Downloaded from orbit.dtu.dk on: Sept 08, 2026

DTU Library

Optimal prediction intervals of wind power generation

Wan, Can; Wu, Zhao; Pinson, Pierre; Dong, Zhao Yang; Wong, Kit Po

Published in:
IEEE Transactions on Power Systems

Link to article, DOI:
10.1109/TPWRS.2013.2288100

Publication date:
2014

Link back to DTU Orbit

Citation (APA):
Wan, C., Wu, Z., Pinson, P., Dong, Z. Y., & Wong, K. P. (2014). Optimal prediction intervals of wind power generation. *IEEE Transactions on Power Systems*, *29*(3). https://doi.org/10.1109/TPWRS.2013.2288100

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Optimal Prediction Intervals of Wind Power Generation

Can Wan, *Student Member, IEEE*, Zhao Xu, *Senior Member, IEEE*, Pierre Pinson, *Senior Member, IEEE*, Zhao Yang Dong, *Senior Member, IEEE*, and Kit Po Wong, *Fellow, IEEE*

***Abstract*—Accurate and reliable wind power forecasting is essential to power system operation. Given significant uncertainties involved in wind generation, probabilistic interval forecasting provides a unique solution to estimate and quantify the potential impacts and risks facing system operation with wind penetration beforehand. This paper proposes a novel hybrid intelligent algorithm approach to directly formulate optimal prediction intervals of wind power generation based on extreme learning machine and particle swarm optimization. Prediction intervals with associated confidence levels are generated through direct optimization of both the coverage probability and sharpness to ensure the quality. The proposed method does not involve the statistical inference or distribution assumption of forecasting errors needed in most existing methods. Case studies using real wind farm data from Australia have been conducted. Comparing with benchmarks applied, experimental results demonstrate the high efficiency and reliability of the developed approach. It is therefore convinced that the proposed method provides a new generalized framework for probabilistic wind power forecasting with high reliability and flexibility and has a high potential of practical applications in power systems.**

***Index Terms*—Extreme learning machine, forecasts, particle swarm optimization, prediction intervals, wind power.**

I. Introduction

WIND energy is the most important and efficient renewable energy and is widely utilized for power generation in modern power systems in the past decades. In particular, wind power can supply up to 20% of annual electricity consumption in Denmark. However, wind power also introduces much more uncertainties than conventional generation due to the chaotic nature of the weather system. Accurate and reliable wind power forecasting becomes highly meaningful to optimize the operation cost and improve the reliability for power systems with increased wind penetration [1].

In the past, most researches focused on point forecasting approaches for wind power [2]. Due to the nonstationarity of wind power series, traditional point forecasting can hardly be accurate, and the forecasting errors are unavoidable and significant to some extent. For traditional applications, wind power forecasting errors are statistically analyzed in [3]. With the integration of high penetration of wind generation in deregulated power systems, the development of probabilistic forecasting tools is needed for making decisions in the operational domain to accounting for wind generation uncertainties. Recently, different approaches have been proposed for probabilistic wind power forecasts to obtain prediction intervals (PIs). Meteorological ensembles are used to obtain predictive distribution and estimate the uncertainty of forecasts [4], [5]. The uncertainty of wind power forecasting is investigated based on the nonlinear power curve and statistical analysis of wind speed prediction errors [6]. Quantile regression is used to estimate different forecasting quantiles [7], [8]. Based on the point prediction results of AWPPS, WPPT and Sipreólico, PIs are constructed through a combined nonparametric probability forecasts and adaptive resampling approach [9]. In [10], radial basis function has been implemented to derive quantile forecasts of wind power based on point prediction results, weather conditions, etc. The conditional kernel density (CKD) estimation approach is proposed to estimate of the probability distribution of wind power generation [11]. In general, PIs with associated confidence levels successfully quantify the uncertainties of wind power forecasts, which essentially benefit all participants in power systems to prepare for possible scenarios in advance and significantly reduce risks facing power system operation and control, such as wind farm control, reserve setting, energy storage sizing, unit commitment, wind power trading, and so forth [12]–[17].

In this paper, a hybrid intelligent algorithm (HIA) based interval forecasting approach is newly developed to produce prediction intervals of wind power generation based on the extreme learning machine (ELM) [18] and particle swarm optimization (PSO) [19]. The proposed HIA method aims to obtain optimal PIs without the prior knowledge, statistical inference or distribution assumption of forecasting errors required in most traditional approaches. ELM applied in the proposed approach is a novel learning algorithm proposed for training single-hidden layer feedforward neural networks (SLFNs) featuring extremely fast learning speed and superior generalization capability. ELM successfully avoids the limitations of traditional neural networks (NNs) learning algorithms, such as local

Manuscript received March 27, 2013; revised July 13, 2013 and September 23, 2013; accepted October 23, 2013. This work was supported in part by Hong Kong RGC GRF grants no.515110, 528412, and T23-407/13-N. The work of C. Wan was supported by a Hong Kong Ph.D. Fellowship. Paper no. TPWRS-00371-2013.

C. Wan is with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, and also with Centre for Electric Power and Energy, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark (e-mail: can.wan@connect.polyu.hk).

Z. Xu is with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, and also with the University of Newcastle, Callaghan NSW 2308, Australia (e-mail: eezhaoxu@polyu.edu.hk).

P. Pinson is with the Centre for Electric Power and Energy, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark (e-mail: ppin@elektro.dtu.dk).

Z. Y. Dong is with the School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia (e-mail: zydong@ieee.org).

K. P. Wong is with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, WA 6009, Australia (e-mail: kitpo@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2013.2288100

0885-8950 © 2013 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

minima, overtraining, high computation costs and so forth. Due to its excellent performance, ELM has been used in many different applications including both regression and classification tasks [20]–[22].

Classical NNs based PIs construction methods always assume that prediction errors are normally distributed [23]–[26]. These methods are not applicable to the case of ELM, since the extreme learning process is very different from that for conventional NNs. Lower upper bound estimation (LUBE) method is proposed for PI construction based on traditional NN [27], applied in load forecasting [28] and wind power forecasting [29]. However, traditional NNs employed in the LUBE method would cause several inevitable limitations, such as overtraining, high computation burden, and so forth. Furthermore, the coverage width-based criterion (CWC) cannot accurately measure the overall skill of constructed PIs. Traditional PIs construction methods for wind power rely on quantile analysis of point forecast errors with or without prior distribution assumptions [4]–[11], where the procedures of PIs formulation and final performance assessment are usually separated. E.g. in [9], PIs can be achieved through a conditional probabilistic modeling between point forecast outputs and associated errors. In contrast, the proposed HIA approach integrates the two procedures holistically to formulate the PIs directly to pursue the best quality of resultant PIs, without the need of prior knowledge and distribution assumption of point forecasts errors. As investigated in [30] as early as 1970s, with a properly constructed cost function, PI estimation could be considered as a Bayesian decision-making procedure to acquire an optimal PI that minimizes the expected cost. The objective function of HIA is specially formulated to address both the coverage probability and sharpness of PIs simultaneously, and is optimized through PSO featuring fast convergence and gradient-free optimization. Furthermore, the proposed method is able to generate multiple optimal PIs of different confidence levels in one single optimization process.

Generally, different decision-makers in power systems have different look-ahead time preferences ranging from minutes to days for wind power forecasts according to their own operational requirements. Very short-term wind power prediction is needed to wind farm control [12], [13], the temporal operation of wind storage systems associated with temporal market regulations such as Australian National Electricity Market with 5-min resolution [31], and the transmission system operator (TSO) which aims to optimally dispatch reserves for the continuous balance of the power system [32], [33]. Hourly ahead forecast is crucial for power system and electricity market balance, e.g., Nord pool market [34]. Longer term forecasts up to days ahead are very meaningful for unit commitment [16], day-ahead market trading [17], etc. The proposed HIA method has been tested using the practical data of two wind farms in Australia. Without loss of generality, in the case study we focus on the hourly forecast on an hourly basis though with extendibility. Comparing with benchmarks, the effectiveness of the proposed method has been proved through comprehensive evaluations with respect to both the reliability and overall skill of the forecasting results. By accurate quantification of the uncertainties of wind generation forecasts, the proposed interval forecasting approach has a high potential to support various operation and planning activities in power systems, such as to provide reliable information for dispatching, e.g., the hourly Nord pool market. Particularly, the interval forecasting results can also be used to develop new operation and planning tools for TSO to probabilistically determine the needed reserves in advance [14], [15], and to facilitate Gencos' risk management through strategic biding [17].

The rest of this paper is organized as follows. Section II introduces the ELM and PIs formulation. Section III describes PIs evaluation indices including reliability and sharpness. Objective function modeling and the detailed procedures of the proposed HIA approach are expressed in Section IV. Comprehensive numerical studies are implemented and analyzed in Section V. Finally, the conclusion is drawn in Section VI.

II. Prediction Interval Formulation

A. Extreme Learning Machine

ELM is a recently developed novel algorithm for training a single hidden-layer feedforward neural network [18]. Different from traditional gradient-based training algorithms in supervised batch learning, ELM randomly chooses the input weights and hidden biases and needs not be tuned in the training process, dramatically saving learning time. Given datasets with N arbitrary distinct samples $\{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^{N}$ where the inputs $\mathbf{x}_i \in \mathbf{R}^n$ and the targets $\mathbf{t}_i \in \mathbf{R}^m$, if the ELM with K hidden neurons and activation function $\psi(\cdot)$ can approximate the N samples with zero error, it can be expressed by the following equation:

$$f_K(\mathbf{x}_j) = \sum_{i=1}^{K} \beta_i \psi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, j = 1, \ldots, N \quad (1)$$

where $\mathbf{a}_i = [a_{i1}, a_{i2}, \ldots, a_{in}]^T$ represents the weight vector connecting the ith hidden neuron and the input neurons, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ denotes the weight vector connecting the ith hidden neuron and the output neurons, b_i denotes the threshold of the ith hidden neuron, and $\psi(\mathbf{a}_i \cdot \mathbf{x}_j + b_i)$ is the output of the ith hidden neuron with respect to the input $\mathbf{x}_j$. Equation (1) can be simplified as

$$\mathbf{H}\beta = \mathbf{T} \quad (2)$$

where $\mathbf{H}$ is the hidden layer output matrix of the modeled ELM, expressed as

$$\mathbf{H} = \begin{bmatrix} \psi(\mathbf{a}_1 \cdot \mathbf{x}_1 + b_1) \cdots \psi(\mathbf{a}_K \cdot \mathbf{x}_1 + b_K) \\ \vdots \cdots \vdots \\ \psi(\mathbf{a}_1 \cdot \mathbf{x}_N + b_1) \cdots \psi(\mathbf{a}_K \cdot \mathbf{x}_N + b_K) \end{bmatrix}_{N \times K}. \quad (3)$$

The ith column of $\mathbf{H}$ denotes the output vector of the ith hidden neuron with respect to the inputs $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T$. In addition, β is the matrix of output weights and $\mathbf{T}$ is the matrix of targets, respectively represented as

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_K^T \end{bmatrix}_{K \times m} \text{ and } \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}. \quad (4)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

After the weights $\mathbf{a}_i$ and the hidden layer biases b_i are randomly assigned, the hidden layer output matrix $\mathbf{H}$ can be uniquely determined, and consequently the estimated parameters $\mathbf{a}_i^*$, b_i^* and β_i^* can be obtained such that

$$\begin{aligned}\|\mathbf{H}(\mathbf{a}_1^*,\ldots,\mathbf{a}_k^*,b_1^*,\ldots,b_k^*)\beta^* - \mathbf{T}\| \\ = \min_{\beta}\|\mathbf{H}(\mathbf{a}_1,\ldots,\mathbf{a}_k,b_1,\ldots,b_k)\beta - \mathbf{T}\|\end{aligned} \quad (5)$$

which is equivalent to minimizing the cost function of the traditional gradient-based back-propagation learning algorithm

$$C = \sum_{j=1}^{N}\left[\sum_{i=1}^{K}\beta_i\psi(\mathbf{a}_i\cdot\mathbf{x}_j + b_i) - \mathbf{t}_j\right]^2. \quad (6)$$

With unchanged input weights and the hidden layer biases of ELM, training an SLFN is simply equivalent to finding a unique smallest norm least-squares solution of the linear system in (2), expressed as

$$\beta^* = \mathbf{H}^{\dagger}\mathbf{T} \quad (7)$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix $\mathbf{H}$, which can be derived through singular value decomposition (SVD) method.

For any infinitely differentiable activation function, when the number of hidden neurons equals to the number of training samples, i.e., $K = N$, ELM can exactly learn with zero error. The ELM overcomes many limitations of traditional gradient based NNs training algorithms, such as the local minima, overtraining, high computational burdens, etc. The traditional gradient based NNs leaning algorithms always involve a number of iterations that affect the training speed. The ELM training features extremely fast speed because of the simple matrix computation, and can always guarantee the optimal performance [18].

B. Formulation of PIs

PIs quantify the uncertainty associated with forecasts. Given a set of process pairs

$$D_t = \{(\mathbf{x}_i, t_i)\}_{i=1}^{N} \quad (8)$$

where t_i is the future target to forecast, and $\mathbf{x}_i$ denotes relevant input variables that can include historical wind power and wind speeds, numerical weather predictions and so on for wind power forecasting in the study. PI with nominal confidence $100(1-\alpha)\%$ of the future target t_i , represented as $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$, can be expressed as the following equation:

$$\tilde{I}_t^{(\alpha)}(\mathbf{x}_i) = [\tilde{L}_t^{(\alpha)}(\mathbf{x}_i), \tilde{U}_t^{(\alpha)}(\mathbf{x}_i)] \quad (9)$$

where $\tilde{L}_t^{(\alpha)}(\mathbf{x}_i)$ and $\tilde{U}_t^{(\alpha)}(\mathbf{x}_i)$ denote the lower and upper bounds of PI $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$, respectively, such that the future target t_i is expected to be enclosed by $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$ with coverage probability

$$P\left(t_i \in \tilde{I}_t^{(\alpha)}(\mathbf{x}_i)\right) = 100(1-\alpha)\%. \quad (10)$$

The proposed method aims to directly generate the lower and upper bounds of the expected PIs by ELM. It should be pointed out that the proposed method actually provides an unique framework capable of generating multiple pairs of PI bounds with different nominal coverage probabilities $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n]^T$ simultaneously through a single optimization approach. The overall structure of the proposed ELM model is shown in Fig. 1, where the ELM takes the inputs and outputs the corresponding PI bounds of different confidence levels.

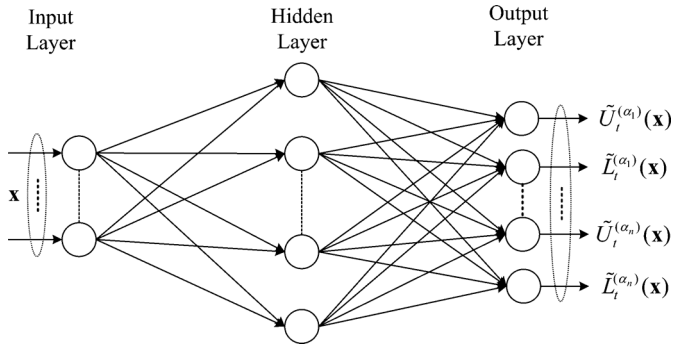

Fig. 1. ELM model for PIs generation by the proposed HIA approach.

III. PIs Evaluation Criteria

In this section, comprehensive PIs evaluation indices are introduced from the perspectives of reliability and sharpness in detail.

A. Reliability

Reliability is regarded as a major property for validating probabilistic forecasting models, due to that low reliability could cause systematic bias involved in following decision-making problems. According to the PIs definition, the future targets t_i are expected to be covered by the constructed PIs with the nominal probability $100(1-\alpha)\%$, termed as PI nominal confidence (PINC). PI coverage probability (PICP), represented by $\hat{P}_t^{(\alpha)}$, is a key measure for the reliability of the constructed PIs [9], [10], defined by

$$\hat{P}_t^{(\alpha)} = \frac{1}{N_t}\sum_{i=1}^{N_t}\kappa_i^{(\alpha)} \quad (11)$$

where N_t is the size of test dataset, and $\kappa_i^{(\alpha)}$ is the indicator of PICP, expressed as

$$\kappa_i^{(\alpha)} = \begin{cases} 1 & t_i \in \tilde{I}_t^{(\alpha)}(\mathbf{x}_i) \\ 0 & t_i \notin \tilde{I}_t^{(\alpha)}(\mathbf{x}_i) \end{cases}. \quad (12)$$

The PICP of derived PIs should asymptotically approach the PINC as closely as possible. Therefore, average coverage error (ACE), represented by $A_t^{(\alpha)}$, can be used to assess the PIs quality [9], [10], defined by

$$A_t^{(\alpha)} = \hat{P}_t^{(\alpha)} - PINC. \quad (13)$$

The value of ACE should diminish towards zero as closely as possible, i.e., the smaller the absolute ACE is, the higher reliability the obtained PIs possess.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

B. Sharpness

Under the condition of high reliability, sharper PIs have higher quality and would be preferred by decision makers. The sharpness is an indispensable aspect of PIs quality and should be considered in the evaluation process [35]. It can be understood that high reliability can be easily achieved by simply increasing or decreasing the distance between the bounds of PI, which can result in the degradations of sharpness. Obviously, the resultant PIs can be far from satisfactory and would be useless in practice since they cannot provide accurate quantifications of uncertainties involved in the real-world processes.

The width of PI $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$ defined in (9), represented by $\vartheta_t^{(\alpha)}(\mathbf{x}_i)$, can be calculated through

$$\vartheta_t^{(\alpha)}(\mathbf{x}_i) = \tilde{U}_t^{(\alpha)}(\mathbf{x}_i) - \tilde{L}_t^{(\alpha)}(\mathbf{x}_i). \tag{14}$$

In the study, we focus on obtaining PIs with two quantiles at particular confidences. Therefore, the interval score can be used to assess the overall skill of wind power PIs to involve the sharpness aspect [30]. The interval score of the specific PI $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$, represented by $S_t^{(\alpha)}(\mathbf{x}_i)$, is defined through

$$S_t^{(\alpha)}(\mathbf{x}_i) = \begin{cases} -2\alpha\vartheta_t^{(\alpha)}(\mathbf{x}_i) - 4[\tilde{L}_t^{(\alpha)}(\mathbf{x}_i) - t_i], & \text{if } t_i < \tilde{L}_t^{(\alpha)}(\mathbf{x}_i) \\ -2\alpha\vartheta_t^{(\alpha)}(\mathbf{x}_i), & \text{if } t_i \in \tilde{I}_t^{(\alpha)}(\mathbf{x}_i) \\ -2\alpha\vartheta_t^{(\alpha)}(\mathbf{x}_i) - 4[t_i - \tilde{U}_t^{(\alpha)}(\mathbf{x}_i)], & \text{if } t_i > \tilde{U}_t^{(\alpha)}(\mathbf{x}_i). \end{cases} \tag{15}$$

The score is calculated for each prediction point and then the overall score value $\bar{S}_t^{(\alpha)}$ can be derived as the average over the entire test dataset

$$\bar{S}_t^{(\alpha)} = \frac{1}{N_t}\sum_{i=1}^{N_t} S_t^{(\alpha)}(\mathbf{x}_i). \tag{16}$$

Obviously, the score awards the narrow PI and penalizes it if the target is not enclosed. Including all aspects of PIs evaluation, the interval score can be used to compare the overall skill of interval forecasts. However, the score cannot quantitatively distinguish the contributions of reliability and sharpness to the overall skill. Though, based on a prior analysis of reliability, a skill score can still be employed to carry out an assessment from the sharpness perspective. Given PIs with the same PINC and similar reliability, the smaller the absolute score $\bar{S}_t^{(\alpha)}$ indicates the higher sharpness and consequently the higher quality.

With the described PIs assessment criteria above, we can see that both the reliability and sharpness should be taken into consideration to comprehensively assess the quality of constructed PIs. Meanwhile, it should be noted that reliability is the primary feature reflecting the correctness of the constructed PIs. Generally, in the evaluation process, the reliability of PIs should be prioritized.

IV. Optimal Construction of PIs

A. Objective Function

The proposed HIA method adopts an ELM to predict the PIs and pursues the optimal quality of produced PIs without statistical inferences and distribution assumptions for forecasting errors. Because of the unique properties of ELM described in Section II-A, training the ELM based forecasters is equivalent to analytically determining the output weights alone. Comprehensive PIs evaluation criteria are well established and described in Section III. To ensure the quality of produced PIs, ELM output weights are optimized to account for both reliability and sharpness of the generated PIs simultaneously, which can be considered as a multi-objective optimization problem (MOOP) [36].

A multi-objective function for training the proposed model is developed based on well-established PIs evaluation criteria introduced in the previous section to produce optimal PIs. It should be highlighted that though the interval score accounts for reliability and sharpness, it cannot quantitatively distinguish the contributions of the two aspects. However, the interval score can provide an evaluation from the perspective of sharpness given a prior analysis of reliability. Under the same nominal confidence and similar reliability, PIs with the smaller the absolute score $\bar{S}_t^{(\alpha)}$ have the higher sharpness and the higher quality. The interval score is not a dedicated index for reliability assessment anyhow. As the primary requirement of probabilistic forecasting, the reliability of PIs should be given a prior analysis in the assessment process. Therefore to specifically quantify and emphasize the reliability aspect, ELM output weights β are optimized with respect to the objective F combining ACE $A_t^{(\alpha)}$ and overall score $\bar{S}_t^{(\alpha)}$ to optimize both reliability and sharpness of PIs at particular confidence levels $100(1-\alpha_i)\%, i = 1, 2, \ldots, n$

$$\min_{\beta} F = \sum_{i=1}^{n}\left[\gamma_i \left|A_t^{(\alpha_i)}\right| + \lambda_i \left|\bar{S}_t^{(\alpha_i)}\right|_{norm}\right] \tag{17}$$

$$\text{s.t } \tilde{L}_t^{(\alpha_i)}(\mathbf{x}) \geq \tilde{L}_t^{(\alpha_j)}(\mathbf{x}) \text{ and } \tilde{U}_t^{(\alpha_i)}(\mathbf{x}) \leq \tilde{U}_t^{(\alpha_j)}(\mathbf{x}), \text{ if } \alpha_i \leq \alpha_j \tag{18}$$

$$\tilde{L}_t^{(\alpha_i)}(\mathbf{x}) \leq \tilde{U}_t^{(\alpha_i)}(\mathbf{x}) \tag{19}$$

where $|\cdot|$ is the absolute value function, $A_t^{(\alpha_i)}$ is ACE of PIs with corresponding PINC $100(1-\alpha_i)\%$, $|\bar{S}_t^{(\alpha_i)}|_{norm}$ denotes the normalized absolute interval score $|\bar{S}_t^{(\alpha_i)}|$ which is normalized over the corresponding maximum score $|\bar{S}_t^{(\alpha_i)}|_{\max}$ and minimum score $|\bar{S}_t^{(\alpha_i)}|_{\min}$, defined by

$$\left|\bar{S}_t^{(\alpha_i)}\right|_{norm} = \frac{\left|\bar{S}_t^{(\alpha_i)}\right| - \left|\bar{S}_t^{(\alpha_i)}\right|_{\min}}{\left|\bar{S}_t^{(\alpha_i)}\right|_{\max} - \left|\bar{S}_t^{(\alpha_i)}\right|_{\min}} \tag{20}$$

and γ_i and λ_i are importance weights of the reliability and overall skill (including sharpness), respectively. With the normalized objectives, the importance weights γ_i and λ_i are set as unit values in the study. The compatibility of the resultant PIs with different confidence levels can be assured through the constraints given in (18) and (19). The minimum value $|\bar{S}_t^{(\alpha_i)}|_{\min}$ is set to 0, which means the perfect condition with exact forecasting results. The maximum value $|\bar{S}_t^{(\alpha_i)}|_{\max}$ is set to 2α, which indicates the most conservative PIs with the maximum width.

B. Particle Swarm Optimization

Particle swarm optimization is a heuristic and population based optimization method and has proved to be an efficient, robust and gradient-free optimization algorithm [19]. PSO also

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

distinguishes itself from other heuristic optimization methods by its fast convergence speed. It can be seen that the objective function in (17) is non-differentiable with respect to the ELM output weights. Therefore PSO is applied for objective function minimization to obtain the optimized ELM.

Given that the search space of PSO is S-dimensional and the size of the particles population is N_P, the ith particle of the swarm can be represented by the S-dimensional vector $x_i = [x_{i1}, x_{i2}, \ldots, x_{iS}]^T$ and the best particle in the swarm, i.e., the particle generating the smallest objective function value, is expressed by P_g^b. The previous best position, i.e., the position with the smallest objective function value of the ith particle, is stored in a vector and expressed as $P_i^b = [P_{i1}^b, P_{i2}^b, \ldots, P_{iS}^b]^T$, and the position velocity of the ith particle is represented as $v_i = [v_{i1}, v_{i2}, \ldots, v_{iS}]^T$. In each iteration of PSO, the velocity of each particle is computed, and the particles are manipulated accordingly

$$v_i = wv_i + c_1R_1(P_i^b - x_i) + c_2R_2(P_g^b - x_i) \quad (21)$$

$$x_i = x_i + \phi v_i \quad (22)$$

where $i = 1, 2, \ldots, N_P$; w is the inertia weight; ϕ is a constriction factor controlling and keeping the velocity within the range $[-v_{\max}, +v_{\max}]$; c_1 and c_2 are two positive constants; R_1 and R_2 are random numbers within $[0, 1]$. The velocity of the ith particle is a function with respect to three components: the particle's previous velocity, the distance between the previous best position of the particle and its current position, and the distance between the swarm's best success and the particle's current location. The performance of each particle is evaluated through the objective function modeled.

C. Hybrid Intelligent Algorithm for PI Optimization

The proposed HIA method aims to achieve the PIs of the best quality through directly optimizing the ELM with respect to the objective function (17) using PSO. The core idea underneath is simply to directly approximate the PIs through a regression procedure using the PSO based optimization, where the objective function strictly measures the quality of resultant PIs including both reliability and sharpness. The major steps of the developed algorithm are described as follows:

Step 1) With the historical data of wind generation, wind speed and numerical weather prediction information and so forth, formulate the dataset $D_t = \{(\mathbf{x}_i, t_i)\}_{i=1}^N$, based on which two training datasets $D_t^+ = \{(\mathbf{x}_i, \hat{t}_i^+)\}_{i=1}^N$ and $D_t^- = \{(\mathbf{x}_i, \hat{t}_i^-)\}_{i=1}^N$, respectively, for the upper and lower bounds of the PI should be prepared for ELM initialization. The targets of bounds including $\hat{t}_i^+$ and $\hat{t}_i^-$ can be generated by slightly increasing or decreasing original t_i by, e.g., $\pm\rho\%$, $0 < \rho < 100$, respectively. This manipulation is based on the knowledge that the actual wind power should be enclosed by the potential PIs.

Step 2) Given the randomly determined the input weights $\mathbf{a}_i$ and biases b_i, establish an ELM to initialize the output weights β_{int} which is S-dimensional, using the modified training datasets obtained in Step 1).

Step 3) Initialize a population array of particles *Pop* with random positions around the output weights β_{int} of the ELM obtained in the Step 2) and velocities V in the S-dimensional search space.

Step 4) Set the iteration counter $L = 0$.

Step 5) WHILE maximum number of iterations or sufficiently good fitness has not been reached, do

a) For each particle in *Pop*, evaluate the objective function according to the PIs generated by ELM with the output weights over the original training data D_t.

b) Compare the particle's evaluation through value of objective function (17) satisfying the constraints (18) and (19) with its previous best position P_i^b. If current value is better than that of P_i^b, then set P_i^b equal to the current location.

c) Identify the particle in the swarm better than the best experience and update the smallest value of objective function (17) and the best position P_g^b.

d) Change the velocities and move the positions of particles according to (21) and (22).

e) Keep the particles in the given search space in case that they exceed their valid boundaries, and when the decision variable is out of its lower or upper boundary, takes the value of its corresponding boundary.

f) Increment the iteration counter $L = L + 1$.

Step 6) END WHILE

Step 7) Based on the test data, evaluate the PIs generated by the ELM with optimized parameters β.

According to the detailed procedures of the proposed algorithm introduced above, the proposed HIA approach can construct an optimized ELM to directly generate the bounds of PIs with different confidences of the best quality, avoiding the efforts needed for statistical inference and distribution assumption of point forecasting errors for traditional approaches. The application of ELM provides an extremely fast initialization procedure and significantly reduces the complexity of optimizing decision variables. The proposed algorithm demonstrates high flexibility due to the high mapping capability of ELM. The proposed HIA approach is indeed performance-oriented, and the quality of constructed PIs can be ensured through optimization on the formulated objective function.

V. Case Studies

A. Introduction of Experiment Data

The highly chaotic climate systems are responsible for the high level of uncertainties in wind power generation. To comprehensively validate the effectiveness of the proposed approach, it is tested by two wind farms the Challicum Hills wind farm and the Starfish Hill wind farm in Australia. The weather conditions and wind speeds vary significantly in the two regions where the wind farms locate. Therefore forecasting models and case studies are separately constructed and conducted for the two wind farms, respectively.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The Challicum Hills wind farm locates near Ararat in western Victoria, Australia, with coordinate latitude -37.38° S and longitude 143.09° E. The wind farm has a combined generating capacity $P_c 52.5$ MW consisting of 35 wind turbines of 1.5 MW. Wind power generation data with one-hour resolution of this wind farm used in the study covers the period from September 2008 to August 2010.

The second wind farm Starfish Hill is near Cape Jervis on the Fleurieu Peninsula, South Australia, with coordinate latitude -35.57° S and longitude 138.16° E. It consists of 23 wind turbines of 1.5 WM each, with a total installed capacity of 34.5 MW. Wind power generation data with one-hour resolution of Starfish Hill wind farm used in the study covers the period from January 2009 to May 2010.

To ensure both forecasting performance and computation efficiency, in the case study the wind power series is used as the inputs alone to the proposed HIA approach to conduct hourly ahead forecasting, of which the results can be significant to generation and ancillary service dispatch and so on in practice, e.g., in the Nord Pool market in Scandinavia, the hourly market plays a key role in maintaining system balance [34].

B. Experimental Results and Analysis

To evaluate the forecast performance of the proposed approach, five other PI forecasting methods including the climatology method, the constant forecast method, the persistence method, the exponential smoothing method (ESM), and the quantile regression (QR) approach are employed to compute PIs using the same training and testing data for benchmarking.

The climatology is the most commonly used benchmark for probabilistic forecasts of meteorological or weather-related processes. It is the unconditional predictive distribution computed from all historical observations available. The constant forecast takes the form of normal distribution, and the mean and variance are derived from the observed wind power data. Since the climatology and constant approaches are fairly easy to outperform for short look-ahead time forecasting, other three methods also are applied for comparisons. For point forecasting, the persistence forecast method is a widely used benchmark and is known to be difficult to outperform for short look-ahead time. The persistence based probabilistic forecast model is used as benchmark herein, of which the forecast error is assumed to be random and normally distributed. Its mean is given by the last available power measurement, and the variance is computed using the latest observations. In addition, a nice benchmark the exponential smoothing method is employed for comparisons as well, which applies a normal predictive density with its conditional mean based on exponential smoothing of past measured values and its conditional variance determined from exponential smoothing of previous squared residuals [33]. It is obvious that both the persistence and ESM approaches are based on the normal assumption of forecasting uncertainty. To better demonstrate the effectiveness of the proposed approach, quantile regression approach is employed as an advanced benchmark, which does not need the assumption of probability distribution for forecasting errors [7], [37].

The proposed model mainly aims to optimally compute reliable PIs with expected confidences. In practice, power system operation always requires accurate information with high confidence levels, e.g., state estimation always pursues higher confidence level like in [38] to ensure operation security. Therefore it is much more practically meaningful to produce high confidence level PIs to satisfy the requirements of power system operation. In our case study, PIs with different PINCs involving 90%, 95%, and 99% are constructed to evaluate the performance of the proposed approach, i.e., $\alpha = [0.1, 0.05, 0.01]^T$ and $n = 3$ in the optimization objective function defined by (17). The parameter ρ in Step 1) of the HIA algorithm is set to 30 in the case studies. The proposed method and applied benchmarks are tested for the two wind farms for detailed analysis and comparisons. For the Challicum Hills wind farm, the wind power generation data from March 2010 to August 2010 are used for testing the forecasting methods. For the Starfish Hill wind farm, the wind power generation data from January 2010 to May 2010 are used for testing the forecasting methods. The rest data of the two wind farms are used for training the applied methods separately.

The detailed testing results from the two wind farms, including the PIs evaluation indices PICP, ACE and overall score, are given in Tables I and II , respectively. It can be observed that the proposed method can provide fairly satisfactory performances for both wind farms from Tables I and II. At all confidence levels in the case studies, the PICPs of the proposed method are close to the corresponding nominal confidences. The absolute ACEs obtained from the proposed method at different nominal confidence levels for the two farms are smaller than 1%, indicating a significantly high reliability of the generated PIs. E.g., at the confidence level with $\mathrm{PINC} = 90\%$, the proposed method produces PICPs of 90.80% and 90.91% for the Challicum Hills wind farm and the Starfish Hill wind farm respectively, which outperform all other methods. As an advanced approach, quantile regression method provides comparable reliability as the proposed approach, better than the other four benchmarks. Nevertheless, the proposed method has the smallest absolute interval scores for all studied cases in the two wind farms, which indicates the best overall skill and the highest sharpness of the PIs generated by the proposed approach compared to other methods. E.g., at the nominal confidence level 90%, the proposed method produces PIs with absolute interval score 6.43% for the Starfish Hill wind farm, which outperforms the applied five benchmarks. Accounting for both reliability and overall skill, the proposed HIA approach produces the best PIs in terms of comprehensive performance against the other five benchmarks.

The climatology and constant approaches are unconditional forecasts and do not take into account the nonstationarity and heteroscedasticity of wind power series. Though PIs derived by the climatology and constant forecasts demonstrate fair reliability at the tested high confidence levels, they are generally too wide with low sharpness and therefore not meaningful for practical applications. ESM and persistence based interval forecasting approaches are difficult to outperform for short-term forecasts. According to the experiment results, the ESM and persistence forecasts cannot generate PIs to best fit the expected confidences especially for the PINCs larger than 95%. According to the experiment results, quantile regression

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
RESULTS OF DIFFERENT METHODS IN CHALLICUM HILLS WIND FARM

PINC	Methods	PICP	ACE	Score
90%	Proposed Method	90.80%	0.80%	-6.61%
	Constant	89.31%	-0.69%	-19.77%
	Climatology	92.57%	2.57%	-18.54%
	Persistence	85.87%	-4.13%	-7.24%
	ESM	89.10%	-0.90%	-7.33%
	QR	91.43%	1.43%	-7.30%
95%	Proposed Method	95.50%	0.50%	-3.94%
	Constant	92.84%	-2.16%	-10.41%
	Climatology	94.81%	-0.19%	-9.76%
	Persistence	89.61%	-5.39%	-4.55%
	ESM	92.43%	-2.57%	-4.57%
	QR	95.83%	0.83%	-4.42%
99%	Proposed Method	98.48%	-0.52%	-1.14%
	Constant	99.91%	0.91%	-1.94%
	Climatology	97.73%	-1.27%	-2.05%
	Persistence	94.14%	-4.86%	-1.65%
	ESM	96.76%	-2.24%	-1.67%
	QR	99.14%	0.14%	-1.23%

TABLE II
RESULTS OF DIFFERENT METHODS IN STARFISH HILL WIND FARM

PINC	Methods	PICP	ACE	Score
90%	Proposed Method	90.91%	0.91%	-6.43%
	Constant	93.48%	3.48%	-16.33%
	Climatology	97.65%	7.65%	-16.47%
	Persistence	86.93%	-3.07%	-7.17%
	ESM	89.06%	-0.94%	-7.12%
	QR	91.60%	1.60%	-7.15%
95%	Proposed Method	94.90%	-0.10%	-4.00%
	Constant	99.01%	4.01%	-8.44%
	Climatology	99.59%	4.59%	-8.62%
	Persistence	90.33%	-4.67%	-4.62%
	ESM	92.65%	-2.35%	-4.48%
	QR	96.05%	1.05%	-4.31%
99%	Proposed Method	99.28%	0.28%	-1.08%
	Constant	100.0%	1.00%	-2.00%
	Climatology	100.0%	1.00%	-1.83%
	Persistence	94.88%	-4.12%	-1.85%
	ESM	96.11%	-2.89%	-1.67%
	QR	99.28%	0.28%	-1.16%

approach can derive relatively comparable PIs to the proposed approach, especially from the aspect of reliability. Generally, it performs better than the other four benchmarks from the perspectives of both reliability and sharpness. This should not be unreasonable since the quantile regression approach does not require any distribution assumption of forecasting errors, as a conditional forecasting approach.

PIs with PINC 90% obtained by the proposed method and the corresponding actual wind power are displayed in Figs. 2–5 where the actual measured wind farm outputs are perfectly covered by the constructed PIs in the tested two wind farms. Figs. 2–5 visually demonstrate the highly satisfactory performance of the proposed approach in different months for the two wind farms. It also can be easily found that the wind power series have different nonstationary characteristics at different time and different regions. In consideration of that some generated PIs may have abnormal values beyond the possible generation range of the wind farms, the resultant predictive densities shown in Figs. 2–5 have been censored to concentrate probability of abnormal conditions mass on the bounds.

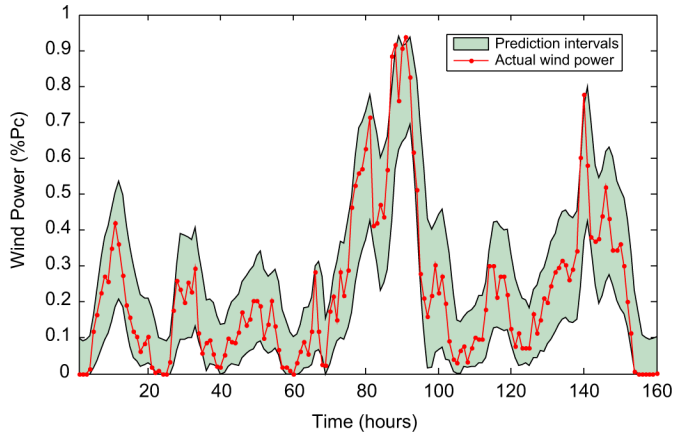

Fig. 2. PIs with PINC 90% in March 2010 of the Challicum Hills wind farm obtained the proposed HIA approach.

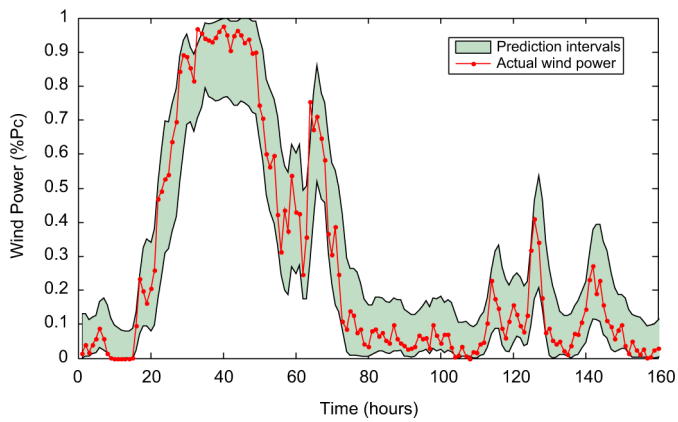

Fig. 3. PIs with PINC 90% in June 2010 of the Challicum Hills wind farm obtained by the proposed HIA approach.

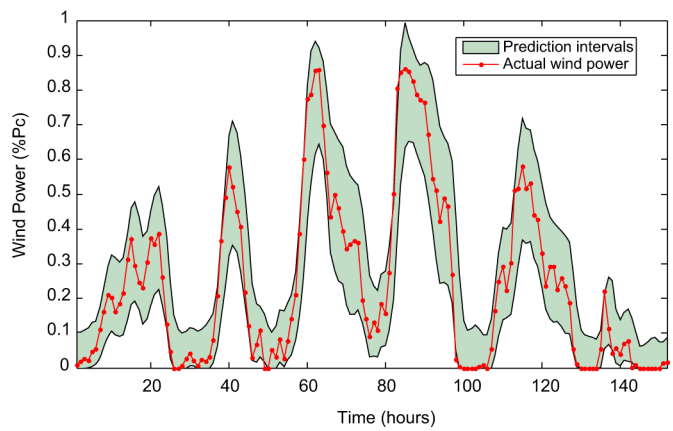

Fig. 4. PIs with PINC 90% in February 2010 of the Starfish Hill wind farm obtained by the proposed HIA approach.

The experimental results demonstrate that the proposed method is highly satisfactory for short-term probabilistic wind power forecasting in comparisons with other five benchmarks including both time series and statistical models. Though wind power series is taken as the input alone to produce hourly ahead PIs in the case study, the proposed HIA approach in this paper

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

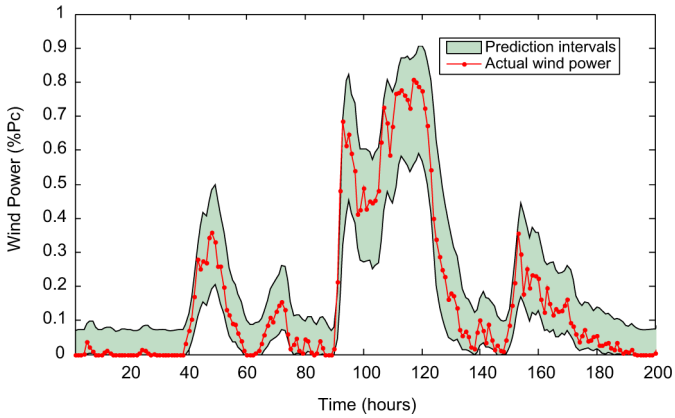

Fig. 5. PIs with PINC 90% in May 2010 of the Starfish Hill wind farm obtained by the proposed HIA approach.

gives a generalized forecasting framework having the advantages of flexible extendibility in terms of inputs, outputs and look-ahead time window, because of the high mapping ability of ELM. It is well known that wind power generation fluctuates due to the volatility of the wind speed, wind direction, etc. For wind power prediction with longer than a few hours look-ahead time, it is necessary to involve numerical weather prediction data as the forecasting model inputs. Certainly, this can be easily included to the proposed model.

In most existing interval forecasting methods, it is necessary to conduct quantile analysis of point forecast errors involving statistical inferences, with or without prior assumption of the forecast error distribution. For instance, in the case study the ESM and persistence rely on the normal assumption of wind power forecasting errors. Comparing with quantile regression without the need of distribution assumption, the proposed approach shows flexible and higher regression ability due to the universal mapping capability of ELM. The HIA approach focuses on PIs quality and offers a novel framework that does not require any information of point forecast results or the associated errors at all. Moreover, since the proposed method provides a performance oriented optimization model, the quality of PIs can be ensured through the optimization directly. Due to the optimization and flexibility, it has high potential practical applications to power systems operation, including reserve determination, wind power trading, wind farm control, unit commitment and so on.

VI. Conclusion

Wind power forecasting is critical to modern power system operation with increased wind penetration. However, wind power forecasting errors are naturally inevitable due to the chaotic nature of weather systems. Traditional probabilistic wind power forecasting approaches are usually based on prior knowledge or assumption of forecasting errors. In this paper, a novel HIA approach combining extreme learning machine and particle swarm optimization is developed and successfully applied for interval forecasting of wind power without the prior knowledge of forecasting errors. A novel objective function accounting for PIs coverage probability and overall skill is constructed to obtain optimal PIs at multiple confidence levels simultaneously through one single performance-oriented optimization process to ensure both reliability and sharpness. The effectiveness of the proposed method for short term forecast has been successfully verified through tests and comparisons with several well-established benchmarks using practical wind farm data. The proposed HIA approach provides a general framework of probabilistic wind power forecasting, with high flexibility. With large scale of wind power integration in modern power systems, the proposed HIA approach indicate high potential in practical applications in power systems operations, e.g., reserve determination by TSO to meet the load and safely and economically operate the systems.

References

[1] R. J. Barthelmie, F. Murray, and S. C. Pryor, "The economic benefit of short-term forecasting for wind energy in the UK electricity market," *Energy Policy*, vol. 36, no. 5, pp. 1687–1696, May 2008.

[2] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renew. Sustain. Energy Rev.*, vol. 12, no. 6, pp. 1725–1744, Aug. 2008.

[3] 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.

[4] P. Pinson and H. Madsen, "Ensemble-based probabilistic forecasting at horns rev," *Wind Energy*, vol. 12, no. 2, pp. 137–155, 2009.

[5] G. Giebel, L. Landberg, J. Badger, K. Sattler, H. Feddersen, T. S. Nielsen, H. A. Nielsen, and H. Madsen, "Using ensemble forecasting for wind power," in *Proc. EWEC'03*, Madrid, Spain, Jun. 2003.

[6] M. Lange, "On the uncertainty of wind power predictions-Analysis of the forecast accuracy and statistical distribution of errors," *J. Solar Energy Eng. (Trans. ASME)*, vol. 127, no. 2, pp. 177–184, 2005.

[7] 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, 2006.

[8] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, 2004.

[9] 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.

[10] 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.

[11] J. Jeon and J. W. Taylor, "Using conditional kernel density estimation for wind power density forecasting," *J. Amer. Statist. Assoc.*, vol. 107, no. 497, pp. 66–79, Mar. 2012.

[12] P. Sørensen, A. D. Hansen, K. Thomsen, T. Buhl, P. E. Morthorst, L. H. Nielsen, F. Iov, F. Blaabjerg, H. A. Nielsen, H. Madsen, and M. H. Donovan, 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.

[13] J. R. Kristoffersen and P. Christiansen, "Horns rev offshore wind farm: its main controller and remote control system," *Wind Eng.*, vol. 27, no. 5, pp. 351–359, Sep. 2003.

[14] M. Matos and R. Bessa, Decision Support Tool for Power Balance and Reserve Management, Tech. Rep., ANEMOS.plus Project Deliverable Report 3.3, 2008.

[15] R. J. Bessa, M. A. Matos, I. C. Costa, L. Bremermann, I. G. Franchin, R. Pestana, N. Machado, H. P. Waldl, and C. Wichmann, "Reserve setting and steady-state security assessment using wind power uncertainty forecast: a case study," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 827–835, Oct. 2012.

[16] Y. V. Makarov, P. V. Etingov, J. Ma, Z. Y. Huang, and K. Subbarao, "Incorporating uncertainty of wind power generation forecast into power system operation, dispatch, and unit commitment procedures," *IEEE Trans. Sustain. Energy*, vol. 2, no. 4, pp. 433–442, Oct. 2011.

[17] P. Pinson, C. Chevallier, and G. 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.

[18] 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, Dec. 2006.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[19] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, Australia, 1995, vol. 4, pp. 1942–1948.
[20] A. H. Nizar, Z. Y. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 946–955, Aug. 2008.
[21] G. B. Huang, X. J. Ding, and H. M. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, no. 1-3, pp. 155–163, Dec. 2010.
[22] G. B. Huang, H. M. Zhou, X. J. 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.
[23] J. T. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *J. Amer. Statist. Assoc.*, vol. 92, no. 438, pp. 748–757, Jun. 1997.
[24] D. J. C. MacKay, "The evidence framework applied to classification networks," *Neural Comput.*, vol. 4, no. 5, pp. 720–736, Sep. 1992.
[25] T. Heskes, *Practical Confidence and Prediction Intervals, Neural Information Processing Systems*, T. P. M. Mozer and M. Jordan, Eds. Cambridge, MA, USA: MIT Press, 1997, vol. 9, pp. 176–182.
[26] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proc. IEEE Int. Conf. Neural Netw.*, Orlando, FL, USA, 1994, vol. 1, pp. 55–60.
[27] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 337–346, Mar. 2011.
[28] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problem," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.
[29] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation forecasts," *IEEE Trans. Sustain. Energy*, to be published.
[30] R. L. Winkler, "A decision-theoretic approach to interval estimation," *J. Amer. Statist. Assoc.*, vol. 67, no. 337, pp. 187–191, Mar. 1972.
[31] Australian Energy Market Operator (AEMO) [Online]. Available: http://www.aemo.com.au/.
[32] P. Sørensen, N. A. Cutululis, A. Vigueras-Rodriguez, L. E. Jensen, J. Hjerrild, M. H. Donovan, and H. Madsen, "Power fluctuations from large wind farms," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 958–965, Aug. 2007.
[33] P. Pinson, "Very short-term probabilistic forecasting of wind power with generalized logit-Normal distributions," *J. R. Statist. Soc., Series C*, vol. 61, no. 4, pp. 555–576, Aug. 2012.
[34] H. Holttinen, "Impact of hourly wind power variations on the system operation in the Nordic countries," *Wind Energy*, vol. 8, no. 2, pp. 197–218, 2005.
[35] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. R. Statist. Soc.: Series B (Statist. Methodol.)*, vol. 69, pp. 243–268, 2007.
[36] K. E. Parsopoulos and M. N. Vrahatis, "Particle swarm optimization method in multiobjective problems," in *Proc. 2002 ACM Symp. Appl. Computing (SAC 2002)*, Madrid, Spain, 2002, pp. 603–607.
[37] J. Juban, L. Fugon, and G. Kariniotakis, "Uncertainty estimation of wind power forecasts: Comparison of probabilistic modelling approaches," in *Proc. Eur. Wind Energy Conf. EWEC'08*, Brussels, Belgium, Apr. 3, 2008.
[38] A. K. Al-Othmana and M. R. Irving, "Analysis of confidence bounds in power system state estimation with uncertainty in both measurements and parameters," *Elect. Power Syst. Res.*, vol. 76, no. 12, pp. 1011–1018, Aug. 2006.

Can Wan (S'10) received the B.Eng. degree from Zhejiang University, China, in 2008. He is now pursuing the Ph.D. degree from The Hong Kong Polytechnic University, with a Hong Kong Ph.D. Fellowship.

He is also a visiting scholar at the Center for Electric Power and Energy, Technical University of Denmark. His research interests include power systems analysis and planning, grid integration of renewable energies, machine learning, and system engineering.

Zhao Xu (M'06–SM'12) received the B.Eng., M.Eng, and Ph.D. degrees from Zhejiang University, China, in 1996, National University of Singapore, Singapore, in 2002, and The University of Queensland, Australia, in 2006, respectively.

He is now with The Hong Kong Polytechnic University. He was previously an Associate Professor with the Centre for Electric Power and Energy, Technical University of Denmark. His research interest includes demand side, grid integration of renewable energies and EVs, electricity market planning and management, and AI applications in power engineering.

Dr. Xu is an Editor of the *Electric Power Components and Systems* journal.

Pierre Pinson (M'11–SM'13) received the M.Sc. degree in applied mathematics from the National Institute for Applied Sciences (INSA Toulouse, France) and the Ph.D. degree in Energy from Ecole des Mines de Paris.

He is a Professor in modeling of electricity markets at the Technical University of Denmark, Centre for Electric Power and Energy, Department of Electrical Engineering. His research interests include among others forecasting, uncertainty estimation, optimization under uncertainty, decision sciences, and renewable energies.

Prof. Pinson acts as an Editor for the IEEE TRANSACTIONS ON POWER SYSTEMS and for *Wind Energy*.

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

He is now Head and Chair Professor with the 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, the University of Queensland, Australia, and Transend Networks, Australia. His research interests include smart grid, power system planning, power system security, load modeling, renewable energy systems, electricity market, and computational intelligence.

Prof. Dong is an editor of the IEEE TRANSACTIONS ON SMART GRID and the IEEE POWER ENGINEERING LETTERS.

Kit Po Wong (M'87–SM'90–F'02) received the M.Sc, Ph.D., and D.Eng. degrees from the University of Manchester, Institute of Science and Technology, Manchester, U.K., in 1972, 1974, and 2001, respectively.

He was with the University of Western Australia, Perth, Australia, from 1974 until 2004 and is now an Adjunct Professor there. Since 2002, he has been Chair Professor, and previously Head, of the Department of Electrical Engineering, The Hong Kong Polytechnic University. His current research interests include computation intelligence applications to power system analysis, planning and operations, as well as power market.

Prof. Wong received 3 Sir John Madsen Medals (1981, 1982, and 1988) from the Institution of Engineers Australia, the 1999 Outstanding Engineer Award from IEEE Power Chapter Western Australia, and the 2000 IEEE Third Millennium Award. He was a Co-Technical Chairman of IEEE Machine Learning and Cybernetics (ICMLC) 2004 Conference and General Chairman of IEEE/CSEE PowerCon2000. He is now Editor-in-Chief of IEEE POWER ENGINEERING LETTERS and was Editor-in-Chief of *IEE Proc. Generation, Transmission & Distribution*. He is also a Fellow of IET, HKIE, and IEAust.