

Very Short-term Probabilistic Prediction Method for Wind Speed Based on ALASSO-nonlinear Quantile Regression and Integrated Criterion

Yan Zhou, Yonghui Sun[✉], *Member, IEEE*, Sen Wang, Linquan Bai, *Senior Member, IEEE*, Dongchen Hou, Rabea Jamil Mahfoud, and Peng Wang

Abstract—To enhance the performance of the prediction intervals (PIs), a novel very short-term probabilistic prediction method for wind speed via nonlinear quantile regression (NQR) based on adaptive least absolute shrinkage and selection operator (ALASSO) and integrated criterion (IC) is proposed. The ALASSO method is studied for shrinkage of output weights and selection of variables. Furthermore, for the better performance of PIs, composite weighted linear programming (CWLP) is proposed to modify the conventional linear programming cost function of quantile regression (QR), by combining it with Bayesian information criterion (BIC) as an IC to optimize the coefficients of PIs. Then, the multiple fold cross model (MFCM) is utilized to improve the PIs performance. Multistep probabilistic prediction of 15-minute wind speed is performed based on the real wind farm data from the northeast of China. The effectiveness of the proposed approach is validated through the performances' comparisons with conventional methods.

Index Terms—Composite weighted linear programming, least absolute shrinkage and selection operator, integrated criterion, wind speed.

I. INTRODUCTION

WITH increasing economic developments, wind energy has recently become an important and efficient renewable energy. However, wind power brings more uncertainties into generation, transmission and distribution than the conventional energies [1]. Thus, it is essential to improve the reliability of wind power generation [2]. So far, the research on uncertainty analysis has been increasing gradually for load forecast, electricity price forecast and so on, based on the effective mechanism and reasonable calculation model [3]–[8]. Wind power output basically relies on wind speed and as

well as the wind turbines' physical characteristics. The strong randomness of wind speed makes its power prediction a very challenging task. Accurate and reliable wind speed or power prediction, especially very short-term prediction is a robust tool to optimize the safety and reliability of the wind power system [9], [10]. Therefore, in the past years, a great deal of research has focused on wind energy prediction.

Recently, a great deal of attention has been given to wind power or wind speed point prediction [11]–[15]. However, the prediction errors are unavoidable, and cannot be eliminated due to the nonstationary and complexity of wind power or wind speed [16]. For solving this problem, prediction intervals (PIs) become very important for identifying potential risks and uncertainties quantification with a prescribed confidence level. Being different from the point prediction which quantifies the deterministic values, PIs can produce not only a range of observations, but also the indication of prediction accuracy [17]. Generally, the main steps of parametric PIs consist of point predictions and the error assumption [18]. Based on the assumption of error distribution, the parametric estimation for prediction errors can be approximately quantified [18]–[23]. However, those parametric prediction approaches suffer from the uncertainty and randomness of wind energy. Recently, the statistical methods of nonparametric assumptions have been proposed for interval forecast [24]. In [25] and [26], the direct interval prediction methods were analyzed for nonparametric PIs by minimizing the cost function. However, they suffered from the low efficiency due to the use of optimization algorithms [27].

To study the uncertainties of wind energy, the nonparametric PIs based on linear quantile regression (LQR) for different proportions have been developed as efficient methods [28], [29]. In [30], a novel nonlinear quantile regression (NQR) method combining LQR and the extreme learning machine (ELM) was proposed to quantify the boundary quantiles of PIs for wind power. The approach was based on linear programming (LP) [31], [32], which had the highest effectiveness compared with the nonlinear optimization algorithms. In [27], sensitivity analysis was utilized to optimize the NQR-based PIs for improving the sharpness, where the experiments demonstrated the high efficiency and effectiveness of machine learning-based linear programming (MLLP).

To further improve the performance, other methods for

Manuscript received October 9, 2020; revised November 10, 2020; accepted January 4, 2021. Date of online publication June 25, 2021; date of current version July 25, 2021. This work was supported by the National Key R&D Program of China (Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDK00KJJS1800266).

Y. Zhou, Y. H. Sun (corresponding author, email: sunyonghui168@gmail.com; ORCID: <https://orcid.org/0000-0002-9507-8673>), S. Wang, D. C. Hou, R. J. Mahfoud, and P. Wang are with the College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China.

L. Q. Bai is with Department of Systems Engineering and Engineering Management, the University of North Carolina, Charlotte NC 28223, USA.

DOI: 10.17775/CSEEJPES.2020.05370

reducing the influence of input variables with low correlation related to the output variables have been proposed in [33] and [34]. In [33], before building the prediction model, the partial autocorrelation function (PACF) was used to analyze the relationship between the variables, then the variables with higher correlations were chosen. The coefficients were optimized with a least absolute shrinkage and selection operator (LASSO) in the regression process without the prior calculation of correlation, unlike the PACF for analyzing the data before training the model. In [35], LASSO was proposed to optimize the cost function of the least squares estimation by adding a penalty to limit the values of output weights. In [36], Bayesian information criterion (BIC) was utilized to choose suitable coefficients for the optimal LASSO regression, where the calculation of penalty coefficients affects the accuracy of the prediction result. In [37], a point prediction method based on LASSO was proposed to predict solar power generation. In the LASSO method, the values of coefficients could be shrunk effectively to reduce the error. The penalty coefficients of variables were shrunk in the same multiple. However, it made the values of the coefficients shrink in an abnormal multiple. In [38], the LASSO penalty was modified to adaptively regulate weights of output variables in the least squares type objective function. This regulation improved the accuracy of point estimation without additional calculation in the linear model. Based on the aforementioned discussions, in this paper, a novel approach of nonparametric PIs for wind speed is proposed. The main contributions in this paper are as follows:

- 1) An adaptive least absolute shrinkage and selection operator (ALASSO) is proposed to shrink the values of output weights for the selection of rational variables, which can reduce the overfitting and influence of insignificant variables.
- 2) The composite weighted linear programming (CWLP) is proposed to optimize the forecasting performance, considering the reliability and sharpness, based on the efficient LP.
- 3) The integrated criterion (IC) is proposed, by combining CWLP and BIC, to optimize the coefficients considering the sparsity for the improved performance of PIs.
- 4) The multiple fold cross model (MFCM) is employed to improve the robustness of the PIs, and reduce the influence caused by coefficients randomness.

The rest of this paper is arranged as follows: In Section II, the conventional PIs formulation based on NQR is described. The proposed ALASSO-NQR method with the IC based on CWLP and BIC is presented in Section III for a better forecasting performance with MFCM. Also, the model construction process of the proposed method is given in this section. In Section IV, case studies are performed to verify the superiority of the proposed approach. Finally, conclusions and future study are provided in Section V.

II. PREVIOUS PIS FORMULATION WITH QR

In this section, the conventional ELM and quantile regression (QR) are demonstrated, respectively. The conventional LQR can effectively quantify the PIs via the LP cost function. The nonlinear method ELM can be used to improve the

accuracy of the linear method. Thus, the combination of the two methods for the construction of PIs is significant.

A. Extreme Learning Machine

ELM is an efficient single hidden-layer feedforward neural network [39]. Being different from the gradient-based approaches that require lots of iterations for the parameters' training, the biases and input weights are randomly chosen in the ELM. Then, according to the input data, the output will be quickly and easily calculated. The ELM approach can be formulated as:

$$g(\mathbf{x}_j, \mathbf{a}_i, b_i, w_i) = \sum_{i=1}^l w_i \varphi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i) = y_j, \quad j = 1, 2, 3, 4 \dots T \quad (1)$$

where $\mathbf{a}_i$ is the input weight vector connecting the i -th hidden neuron and the input neurons, and b_i is the bias of the i -th hidden neuron, $\mathbf{x}_j$ denotes the j -th input vector and y_j denotes the j -th output. T is the size of samples, $\varphi()$ is the activation function, l represents the number of hidden neurons, and w_i is the output weight connecting the i -th hidden neuron and the output. Then, the equation is represented as:

$$\mathbf{H}\mathbf{w} = \mathbf{Y} \quad (2)$$

where $\mathbf{w}$ denotes the output weight vector and $\mathbf{Y}$ denotes the vector of targets. $\mathbf{H}$ is the output of the hidden layer, formulated as follows:

$$\mathbf{H} = \begin{bmatrix} \varphi(\mathbf{a}_1 \cdot \mathbf{x}_1 + b_1) & \dots & \varphi(\mathbf{a}_l \cdot \mathbf{x}_1 + b_l) \\ \vdots & \ddots & \vdots \\ \varphi(\mathbf{a}_1 \cdot \mathbf{x}_T + b_1) & \dots & \varphi(\mathbf{a}_l \cdot \mathbf{x}_T + b_l) \end{bmatrix} \quad (3)$$

To the same input, $\mathbf{H}$ is kept unchanged. The purpose of the training is to quantify $\mathbf{w}$, expressed as:

$$\mathbf{w} = \mathbf{H}^+ \mathbf{Y} \quad (4)$$

$\mathbf{H}^+$ denotes the Moore-Penrose generalized inverse matrix of $\mathbf{H}$. After the activation of the functions, it is easy to be combined with the advanced approaches achieving high computational efficiency.

B. Upper and Lower Bounds with QR

In [40], LQR was studied to evaluate the distribution with quantiles. The quantiles can be quantified by minimizing the sum of the following absolute cost function with proportion α , as:

$$\min_{q_t^a \in R} \sum_{t=1}^T \rho_\alpha(y_t - q_t^a) \quad (5)$$

where $\rho_\alpha()$ is given by:

$$\rho_\alpha(x) = \begin{cases} \alpha x, & \text{if } x \geq 0 \\ (\alpha - 1)x, & \text{if } x < 0 \end{cases} \quad (6)$$

which is the piecewise linear function depicted in Fig. 1. Here, x denotes the difference between observations and prediction values.

q_t^a is replaced by the output of ELM, then, the modified cost function is given by:

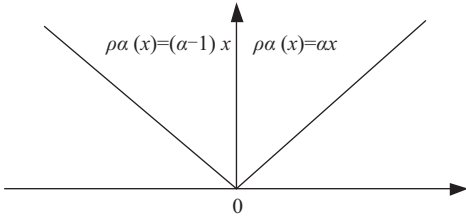


Fig. 1. Piecewise linear function.

$$\min_{\mathbf{w}_\alpha} \sum_{\alpha \in \{\underline{\alpha}, \bar{\alpha}\}} \sum_{t=1}^T \rho_\alpha(y_t - g(\mathbf{x}_t, \mathbf{w}_\alpha)) \quad (7)$$

subject to:

$$g(\mathbf{x}_t, \mathbf{w}_\alpha) - g(\mathbf{x}_t, \mathbf{w}_{\bar{\alpha}}) \leq 0, \quad \forall t \quad (8)$$

$$0 \leq g(\mathbf{x}_t, \mathbf{w}_\alpha) \leq 1, \quad \forall \alpha, t \quad (9)$$

where $g(\mathbf{x}_t, \mathbf{w}_\alpha)$ is the output of ELM and $\mathbf{x}_t$ denotes the t -th input vector. $\mathbf{w}_\alpha$ denotes the output weight vector with the PI nominal confidence (PINC) α . $\bar{\alpha}$ and $\underline{\alpha}$ denote the proportions of the upper and lower boundary quantiles respectively. For training the prediction model, searching for the output weights of quantiles as the bounds of PIs is the cost function's objective.

The quantiles for the PIs in ELM are depicted in Fig. 2. The wind speed time series are utilized as the input. The boundary quantiles can be efficiently calculated by NQR.

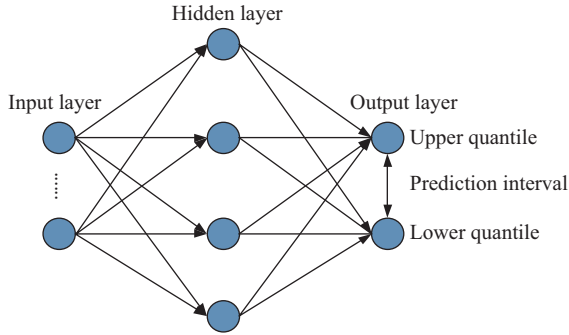


Fig. 2. Framework of ELM for boundary quantiles.

III. MODEL CONSTRUCTION OPTIMIZATION

In this section, the analysis of ALASSO-NQR is formulated for model selection. In addition, the CWLP function is proposed to fit the actual wind speed with preset proportions, where it is combined with BIC for a better PIs overall performance. In addition, MFCM is studied for PIs. With the above analysis, the process of PIs construction is demonstrated.

A. Optimal Prediction Intervals Via ALASSO

Based on NQR, PIs can be directly and efficiently quantified [30]. According to (7), the aim of the training process is to achieve the output weight vector $\mathbf{w}$. In the actual calculation, the result of $\mathbf{w}$ is not unique. The deviation of prediction can be reduced by shrinking the value of output weights [35]. By L_1 regulation, the LASSO penalty can be used to shrink the weights as:

$$\rho_\lambda(|\mathbf{w}|) = \lambda \sum_{j=1}^l |w_j| \quad (10)$$

where ρ_λ represents the penalty term of cost function, λ is the regulation coefficient, w_j denotes each element of the output weight vector. After that, the cost function of QR and LASSO is formulated as:

$$\min_{\substack{\mathbf{w}_\alpha, |\mathbf{w}_{\alpha,j}| \\ \bar{\xi}_t^\alpha, \xi_t^\alpha}} \sum_{\alpha \in \{\underline{\alpha}, \bar{\alpha}\}} \left\{ \sum_{t=1}^T [\alpha \bar{\xi}_t^\alpha + (1 - \alpha) \xi_t^\alpha] + \lambda \sum_{j=1}^l |w_{\alpha,j}| \right\} \quad (11)$$

where $\mathbf{w}_\alpha$, $|w_{\alpha,j}|$, $\bar{\xi}_t^\alpha$ and ξ_t^α can be obtained based on LP. $|w_{\alpha,j}|$ denotes the absolute value of the j -th output weight.

In statistical methods, the analysis between the response variables and the related variables are utilized for the improvement of regression performance [35]–[37]. The penalty of shrinkage in LP is to decrease the absolute values of the output weights for less deviation of prediction.

In the aforementioned LASSO-based QR method, λ in (11) is the penalty function coefficient to control the degree of constraint. In the conventional LASSO, the coefficient λ limits each number in output weights with the same multiple. It means that the weights of larger values shrink more, which reduces the accuracy of the coefficient's shrinkage. Consequently, the estimation of coefficients becomes biased.

The conventional input variables in the linear model are equivalent to the output variables $\mathbf{H}$ from the hidden-layer in (3) which can be easily obtained according to the calculation based on ELM. The output weights of $\mathbf{H}$ are shrunk and the elements of $\mathbf{H}$ are the objective variables of shrinkage in the nonlinear methods. And for each quantile, the target $\mathbf{H}$ is a vector containing the variables to be optimized. λ in (11) regulates the degree of penalty, but variables have different degrees of influence to the prediction result. Thus, it is significant to quantify the shrinkage of each variable in the rational degree.

To balance the shrinkage of different weights, (11) is changed as:

$$\min_{\substack{\mathbf{w}_\alpha, |\mathbf{w}_{\alpha,j}| \\ \bar{\xi}_t^\alpha, \xi_t^\alpha}} \sum_{\alpha \in \{\underline{\alpha}, \bar{\alpha}\}} \left\{ \sum_{t=1}^T [\alpha \bar{\xi}_t^\alpha + (1 - \alpha) \xi_t^\alpha] + \lambda T \sum_{j=1}^l \frac{|w_{\alpha,j}|}{|\beta_{\alpha,j}|} \right\} \quad (12)$$

subject to

$$g(\mathbf{x}_t, \mathbf{w}_\alpha) - g(\mathbf{x}_t, \mathbf{w}_{\bar{\alpha}}) \leq 0, \quad \forall t \quad (13)$$

$$0 \leq g(\mathbf{x}_t, \mathbf{w}_\alpha) \leq 1, \quad \forall \alpha, t \quad (14)$$

$$|w_{\alpha,j}| \geq 0, \quad \forall \alpha, j \quad (15)$$

$$-|\beta_{\alpha,j}| \leq w_{\alpha,j} \leq |\beta_{\alpha,j}|, \quad \forall \alpha, j \quad (16)$$

$$\bar{\xi}_t^\alpha, \xi_t^\alpha \geq 0, \quad \forall \alpha, t \quad (17)$$

$$-\xi_t^\alpha \leq y_t - g(\mathbf{x}_t, \mathbf{w}_\alpha) \leq \bar{\xi}_t^\alpha, \quad \forall \alpha, t \quad (18)$$

where $|\beta_{\alpha,j}|$ represents the absolute value of non-zero output weights in NQR without LASSO. ALASSO is utilized to shrink the non-zero weights of the variables. If the weight of an output variable is zero, the variable itself has nothing

to do with the output, so the weight does not participate in the shrinkage of ALASSO, and the weight is still zero. The coefficients $|\beta_{\alpha,j}|$ are adaptive regulation coefficients for the LASSO penalty to automatically balance the degree of shrinkage. The values of output weights can be shrunk in different multiples with the ALASSO penalty. Each multiple of the variable can be obtained by quantifying the output weights of NQR which reflects the significance degrees of the variables from the hidden layer. The larger absolute values of output weights in NQR should shrink less by multiples than the lower values of coefficients to reduce the deviation of shrinkage. In addition, in the LP of ALASSO-NQR, the weights of variables with less significance reflected by the output weights in NQR are easier to reach the zero value. The ALASSO method contributes to the sparsity of output weights for the accuracy of variables selection.

B. Integrated Criterion

1) Composite Weighted Linear Programming

In the QR method, (5) is often used as the basic cost function [41], [42]. The deviation of PIs coverage probability (PICP) is minimized for the optimal output weights by the efficient LP. Actually, the performance of PIs is evaluated by reliability and sharpness [21]. Thus, the sharpness is neglected in the cost function of LP. The CWLP is formulated by adding the regulation of sharpness to the LP cost function of ALASSO as:

$$\min_{\substack{\mathbf{w}_{\alpha}, |\mathbf{w}_{\alpha,j}| \\ \xi_t^{\alpha}, \underline{\xi}_t^{\alpha}}} \sum_{\alpha \in \{\bar{\alpha}, \underline{\alpha}\}} \left\{ \sum_{t=1}^T [\alpha \bar{\xi}_t^{\alpha} + (1 - \alpha) \underline{\xi}_t^{\alpha}] + \lambda_1 T \sum_{j=1}^l \frac{|w_{\alpha,j}|}{|\beta_{\alpha,j}|} \right\} + \lambda_2 G(\mathbf{w}_{\bar{\alpha}}, \mathbf{w}_{\underline{\alpha}}) \quad (19)$$

$$G(\mathbf{w}_{\bar{\alpha}}, \mathbf{w}_{\underline{\alpha}}) = \frac{1}{T} \sum_{t=1}^T g(\mathbf{x}_t, \mathbf{w}_{\bar{\alpha}}) - g(\mathbf{x}_t, \mathbf{w}_{\underline{\alpha}}) \quad (20)$$

where $G()$ represents the average width of the PIs in the training process. λ_1 denotes the shrinkage coefficient of ALASSO, and λ_2 denotes the regulation coefficient to balance the accuracy of PICP and the sharpness of the PIs for achieving better overall performance.

2) Sparsity Criterion of Output Weights

In ALASSO, the reasonable penalty coefficients shrinkage can shrink the coefficients values for obtaining better performance, and eliminate the variables that have little influence. Hence, the criterion of model construction along with coefficients optimization is significant. BIC is usually used for conventional model selection [43]. The cost function based on the sparsity of output weights can be formulated as:

$$\text{BIC} = \log(\sigma^2) + |S| \times \frac{\log(T)}{T} \quad (21)$$

where σ^2 is the mean square error of model training, $|S|$ denotes the model size. The main aim of BIC for coefficients selection is to integrate the conventional objective function as well as shrinkage of coefficients. BIC is used to balance the coverage accuracy of PIs and sparsity of output weights. Then the sparsity criterion $C(\mathbf{w})$ for ALASSO-NQR is given by:

$$C(\mathbf{w}) = \log \left\{ \sum_{t=1}^T \rho_{\alpha} [y_t - q_t^{\alpha}(\mathbf{w})] \right\} + \frac{\log(T)}{T} \times M \quad (22)$$

where $\mathbf{w}$ denotes the output vectors of ALASSO-NQR, M represents the number of non-zero coefficients in the output weights of the upper and lower bounds.

CWLP and BIC are integrated to improve the performance of PIs including the reliability, sharpness and sparsity. It is difficult to integrate the criteria directly with weighted coefficients. Therefore, CWLP is first utilized to quantify the optimal coefficients of output weights for reliability and sharpness. Then, the output weights are substituted into the BIC formula for the model construction.

C. Multiple Fold Cross Model

In the point prediction model, multiple-prediction for finding the output's mean value is an efficient way for determining true value in the regression [21]. In the proposed approach, MFCM is utilized for output weights training and prediction after the optimization of coefficients. The training samples are divided into multiple groups, and each group in turn is utilized for testing. The multiple ALASSO-NQR processes of different training samples are to quantify the true values of the PIs. The upper and lower quantiles in each prediction process are averaged respectively for the bounds of PIs, as:

$$q_{\alpha}^* = \frac{1}{F} \sum_{i=1}^F q_{\alpha,i}, \quad \alpha \in \{\bar{\alpha}, \underline{\alpha}\} \quad (23)$$

where F denotes the number of folds, $q_{\alpha,i}$ is the i -th upper or lower quantile in the method of folds with proportion α , and q_{α}^* represents the output. Meanwhile, the number of samples T in (19) is changed into T' , as:

$$T' = \frac{F-1}{F} T \quad (24)$$

Here, MFCM is applied to obtain more data information by multiple LPs and predictions to enhance the robustness of the PIs.

D. PIs Construction for the Proposed Method

The aim of the proposed approach is to quantify the boundary quantiles by ALASSO-NQR. In IC, CWLP is used to obtain the optimal performance of PIs considering reliability and sharpness, and BIC is utilized for the sparsity of output weights. There are two optimization methods in the model, LP and particle swarm optimization (PSO) [44]. LP is used for the QR based cost function. PSO is applied to optimize the coefficients with the nonlinear cost function. Following this, the optimal coefficients and output weights can be quantified. Then, the quantiles can be quantified. The main steps of the proposed approach are listed as follows:

Step 1: Normalize dataset for training, initialize the parameters of ELM.

Step 2: By (3), quantify $\mathbf{H}$ of samples. Change the samples $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^T$ into $D = \{(\mathbf{H}_i, y_i)\}_{i=1}^T$ preparing for training.

Step 3: By (7)–(9), apply LP for NQR based on CWLP for $\mathbf{w}_{\alpha}$ and λ_2 . The output weights in (7) are used as the regulation coefficients β_{α} in the ALASSO penalty.

Step 4: FOR each search agent of PSO for optimal λ_1 based on IC.

- 1) By (13)–(20), apply LP for ALASSO-NQR with CWLP criterion for the quantiles of PIs bounds.
- 2) Obtain the number of non-zero coefficients on the upper and lower quantiles in the output weights as M in (22).
- 3) Calculate the value of BIC as the criterion of the model selection. Record the value and the corresponding λ_1 .

Step 5: Based on the result of IC, choose the minimum and the corresponding λ_1 as the results of the coefficients optimization.

Step 6: With each group of samples and the coefficients in MFCM, ALASSO-NQR is constructed for the groups of output weights to end the training. Then, the upper and lower quantiles are quantified with average of prediction values for all groups.

Remark1. In the LASSO method, the output coefficients are shrunk in the same multiple [27], [35]. The output variables with larger coefficients are shrunk too much which causes prediction bias. In the proposed method, the ALASSO penalty is added to the cost function of NQR for the variable selection. With the shrinkage in different degrees for sparsity and preventing overfitting, the variables with low weights are eliminated. In addition, the shrinkage degrees of output weights can be regulated automatically without prior quantification of variables importance.

Remark2. The conventional QR is to minimize the deviation of PICP and PINC for optimal reliability of PIs [30]. However, the performance of PIs is evaluated by reliability and overall performance considering sharpness. Thus, CWLP is proposed to quantify PIs, considering both reliability and sharpness, by the efficient LP. In addition, by combining with BIC, the sparsity is also considered.

Remark3. In order to improve the robustness of PIs, compared with the PIs approach directly studying the training samples [30], the MFCM for ALASSO-NQR is proposed. The multiple folds of reconstructed samples are to diversify the training samples on the basis of the original dataset so that the accuracy of the PIs is improved.

IV. CASE STUDIES

A. Introduction of Dataset

In this part, the wind speed data from northeast of China are studied. The time series of wind speed with 15-minute resolution are utilized for very short-term PIs of wind speed. The PIs are based on the different proportions for a 1-hour look-ahead time. The wind speed time series of the first 1000-point per month represent the current month, and the time series of the last 1000-point from April, July, October, and January represent each season, respectively. Considering the sample size and prediction effectiveness, for ELM, the input length is 8, and the number of hidden nodes is 10. The results of different prediction approaches are used for comparison, and the dataset is used after normalization according to the maximum wind speed. Then the 385-point of the dataset is used for test and the rest are used for training.

B. PIs Evaluation Criteria

The performance of PIs is evaluated considering the reliability and overall performance, including sharpness [21]. The reliability which is evaluated by the absolute value of average coverage error (ACE) is defined as:

$$|ACE| = |PICP - PINC| \quad (25)$$

To comprehensively evaluate the forecasting performance, interval score [25]–[27] which considers the reliability and sharpness is given by:

$$\delta_t^\alpha(\mathbf{x}_i) = U_t^\alpha(\mathbf{x}_i) - L_t^\alpha(\mathbf{x}_i) \quad (26)$$

$$Sc_t^\alpha(\mathbf{x}_i) = \begin{cases} -2(1 - \alpha)\delta_t^\alpha(\mathbf{x}_i) - 4[L_t^\alpha(\mathbf{x}_i) - t_i], & \text{if } t_i < L_t^\alpha(\mathbf{x}_i) \\ -2(1 - \alpha)\delta_t^\alpha(\mathbf{x}_i), & \text{if } t_i \in I_t^\alpha(\mathbf{x}_i) \\ -2(1 - \alpha)\delta_t^\alpha(\mathbf{x}_i) - 4[t_i - U_t^\alpha(\mathbf{x}_i)], & \text{if } t_i > U_t^\alpha(\mathbf{x}_i) \end{cases} \quad (27)$$

$$Sc_t^\alpha = \frac{1}{T} \sum_{i=1}^T Sc_t^\alpha(\mathbf{x}_i) \quad (28)$$

where $Sc_t^\alpha(\mathbf{x}_i)$ is the score of each point, $U_t^\alpha(\mathbf{x}_i)$ and $L_t^\alpha(\mathbf{x}_i)$ are the upper and lower bounds respectively, $\delta_t^\alpha(\mathbf{x}_i)$ represents the width of each interval, t_i denotes the actual value, and Sc_t^α denotes the interval score.

C. Variables Selection of ALASSO

In the proposed method, adaptive regulation is implemented to improve the accuracy of output weights shrinkage for the sparsity. Here, the coefficients of output variables in NQR, LASSO-NQR and ALASSO-NQR are used for comparison.

In Fig. 3, (a) and (b) are the coefficients of output variables for the upper and lower quantiles from September with PINC 90% respectively, whereas (c) and (d) are from October. h_i represents the i -th output variable, which is the output of i -th hidden node and quantified by (3). For reasonable shrinkage and selection of variables, there are two criteria. One is that the larger values of absolute weights cannot be reduced to zero, while the other is that the smaller values of absolute weights, especially the weights close to zero, cannot be increased. The shrinkage performance of LASSO, can be observed through comparison with ALASSO according to the both aforementioned criteria. In (a), the weights of variables h_1 , h_2 , h_7 , and h_{10} are abnormal in the shrinkage of LASSO compared with ALASSO. The weights of h_1 and h_2 do not follow the second criterion, and the weights of h_7 and h_{10} do not follow the first criterion. In (b) for LASSO, the weights of h_8 do not follow the second criterion, and h_{10} does not follow the first criterion. In (c) for LASSO, all the weights are shrunk to the values close to 0, including the larger weights of variables which influence the accuracy of estimation. In (d), both LASSO and ALASSO are presented in the same shrinkage, but the upper quantiles shrinkage of ALASSO shown in (c) is better. Compared with NQR, the LASSO based method can shrink the weights to reduce bias, but the error of variable selection cannot be neglected. Compared with LASSO, the ALASSO based method has high accuracy for variable selection. The

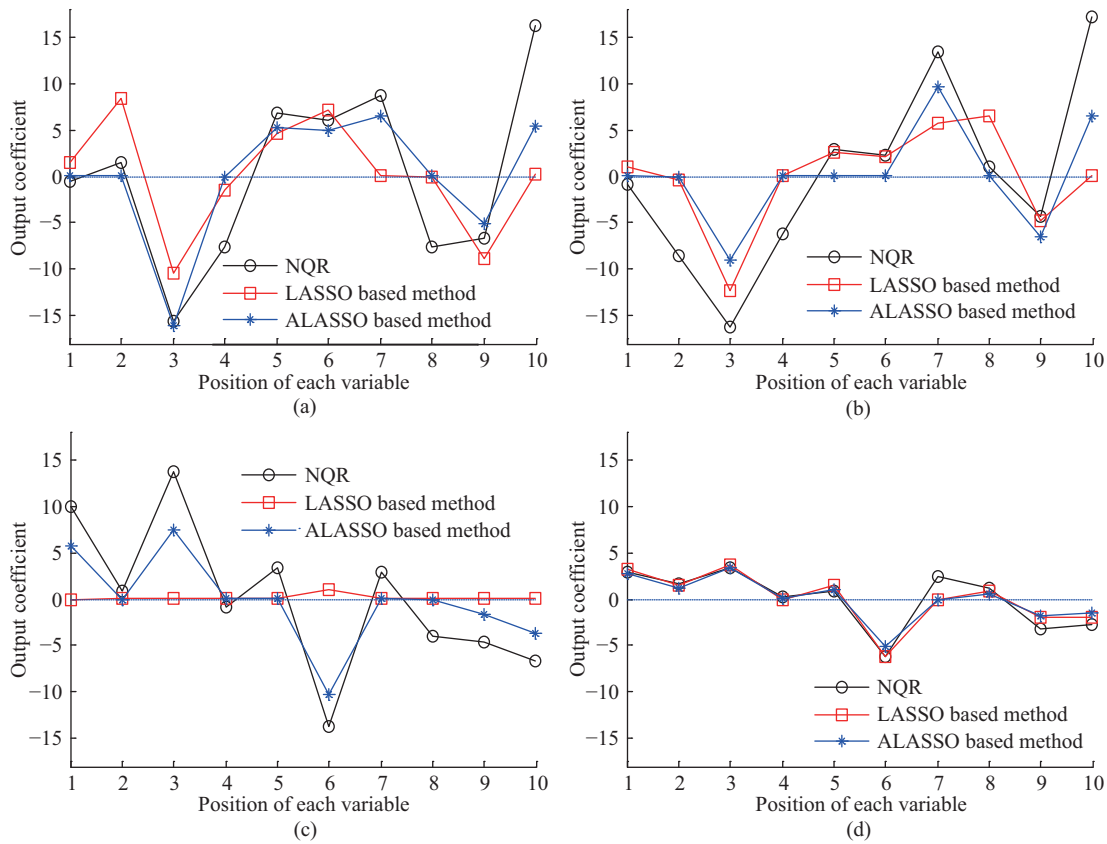


Fig. 3. Output coefficients of boundary quantiles. (a) Upper quantiles in September. (b) Lower quantiles in September. (c) Upper quantiles in October. (d) Lower quantiles in October.

performance of PIs is given in Tables I and II. According to the evaluation criteria of PIs, compared with NQR, the LASSO based method has better forecasting performance, and the ALASSO based method has the best performance.

D. Numerical Analysis of CWLP

To improve the overall performance of PIs, CWLP is proposed to achieve the enhanced interval score. The regulation coefficient λ_2 in (19) is calculated based on the optimal performance considering the interval score and reliability obtained by training. Based on the dataset from January, the interval scores and PICPs with PINC 90% in the training process are presented. As illustrated in Fig. 4 with the resolution of 0.01 to show the trend, the PIs with $\lambda_2 = 0.06$ have

TABLE I
RESULTS OF PIS IN SEPTEMBER

Method	PICP	AW	Score
NQR	84.42%	0.1848	-0.0613
LASSO	86.23%	0.1894	-0.0585
ALASSO	86.75%	0.1821	-0.0555

TABLE II
RESULTS OF PIS IN OCTOBER

Method	PICP	AW	Score
NQR	85.45%	0.1694	-0.0524
LASSO	84.94%	0.1691	-0.0518
ALASSO	86.23%	0.1699	-0.0514

the best performance considering interval score and reliability in training. The comparison between LQR, NQR and NQR-CWLP is given in Table III, where the NQR-CWLP has the best overall performance with similar reliability.

E. Analysis of Prediction Results

To verify the performances of ALASSO, IC and MFCM, the methods NQR, LASSO-NQR, ALASSO-NQR, ALASSO-NQR-IC are compared with the proposed method for numerical study in different months as shown in Table IV. The PIs are based on a 90% proportion for a 1-hour ahead period. According to the comparison, the proposed method has the best forecasting performance considering both reliability and

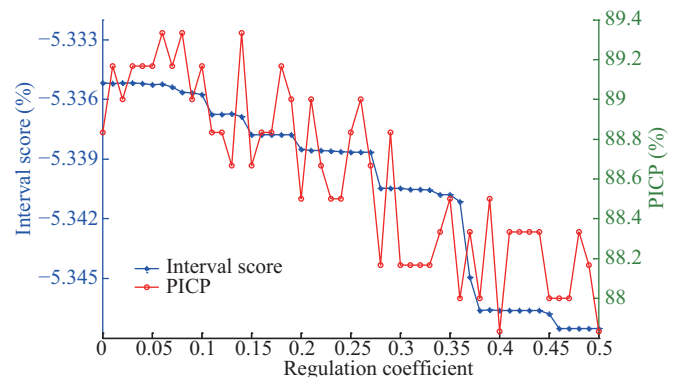


Fig. 4. PIs performances of CWLP coefficients.

TABLE III
RESULTS OF PIS IN JANUARY

Method	PICP	AW	Score
LQR	87.79%	0.2369	-0.0746
NQR	90.91%	0.2327	-0.0614
NQR-CWLP	91.69%	0.2202	-0.0572

overall performance of PIs, followed by ALASSO-NQR-IC. ALASSO-NQR has better performance than the LASSO-NQR, and ALASSO-NQR-IC is better than ALASSO-NQR.

It is well known that high complexity of chaotic climates contributes to the high level of uncertainty in wind speed. The uncertainty of weather conditions leads to extensive variations. To examine the validity and practicability of the proposed method, data from four seasons, spring (March to May), summer (June to August), autumn (September to November), winter (December to February) are considered respectively. Then, different approaches are used for comparison, such as bootstrap-based ELM (BELM) [21], LQR [40], direct quantile regression (DQR) [30], and MLLP [27] for a 1-hour ahead period on PINC 90%, 95% and 99%. In MLLP, the probability mass bias (PMB) selections in the training process help to choose the minimum interval width to improve the interval scores. The optimal PMB is chosen from the range [0, 1] in training. According to the results, it can be observed that the proposed method has the best forecasting performance. According to the results, the MLLP method is typically better than the DQR method, and close to the forecasting performance of the proposed method. The proposed method and MLLP are both considering the reliability and sharpness of PIs, combined with the conventional QR based methods. BELM is based on the error parameter modeling, and the performance of PIs is influenced by the error distribution.

Because of randomness and complexity, it is difficult to estimate the accurate error distribution with the prior assumption. According to the forecasting performances given in Table V, the results show that the performance of the proposed approach is valid with different proportions. Therefore, it is concluded that the proposed method is effective to highly improve the performance of PIs for very short-term forecasts of wind speed.

Figure 5 demonstrates PIs with a proportion of 90% in different seasons where the values of wind speed have been normalized. The PICP is as close to PINC as possible. Furthermore, according to Fig. 5, it can be easily noted that the sample dataset of wind speed time series has serious nonstationary characteristics and high frequency. It is difficult to quantify the PIs in the fixed error assumption. The methods based on NQR are reasonable and practical approaches with the calculation of the lower and upper bounds.

V. CONCLUSION AND FUTURE STUDY

In this paper, a novel nonparametric very short-term probability prediction method for wind speed based on ALASSO-NQR and IC was proposed. This paper basically completed the following studies: First, ALASSO was implemented for the reasonable coefficient shrinkage and variable selection, verified by numerical analysis compared to the LASSO based method and NQR. Secondly, the IC based on CWLP was proposed to optimize the NQR cost function in order to improve the composite performance, which was verified by the numerical comparison. Thirdly, the MFCM was utilized to expand the samples for multiple models so that the accuracy of prediction was improved. Finally, the effectiveness of the proposed approach was validated by the comparison with several previous approaches for the realistic wind speed data

TABLE IV
RESULTS OF PIS IN DIFFERENT MONTHS

Month	NQR			LASSO-NQR			ALASSO-NQR			ALASSO-NQR-IC			Proposed method		
	PICP	AW	Score	PICP	AW	Score	PICP	AW	Score	PICP	AW	Score	PICP	AW	Score
Jan.	90.91%	0.2327	-0.0614	89.35%	0.2414	-0.0610	91.17%	0.2345	-0.0605	91.17%	0.2180	-0.0556	90.65%	0.2001	-0.0551
Feb.	89.09%	0.2728	-0.0823	88.57%	0.2664	-0.0803	88.05%	0.2571	-0.0777	88.57%	0.2396	-0.0750	89.35%	0.2436	-0.0720
Mar.	87.01%	0.2435	-0.0674	88.57%	0.2516	-0.0665	87.79%	0.2438	-0.0647	86.23%	0.2294	-0.0644	88.57%	0.2289	-0.0637
Apr.	87.53%	0.2408	-0.0817	88.05%	0.2487	-0.0794	87.01%	0.3005	-0.0771	88.31%	0.2549	-0.0758	89.09%	0.2462	-0.0739
May	91.95%	0.3334	-0.0905	90.65%	0.3086	-0.0880	94.81%	0.3269	-0.0863	93.25%	0.3179	-0.0847	90.39%	0.2978	-0.0822
Jun.	87.27%	0.2531	-0.0665	87.01%	0.2441	-0.0658	88.57%	0.2573	-0.0650	87.79%	0.2492	-0.0645	89.35%	0.2531	-0.0637

TABLE V
RESULTS OF PIS IN DIFFERENT SEASONS

Season	PINC	BELM [21]			LQR [40]			DQR [30]			MLLP [27]			Proposed method		
		PICP	AW	Score	PICP	AW	Score	PICP	AW	Score	PICP	AW	Score	PICP	AW	Score
Spring	90%	88.31%	0.2293	-0.0606	90.39%	0.2549	-0.0644	93.51%	0.2610	-0.0594	91.69%	0.2301	-0.0549	90.13%	0.2175	-0.0514
	95%	94.55%	0.2726	-0.0349	94.81%	0.3297	-0.0391	97.14%	0.3088	-0.0340	97.66%	0.3019	-0.0320	95.84%	0.2540	-0.0293
	99%	97.66%	0.3532	-0.0090	98.18%	0.7117	-0.0158	100%	0.4548	-0.0091	100%	0.4417	-0.0088	99.74%	0.3799	-0.0078
Summer	90%	92.73%	0.1980	-0.0527	86.49%	0.1753	-0.0542	88.31%	0.1862	-0.0521	91.69%	0.1806	-0.0503	90.65%	0.1794	-0.0498
	95%	95.84%	0.2294	-0.0313	93.51%	0.2328	-0.0324	92.73%	0.2251	-0.0322	94.29%	0.2360	-0.0318	94.55%	0.2249	-0.0303
	99%	98.18%	0.3008	-0.0103	98.70%	0.4332	-0.0106	97.92%	0.3041	-0.0104	97.66%	0.2940	-0.0100	98.44%	0.2831	-0.0091
Autumn	90%	92.47%	0.2480	-0.0695	95.84%	0.3329	-0.0759	91.69%	0.2799	-0.0717	92.73%	0.2643	-0.0687	90.91%	0.2517	-0.0670
	95%	93.25%	0.3034	-0.0458	97.92%	0.4545	-0.0494	94.55%	0.3297	-0.0449	95.84%	0.3305	-0.0436	95.06%	0.3203	-0.0421
	99%	97.92%	0.3831	-0.0136	99.22%	0.7123	-0.0156	98.44%	0.4443	-0.0118	98.70%	0.4052	-0.0105	98.96%	0.4550	-0.0100
Winter	90%	96.36%	0.3051	-0.0713	80.00%	0.1929	-0.0727	91.95%	0.2780	-0.0715	92.21%	0.2731	-0.0698	89.35%	0.2507	-0.0689
	95%	97.66%	0.3523	-0.0426	85.97%	0.2471	-0.0514	93.77%	0.3192	-0.0454	96.10%	0.3362	-0.0435	94.03%	0.3278	-0.0417
	99%	99.48%	0.5164	-0.0122	93.51%	0.3611	-0.0230	96.88%	0.3506	-0.0159	98.18%	0.3948	-0.0125	99.22%	0.4162	-0.0094

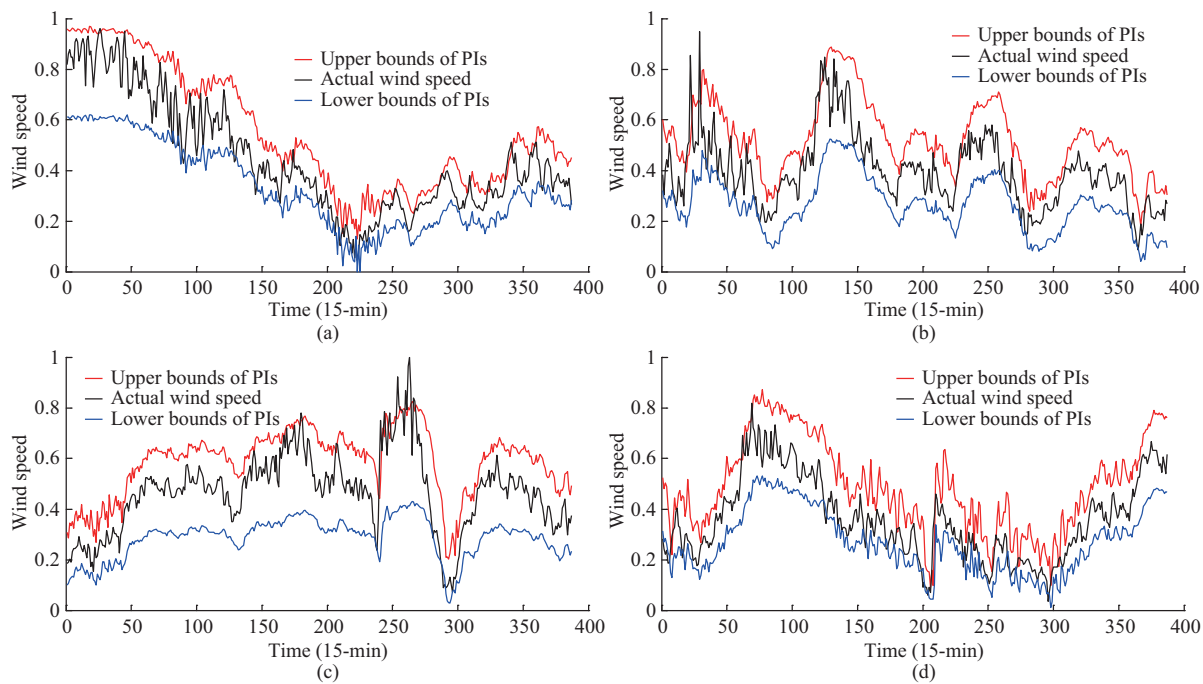


Fig. 5. PIs with PINC 90% in different seasons of 2009 obtained by the proposed method: (a) spring, (b) summer, (c) autumn, (d) winter.

of 15-min resolution for a 1-hour ahead period prediction.

The future study may focus on the advanced optimization of shrinkage coefficients with more effective and efficient approaches for LASSO-based methods, and the reasonable criteria for the performance of PIs.

REFERENCES

- [1] R. Bo and F. X. Li, "Efficient estimation of critical load levels using variable substitution method," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2472–2482, Nov. 2011.
- [2] C. F. Zhao, C. Wan, and Y. H. Song, "An adaptive bilevel programming model for nonparametric prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 424–439, Jun. 2019.
- [3] O. Abedinia, M. Zareinejad, M. H. Doranheghard, G. Fathi, and N. Ghadimi, "Optimal offering and bidding strategies of renewable energy based large consumer using a novel hybrid robust-stochastic approach," *Journal of Cleaner Production*, vol. 215, pp. 878–889, Apr. 2019.
- [4] M. Saeedi, M. Moradi, M. Hosseini, A. Emamifar, and N. Ghadimi, "Robust optimization based optimal chiller loading under cooling demand uncertainty," *Applied Thermal Engineering*, vol. 148, pp. 1081–1091, Jan. 2019.
- [5] W. Gao, A. Darvishan, M. Toghiani, M. Mohammadi, O. Abedinia, and N. Ghadimi, "Different states of multi-block based forecast engine for price and load prediction," *Electrical Power and Energy Systems*, vol. 104, pp. 423–435, Jan. 2019.
- [6] N. Ghadimi, A. Akbarimajd, H. Shayeghi, and O. Abedinia, "Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting," *Energy*, vol. 161, pp. 130–142, Oct. 2018.
- [7] H. Khodaei, M. Hajiali, A. Darvishan, M. Sepehr, and N. Ghadimi, "Fuzzy-based heat and power hub models for cost-emission operation of an industrial consumer using compromise programming," *Applied Thermal Engineering*, vol. 137, pp. 395–405, Jun. 2018.
- [8] H. A. Bagal, Y. N. Soltanabad, M. Dadjuo, K. Wakil, and N. Ghadimi, "Risk-assessment of photovoltaic-wind-battery-grid based large industrial consumer using information gap decision theory," *Solar Energy*, vol. 169, pp. 343–352, Jul. 2018.
- [9] R. Bo and F. X. Li, "Probabilistic LMP forecasting considering load uncertainty," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1279–1289, Aug. 2009.
- [10] T. Y. Hu, W. C. Wu, Q. L. Guo, H. B. Sun, L. B. Shi, and X. W. Shen, "Very short-term spatial and temporal wind power forecasting: A deep learning approach," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 2, pp. 434–443, Jun. 2020.
- [11] Y. S. Xue, C. Chen, K. Li, F. S. Wen, Y. Ding, Q. W. Wu, and G. Y. Yang, "Adaptive ultra-short-term wind power prediction based on risk assessment," *CSEE Journal of Power and Energy Systems*, vol. 2, no. 3, pp. 59–64, Sep. 2016.
- [12] S. X. Wang, X. Zhao, H. Wang, and M. Li, "Small-world neural network and its performance for wind power forecasting," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 2, pp. 362–373, Jun. 2020.
- [13] S. Saroha and S. K. Aggarwal, "Wind power forecasting using wavelet transforms and neural networks with tapped delay," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 2, pp. 197–209, Jun. 2018.
- [14] C. W. Potter and M. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 965–972, May 2006.
- [15] M. Javadi, A. M. Malyschiff, D. Wu, C. Q. Kang, and J. N. Jiang, "An algorithm for practical power curve estimation of wind turbines," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 1, pp. 93–102, Mar. 2018.
- [16] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. Mckeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, Jan. 2012.
- [17] J. T. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 748–757, Jun. 1997.
- [18] H. Bludszuweit, J. A. Dominguez-navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 983–991, Sep. 2008.
- [19] S. Tewari, C. J. Geyer, and N. Mohan, "A statistical model for wind power forecast error and its application to the estimation of penalties in liberalized markets," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2031–2039, Nov. 2011.
- [20] P. Pinson, "Very-short-term probabilistic forecasting of wind power with generalized logit—normal distributions," *Journal of the Royal Statistical Society Series C-Applied Statistics*, vol. 61, no. 4, pp. 555–576, Aug. 2012.
- [21] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1033–1044, May 2014.
- [22] A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton, and J. W. C. Lint, "Prediction intervals to account for uncertainties in travel time prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 537–547, Jul. 2014.
- [23] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," *Proceedings of 1994 IEEE International Conference on Neural Networks*, vol. 1, pp. 55–60, Jan. 1994.

- [24] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renewable and Sustainable Energy Reviews*, vol. 31, no. 2, pp. 762–777, Mar. 2014.
- [25] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Direct interval forecasting of wind power," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4877–4878, Nov. 2013.
- [26] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1166–1174, May 2014.
- [27] C. Wan, J. H. Wang, J. Lin, Y. H. Song, and Z. Y. Dong, "Nonparametric prediction intervals of wind power via linear programming," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1074–1076, Jan. 2018.
- [28] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, Jan. 2010.
- [29] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renewable Energy*, vol. 40, no. 1, pp. 29–39, Apr. 2012.
- [30] C. Wan, J. Lin, J. H. Wang, Y. H. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767–2778, Jul. 2016.
- [31] G. B. Huang, H. M. Zhou, X. J. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, Oct. 2011.
- [32] N. Y. Liang, G. B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, Nov. 2006.
- [33] Y. H. Sun, P. Wang, S. W. Zhai, D. C. Hou, S. Wang, and Y. Zhou, "Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution," *Wind Energy*, vol. 23, no. 1, pp. 63–76, Jan. 2020.
- [34] L. X. Lian, A. Liu, and V. K. N. Lau, "Weighted LASSO for sparse recovery with statistical prior support information," *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1607–1618, Mar. 2018.
- [35] R. Tibshirani, "Regression shrinkage and selection via the LASSO: A retrospective," *Journal of the Royal Statistical Society Series B-Statistical Methodology*, vol. 73, no. 3, pp. 273–282, Apr. 2011.
- [36] W. H. Zhao, R. Q. Zhang, and J. C. Liu, "Sparse group variable selection based on quantile hierarchical Lasso," *Journal of Applied Statistics*, vol. 41, no. 7, pp. 1658–1677, Feb. 2014.
- [37] N. K. Tang, S. W. Mao, Y. Wang, and R. M. Nelms, "Solar power generation forecasting with a LASSO-based approach," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1090–1099, Apr. 2018.
- [38] H. S. Wang and C. L. Leng, "A note on adaptive group lasso," *Computational Statistics and Data Analysis*, vol. 52, no. 12, pp. 5277–5286, Aug. 2008.
- [39] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, May 2006.
- [40] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, Jan. 1978.
- [41] X. G. Agoua, R. Girard, and G. Kariniotakis, "Probabilistic model for spatio-temporal photovoltaic power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 780–789, Apr. 2019.
- [42] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation forecasts," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 602–610, Apr. 2013.
- [43] E. R. Lee, H. Noh, and B. U. Park, "Model selection via Bayesian information criterion for quantile regression models," *Journal of the American Statistical Association*, vol. 109, no. 505, pp. 216–229, Jan. 2014.
- [44] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *1995 IEEE International Conference on Neural Networks Proceedings*, vol. 4, pp. 1942–1948, Nov. 1995.



Yonghui Sun (M'16) received the Ph.D. degree from the City University of Hong Kong, Hong Kong, China, in 2010. He is currently a Professor with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. He has authored more than 80 papers in refereed international journals. His research interests include stability analysis and control of power systems, optimal planning and operation of integrated energy system, optimization algorithms, and data analysis.



Sen Wang is currently pursuing the Ph.D. degree in Electrical Engineering with Hohai University, Nanjing, China. His research interests include data analysis, renewable energy grid-connected control, power system operation optimization, and renewable energy generation forecasting.



Linquan Bai (S'13–M'17–SM'20) received the B.S. and M.S. degrees in Electrical Engineering from Tianjin University, Tianjin, China, in 2010 and 2013, respectively, and Ph.D. degree at the University of Tennessee, Knoxville, USA. He is an Assistant Professor at University of North Carolina at Charlotte, Charlotte, USA. His research interests include power system operation and optimization, electricity markets, and integrated energy systems.



Dongchen Hou received the B.S. degree in Electrical Engineering and Automation from North China University of Water Resources and Electric Power, Henan, China, in 2017. He is currently pursuing the Ph.D. degree in Electrical Engineering with Hohai University, Nanjing, China. His research interests include theoretical and algorithmic studies in power system estimation, parameters identification, power system dynamics, signal processing, and cyber security.



Rabea Jamil Mahfoud received the Ph.D. degree in Electrical Engineering at Hohai University, Nanjing, China, in 2020. His current research interests include optimal operation and planning of power systems, optimization algorithms, and data analysis.



Yan Zhou is currently pursuing the Ph.D. degree in Electrical Engineering with Hohai University, Nanjing, China. His research interests include data analysis, machine learning, and renewable energy generation forecasting.



Peng Wang received the M.S. degree in Electrical Engineering at Hohai University, Nanjing, China, in 2020. His research interests include machine learning and renewable energy generation forecasting.