



Short-term photovoltaic power forecasting based on multiple mode decomposition and parallel bidirectional long short term combined with convolutional neural networks

Qian liu^a, Yulin li^{b,*}, Hang jiang^b, Yilin chen^b, Jiang zhang^b

^a School of Mechanical and Electrical Engineering, Chengdu University of Technology, Chengdu, 610000, China

^b College of Electrical Engineering and New Energy, China Three Gorges University, Hubei, 443002, China

ARTICLE INFO

Keywords:

Multiple mode decomposition
Photovoltaic power forecast
Parallel BiLSTM-CNN

ABSTRACT

Photovoltaic (PV) power generation exhibits significant variability due to the unpredictable nature of solar energy and volatile weather conditions. This paper introduces a novel integrated model that combines parallel Bi-directional Long Short-Term Memory (BiLSTM) and Convolutional Neural Network (CNN), utilizing multi-modal decomposition. The proposed model provides precise photovoltaic (PV) forecasts, essential for optimizing short-term dispatches and scheduling in PV power stations. Firstly, Pearson correlation coefficient is employed to assess the correlation between meteorological data and PV power. Variational mode decomposition (VMD), complementary ensemble empirical mode decomposition (CEEMD), and singular spectrum analysis (SSA) are utilized to decompose the highly correlated features including global radiation and radiation global title. Secondly, employing PV power as output, this study introduces sequences from decomposition methods, temperature, humidity, diffuse radiation, wind direction, and tilted diffuse radiation into the training of the Parallel BiLSTM-CNN (PBiLSTM-CNN) network. Finally, the feasibility of the proposed method is demonstrated by example verification and comparative analysis with alternative methodologies. By employing multiple decomposition methods to extract features, the PBiLSTM-CNN model achieves an average accuracy improvement of approximately 19 % and 37 % in different weather conditions and seasons. Moreover, the implementation of PBiLSTM-CNN results in an enhanced forecasting accuracy of about 48 % and 23 %.

1. Introduction

The overreliance on fossil fuels has contributed to a global energy crisis and environmental deterioration, such as the greenhouse effect. The United Nations has proposed the goal of zero emissions in this context [1]. Renewable energy sources such as solar, wind, hydroelectric, and tidal have gained prominence as prospective candidates to gradually replace conventional fossil fuels. Solar energy, a notable form of renewable energy, possesses various advantages such as easy accessibility, abundant resources, cost-free utilization, and extensive distribution compared to conventional power sources. Therefore, PV is considered to be one of the most promising generation technologies. Solar energy systems have emerged as a primary focus of development in various countries, leading to a global installed photovoltaic (PV) capacity increase to 220 GW since 2022 [2].

However, PV power exhibits a strong correlation with weather

conditions, which possess characteristics such as instability and unpredictability. Therefore, the integration of large-scale photovoltaic systems into the power grid is recognized as a significant source of disruption, bringing considerable challenges to power system management and operation [3]. The integration may lead to increased curtailment of solar energy, which could impede the development of solar energy. Accurate PV power forecasting (PPF) is widely proposed as the solution to the problem, which provides references for operation planning and short-term scheduling, reduces potential operational risks, before improving the penetration level of PV in the power grid and promote the development of the solar energy [4].

According to the forecast timescales, the PPF can be categorized into long-term, medium-term, short-term, and ultra-short-term forecasts [5]. Due to the indeterminacy of environmental factors, medium and long-term forecasts tend to result in significant errors. For the assurance of power system safety, high forecasting accuracy is indispensable [6]. Relatively, short-term forecasting can offer valuable references for

* Corresponding author.

E-mail address: liyulin74@gmail.com (Y. li).

Nomenclature	
BiLSTM	Bi-direction long short-term memory
CNN	Convolutional neural network
VMD	Variational mode decomposition
RNN	Recurrent neural network
GRU	Gate recurrent unit
SVM	Support vector machine
ANN	Artificial neural network
CEEMD	Complementary ensemble empirical mode decomposition
SVD	Singular value decomposition
PV	Photovoltaic power
PPF	PV power forecasting
LSSVM	Least squares support vector machine
NPKDE	Nonparametric and density estimation
GA	Genetic algorithm
RCC	Radiometric classification coordinate
DKASC	Desert knowledge australia solar centre
NWP	Numerical weather forecast
FCM	Fuzzy c-means clustering
WOA	Whale optimization algorithm
WT	Wavelet transform
PSO	Particle swarm optimization
RF	Random forests
DT	Decision tree
SSA	Singular spectral analysis
RMSE	Root means squared error
MAE	Means absolute error
MAPE	Mean absolute percentage error
SDA	Similar day analysis
PBiLSTM-CNN	Parallel BiLSTM-CNN
CLSTM	Hybrid CNN-LSTM
ELM	Extreme learning machine

market players. Generally, the current forecast can be categorized into physical and data-driven methods. The physical method mainly calculates PV power directly based on Numerical Weather Forecast (NWP) and equipment installation characteristics [7]. Yet the modeling process is complex and susceptible to environmental influences, making the physical method unsuitable for short-term PPF [8]. Data-driven method encompasses statistical methods and meta-heuristic learning methods. The statistical method requires the establishment of a stable correlation mapping between input and output, which can result in inaccuracies in forecasting rainy days or extreme weather [9]. The processing of substantial volumes of historical data poses significant complexity and challenges. However, the deep characteristics of feature variables cannot be extracted by the statistical method. Alternatively, meta-heuristic method can overcome this weakness. Meta-heuristic method can extract the high-dimensional features of the PV system and map them to the output [10]. The results of the meta-heuristic method typically exhibit high accuracy. Because of these advantages, the meta-heuristic method has become the predominant approach for short-term PPF.

Many meta-heuristic learning methods have been proven to demonstrate outstanding forecasting performance. GU et al. proposed a day-ahead PPF and uncertainty analysis method based on fuzzy c-means clustering (FCM), whale optimization algorithm (WOA), least squares support vector machine (LSSVM), and nonparametric and density estimation (NPKDE) [11]. FCM is adopted to cluster the sample, which improves the calculation speed and forecasting accuracy of LSSVM. Meanwhile, the penalty factor and kernel function width of LSSVM was optimized by WOA to obtain accurate PPF results. ABinet Tesfaye Eseye et al. proposed a hybrid forecast model combining wavelet transform (WT), particle swarm optimization (PSO), and SVM for short-term forecasts of actual microgrid PV systems [12]. The SVM parameters are optimized by PSO. Finally, the effectiveness of the method is demonstrated through examples. The key concept of the two mentioned research is using SVM to train feature sample mapping output. The models employed in the mentioned literature can be categorized as machine learning, where optimization algorithms are utilized to adjust the model parameters, enabling optimal model training. However, this comes at a high cost in terms of time. Other machine learning models, such as random forests (RF) [13] and decision tree (DT) [14] can also be utilized for forecasting, but they require high data quality and exhibit poor performance when dealing with big data [15]. As a result, an increasing number of scholars have directed their research efforts toward deep learning models. Deep learning has a greater learning capacity than conventional machine learning [16]. It can extract features from complex data and map them to results. At present, the neural

network is the most widely studied and used method in short-term PPF, such as RNN, LSTM, GRU, and CNN [17–20]. The most commonly utilized models are the LSTM and CNN models. Based on LSTM, Gao. M's team established ideal and non-ideal weather forecast models, respectively. For ideal weather conditions, an ideal weather forecast model based on meteorological data of the next day was proposed [21]. Features such as adjacent daytime series are introduced into the non-ideal weather forecast model. This approach improves the forecast accuracy of LSTM in various meteorological conditions. Although deep learning models have a high learning capacity, they extract a limited amount of information from raw historical data. However, richer information can be extracted through data processing. Recent research has put forward the decomposition of historical data using signal processing methods to decrease the complexity of the original series and mitigate model learning difficulties. For example, Wang et al. proposed a CNN-LSTM model based on wavelet decomposition for the day-ahead forecast of PV output and compared the results with those of ANN and CNN [22]. Although the team achieved favorable prediction results, the limited enhancement in model performance arose from the restricted amount of feature information provided by using a single decomposition method, without considering the sequence of feature extraction for CNN and LSTM. Lin et al. proposed a multi-step forecast method base on ensemble empirical mode decomposition (EEMD)-VMD and BiLSTM [23]. The original power series are decomposed by EEMD, then the secondary decomposition of IMF1 component was taken using VMD with meteorological features as input. The decomposed series were predicted using BiLSTM for the final modal reconstruction. The authors explored a two-step decomposition to reduce the complexity of the original PV sequences and improve the forecasting accuracy of BiLSTM. However, the decomposition of the original PV sequences reduces the complexity while decreasing the correlation between the sequences and the features, potentially leading to difficulties in network learning. Deniz Korkmaz proposed a hybrid reliable model based on SolarNet and VMD [24]. The authors decomposed the original energy sequence using VMD and transformed the decomposed subseries with the weather features into RGB images trained in SolarNet. The method attains superior results, but the use of image features increases the complexity of feature processing while compromising the temporal characteristics of the PV sequence. This may not be an ideal choice for practical application and feature engineering. The essence of most predictions involves regression analysis using models, with the two crucial components in regression analysis being feature engineering and the models utilized for training the features. Diverse features play an important role in the learning process of model training which could help model capture potential correlations between data to achieve improved outcomes. While dealing

with rich features, the significance lies not only in selecting an appropriate model but also in making prudent use of the model for feature learning. Most of the aforementioned research employs a feature extraction method that faces challenges in ensuring feature richness, while the trained model remains simplistic to guarantee capacity. Meanwhile, building a model entails considering the types of data and using appropriate methods to extract the features of the data, which can greatly benefit the model training process. Kejun W. et al. suggested the use of LSTM as a pre-model to extract features from raw data. The results demonstrated superior performance compared to CNN in terms of pure data features [25]. However, the authors did not carry out feature engineering which could not guarantee the richness of the features, and a simple model was used for training without generalizability test. These cited literatures are of well inspiration for our work.

In this work, a PBILSTM-CNN network based on multiple modal decompositions is proposed. Firstly, the correlation between meteorological data and PV power is analyzed, and the two characteristic variables with the greatest correlation are decomposed using three different decomposition methods (VMD, CEEMD, SSA) to obtain more diverse and informative features. Secondly, a three-way parallel BILSTM-CNN network is established, incorporating the decomposed sub-feature variables and additional weather features for training. Generalizability tests were eventually carried out with full assurance of feature richness and model capacity. Finally, the validity of the method is tested by example verification, and the advantages of the model are compared with other methods.

2. Background theories

2.1. Variational mode decomposition

VMD is proposed by Dragomiretskiy and Zosso in 2013 [26] as one of the most recent decomposition techniques. It functions as a fully non-recursive signal processing method with strong anti-noise capability. VMD can effectively smooth the complexity of time series, ensuring greater stability of sub-sequences. In addition, VMD is more suitable for nonlinear and non-stationary signal decomposition, effectively reducing the occurrence of modal aliasing [27]. The VMD function decomposes the signal $f(t)$ into K narrow-band eigenmode functions [26]:

$$f(t) = \sum_{k=1}^K u_k(t) \quad (2-1)$$

The decomposition process is a signal processing method that combines Wiener filtering, Hilbert transform, and the alternating direction multiplier method. The fundamental concept of VMD is to minimize the discrete set of $u_k(t)$ and $w_k(t)$ [26]:

$$\begin{cases} \min_{\{u_k\}, \{w_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j w_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^K u_k(t) = y(t) \end{cases} \quad (2-2)$$

where k is the total number of decomposition modes, $u_k(t)$ represents k IMFs decomposed by VMD, $w(k)$ describes the center frequency corresponding to $u_k(t)$, ∂_t is the time derivative, and δ_t is the impulse function. By introducing the Lagrange multiplier operator $\lambda(t)$ into the above expression, the expression is transformed into an unconstrained variational problem [28]:

$$L(\{\{u_k\}, \{w_k\}, \lambda(t)\}) = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j w_k t} \right\|_2^2 + \left\| y(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \langle \lambda(t), y(t) - \sum_{k=1}^K u_k(t) \rangle \quad (2-3)$$

where $\left\| y(t) - \sum_{k=1}^K u_k(t) \right\|_2^2$ is the second penalty term, $\langle \cdot, \cdot \rangle$ is the inner product operation, and α is the penalty factor.

2.2. Complementary ensemble empirical mode decomposition

CEEMD is a signal decomposition method improved upon empirical and ensemble empirical mode decomposition [29]. By adding independent and identically distributed white noise columns with opposite signs to the original signal, CEEMD can reduce the residual auxiliary noise in the original signal and suppress the modal aliasing phenomenon. The decomposition steps are as follows:

- 1) Firstly, a pair of white noise signals are added to the original signal to form a composite signal. Its expression is as follows [30]:

$$\begin{cases} p_i^+(t) = x(t) + n_i^+(t) \\ p_i^-(t) = x(t) + n_i^-(t) \end{cases} \quad (2-4)$$

where $x(t)$ is the original signal to be decomposed, and $n(t)$ is a pair of white noise sequences.

- 2) Through the empirical mode decomposition of the synthetic signal, the k th intrinsic mode components of $p_i^+(t)$ and $p_i^-(t)$ are F_{ik} and F_{ik} , respectively. The average value of each modal and residual component is taken as the decomposition result. It is shown in Equation (2-5) [30]:

$$\begin{cases} I_k(t) = \frac{1}{2k} \sum_{k=1}^K (F_{ik}(t) + F_{-ik}(t)) \\ R_c(t) = x(t) - \sum_{k=1}^K I_k(t) \end{cases} \quad (2-5)$$

$I_k(t)$ is the k th intrinsic modal component, and $R_c(t)$ is the final residual component.

2.3. Singular spectrum analysis

Singular spectral analysis (SSA) is the construction of a trajectory matrix based on the observed time series [31]. The decomposition and reconstruction of the trajectory matrix are taken to extract signals representing different components of the original time series, such as long-term trend signals, periodic signals, and noise signals. The signals obtained through the decomposition are further analyzed. The specific implementation process of the algorithm is as follows:

- 1) Choosing the appropriate L window to transform the 1D time series into a multidimensional series to obtain the trajectory matrix $X = [X_1, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K}$.
- 2) Performing singular value decomposition (SVD) on the trajectory matrix X , let $S = XX^T$, the eigenvalue of S is $\lambda_1, \dots, \lambda_L$, and U_1, \dots, U_L are the standard orthogonal vector of the S matrix corresponding to these eigenvalues. Let $d = \text{rank}(X) = \max\{i, \lambda_i > 0\}$. $V_i = X^T U_i / \sqrt{\lambda_i} (i = 1, \dots, d)$. The SVD of X can be expressed as $X = X_1 + \dots + X_d$, where $X_i = \sqrt{\lambda_i} U_i V_i^T$.

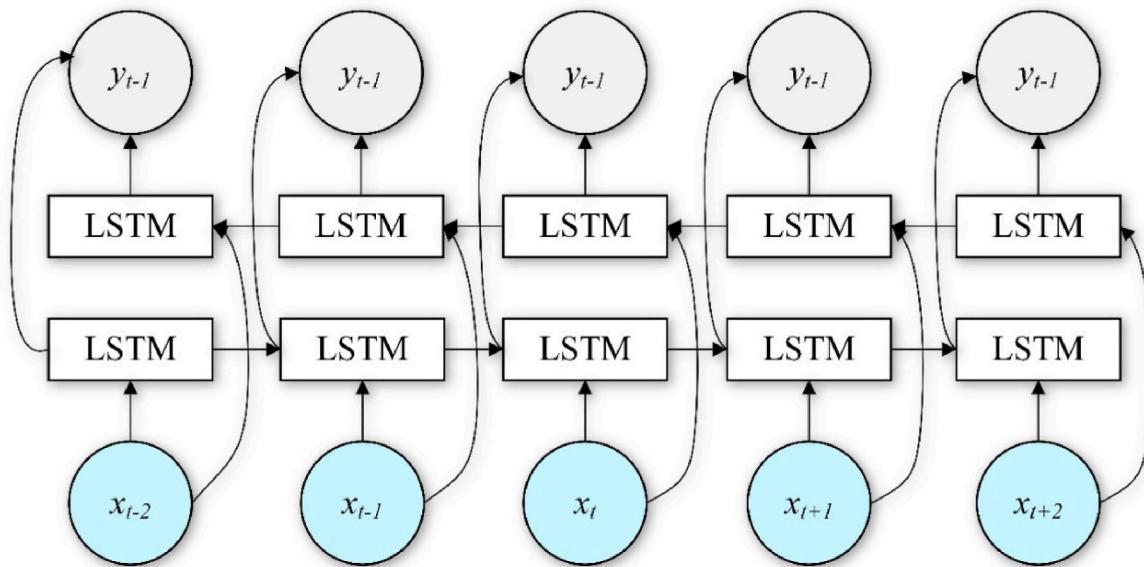


Fig. 1. The typical structure of BiLSTM.

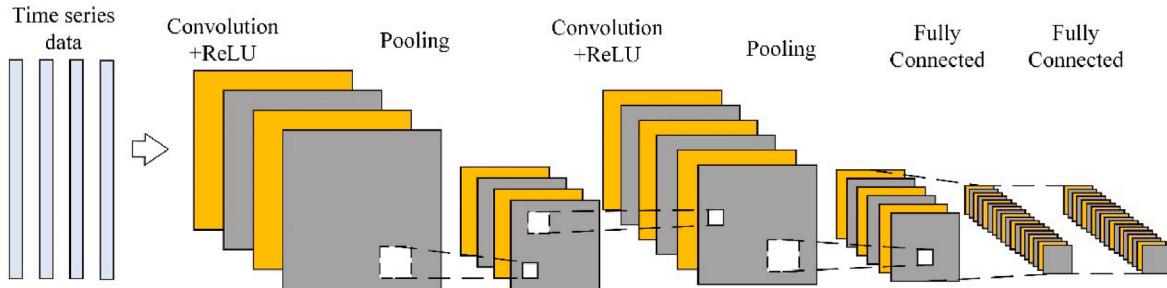


Fig. 2. The typical structure of CNN.

Table 1
Pearson correlation coefficient between photovoltaic output and meteorological factors.

Factor	Pearson	Spearman
Temperature	0.24	0.29
Humidity	-0.43	-0.35
Global Radiation	0.97	0.96
Diffuse Radiation	-0.11	0.37
Wind Direction	-0.12	-0.06
Radiation Global Tilted	0.98	0.99
Radiation Diffuse Tilted	0.15	0.39

- 3) Firstly, divide the set $\{1, \dots, d\}$ into mutually disjoint subsets I_1, \dots, I_m , let $I = \{i_1, \dots, i_p\}$. The synthetic matrix $X_I = X_{i1} + \dots + X_{ip}$ corresponds to I , $X = X_{I_1} + \dots + X_{I_m}$.
- 4) Each matrix X is transformed into a new sequence of length N , and the decomposed sequence will be obtained. Let Y be a matrix of $L^* \times K$, and the elements of the matrix are y_{ij} , $1 \leq i \leq L$, $1 \leq j \leq K$. Let $L^* = \min(L, K)$, $K^* = \max(L, K)$, $N = L + K - 1$. If $L < K$, $y_{ij}^* = y_{ij}$, or $y_{ij}^* = y_{ji}$, calculation of the diagonal mean using the equation, converting the Y matrix into y_1, \dots, y_N . The equations can be expressed as [32]:

$$y_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k y_{m,k-m+1}^*, & 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+1}^*, & L^* \leq k \leq K^* \\ \frac{1}{N-K+1} \sum_{m=k-K+1}^{N-K+1} y_{m,k-m+1}^*, & K^* < k \leq N \end{cases} \quad (2-6)$$

2.4. Bi-direction long short-term memory

The LSTM network is improved by the recurrent neural network (RNN). The LSTM adds three gate logic control units based on the recurrent neural network: the input gate, the forgetting gate, and the output gate [33]. The weight of the self-circulation is changed due to the increase of the input threshold, forgetting threshold, and output threshold. This allows LSTM to avoid the issue of gradient disappearance or gradient expansion, enabling the network to converge more rapidly and significantly enhance forecasting accuracy. The data transfer process can be expressed as follows [34]:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t \odot \tanh(c_t) \end{cases} \quad (2-7)$$

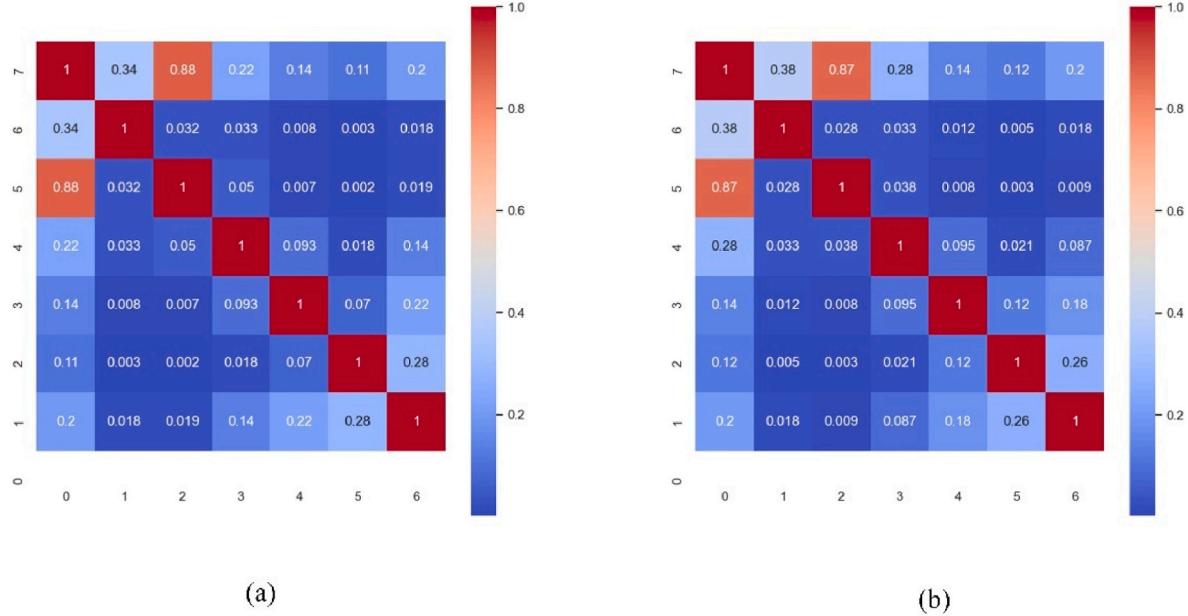


Fig. 3. (a) global radiation, (b) radiation global tilted, Heatmap of pearson correlation coefficients for VMD subsequence with PV power.

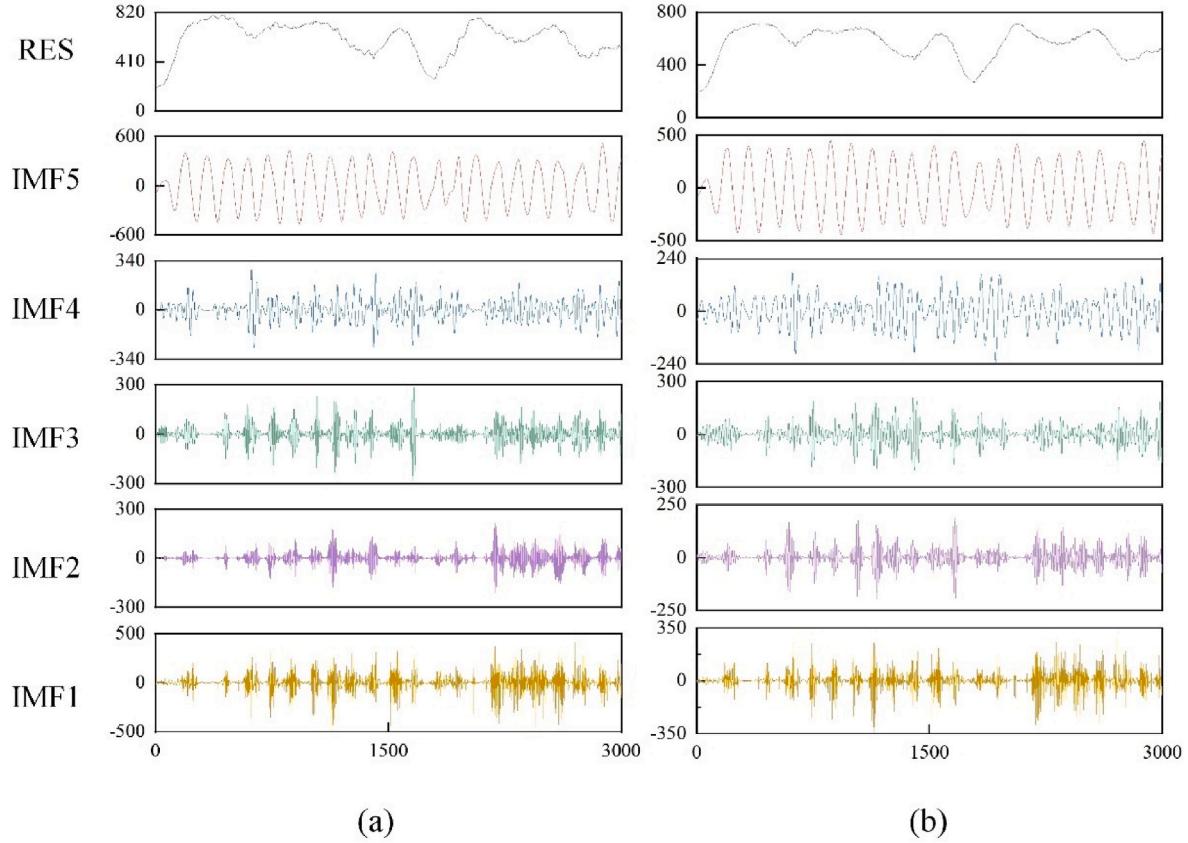


Fig. 4. Decomposition results for A and B using VMD (a) Radiation Global Tilted, (b) Global Radiation.

However, LSTM can only encode the information of time series in a forward direction and cannot encode the information in a backward direction. In this work, BiLSTM is used in the first layer of the network to extract the temporal characteristics of the historical data. The structure is shown in Fig. 1. BiLSTM can simultaneously capture the feature sequence from the future to the current. Meanwhile, the forecast results

are superior to those of LSTM [35,36].

2.5. Convolutional neural network

CNN is a kind of neural network comprising an input layer, convolution layer, pooling layer, fully connected layer, and output layer [37].

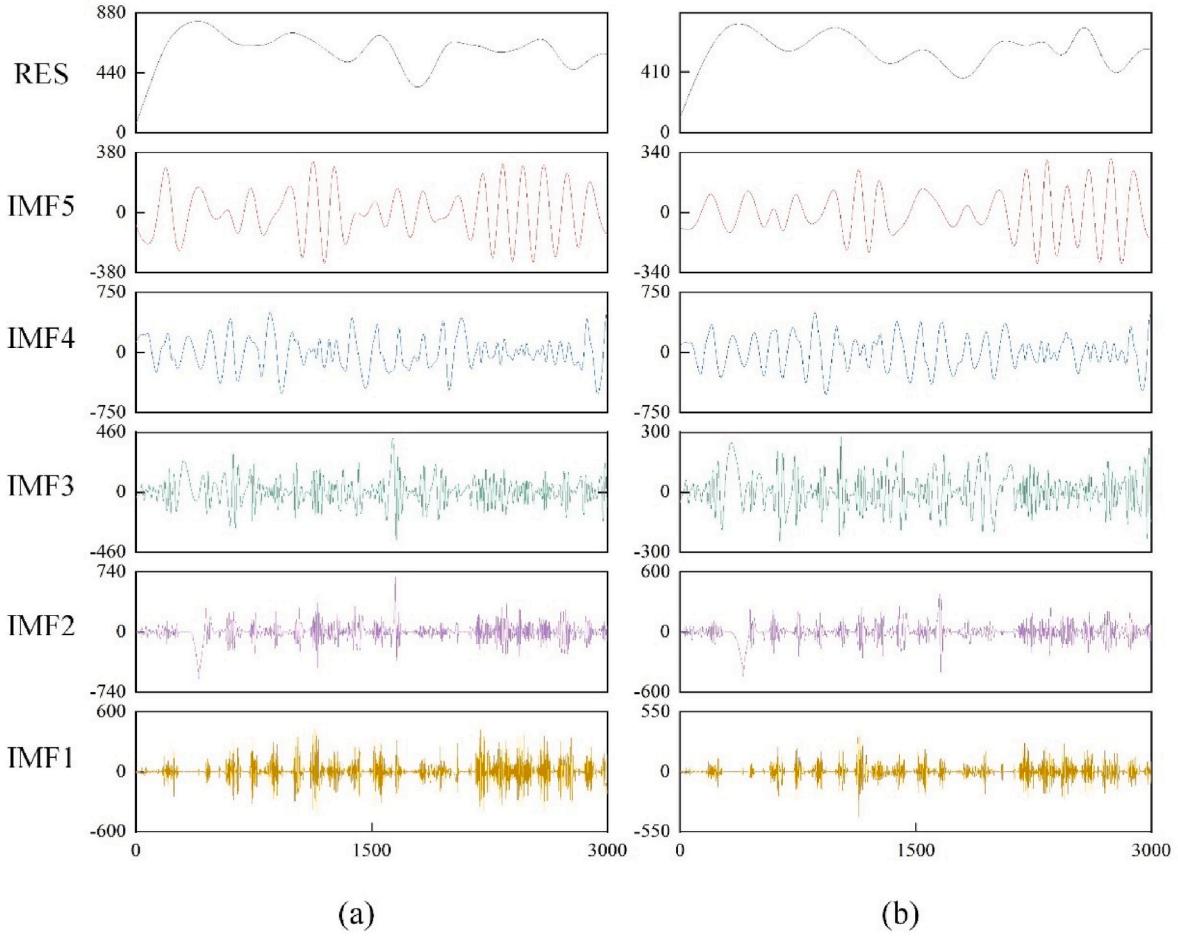


Fig. 5. Decomposition results for A and B using CEEMD (a) Radiation Global Tilted, (b) Global Radiation.

It is specifically designed to process data with a similar grid structure and has found widespread application in various fields [38–40]. The typical CNN structure is shown in Fig. 2. An important characteristic that sets CNN apart from other neural networks is its use of convolution operations in at least one layer, replacing general matrix multiplication. Local connection, weight sharing, and invariance are the three characteristics of CNN. The convolution layer, pooling layer, and output vector can be described as [37]:

$$\begin{cases} x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l \right) \\ x_j^l = f \left(\beta_j^l \text{down} \left(x_j^{l-1} \right) + b_j^l \right) \\ x_j^l = f \left(K^l x_j^{l-1} + b^l \right) \end{cases} \quad (2-8)$$

3. Data analysis and processing

Firstly, the working time of the PV generator is taken in the historical data. Meanwhile, the outliers and missing values will be eliminated and completed. Then the correlation between PV power and meteorological data is analyzed using Pearson and Spearman correlation coefficient. The correlation is shown in the following table. As shown in Table 1, the most correlated variable was Radiation Global Tilted with a correlation of 0.99 and 0.98, followed by Global Radiation with a correlation of 0.97 and 0.96. The sub-feature sequences of other weakly correlated variables are even less correlated for the training network reference. Therefore, the two characteristic variables with the highest correlation are selected for decomposition, and their sub-feature sequences are used

as training features in this work.

PV and meteorological data can be regarded as complex time series signals. Therefore, the concept of signal decomposition can be employed to reduce the complexity of the original signal and increase regularity. Moreover, the subsequence has a specific correlation with the actual power series and can be used as input features of the network as a forecast reference. As shown in Fig. 3, the Pearson correlation coefficient between the VMD-decomposed sub-feature sequence and the original power sequence is 0.34–0.88–0.22–0.14–0.11–0.2. The correlation of the first subsequence is more informative for network learning than features such as Temperature, Diffuse Radiation, and Radiation Diffuse Tilted.

Meanwhile, the two characteristic variables exhibiting the highest correlation are decomposed using VMD, CEEMD, and SSA, which are three signal decomposition methods distinguished by distinct mathematical mechanisms and robust performance. Three distinct modal decomposition algorithms including CEEMD, SSA, and VMD are applied to process the identical original signal, producing three separate modal decomposition outcomes. This approach can avoid the escape of some easily aliasing modal subsequences in a single decomposition method. The modal subsequences of three different mathematical mechanisms are employed to form a more comprehensive description of signal complexity, and a heterogeneous modal subsequence that can better represent the inherent law of complex signals is obtained. The decomposition results are shown in Fig. 4, Fig. 5, and Fig. 6.

The number of subsequences decomposed by VMD is 12. CEEMD and SSA are referenced to VMD. The sub-features of VMD, CEEMD, and SSA are used as features with other weather data and fed into the network in three different channels for training.

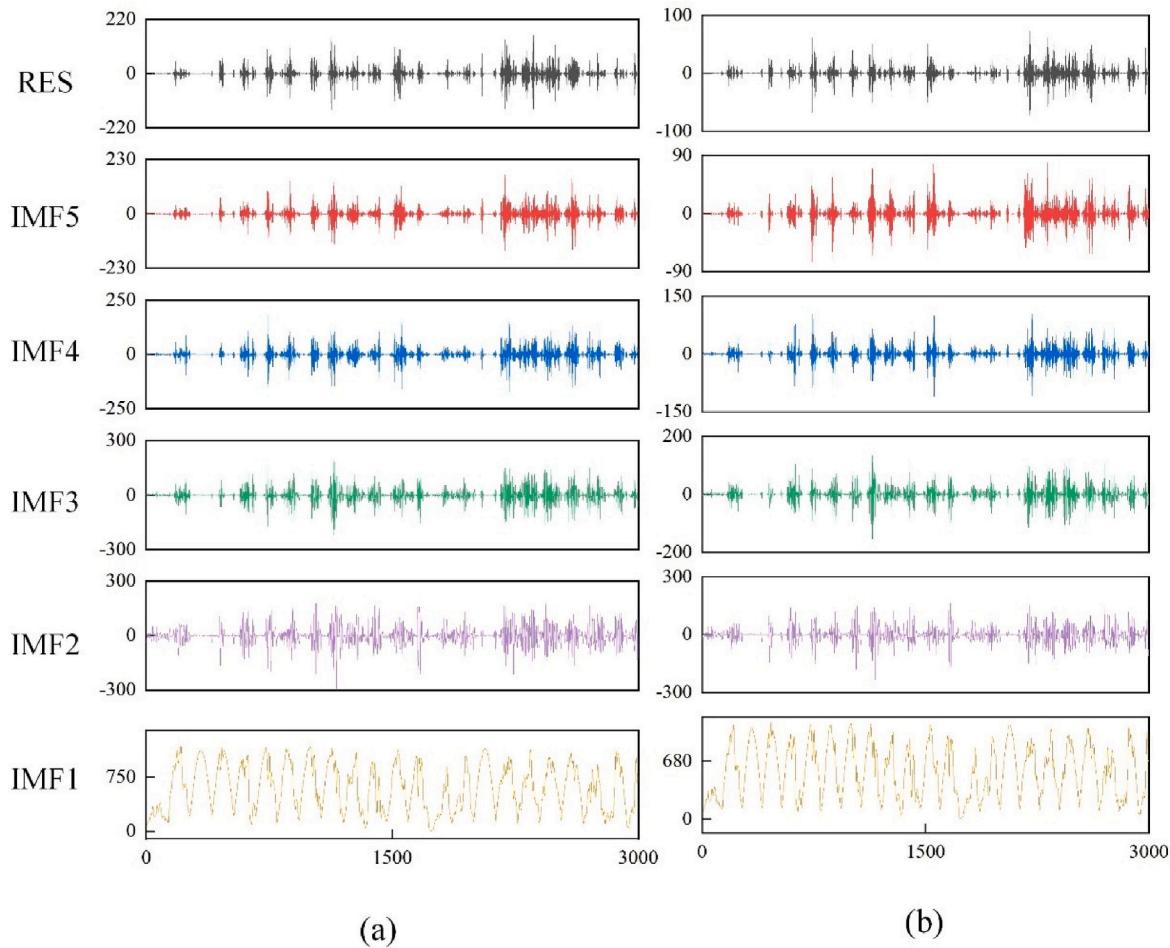


Fig. 6. Decomposition results for A and B using SSA (a) Radiation Global Tilted, (b) Global Radiation.

4. Model construction and parameterization

In this part, a PBiLSTM-CNN network is proposed. The BiLSTM network is employed as a front-end network to extract temporal features from time series. Meanwhile, the CNN network extracts spatial features as a back-end network. Subsequently, the CNN network is employed to extract more profound features, which are then integrated and connected to the fully connected layer. Ultimately, the PV time series prediction is generated from the regression layer utilizing these deep features. The parallel block consists of a BiLSTM layer, a Conv1D layer, and a max pooling layer. The features extracted from the parallel block are concatenated, processed by the Conv1D layer, aggregated, connected to the following Maxpool layer, then linked to the subsequent Con1D-Maxpool layer, and ultimately connected to the fully connected layer. The detail of the PBiLSTM-CNN is shown in Fig. 7. In the parallel block, BN, ReLU, L-ReLU, and fc represent batch normalization, rectified linear unit, leak rectified linear unit, and fully connected layers, respectively. The unit of the BiLSTM layer is 64. The kernel size, stride, and output depth of the Conv1D layer in the parallel block are 3, 1, and 64, respectively. The pool size and stride of the max-pooling are 3 and 2. The constant multiplier of L-ReLU is 0.01. After concatenating the features, the kernel size of the Conv1D layer changes to 128, and after max pooling, the kernel changes to 256. The decomposed time series data is input into the network for training, where the BiLSTM layer extracts time series features initially, followed by the CNN layer extracting spatial features. The three modalities are then combined in the Max-pool layer and integrated with the subsequent CNN layer to derive deeper features. Finally, the predicted PV power is output as a time series. In

this process, L-ReLU is applied to overcome the problem of explosion and disappearance during gradient descent, and BN layers are employed to reduce the noise level during training. The detailed configuration of the parameters is shown in Table 2.

5. Case study

In this section, the experimental studies and evaluation of the proposed forecasting method are presented in details. All the experiments are realized on a workstation with AMD Ryzen7 5800H CPU @3.60 GHz, NVIDIA GTX 3060 GPU, and 16 GB RAM. The performance of the model is evaluated in this section. Meanwhile, the accuracy of the model forecast under various weather conditions is demonstrated, and a longitudinal comparison with other methods is made to verify the superiority of the model.

5.1. Data sets

The DKASC Alice Springs PV system data is selected for the following research in this paper. The data includes weather temperature, humidity, global horizontal radiation, diffuse horizontal radiation, wind direction, radiation horizontally tilted, and radiation diffuse tilted. Historical data comprise the full duration of 2016. The ratio of the training set to the test set is 8:2 during model training. To verify the generalization of PBiLSTM-CNN, the validation set have been set, comprising data that has not been utilized during training.

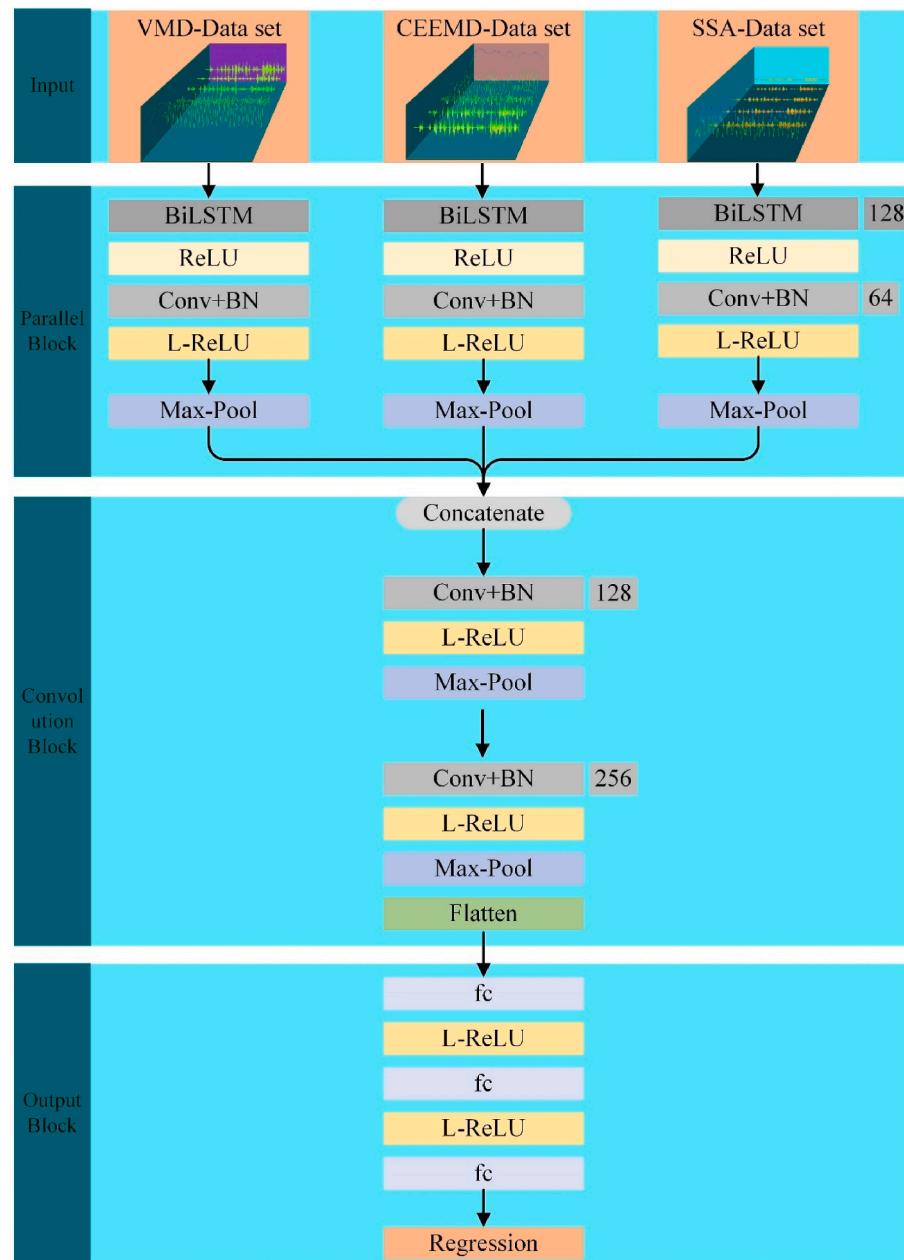


Fig. 7. Details of the network-specific composition of the PBiLSTM-CNN.

5.2. Performance evaluation metrics

In this part, the performance of the proposed method is evaluated. The root means squared error (RMSE) and means absolute error (MAE) are introduced, which are widely used in performance metrics. The RMSE is the arithmetic square root of the mean square error. A lower value of the root mean square error indicates greater model accuracy. The MAE is the average of the absolute error, which reflects the reality of forecast value errors. Hence, the performance of the model can be proved by these two evaluations. These metrics can be given with the following equations [41]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)^2} \quad (5-1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - \hat{P}_i| \quad (5-2)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{P_i - \hat{P}_i}{\bar{P}_i} \right| \quad (5-3)$$

where P_i and \hat{P}_i are the observed and forecasted power values, respectively. \bar{P}_i is the average of the actual power on the measurement day, and n is the total number of samples.

5.3. Experimental

Firstly, time series data of the most highly correlated variable, light intensity, and the second most highly correlated variable, humidity, in the training set are selected and decomposed using VMD, CEEMD, and SSA. There are 12 sets of VMD subsequence components, 12 sets of

Table 2
Detailed layer configuration of the PBiLSTM-CNN architecture.

Models	Configuration		Epoch = 128 Batch size = 128, Optimizer = adam
PBiLSTM-CNN	Parallel Block	Units	Units1 = 128
	BiLSTM		filter = 64; kernel size = 3; stride = 1
	Conv1D + BN + L-ReLU	Convolution	$\alpha = 0.01$
		L-ReLU	kernel size = 3; stride = 2
		Max-Pooling	
	Other Block	Convolution	filter = 128; kernel size = 3; stride = 1
		L-ReLU	$\alpha = 0.01$
		Max-Pooling	kernel size = 3; stride = 2
	Conv1D + BN + L-ReLU	Convolution	filter = 256; kernel size = 3; stride = 1
		L-ReLU	$\alpha = 0.01$
		Max-Pooling	kernel size = 3; stride = 2
	Fully Connected + L-ReLU	neurons = 100	
	L-ReLU	$\alpha = 0.01$	
	Fully Connected + L-ReLU	neurons = 10	
	L-ReLU	$\alpha = 0.01$	
	Fully Connected Output	neurons = 1	
	Regression output		

CEEMD subsequence components, and 12 sets of subsequence components in total. The sub-elements sequences of VMD, CEEMD, SSA, and other features, including temperature, diffuse horizontal radiation, radiation global titled, and radiation diffuse tilted, are introduced into the network model in three different methods. The epochs and batch size of PBiLSTM-CNN are 128 and 128. Adam is selected as the model optimizer. The process of the forecast is shown in Fig. 8.

To test the effectiveness of using multiple modal decompositions and parallel networks, the accuracy of the model trained with features directly input to the network without modal decomposition is compared to that of the model trained with a single modal decomposition feature, as well as to the LSTM-CNN network and the CNN-LSTM network. The structure and hyper-parameter of the LSTM-CNN and CNN-LSTM networks refer to the PBiLSTM-CNN network. Meanwhile, the meteorological data, such as temperature, humidity, global radiation, diffuse radiation, and radiation diffuse tilted are selected as characteristics. The focus was on the accuracy of the model forecasts for the next day. Firstly, the forecasting performance of each trained model is tested under different seasons. Secondly, the forecasting performance of the models is tested under different weather conditions. The results of each model's forecasts for all seasons are shown in the figure. The RMSE and MAE are shown in Table 3 and Table 4.

From the results of the forecasts for the four seasons, the RMSE of the BiLSTM-CNN network and CNN-BiLSTM network are 0.5140-0.6343-0.6249-0.9463 and 1.2393-0.8043-0.9086-0.9780. Hence, the CNN-BiLSTM network is almost impossible to generate realistic forecasts. The forecasts for the BiLSTM-CNN network are more reliable for spring, summer, and autumn. Neither the BiLSTM-CNN nor the CNN-BiLSTM network can be used to forecast the winter accurately. As for the forecasting results in different weather, the RMSE, MAE, and MAPE of the BiLSTM-CNN are 0.3895-0.3121-0.1462 for a sunny day with a smooth output, indicating a more reliable forecast output for sunny days. The forecasts of the RMSE, MAE, and MAPE for cloudy and rainy days are 0.8537-0.5196-0.3320, 3.0369-2.7691-3.3953 for BiLSTM-CNN and

1.1009-0.9334-0.7941, 1.4396-1.2634-1.5973 for CNN-BiLSTM, respectively. As observed in Figs. 9 and 10, the forecast results of these two networks for cloudy and rainy days imply their difficulty in capturing abrupt data fluctuations. In general, it is difficult for both networks to produce accurate and reliable forecasts in various situations. The overall accuracy of BiLSTM-CNN networks is higher than that of CNN-BiLSTM networks, confirming that BiLSTM-CNN networks are better suitable for handling the data type characteristics.

The two characteristic variables with the highest correlation are decomposed using the VMD decomposition. According to the analysis in Section 3, the sub-feature sequence of VMD can provide richer information for the network. The experimental results are shown in Table 3. The RMSE of the VMD-BiLSTM-CNN for the four seasons is 0.2289-0.3179-0.5092-0.6505, which expects 40 % better accuracy than inputting the data directly into the BiLSTM-CNN network.

Meanwhile, the predicting performance of VMD-BiLSTM-CNN for different weather conditions is also studied. The RMSE, MAE, and MAPE of the model for sunny days are 0.1847-0.1377-0.0485, indicating the reliability of the forecasts generated by VMD-BiLSTM-CNN. While the model's performance does not meet the anticipated level for cloudy and rainy days, the forecast accuracy for rainy days exceeds that for cloudy weather. This suggests that the model exhibits greater sensitivity to continuous fluctuations rather than precise responses to abrupt fluctuations. Using VMD on raw features can provide the network with more reference variables and improve the model-learning effect.

Except for VMD, other decomposition methods with different principles, such as CEEMD and SSA, are also applied to global radiation and radiation global title variables to acquire more informative characteristics. Moreover, a parallel multi-channel network structure is adopted to achieve better training results. To verify the effectiveness of the PBiLSTM-CNN, the subseries characteristic variables obtained from the modal decomposition and the rest of the original meteorological features are trained using a single-channel BiLSTM-CNN network. Single-channel BiLSTM-CNN network differ from multi-channel networks in having only one input. The results are compared with that of PBiLSTM-CNN training. Again, the predictive performance of the model is verified in terms of season and weather conditions. The RMSE, MAE, and MAPE of these two models are shown in Table 5, Fig. 11, and Fig. 12.

The single-channel network based on multimodal decomposition exceeds the VMD-BiLSTM-CNN, primarily due to the additional informative features provided to the network by CEEMD and SSA. Compared to the single-channel BiLSTM-CNN network, the PBiLSTM-CNN network exhibits superior forecast performance for all seasons and weather conditions. Regarding seasonal forecast performance, the RMSE of PBiLSTM-CNN and single channel network are 0.1349-0.01593-0.1693-0.2171 and 0.2194-0.2668-0.1755-0.2434. The overall accuracy of PBiLSTM-CNN is higher than that of single-channel networks. For spring and summer forecasts, PBiLSTM-CNN networks are superior. The PBiLSTM-CNN is slightly more accurate for autumn and winter than single-channel networks. Single-channel networks do not perform as well for different weather forecasts as PBiLSTM-CNN, especially for rain and cloudy days. The RMSE of the PBiLSTM-CNN is 0.2727-0.1887, which is 0.4376-0.2442, higher than the single-channel network. This suggests that the PBiLSTM-CNN is sensitive to data changes and exhibits a higher capability for feature extraction. Moreover, the feasibility of using a parallel network architecture is also confirmed. As shown in Fig. 13, PBiLSTM-CNN has shown reliable forecasts for different seasons and weather conditions, demonstrating the good generalization and robustness of the model and the feasibility of using multiple modal decompositions and parallel BiLSTM-CNN.

5.4. Comparison between the proposed method and state-of-the-art methods

In this section, the method proposed in this work is compared with alternative approaches to showcase its superiority. The data used for the

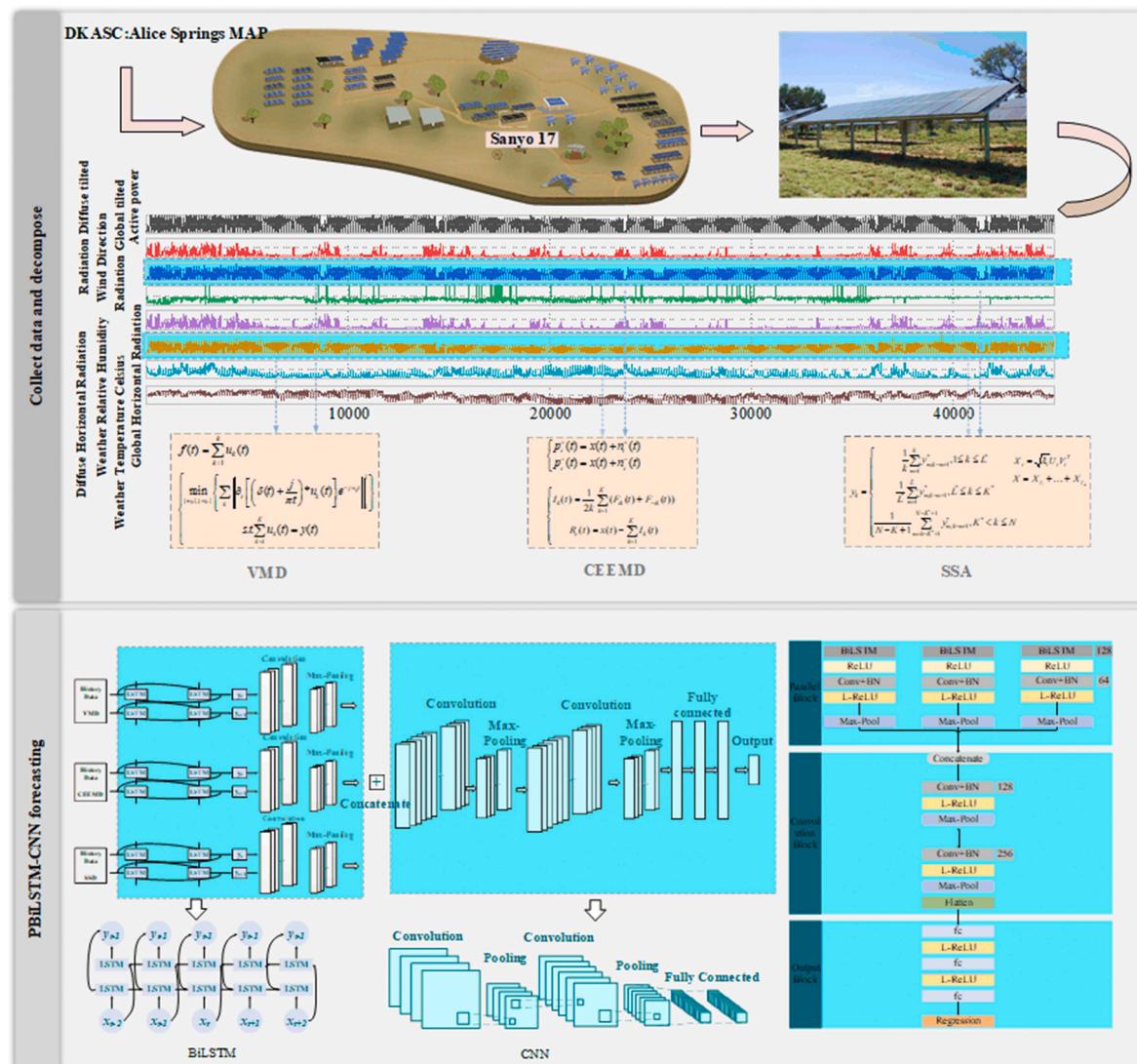


Fig. 8. The complete process and details of forecasting.

Table 3
Forecasting result of models for four seasons.

Season	Model	RMSE	MAE	MAPE
Spring	PBiLSTM-CNN	0.1349	0.1214	0.0540
	VMD-BiLSTM-CNN	0.2289	0.1749	0.0615
	BiLSTM-CNN	0.5140	0.4456	0.1630
	CNN-BiLSTM	1.2393	1.1106	0.5021
Summer	PBiLSTM-CNN	0.1593	0.1100	0.0347
	VMD-BiLSTM-CNN	0.3179	0.2933	0.1000
	BiLSTM-CNN	0.6343	0.4548	0.1721
	CNN-BiLSTM	0.8043	0.6720	0.2487
Autumn	PBiLSTM-CNN	0.1693	0.0956	0.0665
	VMD-BiLSTM-CNN	0.5092	0.4489	0.2094
	BiLSTM-CNN	0.6249	0.4061	0.3184
	CNN-BiLSTM	0.9086	0.7187	1.5535
Winter	PBiLSTM-CNN	0.2171	0.1618	0.1402
	VMD-BiLSTM-CNN	0.6505	0.5777	0.4074
	BiLSTM-CNN	0.9463	0.6949	0.6593
	CNN-BiLSTM	0.9780	0.7988	0.7564

comparative analysis originates from solar-generating plants from the DKASC. Therefore, a general performance comparison with these methods is shown in Table 6.

According to Table 6, Chen et al. proposed a radiometric

Table 4
Forecasting results of models for different weather.

Weather	Model	RMSE	MAE	MAPE
Sunny	PBiLSTM-CNN	0.1199	0.1057	0.0434
	VMD-BiLSTM-CNN	0.1847	0.1377	0.0485
	BiLSTM-CNN	0.3895	0.3121	0.1462
	CNN-BiLSTM	1.0843	0.9435	0.4681
Cloudy	PBiLSTM-CNN	0.2727	0.1915	0.1678
	VMD-BiLSTM-CNN	1.0042	0.6659	0.4242
	BiLSTM-CNN	0.8537	0.5196	0.3320
	CNN-BiLSTM	1.1009	0.9334	0.7941
Rainy	PBiLSTM-CNN	0.1887	0.1438	0.1862
	VMD-BiLSTM-CNN	0.5141	0.4316	0.5367
	BiLSTM-CNN	3.0369	2.7691	3.3953
	CNN-BiLSTM	1.4396	1.2634	1.5973

classification coordinate (RCC) method to categorize the selected similar periods. The selected similar period dataset was reconstructed as a training subset based on PV generation characteristics, and an LSTM model was utilized as the training network. The average RMSE and MAE values for all seasons are 0.9400 and 0.5870, respectively. Wang et al. designed a CNN, LSTM, and hybrid CNN-LSTM (CLSTM). The data were selected with a 5-min resolution and the average RMSE and MAE are

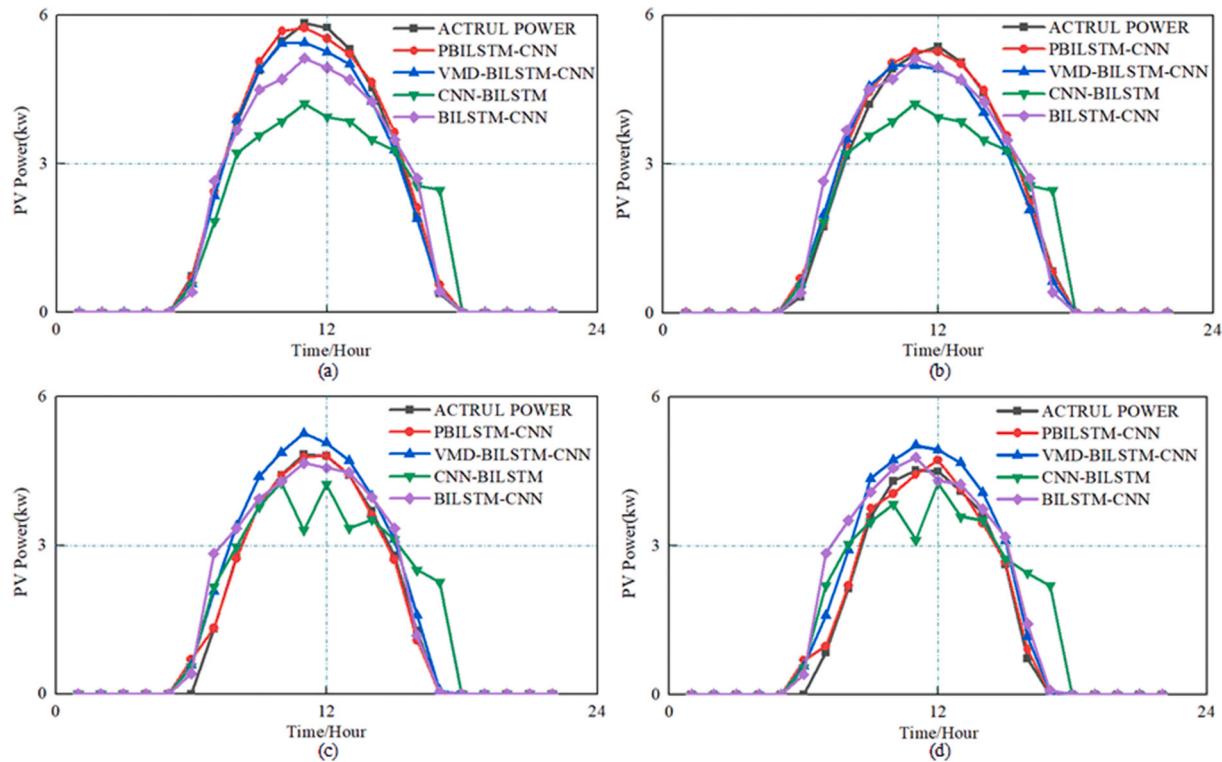


Fig. 9. Forecasting results of different models for four seasons. (a) Spring, (b) Summer, (c) Autumn, (d) Winter.

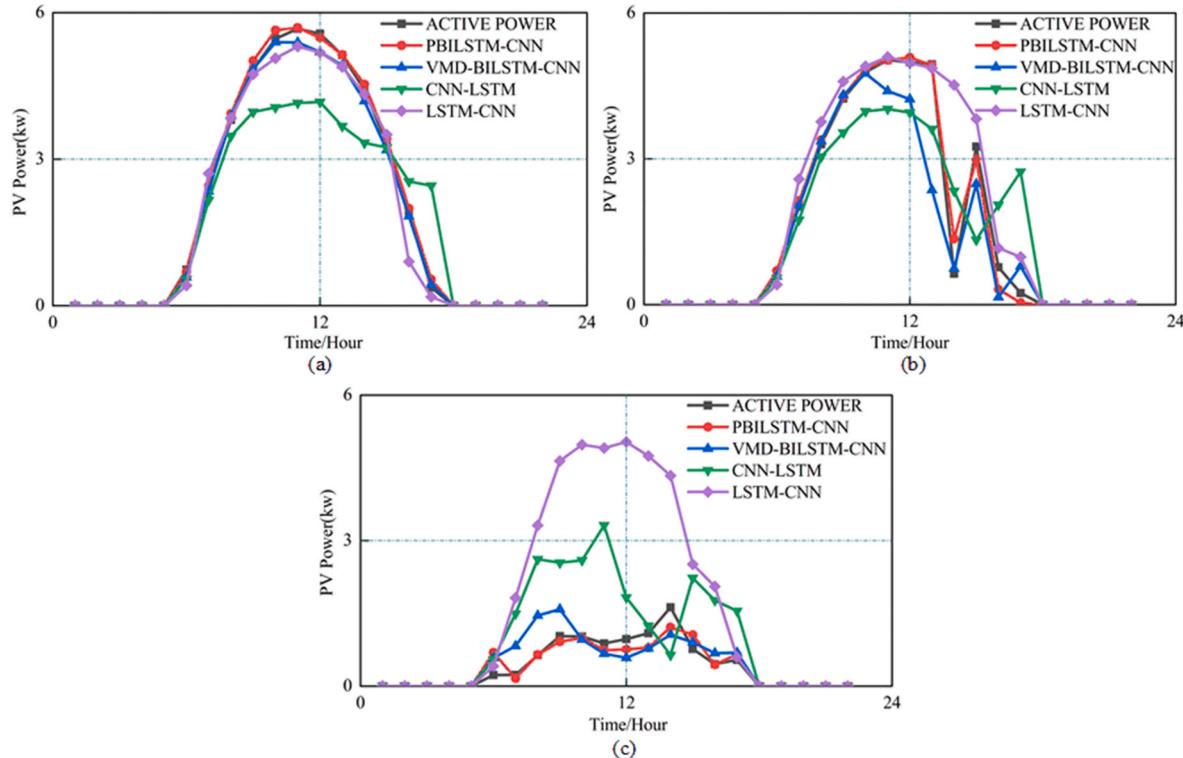


Fig. 10. Forecasting results of models for different weather. (a) Sunny, (b) Cloudy, (c) Rainy.

0.3430 and 0.1260. Zhou et al. proposed a hybrid forecasting model named SDA-GA-ELM based on ELM, genetic algorithm (GA), and customized similar day analysis (SDA). In the SDA, the dataset was employed with a resolution of 1-h. This method reached a higher MAE value of 0.2367. Compared to these studies, the proposed method

provides more accurate PPF and exhibits robust generalization capabilities, enabling precise forecasts across various weather conditions and seasons.

Table 5
Forecasting results of different structure.

Metrics	Single Channel			PBiLSTM-CNN				
	Sunny	Cloudy	Rainy	Sunny	Cloudy	Rainy		
RMSE	0.1624	0.7103	0.4330	0.1199	0.2727	0.1887		
MAE	0.1448	0.4627	0.3462	0.1057	0.1915	0.1438		
MAPE	0.0653	0.2134	0.5240	0.0434	0.1678	0.1862		
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
RMSE	0.2194	0.2668	0.1755	0.2434	0.1349	0.1593	0.1693	0.2171
MAE	0.2000	0.2268	0.1311	0.1827	0.1214	0.1100	0.0956	0.0665
MAPE	0.0828	0.0812	0.1909	0.1748	0.0540	0.0347	0.0665	0.1402

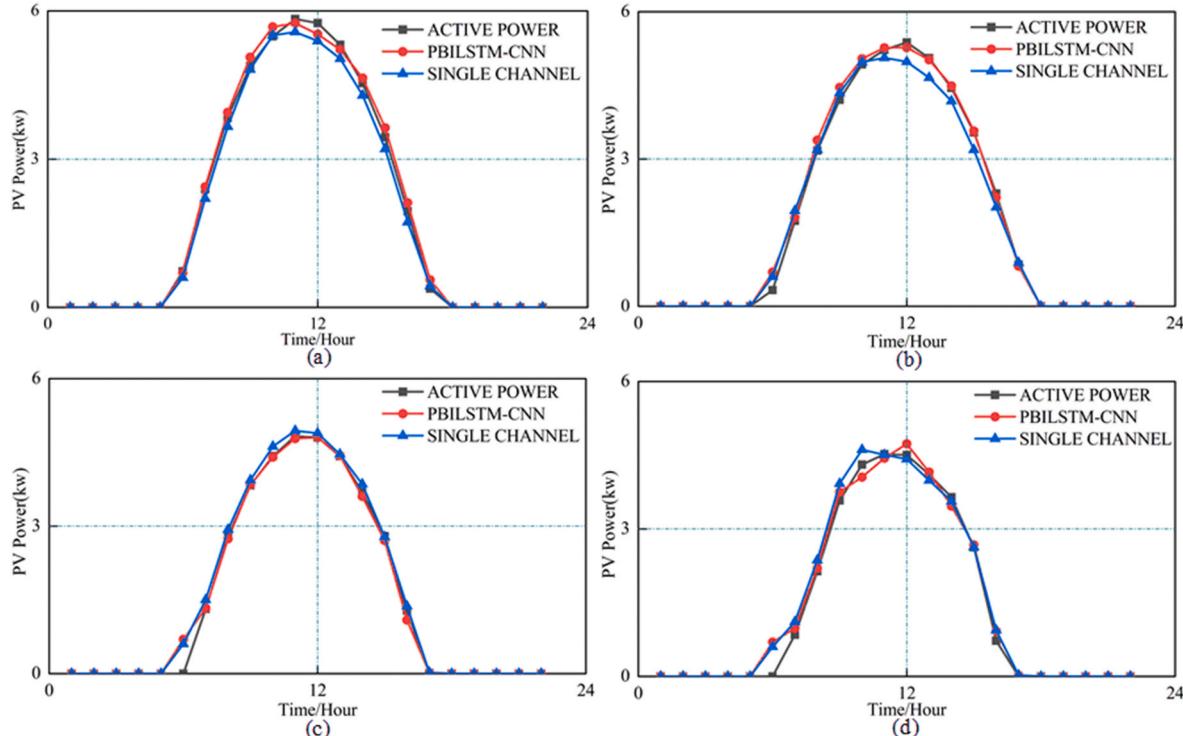


Fig. 11. Single channel network compared to PBiLSTM-CNN in four seasons. (a) Spring, (b) Summer, (c) Autumn, (d) Winter.

6. Conclusion

Solar energy is currently extensively integrated into power grids. Precise photovoltaic data can offer reliable decision support for grid operation and planning. In this study, a method based on multiple modal decompositions with PBiLSTM-CNN is proposed. It is mainly employed to forecast the values of PV power of the next day in advance. The framework of the whole method includes the VMD, CEEMD, and SSA to decompose the two most correlated characteristic variables. Subsequently, the subsequence of each modal decomposition and the remaining relevant variables are allocated to each of the three channels of the PBiLSTM-CNN network for training. The deep spatiotemporal features of each modal decomposition are fully extracted by the network mining and employed to forecast the PV power.

As shown in the test results, the use of multiple modal decompositions brings richer features to the model compared to the model using a single modal decomposition, demonstrating enhanced generalizability and higher predictive accuracy. Meanwhile, the parallel network architecture is employed to construct multiple sub-networks for parallel processing to aid the model in learning more accurate features, which can also improve the expressive ability of model compared to the single-channel network.

All experiments are conducted across two primary scenarios, encompassing sunny, cloudy, and rainy days in all seasons. The feasibility of using multiple modal decompositions is verified by the comparison with BiLSTM-CNN, CNN-BiLSTM, and BiLSTM-CNN networks based on the VMD decomposition, and the comparison with single channel BiLSTM-CNN network proves the effectiveness of using multi-channel networks. The average RMSE, MAE, and MAPE of seasonal and different weather forecasts are 0.1702–0.1938, 0.1222–0.1470, and 0.0985–0.1324 compared to other methods. The accuracy has been considerably improved. The model proposed in this work also has advantages over other models in the literature. It can be reasonably stated that PBiLSTM-CNN can provide reliable references for grid operation and planning.

Credit author statement

Qian liu: Conceptualization, Formal analysis, Writing – review & editing Yulin li: Methodology, Software, Writing – original draft Hang jiang: Writing – review & editing, Data curation Yilin chen Supervision Jiang zhang Supervision.

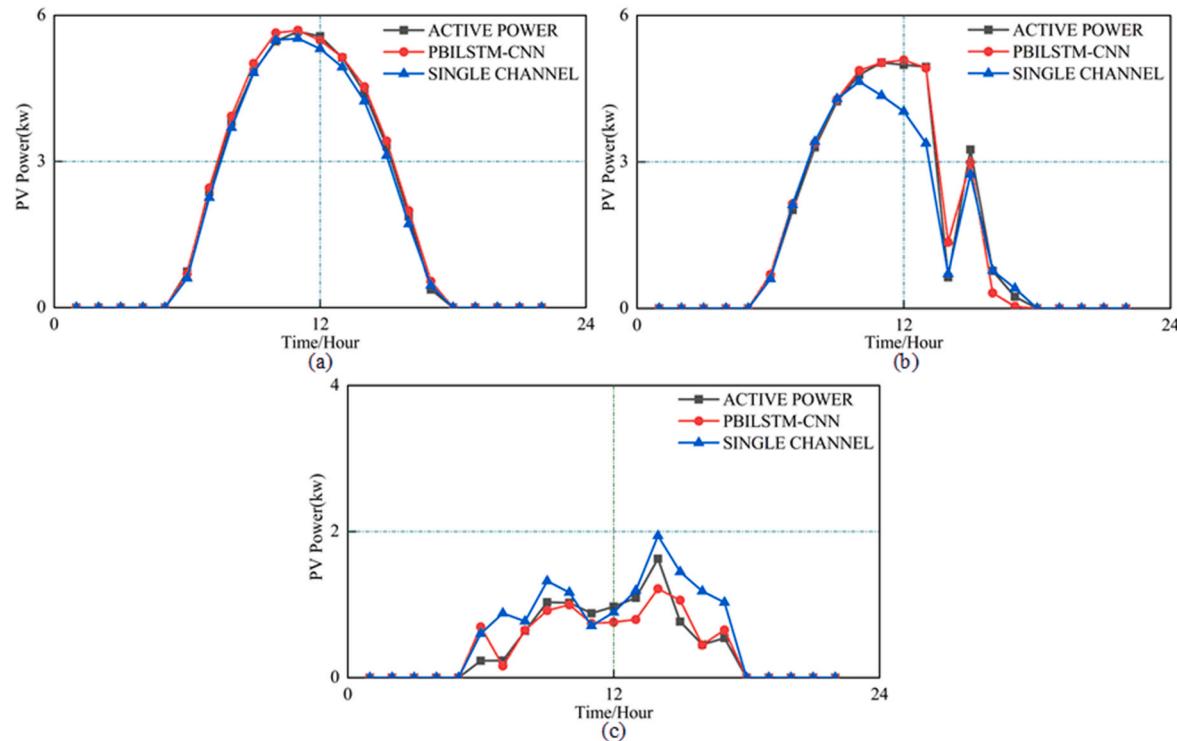


Fig. 12. Single channel network compared to PBiLSTM-CNN in different weather. (a) Sunny, (b) Cloudy, (c) Rainy.

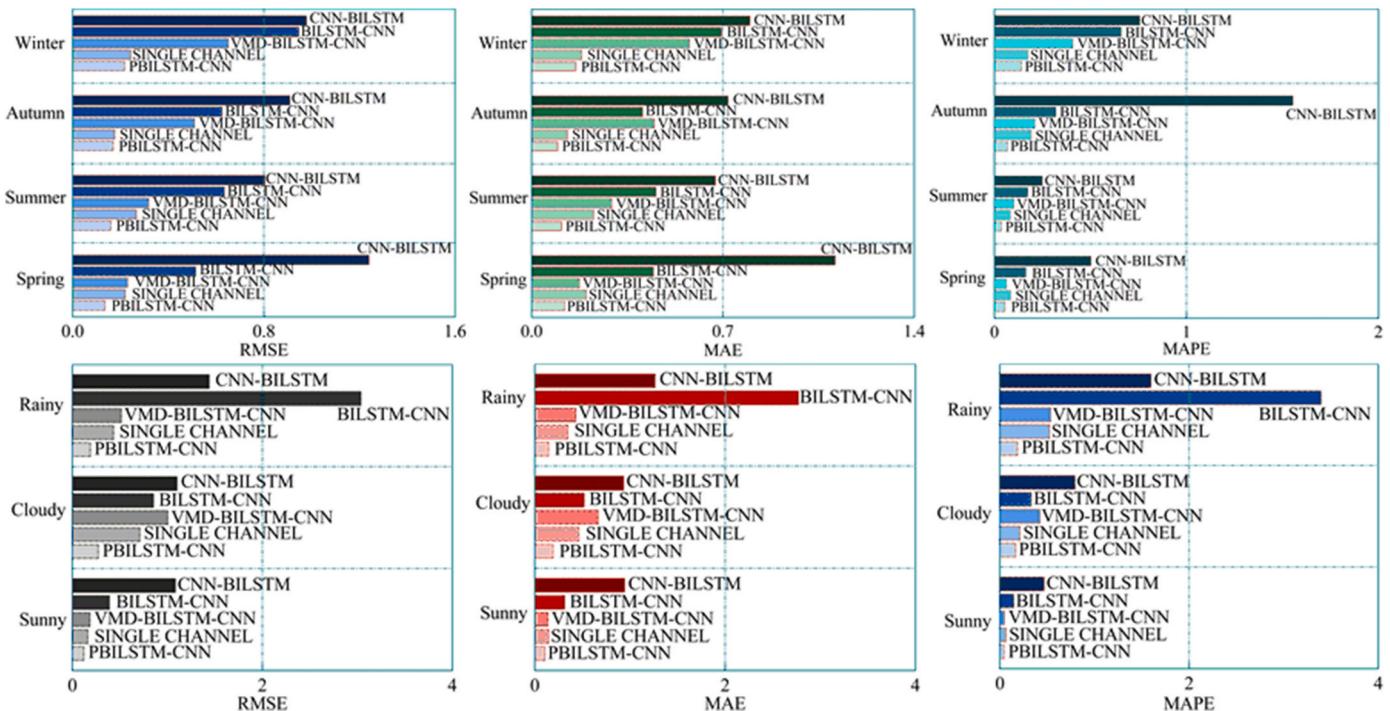


Fig. 13. Overall accuracy comparison.

Declaration of competing interest

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

Data availability

The data that has been used is confidential.

Table 6

General comparison of the proposed method with some state-of-the-art methods.

Study	Method	Dataset	Active Power	Resolution	Forecast Condition	RMSE	MAE
Chen et al.	RCC-LSTM	DKASC, Yulara, 3A	22.56kw	5 min	–	0.9400	0.5870
Wang et al.	CLSTM	DKASC, Alice, Springs, 1B	22.40kw	5 min	–	0.3430	0.1260
Zhou et al.	SDA-GA-ELM	DKASC	4.95kw	1 h	–		0.2367
Present Study	PBiLSTM-CNN	DKASC, Sanyo	6.3kw	5 min	Rainy Cloudy Sunny Average	0.1887 0.2727 0.1199 0.1938	0.1438 0.1915 0.1057 0.1470

References

- [1] Blaga R, Sabadus A, et al. A current perspective on the accuracy of incoming solar energy forecasting. *Prog Energy Combust Sci* 2019;70:119–44. <https://doi.org/10.1016/j.pecs.2018.10.003>.
- [2] Van LP, Hoang LH, et al. A comprehensive review of direct coupled photovoltaic-electrolyser system: sizing techniques, operating strategies, research progress, current challenges, and future recommendations. *Int J Hydrogen Energy* 2023. <https://doi.org/10.1016/j.ijhydene.2023.03.257>.
- [3] Al-Shetwi AQ, Hannan M, et al. Grid-connected renewable energy sources: review of the recent integration requirements and control methods. *J Clean Prod* 2020; 253:119831. <https://doi.org/10.1016/j.jclepro.2019.119831>.
- [4] Van der Meer DW, Widén J, et al. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renew Sustain Energy Rev* 2018;81:1484–512. <https://doi.org/10.1016/j.rser.2017.05.212>.
- [5] de Freitas Visconti G, Alves-Souza SN. A Systematic Literature Review on big data for solar photovoltaic electricity generation forecasting. *Sustain Energy Technol Assessments* 2019;31:54–63. <https://doi.org/10.1016/j.seta.2018.11.008>.
- [6] Ahmed R, Seeram V, et al. A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization. *Renew Sustain Energy Rev* 2020;124:109792. <https://doi.org/10.1016/j.rser.2020.109792>.
- [7] Sobri S, Koohi-Kamali S, et al. Solar photovoltaic generation forecasting methods: a review. *Energy Convers Manag* 2018;156:459–97. <https://doi.org/10.1016/j.enconman.2017.11.019>.
- [8] Mayer MJ, Gróf G. Extensive comparison of physical models for photovoltaic power forecasting. *Appl Energy* 2021;283:116239. <https://doi.org/10.1016/j.apenergy.2020.116239>.
- [9] Ibrahim MS, Dong W, et al. Machine learning driven smart electric power systems: current trends and new perspectives. *Appl Energy* 2020;272:115237. <https://doi.org/10.1016/j.apenergy.2020.115237>.
- [10] Lu P, Ye L, et al. Review of meta-heuristic algorithms for wind power prediction: methodologies, applications and challenges. *Appl Energy* 2021;301:117446. <https://doi.org/10.1016/j.apenergy.2021.117446>.
- [11] Gu B, Shen H, et al. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. *Appl Energy* 2021;299:117291. <https://doi.org/10.1016/j.apenergy.2021.117291>.
- [12] Eseye AT, Zhang J, et al. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. *Renew Energy* 2018;118:357–67. <https://doi.org/10.1016/j.renene.2017.11.011>.
- [13] Benali L, Nottong G, et al. Solar radiation forecasting using artificial neural network and random forest methods: application to normal beam, horizontal diffuse and global components. *Renew Energy* 2019;132:871–84. <https://doi.org/10.1016/j.renene.2018.08.044>.
- [14] Ganti PK, Naik H, et al. Environmental impact analysis and enhancement of factors affecting the photovoltaic (PV) energy utilization in mining industry by sparrow search optimization based gradient boosting decision tree approach. *Energy* 2022; 244:122561. <https://doi.org/10.1016/j.energy.2021.122561>.
- [15] Voyant C, Nottong G, et al. Machine learning methods for solar radiation forecasting: a review. *Renew Energy* 2017;105:569–82. <https://doi.org/10.1016/j.renene.2016.12.095>.
- [16] Schmidhuber J. Deep learning in neural networks: an overview. *Neural Network* 2015;61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- [17] Agga A, Abbou A, et al. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. *Renew Energy* 2021; 177:101–12. <https://doi.org/10.1016/j.renene.2021.05.095>.
- [18] Dai Y, Wang Y, et al. LOWESS smoothing and Random Forest based GRU model: a short-term photovoltaic power generation forecasting method. *Energy* 2022;256: 124661. <https://doi.org/10.1016/j.energy.2022.124661>.
- [19] Wang F, Xuan Z, et al. A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Convers Manag* 2020;212:112766. <https://doi.org/10.1016/j.enconman.2020.112766>.
- [20] Zhen H, Niu D, et al. Photovoltaic power forecasting based on GA improved Bi-LSTM in microgrid without meteorological information. *Energy* 2021;231:120908. <https://doi.org/10.1016/j.energy.2021.120908>.
- [21] Gao M, Li J, et al. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. *Energy* 2019;187:115838. <https://doi.org/10.1016/j.energy.2019.07.168>.
- [22] Wang F, Yu Y, et al. Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting. *Appl Sci* 2018;8(8):1286. <https://doi.org/10.3390/app8081286>.
- [23] Lin W, Zhang B, et al. Multi-step prediction of photovoltaic power based on two-stage decomposition and BiLSTM. *Neurocomputing* 2022;504:56–67. <https://doi.org/10.1016/j.neucom.2022.06.117>.
- [24] Korkmaz D. SolarNet: a hybrid reliable model based on convolutional neural network and variational mode decomposition for hourly photovoltaic power forecasting. *Appl Energy* 2021;300:117410. <https://doi.org/10.1016/j.apenergy.2021.117410>.
- [25] Wang K, Qi X, et al. Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy* 2019;189:116225. <https://doi.org/10.1016/j.energy.2019.116225>.
- [26] Dragomireski K, Zosso D. Variational mode decomposition. *IEEE Trans Signal Process* 2013;62(3):531–44. <https://doi.org/10.1109/TSP.2013.2288675>.
- [27] Zhang X, Miao Q, et al. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery. *Mech Syst Signal Process* 2018;108:58–72. <https://doi.org/10.1016/j.ymssp.2017.11.029>.
- [28] Bertsekas DP. Multiplier methods: a survey. *Automatica* 1976;12(2):133–45. [https://doi.org/10.1016/0005-1098\(76\)90077-7](https://doi.org/10.1016/0005-1098(76)90077-7).
- [29] Torres ME, Colominas MA, et al. A complete ensemble empirical mode decomposition with adaptive noise. *Conference A complete ensemble empirical mode decomposition with adaptive noise*. IEEE, p. 4144–4147.
- [30] Li J, Liu C, et al. GPR signal denoising and target extraction with the CEEMD method. *Geosci Rem Sens Lett IEEE* 2015;12(8):1615–9. <https://doi.org/10.1109/LGRS.2015.2415736>.
- [31] Mi X, Liu H, et al. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Convers Manag* 2019;180:196–205. <https://doi.org/10.1016/j.enconman.2018.11.006>.
- [32] Hassani H, Zhigljavsky A. Singular spectrum analysis: methodology and application to economics data. *J Syst Sci Complex* 2009;22:372–94. <https://doi.org/10.1007/s11424-009-9171-9>.
- [33] Yu Y, Si X, et al. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput* 2019;31(7):1235–70. https://doi.org/10.1162/neco_a_01199.
- [34] Peng T, Zhang C, et al. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* 2021;221:119887. <https://doi.org/10.1016/j.energy.2021.119887>.
- [35] Siami-Namini S, Tavakoli N, et al. The performance of LSTM and BiLSTM in forecasting time series. *Conference the performance of LSTM and BiLSTM in forecasting time series*. IEEE, p. 3285–3292.
- [36] Nizar IM, Adytia D, et al. Forecasting of temperature by using LSTM and bidirectional LSTM approach: case study in Semarang, Indonesia. *Conference Forecasting of temperature by using LSTM and bidirectional LSTM approach: case study in Semarang, Indonesia*. IEEE, p. 146–150.
- [37] Chua LO, Roska T. The CNN paradigm. *IEEE Trans Circ Syst I* 1993;40(3):147–56. <https://doi.org/10.1109/81.222795>.
- [38] Chen Y-C, Huang W-C. Constructing a stock-price forecast CNN model with gold and crude oil indicators. *Appl Soft Comput* 2021;112:107760. <https://doi.org/10.1016/j.asoc.2021.107760>.
- [39] Somu N, Mr Gr, et al. A deep learning framework for building energy consumption forecast. *Renew Sustain Energy Rev* 2021;137:110591. <https://doi.org/10.1016/j.rser.2020.110591>.
- [40] Desai M, Shah M. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). *Clinical eHealth* 2021;4:1–11. <https://doi.org/10.1016/j.ceh.2020.11.002>.
- [41] Fildes R. The evaluation of extrapolative forecasting methods. *Int J Forecast* 1992; 8(1):81–98. [https://doi.org/10.1016/0169-2070\(92\)90009-X](https://doi.org/10.1016/0169-2070(92)90009-X).