	2.89009
MAPE (%)	0.06091	0.05833	0.05308	0.03896	0.03017
MSE	49.7936	41.7683	36.2988	18.8344	13.0243
RMSE (kW)	7.05646	6.46284	6.02484	4.33986	3.60892
NRMSE	7.08E-04	6.48E-04	6.04E-04	4.35E-04	3.62E-04

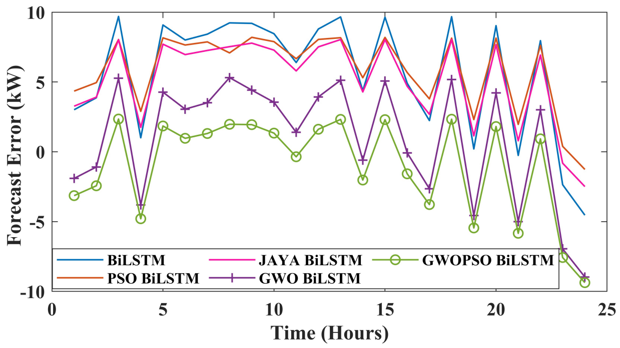

Figure 4. One-day forecast error magnitude comparison.

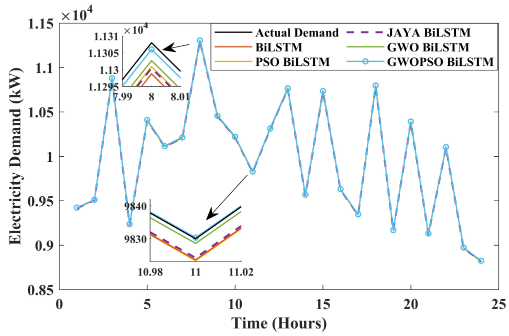

Figure 6. One-day forecast comparison.

one-day and one-week forecasts. The MATLAB environment tool is used on an Intel i5 core processor and 16GB RAM system.

4.1 One-day forecast

Simulation results for one-day forecasts are shown in Fig. 4, which compares all considered techniques in terms of the error magnitude concerning time. The error values are within 10 kW. The suggested method's error magnitude is rarely greater than 5kW but is frequently close to zero. Further, the quantification of error in terms of performance metrics is demonstrated in Table. 3. It can be clearly observed from Table. 3 that the performance of the BiLSTM technique had been improved by optimally calculated hyperparameters using metaheuristic algorithms. The parameter values of the BiLSTM, which several researchers have considered, were randomly chosen. Results of the proposed optimization techniques are compared with existing parameter values in the literature, and it is observed that all the optimized models reduce the error. The best results were obtained by GWO-PSO-BiLSTM, which is given in Table. 3. This improvement is due to the addition of exploitation capability of PSO with GWO. The PSO-BiLSTM, JAYA-BiLSTM, and GWO-BiLSTM outperformed traditional BiLSTM by 4.82%, 13.08%, and 38.07% in MAE, respectively. Compared to PSO-BiLSTM and JAYA-BiLSTM, the prediction ability of GWO-BiLSTM is better due to its exploration capabilities. The performance of various forecast techniques in terms of MAPE and RMSE is also shown by the bar graph in Fig. 5, which also indicates that the error is drastically reduced in the proposed GWO-PSO-BiLSTM technique. It demonstrates that its MAE is almost half that of a non-optimized BiLSTM. In MAE, MAPE, and RMSE, respectively, the recommended GWO-PSO-BiLSTM strategy outperforms PSO-BiLSTM by 51.41%, 48.28%, and 44.16%. In terms of MAE, MAPE, and RMSE, it also leads GWO-BiLSTM by 25.33%, 22.56%, and 16.84%, respectively.

A comparison of forecast demand against actual demand has also been made and is given in Fig. 6. The curves show that all considered methods closely follow actual consumption. Due to the changing hours in the morning and evening between classrooms, offices, and hostels and higher gym usage, security, and cleaning services, the fluctuations particularly tended to be more significant. BiLSTM performance is good during load fall, but prediction errors are more than other techniques during load rise at hours 3, 8, 13, 15, and 18. The standard BiLSTM, PSO-BiLSTM, and JAYA-BiLSTM errors are positive every instant, indicating that the anticipated values are lower than the actual demand. While the errors from the GWO-BiLSTM and proposed GWO-PSO-BiLSTM are positive and negative, they are trying to trace the rise and fall patterns more precisely.

4.2 One week forecast

A comparison of error magnitudes at each time instant is shown in Fig. 7, and it reveals that the maximum error magnitudes of all models are higher than the maximum error of the day forecast. The error magnitudes are less than 10 kW,

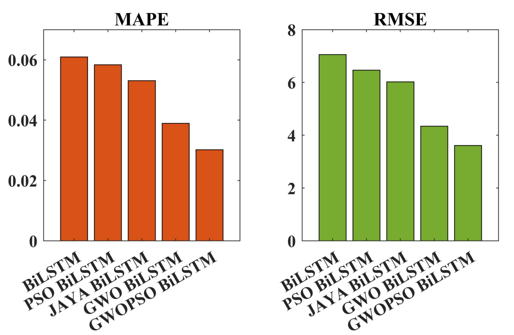

Figure 5. Performance metrics representation for the one-day forecast.

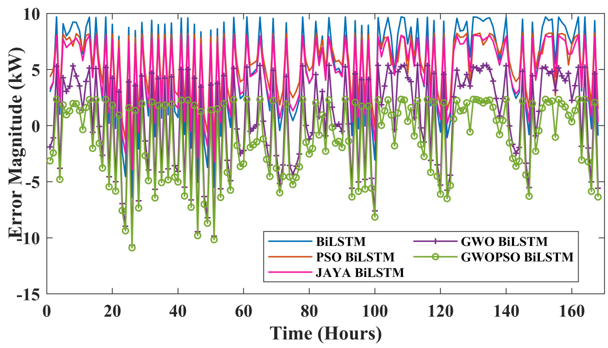

Figure 7. One-week forecast error magnitudes.

which is a minimal amount concerning actual demand.
Except for the proposed technique, all approaches perform slightly better for a one-week forecast than a one-day forecast. Still, the predictions of the proposed method are more accurate, and its error levels are lower than those of other developed models, as shown in Table. 4. The proposed method and other developed models perform similarly to one-day projections, such as the optimization-based BiLSTM networks, which are superior to the traditional BiLSTM technique for over a week forecast.
The performance metrics for the week forecast are shown in Fig. 8 and reveal that the proposed model metrics are almost half of the non-optimized BiLSTM network error metrics when comparing models of MAPE and RMSE. The forecast comparison of the models against actual demand is shown in Fig. 9. Even though the proposed GWO-PSO-BiLSTM technique has a little more discrepancy during a drop in consumption between 20 and 60 hours, it still works well. However, it better interpreted the rising demand and peak conditions over 120 to 168 hours.
The day-wise performance of all models, including the proposed model, are compared and shown in Table. 5. The prediction accuracy of all models during the weekend is slightly less than on weekdays and the overall week forecast due to the high amplitudes of consumption swings. Among the weekdays, starting and ending weekdays (Monday and Friday) have a slightly higher forecast error than mid-weekdays. During mid-weekdays, the performance of the traditional BiLSTM is much better over weekends and slightly outperforms the PSO-BiLSTM in terms of MAE and MAPE. In contrast to the above, the PSO-BiLSTM performance during weekends is much better than traditional BiLSTM and thus improved overall week forecast compared to traditional BiLSTM. Even though the performance of the PSO-BiLSTM is better than the BiLSTM technique, accuracy improvement is less. The JAYA-BiLSTM and GWO-BiLSTM approaches outperform the PSO-BILSTM in one-day and one-week forecasts. The one-week forecast also proves that the optimal set of hyperparameters improves the forecast accuracy.

5. Conclusions

This paper proposed an improved BiLSTM technique to enhance the accuracy of a one-day and one-week forecast for a college building's electricity consumption. The BiLSTM network performance has been enhanced by determining its optimal values for hyperparameters using GWO-PSO-based hybrid optimization. The BiLSTM prediction-based MSE is considered as an objective function of the optimization technique. The performance of the proposed model is validated using metrics MAE, MSE, MAPE, RMSE, and NRMSE against BiLSTM without optimization. Simulation results indicate that the proper tuning of hyperparameters using optimization algorithms outperformed the traditional BiLSTM network, improving accuracy by minimizing errors. The proposed method is flexible and reduces the data collection requirement. The work can be extended to other commercial buildings, like hospitals, restaurants, office buildings, and grocery stores, using a hybrid DL technique optimized by a hybrid optimization technique.

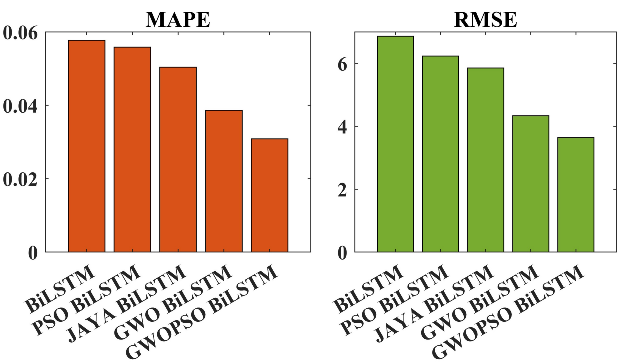

Figure 8. Performance metrics comparison for one week forecast.

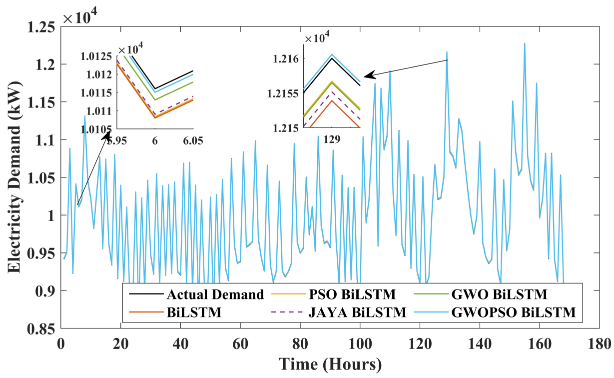

Figure 9. One-week forecast comparison.

Table 4. One-week forecast performance metrics.

Technique/ metrics	BiLSTM	PSO BiLSTM	JAYA BiLSTM	GWO BiLSTM	GWOPSO BiLSTM
MAE (kW)	5.96250	5.70425	5.18374	3.85568	2.95618
MAPE (%)	0.05771	0.05582	0.05035	0.03864	0.03085
MSE	47.08273	38.83751	34.20738	18.79067	13.25987
RMSE (kW)	6.86168	6.23197	5.84870	4.33481	3.64141
NRMSE	6.86E-04	6.23E-04	5.84E-04	4.33E-04	3.64E-04

Table 5. Day-wise performance comparison.

Day	Technique/metrics	BiLSTM	PSO BiLSTM	JAYA BiLSTM	GWO BiLSTM	GWOPSO BiLSTM
Monday	MAE	6.24	5.94	5.43	3.87	2.89
	MAPE	0.060	0.058	0.053	0.038	0.030
	MSE	49.7	41.7	36.2	18.8	13.02
	RMSE	7.05	6.46	6.02	4.33	3.608
	NRMSE	7.08E-04	6.48E-04	6.04E-04	4.35E-04	3.62E-04
Tuesday	MAE	5.27	4.89	4.38	4.95	4.25
	MAPE	0.052	0.048	0.043	0.051	0.045
	MSE	41.6	33.4	29.2	28.1	25.66
	RMSE	6.45	5.78	5.40	5.30	5.06
	NRMSE	6.65E-04	5.96E-04	5.57E-04	5.47E-04	5.22E-04
Wednesday	MAE	4.78	4.94	4.28	3.62	3.59
	MAPE	0.047	0.049	0.042	0.037	0.038
	MSE	33.4	30.3	24.9	19.0	18.2
	RMSE	5.78	5.50	4.99	4.35	4.27
	NRMSE	5.97E-04	5.68E-04	5.15E-04	4.50E-04	4.41E-04
Thursday	MAE	5.39	5.70	4.91	2.80	2.56
	MAPE	0.053	0.056	0.048	0.028	0.026
	MSE	39.1	36.9	30.1	11.8	9.12
	RMSE	6.26	6.07	5.49	3.44	3.02
	NRMSE	6.36E-04	6.17E-04	5.57E-04	3.50E-04	3.07E-04
Friday	MAE	6.20	5.69	5.37	3.98	2.67
	MAPE	0.058	0.054	0.050	0.038	0.027
	MSE	50.2	38.0	35.7	19.1	11.0
	RMSE	7.08	6.172	5.98	4.37	3.32
	NRMSE	6.92E-04	6.02E-04	5.83E-04	4.27E-04	3.24E-04
Saturday	MAE	7.09	6.57	6.11	3.72	2.18
	MAPE	0.067	0.063	0.058	0.035	0.02
	MSE	59.0	47.4	42.7	16.3	6.73
	RMSE	7.68	6.88	6.53	4.04	2.59
	NRMSE	7.47E-04	6.70E-04	6.36E-04	3.93E-04	2.52E-04
Sunday	MAE	6.72	6.16	5.76	4.03	2.52
	MAPE	0.063	0.059	0.054	0.039	0.025
	MSE	56.2	43.8	40.3	18.1	8.98
	RMSE	7.49945	6.62210	6.34895	4.26018	2.99682
	NRMSE	7.28E-04	6.43E-04	6.16E-04	4.14E-04	2.91E-04

Authors Contributions
All the authors have participated sufficiently in the intellectual content, conception and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript.

Availability of data and materials
Data presented in the manuscript are available via request.

Conflict of Interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] I. Hamilton and O. Rapf. "**Executive summary of the 2020 global status report for buildings and construction**.". *United Nations Environment Programme*, , 2020. DOI: https://doi.org/20.500.11822/34572.

[2] M. Santamouris and K. Vasilakopoulou. "**Present and future energy consumption of buildings: challenges and opportunities towards decarbonisation**.". *e-Prime*, 1(8):pp. 100002, 2021. DOI: https://doi.org/10.1016/j.prime.2021.100002.

[3] Inger Andersen. "**Global status report for buildings and construction: towards a zero-emission, efficient and resilient buildings and construction sector**.". *United Nations Environment Programme*, , 2022.

[4] N. Amjady, F. Keynia, and H. Zareipour. "**Short-term load forecast of microgrids by a new bilevel prediction strategy**.". *IEEE Trans. Smart Grid*, 1(3):pp. 286–294, 2010. DOI: https://doi.org/10.1109/TSG.2010.2078842.

[5] EIA. "**CBECS 2018 building characteristics flipbook**.". *EIA*, , 2018.

[6] L. Hernandez et al. "**A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings**.". *IEEE Commun. Surv. Tutorials*, 16.

[7] J. P. Catalã and Ed. "**ELECTRIC POWER SYSTEMS advanced forecasting techniques and optimal generation scheduling, 2012th ed**.". *Boca Raton: CRC PRESS*, :pp. 3.1–3.2., 2017.

[8] R. Weron. "**Modeling and forecasting electricity loads and prices**.". *Chichester: Wiley*, :pp. 67–100, 2006.

[9] K. Li, C. Hu, G. Liu, and W. Xue. "**Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis**.". *Energy Build*, 108:pp. 106–113, 2015. DOI: https://doi.org/10.1016/j.enbuild.2015.09.002.

[10] E. Ofori-Ntow Jnr and Y. Y. Ziggah. "**Electricity demand forecasting based on feature extraction and optimized backpropagation neural network**.". *e-Prime - Adv. Electr. Eng. Electron. Energy.*, 6(9):pp. 100293, 2023. DOI: https://doi.org/10.1016/j.prime.2023.100293.

[11] V. Gundu and S. P. Simon. "**PSO–LSTM for short term forecast of heterogeneous time series electricity price signals**.". *J. Ambient Intell. Humaniz. Comput.*, 12(2):pp. 2375–2385, 2021. DOI: https://doi.org/10.1007/s12652-020-02353-9.

[12] T. Y. Kim and S. B. Cho. "**Particle swarm optimization-based CNN-LSTM networks for forecasting energy consumption**.". *2019 IEEE Congress on Evolutionary Computation, CEC 2019 - Proceedings*, :pp. 1510–1516, 2019. DOI: https://doi.org/10.1109/CEC.2019.8789968.

[13] J. Sauer, V. C. Mariani, L. dos Santos Coelho, M. H. D. M. Ribeiro, and M. Rampazzo. "**Extreme gradient boosting model based on improved Jaya optimizer applied to forecasting energy consumption in residential buildings**.". *Evol. Syst.*, 13(4):pp. 577–588, 2022. DOI: https://doi.org/10.1007/s12530-021-09404-2.

[14] A. M. Althahabi, H. M. Abed, R. Khalid, A. Ryadh, and A. Al. "**Prediction of the electricity demand in the market an application of optimization and machine learning**.". *Majlesi J. Electr. Eng.*, 17(2):pp. 109–115, 2023. DOI: https://doi.org/10.30486/mjee.2023.1986619.1140.

[15] C. Li. "**Designing a short-term load forecasting model in the urban smart grid system**.". *Appl. Energy.*, 266(1):pp. 114850, 2020. DOI: https://doi.org/10.1016/j.apenergy.2020.114850.

[16] U. B. Tayab, A. Zia, F. Yang, J. Lu, and M. Kashif. "**Short-term load forecasting for microgrid energy management system using hybrid HHO-FNN model with best-basis stationary wavelet packet transform**.". *Energy*, 203:pp. 117857, 2020. DOI: https://doi.org/10.1016/j.energy.2020.117857.

[17] S. K. Panda, P. Ray, and S. R. Salkuti. "**Hybrid techniques for short term load forecasting**.". *Majlesi J. Electr. Eng.*, 17(1):pp. 55–61, 2023. DOI: https://doi.org/10.30486/mjee.2023.1970200.0.

[18] A. Irankhah, M. H. Yaghmaee, and S. Ershadi-Nasab. "**Optimized short-term load forecasting in residential buildings based on deep learning methods for different time horizons**.". *J. Build. Eng.*, 84(1):pp. 108505, 2024. DOI: https://doi.org/10.1016/j.jobe.2024.108505.

[19] S. Afzal, A. Shokri, B. M. Ziapour, H. Shakibi, and B. Sobhani. "**Building energy consumption prediction and optimization using different neural network-assisted models; comparison of different networks and optimization algorithms**.". *Eng. Appl. Artif. Intell.*, 127:pp. 107356, 2024. DOI: https://doi.org/10.1016/j.engappai.2023.107356.

[20] K. P. Amber, M. W. Aslam, and S. K. Hussain. "**Electricity consumption forecasting models for administration buildings of the UK higher education sector**.". *Energy Build*, 90:pp. 127–136, 2015. DOI: https://doi.org/10.1016/j.enbuild.2015.01.008.

[21] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer. "**Short-term load forecasting in a non-residential building contrasting models and attributes**.". *Energy Build*, 92:pp. 322–330, 2015. DOI: https://doi.org/10.1016/j.enbuild.2015.02.007.

[22] Y. T. Chae, R. Horesh, Y. Hwang, and Y. M. Lee. "**Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings**.". *Energy Build*, 111:pp. 184–194, 2016. DOI: https://doi.org/10.1016/j.enbuild.2015.11.045.

[23] H. Dagdougui, F. Bagheri, H. Le, and L. Dessaint. "**Neural network model for short-term and very-short-term load forecasting in district buildings**.". *Energy Build*, 203, 2019. DOI: https://doi.org/10.1016/j.enbuild.2019.109408.

[24] C. Ghenai, O. A. A. Al-Mufti, O. A. M. Al-Isawi, L. H. L. Amirah, and A. Merabet. "**Short-term building electrical load forecasting using adaptive neuro-fuzzy inference system (ANFIS)**.". *J. Build. Eng.*, 52:pp. 104323, 2022. DOI: https://doi.org/10.1016/j.jobe.2022.104323.

[25] R. E. Edwards, J. New, and L. E. Parker. "**Predicting future hourly residential electrical consumption: a machine learning case study**.". *Energy Build*, 49:pp. 591–603, 2012. DOI: https://doi.org/10.1016/j.enbuild.2012.03.010.

[26] N. Tsalikidis, A. Mystakidis, C. Tjortjis, P. Koukaras, and D. Ioannidis. "**Energy load forecasting: one-step ahead hybrid model utilizing ensembling**.". *Computing*, 106(1), 2023.

[27] H. Chitsaz, H. Shaker, H. Zareipour, D. Wood, and N. Amjady. "**Short-term electricity load forecasting of buildings in microgrids**.". *Energy Build*, 99:pp. 50–60, 2015. DOI: https://doi.org/10.1016/j.enbuild.2015.04.011.

[28] M. Cai, M. Pipattanasomporn, and S. Rahman. "**Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques**.". *Appl. Energy.*, 236(12):pp. 1078–1088, 2019. DOI: https://doi.org/10.1016/j.apenergy.2018.12.042.

[29] M. Sajjad et al. "**A novel CNN-GRU-Based hybrid approach for short-term residential load forecasting**.". *IEEE Access*, 8:pp. 143759–143768, 2020. DOI: https://doi.org/10.1109/ACCESS.2020.3009537.

[30] P. Koukaras, N. Bezas, P. Gkaidatzis, D. Ioannidis, D. Tzovaras, and C. Tjortjis. "**Introducing a novel approach in one-step ahead energy load forecasting**.". *Sustain. Comput. Informatics Syst.*, 32:pp. 100616, 2021. DOI: https://doi.org/10.1016/j.suscom.2021.100616.

[31] Y. X. Kong, Weicong, Zhao Yang Dong, David J. Hill, and Fengji Luo. "**Short-term residential load forecasting based on resident behaviour learning**.". *IEEE Trans. Power Syst.*, 33(1):pp. 2017–2018, 2018.

[32] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang. "**Short-term residential load forecasting based on LSTM recurrent neural network**.". *IEEE Trans. Smart Grid.*, 10(1):pp. 841–851, 2019. DOI: https://doi.org/10.1109/TSG.2017.2753802.

[33] A. Rahman, V. Srikumar, and A. D. Smith. "**Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks**.". *Appl. Energy.*, 212(10):pp. 372–385, 2018. DOI: https://doi.org/10.1016/j.apenergy.2017.12.051.

[34] A. M. Mahmood et al. "**Electricity demand prediction by a transformer-based model**.". *Majlesi J. Electr. Eng.*, 16(4):pp. 97–102, 2022. DOI: https://doi.org/10.30486/mjee.2022.696520.

[35] J. Wang, S. Zhu, W. Zhang, and H. Lu. "**Combined modeling for electric load forecasting with adaptive particle swarm optimization**.". *Energy*, 35(4):pp. 1671–1678, 2010. DOI: https://doi.org/10.1016/j.energy.2009.12.015.

[36] W. S. McCulloch and W. Pitts. "**A logical calculus of the ideas immanent in nervous activity**.". *Bull. Math. Biophys.*, 5(4):pp. 115–133, 1943. DOI: https://doi.org/10.1007/BF02478259.

[37] S. Hochreiter and J. Schmidhuber. "**Long short-term memory**.". *Neural Comput*, 9(8):pp. 1735–1780, 1997. DOI: https://doi.org/10.1162/neco.1997.9.8.1735.

[38] F. C. Felix, A. Gers, and Jurgen Schmidhuber. "**Learning to forget: continual prediction with LSTM**.". *Neural Comput*, 12(10):pp. 2451–2471, 2000.

[39] N. Somu, G. Raman, M R, and K. Ramamritham. "**A deep learning framework for building energy consumption forecast**.". *Renew. Sustain. Energy Rev.*, 137(4):pp. 110591, 2021. DOI: https://doi.org/10.1016/j.rser.2020.110591.

[40] N. Mughees, S. A. Mohsin, A. Mughees, and A. Mughees. "**Deep sequence to sequence Bi-LSTM neural networks for day-ahead peak load forecasting**.". *Expert Syst. Appl.*, 175(12):pp. 114844, 2021. DOI: https://doi.org/10.1016/j.eswa.2021.114844.

[41] R. Masadeh, A. Alzaqebah, and A. Hudaib. "**Grey wolf algorithm for requirements prioritization**.". *Mod. Appl. Sci.*, 12(2):pp. 54, 2018. DOI: https://doi.org/10.5539/mas.v12n2p54.

[42] M. Schuster and K. K. Paliwal. "**Bidirectional recurrent neural networks**.". *IEEE Trans. Signal Processing.*, 45(11):pp. 2673–2681, 1997.

[43] S. Mirjalili, S. M. Mirjalili, and A. Lewis. "**Grey wolf optimizer**.". *Adv. Eng. Softw.*, 69:pp. 46–61, 2014. DOI: https://doi.org/10.1016/j.advengsoft.2013.12.007.

[44] M. A. M. Shaheen, H. M. Hasanien, and A. Alkuhayli. "**A novel hybrid GWO-PSO optimization technique for optimal reactive power dispatch problem solution**.". *Ain Shams Eng. J.*, 12(1):pp. 621–630, 2021. DOI: https://doi.org/10.1016/j.asej.2020.07.011.

[45] J. Kennedy and R. Eberhart. "**Particle swarm optimization**.". *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4, 1995. DOI: https://doi.org/10.1109/ICNN.1995.488968.

[46] Angizeh Farhad, Ghofrani Ali, and Jafari Mohsen A. "**Dataset on hourly load profiles for a set of 24 facilities from industrial, commercial, and residential end-use sectors**.". *Mendeley Data*, 1, 2020. DOI: https://doi.org/10.17632/rfnp2d3kjp.1.