An Advanced Approach for Construction of Optimal Wind Power Prediction Intervals

Guoyong Zhang, Yonggang Wu, Kit Po Wong, *Fellow, IEEE*, Zhao Xu, *Senior Member, IEEE*, Zhao Yang Dong, *Senior Member, IEEE*, and Herbert Ho-Ching Iu, *Senior Member, IEEE*

***Abstract*—High-quality wind power prediction intervals (PIs) are of utmost importance for system planning and operation. To improve the reliability and sharpness of PIs, this paper proposes a new approach in which the original wind power series is first decomposed and grouped into components of reduced order of complexity using ensemble empirical mode decomposition and sample entropy techniques. The methods for the prediction of these components with extreme learning machine technique and the formation of the overall optimal PIs are then described. The effectiveness of proposed approach is demonstrated by applying it to real wind farms from Australia and National Renewable Energy Laboratory. Compared to the existing methods without wind power series decomposition, the proposed approach is found to be more effective for wind power interval forecasts with higher reliability and sharpness.**

***Index Terms*—Ensemble empirical mode decomposition, extreme learning machine, prediction intervals, sample entropy, wind power.**

I. Introduction

DISTRIBUTED renewable energy resources have rapidly grown in recent years as the large-scale construction of smart grid. Renewable energy generators, such as wind power, are always used to lower the costs of fuel consumption and greenhouse gas emissions. However, the intermittency and variability of wind power brings a high level of uncertainty into electricity generation, transmission, and distribution [1]. In tackling these issues, accurate and reliable wind power forecasting becomes a meaningful tool to ensure the reliability and economy of power systems.

Most of existing wind power forecasting methods focused on the deterministic point prediction. However, point prediction errors always exist and cannot be eliminated due to the non-stationary and highly volatile nature of wind power time series [2], [3]. In [4]–[6], the probability density of the point prediction errors is statistically analyzed by Gaussian, Beta, and Versatile distribution. However, it is difficult to represent the forecast errors of a real wind farm by any one form of these probabilistic distributions. For this problem, prediction interval (PI) with a prescribed confidence level is essential for quantification of the potential uncertainties and risks. Unlike the point prediction method, which only provide the point prediction value but does not indicate the probability for correct prediction, PI is an interval consisting of lower and upper bound, which not only provides a range that targets value will lie within, but also has an indication of their accuracy [7], [8].

Conventional PIs construction methods are often placed after a deterministic forecasting model with special prior assumptions. Bayesian [9], mean-variance estimation [10], and bootstrap [11] are frequently used to obtain PIs. The main disadvantage of these existing methods is the high computational requirement. Recently, a number of statistical methods have been reported for the interval forecasts of wind power. Without a predefined assumption for the probability distribution of point forecast errors, ensemble simulations [12], quantile regression [13], [14], and kernel density estimation [15]–[17] are currently being used to obtain different quantiles for wind power. However, these PIs construction methods rely on quantile analysis of point forecasts results. Without the prior knowledge of point forecasts, lower upper bound estimation (LUBE) method is proposed in [18]. However, conventional neural networks (NNs) employed in the LUBE method has the problem of overtraining and high computational burden. Alternatively, extreme learning machine (ELM) [19], as a new learning algorithm for single hidden layer feed-forward neural networks, has been applied to construct PIs for wind power and shown to has a faster learning rate and a better generalization capability than conventional NNs [20], [21]. Still much further improvements of optimal PIs construction are needed with respect to, e.g., reliability and sharpness.

Moreover, existing PIs construction methods do not examine the inherent characteristics of the observed wind power time series, which are non-stationary and highly noisy due to the climate conditions such as wind speed, wind direction, wind gust

Manuscript received May 19, 2014; revised September 08, 2014; accepted October 14, 2014. Date of publication October 27, 2014; date of current version July 17, 2015. This work was supported by the National Natural Science Foundation of China (Grant No. 51379081). Paper no. TPWRS-00619-2014.

G. Zhang is with the School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China, and he is visiting the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, WA 6009, Australia (e-mail: gyzhang.hust@gmail.com).

Y. Wu is with the School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: ygangwu@163.com).

K. P. Wong and H. H.-C. Iu are with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, WA 6009, Australia (e-mail: kitpo.wong@uwa.edu.au; herbert.iu@uwa.edu.au).

Z. Xu is with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hong Kong (e-mail: eezhaoxu@polyu.edu.hk).

Z. Y. Dong is with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: zydong@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.2014.2363873

and air density [22]. Forecasting wind power with noisy contents can lead to large errors. As investigated in [22]–[25], the properties of the inherent characteristics in historical time series can offer rich information for point prediction. In these earlier works, the time series are decomposed into a set of independent frequency components, and the next point value is predicted for each component. However, owing to the existence of high frequency components in wind power time series, the point forecasting accuracy for noise component can be very low.

This paper proposes not only to decompose the wind power time series into multiple components but also to construct the *noise PIs* for the high frequency components, i.e., high complexity components. To achieve this, the wind power time series is first decomposed into several independent intrinsic mode functions (IMFs) using ensemble empirical mode decomposition (EEMD) [26]. Then, the *sample entropy* (SampEn) theory [27] is used to analyze the complexity of obtained IMFs, which are then grouped to form the *noise component*, *cycle component*, and *trend component*. For the latter two well-behaved components, the next point value is predicted by the ELM method. For the more unstable noise component, since the point prediction method can hardly satisfy the prediction accuracy, a novel problem formulation with quantum-behaved particle swarm optimization (QPSO) [28] algorithm is used to construct the next optimal noise PIs. Finally, the noise PIs and the predicted points are combined to form the overall optimal PIs.

For decomposing the wind power time series, wavelet decomposition (WD) and empirical mode decomposition (EMD) [29] are widely applied in the point prediction, while WD is sensitive to the selection of thresholds and EMD suffers from the mode mixing phenomenon [30], [31]. This paper proposes to utilize EEMD since it has been demonstrated in [32] and [33] that time frequency spectrum obtained by EEMD reflects more realistic spectrum than by EMD.

The rest of this paper is organized as follows: Section II describes the proposed approach. Section III develops a novel formulation to construct optimal PIs for the noise component with QPSO. Case study results on wind farms and comparisons of the proposed approach with other methods are presented in Sections IV. Finally, Section V concludes the paper.

II. Proposed Approach for Optimal Prediction Intervals Construction

The proposed hybrid approach for PIs construction has three primary stages. They are wind power time series decomposition stage using EEMD, the components formation stage using sample entropy and the overall optimal PIs construction stage using ELM. Fig. 1 illustrates the proposed approach. The EEMD and component formation methods are described in Section II-A and B, the construction of overall PIs is described in Section II-C.

A. Stage 1: Wind Power Time Series Decomposition

The wind power time series are first decomposed using the EEMD method, which is based on the EMD technique. The EMD has been widely used to identify the oscillatory modes and

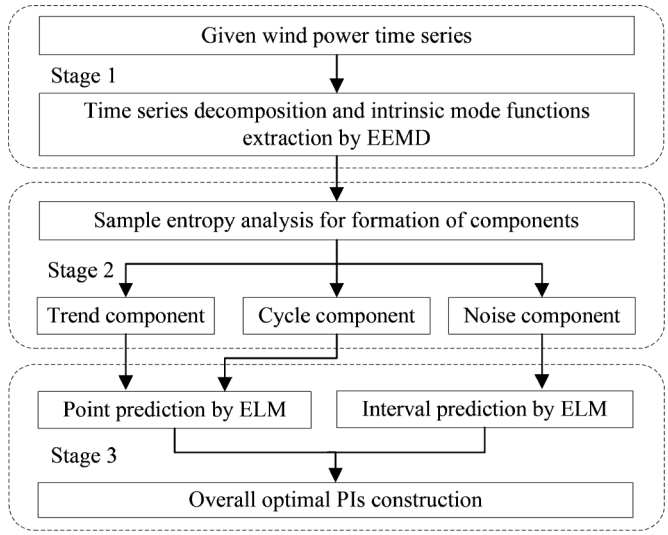


Fig. 1. Diagram of proposed approach for optimal PIs construction.

to decompose time series into IMFs through the sifting process [29]. After sifting processing, the given time series $x(i)$ can finally be expressed as a sum of IMFs and a residual:

$$x(i) = \sum_{q=1}^{Q} c_q(i) + \psi(i) \tag{1}$$

where q denotes the *IMF number*, Q denotes the total number of IMFs, c_q denotes the obtained IMF, and $\psi(i)$ is the residual.

By using the EMD method, the time series can be adaptively decomposed into finite IMFs with progressively decreasing frequencies. However, one of the major drawbacks of EMD is the mode mixing problem, which is defined as a single IMF including oscillations of dramatically disparate scales, or a component of a similar scale residing in different IMFs [31]. As the inheritor of EMD, EEMD extracts the true IMFs by adding finite white noise series to the data and then offset them via ensemble averaging, which can eliminate the mode mixing problem automatically. As investigated in [26], the effect of the added white noise can be controlled by the following rule:

$$\varepsilon_e = \varepsilon/\sqrt{e} \tag{2}$$

where e is the number of ensemble members, ε is the amplitude of the added noise, and ε_e is the standard deviation of error.

B. Stage 2: Formation of Noise Component, Cycle Component and Trend Component

After extracting the IMFs of wind power time series, instead of carrying out prediction for each IMF, in this paper, these IMFs series are then recombined to form the noise component, cycle component and trend component according to their complexities. Approximate entropy (ApEn) [34], as a complexity measure method, has been widely applied to the noise series. However, ApEn is heavily dependent on the record length of the series and it lacks relative consistency. As investigated in [35], sample entropy is a modification of ApEn and it shows good

traits such as data length independence and trouble-free implementation in measuring the complexity of time series. Therefore, the sample entropy theory is used here to analyze the obtained IMFs and to group them respectively into the noise-, cycle-, and trend-component according to their sample entropy values, the calculation of which can be found in [35] and is depicted below.

For a given embedding dimension m, tolerance r and number of data points N, SampEn (m, r, N) is defined as the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, excluding self-matches. For a data sequence $\{x(i)\}$ with N samples, computation of the SampEn value is shown as follows.

1) Form vector sequence of size m, and each of them are defined by

$$X_m(i) = [x(i), x(i+1), \ldots, x(i+m-1)] \quad i = 1, 2, \ldots, N-m+1 \quad (3)$$

where m is the embedded dimension of the vector to be formed.

2) Define the distance between vectors $X_m(i)$ and $X_m(j)$, $d_m[X(i), X(j)]$, as the absolute maximum difference between their scalar components:

$$d_m[X_m(i), X_m(j)] = \max_{k=0\sim m-1} |x(i+k) - x(j+k)|. \quad (4)$$

3) For each $X_m(i)$, define B_i as the number of j with the distance between $X_m(i)$ and $X_m(j)$ is less than or equal to r

$$B_i = \text{num}\{d_m[X(i), X(j)] \le r\}, \quad i \ne j \quad (5)$$

$$B_i^m(r) = \frac{1}{N-m-1} \sum_{i=1}^{N-m} B_i(r) \quad (6)$$

where r is a threshold that serves as a noise filter.

4) By increasing the dimension to $m+1$ and calculate A_i as the number of $X_{m+1}(i)$ within r of $X_{m+1}(j)$, where $1 \le j \le N-m, j \ne i$. Similarly, define A_i as

$$A_i = num\{d_{m+1}[X(i), X(j)] \le r\}, \quad i \ne j \quad (7)$$

$$A_i^m(r) = \frac{1}{N-m-1} \sum_{i=1}^{N-m} A_i(r). \quad (8)$$

5) Define $B^m(r)$ and $A^m(r)$ as

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r) \quad (9)$$

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \quad (10)$$

where $B^m(r)$ and $A^m(r)$ are the probability that two sequences will match for m points and $m+1$ points, respectively. Finally, SampEn for a finite data length of N can be estimated as

$$\text{SampEn}(m, r, N) = -\ln[A^m(r)/B^m(r)]. \quad (11)$$

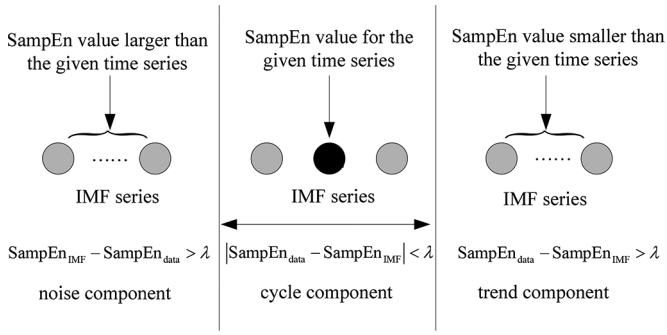


Fig. 2. Grouping rules for the formation of recombination components.

It is clear from the definition that $A^m(r)$ will always smaller than $B^m(r)$, so the value of SampEn will be positive. As investigated in [35], values for m and r are $m = 1$ or $m = 2$ and r between 0.1 and 0.25 times the standard deviation of time series. In the present work, m is set to 2 and the value of r is 0.2 times the standard deviation of $\{x(i)\}$.

According to the theory of EEMD method in [26], with the increment of the IMF number, each IMF contains lower frequency oscillations than the preceding one, indicating a continuous decline in the complexity of IMFs series. This phenomenon lends itself to a grouping rule to group the obtained IMFs into separate components based on the SampEn values. This paper proposes that the IMFs with SampEn values (i.e., $\text{SampEn}_{\text{IMF}}$) close to that of the original series data (i.e., $\text{SampEn}_{\text{data}}$) can be recombined to form component of lower order to the original series. This newly combined component is termed as the cycle component. Elemental IMFs of higher complexities than the original series are grouped and combined to form the noise component. Conversely, the trend component is formed from lower complexity IMFs. Fig. 2 summaries the IMFs grouping rules for the formation of three components with a specified threshold λ value.

C. Stage 3: Prediction for Components and Overall PIs Construction by Extreme Learning Machine

1) Point Prediction by ELM: For the obtained trend component and cycle component, the next point values can be predicted by the ELM method since the variations of these two components are well-behaved and more predictable. As a novel algorithm for training a single hidden- layer feedforward neural network, ELM randomly generates all the input weights and the parameters of hidden layer nodes, and then analytically determines the output weights using simple matrix computation [19]. For N distinct samples $\{(\mathbf{x}_j, \mathbf{t}_j)\}_{j=1}^N$, where the input $\mathbf{x}_j \in \mathbf{R}^n$ and the output $\mathbf{t}_j \in \mathbf{R}^m$, ELM with K hidden nodes and activation function $g(x)$, can approximate these N samples with zero error. This can be modeled as

$$\sum_{i=1}^{K} \beta_i g_i(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, \ldots, N \quad (12)$$

where $\mathbf{w}_i = [w_{i1}, w_{i2}, \ldots, w_{\text{in}}]^T$ is the weight vector connecting the ith hidden neuron and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots \beta_{im}]^T$ is the weight vector connecting the

ith hidden neuron and the output nodes, and b_i is the threshold of the ith hidden neuron.

The above equation can be written compactly as

$$\mathbf{H}\beta = \mathbf{T} \tag{13}$$

$$\mathbf{H} = \begin{bmatrix} g_i(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g_i(\mathbf{w}_K \cdot \mathbf{x}_1 + b_K) \\ \vdots & \cdots & \vdots \\ g_i(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g_i(\mathbf{w}_K \cdot \mathbf{x}_N + b_K) \end{bmatrix}_{N \times K} \tag{14}$$

where $\mathbf{H}$ is the hidden layer output matrix, and $\mathbf{T}$ is the matrix of targets. It has been shown in [19] that ELM with randomly chosen input weights and hidden layer biases can exactly learn N distinct observations, which means that the $\mathbf{H}$ can actually remain unchanged once random values have been assigned. Through this way, the training in ELM is simply equivalent to find a least-squares solution of $\mathbf{H}\beta = \mathbf{T}$, expressed as

$$\hat{\beta} = \mathbf{H}^{\dagger}\mathbf{T} \tag{15}$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of matrix $\mathbf{H}$. Compared to conventional gradient based NN, ELM avoids many difficulties, such as local minima and high computational burdens, and it tends to obtain good generalization performance with a radically increased learning speed.

2) PIs Construction for Noise Component: To construct PIs for the noise component of wind power time series, the ELM-based PIs construction method can be applied, which is illustrated in Fig. 3. It can be seen that the ELM-based interval forecasts model aims to directly generate the upper and lower bounds at the output layer of the neural network. The determination of optimal PIs for the noise component through direct optimization of both reliability and sharpness will be described in Section III.

3) Construction of Overall PIs: After obtaining the point prediction results of the cycle and trend components and the optimal PIs of noise component, the overall optimal PIs of the given wind power time series can be generated in the following way:

$$\begin{cases} P^u = P^u_{\text{noise}} + P_{\text{cycle}} + P_{\text{trend}} \\ P^l = P^l_{\text{noise}} + P_{\text{cycle}} + P_{\text{trend}} \end{cases} \tag{16}$$

where P^u and P^l are the upper and lower bounds of overall PIs, P^u_{noise} and P^l_{noise} are the upper and lower bounds of noise PIs, and P_{cycle} and P_{trend} are the point prediction values of the cycle component and trend component.

III. Construction of Optimal PIs for Noise Component

A. Reliability and Sharpness Evaluation Criteria for PIs

To assess the quality of PIs obtained, the PI coverage probability (PICP) and PI normalized average width (PINAW), which described the reliability and sharpness of PIs are used in [18] and [20]. PICP indicates the probability of the target values within the upper and lower bounds is defined as

$$\text{PICP} = \frac{1}{N_t}\sum_{i=1}^{N_t} \xi_i \tag{17}$$

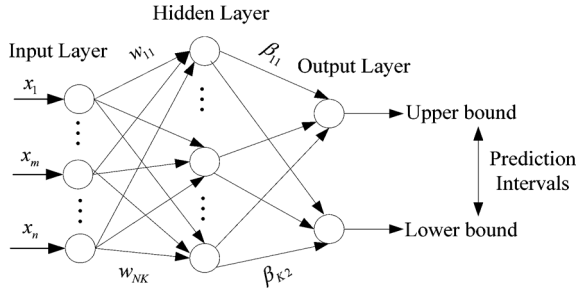


Fig. 3. ELM-based PIs construction method.

where N_t is the total number of samples. If the target value t_i is covered between the upper bound U_i and lower bound L_i, then $\xi_i = 1$; otherwise $\xi_i = 0$.

In the process of interval forecasts, the output is expected to lie within the bounds of PIs with the nominal confidence (PINC) level, which means a large PICP. However, this can be easily achieved by widening PIs from either side. In practice, such PIs are too conservative for decision making. In the literature, a quantitative measure of the average width of PIs is defined as PINAW, which can be defined by

$$\text{PINAW} = \frac{1}{N_t R}\sum_{i=1}^{N_t} (U_i - L_i) \tag{18}$$

where R is the range of the underlying targets, which is used to normalize the PI average width in percentage.

B. Problem Formulation

To ensure the high quality of generated PIs for the noise component, the ELM output weights β are optimized to account for both higher PICP and lower PINAW values. The widely used coverage width-based criterion (CWC) [18] in (19) can be considered for training the ELM model:

$$\text{CWC} = \text{PINAW} \cdot (1 + \gamma_{\text{PICP}} \cdot e^{-\eta(\text{PICP}-\mu)}) \tag{19}$$

where μ denotes the prescribed probability, determined by the nominal confidence level $100(1-\alpha)\%$. The role of η is to magnify any small difference between PICP and μ. γ_{PICP} is a function of PICP. If PICP is less than the assigned μ, $\gamma_{\text{PICP}} = 1$ and the exponential penalty function will be accounted by CWC. Otherwise, $\gamma_{\text{PICP}} = 0$ and CWC will equal to PINAW.

However, by using the CWC function, there is difficulty in deciding the value of η in the exponential penalty term. To avoid this problem, [36] proposed to use the PICP as a hard constraint and the objective is only to minimize the average width of PIs. However, the hard PICP constraint is difficult to be satisfied without sacrificing the width of the PIs, that is the value of PINAW becomes larger than it needs to be.

In this paper, a *deviation information-based criterion* (DIC) is proposed that the exponential penalty term in CWC index is to be replaced by the function *pun* below to form a new objective function in (20) for the construction of noise PIs. The function *pun* takes into account the PIs deviation information for each sample and can be expressed in (21):

$$f(\beta) = \text{PINAW} + \gamma_{\text{PICP}} \cdot \text{pun} \tag{20}$$

$$\text{pun} = \sigma \cdot \left[\sum_{i=1}^{N_L} (L_i - t_i) + \sum_{i=1}^{N_U} (t_i - U_i) \right] \tag{21}$$

where N_L and N_U are the number of samples that the target value t_i is beyond the lower bound and the upper bound, respectively, σ is the penalty coefficient, and in the present work, σ is set to $1/\alpha$. In contrast to η in CWC, the value of *pun* gives a more comprehensive description of the sharpness. For PICP $< \mu$, the amount of penalty is determined by the deviation information of all samples. Otherwise, PICP tends to be equal to PINAW and the optimization concentrate onto narrowing PINAW.

C. QPSO Algorithm for PIs Optimization

The proposed optimal noise PIs construction method in the second part of *Stage 3* aims to achieve the best quality through optimizing the output weights β of ELM model with respect to the objective function in (20). Since the objective function is nonlinear, conventional PSO employed in the PIs optimization has been proved by Bergh [37] to be a local optimization algorithm. As an advanced method, QPSO [28] is a probability searching method and the particles are moved to a quantum space. In quantum time space framework, the state of particles is depicted by wave function, instead of the position and velocity in conventional PSO. Owing to the characteristics of wave function, the particles in QPSO may appear at far distance from current position, which improves the probability to escape from local optimum. Given the search space of QPSO is D -dimensional and the population size of the particles is S, the ith particle x_i is updated according to

$$x_{id} = P_d \pm \delta \cdot |\text{mbest} - x_{id}| \cdot \ln(1/u) \tag{22}$$

$$P_d = (\varphi_{1d} p_{id} + \varphi_{2d} p_{gd})/(\varphi_{1d} + \varphi_{2d}) \tag{23}$$

$$\text{mbest} = \sum_{i=1}^{S} P_i/S = \left[\sum_{i=1}^{S} P_{i1}/S, \sum_{i=1}^{S} P_{i2}/S \cdots \sum_{i=1}^{S} P_{iD}/S \right) \tag{24}$$

where $i = 1, 2, \ldots, S, d = 1, 2, \ldots, D, u$ is random number uniformly distributed in (0, 1), P_d is the local attractor of each particle at d -dimension based on the trajectory analyses in [37], *mbest* is known as the mean best position, p_{id} is the best position of the ith particle, p_{gd} is the position of global best particle, φ_{1d} and φ_{2d} are random vectors, δ is the contraction expansion coefficient, and it is suggested to decrease linearly with the iteration generations from $\delta_{\max}$ to $\delta_{\min}$.

For the implementation of QPSO algorithm in optimizing the ELM-based interval forecasts, the decision variables are the output weight β in ELM. The quality of obtained PIs is reflected by the fitness of each particle, which can be calculated by the objective function in (20). The major steps of the QPSO in PIs optimization are described as follows:

1) *Preparation*: Normalize the training and test data to $[-1, 1]$, and then construct ELM-based interval forecasting model with two outputs. The input weights and hidden biases are randomly generated.
2) *Initialization*: The particles position $[x_{i1}, x_{i2}, \ldots, x_{iD}]$ is randomly initialized and the initial bound for β is $[-1, 1]$.
3) *Construct PIs and evaluate cost function*: With the initialization parameter, construct ELM-based PIs model and calculate the fitness for each particle using (20).
4) *Update position of particles*: The positions of particles are updated according to (22)–(24).
5) *Update* P_i and P_g: With the updated β, construct a new ELM-based model and evaluate the fitness of new particles. If current fitness is better than that of p_i, then p_i is updated. Furthermore, if the particle is better than p_g, then p_g is updated.
6) *Loop*: If the maximum number of iteration has not been reached, go to step 4). Otherwise, the process is terminated and the best output weight is obtained to construct ELM model.

IV. Application Studies

In summary, the proposed three-stage approach above is a combination of ELM, EEMD, DIC, and SampEn techniques. The decomposition of wind power series using WD, EMD, and EEMD methods and the formation of lower order complexity components are described in Section IV-B. The comparison results of various prediction methods are discussed in Section IV-C. To evaluate the performance of the use of CWC and DIC in the problem formulations for noise PIs in Section III, the ELM using DIC (i.e., ELM-DIC method) and CWC (i.e., ELM-CWC method) are first compared in Section IV-C1 and the ELM-DIC method is found to be more superior. For the construction of the noise PIs, the time series decomposition methods in *Stage* 1 of the proposed ELM-DIC-EEMD approach are compared with the ELM-DIC-WD and ELM-DIC-EMD methods in Section IV-C2.

The effectiveness of proposed approach is also demonstrated by comparing it with 1) a recent bootstrap-based probabilistic forecasting method in [11] for the data samples of four seasons in a year, and 2) a neural network-based prediction method in [38] for each month of a year and for three different forecasting horizons. The studies and results are presented in Section IV-C3.

A. Dataset

In the studies, the proposed approach has been tested using the Cathedral Rocks wind farm dataset from [11], the Woolnorth wind farm dataset from [38] and a wind farm data from the National Renewable Energy Laboratory (NREL) [39]. The selected wind farm in NREL is located in Wisconsin State with a combined generating capacity of 20 MW. The whole year of 2006 with hourly observation data is chosen, out of which, the first six months are used for training the ELM method in *Stage 3* and the data between July 2006 and September 2006 are used for testing.

The planning and scheduling in power systems requires the prediction of wind power generation in different forecasting horizons. Short-term (e.g., hourly ahead) and very short-term (e.g., minutes ahead) forecasts are important because of their widely application in load dispatch planning, electricity market clearing, etc. [40]. In the present case studies, hourly ahead PIs are constructed for the Cathedral Rocks wind farm [11] and

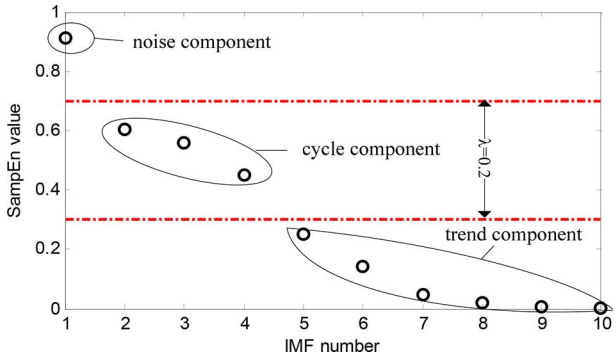


Fig. 4. Distribution of SampEn values of IMFs obtained by EMD.

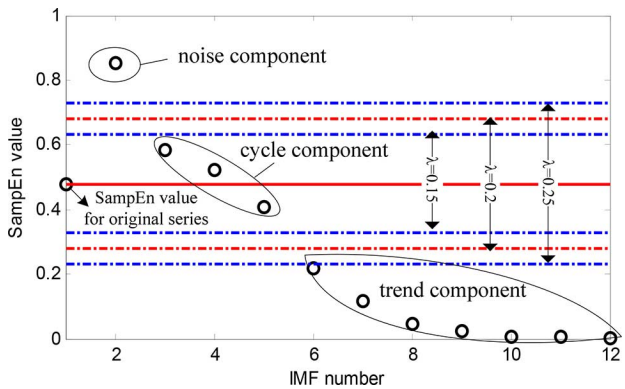


Fig. 5. Distribution of SampEn values of IMFs obtained by EEMD.

the wind farm in NREL, while PIs with 5-, 15-, 30-min forecasting horizons are constructed for the Woolnorth wind farm [38] with 5-min intervals. The detailed processes of the proposed approach are present below using the wind farm dataset in NREL as an example.

B. Decomposition of Wind Power Series and Formation of Lower Order Complexity Components

The given wind power time series are decomposed by EEMD as in *Stage* 1 of the proposed approach, besides they are also decomposed separately by a three-level WD and the EMD methods for comparison purposed. The parameters e and ε in EEMD are set to 20 and 0.15, respectively. Two series of IMFs are obtained by EEMD and EMD. The SampEn values of the two series of IMFs are given in Figs. 4 and 5, respectively.

According to the theory of EEMD, the first IMF in Fig. 5 is the original time series. With the increment of IMF number, i.e., q in Section II-A, it can be seen from Fig. 5 that the SampEn value becomes smaller. This confirms the earlier statements in Section II-B about the relationship between SampEn values and the IMF number. The obtained IMFs are then recombined into the trend-, cycle-, and noise-component according to the grouping rules of IMFs series in Fig. 2 with λ set to 0.2. For comparison purpose, the cases with λ set to 0.15 and 0.25 are also shown in Fig. 5. Here, it can be seen that the small change of the λ value does not affect the grouping since the SampEn for the second IMF series is much larger than the remaining IMFs. The elemental IMFs obtained by EEMD in three recombination components are shown in Table I. The resultant components

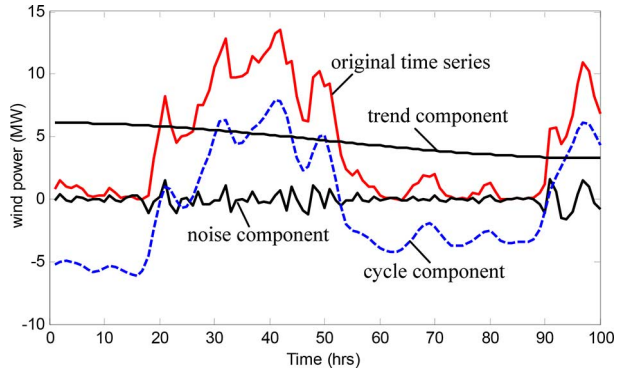


Fig. 6. Resultant noise, cycle, and trend components for wind power time series.

TABLE I
ELEMENTAL IMFS IN NOISE, CYCLE, AND TREND COMPONENTS

Methods	Noise component	Cycle component	Trend component
EMD	IMF 1	IMFs 2-4	IMFs 5-10
EEMD	IMF 2	IMFs 3-5	IMFs 6-12

TABLE II
COMPARISONS OF ELM-DIC, ELM-CWC, AND CONVENTIONAL PREDICTION METHODS

Confidence level	Methods	PICP (%)	PINAW (%)
90%	QR	89.07	37.55
	ARIMA	91.40	60.96
	ELM-CWC	88.84	31.79
	ELM-DIC	90.30	30.25
80%	QR	80.39	29.17
	ARIMA	80.60	40.24
	ELM-CWC	78.78	22.27
	ELM-DIC	81.62	20.17

after combining their elemental IMFs are shown in Fig. 6. It can be seen that the trend component tends to be a straight line and the cycle component can track the changes of the given wind power series with a lower complexity, which facilitates a better performance in point prediction.

C. Study Results and Discussions

1) Comparison of ELM-Based Interval Forecasts Using Criterion DIC and Criterion CWC: For the construction of PIs for the noise component in Section III, the combined ELM/QPSO method is first used to decide whether the CWC in (19) or DIC in (20) should be employed. For the process of QPSO, the population size and the numbers of iterations are set to 100 and 500; the values for $\delta_{\max}$ and $\delta_{\min}$ are 0.9 and 0.5. The activation function in ELM model is sigmoidal and the number of hidden neurons is set to 10. The confidence level at the training phase μ_{train} is set 3%–5% higher than the nominal confidence level as pointed out in [18]. For the given wind power time series, the constructed PIs at confidence levels 90% and 80% are summarized in Table II in terms of the reliability index PICP and the sharpness index PINAW. Also summarized in Table II are the results obtained by conventional methods including autoregressive integrated moving average (ARIMA) and quantile regression (QR). From the results of the PICP and PINAW indices, it can be seen that the combined use of ELM and DIC, that is the

TABLE III
COMPARISONS OF DIFFERENT PREDICTION METHODS FOR THE WIND FARM IN NREL

Confidence level	Methods	RMSE for point prediction			Overall optimal PIs	
		Cycle component	Trend component	Original data	PICP	PINAW
90%	ELM-DIC	-	-	0.196	90.30%	30.25%
	ELM-DIC-WD	-	4.87e-3		90.12%	23.03%
	ELM-DIC-EMD	4.38e-2	1.81e-4		89.37%	17.38%
	Proposed ELM-DIC-EEMD	2.93e-2	1.03e-5		90.89%	14.97%
80%	ELM-DIC	-	-	0.196	81.62%	20.17%
	ELM-DIC-WD	-	5.42e-3		79.48%	17.45%
	ELM-DIC-EMD	4.31e-2	1.73e-4		80.09%	12.94%
	Proposed ELM-DIC-EEMD	2.89e-2	8.87e-6		80.57%	9.78%

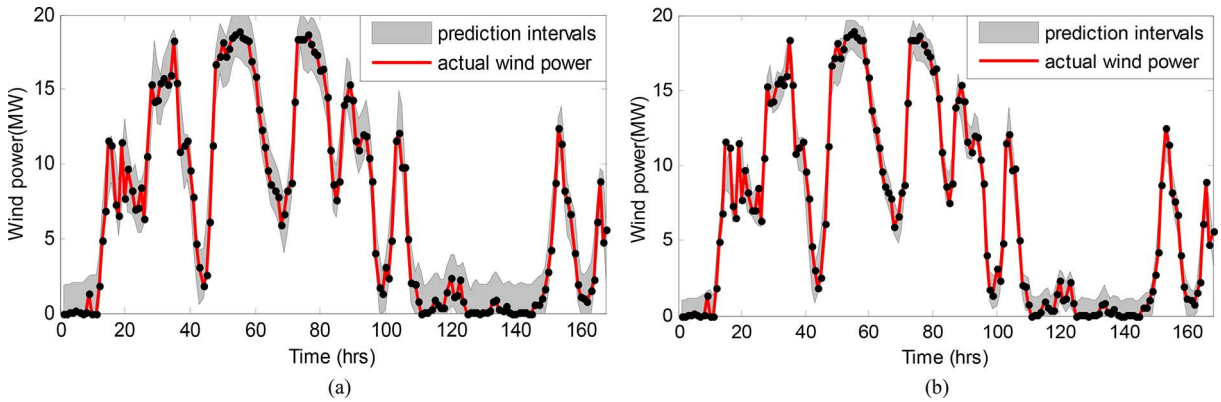


Fig. 7. Overall optimal PIs by proposed ELM-DIC-EEMD approach. (a) 90% confidence level case. (b) 80% confidence level case.

ELM-DIC method, provides much more accurate PIs combined reliability (reflected by higher PICP values) and sharpness (indicated by lower PINAW values) at the two different confidence levels.

2) Comparison of Methods for Determining the Optimal PIs: Based on the superior performance of the ELM-DIC method shown in the last section, to obtain the overall optimal PIs, the ELM-DIC method is employed to generate optimal PIs for the noise component. Combining it with sample entropy analysis and using WD, EMD and EEMD separately, three hybrid series decomposition prediction methods, namely the ELM-DIC-WD, the ELM-DIC-EMD, and the proposed ELM-DIC-EEMD methods are formed. The obtained PICP and PINAW using the above three hybrid methods and the ELM-DIC method, which contains no time series decomposition for wind power series complexity accounting, are given in Table III. It can be seen that at all confidence levels in the case studies, the PICP values of these methods are close to the corresponding confidence levels. However, all the hybrid series decomposition methods have a much higher sharpness than the ELM-DIC method, which does not account for the complexity of the series. This confirms the effectiveness of the decomposition process in *Stage 1* of the proposed approach in Fig. 1.

The comparisons of proposed ELM-DIC-EEMD approach with the other two hybrid methods in Table III show that the proposed method generates results with the highest reliability and sharpness of the PIs. According to the process of proposed approach in Fig. 1, an accurate point prediction is crucial to ensure the reliability of PIs. For the point prediction in the first part of *Stage* 3, the root mean square error (RMSE) index is used to measure the point prediction error of cycle component and trend component. It can be observed from Table III that the proposed hybrid EEMD method has a minimum RMSE value for point prediction against the WD and EMD methods, which also demonstrates the superiority of proposed approach. The overall PIs obtained by the proposed approach at confidence levels 90% and 80% are also shown in Fig. 7(a) and (b), respectively. Considering the bound of generation capacity, the prediction results have been restricted within the generation range of wind farm. It can be found that almost all the target values lie within the lower bound and the upper bound with a narrow average width of PIs.

To validate the proposed approach further, it is compared with the recent bootstrap-based probabilistic forecasting method using the same wind farm dataset in [11]. In this validation, the seasonal difference and diversity are considered. PIs with confidence levels of 90%, 95%, and 99% in the four seasons are constructed, respectively. In [11], the index called *interval score* was employed to evaluate the sharpness of PIs. The reliability and interval score indices of PIs using the above two methods are given in Table IV. It can be observed that the reliability index (reflected by higher PICP values) in the four seasons obtained by the proposed approach compares well with that of [11], and is marginally better in the cases of higher confidence levels of 95% and 99%. For the values of interval score, the proposed approach outperforms the bootstrap-based probabilistic forecasting method with the larger interval score values, indicating a higher sharpness.

3) Comparison of PIs in Different Forecasting Horizons: For the interval forecasts of wind power series in different forecasting horizons, monthly datasets of the Woolnorth wind farm [38] with 5-min intervals in 2010 are used as samples. The op-

TABLE IV
COMPARISONS OF PROPOSED APPROACH AND BOOTSTRAP-BASED PROBABILISTIC FORECASTING METHOD: DIFFERENT SEASONS STUDY

Season	PINC=90%				PINC=95%				PINC=99%			
	Bootstrap-based method [11]		Proposed ELM-DIC-EEMD		Bootstrap-based method [11]		Proposed ELM-DIC-EEMD		Bootstrap-based method [11]		Proposed ELM-DIC-EEMD	
	PICP	Score	PICP	Score	PICP	Score	PICP	Score	PICP	Score	PICP	Score
Summer	92.46%	-7.61%	91.15%	-5.48%	95.09%	-4.59%	96.02%	-3.16%	98.08%	-1.41%	99.23%	-0.87%
Autumn	93.01%	-5.92%	90.83%	-4.54%	95.59%	-3.64%	95.17%	-2.79%	97.67%	-1.19%	99.02%	-0.86%
Winter	91.30%	-6.91%	91.19%	-5.21%	94.41%	-4.12%	95.88%	-3.22%	97.89%	-1.28%	99.34%	-0.81%
Spring	93.19%	-7.10%	91.27%	-5.55%	96.12%	-4.15%	95.74%	-3.09%	98.68%	-1.12%	99.18%	-0.84%

TABLE V
COMPARISONS OF PROPOSED APPROACH AND NEURAL NETWORK-BASED PROBABILISTIC FORECASTING METHOD: DIFFERENT FORECASTING HORIZONS STUDY

Month	5 minutes (one-step ahead)				15 minutes (three-steps ahead)				30 minutes (six-steps ahead)			
	Neural network -based method [38]		Proposed ELM-DIC-EEMD		Neural network -based method [38]		Proposed ELM-DIC-EEMD		Neural network -based method [38]		Proposed ELM-DIC-EEMD	
	PICP	PINAW	PICP	PINAW	PICP	PINAW	PICP	PINAW	PICP	PINAW	PICP	PINAW
Jan.	90.37%	24.44%	92.43%	4.22%	90.61%	89.63%	91.87%	7.06%	90.54%	61.13%	90.80%	7.72%
Feb.	92.70%	30.83%	90.57%	4.47%	90.48%	57.41%	91.17%	7.82%	91.32%	90.54%	91.33%	8.32%
Mar.	93.42%	8.00%	91.05%	5.12%	90.10%	46.71%	91.83%	9.83%	90.97%	55.55%	91.86%	10.94%
Apr.	88.88%	26.46%	90.63%	4.98%	90.10%	27.82%	91.46%	7.54%	84.30%	72.71%	90.17%	8.26%
May	94.88%	27.66%	95.38%	5.10%	94.20%	106.4%	95.08%	7.27%	98.63%	81.93%	94.00%	8.03%
Jun.	92.26%	33.58%	92.17%	5.74%	93.03%	31.43%	91.42%	7.73%	90.03%	101.5%	91.92%	9.13%
Jul.	95.27%	18.70%	91.03%	4.73%	93.38%	42.27%	90.58%	6.45%	95.17%	67.28%	90.58%	8.79%
Aug.	93.41%	16.16%	91.77%	5.24%	92.54%	26.65%	90.71%	10.45%	91.03%	50.80%	90.29%	11.91%
Sep.	91.44%	69.82%	90.25%	8.33%	87.22%	39.65%	90.29%	13.22%	91.13%	57.73%	90.11%	14.58%
Oct.	93.25%	10.23%	94.08%	5.18%	91.17%	53.75%	91.42%	7.39%	89.58%	103.7%	90.25%	7.87%
Nov.	92.56%	21.50%	95.25%	4.89%	92.49%	12.74%	91.50%	6.50%	92.01%	52.09%	92.04%	8.26%
Dec.	91.47%	27.28%	92.92%	5.44%	88.91%	68.17%	90.92%	7.05%	90.66%	93.05%	91.75%	9.77%

timal PIs are constructed for each month of the dataset at a 90% confidence level. As in [38], for each month, 80% of the dataset are used for training of ELM and the remaining 20% for testing. The same study has been repeated for 5-, 15-, and 30-min forecasting horizons, respectively. Table V summarizes the monthly prediction results of the PICP and PINAW indices with three forecasting horizons. As seen in Table V, in all twelve months, the PIs obtained by the proposed approach are not only valid (indicated by higher PICP values), but also more informative (indicated by smaller PINAW values) than the neural network-based method in [38]. It should be noted that a longer forecasting horizon shows more uncertainties and this in return leads to less informative PIs. In the case studies, the results in Table V demonstrate the effectiveness of proposed approach in longer forecasting horizon. The PIs for months of August and September in [38] are unsatisfactory due to abrupt changes in weather conditions and wind speeds. The present proposed approach, however, can decompose the data samples and group both the second IMF series and the third IMF series into the noise component according to sample entropy analysis in Section II-B, which lead to much more accurate PIs. This also demonstrates the robustness and capability of the proposed approach in dealing with wind data sample affected by abrupt weather changes. Therefore, it is reasonable to conclude that the proposed approach is an effective way to improve the quality of PIs for wind power generation with a higher combination quality of reliability and sharpness.

V. CONCLUSION

This paper has proposed an advanced approach for the construction of optimal PIs taking into account the complexity of wind power time series. The approach has been found to be more superior than the existing methods without wind power series decomposition. The proposed EEMD method used in the decomposition has been found better than the WD and EMD methods. The sample entropy technique has been shown to be effective in determining the complexity of the IMFs series. A simple grouping rule to group the IMFs series into lower order complexity components of the wind power series has also been developed. Moreover, the combined use of ELM-DIC and QPSO has been demonstrated to be a more effective method to construct the optimal PIs for the noise component. Study results have illustrated that the proposed hybrid approach can provide the excellent quality of PIs with a high potential for practical applications in systems operation and risk assessment.

ACKNOWLEDGMENT

G. Zhang is grateful to China Scholarship Council for a scholarship enabling him to study at The University of Western Australia with Winthrop Prof. K. P. Wong.

REFERENCES

[1] 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.

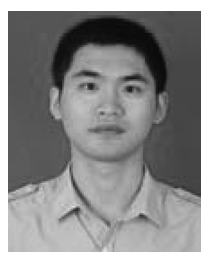
[2] A. M. Foley, P. G. Leahy, A. Marvuqlia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renew. Energy*, vol. 37, no. 1, pp. 1–8, Jan. 2012.

[3] G. Sideratos and N. D. Hatziargyriou, "An advanced statistical method for wind power forecasting," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 258–265, Feb. 2007.

[4] H. Bludszuweit, J. Domínguez-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] M. Lange, "On the uncertainty of wind power predictions—Analysis of the forecast accuracy and statistical distribution of errors," *J. Sol. Energy Eng.*, vol. 127, no. 2, pp. 177–184, May 2005.

[6] Z. S. Zhang, Y. Z. Sun, J. Lin, L. Cheng, and G. J. Li, "Versatile distribution of wind power output for a given forecast value," in *Proc. IEEE PES General Meeting*, San Diego, CA, USA, 2012, pp. 22–26.
[7] 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.
[8] 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.
[9] A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton, and J. van Lint, "Prediction intervals to account for uncertainties in travel time prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 537–547, Jun. 2011.
[10] 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, pp. 55–60.
[11] 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.
[12] 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.
[13] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, Mar. 2004.
[14] 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.
[15] J. Juban, L. Fugon, and G. Kariniotakis, "Probabilistic short-term wind power forecasting based on kernel density estimators," in *Proc. European Wind Energy Conf.*, 2007, pp. 1–8.
[16] 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.
[17] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renew. Energy*, vol. 40, no. 1, pp. 29–39, Apr. 2012.
[18] A. Khosravi, S. Nahavandi, D. Creighton, and A. 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.
[19] 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.
[20] 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.
[21] 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.
[22] K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," *IEEE Trans. Sustain. Energy*, vol. 3, no. 2, pp. 306–315, Apr. 2012.
[23] O. Renaud, J. L. Starck, and F. Murtagh, "Wavelet-based combined signal filtering and prediction," *IEEE Tran. Syst., Man, Cybern. B: Cybern.*, vol. 35, no. 6, pp. 1241–1251, Dec. 2005.
[24] A. O. Boudraa and J. C. Cexus, "EMD-based signal filtering," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 6, pp. 2196–2202, Dec. 2007.
[25] H. Liu, C. Chen, H. Q. Tian, and Y. F. Li, "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks," *Renew. Energy*, vol. 48, pp. 545–556, Dec. 2012.
[26] Z. H. Wu and N. E. Huang, "Ensember empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
[27] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol.-Heat Circulat. Phys.*, vol. 278, no. 6, pp. 2039–2049, Jun. 2000.
[28] J. Sun, B. Feng, and W. B. Xu, "Particle swarm optimization with particles having quantum behavior," in *Proc. IEEE Congr. Evolutionary Computation*, 2004, pp. 325–331.
[29] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proc. R. Soc. London A*, 1998, vol. 454, no. 1971, pp. 903–995.
[30] X. Hu, S. Peng, and W. Hwang, "EMD revisited: A new understanding of the envelope and resolving the mode-mixing problem in AM-FM signals," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1075–1086, Mar. 2012.
[31] B. P. Tang, S. J. Dong, and T. Song, "Method for eliminating mode mixing of empirical mode decomposition based on the revised blind source separation," *Signal Proceess.*, vol. 92, no. 1, pp. 248–258, Jan. 2012.
[32] T. Wang, M. C. Zhang, and Q. H. Yu, "Comparing the applications of EMD and EEMD on time-frequency analysis of seismic signal," *J. Appl. Geophys.*, vol. 83, pp. 29–34, Aug. 2012.
[33] J. Zhang, R. Yan, R. X. Gao, and Z. Feng, "Performance enhancement of ensemble empirical mode decomposition," *Mech. Syst. Signal Process.*, vol. 24, no. 7, pp. 2104–2123, Oct. 2010.
[34] S. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos*, vol. 5, no. 1, pp. 110–117, 1995.
[35] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol. Heart Circulat. Physiol.*, vol. 278, no. 6, pp. 2039–2049, Jun. 2000.
[36] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 303–315, Feb. 2014.
[37] F. van den Bergh and A. P. Engelbrecht, "A study of particle swarm optimization particle trajectories," *Inf. Sci.*, vol. 176, no. 8, pp. 937–971, Apr. 2006.
[38] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation," *IEEE Trans. Sustain. Energy*, vol. 4, no. 3, pp. 602–610, Jul. 2013.
[39] Eastern Wind Integration and Transmission Study, National Renewable Energy Laboratory [Online]. Available: http://www.nrel.gov/electricity/transmission/eastern_wind_dataset.htm
[40] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *Proc. North Amer. Power Symp.*, 2010, pp. 1–8.



Guoyong Zhang received the B.S. degree from Huazhong University of Science and Technology, China, in 2010, where he is pursuing the Ph.D. degree.

Currently, he is a joint Ph.D. student sponsored by China Scholarship Council (CSC) at The University of Western Australia, Australia. His main research interests are grid integration of renewable energies, machine learning, and optimization computation in power systems.

Yonggang Wu received the B.S., M.S., and Ph.D. degrees from Huazhong University of Science and Technology, China, in 1984, 1987, and 1997, respectively.

He is now a Professor at the School of Hydropower and Information Engineering, Huazhong University of Science and Technology. His research interests include renewable energies, automatic generation control, and computational intelligence applications in power system.

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

He is currently a Winthrop Professor with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia. His current research interests include power system analysis, planning and operations.

Prof. Wong received three 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 General Chairman of IEEE/CSEE PowerCon2000 and IEE APSCOM in 2003 and 2009. He was an Editor-in-Chief of *IEE Proceedings in Generation, Transmission and Distribution*. He currently serves as Editor-in-Chief for IEEE PES LETTERS.

Zhao Xu (M'06–SM'12) received the B.Eng. degree from Zhejiang University, China, in 1996, the M.Eng. degree from the National University of Singapore, Singapore, in 2003, and the Ph.D. degree from The University of Queensland, Australia, in 2006.

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

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

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, and computational intelligence and its application in power engineering.

Prof. Dong is an editor of the IEEE TRANSACTIONS ON SMART GRID, IEEE POWER ENGINEERING LETTERS, and *IET Renewable Power Generation*.

Herbert Ho-Ching Iu (M'00–SM'06) received the B.Eng. (Hons.) degree in electrical and electronic engineering from the University of Hong Kong, Hong Kong, in 1997. He received the Ph.D. degree from the Hong Kong Polytechnic University, Hong Kong, in 2000.

In 2002, he joined the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, as a Lecturer. He is currently a Professor at the same school. His research interests include renewable energy, power electronics, nonlinear dynamics, and memristive systems.