



OPEN ACCESS

EDITED BY

Xingxing Zhang,
Dalarna University, Sweden

REVIEWED BY

Salim Heddami,
University of Skikda, Algeria
Vishnupriyan Jegadeesan,
Chennai Institute of Technology, India

*CORRESPONDENCE

Yongquan Zhou,
zhouyongquan@gxun.edu.cn

SPECIALTY SECTION

This article was submitted to Process and Energy Systems Engineering, a section of the journal Frontiers in Energy Research

RECEIVED 09 July 2022

ACCEPTED 11 August 2022

PUBLISHED 12 September 2022

CITATION

Liu H, Zhou Y, Luo Q, Huang H and Wei X (2022), Prediction of photovoltaic power output based on similar day analysis using RBF neural network with adaptive black widow optimization algorithm and K-means clustering. *Front. Energy Res.* 10:990018. doi: 10.3389/fenrg.2022.990018

COPYRIGHT

© 2022 Liu, Zhou, Luo, Huang and Wei. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Prediction of photovoltaic power output based on similar day analysis using RBF neural network with adaptive black widow optimization algorithm and K-means clustering

Hui Liu^{1,2}, Yongquan Zhou^{1,2,3*}, Qifang Luo^{1,2}, Huajuan Huang^{1,2} and Xiuxi Wei^{1,2}

¹College of Artificial Intelligence, Guangxi University for Nationalities, Nanning, China, ²Guangxi Key Laboratories of Hybrid Computation and IC Design Analysis, Nanning, China, ³Xiangsihu College of Guangxi University for Nationalities, Nanning, China

Solar photovoltaic power generation has become the focus of the world energy market. However, weak continuity and variability of solar power data severely increase grid operating pressure. Therefore, it is necessary to propose a new refined and targeted forecasting method to broaden the forecasting channels. In this paper, a hybrid model (KM-SDA-ABWO-RBF) based on radial basis function neural networks (RBFNNs), adaptive black widow optimization algorithm (ABWO), similar day analysis (SDA) and K-means clustering (KM) has been developed. The ABWO algorithm develops adaptive factors to optimize the parameters of RBFNNs and avoid getting trapped in local optima. SDA and K-means clustering determine the similarity days and the optimal similarity day through meteorological factors and historical datasets. Nine models compared forecast accuracy and stability over four seasons. Experiments show that compared with other well-known models on the four indicators, the proposed KM-SDA-ABWO-RBF model has the highest prediction accuracy and is more stable.

KEYWORDS

radial basis function neural networks, adaptive black widow optimization algorithm, similar day analysis, k-means clustering, photovoltaic power prediction, metaheuristic

1 Introduction

The traditional method of generating electricity is to use non-renewable fossil fuels such as coal and oil. However, due to rising electricity prices, the depletion of non-renewable energy sources and the consequent issues of climate change, solar photovoltaic (PV) is one of the most popular renewable energy sources in the utility grid. [Figure 1](#) (Renewable Energy Policy) shows the global capacity and annual increase in PV power system from 2011 to 2021. The solar PV capacity has increased from 70 to 942 Gigawatts

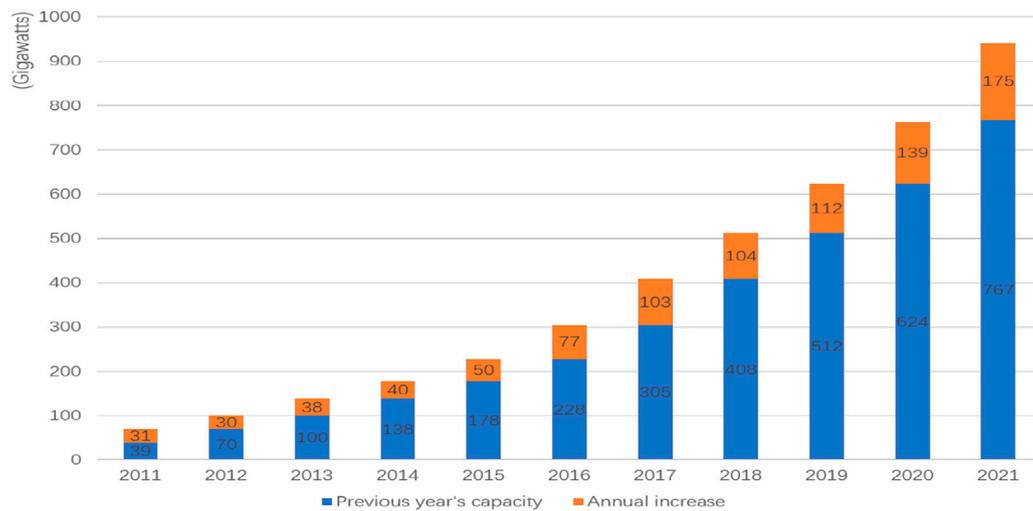


FIGURE 1
Solar PV global capacity and annual additions, 2011–2021.

in the last 10 years. However, the photovoltaic (PV) power generation industry and power system are faced with severe challenges (Mahmud et al., 2021). Due to the influence of the amount of solar radiation by weather factors (temperature, etc.), photovoltaic power generation has the defects of instability, uncertainty, difficulty in control, randomization and intermittency. During grid-connected operation, the stable operation and power quality of the power grid are disturbed. The stability and safety of photovoltaic power generation grid-connected operation not only has an important influence on social and economic development, but also is the focus of current social and economic concerns. Therefore, the method and system research of PV power generation prediction has important academic and application value. In the past, complex physical models and statistical models were used in the field of photovoltaic power generation forecasting. Reikard (Reikard, 2009) predicted solar irradiance using the Autoregressive Integrated Moving Average (ARIMA) methods. To take climatic factors into account, the ARMAX model is proposed, and the prediction accuracy is higher than that of the ARIMA model (Li et al., 2014). However, recent study (Sobri et al., 2018) has demonstrated that machine learning algorithms, such as neural networks (NNs) and support vector machines (SVMs), are superior to other forecasting methods for predicting photovoltaic power generation. A weighted SVM was proposed in Ref. (Xu et al., 2012). Chen et al. (2011) proposed an ANN-based weather classification method and utilized radial basis function neural networks (RBFNNs) for short-term photovoltaic power generation forecasting. Using appropriate algorithms to optimize the parameters of the model is one of the keys to improve the prediction effect (Du

et al., 2018). Several studies have found that the combined approach of parameter optimization algorithms combined with NN or SVM shows higher prediction accuracy (Dolara et al., 2015). In Ref. (Ghimire et al., 2018), adaptive differential evolution extreme learning machine model outperformed all 9 benchmark models. Hossain and Mahmood, (2020) proposed a novel clustering method to classify irradiance values by weather type, improving prediction accuracy by 33%. Dong et al. (2020) proposed a convolutional neural network model based on genetic algorithm and particle swarm algorithm to reduce the complexity of irradiance prediction. Zhang et al. (2015) proposed SDD Engine to construct a Euclidean distance metric for historical data to obtain the similarity days of predicted days. Wang et al. (2018) designed a RBFNN based on the multi-objective dragonfly algorithm and obtained excellent prediction results. Liu et al. (2015) used a BPNN to forecast output power by researching the influence of aerosol index data etc. Combination forecasting method, that is, combining the advantages of different forecasting methods to forecast is a hot research topic at present (Lin et al., 2018). Akhter et al. (2019) demonstrated that the model combining machine learning methods (ANN, SVM, ELM) and metaheuristics achieved the combined advantages of two or more techniques in PV power generation forecasting. The performance of the hybrid method is determined by the performance of each method (Xu et al., 2012). Therefore, several proven methods with better performance are synthesized, which can maximize the performance of the hybrid model.

Compared with other neural networks, RBFNNs have stronger robustness, learning and memory ability. It not only has good classification ability, but also is the optimal network to

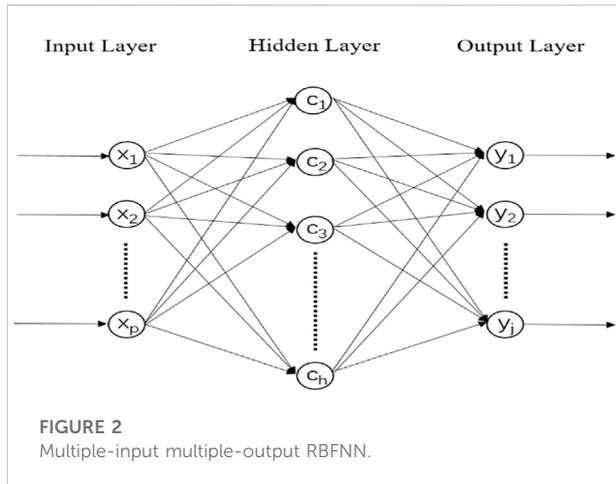
complete the mapping. However, the original RBFNNs are easy to get trapped in local optima and have poor global convergence ability (Han et al., 2017). Therefore, it is necessary to select an appropriate learning algorithm to improve the convergence performance of RBFNNs. The focus is on optimizing the parameters of RBFNNs (i.e. the center, width of the basis functions, and the weights from the hidden layer to the output layer), as well as determining the number of hidden layer neurons. Combined with common training algorithms, RBFNNs have been widely used in a large number of classification and regression problems. For example, Huang et al. (2004) designed a Growing and pruning RBF (GAP-RBF), which was improved and called GGAP-RBF (Huang et al., 2005). The goal of these two models is to make the network more compact to achieve less time consumption. Liu et al. (2005) tested on UCI classification datasets and seismic datasets using original RBFNNs and outperformed other models. However, what leaves these algorithms with little room for improvement is that they do not have a way to adjust parameters and network size simultaneously (Xie et al., 2012). Later, some studies introduced metaheuristic algorithms to simultaneously optimize network parameters and scale, which provided some unified and regular learning algorithms for researchers, broadening the space for improvement and optimization. Such as, Feng (Feng, 2006) proposed the SORBFNN based on particle swarm optimization (PSO) algorithm to solve three nonlinear problems such as approximation of nonlinear functions. Joe et al. (Qiao et al., 2019) combined an improved immune algorithm (IA) to optimize RBFNNs. Li et al. (Lee and Ko, 2009) developed the time-varying evolutionary PSO algorithm and compared it with other improved versions of the PSO algorithm. Duvvuri et al. (Duvvuri and Anmala, 2019) used the genetic algorithm to learn RBFNNs to predict fecal coliforms.

It is statistically known that the PSO algorithm and other improved PSO algorithms are used more for training RBFNNs than other metaheuristic algorithms. However, the PSO algorithm converges faster, but produce a low performance in exploitation search ability, and have low convergence precision (Parsopoulos and Vrahatis, 2002). To overcome this limitation, other metaheuristic algorithms that balance exploration and exploitation search capabilities can be considered, such as cuckoo search (CS) (Yang and Deb, 2009), symbiotic organisms search (SOS) (Cheng and Prayogo, 2014), gravitational search algorithm (GSA) (Rashedi et al., 2009), sine cosine algorithm (SCA) (Mirjalili, 2016), grey wolf optimizer (GWO) (Mirjalili et al., 2014), differential evolution (DE) (Storn and Price, 1997), etc. Although many metaheuristic algorithms have been proposed for many years, however, due to the “No Free Lunch (NFL) Theorem” (Wolpert and Macready, 1997), there is no one-size-fits-all algorithm that can reach the upper limit of the optimization capability, which is sufficient to solve all optimization problems. In terms of training neural network parameters, the black widow

optimization algorithm proposed in the past two years showed excellent talent due to its powerful search and extraction capabilities. Since BWO was proposed, several papers have verified its superiority in optimizing neural networks, such as adaptive network-based fuzzy inference system (ANFIS). For example, Katooli et al. (Katooli and Koochaki, 2020) applied BWO in the training process of ANFIS to improve its accuracy and robustness, and combined the association rule mining technique (ARMT) to select the most necessary input data. Tigtiz et al. (2020) demonstrated that BWO trains faster and has higher classification accuracy than gradient-based learning algorithms. Memar et al. (2021) used BWO to learn ANFIS and SVR and forecast maximum wave height. Panahi et al. (2021) proposed the ANFIS-BWO and SVM-BWO models for hydraulic engineering prediction. It is known from previous research that both BWO and RBFNN have excellent performance. Therefore, the combined method of BWO and RBFNN combines the advantages of both methods, and it is expected to have higher prediction performance than other methods.

In addition, it is necessary to ensure not only the efficiency of the predictive model, but also the accuracy of the input data. This requires the prediction model not only to consider the relationship between power generation and meteorological factors, but also to take historical power generation data as one of the influencing factors (Lin et al., 2018). Similar day analysis (SDA) is a mathematical model used to filter historical datasets, selecting similar days for training from a large amount of historical data (Zhou et al., 2020). Moreover, many literatures have also demonstrated that clustering algorithms can improve the prediction efficiency (Ghayekhloo et al., 2015; Lin et al., 2018; Hossain and Mahmood, 2020). Combining the advantages and disadvantages of previous research works, this paper proposes a combined prediction model based on K -means, SDA, ABWO and RBFNN, namely KM-SDA-ABWO-RBF, for real-life photovoltaic power generation prediction. The innovations and contributions of this paper are summarized as follows:

- 1) The ABWO algorithm is designed. The performance of the original BWO algorithm on optimizing RBFNNs needs to be improved. Numerous studies (Mirjalili et al., 2012; Gonzalez et al., 2015; Justus and Anuradha, 2022) have found that PSO algorithm, gravitational search algorithm (GSA), golden eagle optimizer (GEO) have better results than other algorithms in training neural network parameters. By analyzing the parameter design methods in these algorithms, it is known that they all have adaptive inertia weights that vary with the number of iterations. Therefore, the BWO algorithm is adaptively designed, and the parameters of the crossover stage are changed from random values to adaptively reduce with the increase of the number of iterations. Compared with many algorithms such as PSO, the accuracy of the adaptive black widow optimization algorithm is guaranteed.



- 2) The goal of the metaheuristic algorithm is to find the optimal value of a function. Moreover, the goal of RBFNNs in the PV power generation problem is to minimize the prediction error, which boils down to the problem of solving the function minimum problem. Therefore, the metaheuristic algorithm is used to optimize RBFNNs to solve the PV power prediction problem.
- 3) In photovoltaic power generation forecasting, in order to avoid excessively long forecasting time caused by a large number of irrelevant abnormal historical data, and to synthesize the advantages of different forecasting methods, this paper adopts a combined forecasting method, namely the KM-SDA-ABWO-RBF model. The model uses the Pearson correlation coefficient to analyze the correlation between historical data weather factors and photovoltaic power generation, and then calculates the weighted Euclidean metric. Some days with smaller weighted Euclidean metric results are used as the similarity days useful for forecasting days. The influencing factors of power generation are not only temperature, humidity, irradiance, etc. Under different weather types (sunny, cloudy, rainy), the output power is different. Therefore, this paper conducts in-depth cluster analysis on the input data, extracts refined modalities, and performs classification prediction.
- 4) A limitation of most studies is that they do not take into account the stability of the predicted results. The model proposed in this paper not only has a small prediction error, but also has a relatively stable prediction result.

The remainder of this article is summarized as follows: The specific scheme of the KM-SDA-ABWO-RBF model is proposed in Section 2. The specific implementation process of PV power prediction is introduced and discussed in Section 3. Section 4 presents and discusses the experimental results. Finally, conclusions and future work in Section 5.

2 KM-SDA-adaptive black widow optimization-RBF model

2.1 RBF neural network

RBFNN is a special type of fully connected feedforward network (Er et al., 2002), (Chen et al., 1991), consisting of only three layers: input layer, output layer and hidden layer, each of which has an activation function. RBFNNs have achieved higher convergence accuracy, with more convergence speed for training, less convergence time (Mandal et al., 2012). The network topology of a multiple-input multiple-output RBFNN is drawn in Figure 2 below.

The hidden layer can directly map the input vector to the high dimension space without the need for weight connection. When the center point of the RBFNN is determined, this mapping relationship is also determined. The number of hidden layer neurons determines the network size. The activation function of the hidden layers neuron uses the Gaussian function, which is a nonlinear transformation function. Each hidden layer neuron of the RBFNN obtains the output of the hidden layer by calculating the activation function, and the final output of the output layer is the linear weighted sum of the outputs of the hidden layer. The output equation is described as follows:

$$y_j = \sum_{i=1}^h w_{i,j} e^{-\|x - c_i\|^2 / 2\sigma_i^2} \quad (1)$$

where $i = 1, 2, \dots, h$, h is the number of hidden layer; $j = 1, 2, \dots, n$, n is the number of output neurons; $w_{i,j}$ is the output weight between the i th hidden neuron and the j th output neuron. The input vector is expressed as $x = [x_1, x_2, \dots, x_p]^T$, and $x \in \mathcal{R}^{p \times 1}$; p is the number of input neurons; c_i is the center vector of the i th hidden neuron; $\|x - c_i\|$ is the Euclidean distance between the input vector and the center vector; σ_i is the width of the i th hidden neuron. The proposed ABWO-based method simultaneously searches all parameters of RBFNNs.

2.2 BWO and adaptive black widow optimization algorithm

Metaheuristic algorithms can be divided into three broad categories: evolutionary algorithms, physicochemical-based algorithms, and swarm intelligence algorithms (Wang et al., 2020). Swarm intelligence algorithms are based on the natural biological proliferation of social connections and the behavior of natural animals and other people in their lives (Abualigah and Diabat, 2021). As a swarm intelligence algorithm, the BWO algorithm (Hayyolalam and Kazem, 2020) is based on the mating and reproduction behavior of black widow spiders. Each individual spider corresponds to a candidate solution to

the optimization problem, and survivability represents a fitness function. Getting the strongest spider is getting the optimal solution to the problem. The mating and reproduction behaviors of black widow spiders include crossover, cannibalism and mutation.

1) *Crossover*: The crossover phase is the exploration search phase. Firstly, combined with the parameter optimization problem of RBFNNs, the solution of root mean square error (RMSE) is used as the fitness function. Secondly, the populations are sorted according to their fitness. According to the fertility rate, spiders with high fitness values in the population were selected to participate in mating, and a pair of parents (male and female spiders) were randomly selected from them for mating and reproduction each time. Each black widow spider displays the value of the problem variable. The structure is treated as a one-dimensional array, where n is the dimension of the feature:

$$\text{spider} = [s_1, s_2, \dots, s_n] \quad (2)$$

Each pair of parents simulates the reproductive process with the help of an α array:

$$\begin{cases} y_1 = \alpha * s_1 + (1 - \alpha) * s_2 \\ y_2 = \alpha * s_2 + (1 - \alpha) * s_1 \end{cases} \quad (3)$$

where α_i representative is a random value in [0,1], s_1 and s_2 representative is parent, y_1 and y_2 representative is offspring, and each pair of parents mating is repeated $n/2$ times.

2) *Cannibalism*: The behavior of cannibalism refers to strong spiders eating weak spiders, that is, keep spiders with high fitness values and eliminating spiders with low fitness values. It includes sexual cannibalism and sibling cannibalism. Sexual cannibalism is when female spiders eat male spiders due to their smaller size, gaining nutrients that give offspring a higher chance of surviving. After mating, this behavior is achieved by destroying the mated parent. Corresponding to the optimization problem, this means that those with high fitness values are females, those with low fitness values are males, and females with higher fitness values are retained. Sibling cannibalism is when the offspring hatch and the stronger offspring eats the weaker siblings, thereby increasing the survival rate of the survivors. It achieves its goal by destroying a portion of offspring with weak fitness values according to the cannibalism rate. Each time through the crossover and cannibalistic stages, only one parent and one offspring are retained for each pair of parent spiders that mate.

3) *Mutation*: The mutation phase is the exploitation search phase. At this stage, the algorithm selects multiple spiders with higher fitness values according to the mutation rate, changes the attribute information of these spiders, and

obtains new candidate solutions, so as to increase the diversity of the population. The mutation method randomly swaps the two eigenvalues in the array for each spider. e.g., $[s_1, \dots, s_j, \dots, s_j, \dots, s_n]$ mutates to $[s_1, \dots, s_j, \dots, s_j, \dots, s_n]$.

The advantage of the BWO algorithm is that the equation is simple, easy to improve, and the principle is obvious, easy to understand, and even has a fast convergence speed. In order to increase the diversity and difference of the population and avoid getting trapped in local optima, only the equation of the crossover stage is improved, and the other stages remain unchanged. The parameter optimization of the improved algorithm on RBFNNs will be more accurate. Since the improved algorithm is adaptive, it is called adaptive black widow optimization (ABWO) algorithm.

The specific improvements are as follows: Eq. 3 shows that in the crossover stage, the α array is used to represent the process of generating offspring from the parent generation. In Eq. 4, the β array is introduced: $\beta = [\beta_1, \beta_2, \dots, \beta_n]$. Each component of this array is not randomly selected in the range [0, 1] like the α array. It is randomly given in the range [0, 1] at the beginning, and then it adaptively decreases as it increases. The detailed description of the β array is as follows:

$$\beta_i = \frac{5 \times (\text{Maxiter} - it + 1)}{\text{Maxiter}} \quad (4)$$

Eq. 3 is changed to

$$\begin{cases} y_1 = \alpha * s_1 + (1 - \alpha) * s_2 \\ y_2 = \beta * s_2 + (1 - \beta) * s_1 \end{cases} \quad (5)$$

where s_1 and s_2 are parents, y_1 and y_2 are descendants, Maxiter is the maximum number of iterations, and it is the current number of iterations.

2.3 Similar day analysis

In order to avoid training numerous irrelevant abnormal data, which leads to a long prediction time, this paper introduces the similar day analysis to obtain historical data that is instructive for the prediction results. The weather characteristics of certain days in the historical data are very similar to the weather characteristics of the forecasting days, and these days are called the similarity days (Zhou et al., 2020). The similarity days are obtained from the day the dataset starts to the day before the prediction day. There are many machine learning methods to obtain the similarity days: Euclidean metric (Mandal et al., 2007), PCC and cosine similarity, etc. When comparing the similarity of a certain meteorological factor between two days, more attention is paid to the numerical difference of weather characteristics between two days, rather than the similarity of change trends and the difference of numerical direction.

Therefore, a weighted Euclidean metric based on the influence of multiple meteorological factors was chosen instead of PCC and cosine similarity.

When predicting photovoltaic power generation, only the influence of meteorological factors on the output power needs to be considered. Different meteorological factors will lead to different output photovoltaic power. In order to obtain the influence weight of each meteorological factor on the output power, this paper uses the Pearson correlation coefficient (PCC) to conduct a correlation analysis on the relationship between weather factors and photovoltaic power generation. PCC (Bugala et al., 2018) is used to measure the degree of correlation between two variables, and its absolute value ranges from [0,1]. The closer the calculated PCC value is to 1 or -1, the stronger the correlation between the two variables. Conversely, the closer the value of PCC is to 0, the weaker the correlation. This study randomly selects a day from the dataset and calculates the PCC between the output power and each meteorological factor, as the weight coefficient of the weighted Euclidean metric. The greater the influence of the meteorological factor on the output power, the higher the weight coefficient of the meteorological factor. PCC is calculated as follows:

$$P(k) = \frac{\sum_{i=1}^n (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^n (u_i - \bar{u})^2} \sqrt{\sum_{i=1}^n (v_i - \bar{v})^2}} \quad (6)$$

where n represents the number of power data of the day; k represents the number of weather factors of the day; u_i represents the k th weather factor at the i th hour of the day; v_i represents the photovoltaic power generation at the i th hour of the day. $P(k)$ is the PCC between the k th weather factor and the output power, that is, the degree of influence of the weather factor on the output power.

The Euclidean distance for each meteorological factor between two days is described as follows:

$$ED(k) = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (7)$$

where X_i and Y_i represent the k th weather factor at the i th hour of the forecasting day and the history day, respectively; n represents the number of power data per day; k represents the number of weather factors per day. The smaller the value of $ED(k)$, the higher the similarity of the k th weather factor between two days.

The similarity between a historical day and a forecasting day is defined as the sum of the weighted Euclidean metric, described as follows:

$$S = \sum_{k=1}^K P(k)ED(k) \quad (8)$$

where $P(k)$ represents the weight coefficient of the k th weather factor; $ED(k)$ represents the Euclidean distance of the k th weather factor between two days. K represents the number of weather factors per day. The smaller the value of S , the higher the similarity between the historical day and the predicted day. The historical day corresponding to the minimum value of S is considered as the best similar day.

2.4 K-means clustering

In addition to meteorological factors, the type of weather can also affect the output power value. However, there is no information on weather types such as sunny, cloudy, rainy, etc. In the dataset used in this paper. In order to determine the weather type for each day of each season, the K-means algorithm is used to classify the input meteorological factors. Meteorological factors such as temperature and humidity in the same season are not very different. On the contrary, the irradiance of each day varies greatly. Therefore, the proposed clustering algorithm classifies the irradiance within the forecast horizon (similar days and the forecasting day). After the cluster center is obtained, the original irradiance value is replaced with the cluster center of the class to which the irradiance belongs. Different irradiance values represent different weather types. The flowchart of the clustering process is shown in Figure 3.

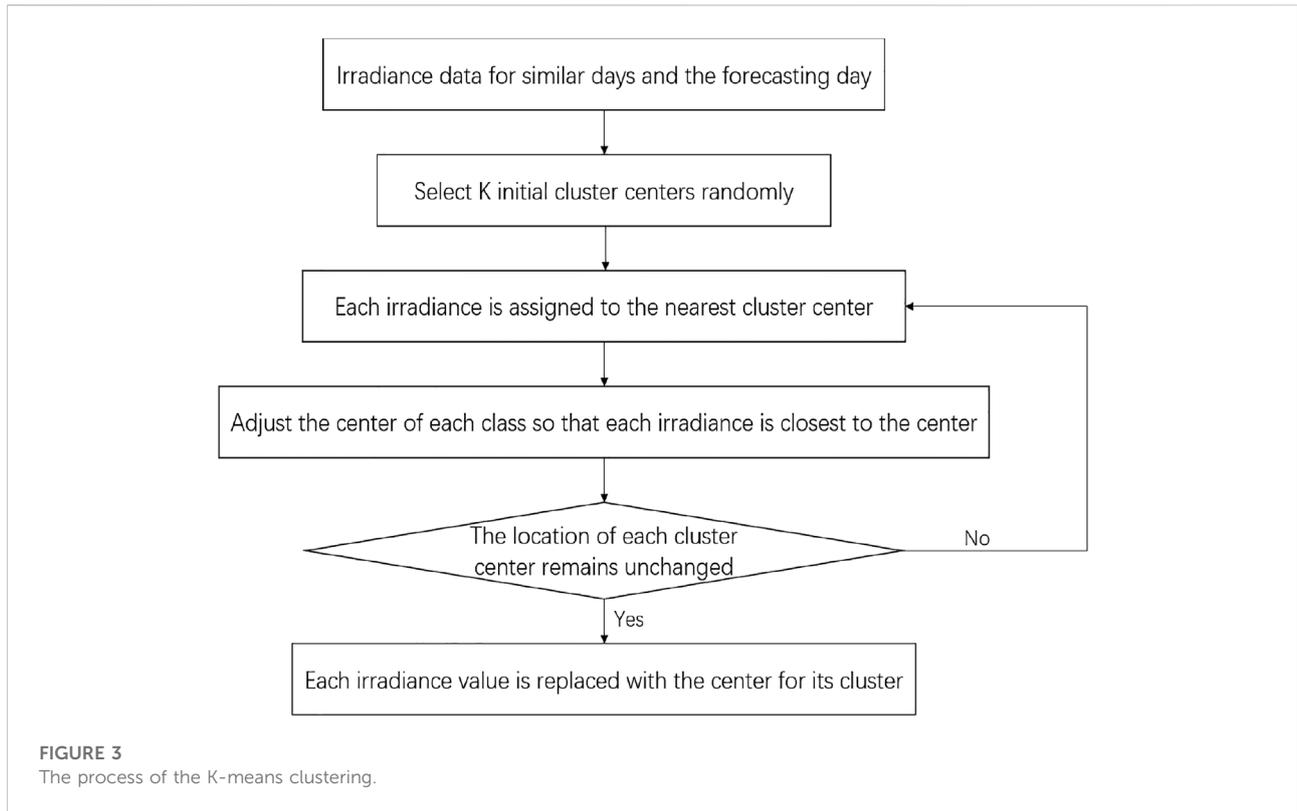
2.5 adaptive black widow optimization-RBF neural network

ABWO-RBF neural network can adjust all parameters during the learning process. Taking the structure of multiple-input and multiple-output as an example, in the widow initialization stage, the i th widow is described as

$$widow_i = [c_{i,1}^T, \sigma_{i,1}, w_{i,1}^T, c_{i,2}^T, \sigma_{i,2}, w_{i,2}^T, \dots, c_{i,h}^T, \sigma_{i,h}, w_{i,h}^T] \quad (9)$$

where $c_{i,h}$, $\sigma_{i,h}$, and $w_{i,h}$ are the center vector, width, and output vector of the h th hidden neuron in the i th widow, respectively; h is the network size; D is the dimension of the widow satisfying $D = p + 1 + j$, with p being the number of input variable and j being the number of output variable. The fitness value of each widow represents the accuracy of the ABWO-RBF neural network. In order to optimize the parameters of RBFNNs, the fitness value of each widow is the RMSE. Taking the structure of multi-input and multi-output as an example, the specific definitions are as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - f(x_j))^2} \quad (10)$$



where n is the number of output neuron, y_j and $f(x_j)$ are the real output of the network and the desired output, respectively.

2.6 KM-SDA-adaptive black widow optimization-RBF model

The system architecture of the KM-SDA-ABWO-RBF model for PV power generation forecast is shown in Figure 4. The procedure is elaborated as follows:

Step 1. Obtain historical datasets, including attribute information such as several meteorological factors and photovoltaic power generation.

Step 2. The data is preprocessed for outliers and normalized.

Step 3. According to the equation of weighted Euclidean metric, several days with more similar meteorological characteristics to the predicted day are selected from the historical data as similar days, that is, days with smaller weighted Euclidean metric value. The day with the smallest weighted Euclidean metric is the optimal similarity day.

Step 4. The K-means algorithm classifies the daily average global horizontal radiation and the daily average diffuse horizontal radiation respectively, and replaces the original radiation value with the corresponding cluster center.

Step 5. Initialize the center, width and weights of the RBFNN using the ABWO algorithm.

Step 6. The data of similar days are input into RBFNN as training samples, including daily average weather temperature, daily average weather relative humidity, daily average global horizontal radiation, daily average diffuse horizontal radiation and active power; ABWO algorithm updates the parameters of RBFNN to obtain the optimal parameters of the KM-SDA-ABWO-RBF model.

Step 7. Input the data of the optimal similarity day and the forecasting day as the test sample into RBFNN to obtain the photovoltaic power generation on the forecasting day.

3 Specific implementation process of PV power prediction

All the schemes were programmed in MATLAB version 2019 and run on a normal personal computer with an Intel(R) Core(TM) i7-9700 CPU and 16.0 GB RAM, under a Microsoft Windows 10.0 environment.

3.1 Data description

Effective forecasting needs to meet two requirements, one is the validity of forecasting methods and software, and the other is the reliability of historical meteorological data. The dataset for

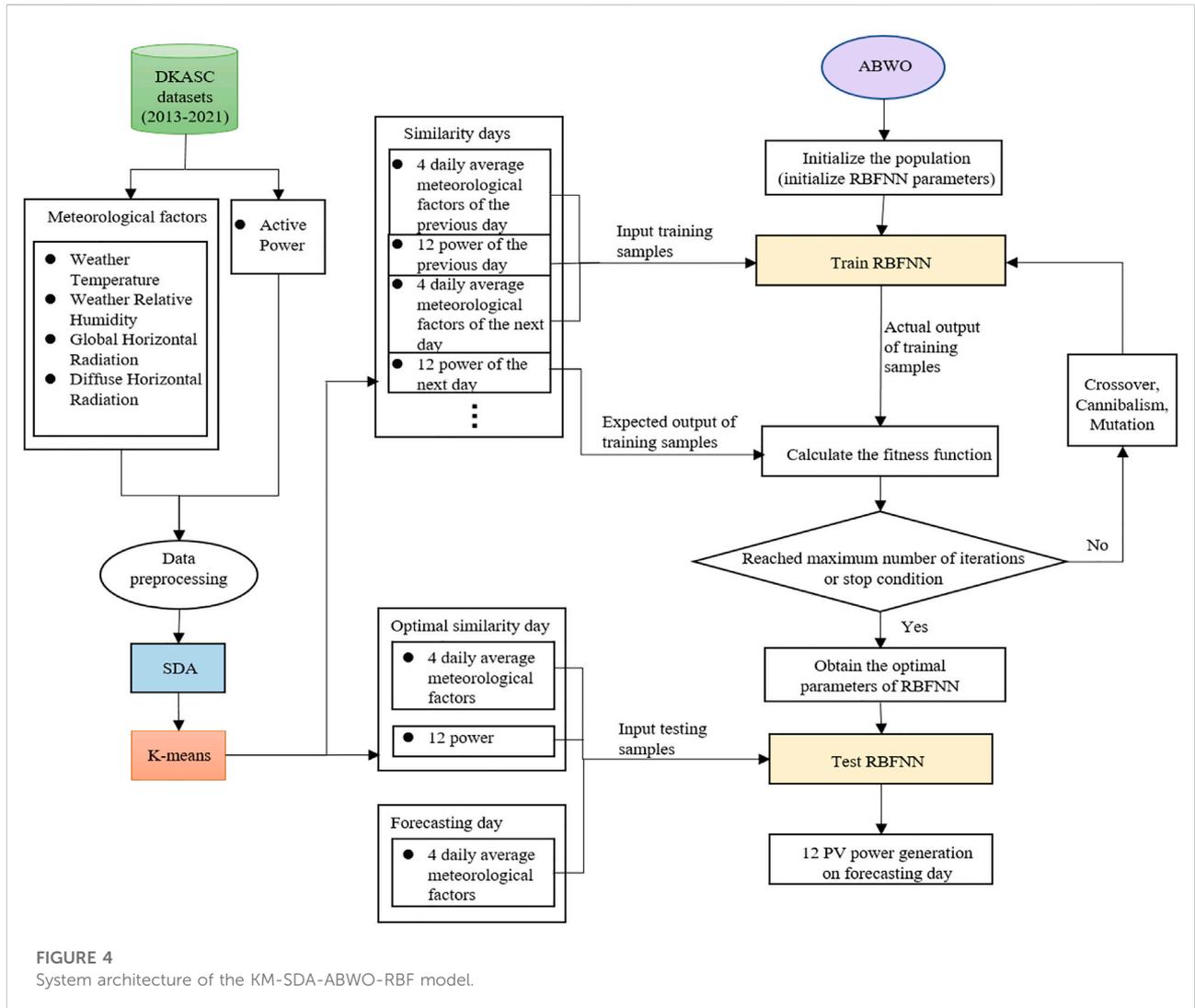


FIGURE 4 System architecture of the KM-SDA-ABWO-RBF model.

the experiment is from the Desert Knowledge Australia Solar Centre (DKASC), Alice Springs, (<http://dkasolarcentre.com.au/download?location=alice-springs>). The photovoltaic power plants use 60 monocrystalline panels and were installed in 2009 with an array rating of 10.5 kW. The dataset is from August 14, 2013 to November 28, 2021 with a time interval of 5 min. A large number of original datasets are inconvenient to organize, so this section extracted data from 7:30 to 18:30 with an interval of 1 h. The properties of the dataset include weather temperature, weather relative humidity, global horizontal radiation, diffuse horizontal radiation, and active power. The four seasons in Australia: spring days are from September to November; summer days are from December to February; autumn days are from March to May; winter days are from June to August. According to the time of the four seasons, the datasets from 2013 to 2021 are divided into four categories by season. Four days were arbitrarily selected from the four seasons

in the dataset, that is, November 19, 2021, February 22, 2020, May 13, 2021 and June 29, 2021 as the forecasting day, respectively. The similar days for four forecasting days were selected from the day from the beginning of the four types of datasets to the day before the forecasting day, respectively.

Data mining methods are very common in classification and regression problems (Hossain and Mahmood, 2020). The data mining methods used in this section are dataset selection, outlier processing, data normalization and correlation analysis. There may be errors in the process of data collection and transmission, such as metering failure, system disconnection for cabling works, system outage for new array connection and so on. In a word, there are incomplete and abnormal data in a large number of original datasets, so the data needs to be preprocessed. So when the data at a certain sample is empty, the empty value is replaced with the data at the previous sample. When data missing is severe, consider removing missing data for a whole day. Data

TABLE 1 PCC of meteorological factors and active power.

Meteorological factors	PCC
Weather temperature	0.22523
Weather relative humidity	-0.13469
Global horizontal radiation	0.91518
Diffuse horizontal radiation	0.87048

normalization refers to scaling each meteorological attribute and output power respectively and mapping them to the interval [-1, 1]. Through outlier processing and normalization, the quality of the data is improved, and can be better adapted to algorithms and neural network models.

In order to study the correlation between the four weather factors and active power, an arbitrary day was selected from the dataset to calculate the PCC of weather factors and active power.

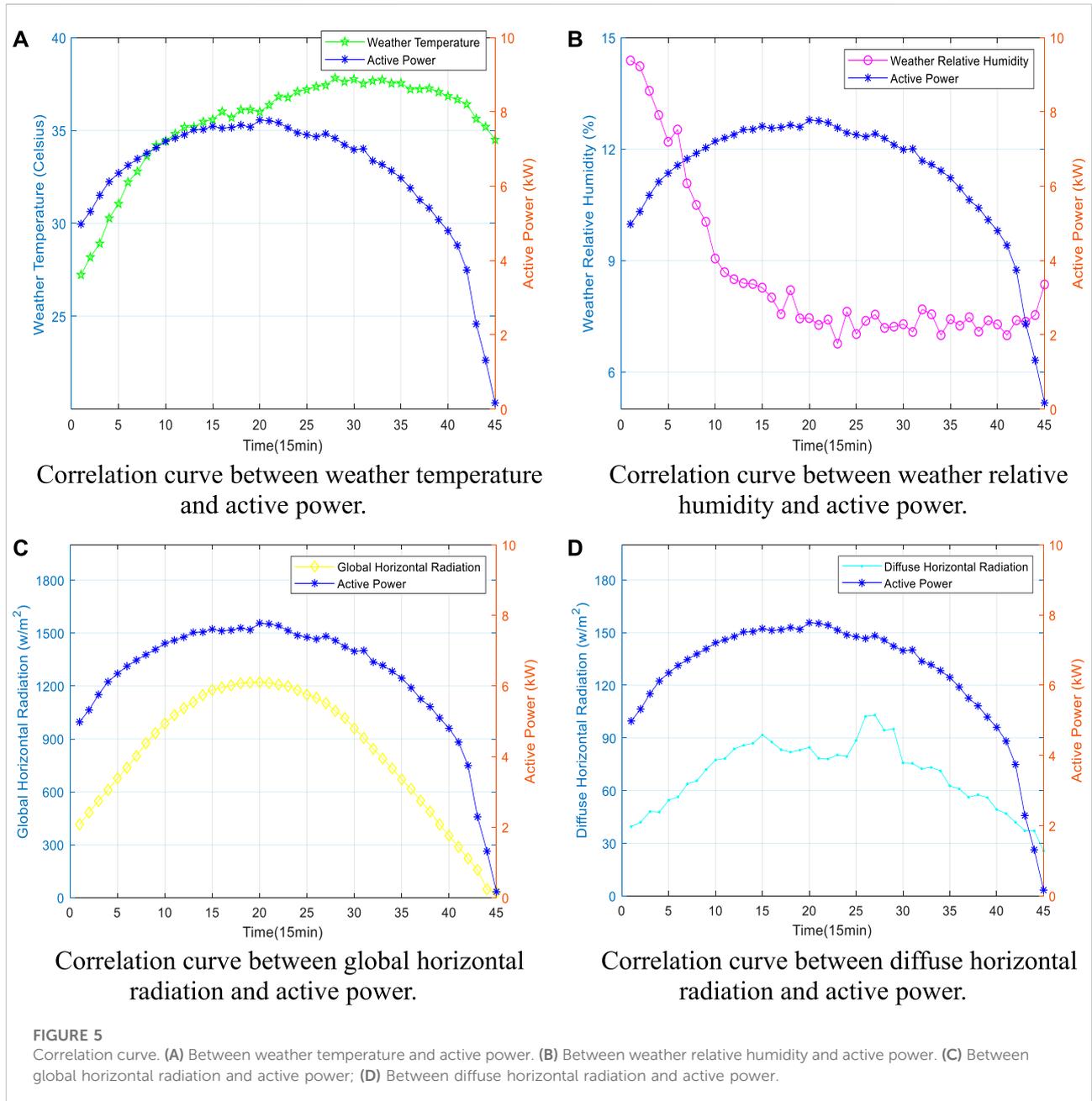


TABLE 2 The optimal similarity day.

Season	Spring	summer	autumn	winter
Optimal similarity day	November 18, 2021	February 19, 2020	May 13, 2019	June 30, 2020

The specific calculation results are shown in Table 1. Moreover, Figures 5A–D shows the correlation curves of weather temperature, weather relative humidity, global horizontal radiation, diffuse horizontal radiation and photovoltaic power, respectively. The time point interval for this day is 15 min. From the correlation analysis, it is known that among the multivariate meteorological factors affecting photovoltaic power generation, the correlation coefficient of global horizontal radiation is the highest, while the correlation coefficient of weather temperature is the lowest. In descending order of influence on power generation are global horizontal radiation, diffuse horizontal radiation, weather relative humidity and weather temperature. Therefore, the predicted days and the similarity days have the most consistency in radiation intensity, followed by temperature and humidity.

After determining the correlation between meteorological factors and power generation, the SDA model is used to determine the similarity days and the optimal similarity day. For each forecasting day, the weighted Euclidean metric values are sorted in ascending order, and the first 56 days are selected as the similarity days and the first 1 day as the optimal similarity day. The optimal similarity days of the 4 forecasting days are shown in Table 2. The K-means clustering method sets three cluster centers for the two irradiances respectively. It takes temperature, humidity, sunny/cloudy/rainy, irradiance all into account to get complete weather information. Thereafter, taking the weather information of the similarity days as the input value, the KM-SDA-ABWO-RBF model was used to establish the functional relationship between the weather factors and the output power of photovoltaic power generation, and finally the output power value was obtained.

3.2 Experimental setup and evaluation metrics

This paper compares the proposed SDA-ABWO-RBF model with SDA-BWO-RBF, SDA-PSO-RBF, SDA-RBF, SDA-ELM, SDA-BPNN, RBF, ELM, BPNN. The population size of all metaheuristic algorithms is 50 and the number of iterations is 250. Acceleration constants of PSO is [2.1, 2.1], and inertia weights is [0.9, 0.6]. In addition, the number of hidden neurons is 4. The number of input neurons and output neurons are 16 and 12, respectively.

The performance of the SDA-ABWO-RBF model is evaluated by four evaluation indicators, namely RMSE, MAE,

the standard deviation of error (SDE) (Eseye et al., 2018) and the coefficient of determination (R^2). Four metrics are expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{p,i} - y_{a,i})^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{p,i} - y_{a,i}| \quad (12)$$

$$SDE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{p,i} - y_{a,i} - \bar{e})^2} \quad (13)$$

$$R^2 = \frac{\left(\sum_{i=1}^n y_{p,i} y_{a,i} - \frac{\sum_{i=1}^n y_{p,i} \sum_{i=1}^n y_{a,i}}{n} \right)^2}{\left(\sum_{i=1}^n y_{p,i}^2 - \left(\frac{\sum_{i=1}^n y_{p,i}}{n} \right)^2 \right) \left(\sum_{i=1}^n y_{a,i}^2 - \left(\frac{\sum_{i=1}^n y_{a,i}}{n} \right)^2 \right)} \quad (14)$$

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n (y_{p,i} - y_{a,i}) \quad (15)$$

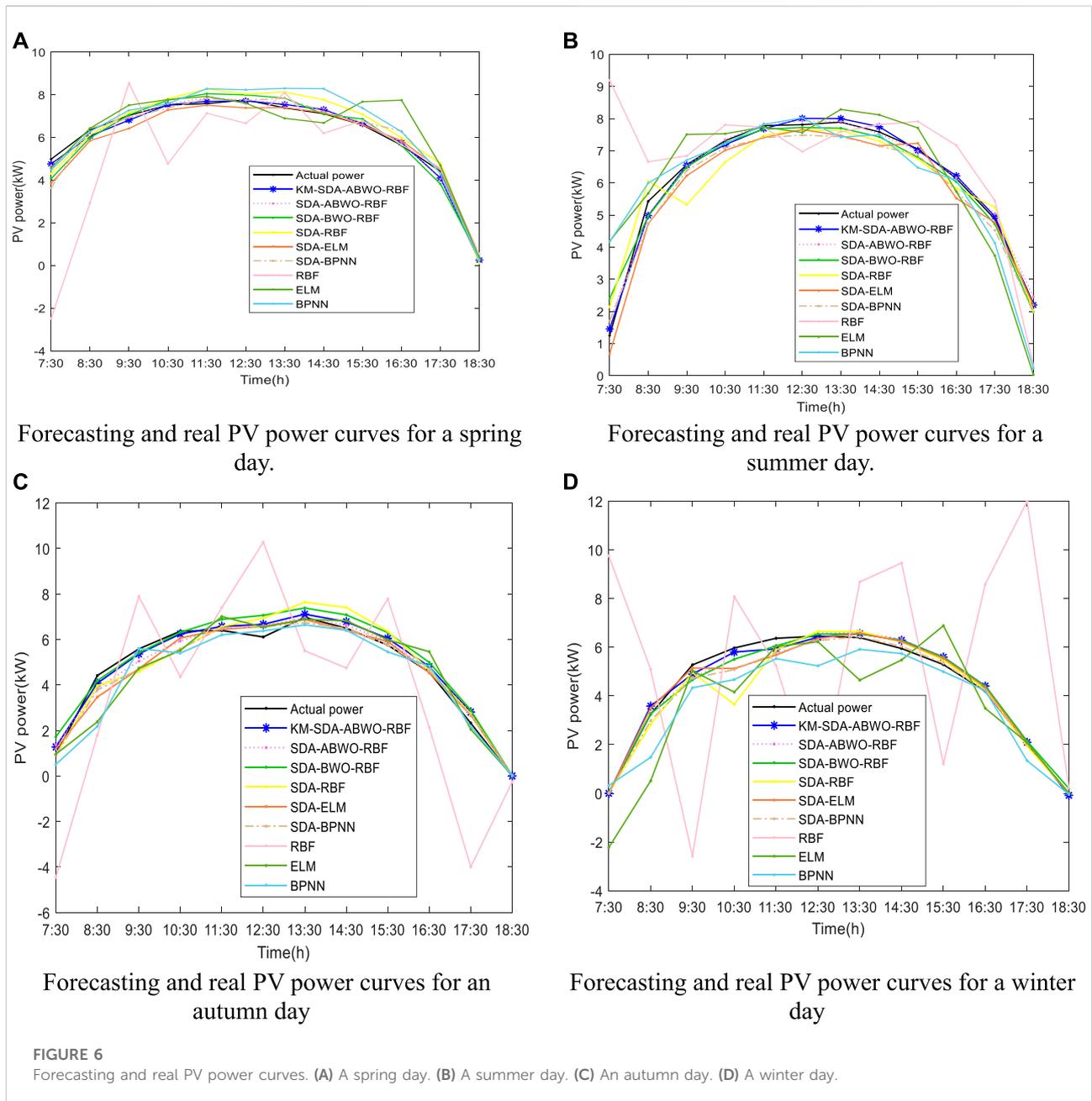
where $y_{p,i}$ and $y_{a,i}$ are the predicted output and actual output at i th hour; n is the number of predicted samples, which is 12.

4 Prediction results and discussion

In order to verify the superiority of the proposed KM-SDA-ABWO-RBF model in PV power generation prediction, this paper compares the model with 8 prediction models (SDA-ABWO-RBF, SDA-BWO-RBF, SDA-RBF, SDA-ELM, SDA-BPNN, RBF, ELM and BPNN). The experiment selects 4 days, spring, summer, autumn and winter, and 12 forecast points per day for forecasting. The optimized number of neurons in the hidden layer is 4 in both ABWO-RBF and BWO-RBF in four seasons. The prediction results obtained from the experiments are discussed from two aspects. On the one hand, four metrics (RMSE, MAE, SDE and R^2) are used to evaluate the accuracy of the predicted values; on the other hand, each prediction model is run 10 times independently, and the stability of the prediction results of the 10 runs is compared on the four metrics sex.

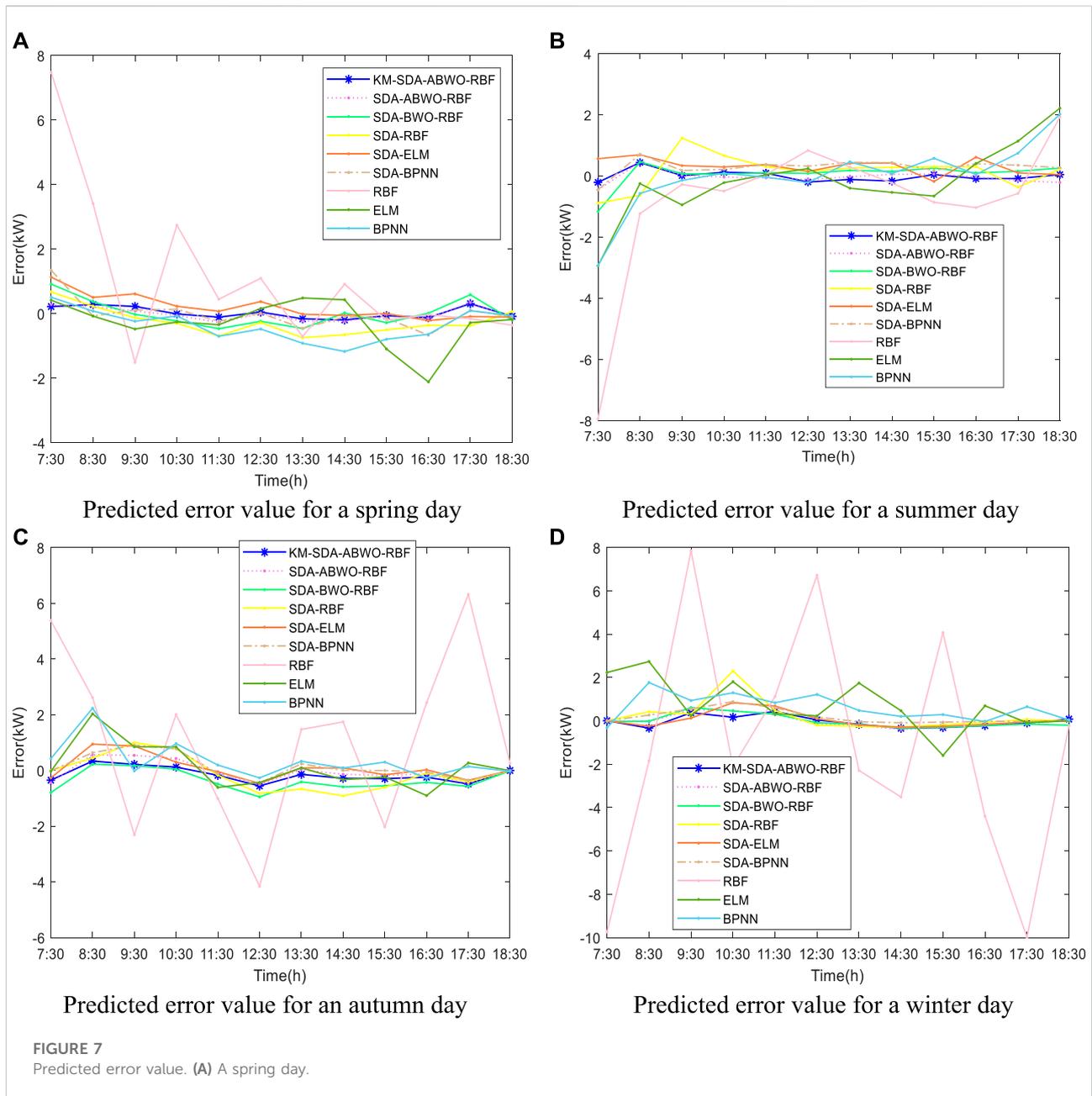
4.1 The accuracy of the forecast results

Figures 6A–D shows the predicted power output curve of the 9 models in different seasons. It can be seen that the predicted values of the KM-SDA-ABWO-RBF model and the ABWO-RBF model are closer to the actual values, while the predicted outputs



of the other models are far from the actual values. Figures 7A–D shows the predicted error curve of 9 models in different seasons. Similarly, the error of the proposed model is less volatile around 0. Figure 8 shows the correlation between the predicted power of the proposed model and the actual power in four seasons. It is obvious from Figure 8 that R^2 of the proposed model in the four seasons is 0.9929, 0.9934, 0.9862 and 0.9881, respectively. Its value is very close to 1, showing the superiority of the model for prediction. summer in the selected dataset has the largest R^2 and the most accurate forecast result.

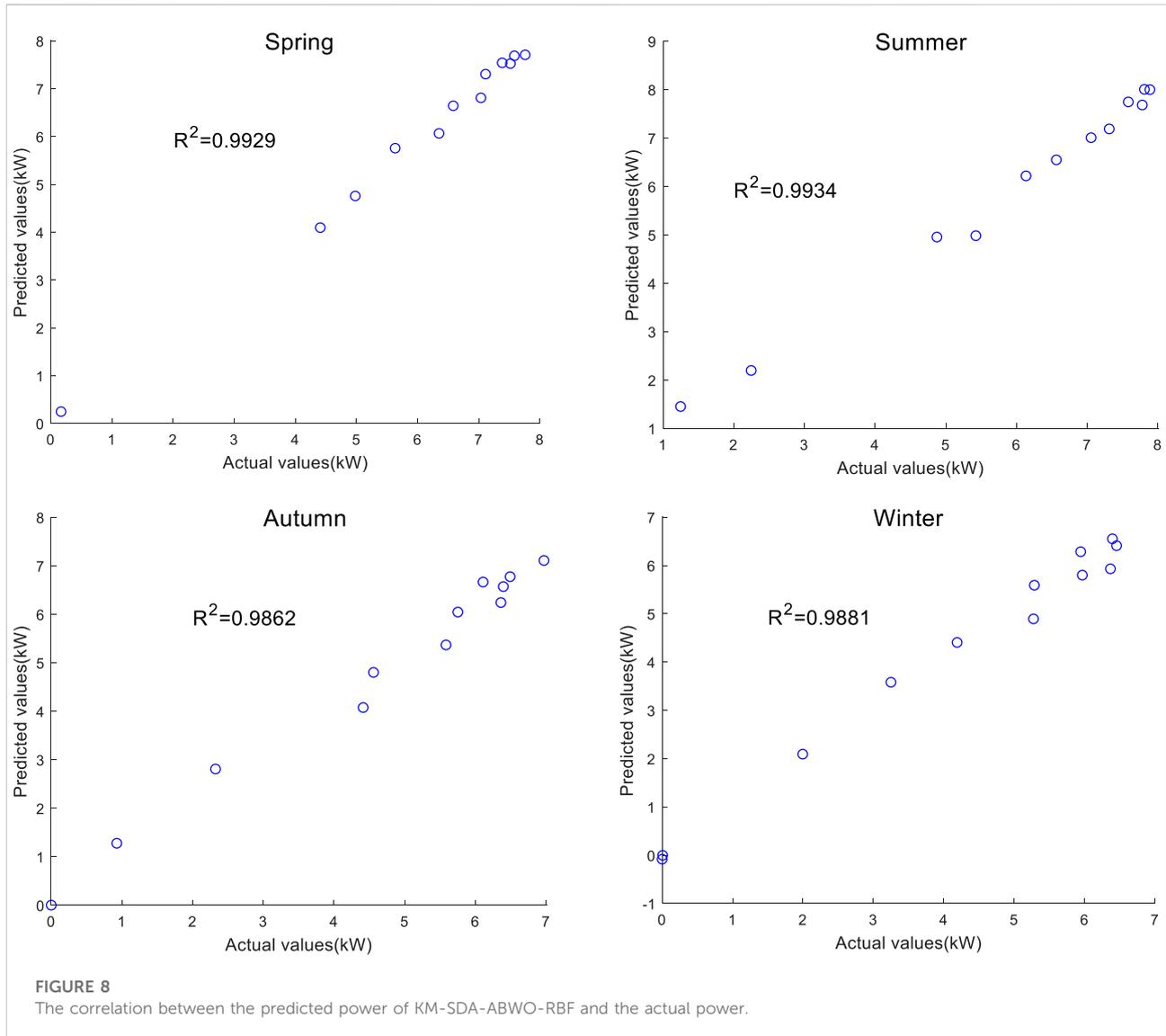
Figure 9A shows predicted RMSE of 9 models for 4 forecasting days. Figure 9B shows predicted MAE of 9 models for 4 forecasting days. Figure 9C shows predicted SDE of 9 models for 4 forecasting days. Figure 9D shows predicted R^2 of 9 models for 4 forecasting days. RMSE, MAE, SDE are several types of errors; the smaller the predicted error value, the higher the prediction accuracy of the model. R^2 represents the correlation between two variables; a larger value of R^2 indicates a more reliable model. From the figure it can be concluded that the proposed model achieved the smallest RMSE,



MAE, SDE, and the largest R^2 for each forecasting day. The prediction accuracy of SDA-ABWO-RBF and SDA-BWO-RBF is second only to the proposed model and higher than the other 6 traditional models that lack metaheuristic optimization. It proves that the metaheuristic algorithm to optimize neural network parameters greatly improves the prediction performance of traditional neural network models. In addition, SDA-ELM outperforms SDA-BPNN and SDA-RBF; SDA-RBF and RBF exhibit the worst performance. The reason for this result is that the parameters of the ELM are random so that the predicted output has a wide range. Therefore, it is

possible to obtain more accurate output power values. Traditional BPNN and RBF have a high probability of leading to early convergence. Overall, it can be seen that on each evaluation metric, SDA-RBF performed better than RBF, SDA-ELM performed better than ELM, and SDA-BPNN performed better than BPNN. It can be shown that the SDA method greatly improves the prediction performance.

Figures 10A–D shows Taylor diagram of 9 models in different seasons. It visually presents the performance of 9 models on Standard Deviation, RMSE and R^2 . KM-SDA-ABWO-RBF is closest to the observation in all four seasons,



while RBF is farthest from observation. Accordingly, it validates the accuracy of KM-SDA-ABWO-RBF model on three metrics.

Table 3 specifically describes the mean values of RMSE, MAE, SDE and R^2 for the four seasons. The mean values of RMSE, MAE, SDE and R^2 of the proposed model were 0.227, 275, 0.19175, 0.215, 725 and 0.99015, respectively. KM-SDA-ABWO-RBF has a 13%, 8% and 17% decrease in RMSE, MAE and SDE, and a 0.4% increase in R^2 compared to SDA-ABWO-RBF. It shows that adding K-means algorithm can further improve the prediction accuracy. In addition, the errors of KM-SDA-ABWO-RBF, SDA-ABWO-RBF and SDA-BWO-RBF were smaller than those of SDA-RBF, and the correlation with the actual value was greater. It illustrates the reliability of the combination of metaheuristics and machine learning methods. Again, the prediction results of KM-SDA-ABWO-RBF and SDA-ABWO-RBF were more accurate than SDA-BWO-RBF. The effectiveness

of the improved BWO algorithm is proved. In a word, KM-SDA-ABWO-RBF achieves the best value for all forecasting days on all evaluation metrics. In the PV power prediction, the experimental results verified that the KM-SDA-ABWO-RBF model outperforms the other models in terms of robustness, effectiveness and accuracy.

4.2 The stability of the forecast results

Figures 11A,B show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for a spring day. Figures 12A,B show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for a summer day. Figures 13A–D show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for an autumn day. Figures 14A–D show the RMSE, MAE, SDE, and R^2 values for 10 runs of

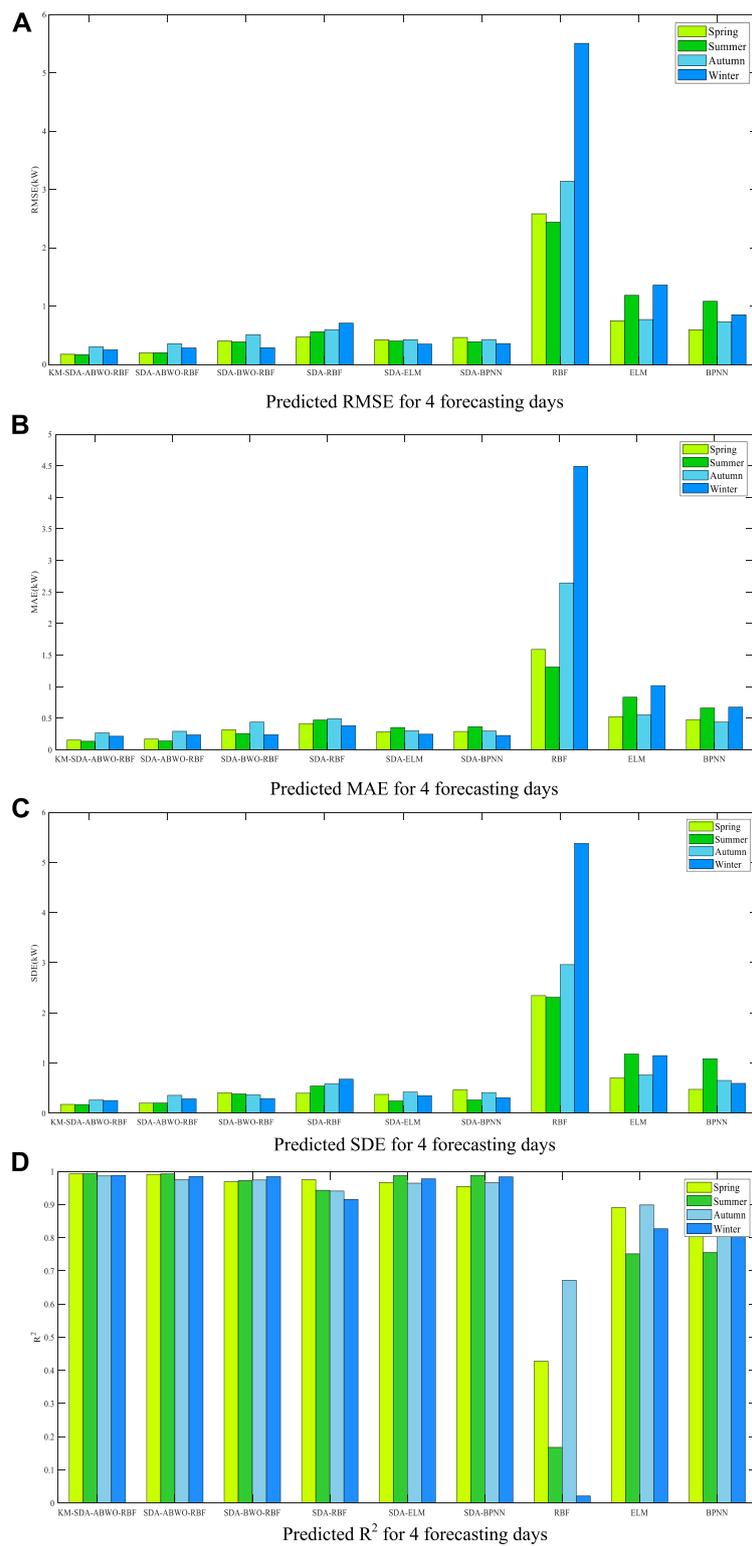


FIGURE 9
 Predicted (A) RMSE. (B) MAE. (C) SDE. (D) R^2 for 4 forecasting days.

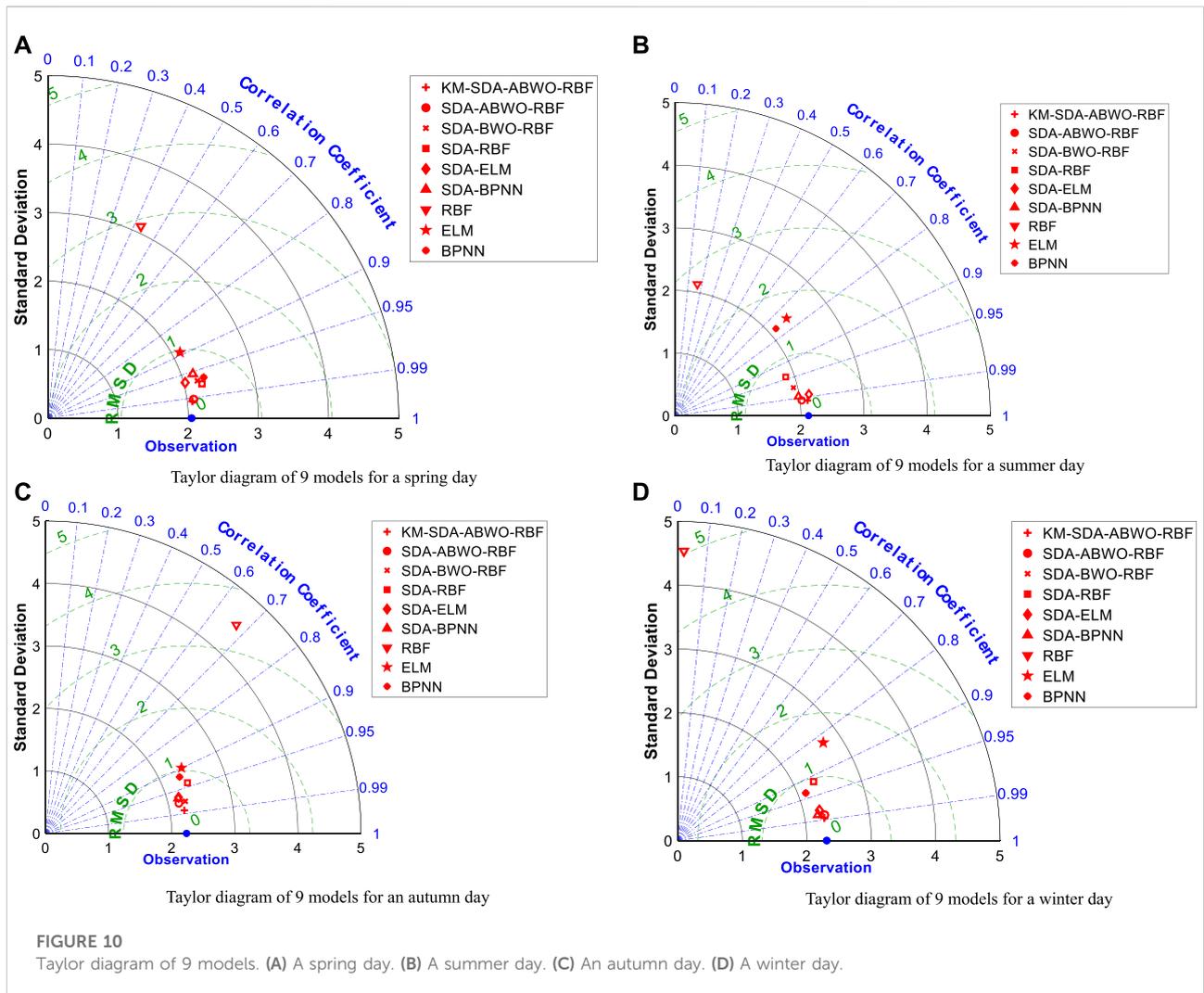


TABLE 3 Average value of RMSE, MAE, SDE and R² in four seasons.

Model	RMSE	MAE	SDE	R ²
KM-SDA-ABWO-RBF	0.227,275	0.19175	0.215,725	0.99015
SDA-ABWO-RBF	0.261,695	0.20815	0.2603	0.98595
SDA-BWO-RBF	0.400,383	0.312,025	0.361,625	0.975,025
SDA-RBF	0.586,443	0.43995	0.550,675	0.943,675
SDA-ELM	0.402,948	0.29595	0.346,525	0.974
SDA-BPNN	0.412,488	0.294,275	0.361,975	0.973
RBF	3.421	2.508,375	3.2487	0.32185
ELM	1.01976	0.73165	0.9475	0.841,875
BPNN	0.818,888	0.56285	0.701,475	0.894,425

9 models for a winter day. Taking a spring day as an example, it can be seen from Figures 26, 27 that the violin plots of SDA-ELM, SDA-BPNN, ELM and BPNN are narrower and longer, while

KM-SDA-ABWO-RBF, SDA-ABWO-RBF and SDA-BWO-RBF are flatter and wider. It shows that the RMSE and MAE values of KM-SDA-ABWO-RBF, SDA-ABWO-RBF and SDA-BWO-RBF are closer and more stable. The SDA-RBF and RBF models have the highest stability, with the same values for 10 runs but the lowest accuracy. Roughly, the stability differences of KM-SDA-ABWO-RBF, SDA-ABWO-RBF and SDA-BWO-RBF are small, and the prediction results of the three models all have high stability and accuracy. However, SDA-BWO-RBF has higher stability than SDA-ABWO-RBF, which is more stable than KM-SDA-ABWO-RBF. Stability is reflected in the length of the box in Figures 11A,B. It can be clearly found from Figures 11A,B that SDA-ELM and SDA-BPNN are more stable than ELM and BPNN, respectively. It verifies that SDA can not only improve the prediction accuracy of PV generation power prediction, but also enhance the stability. Similarly, the same conclusion is reached for summer, autumn and winter.

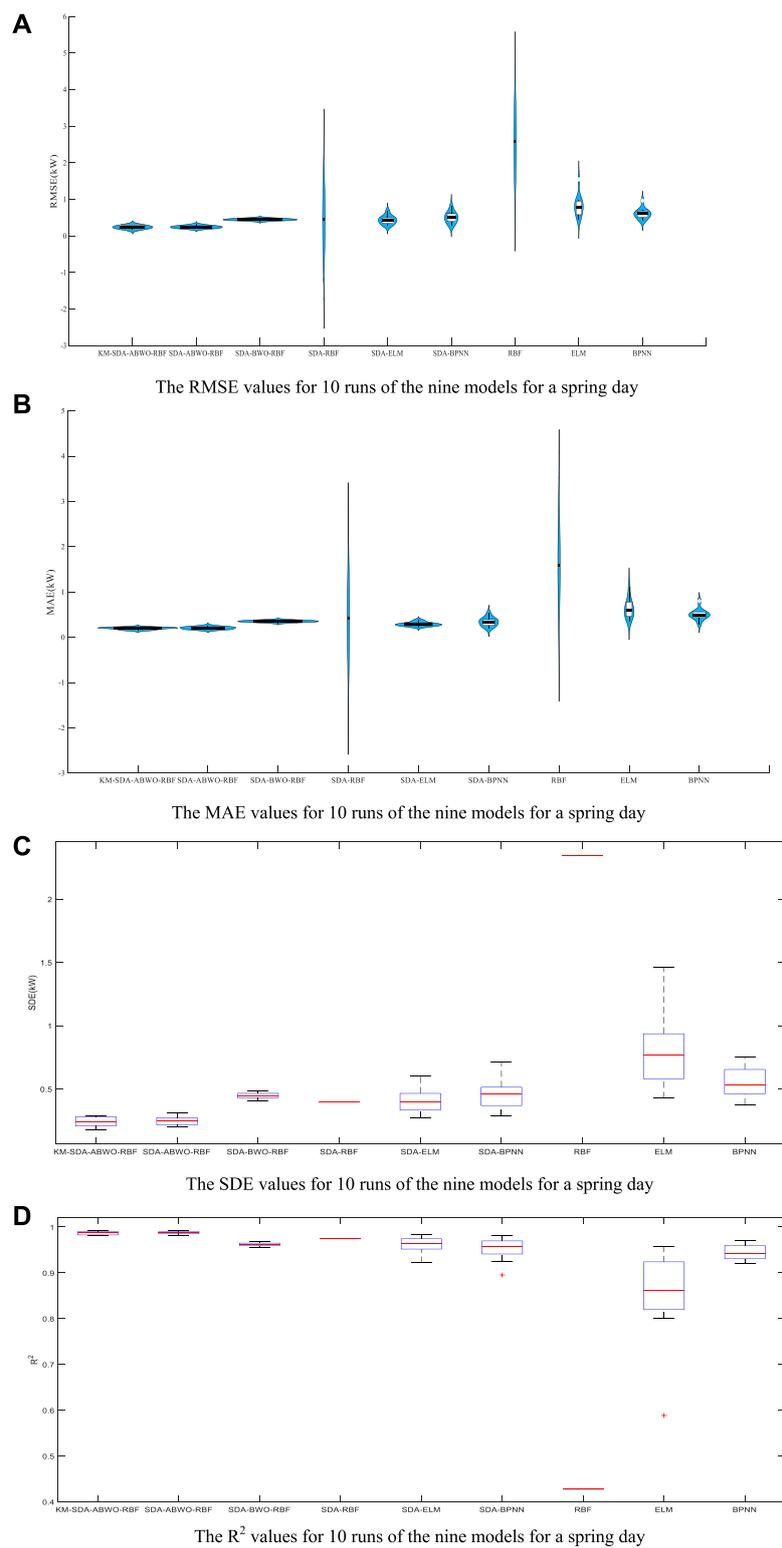


FIGURE 11
 (A–D) shows the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for a spring day.

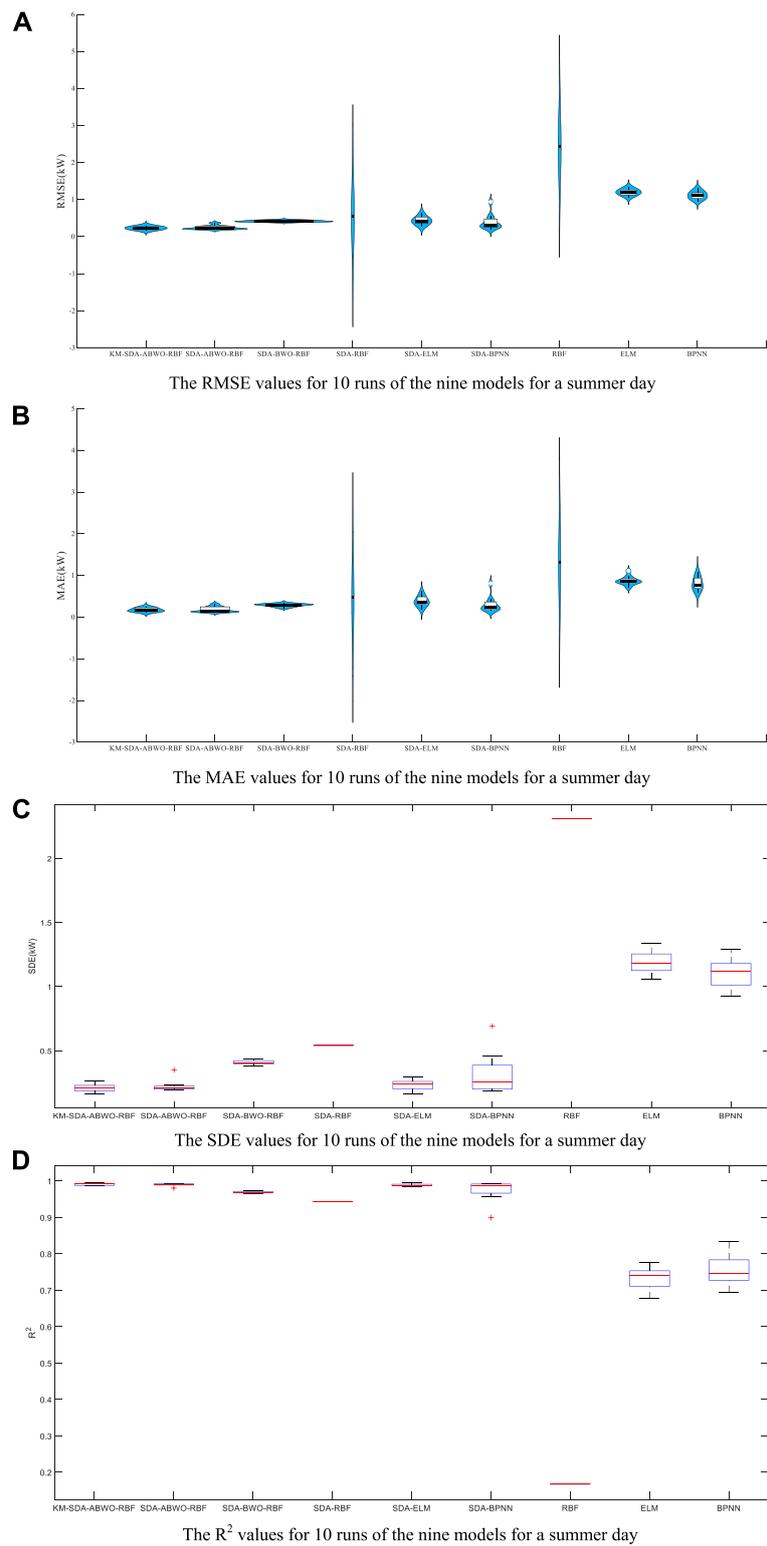


FIGURE 12 (A–D) show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for a summer day.

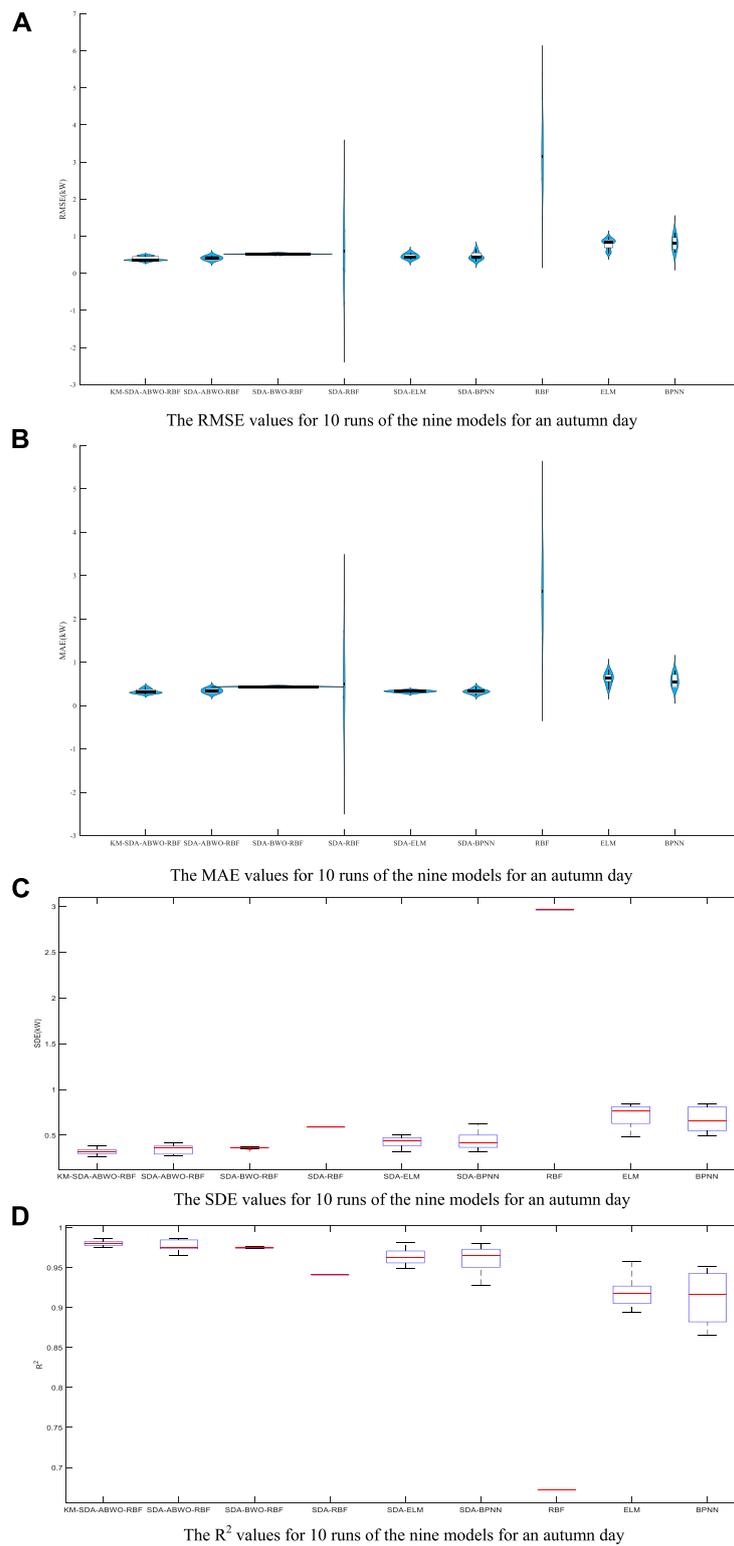


FIGURE 13
 (A–D) show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for an autumn day.

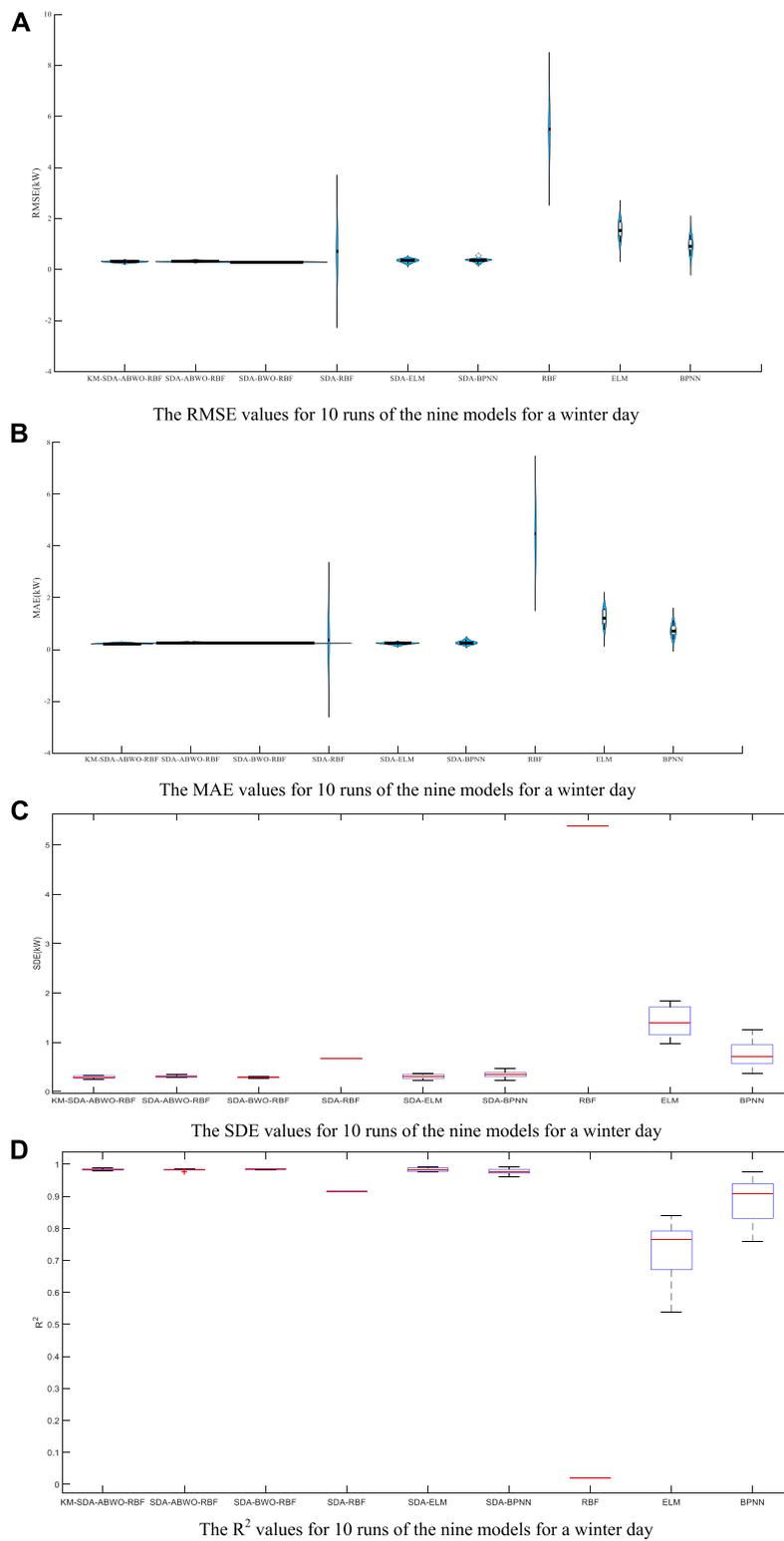


FIGURE 14
(A–D) show the RMSE, MAE, SDE, and R^2 values for 10 runs of 9 models for a winter day.

5 Conclusion and future work

This paper develops a KM-SDA-ABWO-RBF model, and utilizes the ABWO algorithm to train the parameters of RBFNNs to improve the generalization performance of RBFNNs and reduce the computational complexity. The ABWO algorithm enhances the search ability by introducing an adaptive factor, that is, the ability to search for optimal neural network parameters. Since the structure and parameters of the network determine the performance of RBFNNs, constructing the ABWO-RBF algorithm improves the performance and stability of the neural network. In addition, SDA and K-means methods are introduced to obtain historical data with strong correlation with the forecast day, which improves the overall accuracy of ABWO-RBF neural network in photovoltaic power prediction. The experimental results demonstrate the simplicity and high efficiency of the proposed model in the prediction of PV power generation. In the future, the hybrid model is still valid for forecast of the PV power output. Deep learning methods used in this field are expected to have higher prediction accuracy. The model in this paper can be used can be used in power plants to stably and efficiently regulate photovoltaic power generation in the future.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

References

- Abualigah, L., and Diabat, A. (2021). Advances in sine cosine algorithm: A comprehensive survey. *Artif. Intell. Rev.* 54 (4), 2567–2608. doi:10.1007/s10462-020-09909-3
- Akhter, M. N., Mekhilef, S., Mokhlis, H., and Mohamed Shah, N. (2019). Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renew. Power Gener.* 13 (7), 1009–1023. doi:10.1049/iet-rpg.2018.5649
- Bugala, A., Zaborowicz, M., Boniecki, P., Janczak, D., Koszela, K., Czekala, W., et al. (2018). Short-term forecast of generation of electric energy in photovoltaic systems. *Renew. Sustain. Energy Rev.* 81, 306–312. doi:10.1016/j.rser.2017.07.032
- Chen, C., Duan, S., Cai, T., and Liu, B. (2011). Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Sol. energy* 85 (11), 2856–2870. doi:10.1016/j.solener.2011.08.027
- Chen, S., Cowan, C. F. N., and Grant, P. M. (1991). Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Trans. Neural Netw.* 2 (2), 302–309. doi:10.1109/72.80341
- Cheng, M. Y., and Prayogo, D. (2014). Symbiotic organisms search: A new metaheuristic optimization algorithm. *Comput. Struct.* 139, 98–112. doi:10.1016/j.compstruc.2014.03.007
- Dolara, A., Grimaccia, F., Leva, S., Mussetta, M., and Ogliari, E. (2015). A physical hybrid artificial neural network for short term forecasting of PV plant power output. *Energies* 8 (2), 1138–1153. doi:10.3390/en8021138
- Dong, N., Chang, J. F., Wu, A. G., and Gao, Z. K. (2020). A novel convolutional neural network framework based solar irradiance prediction method. *Int. J. Electr. Power & Energy Syst.* 114, 105411. doi:10.1016/j.ijepes.2019.105411
- Du, P., Wang, J., Yang, W., and Niu, T. (2018). Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. *Renew. energy* 122, 533–550. doi:10.1016/j.renene.2018.01.113
- Duvvuri, S. P., and Anmala, J. (2019). Fecal coliform predictive model using genetic algorithm-based radial basis function neural networks (GA-RBFNNs). *Neural comput. Appl.* 31 (12), 8393–8409. doi:10.1007/s00521-019-04520-2
- Er, M. J., Wu, S., Lu, J., and Hock Lye Toh (2002). Face recognition with radial basis function (RBF) neural networks. *IEEE Trans. Neural Netw.* 13 (3), 697–710. doi:10.1109/tnn.2002.1000134
- Eseye, A. T., Zhang, J., and Zheng, D. (2018). Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. *Renew. energy* 118, 357–367. doi:10.1016/j.renene.2017.11.011
- Feng, H. M. (2006). Self-generation RBFNs using evolutionary PSO learning. *Neurocomputing* 70 (1-3), 241–251. doi:10.1016/j.neucom.2006.03.007
- Ghayekhloo, M., Ghofrani, M., Menhaj, M. B., and Azimi, R. (2015). A novel clustering approach for short-term solar radiation forecasting. *Sol. Energy* 122, 1371–1383. doi:10.1016/j.solener.2015.10.053
- Ghimire, S., Deo, R. C., Downs, N. J., and Raj, N. (2018). Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. *Remote Sens. Environ.* 212, 176–198. doi:10.1016/j.rse.2018.05.003
- Gonzalez, B., Valdez, F., Melin, P., and Prado-Arechiga, G. (2015). Fuzzy logic in the gravitational search algorithm for the optimization of modular neural networks in pattern recognition. *Expert Syst. Appl.* 42 (14), 5839–5847. doi:10.1016/j.eswa.2015.03.034

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This research is funded by National Natural Science Foundation of China under Grant No. 62066005, U21A20464, Project of the Guangxi Science and Technology under Grant No. AD21196006.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Han, H., Wu, X., Zhang, L., Tian, Y., and Qiao, J. (2017). Self-organizing RBF neural network using an adaptive gradient multiobjective particle swarm optimization. *IEEE Trans. Cybern.* 49 (1), 69–82. doi:10.1109/tcyb.2017.2764744
- Hayyolalam, V., and Kazem, A. (2020). Black widow optimization algorithm: A novel meta-heuristic approach for solving engineering optimization problems. *Eng. Appl. Artif. Intell.* 87, 103249–103249. doi:10.1016/j.engappai.2019.103249
- Hossain, M. S., and Mahmood, H. (2020). Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. *IEEE Access* 8, 172524–172533. doi:10.1109/access.2020.3024901
- Huang, G. B., Saratchandran, P., and Sundararajan, N. (2005). A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation. *IEEE Trans. Neural Netw.* 16 (1), 57–67. doi:10.1109/tnn.2004.836241
- Huang, G. B., Saratchandran, P., and Sundararajan, N. (2004). An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks. *IEEE Trans. Syst. Man, Cybern. B* 34 (6), 2284–2292. doi:10.1109/tsmcb.2004.834428
- Justus, J. J., and Anuradha, M. (2022). A golden eagle optimized hybrid multilayer perceptron convolutional neural network architecture-based three-stage mechanism for multiuser cognitive radio network. *Int. J. Commun.* 35 (4), e5054. doi:10.1002/dac.5054
- Katooli, M. S., and Koochaki, A. (2020). Detection and classification of incipient faults in three-phase power transformer using DGA information and rule-based machine learning method. *J. Control Autom. Electr. Syst.* 31 (5), 1251–1266. doi:10.1007/s40313-020-00625-5
- Lee, C. M., and Ko, C. N. (2009). Time series prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm. *Neurocomputing* 73 (1–3), 449–460. doi:10.1016/j.neucom.2009.07.005
- Li, Y., Su, Y., and Shu, L. (2014). An ARMAX model for forecasting the power output of a grid connected photovoltaic system. *Renew. Energy* 66, 78–89. doi:10.1016/j.renene.2013.11.067
- Lin, P., Peng, Z., Lai, Y., Cheng, S., Chen, Z., and Wu, L. (2018). Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. *Energy Convers. Manag.* 177, 704–717. doi:10.1016/j.enconman.2018.10.015
- Liu, J., Fang, W., Zhang, X., and Yang, C. (2015). An improved photovoltaic power forecasting model with the assistance of aerosol index data. *IEEE Trans. Sustain. Energy* 6 (2), 434–442. doi:10.1109/tste.2014.2381224
- Liu, Y., Li, Y., Li, G., Zhang, B., and Wu, G. (2005). “Constructive ensemble of RBF neural networks and its application to earthquake prediction,” in International Symposium on Neural Networks, Chongqing, China, May 30–June 1, 2005 (Berlin, Heidelberg: Springer), 532–537.
- Mahmud, K., Azam, S., Karim, A., Zobaed, S., Shanmugam, B., and Mathur, D. (2021). Machine learning based PV power generation forecasting in alic springs. *IEEE Access* 9, 46117–46128. doi:10.1109/access.2021.3066494
- Mandal, P., Madhira, S. T. S., Meng, J., and Pineda, R. L. (2012). Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques. *Procedia Comput. Sci.* 12, 332–337. doi:10.1016/j.procs.2012.09.080
- Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T., and Srivastava, A. K. (2007). A novel approach to forecast electricity price for PJM using neural network and similar days method. *IEEE Trans. Power Syst.* 22 (4), 2058–2065. doi:10.1109/tpwrs.2007.907386
- Memar, S., Mahdavi-Meymand, A., and Sulisz, W. (2021). Prediction of seasonal maximum wave height for unevenly spaced time series by black widow optimization algorithm. *Mar. Struct.* 78 (10), 103005. doi:10.1016/j.marstruc.2021.103005
- Mirjalili, S. A., Hashim, S. Z. M., and Sardroudi, H. M. (2012). Training feedforward neural networks using hybrid particle swarm optimization and gravitational search algorithm. *Appl. Math. Comput.* 218 (22), 11125–11137. doi:10.1016/j.amc.2012.04.069
- Mirjalili, S., Mirjalili, S. M., and Lewis, A. (2014). Grey wolf optimizer. *Adv. Eng. Softw.* 69, 46–61. doi:10.1016/j.advengsoft.2013.12.007
- Mirjalili, S. (2016). Sca: A sine cosine algorithm for solving optimization problems. *Knowledge-based Syst.* 96, 120–133. doi:10.1016/j.knsys.2015.12.022
- Panahi, F., Ehteram, M., and Emami, M. (2021). Suspended sediment load prediction based on soft computing models and Black Widow Optimization Algorithm using an enhanced gamma test. *Environ. Sci. Pollut. Res.* 28 (35), 48253–48273. doi:10.1007/s11356-021-14065-4
- Parsopoulos, K. E., and Vrahatis, M. N. (2002). Recent approaches to global optimization problems through Particle Swarm Optimization. *Nat. Comput.* 1 (2–3), 235–306. doi:10.1023/a:1016568309421
- Qiao, J., Li, F., Yang, C., Li, W., and Gu, K. (2019). A self-organizing RBF neural network based on distance concentration immune algorithm. *IEEE/CAA J. Autom. Sin.* 7 (1), 276–291. doi:10.1109/jas.2019.1911852
- Rashedi, E., Nezamabadi-Pour, H., and Saryzadi, S. (2009). Gsa: A gravitational search algorithm. *Inf. Sci.* 179 (13), 2232–2248. doi:10.1016/j.ins.2009.03.004
- Reikard, G. (2009). Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Sol. Energy* 83 (3), 342–349. doi:10.1016/j.solener.2008.08.007
- Renewable Energy Policy (2017). Renewable energy policy network for the 21st century (REN21), renewables global status report (paris: Various editions). <http://www.ren21.net/status-of-renewables/global-status-report/> (Accessed 27 January 2017).
- Sobri, S., Koohi-Kamali, S., and Rahim, N. A. (2018). Solar photovoltaic generation forecasting methods: A review. *Energy Convers. Manag.* 156, 459–497. doi:10.1016/j.enconman.2017.11.019
- Storn, R., and Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* 11 (4), 341–359. doi:10.1023/a:1008202821328
- Tightiz, L., Nasab, M. A., Yang, H., and Addeh, A. (2020). An intelligent system based on optimized ANFIS and association rules for power transformer fault diagnosis. *ISA Trans.* 103, 63–74. doi:10.1016/j.isatra.2020.03.022
- Wang, J., Yang, W., Du, P., and Li, Y. (2018). Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. *Energy* 148, 59–78. doi:10.1016/j.energy.2018.01.112
- Wang, P., Zhou, Y., Luo, Q., Han, C., Niu, Y., and Lei, M. (2020). Complex-valued encoding metaheuristic optimization algorithm: A comprehensive survey. *Neurocomputing* 407, 313–342. doi:10.1016/j.neucom.2019.06.112
- Wolpert, D. H., and Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* 1 (1), 67–82. doi:10.1109/4235.585893
- Xie, T., Hao, Y., Hewlett, J., Rozycki, P., and Wilamowski, B. (2012). Fast and efficient second-order method for training radial basis function networks. *IEEE Trans. Neural Netw. Learn. Syst.* 23 (4), 609–619. doi:10.1109/tnnls.2012.2185059
- Xu, R., Chen, H., and Sun, X. (2012). “Short-term photovoltaic power forecasting with weighted support vector machine,” in 2012 IEEE International Conference on Automation and Logistics (IEEE), 248
- Yang, X. S., and Deb, S. (2009). *Cuckoo search via Lévy flights*. 2009 World congress on nature & biologically inspired computing (NaBIC). Ieee, 210
- Zhang, Y., Beaudin, M., Taheri, R., Zareipour, H., and Wood, D. (2015). Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators. *IEEE Trans. Smart Grid* 6 (5), 2253–2262. doi:10.1109/tsg.2015.2397003
- Zhou, Y., Zhou, N., Gong, L., and Jiang, M. (2020). Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine. *Energy* 204, 117894. doi:10.1016/j.energy.2020.117894