Received: 8 December 2022 | Revised: 2 February 2023 | Accepted: 17 March 2023

IET Renewable Power Generation

DOI: 10.1049/rpg2.12727

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A Deep Learning-to-learning Based Control system for renewable microgrids

Hossein Mohammadi[1] | **Shiva Jokar**[2] | **Mojtaba Mohammadi**[3] | **Abdollah Kavousi Fard**[3] | **Morteza Dabbaghjamanesh**[4] | **Mazaher Karimi**[5]

[1]Department of Electrical and Electronics Engineering, Darion Branch, Islamic Azad University, Shiraz, Iran

[2]Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

[3]Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran

[4]Smart Power Tech LLC, Dallas, Texas, USA

[5]School of Technology and Innovations, University of Vaasa, Vaasa, Finland

Correspondence
Abdollah Kavousi Fard, Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz, Iran.
Email: kavousi@sutech.ac.ir

Abstract

In terms of microgrids (MGs) operation, optimal control and management are vital issues that must be addressed carefully. This paper proposes a practical framework for the optimal energy management and control of renewable MGs considering energy storage (ES) devices, wind turbines, and microturbines. Due to the non-linearity and complexity of operation problems in MGs, it is vital to use an accurate and robust optimization technique to control the power flow of units efficiently. To this end, in the proposed framework, teacher learning-based optimization (TLBO) is utilized to solve the power flow dispatch in the system efficiently. Moreover, a novel hybrid deep learning model based on principal component analysis (PCA), convolutional neural networks (CNN), and bidirectional long short-term memory (BLSTM) is proposed to address the short-term wind power forecasting problem. The feasibility and performance of the proposed framework and the effect of wind power forecasting on operation efficiency are examined using the IEEE 33-bus test system. Also, the Australian Woolnorth wind site data is utilized as a real-world dataset to evaluate the performance of the forecasting model. The results show that the proposed framework can be used to schedule MGs in the best way possible.

1 | INTRODUCTION

Smart grids, which include multiple microgrids (MGs) at the distribution level, are a significant interest in future electrical grids. In to reduce costs and have a reliable operation in smart grids, MGs' energy management and optimal scheduling must be considered [1]. Generally, there are three classes of MGs in terms of voltage: AC MGs, DC MGs, and hybrid AC-DC MGs. Given that MGs are the link between the end-users and utilities, energy management is one of the main issues in these systems. MGs' operation is categorized into two different modes: Grid-connected and isolated. Isolated MGs can supply small loads and maintain the system's performance in the event of a disturbance in the upstream grid. However, one of the main obstacles to the isolated method is that they are not connected to the utility grid. Therefore, loads must be supplied using only the energy provided by distributed generation units, which is the greatest challenge to manage. On the other hand, MG is connected to the utility grid through a common coupling point in the grid-connected method. This allows power to be exchanged with the utility grid for costs and benefits. Renewable energy sources (RESs) are among the technologies that are becoming more popular due to environmental conditions, greenhouse gas emissions, and limited fossil fuels. Integrating RES and ESs with power grids can cause positive changes and benefits. That is why MGs are so important as the next generation of the power grid. One of the most critical issues associated with their operation is their optimal energy management and scheduling. Today, wind turbines have become one of the most popular technologies. Due to the many benefits of this technology, scientists have done much research to eliminate the challenges in this technology, including increasing the accuracy of forecasting in the system as much as possible. In this regard, wind power forecasting is one of the main topics of this study.

In recent years, MGs' energy management and optimal scheduling have received considerable attention. Authors in [2] have proposed a probabilistic energy management scheme based on the crow search method to minimize the cost of

the hybrid MGs. That paper has focused on using unscented transform to deal with uncertain variables. In [3], an energy management scheme is developed for combined heat power (CHP) based MGs and has formulated the operation of these systems as a multi-objective optimization problem. Authors in [4] suggested the 2 m point estimation algorithm for MGs and developed a stochastic management framework. Authors in [5] presented a management plan considering electric vehicles and ESs. Paper [6] proposes an energy management method for combined heat power-based MGs that considers demand response technology and various generating units. In that paper, demand response is utilized for improving energy management and generative adversarial nets for renewable power forecasting. A machine learning-based energy management system for optimal scheduling of hybrid AC-DC MGs has been suggested in [7]. Ref. [8] has investigated the components and basics of hybrid microgrids (HMGs). In that paper, the energy management of CHP-based MGs is formulated as a multi-objective optimization problem, and different types of units and demand response technology, are utilized. In Article [9], different generation units and uncertainties in the system are considered, and a statistical energy policy using odourless conversion is proposed. Article [10] has proposed a stochastic energy management method for AC-DC MGs. Various generating units are addressed in that work and system uncertainties are recorded using the unscented transform. The paper [11] proposed an energy management method for CHP-based MGs that considers the need for reaction power and RESs. Ref. [5] has suggested a management scheme for HMGs that thought CHP models, plug-in electric vehicles, ESs, RESs, and a strategy for feeder reconfiguration. A secure distributed cloud-fog-based optimal scheduling system for MGs was presented in [12]. A two-layer energy management method for MGs with high penetration of RESs and ESs considering battery degradation is proposed in [13]. Ref. [14] provides an efficient energy management technique for renewable MGs considering ESs and tidal power generation units. While the above works addressed essential topics regarding the energy management of MGs, the research in this area is still in its infancy. In this regard, this paper proposes robust energy management for MGs considering RESs.

Another vital concept in modern power grids is renewable energy. This concept, which is essential to power grids creates several challenges including power forecasting. Due to the crucial role of wind power in the optimal scheduling and operation of power grids, wind power forecasting has attracted many researchers in recent years. In the following, several studies about wind power forecasting are investigated. Ref. [15] has suggested a three-layer backpropagation neural network prediction method for one-step-ahead wind speed prediction. In [16], a prediction technique based on the wavelet transform for wind power forecasting is proposed. Future wind data is forecasted using this strategy by analyzing hypothetical symptoms. In [6], a renewable forecasting model based on generative adversarial networks is proposed, and the results of the proposed method are compared with several well-known methods. Ref. [17] has suggested a recurrent neural network-based approach to estimate wind power by learning the temporal relationship of the time series data. Ref. [18] has proposed a forecasting model using long short-term memory (LSTM) to address the gradient vanishing problem. As one of the advanced variations of recurrent neural networks, LSTM can learn the information stored in time series data more successfully. To improve load forecasting accuracy, this paper proposes a novel method based on PCA, CNN, and BLSTM. It is worth noting that the PCA algorithm is used to reduce the dimensions of the data.

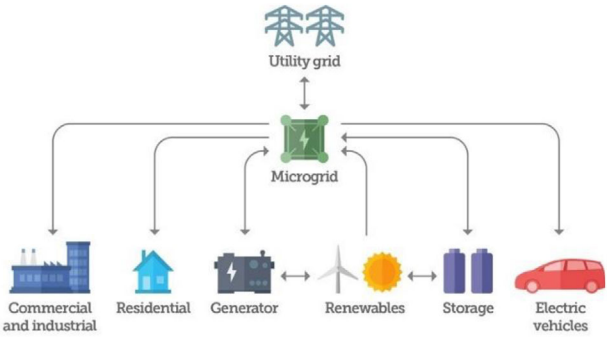

FIGURE 1 Schematic construction of MG's structure.

The main contribution of this work can be summarized as follows:

- Proposing a robust optimal scheduling framework for AC microgrids is presence of ESs and renewable energy sources
- Proposing a novel deep learning method for wind power forecasting considering environmental features.
- Assessing the impact of wind power forecasting in on the efficiency of MG scheduling.

The rest of the paper is organized as follows: In Section 2, the optimal scheduling and operation of the MG are formulated as an optimization problem, and the details of the TLBO are presented. Section 3 is focused on the proposed deep-learning-based forecasting model. The simulation results and conclusion are provided in Sections 4 and 5.

2 | MG OPERATION AND OPTIMAL SCHEDULING

MGs are designed to facilitate electric power supply at the distribution level. Figure 1 illustrates the schematic construction of typical MGs. In smart MGs, physical layer parameters (e.g. voltage, active power, reactive power, etc.) are measured and sent to the MG central control at regular intervals through advanced metering infrastructures (AMI). Each node in these systems is equipped with a smart metering device that communicates with central control through communication channels. MG central control schedules the system's components one day ahead of operation based on the system's characteristics, energy market price data, RESs' forecasted output power, and load demand. In the following, the operation of MG is formulated as a constraint optimization problem, and then the TLBO is described in detail.

2.1 | MG scheduling problem formulation

In this section, the operation of renewable MGs is modeled as a single-objective optimization problem. The main cost objective function incorporates the cost of generation units, ESs, and power sold/ purchased to/ from the utility grid as follows [19]:

$$\begin{aligned} Minh(X) = \sum_{t=1}^{N_T} \Bigg(& \sum_{i=1}^{N_g} [u_i^t P_{Gi}^t B_{Gi}^t + S_{Gi}^{on} \max\{0, u_i^t - u_i^{t-1}\} \\ & + S_{Gi}^{off} \max\{0, u_i^{t-1} - u_i^t\}] \\ & + \sum_{j=1}^{N_S} [u_j^t P_{sj}^t B_{sj}^t + S_{sj}^{on} \max\{0, u_j^t - u_j^{t-1}\} \\ & + S_{sj}^{off} \max\{0, u_j^{t-1} - u_j^t\}] + P_{Grid}^t B_{Grid}^t \Bigg) \end{aligned} \quad (1)$$

The variable X in Equation (1) indicates the optimum operation point of grid components and their binary OFF/ON status as follows:

$$\begin{aligned} X &= [P_g, U_g]_{1\times(2\times n\times N_T)}, n = N_d + N_s + 1; \forall t \in N_T \\ P_g^t &= [P_G^t, P_s^t, P_{Grid}^t]; P_G^t = [P_{G1}^t, P_{G2}^t, \dots, P_{GN_g}^t] \\ U_g^t &= [u_1^t, u_2^t, \dots, u_{N_d}^t]; P_s^t = [P_{s1}^t, P_{s2}^t, \dots, P_{sN_s}^t] \\ P_{Grid}^t &= [P_{Grid}^t], u_k^t \in \{0, 1\} \end{aligned} \quad (2)$$

The primary optimization challenge is solved by taking into account various practical limitations. The balance between consumption and generation in the system is maintained through the following load flow equations:

$$P_j^{inj,t} = \sum_{n=1}^{N_b} V_j^t V_n^t Y_{jn} \cos(\theta_{jn} + \delta_j - \delta_n) \quad (3)$$

$$Q_j^{inj,t} = \sum_{n=1}^{N_b} V_j^t V_n^t Y_{jn} \sin(\theta_{jn} + \delta_j - \delta_n) \quad (4)$$

The restrictions on the generation units, utility grid, etc., are presented in the rest.

$$\begin{aligned} P_{Gi,\min}^t \le P_{Gi}^t \le P_{Gi,\max}^t \\ P_{Grid,\min}^t \le P_{Grid}^t \le P_{Grid,\max}^t \end{aligned} \quad (5)$$

Feeder capacity:

$$|P_i^{Line,t}| \le P_{i,\max}^{Line} \quad (6)$$

Spinning reserve:

$$\sum_{i=1}^{N_d} u_i^t P_{Gi,\max}^t + P_{Grid,\max}^t \ge \sum_{k=1}^{N_{Load}} P_{Load,k}^t + P_{loss}^t + \mathrm{Re}s^t \quad (7)$$

Bus voltage:

$$V_m^{\min} \le V_m^t \le V_m^{\max} \quad (8)$$

Generation units' ramp rate:

$$|P_{Gi}^t - P_{Gi}^{t-1}| < UR_i, DR_i \quad (9)$$

ES constraints:

$$W_{ess}^t = W_{ess}^{t-1} + \eta_{charge} P_{charge} \Delta t - \frac{1}{\eta_{discharge}} P_{discharge} \Delta t \quad (10)$$

$$\begin{cases} W_{ess,\min} \le W_{ess}^t \le W_{ess,\max} \\ P_{charge,t} \le P_{charge,\max} \\ P_{discharge,t} \le P_{discharge,\max} \end{cases} \quad (11)$$

2.2 | TLBO algorithm

As mentioned before, the TLBO is utilized to solve the scheduling problem of MGs and minimize the operation cost. Due to the high ability of evolutionary optimization techniques in solving constraint/non-constraint single/multi-objective optimization problems, these algorithms have attracted much attention in recent years. For instance, authors in [2] used the crow search optimization to solve the optimal scheduling problem of MGs. This paper employs the TLBO algorithm to solve the optimal scheduling problem of MGs. The TLBO algorithm is a population-based algorithm that models the learning process of students in the class [20]. The algorithm consists of two phases: The teacher and student phase. The main idea behind the teacher's phase is that during the learning process, the teacher attempts to move students' mean grades toward the best student. It is worth noting that in each iteration, the teacher is considered the best current solution. The mathematical model of the teacher phase is described as follows:

$$\Delta X_{iT} = r_i (T_i - F_i M_i) \quad (12)$$

$$X_i^{new} = X_i^{old} + \Delta X_{iT} \quad (13)$$

where T_i, M_i, r_i, and F_i present the teacher, the mean grade of students, a random number in the range of [0,1], and a random integer between 1 and 2 in the ith iteration, respectively. The second phase (i.e. the student phase) simulates the learning process among students in which students increase their knowledge by discussing with each other. For two randomly chosen students, the student phase is expressed as follows:

$$\begin{cases} X_i^{new} = X_i^{old} + r_i (X_i - X_j) \;\; if\; F(X_i) < F(X_j) \\ X_j^{new} = X_i^{old} + r_i (X_j - X_i) \;\; if\; F(X_i) > F(X_j) \end{cases} \; i \ne j \quad (14)$$

Note that $F(X)$ represents the objective cost function of students in Equation (15). The advantages of TLBO over other well-known optimization methods (e.g. PSO, CSA) are its

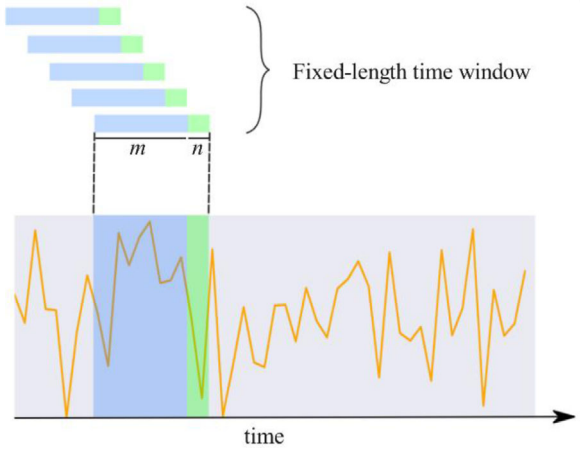

FIGURE 2 Dataset generation stage.

simplicity, two-step movement in search space, and no dependence on regulatory parameters.

3 | WIND POWER FORECASTING MODEL BASED ON PCA-CNN-BLSTM

This section presents the detailed and mathematical model of the different parts of the proposed forecasting method. The proposed method forecasts future wind power using previously measured values and environmental features. In this model, first, the PCA is utilized as a feature selection and dimensionality reduction method, and then the CNN-BLSTM model is trained using the output of the PCA. This section presents three different concepts as follows:

A. Data generation stage
B. PCA
C. CNN-BLSTM model

3.1 | Data generation stage

The first stage of the proposed model is the dataset generation stage. The historical data is first processed at this stage to generate the proper dataset for training the model. Figure 2 shows the dataset generation stage. In this stage, two windows of sizes m and n are utilized to generate the input and the corresponding outputs from the historical data sequence. For instance, if m is 4 and n is 1, the input related to each data sample is its four previous samples. In Figure 2, the data generation stage is depicted.

3.2 | Feature selection using PCA

PCA can reduce the dimensionality of data by decreasing the number of feature vectors. It can be used to improve the computation rate of artificial neural networks by diminishing their calculation times. The procedure is described in detail in the following paragraphs. The PCA simplifies the input features into a complete output element. To separate the similarity between the data vectors and maintain the initial variable details, the PCA mostly discovers a small set of linear combination variables to replace the actual variables.

If i was the temporal variable of observation and j was the environmental variable in an initial dataset $\mathrm{X} = \{x_{11}, \mathrm{x}_{12}, \mathrm{x}_{\mathrm{ij}}, \ldots, x_{\mathrm{mn}}\}$, the matrix can be constructed as:

$$\mathrm{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{15}$$

Stage 1. The normalization matrix (X^*) is produced by normalizing the data set:

$$\mathrm{X}^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{16}$$

Stage 2. The covariance matrix (R) is created by using a linear transformation:

$$\mathrm{R} = \frac{1}{n}(X^*)^T X^* \tag{17}$$

Stage 3. The feature matrix, λ; is obtained by solving $|\lambda \mathrm{I} - \mathrm{R}| = 0$, and finally, it leads to calculating the collected contribution portion, β_i:

$$\beta_i = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{18}$$

A considerable amount of data is within the k main features with a 75–95% contribution, which the n initial variables can deliver.

3.3 | The CNN-BLSTM model

In this section, first, the basic details of the CNN and BLSTM and then the hybrid model of CNN-BLSTM are described in detail. The CNN-BLSTM has been extensively used for video frame prediction, speech recognition, text understanding, etc. The main advantage of this architecture is that the CNN can extract high-level features from data and feed them to the BLSTM, which has an excellent capability for capturing patterns in sequential data like time series. Also, the bidirectional structure of BLSTM allows the model to learn in two directions (forward and backward). Hence, combining two architectures can benefit from the advantages of both models.

3.3.1 | CNN architecture

The CNN architecture was initially proposed for the image classification task. The input of the model is usually a 2D matrix with some channels. The CNN model has a sliding window with some trainable weights called the kernel, which moves on the

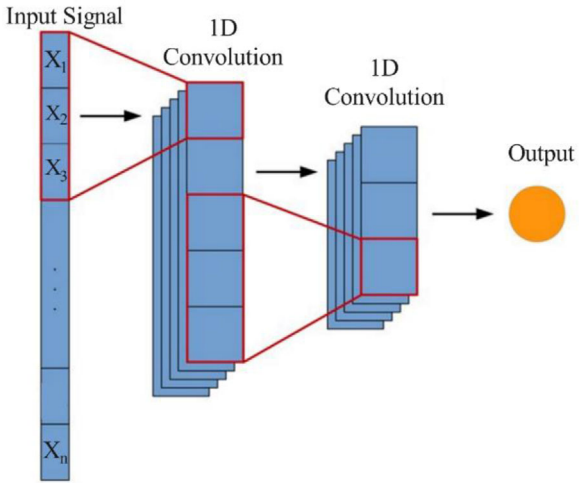

FIGURE 3 Schematic illustration of the 1D convolution.

information and learns the features of the input through the backpropagation process. The shape of the kernel depends on the input. For example, in the case of RGB images, the kernel is 3D. Here, since the input data is a time series, 1D convolutional layers are utilized. In 1D convolutional layer, the kernel has only one dimension and moves on the input data in only one direction. The schematic illustration of the one-dimensional convolutional layer is presented in Figure 3. Additionally, the convolutional neural network has a max pooling layer and a dropout layer to prevent overfitting. It is worth noting that by moving from the input layer to the output layer, the size of the data decreases, and the number of channels increases.

3.3.2 | BLSTM architecture

LSTMs were first proposed to solve the gradient vanishing issue in regular recurrent neural networks (RNN) [21]. The LSTM records and stores information using the cell state and three cell gates. The input and forget gates tell the cell state what information needs to be added or removed. Additionally, the output gate determines which section of the cell state should be output. Note that the LSTM network's training procedure is based on backpropagation [22]. The LSTM's transition equations are as follows:

$$h_t = H(W_{x,h}x_t + W_{h,h}h_{t-1} + b_h) \tag{19}$$

$$y_t = W_{h,y}h_t + b_y \tag{20}$$

The following equations implement the function H:

$$i_t = \partial(W_{i,x}x_t + W_{i,h}h_{t-1} + b_i) \tag{21}$$

$$f_t = \partial(W_{f,x}x_t + W_{f,h}h_{t-1} + b_f) \tag{22}$$

$$\bar{c}_t = \tanh(W_{\bar{c},x}x_t + W_{\bar{c},x}h_{t-1} + b_{\bar{c}}) \tag{23}$$

$$o_t = \partial(W_{o,x}x_t + W_{o,h}h_{t-1} + b_o) \tag{24}$$

$$C_t = f_t \cdot c_{t-1} + i_t \cdot \bar{c}_t \tag{25}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{26}$$

The notion of memory cells with regulating gates enables LSTMs to significantly outperform RNNs in overcoming long-term reliance and vanishing gradients [21]. The learning process calculates the weights that determine whether the cells store or delete data through backpropagation.

Another variant on RNNs is presented in [23], in which a single RNN layer is composed of two RNN blocks simultaneously processing temporal input in opposing directions. At each step, the ultimate output is a sum of the results of RNN blocks. The bidirectional structure could be applied to LSTM to create BLSTM. In the following, the mathematical relation between forwarding and backward processes and the network's output is described:

$$\vec{h}_t = H(W_{x,\vec{h}}x_t + W_{\vec{h},\vec{h}}\vec{h}_{t-1} + b_{\vec{h}}) \tag{27}$$

$$\overleftarrow{h}_t = H(W_{x,\overleftarrow{h}}x_t + W_{\overleftarrow{h},\overleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}) \tag{28}$$

$$y_t = W_{\vec{h},y}\vec{h}_t + W_{\overleftarrow{h},y}\overleftarrow{h}_t + b_y \tag{29}$$

3.3.3 | CNN-BLSTM

In this section, the details of the CNN-LSTM architecture are presented in detail. In order to extract high-level features from the input data, the first layer of the model in the CNN-BLSTM structure is a one-dimensional convolutional layer with several kernels where the shape of the kernels is usually 3×1. The output of this layer goes through the RELU activation function to make the system non-linear. Afterward, to provide an abstract from the representation and reduce the overfitting, the output of the activation function goes through a max pooling layer. Also, the max pooling layer reduces the computation cost by reducing the number of learnable parameters. Several BLSTM layers continue the network. After the BLSTM layers, there is a dropout layer to prevent overfitting. The output of the dropout layer is fed to a fully connected layer before going to the output layer. The size of the output layer depends on the application. Since the wind speed is a regression problem, the model's loss function is mean square error. The structure of the CNN-BLSTM network is depicted in Figure 4.

4 | SIMULATION RESULTS

In this section, the performance and feasibility of the proposed strategies are examined on practical test cases constructed based on the IEEE 33-bus test system. The test system includes

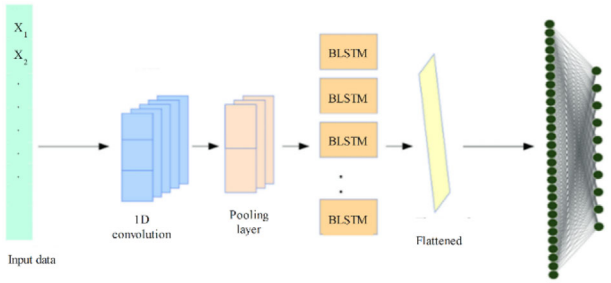

FIGURE 4 The structure of the CNN-BLSTM network.

a wind turbine, two microturbines, and one ES. The complete details of the characteristics of the components and a schematic illustration of the test system can be found in Table 1 and Figure 5 respectively. The test system operates based on the 12 kV voltage level, and the MG is connected to the upstream grid through the point of standard coupling at bus 1. MG and the upstream grid can exchange power based on cost and benefits. In order to respect the idea of green energy, the wind turbine unit is considered non-dispatchable, meaning that all of its generation is consumed without considering the cost-efficiency. This study considers the complete model for transmission lines (impedance). The hourly electricity market price is provided in Figure 9, and the load factor is from [24].

Moreover, the historical data of the Woolnorth wind site [25], which is located at Woolnorth on Cape Grim at the northwest tip of Tasmania, is utilized to evaluate the performance of the proposed wind power forecasting model. It is worth noting that the Woolnorth wind site is one of Australia's most difficult situations for wind power forecasting due to its location on the edge of a cliff exposed to the Southern Ocean. The data set contains the wind turbine's hourly power measurements for five years from 2009 to 2014. The dataset includes five environmental features: Wind speed and direction, temperature, atmospheric pressure, and solar irradiance. In the following, the proposed wind power forecasting model is examined, and then its results are utilized to evaluate the performance of the proposed optimal scheduling strategy.

4.1 | Wind power forecasting results

As mentioned before, the hourly historical data of the Woolnorth wind site is utilized to evaluate the performance of the proposed forecasting model. In the first stage, the PCA algorithm is used to identify less important features and decrease the dimensions of the data. The total contribution rate and the cumulative contribution rate of the input features corresponding to eigenvalues are depicted in Figure 6.

As shown in Figure 6, the first two features (i.e. wind speed and direction) include 86% of the information. This value goes

TABLE 1 Characteristics of the components of the system.

Type	Min power (kW)	Max power (kW)	Bid ($/kWh)	Startup/shutdown cost ($ct)	Ramp rate	Location (Bus number)
Micro-turbine 1	100	1300	0.645	75	220	12
Micro-turbine 2	90	1100	0.675	70	180	25
Wind-turbine 1	0	2000	1.073	0	–	30
ES	−250	250	0.318	0	–	18

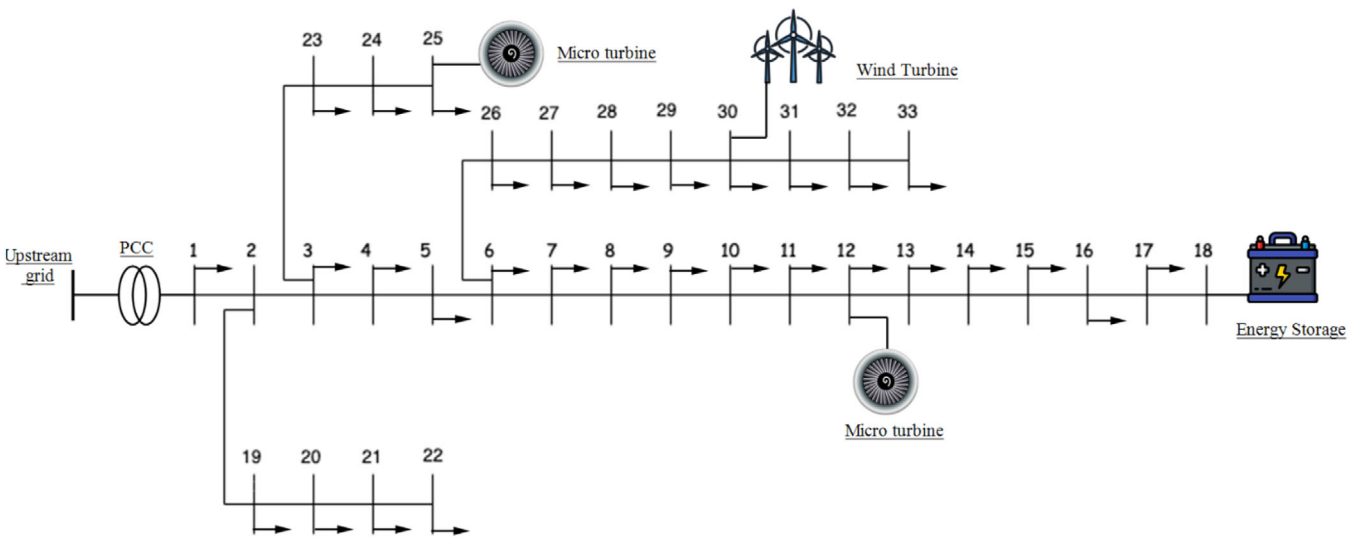

FIGURE 5 Schematic illustration of the test system.

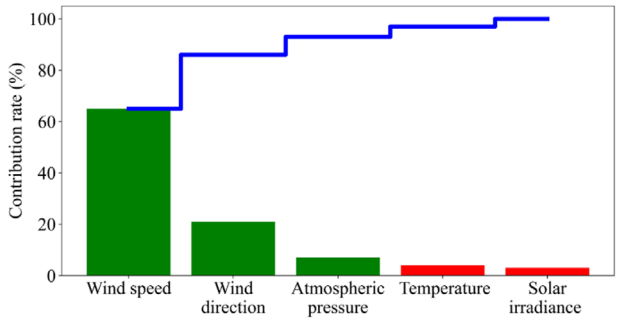

FIGURE 6 Absolute contribution rate and cumulative contribution rate of different features.

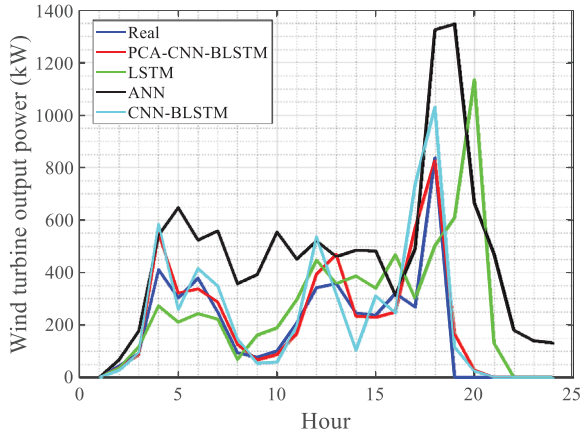

FIGURE 7 Regression diagram of wind data.

up to 93% when the atmospheric pressure is included. So the other features only contribute 7% of the data. In this regard, only wind speed, wind direction, and atmospheric pressure are considered from the output of the PCA. In the following, this policy's results are compared with the case that all features are included.

After the feature selection stage using PCA, it is time for the data generation stage. The dataset generation stage is performed on data considering $m = 48$ and $n = 24$, meaning that the proposed model can simultaneously forecast the next 24 h of data. After the input data is generated, it is time to train the model. For this purpose, the data is divided into test and training datasets where the split ratio is 80%. The proposed model is implemented in Python 3.8 using Tensorflow.Keras library. The optimizer of the model is ADAM, and the number of epochs is 250. The network architecture includes two 1D convolutional layers of size 3 with RELU activation functions, followed by one max pooling layer and a dropout layer with a 0.3 rate. Afterward, there are two BLSTM layers with 128 cells, followed by one dropout layer with a ratio of 0.3 and one fully connected layer that maps the information to the output. The activation function of the last layer is linear. Since the whole problem is a regression problem, the model's loss function is the mean square error.

Figure 7 shows the regression diagram and compares the actual output power with the predicted values for 24 h related

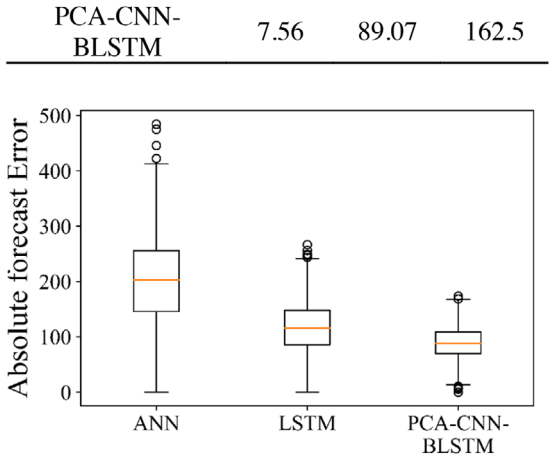

FIGURE 8 Boxplot of the absolute error of different methods.

TABLE 2 Prediction metrics.

Method	MAPE (%)	MAE	RMSE
ANN	17.868	209.36	362.09
CNN-BLSTM	10.95	117.3	187.3
LSTM	12.345	115.91	198.88
PCA-CNN-BLSTM	7.56	89.07	162.5

to 22/11/2009. Note that for better evaluation, the proposed methodology is compared with two well-known methods called Artificial neural network (ANN) and LSTM.

As can be seen from Figure 7, the proposed PCA-CNN-BLSTM model has the best performance. The superiority of the proposed method over other models can be seen at sharp edges. This shows the effectiveness and positive role of PCA and the deep structure of the CNN-BLSTM in forecasting. In comparison with other models, it can be seen in Figure 7 that the actual output power of the wind turbine at the 5th hour is 303.76, and the forecasted value of PCA-BLSTM is 321.41, where the output of LSTM and ANN are 210.84 and 647.68, respectively.

For better comparison, the boxplot of the absolute forecast error of baseline models is presented in Figure 8. As seen from that figure, not only does the PCA-CNN-LSTM have a lower mean absolute error, but also the diversity of the error is lower for this model. Additionally, the regression metrics of the models are presented in Table 2. The regression metrics in this study are mean absolute percentage error (MAPE), root mean square error (RMSE), and mean absolute error (MAE) [6]. By comparing the regression metrics of the CNN-BLSTM with PCA-CNN-BLSTM, we can see the PCA's effectiveness in improving the forecasting model's accuracy.

4.1.1 | MG optimal scheduling

This section is focused on the optimal scheduling of the MG for 24 h. Four different scenarios are looked at to show how

TABLE 3 Power loss, maximum voltage deviation, and total cost of the grid for different scenarios.

Case study	Power loss (kW)	Maximum voltage deviation (pu)	Total cost ($)
Scenario 1	19837	0.0042	55440
Scenario 2	20945	0.0053	55316
Scenario 3	19777	0.0029	56758
Scenario 4	19204	0.0023	58023

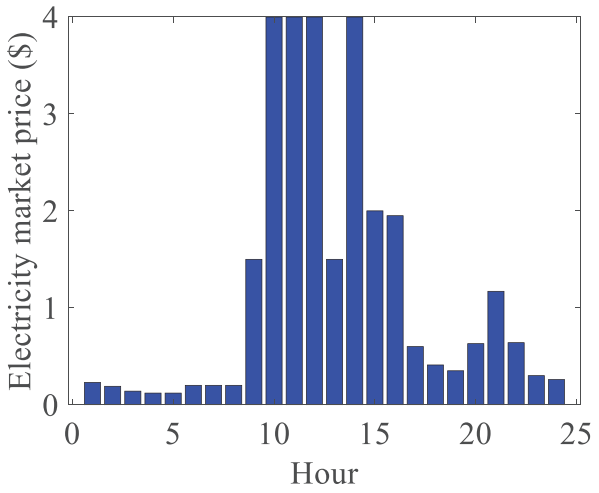

FIGURE 9 Hourly electricity market price.

the wind power forecasting model affects the accuracy and efficiency of the best scheduling:

- Scenario 1: optimal scheduling is performed using actual wind data.
- Scenario 2: optimal scheduling is performed using PCA-CNN-BLSTM results.
- Scenario 3: optimal scheduling is performed using LSTM results.
- Scenario 4: optimal scheduling is performed using ANN results.

The power loss, maximum voltage deviation, and total cost of the grid for all scenarios are presented in Table 3. As seen from that table, the grid's total cost for scenario 2 is minuscule different from scenario 1 (124$). These values for scenarios 3 and 4 are (1318$) and (2583$). This shows the high impact of the wind power forecasting model on the optimal operation and efficiency of the grid. It is worth noting that the maximum voltage deviation in Table 3 is the highest value among all buses during the operation day. As can be seen, the maximum voltage deviation for all scenarios is less than the corresponding threshold. Since scenario 2 had the best results, only the results of scenario two are analyzed.

According to Figure 9 and Table 1, the electricity price varies at different day hours, and the generation cost of different units are not the same. Therefore, purchasing power from the upstream grid and turning off expensive units is more beneficial in the early hours of the day. As can be seen from Table 4, dispatchable units (i.e. microturbines 1 and 2) are turned off in the early hours of the day when the energy market price is low. On the other hand, in the middle of the day, when the electricity price is very high, all generation units operate at their maximum capacity to keep the operation cost as low as possible. Table 4 shows the MG optimal scheduling for 24 h using PCA-CNN-BLSTM results considering 1 h intervals.

TABLE 4 MG optimal scheduling for 24 h.

Hour	ES	Micro turbine 1	Micro turbine 2	Wind turbine
1	−250	0	0	0.0
2	−250	0	0	36.5
3	−250	0	0	87.4
4	−250	0	0	549.5
5	−250	0	180	321.4
6	−250	220	360	337.7
7	−250	440	540	287.4
8	−250	660	720	125.8
9	250	880	900	66.5
10	250	1100	1080	87.3
11	250	1300	1100	167.2
12	250	1300	1100	394.0
13	250	1300	1100	469.2
14	250	1300	1100	233.2
15	250	1300	1100	228.7
16	250	1300	1100	249.3
17	−250	1080	1100	369.5
18	250	860	920	827.0
19	−250	640	740	165.8
20	0	440	920	27.6
21	250	660	740	0.0
22	−250	440	560	0.0
23	−250	220	380	0.0
24	−250	0	200	0.0

Another beneficial policy for the grid would be storing energy in ES at low electricity prices and injecting that energy into the grid in the middle of the day when the energy price is high. As can be seen from Table 4 and Figure 9, the ES is charged at the low-price hours (i.e. 12:00 AM to 9:00 AM and 5:00 PM to 7:00 PM) and discharged at high price hours. This policy could decrease the total operation cost of the system significantly. Therefore, the presence of ESs in the grid can improve the system's efficiency.

According to Table 4, the amount of power exchanged with the upstream grid has a reverse relation with the energy market price, meaning that in low price hours, the priority is given to the upstream grid rather than generation units. On the other hand,

in the middle of the day, when the electricity price is high, most of the grid's demand is supplied through distributed generation units.

5 | CONCLUSION

This paper focuses on wind power forecasting and the optimal management and control of renewable MGs in the realistic environment. For robust and efficient scheduling of MG, an optimal framework based on TLBO was proposed. The first part of the proposed methodology, which is for optimal scheduling, takes the practical limitation of the system, such as ramp rate etc., into account and presents an optimal framework for the efficient operation of MGs. The performance of the proposed methodology was tested on the IEE 33-bus test system, including three generation units and an ES. The results showed that the proposed methodology could efficiently minimize operation costs by managing the generation units and batteries. Also, the impact of wind power forecasting accuracy on the efficiency of the grid was examined. The second part of the proposed framework was focused on wind power forecasting using a novel method based on PCA-CNN-BLSTM. The performance of the proposed model was evaluated using a real-world test case. The results showed the acceptable performance of the proposed model and its priority over two other well-known forecasting models.

NOMENCLATURE

σ	Logistic sigmoid function
$\eta_{discharge,\ charge}$	ES charging/discharging efficiency
$B_{Gi}^{\ t}$ / $B_{si}^{\ t}$	Cost of generation units /ESs
$B_{Grid}^{\ t}$	Cost of upstream grid power
$b_i, b_f, b_o, b_h, \ldots$	Bias vectors
c_t	Cell state
DR_i/UR_i	ramp down/up rate
f_t	Forget gate
$h(X)$	The cost function
i_t	Input gate at time t
N_b	Number busses
N_d	Number of generation units in the MG
N_{Load}	number of loads
N_s	Number of ESs in the grid
N_s	Number of batteries in MG
N_T	Number of time intervals
o_t	Output gate
$P_{charge/discharge}$	Permitted charge/discharge rate within a definite period Δt
$P_{charge/discharge,max}$	The maximum value of the charging/discharging rate within a defined period of time Δt
$P_{Gi}^{\ min}$ / $P_{Gi}^{\ max}$	Max/min ith generation units' output power
$P_{Grid,min}^{\ t}$ / $P_{Grid,max}^{\ t}$	Min/max power of the upstream grid
$P_{Grid}^{\ t}$	The power exchanged with the upstream grid
$P_{i,max}^{\ line,t}$	Feeder capacity
$P_j^{inj,t}$/ $Q_j^{inj,t}$	Injected active/reactive power
$P_{loss}^{\ t}$	Total power loss
$P_{si}^{\ t}$ / $P_{Gi}^{\ t}$	Power of batteries/generation units
RES^t	The spinning reserve
$S_{Gi}^{\ on}/S_{Gi}^{\ off}$	Generation units ' Startup/shutdown cost
$S_{sj}^{\ on}/S_{sj}^{\ off}$	ESs' Startup/shutdown cost
$u_i^{\ t}$	ON/OFF Status of the ith generation units/ES
V/δ	Magnitude/phase of the voltage
$V_{min}^{\ i}/V_{max}^{\ i}$	Voltage limitations of the bus
$W_{ess,\ max/min}$	Capacity of ES
$W_{ess}^{\ t}$	battery
$W_{f,x}, W_{i,x}, W_{o,h}, \ldots$	Weighting matrices
X	Control variables
Y/θ	line impedance magnitude/phase
y_t	LSTM cells' output at time t

AUTHOR CONTRIBUTIONS

Hossein Mohammadi: Conceptualization; Data curation; Formal analysis; Resources; Software; Supervision. Shiva Jokar: Validation; Visualization; Writing—original draft; Writing—review and editing. Mojtaba Mohammadi: Formal analysis; Investigation; Methodology; Project administration. Mazaher Karimi: Resources; Software; Writing—original draft; Writing—review and editing.

FUNDING INFORMATION

The authors received no specific funding for this article.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data would be provided as per request.

ORCID

Abdollah Kavousi Fard https://orcid.org/0000-0001-8316-5588

Morteza Dabbaghjamanesh https://orcid.org/0000-0001-5224-9973

Mazaher Karimi https://orcid.org/0000-0003-2145-4936

REFERENCES

1. Mohammadi, M., Kavousi-Fard, A., Dehghani, M., Karimi, M., Loia, V., Alhelou, H.H., Siano, P.: Reinforcing data integrity in renewable hybrid AC-DC microgrids from social-economic perspectives. ACM Trans. Sens. Networks 19(2), 1–19 (2022)
2. Papari, B., et al.: Effective energy management of hybrid AC–DC microgrids with storage devices. IEEE Trans. Smart Grid 10(1), 193–203 (2017)
3. Bornapour, M., et al.: Probabilistic optimal coordinated planning of molten carbonate fuel cell-CHP and RESs in microgrids considering hydrogen storage with point estimate method. Energy Convers. Manage. 206, 112495 (2020)
4. Baziar, A., Kavousi-Fard, A.: Considering uncertainty in the optimal energy management of renewable micro-grids including storage devices. Renewable Energy 59, 158–166 (2013)

5. Gong, X., et al.: A secured energy management architecture for smart hybrid microgrids considering PEM-fuel cell and electric vehicles. IEEE Access 8, 47807–47823 (2020)
6. Mobtahej, M., et al.: Effective demand response and GANs for optimal constraint unit commitment in solar-tidal based microgrids. IET Renewable Power Gener. (2021)
7. Lei, M., Mohammadi, M.: Hybrid machine learning based energy policy and management in the renewable-based microgrids considering hybrid electric vehicle charging demand. Int. J. Electr. Power Energy Syst. 128, 106702 (2021)
8. Pourbehzadi, M., et al.: Optimal operation of hybrid AC/DC microgrids under uncertainty of renewable energy resources: A comprehensive review. Int. J. Electr. Power Energy Syst. 109, 139–159 (2019)
9. Hossain, M.A., et al.: Energy management of community energy storage in grid-connected microgrid under uncertain real-time prices. Sustainable Cities Soc. 66, 102658 (2021)
10. Pourbehzadi, M., et al.: Stochastic energy management in renewable-based microgrids under correlated environment. In: *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, Madrid, Spain (2020)
11. Gitizadeh, M., Farhadi, S., Safarloo, S.: Multi-objective energy management of CHP-based microgrid considering demand response programs. In: *2014 Smart Grid Conference*, Tehran, Iran (SGC) (2014)
12. Tajalli, S.Z., et al.: DoS-resilient distributed optimal scheduling in a fog supporting IIoT-based smart microgrid. IEEE Trans. Ind. Appl. 56(3), 2968–2977 (2020)
13. Ju, C., et al.: A two-layer energy management system for microgrids with hybrid energy storage considering degradation costs. IEEE Trans. Smart Grid 9(6), 6047–6057 (2017)
14. Kavousi-Fard, A., Su, W., Jin, T.: A Machine-Learning-Based Cyber Attack Detection Model for Wireless Sensor Networks in Microgrids. IEEE Trans. Ind. Inf. 17(1), 650–658 (2020)
15. Xiao, Y.-S., Wang, W.-Q., Huo, X.-P.: Study on the time-series wind speed forecasting of the wind farm based on neural networks. Energy Conserv. Technol. 25(2), 106–108 (2007)
16. Wang, L.-J., et al.: Short-term power prediction of a wind farm based on wavelet analysis. Proc. CSEE 29(28), 30–33 (2009)
17. Singh, S., Bhatti, T., Kothari, D.: Wind power estimation using artificial neural network. J. Energy Eng. 133(1), 46–52 (2007)
18. Schmidhuber, J., Hochreiter, S.: Long short-term memory. Neural Comput. 9(8), 1735–1780 (1997)
19. Mohammadi, M., et al.: Effective management of energy internet in renewable hybrid microgrids: A secured data driven resilient architecture. IEEE Trans. Ind. Inf. 18(3), 1896–1904 (2021)
20. Rao, R.V., Savsani, V.J., Vakharia, D.: Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems. Comput. Aided Des. 43(3), 303–315 (2011)
21. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. 9(8), 1735–1780 (1997)
22. Werbos, P.J.: Backpropagation through time: What it does and how to do it. Proc. IEEE 78(10), 1550–1560 (1990)
23. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 45(11), 2673–2681 (1997)
24. Cheng, T., et al.: Stochastic energy management and scheduling of microgrids in correlated environment: A deep learning-oriented approach. Sustainable Cities Soc. 69, 102856 (2021)
25. Kavousi-Fard, A., Khosravi, A., Nahavandi, S.: A new fuzzy-based combined prediction interval for wind power forecasting. IEEE Trans. Power Syst. 31(1), 18–26 (2015)

How to cite this article: Mohammadi, H., Jokar, S., Mohammadi, M., Kavousi Fard, A., Dabbaghjamanesh, M., Karimi, M.: A Deep Learning-to-learning Based Control system for renewable microgrids. IET Renew. Power Gener. 1–10 (2023).
https://doi.org/10.1049/rpg2.12727