

Article

A New Hybrid Short-Term Interval Forecasting of PV Output Power Based on EEMD-SE-RVM

Sen Wang, Yonghui Sun *, Yan Zhou, Rabea Jamil Mahfoud  and Dongchen Hou

College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China; senwang@hhu.edu.cn (S.W.); zhouyan@hhu.edu.cn (Y.Z.); rabea7mahfoud@hotmail.com (R.J.M.); hdc190406030001@hhu.edu.cn (D.H.)

* Correspondence: sunyonghui168@gmail.com; Tel.: +86-139-0516-9126

Received: 30 October 2019; Accepted: 23 December 2019; Published: 23 December 2019



Abstract: The main characteristics of the photovoltaic (PV) output power are the randomness and uncertainty, such features make it not easy to establish an accurate forecasting method. The accurate short-term forecasting of PV output power has great significance for the stability, safe operation and economic dispatch of the power grid. The deterministic point forecast method ignores the randomness and volatility of PV output power. Aiming at overcoming those defects, this paper proposes a novel hybrid model for short-term PV output power interval forecasting based on ensemble empirical mode decomposition (EEMD) as well as relevance vector machine (RVM). Firstly, the EEMD is used to decompose the PV output power sequences into several intrinsic mode functions (IMFs) and residual (RES) components. After that, based on the decomposed components, the sample entropy (SE) algorithm is utilized to reconstruct those components where three new components with typical characteristics are obtained. Then, by implementing RVM, the forecasting model for every component is developed. Finally, the forecasting results of every new component are superimposed in order to achieve the overall forecasting results with certain confidence level. Simulation results demonstrate, by comparing them with some previous methods, that the hybrid method based on EEMD-SE-RVM has relatively higher forecasting accuracy, more reliable forecasting interval and high engineering application value.

Keywords: photovoltaic output power forecasting; hybrid interval forecasting; relevance vector machine; sample entropy; ensemble empirical mode decomposition

1. Introduction

With the development of industrialization, traditional fossil fuels are faced with the increased depletion and the environmental pollution problems brought by fossil fuels' combustion become the main obstacle to global economic development. To solve this problem, in the past few decades, more and more attentions have been paid on the renewable energy sources, such as biomass energy, tidal energy, wind energy, solar energy, etc. [1]. However, due to the intermittency and variability of those renewable energies, they would cause unavoidable fluctuations and instability if they are highly integrated in the power grid. Therefore, how to obtain the accurate forecast of renewable energy sources is massively important for the safe, steady and reliable operation of power grid [2].

Regarding the short-term renewable energy generation forecasting, the existing models are roughly divided into four categories: artificial intelligence based models (AIBM), statistical models, physical models and hybrid models [3]. In [4], the statistical smoothing techniques were utilized to create a statistical normalization of the solar energy, which was beneficial to implement the online short-term power forecasting of photovoltaic (PV). In [5], the ARIMA model was taken as a statistical model to realize the output power forecasting of a PV-grid-connected system. As a method of statistical and

machine learning, ensemble approach also played a crucial role in short-term load forecasting [6,7]. In [6], an ensemble approach was combined together with extreme learning machine (ELM) and wavelet for short-term load forecasting in solar power system. In [7], a solar forecasting model was proposed based on multiple satellite images and support vector machine (SVM). The motion vector of the cloud was predicted by the satellite atmosphere motion vector (AMVS) image, then, the output prediction of the PV power was realized. In [8], the ANN techniques were combined with spatial modes to forecast the daily global horizontal irradiance. Physical models used physical factors to construct the required models [9–15], and in most cases, there were no distinct boundaries within different models, thus the hybrid models [4,5,11,16–18] have become the most frequently used models to forecast PV generation. For example, in [4], the statistical models and AIBM were combined to implement short-term solar power forecasting. In [11,16–18], the AIBM and physical models were integrated to obtain the forecasting of PV systems output power. On the other hand, by taking the randomness and uncertainty of solar energy into account, in recent years, there have been lots of results discussing the short-term forecasting problem of PV output power [5,16,19]. Besides the one-day-ahead time horizons [16], other forecasting time scales have also been considered, such as one-hour-ahead, 15-, 30- and 45-min-ahead time horizons [20].

However, in most of the aforementioned models, only point forecast problems were concerned, with few determined values were achieved. Nevertheless, many forecasting errors were detected in the results [21]. Besides that, those models lacked the ability to describe the non-stationary with a probable range of fluctuation. Different from the specific value of conventional point forecast, prediction interval (PI) can deliver a quantification of uncertainty with a prescribed confidence level, which indicates the probable prediction. Due to the uncertainty of the forecasting, a range consisting of upper and lower bounds with the indication of accuracy is more credible than the conventional prediction points [22]. Interval forecasting can provide more information about changeability of the target variable, which is more suitable to predict the renewable energy generation [23,24]. According to the results of point forecasting, if the probability distribution of model error is known exactly, the prediction interval can be calculated accurately. In [25], a method was established based on ELM and the pairs bootstrap and then applied to obtain the probabilistic interval forecasting of wind power, where the prediction error was assumed to obey Gaussian distribution. In [26], the prediction error was analyzed and assumed to obey Beta distribution, and then the interval forecasting model was developed. The conventional prediction interval methods mainly depend on the accuracy of point forecasts and error assumptions, but it is difficult to quantify a special prior error assumption, which influences the performance of prediction interval.

Up to now, several forecasting methods have been proposed for forecasting renewable energy power [27–31]. For data with strong randomness, the preprocessing of data is especially important to improve the prediction accuracy. The common data processing methods include EMD, ensemble empirical mode decomposition (EEMD) and wavelet decomposition. For example, EMD can decompose complex sequences and then predict them separately. In order to obtain better performance of wind forecasting, in [32], the prediction interval is optimized by combining the conditional probability. In [33], the EEMD method was used to solve the model mixing problems. However, the relativities among the decompositions were usually ignored in the conventional EEMD methods, where some complexity was also added. In [34], a kind of ELM was proposed to realize the probabilistic interval forecasting of wind power, where the authors used a two-layer integrated machine learning method. In [35], the random forest model of different meteorological conditions was established and the components were predicted, then the weighted output was carried out on the prediction results. To obtain better performance of short-term forecasting, EEMD method based on sample entropy (SE) was proposed, which was more effective and accurate than the conventional EEMD.

Nevertheless, there have been few interval prediction methods of solar power based on EEMD, which decomposed the time series into diverse frequency components and forecasting each component to improve the accuracy. Thus, the method involving EEMD and SE was used to decompose the

original sequence into different new components. That method was also used to construct the different components in order to analyze the complexity. Then, the characteristic of EEMD method was optimized. The results considering the interval forecasting methods by the hybrid method including EEMD, SE and relevance vector machine (RVM), which have great challenge and importance can enhance the accuracy of the conventional RVM method.

Based on the above discussions, this paper proposes a new hybrid model based on EEMD-SE-RVM for short-term interval forecasting of PV output power. Several intrinsic mode functions (IMFs) and residual (RES) components can be obtained by using the EEMD to decompose the original PV power output sequences. Consequently, three new components with typical characteristics are obtained based on the SE algorithm. Then, for each new component, a prediction model is established using RVM, respectively, and, the forecasting results of every new component are superimposed so that the overall forecasting result with a certain confidence level is obtained. Considering the simulated case study, the results show that this hybrid approach is very effective and has a robust generalization ability as well as a strong practical application value.

The rest of this article is organized as follows. Section 2 introduces the basic models of EEMD, SE and RVM algorithms, respectively. Section 3 develops the hybrid model interval forecasting of PV output power. Case studies and numerical results are given in Section 4. Finally, conclusions are drawn in Section 5.

2. Methodology

2.1. EEMD Principle

The most obvious drawback of conventional EMD is that it will produce mode mixing, which indicates that either a single IMF consisting of obvious different proportion or composed of signals of the same proportion in different IMF components, and it usually leads to signal instability. Aiming at solving this drawback, a new method named EEMD was proposed, which is basically a noise-assisted data analysis method. This demonstrates that noise can be performed using in the EMD method.

In EEMD, there are two important parameters. One is the amplitude k of the white noise and the other is maximum number of iterations M of EMD. Usually, the values of M and k are chosen according to the characteristics of personal experience and data. Without loss of generality, in this paper, M was taken as 100 and the range of k was 0.05–0.5.

The detailed steps of EEMD can be highlighted in the following five points [18]:

- (1) Set both values of k and M .
- (2) The white noise sequence is added to the signal.
- (3) EMD is used to decompose the signal that has been added with white noise to IMFs.
- (4) Repeat steps (2) and (3) for a certain amount of white noise each time and the decomposition of corresponding IMF components is obtained. The average of all the corresponding IMFs was calculated where it is the final result of each IMF. Then, the average value of all residual components was calculated, and the average value was taken as the final result of the residual.

$$\bar{c}_i(t) = \sum_{n=1}^N c_{i,n}(t)/N, \quad \bar{r}_m(t) = \sum_{n=1}^N r_{m,n}(t)/N. \quad (1)$$

- (5) Output $\bar{c}_i(t)$ ($i = 1, \dots, m$) represents IMF components and $\bar{r}_m(t)$ represents the RES component.

2.2. SE Principle

For the IMF components and the RES component that are decomposed by the EEMD, if the forecasting model is developed individually, the calculation will be greatly increased, and the correlation between different components will be ignored. In this paper, the sample entropy theory was used for recombination of these components with relevant characteristics.

For a given k, r and N , where k represents embedding dimension, r denotes tolerance, N represents number of data points. $SampEn(N, k, r)$ is the negative logarithm of the conditional probability. For a data sequence $\{x_i\} = \{x(1), \dots, x(N)\}$, the specific algorithm of sample entropy is expressed as follows:

(1) Construct the sequence $\{x_i\}$ constitute m-dimensional vector

$$X(i) = [x(i), x(i+1), \dots, x(i+k-1)] \quad (2)$$

(2) Define the distance $d_k(X(i), X(j))$ between vectors $X(i)$ and $X(j)$ as the absolute maximum difference between their scalar components

$$d_k(X(i), X(j)) = \max_{0 \sim k-1} |x(i+k) - x(j+k)| \quad (3)$$

(3) For a given value of r , count the number of $d_k(X(i), X(j)) \leq r$, and then calculate the ratio of $N - k$. Be defined as

$$B_i^k(r) = \frac{1}{N-k} \text{num}\{d_k(X(i), X(j)) \leq r\} \quad (4)$$

where r denotes the threshold, which serves as a noise filter, $r > 0; i = 1, \dots, N - k + 1$.

(4) The mean value of $B_i^k(r)$ can be represented as

$$B^k(r) = \frac{1}{N-k+1} \sum_{i=1}^{N-k+1} B_i^k(r) \quad (5)$$

(5) By increasing the iteration to $k + 1$, repetition step (1) to step (4), the mean value of $B_i^{k+1}(r)$ can be represented as

$$B^{k+1}(r) = \frac{1}{N-k} \sum_{i=1}^{N-k} B_i^{k+1}(r) \quad (6)$$

(6) Finally, $SampEn$ for a finite data length of N can be estimated as

$$SampEn(N, k, r) = -\ln[B^{k+1}(r) / B^k(r)] \quad (7)$$

In general, r is between 0.1 and 0.25 SD, k equals to 1 or 2, among them SD represents the standard deviation of time series. Here k is set as 2 and r is 0.15 SD.

2.3. RVM Principle

Comparing with other forecasting algorithms, RVM not only has the characteristics of modeling highly sparse, less optimized parameters, flexible kernel selection and strong generalization ability, but also can directly implement the interval forecasting. Therefore, RVM is used to develop the interval forecasting model for those new components reconstructed by SE.

For a specified input training sample $\{x_n\}_{n=1}^N$ and the corresponding output set $\{t_n\}_{n=1}^N$, the relevance vector machine regression model can be defined as follows

$$t_i = \sum_{i=1}^N \omega_i K(x, x_i) + \omega_0 + \varepsilon \quad (8)$$

where $\varepsilon \sim N(0, \sigma^2)$ is the error of the independent sample, ω_i are the model weights, N is the sample size and $K(x, x_i)$ is a nonlinear kernel function.

Given a training sample set $\{x_i, t_i\}_{i=1}^N$, suppose the target value t_i is independent and the noise in data follows the Gaussian distribution with the variance σ^2 , then the likelihood function of the training sample set can be described as

$$\begin{aligned} p(t|\omega, \sigma^2) &= \prod_{n=1}^N p(t_i|\omega, \sigma^2) \\ &= (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{\|t-\Phi\omega\|^2}{2\sigma^2}\right\} \end{aligned} \quad (9)$$

where $t = (t_1, \dots, t_n)^T$, $\omega = (\omega_0, \omega_1, \dots, \omega_n)^T$ and Φ is the design matrix defined by

$$\Phi = \begin{bmatrix} 1 & K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_N) \\ 1 & K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_N, x_1) & K(x_N, x_2) & \cdots & K(x_N, x_N) \end{bmatrix} \quad (10)$$

Based on the priori probabilities distribution and likelihood distribution, the posterior distribution over the weight form Bays rule can be written as

$$\begin{aligned} p(\omega|t, \alpha, \sigma^2) &= \frac{p(t|\omega, \sigma^2)p(\omega|\alpha)}{p(t|\alpha, \sigma^2)} \\ &= (2\pi)^{-(N+1)/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\omega - \mu)^T \Sigma^{-1}(\omega - \mu)\right\} \end{aligned} \quad (11)$$

where $\Sigma = (\sigma^{-2}\Phi^T\Phi + A)^{-1}$, $\mu = \sigma^{-2}\Sigma\Phi^T t$ and $A = \text{diag}(a_0, a_1, \dots, a_N)$.

At last, the hyper parameter α and the variance σ^2 can be estimated by using the maximum likelihood algorithm.

The input value is x_i^* , then the corresponding forecasting value can be described as [13]

$$\begin{cases} y_* = \varphi(x_i^*)\mu \\ \sigma_*^2 = \sigma_{MP}^2 + \varphi(x_i^*)^T \Sigma \varphi(x_i^*) \end{cases} \quad (12)$$

Under the confidence level of α , the interval forecasting value results can be described as [25]

$$[L_b^\alpha, U_b^\alpha] = [y_* - z_{\alpha/2}\sigma_*, y_* + z_{\alpha/2}\sigma_*] \quad (13)$$

where L_b^α and U_b^α represents lower and upper bound of forecasting value. $z_{\alpha/2}$ represents standard Gaussian distribution, which depends on the confidence level.

3. Hybrid Forecasting Model

The proposed hybrid method mainly has three stages in PI construction. Those stages are historical PV output power series decomposition stage by using EEMD, the components construction stage utilizing SE and the construction stage by RVM. This part is divided into five sections. The first section is to introduce the principle of sample selection. The second section is to describe the decomposition of the data using EEMD and the third section is to demonstrate the reconstruction of components using SE. In the last section, the analysis of the RVM method and the corresponding flow chart as well as the pseudo-code program are given.

3.1. Sample Selection

For the sake of validating the forecasting ability of the method proposed in this paper, the PV output power simulation data of a PV power plant in Jiangsu province from July 2011 to June 2012 was obtained. Considering the different sunrise and sunset time in each season, and in order to ensure that the data obtained has value, only 10 h data from 8:00 to 17:00 was taken. If different seasons are

selected, then the sunrise and sunset time of different seasons are different. In order to unify the data, 8:00–17:00 time period was selected. Otherwise, the changes of weather have massive impacts on the PV output power. By comparing the historical output power curve with the meteorological curve, it can be found that the meteorological conditions have a great influence on the PV power output. In order to ensure the consistency of the same kind of data and to predict the PV output power more accurately, the PV historical output power data was divided into three types (sunny days, cloudy days and rainy days) according to the numerical weather prediction (NWP). The photovoltaic historical output power was classified according to the weather type, and the model prediction was respectively carried out on the photovoltaic historical output power. Using the EEMD to decompose the historical PV output power. The forecasting model was developed respectively. The historical photovoltaic output power data of 6 h to be predicted and the NWP at the time to be predicted were used as the input of the model. The model in this paper was a rolling prediction model. For different time to be predicted, the input data was updated online and in real time.

3.2. Decomposing the Classified PV Output Power Using EEMD

While PV output power contains randomness and volatility with the influence of weather changes and other factors, the result of direct forecasting would have a large error. For the sake of enhancing the forecasting results, it is essential to preprocess the original data. In the performed comparison, the EEMD shows better noise robustness and decomposing result than other decomposition algorithms. In this paper, the PV output power was decomposed by using EEMD, and some new components were achieved. For example, Figure 1 shows the decomposition results of a sunny-day PV output power data by applying EEMD.

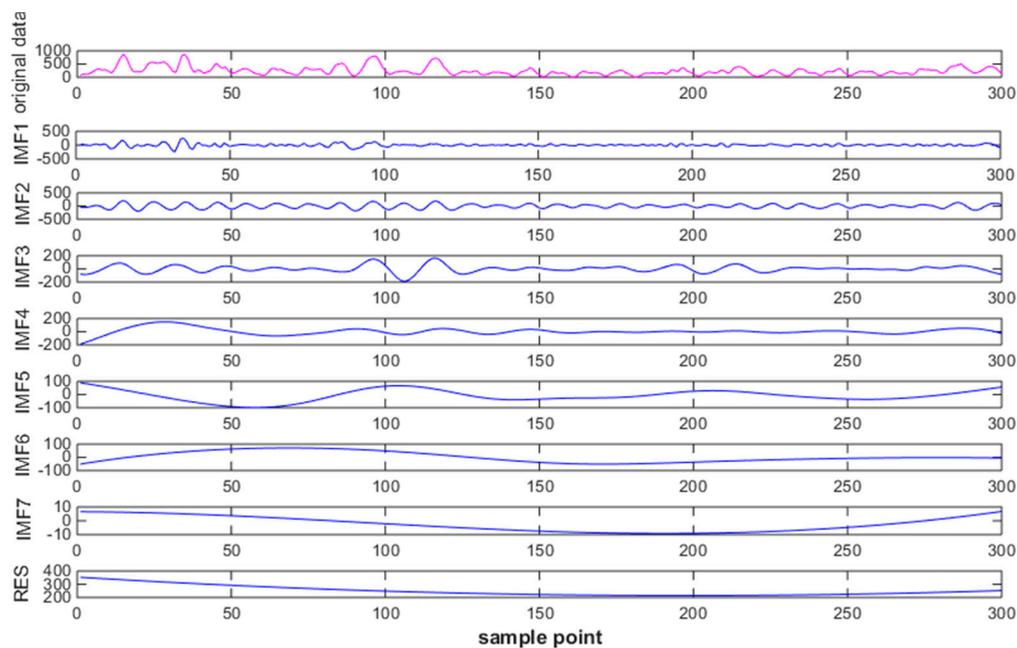


Figure 1. Decomposition results by ensemble empirical mode decomposition (EEMD).

3.3. Reconstructing the New Components Using SE

As it can be seen from Figure 1, there was a similar trend for some components. If these components are highly similar, the value of the sample entropy between them will be small. Therefore, the rules to reconstruct the new components based on SE are as presented as follows:

(1) The sample entropy of the given PV data sequence, IMF components and RES component were calculated.

(2) The components with obviously lower sample entropy value than that of the given sequence could form the trend component.

(3) The components with obviously higher sample entropy value than that of the given sequence could form the random component.

(4) The detail component's sample entropy value was within a given threshold of θ around the given sequence. In this paper, we chose $\theta = 0.7$.

Figure 2 gives the trend graph of the new components after reconstruction.

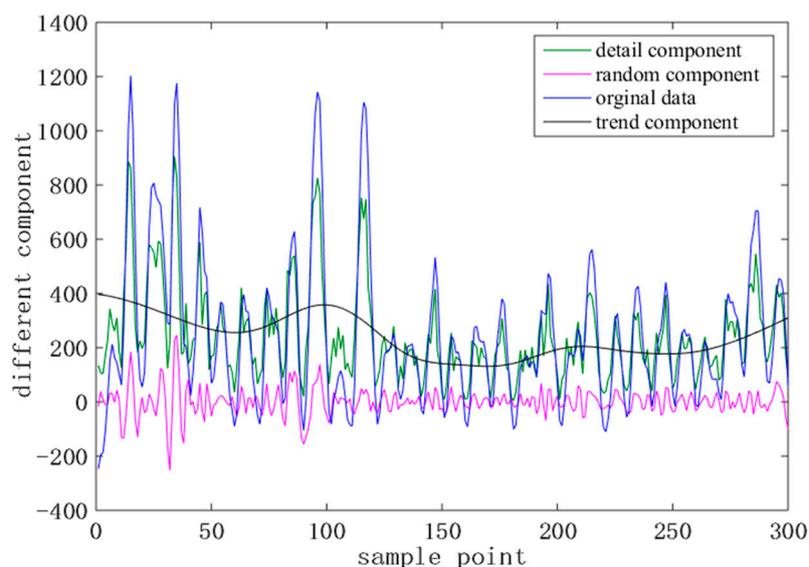


Figure 2. Trend graph of each new component.

It can be obviously noticed from Figure 2 that the three components (trend, detail and random) have their own typical features. With respect to the trend component, it can roughly reflect the overall fluctuation of the original PV power sequence. Similarly, for the detail component, it can characterize the detailed fluctuations of the original PV power sequence. Considering the random component, it represents the fluctuations caused by other factors, which cannot to be explicitly described. Table 1 shows the composition of each new component.

Table 1. Composition of each component.

New Component	Trend Component	Detail Component	Random Component
IMFs and RES	IMF6, IMF7, IMF8, RES	IMF1, IMF3, IMF4, IMF5	IMF2

For further simplification of the calculation, the forecasting interval was reduced. The trend component was selected for point forecasting, the detail and random components were selected for interval forecasting. Then, the result of the different component forecasting was superposed, the interval forecasting at a certain degree of confidence was obtained and the optimal prediction was realized.

3.4. Kernel Function of RVM

RVM is a pattern recognition as well as regression forecasting method, which is based on kernel function, the kernel implements non-linear transformation among plurality of feature spaces. The basic method of mixed kernel is to combine plurality of kernels having different characteristics together with a certain proportion, and optimizes the combined kernel function so as to have better performance. Considering that RVM has the advantages of less limitation of kernel function selection and the excellent properties of RBF kernel in solving local fluctuations and polynomial kernel in dealing with

global fluctuations, the combination of the global kernel of polynomial kernel and the typical local kernel of RBF kernel is used for short-term PV output power interval forecasting so as to obtain better forecasting results. The hybrid kernel is shown as [13,28]

$$K(x, y) = \theta G(x, y) + (1 - \theta)P(x, y) \quad (14)$$

$$G(x, y) = \exp\left(-\frac{\|x - y\|^2}{\sigma^2}\right) \quad (15)$$

$$P(x, y) = (x \cdot y) = (x \cdot y + 1)^2 \quad (16)$$

where $G(x, y)$ is the Gaussian kernel function, $P(x, y)$ is the binomial kernel function, θ is the weight of the kernel function and σ is the kernel width. θ and σ are the parameters that need to be optimized. In this paper, the optimal values of θ and σ are obtained by using the method of grid search [36].

3.5. Evaluating Indicator

There are many evaluation indicators for the forecasting, an evaluation index different from the well-known point forecasting, such as MAPE and RMSE. The following evaluation indicators were used in this paper.

(1) Mean absolute percentage error

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_{for} - y_{tru}}{y_{tru}} \right| \times 100\% \quad (17)$$

where y_{for} is the value of forecasting, y_{tru} is the actual value of sample and N represents the number of the sample.

(2) Forecasting interval coverage percentage

$$\text{FICP}^{(1-\beta)} = \frac{1}{N} \xi^{(1-\beta)} \times 100\% \quad (18)$$

where N denotes the number of the sample, ξ is the number of the actual PV output power within the interval under the level $1 - \beta$.

(3) Forecasting interval average width

$$\text{FI AW}^{(1-\beta)} = \frac{1}{N} \sum_{i=1}^N \frac{U^\beta - L^\beta}{y_{tru}} \quad (19)$$

where N represents the number of the sample, y_{tru} is the actual value of the sample, U^β is the upper boundary and L^β is the lower boundary under the level $1 - \beta$.

This paper proposed a new EEMD-SE-RVM method used for the PV output power short-term interval forecasting. A simplified pseudo-algorithm that summarizes this process is provided in Algorithm 1.

Algorithm 1. PV power forecast

```

1: PV power data and climate data
2: Divide the data into three categories (sunny, cloudy and rainy)
3: Using EEMD decompose PV output power into IMFs and RES
4: Input: IMFs and RES;
5: if SE  $\theta < 0.7$  then,
6:   constitute the trend component
7: end if
8: if SE  $\theta > 0.7$  then,
9:   constitute the random component
10:  else constitute the detail component
11: end if
12: while trend component do
13:   RVM point forecasting
14: while detail component and random component do
15:   RVM interval forecasting
16: end while

```

The EEMD method has better performance used in the interval forecast by eliminating the mode mixing problem, which exists in the EMD method. However, prediction interval forecast based on conventional EEMD is still influenced by the high complexity. The proposed method uses SE to analyze the decompositions so that the complexity is reduced. According to the analysis above, SE recombined the decomposition into trend, detail and random components to optimize the forecasting method. The trend component, which is smoother and steadier, was used to achieve point forecasting, and the detail component and random component were difficult to be used in the conventional point forecast method because of the uncertainty and non-stationary. The method that achieved point and interval forecasts respectively could guarantee better performance by reducing the numerical value fluctuation.

4. Case Study

In this part, the PV data of Jiangsu photovoltaic power station from July 2011 to June 2012 was used to test the accuracy and effectiveness of the EEMD-SE-RVM model proposed in this paper. The installed capacity of this PV plant was 30 MW, consisting of 28 PV arrays of 1.09 MW. The data were collected once an hour and 24 times a day. What is collected is the instantaneous value of PV output power at the current time. The prediction date was randomly selected and the data before the prediction date was used as the training data of the model.

For the sake of validating the interval forecasting effect of the model proposed in this paper under different confidence levels, two confidence levels of 90% and 60% were chosen as examples. Figures 3–8 depict the results in different case interval forecasting. In this paper, three common indices forecasting interval coverage percentage (FICP), forecasting interval average width (FIAW) and mean absolute percentage error (MAPE) were used to assess the effect of the interval forecasting [24,27]. Tables 2–4 give the different case interval forecasting results and analysis.

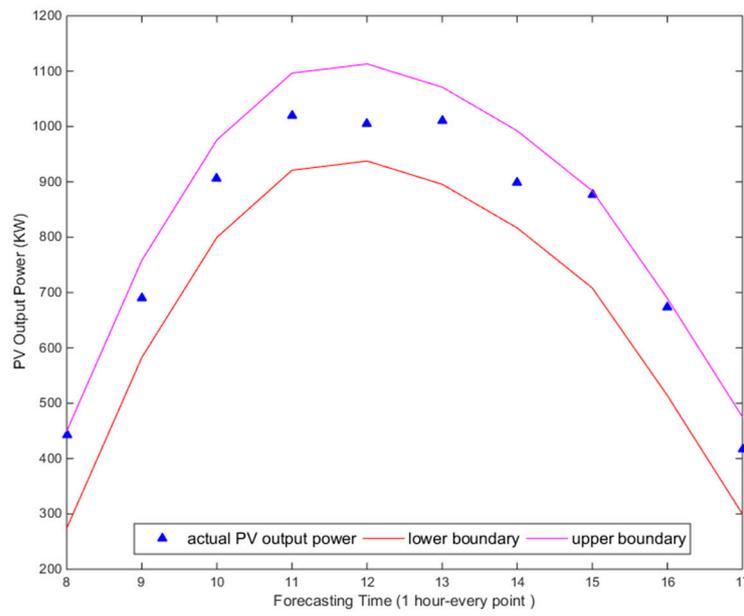


Figure 3. Interval forecasting results in a sunny day with the 90% confidence level.

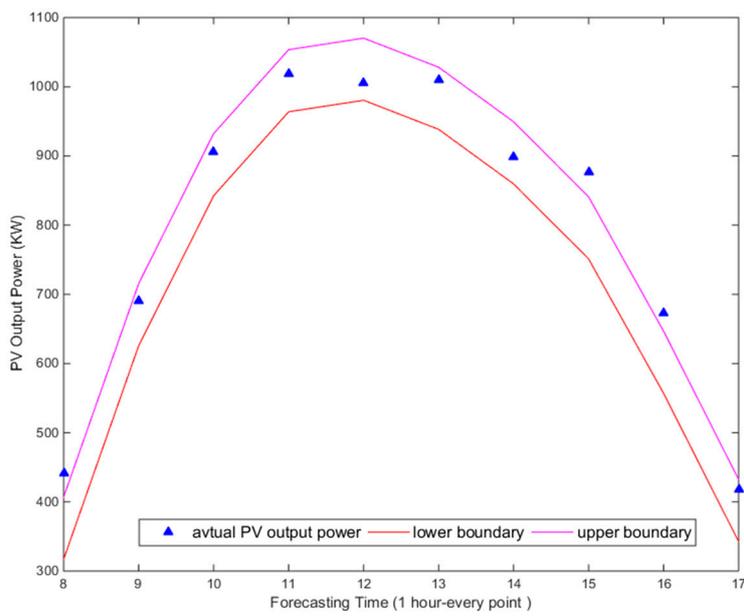


Figure 4. Interval forecasting results in a sunny day with the 60% confidence level.

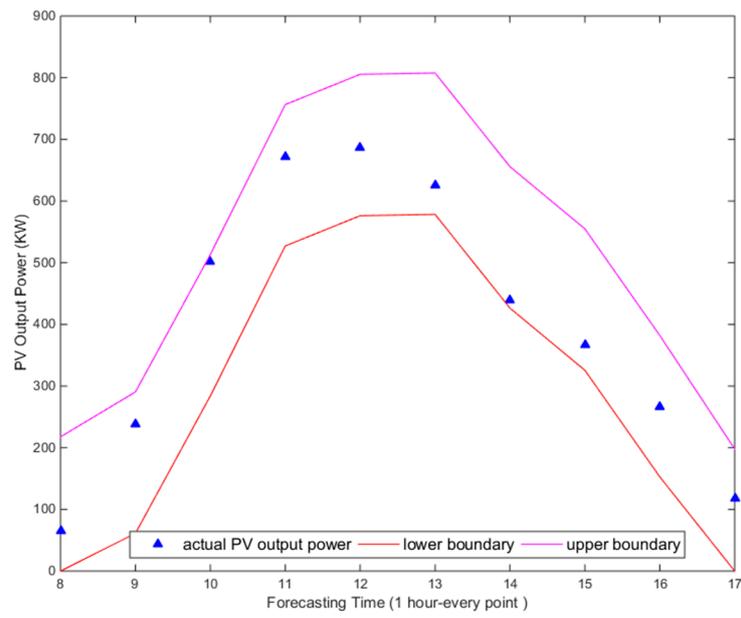


Figure 5. Interval forecasting results in a cloudy day with the 90% confidence level.

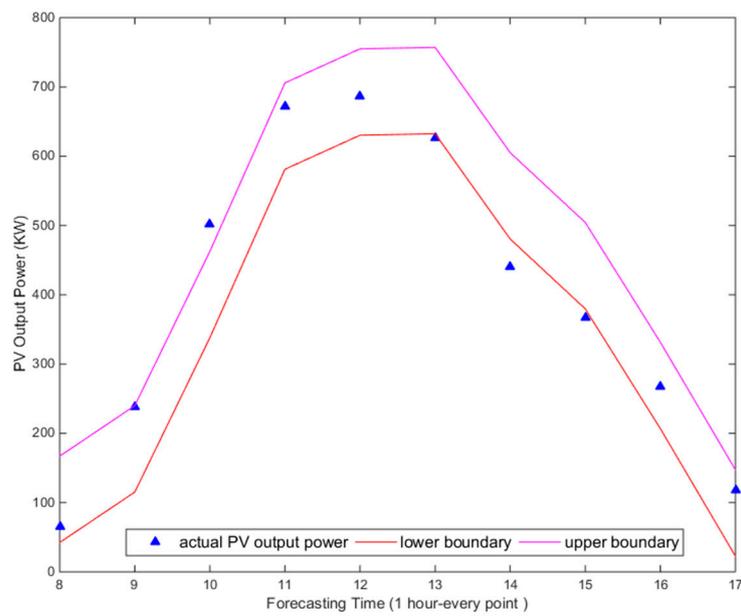


Figure 6. Interval forecasting results in a cloudy day with the 60% confidence level.

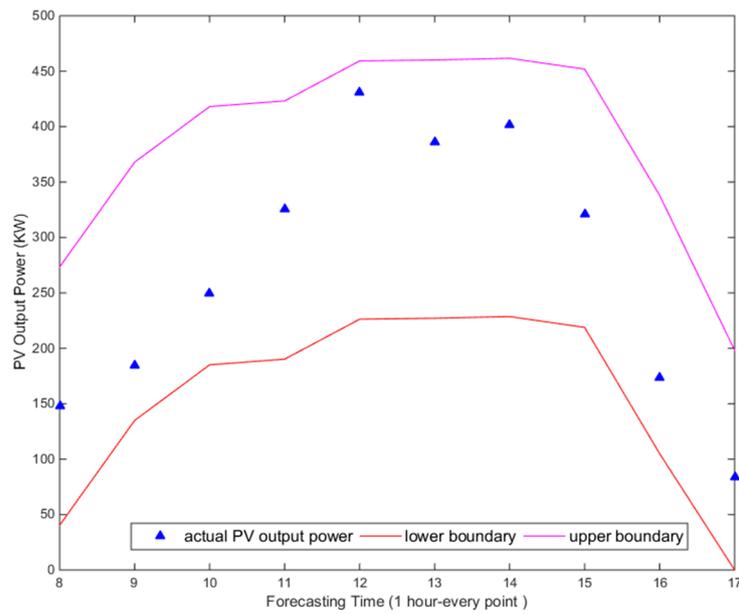


Figure 7. Interval forecasting results in a gloomy day with the 90% confidence level.

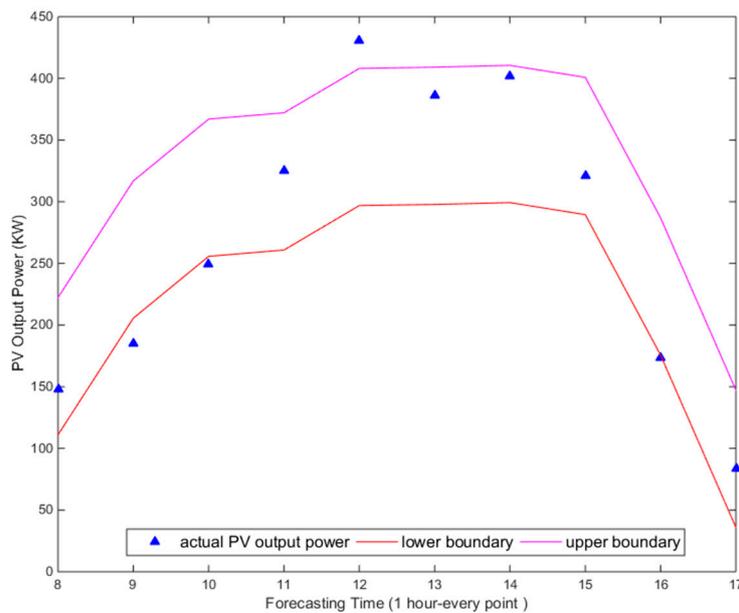


Figure 8. Interval forecasting results in a gloomy day with the 60% confidence level.

Table 2. Interval forecasting results of EEMD-sample entropy (SE)-relevance vector machine (RVM) mode in a sunny day.

	90% Confidence Level	60% Confidence Level
MAPE	4.05%	4.05%
FICP	100%	70%
FIAW	0.2441	0.1249

Table 3. Interval forecasting results of EEMD-SE-RVM model in a cloudy day.

	90% Confidence Level	60% Confidence Level
MAPE	19.08%	19.08%
FICP	100%	60%
FIAW	0.9904	0.5383

Table 4. Interval forecasting results of EEMD-SE-RVM model in a gloomy day.

	90% Confidence Level	60% Confidence Level
MAPE	14.91%	14.91%
FICP	100%	60%
FIAW	1.1051	0.5279

To prove the superiority of the method proposed in this paper, the RVM model, EMD-RVM model and EEMD-RVM model were also used for the same PV output power short-term interval forecasting, respectively. In this case, the forecasting results in the sunny day were chosen for example. In this paper, three evaluation indexes FICP, FIAW and MAPE and model running time were used to evaluate the effect of interval prediction. In Table 5, the results at 90% confidence level of four different models were provided.

Table 5. Comparison of forecasting effect among four models.

Forecasting Model	FICP	FIAW	MAPE	Running Time
RVM [7]	90%	0.2595	7.08%	19 s
EMD-RVM	90%	0.2502	5.19%	126 s
EEMD-RVM [33]	100%	0.2513	4.96%	97 s
EEMD-SE-RVM	100%	0.2441	4.05%	58 s

On the other hand, for more evaluation of the adaptability to different PV output power data of this proposed model, the other forecast days in different seasons were considered. For example, the date of 6 August 2011, 30 October 2011, 14 May 2012 and 19 March 2012 were selected stochastically. According to the four days original PV output power data, the probability of one hour-ahead of the PV output power in these four days at 90% confidence was predicted, and the results are illustrated in Figure 9. At the same time, in Table 6, the results of evaluating the indicators FICP, FIAW and MAPE are given.

Table 6. Comparison of indices results among four different days.

Forecast Day	FICP	FIAW	MAPE
6 August 2011	100%	0.2871	5.73%
30 October 2011	89.76%	0.2249	3.99%
14 March 2012	90%	0.1840	5.05%
19 May 2012	90%	0.1825	3.70%

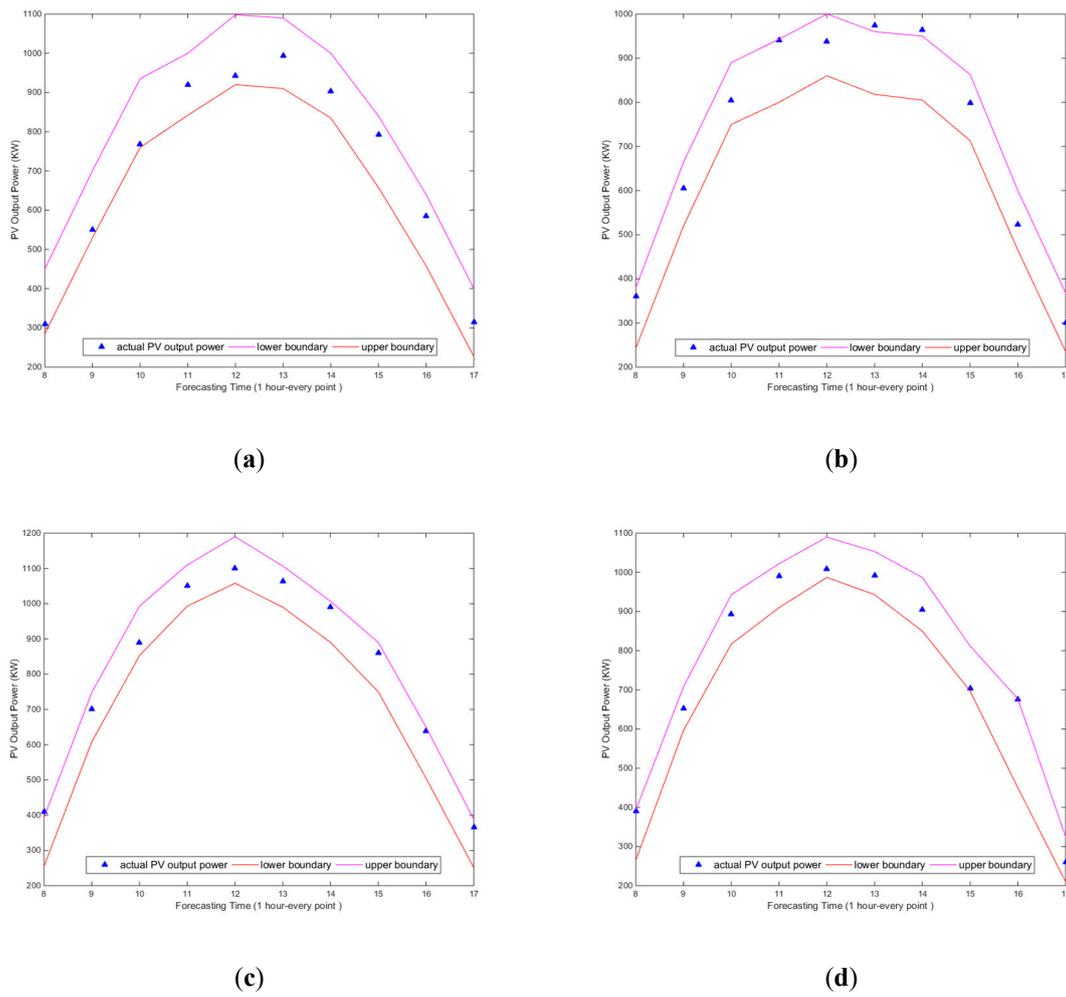


Figure 9. Interval forecasting results of four days for a PV power plant. (a) 6 August 2011; (b) 30 October 2011; (c) 14 March 2012 and (d) 19 May 2012.

Taking sunny days as an example, the short-term interval prediction of two different time scales was carried out under a 90% confidence level are depicted in Figure 10. At the same time, in Table 7, the results of evaluating the indicators FICP, FIAW and MAPE are given.

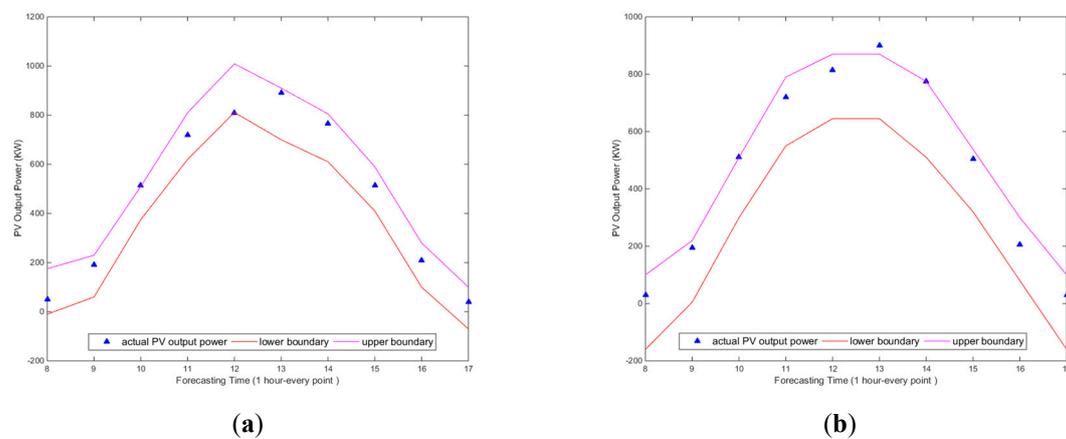


Figure 10. Results under two different circumstances (a) Hour-ahead and (b) day-ahead.

Table 7. Comparison of indices results between two different days.

Circumstance	FICP	FIAW	MAPE
Day-ahead	90%	0.7208	7.90%
Hour-ahead	90%	0.5097	4.86%

It can be clearly noted from the comparison results that the forecasting effects obtained by the proposed method were better than the other methods. Furthermore, the superiority and wide adaptability of this proposed model were fully confirmed based on the above comparison.

5. Conclusions

Firstly, considering the influence of different meteorological conditions on the output power of PV, the original PV output power data has been classified into three categories. Strong theoretical basis in addition to noise robustness are some of the advantages of EEMD. Those features overcome the drawbacks that the wavelet analysis requires, which are the artificial selection of the basic functions and the mode mixing phenomenon of EMD. Consequently, the original PV output power achieves better decomposition by the use of EEMD. Secondly, the use of SE excavates the correlation among the components as well as reduces the model complexity, which creates contributions to enhance the running efficiency. Thirdly, the hybrid kernel RVM method was implemented to achieve the PV output power short-term interval forecasting. In the part of illustrative results, comparing the EEMD-SE-RVM with other models, the obtained MAPE and FIAW indices had better values than other models, and the FICP of the proposed model was higher than that obtained from the compared models. In this paper, the proposed hybrid model not only improved the forecasting precision, but also enhanced the interval coverage rate, and at the same time, reduced the width of the prediction interval, which made it suitable for practical application on other renewable energies output power forecasting.

Author Contributions: S.W. conceived and designed the experiment and wrote the original manuscript. Y.S. performed the experiments and evaluated the data. Y.Z., R.J.M. and D.H. reviewed and proofread the manuscript. All authors read and proofread the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Key R&D Program of China under Grant 2018YFB0904200 (Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption), and in part by the eponymous Complement S&T Program of State Grid Corporation of China under Grant SGLNDK00KJJS1800266.

Conflicts of Interest: The authors state that there is no conflict of interest.

References

- Asrari, A.; Wu, T.; Ramos, B. A hybrid algorithm for short-term solar power prediction-sunshine state case study. *IEEE Trans. Sustain. Energy* **2017**, *8*, 582–591. [[CrossRef](#)]
- Yang, M.; Zhang, L.; Cui, Y.; Zhou, Y.; Chen, Y.; Yan, G. Investigating the wind power smoothing effect using set pair analysis. *IEEE Trans. Sustain. Energy* **2019**. [[CrossRef](#)]
- Wan, C.; Zhao, J.; Song, Y.; Xu, Z.; Lin, J.; Hu, Z. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE J. Power Energy Syst.* **2015**, *1*, 38–46. [[CrossRef](#)]
- Bacher, P.; Madsen, H.; Nielsen, H.A. Online short-term solar power forecasting. *Sol. Energy* **2009**, *83*, 1772–1783. [[CrossRef](#)]
- Li, Y.; Su, Y.; Shu, L. An ARMAX model for forecasting the power output of a grid connected photovoltaic system. *Renew. Energy* **2014**, *66*, 78–89. [[CrossRef](#)]
- Li, S.; Goel, L.; Wang, P. An ensemble approach for short-term load forecasting by extreme learning machine. *Appl. Energy* **2016**, *170*, 22–29. [[CrossRef](#)]
- Jang, H.S.; Bae, K.Y.; Park, H.-S.; Sung, D.K. Solar Power Prediction Based on Satellite Images and Support Vector Machine. *IEEE Trans. Sustain. Energy* **2016**, *7*, 1255–1263. [[CrossRef](#)]

8. Amrouche, B.; Le, P. Artificial neural network based daily local forecasting for global solar radiation. *Appl. Energy* **2014**, *130*, 333–341. [[CrossRef](#)]
9. Bai, W.; Zhou, Q.; Li, T.; Li, H. Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation. *IEEE Trans. Cybern.* **2019**. [[CrossRef](#)]
10. Mellit, A.; Pavan, A.M.; Lughi, V. Short-term forecasting of power production in a large-scale photovoltaic plant. *Sol. Energy* **2014**, *105*, 401–413. [[CrossRef](#)]
11. Shi, J.; Lee, W.-J.; Liu, Y.; Yang, Y.; Wang, P. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1064–1069. [[CrossRef](#)]
12. Chen, Y.; Xu, P.; Chu, Y.; Li, W.; Wu, Y.; Ni, L.; Bao, Y.; Wang, K. Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings. *Appl. Energy* **2017**, *195*, 659–670. [[CrossRef](#)]
13. He, Y.; Liu, R.; Li, H.; Wang, S.; Lu, X. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. *Appl. Energy* **2017**, *185*, 254–266. [[CrossRef](#)]
14. Zhukov, A.; Tomin, N.; Sidorov, D.; Panasetsky, D.; Spirayev, V. A hybrid artificial neural network for voltage security evaluation in a power system. In Proceedings of the 2015 5th International Youth Conference on Energy (IYCE), Pisa, Italy, 27–30 May 2015; pp. 1–7.
15. Kurbatsky, V.G.; Sidorov, D.N.; Spiryaev, V.A.; Tomin, N.V. The hybrid model based on Hilbert-Huang Transform and neural networks for forecasting of short-term operation conditions of power system. In Proceedings of the IEEE Trondheim PowerTech, Trondheim, Norway, 19–23 June 2011; pp. 1–7.
16. Huang, C.; Kuo, C.; Chen, S.; Yang, S. One-day-ahead hourly forecasting for photovoltaic power generation using an intelligent method with weather-based forecasting models. *IET Gener. Transm. Distrib.* **2015**, *9*, 1874–1882. [[CrossRef](#)]
17. Liu, J.; Fang, W.; Zhang, X.; Yang, C. An improved photovoltaic power forecasting model with the assistance of aerosol index data. *IEEE Trans. Sustain. Energy* **2015**, *6*, 434–442. [[CrossRef](#)]
18. Zheng, H.; Yuan, J.; Chen, L. Short-term load forecasting using EMD-LSTM neural networks with a XGboost algorithm for feature importance evaluation. *Energies* **2017**, *10*, 1168. [[CrossRef](#)]
19. Li, Y.; Zou, Y.; Tan, Y.; Cao, Y.; Liu, X.; Shahidehpour, M.; Tian, S.; Bu, F. Optimal Stochastic Operation of Integrated Low-Carbon Electric Power, Natural Gas, and Heat Delivery System. *IEEE Trans. Sustain. Energy* **2018**, *9*, 273–283. [[CrossRef](#)]
20. Yang, C.; Thatte, A.; Xie, L. Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. *IEEE Trans. Sustain. Energy* **2015**, *6*, 104–112. [[CrossRef](#)]
21. Yu, X.; Zhang, W.; Zang, H.; Yang, H. Wind power interval forecasting based on confidence interval optimization. *Energies* **2018**, *11*, 3336. [[CrossRef](#)]
22. Nam, S.; Hur, J. Probabilistic forecasting model of solar power outputs based on the naive Bayes classifier and kriging models. *Energies* **2018**, *11*, 2982. [[CrossRef](#)]
23. Wan, C.; Lin, J.; Song, Y.; Xu, Z.; Yang, G. Probabilistic forecasting of photovoltaic generation: An efficient statistical approach. *IEEE Trans. Power Syst.* **2017**, *32*, 2471–2472. [[CrossRef](#)]
24. Hu, M.; Hu, Z.; Yue, J.; Zhang, M.; Hu, M. A novel multi-objective optimal approach for wind power interval prediction. *Energies* **2017**, *10*, 419. [[CrossRef](#)]
25. Wan, C.; Xu, Z.; Pinson, P.; Dong, Z.Y.; Wong, K.P. Probabilistic forecasting of wind power generation using extreme learning machine. *IEEE Trans. Power Syst.* **2014**, *29*, 1033–1044. [[CrossRef](#)]
26. Bludszuweit, H.; Dominguez-Navarro, J.; Lombart, A. Statistical analysis of wind power forecast error. *IEEE Trans. Power Syst.* **2018**, *23*, 983–991. [[CrossRef](#)]
27. Wang, S.; Sun, Y.; Zhai, S.; Hou, D.; Wang, P.; Wu, X. Ultra-short-term wind power forecasting based on deep belief network. In Proceedings of the 38th Chinese Control Conference, Guangzhou, China, 27–30 July 2019; pp. 7479–7483.
28. Yang, M.; Chen, X.; Du, J.; Cui, Y. Ultra-short-term multistep wind power prediction based on improved EMD and reconstruction method using Run-Length analysis. *IEEE Access* **2018**, *6*, 31908–31917. [[CrossRef](#)]
29. Li, Y.; Xiao, J.; Chen, C.; Tan, Y.; Cao, Y. Service Restoration Model with Mixed-Integer Second-Order Cone Programming for Distribution Network with Distributed Generations. *IEEE Trans. Smart Grid* **2019**, *10*, 4138–4150. [[CrossRef](#)]

30. Zhukov, A.; Sidorov, D.N.; Foley, A.M. Random Forest Based Approach for Concept Drift Handling. In Proceedings of the 5th International Conference on Analysis of Images, Social Networks, and Texts (AIST), Yekaterinburg, Russia, 7–9 April 2016; Volume 661, pp. 69–77.
31. Yang, M.; Huang, X. Ultra-short-term prediction of photovoltaic power based on periodic extraction of PV energy and LSH algorithm. *IEEE Access* **2018**, *6*, 51200–51205. [[CrossRef](#)]
32. Sun, Y.; Wang, P.; Zhai, S.; Hou, D.; Wang, S.; Zhou, Y. Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution. *Wind Energy* **2019**. [[CrossRef](#)]
33. Wang, T.; Zhang, M.; Yu, Q.; Zhang, H. Comparing the applications of EMD and EEMD on time–frequency analysis of seismic signal. *J. Appl. Geophys.* **2012**, *83*, 29–34. [[CrossRef](#)]
34. Feng, C.; Cui, M.; Hodge, B.; Zhang, J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. *Appl. Energy* **2017**, *190*, 1245–1257. [[CrossRef](#)]
35. Pan, C.; Tan, J. Day-ahead hourly forecasting of solar generation based on cluster analysis and ensemble model. *IEEE Access* **2019**, *7*, 112921–112930. [[CrossRef](#)]
36. Huang, Q.; Mao, J.; Liu, Y. An improved grid search algorithm of SVR parameters optimization. In Proceedings of the 2012 IEEE 14th International Conference on Communication Technology, Chengdu, China, 9–11 November 2012; pp. 1022–1026.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).