IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 24 April 2024, accepted 3 May 2024, date of publication 8 May 2024, date of current version 15 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3398033

RESEARCH ARTICLE

RSMD-RF-BGSkip Based PV Generation Prediction Method

GUOMIN XIE AND ZHONGBAO LIN
Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China
Corresponding author: Zhongbao Lin (lzblntu@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51974151.

ABSTRACT With the increasing use of photovoltaic (PV) power generation, power forecasting has become equally important to maintain stable and economic operation of the power system. However, the high frequency component of the PV power time series reduces the accuracy of the model predictions. Therefore, this paper proposes a short-term prediction model for PV power based on Row Secondary Modal Decomposition (RSMD), Random Forest (RF), and BGSkip neural network. Firstly, modal features with different complexity are obtained by RSMD. Secondly, RF is used to select different modal features. Then, to improve the prediction performance of the model, the BGSkip model employs a hybrid neural network to accurately predict the nonlinear part, while the linear part is handled by an autoregressive model. The prediction results of these two parts are integrated through the BGSkip model to output more accurate prediction values. Finally, the historical data of PV power plants in a region of Liaoning is utilized for experimental validation. The experimental results show that R^2, E_{MAPE} and E_{RMSE} of the RSMD-RF-BGSkip short-term forecasting model are improved by 7.55%, 0.261% and 16.01% respectively compared with the most advanced models, which has higher forecasting accuracy.

INDEX TERMS PV power, row secondary modal decomposition, random forest, neural network.

NOMENCLATURE

This article uses the following nomenclature throughout.

RNN	Recurrent Neural Network.
LSTM	Long Short-Term Memory.
GRU	Gated Recurrent Unit Network.
EEMD	Ensemble Empirical Mode Decomposition.
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise.
PSO	Particle Swarm Optimization.
VMD	Variational Mode Decomposition.
TransNN	Transform Neural Networks.
CNN	Convolutional Neural Networks.
PE	Permutation Entropy.
RF	Random Forest.
IMF	Intrinsic Modal Function.
NMD	No Modal Decomposition.
EMD	Empirical Model Decomposition.
PMD	Primary Modal Decomposition.
SMD	Secondary Modal Decomposition.
RSMD	Row Secondary Modal Decomposition.
RSMD-RF	Row Secondary Mode Decomposition combined with Random Forest.
$*$	Convolution operation.
K	The number of components obtained by decomposition.
$X(t)$	Original sequence signal.
K	The number of categories.
k	Feature category.
p	The number of jump BiGRU models.
h_t^R	Output of the cyclic layer.
h_{t-p}^S	Output of the cyclic skip laye.
Y_t	Final prediction result.
R^2	Coefficient of Determination.
E_{MAPE}	Mean Absolute Percentage Error.
E_{RMSE}	Root Mean Square Error.

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

I. INTRODUCTION

In the context of global warming, renewable energy sources such as solar, wind, hydropower and geothermal energy have become crucial [1]. Although photovoltaic (PV) power generation technology is widely used in smart grids, however, the output power of the PV power generation system suffers from instability and volatility because of environmental and meteorological factors [2]. Therefore, accurate prediction of PV power generation has become a hot issue at present [2].

Currently, short-term PV power prediction methods are mainly categorized into two main groups: statistical learning algorithms and machine learning algorithms. In statistical learning algorithms, such as multivariate linear regression [3], exponential smoothing [4], etc., although these methods are computationally simple and fast, they are poorly robust and easily interfered with by stochastic factors, which reduces their reliability in PV power prediction. Machine learning algorithms include Support Vector Machine [5], Extreme Gradient Augmentation [6], Random Forest [7], etc., which are more suitable for dealing with nonlinear and multidimensional data but are sensitive to outliers and are prone to performance degradation in the face of sudden data changes or breakpoints [8].

In recent years, deep learning has attracted much attention in the field of PV power prediction. Recurrent Neural Network [9] (RNN), Long Short-Term Memory [10] (LSTM) and Gated Recurrent Unit Network [11] (GRU) have achieved remarkable results in time series prediction. Although in paper [12] achieved high accuracy in PV power prediction using RNN models, it faced the problems of gradient vanishing and gradient explosion. In paper [13], LSTM is used for PV prediction, but the forecasting accuracy is not high under special weather conditions and there is an overfitting problem. In addition, the LSTM model has a complex structure, many parameters and long training time. In paper [14], a GRU model is proposed, who is another RNN gating architecture after inheriting LSTM, which is due to its simple structure and few training parameters. GRU not only has high prediction accuracy, but also alleviates the LSTM overfitting problem [15], [16], [17]. Compared with statistical models and machine learning models, the prediction accuracy of GRU model has been greatly improved, but it still cannot tap the fully mined sequence features [18].

Given the high complexity and nonlinear characteristics of PV series, single algorithm prediction is still difficult. Signal decomposition techniques such as Wavelet Transform (WT) [19], Ensemble Empirical Mode Decomposition (EEMD) [20], Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [21], and Variational Modal Decomposition (VMD) [22] are widely used in the field of PV power prediction. In paper [23], EEMD algorithm is used to decompose the weather sequence into components of different frequencies, which improves the prediction accuracy. However, the high-frequency components generated by the first decomposition are non-smooth components, and there are large errors in direct prediction. In paper [24], CEEMDAN is used to decompose the sequence, which effectively reduces the residual noise of the data, but it may ignore some of the high-frequency components, which leads to an increase in the prediction error. In paper [25], CEEMDAN is used in combination with alignment entropy to determine the complexity of each component, which effectively improves the prediction accuracy. In paper [26], Particle Swarm Optimization (PSO) is used to search for the optimal parameter combinations of VMD, so that it can automatically adjust the parameters for energy data with different characteristics, reduce the human error, and improve the stability and accuracy of the prediction model. In paper [27], a combined deep learning prediction method is proposed. The method first extracts the trend features of PV power using VMD technique, and then selects the optimal input feature set by Fast Correlation Based Filter (FCBF) to reduce the prediction error caused by redundant features. Experimental results show that the method significantly improves the prediction results. In paper [28], Transform Neural Networks (TransNN) and Convolutional Neural Networks (CNN) are combined and VMD decomposition is used in the data preprocessing stage. The combination of this model can significantly improve the accuracy of the prediction model. Although papers [23], [24], [25], [26], [27], [28] have achieved better results by combining signal decomposition techniques with deep learning, there is still the problem of high-frequency modal components generated by a single decomposition technique, as well as the troubling problem of low prediction accuracy of the model itself, which are yet to be solved. Therefore, the comprehensive use of multiple algorithms and signal decomposition techniques may be an effective way to improve the accuracy of PV power prediction [29].

To solve the problem of insufficient prediction accuracy caused by high-frequency modal components in the modal decomposition of PV power generation sequences, this paper innovatively proposes the **RSMD-RF-BGSkip** model, which is a short-term PV power prediction model. Firstly, the original data are decomposed twice consecutively by Variational Mode Decomposition (VMD), and the complexity is calculated using Permutation Entropy (PE), which skillfully fuses similar features. Subsequently, Random Forest (RF) is applied for feature screening to further optimize the input features of the model. Finally, the high coupling feature matrix is constructed for each subsequence, which is input into the BGSkip model for accurate prediction. The accuracy of PV power prediction is effectively improved by integrating the prediction results of each component through the fully connected layer.

The main contributions of this paper are summarized below:

(1) In this paper, a short-term forecasting model of photovoltaic power based on Row Secondary Modal Decomposition (RSMD), Random Forest (RF) and BGSkip Neural Network is proposed.

(2) Firstly, modal features with different complexity are obtained by RSMD. Secondly, RF is used to select different modal features. Then, to improve the prediction performance of the model, the BGSkip model employs a hybrid neural network to accurately predict the nonlinear part, while the linear part is handled by an autoregressive model.

(3) The historical data of a photovoltaic power station in a certain area of Liaoning Province is used for experimental verification. The experimental results show that RSMD-RF-BGSkip short-term forecasting model has higher forecasting accuracy than the most advanced mode.

II. DATA PREPROCESSING

A. VARIATIONAL MODE DECOMPOSITION

Variational mode decomposition (VMD) is a new method for non-recursive signal adaptive processing [30]. The algorithm decomposes the PV sequence into Intrinsic Modal Function (IMF) with frequencies ranging from high to low, and each IMF embodies the characteristics of the original data in different frequency ranges, which effectively reduces the complexity of the sequence and improves the accuracy of the subsequent model training. The VMD decomposition process is as follows [30]:

(1) Constructing constrained variational optimization problems for intrinsic modal functions:

$$\begin{cases} \min\limits_{\{u_k\},\{w_k\}} \left\{ \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta_t + \dfrac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k^t} \right\|_2^2 \right\} \\ s.t. \sum\limits_{k=1}^{K} u_k(t) = X(t) \end{cases} \tag{1}$$

where: $*$ is the convolution operation; K is the number of components obtained by decomposition; u_k and w_k are the modal components and center frequency, respectively; $X(t)$ represents the original sequence signal.

(2) The problem is turned into an unconstrained variational model by means of penalty factors α and Lagrange multiplier operators λ:

$$\begin{aligned} &L(\{u_k\},\{w_k\},\{\lambda(t)\}) \\ &= \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta_t + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k^t} \right\|_2^2 \\ &\quad + \left\| G_{CIMF}(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 \\ &\quad + \left\langle \lambda(t), G_{CIMF}(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \end{aligned} \tag{2}$$

(3) The optimized objective function is iteratively updated with u_k and w_k to obtain a specified number of modal components IMF.

B. PERMUTATION ENTROPY

Permutation Entropy (PE) can accurately represent the degree of complexity and mutability of time series [31]. With its fast calculation speed and strong noise resistance, it has good applicability and strong sensitivity to the dynamic nonlinear data of photovoltaic power generation. In this paper, we use the alignment entropy to calculate the complexity of each subsequence obtained by VMD decomposition, and the similar features are fused. The PE calculation process is as follows [31]:

(1) Reconstruct the time series feature matrix $\{x(i), i = 1, 2 \cdots n\}$ to obtain the reconstruction matrix X.

$$\left\{ \begin{array}{l} X(1) = \{x(1), x(1+\tau), \cdots, x(1+(m-1)\tau)\} \\ \vdots \\ X(i) = \{x(i), x(i+\tau), \cdots, x(i+(m-1)\tau)\} \\ \vdots \\ X(N) = \{x(N), x(N+\tau), \cdots, x(N+(m-1)\tau)\} \end{array} \right\} \tag{3}$$

where: m is the embedding dimension; τ is the delay time; N is the number of reconstruction components.

(2) Each row in the matrix X is a reconstructed component, and a new set of time series is obtained for each row x in ascending order.

(3) For any reconstructed component a new sequence can be obtained.

$$s(k) = \{j_1, j_2, \cdots, j_m\} \tag{4}$$

where: $k = 1, 2, \ldots, g$, $g < m!$ will produce a sequence of $m!$ species $s(k)$. Calculating the probability distribution P_k of the labeled sequence at this point yields a probability sum of 1.

(4) The entropy of the permutation of the time series $x(i)$ can be defined as

$$H_p(m) = -\sum_{k=1}^{g} P_k \ln P_k \tag{5}$$

where: P_k is the probability of occurrence of one of the $s(k)$ sequences.

C. RANDOM FOREST

Randomized Forest (RF) is an integrated learning method, and its core idea is to form a forest by constructing multiple weakly categorized regression trees, and finally get the comprehensive results by voting [32]. Compared with traditional feature screening methods such as Pearson correlation coefficient and chi-square test, randomized deep forest performs more powerfully in dealing with nonlinear features. Its advantage lies in its ability to better capture nonlinear patterns and its excellent scalability, which makes it suitable for handling large-scale data and many features. Feature select by Random Forest can reduce the interference to the model, especially reduce the influence of features with weak correlation, to improve the generalization ability of the model.

Where the Gini index G_m is commonly used to measure the purity of the dataset after feature partitioning as a way of

assessing the importance of the features, which is expressed as [32]:

$$G_m = 1 - \sum_{k=1}^{|K|} p_{mk}^2 \tag{6}$$

where: K represents the number of categories; k is a feature category; p_{mk} indicates the proportion of category k in node m.

III. BGSKIP NEURAL NETWORK MODEL

BGSkip consists of a linear part and a nonlinear part. In this case, the linear part is modeled by autoregressive modeling, while the nonlinear part consists of convolutional, cyclic, and cyclic skip layers. After the results of the nonlinear part are obtained, they are fused through full connect layers. Finally, the linear prediction is superimposed with the nonlinear results to get the final prediction. The BGSkip neural network model is shown in Figure 1.

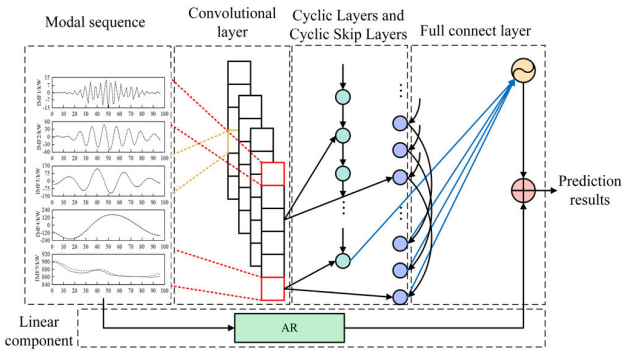


FIGURE 1. BGSkip neural network model.

A. CONVOLUTIONAL LAYER

CNN is constructed by mimicking biological visual perception mechanisms and is capable of both supervised and unsupervised learning [33]. The sharing of convolutional kernel parameters in the implicit layers and the sparsity of inter-layer connections enable CNN to extract deep local features from high-dimensional data with less computational effort and to obtain effective representations through the convolutional and pooling layers. Therefore, to fully extract modal sequence features as well as capture local dependencies between variables, the first layer of the BGSkip model uses a CNN model without pooling layers. This design aims to maximize the advantages of CNN while ensuring effective extraction of modal sequence features and local dependencies.

B. CYCLIC LAYER

When the number of time steps is large, the historical gradient information of RNN cannot be maintained in a reasonable range all the time, so gradient decay or explosion is almost inevitable, which leads to the RNN will be very difficult to capture effective information from long distance sequences [7]. LSTM, as a special kind of RNN, is proposed to solve the problem of gradient vanishing in RNN very well [34]. GRU, on the other hand, is proposed based on LSTM, which has a simpler structure, fewer parameters, shorter training time, and faster training speed than LSTM.

BiGRU is a deformation of recurrent neural network, which is especially suitable for processing time series with long intervals [35]. In this paper, the output of the convolutional layer is input to both cyclic layer and cyclic skip layer. In addition, to realize the capture of long-term dependencies, BiGRU is used as the recurrent unit of cyclic layer. The computation of the cyclic layer is as follows:

$$\begin{cases} \overrightarrow{h}_t = GRU(X_t, \overleftarrow{h}_{t-1}) \\ \overleftarrow{h}_t = GRU(X_t, \overrightarrow{h}_{t-1}) \\ h_t = \omega_t \overrightarrow{h}_t + v_t \overleftarrow{h}_t + b_t \end{cases} \tag{7}$$

where: h_t is the hidden layer state at time t; $\overrightarrow{h}$ is the forward hidden layer output at time t; $\overleftarrow{h}$ is the reverse hidden layer output at time t; ω_t and v_t represent the corresponding weights; and b_t denotes the bias.

C. CYCLIC SKIP LAYER

By introducing a cycle layer with BiGRU units, relationships can be effectively captured in historical information. The BiGRU model suffers from the problem of vanishing gradient, which leads to low model prediction accuracy [36]. To solve this problem, this paper mitigates it by adding a cyclic skip layer. The cyclic skip layer updating process is represented as follows:

$$\begin{cases} \overrightarrow{h}_t = GRU(X_t, \overleftarrow{h}_{t-p}) \\ \overleftarrow{h}_t = GRU(X_t, \overrightarrow{h}_{t-p}) \\ h_t = \omega_t \overrightarrow{h}_t + v_t \overleftarrow{h}_t + b_t \end{cases} \tag{8}$$

where: p is the number of jump BiGRU models.

The nonlinear part of the prediction is obtained by using a fully connected layer that combines the outputs of the cyclic and cyclic skip layers.

$$h_t^D = W^R h_t^R + \sum_{i=0}^{p-1} W_i^S h_{t-p}^S \tag{9}$$

where: h_t^R is the output of the cyclic layer; h_{t-p}^S is the output of the cyclic skip layer; h_t^D is the nonlinear part of the prediction at time t.

D. AUTOREGRESSIVE LAYER

Since the nonlinear nature of the convolutional and cyclic layers can make the output data insensitive to the input data and reduce the prediction accuracy of the neural network for the non-periodic variation data in the sequence, the Autoregressive (AR) model is used to predict the linear part of

the load data. The prediction results of the AR layer are as follows:

$$h_{t,i}^{L} = \sum_{k=0}^{q^{ar}-1} W_k^{ar} y_{t-k,i} + b^{ar} \tag{10}$$

where: q^{ar} is the number of windows; h_t^L is the autoregressive model outputs.

The advantages of neural networks and AR models are considered, and the integration of both is carried out and the final prediction results are obtained.

$$Y_t = h_t^D + h_t^L \tag{11}$$

where: Y_t is the final prediction result at time t.

IV. PREDICTION MODELING BASED ON RSMD-RF-BGSKIP

As the PV power is greatly affected by meteorological factors, it shows strong volatility and randomness. The traditional primary modal decomposition method may have a problem of too many high-frequency components, which makes it difficult to analyze the PV power generation accurately for prediction. Therefore, the RSMD-RF-BGSkip prediction model is proposed in this paper. The prediction process of this model is shown in Figure 2.

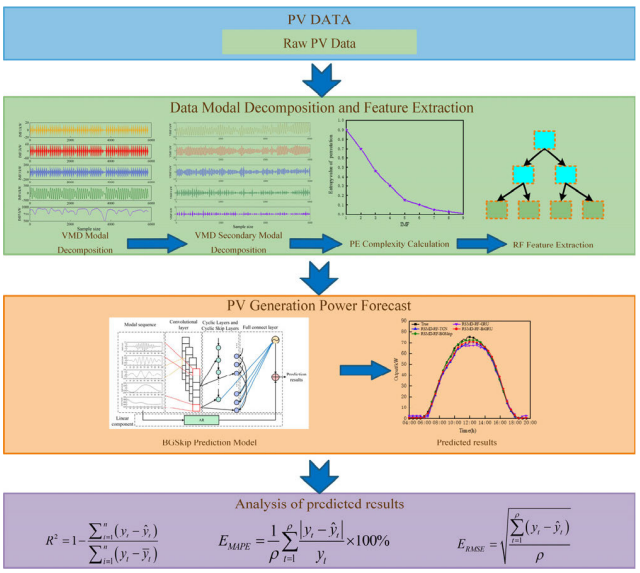


FIGURE 2. RSMD-RF-BGSkip prediction flow.

V. CASE STUDY

In this paper, the PV power generation data of a PV power station of the national grid in Liaoning region from September 14, 2018 to August 17, 2020 at the time period of 4:00-20:00 every day is used as the input data, and the step size is 15min. All experiments are completed on the Pycharm platform with a computer configuration of NVDIA RTX 3060 GPU; CPU AMD 5800H; RAM 16GB.

A. EVALUATION PERFORMANCE INDICATORS

The predictive effectiveness of the model was assessed using Coefficient of Determination R^2, Mean Absolute Percentage Error E_{MAPE} and Root Mean Square Error E_{RMSE} [37], which are given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_t - \hat{y}_t)}{\sum_{i=1}^{n} (y_t - \bar{y}_t)} \tag{12}$$

$$E_{MAPE} = \frac{1}{\rho} \sum_{t=1}^{\rho} \frac{|y_t - \hat{y}_t|}{y_t} \times 100\% \tag{13}$$

$$E_{RMSE} = \sqrt{\frac{\sum_{t=1}^{\rho} (y_t - \hat{y}_t)}{\rho}} \tag{14}$$

where: ρ is the number of samples; y_t, $\bar{y}_t$, $\hat{y}_t$ are the real value of the data at the moment of t, the average value of the real value and the predicted value, respectively. The larger the value of R^2, the better the prediction, and the smaller the values of E_{MAPE} and E_{RMSE}, the better the prediction.

B. MODAL DECOMPOSITION AND FEATURE SELECTION

In short-term PV power generation forecasting, the characteristic factors such as temperature and radiation intensity both have an important impact on the forecasting results. Radiation intensity has the most significant impact on PV power generation. The geographical location of the data selected in this paper is special and has a large weather change, which makes the PV power generation power have a large volatility. Table 1 describes the specific information of the input features, which include five climate factors and historical power generation. Considering the importance of historical power generation for future PV power prediction, this paper adds it to the model as a characterization factor.

TABLE 1. Input feature information.

Feature category	Feature	Characterization
Meteorological feature	Temperature	Average temperature per 15 minutes
	Humidity	Average humidity per 15 minutes
	Pneumatic	Average pneumatic per 15 minutes
	Quantity of rainfall	Average rainfall per 15 minutes
	Radiation intensity	Average radiation intensity per 15 minutes
Other feature	Historical PV generation	Average power generation per 15 minutes

Before decomposing the original PV sequence using VMD, the number of IMF of the decomposition needs to be specified. In this paper, the K value is determined by the center frequency observation method [38]. The best results are obtained when $K = 5$ at the first modal decomposition. The results of one VMD modal decomposition are shown in Figure 3.

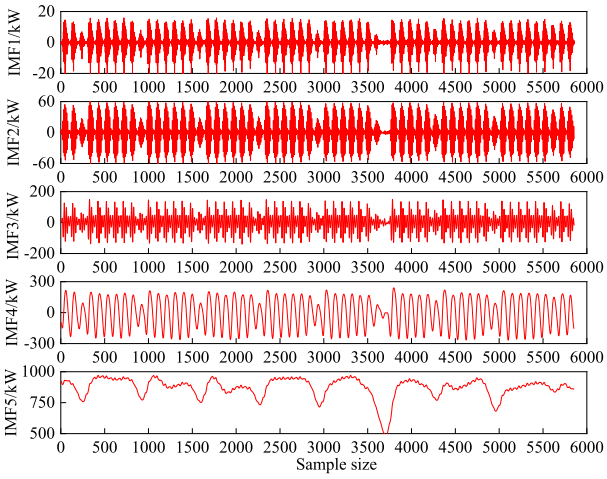


FIGURE 3. Results of one VMD modal decomposition.

The analysis in Figure 3 shows that the complexity of the modal components decomposed by the primary VMD is high, and it is difficult to maintain high prediction accuracy by direct prediction, so the VMD is used again to decompose them in the secondary decomposition, and the modal components generated by the secondary decomposition are VIMF1-VIMF9. The second modal decomposition is still determined by using the center frequency observation, and the best results are obtained when using the center frequency observation K =9. The results of the second VMD modal decomposition are shown in Figure 4.

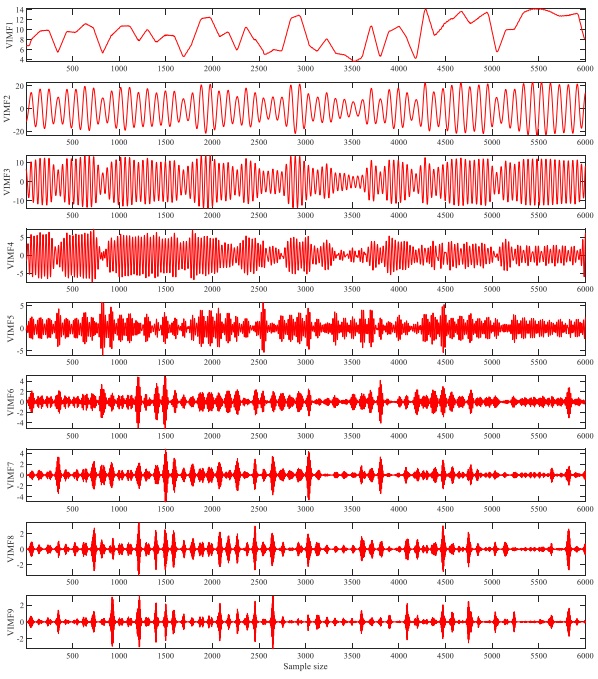


FIGURE 4. Second VMD modal decomposition results.

Due to the excessive number of VIMF components, inputting each component into the BGSkip model for prediction analysis will increase the computation of the model. Therefore, to reduce the model computation, the complexity of all modes is analyzed by using the permutation entropy, and the similar arrangement entropy values are superimposed, and the permutation entropy values of each modal component are shown in Figure 5.

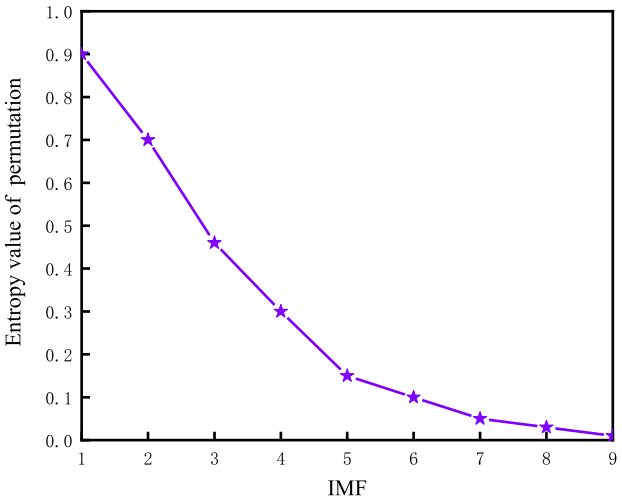


FIGURE 5. Entropy value of each VIMF component permutation.

As can be seen from Figure 5, the entropy value of the samples of each modal component decreases gradually. After the calculation of arrangement entropy, those with similar entropy values are superimposed to achieve the purpose of reducing the modal components. The superimposed modal components are shown in Figure 6.

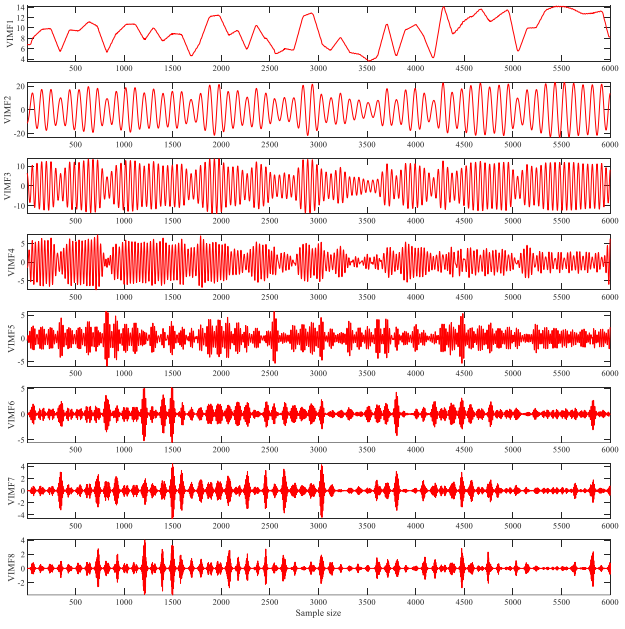


FIGURE 6. Modal fractional energy results after superposition.

After the original time series are subjected to row second modal decomposition, the correlation information between the meteorological factors and each modal component changes. Therefore, feature selecting of individual modal components is required to reduce unnecessary features and improve forecast accuracy. The results of the Random Forest feature selecting are shown in Figure 7, and VIMF1-VIMF8 represent the components obtained by the secondary modal decomposition of IMF1-IMF5 rows. The Random Forest feature selecting method can more closely correlate meteorological factors and modal components, thus improving forecast performance and reducing information loss.

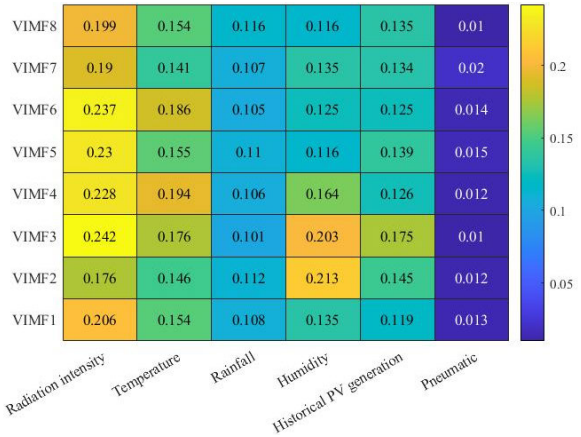


FIGURE 7. Feature selection results.

From Figure 7, Radiation intensity, temperature and humidity all have a large impact on the components. Among them, radiation intensity has the highest correlation coefficient, indicating that radiation intensity has a more significant effect on PV power prediction. The correlation coefficients of temperature, humidity, historical PV power generation and rainfall are mostly between 0.1 and 0.2, which also have a strong correlation on PV power. While pneumatic has low correlation, this feature is removed to reduce the computational effort of the model.

C. ROW SECONDARY MODAL DECOMPOSITION COMPARISON TEST

To analyze the effect of each module in the row secondary modal decomposition framework on the overall prediction performance, the last 48 hours of data in the 2018 dataset are selected for the comparison experiment. In comparing the effects of different modal decomposition methods, this paper comparatively analyzes seven scenarios: No Modal Decomposition (NMD), Empirical Model Decomposition (EMD), CEEMDAN, Primary Modal Decomposition (PMD), Secondary Modal Decomposition (SMD), Row Secondary Modal Decomposition (RSMD) and Row Secondary Mode Decomposition combined with Random Forest (RSMD-RF). Their prediction results are compared, and the experimental results are shown in Table 2 and Figure 8.

TABLE 2. Experimental errors for different modal decompositions.

Prediction method	R^2	E_{MAPE} / %	E_{RMSE} / kW
NMD	0.833	1.116	8.311
EMD	0.887	0.839	6.565
CEEMDAN	0.890	0.812	6.411
PMD	0.862	0.932	7.513
SMD	0.891	0.796	6.385
RSMD	0.924	0.691	5.846
RSMD-RF	0.995	0.630	5.133

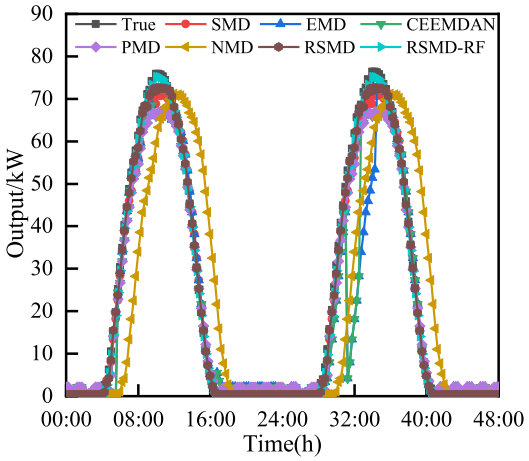


FIGURE 8. Comparative prediction results for different modal decompositions.

Table 2 shows that the prediction error of no modal decomposition is the largest; after the primary modal decomposition, the prediction accuracy is significantly improved, R^2 is improved by 3.25%, and E_{MAPE} and E_{RMSE} are reduced by 14.60% and 15.01%, respectively; after secondary modal decomposition, the prediction accuracy is further improved; and the prediction accuracy is the highest when RSMD-RF is performed. Moreover, RSMD-RF method compared to other modal decomposition methods VMD, CEEMDAN its R^2 is improved by at least 11.79%; E_{MAPE} and E_{RMSE} are reduced by 28.89% and 24.99%, respectively.

Analyzing Figure 6, due to the strong volatility and randomness of PV power generation, the direct prediction error is large. Although performing primary modal decomposition can effectively improve the prediction accuracy, it is often difficult to predict the high-frequency stochastic components generated by the decomposition and has limited ability to deal with the prediction of small-scale fluctuations in the sequence. By further decomposing the high-frequency components generated by the one-time decomposition, multiple smooth components are obtained. Then the overall model prediction accuracy is significantly improved after using permutation entropy for complexity analysis and combining with Random Forest for feature selection.

D. CONTRASTIVE OF DIFFERENT MODEL PREDICTIONS

To verify the performance advantages of the RSMD-RF-BGSkip prediction model, the prediction model proposed in

this paper is compared with BiGRU, GRU, and TCN, and the 2018-2019 data in the dataset is intercepted for experimentation, with the length of 17,542 data, and the last 2 days of 2019 is used as the test set. To control the variables, the input data of all models are subjected to the same preprocessing and feature screening methods and are trained with the optimal parameters shown in Table 3. The prediction results of each model combined with the framework of this paper are shown in Figure 9 and the prediction errors are shown in Table 4.

From Table 4, the E_{MAPE} of this method is reduced by 64.91%, 60.03%, and 57.83%, and the E_{RMSE} is reduced by 65.28%, 60.20%, and 56.23% compared to GRU, TCN, and BiGRU, respectively, and the prediction accuracy is improved by a factor of 1-2. The prediction effect of each model is shown in Figure 9, when the RSMD-RF decomposed features show the best results when combined with the BGSkip model, while the BiGRU, GRU, and TCN models are somewhat less capable in PV power prediction.

TABLE 3. Parameter settings for each model.

Model	Parameter setting	Value
BGSkip	Input step length	8
	Output step length	1
	Number of hidden layer	64
	Learning rate	0.0001
	Batch size	64
	Dropout rate	0.05
	Optimizer	Adam
	p	4
BiGRU	Input step length	8
	Output step length	1
	Number of hidden layer	64
	Learning rate	0.0001
	Batch size	64
	Dropout rate	0.05
	Optimizer	Adam
GRU	Input step length	8
	Output step length	1
	Number of hidden layer	64
	Learning rate	0.0001
	Batch size	64
	Dropout rate	0.05
	Optimizer	Adam
TCN	Input step length	8
	Output step length	1
	Number of filters	64
	Kernel size	3
	Learning rate	0.0001
	Batch size	64
	Optimizer	Adam

TABLE 4. Prediction errors of different models.

Prediction method	R^2	E_{MAPE} / %	E_{RMSE} / kW
RSMD-RF-GRU	0.806	1.311	9.972
RSMD-RF -TCN	0.857	1.135	8.382
RSMD-RF -BiGRU	0.918	0.873	7.339
RSMD-RF -BGSkip	0.993	0.612	6.164

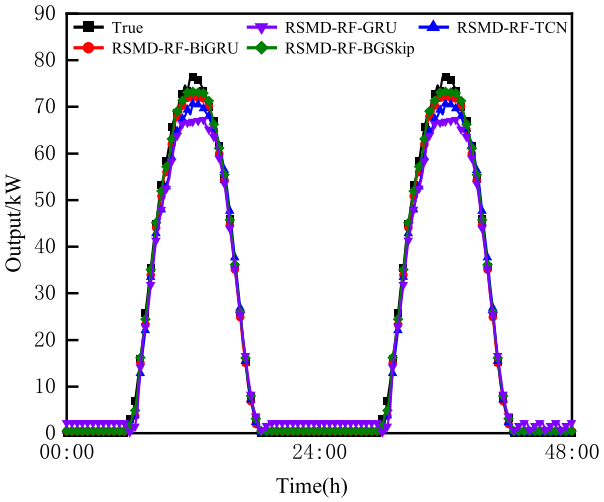


FIGURE 9. Prediction results of different models.

E. COMARATIVE ANALYSIS OF FORECAST RESULTS FOR DIFFERENT WEATHER CONDITIONS

In a bid to more comprehensively verify the applicability of the short-term PV power generation forecasting method proposed in this paper, sunny and rainy days with large weather fluctuations are selected to validate the model. To exclude the influence of different date types on the PV power generation prediction results, 2 consecutive days (September 27, 2019, and September 28, 2019) of weekdays were purposely selected for the study. Among these two days, day 1 is sunny and day 2 is rainy. Figure 10 and Figure 11 shows the prediction results under different weather conditions. The prediction errors under different weather conditions are shown in Table 5.

Analyzing Figure 10, the PV power curve fluctuates less under sunny weather conditions, and the PV power changes with a certain regularity, and the four models have a more stable prediction effect. In terms of the error index, the proposed prediction models are smaller than the other prediction models, and in the midday, the power curve has a slight fluctuation phenomenon, compared with the other three models, the proposed prediction model has the best fitting effect overall, and it can be better close to the actual curve.

TABLE 5. Prediction errors of different models.

Weather	Prediction method	R^2	E_{MAPE} / %	E_{RMSE} / kW
Sunny	RSMD-RF -GRU	0.851	0.911	7.693
	RSMD-RF -TCN	0.881	0.806	6.448
	RSMD-RF -BiGRU	0.925	0.690	5.746
	RSMD-RF -BGSkip	0.994	0.633	5.139
Rainy	RSMD-RF -GRU	0.801	1.329	9.979
	RSMD-RF -TCN	0.853	1.165	8.382
	RSMD-RF -BiGRU	0.918	0.873	7.339
	RSMD-RF -BGSkip	0.991	0.622	6.233

Figure 11 in rainy weather conditions, the PV power curve fluctuates relatively large due to a variety of factors,

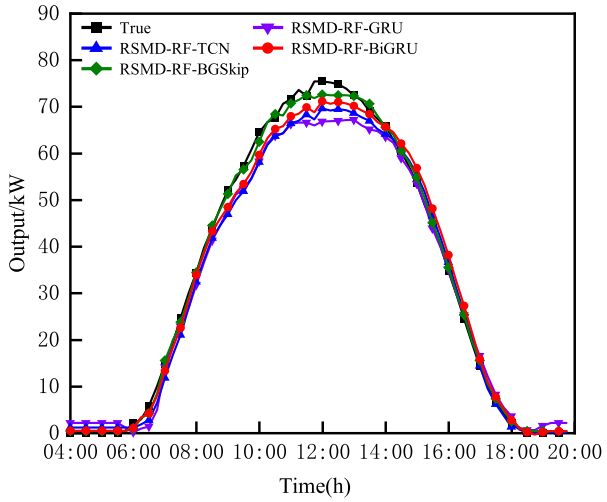


FIGURE 10. September 27th sunny day forecast results.

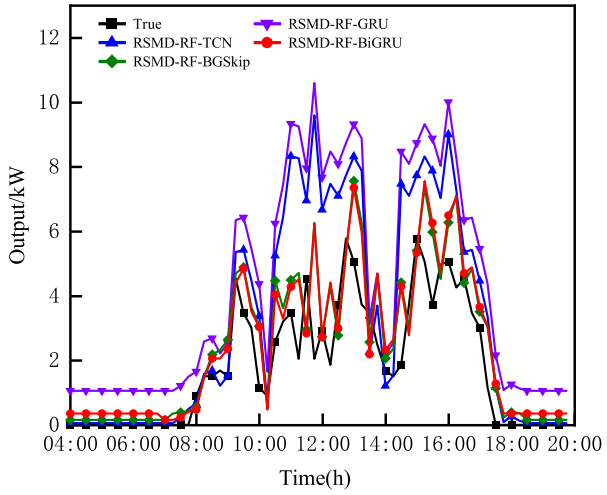


FIGURE 11. September 28th rainy day forecast results.

and the prediction curves of the models produce a certain deviation from the actual curve, in which the prediction errors of RSMD-RF-TCN, RSMD-RF-GRU and RSMD-RF-BiGRU are obviously increased, and the linear and nonlinear characteristics of data are fully extracted using the BGSkip model, and the linear and nonlinear characteristics of data are fully extracted in the event of drastic weather conditions. characteristics, the prediction performance of the model is improved during the periods of dramatic fluctuations in weather conditions. Figure 12 and Figure 13 respectively show the prediction of different methods in sunny and rainy days, and the model evaluation index values R^2, E_{MAPE}, E_{RMSE} are displayed visually, as shown below.

In the comprehensive analysis, the RSMD-RF-TCN, RSMD-RF-GRU and RSMD-RF-BiGRU models have large deviations between the predicted values and the actual values, especially the tracking of the sudden change points is not strong, while the RSMD-RF-BGSkip model has the smallest

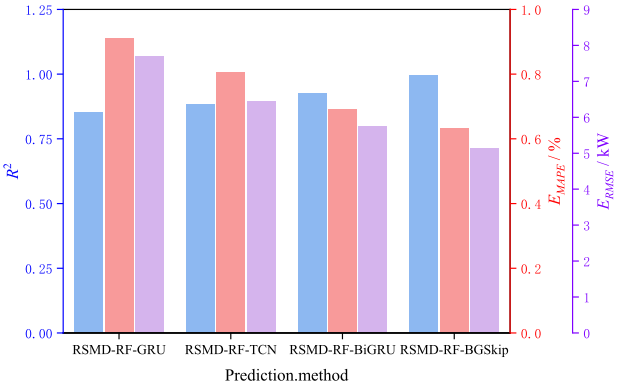


FIGURE 12. Sunny day forecast error for september 27.

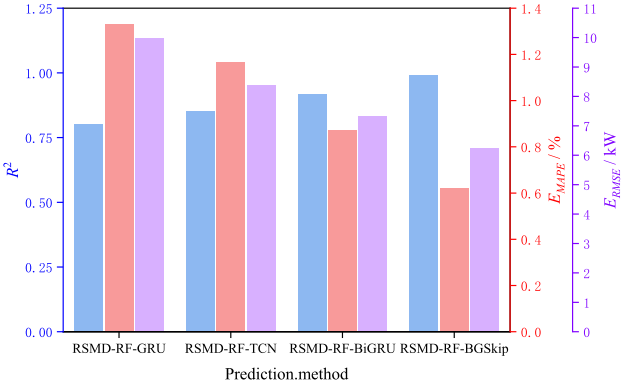


FIGURE 13. Rainy day forecast error for september 27.

deviation, and is able to make good predictions of the sudden changes in the PV power generation capacity, which shows a better prediction performance.

VI. CONCLUSION

Aiming at the problem of low prediction accuracy due to high-frequency components in the modal decomposition of short-term PV power prediction, this paper proposes a new method of RSMD-RF-BGSkip. After experimental validation, the method achieves the following conclusions:

(1) The complexity of the high-frequency component is reduced by using the row second modal decomposition and feature extraction using Random Forest for smoother and stable subsequences. The performance differences in the RSMD-RF method with the NMD, PMD, SMD and RSMD methods were analyzed through experimental comparisons. The results show that the RSMD-RF method improves the prediction accuracy by 19.45%, 15.43%, 10.4% and 7.68%, respectively, compared to these methods.

(2) The BGSkip model is applied to the subsequence to divide the sequence into linear and nonlinear parts and forecast the next 48h. Compared with the three models GRU, TCN and BiGRU, R^2 increases by 15.75%, 8.86% and 8.17%, respectively. The results show that the BGSkip

prediction model has excellent robustness and good prediction accuracy.

(3) Comparing with RSMD-RF -GRU, RSMD-RF -TCN, and RSMD-RF -BiGRU models under different weather conditions, the RSMD-RF-BGSkip model has higher prediction accuracy. In addition, by calculating the evaluation parameter R^2, at least 7% improvement of forecasting accuracy occurs.

The PV power data used in this paper did not include other energy types, and in subsequent work, the joint prediction of integrated energy sources such as wind, PV, geothermal, and biomass can be considered to further improve the applicability of the model.

REFERENCES

[1] X. Fu, X. Wang, Y. Gong, Y. Wang, and Y. Zhang, "Impact of snow weather on PV power generation and improvement of power forecasting," in *Proc. Int. Conf. Power Energy Syst. Appl. (ICoPESA)*, Nanjing, China, Feb. 2023, pp. 448–453, doi: 10.1109/ICoPESA56898.2023.10140199.

[2] D. P. Simatupang and J. Choi, "Integrated photovoltaic inverters based on unified power quality conditioner with voltage compensation for submarine distribution system," *Energies*, vol. 11, no. 11, p. 2927, Oct. 2018, doi: 10.3390/en11112927.

[3] Y. Su, Y. Sun, H. Yang, H. Ren, Z. Zhen, and F. Wang, "A short-term PV power forecasting method based on NWP correction considering meteorological coupling correlation," in *Proc. IEEE/IAS 59th Ind. Commercial Power Syst. Tech. Conf.*, Las Vegas, NV, USA, May 2023, pp. 1–8, doi: 10.1109/ICPS57144.2023.10142105.

[4] S. Kallio and M. Siroux, "Photovoltaic power prediction for solar microgrid optimal control," *Energy Rep.*, vol. 9, pp. 594–601, Mar. 2023, doi: 10.1016/j.egyr.2022.11.081.

[5] H. Wang, S. Yan, D. Ju, N. Ma, J. Fang, S. Wang, H. Li, T. Zhang, Y. Xie, and J. Wang, "Short-term photovoltaic power forecasting based on a feature rise-dimensional two-layer ensemble learning model," *Sustainability*, vol. 15, no. 21, p. 15594, Nov. 2023, doi: 10.3390/su152115594.

[6] M. J. Mayer and G. Gróf, "Extensive comparison of physical models for photovoltaic power forecasting," *Appl. Energy*, vol. 283, Feb. 2021, Art. no. 116239, doi: 10.1016/j.apenergy.2020.116239.

[7] L. Liu, K. Guo, J. Chen, L. Guo, C. Ke, J. Liang, and D. He, "A photovoltaic power prediction approach based on data decomposition and stacked deep learning model," *Electronics*, vol. 12, no. 13, p. 2764, Jun. 2023, doi: 10.3390/electronics12132764.

[8] M. Massaoudi, I. Chihi, L. Sidhom, M. Trabelsi, S. S. Refaat, and F. S. Oueslati, "Enhanced random forest model for robust short-term photovoltaic power forecasting using weather measurements," *Energies*, vol. 14, no. 13, p. 3992, Jul. 2021, doi: 10.3390/en14133992.

[9] A. I. Khalyasmaa, S. A. Eroshenko, V. A. Tashchilin, H. Ramachandran, T. P. Chakravarthi, and D. N. Butusov, "Industry experience of developing day-ahead photovoltaic plant forecasting system based on machine learning," *Remote Sens.*, vol. 12, no. 20, p. 3420, Oct. 2020, doi: 10.3390/rs12203420.

[10] C. Keerthisinghe, E. Mickelson, D. S. Kirschen, N. Shih, and S. Gibson, "Improved PV forecasts for capacity firming," *IEEE Access*, vol. 8, pp. 152173–152182, 2020, doi: 10.1109/ACCESS.2020.3016956.

[11] Q. Liu, Y. Li, H. Jiang, Y. Chen, and J. Zhang, "Short-term photovoltaic power forecasting based on multiple mode decomposition and parallel bidirectional long short term combined with convolutional neural networks," *Energy*, vol. 286, Jan. 2024, Art. no. 129580, doi: 10.1016/j.energy.2023.129580.

[12] E. Sauter, M. Mughal, and Z. Zhang, "Evaluation of machine learning methods on large-scale spatiotemporal data for photovoltaic power prediction," *Energies*, vol. 16, no. 13, p. 4908, Jun. 2023, doi: 10.3390/en16134908.

[13] M. S. Hossain and H. Mahmood, "Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast," *IEEE Access*, vol. 8, pp. 172524–172533, 2020, doi: 10.1109/ACCESS.2020.3024901.

[14] F. Mei, J. Gu, J. Lu, J. Lu, J. Zhang, Y. Jiang, T. Shi, and J. Zheng, "Day-ahead nonparametric probabilistic forecasting of photovoltaic power generation based on the LSTM-QRA ensemble model," *IEEE Access*, vol. 8, pp. 166138–166149, 2020, doi: 10.1109/ACCESS.2020.3021581.

[15] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1724–1734.

[16] R. Wang, C. Li, W. Fu, and G. Tang, "Deep learning method based on gated recurrent unit and variational mode decomposition for short-term wind power interval prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 3814–3827, Oct. 2020, doi: 10.1109/TNNLS.2019.2946414.

[17] B. Liu, C. Fu, A. Bielefield, and Y. Liu, "Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network," *Energies*, vol. 10, no. 10, p. 1453, Sep. 2017, doi: 10.3390/en10101453.

[18] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (GRU) neural networks," in *Proc. IEEE 60th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2017, pp. 1597–1600, doi: 10.1109/MWSCAS.2017.8053243.

[19] A. Mellit, A. M. Pavan, and V. Lughi, "Deep learning neural networks for short-term photovoltaic power forecasting," *Renew. Energy*, vol. 172, pp. 276–288, Jul. 2021, doi: 10.1016/j.renene.2021.02.166.

[20] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and meteorological information," *Renew. Energy*, vol. 118, pp. 357–367, Apr. 2018, doi: 10.1016/j.renene.2017.11.011.

[21] N. Nayak and A. K. Pani, "Short term PV power forecasting using empirical mode decomposition based orthogonal extreme learning machine technique," *Indian J. Public Health Res. Develop.*, vol. 9, no. 11, p. 2170, 2018, doi: 10.5958/0976-5506.2018.01771.0.

[22] H. Feng and C. Yu, "A novel hybrid model for short-term prediction of PV power based on KS-CEEMDAN-SE-LSTM," *Renew. Energy Focus*, vol. 47, Dec. 2023, Art. no. 100497, doi: 10.1016/j.ref.2023.100497.

[23] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014, doi: 10.1109/TSP.2013.2288675.

[24] H. Wang, J. Sun, and W. Wang, "Photovoltaic power forecasting based on EEMD and a variable-weight combination forecasting model," *Sustainability*, vol. 10, no. 8, p. 2627, Jul. 2018, doi: 10.3390/su10082627.

[25] L. Ilias, E. Sarmas, V. Marinakis, D. Askounis, and H. Doukas, "Unsupervised domain adaptation methods for photovoltaic power forecasting," *Appl. Soft Comput.*, vol. 149, Dec. 2023, Art. no. 110979, doi: 10.1016/j.asoc.2023.110979.

[26] V. H. Wentz, J. N. Maciel, J. J. Gimenez Ledesma, and O. H. Ando Junior, "Solar irradiance forecasting to short-term PV power: Accuracy comparison of ANN and LSTM models," *Energies*, vol. 15, no. 7, p. 2457, Mar. 2022, doi: 10.3390/en15072457.

[27] Y. Zhang, Z. Pan, H. Wang, J. Wang, Z. Zhao, and F. Wang, "Achieving wind power and photovoltaic power prediction: An intelligent prediction system based on a deep learning approach," *Energy*, vol. 283, Nov. 2023, Art. no. 129005, doi: 10.1016/j.energy.2023.129005.

[28] K. Wu, X. Peng, Z. Li, W. Cui, H. Yuan, C. S. Lai, and L. L. Lai, "A short-term photovoltaic power forecasting method combining a deep learning model with trend feature extraction and feature selection," *Energies*, vol. 15, no. 15, p. 5410, Jul. 2022, doi: 10.3390/en15155410.

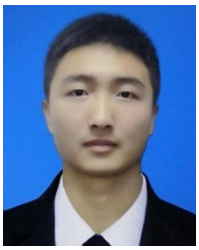
[29] T. Nguyen Trong, H. Vu Xuan Son, H. Do Dinh, H. Takano, and T. Nguyen Duc, "Short-term PV power forecast using hybrid deep learning model and variational mode decomposition," *Energy Rep.*, vol. 9, pp. 712–717, Oct. 2023, doi: 10.1016/j.egyr.2023.05.154.

[30] L. Liu and Y. Li, "Research on a photovoltaic power prediction model based on an IAO-LSTM optimization algorithm," *Processes*, vol. 11, no. 7, p. 1957, Jun. 2023, doi: 10.3390/pr11071957.

[31] X. Chen, K. Ding, J. Zhang, W. Han, Y. Liu, Z. Yang, and S. Weng, "Online prediction of ultra-short-term photovoltaic power using chaotic characteristic analysis, improved PSO and KELM," *Energy*, vol. 248, Jun. 2022, Art. no. 123574, doi: 10.1016/j.energy.2022.123574.

[32] R. Liu, J. Wei, G. Sun, S. M. Muyeen, S. Lin, and F. Li, "A short-term probabilistic photovoltaic power prediction method based on feature selection and improved LSTM neural network," *Electric Power Syst. Res.*, vol. 210, Sep. 2022, Art. no. 108069, doi: 10.1016/j.epsr.2022.108069.
[33] K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, "Machine learning based PV power generation forecasting in Alice springs," *IEEE Access*, vol. 9, pp. 46117–46128, 2021, doi: 10.1109/ACCESS.2021.3066494.
[34] Y. Peng, S. Wang, W. Chen, J. Ma, C. Wang, and J. Chen, "LightGBM-integrated PV power prediction based on multi-resolution similarity," *Processes*, vol. 11, no. 4, p. 1141, Apr. 2023, doi: 10.3390/pr11041141.
[35] X. Huang, J. Liu, S. Xu, C. Li, Q. Li, and Y. Tai, "A 3D ConvLSTM-CNN network based on multi-channel color extraction for ultra-short-term solar irradiance forecasting," *Energy*, vol. 272, Jun. 2023, Art. no. 127140, doi: 10.1016/j.energy.2023.127140.
[36] C. Zhang, T. Peng, and M. S. Nazir, "A novel integrated photovoltaic power forecasting model based on variational mode decomposition and CNN-BiGRU considering meteorological variables," *Electric Power Syst. Res.*, vol. 213, Dec. 2022, Art. no. 108796, doi: 10.1016/j.epsr.2022.108796.
[37] X. Wang, Q. Zhao, K. Li, T. Yang, and C. Fan, "Short term photovoltaic power prediction based on FWA-CNN-GRU neural network," *J. Phys., Conf.*, vol. 2303, no. 1, Jul. 2022, Art. no. 012024, doi: 10.1088/1742-6596/2303/1/012024.
[38] X. Xu, D. Nie, W. Xu, K. Wang, S. Chen, Y. Nie, X. Fu, and W. Xu, "Design of photovoltaic power generation servo system based on discrete adaptive network dynamic surface control technology," *Processes*, vol. 11, no. 6, p. 1667, May 2023, doi: 10.3390/pr11061667.

GUOMIN XIE received the B.Sc., M.Sc., and Ph.D. degrees from Liaoning Technical University, in 1991, 2003, and 2012, respectively. He is currently a Professor with Liaoning Technical University. His research interests include power system information detection and fault diagnosis.



ZHONGBAO LIN received the B.Sc. degree from Liaoning Technical University, in 2022, where he is currently pursuing the M.Sc. degree. His research interests include power system information detection and fault diagnosis.