

Article

Wind Power Interval Prediction with Adaptive Rolling Error Correction Based on PSR-BLS-QR

Xu Ran , Chang Xu *, Lei Ma and Feifei Xue 

College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China; rx15850680102@163.com (X.R.); 191606030012@hhu.edu.cn (L.M.); xuefeifeihhu@163.com (F.X.)

* Correspondence: zhuifengxu@163.com

Abstract: Effective prediction of wind power output intervals can capture the trend of uncertain wind output power in the form of probability, which not only can avoid the impact of randomness and volatility on grid security, but also can provide supportable information for grid dispatching and grid planning. To address the problem of the low accuracy of traditional wind power interval prediction, a new interval prediction method of wind power is proposed based on PSR-BLS-QR with adaptive rolling error correction. First, one-dimensional wind power data are mapped to high-dimensional space by phase space reconstruction (PSR) to achieve data reconstruction and the input and output of the broad learning system (BLS) model are constructed. Second, the training set and the test set are divided according to the input and output data. The BLS model is trained by the training set and the initial power interval of training data is constructed by quantile regression (QR). Then, the error distribution of nonparametric kernel density estimation is constructed at different power interval segments of the interval upper and lower boundaries, respectively, and the corresponding error-corrected power is found. Next, the optimal correction index is used as the objective function to determine the optimal error correction power for different power interval segments of the interval upper and lower boundaries. Finally, a test set is used for testing the performance of the proposed method. Three wind power datasets from different regions are used to prove that the proposed method can improve the average prediction accuracy by about 6–14% with the narrower interval width compared with the traditional interval prediction methods.

Keywords: wind power interval prediction; PSR; BLS; QR; error correction



Citation: Ran, X.; Xu, C.; Ma, L.; Xue, F. Wind Power Interval Prediction with Adaptive Rolling Error Correction Based on PSR-BLS-QR. *Energies* **2022**, *15*, 4137. <https://doi.org/10.3390/en15114137>

Academic Editor: Adrian Ilinca

Received: 16 April 2022

Accepted: 2 June 2022

Published: 4 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The effective use of new energy can overcome the pollution of traditional fossil energy to the environment and ecology. The installed capacity of new energy is increasing year by year [1], and the most typical one is wind energy. Under the background of the national new energy policy, the power grid will be transformed from a traditional grid to a new energy grid with a large number of new energy units. However, wind power is affected by environmental and meteorological factors [2]. The fluctuation and randomness of output power is relatively large, and the large-scale integration into the grid will affect the security and stability of the grid. Thus, the more accurate wind power prediction is, the safer the entire grid system is. Wind power prediction consists of point prediction [3,4] and interval prediction [5,6]. Interval prediction can provide more information for decision-makers by predicting the upper boundary and the lower boundary of power [7]. How to improve the interval prediction accuracy of wind power has become a popular direction of research.

Wind power interval prediction can be divided into upper and lower boundary theory [8] and probabilistic interval prediction [9]. The upper and lower boundary theory is used to predict the interval of power output directly, and the interval prediction model is relatively simple, without a complex structure. Mei directly constructed the output of the model with different confidence intervals by introducing confidence interval parameters,

and directly output the power intervals through the input data [10]. The overall model is simple and feasible, and the effect is improved compared to the traditional interval prediction method. In addition to the construction of the model output, it is also feasible to construct the structure of the model itself. Reference [11] completes the model power interval output by constructing two loss functions of the long short term memory (LSTM) network. However, the quality of interval prediction depends on the construction of the loss functions, and it is difficult to construct good loss functions. In addition to constructing the model interval output directly by using the upper and lower boundary theory, the data interval can also be constructed indirectly in a probabilistic way based on point prediction. They are mainly divided into statistical probability interval prediction [12], quantile regression (QR) [13], bootstrap [14], and other probability predictions [15]. Statistical probability interval prediction is based on the known probability distribution and on construct interval predictions with confidence intervals according to the quantile. Zhang et al. used the Fourier distribution to fit the wind speed error. The different confidence intervals of wind speed are constructed [16]. Although the use of distribution fitting can construct a good confidence interval and obtain a certain accuracy of the interval distribution, it is very dependent on the model of the fitting error and the accuracy of point prediction. Traditional Bayes [17], bootstrap [18], quantile regression [19], and Gaussian regression [20] are also used to construct prediction intervals. In [19], Mei et al. obtained the prediction intervals under different quantiles by using a quantile regression model. Quantile regression does not suppose data distribution; however, a great amount of relevant historical data is required, and the solution process is more complicated. Various neural network algorithms with the quantile regression model are also used to make interval predictions. Antonio et al. established a two-stage quantile cooperation system for load interval prediction, and a quantile regression forest was used to improve the accuracy of forecasting load compared with traditional quantile regression [20]. Although the prediction interval network based on neural network and quantile regression can directly output the interval of the data, the structure of a neural network is relatively complex, with a large number of parameters and hyper-parameters, and the training process is very time-consuming. It is also easy to fall into the local optimum. Apart from the adjustment of the model, the accuracy can be further improved by changing the learning rate, training batch, dropout, and solving algorithms [21]. He et al. decomposed the predicted load power into different layers by variational mode decomposition (VMD) and performed quantile convolutional neural network (QCNN) interval prediction for different decomposition layers. The decomposed effect is better than the undecomposed effect [22]. The above is based on changing the prediction model and adjusting the internal structure of the model. It is necessary to proceed from the inherent nature of the data and further analyze their characteristics [23]. Mining the data character can be considered from data dimensionality reduction [24], clustering [25], and data reconstruction [26]. In [26], Wang mined the correlation of the high-dimensional space of the data and used phase space reconstruction (PSR) to realize the high-dimensional reconstruction of one-dimensional data. Clustering is also used. Fuzzy C-mean clustering is used to divide the initial dataset into several clusters and prediction models for different clusters are built separately. The result verifies that the prediction effect is higher than the unclassified effect [27]. In [28], the typical weather factors were extracted using principal analysis, and the high-dimensional feature information was reduced to the low-dimensional space. The reduction of dimensionality is beneficial to avoid redundant information.

Since wind power output is affected by various factors and its own output has nonlinear characteristics, it is more complicated to establish a nonlinear model and more difficult to predict when correlations between data are not explored. Moreover, it is necessary to find a new prediction model to avoid a series of problems such as a large amount of parameter training and time-consuming traditional neural networks' prediction models. What is more, the method for improving the accuracy of interval prediction needs to be further explored. Based on this, the characteristics of data are mined, and correlation of the

data is found by PSR; the prediction model effect is improved by the broad learning system (BLS) model, and the accuracy of interval prediction can be further improved by adaptive rolling error correction proposed in this paper. Therefore, a new interval prediction method based on PSR-BLS-QR with adaptive rolling error correction is proposed. One-dimensional wind power data is mapped to high-dimensional phase space by PSR, and the correlation between one-dimensional wind power data is mined to determine which data have correlations. The related data are established in the form of input and output data. The BLS model is used to find the specific relationship of input and output data. Wind power is predicted by the BLS model, and the initial wind power prediction intervals are constructed by quantile regression model. The initial wind power prediction intervals are corrected by adaptive error rolling correction model to obtain the final wind power prediction intervals.

2. Core Model of Interval Prediction

The interval prediction models include the phase space reconstruction model, the broad learning system prediction model, and the quantile regression model.

2.1. Phase Space Reconstruction

Since the one-dimensional time series itself has a certain correlation, the phase space reconstruction is used to mine the points in the one-dimensional time series that are correlated [29]. When the correlation of the one-dimensional time series is mined, the data will change from one-dimensional space to high-dimensional space, which is called phase space.

Thus, one-dimensional time series is input into the PSR model, and high-dimensional matrices are output. The embedding dimension m and the delay amount t are most important parameters to reflect whether the reconstructed phase space can well reflect the original information of one-dimensional time series and determine the degree of correlation between points. Inappropriate delay amount and embedding dimension are difficult to reproduce in the original one-dimensional time series. Thus, the best embedding dimension m_{op} and the delay amount t_{op} need to be found. C-C phase method is used to find m_{op} and t_{op} . The process of obtaining the best parameters m_{op} and t_{op} can be explained as follows.

$\{x(i), i = 1, 2, \dots, N\}$ is a one-dimensional time series of wind power, whose length is N . The i -th phase point can be expressed in Equation (1) after mapping the phase space.

$$X(i) = \{x(i), x(i+t), \dots, x(i+(m-1)t)\} \quad (1)$$

where $M = N - (m-1)t$.

The correlation integral of wind power output time series is introduced in Equation (2).

$$C(m, N, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i < j < M} \theta(r - \|X(i) - X(j)\|_{\infty}) \quad (r > 0) \quad (2)$$

where r is the search radius, θ function can be expressed in Equation (3).

$$\theta(*) = \begin{cases} 0 & * \leq 0 \\ 1 & * > 0 \end{cases} \quad (3)$$

Considering the limitation of time series length and the possible correlation between the data, the length of time series is divided into t subsequences of length N/t . Test statistics $S(m, N, r, t)$ can be defined in Equation (4) combined with Equation (2).

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, N/t, r, t) - C_s^m(1, N/t, r, t)] \quad (4)$$

When N is infinite, test statistics $S(m, N, r, t)$ can be modified into $S(m, r, t)$.

$$S(m, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, r, t) - C_s^m(1, r, t)] \quad (m = 2, 3, \dots) \tag{5}$$

Combined with $S(m, r, t)$, test statistics $\Delta S(m, t)$ can be defined in Equation (6).

$$\Delta S(m, t) = \max\{S(m, r_i, t)\} - \min\{S(m, r_i, t)\} \tag{6}$$

The best delay amount t_{op} can be obtained from first local minimum point of $\Delta S(m, t)$.

If the length of the time series N is more than 3000 according to Brock–Dechert–Scheinkman (BDS) statistics theorems, the range of the embedding dimension m is between 2 and 5, and the value of the search radius r can be obtained by the equation $r = k \times 0.5\sigma$. The range of k is between 1 and 4. σ represents the standard deviation of the wind power time series.

Combined with the related parameter above, the means of $S(m, N, r, t)$ and $\Delta S(m, t)$ are calculated in Equation (7).

$$\begin{aligned} \bar{S}(t) &= \frac{1}{16} \sum_{m=2}^5 \sum_{i=1}^4 S(m, r_i, t) \\ \Delta \bar{S}(t) &= \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t) \end{aligned} \tag{7}$$

The new test statistic $S_{cor}(t)$ is constructed by Equation (7)

$$S_{cor}(t) = \Delta \bar{S}(t) + |\bar{S}(t)| \tag{8}$$

The best embedding dimension m_{op} can be calculated by t_W in Equation (9). t_W is the global minimum point of $S_{cor}(t)$.

$$m_{op} = \text{floor}(t_w/t_{op}) + 2 \tag{9}$$

where $\text{floor}(\ast)$ is round-down function.

After phase space reconstruction, one-dimensional time series is converted into high-dimensional space. The matrix of high-dimensional space with m_{op} and t_{op} is shown in Equation (10).

$$X = \begin{bmatrix} x(1) & x(1+t_{op}) & \cdots & x(1+(m_{op}-1)t_{op}) \\ x(2) & x(2+t_{op}) & \cdots & x(2+(m_{op}-1)t_{op}) \\ \vdots & \vdots & \ddots & \vdots \\ x(M) & x(M+t_{op}) & \cdots & x(M+(m_{op}-1)t_{op}) \end{bmatrix} = \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(M) \end{bmatrix} \tag{10}$$

where $M = N - (m_{op} - 1)t_{op} - 1$.

Each row of Equation (10) represents each phase point and maintains equivalence with the original one-dimensional time series. Moreover, there exists a nonlinear mapping relation between X and Y according to the trend change of the chaotic attractor trajectory in [30], which can be expressed as:

$$Y = \begin{bmatrix} x(2+(m_{op}-1)t_{op}) \\ x(3+(m_{op}-1)t_{op}) \\ \vdots \\ x(M+1+(m_{op}-1)t_{op}) \end{bmatrix} = f \left(\begin{bmatrix} x(1) & x(1+t_{op}) & \cdots & x(1+(m_{op}-1)t_{op}) \\ x(2) & x(2+t_{op}) & \cdots & x(2+(m_{op}-1)t_{op}) \\ \vdots & \vdots & \ddots & \vdots \\ x(M) & x(M+t_{op}) & \cdots & x(M+(m_{op}-1)t_{op}) \end{bmatrix} \right) \tag{11}$$

where $f(\ast)$ is the nonlinear mapping function.

The following BLS prediction model is to find the nonlinear mapping relationship between X and Y . X is the input of BLS model, Y is the output of BLS model.

2.2. Broad Learning System

Compared with the traditional deep neural network [31], the structure of BLS is based on the random vector function-link neural network (RVFLNN) and is relatively simple [10]. It consists of an input layer, feature layer, enhancement layer, and output layer, as shown in Figure 1.

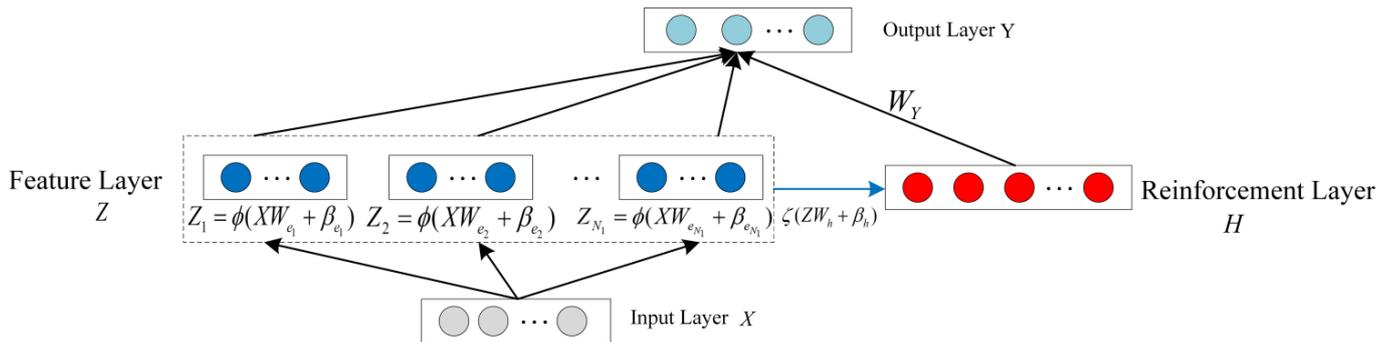


Figure 1. The network structure of BLS.

The feature layer is composed of N_1 feature windows, and each feature window contains N_2 feature nodes for feature extraction. The enhancement layer is only one layer and consists of N_3 strong nodes. Input data X are extracted by the feature layer into a series of feature nodes. The feature vector Z_i of the i -th feature window is shown in Equation (12).

$$Z_i = \phi(XW_{ei} + \beta_{ei}) \quad i = 1, 2, \dots, N_1 \tag{12}$$

where W_{ei} is feature weight part of the i -th feature window, β_{ei} is a bias part of the i -th feature window, and ϕ is the feature function. In this paper, ϕ is linear function [10].

The sparse property of data is beneficial not only to reduce the linear correlation of newly generated data, but also to improve the speed and efficiency of data operation. Therefore, sparse refactoring of each feature window is carried out based on least absolute shrinkage and selection operator (LASSO) in Equation (13).

$$W_{ei} = \underset{W_{ei}}{\operatorname{argmin}} \|XW_{ei} - Z_i\| + \lambda \|W_{ei}\|_1 \tag{13}$$

where λ is the regularization parameter, which is used to control the relative size of the second item. This paper uses cross-validation to obtain λ .

The first item is used to solve the least square estimation of linear equations $Z_i = XW_{ei}$. The second item uses the norm l_1 to limit the sparsity.

The feature vectors of all feature windows are combined into the feature layer vector $Z = [Z_1, Z_2, \dots, Z_{N_1}]$ after refactoring. The vector Z is mapped into the enhancement layer, and the enhancement layer vector H can be obtained through non-linear functions ζ in Equation (14).

$$H = \zeta(ZW_h + \beta_h) \tag{14}$$

where W_h is the weight of the connection between the feature layer and the enhancement layer, β_h is the bias term of the connection between the feature layer and the enhancement layer, and ζ is the non-linear mapping function. In this paper, the hyperbolic tangent function is adopted as the non-linear mapping function in Equation (15).

$$\zeta(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \tag{15}$$

Since the output layer is connected with the feature layer and the enhancement layer by weights W_Y , the feature layer vector Z and the enhancement layer vector H are concatenated into a new vector $E = [Z|H]$. Their connection relationship of output layer, feature layer, and the enhancement layer is shown in Equation (16).

$$Y = [Z|H] W_Y = E W_Y \quad (16)$$

In the process of BLS network training, W_{ei} , β_{ei} , W_h , and β_h are randomly generated and do not change during the training process. Therefore, it only calculates W_Y , which is solved in Equation (17).

$$W_Y = \underset{W_Y}{\operatorname{argmin}} \|E W_Y - Y\| = E^+ Y \quad (17)$$

where E^+ is the Moore–Penrose generalized inverse matrix of the matrix E .

2.3. Quantile Regression

Quantile regression does not need the best error probability distribution and is used to construct intervals of data [17]. Combined with the BLS model, the input data X and the output data Y are used to train the BLS model. After the BLS model is trained, the input data X is input into the model again; the predicted data Y_{pre} of the BLS model are output. The quantile regression model of the output data Y and predicted data Y_{pre} can be expressed in Equation (18).

$$Q_{F_Y}(\tau|F_{\text{pre}}) = \beta_0(\tau) + \beta_1(\tau)F_{\text{pre},1} + \beta_2(\tau)F_{\text{pre},2} + \dots + \beta_K(\tau)F_{\text{pre},K} = \beta(\tau)^T F_{\text{pre}} \quad (18)$$

where $Q_{F_Y}(\tau|F_{\text{pre}})$ represents the conditional quantile of the output variable F_Y given the explanatory variables $F_{\text{pre}} = [F_{\text{pre},1}, F_{\text{pre},2}, \dots, F_{\text{pre},K}]$, F_{pre} is variable of predicted data Y_{pre} , F_Y is variable of output data Y , τ is quantile, and $\beta(\tau)$ is the vector of quantile regression coefficients, which varies with the quantile τ . Reference [17] expands on the optimization solution of $\beta(\tau)$.

When $\beta(\tau)$ is determined, the confidence intervals of Y can be constructed. For example, the 90% confidence interval can be constructed by 0.05 quantile and 0.95 quantile. $\beta(0.05)^T Y_{\text{pre}}$ is the lower boundary and $\beta(0.95)^T Y_{\text{pre}}$ is the upper boundary.

3. Adaptive Rolling Error Correction Model Based on PSR-BLS-QR

The second part of this paper describes the models of PSR, BLS, and QR. The original wind power data is reconstructed by PSR to obtain the input and output data of BLS model. Then, the interval construction of the output data is realized by the BLS and QR models, and the initial wind power prediction interval is obtained. The adaptive rolling error correction model is mainly used to correct the initial wind power prediction interval and improve the accuracy of the prediction interval. The adaptive rolling error correction part will be further elaborated below.

3.1. Adaptive Rolling Error Correction

The one-dimensional wind power is transformed into high-dimensional phase space by PSR, and nonlinear mapping relationship of the wind power is established. Then, the input and output part of the BLS model is determined in Equation (11). BLS is used to train the input data and obtain the corresponding predicted data. Moreover, the power interval of the predicted data is constructed with the predicted data and the actual output data by the quantile regression model. Different quantile coefficients constitute different confidence intervals of the predicted data. For example, the 90% confidence interval is composed of the 0.05 quantile and 0.95 quantile prediction curves, and the upper and lower boundaries of the prediction are formed. The input and output results after BLS and QR are as follows:

$$X = \begin{bmatrix} x(1) & x(1+t_{op}) & \cdots & x(1+(m_{op}-1)t_{op}) \\ x(2) & x(2+t_{op}) & \cdots & x(2+(m_{op}-1)t_{op}) \\ & \vdots & & \vdots \\ x(M) & x(M+t_{op}) & \cdots & x(M+(m_{op}-1)t_{op}) \end{bmatrix} Y_{pr} = \begin{bmatrix} y_1(\tau_i), y_1(1-\tau_i) \\ y_2(\tau_i), y_2(1-\tau_i) \\ \vdots \\ y_M(\tau_i), y_M(1-\tau_i) \end{bmatrix} = \begin{bmatrix} L_1, U_1 \\ L_2, U_2 \\ \vdots \\ L_M, U_M \end{bmatrix} \quad (19)$$

where Y_{pr} is predicted interval power output at different confidence intervals, $y_i(\tau_i)$ is i -th lower boundary prediction value at quantile τ_i , which can be expressed simply by L_i ; $y_i(1-\tau_i)$ is i -th upper boundary prediction value at quantile $1-\tau_i$, which can be expressed simply by U_i .

Since the prediction interval is composed of upper and lower boundaries, the upper and lower boundaries are subtracted from actual output data to obtain the errors of the upper and lower boundaries. The error of the upper and lower boundaries for different confidence intervals are shown in Equation (20).

$$\begin{aligned} E_{up}(\tau_i) &= Y_U - Y_{ac} \\ E_{lo}(\tau_i) &= Y_{ac} - Y_L \end{aligned} \quad (20)$$

where $E_{up}(\tau_i)$ and $E_{lo}(\tau_i)$ are upper boundary and lower boundary error at quantile τ_i respectively, Y_U is prediction power of upper boundary, Y_L is prediction power of lower boundary, and Y_{ac} is actual power.

The traditional error correction is to subtract the predicted and actual values to obtain the error, establish the non-parametric kernel density cumulative distribution of all power error, and to find out the error correction power by the cumulative distribution. However, the error distributions of different power prediction range segments are different, such as the power prediction range segment between 0 kW and 200 kW and the power prediction range segment between 200 kW and 400 kW. Therefore, the error distribution of overall power cannot be used to replace the error distribution of different prediction power range segments, and the best error correction power of different power prediction range segments needs to be found respectively. Moreover, the error distributions of the upper and lower boundaries are also different. In this paper, the error distributions of different predicted power interval segments at the upper and lower boundaries will be constructed separately, and the error power corresponding to the quantile points of the distributions will be determined, and the optimal correction power for different power interval segments at the upper and lower boundaries will be chosen from the error power by the optimal correction function. In this paper, these quantile points of error distribution are 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, and 95%, respectively.

Based on the above, the adaptive rolling error correction can be understood as follows. The first phase point $X(1) = \{x(1), x(1+t_{op}), \dots, x(1+(m_{op}-1)t_{op})\}$ is input to the BLS model, and the predicted value y_1 can be obtained. At the same time, the prediction interval $[y_1(\tau_i), y_1(1-\tau_i)]$ of y_1 is also obtained under the set confidence interval by quantile regression model. Based on the predicted values of the upper and lower boundaries, the optimal error corrected power of the upper boundary p_{eup} and the optimal error corrected power of the lower boundary p_{elo} summed with the original prediction interval power to obtain the final result $[y_2(\tau_i) + p_{elo}, y_2(1-\tau_i) + p_{eup}]$. When reaching the time of $x(2+(m_{op}-1)t_{op})$, the actual values are added to the original time series and the point furthest away from the time series is removed to form the second phase point $X(2) = \{x(2), x(2+t_{op}), \dots, x(2+(m_{op}-1)t_{op})\}$. The second phase point is predicted to be completed in the manner described above. Based on this rolling pattern, the final prediction is completed.

3.2. Interval Evaluation Indexes

In this paper, two evaluation indexes are introduced to evaluate the performance of the proposed prediction method.

1. Prediction Interval Coverage Probability (PICP)

PICP calculates the probability of falling into the prediction interval and reflects the reliability of the prediction interval in Equation (21).

$$e_{\text{PICP}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \gamma_i \times 100\% \quad (21)$$

where N_p is the number of predicted samples and γ_i is represented in Equation (22).

$$\gamma_i = \begin{cases} 0 & y_{i,\text{ac}} \notin [L_i, U_i] \\ 1 & y_{i,\text{ac}} \in [L_i, U_i] \end{cases} \quad (22)$$

where $y_{i,\text{ac}}$ is the actual power of the sample.

2. Prediction Interval Normalized Average Width (PINAW)

Average bandwidth of prediction interval reflects the sharpness of the prediction in Equation (23). The wider the prediction width is, the less reasonable and effective information can be given.

$$e_{\text{PINAW}} = \frac{1}{N_p} \sum_{i=1}^{N_p} ((U_i - L_i) / y_{\text{max}}) \quad (23)$$

where y_{max} is the maximum value of wind power.

3. Optimal Correction Index

The optimization correction index is defined as Equation (24) to find the most optimal correction power of different prediction power ranges.

$$f = (e_{\text{PICP}}' - e_{\text{PICP}}) / e_{\text{PICP}} + (e_{\text{PINAW}} - e_{\text{PINAW}}') / e_{\text{PINAW}} \quad (24)$$

where e_{PICP} , e_{PINAW} are interval prediction evaluation indexes before correction e_{PICP}' , e_{PINAW}' are interval prediction evaluation indexes after correction.

3.3. The Process of Interval Prediction

Data preprocessing is also important before interval prediction. Data preprocessing avoids the influence of bad data on model accuracy and improves prediction accuracy. Therefore, the data needs to be cleaned before the data is divided and trained. The flowchart of the wind power interval prediction method with adaptive rolling error correction is shown in Figure 2, and the process is as follows:

1. The abnormal data of different datasets need be checked and corrected before training and testing;
2. PSR is used to reconstruct the wind power time series and the input and output data of BLS model are constructed;
3. Eighty percent of the input and output data is selected as training data and the rest is used for testing;
4. The optimal parameters of the BLS model are found by the grid search method and the training data is used for training the model of BLS;
5. Based on the training predicted value and the training actual value, the quantile regression model is used to determine the quantile coefficients of different confidence intervals, and the original wind power interval can be determined;
6. The nonparametric kernel density error distributions of different power interval segments of upper and lower prediction boundaries under different confidence intervals are established respectively. Moreover, the optimal error correction power of different power interval segments of upper and lower prediction boundaries under different confidence intervals can be found by the optimal error correction index;

7. BLS and QR are used to predict the new wind power curves to obtain the original power interval of wind power. The original power interval is corrected to obtain the final wind power interval according to the prediction power value.

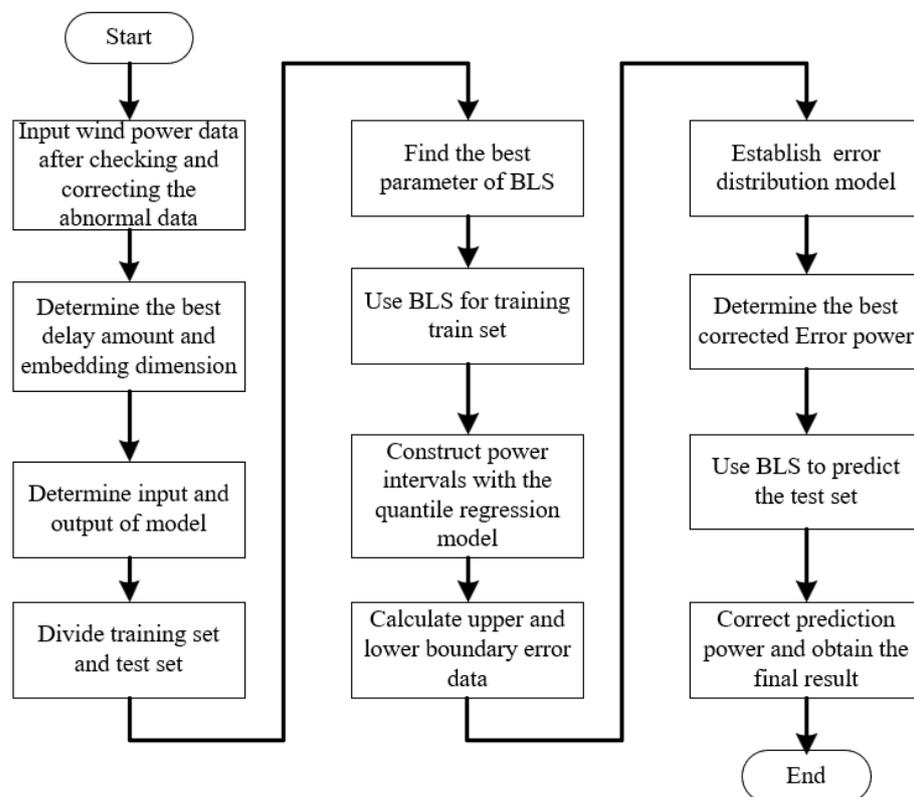


Figure 2. The process of the rolling prediction correction model.

4. Analysis of Examples

Prediction and algorithmic models used in this paper are based on Matlab 2018a. The operating system is a 64-bit Windows 10 system, the RAM is 12.0 GB, the CPU is Intel Core I7-5500U, 2.4 GHz.

4.1. Datasets and Data Division

In this paper, three wind power datasets labeled Set 1–3 from different regions are used to verify the effectiveness of the proposed method. The wind power curves of different datasets are shown in Figure 3.

Set 1: The first dataset is total wind power generation from Jiangsu Province. The rated power is 3.5 MW. The wind power data is selected from 1 January 2020 to 30 April 2020, with a sampling interval of 15 min, for a total of 121 days.

Set 2: The second dataset is total wind power generation from Hubei Province. The wind turbine model is EN-141/2.5. The rated power is 2.5 MW, the rated wind speed is 8.9 m/s, and the unit class is IEC 61400-1:2005S. The wind power data is selected from 1 January 2021 to 31 December 2021, with a sampling interval of 10 min, for a total of 365 days.

Set 3: The third database is the total wind power generation offshore from Elia, located in Belgium [32]. Set 3 involves the data collected once every 15 min from 1 January 2020 to 31 December 2020.

Wind data should be preprocessed before dividing the training set and testing set. When there are bad data such as null values, negative values, and abnormal values in the data sets, the data needs to be cleaned.

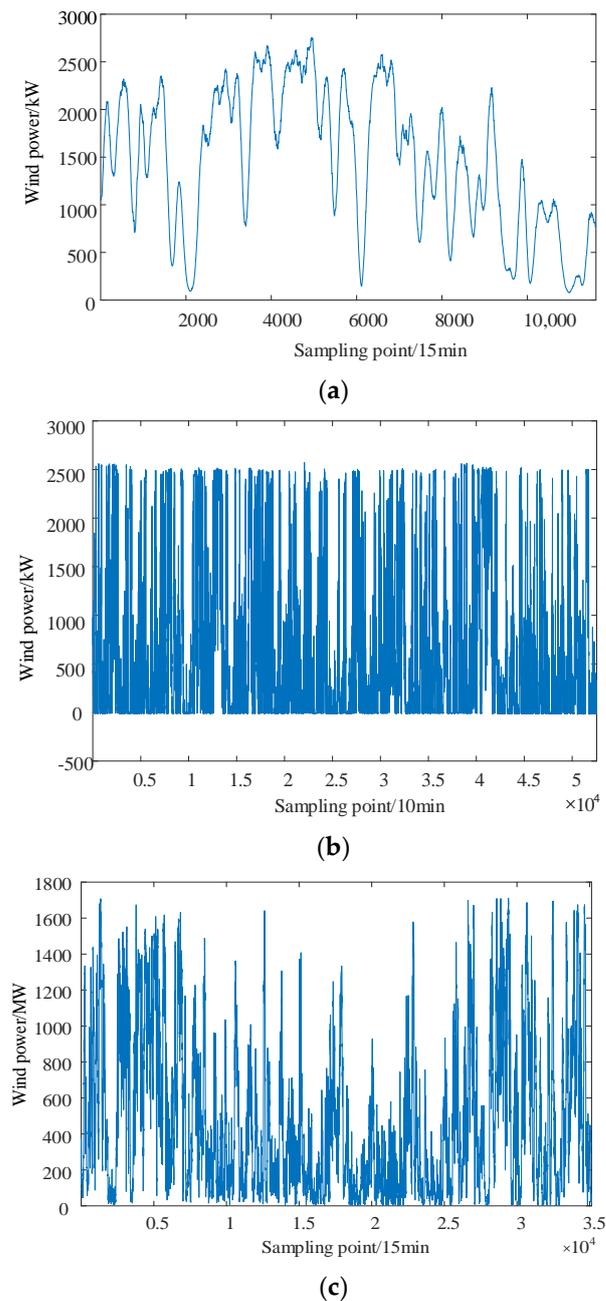


Figure 3. Wind power of different datasets: (a) Set 1; (b) Set 2; (c) Set 3.

In this paper, the data values of Set 1 and Set 3 are positive and are not null. Moreover, the data values of Set 1 and Set 3 are normal values by 3σ criteria test. Therefore, it can be considered that the data in Set 1 and Set 3 are normal values and can be directly used for training and testing. However, there is a negative value of low power at a certain moment in Set 2, which is due to the fact that the turbine itself did not cut into the wind speed, and power is provided for the wind turbine. Therefore, these negative power data values are directly assigned to 0 and data in Set 2 is normal by 3σ criteria test after data correction.

To fully consider the influence of seasonal characteristics on the prediction model, this paper constructs the probability density distribution of different seasons in different datasets and judges whether to construct different seasons' prediction models by the probability density distribution of different seasons. Since the data in Set 1 is from January to April, the seasonality of data is not strong. Therefore, the probability density of different seasons in Set 2 and Set 3 are explored, which is shown in Figure 4.

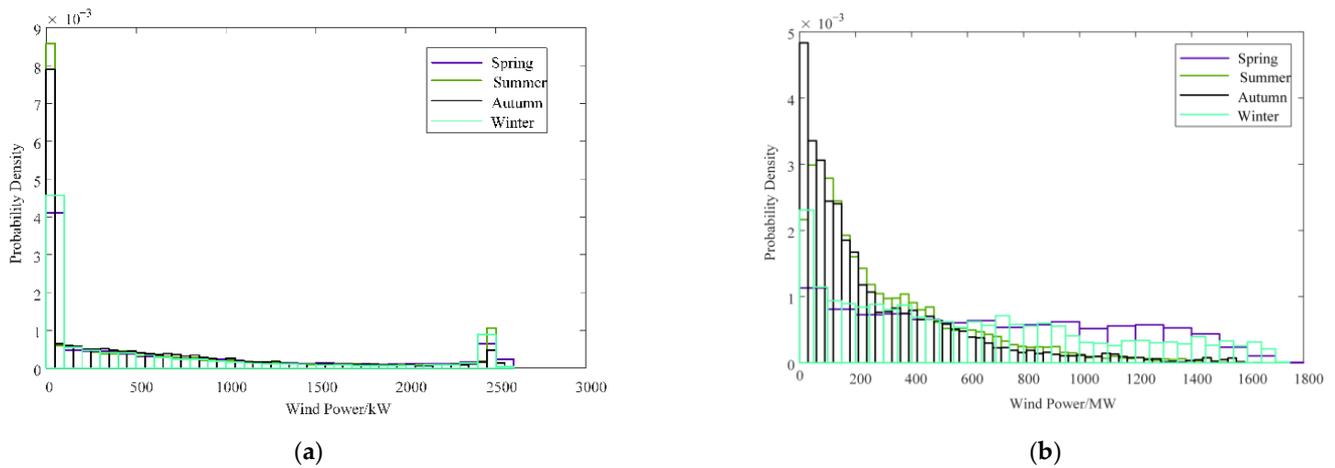


Figure 4. Probability density distributions of different seasons in different datasets: (a) Set 2; (b) Set 3.

As can be seen from Figure 4, the obvious changes in different seasons of Set 2 are not strong, and different seasons maintain a certain trend as a whole. Therefore, it is not necessary to carry out model predictions for Set 2 in different seasons. However, the seasonal changes of Set 3 are more obvious, and the change trends of different seasons are different.

Considering the characteristics of different datasets, the training and testing sets of different datasets are determined respectively. The first 80 days are selected as training data and the last 41 days are selected as test data for validation in Set 1. Twenty days of each month are selected to form a training set for model training, and the remaining days of each month form a test set for model testing in Set 2. The wind power in March, June, September, and December are used as the test set, and other months are used as the training set in Set 3. Models in different seasons are trained separately.

4.2. The Best Delay Amount and Embedding Dimension

Based on the phase space reconstruction theory, the C-C phase method is used to obtain the best delay amount and embedding dimension to realize the phase space reconstruction of the original one-dimensional time series. Taking the wind power of Set 1 as an example, $\Delta S(m, t)$ and $S_{cor}(t)$ of Set 1 are shown in Figure 5.

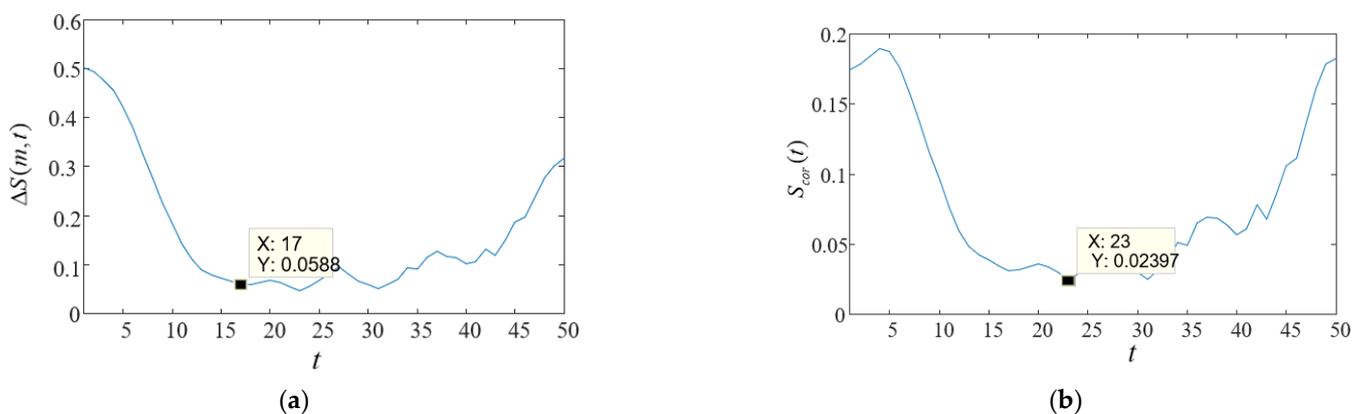


Figure 5. Statistics curves with C-C phase method: (a) $\Delta S(m, t)$; (b) $S_{cor}(t)$.

It can be seen that the best delay amount is 17. Combined with the sampling interval, the best delay time is 255 min. From Figure 5, we can see that the best delay window is 23. Therefore, the best embedding dimension is 3.

The best delay amount and the best embedding dimension of Set 2 are 8 and 4, respectively. The best delay amount and the best embedding dimension of Set 3 are 15 and 5, respectively. Based on the best delay amount and the best embedding dimension, we use Equations (10) and (11) to obtain the input and output of the BLS model.

4.3. The Effect of BLS Model

BLS can greatly reduce the complex structure of nonlinearities as well as reduce parameter tuning. To verify the effectiveness of the BLS model used in this paper, the BLS model is compared with traditional models, which include the traditional backpropagation neural network (BPNN), long short term memory (LSTM) network, radial basis function (RBF) neural network, kernel extreme learning machine (KELM), and extreme gradient boosting (XGBoost), respectively.

To verify the validity of the model used only, each model uses the same training and test set. Quantile regression is performed after a training model, and the initial interval of the model is constructed for model performance comparison. The adaptive error correction is not used. Optimal parameters of different models for different datasets are found by grid search method. A comparison of the average prediction results between different models in different datasets is shown in Table 1.

Table 1. Average accuracy comparison between different models in different datasets.

Model	Average Prediction Index	Set 1	Set 2	Set 3			
				March	June	September	December
BPNN	PICP	56.1	53.6	56.7	57.5	57.9	55.9
	PINAW	0.101	0.104	0.102	0.103	0.103	0.101
	Time/s	23.2	45.6	10.2	11.3	11.6	10.9
LSTM	PICP	56.1	54.1	57.1	57.8	57.6	56.8
	PINAW	0.101	0.105	0.102	0.103	0.103	0.102
	Time/s	44.5	90.1	17.3	17.6	15.1	16.0
RBF	PICP	58.2	54.5	56.9	57.1	57.3	56.8
	PINAW	0.101	0.105	0.101	0.102	0.103	0.101
	Time/s	162.2	302.5	50.1	50.2	50.1	50.1
BLS	PICP	60.8	59.4	62.3	63.2	63.3	61.4
	PINAW	0.099	0.103	0.101	0.102	0.102	0.101
	Time/s	0.7	1.9	0.4	0.4	0.4	0.4
KELM	PICP	56.6	55.1	57.6	58.1	58.6	56.9
	PINAW	0.100	0.103	0.101	0.103	0.103	0.100
	Time/s	1.5	3.5	0.8	0.9	0.8	0.8
XGBoost	PICP	57.4	56.3	58.6	59.3	60.0	57.3
	PINAW	0.102	0.104	0.102	0.104	0.104	0.102
	Time/s	7.5	20.3	6.2	5.9	6.7	6.3

It can be seen from Table 1 that, under the condition of only comparing the model effects, the prediction interval coverage probability and running time of BLS model are obviously better than the other five models with the same interval width or the narrower interval width in three different datasets. In Set 1 and Set 2, the predicted interval width of BLS model is 0.001–0.002 lower than that of the other models. However, the average accuracy of interval predictions improved by 2% to 6%. In Set 3, the width and coverage probability of the model interval are not the same in different seasons due to seasonal characteristics. However, no matter which season it is in, the width of the interval predicted

with the BLS model is still the narrowest, the prediction effect of the interval is also the best, and the running time is much lower than that of other models. This is mainly determined by the algorithm and structure of the model. Compared with other model structures, the weights and biases between the input layer and the middle layer of the model are randomly set, while the weights and biases can be directly solved by the Moore–Penrose generalized inverse matrix without using the loss function to modify the weights and biases. This structure and algorithm can ensure the accuracy and greatly reduce the running time of model training and testing. At the same time, there are only three parameters to be adjusted, and the model adjustment parameters are also greatly reduced compared with the other five models.

To further explore the detailed change of interval width, interval coverage probability, and the running time of different intervals in different datasets, this paper takes Set 1 as an example to show them in Tables 2 and 3. Table 2 shows the parameter settings of different models with different confidence intervals, and Table 3 shows an accuracy comparison between different models with different confidence intervals.

The BPNN in this paper adopts a three-layer hidden layer structure. The column of the BPNN model is the number of nodes in the first, second, and third hidden layers in order. The number of iterations is set to 300 and the activation functions are “tansig”, “logsig”, and “tansig” from the first hidden layer to the third hidden layer, respectively. The learning rate is set to 0.01, the target error is set to 0.001, and training function “trainlm” is used. Other parameters are default parameters. The LSTM model in this paper adopts the structure of a single-layer hidden layer. The number of nodes in the hidden layer with different confidence intervals is shown in the table. The number of iterations is set to 100, and other parameters are default parameters. The expansion speed of RBF model with different confidence intervals is shown in the Table 2, the target error is set to 0.01, and other parameters are default parameters. N_1 , N_2 , and N_3 of the BLS model with different confidence intervals are shown in the table in order. Because KELM is a single-layer structure, the number of single-layer nodes, kernel parameters, and penalty coefficients are shown in order. The XGBoost model parameter is the maximum depth of the tree in Table 2. The learning rate of XGBoost with different confidence intervals are all 0.1. The number of the generated maximum trees is set to 500. Other parameters are default parameters.

Table 3 shows the accuracy comparison results between different models with the model parameters of Table 2. It can be found that the interval width of the prediction interval increases as the confidence interval increases, and the prediction interval coverage probability also increases among all the models. Comparing the average values of the six models, it can be seen that the prediction using the BLS model can have more prediction points falling into the interval band in a narrower interval width, that is, the curve predicted using the BLS model can fit the actual power value better, meaning the interval band is well around the actual power value after using the quantile regression model. In contrast, the width of prediction interval is much wider than the other models when using XGBoost. When using the BPNN model, the prediction interval coverage probability is much less than other models, especially for low confidence intervals, such as the 10% confidence interval. The prediction interval coverage probability of 10% confidence interval is 4–6% lower than the proposed model within roughly the same interval width. In contrast, the interval width is much wider than the BLS method when using RBF, LSTM, and KELM. Taking 60% confidence interval as example, the interval coverage probability of three models is almost similar to the BLS model; however, the interval width is wider than the BLS model. This means that three models sacrifice interval width to improve prediction accuracy. Overall, the BLS model can improve 2–4% accuracy on average, with a narrower interval width when comparing with other models. The effect of the model is better than other traditional prediction models. In addition, due to its own structural characteristics, the running time of the model also has an absolute advantage. For short-term interval prediction, information can be obtained in time for faster processing.

Table 2. Parameter settings of different models with different confidence intervals in Set 1.

Confidence Interval	BPNN	LSTM	RBF	BLS	ELM	XGBoost
10%	[20, 15, 10]	30	0.1	[5, 10, 19]	[500, 0.5, 0.5]	25
20%	[17, 11, 8]	35	0.1	[4, 15, 21]	[500, 0.5, 0.5]	25
30%	[18, 14, 10]	32	0.1	[6, 3, 18]	[500, 0.5, 0.5]	25
40%	[15, 7, 3]	30	0.1	[10, 8, 18]	[500, 0.5, 0.5]	20
50%	[18, 15, 7]	35	1	[15, 7, 10]	[600, 1, 0.5]	20
60%	[14, 9, 7]	37	1	[8, 8, 12]	[600, 1, 1]	30
70%	[15, 12, 10]	35	1	[12, 12, 20]	[600, 1, 1]	30
80%	[18, 14, 10]	30	1	[4, 9, 25]	[600, 1, 1]	20
90%	[18, 15, 8]	32	1	[7, 14, 19]	[600, 1, 1]	20

Table 3. Accuracy comparison between different models in Set 1.

Confidence Interval	BPNN			LSTM			RBF		
	PICP	PINAW	Running Time/s	PICP	PINAW	Running Time/s	PICP	PINAW	Running Time/s
10%	21.9	0.017	25.3	23.3	0.018	38.5	24.2	0.018	164.5
20%	32.5	0.035	23.1	33.4	0.036	49.6	31.3	0.035	164.6
30%	43.8	0.068	24.2	45.7	0.067	48.7	46.7	0.068	160.1
40%	52.1	0.084	20.5	55.8	0.083	40.3	55.1	0.082	160.3
50%	59.5	0.103	23.9	61.3	0.104	36.1	65.3	0.102	168.6
60%	68.7	0.124	20.9	69.5	0.129	55.2	70.0	0.132	160.4
70%	69.9	0.138	21.7	72.9	0.131	50.1	73.3	0.137	160.1
80%	76.4	0.151	24.4	78.3	0.154	36.2	76.5	0.152	160.1
90%	80.3	0.181	25.1	80.9	0.179	45.7	81.6	0.180	161.2
Average	56.1	0.101	23.2	56.1	0.101	44.5	58.2	0.101	162.2
Confidence Interval	BLS			KELM			XGBoost		
	PICP	PINAW	Running Time/s	PICP	PINAW	Running Time/s	PICP	PINAW	Running Time/s
10%	28.1	0.018	0.9	22.2	0.018	1.1	23.9	0.019	7.6
20%	36.5	0.036	0.8	30.1	0.036	1.3	32.8	0.035	7.4
30%	49.1	0.069	0.5	44.3	0.068	1.4	45.1	0.069	7.4
40%	57.8	0.081	0.7	53.2	0.080	1.1	51.9	0.084	6.9
50%	65.6	0.096	0.7	60.1	0.098	1.7	61.5	0.102	8.1
60%	70.1	0.125	0.8	69.2	0.131	1.6	68.8	0.134	7.6
70%	76.1	0.136	0.9	71.2	0.139	1.7	72.3	0.139	7.8
80%	79.3	0.151	0.8	77.8	0.155	1.8	78.0	0.157	7.3
90%	84.6	0.178	0.5	81.2	0.179	1.6	82.3	0.183	7.3
Average	60.8	0.099	0.7	56.6	0.100	1.5	57.4	0.102	7.5

To verify the validity of the BLS model used in this paper, a curve is randomly selected as the predicted curve. The prediction results of 80% confidence intervals for different models are shown in Figure 6.

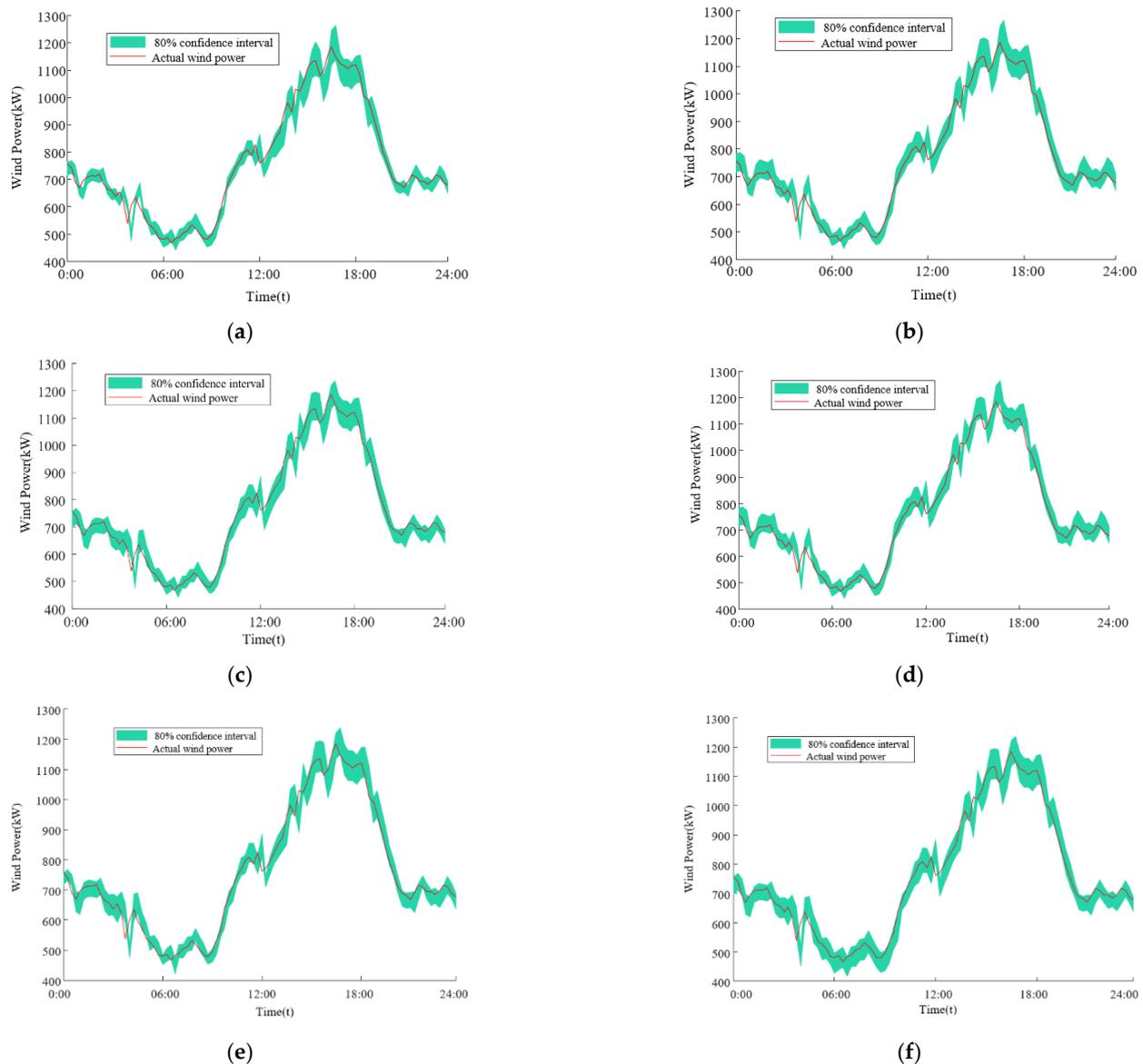


Figure 6. Prediction results of different models under 80% confidence interval: (a) BPNN; (b) LSTM; (c) BLS; (d) RBF; (e) KELM; (f) XGBoost.

As can be seen from the Figure 6, the predicted power intervals of all six models capture the trend of wind power output; however, there are some differences in the prediction accuracy. Compared with the six graphs, it can be found that the actual power points of wind power falling into the predicted interval band of BPNN are significantly lower than the other models when using BPNN, especially between 0:00 and 6:00. At that time, the narrow width of the formed interval results in the poor coverage of the predicted interval. However, the interval width of the XGBoost model is significantly wider than that of other models, which means that the interval width is sacrificed to ensure the accuracy of the prediction.

Although the width of the interval is basically the same as the width of the other four models, the prediction accuracy of the RBF model is not very high when the predicted power is relatively high such as 1100 kW and 1300 kW. While using LSTM and KELM, the effect is similar to BLS when predicting low power. However, the prediction width is significantly wider when predicting high power. The BLS ensures suitable prediction results in both low and high power intervals, with the best prediction coverage as well as interval width.

4.4. Error Correction

To further improve the prediction accuracy, this section uses error correction to improve the effect. The error correction is made separately for the predicted power interval segments of upper and lower boundaries. Moreover, 200 kW is the reference point for dividing the power segment. In other words, the entire power is divided into 0–200 kW, 200–400 kW, 400–600 kW and other interval segments. The cumulative error distribution of each power interval segment is established respectively.

Figure 7 shows the non-parametric kernel density probability density error distribution for the predicted power interval segment of 200–400 kW and 400–600 kW of the lower boundary at 80% confidence interval. It can be seen that the probability density error distributions of two power interval segments are different, therefore it is also necessary to obtain different error correction values for different predicted power ranges.

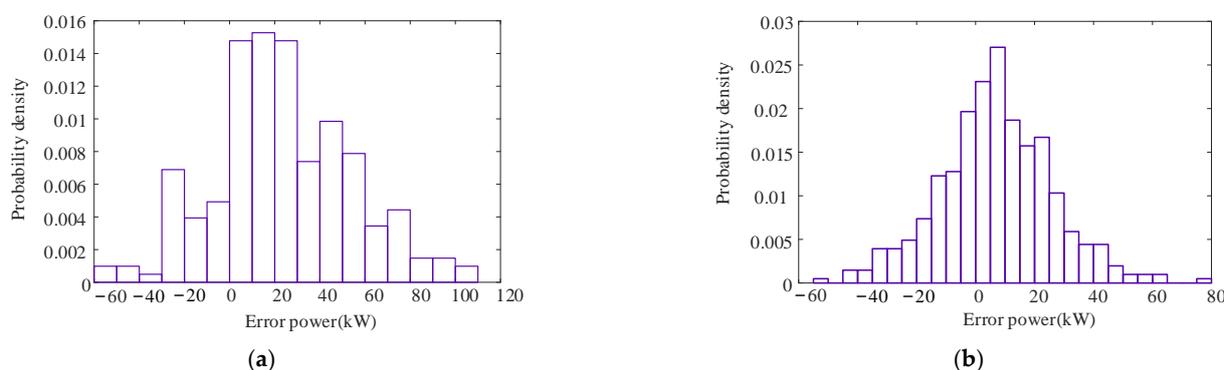


Figure 7. Probability density error distribution between different power interval segments: (a) 200–400 kW; (b) 400–600 kW.

It is worth stating that 200 kW is the reference point for dividing the power range in this paper. This is because the probability density function is different for each 200 kW interval in this paper. Of course, other values can be used for the division.

The optimal error correction for different power interval segments is determined by the optimal correction index, which is superimposed with the above predictions to obtain the final prediction value. A comparison of the before and after correction in different datasets is shown in Table 4.

Table 4. Accuracy comparison before and after error correction in different datasets.

Datasets	Before Error Correction		After Error Correction		Correction Improvement	
	PICP	PINAW	PICP	PINAW		
Set 1	60.8	0.099	68.9	0.100	12.3%	
Set 2	59.4	0.103	67.2	0.103	13.1%	
Set 3	March	62.3	0.101	70.5	0.102	12.2%
	June	63.2	0.102	71.4	0.102	13.0%
	September	63.3	0.102	72.3	0.103	13.3%
	December	61.4	0.101	69.3	0.102	11.9%

It can be seen from Table 4 that the prediction interval coverage probability has been significantly improved after the error correction in different datasets, while the interval width of the prediction remains unchanged or is wider after error correction. To verify that the effect after error correction is greatly better than that before correction, two indexes' values before and after correction are brought into Equation (24), which is shown as the correction improvement in the Table 4.

It can be seen from Table 4 that, no matter which dataset it is, the correction effect after error correction has been improved by more than 10%. Therefore, the effect after the correction is much better than the effect before the correction.

To further demonstrate the detailed effects of different datasets before and after error correction under different confidence intervals, this paper takes Set 1 as an example to show the change of two indexes before and after correction under different confidence intervals in Table 5.

Table 5. Prediction accuracy before and after error correction in Set 1.

Confidence Interval	Before Error Correction		After Error Correction	
	PICP	PINAW	PICP	PINAW
10%	28.1	0.018	35.8	0.023
20%	36.5	0.036	45.9	0.042
30%	49.1	0.069	56.7	0.072
40%	57.8	0.081	66.7	0.088
50%	65.6	0.096	75.9	0.118
60%	70.1	0.125	79.9	0.123
70%	76.1	0.136	82.3	0.131
80%	79.3	0.151	86.9	0.145
90%	84.6	0.178	90.3	0.168
Average	60.8	0.099	68.9	0.100

As can be seen from Table 5, the overall prediction after correction is slightly higher than the prediction before error correction. Although the interval width is 0.001 higher after correction than the interval width before correction, the prediction interval coverage probability is, on average, 8% higher than before correction, which means that the accuracy is 8% higher at any confidence interval. It can be found that the interval width is significantly wider than that before the correction between 10% and 50% low confidence interval, while the prediction coverage can improve by about 10%. This means that, at the low confidence interval, the increased interval width makes the prediction points that were outside the interval band fall into the interval band, making the prediction accuracy improve further. The interval width is narrower, and the prediction coverage is more accurate than that before the correction at the 60–90% confidence interval. In other words, the correction of the high confidence interval is to reduce the width of the interval that is relatively wide and enhance the width of the interval that is relatively narrow. Therefore, the accuracy of the prediction interval is improved while the whole width of the interval is reduced.

As can be seen from Figure 8, the power band with a narrower interval width is increased so that the narrower power band becomes a little wider, which makes the actual power closer to the interval band and more points fall into the interval. The prediction effect can be improved further. Taking region 1 and region 2 as an example, the interval width is wider after correction than that before correction, the points that originally did not fall into the interval can also fall into the interval. At the same time, when the predicted power interval can cover the actual power, it is also able to reduce the power part of the comparative width of the interval. In this way, the accuracy of the prediction effect can be guaranteed, and the width of the interval can be reduced. As in region 3, it is possible to cover the actual power with a narrower interval width.

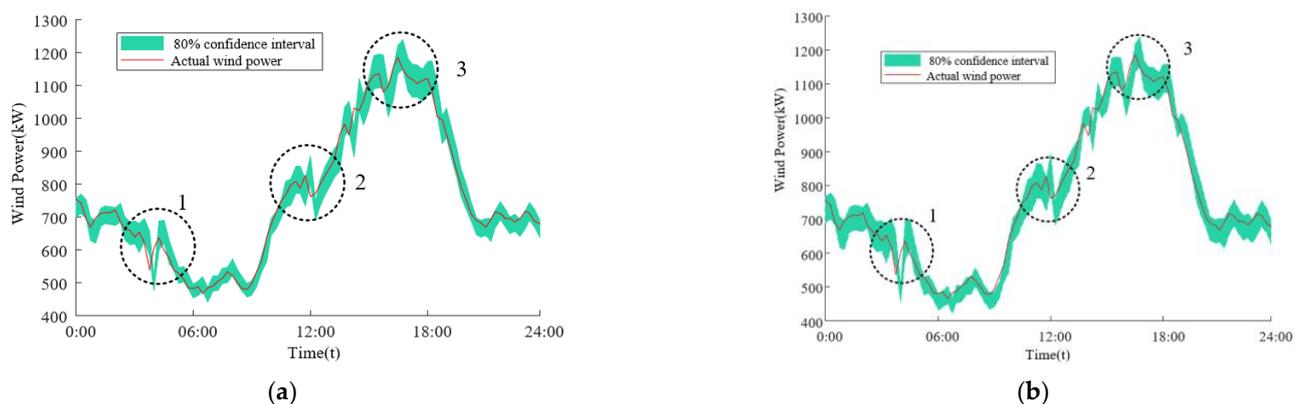


Figure 8. Power interval before correction and after correction. (a) Power interval before correction; (b) Power interval after correction.

4.5. Comparison of Interval Prediction Methods

In this paper, the initial wind power time series output intervals are constructed by the BLS-QR method, and then the final wind power intervals are obtained by using the adaptive rolling error correction model. To verify the effectiveness of the proposed method in this paper, it will be compared with the traditional interval prediction methods. The traditional interval prediction methods chosen in this paper are statistical probability interval prediction and quantile neural network, respectively. The statistical probability interval prediction method uses BLS for point prediction to obtain the corresponding error, and then constructs the interval directly by nonparametric kernel density, which is denoted as method 1. The quantile neural network is to use the quantile as the last layer of the LSTM, which forms LSTMQR, and directly outputs the data interval. The method with quantile neural network is denoted as method 2. The interval prediction method in this paper is denoted as The Proposed Method. The accuracy comparison using different methods in different datasets is shown in Table 6.

Table 6. Accuracy comparison with different interval prediction methods in different datasets.

Method	Average Prediction Index	Set 1	Set 2	Set 3			
				March	June	September	December
Method 1	PICP	55.7	53.4	60.1	62.8	63.1	60.5
	PINAW	0.101	0.106	0.103	0.104	0.104	0.103
Method 2	PICP	60.9	55.9	62.3	63.6	66.1	63.1
	PINAW	0.100	0.103	0.102	0.103	0.103	0.102
The Proposed Method	PICP	68.9	67.2	70.5	71.4	72.3	69.3
	PINAW	0.100	0.103	0.102	0.102	0.103	0.102

The interval width of method 1 is obviously wider than that of the other two methods no matter which dataset it is in; however, the prediction coverage probability is obviously lower than that of other methods in Table 6. In other words, the interval prediction effect with method 1 is obviously inferior to that of two other methods. The average width of the interval is the closest to that of the proposed method when using method 2. However, the prediction interval coverage probability is much worse than that of the proposed method. The interval prediction method proposed in this paper can improve the average accuracy of the interval by about 6–14% with the narrowest interval width. On the whole, the effect of the interval prediction method proposed in this paper is obviously better than that of the traditional interval prediction method.

To further show the different confidence intervals of different methods in different datasets in detail, this paper takes Set 1 as an example to show the comparison of indexes in different confidence intervals with different methods in Table 7.

Table 7. Accuracy comparison with different interval prediction methods in Set 1.

Confidence Interval	Method 1		Method 2		The Proposed Method	
	PICP	PINAW	PICP	PINAW	PICP	PINAW
10%	21.1	0.014	27.5	0.017	35.8	0.023
20%	29.6	0.032	34.1	0.034	45.9	0.042
30%	37.9	0.063	48.1	0.066	56.7	0.072
40%	42.9	0.075	56.3	0.079	66.7	0.088
50%	56.3	0.097	66.7	0.104	75.9	0.118
60%	69.6	0.127	72.1	0.129	79.9	0.123
70%	73.8	0.139	76.9	0.137	82.3	0.131
80%	81.5	0.167	80.3	0.156	86.9	0.145
90%	93.1	0.191	86.8	0.181	90.3	0.168
Average	55.7	0.101	60.9	0.100	68.9	0.100

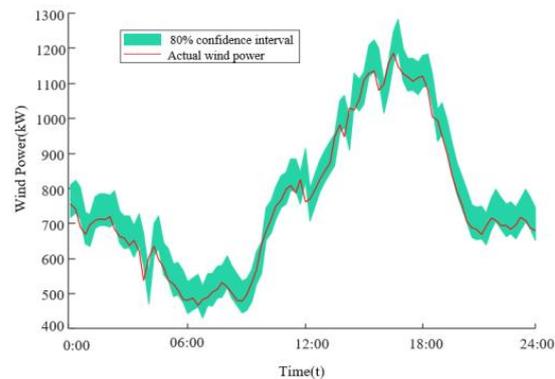
As can be seen from Table 7, the prediction interval coverage probability and interval width keep increasing as the confidence interval increases among all methods. In other words, the power point falls into the interval band by sacrificing the interval width. On the whole, the interval width of method 1 is the largest interval width among the three methods, but the interval coverage probability is much lower than the other methods. The use of the proposed method can improve the accuracy by about 8–13%, on average, with the narrowest interval width when compared with the other interval prediction methods.

The interval width of method 1 and method 2 is significantly narrower than the interval of the proposed method at the confidence interval between 10% and 50%; however, interval coverage probability is about 10–20% lower than the proposed method, especially for method 1. When the confidence interval is between 60% and 90%, the interval width of method 1 and method 2 gradually become wider, which significantly improves the prediction accuracy compared with the low confidence interval. However, the interval prediction accuracy is still lower than that of the proposed method. For example, the interval width of method 1 and method 2 are much wider than those of the proposed method at the 80% confidence interval, though the prediction results are not as good as those of the proposed method. This means that, although the interval width can be sacrificed to improve the interval prediction effect, the accuracy remains low due to the failure to increase the interval width at a reasonable location.

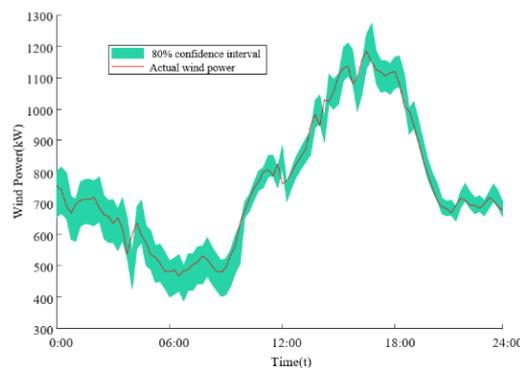
Therefore, the interval prediction method proposed in this paper can increase the interval width and reduce the interval width at a reasonable location, which can better fit the data itself and improve the prediction accuracy effectively.

Figure 9 shows the wind power interval predictions with different methods at 80% confidence interval. It is obvious from Figure 9 that the interval width of the proposed method is narrower than that of the other two methods. When the prediction power is low, between 0:00 and 6:00, it is able to ensure the same prediction coverage probability while the interval width is narrower. Moreover, the width of power interval is also narrower at high power, and the actual power is covered by predicted power interval and is closer to the middle of the interval. Method 1 is closer to the lower bound of the interval at high power and does not cover the actual power well. Therefore, the prediction is not as good as the proposed method. The prediction effect of method 2 is worse than the proposed method in some power prediction power, such as power between 700 and 800 kW. The

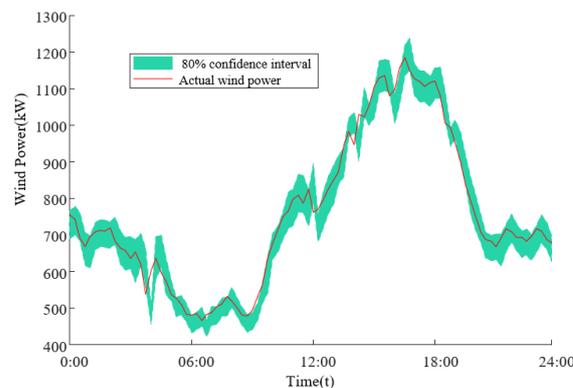
proposed method can make adaptive optimal corrections at the most suitable position. Therefore, the corrected interval can better cover the actual power.



(a)



(b)



(c)

Figure 9. Prediction results of 80% confidence intervals for different methods: (a) method 1, (b) method 2, (c) the proposed method.

5. Conclusions

Aiming at the solving problem of the low accuracy of traditional wind power interval prediction, this paper proposes a wind power interval prediction method with adaptive rolling error correction based on PSR-BLS-QR, which can accomplish adaptive error correction and improve the accuracy of wind power interval prediction. The conclusions are as follows:

1. The implicit characteristic information of the one-dimensional wind power is mined, and the correlation between the data can be constructed;

2. Due to the superiority of the BLS model compared with other models, it improves the interval prediction by about 4% accuracy at a narrower interval width compared with the traditional prediction model, and the running time of BLS has obvious advantages;
3. The adaptive error rolling correction model is used to make adaptive error corrections to further improve the interval prediction accuracy at the same or narrower interval width. Compared with the traditional interval prediction methods, the interval prediction accuracy can be improved by about 6–14%.

Author Contributions: X.R. analyzed the data and wrote the paper. C.X. performed the simulation and modeling; L.M. analyzed the results and reviewed the modeling and text; F.X. prepared the data. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by research on smart operation control technologies for offshore wind farms (Grant No. 2019YFE0104800), the special fund project of Central universities basic research business expenses (Grant No. B210201018), the National Natural Science Youth Foundation of China (Grant No. 52106238) and the National Natural Science Foundation of China (NSFC) (U1865101).

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

References

1. Niu, X.; Wang, J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. *Appl. Energy* **2019**, *241*, 519–539. [[CrossRef](#)]
2. Bhaskar, K.; Singh, S.N. AWNN-assisted wind power forecasting using feed-forward neural network. *IEEE Trans. Sustain. Energy* **2012**, *3*, 306–315. [[CrossRef](#)]
3. Alkesaiberi, A.; Harrou, F.; Sun, Y. Efficient wind power prediction using machine learning methods: A comparative study. *Energies* **2022**, *15*, 2327. [[CrossRef](#)]
4. Hong, D.Y.; Ji, T.Y.; Li, M.S.; Wu, Q.H. Ultra-short-term forecast of wind speed and wind power based on morphological high frequency filter and double similarity search algorithm. *Int. J. Electr. Power Energy Syst.* **2019**, *104*, 868–879. [[CrossRef](#)]
5. Yang, X.Y.; Xing, G.; Ma, X.; Fu, G. A model of quantile regression with kernel extreme learning machine and wind power interval prediction. *Acta Energ. Sol. Sin.* **2020**, *41*, 300–306.
6. Wang, H.Z.; Wang, G.B.; Li, G.Q.; Peng, J.C.; Liu, Y.T. Deep belief network based deterministic and probabilistic wind speed forecasting approach. *Appl. Energy* **2016**, *182*, 80–93. [[CrossRef](#)]
7. Naik, J.; Dash, P.K.; Dhar, S. A multi-objective wind speed and wind power prediction interval forecasting using variational modes decomposition based Multi-Kernel Robust Ridge regression. *Renew. Energy* **2019**, *136*, 701–731. [[CrossRef](#)]
8. Shrivastava, N.A.; Khosravi, A.; Panigrahi, B.K. Prediction Interval Estimation of Electricity Prices Using PSO-Tuned Support Vector Machines. *IEEE Trans. Ind. Inform.* **2015**, *11*, 322–331. [[CrossRef](#)]
9. Han, L.; Jing, H.; Zhang, R.; Gao, Z. Wind power forecast based on improved Long Short Term Memory Network. *Energy* **2019**, *189*, 116300. [[CrossRef](#)]
10. Mei, F.; Gu, J.Q.; Pei, X.; Zheng, J.Y. Photovoltaic interval prediction based on adaptive rolling matching prediction correction mode. *Electr. Power Autom. Equip.* **2022**, *42*, 92–98.
11. Li, C.; Tang, G.; Xue, X.; Chen, X.; Wang, R.; Zhang, C. The short-term interval prediction of wind power using the deep learning model with gradient descend optimization. *Renew. Energy* **2020**, *155*, 197–211. [[CrossRef](#)]
12. Liu, L.Y.; Meng, S.L.; Wu, J.J. Dynamic economic dispatch based on wind power forecast error interval. *Electr. Power Autom. Equip.* **2016**, *36*, 87–92.
13. Zhang, Z.; Qin, H.; Liu, Y.; Yao, L.; Yu, X.; Lu, J.; Jiang, Z.; Feng, Z.-K. Wind speed forecasting based on quantile regression minimal gated memory network and kernel density estimation. *Energy Convers. Manag.* **2019**, *196*, 1395–1409. [[CrossRef](#)]
14. Zhang, Y.; Pan, G.; Zhao, Y.; Li, Q.; Wang, F. Short-term wind speed interval prediction based on artificial intelligence methods and error probability distribution. *Energy Convers. Manag.* **2020**, *224*, 113346. [[CrossRef](#)]
15. Khosravi, A.; Mazloumi, E.; Nahavandi, S.; Creighton, D.; Van Lint, J.W.C. Prediction intervals to account for uncertainties in travel time prediction. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 537–547. [[CrossRef](#)]
16. Chen, X.; Dong, Z.; Meng, K.; Xu, Y.; Wong, K.P.; Ngan, H.W. Electricity price forecasting with extreme learning machine and bootstrapping. *IEEE Trans. Power Syst.* **2012**, *27*, 2055–2062. [[CrossRef](#)]
17. Mei, F.; Gu, J.; Lu, J.; Lu, J.; Zhang, J.; Jiang, Y.; Shi, T.; Zheng, J. Day-ahead nonparametric probabilistic forecasting of photovoltaic power generation based on the LSTM-QRA ensemble model. *IEEE Access* **2020**, *8*, 166138–166149. [[CrossRef](#)]
18. Van der Meer, D.W.; Shepero, M.; Svensson, A.; Widén, J.; Munkhammar, J. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian processes. *Appl. Energy* **2018**, *213*, 195–207. [[CrossRef](#)]
19. Bracale, A.; Caramia, P.; De Falco, P.; Hong, T. Multivariate quantile regression for short term probabilistic load forecasting. *IEEE Trans. Power Syst.* **2020**, *35*, 628–638. [[CrossRef](#)]

20. 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](#)]
21. Zhang, W.; Quan, H.; Srinivasan, D. An improved quantile regression neural network for probabilistic load forecasting. *IEEE Trans. Smart Grid* **2019**, *10*, 4425–4434. [[CrossRef](#)]
22. He, F.; Zhou, J.; Mo, L.; Feng, K.; Liu, G.; He, Z. Day-ahead short-term load probability density forecasting method with a decomposition-based quantile regression forest. *Appl. Energy* **2020**, *262*, 114396. [[CrossRef](#)]
23. Zhang, W.; Quan, H.; Srinivasan, D. Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination. *Energy* **2018**, *160*, 810–819. [[CrossRef](#)]
24. Huang, H.; Xu, Q.G.; Zhang, Y.Q.; Shi, H.W. Weight-LSSVM Financial crisis warning model based on KPCA dimension reduction. *Tatistics Decis. Mak.* **2020**, *36*, 180–184.
25. Li, C.; Cai, W.; Yu, C.; Zhao, R.; Zhang, Q. Electricity consumption behaviour analysis based on adaptive weighted-feature K-means-AP clustering. *IET Gener. Transm. Distrib.* **2019**, *13*, 2352–2361. [[CrossRef](#)]
26. Wang, Y.F.; Fu, Y.C.; Xue, H. Improved prediction method of PV output power based on optimised chaotic phase space reconstruction. *IET Renew. Power Gener.* **2020**, *14*, 1831–1840. [[CrossRef](#)]
27. Yang, G.; Zhang, H.; Zheng, H. Short-term photovoltaic power forecasting based on similar weather clustering and IHGWO-WNN-AdaBoost model. *High Volt. Eng.* **2021**, *47*, 1185–1194.
28. Wang, X.; Huang, K.; Zheng, Y.; Li, L.X.; Lang, Y.B.; Wu, H. Short-term forecasting method of photovoltaic output power based on PNN/PCA/SS-SVR. *Autom. Electr. Power Syst.* **2016**, *40*, 156–162.
29. Shi, T.; Mei, F.; Lu, J.; Lu, J.; Pan, Y.; Zhou, C.; Wu, J.; Zheng, J. Phase Space Reconstruction Algorithm and Deep Learning-Based Very Short-Term Bus Load Forecasting. *Energies* **2019**, *12*, 4349. [[CrossRef](#)]
30. Wang, J.; Zhou, J.; Peng, B. A hybrid neural genetic method for load forecasting based on phase space reconstruction. *Kybernetes* **2010**, *39*, 1291–1297.
31. Zhou, D.X. Universality of deep convolutional neural networks. *Appl. Comput. Harmon. Anal.* **2020**, *48*, 787–794. [[CrossRef](#)]
32. Elia, Wind Power Data. Available online: <http://www.elia.be/en/griddata/power-generation/wind-power> (accessed on 1 January 2020).