energies

MDPI

Article

Optimal Kernel ELM and Variational Mode Decomposition for Probabilistic PV Power Prediction

Xiaomei Wu [1], Chun Sing Lai [1,2,3,*], Chenchen Bai [4], Loi Lei Lai [1,*], Qi Zhang [1,5] and Bo Liu [1]

[1] Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou 510006, China; epxm_wu@gdut.edu.cn (X.W.); zhangqi@gdzq.csg.cn (Q.Z.); 2111704170@mail2.gdut.edu.cn (B.L.)
[2] Brunel Institute of Power Systems, Department of Electronic and Computer Engineering, Brunel University London, London UB8 3PH, UK
[3] School of Civil Engineering, Faculty of Engineering and Physical Sciences, University of Leeds, Leeds LS2 9JT, UK
[4] Department of Mathematics, Southern University of Science and Technology, Shenzhen 518055, China; 11811833@mail.sustech.edu.cn
[5] Zhaoqing Power Supply Bureau, Guangdong Power Grid Co., Ltd., Zhaoqing 526060, China
* Correspondence: chunsing.lai@brunel.ac.uk (C.S.L.); l.l.lai@gdut.edu.cn (L.L.L.)

Received: 29 May 2020; Accepted: 10 July 2020; Published: 13 July 2020

Abstract: A probabilistic prediction interval (PI) model based on variational mode decomposition (VMD) and a kernel extreme learning machine using the firefly algorithm (FA-KELM) is presented to tackle the problem of photovoltaic (PV) power for intra-day-ahead prediction. Firstly, considering the non-stationary and nonlinear characteristics of a PV power output sequence, the decomposition of the original PV power output series is carried out using VMD. Secondly, to further improve the prediction accuracy, KELM is established for each decomposed component and the firefly algorithm is introduced to optimize the penalty factor and kernel parameter. Finally, the point predicted value is obtained through the summation of predicted results of each component and then using the nonlinear kernel density estimation to fit it. The cubic spline interpolation algorithm is applied to obtain the shortest confidence interval. Results from practical cases show that this probabilistic prediction interval could achieve higher accuracy as compared with other prediction models.

Keywords: photovoltaic power output prediction; variational mode decomposition; firefly algorithm; kernel extreme learning machine; probabilistic prediction interval

1. Introduction

Driven by the global severe condition of fossil fuel depletion and growing environmental pollution, as an environmentally friendly renewable energy, photovoltaic (PV) generation is an exemplar of widely used power generation methods in the renewable energy industry. PV generation is susceptible to surface solar irradiance, and its output is strongly random, which challenges frequency regulation, peak load regulation, and system reserve. With the increase in grid integration capacity, the randomness of PV generation brings more and more risks to power system scheduling and operation. More accurate prediction of PV power can provide a reliable basis for power grid dispatching decisions [1]. It is of significant importance to ensure system security, stability, and optimal operation.

There are four main techniques to predict PV power output, namely physical, artificial intelligence (AI), statistical, and hybrid approaches [2–4]. The physical method uses numerical weather prediction (NWP) data and measured data. The statistical method establishes a relationship between historical data and forecasted variables based on data-driven formulations such as regression models [5,6], time series [7], and cluster analysis of clearness index [8,9]. For AI methods, there are artificial

neural network (ANN) [10–13], support vector machine (SVM) [14], and fuzzy logic. A hybrid approach is a mixture of the above approaches. To determine the greatest influencing variables for PV power prediction, multiple linear regression (MLR) and ANN models were established in [15]. A dynamic artificial neural network is presented in [16]. Reference [17] proposed a method based on principal component analysis, support vector regression (SVR), and weather forecast data to obtain one-day-ahead hourly regional forecasts of PV power when regional information is available. Group least square support vector machine (GLSSVM) is an hybrid algorithm that was developed to forecast the PV output power at multiple time horizons up to 24 h [18]. An extreme learning machine (ELM) could be developed by a single-hidden layer feedforward neural network (SLFN) structure, whose input weights and hidden layer biases are set arbitrarily. The random process of parameters initialization impacts on the forecasting model. Reference [19] added a regularization term to enhance the forecasting accuracy of the ANN and greatly reduced the randomness by combining multiple ELMs. The authors have demonstrated that the randomness is significantly minimized which led to an increased forecasting accuracy as compared with a single ELM.

These methods are deterministic, that is, to give an absolute value of a given prediction time. However, the PV power generation can be influenced by meteorological features and has substantial variability and intermittency. When the meteorological factors change obviously during the forecasting period, the deterministic single point prediction may not be able to achieve the ideal precision when PV output fluctuates greatly. More importantly, deterministic methods cannot express the probabilistic reliability of the predicted results. With the forecast result, the dispatcher has no idea about the possible situation and its possibility in the future. In contrast, the probabilistic prediction interval (PI) can give all possible PV outputs and the probability of their occurrence at the next instance, which covers additional prediction information. Therefore, a probabilistic prediction is more valuable in arranging power system operation and regulation at a reasonable risk level. There are works of literature for wind power forecasting based on PI forecasting [20–22]. However, there are limited works of literature available on PI forecasting for PV power [23].

It is observed from the literature survey that improved models are needed for a PV output prediction interval with more accuracy. A single PV power prediction model has limitations, and it is hard to obtain the best prediction performance. Table 1 depicts a summary of the state-of-the-art research. The hybrid forecast model is established in this work to enhance the overall prediction performance. As the PV output is closely related to weather conditions, the authors introduced a novel prediction interval method for PV power output where the clustering algorithm is used to classify the weather type, the models of PV output are established for various weather conditions, and the prediction intervals can be calculated under different weather types. The paper follows the following structure: Section 2 gives brief descriptions of variational mode decomposition (VMD) and kernel extreme learning machine using the firefly algorithm (FA-KELM). Section 3 presents the influential factors for PV output and chooses the most important ones for prediction. Section 4 formalizes the problem, how kernel density is estimated, and how a probabilistic prediction interval model is formed. The PV generation for various weather conditions is forecasted in Section 5, in which data collection is illustrated, and forecasting results with future work are discussed. Finally, Section 6 gives the conclusions.

Table 1. Literature review on recent works.

	This Work	Majumder, et al. [3]	YutongHan, et al. [24]	LI G H, et al. [25]	Inés M. Galván, et al. [26]	Qiang Ni, et al. [23]	Xiuyuan Yang, et al. [27]	Trapero J R. [28]	Can Wan, et al. [29]
Country	USA	India and USA	China	Belgium	USA	Singapore	China	Spain	Denmark
Research context	Presenting a hybrid technique implementing VMD and KELM improved by the firefly algorithm for PV power prediction interval.	Presenting a hybrid technique for implementing VMD and different low-level robust KELMs: for solar radiation prediction under different weather conditions such as sunny, rainy, and foggy days in different time horizons.	Building a seasonal multi-model for deterministic prediction of PV power based on an ELM. The deterministic prediction error is fitted by KDE to complete the PV power interval pre -diction.	Proposing a new type of hybrid forecasting model: firefly algorithm and VMD to optimize BP neural network. The model is used to forecast sunspot numbers.	Proposing a multi-objective particle swarm optimization (PSO) for optimizing both coverage probability of the interval and prediction intervals simultaneously. The ANN is proposed to develop complex non-linear models with the outputs to be the prediction interval's lower and upper limits.	Proposing a new ensemble method: Lower upper bound estimation (LUBE) and ELM for short- term PV power forecasting.	Proposing a day-ahead PV power forecasting method with alike cloud space fusion according to missing historical data mining.	Presenting KDE methods, volatility forecasting models and the hybrid of the aforementioned models to enhance the prediction intervals results.	Using ELM and quantile regression to develop a PV power generation prediction interval model
Prediction interval	Day-ahead and intra-day-ahead	15-min-, 1-h-, and 1-day-ahead	A season	Monthly average	1 day	5 min	1 day	1 h	5 min
Interval forecasting	Yes	No	Yes	No	Yes	Yes	No	Yes	Yes
Model considered	VMD KELM Firefly algorithm	VMD Low-rank robust KELM Morlet wavelet kernel	ELM KDE	VMD Backpropagation neural network and firefly algorithm	Multi-objective PSO Neural networks	ELM LUBE	Markov chain Cloud interval fusion	Kernel density estimation Volatility forecasting models	ELM

Table 1. *Cont.*

	This Work	Majumder, et al. [3]	YutongHan, et al. [24]	LI G H, et al. [25]	Inés M. Galván, et al. [26]	Qiang Ni, et al. [23]	Xiuyuan Yang, et al. [27]	Trapero J R. [28]	Can Wan, et al. [29]
Findings	The hybrid use of VMD and FA-KELM shows to be an enhanced technique to develop optimal prediction intervals.	The solar prediction error reduces in chaotic variations when employing a weighted kernel matrix formulation based on the non-parametric kernel density estimation (KDE) technique.	KDE can successfully reflect the forecasting errors distribution. Analysis of PV power and variation features show seasonal differences, so it is feasible to build multiple models by seasons and enhance their forecasting accuracy.	VMD has greater robustness than EMD to study noisy signals including sunspot data. The model is effective for forecasting challenges including short-time traffic flow.	A multi-objective PSO is developed to build PIs with a multi-layer perceptron (MLP) gives comparable or enhanced results than single-objective PSO and gradient boosted regression.	An improved differential evolution (DE) algorithm is created to gain an optimal collection of weighting coefficients, and the proposed model is better than the traditional DE and PSO methods.	The proposed model has improved forecasting accuracy for incomplete historical data.	Proposed approach is a combination of the non-parametric and parametric approaches. The results show a good compromise between coverage (unconditional and conditional) and average interval width.	The proposed ELM approach performs 300 times and 140 times quicker compared to common bootstrap based neural network and non-parametric granule computing techniques, respectively.
Drawback	Future work is carried out on sensitivity study to search for a near global optimal solution.	The method is point prediction which is deterministic. However, solar radiation has strong variability and intermittence. The prediction interval may give better results.	The robustness and stability can be improved for the ELM by employing signal decomposition.	The firefly algorithm optimizes the weights and thresholds outside the neural network loop, which cannot modify the weights and thresholds by every training. It may be better to do the optimization within the BP neural network loop.	The problem of slow training speed and poor generalization ability of neural networks can be solved by using ELM.	ELM data samples have complex collinearity problems, resulting in random fluctuations in the output of the ELM model and poor stability and generalization capabilities. KELM uses the kernel function to map the input samples to the high -dimensional kernel space, which overcomes the random fluctuation of the model output caused by the random hidden layer output matrix H in ELM, and makes the model have better generalization ability.	Proposed model is for deterministic point predict -ion. Interval prediction ensures that the actual value falls within the prediction interval at a certain confidence level. The results of probabilistic interval prediction under a certain confidence level are given, and the range of possible future load values can be described.	In view of the non-stationary and non-linear characteristics of solar radiation, this work does not decompose the original solar radiation signal, such as the VMD.	The amount of data is small (three months) and does not consider the impact of the season on PV power.

2. Fundamentals of VMD and FA-KELM

2.1. Variational Mode Decomposition Principle

Signal decomposition is often used in hybrid prediction methods. The common signal decomposition techniques are wavelet packet decomposition (WPD) [30] and empirical mode decomposition (EMD) [31]. EMD can decompose signals into different frequency characteristics and solve the envelope and instantaneous frequency decomposition based on Hilbert–Huang transform.

Although Hilbert–Huang transform is an effective decomposition method, EMD also has some drawbacks, such as unavailability of a rigorous mathematical model, interpolation option, and is sensitive to sampling and noise. The authors in [31] described that EMD exhibits a lack of "sparseness" characteristics, whereas VMD could probably render a slightly higher degree of sparseness than EMD. To overcome these limitations, in 2014, an alternative multiresolution technique named variational mode decomposition (VMD) was presented [32]. VMD is a model of entirely non-recursive variational, and the modes are simultaneously determined. The VMD model looks for multiple modes and their center frequencies. The band-limited modes recreate the input signal precisely or in a least squares manner. The instantaneous frequency of each analytic signal has practical physical significance in VMD, whereas this is not the case for EMD. In this paper, the VMD is selected as the signal decomposition method. VMD aims to decompose the original series u into a series of band-limited modes u_k, where individual mode compacts a center frequency ω_k identified in the decomposition. Individual mode u_k for the bandwidth can be calculated with the procedures presented in [33].

2.2. Kernel Extreme Learning Machine

A single-hidden layer feedforward neural network (SLFN) structure is an extreme learning machine (ELM), whose input weights and hidden layer biases are chosen arbitrarily. Well-known neural network learning methods including the back propagation (BP) algorithm require the user to manually establish a significant number of training parameters, and such procedure gives a prediction output that can be of local optimum. Differently, the ELM requires to establish the number of hidden layer nodes for the model without the task to modify the bias of the hidden layer units and the network's input weights. The Moore–Penrose generalized inverse matrix theory is used to obtain the optimal output weights for ELM. Moreover, to minimize the training error, its output weights can be solved by only one step. It has the characteristic of rapid learning and great generalization capability [34–36]. However, the number of its hidden nodes is difficult to be determined and its output is of stochastic volatility. By introducing the kernel function to ELM and comparing with SVM theory, the kernel extreme learning machine (KELM) algorithm was developed. KELM enhances the algorithm's learning accuracy [37]. The detailed proof process can be referred to references [38,39].

For N random sample (x_i,t_i), where t_i is the relating target class label of the training sample x_i, $x_i \in R^n$, $t_i \in R^n$, $i = 1, 2, \ldots, n$. The ELM's output function is [39]

$$f(x) = \mathbf{h}(x)\boldsymbol{H}^{\mathrm{T}}\left(I/C + \boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\right)^{-1}\boldsymbol{T} \tag{1}$$

$\mathbf{h}(x) = [\mathbf{h}_1(x), \ldots, \mathbf{h}_L(x)]$ is the output, a row vector of the hidden layer concerning the input x, which for the relationship for the samples from the input space to the hidden layer feature space, L is the number of hidden layers, $\boldsymbol{H}$ is the hidden layer output matrix, $\boldsymbol{I}$ is unit sparse, C is penalty coefficient, and $\boldsymbol{T}$ is the output matrix of the SLFN.

If the feature mapping $\mathbf{h}(x)$ is unknown, then a kernal matrix for ELM needs to be defined. The output of the KELM model can be calculated by Equation (2) below.

$$f(x) = \mathbf{h}(x)H^{\mathrm{T}}\left(\boldsymbol{I}/C + \boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\right)^{-1}\boldsymbol{T} = \begin{bmatrix} k(x,x_1) \\ \cdots \\ k(x,x_n) \end{bmatrix}^{\mathrm{T}} (\boldsymbol{I}/C + \Omega_{\mathrm{ELM}})^{-1}\boldsymbol{T} \tag{2}$$

where $\Omega_{\mathrm{ELM}} = h(x_i) \cdot h\left(x_j\right) = k\left(x_i, x_j\right)$ is a kernel function. In Equation (2), the hidden layer feature mapping $\mathbf{h}(x)$ does not required to be defined and the function does not identify the number of hidden layers L. The kernel $k(u, v)$ that substitutes $\mathbf{h}(x)$ and L needs to be defined. The stable kernel function substitutes the ELM's arbitrate mapping and the output weight becomes robust. KELM constitutes an enhanced generalization capability than ELM.

Various kernel functions exist such as polynomial kernel, linear kernel, and radial basis function (RBF) kernel. RBF kernel shows great learning capability in practical challenges and the amount of unknown parameters is less compared to polynomial kernel. Therefore, the RBF kernel function is considered in this work.

The RBF kernel function and output weight of the KELM model is written as

$$k\left(x_i, x_j\right) = \exp\left(-g\|x_i - x_j\|^2\right) \tag{3}$$

$$\beta = \left(\boldsymbol{I}/C + \Omega_{\mathrm{ELM}}\right)^{-1}\boldsymbol{T} \tag{4}$$

where g is the kernel parameter, and β is the output weight between the output and hidden layers.

The ELM with the kernel function overcomes the shortcomings of the traditional ELM [38], but its performance is affected by the penalty coefficient C and g. Therefore, the firefly algorithm (FA) is chosen to optimize the parameters.

2.3. Optimization of KELM's Parameters through FA

Since the kernel parameter and penalty coefficient of KELM are the two major factors affecting the prediction performance, this paper selects the firefly algorithm which is characterized with fewer parameters, better stability, and convergence to optimize these two parameters for further enhancing the prediction model's accuracy. FA is a swarm optimization technique developed by simulating the flashing characteristics of fireflies in nature [40]. In the algorithm, each firefly represents a solution of the solution space and is randomly distributed in the searching space. Each firefly individual has its own brightness and neighborhood, known as the radius of the decision domain, $\left\{r_d^i \middle| 0 < r_d^i < r_s\right\}$, where r_s is the visual range of the firefly. To simplify the problem, the following three hypotheses are considered to describe FA:

(1) All fireflies are unisex. A firefly can be attracted to other fireflies of either sex;

(2) The firefly's brightness is relevant to the analytical expression of the objective function. To solve the maximization problem, brightness is considered to be a proportion to the objective function's value. Alternative forms of brightness could be established in a similar approach to the fitness function in some optimization techniques;

(3) Attractiveness is related to the firefly's brightness. The dimmer firefly will approach the brighter firefly for two flashing fireflies. Attractiveness is related to the brightness of the fireflies and will reduce with increasing distance between fireflies. A firefly will move arbitrarily in the space if there are no fireflies brighter than itself.

The details for the FA are given in [41]. The solution to the KELM based on FA optimization is shown in Algorithm 1.

Algorithm 1. Algorithm of firefly algorithm (FA) optimization for kernel extreme learning machine (KELM).

Initialize: population of fireflies x_i ($i = 1, 2, \ldots, n$), the number of iterations $t \leftarrow 1$
Set objective function $f(x)$, $x = \left(x_1, \ldots, x_d \right)^T$
Brightness $\boldsymbol{I}_i$ at x_i is determined by $f(x_i)$
Set light absorption coefficient γ, *MaxGeneration*, calculation dimension d, the attractiveness β, the randomization parameter α
while (t<*MaxGeneration*)
  for i = 1:n all n fireflies
    for j = 1:n all n fireflies
      if ($\boldsymbol{I}_j > \boldsymbol{I}_i$) **then**
      Update x_i
      $x_i \leftarrow x_i + \beta\left(x_j - x_i\right) + \alpha(rand - 0.5)$
else
end
    end for j
  end for i
  Rank the fireflies and determine the optimal solution
end while
Output the optimum C, g

3. Influential Factors of PV Output

PV output forecasting is a complex nonlinear problem. Therefore, environmental factors and weather conditions should be considered. In this paper, different weather conditions including solar radiation, wind speed, and temperature are studied. The full extent of the impact of these weather conditions on PV power output is analyzed [42–44].

3.1. PV Output for Different Weather Conditions

The PV outputs as shown in Figure 1 under the different weather conditions, such as three main weather conditions, sunny, cloudy, and rainy day in 2015, are selected from the Ashland station in Oregon, USA. We can conclude that, from Figure 1, the PV output is stable: basically, there is normal distribution on sunny days. On cloudy days, the PV output forecasting becomes more challenging with greater randomness and volatility of solar radiation. The PV output is changing all the time on rainy days, and the average output of PV power is very small. This situation will affect the safety and stability of PV power plant operation [45]. Therefore, it is essential to classify a large number of PV output historical data according to their weather conditions to enhance the prediction's accuracy.

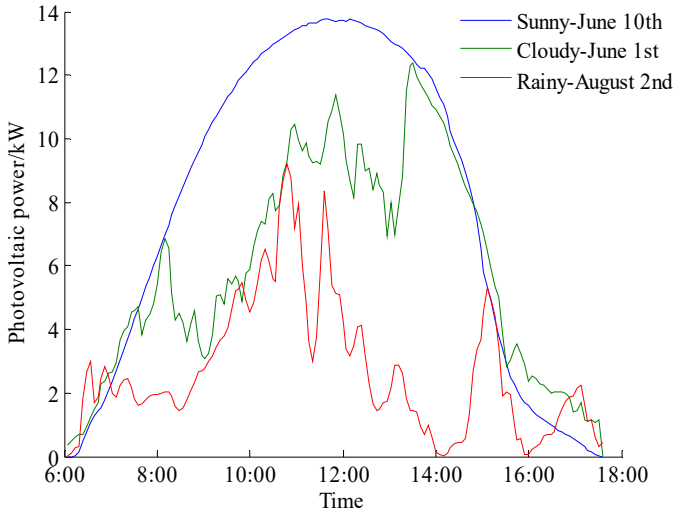


Figure 1. Photovoltaic (PV) power curves.

3.2. The Influences of Different Meteorological Factors on PV Power Output

Different meteorological factors can change the PV power output. If we take all terms into account, the complexity of prediction will be increased. Historical meteorological data encompassing solar radiation, temperature, and wind speed, and PV power outputs for the whole year of 2015 are selected, and the Kendall correlation coefficient method is applied to analyze the factors affecting PV power [46,47]. The result of the Kendall rank correlation coefficients is provided in Table 2.

Table 2. Kendall rank correlation coefficients.

Variables	Rank Correlation Coefficient			
	PV Power	**Solar Radiation**	**Temperature**	**Wind Speed**
PV power	1	0.977	0.265	0.094
Solar radiation	0.977	1	0.328	0.094
Temperature	0.265	0.328	1	0.421
Wind speed	0.094	0.094	0.421	1

As shown in Table 2, the PV power output has a maximum correlation with solar radiation, which is 0.977. The correlation between PV power output and the temperature is 0.265, which is a weak correlation. Further, the correlation between PV power output and wind speed is 0.094, which is irrelevant. These three factors are positively related to PV power output. As a result, solar radiation and temperature are chosen as the input variables of the FA-KELM model.

4. The Proposed Method

4.1. Prediction Interval Evaluation Indices

The prediction interval results are composed of the upper and lower boundaries and correspond to certain expected confidence levels, which are different from point forecasts. The common predictive indices used here are given in Appendix A.

4.2. Kernel Density Estimation (KDE)

In order to get the confidence interval as high as possible and the average interval bandwidth as small as possible, the kernel density distribution interval is estimated by the predicted error between the point predicted value and actual value of PV power. For a specific predicted error e, the formula of its probability density function can be written as follows [30,48,49]:

$$f(e) = \frac{1}{Nh}\sum_{i=1}^{N} k\left(\frac{e-\hat{e}_i}{h}\right) \tag{5}$$

where $k(x)$ is the kernel function, and Gauss kernel function is used, $\hat{e}_i$ represents the point predicted error samples for PV power, N is forecast sample size, and h is the PV interval bandwidth. In this paper, the bandwidth is calculated by the following equation [50]:

$$h = 0.9\min\left(\sigma, F_{IQR}/1.34\right)N^{-1/5} \tag{6}$$

where σ stands for sample standard deviation, and FIQR is the interquartile sample range.

If the kernel density estimation is used for each data point, the computation will increase when the sample number is increasing. This work allocates the power prediction errors into equal intervals in hours. Considering that the length of the power section is ΔP and the range of power fluctuation is $[P_1, P_h]$, the interval can be calculated as

$$D_i = [P_1 + (i-1)\Delta P, P_1 + i\Delta P] \tag{7}$$

where $i = 1, 2, \ldots, l$, l is the number of sectors. Moreover, l is calculated as

$$l = [(P_h - P_1)/\Delta P] + 1 \tag{8}$$

According to the PV power point prediction value, the probability density curve of the power prediction error is computed by a kernel density estimation, then the PV power prediction interval is obtained under a certain confidence level. Suppose that the probability distribution function of a PV power prediction error e is $F(\varepsilon)$, where ε is a random variable of prediction error, then the probability prediction range of the actual value of PV power under a given confidence level $1 - \alpha$ is

$$\left[\hat{P} + G(\alpha_1), \hat{P} + G(\alpha_2)\right] \tag{9}$$

where $\hat{P}$ is the point prediction value, and $G(\alpha_1)$ and $G(\alpha_2)$ are both the inverse functions of the probability distribution function $F(\varepsilon)$. Through the sensitivity analysis, it is found that when $\alpha_1 = \alpha/2$ and $\alpha_2 = 1 - \alpha_1$, the confidence interval is minimum.

The practical implementation of the prediction interval under a certain confidence level is summarized as follows:

Step 1: Determine the corresponding power interval for a predicted value;

Step 2: Find the error probability density curve corresponding to the above interval, and solve the corresponding probability distribution function with the integral;

Step 3: Adopt the cubic spline interpolation method to fit the probability distribution curve of the prediction error. Then, solve the $\alpha/2$ and $1 - \alpha/2$ intervals of the prediction error for the PV output;

Step 4: Calculate the prediction interval for the power output according to Equation (8).

4.3. Probabilistic Prediction Interval Model

A probabilistic prediction interval is formed under a certain confidence level, and the interval predicted value is calculated by determining the lower and upper bounds of the confidence interval, which will provide more comprehensive information for power system decision-makers. The flowchart is shown in Figure 2 below.

The procedure can be explained as follows:

Step 1: Preprocess the power data of the PV power plant with normalization.

$$u = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{10}$$

where x_i represents the original input data, $x_{\max} = \max(x)$, $x_{\min} = \min(x)$, $u \in [0, 1]$ is the normalized data, and $i = 1, 2, \ldots, m$, m is the total input number;

Step 2: Use the clustering algorithm to classify the original PV power data into sunny, cloudy, rainy, and other weather conditions;

Step 3: Use the VMD algorithm to decompose different weather conditions data, then the subseries is divided into training set and testing set;

Step 4: Add the corresponding meteorological data into the data set from Step 3, and the FA-KELM algorithm is adopted to create the forecasting model of intra-day-ahead. Finally, the subseries will be added to obtain the final point prediction value;

Step 5: According to the error between the point prediction value and the actual value used to estimate the kernel density distribution, the confidence interval under a certain confidence level will be obtained by the cubic spline interpolation method. Then, the prediction interval model is built;

Step 6: Analyze the predicted results according to the prediction interval evaluation indices.

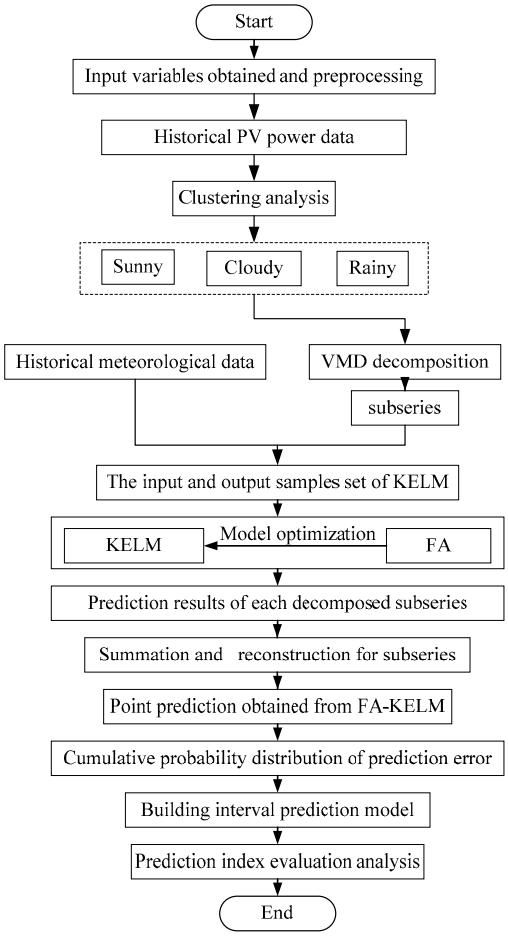


Figure 2. The structure of the proposed prediction model.

5. Case Study and Discussion

5.1. Data Collection

The data from the Ashland PV power plant with the capacity of 15 kW per unit in Oregon, USA (latitude: 42.19; longitude: 122.70; altitude: 595 m) are selected for modeling and evaluating the prediction performance. The research duration time is between 6 a.m. and 6 p.m. from 1 January 2015 to 31 December 2015, in which some "incomparable" data of 11 days are found by comparing and analyzing the original data [51]. Then, these data are removed in the prediction. The database from the website contains data for every 5 min, and we derive the average data value for every 15 min. We collect data for 12 h per day, that is, there are 48 data points per day with a sampling rate of 15 min per sample. The corresponding weather information per day was obtained from the historical weather data through the website given in reference [52]. The data which fluctuate excessively and irregularly are eliminated. The data sets are normalized to the interval of [0, 1]. In the meantime, 75% of the data are treated as the training set, while the remaining 25% is regarded as the testing set. The original sample data of the PV power series in April 2015 are shown in Figure 3 below.

5.2. Cluster Analysis

The self-organizing map (SOM) neural network [53,54] consists of an input layer and a competition layer. The main idea is that neurons in the competition layer of the network will compete with each other in order to gain the response opportunity to input variables. Finally, only one neuron will win. Usually, it is mainly used for clustering.

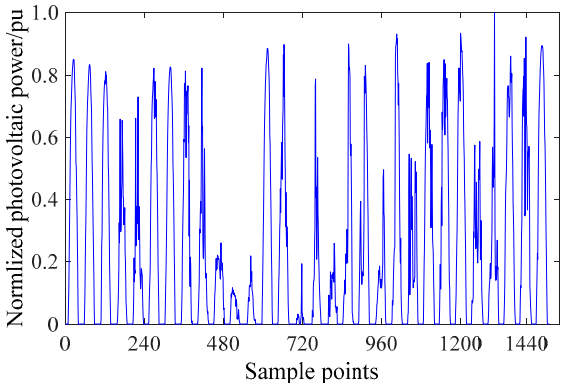


Figure 3. The original PV power series in April 2015.

Considering the randomness and intermittent nature of PV power, it will generate significant errors if we make predictions directly through the original data. Therefore, according to the already known weather conditions, cluster analysis for the original data sample in the year 2015 is conducted through the SOM neural network. Its clustering distribution is shown in Figure 4, and the clustering results are shown in Table 3.

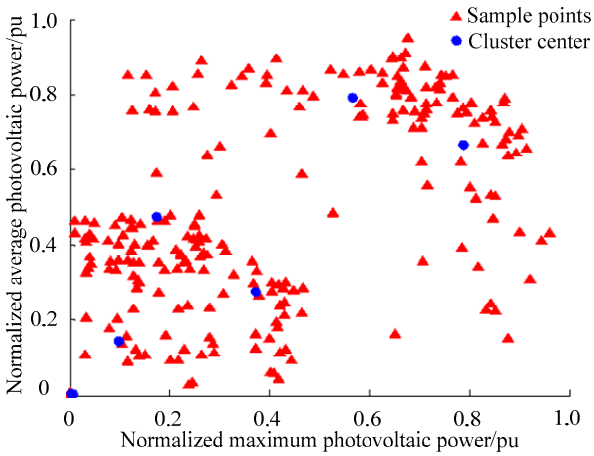


Figure 4. Sample points distribution.

Table 3. Clustering results.

Weather Conditions	Clustering Days/Day
Sunny	160
Cloudy/heavily overcast	85
Rainy/snow	40
Cloudy to rainy	24
Rainy to cloudy	28
Fog	17

5.3. VMD Decomposition

In this paper, VMD is used to decompose the original PV power data into finite subseries, then build the prediction model for each subseries, which can reduce the non-stationary characteristic of prediction data. VMD transfers signal decomposition to the framework of variational theory, and the optimal solution of the variational model is obtained by iterative calculation to determine the center frequency and bandwidth of each mode. The sum of all modes is the source signal. Empirical mode decomposition (EMD) is another popular decomposition method. It decomposes signals into

characteristic modes. Further, it has the advantage of not using any defined functions as a basis, but instead adaptive generation of intrinsic modal functions based on the analyzed signals.

On sunny days, the data collected from the first 125 days out of 160 days are treated as the training sets, and the data from the last 35 days are regarded as testing sets, which are obtained from the clustering algorithm of Section 5.2. The main parameters need to be set for the VMD program including penalty factor α, discriminant precision e, and mode number K. K has an influence on the decomposition process, and it will be under decomposition when K is too small; otherwise, it will be over decomposition. A simple and effective process is adapted to determine the mode number. A sensitivity analysis is conducted on the K value ranging from 3 to 10 with the parameters in the VMD-FA-KELM algorithm presented in Table 4. The analysis shows that while K is greater than 6, the central frequency of the modes will be similar. Therefore, choosing $K = 6$ will be the optimal choice. VMD decomposes the PV power series with the decomposition results for three weather conditions, i.e., sunny, cloudy, and rainy, presented in Figure 5. IMF1 to IMF6 represent the subseries. It is worth noting that only six IMFs have been plotted as determined by the algorithm. Table 3 depicts the number of days for other weather conditions including cloudy, rainy, and fog.

Table 4. Parameters in the VMD-FA-KELM algorithm.

Method	Main Parameter	Numerical Value
VMD	Penalty factor α	2000
	Discriminant precision	e^{-7}
FA	Maximum iterations	200
	Population number	60
	Dimensions to be optimized d	2
	Attractiveness β_0	0.2
	Light absorption factor γ	1
	Step size	0.5
KELM	Kernel function	RBF
	Initialized penalty coefficient C	$[10^{-2}, 10^{10}]$
	Initialized kernel parameter	$[10^{-2}, 10^{10}]$

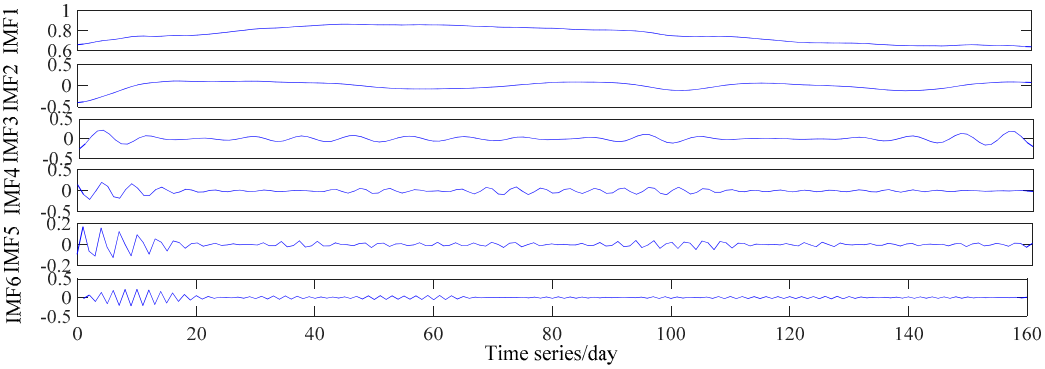
5.4. Simulation Results under Different Confidence Levels

To validate the proposed approach, the simulation results under different confidence levels are compared in this section. Select sample data of 16 sunny days, 12 cloudy days, and 7 rainy days were obtained from the cluster analysis of Section 5.2 to get the corresponding point prediction value. Based on this, the prediction error is obtained, and the prediction interval model based on the kernel density estimation is then produced. The relevant parameter settings of the proposed model are shown in Table 4 below.

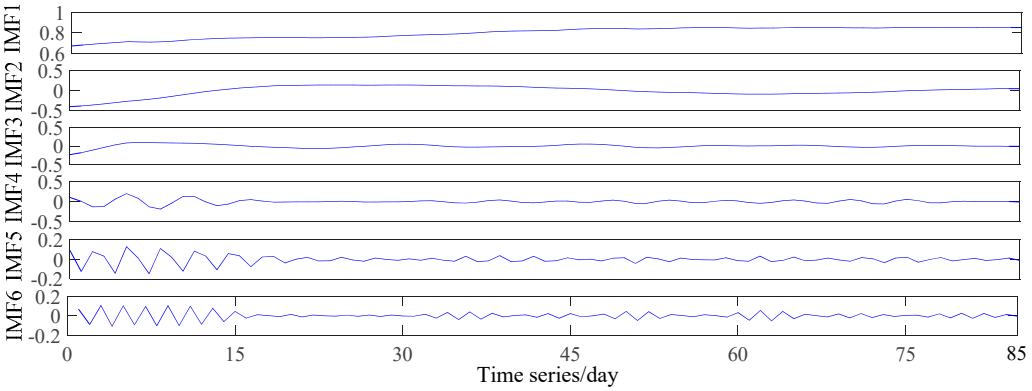
Based on the VMD-FA-KELM algorithm and kernel density estimation, the prediction interval results of sunny, cloudy, and rainy weather for each two-day period at 90% and 70% confidence levels, respectively, are shown in Figures 6 and 7, respectively, where the daily time interval is from 6 a.m. to 6 p.m. Further, its corresponding prediction interval indices are given in terms of the prediction interval coverage probability (PICP) and prediction interval normalized averaged width (PINAW).

As shown in Figures 6 and 7, we can see that the actual PV power almost falls within the 90% confidence level, and a small part is outside the 70% confidence level. This shows that the smaller the confidence level is, the narrower the interval width is. In order to satisfy the corresponding confidence level, it is shown that when the confidence level is high, the average width of the interval gets wider. As such, the probability of real PV power falling into the prediction interval is greater. Moreover, when the confidence level is low, the average width of the interval is narrower. Then, the probability of real PV power falling into the prediction interval is smaller. As the confidence level decreases gradually, the corresponding prediction interval normalized averaged width will decrease, and the

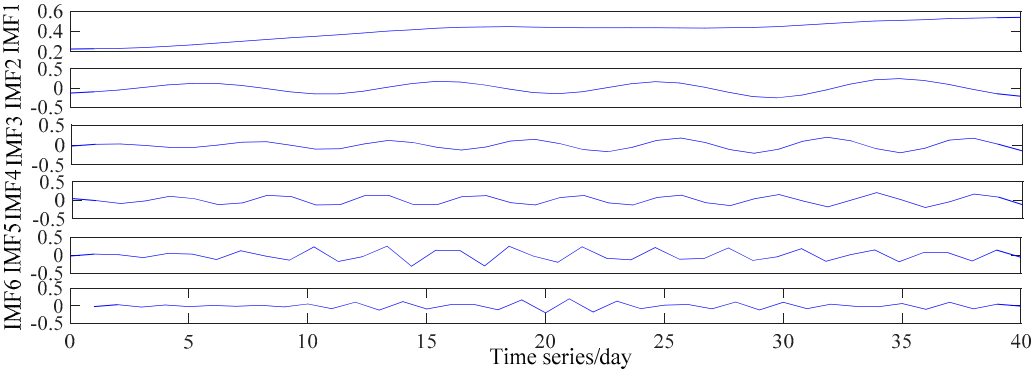
coverage probability will decrease as well. The width of the sunny prediction interval is relatively narrow, which indicates that the sunny data are more stable, and the prediction interval accuracy is higher than cloudy weather and rainy weather. Further, on the same day, the larger the point predicted error is, the greater the range of error fluctuation is. Moreover, at noon, the error is the largest, and the interval is the widest.



(**a**) Sunny weather



(**b**) Cloudy weather



(**c**) Rainy weather

Figure 5. Variational mode decomposition (VMD) of PV power at noon each day.

The comparisons in Table 5 show that whether it is sunny, cloudy, or rainy, the interval width of the 90% confidence level is wider than that under the 70% confidence level, and the corresponding prediction interval coverage probability is also higher than the 70% confidence level by more than

10%. The width of the prediction interval is minimum in sunny weather, and the prediction interval coverage probability is higher than cloudy and rainy days.

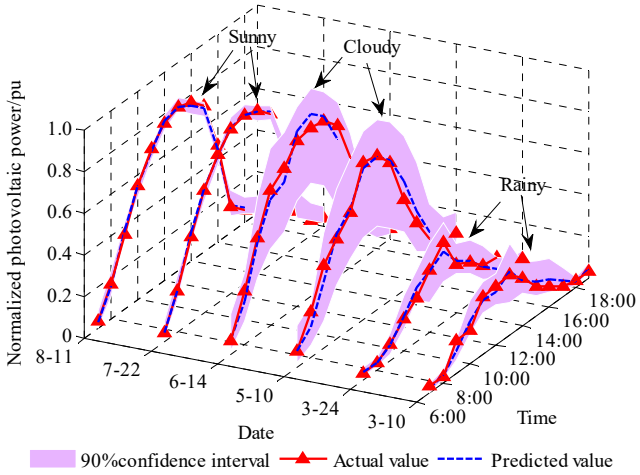


Figure 6. The predicted results under 90% confidence interval.

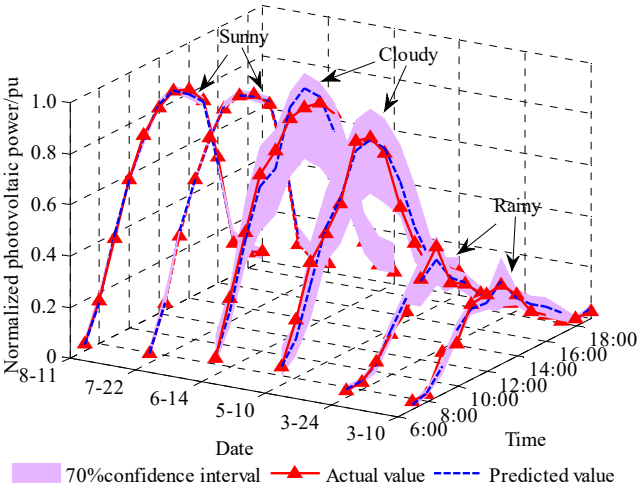


Figure 7. The predicted results under 70% confidence interval.

Table 5. Prediction interval indices of proposed model at different confidence levels.

Confidence Level/%	PICP/%			PINAW/%		
	Sunny	**Cloudy**	**Rainy**	**Sunny**	**Cloudy**	**Rainy**
90	97.96	97.96	95.92	9.98	43.75	30.28
70	85.71	84.69	76.53	4.87	27.62	18.61

5.5. Comparison of the Proposed Model with Other Models

The proposed VMD-FA-KELM model is compared with three other models, namely VMD-PSO-KELM, WPD-FA-KELM, and VMD-KELM. In the PSO algorithm, the number of iterations is 200, the group size is 60, and the values of learning parameters c_1 and c_2 are both 2. The several case studies of point prediction on 11 August 2015 sunny day, 10 May cloudy day, and 14 September rainy weather are shown in Figures 8–10, respectively.

Figures 8–10 show that all models have high prediction accuracy in sunny weather, and the PV power forecasting output can fit the actual output better. In cloudy and rainy weather, the prediction results fluctuate violently because the PV power output has more randomness and uncertainty.

As shown in Figure 9, the prediction value obtained by the FA-KELM algorithm is closer to the actual value. This means that FA is more optimal for KELM.

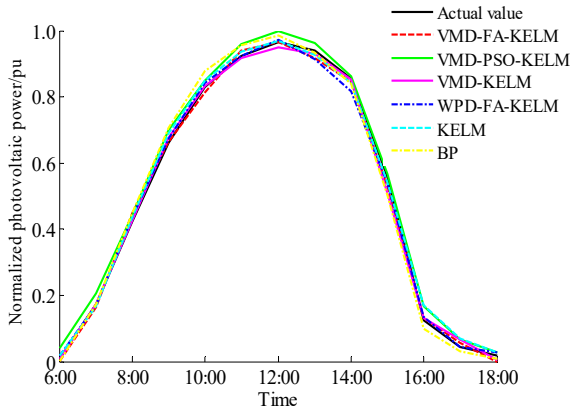


Figure 8. Point prediction of the sunny day under different optimization algorithms.

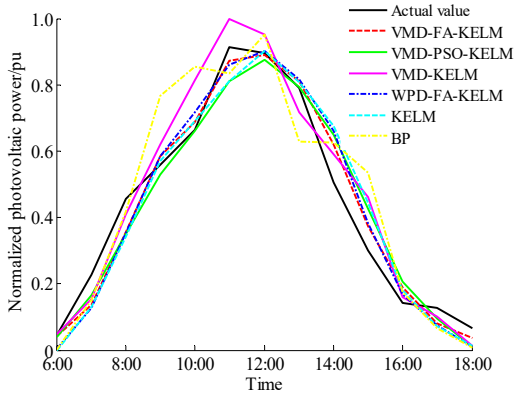


Figure 9. Point prediction of the cloudy day under different optimization algorithms.

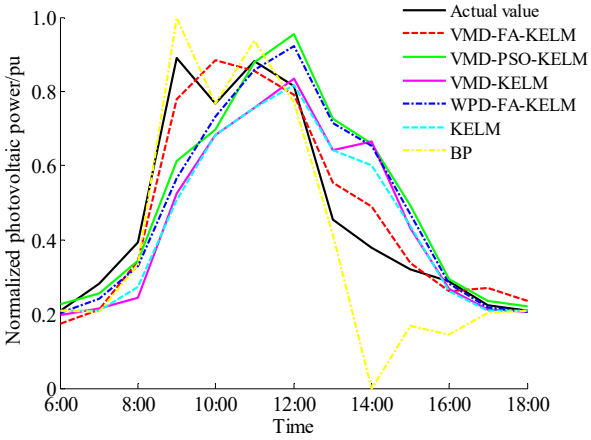


Figure 10. Point prediction of rainy day under different optimization algorithms.

In order to further compare the proposed method with other methods, the normalized root mean square error index e_{NRMSE} and normalized mean absolute error index e_{NMAE} are used to measure the point prediction error in Table 6, while the constructed PI at the 90% confidence level is summarized in Table 7 in terms of the reliability index, prediction interval coverage probability (PICP), sharpness index, and prediction interval normalized averaged width (PINAW).

Table 6. Point prediction indices of different optimization algorithms.

Model	e_{NRMSE}%			e_{NMAE}%		
	Sunny	**Cloudy**	**Rainy**	**Sunny**	**Cloudy**	**Rainy**
VMD-FA-KELM	1.37	5.29	2.71	1.13	4.28	2.31
VMD-PSO-KELM	2.97	6.58	5.84	2.55	5.07	4.16
WPD-FA-KELM	1.48	6.29	5.87	1.35	5.20	4.08
VMD-KELM	1.43	7.22	6.09	1.13	6.03	4.50
KELM	1.80	7.48	5.84	1.51	5.94	4.26
BP	2.61	11.13	6.78	2.13	9.39	4.80

Table 7. Prediction interval indices of different optimization algorithms under the 90% confidence level.

Model	PICP/%			PINAW/%		
	Sunny	**Cloudy**	**Rainy**	**Sunny**	**Cloudy**	**Rainy**
VMD-FA-KELM	97.96	97.96	95.92	9.98	43.75	30.28
VMD-PSO-KELM	81.63	96.94	80.77	24.16	62.82	33.31
WPD-FA-KELM	95.92	91.84	87.76	18.40	36.35	41.74
VMD-KELM	90.81	87.76	94.90	14.97	58.88	44.81
KELM	100.00	92.31	88.46	17.14	41.73	46.03
BP	96.15	80.77	92.31	20.61	52.32	86.77

From the indices results of Table 7, we can see that the proposed model has the smallest point prediction error, and the values of both e_{NRMSE} and e_{NMAE} are below 10% in all weather conditions. This indicates that the proposed model gives the best performance in all models. From Table 6, the proposed prediction interval model at the 90% confidence level achieves a PICP value of 97.96% and a PINAW value of 9.98% in sunny weather. The PICP of the prediction interval for the proposed method meets the corresponding confidence level, that is, the confidence level is greater than 90%, and the width of the prediction interval is the narrowest, which indicates that the proposed prediction interval model can construct the prediction interval effectively and more practically.

From the indices results of Table 6, we can see that the proposed model has the smallest point prediction error, and the values of both e_{NRMSE} and e_{NMAE} are below 10% in all weather conditions. This indicates that the proposed model gives the best performance of all the models. From Table 5, the proposed prediction interval model at the 90% confidence level achieves a PICP value of 97.96% and a PINAW value of 9.98% in sunny weather. The PICP of the prediction interval for the proposed method meets the corresponding confidence level, that is, the confidence level is higher than 90%, and the width of the prediction interval is the narrowest, which indicates that the proposed prediction interval model can construct the prediction interval effectively and more practically.

Further, from Tables 6 and 7, the forecast errors of sunny days are smaller than those of cloudy days and rainy days. Based on the same VMD decomposition technique, the prediction error of the single KELM model is larger than that from the PSO-KELM model and FA-KELM model. This means that after optimizing the penalty parameter C and kernel parameter g, a better KELM model can be obtained. It effectively avoids the random selection of the two parameters of KELM. The FA- KELM prediction is better than PSO-KELM: this implies that the FA algorithm has a stronger global searching ability and generalization ability than the PSO algorithm. Based on the same FA-KELM algorithm, the prediction error decomposed by VMD is less than WPD. This shows that the VMD algorithm can effectively overcome the disadvantages of the selection of wavelet bases and the number of decomposition layers in WPD decomposition. Compared with the undecomposed KELM prediction method, the prediction effect of the decomposed ones is better, which indicates that the VMD algorithm can effectively reduce the non-stationarity of the sequence. At the same time, compared with the common BP neural network, the prediction error of KELM is smaller. In summary, the proposed FA algorithm produces an overall

optimum. The above analysis coincides with results reported from Figures 6–10, which can provide more accurate decision-making and ensure better security of a power supply.

5.6. Potential Application to Power Systems

The statistical analysis and comparative study of a large amount of irradiance data under different weather types show that sunny, cloudy, and rainy days are typical weather types. The irradiance variation of these three types is different and has their distinct characteristics, and the probability of these three weather types is higher and covers the weather states corresponding to most of the dates. Without a loss of generality, it aims to explain the framework in a simplified way by reducing the calculation scale of the interval prediction but to demonstrate the benefit gained from this approach. With the advancement of computers, it can be foreseen that the number of clusters could be increased significantly to get a practical solution within the allowable time constraints. It is possible to compare the probability forecast of normal distribution and the proposed probability prediction of kernel density distribution to get a better prediction interval.

As reported in [55], microgrids consist of smart buildings with solar panels. The solar energy can be trade with each other in a peer-to-peer approach. This potential application aims to optimize the energy consumption of PV energy merged with energy storage systems (ESSs), such as electric vehicles for a microgrid community of multiple buildings.

Commonly, a community may be equipped with PV systems, heat pumps, and multiple sensors, etc. Energy production prediction based on machine learning and short-term weather forecasts can help identifying possible management and optimal usage of various systems (e.g., heating and cooling) to enhance the system operation. For example, neighbors in the Brooklyn Microgrid project produce, consume, and buy power in their community with a transactive energy platform based on blockchain [56]. The platform facilitates distributed energy supply systems that is highly based on renewable-based sources such as solar energy generation for a more resilient, low-carbon, and customer-driven economy to deploy smart cities [57]. Energy harvesting from solar energy will be important to have a good prediction of solar irradiance.

6. Conclusions

This paper proposed a novel hybrid model for the day-ahead or intra-day-ahead PV power output prediction interval considering the principle of VMD and FA-KELM. The proposed approach shows promising results as compared to existing methods without PV power series decomposition. VMD decomposition has been used for the first time to decompose the PV power series of different weather conditions, which overcomes the disadvantages of the selection of wavelet bases and the number of decomposition layers in WPD decomposition. The decomposing technique is useful in identifying the complexity of the IMFs series. In addition, the hybrid use of VMD and FA-KELM has shown to be an effective method to construct the optimal PIs. In addition to this, to the best knowledge of the authors, it is the first time in applying this integrated approach to solar irradiance prediction. Study results show that the presented hybrid method can give excellent quality of PIs, with significance for practical applications in system operation, planning, and risk assessment.

Solar irradiance prediction is a non-linear and non-deterministic problem and the mathematical model will not be obtained easily. As such in this instance, the authors focus on the artificial intelligence approach. With many simulations done, the authors have confidence that they are getting a good solution as compared with other mainstream methods. Further work will be carried out in a sensitivity study to search for a near global optimal solution.

Author Contributions: Conceptualization, X.W., C.S.L.; methodology, X.W., C.S.L.; validation, Q.Z., and B.L.; formal analysis, X.W. and Q.Z.; investigation, C.B.; resources, Q.Z. and C.B.; writing—original draft preparation, X.W.; writing—review and editing, C.S.L. and L.L.L.; supervision, L.L.L.; funding acquisition, L.L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Department of Finance and Education of Guangdong Province 2016 [202]: Key Discipline Construction Program, China; and the Education Department of Guangdong Province in China: New and Integrated Energy System Theory and Technology Research Group [Project Number 2016KCXTD022]; Brunel University London BRIEF Funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The prediction interval coverage probability (PICP) is calculated based on the number of occurrences that the output value falls inside the PI for a given confidence level α, is given by [46]

$$\begin{aligned} PICP &= \tfrac{1}{N}\sum_{i=1}^{N} a_i \\ a_i &= \begin{cases} 1, t_i \in [L_i, U_i] \\ 0, t_i \notin [L_i, U_i] \end{cases} \end{aligned} \tag{A1}$$

where a_i is a Boolean value, t_i represents the predicted target value, and U_i and L_i are upper and lower bounds of prediction interval, respectively. Moreover, N is the number of the prediction sample. Based on satisfying the confidence level α, the larger the PICP value is, the greater the confidence level is, where the number of actual PV power falling into the prediction interval is larger. The larger the PICP value is, the greater the confidence level is.

Then, to account for the fact that,

The PI normalized averaged width (PINAW) considers the output value will easily fall inside the wider PI as follows [58]:

$$PINAW = \frac{1}{NR}\sum_{i=1}^{N}(U_i - L_i) \tag{A2}$$

R depicts the range of the predicted targets. The PINAW value is used to measure the ability of predicted results for describing uncertain information. When PICP is constant, the smaller the value of PINAW is, the better the predicted results are.

Normalized mean absolute error (NMAE) and normalized root mean square error (NRMSE) are considered to evaluate the deterministic prediction [59]:

$$e_{NMAE} = \frac{1}{P_{rated}N}\sum_{i=1}^{N}\left|\hat{Y}_i - Y_i\right| \times 100\% \tag{A3}$$

$$e_{NRMSE} = \frac{1}{P_{rated}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2} \times 100\% \tag{A4}$$

where P_{rated} is the rated power of the PV unit, $\hat{Y}_i$ is the point predicted value, and Y_i is the actual point value at time i.

References

1. Ceci, M.; Corizzo, R.; Fumarola, F.; Malerba, D.; Rashkovska, A. Predictive modeling of PV energy production: How to set up the learning task for a better prediction? *IEEE Trans. Ind. Inform.* **2017**, *13*, 956–966. [CrossRef]
2. Sobrina, S.; Koohi-Kamali, S.; Nasrudin, A.R. Solar photovoltaic generation forecasting methods: A review. *Energy Convers. Manag.* **2018**, *156*, 459–497. [CrossRef]
3. Majumder, I.; Dash, P.K.; Bisoi, R. Variational mode decomposition based low rank robust kernel extreme learning machine for solar irradiation forecasting. *Energy Convers. Manag.* **2018**, *171*, 787–806. [CrossRef]
4. Lai, C.S.; Jia, Y.; Lai, L.L.; Xu, Z.; McCulloch, M.D.; Wong, K.P. A comprehensive review on large-scale photovoltaic system with applications of electrical energy storage. *Renew. Sustain. Energy Rev.* **2017**, *78*, 439–451. [CrossRef]

5. Wolff, B.; Kühnert, J.; Lorenz, E.; Kramer, O.; Heinemann, D. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. *Sol. Energy* **2016**, *135*, 197–208. [CrossRef]
6. Massidda, L.; Marrocu, M. Use of multilinear adaptive regression splines and numerical weather prediction to forecast the power output of a PV plant in Borkum, Germany. *Sol. Energy* **2017**, *146*, 141–149. [CrossRef]
7. Prema, V.; Rao, K.U. Development of statistical time series models for solar power prediction. *Renew. Energy* **2015**, *83*, 100–109. [CrossRef]
8. Lai, C.S.; Jia, Y.; McCulloch, M.D.; Xu, Z. Daily clearness index profiles cluster analysis for photovoltaic system. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2322–2332. [CrossRef]
9. Lai, C.S.; Li, X.; Lai, L.L.; McCulloch, M.D. Daily clearness index profiles and weather conditions studies for photovoltaic systems. *Energy Procedia.* **2017**, *142*, 77–82. [CrossRef]
10. Mellit, A.; Pavan, A.M. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Sol. Energy* **2010**, *84*, 807–821. [CrossRef]
11. İzgi, E.; Öztopal, A.; Yerli, B.; Kaymak, M.K.; Sahin, A.D. Short-mid-term solar power prediction by using artificial neural networks. *Sol. Energy* **2012**, *86*, 725–733. [CrossRef]
12. Liu, Z.; Liu, X.; Wang, K.; Liang, Z.; Correia, J.A.F.O.; De Jesus, A.M.P. GA-BP neural network-based strain prediction in full-scale static testing of wind turbine blades. *Energies* **2019**, *12*, 1026. [CrossRef]
13. Barbosa, J.F.; Correia, J.A.F.O.; Júnior, R.C.S.F.; De esus, A.M.P. Fatigue life prediction of metallic materials considering mean stress effects by means of an artificial neural network. *Int. J. Fatigue* **2020**, *135*, 105527. [CrossRef]
14. Shamshirband, S.; Mohammadi, K.; Khorasanizadeh, H.; Yee, P.L.; Lee, M.; Petković, D.; Zalnezhad, E. Estimating the diffuse solar radiation using a coupled support vector machine-wavelet transform model. *Renew. Sustain. Energy Rev.* **2016**, *56*, 428–435. [CrossRef]
15. Yadav, A.K.; Chandel, S.S. Identification of relevant input variables for prediction of 1-minute time-step photovoltaic module power using artificial neural network and multiple linear regression models. *Renew. Sustain. Energy Rev.* **2017**, *77*, 955–969. [CrossRef]
16. Almonacid, F.; Pérez-Higueras, P.J.; Fernández, E.F.; Hontoria, L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. *Energy Convers. Manag.* **2014**, *85*, 389–398. [CrossRef]
17. Junior, J.G.D.S.F.; Oozeki, T.; Ohtake, H.; Shimose, K.; Takashima, T.; Ogimoto, K. Regional forecasts and smoothing effect of photovoltaic power generation in Japan: An approach with principal component analysis. *Renew. Energy* **2014**, *68*, 403–413. [CrossRef]
18. Giorgi, M.G.D.; Malvoni, M.; Congedo, P.M. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. *Energy* **2016**, *107*, 360–373. [CrossRef]
19. Teo, T.T.; Logenthiran, T.; Woo, W.L.; Abidi, K. Forecasting of photovoltaic power using regularized ensemble extreme learning machine (RE-ELM). In Proceedings of the International Technical Conference of IEEE Region 10 (TENCON), Singapore, 22–25 November 2016; pp. 455–458.
20. Shrivastava, N.A.; Lohia, K.; Panigrahi, B.K. A multiobjective framework for wind speed prediction interval forecasts. *Renew. Energy* **2016**, *87*, 903–910. [CrossRef]
21. Khosravi, A.; Nahavandi, S.; Creighton, D. Prediction intervals for short-term wind farm power generation forecasts. *IEEE Trans. Sustain. Energy* **2013**, *4*, 602–610. [CrossRef]
22. Khosravi, A.; Nahavandi, S. Combined nonparametric prediction intervals for wind power generation. *IEEE Trans. Sustain. Energy* **2013**, *4*, 849–856. [CrossRef]
23. Ni, Q.; Zhuang, S.; Sheng, H.; Kang, G.; Xiao, J. An ensemble prediction intervals approach for short-term PV power forecasting. *Sol. Energy* **2017**, *155*, 1072–1083. [CrossRef]
24. Han, Y.; Wang, N.; Ma, M.; Zhou, H.; Dai, S.; Zhu, H. A PV power interval forecasting based on seasonal model and nonparametric estimation algorithm. *Sol. Energy* **2019**, *184*, 515–526. [CrossRef]
25. Li, G.; Ma, X.; Yang, H. A hybrid model for forecasting sunspots time series based on variational mode decomposition and backpropagation neural network improved by firefly algorithm. *Comput. Intell. Neurosci.* **2018**, 3713410. [CrossRef]
26. Galván, I.M.; Valls, J.M.; Cervantes, A.; Aler, R. Multi-objective evolutionary optimization of prediction intervals for solar energy forecasting with neural networks. *Inf. Sci.* **2017**, *418–419*, 363–382. [CrossRef]

27. Yang, X.; Xu, M.; Xu, S.; Han, X. Day-ahead forecasting of photovoltaic output power with similar cloud space fusion based on incomplete historical data mining. *Appl. Energy* **2017**, *206*, 683–696. [CrossRef]
28. Trapero, J. Calculation of solar irradiation prediction intervals combining volatility and kernel density estimates. *Energy* **2016**, *114*, 266–274. [CrossRef]
29. 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]
30. Chen, H.; Assala, P.D.S.; Cai, Y.; Yang, P. Intelligent transient overvoltages location in distribution systems using wavelet packet decomposition and general regression neural networks. *IEEE Trans. Ind. Inform.* **2016**, *12*, 1726–1735. [CrossRef]
31. Gao, B.; Woo, W.L.; Dlay, S.S. Single-Channel Source Separation Using EMD-Subband Variable Regularized Sparse Features. *IEEE Trans. Audio Speech Lang. Process.* **2011**, *19*, 961–976. [CrossRef]
32. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
33. Wang, Y.; Liu, F.; Jiang, Z.; He, S.; Mo, Q. Complex variational mode decomposition for signal processing applications. *Mech. Syst. Signal Process.* **2017**, *86*, 75–85. [CrossRef]
34. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
35. Shamshirband, S.; Mohammadi, K.; Chen, H.L.; Samy, G.N.; Petković, D.; Ma, C. Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: A case study for Iran. *J. Atmos. Sol. Terr. Phys.* **2015**, *134*, 109–117. [CrossRef]
36. Shamshirband, S.; Mohammadi, K.; Yee, P.L.; Petković, D.; Mostafaeipour, A. A comparative evaluation for identifying the suitability of extreme learning machine to predict horizontal global solar radiation. *Renew. Sustain. Energy Rev.* **2015**, *52*, 1031–1042. [CrossRef]
37. Wang, M.; Chen, H.; Yang, B.; Zhao, X.; Hu, L.; Cai, Z.; Huang, H.; Tong, C. Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses. *Neurocomputing* **2017**, *267*, 69–84. [CrossRef]
38. Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2012**, *42*, 513–529. [CrossRef]
39. Zhang, Y.; Li, C.; Li, L. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Appl. Energy* **2017**, *190*, 291–305.
40. Olatomiwa, L.; Mekhilef, S.; Shamshirband, S.; Mohammadi, K.; Petković, D.; Sudheer, C. A support vector machine-firefly algorithm-based model for global solar radiation prediction. *Sol. Energy* **2015**, *115*, 632–644. [CrossRef]
41. Mishra, A.; Agarwal, C.; Sharma, A.; Bedi, P. Optimized gray-scale image watermarking using DWT-SVD and firefly algorithm. *Expert Syst. Appl.* **2014**, *41*, 7858–7867. [CrossRef]
42. Barbieri, F.; Rajakaruna, S.; Ghosh, A. Very short-term photovoltaic power forecasting with cloud modeling: A review. *Renew. Sustain. Energy Rev.* **2017**, *75*, 242–263. [CrossRef]
43. Yang, H.T.; Huang, C.M.; Huang, Y.C.; Pai, Y.S. A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. *IEEE Trans. Sustain. Energy* **2014**, *5*, 917–926. [CrossRef]
44. Felice, M.D.; Petitta, M.; Ruti, P.M. Short-term predictability of photovoltaic production over Italy. *Renew. Energy* **2015**, *80*, 197–204. [CrossRef]
45. 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]
46. Xu, W.; Hou, Y.; Hung, Y.S.; Zou, Y. Comparison of Spearman's rho and Kendall's tau in normal and contaminated normal models. *Signal Process.* **2010**, *93*, 261–276. [CrossRef]
47. 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]
48. Lai, C.S.; Xu, F.; Tao, Y.; Ng, W.W.Y.; Jia, Y.; Yuan, H.; Huang, C.; Lai, L.L.; Xu, Z.; Locatelli, G. A robust correlation analysis framework for imbalanced and dichotomous data with uncertainty. *Inf. Sci.* **2019**, *470*, 58–77. [CrossRef]
49. Taylor, J.W.; Jeon, J. Forecasting wind power quantiles using conditional kernel estimation. *Renew. Energy* **2015**, *80*, 370–379. [CrossRef]

50. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman and Hall: London, UK, 1986; pp. 34–60.
51. Data for Ashland PV Station in 2015. Available online: http://solardat.uoregon.edu/download/Archive/ASRF1501.txt (accessed on 25 April 2020).
52. Meteorological Data for Ashland. Available online: https://www.wunderground.com/weather/us/or/ashland/97520 (accessed on 20 May 2020).
53. Liu, H.; Ban, X.J. Clustering by growing incremental self-organizing neural network. *Expert Syst. Appl.* **2015**, *42*, 4965–4981. [CrossRef]
54. Llanos, J.; Morales, R.; Núñez, A.; Sáez, D.; Lacalle, M.; Marín, L.G.; Hernández, R.; Lanas, F. Load estimation for microgrid planning based on a self-organizing map methodology. *Appl. Soft Comput.* **2017**, *53*, 323–335. [CrossRef]
55. Li, Z.; Kang, J.; Yu, R.; Ye, D.; Deng, Q.; Zhang, Y. Consortium blockchain for secure energy trading in industrial internet of things. *IEEE Trans. Ind. Inform.* **2018**, *14*, 3690–3700. [CrossRef]
56. A Microgrid Grows in Brooklyn. Available online: https://new.siemens.com/global/en/company/stories/research-technologies/energytransition/a-microgrid-grows-in-brooklyn.html (accessed on 20 May 2020).
57. Lai, C.S.; Li, X.; Lai, L.L.; Lin, X. Blockchain for Smart Cities. *IEEE Smart Cities Newsletter*, 22 October 2018. Available online: https://resourcecenter.smartcities.ieee.org (accessed on 20 May 2020).
58. Scolari, E.; Sossan, F.; Paolone, M. Irradiance prediction intervals for PV stochastic generation in microgrid applications. *Sol. Energy* **2016**, *139*, 116–129. [CrossRef]
59. Zhang, Y.; Liu, K.; Qin, L.; An, X. Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods. *Energy Convers. Manag.* **2016**, *112*, 208–219. [CrossRef]

© 2020 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/).