LSTM-Attention-Embedding Model-Based Day-Ahead Prediction of Photovoltaic Power Output Using Bayesian Optimization

N.B. When citing this work, cite the original published paper.

(article starts on next page)

LSTM-Attention-Embedding Model-Based Day-Ahead Prediction of Photovoltaic Power Output Using Bayesian Optimization

TONGGUANG YANG[1], **BIN LI**[1], **AND QIAN XUN**[2]

[1]College of Mechanic and Electrical Engineering, Hunan City University, Yiyang 413000, China
[2]Department of Electrical Engineering, Chalmers University of Technology, Gothenburg 41279, Sweden

Corresponding author: Qian Xun (qian.xun@chalmers.se)

ABSTRACT Photovoltaic (PV) output is susceptible to meteorological factors, resulting in intermittency and randomness of power generation. Accurate prediction of PV power output can not only reduce the impact of PV power generation on the grid but also provide a reference for grid dispatching. Therefore, this paper proposes an LSTM-attention-embedding model based on Bayesian optimization to predict the day-ahead PV power output. The statistical features at multiple time scales, combined features, time features and wind speed categorical features are explored for PV related meteorological factors. A deep learning model is constructed based on an LSTM block and an embedding block with the connection of a merge layer. The LSTM block is used to memorize and attend the historical information, and the embedding block is used to encode the categorical features. Then, an output block is used to output the prediction results, and a residual connection is also included in the model to mitigate the gradient transfer. Bayesian optimization is used to select the optimal combined features. The effectiveness of the proposed model is verified on two actual PV power plants in one area of China. The comparative experimental results show that the performance of the proposed model has been significantly improved compared to LSTM neural networks, BPNN, SVR model and persistence model.

INDEX TERMS LSTM-attention-embedding model, features extraction, deep learning, Bayesian optimization, residual connection.

I. INTRODUCTION

With the global concern about environmental issues, it has become the consensus of the world to develop renewable energy resources, such as wind [1], hydro [2], fuel cell [3], photovoltaic (PV) [4], [5]. The International Energy Agency (IEA) estimates that the proportion of new energy resources will reach 60% in 2040, and among them, PV and wind energy will account more than 50%. PV, also known as solar PV, has developed from a niche market of small-scale applications to a mainstream electricity source since 1992. In 2017, the cumulative global PV power generation reached nearly 460 TWh, accounting for about 2% of the total global energy, and 60% of them is for utility-scale applications, and

The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Savaghebi[ID].

the remaining 40% is for distributed applications [6]. By the end of 2018, the total cumulative global PV capacity reached up to 512GW, and it is estimated to be enough to supply 2.55% of the global electricity demand [7].

However, the PV power generation is of an intermittent nature, and considerable fluctuations will be induced due to meteorological factors [8], such as ambient temperature, relative humidity, wind speed, and clearness index. The neural networks have been used in [9] to verify that the PV power output is strongly related to the temperature, wind speed and relative humidity. It is proved that the average photon energy (APE) and the temperature have a large influence on the performance of PV power output [10]. Since the PV power output is greatly affected by meteorological factors, and strong fluctuation and intermittency will cause great impact on the system operation and grid-connected system.

It will cause the grid oscillation when a large proportion of PV system is connected to the grid [11]. Therefore, accurate prediction of PV power output can significantly improve the operation of the power system and increase the penetration of the PV system.

Several studies have been conducted to predict PV power output and have achieved significant results. The methods for PV power prediction can be summarized as physical methods, direct prediction methods, and indirect prediction methods. The physical methods are mainly dependent on a physical model obtained from the theoretical analysis of energy conversion from solar energy into electric energy. The physical model is usually based on numerical weather prediction (NWP) [12], satellite imagery TSI, and cloud movement model [13]. These models can predict the PV power output with high accuracy [14]. However, such prediction methods require additional information on satellite cloud maps, resulting in higher operating cost and computational cost.

To address this issue, the direct prediction methods, also known as time-series methods, are widely used in the PV prediction. These methods establish a mathematical relationship between the historical power series and the future PV power according to the periodicity, tendency, and other properties of PV power profiles. In [15], the rule of PV historical output curves under different weather patterns is analyzed, and the prediction of PV power output is realized by superimposing fluctuations of PV power output at different scales. In [16], the chaotic characteristics of PV power output is presented, and RBF-based neural networks are used to fit the local variation of the phase space trajectory of PV power output. These methods are based on the historical PV output data, and the PV prediction is realized by finding out the PV output law in a certain time period. However, the meteorological information, such as irradiation, temperature, humidity, wind speed and wind direction are missing in such methods, and incomplete weather information is difficult to guarantee the accuracy of the prediction.

To improve the prediction accuracy, indirect prediction methods, also known as regression methods, are proposed. With consideration of meteorological factors, the prediction performance of PV power output at certain time period has been significantly improved. The indirect prediction methods are divided as artificial neural networks [17], [18], support vector machine [21], and Markov chain [22]. Due to the obvious similarity of same weather patterns, the research on PV power prediction based on weather type is conducted, and the improvement has been done in [19]. The Kohonen weather clustering model is improved to identify the weather type of the predicted date, and the meteorological data of the predicted date and clustering historical data are used as the input of neural networks to predict the PV power output. According to different weather types, Artificial neural networks are established in [20], the prediction of PV power output at daytime shows good performance due to the individual weather type, but the prediction performance for time-varying weather is poor. Compared with the previous two methods, artificial neural networks have an advantage of higher accuracy due to that multiple weather information is considered. However, NWP data does not respond effectively to meteorological factors, although further weather division strategy has been applied, the meteorological features still need to be further explored.

Deep learning has developed rapidly in computer vision and natural language processing, and it is also used in the regression and classification of one-dimensional data [23]. Deep learning is categorized as long-short term memory (LSTM) [24], convolution neural networks [25], attention [26], and embedding [27]. Several researchers have worked on PV power prediction based on deep learning. LSTM [28] and LSTM-based self-encoder [29] perform prediction directly without considering any other additional characteristics; although with deep belief networks used in [30], the factors that affect PV power output have been analyzed, and the relevant meteorological characteristics are also used, but the characteristics are not further explored. In [31], restricted Boltzmann machine is proposed, and only original characteristics related to PV power output are explored with only temperature characteristics included, and no more weather information are used; wavelet-based decomposing CNN-LSTM [32], [33] only uses direct prediction method, and no more additional information has been explored; and Bayesian deep learning model [34] is a method of uncertainty prediction. Also, recurrent neural network model is used to learn the nonlinear characteristics of PV sequence in [35], and the PV time-series data is divided into inter-day and intra-day data. This model is superior to the classical persistence method (Persistence), back propagation neural network (BPNN), radial basis function (RBF) neural network and support vector machine (SVM), and long short-term memory (LSTM) network.

Among them, different deep learning structures have been used, and both have improved the prediction accuracy of PV power output and PV irradiance. In [36], the additional meteorological information is used as the input, and nonlinear autoregressive recurrent neural networks (NARX), a feed-forward neural network (FFNN) based model, is proposed to predict the solar irradiance; In [37], dry-bulb temperature, dew-point temperature, and relative humidity are used as the characteristics, and LSTM is used as the model to predict solar irradiance. Since the additional meteorological information can effectively improve the prediction accuracy, the better prediction performance is obtained compared with FFNN network model.

To the best knowledge of the authors, the deep learning methods can fit the PV power output very well, but the characteristics related to PV power output needs to be further explored, and the fitting ability of models should also be further improved. Therefore, to make full advantages of different deep learning models, the appropriate deep learning models are required and combined to improve the performance of the PV prediction. An improved deep learning model based on LSTM-attention-embedding algorithm is proposed in

this paper. The tendency of the meteorological factors related to PV power output under each time window and the combined features of various meteorological factors are explored. An LSTM-attention-embedding model combined with LSTM-attention and embedding mechanism is built up to optimize the meteorological characteristics. The LSTM-attention is used to learn the time-dependent meteorological numerical characteristics, and the embedding is used to describe the categorical features such as wind direction and time. The Bayesian optimization is used to select the optimal combined characteristics for LSTM-attention-embedding model.

This paper presents two original contributions that distinguish our work from existing schemes. First, the meteorological information related to PV power output is further explored, and the redundant characteristics are removed by using Bayesian optimization. Second, the different deep learning models are combined to better represent the correlation between characteristics. In the model, LSTM extracts the sequence information between characteristics, attention focuses on the important information, and embedding is to encode the categorical features.

The remainder of this paper is organized as follows. In section II, the characteristics related to PV power output are constructed, and the improved deep learning model based on Bayesian optimization is proposed. In section III, the prediction model of PV power output is proposed. The experimental results are analyzed in section IV, and finally, the conclusions are drawn in section V.

II. FEATURES EXPLORING MODEL-BASED PREDICTION OF PV POWER OUTPUT

A. FEATURES CONSTRUCTION

The prediction accuracy of the model is dependent on the features of meteorological factors, and more effective meteorological features are required to improve the prediction performance. For the meteorological characteristics $X = [X_1, X_2, \ldots, X_N]^\mathrm{T}$, X_i $(i = 1, 2, \ldots, N)$ is the features at sample time i, where $X_i = [x_{i1}, x_{i2}, \ldots, x_{iM}]$, x_{ik} $(k = 1, 2, \ldots, M)$ is the kth features among samples, and M is the total numbers of meteorological factors. The following three kinds of features are explored.

1) The original features, that is $X = [X_1, X_2, \ldots, X_N]^\mathrm{T}$.

2) The statistical features

The statistical features are used to reflect the data fluctuations within the time window t, and they can be expressed as

$$\begin{cases} \bar{x}_{ikt} = \dfrac{1}{2t+1} \sum_{v=t}^{t} x_{(i+v)k} \\ S_{ikt} = \left[\dfrac{1}{2t+1} \sum_{v=t}^{t} \left(x_{(i+v)k} - \bar{x}_{ikt} \right)^2 \right]^{\frac{1}{2}} \\ x_{\max kt} = \max\{x_{vk}, x_{v+1}, \ldots, x_{v+tk}\} \end{cases} \quad (1)$$

where, $\bar{x}_{ikt}$, S_{ikt}, x_{maxkt} are the mean value, standard deviation, and maximum value within the time window of $[i\text{-}t, i + t]$.

3) The combined features

The combined features, including the linear-combined features and nonlinear-combined features, are expressed as II-B. The linear-combined features are used to explore the linear relationship between each feature, while the nonlinear-combined features are used to explore the nonlinear relationship.

$$\begin{cases} z_{ia} = x_{ij} + x_{ik} \\ z_{is} = x_{ij} - x_{ik} \\ z_{im} = x_{ij} \times x_{ik} \\ z_{id} = x_{ij}/x_{ik} \end{cases} \quad (2)$$

where, z_{ia}, z_{is}, z_{im}, and z_{id} are the additive, subtractive, multiplicative, and divisive features at sample time i. The original features X and statistical features in (1) are included in x_{ij}.

4) Time features

The time features, such as the month of the year, the day of the week, the day of the month, the hour of the day, and the minute of the hour are extracted.

B. IMPROVED DEEP LEARNING MODEL

1) LSTM NEURAL NETWORKS

LSTM is currently a widely-used deep learning approach in machine learning area, and it is proposed by Hochreiter and Schmidhuber in 1997 [24]. The LSTM can solve the long-term dependent sequences, and the main structure is shown in FIGURE 1.

The LSTM can be represented as a chain structure by time expansion, LSTM-neural network architecture shows the time expansion form of LSTM from sample time 0 to sample time t. In the figure, A represents the LSTM unit, for each time instance, there is one corresponding input x_t and output h_t, and the output h_t at sample time t will selectively attend the input information from sample time 0 to sample time t-1.

There are four neural network layers in the repeating module. Neuron unit structure of LSTM shows the specific time connection of each LSTM unit, and it transfers the state (C_{t-1}) and output (h_{t-1}) of the cell at sample time t-1 to LSTM unit at sample time t, then the state (C_t) and output (h_t) of the cell at sample time t are transferred to LSTM unit at sample time t+1 in the same way until to the last sample time.

The detailed structure of each LSTM unit is shown at the bottom of FIGURE 1. The unit is mainly divided into forget gate, input gate, tanh layer and output gate. The input of each gate and layer are h_{t-1} and x_t, and the output f_t of the forget gate is obtained by the activation function σ. The output i_t of the input gate is also obtained by function σ, while the output $\tilde{C}_t$ of the tanh layer is obtained by tanh function. The output O_t of the output gate is calculated from function σ, and the output h_t of the LSTM is calculated by c_t and O_t.

For one LSTM cell, it has four gates. The first layer of neurons is a sigmoid control layer of the forget gate, and it
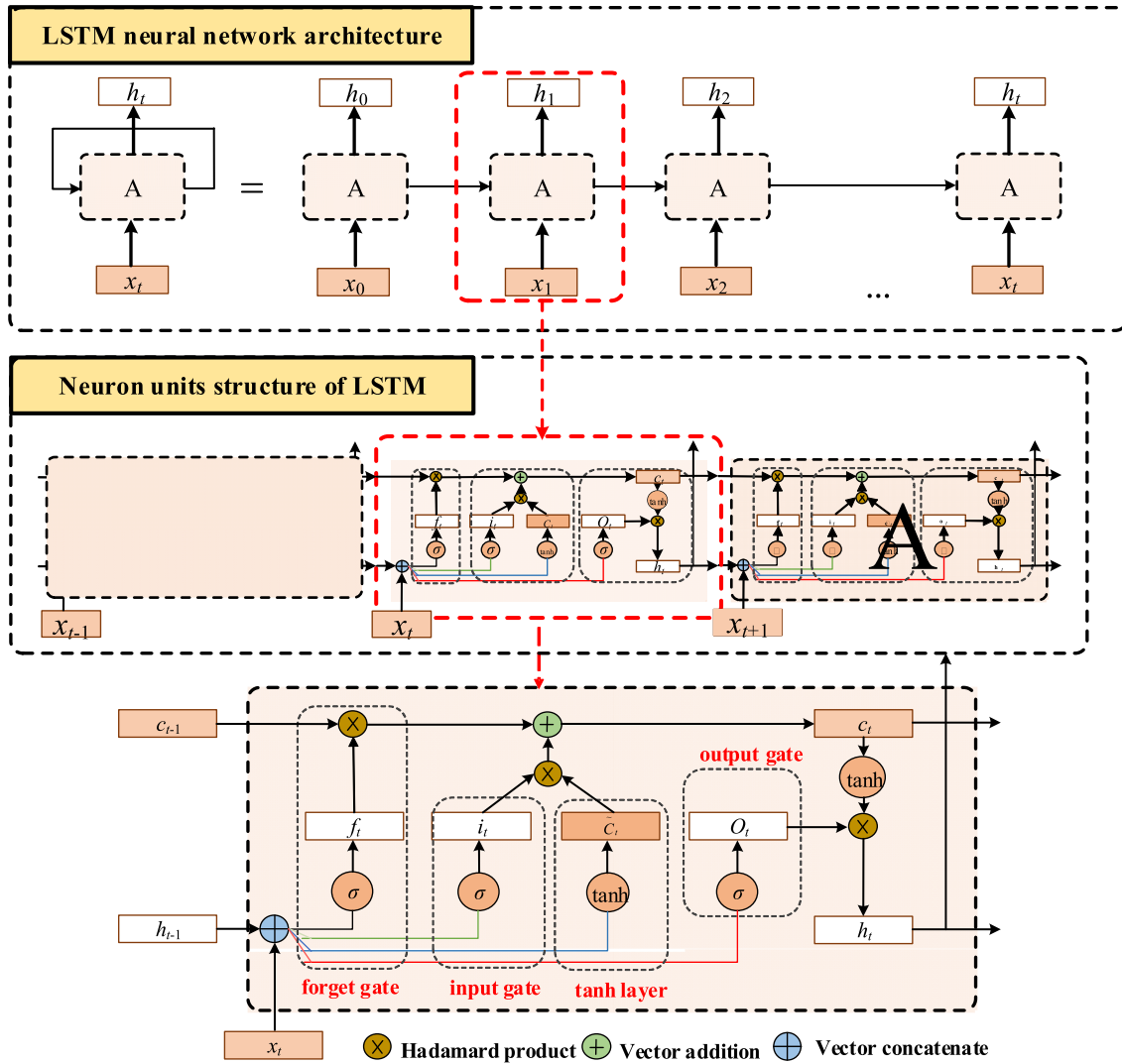
FIGURE 1. LSTM structure.

can be expressed as

$$f_t = \sigma \left(W_f \cdot [h_{t-1}, x_t] + b_f \right) \tag{3}$$

where f_t is the output of the forget gate, σ is the sigmoid activation function, W_f and b_f are denoted as the weight coefficient and offset of the forget gate, respectively. h_{t-1}, and x_t are the output at sample time t-1 and the input at sample time t, respectively.

The sigmoid layer of the input gate determines which information to update, the tanh layer is used to create a new candidate value that might be added to the states of the cell.

The second and third layer are the input gate and tanh layer, respectively. They can be expressed as

$$i_t = \sigma \left(W_i[h_{t-1}, x_t] + b_i \right) \tag{4}$$

where i_t is the output of the input gate, W_i and b_i are the weight coefficient and offset of the input gate, respectively.

$$\tilde{C}_t = \tanh \left(W_C[h_{t-1}, x_t] + b_C \right) \tag{5}$$

where $\tilde{C}_t$ is candidate status, W_C and b_C are the weight coefficient and offset of the updated variables, respectively.

Then, the old state of the cell C_{t-1} is updated as

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{6}$$

Finally, an output gate (O_t at sample time t) including the sigmoid layer, shown as II-B.2, and the previous state of the cell, shown as (6), are multiplied together through a tanh layer to get the final output of the LSTM. The tanh layer is used to guarantee the output of the previous state of the cell to be bonded between -1 and 1.

$$O_t = \sigma \left(W_o[h_{t-1}, x_t] + b_o \right) \tag{7}$$

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

2) ATTENTION MECHANISM

When a large amount of input information is input into neural networks, the influence of the different inputs on the output

value is different. To improve the computational efficiency of neural networks, more computing power is allocated to the important input information for the same output, and this is called the attention mechanism.

The attention mechanism is calculated in two steps, the first step is to calculate the attention distribution value α_i on all input values; and another step is to calculate the weighted average of the input information for a single output value. The attention mechanism is shown in FIGURE 2.
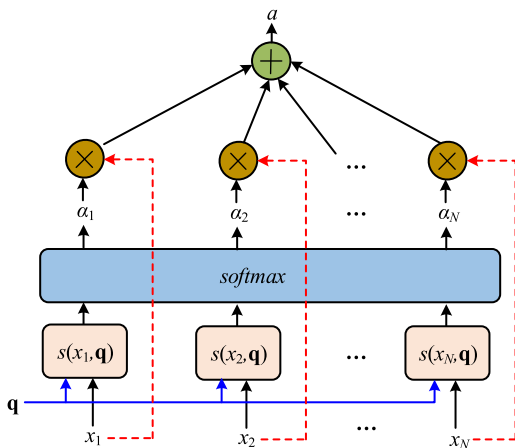

FIGURE 2. Attention structure.

where x_i is the input value, $\mathbf{q}$ is the query vector of the neural network, s is the score function of the attention, α_i is the attention distribution value of all input values for the query vector $\mathbf{q}$.

The score function can be expressed as

$$s(x_i, \mathbf{q}) = x_i^{\mathrm{T}} W \mathbf{q} \qquad (9)$$

where, W is neural network parameters that can be learned autonomously. Taking the *softmax* value of (9), one can get the attention distribution α_i, shown as

$$\alpha_i = softmax\left(s\left(x_i, \mathbf{q}\right)\right) \qquad (10)$$

a is the weighted average of the attention distribution value α and the input value x_i, and it can be expressed as

$$a = \sum_{i=1}^{N} \alpha_i x_i \qquad (11)$$

3) EMBEDDING

To ensure that the network can learn the categorical features such as time and wind direction, the embedding, a widely used structure in natural language processing, is applied here. The embedding is used to map the discrete features, such as the category number of the wind direction D, and a matrix $D \times n_{emb}$ is obtained after embedding. In the matrix, n_{emb} is the potential features, and the same processing is performed for other time features.
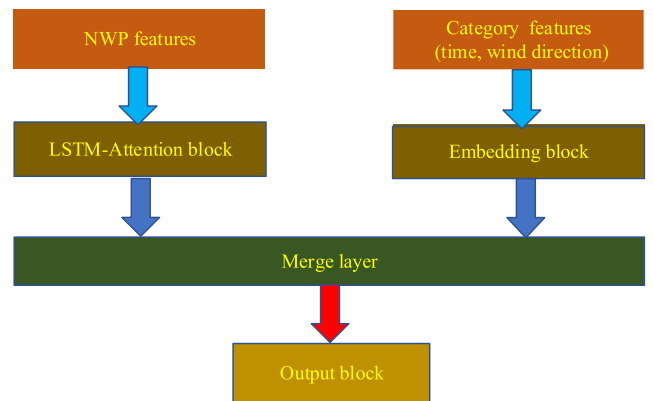

FIGURE 3. Improved LSTM network.

4) THE IMPROVED DEEP LEARNING NETWORK

To better learn the numerical features and categorical features for a long time, and an improved LSTM network for PV output prediction is shown in FIGURE 3.

The improved LSTM network is made up by several blocks, and they include continuous features (NWP features) and categorical features (time and wind direction), LSTM-attention block, embedding block, merge layer and the output block. NWP features and categorical features are the input of the improved LSTM network. LSTM-attention block is used to memorize and attend the historical information, and embedding block is used to encode the categorical features. The merge layer is used to connect the LSTM-attention block and embedding block together, and the output block is to output the prediction results of PV power output. A detailed connection of each block of the improved LSTM network is shown in FIGURE 4.

1) The LSTM block consists of two layers of LSTM followed by two layers of fully connected (FC) layers. A residual connection is added for LSTM layer and FC layer to improve the gradient disappearance problem in the multilayer neural networks [35]. The attention layer is connected between the LSTM layer and the FC layer, and it is used to focus on the important information.

2) The embedding block is used to encode each category feature, such as wind direction, the month of the year, and the minute of the hour. Then, these features will be merged, and finally they are connected to two layers of FC. A residual connection is also added to the FC layer.

3) The output block is used to output the prediction results with two layers of FC.

C. BAYESIAN OPTIMIZATION

Bayesian optimization is usually used to select the optimal parameters in the machine learning process. Compared with the grid search method, the Bayesian optimization has the following advantages: 1) the prior distribution is continuously updated based on the previous parameters; 2) due to the small number of iterations, it is more efficient when the parameter
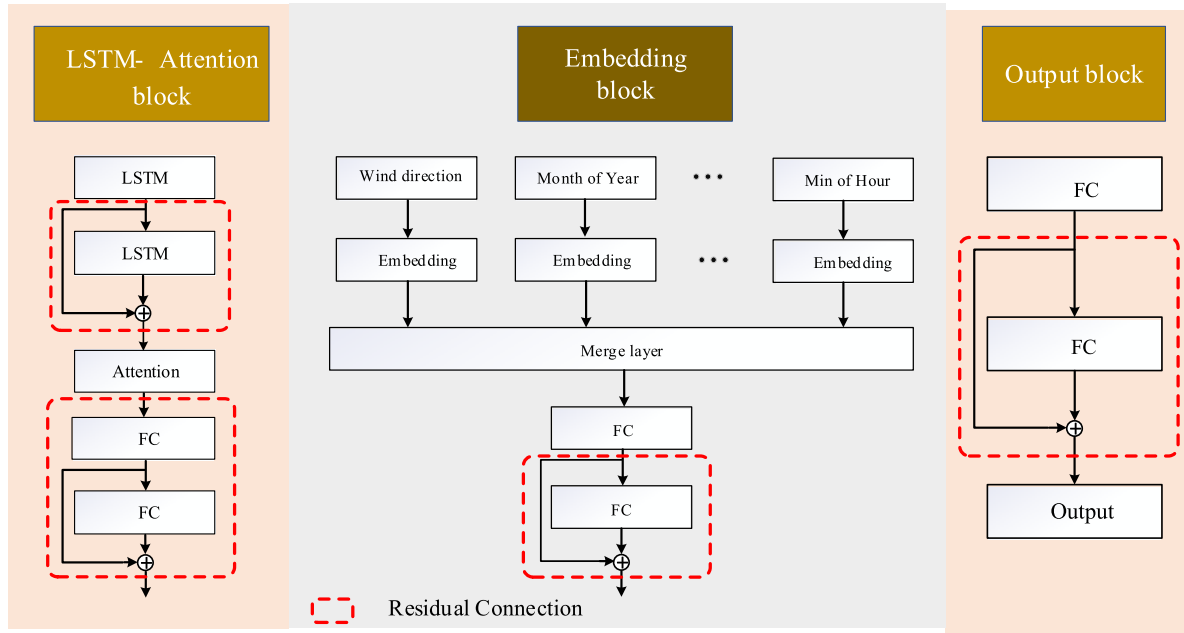
FIGURE 4. The detailed improved LSTM structure.

dimension is large; 3) it is robust to the non-convex problem, while the grid search is easy to fall into the local optimal solution.

In the process of Bayesian optimization, the loss function of the parameter fitting model is a Gaussian process, and parameters are updated to get the posterior probability of this function. To reduce the computational cost of the model, the optimal combined features are obtained in Bayesian optimization where the features presented in section 1.1 are selected as the parameters.

The following parameters, time window t, the number of statistical features $num1$, and the number of the combined features $num2$, are optimized by Bayesian optimization. To constrain the features dimension and optimize the model parameters, the combined parameters of the sample is constructed as combined features $[p_1, p_2, \ldots, p_q]$. The process of Bayesian optimization is described as

1) According to the combined parameters, the loss function is estimated as a Gaussian process;

Part of the combined parameters $[p_1, p_2, \ldots, p_r]$ are selected to calculate their corresponding loss functions $[L(p_1), L(p2), \ldots, L(p_r)]$, then the parameters and the loss functions are combined to get parameters-loss functions set.

$$D_{1r} = \{(p_s, L(p_s)), s = 1, \ldots r\} \quad (12)$$

The estimated loss function obeys Gaussian distribution, and Gaussian distribution of $L(p_{1:r})$ can be expressed as

$$\begin{cases} L(p_{1:r}) \sim N(0, K) \\ K = \begin{bmatrix} k(p_1, p_1) & \cdots & k(p_1, p_r) \\ \vdots & \ddots & \vdots \\ k(p_r, p_1) & \cdots & k(p_r, p_r) \end{bmatrix} \end{cases} \quad (13)$$

where, k is the core function, and it is used to represent the covariance.

2) Sample selection using expected improvement (EI) function;

The new sampling point is determined by comparing the loss function of the new sampling point with the current sampling point, and this mathematical expectation is called EI function. The Three-structured Parzen Estimator (TPE) is used to establish the EI function [39], and the present p^+ is selected as the optimal combined parameters, with y^*, greater than $L(p^+)$, as the threshold value, and the EI function is expressed as

$$\begin{cases} EI_{y^*}(p) = \int_{-\infty}^{y^*} (y^* - y) p(y|p) \, dy \\ = \dfrac{\gamma y^* \ell(p) - \ell(p) \int_{-\infty}^{y^*} p(y) \, dy}{\gamma \ell(p) + (1 - \gamma) g(p)} \\ p(p|y) = \begin{cases} \ell(p) & \text{if } y \le y^* \\ g(p) & \text{if } y \ge y^* \end{cases} \end{cases} \quad (14)$$

where, $\ell(p)$ is the distribution function which is smaller than y^*, $g(p)$ is the distribution function which is greater than y^*. The point that makes the EI function the largest is selected as the sampling point.

$$p_{new} = \arg\max_p EI(p) \quad (15)$$

3) Update Gaussian distribution;

After adding the new sample, the new Gaussian distribution can be expressed as

$$\begin{cases} \begin{bmatrix} L(p_{1:r}) \\ L(p_{new}) \end{bmatrix} \sim N\left(0, \begin{bmatrix} K & k' \\ k'^T & k(p_{new}, p_{new}) \end{bmatrix}\right) \\ k' = [k(p_{new}, p_1), \cdots, k(p_{new}, p_r)] \end{cases} \quad (16)$$

The Gaussian distribution of $L(p_{r+1})$ is calculated as [40]

$$\begin{cases} P\left(L\left(p_{new}|D_{1:r}, p_{new}\right)\right) \sim N\left(\mu, \sigma^2\right) \\ \mu = k'^T K^{-1} L\left(p_{1:r}\right) \\ \sigma^2 = k\left(p_{new}, p_{new}\right) - k'^T K^{-1} k' \end{cases} \quad (17)$$

4) Loop 2) and 3) until the maximum number of the iteration is reached and the optimal output point at the present state is calculated.
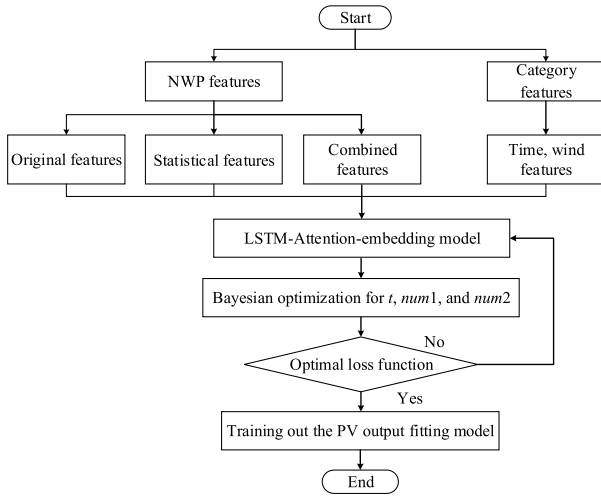

FIGURE 5. The flowchart of PV power prediction.

III. PREDICTION OF PV POWER OUTPUT

A. THE FLOWCHART OF PV PREDICITION

The flowchart of the PV prediction is shown in FIGURE 5. According to the meteorological factors, the statistical features and the combined features are constructed, and the time features are also extracted. Then, the LSTM-attention-embedding model is built up. To reduce the redundancy of the features, the Bayesian optimization is used to optimize the parameters. With the optimization method, the optimal time window, the number of statistical features and the number of the combined features are obtained. Finally, the prediction model of the PV power output is trained based on the available data.

B. EVALUATION INDEX

To quantify how well a prediction model is, mean squared error (MSE), mean absolute error (MAE), coefficient of determination (R^2) and root mean squared error (RMSE) are used to analyze the experimental results. All evaluation indexes are based on the difference between each predicted value and the actual value.

1) Mean squared error

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 \quad (18)$$

2) Mean absolute error

$$MAE\left(y, \hat{y}\right) = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \quad (19)$$

3) Coefficient of determination

$$R^2\left(y, \hat{y}\right) = 1 - \frac{\sum\limits_{i=0}^{n-1} \left(y_i - \hat{y}_i\right)^2}{\sum\limits_{i=0}^{n-1} \left(y_i - \bar{y}\right)^2} \quad (20)$$

4) Root mean squared error

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2} \quad (21)$$

IV. EXPERIMENT ANALYSIS

A. DATA ACQUISITION AND SIMULATION CONFIGURATION

The dataset used for the experiment was collected from a PV power output dataset within 25 months (from 1st April 2016 to 30th April 2018) from two PV stations located in one area of China, and the sampling frequency of the data is 1.1×10^{-3} Hz (every 15min for one point). From the PV stations, we can get day-ahead historical meteorological data, which is same as NWP data. The information of each PV station is shown in TABLE 1, and the sampling data of each station includes time, irradiance, wind direction, temperature, pressure and humidity, with the resolution of 15 mins. Detailed meteorological factors are predicted by in-house model, which are listed as IV-B.

TABLE 1. Information of each PV station.

PV station	The number of datasets	Time window
Station 1	Training set 53488	2016/4/1 0:15:00—2017/12/12 17:00:00
	Test set 13371	2017/12/12 17:15:00—2018/4/30 23:45:00
Station 2	Training set 35005	2017/1/1 0:00:00—2018/1/29 20:15:00
	Test set 8750	2018/1/29 20:30:00—2018/4/30 23:45:00

Due to advantages of being easy and simple, Keras toolkit is used to build up the deep learning network. In the network, python is used to program, NumPy and pandas are used for data processing, and matplotlib and seaborn are used for drawing. Our computer resources consist of an i7-9000 CPU, 32GB RAM, and GPU 2080Ti*2. The learning and experimental phase of the proposed model are conducted in the GPU environment.

B. PARAMETERS OPTIMIZATION

The detailed training parameters for the improved deep learning model is shown in TABLE 3. The method for the parameter initialization can be found in [32].

TABLE 2. The detailed meteorological parameters.

Time	Irradiance (W/m²)	Wind speed (m/s)	Wind angle (°)	Temperature (°C)	Pressure (kPa)	Humidity (%)
2017-01-01 23:59:59.994	-1.0	-0.811321	340	-0.486869	0.393939	-0.062500
2017-01-02 00:14:59.994	-1.0	-0.811321	342	-0.490909	0.393939	-0.062500
2017-01-02 00:29:59.994	-1.0	-0.801887	346	-0.494949	0.393939	-0.041667
2017-01-02 00:44:59.994	-1.0	-0.801887	349	-0.498990	0.393939	-0.041667
2017-01-02 00:59:59.994	-1.0	-0.801887	349	-0.503030	0.393939	-0.041667

TABLE 3. Training parameters used in the model.

Parameters	Value
Learning rate	0.001
Batch size	100
Loss Function	MSE
Epoch	300
Optimization method	Adam Optimizer

TABLE 4. Parameters for different models.

Model	Parameters
LSTM	[LSTM (128)-LSTM (64)-FC (32)-FC (1)]
BPNN	[FC (128) FC (64)-FC (32)-FC (1)]
SVR	kernel='rbf', C=100, gamma=0.01
Persistence	$\Delta t = 96$

TABLE 5. Search Space.

Parameters	Parameter description	The range of parameters	Sampling method
t	Time window	[1, 96]	Discrete random sampling
$num1$	The number of statistical features	[1, 5]	Discrete random sampling
$num2$	The number of combined features	[2, 15]	Discrete random sampling

To verify the effectiveness of the proposed model, LSTM, BPNN, SVR, and Persistence [41] are compared. The detailed parameters are shown in IV-C. LSTM and BPNN model are neural networks that are used to layer stack structures, the parameters show the connection between each layer, the word before () represents the neural network layer, and the number in () represents the number of neural units of one layer.

C. DATA ANALYSIS

According to features construction method of meteorological factors, such as irradiance, temperature, humidity, pressure and wind speed presented in section 1.1, three statistical features are extracted for each factor. Then the number of statistical features is 15 (5×3). Among the combined features, $2 \times C_{(15+5)}^2$ features are added due to the exchangeability of the addition and multiplication, and $2 \times 2 \times C_{(15+5)}^2$ features are added due to the unchangeability of the subtraction and the division, an additional 5 time features and one wind direction features are added, and the total number of features is $5 + 15 + 2 \times C_{(15+5)}^2 + 2 \times 2 \times C_{(15+5)}^2 + 6 = 1166$.

Since excessive features will significantly increase the training time of the model, the computational cost will also be increased. To reduce the number of features, the effective features are selected by Bayesian optimization, and search space is shown in TABLE 5.

In our experimental process, LSTM-attention-embedding model needs 433s to find optimal features, while BLSTM-attention-embedding model requires 73610s. Obviously seeing that the time required for BLSTM-attention-embedding model is multiplied compared to LSTM-attention-embedding, and this is the disadvantage of the

proposed method. However, the training and prediction cost after finding the optimal features are equivalent to other models. Also, once the optimal features are selected by Bayesian optimization, and a more accurate prediction can be obtained without demand to optimize features every time. This is the advantage of BLSTM-attention-embedding.

The experimental dataset of PV station 1 is divided into a training dataset and a test dataset in chronological order of 8:2, and the test error is selected as the target of Bayesian optimization. The iteration steps are 200, and the optimal parameters appear at the step of 199, and the optimal parameters are $\{t = 18, num_1 = 3, num_2 = 12\}$ with MSE of 0.66MW. Compared with the model $\{t = 48, num_1 = 5, num_2 = 15\}$ without any optimization, the MSE is decreased by 0.142MW, and the features dimension is also effectively reduced. Due to the redundant features and the invalid features are eliminated, the computational cost of the model is significantly reduced. The joint distribution map of the optimization process is shown in FIGURE 6. The selected optimal features are listed in TABLE 6, and num1 and num2 are statistical features and combined features, respectively. Among them, the most important feature is the irradiance, and this is in line with the real situation. That is because the fundamental of photovoltaic power generation is to convert solar energy into electrical energy.

D. DAY-AHEAD PREDICTION RESULTS

To demonstrate the effectiveness of the proposed method, the following models are compared.

1) Model 1: The BLSTM-attention-embedding model is the prediction model, and the features obtained from section 1.1 are the model input;

2) Model 2: The LSTM-attention-embedding model is the prediction model, and the features obtained from section 1.1 are the model input;

TABLE 6. Features selection result.

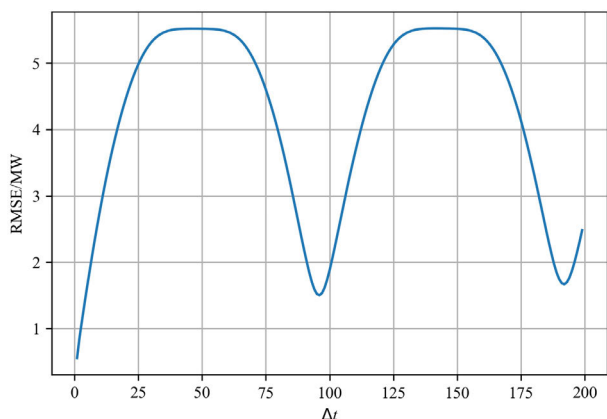| num1	num2			
Irradiance	Irradiance	Pressure	Temperature-std	Irradiance-std
Temperature	Humidity	Wind speed	Temperature-mean	Irradiance-max
Pressure	Temperature	Wind angle	Humidity-mean	Pressure-std

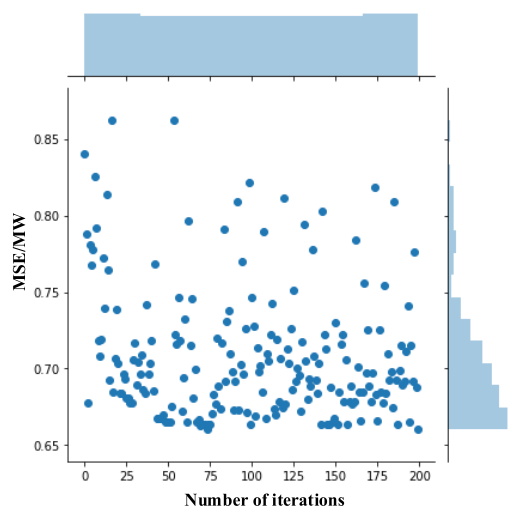**FIGURE 6.** Parameter of Persistence and RMSE.

FIGURE 7. The joint distribution map of errors with the number of iterations.

3) Model 3: The LSTM model is the prediction model, and the original features are the model input;

4) Model 4: The BPNN model is the prediction model, and the original features are the model input;

5) Model 5: The SVR model is the prediction model, and the original features are the model input;

6) Model 6: The persistence model is the prediction model, and historical time series of PV power are the model input.

FIGURE 9 compare the scatter plots of the predicted PV data with the true data by using different models for station 1 and station 2, respectively. The solid lines indicate a perfect forecast, while the star-marks indicate an instance of prediction. The closer a star-mark is to the solid line, the more

accurate the prediction is. From **FIGURE 8**, we can see that BLSTM-attention-embedding model has the best symmetry about the solid line, and the scatter points in LSTM-attention-embedding model, LSTM model, BPNN model and SVR model are more biased toward true PV data, while persistence model is symmetrical but the points spread too widely. The situation of station 2, shown in **FIGURE 9**, is similar to that of **FIGURE 8**, but the overall performance is significantly worse due to less training samples.

TABLE 7. Errors Comparison with Different Models

Models (PV station 1)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.287	0.621	0.879	1.135
LSTM-Attention-Embedding	1.289	0.651	0.875	1.144
LSTM	1.309	0.685	0.837	1.195
BPNN	1.427	1.002	0.822	1.229
SVR	1.511	1.043	0.812	1.503
Persistence	2.259	1.022	0.780	1.503
Models (PV station 2)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.235	0.613	0.843	1.111
LSTM-Attention-Embedding	1.401	0.726	0.822	1.184
LSTM	1.498	1.015	0.810	1.224
BPNN	1.594	1.059	0.798	1.263
SVR	1.720	1.108	0.782	1.311
Persistence	2.545	1.154	0.676	1.596

The errors of the prediction results of PV station 1 and PV station 2 with different models are shown in TABLE 7. It shows BLSTM-attention-embedding model has the best MSE, MAE, R^2 and RMSE, and followed by LSTM model, BPNN model, SVR, and Persistence model. The prediction accuracy of PV power output can be effectively improved by using LSTM-attention-embedding model, and after Bayesian optimization, the prediction accuracy can be further improved since more diverse effective features can provide more useful information for training the model. The proposed LSTM-attention-embedding model can learn the long-term tendency and short-term mutation tendency of PV power generation, and Bayesian optimization can remove the redundant features and select the optimal effective features to improve the prediction accuracy of PV power output.

Regarding the prediction performance between PV station 1 and PV station 2, index R^2 is selected as evaluation index, since R^2 is the only index to show the fit degree between the predicted data and the true data without any limitations to training samples. However, index MSE, MAE and RMSE are positively correlated with the true data, and they are used to evaluate the prediction performance based on the same dataset. We can see that PV station 1 has higher value on index R^2, and its overall prediction
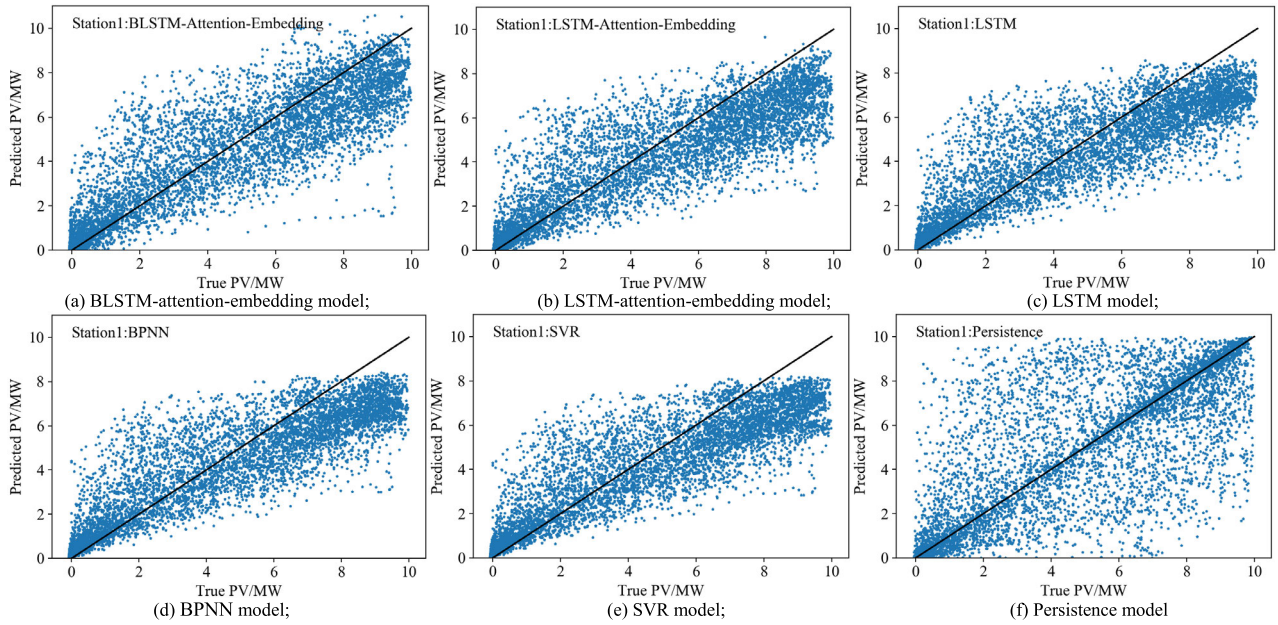
FIGURE 8. Scatter plots of the true and predicted PV data for all models on station 1.
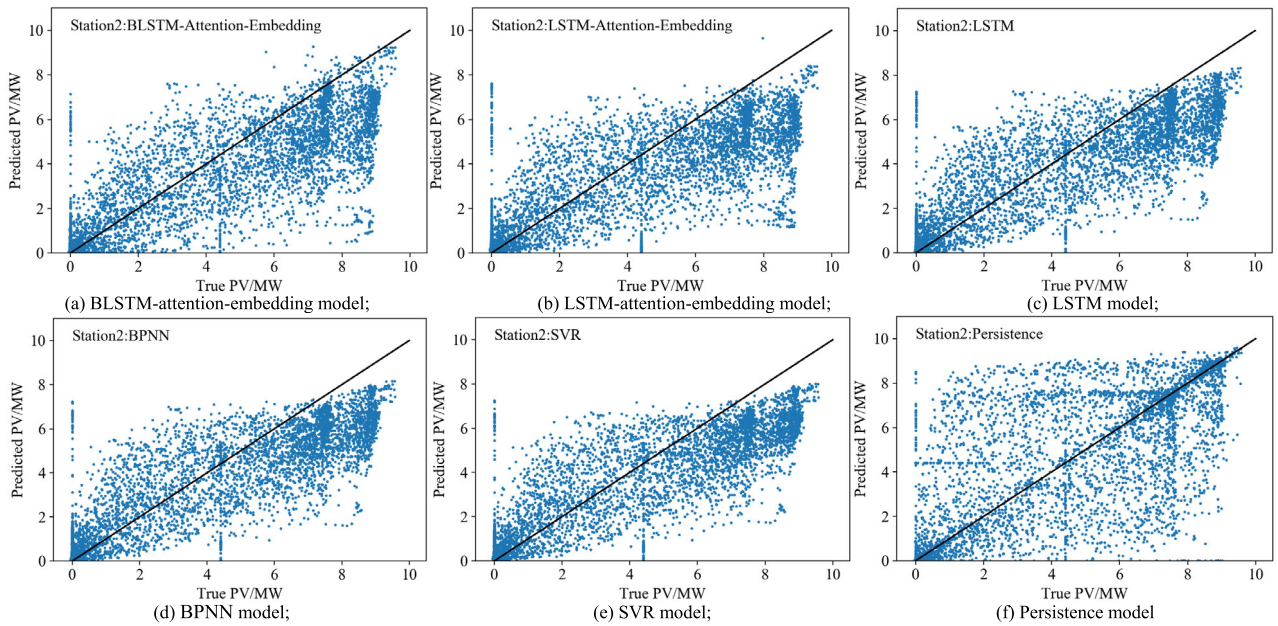

FIGURE 9. Scatter plots of the true and predicted PV data for all models on station 2.
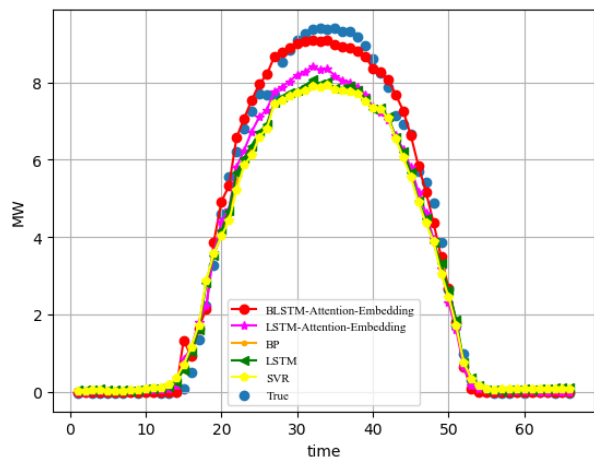
performance is better than the PV station 2 due to that station 1 has more dataset. The prediction performance of LSTM-attention-embedding model with Bayesian optimization can be significantly improved compared with other models. The effectiveness of features construction and Bayesian optimization has been verified, and both can effectively improve the prediction accuracy of PV power output.

To further verify the superiority of the proposed model, the prediction results of PV power output under three typical weather conditions are analyzed on two PV stations.
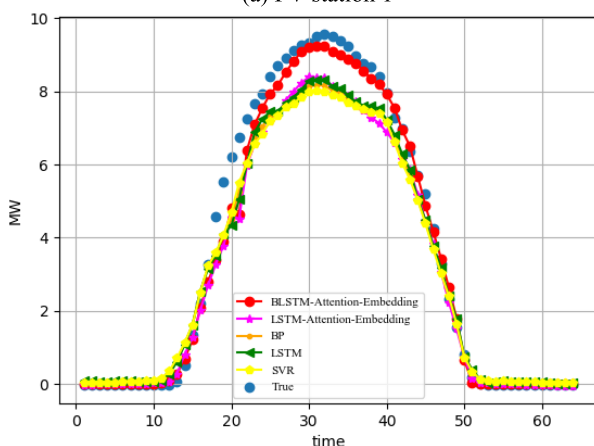
Since there is no PV output at night, 64-time samples at daytime are selected as the experimental results of PV output.

1) The prediction results at sunny weather

The real-time prediction curves of the PV power output of two stations at sunny weather are shown in FIGURE 10. The PV output prediction of each model is relatively accurate at sunny weather, and this is mainly because the PV power curve at sunny days is smooth and less fluctuating. Compared with other models, the proposed BLSTM-attention-embedding model is closer to experimental results.

(a) PV station 1

(b) PV station 2

TABLE 8. Prediction errors at sunny weather.

Models (PV station 1)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	0.130	0.265	0.990	0.361
LSTM-Attention-Embedding	0.513	0.587	0.961	0.717
LSTM	0.721	0.698	0.946	0.849
BPNN	0.792	0.740	0.940	0.890
SVR	0.871	0.783	0.934	0.933
Persistence	1.179	0.827	0.911	1.086
Models (PV station 2)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	0.221	0.258	0.985	0.470
LSTM-Attention-Embedding	0.583	0.559	0.961	0.821
LSTM	0.639	0.589	0.957	0.799
BPNN	0.674	0.562	0.955	0.763
SVR	0.707	0.618	0.953	0.841
Persistence	0.808	0.567	0.945	0.899

The prediction errors of different models at sunny weather are summarized in TABLE 8. Compared to LSTM, BPNN, SVR, and Persistence model, the prediction accuracy of LSTM-attention-embedding model is slightly improved. However, the performance of BLSTM-attention-embedding model with Bayesian optimization has been significantly improved, showing high fitting degree due to the optimal

characteristics used. We can conclude that the proposed prediction model can improve the prediction accuracy of PV power output and can reflect the tendency of the PV power generation at sunny weather.

2) The prediction results at cloudy weather

The real-time prediction curves of the PV power output of two stations at cloudy weather are shown in FIGURE 10. Since more effective features are obtained from BLSTM-attention-embedding model and LSTM-attention-embedding model, and these models can capture the fluctuation of PV power output very well. BLSTM-attention-embedding model and LSTM-attention-embedding model are closer to experimental results compared with other models. However, the prediction accuracy of each model at cloudy weather is lower than that at sunny weather.

TABLE 9. Prediction errors at cloudy weather.

Models (PV station 1)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.193	0.599	0.909	1.092
LSTM-Attention-Embedding	1.454	0.859	0.889	1.206
LSTM	1.332	0.827	0.898	1.154
BPNN	1.331	0.806	0.898	1.153
SVR	1.361	0.864	0.896	1.167
Persistence	1.925	0.806	0.865	1.388
Models (PV station 2)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.192	0.745	0.888	1.092
LSTM-Attention-Embedding	1.279	0.87	0.880	1.131
LSTM	1.215	0.84	0.886	1.102
BPNN	1.541	0.978	0.855	1.241
SVR	1.363	0.901	0.872	1.167
Persistence	2.249	1.001	0.846	1.500

The prediction errors of different models at cloudy weather are summarized in TABLE 9. Compared to LSTM, BPNN, SVR, and Persistence model, the prediction accuracy of LSTM-attention-embedding model is slightly improved. BLSTM-attention-embedding model with Bayesian optimization has a slight improvement in accuracy based on LSTM-attention-embedding model since attention mechanism in LSTM unit can capture the tendency of PV power generation. Therefore, the proposed model can improve the prediction accuracy of PV output at cloudy weather, but the model has a lower accuracy than that at sunny weather.

3) The prediction results at rainy weather

The real-time prediction curves of the PV power output of two stations at rainy weather are shown in FIGURE 12. It is difficult to capture PV power generation due to more fluctuate and violent at rainy weather. Therefore, the prediction performance of PV power output of each model is not as good as sunny and cloudy weather, and even the prediction accuracy of BLSTM-attention-embedding model and LSTM-attention-embedding model are not close to experimental results. However, we can still see that LSTM-attention-embedding model can represent the general tendency of PV power generation, and BLSTM-attention-embedding model is more easily to capture the sudden variation of the PV power, since the features after Bayesian optimization can effectively predict the PV power output, and it allows the model to capture the PV power more accurately.
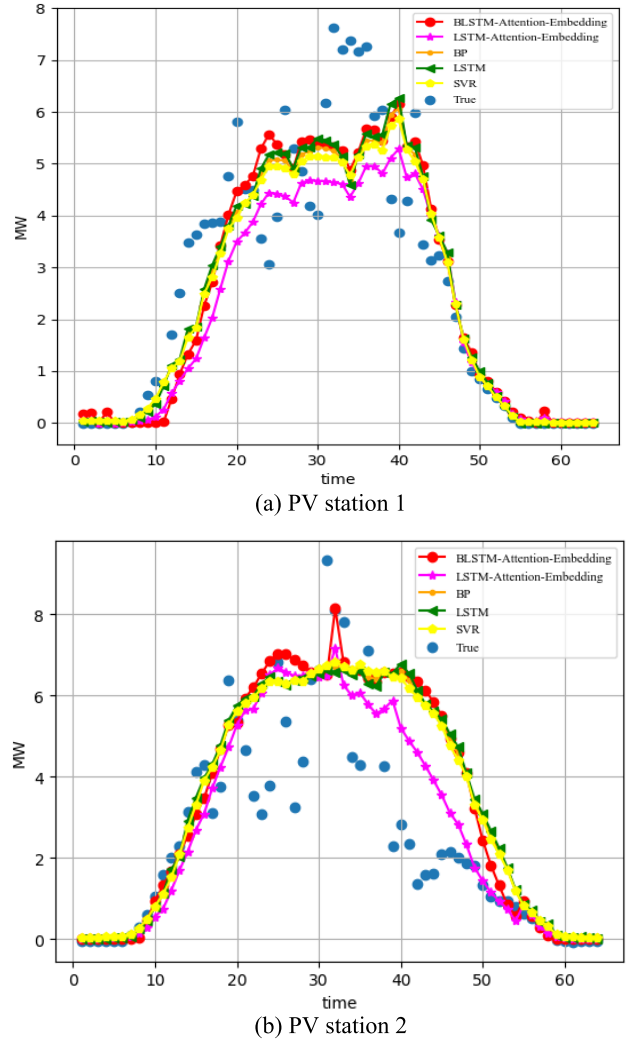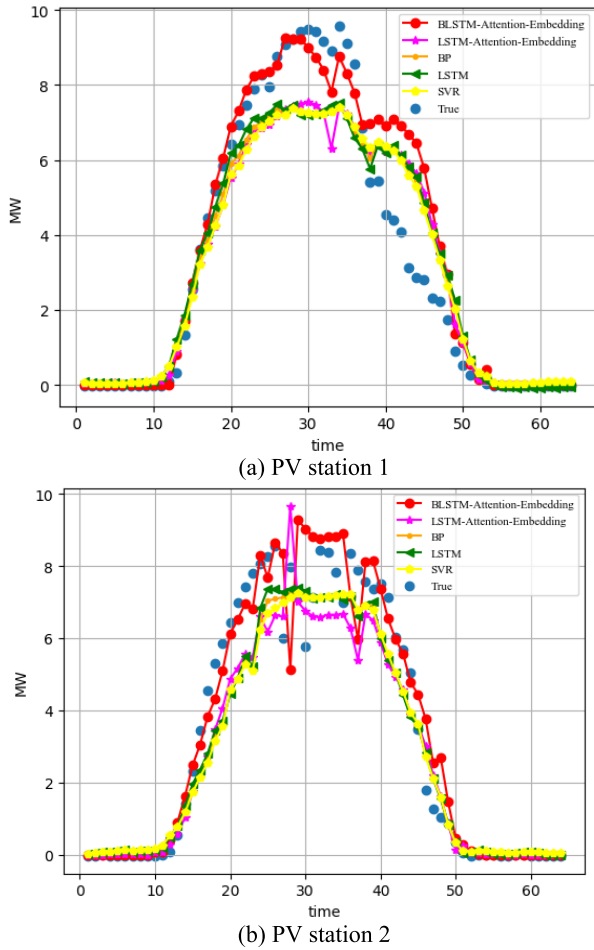
(a) PV station 1

(b) PV station 2

FIGURE 11. The prediction results at cloudy weather.

TABLE 10. Prediction errors at RAINLY weather.

Models (PV station 1)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.906	0.774	0.771	1.381
LSTM-Attention-Embedding	1.875	0.762	0.773	1.369
LSTM	2.121	0.850	0.743	1.456
BPNN	2.209	0.844	0.732	1.486
SVR	2.086	0.869	0.747	1.444
Persistence	3.099	1.322	0.438	1.760
Models (PV station 2)	MSE	MAE	R^2	RMSE
BLSTM-Attention-Embedding	1.906	0.944	0.683	1.380
LSTM-Attention-Embedding	3.195	1.221	0.468	1.787
LSTM	3.474	1.272	0.422	1.864
BPNN	3.330	1.245	0.446	1.825
SVR	3.588	1.235	0.403	1.894
Persistence	4.126	1.467	0.314	2.031

The prediction errors of different models at rainy weather are summarized in TABLE 10. The prediction accuracy of BLSTM-attention-embedding model and LSTM-attention-embedding model for the PV station 1 is higher than that for PV station 2 due to more training datasets of PV station 1. Due to the randomness of PV power on rainy weather, the LSTM-attention-embedding model shows better accuracy on some individual rainy days, and this is the part that needs to be improved in the future work. The proposed method can

(a) PV station 1

(b) PV station 2

FIGURE 12. The prediction results at rainy weather.

improve the prediction accuracy of PV output on rainy days, but the accuracy is limited. This is the present challenge in predicting the PV power. In [33] and [42], the prediction accuracy on rainy days is also relatively poorer than the cloudy and sunny days.

V. CONCLUSION
PV power output is strongly related to the meteorological factors, and it shows the intermittency and volatility. Large-scales of PV fed to the grid will challenge the power balance of the grid, leading to a series of disturbances. Accurate prediction of PV power output is an effective way to eliminate the problems caused by a high proportion of PV grid. Therefore, a BLSTM-attention-embedding model is proposed for the prediction of PV power output, with features construction of meteorological factors.

Several blocks are included in BLSTM-attention-embedding model, and they are: 1) LSTM, which is used to input meteorological factors and attend historical information; 2) Attention mechanism, which is used to focus on the

import features; 3) Embedding, which is used to encode each feature; 4) Bayesian optimization, which is used to remove the redundant features and select more effective combined features.

The prediction results show the effectiveness of the proposed method. The LSTM-attention-embedding model can effectively improve the prediction accuracy of PV power generation, and after Bayesian optimization, the prediction accuracy can be further improved. Due to more effective diverse features are obtained, LSTM-attention-embedding model can learn the long-term and short-term tendency of PV power generation, and Bayesian optimization can optimize combined features to improve the prediction accuracy of PV power generation.

The proposed method in this paper shows that the forecast performance on rainy days is relatively poorer than that on cloudy and sunny days. Therefore, the prediction performance of PV power on rainy days should be further improved in the future research work.

REFERENCES

[1] D. Gielena, F. Boshella, and D. Saygin, "The role of renewable energy in the global energy transformation," *Energy Strategy Rev.*, vol. 24, pp. 38–50, Apr. 2019.

[2] C. S. Kaunda, C. Z. Kimambo, and T. K. Nielsen, "Hydropower in the context of sustainable energy supply: A review of technologies and challenges," *Int. Scholarly Res. Netw.*, vol. 2012, pp. 1–15, Dec. 2012.

[3] Q. Xun and Y. Liu, "Evaluation of fluctuating voltage topology with fuel cells and supercapacitors for automotive applications," *Int. J. Energy Res.*, vol. 43, no. 9, pp. 1–13, Jun. 2019.

[4] S. Abdelhady, M. S. Abd-Elhady, and M. M. Fouad, "An understanding of the operation of silicon photovoltaic panels," *Energy Proc.*, vol. 113, pp. 466–475, May 2017.

[5] F. Wang, Z. Zhen, Z. Mi, H. Sun, S. Su, and G. Yang, "Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting," *Energy Build*, vol. 86, pp. 427–438, Jan. 2015.

[6] *Solar Energy*. Accessed: Aug. 12, 2019. [Online]. Available: https://www.iea.org/topics/renewables/solar/

[7] R. Zhu, W. Guo, and X. Gong, "Short-term photovoltaic power output prediction based on k-fold cross-validation and an ensemble model," *Energies*, vol. 12, no. 7, pp. 1220–1234, Mar. 2019.

[8] A. Şağlam, B. Oral, and S. Görgülü, "Measurements of meteorological parameter effects on photovoltaic energy production," *Int. J. Circuits, Syst. Signal Process.*, vol. 9, pp. 240–246, Jan. 2015.

[9] W. Xiao, J. Dai, H. Wu, G. Nazario, and F. Cheng, "Effect of meteorological factors on photovoltaic power forecast based on the neural network," *RSC Adv.*, vol. 7, no. 88, pp. 55846–55850, 2017.

[10] C. Sirisamphanwong, N. Ketjoy, and C. Sirisam-phanwong, "The effect of average photon energy and module temperature on performance of photovoltaic module under Thailand's climate condition," *Energy Procedia*, vol. 56, pp. 359–366, Jan. 2014.

[11] S. You, G. Kou, Y. Liu, X. Zhang, Y. Cui, M. J. Till, W. Yao, and Y. Liu, "Impact of high PV penetration on the inter-area oscillations in the U.S. eastern interconnection," *IEEE Access*, vol. 5, pp. 4361–4369, 2017.

[12] C. Arbizu-Barrena, J. A. Ruiz-Arias, F. J. Rodríguez-Benítez, D. Pozo-Vázquez, and J. Tovar-Pescador, "Short-term solar radiation forecasting by advecting and diffusing MSG cloud index," *Solar Energy*, vol. 155, pp. 1092–1103, Oct. 2017.

[13] F. Wang, Z. Zhen, C. Liu, Z. Mi, B.-M. Hodge, M. Shafie-Khah, and J. P. S. Catalão, "Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting," *Energy Convers. Manage.*, vol. 157, pp. 123–135, Feb. 2018.

[14] F. Wang, K. Li, X. Wang, L. Jiang, J. Ren, Z. Mi, M. Shafie-khah, and J. P. S. Catalã, "A distributed PV system capacity estimation approach based on support vector machine with customer net load curve features," *Energies*, vol. 11, no. 7, pp. 1750–1768, Jul. 2018.

[15] K. Brecl and M. Topič, "Photovoltaics (PV) system energy forecast on the basis of the local weather forecast: Problems, uncertainties and solutions," *Energies*, vol. 11, no. 5, p. 1143, 2018.

[16] M. Awad, "Forecasting of chaotic time series using RBF neural networks optimized by Genetic algorithms," *Int. Arab J. Inf. Technol.*, vol. 14, no. 6, pp. 826–834, Jan. 2017.

[17] H. R. Maier and G. C. Dandy, "Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications," *Environ. Model. Softw.*, vol. 15, pp. 101–124, Jan. 2000.

[18] F. Wang, Z. Mi, S. Su, and H. Zhao, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters," *Energies*, vol. 5, no. 5, pp. 1355–1370, May 2012.

[19] X. Zhang, B. Jiang, X. Zhang, F. Fang, Z. Gao, and T. Feng, "Solar photovoltaic power prediction based on similar day approach," in *Proc. 36th Chin. Control Conf. (CCC)*, Dalian, China, Jul. 2017, pp. 10634–10639.

[20] L. P. David, L. Nonnenmacher, and C. F. M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renew. Energy*, vol. 91, pp. 11–20, Jun. 2016.

[21] J. Zeng and W. Qiao, "Short-term solar power prediction using a support vector machine," *Renew. Energy*, vol. 52, pp. 118–127, Apr. 2013.

[22] A. Shakya, S. Michael, C. Saunders, D. Armstrong, P. Pandey, S. Chalise, and R. Tonkoski, "Using Markov switching model for solar irradiance forecasting in remote microgrids," in *Proc. IEEE Energy Convers. Congr. Expo. (ECCE)*, Sep. 2016, pp. 895–905.

[23] A. Almeida and G. Azkune, "Predicting human behaviour with recurrent neural networks," *Appl. Sci.*, vol. 8, no. 2, pp. 305–317, 2018.

[24] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications*, vol. 60, no. 6, pp. 84–90, 2012.

[26] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. NIPS*, 2017, pp. 1–11.

[27] S. Rasp and S. Lerch, "Neural networks for postprocessing ensemble weather forecasts," *Monthly Weather Rev.*, vol. 146, no. 11, pp. 3885–3900, 2018.

[28] J. Zhang, Y. Chi, and L. Xiao, "Solar power generation forecast based on LSTM," in *Proc. IEEE 9th Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Nov. 2018, pp. 869–872.

[29] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep learning for solar power forecasting—An approach using AutoEncoder and LSTM neural networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Budapest, Hungary, Oct. 2016, pp. 2858–2865.

[30] G. W. Chang and H. J. Lu, "Integrating grey data preprocessor and deep belief network for day-ahead PV power output forecast," *IEEE Trans. Sustain. Energy*, to be published, doi: 10.1109/TSTE.2018.2888548.

[31] S. Ogawa and H. Mori, "Application of evolutionary deep neural netwok to photovoltaic generation forecasting," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Sapporo, Japan, May 2019, pp. 1–4.

[32] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–15.

[33] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Appl. Sci.*, vol. 8, no. 8, p. 1286, 2018.

[34] M. Sun, T. Zhang, Y. Wang, G. Strbac, and C. Kang, "Using Bayesian deep learning to capture uncertainty for residential net load forecasting," *IEEE Trans. Power Syst.*, to be published, doi: 10.1109/TPWRS.2019.2924294.

[35] G. Li, H. Wang, S. Zhang, J. Xin, and H. Liu, "Recurrent neural networks based photovoltaic power forecasting approach," *Energies*, vol. 12, no. 13, p. 2538, 2019.

[36] A. Ahmad, T. N. Anderson, and T. T. Lie, "Hourly global solar irradiation forecasting for New Zealand," *Sol. Energy*, vol. 122, pp. 1398–1408, Dec. 2015.

[37] M. Husein and I. Y. Chung, "Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach," *Energies*, vol. 12, no. 10, p. 1856, 2019.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[39] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyperparameter optimization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Granda, Spain, 2011, pp. 1–9.

[40] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2005.

[41] C. Yang, A. A. Thatte, and L. Xie, ''Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation,'' *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 104–112, Jan. 2015.

[42] W. Jidong, R. Ran, S. Zhilin, and J. Sun, ''Short-term photovoltaic power generation forecasting based on environmental factors and GA-SVM,'' *J. Elect. Eng. Technol.*, vol. 12, no. 1, pp. 64–71, 2017.

BIN LI received the masters' degree from the Changsha University of Science and Technology, China, in 2019. He is currently with the School of Mechanical and Electrical Engineering, Hunan City University. His current research interests include power system informatization, planning, and new energy.

TONGGUANG YANG received the Ph.D. degree from Central South University, China, in 2013. He is currently a Professor with the School of Mechanical and Electrical Engineering, Hunan City University. His current research interests include the grid-connected inverters control for new energy and fault diagnosis for induction motor.
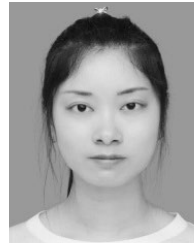
QIAN XUN received the B.Eng. degree from Hohai University, in 2012, and the M.Sc. degree from the Nanjing University of Aeronautics and Astronautics, China, in 2015. She is currently pursuing the Ph.D. degree with the Chalmers University of Technology, Sweden. She was a Lecturer with Huzhou University, China, for two years. Her doctoral project is about the research on the drivetrain design for fuel cell electric vehicle.

• • •