Direct Quantile Regression for Nonparametric Probabilistic Forecasting of Wind Power Generation

Can Wan, *Member, IEEE*, Jin Lin, *Member, IEEE*, Jianhui Wang, *Senior Member, IEEE*, Yonghua Song, *Fellow, IEEE*, and Zhao Yang Dong, *Senior Member, IEEE*

***Abstract*—The fluctuation and uncertainty of wind power generation bring severe challenges to secure and economic operation of power systems. Because wind power forecasting error is unavoidable, probabilistic forecasting becomes critical to accurately quantifying the uncertainty involved in traditional point forecasts of wind power and to providing meaningful information to conduct risk management in power system operation. This paper proposes a novel direct quantile regression (DQR) approach to efficiently generate nonparametric probabilistic forecasting of wind power generation combining extreme learning machine (ELM) and quantile regression. Quantiles with different proportions can be directly produced via an innovatively formulated linear programming optimization model, without dependency on point forecasts. Multi-step probabilistic forecasting of 10-min wind power is newly carried out based on real wind farm data from Bornholm Island in Denmark. The superiority of the proposed approach is verified through comparisons with other well-established benchmarks. The proposed approach forms a new artificial neural network-based nonparametric forecasting framework for wind power with high efficiency, reliability, and flexibility, which can be beneficial to various decision-making activities in power systems.**

***Index Terms*—Wind power, probabilistic forecasting, extreme learning machine, quantile regression, linear programming, uncertainty.**

NOMENCLATURE

$\mathbf{a}$	Input weights of ELM.
b	Biases of ELM.
C_N	Cost function of traditional neural networks.
D	Dataset of wind power.
f, $\hat{f}$	Probability density function (PDF), estimated PDF.
F, $\hat{F}$	Cumulative distribution function (CDF), estimated CDF
$g(\ )$	Output function of ELM.
$\mathbf{H}$	Hidden layer output matrix of ELM.
$\mathbf{H}^{\dagger}$	Moore-Penrose generalized inverse of ELM's hidden layer output matrix.
i, j	Common indices.
k	Prediction horizon.
L	Number of ELM's hidden nodes.
m	Dimension of the output vector of ELM.
n	Dimension of the input vector of ELM.
N	Number of training dataset for ELM.
$\hat{P}_{\mathrm{Q}}^{\alpha}$	Empirical proportion of estimated quantile.
PD_{Q}^{α}	Average proportion deviation.
$\Pr(\ )$	Probability operator.
P_n	Nominal wind capacity.
q, $\hat{q}$	Quantile, quantile forecast.
r	Number of estimated quantiles.
$S_{Q,t}$	Score for nonparametric probabilistic forecasting evaluation at a single time point.
S_Q	Score for nonparametric probabilistic forecasting evaluation.
t	Time index.
T	Size of time series.
w_{α}	Decision variable vector.
y, $\hat{y}$	Wind power measurement, prediction.
$\mathbf{Y}$,$\mathbf{y}$	ELM outputs/targets.
x	Explanatory variables of wind power prediction.
$\mathbf{x}$	ELM input variables.
α	Nominal proportion of the quantile.
γ	Parameters of conditional quantile function.
ε	Wind power forecasting error.

This work was partially supported by National High-Tech Research and Development Program (863 Program) of China (2014AA051901), International S&T Cooperation Program of China (2014DFG62670), China Postdoctoral Science Foundation (2015M580097), National Natural Science Foundation of China (51261130472, 51577096), and Hong Kong RGC Theme Based Research Scheme (T23-407/13-N, T23-701/14N). J. Wang's work is supported by the U.S. Department of Energy (DOE)'s Office of Electricity Delivery and Energy Reliability.

C. Wan and Y. Song are with the College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: can.wan@ieee.org).

L. Jin is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China.

J. Wang is with Argonne National Laboratory, 9700 S. Cass Avenue, Lemont, IL, USA (e-mail: jianhui.wang@anl.gov).

Z. Y. Dong is with the School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia.

η	Indicator of quantiles.
$\rho_\alpha(\,)$	Nominal absolute function.
ϕ	Auxiliary variables.
$\varphi(\,)$	Activation function of ELM.
$\omega, \hat{\omega}$	Output weights of ELM, approximated output weights.

I. Introduction

Because of the increasingly serious energy crisis and climate change, renewable energy has been widely explored for electricity power generation in recent years. Wind power is regarded as one of the most important renewable resources for power generation, of which the installed capacity worldwide increased about 30 times in the past decade and reached about 320 GW by the end of 2013 [1]. The higher variability and lower controllability of wind power production than conventional generations introduce considerable uncertainties to power systems and, consequently, raise severe challenges to power system balance and stability [2]. Therefore, accurate and reliable wind power forecasting becomes highly meaningful to reduce operation cost and improve reliability for power systems that have high wind penetration.

Wind power production is closely dependent on natural factors and suffers from the chaotic nature of weather systems. Therefore, traditional point forecasting cannot avoid forecasting errors [3]. Forecasting errors have been approximated via various parametric probability distribution models, including normal, log-normal, beta, mixed distribution models [4]-[6], Much attention is given to developing probabilistic forecasting methodologies to quantify the uncertainty involved in wind power forecasting, which can be modeled in the form of quantile, prediction intervals (PIs), or probability density forecasts. In [5], very short-term probabilistic forecasting of wind power was studied based on auto-regressive models and generalized logit-normal distributions. A novel bootstrap-based extreme learning machine (BELM) approach was developed for probabilistic forecasting of wind power [7], which has advantages of extremely high computational efficiency and online application potential. An adaptive conditional PI construction approach was proposed based on fuzzy inference and adaptive sampling technique [8]. Radial basis function neural network (RBFNN) was utilized for probabilistic forecasting of wind power via integrating point prediction and weather information [9], without a parametric probability distribution model of point forecasting error. A direct interval forecasting (DIF) approach was preliminarily proposed to directly generate nonparametric PIs of a specific high confidence [10]. Then a generalized hybrid intelligent approach was developed to produce optimal PIs based on the formulated performance-oriented cost function [11], without the needs of the information of wind power prediction errors. Based on the works in [10] and [11], a nonparametric density forecasting approach was proposed for solar energy prediction [12]. It suffers from high computational burden and severer local optimality, and incorrectly formulates the benchmark hybrid intelligent approach proposed in [11]. Quantile regression was employed to approximate quantiles of the volatility and uncertainty of wind power generation [13], [14]. A hybrid neural network approach was proposed to conduct multi-step probabilistic forecasting of wind power [15]. An ensemble-based probabilistic forecasting approach was developed for short-term wind power generation based on ensemble forecasts of meteorological variables [16]. A time-adaptive kernel density approach based on the Nadaraya-Watson estimator was proposed to estimate the uncertainty of wind power forecasting [17]. Besides, probabilistic forecasting was also applied for electricity price prediction in electricity markets [18],[19].

In general, probabilistic forecasting of wind power can provide meaningful information to participants in power systems if it is appropriately incorporated into decision-making models. Accordingly, decision makers can make proper operation and control schemes in advance to cope with risks introduced by the uncertainty of wind power production. This would effectively assist in transforming various decision-making activities in power systems from traditionally deterministic to probabilistic, and consequently increase the value of wind generation, such as wind power trading in electricity markets [20]-[22], wind turbine control [23], unit commitment [24]-[26], economic dispatch [27], optimal power flow [28], reserve allocation and setting [29], [30], demand response [31], security region assessment [32], frequency control [33], and so on.

Though neural networks (NNs) have excellent nonlinear mapping capability [34], they cannot precisely predict wind power generation. It is noted that classical NNs based PI construction methods are formulated on the basis of the serious assumption of normally distributed prediction errors [35]-[37], meanwhile succumbing to complicated mathematical formulation and high computational burden. Because of the high complexity of the statistical characteristics of wind power forecasting errors, it would be extremely difficult to derive parametric models to precisely describe the uncertainty of wind power prediction.

This paper proposes a novel direct quantile regression (DQR) approach for nonparametric probabilistic forecasting of wind power generation which takes advantage of extreme learning machine (ELM) and quantile regression. The ELM is a novel learning algorithm for training single-hidden layer feed-forward neural networks (NNs) [38], [39], demonstrating an extremely fast learning speed and overcomes the limitations of traditional NNs such as local minima, overtraining, and high computational burden. The quantile regression aims at estimating quantiles with different proportions and does not depend on any model assumption of probability distribution [40], [41], which has been applied, for example, in econometrics and social sciences. The proposed DQR method successfully generates nonparametric quantiles without the need of prior knowledge, statistical inference, or distribution assumption of point forecasts errors. Utilizing the unique merits of the ELM model, the complicated artificial NN-based nonparametric probabilistic forecasting is formulated as a linear programming (LP) problem in this paper. A variety of quantiles with incremental proportions could be accurately approximated, while the crossing of quantiles can be successfully avoided via the constraints of the formulated problem.

Realistic wind generation data from Bornholm Island in Denmark are used to verify the proposed approach. In practice, a few hours before real-time, wind power fluctuations of 10-min measurements are critical for power system operation and control [42], [43]. Different from classical multi-step wind power forecasting based on 1-hour wind power, multi-step 10-min resolution wind power forecasting is newly investigated in this paper. The effectiveness and advantages of the proposed DQR approach are validated and compared with state-of-the-art parametric and nonparametric approaches. With accurate quantification of the uncertainties of wind power prediction and high computational efficiency, the proposed novel nonparametric probabilistic forecasting model has a high potential to support various decision-making problems in power systems.

The remainder of this paper is organized as follows. The definition and evaluation indices are introduced in Section II. Section III presents the mathematical formulation of the proposed direct quantile regression approach. Comprehensive case studies are conducted to verify the proposed approach in Section IV. Finally, Section V gives the summary of the paper.

II. Definition and Evaluation of Nonparametric Probabilistic Forecasting

A. Nonparametric Probabilistic Forecasting

Denote the prediction target of wind power production at time t by y_t, and let x_t be the input vector of the prediction model, the data set can be expressed as

$$D=\{x_t, y_t\}_{t=1}^{T} \tag{1}$$

where x_t can consist of historical wind power, wind speed and wind direction, numerical weather prediction (NWP), etc., and T represents the size of the data set. Note that wind generation used in the study is normalized by nominal capacity P_n, which is accordingly bounded in the range [0, 1].

Let f_t and F_t denote the probability density function and the corresponding cumulative distribution function of y_t, respectively. Assuming that F_t is strictly increasing function, the quantile q_t^{α} with nominal proportion $\alpha\in$ [0, 1] of the random variable y_t can be obtained as the minimum value satisfying that,

$$\Pr(y_t \le q_t^{\alpha})=\alpha \tag{2}$$

which can be equivalently expressed as

$$q_t^{\alpha}=F_t^{-1}(\alpha) \tag{3}$$

The quantile forecast $\hat{q}_{t+k|t}^{\alpha}$ with nominal proportion α is regarded as an approximation of quantile $q_{t+k|t}^{\alpha}$ generated at time t for lead time $t + k$. It can easily be understood that a single quantile cannot meet the requirements of decision-making activities associated with wind power integration in power systems. Furthermore, because of the high complexity of wind power prediction uncertainty, it would be difficult to present the whole information of probability density distribution. Based on the definition of the quantile in (2) and (3), the probability density function $f_{t+k|t}$ and cumulative distribution function $F_{t+k|t}$ of the wind power production can be approximated by means of a set of quantiles produced by the nonparametric probabilistic forecasting, expressed as

$$\hat{F}_{t+k|t}=\left\{\hat{q}_{t+k|t}^{\alpha_i} \mid 0\le\alpha_1<\cdots<\alpha_i<\cdots<\alpha_r\le 1\right\} \tag{4}$$

It can be found that nonparametric probabilistic forecasting has the advantage that it does not rely on parametric probability distribution of prediction errors.

B. Classic Parametric Probabilistic Forecasting

Traditional parametric probabilistic forecasting is based on parametric statistical analysis of point forecasting errors [7]. Firstly, the point forecasting model needs to be constructed based on training dataset given in (1). It can be obtained that

$$y_{t+k}=\hat{y}_{t+k|t}+\varepsilon_{t,k} \tag{5}$$

where y_{t+k} is prediction target at time $t+k$, $\hat{y}_{t+k|t}$ denotes the prediction value at time t, $\varepsilon_{t,k}$ denotes the prediction error.

Then prediction error set can be collected under the training dataset. And the estimate PDF $\hat{f}_{t,k}^{\varepsilon}$ and CDF $\hat{F}_{t,k}^{\varepsilon}$ of prediction errors can be determined under certain probability distribution assumption,

$$\hat{f}_{t,k}^{\varepsilon}\to\{\varepsilon_{t,k}\}_{t=1}^{T} \tag{6}$$

Therefore the predictive distribution of wind generation can be estimated through

$$\hat{q}_{t+k|t}^{\alpha}=\hat{y}_{t+k|t}+\hat{F}_{t,k}^{\varepsilon}(\alpha) \tag{7}$$

Due to the high complexity of wind power forecasting error, the parametric models can be hard to precisely describe the predictive distribution of wind generation.

C. Evaluation Criteria

The quantiles obtained by nonparametric probabilistic forecasts need to be comprehensively assessed from the perspective of both reliability and sharpness. Improper evaluation index applied would result in inaccurate PIs and systematic biases to decision makers [44].

1) Reliability

Reliability is regarded as a major property for validating the correctness of nonparametric probabilistic forecasting methodologies [8]. Low reliability could introduce systematic prejudice to the subsequent decision-making problems. Based on the definition of quantile, the empirical proportion of quantile $\hat{q}_t^{\alpha}$ is defined as,

$$\hat{P}_{\mathrm{Q}}^{\alpha}=\frac{1}{T}\sum_{t=1}^{T}\eta_t \tag{8}$$

where T is the total number of observations, and η_t denotes the indicator of quantile $\hat{P}_{\mathrm{Q}}^{\alpha}$ defined by,

$$\eta_t=\begin{cases}1, & y_t\le\hat{q}_t^{\alpha}\\ 0, & y_t>\hat{q}_t^{\alpha}\end{cases} \tag{9}$$

The reliability diagram will provide the deviation between empirical proportion $\hat{P}_{\mathrm{Q}}^{\alpha}$ and nominal proportion α of predicted quantiles. The average proportion deviation (APD), denoted as PD_{Q}^{α}, of a quantile is defined as,

$$PD_{\mathrm{Q}}^{\alpha} = \hat{P}_{\mathrm{Q}}^{\alpha} - \alpha \qquad (10)$$

Generally, the closer the APD is to zero, the higher the reliability of the estimated quantile $\hat{q}_t^{\alpha}$.

2) Sharpness

In addition to reliability, the sharpness of the nonparametric probabilistic forecasts needs to be considered. It would be critical to utilize proper scoring rules to evaluate the overall skill of nonparametric probabilistic forecasts to include all aspects of the forecasts quality.

When the nonparametric probabilistic forecasts take the form of predictive quantiles, the scoring rule integrating a set of quantiles on a single time point is defined by [8],

$$S_{\mathrm{Q},t} = \sum_{i=1}^{r} (\mathbf{1}\{y_t \le \hat{q}_t^{\alpha_i}\} - \alpha_i)(y_t - \hat{q}_t^{\alpha_i}) \qquad (11)$$

Given a test set with T prediction series, the average score based on the above scoring rule can be calculated to evaluate the skill of predictive quantiles, expressed as,

$$S_{\mathrm{Q}} = \frac{1}{T}\sum_{t=1}^{T} S_{\mathrm{Q},t} \qquad (12)$$

The scoring rule is positively oriented, which implies that the larger score value S_{Q} indicates the better predictive quantiles.

III. Formulation of Direct Quantile Regression

A. *Extreme Learning Machine*

ELM is an efficient learning technique for training a single hidden-layer feed-forward NN [38]. It has unique advantages of simple formulation and extremely fast training speed. The typical structure of ELM is shown in Fig. 1. For a set of training dataset $\{(\mathbf{x}_i, \mathbf{y}_i)\}, i = 1, 2, \cdots, N$ with the input $\mathbf{x}_i \in \mathbf{R}^n$ and the output $\mathbf{y}_i \in \mathbf{R}^m$, suppose that there exists ELM with L hidden neurons and infinitely differentiable activation function $\varphi(\cdot)$ can approximate these N samples with zero error. Then the ELM can be represented as,

$$g(\mathbf{x}_j; \mathbf{a}_i, b_i, \boldsymbol{\omega}_i) = \sum_{i=1}^{L} \boldsymbol{\omega}_i \varphi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i) = \mathbf{y}_j, \;\; j = 1, \ldots, N. \qquad (13)$$

where $\mathbf{a}_i \in \mathbf{R}^n$ is the input weight vector connecting the ith hidden neuron and the input neurons, $\boldsymbol{\omega}_i \in \mathbf{R}^m$ represents the output weight vector connecting the ith hidden neuron and the output neurons, b_i is the bias of the ith hidden neuron, and $\varphi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i)$ denotes the output of the ith hidden neuron. The above equations can be rewritten in matrix form,

$$\mathbf{H}\boldsymbol{\omega} = \mathbf{Y} \qquad (14)$$

where $\mathbf{H}$ denotes the hidden layer output matrix, defined as,

$$\mathbf{H} = \begin{bmatrix} \varphi(\mathbf{a}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & \varphi(\mathbf{a}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \cdots & \vdots \\ \varphi(\mathbf{a}_1 \cdot \mathbf{x}_N + b_1) & \cdots & \varphi(\mathbf{a}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N\times L}, \qquad (15)$$

$\boldsymbol{\omega}$ and $\mathbf{Y}$ denote the matrices of output weights and targets, respectively,

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_1^T \\ \vdots \\ \omega_L^T \end{bmatrix}_{L\times m} \text{ and } \mathbf{Y} = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N\times m} \qquad (16)$$

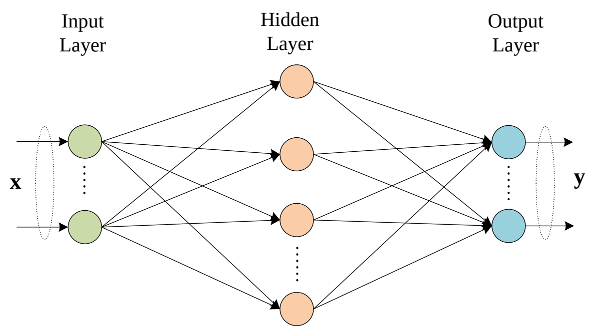


Fig. 1. Typical structure of ELM.

The hidden neuron parameters $\mathbf{a}_i$ and b_i are randomly generated and remain unchanged, and then the hidden layer output matrix $\mathbf{H}$ becomes a constant. Accordingly, training a single hidden-layer feed-forward NN is just equivalent to reaching a norm least-squares solution $\hat{\boldsymbol{\omega}}$ of the linear system (14) such that

$$\|\mathbf{H}\hat{\boldsymbol{\omega}} - \mathbf{Y}\| = \min_{\boldsymbol{\omega}} \|\mathbf{H}\boldsymbol{\omega} - \mathbf{Y}\| \qquad (17)$$

which is equivalent to the minimization of the cost function of the classical gradient-based learning algorithms defined as,

$$C_N = \sum_{j=1}^{N} \left[\sum_{i=1}^{L} \omega_i \varphi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i) - y_j \right]^2 \qquad (18)$$

The optimal solution $\hat{\boldsymbol{\omega}}$ in (17) can be derived through simple matrix operation and expressed as,

$$\hat{\boldsymbol{\omega}} = \mathbf{H}^{\dagger}\mathbf{Y} \qquad (19)$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of the matrix $\mathbf{H}$, and can be calculated by different methods, including the singular value decomposition method, orthogonal projection method, and orthogonalization method.

The traditional gradient-based leaning techniques for NNs require a number of iterations to tune the parameters. However, based on a formulated linear system, ELM can obtain the parameters through just a single matrix computation, with extremely fast speed. It can overcome several limitations involved in traditional gradient-based learning techniques, including local minima, overtraining, and high computational costs.

B. *Quantile Regression*

Quantile regression introduced by [40] aims to approximate the conditional distribution of the random variable by means of quantiles. The conditional quantile functions are modeled as functions of explanatory variables. More specifically, the quantile of proportion α defined in (2) can be approximated through minimizing a sum of asymmetrically weighted absolute residuals cost function [41], expressed as,

$$\min_{q_t^\alpha \in R} \sum_{t=1}^{T} \rho_\alpha \left(y_t - q_t^\alpha \right) \tag{20}$$

where T is the size of datasets for model construction, $\rho_\alpha(\cdot)$ is the nominal absolute function defined by,

$$\rho_\alpha(x) = \begin{cases} \alpha x & \text{if } x \geq 0 \\ (\alpha - 1)x & \text{if } x < 0 \end{cases} \tag{21}$$

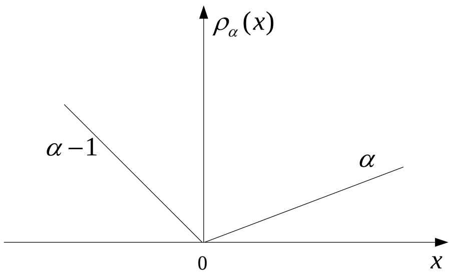


Fig. 2. Cost function of quantile regression.

The cost function of quantile regression is illustrated in Fig. 2, to more explicitly explain the definition. It should be noted that unconditional quantiles are yielded based on the optimization problem in (20). Conditional quantiles can be obtained by replacing the quantitative q_t^α by parametric function $q_t^\alpha(x_t, \gamma)$, expressed as,

$$\min_{\gamma \in R} \sum_{t=1}^{T} \rho_\alpha \left(y_t - q_t^\alpha(x_t, \gamma) \right) \tag{22}$$

C. Direct Quantile Regression

The proposed DQR approach is to fulfill nonparametric probabilistic forecasting of wind power production via producing nonparametric conditional quantiles. In addition to the excellent nonlinear regression ability, the ELM becomes a linear system after randomly determining the input weights, which further motivates the application of ELM in the study.

1) ELM-based Single Conditional Quantile: According to the quantile regression expressed in (22), ELM is applied as the function of the conditional quantile of wind power, which can be expressed as the following cost function,

$$\min_{w_\alpha} \sum_{t=1}^{T} \rho_\alpha \left(y_t - g(x_t, w_\alpha) \right) \tag{23}$$

subject to

$$0 \leq g(x_t, w_\alpha) \leq 1 \tag{24}$$

where $g(x_t, w_\alpha)$ denotes the linear function of ELM, w_α is the decision variable vector of the optimization problem, that is, the output weight of ELM, the constraint (24) aims to guarantee the ELM outputs within the wind generation capacity range [0, 1]. Logic values exist in (23) which cannot be solved using traditional algorithms. It can be transformed to the equivalent optimization problem by introducing auxiliary variables ϕ_t^α, expressed as,

$$\min_{w_\alpha, \phi_t^\alpha} \sum_{t=1}^{T} \phi_t^\alpha \tag{25}$$

subject to

$$\phi_t^\alpha \geq \alpha [y_t - g(x_t, w_\alpha)], \ \forall t \tag{26}$$

$$\phi_t^\alpha \geq (\alpha - 1)[y_t - g(x_t, w_\alpha)], \ \forall t \tag{27}$$

$$(24) \tag{28}$$

It can be found that the optimization problem in (25)-(28) is a LP problem, which can be efficiently solved. In practice, the single quantile cannot provide sufficiently useful information on wind power uncertainty to decision makers in power systems.

2) ELM-based Multiple Conditional Quantiles: To accurately describe the predictive density of wind power, it would be necessary to estimate the whole probability as far as possible with a series of conditional quantiles of different proportions $0 \leq \alpha_1 < \cdots < \alpha_i < \cdots < \alpha_r \leq 1$. Based on the formulation of a single conditional quantile, the cost function of ELM-based multiple condition quantiles can be obtained by summing the cost functions of quantiles of different nominal proportions. The non-crossing constraints can be introduced to obtain multiple quantiles. Then the optimization problem can be formulated as,

$$\min_{w_{\alpha_i}, \phi_t^{\alpha_i}} \sum_{i=1}^{r} \sum_{t=1}^{T} \phi_t^{\alpha_i} \tag{29}$$

subject to

$$\phi_t^{\alpha_i} \geq \alpha_i [y_t - g(x_t, w_{\alpha_i})], \ \forall i, t \tag{30}$$

$$\phi_t^{\alpha_i} \geq (\alpha_i - 1)[y_t - g(x_t, w_{\alpha_i})], \ \forall i, t \tag{31}$$

$$g(x_t, w_{\alpha_i}) \leq g(x_t, w_{\alpha_{i+1}}), \ 1 \leq i \leq r-1, \forall t \tag{32}$$

$$0 \leq g(x_t, w_{\alpha_i}) \leq 1, \ 1 \leq i \leq r, \forall t \tag{33}$$

where the non-crossing between quantiles of different proportions can be guaranteed via the constraint in (32), and the constraint in (33) ensures that the output quantiles lie in the nominal range of wind farm output. Furthermore, it is formulated as a linear optimization problem and can be efficiently solved by LP algorithms. The global optimum can be guaranteed [45].

In general, the proposed DQR method incorporating ELM and quantile regression successfully formulates the nonparametric probabilistic forecasting as a LP model. The correctness of the formulated model can be ensured as the cost function is based on the mature quantile regression technology. Though a feed-forward NN-based forecaster has very complicated parameters, it can be easily solved by LP algorithms, due to the application of ELM. Therefore, the proposed DQR model can demonstrate excellent accuracy and computational efficiency.

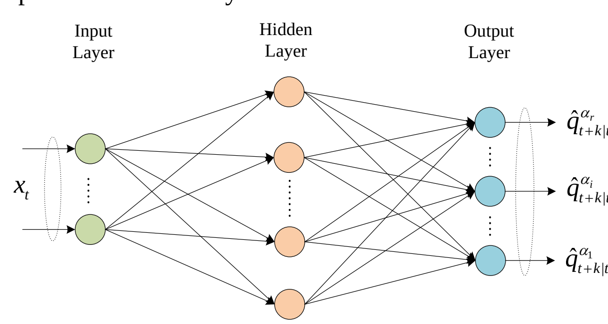


Fig. 3. Framework of the proposed DQR model for nonparametric probabilistic forecasting of wind power.

The framework of the proposed DQR model is visually demonstrated in Fig. 3. The proposed DQR approach establishes an ELM-based forecaster to directly produce the conditional quantiles of wind power production with different proportions, without the assumption of probability distribution of point forecasting errors needed in traditional approaches. A set of produced quantiles can give a full estimation of the probability density of wind power prediction uncertainty. The simple model for nonparametric forecasting based on ELM would dramatically improve computational efficiency and have a large space for practical applications. Moreover, the proposed DQR approach has significantly high flexibility because of the excellent regression capability of ELM.

IV. Case Studies

A. Description of Experiment Data

The realistic wind power data from Denmark are used for the verification of the proposed DQR approach. The wind farm is located in Bornholm Island that is utilized as an experiment base for smart grid technologies in Denmark. The total installed capacity of wind power P_n is about 30 MW. The backbone of electricity network in Bornholm Island is illustrated in Fig. 4.

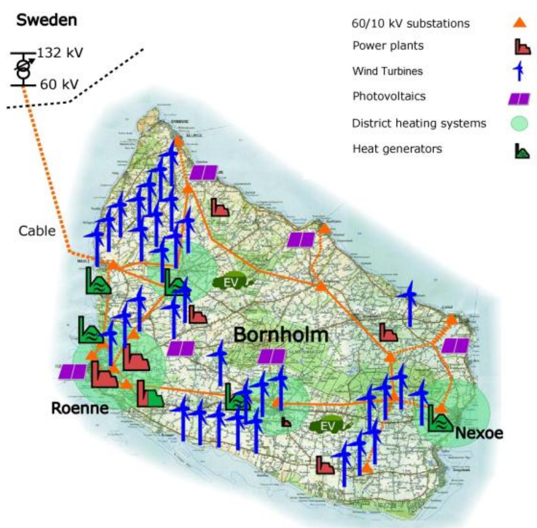


Fig. 4. The wind farm of Bornholm Island (courtesy of Centre for Electric Power and Energy, Technical University of Denmark).

Traditionally, wind power data of hourly resolution are employed for wind power production forecasting to satisfy the requirements of practical power system operation [8]. With increasing high penetration of wind power, it becomes crucial to forecast high-resolution wind generation to provide meaningful information for operation and control of smart grids. It has been revealed that intra-hour resolution (e.g., 10-min) wind power demonstrates much higher volatility than hourly wind power [46]. In particular, the fluctuations of wind power at this time scale have severe impacts on the balance of power systems via electricity market mechanism, defined by the transmission system operator in Denmark [42]. Further, high resolution wind power prediction becomes more and more important due to the rapid development of smart grid. Therefore, 10-min resolution wind power data from the Bornholm Island wind farm are utilized in the study covering the periods from June to July 2012, and from November to December 2012, to involve the seasonality of wind power generation to a certain extent.

B. Description of Case Studies

To verify the forecast performance of the proposed DQR approach, four other probabilistic forecasting methods, including the parametric approaches persistence, BELM based on normal distribution (BELM-Normal) and BELM based on beta distribution (BELM-Beta) [7], and the nonparametric approach RBFNN [9], are employed to estimate PIs using the same training and testing data for benchmarking.

As a naïve model, the persistence approach is commonly used as a basic benchmark for point forecasts of wind power. The persistence-based probabilistic forecast model is used as reference model in the study, of which the predictive density follows the normal distribution. The mean value is given by the average of the latest available measurements of wind power, and the variance is correspondingly determined by the latest observations. In addition, the BELM-Normal approach that obtains bootstrapped ELMs is utilized as an advanced parametric model relying on normal predictive density, which focuses on estimating prediction intervals of high confidence and has extremely high computation efficiency [7]. The BELM-Beta approach based on BELM and beta distribution assumption of prediction uncertainty is applied in the case study to investigate the influences of different distribution models. To further demonstrate the effectiveness of the proposed approach, the RBFNN-based nonparametric probabilistic forecasting approach is used as a mature benchmark that does not need the assumption of probability distribution for forecasting errors. In this approach, the RBFNNs are optimized through the cost function formulated by integrating Bayesian interpolation, the Stein unbiased risk estimator, and the bias of predictive density.

In practice, different wind power prediction horizons, ranging from minutes to days, respond to associated requirements of actual decision-making activities in power systems. Previous studies focus on forecasting wind power with an hourly resolution. In Denmark, the 10-min lead time is considered the most important by the transmission system operator Energinet.dk due to the wind power fluctuations at this scale having the most serious effects on power system balance [42]. For a look-ahead time shorter than 3 to 6 hours, a statistical approach using wind power historical measurements would perform better than NWP-based approaches [3]. Therefore, we investigate multi-step forecasting of wind power from 10 min to 3 hours, which is very important for grid operations (e.g., reserve dispatch and system control in practice), but rarely examined in the state-of-the-art research. For such short look-ahead times in the study, it takes only historical wind power data as inputs of the proposed model. More exogenous variables, such as NWP information, can be easily considered in the proposed model due to the application of ELM with high regression capability.

To comprehensively estimate the predictive density of wind power, quantiles of proportions ranging from 5% to 95% with 5% increments are produced, except for the proportion 50% according to (29)-(33). The symmetric quantiles form prediction intervals with nominal coverage rates 10%, 20%,…, 90%. The proposed method and the four benchmarks applied are tested on the Bornholm Island wind farm for systematical analysis and comparisons. In the study, 60% of wind power production data are used for constructing the forecasting models, and the remaining data are used for model test. Due to different season patterns, the prediction models are separately constructed on the two datasets.

C. *Analysis of Experiment Results*

The average proportion deviations between nominal and empirical proportions in July and December are depicted in Fig. 5 (a) and (b), respectively, for the evaluation of reliability. The reliability diagrams are obtained via the average of probabilistic forecasts for the various look-ahead times, ranging from 10 min to 3 hours.

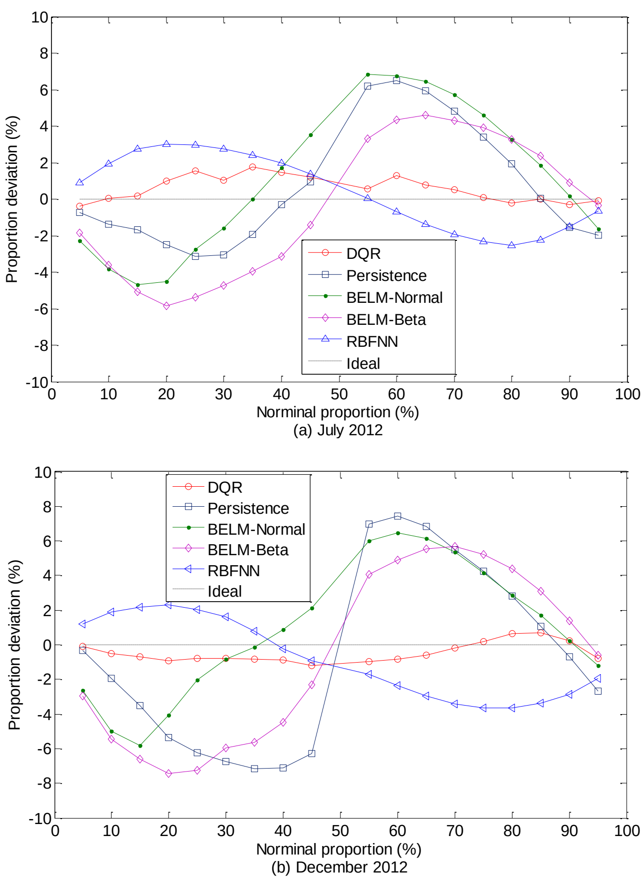


Fig. 5. Reliability diagrams of probabilistic forecasts obtained by proposed DQR approach and the applied four benchmarks.

Fig. 5 shows that the proposed DQR method can provide excellent reliability performance and significantly outperforms the applied four benchmarks in both test periods July and December. For all confidence levels from 5% to 95%, the empirical proportions of quantiles obtained by the proposed method are close to the corresponding nominal proportions. The absolute deviations obtained from the proposed method at different nominal confidence levels are smaller than 2% and 1.5% in July and December respectively, indicating high reliability. In general, the quantiles are slightly overestimated and underestimated by the DQR approach in July and December respectively, which implies that reliability can be improved by adjusting the obtained quantiles. It indicates the difference performance in different season patterns. From Fig. 5 (a), the parametric approaches including persistence, BELM-Beta and BELM-Normal can generate about 6% proportion deviation in July 2012. It can be found from Fig. 5 (b) that the persistence and BELM-Beta can produce proportion deviation of about 7.5% at the nominal proportions 60% and 20%, respectively, in December 2012. The BELM-Normal approach can generate proportion deviation of 6.5% at the nominal proportion 60%. Even though the BELM-Normal approach can generate satisfactory PIs for high confidence; generally, these parametric approaches are inclined to be underestimated for proportions below 50%, and tend to reach overestimation for the higher proportions. In comparison with parametric approaches, RBFNN can generate relatively reliable probabilistic forecasts in the two studied cases, as the probability density bias is used as an objective of the cost function. However, it suffers from high computation burden during the optimization process.

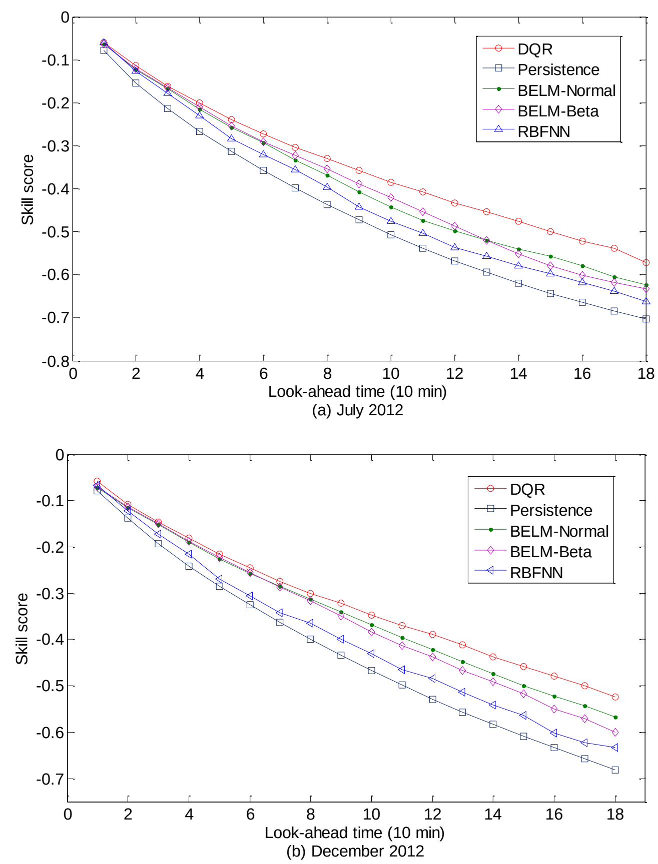


Fig. 6. Skill score of sharpness evaluation of probabilistic forecasts obtained from the proposed DQR approach and the applied four benchmarks.

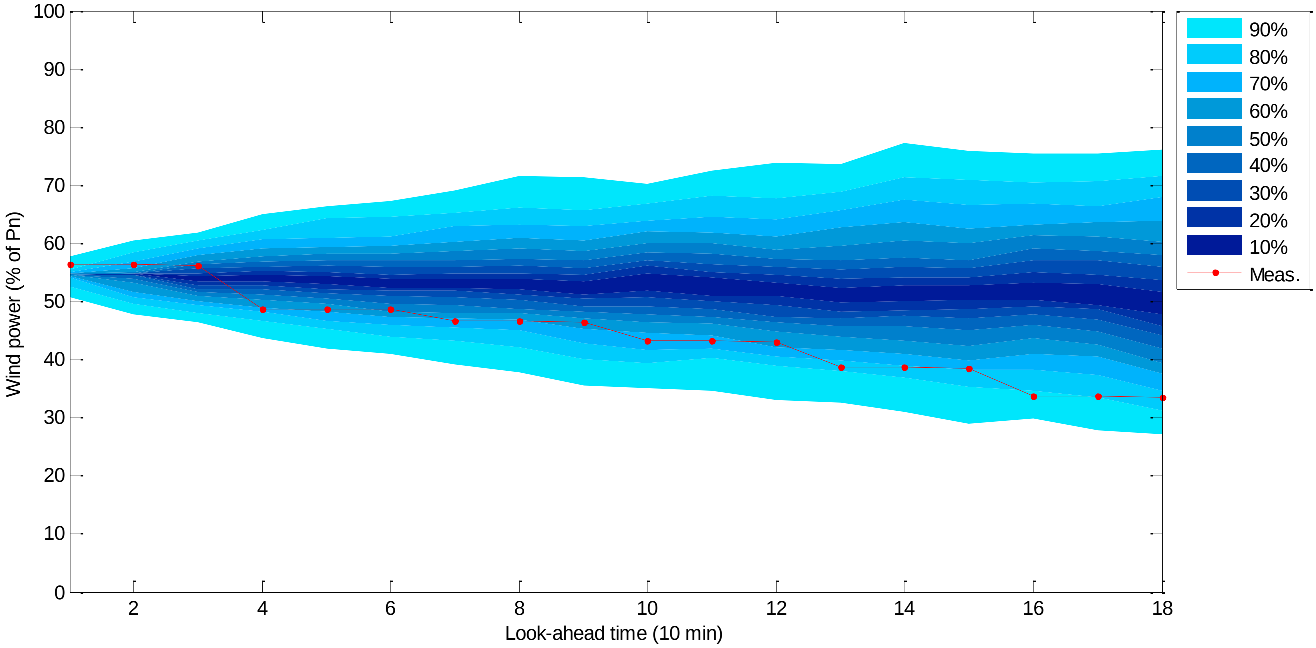


Fig. 7. Multi-step probabilistic forecasting results issued on December 11, 2012, at 18:10, obtained from the proposed DQR approach.

The resultant skill score values of probabilistic forecasts of the proposed method and the four benchmarking methods in July and December 2012 are displayed in Fig. 6 (a) and (b) respectively for the evaluation of sharpness. The *x*-axis gives the look-ahead times of probabilistic forecasts, and the various curves show the sharpness indices. The quantiles obtained by the proposed DQR method have the largest skill score values, indicating the best overall skill and sharpness in the two cases, compared with the employed benchmarks. In particular, the DQR method performs about 25% better than the persistence approach and about 20% better than the RBFNN approach from the perspective of sharpness. Persistence has the lowest sharpness, as it cannot well capture the nonlinearity of wind power series. BELM-Normal and BELM-Beta have similar overall skill in the two cases. RBFNN cannot effectively ensure the sharpness of probabilistic forecasts, since the cost function only includes the reliability index. Furthermore, we can see that the skill score decreases with the look-ahead time. This is because it is more difficult to predict wind power for a longer lead time, which then introduces larger uncertainty.

Predictive distribution of wind generation is relevant to approximation of the error distributions of the point forecasts, and extremely difficult to be precisely estimated by parametric models. The proposed nonparametric probabilistic forecast approach has significant merits compared with parametric approaches with the probability distribution assumption. Considering the comprehensive performance, including reliability and overall skill, the proposed DQR approach produces the best predictive quantiles compared with the applied four benchmarks.

Multi-step prediction intervals at a specific time point on December 11, 2012, at 18:10 are displayed in Fig. 7. In the 3 hours, the wind power generation decreased from 56.4% to 33.5%. Though with a similar ramp rate, the width of prediction intervals apparently increases with the length of look-ahead time. Particularly, the prediction intervals are very tight in the first few steps. This is consistent with the sharpness results in Fig. 6 that the longer prediction horizon will bring the larger uncertainty. Moreover, PIs of lower confidence can be well encapsulated by that of higher confidence. In practice, the multi-step probabilistic forecasting of high-resolution wind power can provide meaningful information for decision-making activities in power systems, including economic dispatch, optimal power flow, demand side, wind farm control, and electricity market clearing. Based on the multi-step forecasting, decision makers can continuously adjust the decision to optimally respond to the fluctuations of wind generation.

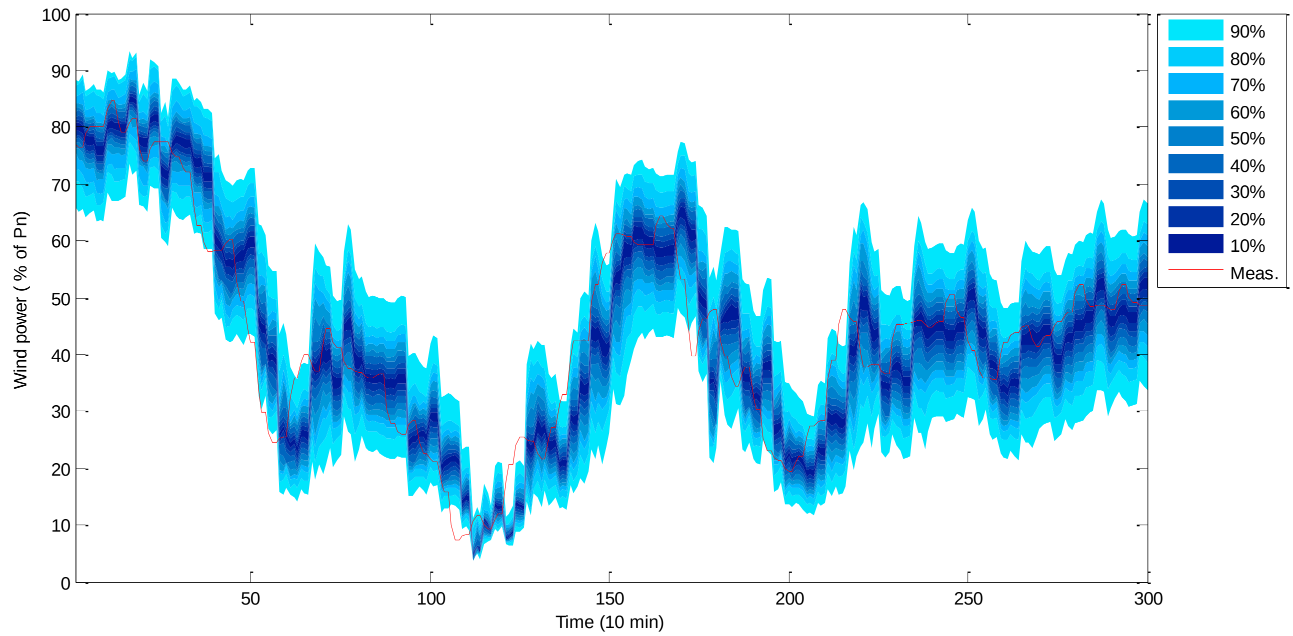


Fig. 8. PIs with look-ahead time of 1 hour of Bornholm Island wind farm in July 2012, obtained from the proposed DQR approach.

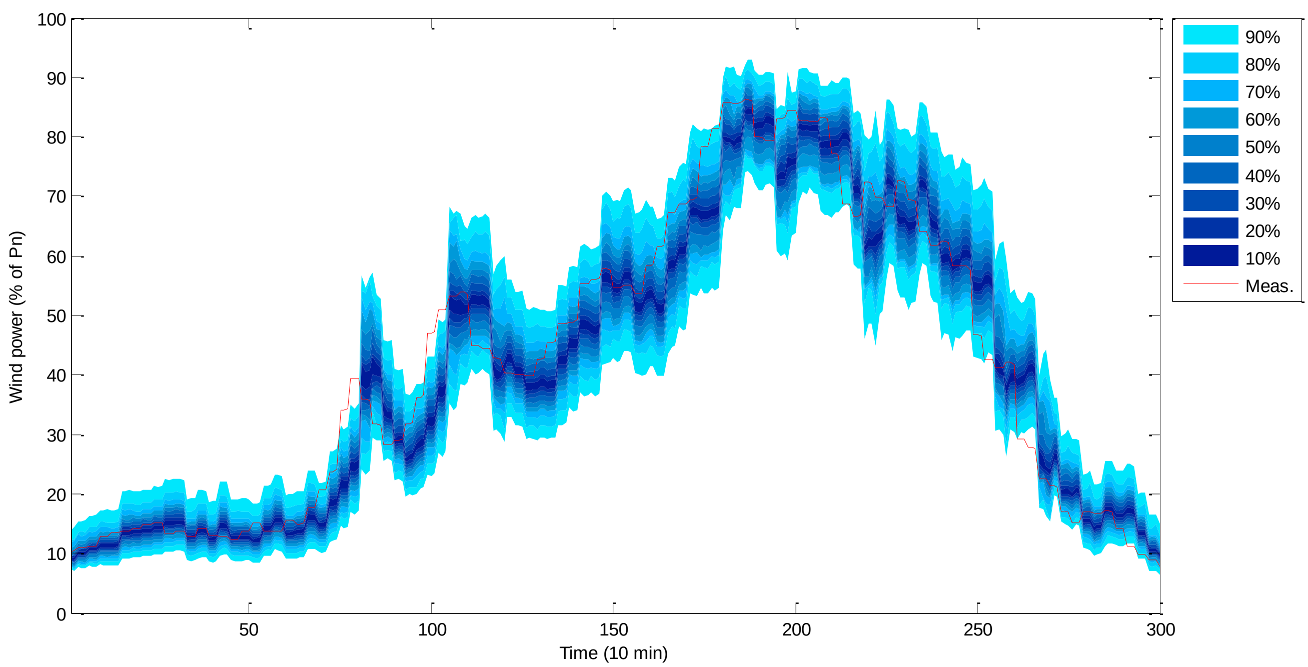


Fig. 9. PIs with look-ahead time of 1 hour of Bornholm Island wind farm in December 2012, obtained from the proposed DQR approach.

Forecasted quantiles of wind power generation in July and December 2012 with 1-hour look-ahead time are displayed in Figs. 8 and 9, respectively, along with the actual measured wind generation. The displayed wind power covers the period of about 2 days. It can be seen that the actual wind power can be well enclosed by the estimated quantiles, demonstrating the high performance of the proposed approach. This can solidly prove the excellent performance of the proposed DQR approach in different season patterns. As the application of ELM with high learning ability, the proposed approach has high adaptability to meet the requirements of wind power prediction in different situations. It is apparent that the PIs of lower confidence are embedded very well by those of higher confidence, which solidly proves that the constraints in (32) formulated for the proposed DQR approach can effectively avoid the crossing of quantiles of different proportions.

From the examples of Figs. 8 and 9, we can clearly find that the PIs are not symmetric on the center point that relates to point forecasts. The asymmetry of predictive distribution is more significant for low and high values of wind power. This also further explains why normal and beta distributions cannot precisely approximate the predictive distribution of wind power. Moreover, the sharpness of PIs varies very much in different conditions, according to the examples given in Fig. 9. The PIs for low wind power value are relatively tight, which can be explained by the wind power fluctuation being fairly smaller at this wind generation level. PIs become much wider, especially when wind power is at the intermediate level. It should not be unreasonable that wind power at the intermediate level has the high potential to be up and down, leading to relatively larger uncertainty, which can be easily found in Fig. 8. Though with excellent performance, there exist some conditions in Fig. 9 in which the actual wind power is beyond the estimated PIs for the pretty large down/up ramp. As the nominal confidence is 90%, theoretically, there should be nearly 10% of wind generation not covered by the PIs. Nevertheless, the estimated PIs can shorten the estimation error and reduce the risk involved in decision making as far as possible.

TABLE I
COMPUTATION TIME OF MODEL CONSTRUCTION

Method	Time (s)
DQR	63.89
BELM	4.43
RBFNN	49780.04

The computational efficiency of the proposed DQR approach is investigated via comparisons with BELM and RBFNN, as listed in Table I. The three methods are executed on a PC with Intel Core Duo i7-4790 CPU @ 3.6GHz CPU and 16GB RAM. The average training time of these models are obtained based on wind power data from November to December 2012 and 1-hour look-ahead time.

From Table I, the BELM has extremely fast speed for model construction with only about 4s training time. As a parametric approach, it can have satisfactory performance for estimation of PIs with high confidences, and provide a very useful tool for online application. It can be seen that the proposed DQR approach implements about 780 times faster than the RBFNN method, indicating a significantly high computational efficiency. The training time of RBFNN can be too long for the prediction horizon shorter than 3 hours. In general, traditional NN based nonparametric probabilistic forecasting techniques reply on complicated cost function, and has to be solved by heuristic algorithm. As the 18 quantiles are optimized simultaneously, the dimension of decision variables can be very high. Consequently, it always suffers from the disadvantage of high computational burden. The proposed approach needs only 64s training time, which can satisfy the online model update to some extent. With the fast training speed and excellent performance, the proposed DQR method has high potential for practical applications in power systems.

D. Discussion

The experimental results illustrate that the proposed method has superior performance for short-term probabilistic wind power forecasting compared with the four well-established parametric and nonparametric benchmarks. The existing probabilistic forecasting approaches are usually based on the results of point forecasts. Generally, the parametric approaches rely on the distribution assumption of the forecast error; for example, the persistence, BELM-Normal and BELM-Beta approaches in the case study. Traditional artificial NN-based nonparametric models are built on the complicated cost function, which has high computational complexity. Integrating the advantages of ELM and quantile regression, the proposed DQR establishes a simple LP model to efficiently execute artificial NN-based nonparametric probabilistic forecasting of wind power generation. This enables the online construction of the proposed prediction model. Meanwhile, it does not need any information on point forecasts and the associated errors.

The DQR approach presents a generalized nonparametric probabilistic forecasting framework owing to the significant merits adaptive to various inputs, outputs, and look-ahead time, under the assistance of the high mapping capability of ELM. Multi-step probabilistic forecasting of high-resolution wind power is focused in the case studies, and the prediction horizon is shorter than 3 hours. Wind generation is closely related to meteorological variables, such as wind speed and wind direction. For a look-ahead time longer than a few hours, numerical weather prediction information needs to be utilized as the inputs of the proposed prediction model to ensure performance.

In practice, wind power forecasting information is a crucial base to various decision-making activities in power systems. Meanwhile, different applications may need forecasting values of different look-ahead times. For instance, it has been studied that high-resolution (e.g., 10 min) wind power has serious influences on power system balance. Wind power forecasting at this resolution would be meaningful for real-time power system operation, wind farm control, and electricity market running. Because of its high reliability and flexibility, the proposed DQR approach can have high potential in other practical applications, including unit commitment, reserve determination, and wind power trading.

V. CONCLUSION

Because of the high penetration and intermittency of wind power, accurate wind power prediction is crucial to the economy and security of modern power systems. With the unavoidability of wind power forecasting errors, probabilistic forecasting of wind power becomes increasingly popular and demonstrates its superiority over traditional point forecasts. A novel direct quantile regression approach has been presented to produce nonparametric probabilistic forecasting of wind power combining ELM and quantile regression. The complicated artificial NN nonparametric probabilistic forecasting is innovatively formulated as a simple LP problem, with high computation efficiency. Moreover, the proposed approach does not rely on the information and distribution assumption of point forecasts error. The predictive quantiles of wind power with multiple proportions can be generated simultaneously via one single linear optimization process, meanwhile ensuring the nesting of PIs among different confidences.

The effectiveness of the proposed DQR method is validated based on realistic wind power data from Bornholm Island in Denmark and comprehensive comparisons with several state-of-the-art benchmarks. The first multi-step probabilistic forecasting of 10-min resolution wind power has been presented in this paper. In general, the proposed DQR approach provides an efficient and flexible framework for nonparametric probabilistic forecasting of wind power generation with highly reliable performance. Under the background of large-scale penetration of wind power, the probabilistic forecasts of wind power can be widely integrated into various decision-making activities in power systems, for example, reserve determination, economic dispatch, and wind power trading, etc.

REFERENCES

[1] M. Bilgili, A. Ozbek, B. Sahin, and A. Kahraman, "An overview of renewable electric power capacity and progress in new technologies in the world," *Renew. Sust. Energ. Rev.*, vol. 49, pp. 323-334, Sep. 2015.

[2] Y. V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on California power systems," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 1039-1050, May 2009.

[3] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power: A literature review," Tech. Rep., ANEMOS.plus Project Deliverable Report D1.2, 2011.

[4] H. Bludszuweit, J. A. Dominguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 983-991, Aug. 2008.

[5] P. Pinson, "Very-short-term probabilistic forecasting of wind power with generalized logit–normal distributions," *J. R. Stat. Soc., Ser. C Appl. Stat.*, vol. 61, no. 4, pp. 555-576, Aug. 2012.

[6] 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 Trans. Power Syst.*, vol. 26, no. 4, pp. 2031-2039, Nov. 2011.

[7] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033 - 1044, May 2014.

[8] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1845-1856, Nov. 2010.

[9] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1788-1796, Nov. 2012.

[10] C. Wan, Z. Xu, and P. Pinson, "Direct interval forecasting of wind power," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4877-4878, Nov. 2013.

[11] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1166 - 1174, May 2014.

[12] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation — With application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850-3863, Sep. 2016.

[13] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47-54, Jan./Mar. 2004.

[14] H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, no. 1-2, pp. 95-108, Jan./Apr. 2006.

[15] C. Wan, Y. Song, Z. Xu, G. Yang, and A. H. Nielsen, "Probabilistic wind power forecasting with hybrid artificial neural networks," *Electr. Power Compon. Syst.*, vol. 44, no. 15, pp. 1656-1668, 2016.

[16] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *IEEE Trans. Energy Convers.*, vol. 24, no. 3, pp. 775-782, Sep. 2009.

[17] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660-669, Oct. 2012.

[18] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A hybrid approach for probabilistic forecasting of electricity price," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463-470, Jan. 2014.

[19] C. Wan, M. Niu, Y. Song, and Z. Xu, "Pareto optimal prediction intervals of electricity price," *IEEE Trans. Power Syst.*, vol. PP, no. 99, pp. 1-3, 2016. doi: 10.1109/TPWRS.2016.2550867.

[20] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1148-1156, Aug. 2007.

[21] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Short-term trading for a wind power producer," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 554-564, Feb. 2010.

[22] Y. Cai, J. Lin, C. Wan, and Y. Song, "A stochastic bi-level trading model for an active distribution company with distributed generation and interruptible loads," *IET Renew. Power Gener.*, 2016, to be published. doi:10.1049/iet-rpg.2016.0364.

[23] L. Yang, Z. Xu, J. Østergaard, Z. Y. Dong, and K. P. Wong, "Advanced control strategy of DFIG wind turbines for power system fault ride through," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 713-722, May 2012.

[24] C. Zhao, J. Wang, J. P. Watson, and Y. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2708-2717, Aug. 2013.

[25] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 206-215, Feb. 2012.

[26] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1319-1327, Aug. 2008.

[27] Á. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1702-1713, Jul. 2015.

[28] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4742-4751, Nov. 2013.

[29] M. A. Matos and R. J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 594-603, May 2011.

[30] A. Ahmadi-Khatir, A. J. Conejo, and R. Cherkaoui, "Multi-area unit scheduling and reserve allocation under wind power uncertainty," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1701-1710, Jul. 2014.

[31] C. Sahin, M. Shahidehpour, and I. Erkmen, "Allocation of hourly reserve versus demand response for security-constrained scheduling of stochastic wind energy," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 219-228, Jan. 2013.

[32] C. Wan, J. Lin, W. Guo, and Y. Song, "Maximum uncertainty boundary of volatile distributed generation in active distribution network," *IEEE Trans. Smart Grid*, vol. PP, no. 99, pp. 1-1, 2016. doi: 10.1109/TSG.2016.2623760.

[33] H. Jiang, J. Lin, Y. Song, and D. J. Hill, "MPC-based frequency control with demand-side participation: A case study in an isolated wind-aluminum power system," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3327-3337, Nov. 2015.

[34] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359-366, Jan. 1989.

[35] J. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *J. Am. Stat. Assoc.*, vol. 92, no. 438, pp. 748-757, Jun. 1997.

[36] G. Papadopoulos, P. J. Edwards, and A. F. Murray, "Confidence estimation methods for neural networks: A practical comparison," *IEEE Trans. Neural Netw.*, vol. 12, no. 6, pp. 1278-1287, Nov. 2001.

[37] D. J. C. MacKay, "A practical Bayesian framework for backpropagation networks," *Neural Comput.*, vol. 4, no. 3, pp. 448-472, May 1992.

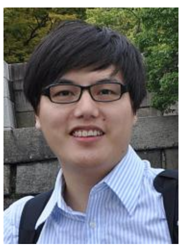
[38] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Trans. Neural Netw.*, vol. 17, no. 6, pp. 1411-1423, Nov. 2006.

[39] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513-529, Apr. 2012.

[40] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33-50, Jan. 1978.

[41] R. Koenker, *Quantile Regression*: Cambridge University Press, 2005.

[42] V. Akhmatov, "Influence of wind direction on intense power fluctuations in large offshore windfarms in the North Sea," *Wind Eng.,* vol. 31, no. 1, pp. 59-64, Jan. 2007.
[43] P. Sørensen, A. D. Hansen, K. Thomsen, T. Buhl, P. E. Morthorst, L. H. Nielsen*, et al.*, "Operation and control of large wind turbines and wind farms - Final report," Tech. Rep., Risø National Laboratories for Sustainable Energy, Technical University of Denmark, Risø-R-1532(EN), 2005.
[44] C. Wan, Z. Xu, J. Østergaard, Z. Y. Dong, and K. P. Wong, "Discussion of "Combined nonparametric prediction intervals for wind power generation"," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, p. 1021, Jul. 2014.
[45] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*: Springer US, 2008.
[46] P. Sorensen, N. A. Cutululis, A. Vigueras-Rodriguez, L. E. Jensen, J. Hjerrild, M. H. Donovan*, et al.*, "Power fluctuations from large wind farms," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 958-965, Aug. 2007.

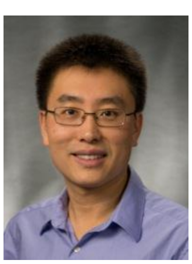


Can Wan (M'15) received his B.Eng. and Ph.D. degrees from Zhejiang University, China, in 2008, and The Hong Kong Polytechnic University in 2015, respectively.

He is with College of Electrical Engineering, Zhejiang University, Hangzhou, China. He was a Postdoc Fellow at Department of Electrical Engineering, Tsinghua University, Beijing, China, and a Research Associate at Department of Electrical Engineering, The Hong Kong Polytechnic University. He was a visiting scholar at the Center for Electric Power and Energy, Technical University of Denmark, and Argonne National Laboratory, IL, USA. His research interests include forecasting, power system uncertainty analysis and operation, renewable energy, active distribution network, demand side, and machine learning.

Jin Lin (S'11-M'12) was born in 1985. He received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2007 and 2012, respectively.

He is currently a Lecturer with the Department of Electrical Engineering, Tsinghua University, Beijing, China. His research interests are in grid integration technology of renewable energy and dynamic power systems.



Jianhui Wang (M'07-SM'12) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, IL, USA, in 2007.

Presently, he is the Section Lead for Advanced Power Grid Modeling at the Energy Systems Division at Argonne National Laboratory, Argonne, IL, USA. Dr. Wang is the secretary of the IEEE Power & Energy Society (PES) Power System Operations Committee. He is an associate editor of *Journal of Energy Engineering* and an editorial board member of *Applied Energy*. He is also an affiliate professor at Auburn University and an adjunct professor at University of Notre Dame. He has held visiting positions in Europe, Australia and Hong Kong including a VELUX Visiting Professorship at the Technical University of Denmark (DTU).

Dr. Wang is the Editor-in-Chief of the *IEEE TRANSACTIONS ON SMART GRID* and an IEEE PES Distinguished Lecturer. He is also the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015.

Yonghua Song (M'90–SM'94–F'08) received the B.Eng. and Ph.D. degrees from Chengdu University of Science and Technology, Chengdu, China, and China Electric Power Research Institute, Beijing, China, in 1984 and 1989, respectively.

He is Executive Vice-President of Zhejiang University, and also an Adjunct Professor at Department of Electrical Engineering, Tsinghua University. From 1989 to 1991, he was a Post-Doctoral Fellow with Tsinghua University, Beijing. He then held various positions at Bristol University, Bristol, U.K.; Bath University, Bath, U.K.; and John Moores University, Liverpool, U.K., from 1991 to 1996. In 1997, he was a Professor of power systems with Brunel University, Uxbridge, U.K., where he has been the Pro-Vice Chancellor of Graduate Studies since 2004. In 2007, he took up the Pro-Vice Chancellorship and Professorship of Electrical Engineering with the University of Liverpool, Liverpool. He was a Professor with the Department of Electrical Engineering, Tsinghua University, where he was an Assistant President and the Deputy Director with the Laboratory of Low-Carbon Energy in 2009. His current research interests include smart grid, electricity economics, and operation and control of power systems.

Prof. Song was a recipient of the D.Sc. Award from Brunel University in 2002 for his original achievements in power systems research. He was elected Vice-President of the Chinese Society for Electrical Engineering (CSEE) and appointed Chairman of the International Affairs Committee of the CSEE in 2009. In 2004, he was elected as a Fellow of the Royal Academy of Engineering, U.K.

Zhao Yang Dong (M'99–SM'06) obtained his Ph.D. degree from the University of Sydney, Australia in 1999.

He is now Head and Chair Professor with School of Electrical and Information Engineering, The University of Sydney, Australia. He previously held academic and industrial positions with The University of Newcastle, The Hong Kong Polytechnic University and Transend Networks (Now TASNetworks), Australia. His research interests include smart grid, power system planning, power system security, load modeling, electricity market, and computational intelligence.

He is an editor of *IEEE TRANSACTIONS ON SMART GRID*, *IEEE TRANSACTIONS ON SUSTAINABLE ENERGY*, *IEEE POWER ENGINEERING LETTERS*, and IET RENEWABLE POWER GENERATION.