Energy 232 (2021) 120996

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Day-ahead hourly photovoltaic power forecasting using attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern

Jiaqi Qu [a], Zheng Qian [a, *], Yan Pei [b]

[a] Beihang University, Beijing, China
[b] State Key Laboratory of Operation and Control of Renewable Energy & Storage Systems, China Electric Power Research Institute, Beijing, China

A R T I C L E I N F O

Article history:
Received 30 November 2020
Received in revised form
30 April 2021
Accepted 17 May 2021
Available online 20 May 2021

Keywords:
Photovoltaic power forecasting
Attention based CNN-LSTM neural Network
Multiple relevant and target variables prediction pattern
Time series forecasting

A B S T R A C T

Accurate forecasting of photovoltaic power plays a pivotal role in the integration, operation, and scheduling of smart grid systems. Notably, volatility and intermittence of solar energy are the primary constraints influencing the accuracy of photovoltaic power prediction. This work proposes, an attention-based long-term and short-term temporal neural network prediction model (ALSM) assembled using the convolutional neural network (CNN), long short-term memory neural network (LSTM), and attention mechanism under the multiple relevant and target variables prediction pattern (MRTPP). This is geared towards capturing the short-term and long-term temporal modes and achieving the day-ahead hourly photovoltaic power forecasting. The proposed method is verified by the historical data of the photovoltaic system downloaded from the DKASC website. Consequently, the results indicate that the forecasting accuracy using the MRTPP pattern is better than those common input-output prediction patterns. Moreover, the proposed ALSM model under the MRTPP pattern demonstrates more superiority compared to a few PV power forecasting methods including the statistical methods as well as artificial intelligence methods. Subsequently, different important parameters affecting the accuracy of forecasting range of the model are analyzed, and suggestions on memory lengths corresponding to the divergent prediction range are provided.

1. Introduction

As one of the most popular renewable energy sources, solar energy is characterized by abundant resources, easy access, and low cost of use [1,2]. With the rising demand for energy in developing countries, photovoltaic power generation is annually growing at an increasing rate, as such, it presents a great challenge to the generation capacity of the power system [3]. Nonetheless, due to the effect of solar radiation and environment temperature, the output power of the photovoltaic grid-connected system harbors the essence of intermittence, fluctuation, and randomness, which causes significant confusion in the operation, scheduling, and planning of the power system [4]. In addition, high-precision photovoltaic power forecasting effectively improves the utilization of solar energy, thereby increasing the returns on investment ratio of power stations, and reducing the economic losses attributed to power restrictions [5]. Therefore, precise prediction is often prescribed for photovoltaic power.

Notably, the forecasting of photovoltaic power belongs to the category of time series forecasting. Based on the classification of forecasting variables, time series forecasting can be subdivided into univariate forecasting and multivariate forecasting [6,7]. Also, it can be divided into direct forecasting and indirect forecasting according to the classification of the forecasting process [8,9]. At present, research on photovoltaic power forecasting primarily underscores on 2-time scales, including, ultra-short-term photovoltaic power forecasting i.e., hour-ahead, and short-term photovoltaic power forecasting i.e., day-ahead [10,11]. The former is majorly used to guide real-time scheduling of the power grid, whereas the latter provides significant data support for the day-ahead generation plan [12]. For any kind of classification, the methods of PV forecasting research can be implemented by different forecasting methods including physical methods, statistical methods, and deep learning

* Corresponding author. Beihang University, Xueyuan Road No.37, Haidian District, Beijing, 100191, China.
E-mail address: qianzheng@buaa.edu.cn (Z. Qian).

https://doi.org/10.1016/j.energy.2021.120996
0360-5442/

methods [13].

Physical forecasting methods do not need historical data but instead rely on exclusive geographic information, precise meteorological data, and complete photovoltaic battery information [14]. The meteorological parameters at the prediction time generated via numerical weather prediction (NWP), sky images, or satellite images of cloud associate with the installation angle, conversion efficiency, among other parameters in establishing a physical model, before directly calculating the photovoltaic power. Nevertheless, physical methods primarily rely on NWP, which takes a relatively long time to calculate and only produces meteorological data after 6 h, hence, its application to ultra-short-term targets remains limited [15]. Sky images and satellite images of the cloud are used to predict ultra-short-term solar irradiance [16,17]. However, due to the low resolution of geographic data and the small coverage of ground-sky images, its prediction accuracy and practicability warrant further improvement. Additionally, the parameters provided by the manufacturers of PV modules deviate from the actual operation, with varying understanding of the empirical parameters based on regions. Thus, there will be deviations in the physical models resulting in low precision [18].

Statistic forecasting methods establish the mapping relationship between historical data and target forecasting data via fitting and completing the prediction of future photovoltaic power [19]. The commonly used statistical methods possess the advantages of easy modeling and strong interregional versatility. Nevertheless, the photovoltaic power generation time series remains a complex time series with dynamic and non-periodic modes which can weaken the precondition of a large amount of correct historical data for the application of statistical methods. Therefore, the collection and computation of precise data during the process of an actual implementation remain challenging.

The development of artificial neural network technology has significantly improved the prediction accuracy of photovoltaic power [20–22], its ability of non-linear processing fits the variation law of photovoltaic power. However, the main demerit of these photovoltaic power forecasting methods is that their learning models are relatively shallow [23]. Moreover, due to the complexity of the weather condition, the shallow models may not be able to fully extract the corresponding deep nonlinear characteristics and time series dynamic characteristics of photovoltaic power data [24]. And the tasks of feature extraction and nonlinear mapping of photovoltaic power forecasting are much more challenging. Therefore, a powerful way to deal with the deficiencies of shallow models is to adopt deep learning methods, on account of the sufficient ability of feature extraction and feature transformation [25].

In recent years, the application of deep learning in photovoltaic power prediction has a very significant effect. In Ref. [25], the wavelet transform (WT) is integrated into the deep convolutional neural network (DCNN) models, and the proposed hybrid WT-DCNN method is applied to predict the PV power in the various horizons. The proposed method only absorbs the photovoltaic power data to establish the model, and other variables affecting the power are not inspected as the input of the forecasting model. In addition, dealing with incomplete data sets becomes the obstacle of the method, because it relies on the complete decomposition of the PV data sequence. Another example is proposed in Ref. [26], which captures the solar irradiance behavior with the deep learning method relying on long short-term memory (LSTM) networks using day-ahead weather forecast data as input for prediction. Then, the mathematical model between irradiance and PV power is established by physical theory to realize the indirect forecasting. However, the prediction results of PV power are greatly affected by the accuracy of weather forecast data, and a large number of calculation formulas and complicated processes need to be carried out. In addition, the error of hourly day-ahead weather forecast variables issued by meteorological service agencies will directly lead to the error of solar irradiance forecast. In Ref. [27], a time series prediction model based on evolutionary attention-based LSTM is proposed. The attention mechanism can assign weights to the features in time series according to time steps, so as to cope with the dilemma of attention dispersion in the traditional LSTM method. However, the forecasting method based on LSTM which only depends on weights has the finite ability to extract complex features between variables. In Ref. [28], a deep learning model based on long-short-term memory recurrent neural network (LSTM-RNN) under the framework of partial daily pattern prediction (PDPP) is proposed for day-ahead PV power forecasting. However, the prediction model based on similar weather types breaks the law essence of the original time series of power data, and the method temporarily verifies the effect of 1-step ahead, and the prediction ability for further horizons is unknown. In Ref. [29], a hybrid time series prediction model based on convolution neural network (CNN) and long short-term memory network (LSTM) is proposed. In fact, this method ignores the gravity of feature selection, that is, all the related environmental variables that may affect the PV power in different degrees are taken as input. Therefore, the introduction of too much redundant information and strict variables requirements will confine the application of the model.

According to existing forecasting methods, there exist a few handles that need further conquering. First of all, the prediction range of weather forecast data is limited, and the historical time series of photovoltaic power is non-stationary, dynamic, and non-periodic, which cannot be interpreted by traditional artificial intelligence methods. Secondly, input-output prediction patterns of the existing modes explore from the perspective of statistical analysis, ignoring the influence of other related factors, or require high-quality data of multiple related factors, limiting the practical application [30,31]. Finally, the complex non-linearity of the photovoltaic time series for various forecasting horizons exists between univariate time steps and among the relevant variables.

Specifically, to deal with the existing obstacles and fulfill the accurate day-ahead hourly PV power forecasting, the hypothesis of the proposed method in this paper is mainly derived from the following aspects: from the current research, it can be seen that CNN has a good ability to extract spatial features, LSTM can draw out the temporal features, and attention mechanism can elude the flaws of distraction. Therefore, we design the attention-based long-term and short-term temporal neural network prediction model (ALSM), a well-designed deep learning network with better forecasting performance. It contains two temporal modules based on CNN and LSTM considering the timing characteristics of PV power series, which focus on the short-term periodicity and long-term periodicity of output power respectively, and extracts the spatial and temporal correlation between multiple relevant variables and target variables in continuous time slots. Moreover, combined with forecasting results of the long-term temporal module and short-term temporal module, the attention mechanism of weight allocation is also introduced to form the ALSM model. In addition, Since the irradiance, temperature and other environmental factors have a negligible influence on the mapping with PV power, as well as the different spatial-temporal correlation between external factors on each prediction horizon. Then, the input-output prediction pattern of multiple relevant and target variables prediction pattern (MRTPP) is proposed in this paper, which absorbs the historical operation data of relevant variables and target variables as input variables into a deep learning model and predicts the target variables. Based on the above analysis, the high precision forecasting of photovoltaic power in different horizons is realized by exerting the proposed ALSM model under the MRTPP pattern.

In general, the primary research contents explored in this paper include: firstly, the ALSM model based on deep learning technology and the input-output pattern of MRTPP suitable for photovoltaic power prediction are studied. Secondly, the forecasting performance of the proposed method is compared with different methods, among them, statistical models including ARMA and ARIMA, and artificial intelligence models including LSTM and CLSTM. Finally, the parameters of the proposed method are analyzed for various forecasting horizons. In addition, the main contributions of this paper can be summarized as follows:

(1) We propose the attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern as an accurate prediction method of day-ahead hourly photovoltaic power forecasting.
(2) We analyze and compare the prediction results of the combined structure of the ALSM model under the MRTPP pattern. It can further illustrate the effectiveness of each structure of the proposed method in PV power time series forecasting.
(3) We study the reliable memory length parameters of the proposed method in different prediction range to ensure the optimal prediction effect.

In this paper, Section 2 introduces the structure of the proposed ALSM model under the MRTPP pattern; Section 3 explores the experimental designs and evaluation indicators; Section 4 analyzes and compares the performance of the models and presents the results and discussion; Finally, Section 5 gives the conclusion of the paper.

2. Structure of attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern

2.1. Structure of proposed ALSM model

The overall structure of the proposed forecasting model (ALSM) based on convolutional neural network (CNN), long short-term memory neural network (LSTM) and attention mechanism is shown in Fig. 1. It is majorly made of a short-term temporal module, a long-term temporal module, and an attention module. For a profound understanding of the model, its detailed description is as follows:

(1) Short-term temporal module.

This module takes the short-term (e.g. a few hours) time mode existing in the photovoltaic power into account to make the prediction. Specifically, it applies the CNN + LSTM structure (CLSTM) to learning short-term periodicity, which takes continuous data sequences with length of $\boldsymbol{p}$ slots as input, given by

$$X^s = \left[X_{\tau+1-p\Delta}, X_{\tau+1-(p-1)\Delta}, \ldots, X_{\tau+1-\Delta}\right] \in R^p \tag{1}$$

where $[X_{\tau+1-p\Delta}, X_{\tau+1-(p-1)\Delta}, \ldots, X_{\tau+1-\Delta}]$ is a sequential series of observed measurements; $\boldsymbol{\tau}$ denotes the time stamp of current; $\boldsymbol{\Delta}$ denotes a fixed 1-h interval of time series; $\boldsymbol{p}$ represents the number of input length named memory length.

(2) Long-term temporal module.

This module takes the long-term (e.g. a few days) time mode existing in the photovoltaic power into account to make a forecast, which is similar to the short-term time module. Nevertheless, due to the influence of gradient disappearance, recursive units of LSTM hardly capture long-term time style. Given the photovoltaic power generated by PV modules having a cycle mode of no power generation at night, the link between the currently hidden cell and the hidden cell in the adjacent period is designed to skip in a similar stage. For the photovoltaic power data with fixed cycle mode, the determined data interval period is 1 day noted as Δl. Therefore, the data sequences with length of $\boldsymbol{p}$ slots is considered as input, i.e., given by

$$X^L = \left[X_{\tau+1-p\Delta l}, X_{\tau+1-(p-1)\Delta l}, \ldots, X_{\tau+1-\Delta l}\right] \in R^p \tag{2}$$

where $\boldsymbol{\tau}$ denotes the time stamp of current; Δl represents the exact number of hidden cells to be skipped; $\boldsymbol{p}$ represents the number of input length named memory length.

(3) Attention module.

Due to the strong feature extraction ability of the attention mechanism, the attention module is designed to capture the dependencies between different sequences and the relationships between different time steps of the input time series. Then, a prediction of PV power $Y \in R^1$ is obtained.

The above three modules are mainly composed of CNN, LSTM,

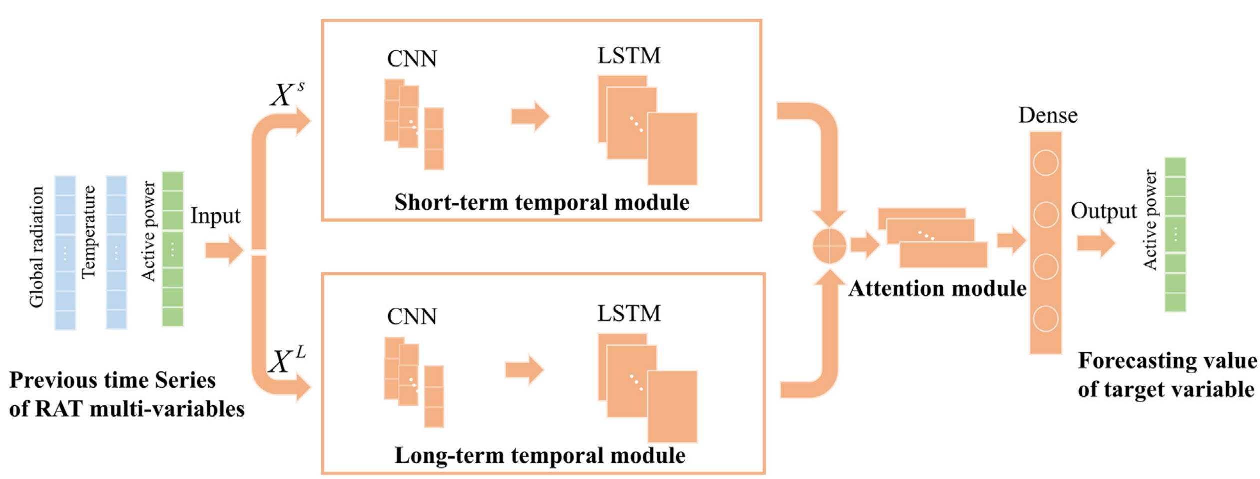


Fig. 1. The overall structure of the proposed forecasting model.

and attention mechanism, their specific principles are introduced in detail as follows:

(1) Convolutional neural networks (CNN)

Of note, convolutional neural networks (CNN) is a type of feedforward neural network. In contrast with other network models, the parameter sharing property of convolution reduces the number of parameters to be optimized, hence, improving the training efficiency and scalability of the model. The convolution operation is primarily used to process data in Euclidean space, thus, it has significant advantages for time-series prediction and image recognition [32].

In this paper, 1-D CNN is used in time series forecasting through the convolution kernel to extract features of short-term and long-term series patterns shown in Fig. 2. Each filter of 1-D CNN swept through the short-term matrix of $\triangle$-skip or long-term matrix of $\triangle l$-skip resampled from the input matrix of size $N \times M$, and produces a vector with size of $1 \times M$, where the filter size is k, and the ReLu activation function is adopted.

(2) Long short-term memory neural network (LSTM)

Due to the problem of gradient disappearance and gradient explosion, the learning ability of cyclic neural networks remains limited and the actual effect of cyclic neural networks is often unsatisfactory. The advantage of long short-term memory neural network (LSTM) in the relatively long-term memory of valuable information makes it broadly used in time series prediction [33,34].

The improvement of LSTM is to add three gates not only calculating h_t based on x_t and h_{t-1}, in Fig. 3. The i_t represents the output of the input gate, f_t represents the output of the forgetting gate, and the output of the output gate is o_t. In addition the c_t is the output of the memory cell [35]. The input gate controls whether the new state of the calculation can be updated to the memory unit. The forgetting gate controls the forgotten information in the previous memory unit. The output gate regulates the extent to which the current output depends on the current memory unit.

The calculation formula of step t is as follows:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{3}$$

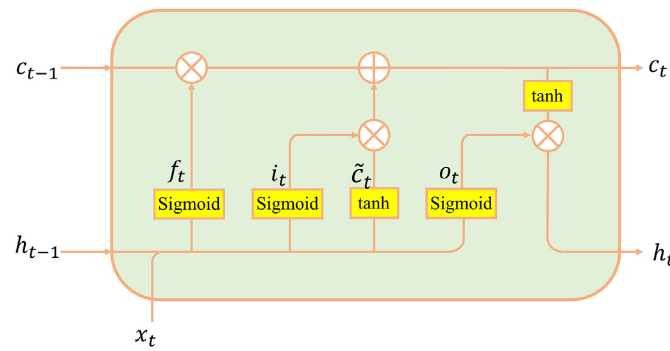


Fig. 3. The structure of LSTM cell.

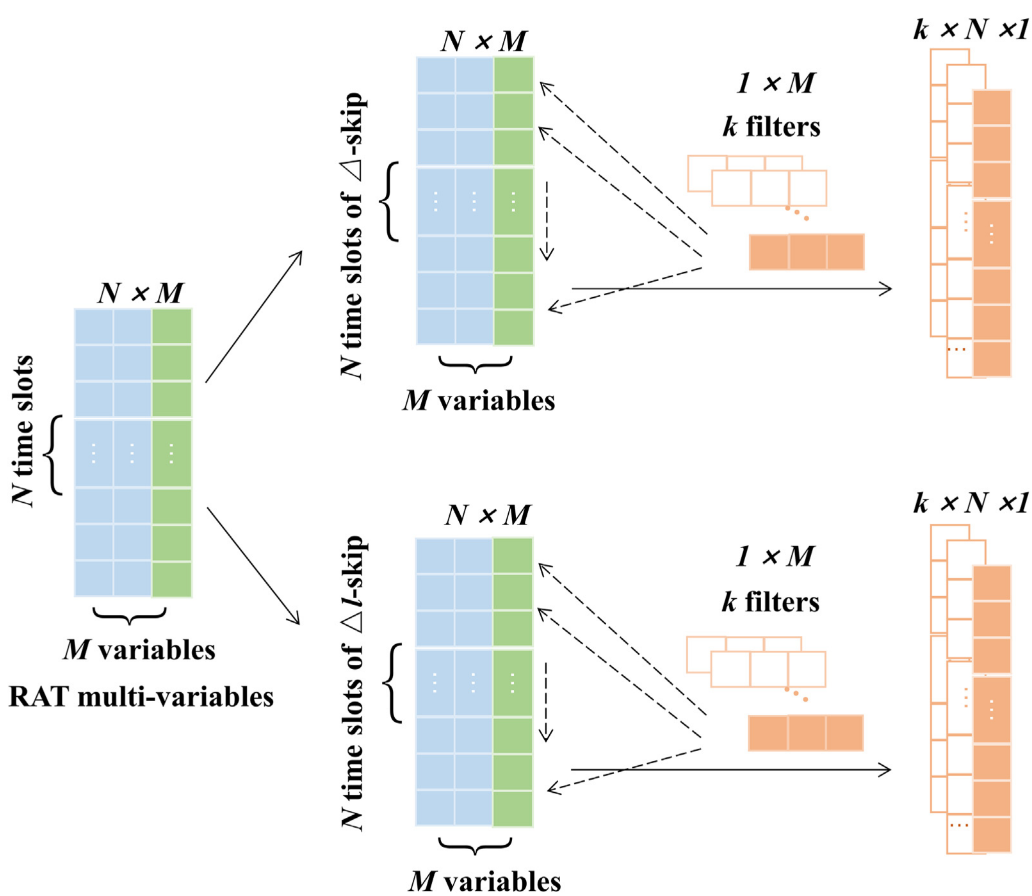


Fig. 2. The 1-D CNN extracting features of short-term and long-term series patterns.

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{4}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{5}$$

$$\tilde{c}_t = tanh(W_c x_t + U_c h_{t-1}) \tag{6}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_{t-1} \tag{7}$$

$$h_t = o_t \odot tanh(c_t) \tag{8}$$

where W_i, W_f, W_o, W_c and U_i, U_f, U_o, U_c represent the weight matrices for input matrices hidden layer matrices of the three gates as mentioned previously; b_i, b_f, and b_o represent the bias vectors of the three gates; σ represents the activation of sigmoid; tanh represents the activation of hyperbolic tangent.

Unlike the traditional recurrent neural network, the transition from state c_{t-1} to current state c_t is determined by the calculated state of activation function and controlled by the input gate and forgetting gate. When the value of the forgetting gate of LSTM is close to 1 and the value of the input gate is close to 0, the long-term memory function is realized. Besides, when the value of the input gate is close to 1, and the value of the forgetting gate is close to 0, the old memory is forgotten and the new important information is retained. This network structure makes it easier for LSTM to learn the long-term dependence between sequences compared to RNN [26].

(3) Attention mechanism

CNN and LSTM have been extensively used in research, however, they are limited by the computing power and optimization algorithm [36]. The attention mechanism imitates how the human brain processes information, which improves the ability of the neural network to process information [37].

In this paper, we construct a perceptron to learn the problems that should be paid attention to in prediction. The attention network in Fig. 4 comprise two full connection layers and a Softmax function that outputs a 1×2 vector represented by$w_i = [w_i^s, w_i^l]$.

The two elements of the vector represent the weights assigned to the predicted value. Therefore, by calculating the weighted value of the two vectors, the final value of predicted PV power can be obtained.

$$y_i = w_i^s h_i^s + w_i^l h_i^l \tag{9}$$

where the h_i^s and h_i^l represent the hidden state of short-term and long-term temporal modules; the w_i^s and w_i^l represent the weights assigned to the target forecasting of y_i.

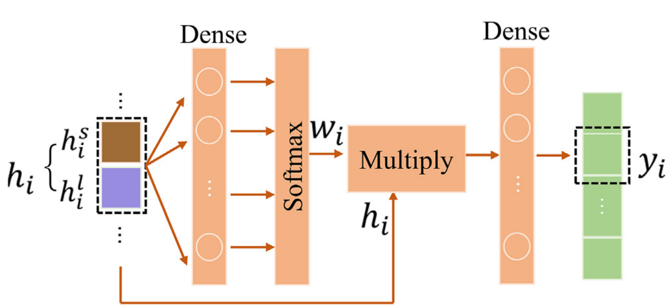


Fig. 4. The structure of the attention network.

2.2. Structure of proposed MRTPP pattern

Different input-output prediction patterns exist for the forecasting of time series based on the forms of input and output variables. At present, the single-input single-output of target variable pattern is widely used [38]. However, since the predicted target variable only relies on its previous time sequence, the ability to capture variation remains limited, specifically, the photovoltaic power is influenced by factors of the external environment. Fig. 5 depicts the trends of photovoltaic power and external environmental factors over time. It can be roughly seen that there is a certain influence relationship between them, especially the intensity of irradiance and temperature is close to the changing trend of PV power. Fig. 6 shows the correspondence between the different external environmental factors and PV power. Among them, irradiance has the most obvious influence on active power, followed by temperature and humidity. The existence of correlations provides a basis for the introduction of relevant variables into the model as input variables.

The input-output prediction patterns considering relevant variables include multiple-input multiple-output of relevant and target variables pattern [39] and multiple-input of relevant variables with single-output of target variables pattern [24], respectively, as shown in Fig. 7. To achieve high forecasting accuracy, this paper combines the advantages of the above 3 prediction models and proposes multiple relevant and target variables prediction pattern (MRTPP) in Fig. 8. Specifically, the target variable and its relevant multivariate time series all are considered as the input, corresponding to the photovoltaic power and the radiation, temperature, and other variables that affect the generation of photovoltaic power. Then the target variable, i.e., photovoltaic power, is regarded as the output.

2.3. Day-ahead PV power ensemble forecasting model

The common strategies for multi-step time series forecasting include direct multi-step forecast strategy and recursive multi-step forecast strategy [40]. The direct multi-step forecast strategy generates 24 data of all hours in a day at one time, hence, the relevance between the two adjacent values is constrained. Whereas. Recursive multi-step forecast strategy only generates photovoltaic power of the next hour and rolls 23 times, thus, the prediction error rapidly accumulates in the future, and the forecasting data become meaningless. This work uses direct-recursive hybrid multi-step forecast strategies which are combined with the former strategies to control the range of time steps of prediction by adding a flexible parameter, i.e., prediction horizon, to realize the forecasting of hourly photovoltaic power in a day.

The method proposed in this paper is applied to the attention-based long-term and short-term convolutional neural network forecasting model (ALSM) under multiple relevant and target variables prediction pattern (MRTPP). Specifically, the original input is the historical sequences of photovoltaic power and relevant environmental variables including global radiation, temperature, humidity, diffuse radiation, wind speed, rainfall, and wind direction, while the output range is the hourly photovoltaic power series that can be regulated by the horizon [41] as shown in Fig. 9. The primary implementation process includes: determining the relevant variables of global radiation, temperature, and PV power as input based on the Pearson correlation coefficients; dividing the training set according to the proposed MRTPP; training the ALSM model for different forecasting range; and predicting the hourly photovoltaic power of one day.

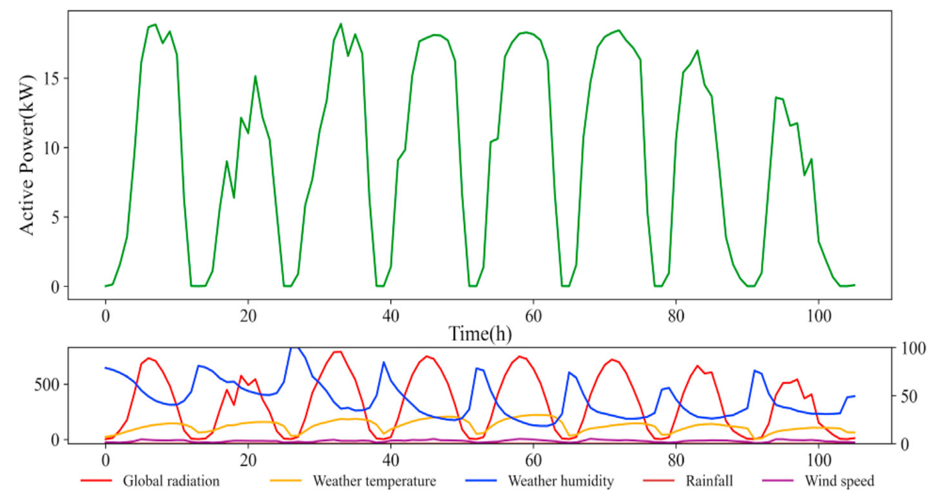


Fig. 5. The curve (time) of photovoltaic power with external environment factors of global radiation, weather temperature, weather humidity, rainfall, and wind speed for 8 consecutive days.

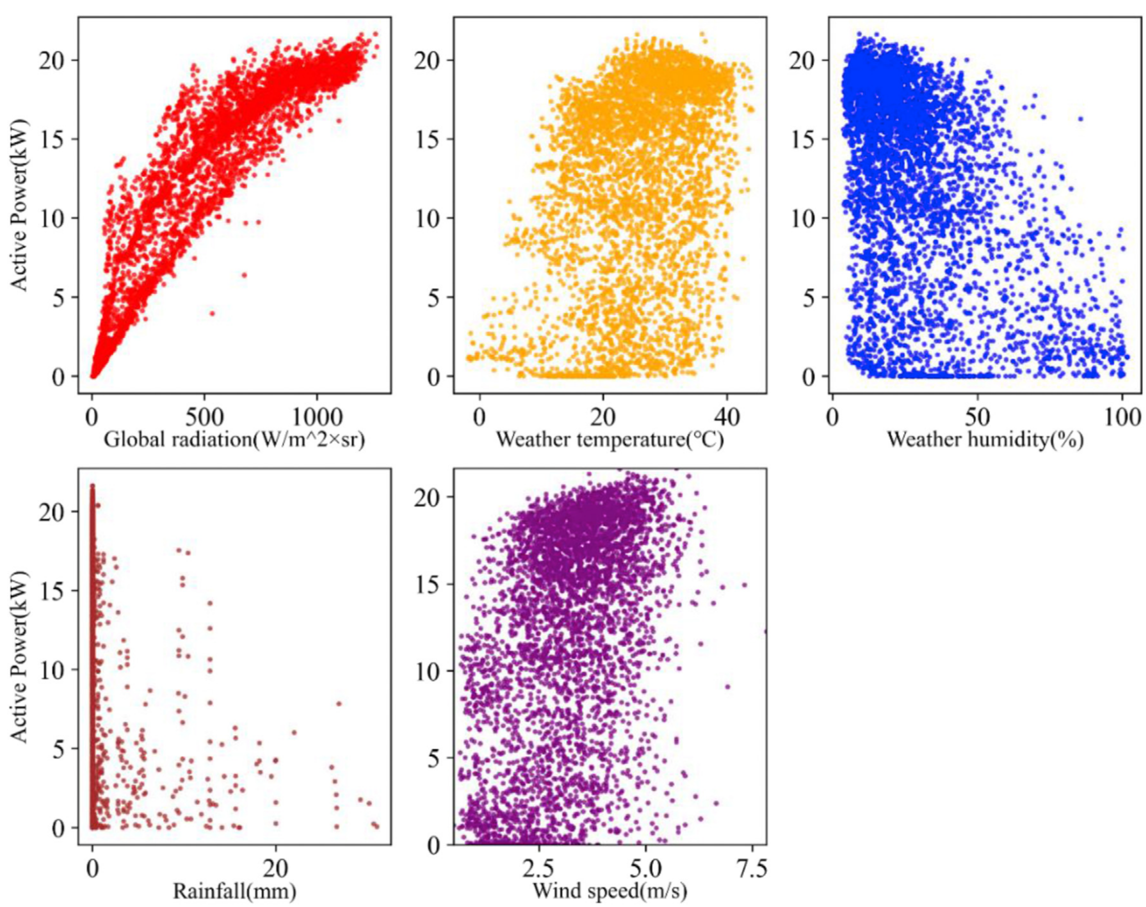


Fig. 6. Scatter diagram of the relationship between different environmental factors and PV power.

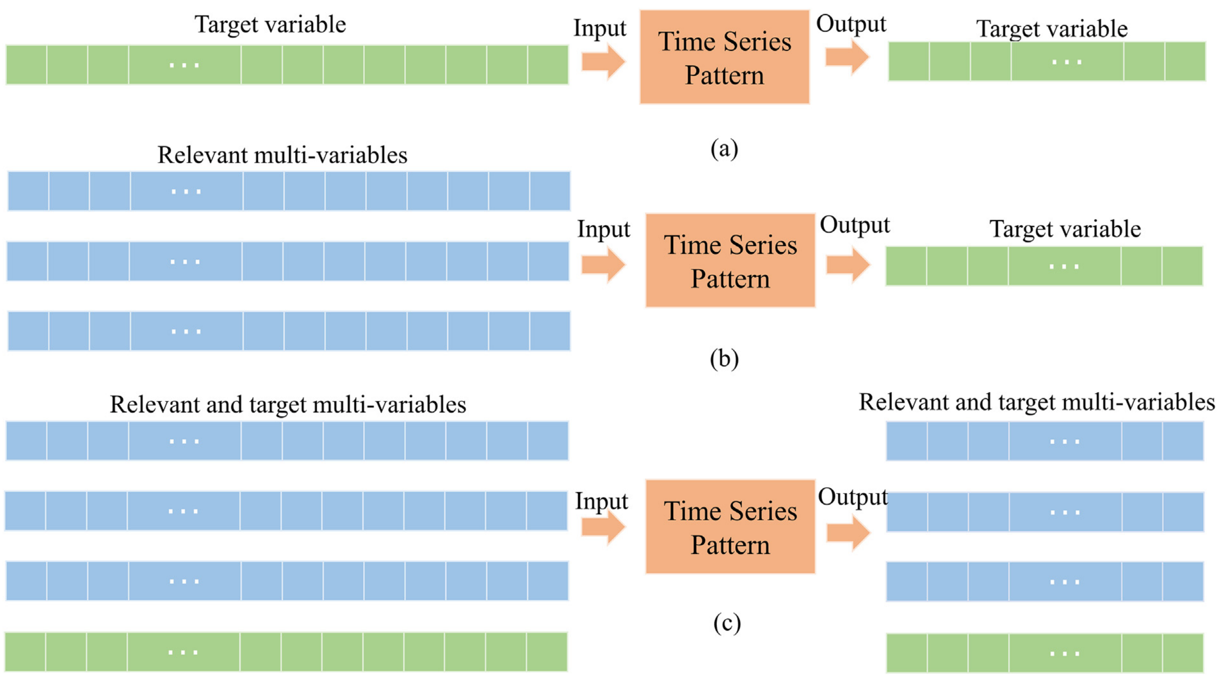


Fig. 7. Three forms of input-output patterns: (a) single-input single-output of target variable pattern; (b) multiple-input multiple-output of relevant and target variables pattern; (c) multiple-input of relevant variables with single-output of target variables pattern.

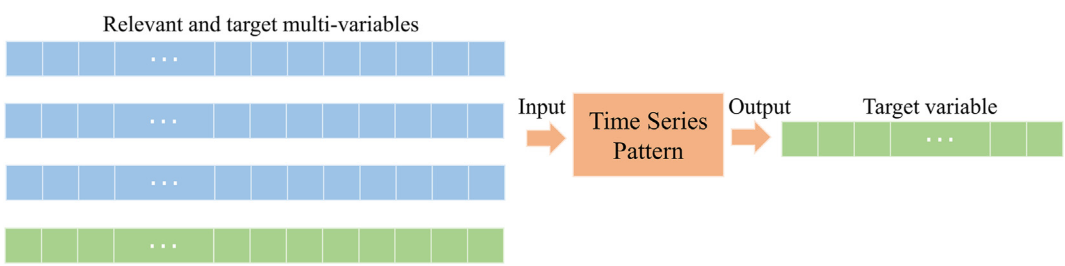


Fig. 8. The proposed input-output prediction pattern of multiple relevant and target variables prediction pattern (MRTPP).

3. Case study

3.1. Experimental data

The data in this work are downloaded from the public website of Alice Springs photovoltaic power system of DKASC, and the data of Site-1B between 2014 and 2017 are considered as a case study [42]. The specific information of this PV system is shown in Table 1. The original resolution of this dataset is 5 min, and to make the data fully reflect the real situation and avoid the impact of violent fluctuations including data transmission on the accuracy of the algorithm, they are hourly averaged. Besides, the power output of the PV modules in the morning and evening is significantly low, i.e., zero most of the time or close to zero. Therefore, we only consider the power between 6 a.m. and 7 p.m. Subsequently, the data are standardized, outliers are removed, and the missing values are replaced by the average PV power values of the previous hour and the next hour.

The training data are the data processed based on the above methods for 3 consecutive years from 2014 to 2016, and 20% of them are used as validation data., and the data in 2017 serve as the testing data. Specifically, the time range, the data volume, and the percentage of each dataset are shown in Table 2.

3.2. Performance indices

In this paper, the 3 indicators are used to evaluate the predictive performance and data characteristics of the neural network, including normalized root mean square error (NRMSE), normalized mean absolute error (NMAE) [43], and coefficient of determination (R^2) [44]. Definition of them are as follows:

$$MSE = \left(\sum_{i}^{m}\left(y_i - \hat{y}_i^2\right)\right) \Big/ m \tag{10}$$

$$RMSE = \sqrt{\left(\sum_{i}^{m}\left(y_i - \hat{y}_i^2\right)\right) \Big/ m} \tag{11}$$

$$NRMSE = (\mathrm{RMSE}/(y_{max} - y_{min})) \times 100\% \tag{12}$$

$$MAE = \left(\sum_{i}^{m}\left|(y_i - \hat{y}_i)\right|\right) \Big/ m \tag{13}$$

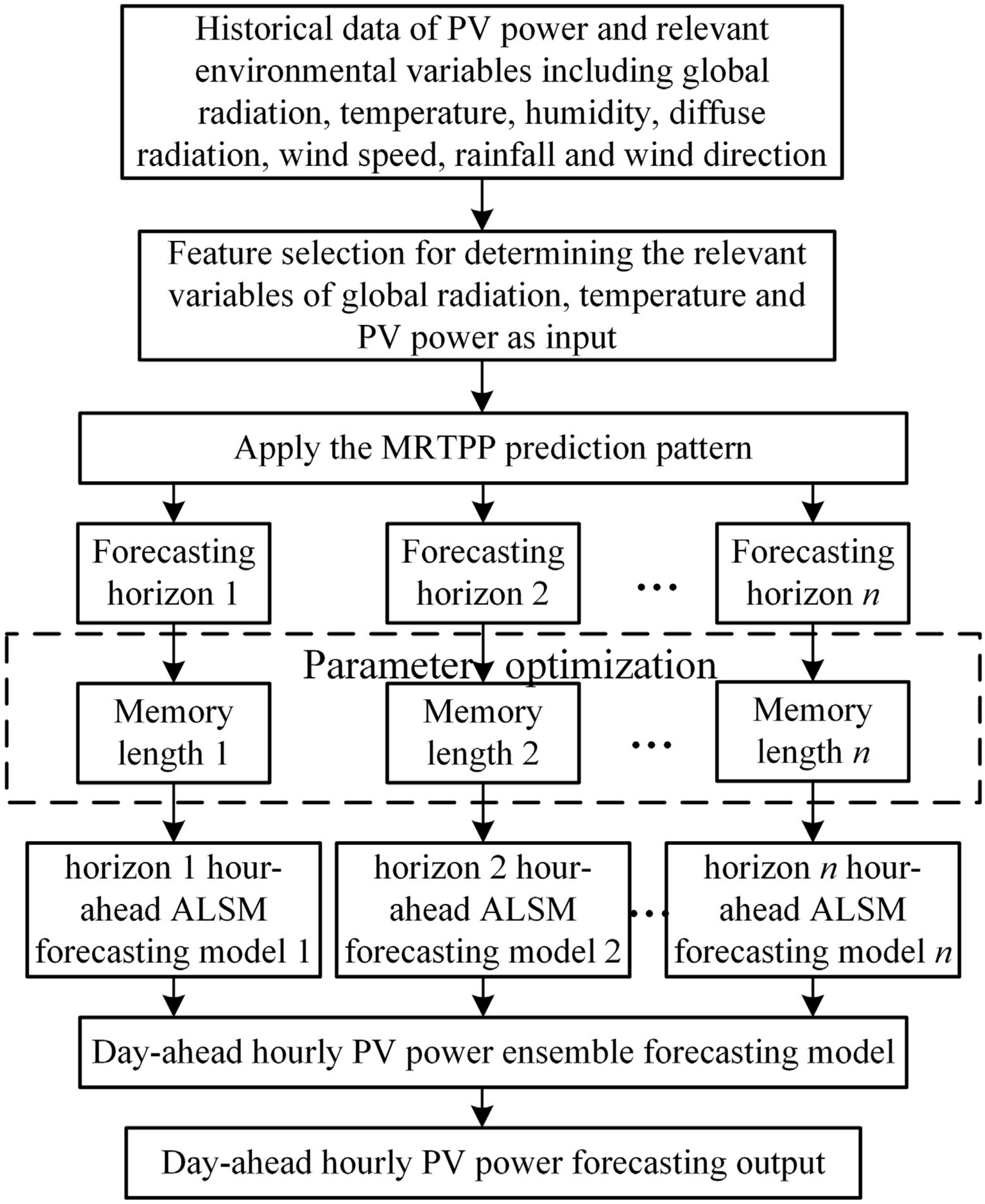


Fig. 9. The proposed day-ahead PV power ensemble forecasting model.

Table 1
The specific information of the Site-1B PV system.

Detailed information	
Latitude	−23.7625
Longitude	133.8753
Array rating	23.4 kW
PV Technology	Mono-Si
Array Area	$4 \times 38.37m^2$
Tracker	DEGERenergie 5000NT, dual axis
Inverter	4×6 kW, SMA SMC 6000A
Installation Time	Thu, Jan 8, 2009
Array Tilt/Angle	Variable: Dual axis tracking

Table 2
The specific information of the different dataset.

Dataset	Number of data	Time range
Training data (80%)	11378	2014.01.01−2016.12.31
Validation data (20%)	2844	
Testing data	4745	2017.01.01−2017.12.31

$$NMAE = (\mathrm{MAE}/(y_{max} - y_{min})) \times 100\% \tag{14}$$

$$R^2 = 1 - \left(\sum_i^m \left(y_i - \widehat{y}_i \right)^2 \right) \Big/ \left(\sum_i^m \left(y_i - \frac{1}{m} \sum_i^m y_i \right)^2 \right) \quad (15)$$

where y_i is the real PV power data; $\widehat{y}_i$ is the forecasting value; m denotes the number of testing set; y_{max} and y_{min} represent the maximum value and minimum value of PV power in the testing set respectively.

RMSE and MAE are commonly used to evaluate the performance of regression models. In fact, a small number of outliers can make RMSE and MAE worse. The NRMSE and NMAE are obtained by normalizing RMSE and MAE, which can reduce the influence of outliers on absolute error. R^2 is generally used in regression model to evaluate the goodness of fit between predicted value and actual value of regression model. Generally speaking, the closer R^2 is to 1, the better the explanation of independent variable to dependent variable in regression analysis.

3.3. Experimental design

To illustrate the efficacy of the proposed method, the following 4 aspects will be analyzed.

(i) The ALSM model under traditional input-output prediction patterns and the proposed multiple relevant and target variables prediction pattern (MRTPP) are tested to predict the photovoltaic power, and the performance of these patterns are evaluated. (ii) Based on the comparison results of the first part of the input-output patterns, the proposed ALSM model is compared with the traditional statistical methods ARMA, ARIMA, and neural network methods LSTM, CNN-LSTM. (iii) The critical parameter, memory length, is analyzed to evaluate the performance of combining models with different structures of the ALSM. (iv) The flexible forecasting horizons with the best memory lengths are analyzed to evaluate the performance of combining models with different structures of the ALSM.

The specific training and prediction process follows in Fig. 10, firstly, the prediction patterns are analyzed and selected, then the forecasting range and important parameters are determined, finally, the different types of PV power forecasting models are applied. All models are implemented by Tensorflow 1.9.0 with Python 3.7. Moreover, the Adam optimizer and loss of mean square error (MSE) are conducted with a grid search approach. The parameter settings of the proposed ALSM network in this paper are were evaluated using trial and error method. The training and validation epoch is set as 50.

4. Results and discussion

4.1. Performance evaluation for ALSM under MRTPP and traditional prediction pattern

Table 3 Demonstrates the Pearson correlation coefficients calculated based on formula (16) between photovoltaic power and the relevant variables, in which correlations were calculated and the Pearson correlation coefficient with a p-value < 0.05 was considered significant [45,46]. In fact, Pearson correlation coefficients and p-values were calculated using python's “scipy.stats.pearsonr” function. The predicted target variable represents photovoltaic power, which is defined as Y. Base on the significance levels of the correlation, the top 3 are global radiation, weather temperature, and weather relative humidity, which are defined as X1, X2, X3. The input-output forecasting patterns considering relevant variables include multiple-input of relevant variables with single-output of target variables pattern (simply defined as X−Y), multiple-input multiple-output of relevant and target variables pattern (simply defined as XY-XY), as well as the proposed MRTPP of multiple-input of relevant and target variables with single-output of target variables pattern (simply defined is XY-Y). The top 3 variables with the highest Pearson correlation coefficient $\rho_{Xi,Y}$ are added as the input relevant variables [47], in turn, and the forecasting results of each prediction mode are shown in Fig. 11, in where, the single-input single-output of target variable pattern (simply defined as Y−Y) only considering the univariate of the target variable is also included.

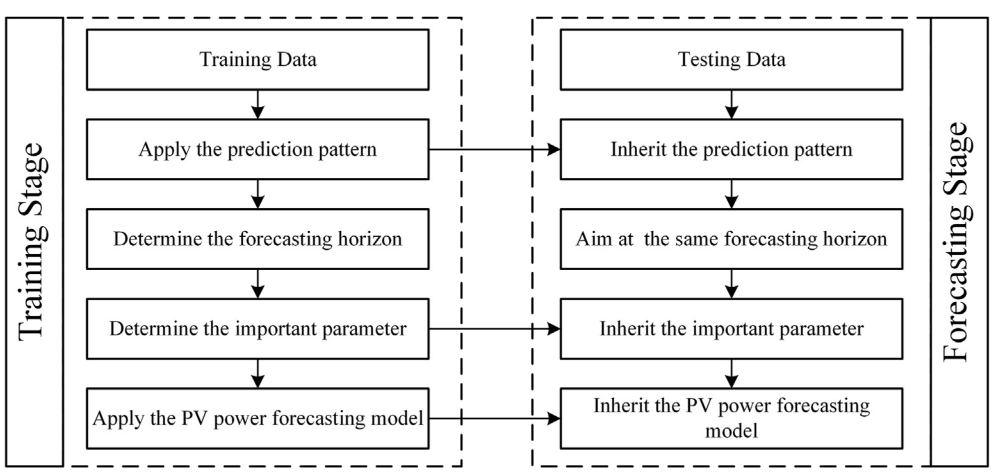


Fig. 10. The training and forecasting process of experiments.

$$\rho_{Xi,Y} = \frac{\mathrm{cov}(Xi, Y)}{\sigma_{Xi}\sigma_{Y}} \tag{16}$$

where $\mathrm{cov}(Xi, Y)$ represents the covariance between Xi and Y; σ_{Xi} and σ_{Y} represent the standard deviation for Xi and Y respectively.

The input is the hourly data sequences with a length of 3 days before the target time, and the output is the predicted data of the target time. The forecasting impact of each prediction pattern with the ALSM model on the predicted target for 7 consecutive days is considered as an example. In Fig. 11, the photovoltaic power curves of the different relevant variables of prediction patterns are represented by different groups and colors. For the pattern of regarding relevant variables as input, the strategy of absorbing the X1X2 as input exhibit higher accuracy compared to using only the highest relevant X1 and the top three highest relevant variables X1X2X3. This may potentially be the reason for the limited information contained in a single relevant variable. Conversely, drawing

Table 3
Pearson correlation coefficients and p-values between photovoltaic power and the relevant variables.

Variable	Symbol	Relevant coefficient $\rho_{Xi,Y}$	P-value
Active Power (kW)	Y	1	0
Global Radiation (w/ $m^2 \times sr$)	X1	0.85	0
Weather Temperature (°C)	X2	0.38	0
Weather Relative Humidity (%)	X3	0.36	0
Wind Speed (m/ s)	X4	0.18	1.4e-248
Diffuse Radiation (w/ $m^2 \times sr$)	X5	0.23	7.2e-115
Daily Rainfall (mm)	X6	0.08	0
Wind Direction ($\hat{A}^{o}$)	X7	0.03	1.2e-43

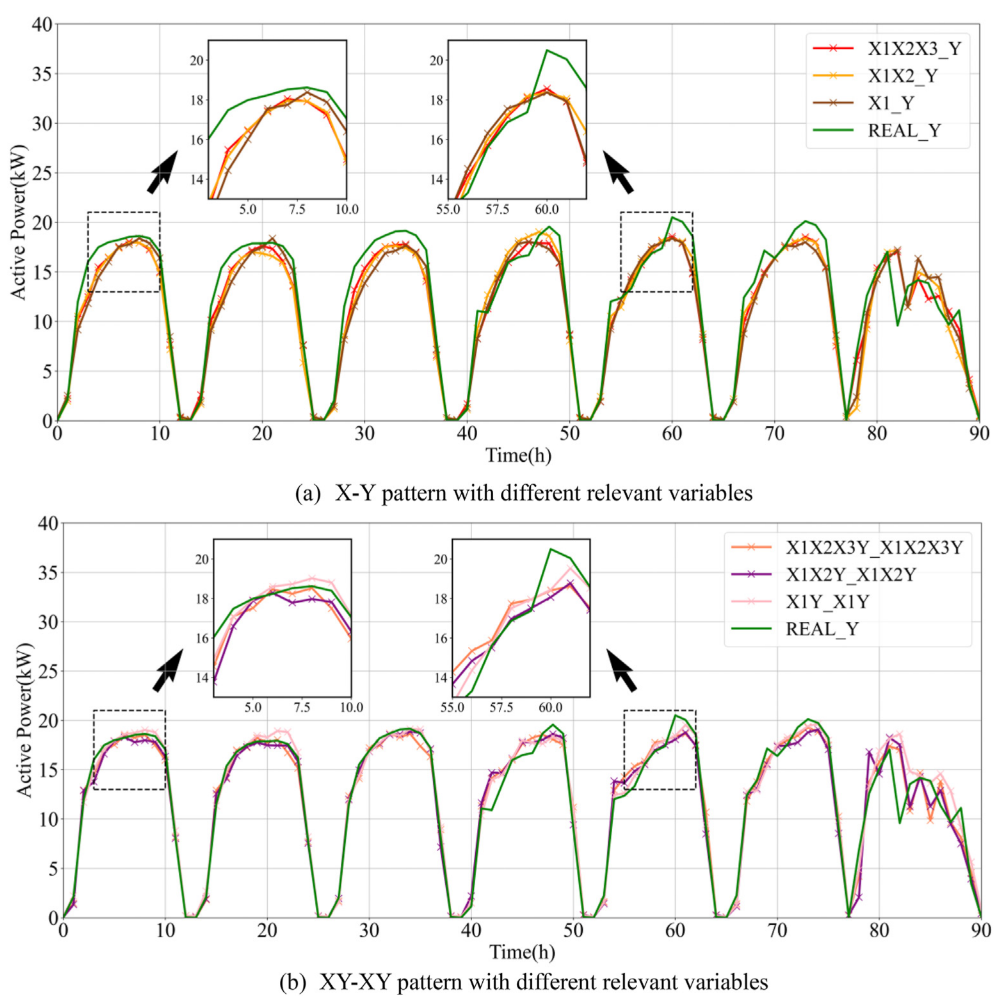
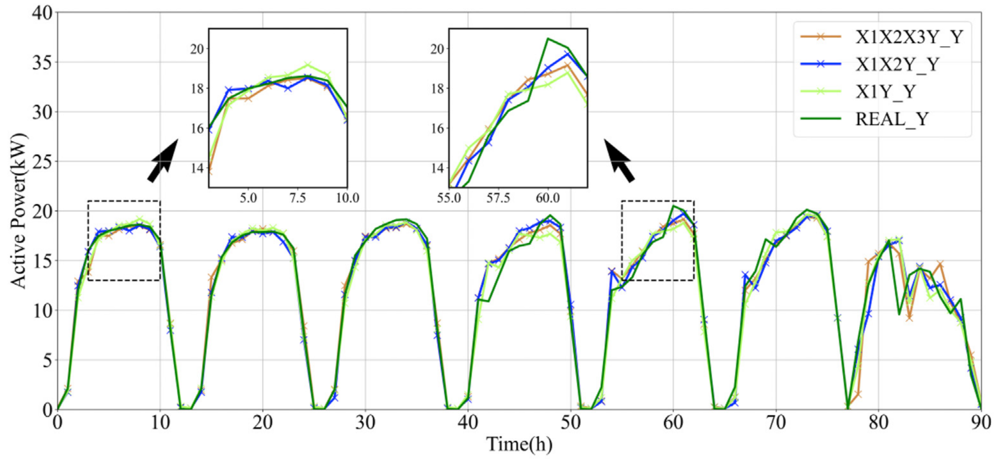
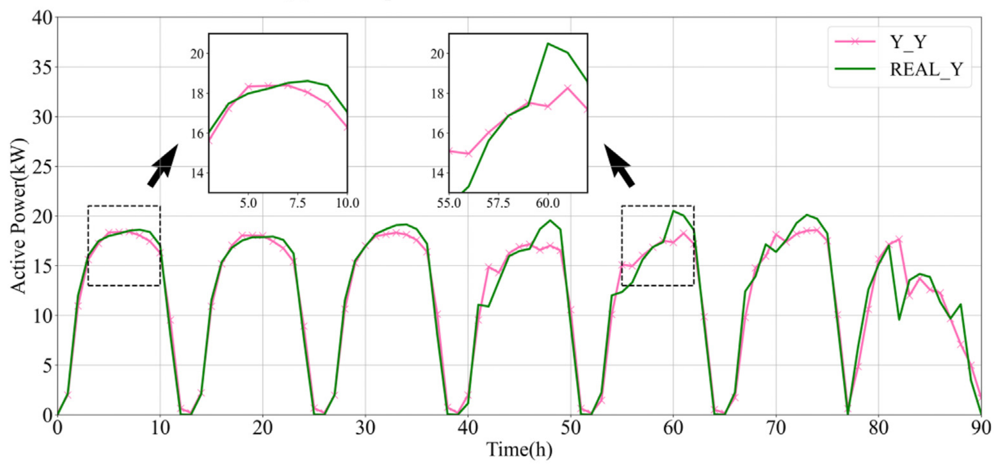


Fig. 11. Comparisons of four input-output prediction patterns with different relevant variables for photovoltaic power forecasting.



(c) XY-Y pattern with different relevant variables



(d) Y-Y pattern with different relevant variables

Fig. 11. (*continued*).

too many relative secondary factors may cause information pollution and necessitates complex learning ability of the model.

It can be seen from Table 4 That the Y−Y prediction pattern is more precise than the X−Y prediction pattern, perhaps because the X−Y model is proposed only from the function mapping ignoring nature. Additionally, the optimal relevant variables strategy (X1X2) for every prediction pattern is selected to compare the forecasting effect between the proposed MRTPP in this paper and the conventional prediction patterns, as shown in Fig. 12. Further, the NRMSE of the XY-Y prediction pattern (MRTPP) is 6.34%, and that of the XY-XY prediction pattern is 6.54%. Unlike Y−Y prediction pattern, the indices including NRMSE, NMAE and R^2 of MRTPP increased by 23.2%, 26.53% and 2.58%, respectively. Therefore, for forecasting photovoltaic power, the validity of the MRTPP is verified in contrast to the above prediction patterns.

4.2. Performance evaluation for ALSM and existing prediction models under MRTPP pattern

To further evaluate the performance of the proposed model, this section compares the proposed ALSM model with the commonly used models in photovoltaic power forecasting. Based on the

Table 4
1-h ahead forecasting using the common input-output prediction patterns and the proposed pattern with different relevant variables.

Prediction patterns	X−Y			XY-XY			XY-Y			Y−Y
Relevant variables	X1	X1X2	X1X2X3	X1	X1X2	X1X2X3	X1	X1X2	X1X2X3	–
NRMSE(%)	9.52	**9.02**	9.43	6.78	**6.54**	6.68	6.77	**6.34**	6.50	8.52
NMAE(%)	6.73	**6.33**	6.67	4.54	**4.43**	4.51	4.54	**4.20**	4.38	6.03
R^2(%)	94.01	**94.38**	94.20	97.23	**97.37**	97.31	97.24	**97.50**	97.41	94.92

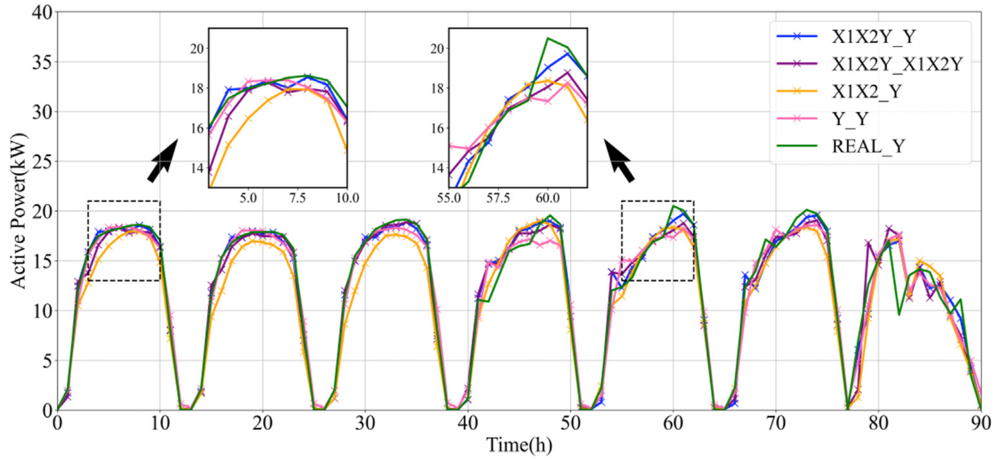


Fig. 12. Comparisons of the proposed MRTPP (XY-Y) in this paper and the conventional forecasting patterns using the optimal relevant variable strategy (X1X2).

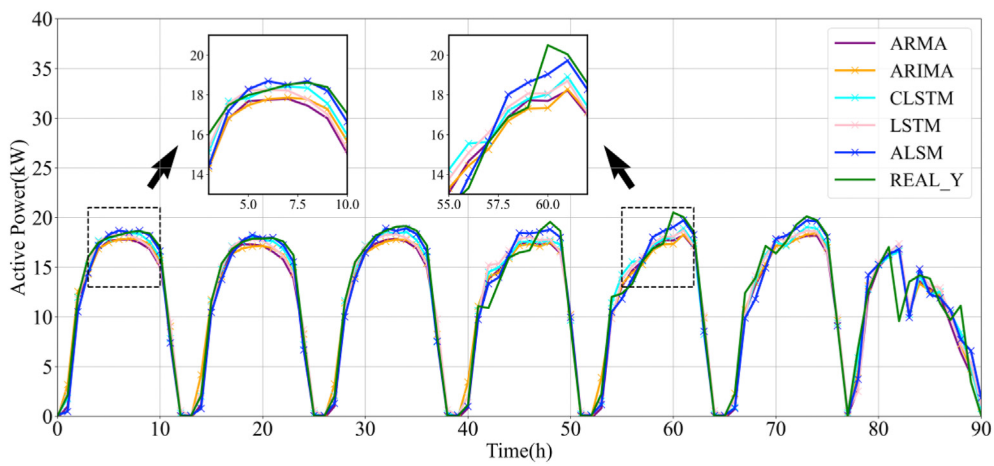


Fig. 13. Forecasting results for 1-h ahead using the proposed ALSM and the existed prediction models.

analysis results of relevant variables and comparison of prediction patterns in the previous section, the optimal relevant variables strategy, i.e., X1X2, is determined and applied to the neural network of LSTM, CNN-LSTM [48,49], and the ALSM. Nevertheless, the traditional time series statistical models of ARMA and ARIMA are only suitable for univariate input-output prediction [50,51], hence, the prediction pattern of Y−Y is applied to them. From the prediction results shown in Fig. 13 and Table 5, the ALSM scored the highest R^2 and the lowest NRMSE in the performance of each model on the issue of 1-h ahead photovoltaic power forecasting. Conclusively, compared with artificial intelligence models e.g. LSTM and CLSTM, statistical models e.g. ARMA and ARIMA have limited accuracy, the ALSM embedded with X1X2 relative variables strategy is the most precise algorithm when applying to PV power forecasting.

4.3. Performance evaluation of important parameters for ALSM and combined structure models

The previous section validates the performance advantages of the proposed ALSM model. In this section, we will further evaluate the various combination structures of ALSM, and the effect of the important parameter memory length, i.e., the input length of the data sequence before the target forecasting time. Moreover, as previously mentioned, the ALSM model mainly includes the short-

Table 5
1-h ahead forecasting using the proposed ALSM and the existing prediction models.

Methods	NRMSE(%)	NMAE(%)	R2(%)
ARMA	10.76	7.72	92.53
ARIMA	10.30	7.56	92.65
LSTM	8.28	5.74	95.43
CLSTM	7.26	4.93	96.72
ALSM	**6.34**	**4.20**	**97.50**

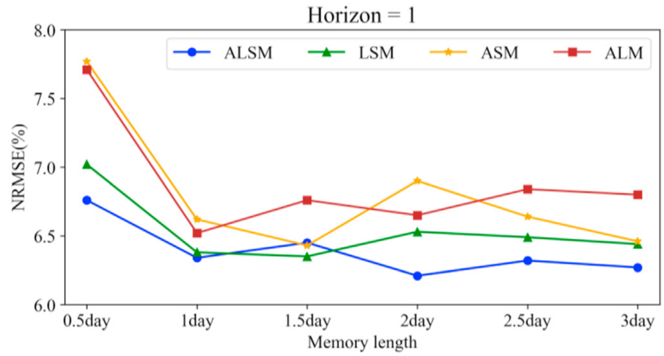


Fig. 14. The NRMSE indices of different models with different memory lengths for 1-h ahead forecasting.

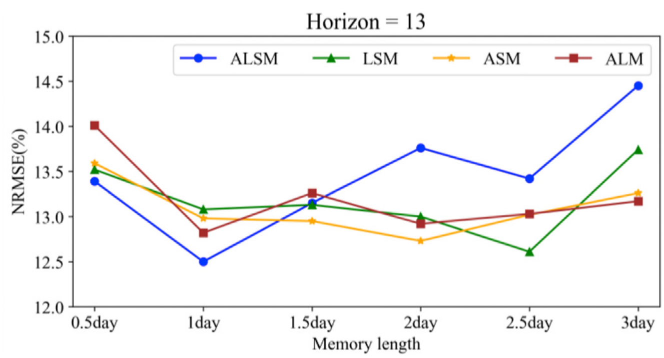


Fig. 15. The NRMSE indices of different models with different memory lengths for 13-h ahead forecasting.

term temporal module, long-term temporal module, and attention module. Hence, we perform some experiments according to various combinations of modules to provide a further indication for the accurate forecasting of our proposed model. The trial and error method is used to determine the corresponding optimal memory length parameters under different prediction horizon targets, and the optimal parameter of memory length is evaluated by the overall forecasting effect of the testing data. The LSM model composed of the long-term temporal module and the short-term temporal module is analyzed for the role of the attention module. And in order to analyze the effect of the long-term temporal module of skipping the hidden unit, we compare the ASM model which only contains the short-term temporal module and attention module. Also, the ALM model comprising the long-term temporal module and attention module is analyzed for the sense of the short-term temporal module.

Figs. 14 and 15 show the results of the proposed integrated model and combining models with different structures of ALSM and different memory lengths, which are represented by diverse colors and symbols. The parameters of the above models are shown in Table 6. *Horizon* = 1 represents the PV power forecasting range of 1-h ahead. Under this target, the lengths of memory length with the best prediction accuracy of models vary based on the structures they possess. The best memory length of the ALSM model is 2 days, and its prediction accuracy is better than other structural models. Specifically, from the accuracy of the LSM model, we can see that the attention module plays an essential role, while the ASM model and the ALM model have limited performance compared to the ALSM model. Similarly, *Horizon* = 13 represents the PV power forecasting range of 13-h ahead. Under this goal, the input data series with memory length of 1 day in the ALSM model is the best, with its accuracy being better than other structural models. The

Table 6
Forecasting accuracy of different models with different memory lengths for various prediction range.

Method	Memory length (day)	NRMSE(%)				NMAE(%)				R2(%)			
		Horizon(hour)				Horizon(hour)				Horizon(hour)			
		1	3	6	13	1	3	6	13	1	3	6	13
ALSM	0.5	6.76	**11.06**	**11.51**	13.39	4.52	**7.31**	**7.70**	9.17	97.26	**90.75**	**89.61**	85.96
	1	6.34	11.21	12.10	**12.50**	4.17	7.53	7.95	**8.45**	97.64	90.52	88.90	**87.47**
	1.5	6.45	11.82	12.68	13.15	4.37	8.95	8.22	9.09	97.46	89.38	87.85	86.82
	2	**6.21**	11.44	12.83	13.76	**3.95**	7.73	8.31	9.61	**97.92**	90.42	87.81	85.32
	2.5	6.32	11.09	12.89	13.42	4.15	7.32	8.49	9.18	97.76	90.73	87.78	85.81
	3	6.27	11.29	12.27	14.45	4.11	7.62	7.96	10.30	97.82	90.47	88.75	84.3
LSM	0.5	7.02	**11.20**	**12.33**	13.52	4.64	**7.45**	**7.98**	9.25	97.15	**90.55**	**88.68**	85.63
	1	6.38	11.27	12.50	13.08	4.20	7.57	8.16	8.97	97.53	90.49	88.37	86.88
	1.5	**6.35**	11.72	12.41	13.13	**4.18**	8.08	8.07	9.03	**97.57**	89.40	88.53	86.82
	2	6.53	11.62	12.73	13.00	4.43	7.82	8.24	8.84	97.38	90.19	87.83	86.98
	2.5	6.49	11.66	12.36	**12.61**	4.38	8.02	8.01	**8.57**	97.42	89.89	88.56	**87.36**
	3	6.44	11.29	13.14	13.74	4.31	7.64	9.55	9.43	97.48	90.47	87.04	85.51
ASM	0.5	7.77	12.86	12.93	13.59	5.35	9.44	8.55	9.27	95.83	87.07	87.76	85.67
	1	6.62	11.56	12.44	12.98	4.45	7.73	8.14	8.79	97.36	90.38	88.48	87.07
	1.5	**6.43**	11.69	12.26	12.95	**4.20**	8.07	7.96	8.74	**97.50**	89.79	88.80	87.10
	2	6.90	11.41	12.00	**12.73**	4.60	7.72	7.94	**8.62**	97.18	90.46	88.98	**87.34**
	2.5	6.64	11.24	11.93	13.02	4.51	7.54	7.93	8.88	97.34	90.50	89.15	86.97
	3	6.46	**11.13**	**11.57**	13.26	4.37	**7.40**	**7.74**	9.14	97.44	**90.62**	**89.55**	86.13
ALM	0.5	7.71	12.81	12.63	14.01	5.26	9.11	8.17	9.7	96.16	87.53	87.94	84.98
	1	**6.52**	11.63	12.47	**12.82**	**4.39**	7.93	8.15	**8.66**	**97.40**	90.13	88.4	**87.23**
	1.5	6.76	11.56	12.01	13.26	4.53	7.73	7.94	9.15	97.26	90.37	88.96	86.13
	2	6.65	11.14	**11.82**	12.92	4.51	7.45	**7.79**	8.72	97.33	90.59	**89.28**	87.15
	2.5	6.84	**11.11**	11.96	13.03	4.60	**7.37**	7.93	8.90	97.21	**90.63**	89.11	86.91
	3	6.80	11.58	11.89	13.17	4.59	7.76	7.83	9.11	97.22	90.29	89.16	86.79

prediction results by weather types are shown in Figs. 16 and 17, it is clear that the observations and the forecastings are in agreement. Overall, the reasonable determination of memory lengths for different forecasting targets gives full play to the role of the ALSM model and improves the prediction accuracy.

4.4. Performance evaluation of prediction range for ALSM and combined structure models

The performance of the ALSM model and the combined structure models under divergent forecasting range are further analyzed. Fig. 18 shows the forecasting accuracy results of the proposed integrated model and the models with the combined structure in the different prediction stages under the optimal parameters analyzed in the previous section. Specifically, for the same forecasting horizon, the best memory length is selected to evaluate the performance of combination methods of ALSM. It can be seen that the combined models are sensitive to prediction horizons from Table 6. For instance, LSM performs better at *horizon* = 1 and *horizon* = 13, while it is general once the horizon is 3 and 6. This may be due to the lack of attention mechanism, hence, less ability to capture the whole stage of the forecasting range. For most of the forecasting horizons, the performance of ASM is better compared to ALM, we can also understand that the information contained in the pure long-term temporal module is limited. In contrast with LSM, ASM, and ALM, the NRMSE of ALSM for 1-h ahead has been improved by 2.20%, 3.42%, and 4.75% respectively.

The statistical results reveal that all 4 models perform effectively and their accuracies are acceptable. With the increase of memory lengths, the prediction accuracies of models are not completely positively correlated. While, when the length of input data reaches a certain length, there is a potential decrease in accuracy. For the combined structure models which are not composed of both short-term temporal module and long-term temporal module, the memory length required by the model with a far prediction horizon is often longer than that of the proposed model. This is major because the hybrid model uses the long-term and short-term temporal modules of the dual-channel CNN (responsible for extracting the spatial characteristics of the data) and the LSTM (responsible for extracting the temporal characteristics of the data) to comprehensively obtain the feature information. For our ALSM model proposed in this paper, when the far horizon is aimed as a predicted target such as 13-h ahead forecasting, the long memory data will cause information redundancy and increase the prediction error. Therefore, for various prediction horizons, the memory lengths corresponding to the optimal performance of the model is different.

It can be seen that the ALSM model proposed in this paper is superior to other combined structure models in various stages of photovoltaic power forecasting. On the one hand, it integrates CNN with good spatial feature extraction ability and LSTM with good temporal feature extraction ability, and forms the combined

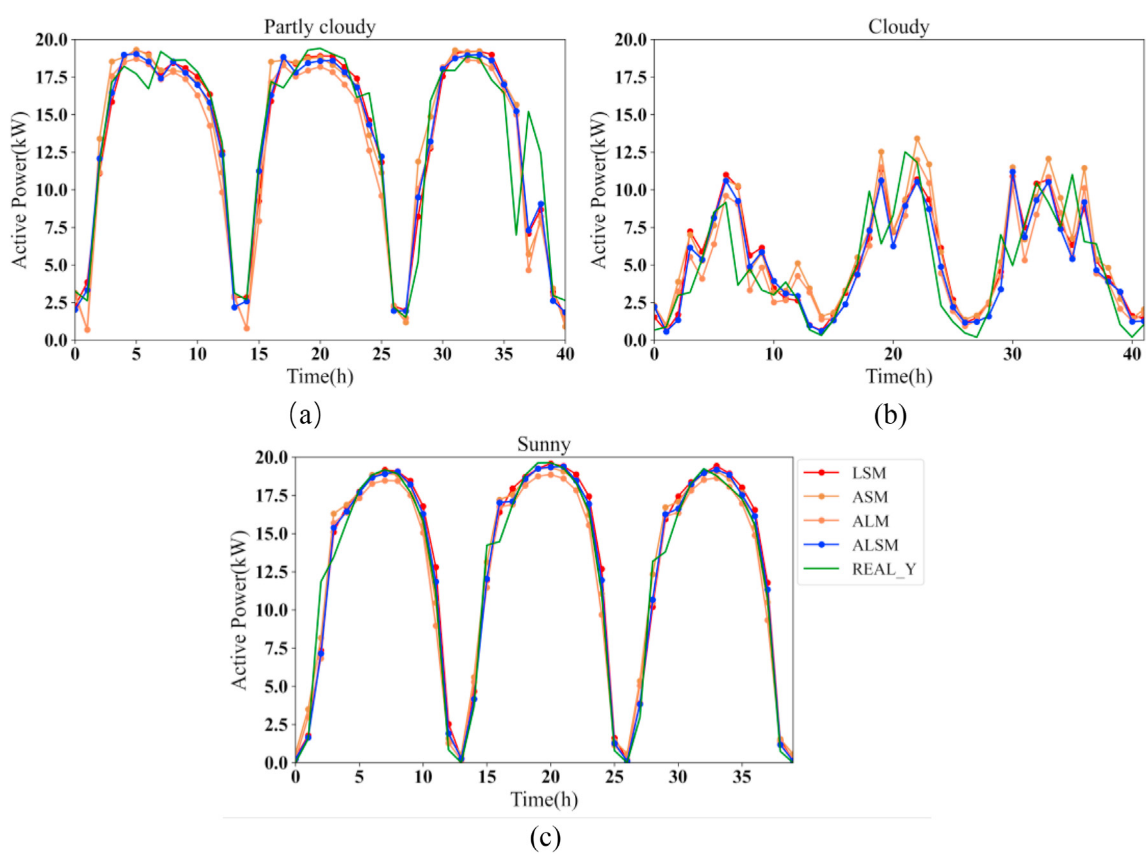


Fig. 16. 1-hour ahead forecasting results using the proposed ALSM and the combined structure models for (a) Partly cloudy day, (b) Cloudy day, and (c) Sunny day.

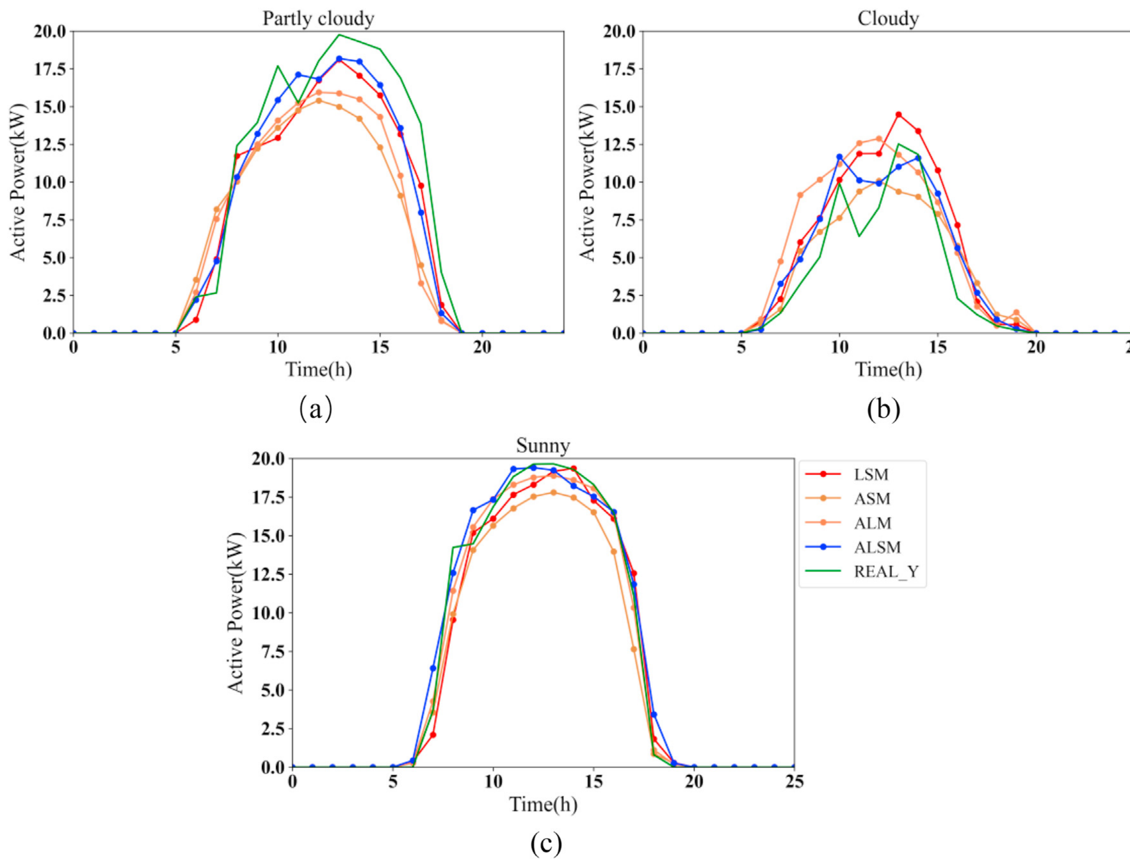


Fig. 17. 13-hour ahead forecasting results using the proposed ALSM and the combined structure models for (a) Partly cloudy day, (b) Cloudy day, and (c) Sunny day.

structure of long-term temporal module and short-term temporal module to extract the spatiotemporal correlation between multiple relevant variables and target variables in continuous time slots. On the other hand, the application of the attention mechanism enables the hybrid model to reasonably allocate attention and assign a high weight to the data with a high correlation with the prediction target, thereby making the forecasting results accurate. It seems that, for forecasting of photovoltaic power, reasonable allocation of weights for long-term and short-term temporal modes has greater advantages than a direct injection of data. Altogether, the value of the hybrid structure forecasting model is further confirmed.

The experimental results were completed in Python 3.7 and a personal computer with a 64-bit operating system, Intel (R) Core (TM) i5-8265U CPU@1.60GHZ 1.80GHZ and 8.00 GB of RAM. The running times of different structural models with different memory length parameters under various forecasting horizons are shown in Fig. 19. As can be seen, for the same forecasting horizon target, the training time of the four models increases with the growth of memory lengths. And the training time of LSM and ALSM models are both longer than that of ASM and ALM models with only a single CNN-LSTM module because they contain two. It is worth noting that when the predicting horizon target is rather far away, the running time required by LSM and ALSM models with complex structures increases obviously, while the training time of simple ASM and ALM models is relatively stable. This is also consistent with the fact that complex networks have stronger learning ability when the forecasting target is further. Compared with the shallow network structural models, the more parameters need training, the longer the training time is. Actually, when improving the hardware environments or optimizing code, the runtime will be much lower, which is acceptable in practical applications.

5. Conclusion

In conclusion, a hybrid forecasting model (ALSM) is proposed through a combination of the long-term and temporal module short-term temporal module based on CNN-LSTM, and attention mechanism in deep learning. Also, a multiple relevant and target variables prediction pattern (MRTPP) is proposed based on univariate and multivariate time series prediction patterns. The innovation of this paper lies in the combination of modern time series forecasting patterns and deep learning networks, which addresses the problem that the traditional time series method cannot fit the composite relationship of time-steps and multi-variables, as well as the conventional artificial intelligence modeling methods have the deficiency of overfitting. Furthermore, our proposed method, an attention-based CNN-LSTM neural network embedded with MRTPP, can simultaneously capture the short-term and long-term temporal changes of time series, and attain the day-ahead hourly forecasting of photovoltaic power.

Specifically, through the experimental analysis and findings, the proposed MRTPP prediction pattern in this paper has higher

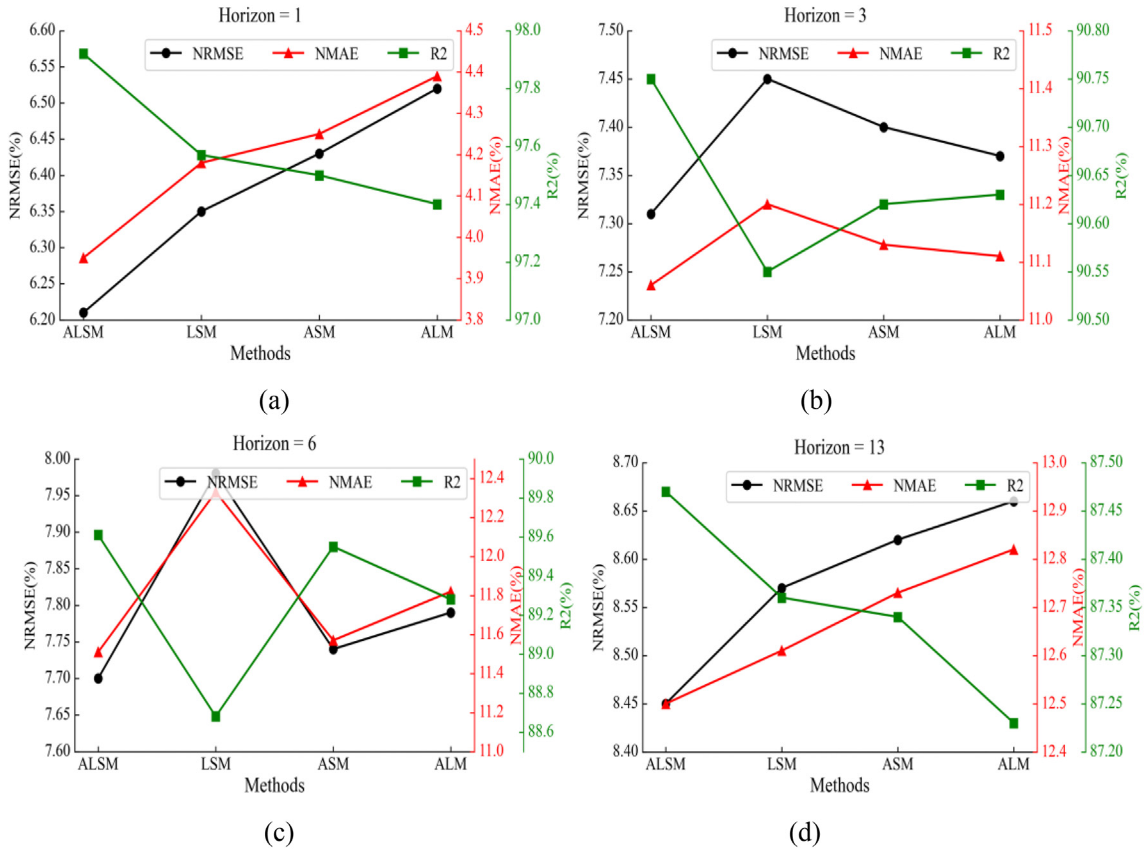


Fig. 18. The performance indices of different models with different forecasting range for (a) *horizon* = 1, (b) *horizon* = 3, (c) *horizon* = 6 and (d) *horizon* = 13.

accuracy compared to the currently used prediction pattern. Moreover, the optimal combination strategy of related variables in MRTPP prediction mode has been carried out, that is, the combination strategy of irradiance, ambient temperature, and target output power as input data. In contrast with statistical time series forecasting models (ARMA and ARIMA) and the extensively used neural network models (CNN and CLSTM), the MRTPP optimal combination strategy prediction pattern applied to the ALSM model exhibits higher accuracy. To analyze the role of the structures of the proposed ALSM model under the MRTPP prediction pattern, a large number of experiments are performed on time series data with different memory lengths (ranging from 0.5 days to 3 days) as input data at different forecasting horizons (ranging from 1 h to 13 h). As a consequence, we find that the application of long-term and short-term temporal pattern modules which are composed of dual-channel CNN-LSTM (convolution neural network and long short-term memory neural network) and attention mechanism provide the proposed model with better stability and accuracy. Besides, the corresponding memory lengths for various forecasting horizons are analyzed and suggested, inappropriate length of input can decrease the prediction performance.

In summary, based on convolutional neural network, long short-term memory neural network, and attention mechanism, the ALSM model is proposed. It harbors a higher forecasting accuracy compared to other models under the MRTPP pattern. Therefore, the efficacy of the proposed forecasting pattern and model for day-ahead hourly PV power forecasting is verified. The hybrid method analyzes characteristics of time series data and considers the influence factors of photovoltaic power to extract dependencies between time steps and multivariate time series. The suggestions on optimal memory lengths corresponding to the divergent prediction range have guiding significance for PV power forecasting in the application of practical engineering systems. However, as one of the time series forecasting methods, the input data referenced by the ALSM model is still the previous historical data, so there will be relatively obvious prediction errors during the short period changes of the point cloud. And the ALSM model is suitable for a certain range of forecasting horizons, that is, the prediction accuracy of day-ahead hourly is acceptable, the prediction accuracy of a much longer prediction range is limited. In the future, the method of combining the hybrid deep learning model ALSM based on MRTPP and NWP data can be considered to forecast the much longer horizon.

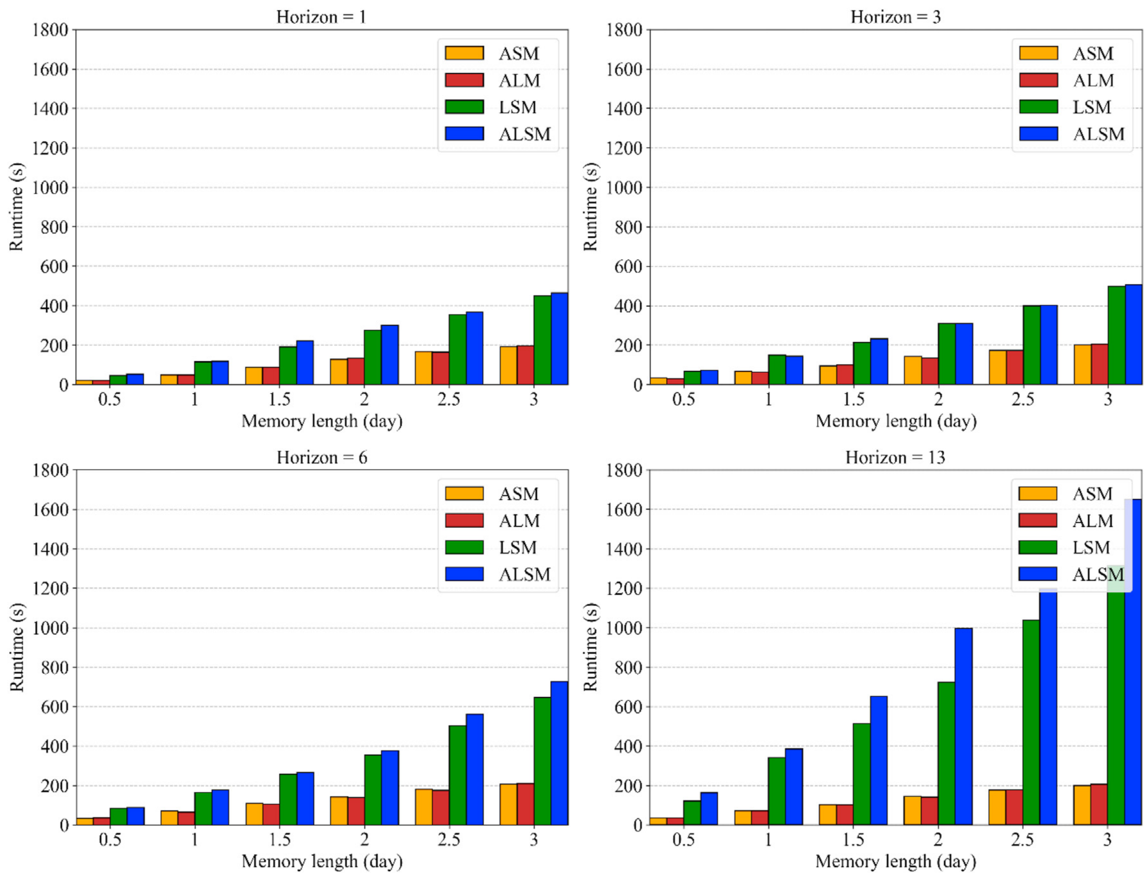


Fig. 19. The runtime of different structural models with different memory lengths parameters under various forecasting horizons.

Credit statement

Jiaqi Qu: Conceptualization, Methodology, Software, Writing – original draft Zheng Qian: Conceptualization, Supervision Yan Pei: Validation, Writing – review & editing.

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.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No. 61573046) and Program for Changjiang Scholars and Innovative Research Team in University (No. IRT1203).

References

[1] Inci M. Design and analysis of dual level boost converter based transformerless grid connected PV system for residential applications. In: 2019 4th international conference on power electronics and their applications (ICPEA); 2019.
[2] Elik Z, Tan A, Inci M, Teke A. Improvement of energy harvesting capability in grid-connected photovoltaic micro-inverters. Energy Sources Part A Recovery Utilization and Environmental Effects; 2020. p. 1–25.
[3] Hodge BM, Brancucci Martinez Anido C, Wang Q, Chartan E, Florita A, Kiviluoma J. The combined value of wind and solar power forecasting improvements and electricity storage. Appl Energy 2018;214:1–15.
[4] Dolara A, Leva S, Manzolini G. Comparison of different physical models for PV power output prediction. Sol Energy 2015;119:83–99.
[5] Wang F, Zhang Z, Liu C, Yu Y, Pang S, Duić N, et al. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. Energy Convers Manag 2019;181:443–62.
[6] Huang S, Wang D, Wu X, Tang A. DSANet: dual self-attention network for multivariate time series forecasting. The 28th ACM International Conference on Information and Knowledge Man-agement (CIKM 2019;19:2129–32.
[7] Harrison L, Penny WD, Friston K. Multivariate autoregressive modeling of fMRI time series. Neuroimage 2003;19:1477–91.
[8] Guermoui M, Melgani F, Danilo C. Multi-step ahead forecasting of daily global and direct solar radiation: a review and case study of Ghardaia region. J Clean Prod 2018;201:716–34.
[9] Kalate Ahani I, Salari M, Shadman A. Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas. Atmospheric Pollution Research 2019;10:689–700.
[10] Almonacid F, Pérez-Higueras PJ, Fernández EF, Hontoria L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. Energy Convers Manag 2014;85:389–98.
[11] Jiang H, Dong Y. Forecast of hourly global horizontal irradiance based on structured Kernel Support Vector Machine: a case study of Tibet area in China. Energy Convers Manag 2017;142:307–21.
[12] Sobri S, Sam KK, Rahim NA. Solar photovoltaic generation forecasting methods: a review. Energy Convers Manag 2018;156:459–97.
[13] Wang K, Qi X, Liu H. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. Appl Energy 2019;251:113315.
[14] Gu W, Ma T, Song A, Li M, Shen L. Mathematical modelling and performance evaluation of a hybrid photovoltaic-thermoelectric system. Energy Convers Manag 2019;198:111800.
[15] Wang F, Zhen Z, Liu C, Mi Z, Hodge BM, Shafie-khah M, et al. Image phase shift

invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. Energy Convers Manag 2018;157:123–35.
[16] Ahmed R, Sreeram V, Mishra Y, Arif MD. 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.
[17] Crisosto C, Hofmann M, Mubarak R, Seckmeyer G. One-hour prediction of the global solar irradiance from all-sky images using artificial neural networks. Energies 2018;11:2906.
[18] Chin VJ, Salam Z, Ishaque K. An accurate modelling of the two-diode model of PV module using a hybrid solution based on differential evolution. Energy Convers Manag 2016;124:42–50.
[19] Sharadga H, Hajimirza S, Balog RS. Time series forecasting of solar power generation for large-scale photovoltaic plants. Renew Energy 2020;150: 797–807.
[20] Ding M, Wang L, Bi R. An ANN-based approach for forecasting the power output of photovoltaic system. Procedia Environmental Sciences 2011;11: 1308–15.
[21] Cervone G, Clemente Harding L, Alessandrini S, Delle Monache L. Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble. Renew Energy 2017;108:274–86.
[22] Mellit A, Sağlam S, Kalogirou SA. Artificial neural network-based model for estimating the produced power of a photovoltaic module. Renew Energy 2013;60:71–8.
[23] Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. Energy Convers Manag 2019;198:111799.
[24] Yu W, Liu G, Zhu L, Yu W. Convolutional neural network with feature reconstruction for monitoring mismatched photovoltaic systems. Sol Energy 2020;212:169–77.
[25] Wang H, Yi H, Peng J, Wang G, Liu Y, Jiang H, et al. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. Energy Convers Manag 2017;153:409–22.
[26] Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 2018;148:461–8.
[27] Li Y, Zhu Z, Kong D, Han H, Zhao Y. EA-LSTM: evolutionary attention-based LSTM for time series prediction. Knowl Base Syst 2019;181:104785.
[28] Wang F, Xuan Z, Zhen Z, Li K, Wang T, Shi M. 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.
[29] Li P, Zhou K, Lu X, Yang S. A hybrid deep learning model for short-term PV power forecasting. Appl Energy 2020;259:114216.
[30] Hirata Y, Aihara K, Suzuki H. Predicting multivariate time series in real time with confidence intervals: applications to renewable energy. Eur Phys J Spec Top 2014;223:2451–60.
[31] Hirata Y, Aihara K. Improving time series prediction of solar irradiance after sunrise: comparison among three methods for time series prediction. Sol Energy 2017;149:294–301.
[32] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. Pattern Recogn 2018;77:354–77.
[33] Srivastava S, Lessmann S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol Energy 2018;162:232–47.
[34] Gao M, Li J, Hong F, Long D. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. Energy 2019;187:115838.
[35] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–80.
[36] Kim TY, Cho SB. Web traffic anomaly detection using C-LSTM neural networks. Expert Syst Appl 2018;106:66–76.
[37] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, et al. Attention is all you need. Neural Information Processing Systems 2017;30. NIPS 2017.
[38] Panigrahi S, Behera HS. A study on leading machine learning techniques for high order fuzzy time series forecasting. Eng Appl Artif Intell 2020;87:103245.
[39] Lai G, Chang WC, Yang Y, Liu H. Modeling long- and short-term temporal patterns with deep neural networks. The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval 2018:95–104.
[40] Lan H, Zhang C, Hong Y, He Y, Wen S. Day-ahead spatiotemporal solar irradiation forecasting using frequency-based hybrid principal component analysis and neural network. Appl Energy 2019;247:389–402.
[41] Chen T, Yin H, Chen H, Wu L, Wang H, Zhou X, et al. TADA: trend alignment with dual-attention multi-task recurrent neural networks for sales prediction. In: 2018 IEEE international conference on data mining (ICDM); 2018. p. 49–58.
[42] DKASC. http://dkasolarcentre.com.au/historical-data/download.
[43] Wang J, Zhong H, Lai X, Xia Q, Wang Y, Kang C. Exploring key weather factors from analytical modeling toward improved solar power forecasting. IEEE Transactions on Smart Grid 2019;10:1417–27.
[44] Jiang H. A novel approach for forecasting global horizontal irradiance based on sparse quadratic RBF neural network. Energy Convers Manag 2017;152: 266–80.
[45] Leloux J, Narvarte L, Trebosc D. Review of the performance of residential PV systems in Belgium. Renew Sustain Energy Rev 2012;16:178–84.
[46] Sanjari MJ, Gooi HB, Nair NC. Power generation forecast of hybrid PV–wind system. IEEE Transactions on Sustainable Energy 2020;11:703–12.
[47] Demirhan H, Kayhan Atilgan Y. New horizontal global solar radiation estimation models for Turkey based on robust coplot supported genetic programming technique. Energy Convers Manag 2015;106:1013–23.
[48] Kim TY, Cho SB. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 2019;182:72–81.
[49] Gensler A, Henze J, Sick B, Raabe N. Deep Learning for solar power forecasting — an approach using AutoEncoder and LSTM Neural Networks. In: 2016 IEEE international conference on systems, man, and cybernetics (SMC); 2016. 002858-002865.
[50] Sun H, Yan D, Zhao N, Zhou J. Empirical investigation on modeling solar radiation series with ARMA–GARCH models. Energy Convers Manag 2015;92: 385–95.
[51] Tunckaya Y, Koklukaya E. Comparative prediction analysis of 600 MWe coal-fired power plant production rate using statistical and neural-based models. J Energy Inst 2015;88:11–8.