An Efficient Renewable Power Prediction Model Using IGOA-based Enhanced BiLSTM Deep Neural Network for Microgrid Smart Energy Management

Safaa Ahmed Jawad Alshareefi

Ministry of Health, Iraq
Email: electr.engineer1977@gmail.com

Article History

Received: Apr. 24, 2026

Revised: Jul. 09, 2026

Accepted: Aug. 05, 2026

Abstract

Accurate forecasting of renewable energy generation and operational variables is a critical requirement for effective smart energy management in microgrid systems. This paper presents an integrated forecasting framework based on a Bidirectional Long Short-Term Memory network optimized using an Improved Grasshopper Optimization Algorithm (IGOA-BiLSTM). The proposed method exploits the bidirectional learning capability of BiLSTM to model complex nonlinear temporal dependencies, while an enhanced Grasshopper Optimization Algorithm is introduced to optimally tune the network's hidden layer size. The IGOA incorporates a dynamic inertia weight updating mechanism and an adaptive triangular mutation strategy to improve exploration–exploitation balance, enhance convergence stability, and avoid premature convergence. The model is evaluated using real-world hourly data from the PJM West Zone, including photovoltaic power, wind turbine power, load demand, and day-ahead electricity prices. Experimental results demonstrate that the proposed IGOA-BiLSTM achieves superior forecasting accuracy compared with conventional LSTM and state-of-the-art hybrid models. Specifically, the proposed method attains coefficients of determination (R^2) of 0.99 for photovoltaic power, wind power, load demand, and day-ahead price forecasting. For photovoltaic and wind power prediction, RMSE values of 0.019 and 0.015 are achieved, respectively, with corresponding MSE values as low as 0.0003 and 0.0002. In load and day-ahead price forecasting, the model achieves RMSE values of 0.019 and 0.009, respectively. These results confirm the robustness, high accuracy, and computational efficiency of the proposed framework, making it well suited for practical microgrid energy management, renewable energy planning, and electricity market decision-support applications.

Keywords- Renewable Power Prediction, Microgrid Energy Management, Smart Grid, Improved Grasshopper Optimization Algorithm (IGOA), Grasshopper Optimization Algorithm (GOA).

I. INTRODUCTION

The growing infiltration of renewable energy sources like photovoltaic (PV) systems and wind turbines has radically altered the design and the workings of the contemporary power systems [1]. Microgrids, which are a combination of distributed generation units, energy storage systems, controllable loads, and market interaction mechanisms, have played a major role in improving flexibility, reliability, and sustainability of the future smart grids [2,3]. Nevertheless, the stochastic and intermittent characteristics of renewable energy production, the dynamic behavior of loads, and unstable electricity prices are major challenges to the effective operation and energy management of microgrid systems. Precise prediction of renewable energy production, load demand and market prices of electricity is hence an essential requirement for dependable scheduling, optimal dispatch, and economical decision-making in microgrid settings [4].

The common traditional forecasting models, which are statistical and time series models, such as autoregressive integrated moving average (ARIMA), exponential smoothing, and linear regression have been extensively utilized in power systems [5]. Although these techniques are computationally efficient and interpretable, they tend to be not very effective in the context of highly nonlinear, nonstationary and complex temporal patterns that are characteristic of renewable energy and electricity market data. Consequently, these models find it difficult to reflect sudden changes due to weather variability, human behavior and market dynamics and thus have lower predictive accuracy and low resilience in practice [6].

Energy forecasting has increasingly relied on data-driven approaches, with machine learning and deep learning methods gaining prominence as robust predictive tools in recent years [7]. These models have proven to be very useful in the modeling of nonlinear relationships and the extraction of hidden relationships in large-scale data. Long Short-Term Memory (LSTM) networks have now become one of the most useful time-series prediction architectures by virtue of being able to support long-term temporal correlations and overcome the vanishing gradient issue that is typical of traditional RNNs. As a result, LSTM-based models have been effectively used in the field of renewable power prediction, load prediction, and electricity price forecasting [8,9].

Although they have benefits, typical LSTM models are unidirectional processors of temporal information, which only use past observations to make predictions. However, in most cases of energy forecasting, contingent information in the future within a time window may be very useful in understanding the temporal aspects as well as enhancing prediction quality [10]. BiLSTM networks overcome this drawback by learning both backward and forward sequences, thus providing more temporal representations as well as modeling power capability. However, the predictive performance of BiLSTM networks is extremely vulnerable to the choice of hyperparameters, especially the size of the hidden layer, which directly influences model capacity, convergence, and generalization performance. Practically, these parameters are frequently chosen either by trial and error or by experiment, and can result in under-performance or overfitting [11,12].

In order to address these limitations, automated hyperparameter tuning of deep learning models has been treated with increasing popularity using metaheuristic-based optimization algorithms [13]. These algorithms have demonstrated good outcomes of enhancing the accuracy of the forecast due to their ability to search large and multifaceted solution spaces efficiently. Nevertheless, the common traditional metaheuristic algorithms have problems of premature convergence, lack of balance between exploration and exploitation as well as parameter initialization sensitivity, which can impede their efficiency in high-dimensional optimization problems [14].

Based on the challenges mentioned above, this paper presents an improved forecasting architecture that combines a Bidirectional Long Short-Memory network and an Optimized Grasshopper Optimization Algorithm (IGOA) to manage smart microgrid energy. The suggested IGOA presents a dynamic inertia weight updating scheme and adaptive triangular mutation policy to enhance exploration ability, uphold population diversity and enhance convergence stability. Using the enhanced optimization strategy, the suggested framework automatically detects the optimal number of hidden layers of the BiLSTM network, thus improving the prediction accuracy and strength without human involvement.

The efficiency of the suggested IGOA-BiLSTM model is confirmed with the help of actual hourly data of the PJM West Zone which comprises photovoltaic power production, wind turbine power generation, load demand, and day-ahead electricity prices. Extensive experiments and comparisons with traditional LSTM-based models and various state-of-the-art hybrid strategies prove that the proposed method has the best performance in all forecasting tasks with considerably lower error of predictions and coefficients of determination that approach unity.

The main contributions of this paper can be summarized as follows:

- This study proposes an integrated deep learning and metaheuristic optimization framework that autonomously performs optimal parameter tuning and selection for the BiLSTM neural network by exploiting enhanced exploration capabilities. The proposed framework leads to a substantial improvement in prediction accuracy and overall performance of renewable energy forecasting and smart energy management systems in microgrids.
- An improved version of the Grasshopper Optimization Algorithm (IGOA) is developed, incorporating a dynamic inertia weight updating mechanism and an adaptive mutation factor to effectively identify optimal BiLSTM network parameters for accurate electrical power prediction in microgrid environments. This design enables a well-balanced trade-off between exploration and exploitation phases, significantly reducing the likelihood of premature convergence to local optima.
- A dynamic and reinforced exploration mechanism is introduced by integrating a dynamic mutation coefficient with a triangular mutation strategy. This mechanism prevents solution updates from being confined to the vicinity of the current local best solution in the GOA optimization process and enhances global search capability and convergence stability by maintaining population diversity.

The remainder of this paper is organized as follows. Section 2 reviews related work on renewable energy, load, and electricity price forecasting. Section 3 presents the method, including data preprocessing, the BiLSTM model, and the IGOA-based optimization strategy. Section 4 discusses the evaluation results. The paper is concluded in Section 5, where future research directions are also identified.

II. RELATED WORKS

In [15], the authors explore the concept of microgrids as a powerful tool for enhancing the sustainability and resiliency of local power systems, with particular emphasis on the significance of precise energy management and load forecasting. In the study, the authors propose a new energy management paradigm using machine learning methods to forecast solar and wind production, load demand, and battery State of Charge (SoC). In combination with several ML models, historical load data and time-series analysis are used to facilitate commercial microgrid operations. The LSTM and Linear Regression are compared on energy prediction, load forecasting, and SoC estimation tasks, and in each of the tasks, the Random Forest, LSTM, and ANN prove to be better.

In [16], the authors highlight the significance of precise solar and wind power prediction to ensure quality integration into the grid and present the weaknesses of the traditional, machine learning, and hybrid models regarding adaptability, generalization, and the ability to predict multiple sources. To address these problems, the paper suggests two hybrid deep learning frameworks CNN-ABiLSTM and CNN-Transformer-MLP. The CNNs in both architectures are employed to extract short-term temporal features, whereas ABiLSTM and Transformer-MLP modules learn long-term dependencies. The hybrid methods always perform better than standalone CNN, BiLSTM and Transformer baselines using quarter-hour real-time data. The CNN-Transformer-MLP has better short-term forecasting performance, but CNN-ABiLSTM has higher performance in long term prediction of wind power.

The authors in [17] introduce a new coupled forecasting model to enhance the predictive performance of short-term and medium-term renewable energy load demand. The paper improves the Whale Optimization Algorithm (WOA) by adding Tent chaotic mapping and a nonlinear convergence factor to create an Improved WOA (IWOA), which optimizes BiLSTM network weights. Simulated Annealing (SA) is also used to optimize some important hyperparameters, such as learning rate, number of nodes in the hidden layer, and the number of training iterations to create an IWOA-SA-BiLSTM model. The findings of the case study illustrate that it has resulted in a high level of performance with improvements of 76.7, 74.5 and 45.9 in MAE, RMSE and R 2, respectively, compared to the benchmark methods.

In [18], the authors address the idea of the relevance of accurate renewable energy forecasting to the stability and efficiency of power systems by comparing a variety of deep learning models (LSTM, BiLSTM, GRU, and CNN-LSTM) to predict solar and wind power. These models are compared to the conventional ARIMA model. The research hypothesizes a designed DNN model to predict renewable energy with consideration of advanced hyperparameter optimization and incorporation of important meteorological and time-related features. The optimized LSTM is the most accurate model among the assessed models, but GRU, CNN-LSTM, and BiLSTM are also very promising. The findings indicate that feature selection and model tuning are effective in the forecasting of renewable energy.

In [19], the authors discuss the issues of large-scale integration of renewable energy and place the microgrids in an effective framework of coordinating distributed energy resources. The work highlights the importance of an Energy Management System to achieve stable operation of the microgrids and suggests a real-time energy management strategy based on deep reinforcement learning. Electric vehicles are also included as controllable resources together with intermittent photovoltaic and wind generation to enhance flexibility and self-balancing of the system. The suggested solution proves to be better than traditional scheduling approaches in terms of self-balancing rate improvement by up to 22.97% and operator profit up to 33.74, which proves its usefulness in practice as a microgrid.

In [20], the authors discuss the difficulties of managing microgrid energy during islanded operation and suggest a hybrid model of energy forecasting with deep learning and adaptive reinforcement learning to plan in real time. The strategy takes advantage of the fact that deep learning is highly predictive, and reinforcement learning is flexible enough to dynamically trade off the supply and demand of energy. The combined framework improves the stability of the grids, efficiency of operations and cost efficiency during uncertain and time-varying situations. The experimental analyses show significant improvements in performance when compared to the current approaches such as PLECS, DRMPC, and SCADA, attaining high performance improvements in SMAPE, energy, and operational costs, which highlights the efficiency of intelligent computational methods in resilient microgrid control.

In [21], the authors present an integrated hybrid energy management system of interconnected microgrid systems with deep neural networks, model-free reinforcement learning, and cooperative game theory. The method maintains the privacy of microgrids since it uses deep learning models to determine the operating behavior without necessarily accessing the internal data. The pricing mechanism based on reinforcement learning is dynamic in that it will adjust the retail electricity price to the current system conditions, and a Shapley value-based allocation scheme will be used to make sure that the profits are distributed fairly among participating microgrids. The framework allows two-way energy flow within the PCC capacity limits. The outcomes of the simulation show significant decreases in the average retail prices, which proves the efficiency of the framework to reach cost efficiency, flexibility in pricing, and fairness of cooperation.

In [22], authors discuss the issues of running microgrids that contain integrated renewable energy sources because of unpredictable changes in supply and demand. The paper suggests an ideal energy management system (using deep learning) that uses a pre-trained model to predict the hybrid renewable energy production. The data of solar and wind is gathered through open sources and characteristics are identified based on a soft attention Visual Geometry Group 16 (SAVGG16) method. Time-series prediction is carried out using an optimized LSTM model, whose tuning parameters are optimized using a chaotic oppositional and shrinkage control-based rat swarm optimization algorithm. The performances of simulations are higher, and the R 2 values of 0.9929 and 0.9922 of solar and wind power forecasts with a small error.

Authors in [23] discuss the operational issues of microgrids where the renewable energy is high, and the uncertainties in the output of renewable energy, load and day-ahead pricing. The research suggests a deep learning forecasting model, which is an LSTM network with a global attention mechanism and its hyperparameters are optimized through a genetic algorithm-adaptive weight particle swarm optimization (GA-AWPSO) algorithm to enhance prediction accuracy. In order to address demand-side uncertainty, a DM-CIDR program is created, in which incentive rate strategies are designed using the k-NN classification and the OPTICS clustering methods to address the needs of the customers. The simulation findings based on the historical PJM data indicate that the proposed method can be successfully used to predict and address uncertainties in the microgrid functioning.

In [24], the authors examine the concept of zero-carbon multi-energy systems (ZCMES), and suggest a deep learning-based and optimization-oriented model of the optimal day-ahead scheduling of virtual power plants. The framework incorporates a carbon capture system, electric vehicle flexibility, and clean energy marketer strategy in order to achieve reliability of the systems that will be, sustainable. In the case of multivariable time-series forecasting, a GRU-BiLSTM is used, and scenario generation and reduction are implemented through an autoencoder with a fast forward reduction algorithm. There is a powerful stochastic optimization model that directs the operations. Historical multi-energy case studies of Arizona show considerable self-consumption, load coverage and cost reduction improvements, which indicate the potential of the approach in improving carbon-neutral urban energy systems, as well as in informing sustainable energy policy.

In [25], the authors create a powerful algorithm to predict the energy demand and renewable energy generation of microgrids with the help of deep learning and metaheuristic optimization. The paper uses an LSTM network to learn nonlinear dynamics over time, with the hyperparameters being optimized using a new Improved Dynamic Arithmetic Optimization Algorithm (IDAOA) that includes inertial weights, mutation factors and a triangle mutation operator to balance between exploration and exploitation. Wind turbine, photovoltaic, load demand, and day-ahead price data are tested with the IDAOA-LSTM model with an RMSE of 0.021 and 0.031 in terms of wind and solar power, respectively, and a R 2 of 0.98. Findings indicate that the model is useful in enhancing the accuracy of predictions and the sustainability of energy-efficient microgrids.

III. Methodology

The primary objective of this research is to accurately predict renewable energy generation for effective energy management in microgrid systems. To achieve this goal, a deep Bidirectional Long Short-Term Memory (BiLSTM) neural network is employed. BiLSTM exhibits strong capability in modeling and analyzing time-series data. By mitigating the challenges of conventional RNNs, BiLSTM can retain long-term temporal dependencies, resulting in high prediction accuracy for nonlinear and complex time-series datasets. The predictive capability of the BiLSTM architecture is governed to a large extent by the choice of hyperparameters, which are often selected empirically or randomly. Among these, the hidden layer size is one of the most influential parameters. In this paper, to enhance the effectiveness of the BiLSTM framework, this parameter is optimally determined using an Improved Grasshopper Optimization Algorithm (IGOA). The proposed model is depicted in Fig.1 (Algorithm 1). The suggested model employs a sequential workflow consisting of four main stages to ensure effective synergy among its components. These stages are executed as follows:

1. Data Preprocessing: In this phase, raw time-series data related to microgrid energy generation and demand are first collected. The data are then scaled using a normalization technique to eliminate the adverse effects of scale differences among variables on the learning process. Finally, the normalized data are transformed into input–output sequence formats required for training the BiLSTM network.
2. Base BiLSTM Model Design: At this stage, the baseline architecture of the BiLSTM neural network is defined and configured. This includes determining the overall network structure, specifying the dimensions of the input and output layers based on the preprocessed time-series features, selecting appropriate activation functions, and initializing hyperparameter values before optimization. The baseline model design establishes the search space for the IGOA algorithm.
3. Hyperparameter Optimization Using IGOA: In this stage, the IGOA algorithm initiates and focuses on fine-tuning the dimension (i.e., the number of hidden units) within a single hidden BiLSTM layer the search process to identify the optimal hyperparameters of the BiLSTM model. At each iteration, a candidate set of hyperparameters is generated by IGOA, and the BiLSTM model is trained using these parameters on the training dataset. The prediction error (e.g., Root Mean Square Error (RMSE) eq.(3)) is employed as the fitness function to guide and update the search mechanism of IGOA. This iterative process continues until the hyperparameter configuration yielding the minimum prediction error is obtained.
4. Final Modeling and Prediction: After the convergence of IGOA and the determination of optimal hyperparameters, the final BiLSTM model is fully trained using these optimized settings on the entire training dataset. Final power outputs are generated by deploying the trained model on the unseen test dataset. These predicted values are subsequently used for performance evaluation and practical implementation in the microgrid energy management system.

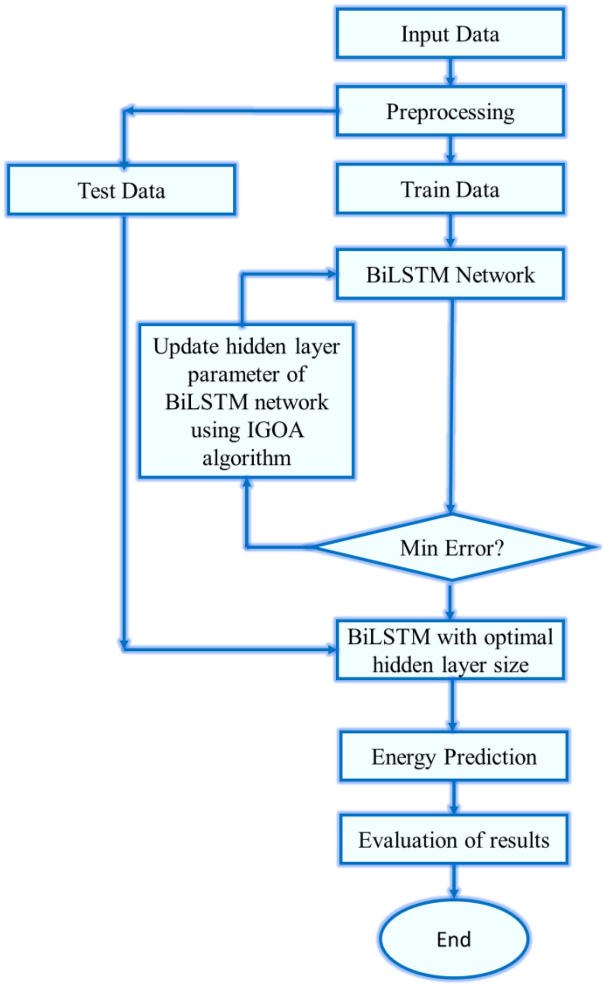

Figure 1. diagram of the proposed method

1) Data Preprocessing

Data preprocessing in the current study involves three key steps, namely, data cleaning, data normalization, and partitioning the dataset into two parts, the training and testing ones.

- **Data Cleaning**

Data cleaning is mainly done to eliminate anomalous records of the original data, such as duplicate, missing and invalid data. The values that are missing in this work are imputed with the help of the average of their closest neighboring values, as presented in (1):

$$x_i = \frac{x_{i-1} + x_{i+1}}{2} \qquad (1)$$

in which x_i is the missing data point to be imputed, and x_{i-1} and x_{i+1} are the data values recorded at the time one hour preceding the missing timestamp and one hour following it, respectively.

The rate of missing data in the dataset utilized in this research is very low and most of the missing data is at single points or over a very short time span (e.g., one or two time steps). In this case, when the temporal separation among observed samples is small, the performance of simple linear interpolation versus more advanced imputation methods, including spline interpolation or KNN-based imputation, is insignificant with respect to the ultimate model performance. That is, there is a restriction on nonlinear effects across these short time gaps. Linear interpolation is thus a good local estimation that is accurate enough, but is simpler and more efficient to compute.

- **Data Normalization**

The large spread in magnitudes of various variables implies that characteristics that have lower numerical values can still play an important role in predicting energy consumption. To ensure the scale differences do not negatively influence the learning process, the data are normalized to the range [0,1] to ensure that all the variables fall within the same range.

Minimum-maximum normalization is used on each feature (feature-wise scaling) to achieve reproducibility and ensure that features are scaled similarly. The normalization process for each feature x_j is defined by the expression below:

$$x_j^{norm} = \frac{x_j - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (2)$$

where $\min(x_j)$ and $\max(x_j)$ are the minimum and maximum values of the j-th feature calculated on the full training set. This minmax scaling on features eliminates the risk that the variables with wider ranges of numbers dominate the learning process, and all the input features have their proportional influence on the model optimization.

- **Data Partitioning**

In this paper, the data are separated into two subsets to be used in training and testing the proposed model. The BiLSTM network parameters are learned using the training dataset and the hidden layer size is optimized using the same dataset and the testing dataset is only used to evaluate the performance of the proposed approach. On the whole, training and testing account for 70 percent and 30 percent of the data, respectively. In order to achieve methodological rigor in time-series forecasting and to avoid information leakage, temporal information in the data is rigorously maintained during the partitioning process and no random shuffling is done. The data is divided by time (the first 70 percent of observations (older time stamps) are in the training set, and the remaining 30 percent (newer time stamps) are in the testing set). This time-sensitive partition is necessary to assess the model using unobserved future data, and it is realistic for forecasting.

2) BiLSTM Network

The Bidirectional Long Short-Term Memory (BiLSTM) network is used for energy prediction in smart grid systems in this study because it is highly effective in temporal modeling and handling abrupt changes in the time-series data. By preserving long-range temporal dependencies and overcoming gradient degradation during training, BiLSTM architectures have proven highly effective for sequential data modeling. The entire design of the BiLSTM recurrent neural network is shown in Figure 2.

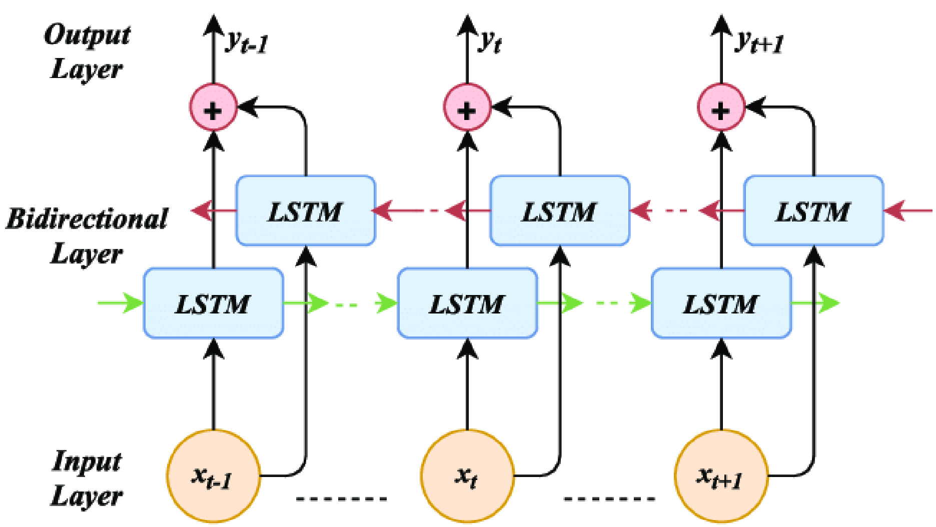

Fig. 2. Structure of the BiLSTM recurrent neural network.

The LSTM cell is a core part of the LSTM architecture and it has four major components: the forget gate, input gate, output gate and the cell state. Information flow within the network is governed by gating units whose sigmoid-based activations decide what content is suppressed, adjusted, or transmitted forward in the temporal sequence. By contrast, the trainable weights of these gates are real numbers that do not have these constraints and are trained during the training process to maximize network performance. The input gate involves the process of choosing and integrating the new information into the cell state, which elements of the input are to be stored. The output gate produces the current output depending on the altered cell state through the regulation of the part of the stored information that is revealed. Furthermore, the cell state is the main memory unit of LSTM, which allows the long-term dependencies to be maintained. It is necessary to mention that the cell state is not directly changed, but it is regulated adaptively by the combined work of the forget, input and output gates. In sequence-based architectures, sophisticated recurrent units, including LSTM, are important in the process of detecting temporal dependencies. But in most of the real-world applications, unidirectional processing is not enough and a model that can capture the past and future context simultaneously is needed. In a Bidirectional LSTM, the two independent layers of LSTM are run in parallel: one is fed with the input sequence in forward mode, and the other with the same sequence reversed from forward to backward. The results of these two layers are then combined at each step in time to create more valuable and detailed temporal representations. Such a bidirectional structure allows exploration of both past and future dependencies, which leads to better performance in those tasks that need to be understood in the full context.

3) Optimization of the BiLSTM Neural Network Using the Improved Grasshopper Optimization Algorithm (IGOA)

Because the size of hidden layers directly affects the learning ability of the BiLSTM neural network, this paper uses the Improved Grasshopper Optimization Algorithm (IGOA) to optimally tune the number of hidden layers of the BiLSTM network. The larger the size of the hidden layer, the more complex features the model can learn, but it is also possible that the model will overfit. Thus, the correct modification of this hyperparameter is of paramount importance in improving the final performance and the generalization capability of the model. In addition, the IGOA algorithm has a high ability to explore large solution spaces and is problem-specific, making it highly effective in searching large solution spaces and having a higher likelihood of finding a globally optimal solution, leading to faster and more efficient hyperparameter optimization. The key stages of the IGOA-based optimization process of finding the optimal hidden layer size of the BiLSTM network can be outlined as follows:

Step 1: Initialization of Parameters

The initial values of the basic parameters of the grasshopper optimization algorithm maximum iteration count, the size of the population and additional control parameters) are set in this step.

Step 2: Initialization of the Grasshopper Population.

Each grasshopper presents a possible answer to the optimization problem and is a candidate configuration of the BiLSTM network parameters. In particular, every grasshopper (particle) represents a potential value of the hidden layer size that is updated and steered to the optimal solution as a result of the IGOA search mechanism. Each execution of the algorithm begins with a randomly initialized population of grasshoppers confined to the specified search space, which is regenerated every run.

Step 3: Fitness Function Evaluation

The definition of a reasonable objective (fitness) function is a critical point of any optimization algorithm. Root Mean Square Error (RMSE) is used as the IGOA fitness function in this study. Based on every candidate hidden layer size that IGOA identifies, a BiLSTM model is trained and applied to make energy predictions. The resulting error between the forecasted and the actual values is then computed as the RMSE, which is used to gauge the quality of the candidate solution. The definition of the RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n} (X_{obs.i} - X_{\text{model }.i})^2} \tag{3}$$

Here, (N) refers to the number of observations, $(X_{obs.i})$ is the actual recorded value, and $(X_{\text{model }.i})$ represents the model's prediction.

Step 4: Selection of the Optimum Answer $\widehat{T_d}(t)$

The solution with the lowest fitness value, which is the best solution in the current iteration, is determined at this stage of the proposed IGOA-based optimization process. Each solution is assessed by comparing it to the best-performing solution identified in previous iterations. In case the current solution has a lower fitness measure, it is adopted as the new global best solution.

Step 5: Updating the Positions of Grasshoppers

The positions of the grasshoppers are changed in the next step by incorporating the following equation:

$$X_i = S_i + G_i + A_i \tag{4}$$

Where (S_i) is the social interaction force of the i-th grasshopper and is determined as:

$$S_i = \sum_{j=1(j\neq i)}^{nPop} s(d_{ij})\hat{d}_{ij} \tag{5}$$

Here, d_{ij} denotes the distance between the i-th and j-th grasshoppers, $\hat{d}_{ij}$ denotes the unit vector between them, and s denotes a function that describes the intensity of social interactions.

where (G_i) is the gravitational force on the i-th grasshopper and is defined as:

$$G_i = -ge_g \tag{6}$$

where g is the gravitational constant and e_g is the unit direction facing towards the center of the earth.

A_i is the force of advection of the wind on the i-th grasshopper and is denoted as:

$$A_i = -ue_w \tag{7}$$

and (u) is a drift constant, and e_w is the unit wind direction.

The combination of the terms above allows the expression of the grasshopper motion equation in the following form:

$$X_i = \sum_{j=1(j\neq i)}^{nPop} s(|x(i) - x(j)|)\frac{x(i)-x(j)}{d_{ij}} - ge_g - ue_w \tag{8}$$

Nevertheless, the initial mathematical representation of the GOA is not quite applicable to optimization problems because grasshoppers very quickly find a comfort zone and the swarm does not always find a single solution. Thus, the position- updating equation is adjusted to solve optimization problems in this study, especially to optimize the position update mechanism of grasshoppers and increase the convergence behavior.

$$x_i^d(t+1) = c_{D\_GOA}\left\{\sum_{j=1(j\neq i)}^{nPop} c_{D\_GOA}\frac{ub_d - lb_d}{2} s(|x(i) - x(j)|)\frac{x(i)-x(j)}{d_{ij}}\right\} + \hat{T}_d(t) \tag{9}$$

Here:

$x_i^d(t+1)$ refers to the new location of the (i)-th grasshopper in the (d)-th dimension,

$x_j(t)$ is the location of the other grasshoppers,

$x_i(t)$ denotes the current location of the (i)-th grasshopper,

c_{D_GOA} represent the dynamic inertia motion weight,
The dynamic inertia weight c_{D_GOA} is critical in balancing exploration and exploitation. By continuously decreasing the sizes of the attraction, neutral, and repulsion regions over successive iterations, this parameter ensures that the population converges. $\hat{T}_d(t)$ refers to the target (best global) position.
Step 6: Dynamic Inertia Weight Updating.
Within the traditional GOA, grasshopper movements are directed mainly by the current global best solution, thus causing the algorithm to be susceptible to premature convergence and getting stuck in local optima, and slow search behavior. In order to overcome this shortcoming, a dynamic inertia weight approach is presented in the proposed IGOA to improve the speed of convergence and search efficiency.
The inertia weight directly regulates the degree of movement of grasshoppers in each step and also the degree to which agents are maintained in their former motion. The larger the inertia weight, the more the exploration of the entire search space is promoted and the smaller the inertia weight, the more is the local exploitation as the search concentrates on potential solutions. Therefore, exploration is better served with greater inertia weights and exploitation with the lower ones.
The inertia weight in the proposed IGOA is varied adaptively and it reduces exponentially and non-linearly with the number of iterations. This adaptive mechanism has a considerable effect on enhancing search performance and quick convergence. To this end, the dynamic inertia weight c_{D_GOA} is updated as shown in the following equation:

$$c_{D\_GOA} = b \times c_{min} \left(\frac{c_{min}}{c_{max}}\right)^{\left(\frac{\frac{1}{1+t}}{t_{max}}\right)} \quad (10)$$

where:
(c_{max}) is the largest value of (c_{D_GOA}) (usually near 1),
(c_{min}) = the minimum value of (c_{D_GOA}) (a small positive value near zero),
b is a random number around 1,
(t_{max})) refers to the maximum number of iterations,
t is the latest iteration.
Step 7: Dynamic Mutation Probability and Triangular Mutation Strategy
To further improve the performance of the proposed IGOA, a dynamic coefficient of mutation is added, which increases search diversity. The mechanism enhances the probability of grasshoppers making larger search moves, thus enhancing the capability of the algorithm to escape local optima. Dynamic mutation probability is given as:

$$p_m = 0.2 + 0.5 \times \frac{t}{T} \quad (11)$$

Moreover, in the mutation strategy, a triangular mutation strategy is used to carry out the mutation operation. The strategy increases population diversity and has lower chances of premature convergence. In the first step, three grasshoppers are chosen randomly, and their data are aggregated in the form below:

$$X(t) = \frac{X_{rg1}+X_{rg2}+X_{rg3}}{3} + (p2 - p1) \times (X_{rg1} - X_{r2}) + (p3 - p2) \times (X_{rg2} - X_{rg3}) + (p1 - p3) \times (X_{rg3} - X_{rg1})$$
(12)

where (X_{rg1}), (X_{rg2}), and (X_{rg3}) are three randomly chosen grasshoppers. The perturbed weights ($p1$), ($p2$) and ($p3$) are obtained as:

$$p1 = \frac{|f(X_{rg1})|}{\acute{p}} \quad (13)$$

$$p2 = \frac{|f(X_{rg2})|}{\acute{p}} \quad (14)$$

$$p3 = \frac{|f(X_{rg3})|}{\acute{p}} \quad (15)$$

Here, $f(.)$ is the fitness function and the normalization constant ($\acute{p}$) is defined as:

$$\acute{p} = |f(X_{r1})| + |f(X_{r2})| + |f(X_{r3})| \quad (16)$$

The use of a triangular mutation strategy allows the algorithm to integrate knowledge from multiple grasshoppers, avoiding dependence on individual local optima when updating positions. This means that the chances of the IGOA getting stuck in local optima are minimized.
Step 8: Stopping Criterion
Steps 3 through 7 are iteratively executed until the algorithm meets its termination condition. In this work, termination occurs when the maximum number of iterations is reached. Lastly, the next section gives the pseudocode of the suggested IGOA.

Algorithm 1- The Pseudocode of IGOA

The IGOA Algorithm pseudocode
Input: Maximum Iteration Number, initial value of Parameters that must be optimized (Number of Hidden Layers), Permissible range of variables. **Output:** Optimized Number of Hidden Layers (H_Best).

Set Initial parameters: cmax, cmin, max_iter, population size, dimension.
Generate Initial populations of Grasshoppers xini
Compute the fitness function **Eq.** Error! Reference source not found.
$\widehat{\mathbf{T}}_{\mathbf{d}}$ = the best search agent
While (l<maximum number of iterations)
  Update c_{D_GOA} Parameter **Eq.** Error! Reference source not found.
  for each search agent
    Update Grasshoppers Position **Eq.** Error! Reference source not found.
    Bring the current search agent back if it goes outside the bounds
  end for
  Update $\widehat{\mathbf{T}}_{\mathbf{d}}$ if there is a better solution
  l=l+1
End while
Return $\widehat{\mathbf{T}}_{\mathbf{d}}$
Set H_Best = $\widehat{\mathbf{T}}_{\mathbf{d}}$

I. Results Evaluation

This section is a detailed analysis of the proposed approach through the performance of four important prediction parameters in the system of smart energy management. These parameters are the power output of photovoltaic (PV) systems, power output of wind turbines (WTs), system load demand (Load) and the day-ahead electricity price (DAP). The choice of these variables is informed by their important functions in balancing supply and demand, optimization of renewable energy resources, as well as economic decision-making in the management of microgrids. Parallel prediction of these parameters gives the complete picture of the future operation of the system and enables the development of effective control mechanisms and schemes of efficient energy consumption. Moreover, taking into account the environmental uncertainty and the natural variation, precise forecasting of every variable improves the reliability of microgrid planning in the short and medium term.

In order to test the proposed model, the computational cost of recurrent training of the LSTM network in the fitness function of the GA-AWPSO algorithm was controlled. The key parameters of the GA-AWPSO algorithm were set in such a way that the population size (15 particles) and the number of iterations (not more than 20) were configured to provide the optimal trade-off between prediction and practical execution time, i.e., 300 fitness evaluations in total during the procedure. It should be noted that the process of hyperparameter optimization is performed once; when convergence is achieved, the best parameters are then fixed in the model. Therefore, at the stage of evaluation (testing), only one application of the model to previously unseen data with an optimized set of parameters is used, which significantly saves the time of testing.

All the simulations were performed in the MATLAB platform on a hardware platform with a 10th-generation Intel Core i7 processor and an NVIDIA GTX 1660 Ti card. In this case, the mean training time of the LSTM model with optimized parameters was approximately 25 minutes and the overall prediction time on the entire test dataset was estimated to be 3 seconds. Since there are only a few fitness evaluations (300) and the inference time in the testing phase is very low, the proposed model can be regarded as computationally efficient and operationally feasible in the context of practical forecasting.

Furthermore, quantitative evaluation of the prediction accuracy was conducted with three common statistical measures of the performance: Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and the coefficient of determination R^2. These measures are commonly used in forecasting models and they are important in the analysis of errors, generalization performance and correlation measurements between observed and predicted values. When these metrics are used together, they allow a multidimensional analysis of the performance of the model, with MSE and RMSE used to measure the numerical accuracy of predictions and the coefficient of determination R^2 used to measure the level of agreement between the patterns that are predicted and the patterns observed. This multi-metric assessment method will help to analyze the results of the performance more accurately and thoroughly and prevent such one-dimensional appraisal. These metrics are formulated as shown in the equations below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n} (X_{obs.i} - X_{\text{model }.i})^2} \tag{17}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{n} (X_{obs.i} - X_{\text{model }.i})^2 \tag{18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (X_{obs.i} - X_{\text{model }.i})^2}{\sum_{i=1}^{n} (X_{obs.i} - \bar{X}_{obs.i})^2} \tag{19}$$

In these equations, (N) represents the number of observations, $(X_{obs.i})$ identifies the actual value, $(X_{\text{model }.i})$ denotes the model's prediction and $(\bar{X}_{obs.i})$ is the average observed value.

1) Dataset

In order to test the performance of the suggested forecasting model, a real-life dataset concerning microgrid energy management is used. The data incorporate load demand, photovoltaic (PV) power generation and wind turbine (WT) power generation. The quality of the data is managed by implementing the right methods of preprocessing so that the training of the model is accurate and reliable. The PJM Interconnection database is used to obtain the dataset to evaluate the model. In particular, the data are related to the PJM West Zone, which is one of the key areas of operation of PJM Interconnection, encompassing the western areas of the PJM power system.

The states of West Virginia, Western Pennsylvania, and Ohio are part of this zone. The PJM West Zone is regarded as a problematic area in terms of load management and power generation analysis because of its geographical features and variety of energy resources. The experiments were chosen to include 800 hourly data samples, which represent a 34-day period in the year 2020. The data are between January 1, 2020, and February 3, 2020, and was sampled at an hourly resolution. The data were split into training and testing groups, where 30 percent of all data (which is also 240 hours) was set to the test group to assess the model's generalization performance [23]. Before any preprocessing, the raw data had a few missing values and noise, which was mostly due to imprecision of sensors or temporary disruption of data recording. A preliminary data cleaning was done before the proposed algorithms were applied. The simple linear interpolation technique was used to identify and reconstruct the missing data points in line with Equation (1). It should be mentioned that the information provided by PJM tends to be structured quite clearly in a chronological sequence and devoid of significant time discrepancies; however, the initial data cleaning was needed to provide strength and uniformity. Figure 3 shows the data used in this study.

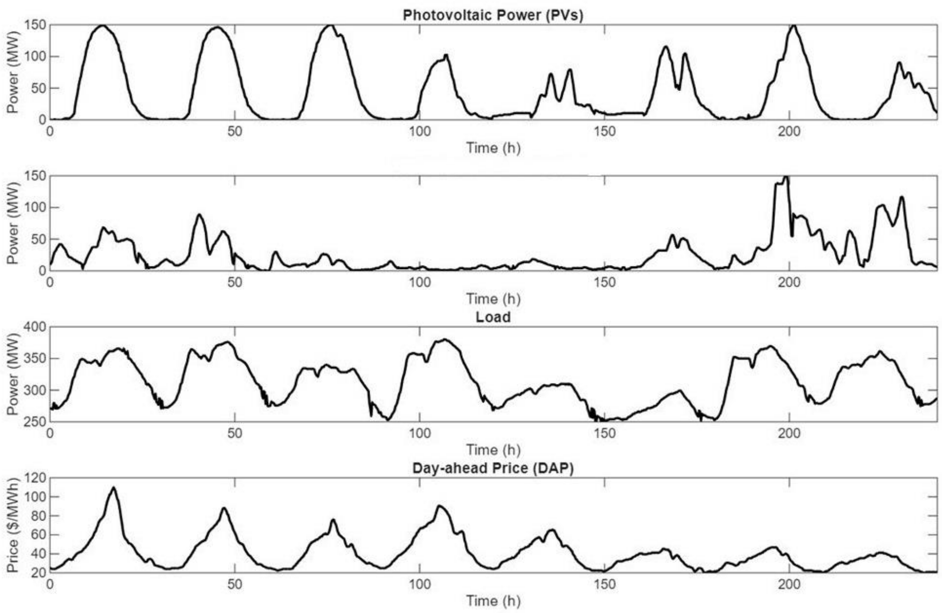

Fig. 3. Dataset utilized in this work

2) Evaluation of the Proposed Method During the Training Phase

Figure 4 demonstrates the convergence pattern of the suggested approach during the training stage. In this figure, the blue curve indicates the value of RMSE at each training iteration, while the red curve indicates the network loss at each training iteration. The loss function and RMSE exhibit a declining pattern as iterations progress, which is a sign of a stable and efficient learning process. The loss of the network reaches a stable point when the number of training iterations exceeds 3000 and the difference between the two becomes insignificant. This affirms the fact that the optimized LSTM network has learned the underlying patterns of the data and has converged well. The low oscillations of the loss function after convergence also demonstrate the strength and good generalization of the trained model.

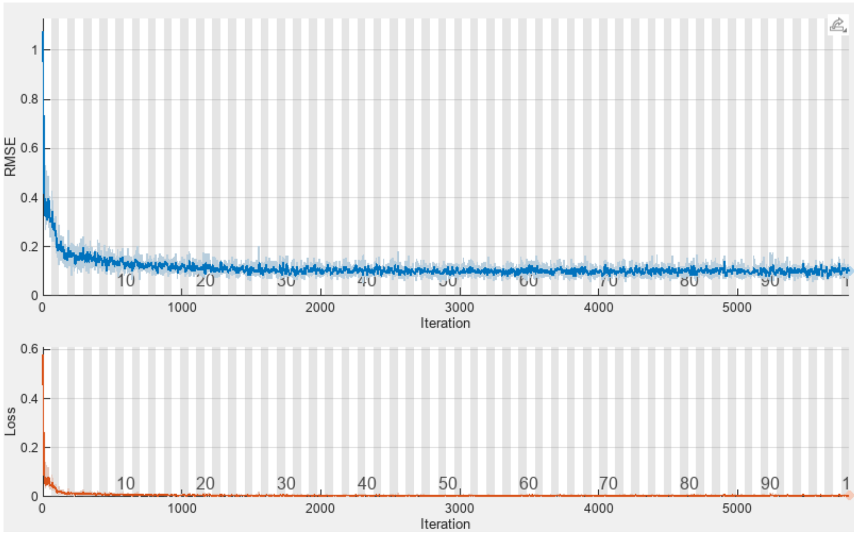

Fig. 4. Performance curve of the LSTM model improved through IGOA-based optimization

3) Performance Evaluation of the Proposed Method in Photovoltaic Power Forecasting

Figure 5 shows the results of the proposed approach for predicting the power output of photovoltaic (PV) systems. The upper plot is the comparison of the actual PV power values (black curve) with the values calculated by the proposed model (green dashed curve). It is also found that the predicted output is very close to both the trend and short-term variation of the actual signal, indicating a very high degree of prediction accuracy and the ability of the model to successfully learn the dynamic behavior of solar power generation over varying periods of time. The bottom plot (cyan curve) shows the normalized difference between the actual and predicted values of PV power with time. As depicted, the error is kept near zero throughout the time steps except for a few areas where there are slight deviations. These minor variations may be explained by environmental noise or sudden weather alterations. It is worth noting that the magnitude of the error lies within the range of [-0.05, 0.05], which indicates the stability, robustness and high generalization capability of the proposed method of operation under different operating conditions. Altogether, the findings presented in Fig. 5 confirm that the proposed model has a high predictive power in terms of photovoltaic power generation and can be used as a stable and predictive instrument in the context of the energy management system based on renewable energy.

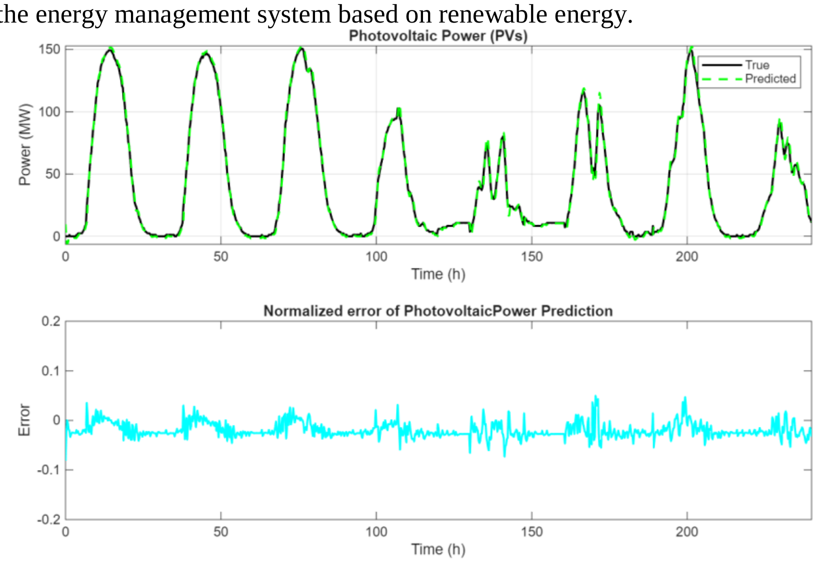

Fig. 5 Performance evaluation of the proposed method in photovoltaic power forecasting.

Figure 6 shows a regression curve of the estimated and actual values of photovoltaic (PV) power that is used to evaluate the accuracy of the proposed model. In this plot, the horizontal axis represents the actual target values and the vertical axis represents the model output. The dots are depicted by black circles and the blue line is the regression (fitting) line between the predicted and actual values. Furthermore, the dotted line (Y = T) is also added as the reference line which represents the perfect fit. As can be seen, the distribution of almost all the data points around the (Y = T) line is close to it, and the regression line almost coincides with the reference line. This means that the proposed model can be used to predict the real PV power values with great accuracy. Moreover, the correlation coefficient indicated at the top of the figure, R = 0.99851, is good testimony to the fact that the linear correlation between the model outputs and the actual values is extremely high. This value is very near one which confirms that the learning model has been able to extract a very precise linear association among the input features and the target outputs. Figure 6 shows that the regression outcomes validate the proposed model's strong predictive performance for photovoltaic power, supporting its applicability as a robust and efficient tool in real-world energy management and renewable energy prediction.

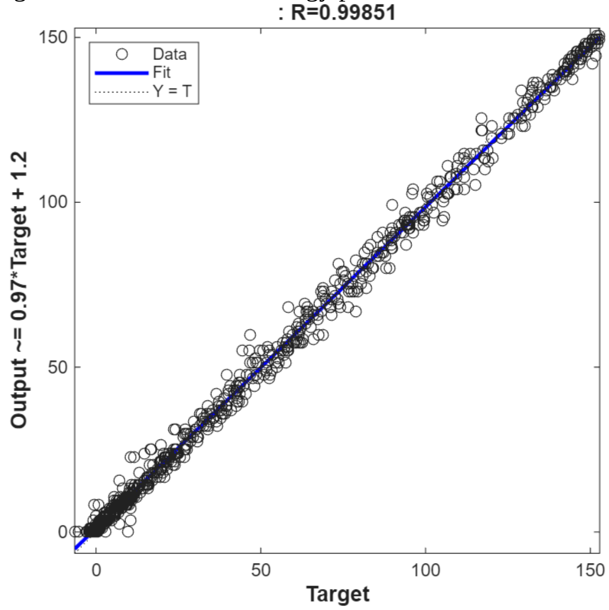

Fig. 6. Regression curve between predicted and actual photovoltaic power values.

4) Performance Evaluation of the Presented Model in Wind Turbine Power Forecasting

Figure 7 shows the findings of power output forecasting of a wind turbine (WT) with the help of the proposed model. The measured power values and the model's predictions are compared in the upper panel of the figure, where the actual values are represented by the black solid line and the predicted values are represented by the green dashed line. As can be seen, the suggested model can successfully reproduce the intricate time-dependent trends and oscillations of wind power production during various periods of time. One of the remarkable features of this figure is that the model has the ability to track sudden and nonlinear changes in power output, which is quite significant for wind energy systems, which are naturally highly volatile. Fig. 7, lower panel, represents the normalized prediction error versus time. Although there are certain localized deviations, the error distribution is limited in general, with most values lying within the range [-0.12, 0.2]. This trend shows that the given model is stable and that it functions in the same manner over the period under consideration. In addition, the fact that no systematic error can be observed implies that the model has generalized well and is not over-reliant on the training data.

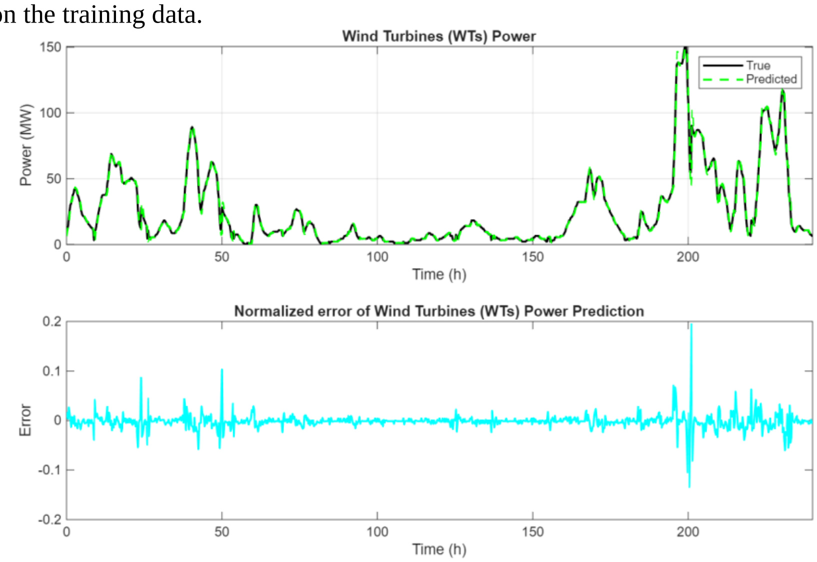

Fig. 7. Performance evaluation of the presented model in wind turbine power forecasting.

The relationship between the predicted and measured wind turbine power values is illustrated in Figure 8. As indicated in the diagram, most of the values (black circles) lie along the ideal (Y = T) line, indicating a strong alignment between predicted and actual values. This agreement is also verified by the fitted regression line (blue), which has a slope of nearly unity with a very small intercept (around 0.26). The prevalence of the data in the neighborhood of the ideal line and the lack of substantial variation at the low or high power levels proves that the given model is stable in its operation throughout the whole span of target values. This consistency is especially significant to applications like wind power generation planning and energy management where it is essential to have the correct forecasting over the entire range of operation. The high correlation between the model's prediction and measured values is an additional indication of the structural precision of the model and its efficiency in acquiring the underlying patterns of wind power variability.

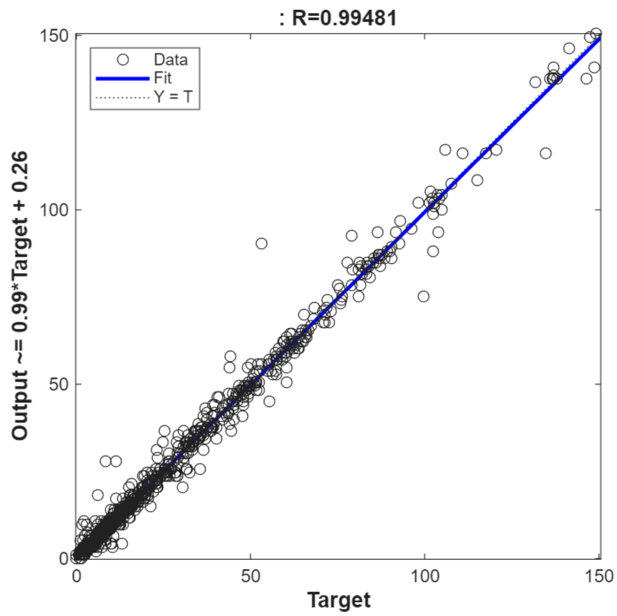

Fig. 8. Regression curve between predicted and actual wind turbine power values.

5) Performance Evaluation of the Proposed Method in Load Forecasting

Figure 9 depicts how effectively the proposed model forecasts the electrical load demand within the power system. The curves presented in this figure clearly indicate a strong correlation between the real load data (black solid line) and the predicted data (green dashed line). The model can also precisely trace the changes in electrical load across various time periods, especially when there is peak and off-peak demand. The ability of the model to learn nonlinear and temporal characteristics of load demand signals is reflected in the high accuracy of the model in capturing short-term fluctuations and recurrent consumption patterns. Fig. 9 (lower plot) shows the normalized difference between the real and forecasted load values. The comparatively homogeneous and symmetrical distribution of the error around the zero point, with no apparent systematic deviation, implies that the given model does not have the issue of systematic bias and behaves predictably.

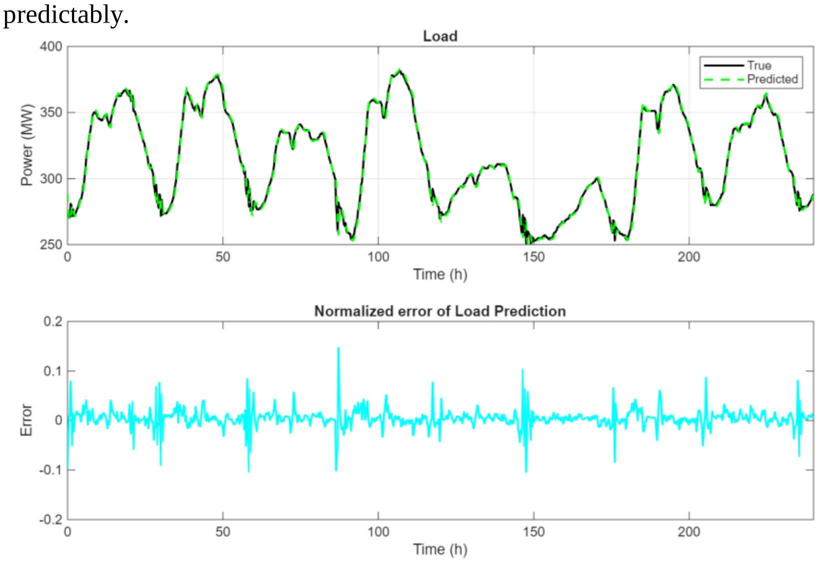

Fig. 9. Performance evaluation of the proposed method in load demand forecasting.

In Figure 10, a statistical evaluation of the association between the outputs of the suggested model and the real load demand values has been given. The fitted linear regression curve has the value of correlation coefficient R = 0.99636, which is very high and the relationship between the predicted and actual values is very strong. This large regression coefficient is a good indication of the reliability and stability of the model in replicating the complicated patterns in energy consumption behavior. The overall result of the regression proves that the suggested model has a high predictive power across various loads and is statistically consistent with actual data. These features indicate the effectiveness of the presented approach in real-life energy consumption prediction and demand-side energy management tasks.

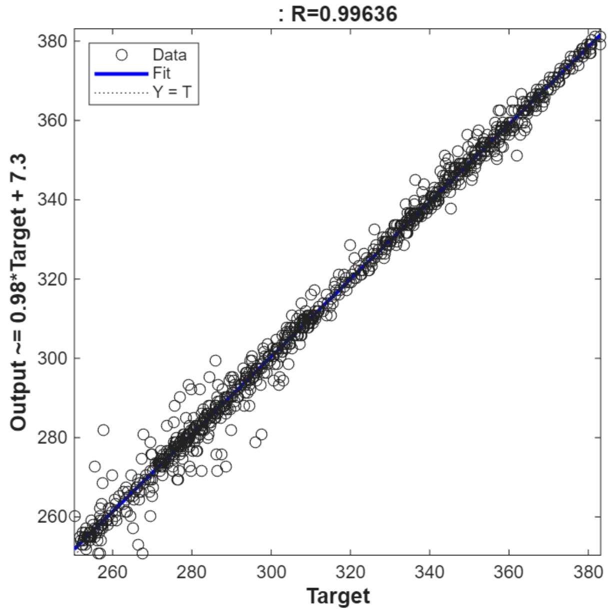

Fig. 10. Regression curve between predicted and actual load demand values

6) Performance Evaluation of the Proposed Method in Day-Ahead Price (DAP) Forecasting

Figure 11 describes the effectiveness of the suggested model in predicting the day-ahead electricity price (DAP). The comparison of the model predictions with the real price values indicates that the suggested strategy can reproduce the general direction of prices as well as local price variations. This is especially relevant in competitive electricity markets, where minor changes in prices can have a significant effect on operational and economic decision-making. The other significant aspect of the suggested model is that it closely captures repeated patterns and nonlinear trends in the price time series, which means that it can be trained to learn more complex time- and economic-related properties of electricity price signals. Fig. 11 (lower half) illustrates the normalized error in the prediction between the predicted and actual value of price. The small magnitude and equal distribution of the errors, particularly in the highly volatile periods of prices, prove that the model is stable and does not show bias toward certain prices. These characteristics emphasize the applicability and usefulness of the suggested approach for real-world price forecasting and market-oriented decision support systems.

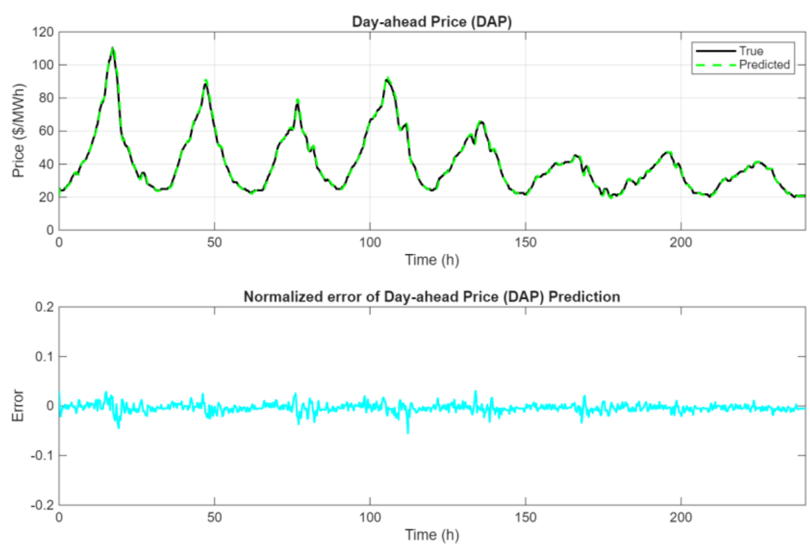

Fig. 11. Performance evaluation of the proposed method in day-ahead electricity price forecasting.

Figure 12 studies the statistical correlation of the outputs predicted by the model and the actual values of the day-ahead price. The fitted regression line has a very high correlation coefficient R = 0.99857, which indicates an excellent level of accuracy of the proposed model in the reconstruction of the actual market prices. This finding suggests that the model not only reflects the overall dynamics of prices but also gives accurate estimates of absolute prices. The regression line intercept value is about +0.91, which is negligible compared to the normal price range in electricity markets, thus showing once again that the proposed model is not plagued by the problem of large systematic bias. Such a high degree of accuracy in day-ahead price forecasting, which is a volatile and uncertain task in itself, proves the strength and efficiency of the offered solution when applied in real-time and operational conditions. On the whole, the results of the analyses provided in Figs. 11 and 12 prove that the suggested model has an extremely high forecasting precision, high adaptability to extreme market changes, and impartial forecasting. These features enable it to be a potent and effective instrument of maximizing bidding plans and assist in the decision-making procedure within day-ahead electricity markets.

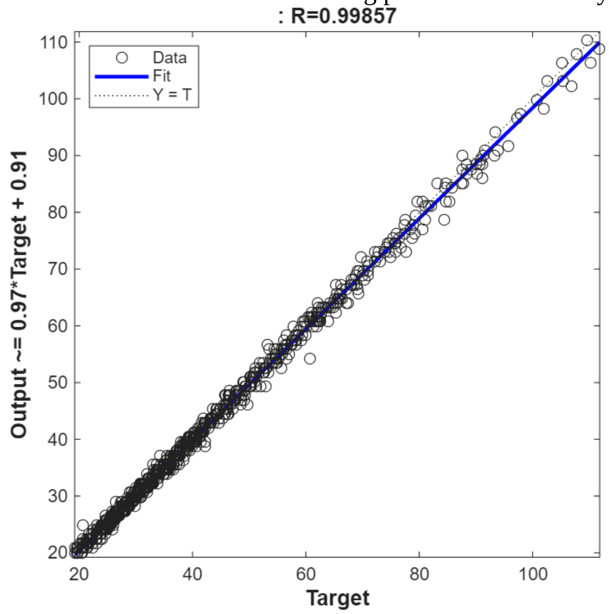

Fig. 12. Regression curve between predicted and actual day-ahead electricity price values.

7) Comparison of Results

Table 1 is a quantitative comparison of the forecasting performance of the suggested IGOA-BiLSTM method to a collection of benchmark and state-of-the-art techniques in power prediction in photovoltaic (PV) and wind turbine (WT). Three popular performance measures are used to do the evaluation, which are the coefficient of determination R^2, Mean Square Error (MSE), and Root Mean Square Error (RMSE), representing respectively the accuracy of the prediction made, the dispersion of errors, and the ability of the model to account for the variability of actual values. Based on the published findings, the initial LSTM-based models have rather low performance. However, the addition of optimization and hybrid methods including PSO, GA, and AWPSO results in a slow increase in the evaluation indices. Moreover, more sophisticated architectures, such as GRU-BiLSTM and GAN-based models, lead to a significant decrease in the value of the RMSE and MSE and a similar increase in the coefficient of determination in comparison with the classical methods. The proposed IGOA-BiLSTM has the best performance in PV and WT power forecasting conditions among all the considered methods. In particular, it yields the best RMSE and MSE as well as the largest value of R^2(equivalent to 0.99) in both renewable energy systems, indicating its good prediction and high generalization. It is further validated as compared to more sophisticated techniques like the IDAOA-LSTM, that the combination of both the bidirectional BiLSTM architecture and the IGOA optimization algorithm is very critical in the optimal capture of the complex temporal dependencies and optimal tuning of the network parameters. In general, the findings presented in Table 1 definitely demonstrate the excellence of the proposed IGOA-BiLSTM model compared to the chosen benchmark models and its high efficiency in the correct prediction of the power output of photovoltaic and wind energy systems.

Table 1. Performance analysis for Renewable Power Prediction (Photovoltaic and Wind Turbines Power)

Method	**WTs**			**PVs**		
	R^2	RMSE	MSE	R^2	RMSE	MSE
GA-AWPSO-LSTM-GAM [23]	0.91	0.064	0.037	0.96	0.055	0.031
GRU-BiLSTM [24]	0.89	0.067	0.039	0.94	0.058	0.034
LSTM [25]	0.80	0.112	0.077	0.85	0.099	0.065
LSTM-GAM [25]	0.83	0.083	0.051	0.89	0.071	0.047
PSO-LSTM [25]	0.86	0.073	0.045	0.91	0.066	0.042
GA-AWPSO-LSTM [25]	0.88	0.071	0.042	0.93	0.062	0.039
GAN [25]	0.89	0.069	0.040	0.94	0.057	0.033

IDAOA-LSTM [25]	0.98	0.021	0.0004	0.98	0.031	0.0009
IGOA-BiLSTM (Proposed)	0.99	0.015	0.0002	0.99	0.019	0.0003

The quantitative assessment of the proposed IGOA-BiLSTM model is devoted to two other and more critical forecasting tasks, i.e., microgrid load demand (Load) prediction and day-ahead electricity price (DAP) forecasting, which are presented in Table 2. To make performance comparisons, the use of the same three standard measures (RMSE, MSE and R^2) is used. The proposed model is contrasted with a group of reference methods, such as LSTM, LSTM-GAM, PSO-LSTM, GA-AWPSO-LSTM, GRU-BiLSTM, GAN, GA-AWPSO-LSTM-GAM, and IDAOA-LSTM. The findings reveal that the suggested IGOA-BiLSTM model is always more effective than all other existing approaches both in Load and DAP forecasting. The model is very effective in improving prediction accuracy by obtaining the lowest values of RMSE and MSE and the highest (R^2). As an example, in the Load forecasting task, the value of RMSE is 0.019 and the value of (R^2) is 0.99, which indicates that the model is highly able to accurately model the actual consumption trends. In the same way, in the DAP forecasting experiment, the suggested approach has the lowest RMSE of 0.009 and the lowest MSE of 0.00008, compared to all other methods, and the (R^2) value of 0.99 indicates an excellent fit. Overall, the discussion of Table 2 indicates that the IGOA optimization algorithm together with the BiLSTM structure helps to significantly increase the accuracy of prediction of both tasks. These findings make the proposed IGOA-BiLSTM a strong, consistent, and dependable framework to predict the use of the model in energy systems, such as load management and electricity market analysis.

Table 2. Performance analysis for Micro Grid Load and Daily-ahead Price (DAP)

Method	**DAP**			**LOAD**		
	R^2	RMSE	MSE	R^2	RMSE	MSE
GA-AWPSO-LSTM-GAM [23]	0.95	0.039	0.028	0.97	0.029	0.021
GRU-BiLSTM [24]	0.93	0.049	0.037	0.96	0.035	0.024
LSTM [25]	0.85	0.066	0.047	0.84	0.058	0.043
LSTM-GAM [25]	0.88	0.059	0.042	0.86	0.67	0.049
PSO-LSTM [25]	0.91	0.058	0.042	0.91	0.060	0.045
GA-AWPSO-LSTM [25]	0.93	0.052	0.040	0.84	0.043	0.033
GAN [25]	0.94	0.041	0.033	0.95	0.039	0.030
IDAOA-LSTM [25]	0.99	0.013	0.0001	0.99	0.024	0.0005
IGOA-BiLSTM (Proposed)	0.99	0.009	0.00008	0.99	0.019	0.0003

To further evaluate the computational efficiency of the proposed Improved Grasshopper Optimization Algorithm (IGOA), a comparative analysis of runtime and convergence speed was conducted against standard GOA and IDAOA under identical hardware configurations. While standard GOA converged in ~45 iterations with an average training runtime of 32 minutes, IDAOA required ~35 iterations and 28 minutes. In contrast, IGOA demonstrated superior computational efficiency by achieving optimal convergence within ~25 iterations and reducing the total training runtime to ~25 minutes, with an inference time of only ~3 seconds. The integration of the dynamic inertia weight and adaptive mutation dynamic in IGOA not only prevented premature convergence but also significantly accelerated the search mechanism towards global optima, validating its suitability for real-time energy management applications, as shown in Table 3.

Table 3. Comparative analysis of the computational overhead (runtime and convergence speed) between standard GOA, IDAOA, and the proposed IGOA

Algorithm	Average Convergence (Iterations)	Total Training Time (min)	Inference Time (s)
Standard GOA	~45	~32	~3.2
IDAOA	~35	~28	~3.1
Proposed IGOA	~25	~25	~3.0

IV. Conclusion

This paper offered an effective and strong forecasting framework of smart microgrid energy management using an Improved Grasshopper Optimization Algorithm-optimized Bidirectional Long Short-Term Memory network (IGOA-BiLSTM). The proposed framework can automatically and effectively tune the key network hyperparameters by combining the temporal learning property of BiLSTM with a more effective metaheuristic optimization strategy, thus substantially enhancing the accuracy of prediction and model generalization. An improved version of the Grasshopper Optimization Algorithm was also developed using a dynamic inertia weight updating mechanism and adaptive triangular mutation strategy. These improvements improved the tradeoff between exploration and exploitation, maintained the diversity of the populations, and minimized the chances of early convergence to local optima. Consequently, the IGOA showed high potential in determining the best BiLSTM hidden layer parameters, and the convergence behavior was steady and better forecasting performance was achieved on all the tasks considered. Hourly real-world experiments were carried out on the PJM West Zone that has photovoltaic power, wind turbine power, load demand, and day-ahead electricity price forecasting. The experiment proved the success of the presented IGOA-BiLSTM model, which was always more successful than traditional LSTM-based models and various state-of-the-art hybrid and optimized models. The proposed approach had coefficients of determination near to 0.99 in all forecasting activities and significantly smaller values of RMSE and MSE, which proves its high accuracy, robustness, and high generalization ability. Further, the computational analysis revealed that the model is efficient enough and can be deployed practically, and the inference time would be low following the optimization phase. In general, the results of this paper demonstrate the possibility of using enhanced deep learning architectures with better metaheuristic optimization algorithms to ensure reliable and accurate predictability in smart energy systems. The suggested IGOA-BiLSTM framework offers an effective and adaptable application for managing the energy of a microgrid, planning renewable resources, and decision support in electricity markets. Future research will be aimed at extending the suggested model to multi-step and probabilistic forecasting, integrating other sources of uncertainty like weather and demand-side flexibility, and testing the framework on larger and more diverse datasets on various microgrid setups. While the current framework was evaluated using an 800-hour continuous dataset (34 days) from early 2020, providing a solid proof-of-concept for the IGOA-BiLSTM model, we acknowledge the potential impact of seasonal variations. Temporal patterns, such as summer peak load demand and seasonal irradiance changes, may differ from winter conditions. Future extensions of this work will involve validating the proposed architecture across full multi-seasonal and multi-year datasets to ensure its generalized robustness against dynamic seasonal shifts.

References

[1] Al-Shetwi, A. Q., Hannan, M. A., Al-Masri, H. M., & Sujod, M. Z. (2025). Latest advancements in smart grid technologies and their transformative role in shaping the power systems of tomorrow: An overview. *Progress in Energy*, *7*(1), 012004.
[2] Meena, S. B., Patil, P. R., Kandharkar, S. R., Hemalatha, N., Khade, A., Dixit, K. K., & Chinthamu, N. (2024). The evolution of smart grid technologies: Integrating renewable energy sources, energy storage, and demand response systems for efficient energy distribution. *Nanotechnology Perceptions*, 1098-1109.
[3] Alhasnawi, B. N., Hashim, H. K., Zanker, M., & Bureš, V. (2025). The rising, applications, challenges, and future prospects of energy in smart grids and smart cities systems. *Energy Conversion and Management: X*, 101162.
[4] Gayen, D., Chatterjee, R., & Roy, S. (2024). A review on environmental impacts of renewable energy for sustainable development. *International Journal of Environmental Science and Technology*, *21*(5), 5285-5310.
[5] Kabeyi, M. J. B., & Olanrewaju, O. A. (2023). Smart grid technologies and application in the sustainable energy transition: a review. *International Journal of Sustainable Energy*, *42*(1), 685-758.
[6] Teixeira, R., Cerveira, A., Pires, E. J. S., & Baptista, J. (2024). Advancing renewable energy forecasting: A comprehensives review of renewable energy forecasting methods. *Energies*, *17*(14), 3480.
[7] Amor, S. B., Möbius, T., & Müsgens, F. (2024). Bridging an energy system model with an ensemble deep-learning approach for electricity price forecasting. *arXiv preprint arXiv:2411.04880*.
[8] Sarathkumar, T. V., Goswami, A. K., Shuaibu, H. A., & Ustun, T. S. (2025). Wind power forecasting and economics of energy arbitrage in electricity market using machine learning techniques. *Discover Sustainability*, *6*(1), 974.
[9] Zia, M. Z. B., Janjua, R. M. A. K., Yasir, M. T., Ali, M. H., Masood, H., & Malik, H. S. (2025). Short-Term Load Forecasting In Smart Grids Using Machine Learning: A Comparative Analysis Of Regression-Based Models. *Spectrum of Engineering Sciences*, *3*(5), 760-772.
[10] Shahinzadeh, H., Sadrarhami, H., Hayati, M. M., Majidi-Gharehnaz, H., Abapour, M., & Gharehpetian, G. B. (2024, February). Review and comparative analysis of deep learning techniques for smart grid load forecasting. In *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)* (pp. 1-9). IEEE.
[11] Zhang, D. (2024). Optimization and research of smart grid load forecasting model based on deep learning. *International Journal of Low-Carbon Technologies*, *19*, 594-602.
[12] Gangwar, H., Parambil Gangadharan, S. M., Daniel, L., Srinivasa Kumar, B., Hariramakrishnan, P., Ramkumar, G., ... & Ifseisi, A. A. (2023). Micro-Grid Renewable Energy Integration and Operational Optimization for Smart Grid Applications Using a Deep Learning. *Electric Power Components and Systems*, 1-16.
[13] Biswal, B., Deb, S., Datta, S., Ustun, T. S., & Cali, U. (2024). Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques. *Energy Reports*, *12*, 3654-3670.
[14] Devi, S. L. R. P. S., & Krishna, C. (2025, February). Towards sustainable energy management: Load forecasting techniques in smart grids. In *AIP Conference Proceedings* (Vol. 3280, No. 1, p. 030005). AIP Publishing LLC.
[15] Onteru, R. R., & Sandeep, V. (2024). An intelligent model for efficient load forecasting and sustainable energy management in sustainable microgrids. *Discover Sustainability*, *5*(1), 170.
[16] Bashir, T., Wang, H., Tahir, M., & Zhang, Y. (2025). Wind and solar power forecasting based on hybrid CNN-ABiLSTM, CNN-transformer-MLP models. *Renewable Energy*, *239*, 122055.
[17] Wang, M., Xia, Y., & Zhang, X. (2024). Research on renewable energy power demand forecasting method based on IWOA-SA-BiLSTM modeling. *Frontiers in Energy Research*, *11*, 1331076.
[18] khan, S., Mazhar, T., Shahzad, T., Waheed, W., Waheed, A., Saeed, M. M., & Hamam, H. (2024). Optimizing deep neural network architectures for renewable energy forecasting. *Discover Sustainability*, *5*(1), 394.
[19] Xiong, B., Zhang, L., Hu, Y., Fang, F., Liu, Q., & Cheng, L. (2025). Deep reinforcement learning for optimal microgrid energy management with renewable energy and electric vehicle integration. *Applied Soft Computing*, 113180.

[20] Ramya, M., & Devaraju, T. (2025, March). Dynamic Scheduling and Islanding Energy Management in microgrids using Deep Learning and Adaptive Reinforcement Learning Techniques. In *2025 7th International Conference on Intelligent Sustainable Systems (ICISS)* (pp. 1678-1682). IEEE.
[21] Sarmadi, B. K., Shayeghi, H., Seyedshenava, S., & Shafie-khah, M. (2025). Multiple microgrids intelligent energy management with capacity constraint using hybrid deep neural network and reinforcement learning. *International Journal of Electrical Power & Energy Systems*, *172*, 111179.
[22] Boopathi, G., & Palanisamy, K. (2025). Energy management in hybrid renewable microgrids using an optimal deep learning model with attention-based transfer learning mechanism. *Waves in Random and Complex Media*, 1-30.
[23] Kim, H. J., and M. K. Kim. "A novel deep learning-based forecasting model optimized by heuristic algorithm for energy management of microgrid." *Applied Energy* 332 (2023): 120525.
[24] Alabi, Tobi Michael, Lin Lu, and Zaiyue Yang. "Data-driven optimal scheduling of multi-energy system virtual power plant (MEVPP) incorporating carbon capture system (CCS), electric vehicle flexibility, and clean energy marketer (CEM) strategy." *Applied energy* 314 (2022): 118997.
[25] Almousawi, Ali Q., Nabil Jalil Aklo, and Zaid Alhadrawi. "An efficient novel deep learning-based predicting model optimized by an improved DAOA algorithm for microgrid energy management." *Future Technology* 5.1 (2026): 303-313.