A Stochastic Sensitivity-Based Multi-Objective Optimization Method for Short-Term Wind Speed Interval Prediction

Xuanqun Chen [a], Chun Sing Lai [b,c,*], Wing W. Y. Ng [a], Keda Pan [b], Loi Lei Lai [b,*], Cankun Zhong [a,*]

[a] Guangdong Provincial Key Lab of Computational Intelligence and Cyberspace Information, School of Computer Science and Engineering, South China University of Technology, Guangzhou 510630, China

[b] Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou 510006, China

[c] Brunel Interdisciplinary Power Systems Research Centre, Brunel University London, London, UB8 3PH, UK

* Corresponding authors: chunsing.lai@brunel.ac.uk (C. S. Lai); l.l.lai@ieee.org (L. L. Lai); curran.z@qq.com (C. Zhong)

Abstract—With the increasing penetration of wind power in renewable energy systems, it is important to improve the accuracy of wind speed prediction. However, wind power generation has great uncertainties which make high-quality interval prediction a challenge. Existing multi-objective optimization interval prediction methods do not consider the robustness of the model. Thus, trained models for wind speed interval prediction may not be optimal for future predictions. In this paper, the prediction interval coverage probability, the prediction interval average width, and the robustness of the model are used as three objective functions for determining the optimal model of short-term wind speed interval prediction using multi-objective optimization. Furthermore, a new Stochastic Sensitivity for Prediction Intervals (SS_PIs) is proposed in this work to measure the stability and robustness of the model for interval prediction. Using wind farm data from countries on two different continents as case studies, experimental results show that the proposed method yields better prediction intervals in terms of all metrics including prediction interval coverage probability (PICP), prediction interval normalized average width (PINAW) and SS_PIs. For example, at the prediction interval nominal confidence (PINC) of 85%, 90% and 95%, the proposed method has the best performance in all metrics of the USA wind farm dataset.

Index Terms—Wind Speed, Prediction Intervals, Multi-Objective Optimization, Stochastic Sensitivity, Neural Network.

LIST OF ABBREVIATIONS

AI	Artificial intelligence
ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average model
ARMA	Autoregressive moving average model
BP	Back propagation
EEMD	Electromagnetic mode decomposition
EMD	Empirical mode decomposition
GRU	Gated recurrent unit
ICEEMDAN	Improved ceemdan
LSTM	Long short-term memory
LUBE	Lower upper bound estimation
MLP	Multilayer perceptron
MOGA	Multi-objective genetic algorithm
MSE	Mean square error
NSGA-III	Non-dominated sorting genetic algorithm-III
NWP	Numerical weather prediction
PICP	Prediction interval coverage probability
PINAW	Prediction interval normalized average width
PINC	Prediction interval nominal confidence
PIs	Prediction intervals
RNN	Recurrent neural network
SS_PIs	Stochastic sensitivity for prediction intervals
SSMOO	Stochastic sensitivity-based multi-objective optimization
WNN-PICEA-g	Wavelet neural network – preference inspired co-evolutionary algorithm - goal vectors

I. INTRODUCTION

Traditionally, electricity is produced by burning fossil fuels such as coal, natural gas, and oil. However, burning fossil fuels release a large number of toxic substances to the environment, which is not conducive to the sustainable development of mankind [1]. With the development of advanced technology, energy consumers pay increasing attention to the use of renewable energy in recent years. At present, wind energy has the characteristics of large reserves, wide distribution and inexhaustible supply. Wind energy can provide abundant of clean electricity to decarbonize the society [2]. For example in the UK, wind energy is the most common form of renewable energy [3]. Although wind power has many advantages, it has intermittent and random fluctuations. Meteorological conditions affect the operation of wind farms, making wind energy changes in a very short time. This also makes wind energy prediction a challenge and brings great difficulties to the stability and safety of a wind power system [4]. Accurate wind power forecasting reduces the impact of wind energy uncertainty on the power system and helps to formulate corresponding plans to ensure stable operation of the power system [5].

Many wind power forecasting techniques are proposed in the existing research [6], [7]. They are usually divided into physical, statistical and artificial intelligence (AI) methods [8], [9]. Physical methods need to collect meteorological data, including humidity, temperature, pressure, wind speed, wind direction and terrain data, which are usually called numerical weather prediction (NWP) model [10]. The advantages of these methods are high prediction accuracy and strong interpretability. The disadvantage is that it is difficult to collect data and need lots of computation and detailed description of atmosphere [11], [12]. Statistical methods are data-driven by using historical time series data collected from wind farms to predict future value, for example, there are autoregressive moving average model (ARMA) [13] and autoregressive integrated moving average model (ARIMA) [14]. However, these time series models are linear and cannot accurately predict the nonlinear and non-stationary fluctuations of wind speed time series [12]. In recent years, many machine learning technologies have been applied. Among them, artificial neural network (ANN) has become a common method for wind speed prediction because it can capture the nonlinear relationship between historical data [15]. Many of the studies use shallow ANNs, and some use deep learning (DL) to capture complex nonlinear characteristics [16], [17]. In recent years, wind speed data preprocessing is also considered to filter noise, such as empirical mode decomposition (EMD) [18], electromagnetic mode decomposition (EEMD) [19] and improved ceemdan (ICEEMDAN) [20]. The wind speed is periodic and the peak point changes rapidly, so [21] uses sinusoidal activation function to replace sigmoid activation function. Some studies combine ANN with statistical methods to capture both linear and nonlinear characteristics of wind speed [20], [22].These methods improve the accuracy of short-term wind speed prediction. However, there are some shortcomings in traditional point prediction methods, such as large prediction errors and large changes in prediction accuracy, and only a certain prediction value can be generated within a certain time step, without the related uncertainties [23], [24]. Therefore, in recent years, the focus of wind speed prediction research is mainly on interval prediction. Interval prediction can estimate the potential uncertainty and risk level more reasonably and provide a more comprehensive reference for the planning and operation of power systems [25].

Traditional interval prediction methods for wind speed first train the model by minimizing the loss function based on point prediction, and then construct the prediction intervals (PIs) according to the training result [2]. However, the PIs constructed in this way will be accompanied by some problems. Firstly, there is a need to assume the distribution of parameters [26] for the methods such as Gaussian process [27] and kernel density forecast method [28]. However, the actual data distribution often does not satisfy the hypothetical distribution, therefore this could bring unacceptable errors. Secondly, the main strategy of this traditional PIs construction method is to minimize the prediction error rather than improving the quality of the PIs [29]. To overcome these problems, Khosravi proposed a lower upper bound estimation (LUBE) method based on a neural network for PIs in [29]. Compared with other traditional PIs construction methods based on neural network, the LUBE does not need to assume the distribution of prediction errors, and the computation speed is greatly accelerated. More importantly, the LUBE directly optimizes the quality of the PIs. In addition, the existing research also studies PIs from many aspects. The methods for wind speed interval prediction based on a single objective framework (e.g. LUBE) may need to combine multiple objectives into one objective, but whether it is a weighted or exponential combination, it needs some prior knowledge and will introduce too many hyper-parameters. There are also some works regarding it as a constrained single objective optimization problem [30]. Some literatures have proposed multi-objective frameworks for PIs [1], [31], [32]. Reference [33] reported a fuzzy-based cost function, which makes the adjustment of neural network parameters with more freedom and flexibility. In addition to the prediction interval coverage probability (PICP) and the prediction interval normalized average width (PINAW), the cost function can also include coverage dependent and width dependent components [34], [35]. To improve the stability of the optimization model, [36] proposed a method based on optimal aggregation. To avoid an exponential cost function, [37] proposed a deviation information-based criterion. Some works focus on the structure of neural networks, such as RNN [23], LSTM [11] and GRU network [38]. Although these studies improve the quality of PIs from various aspects, as far as we know, there is no work to directly optimize the robustness and stability of PIs.

Most of the existing wind speed interval prediction methods based on multi-objective optimization take the coverage probability and the width of the PIs or the variants of them as two objectives. Although this can directly optimize the two most important indices of the PIs, the model for wind speed interval prediction may lack of stability. That is to say, although the trained model performs well in the training set, the performance may be greatly reduced for unknown samples similar to training samples. In this paper, a stochastic sensitivity-based multi-objective optimization method for short-term wind speed interval prediction is proposed, and the neural network is chosen as the model for interval prediction. Specifically, an optimal solution (i.e., weights of the neural network) is obtained by simultaneously considering the PICP, PINAW, and stochastic sensitivity which is the measurement of the robustness and the stability of neural networks [39]. The latest multi-objective optimization algorithm: Non-dominated Sorting Genetic Algorithm-III (NSGA-III) [40], [41] is used as the tool for finding the optimal solution. In this paper, the actual wind farm data are used to carry out the experiments.

Major contributions of this work are:

1) The traditional stochastic sensitivity for point prediction is extended to the interval prediction. The stochastic sensitivity is introduced to the modeling of short-term wind speed interval prediction. Such that, prediction intervals of the trained model are expected to be accurate and narrow with high robustness. By improving the stability of the model via a minimization of the stochastic sensitivity, the generalization capability of the model for future unknown samples can also be improved.
2) Multi-objective optimization is introduced. Compared with the traditional single-objective-based optimization, the multi-objective optimization does not require the selection of regularization parameter beforehand and the final solution is selected from the Pareto front at the end of optimization. In addition to the relief of hyper-parameter selection, multi-objective optimization usually finds better solutions by treating multiple objectives in a vector form instead of a linear combination of them. This prevents anyone of the objectives from dominating the linear combination of objectives.
3) For the case studies of two countries on different continents,

experimental results show that the proposed method outperforms existing methods for wind speed interval predictions.

The rest of this paper is organized as follows. The terminologies for neural network-based interval prediction and the NSGA-III are introduced in Section II. Section III introduces the proposed method for short-term wind speed interval prediction. Experimental setup and results are discussed in Section IV. Section V concludes this work.

II. Terminologies

II-A. Neural Network Structure for Prediction Intervals

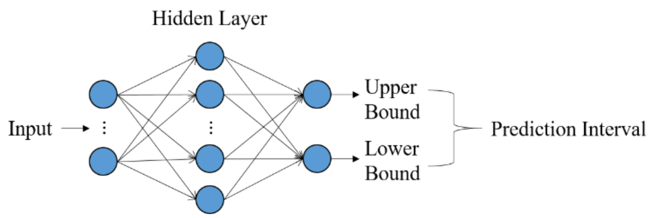

Fig. 1 ANN model developed for PIs.

To directly optimize the quality of PIs by using neural network, Fig 1 presents a neural network structure with two outputs. The number of input neurons and hidden layer neurons can be arbitrary. Among the two outputs of the neural network, the first one corresponds to the upper bound of the PIs and the second one corresponds to the lower bound of the PIs. This kind of neural network structure can make the neural network optimizes the PIs directly.

II-B. Evaluation Indices of PIs

There are two most important indices for evaluating PIs in traditional methods, namely the predicted interval coverage probability (PICP) and the predicted interval normalized average width (PINAW) [42].

In general, PICP is considered as a very important index of PIs, which represents the accuracy of PIs, that is, the probability that the target value is covered by the upper and lower bounds of PIs. A larger PICP means that there are more target values within the built PIs. The definition of the PICP is as follows [42]:

$$\text{PICP} = \frac{1}{N}\sum_{i=1}^{N} c_i \tag{1}$$

where N is the number of samples, c_i is a boolean variable, representing the coverage behavior of the i^{th} PIs. If the target value y_i is between the upper bound U_i and the lower bound L_i of the i^{th} PI, then $c_i = 1$; otherwise, $c_i = 0$. Mathematically, c_i can be defined as:

$$c_i = \begin{cases} 1, & y_i \notin [L(x_i), U(x_i)] \\ 0, & y_i \in [L(x_i), U(x_i)] \end{cases} \tag{2}$$

It is not necessarily better with a higher PICP because this may be accompanied by a wider PIs. Assuming that the width of PIs is infinite, the PICP must be 100%, but this is not the high-quality PIs that we want. So besides considering the PICP, the width of PIs should also be considered. In previous work, the width of PIs is defined as the predicted interval normalized average width (PINAW), which is defined as follows [42]:

$$\text{PINAW} = \frac{1}{NR}\sum_{i=1}^{N}(U_i - L_i) \tag{3}$$

where R is the range of underlying targets (maximum minus minimum). R can normalize the average width of PIs (%) so that it can be compared objectively for different scenarios.

In practice, it is important to have large PICP and narrow PINAW. In theory, these two goals conflict with each other. Reducing the width of PIs usually leads to the decrease of PICP, which is due to the loss of some PIs observations [29]. If the PICP is much smaller than the PINC, the constructed PIs is completely unreliable [29]. Therefore, excellent PIs should have the PICP as close as possible to the PINC with level (1-α)%, where α indicates the probability of error, and PINAW should be as small as possible at the same time.

II-C. Multi-Objective Optimization by NSGA-III

NSGA-III [40], [41] is a multi-objective optimization algorithm based on a genetic algorithm, which is inspired by the principles of genetics and natural selection. It is widely used in various practical optimization problems, and is also one of the latest multi-objective optimization algorithms.

Before the multi-objective optimization, the multi-objective problem must be modeled first. A multi-objective optimization problem can be considered as consisting of M optimization objectives, K equality constraints, J inequality constraints and upper and lower bounds of I decision variables. Mathematically, the problem can be expressed as follows [43]:

$$\text{Minimize/Maximize:} \quad f_m(x), \quad m = 1, 2, ..., M \tag{4}$$

$$\text{subject to:} \quad h_k(x) = 0, \quad k = 1, 2, ..., K \tag{5}$$

$$g_j(x) \geqslant 0, \quad j = 1, 2, ..., J \tag{6}$$

$$x_i^{(l)} \leqslant x_i \leqslant x_i^{(u)}, \quad i = 1, 2, ..., I \tag{7}$$

where, $x = \{x_1, x_2, ..., x_I\}$ is a vector of I-dimensional decision variables in solution space R^I, and Equation (7) restricts the upper and lower bounds of decision variables. Equations (5) and (6) give constraints that decision variables must satisfy, where (5) represent equality constraints and (6) represent inequality constraints. The objective of the optimization is to minimize/maximize M objective functions.

The search process is to optimize x according to the objective functions, and the comparison of solutions is carried out through the concept of dominance [43]. In the minimization problem, for solution x_a and x_b, if the performance of x_a is not lost to x_b for all objective functions, and there exists an objective function x_a outperforms x_b, then the solution x_a dominates x_b. Mathematically, it can be expressed as follows:

$$f_i(x_a) \leq f_i(x_b), \forall i \in \{1, 2, ..., M\} \cap$$
$$f_j(x_a) < f_j(x_b), \exists j \in \{1, 2, ..., M\} \tag{8}$$

If the above conditions are not fully satisfied, x_a does not dominate x_b, or x_b is not dominated by x_a. Ultimately, the goal of NSGA-III algorithm is to determine a set of optimal solutions, that do not dominate each other and are superior to any other solution in search space as compared to all objective functions. This set of optimal solutions is called Pareto optimal set, and the values of the corresponding objective functions constitute the Pareto optimal front.

III. SSMOO FOR WIND SPEED INTERVAL PREDICTION

The proposed Stochastic Sensitivity-based Multi-Objective Optimization (SSMOO) method is a multi-objective optimization problem with three objectives, namely the PICP, PINAW, and proposed stochastic sensitivity for PIs (SS_PIs). We aim to improve stability and robustness of the model to yield better prediction intervals by adding SS_PIs to the opimization. NSGA-III is used to optimize our multi-objective optimization problem. The SSMOO uses the Multilayer Perceptron (MLP) neural network for prediction and can be extended to other neural network models.

The rest of this section will be divided into two parts. The first part introduces the proposed SS_PIs, and the second part introduces the SSMOO. They will be introduced in Sections III-A and III-B, respectively.

III-A. SS for Prediction Intervals

Stochastic sensitivity (SS) is calculated by the average output deviations of the model by a small disturbance to features [39]. If the output of the model is strongly disturbed by small perturbations, then stability and robustness of the model are weak and this usually leads to a weak generalization capability to future unknown samples. The model is more likely to fail to predict future unseen samples.

SS is defined as the average difference of the predicted values of the random perturbed samples to the label which is formulated as follows:

$$SS(x,h) = \frac{\sum_{p=1}^{\beta} |y-h(x_p)|}{\beta} \tag{9}$$

where x, x_p, y, β, and $h(\cdot)$ denote a given training sample, the p^{th} perturbed samples around x, the label of x, the number of perturbed samples, and the predicted value by the model h, respectively. Disturbance samples are created by adding small perturbations to the input of training samples, which are located in the same domain, i.e. Q-neighborhood. The Q-neighborhood of x is defined as follows:

$$S_Q(x) = \{x_p | x_p = x + \Delta x, |\Delta x_i| \leqslant Q, i = 1, 2, \ldots, n\} \tag{10}$$

where Δx, Δx_i, Q, n denote the degree of perturbations to the training sample, the degree of perturbation to the i^{th} dimension of the training sample, the maximum degree of perturbation, and the dimension of sample x, respectively.

For a dataset normalized to [0, 1], $Q = 0.1$ means that the maximum perturbation can deviate from the training sample is by 10%. The samples in the Q-neighborhood of training samples should have the same labels as the training samples, because the model with good generalization ability is robust to small disturbances.

The above-mentioned SS only applies to the traditional point prediction. To apply SS to interval prediction, we need to extend SS to make it naturally applied to the PIs. Thus, the SS for PIs (SS_PIs) is proposed in this paper and defined as follows:

$$\text{SS\_PIs}(x) = \frac{\sum_{p=1}^{\beta} s(x,x_p)}{\beta} \tag{11}$$

where x represents a training sample, the definitions of x_p and β are the same as above, and the definitions of $s(x,x_p)$ are as follows:

$$s(x,x_p) = \begin{cases} 1, & c(x) \neq c(x_p) \\ 0, & c(x) = c(x_p) \end{cases} \tag{12}$$

$$c(x) = \begin{cases} 1, & y \in [L(x), U(x)] \\ 0, & y \notin [L(x), U(x)] \end{cases} \tag{13}$$

where y denotes the label of sample x, $L(x)$ and $U(x)$ denote the lower and upper bounds of the PI of sample x, respectively. The perturbed sample x_p shares the same label y as sample x.When $c(x)$ is equal to $c(x_p)$, $s(x, x_p) = 0$, otherwise $s(x, x_p) = 1$. c is defined in Equation (13). When y falls in the PI, $c(x) = 1$, otherwise $c(x) = 0$.

For a training sample x, SS_PIs generates β perturbed samples x_p ($p \in [1, \beta]$). By comparing the coverage behavior of x and x_p, the stochastic sensitivity of the model in sample x is the number of perturbed samples whose coverage behavior is inconsistent with x divided by the total number of perturbed samples β. SS_PIs measures the stability and robustness of the model in the interval prediction problem. The larger the SS_PIs (x) is, the worse the robustness of the model to small disturbances will be.

SS_PIs is used to evaluate the robustness and stability of wind speed interval prediction model. The detailed flow of SS_PIs for a training sample x is shown in Algorithm 1 below:

Algorithm 1 SS_PIs
Given: A training sample x, label of training sample y, the dimension of training sample n, disturbance degree Q, the number of perturbed samples β, neural network model M
Output: SS_PIs of training sample x

1. SS_PIs(x) = 0
2. For $p = 1$ to β do
   a) An n-dimensional perturbation vector Δx is generated randomly, in which each component is in the range of $[-Q, Q]$.
   b) $x_p = x + \Delta x$
   c) Using x_p as the input of M, the upper bound $U(x_p)$ and lower bound $L(x_p)$ of PIs are obtained.
   d) If $y \in [L(x), U(x)]$, $y \in [L(x_p), U(x_p)]$ or $y \notin [L(x), U(x)]$, $y \notin [L(x_p), U(x_p)]$, $s(x,x_p) = 0$, otherwise $s(x,x_p) = 1$.
   e) SS_PIs(x) = SS_PIs(x) + $s(x,x_p)$
   End for
3. SS_PIs(x) = SS_PIs(x) / β

III-B. Stochastic Sensitivity-based Multi-Objective Optimization Method (SSMOO)

In the previous studies, the PICP and PINAW have been used as two objectives in the multi-objective optimization of interval prediction problems. The optimization equation can be expressed as follows:

$$\text{Objectives:} \quad \text{Finding optimal weights } \omega^*$$

$$\text{Minimize:} \quad \text{PINAW } (\omega) \qquad (14)$$
$$1 - \text{PICP } (\omega)$$

where ω is the weights of the model (i.e., the neural network model in this paper) for wind speed interval prediction. The PICP and PINAW are the two most important metrics for evaluating the quality of PIs. Therefore, this method can construct a PI that performs the best in the training set, but for unknown samples, the performance is not necessarily the best. That is to say, this method may appear to have overfitting phenomenon, and have little generalization ability to unknown samples. Therefore, the Stochastic Sensitivity-based Multi-Objective Optimization Method (SSMOO) is proposed in this paper. It considers not only the PICP and PINAW, but also the SS_PIs as the third optimization objective, aiming for an optimal solution with both PIs quality and generalization ability. The optimization equation can be expressed as follows:

$$\text{Objectives:} \quad \text{Finding optimal weights } \omega^*$$
$$\text{Minimize:} \quad \text{PINAW } (\omega)$$
$$1 - \text{PICP } (\omega) \qquad (15)$$
$$\text{SS\_PIs } (\omega)$$

In addition to the optimization equation, an initial solution, i.e. the initial parameters of the neural network, need to be determined. Because the multi-objective optimization algorithm such as NSGA-III is based on the initial solution for exploratory iteration update, and gradually finds a better solution. Therefore, an excellent initial solution is helpful to find a better final solution. The algorithm for determining the initial solution of the SSMOO is as follows:

Algorithm 2 Determine the initial solution of the SSMOO
Given: Training dataset $\mathcal{D}_{train}$, neural network model M
Output: Initial solution ω_0 of the SSMOO

1. Initialize the parameters of M randomly.
2. Set both target outputs of M to label of D_{train}
3. M is trained by D_{train} with traditional optimization algorithms (such as back propagation). The obtained parameter set is called ω_0.

With the multi-objective optimization equation and initial solution, the optimization objective can be optimized by NSGA-III. It should be noted that NSGA-III does not directly optimize ω_0, but the coefficient of w. The advantage of this method is that the optimization variables of NSGA-III can be controlled in a fixed range, such as [-1, 1], without considering the size of ω_0 itself. At the same time, it can reduce the search space of the solution and facilitate the algorithm to find the optimal solution. As shown in the following equation:

$$\omega = \omega_0 w \qquad (16)$$

where w represents the coefficient of the initial solution ω_0, it is also the optimization variable of NSGA-III.

After NSGA-III optimization, a set of Pareto optimal solutions Ω can be obtained. The last thing the SSMOO needs to do is to select a solution from Ω that best meets the current requirements. SS_PIs is used to improve the robustness and stability of the model in the training phase, but for the testing phase, the PICP and PINAW are directly related to the quality of PIs. Therefore, when choosing the optimal solution from Ω, we mainly consider the PICP and PINAW. Ω is applied to the validation set to obtain the average PICP and PINAW, and then the two values are used to obtain the optimal solution. To avoid introducing parameters, this method does not use the method of combining the PICP and PINAW into a single objective and then selecting the optimal solution from Ω according to the single objective. A more direct and effective approach is adopted, as follows:

$$\underset{\omega \in \Omega}{argmin} \quad \text{PINAW}(\omega)$$
$$\text{s.t.} \quad \text{PICP}(\omega) >= \text{PINC} \qquad (17)$$

PICP (ω) and PINAW (ω) represent the average PICP and PINAW on the validation set with parameter ω. The high-quality PI is based on satisfying the PINC as far as possible, and the interval width reaches the minimum. The final optimal solution is w^*. The best parameters of MLP obtained by the SSMOO method can be obtained by multiplying the w^* with ω_0.

The detailed flow of the SSMOO is shown in Algorithm 3 below:

Algorithm 3 SSMOO
Given: Preprocessed dataset $\mathcal{D}$, neural network model M
Output: Optimal parameters ω^* of neural network model M

1. Dataset D is divided into a training set D_{train}, validation set D_{val} and testing set D_{test}.
2. Obtain the initial solution ω_0 of the SSMOO using Algorithm 2 with D_{train}).
3. Take 1 - PICP, PINAW and SS_PIs (Algorithm 1) as optimization objectives and use NSGA-III to optimize Equation (15), where the decision variable is w ($w \in$ [-1, 1]) and $\omega = \omega_0 w$. The set of solutions obtained is called Pareto optimal set Ω.
4. Use Equation (17) and D_{val} to select the most suitable solution w^* from Ω.
5. The optimal parameter set of M is $\omega^* = \omega_0 w^*$.
6. Using M and ω^* to construct PIs for D_{test}.

IV. EXPERIMENTAL STUDIES

In this experiment, we tested our method on wind speed datasets from two different countries. The first dataset is the wind speed data of wind farm in Colorado, USA in 2004 [44] and the second dataset is the wind speed data of Sotavento wind farm in Galicia, Spain in 2016 [45]. These two datasets are selected to demonstrate that our proposed methods are applicable in different countries, even in different continents. Table I shows details of these two datasets.

TABLE I
Details of Datasets

	Dataset 1	Dataset 2
Location	Colorado, USA	Sotavento, Spain
Coordinates	39.76° N, 105.23° W	43.35° N, 7.88° W
Data Recorded	Wind Speed	
Sampling Rate	10 minutes per sample	
Unit	m/s	
Sampling Period	Jan 1 to Dec 31, 2004	Jan 1 to Dec 31, 2016

Fig. 2 shows the wind speed data of two datasets in January. The wind speed fluctuates from 0 to 30 m / s with the unstable behavior.

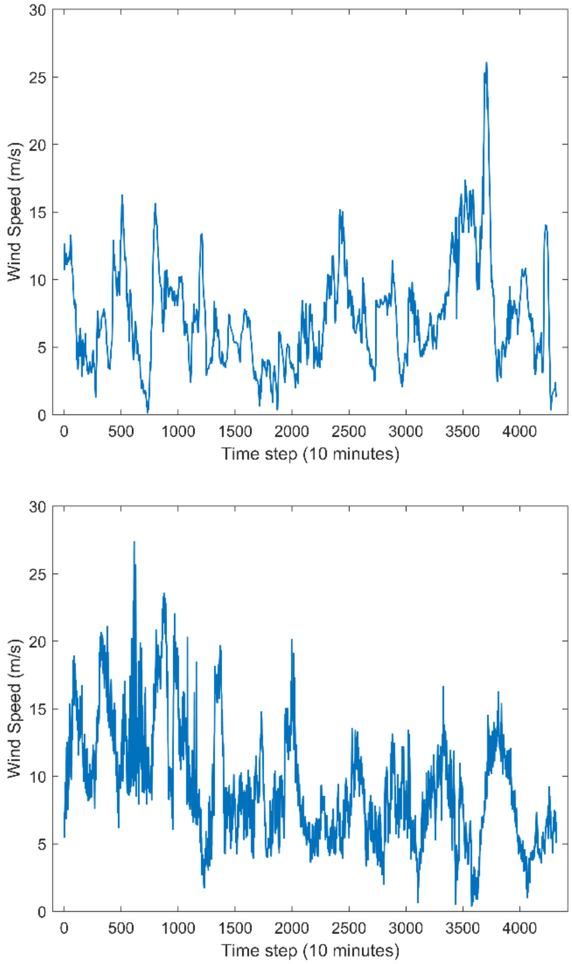

Fig. 2 Wind speed data from a) USA and b) Spain in January.

Data pre-processing: The sampling rate of these two datasets is 1 sample/10 min. The original data contains two columns, one is time (s), the other is current wind speed (m/s). In the actual acquisition process, some data points may be lost. In order to improve the prediction accuracy, these missing data need to be removed. Wind speed prediction is a time series prediction problem. The current wind speed is highly correlated with the recent wind speed. Referring to [32], we use the data of the last one hour (six wind speed values) as input features to predict the current wind speed. To further improve the prediction accuracy, some statistical features (i.e., the mean, variance, median, maximum and minimum values) of the last one hour wind speed values are also utilized. Therefore, each sample has 11 input features. After the data processing, there are N-6 samples. N is the number of data points in the original data. According to the wind speed time series, the proportion of training set, validation set and testing set is 6:2:2. We use the training set to train the model, use the validation set to select the hyper-parameters of the wind speed prediction interval model, and use the testing set to validate the performance of models.

Network structure and pre-training: The proposed method can be easily applied to other network structures. In order to facilitate the experimental comparison, we select the most widely used multi-layer perceptron (MLP). The network structure is shown in Fig.1. In order to improve the prediction effect, we pre-train the model. The process is shown in Algorithm 2. The label corresponding to the sample is set as the ground truth of the two output neurons, and then the mean square error (MSE) loss function of the two output neurons is optimized by back propagation (BP) algorithm. It should be noted that the hyper-parameters of the MLP structure and the number of pre-training iterations are determined by the MSE loss on the validation set. The hyper-parameters of the SS_PIs and the NSGA-III are determined according to references [39] and [41], respectively. After the pre-training, different methods are used to optimize the network parameters.

Experimental hyper-parameters are shown in Table II below:

TABLE II
Parameters for simulation

	Parameter	Value
MLP	Number of neurons in the hidden layer	30
	Number of hidden layers	1
	Activation function	tanh
	Number of neurons in the input layer	11
	Number of neurons in the output layer	2
	Maximum number of pre-training iterations	500
SS_PIs	Disturbance degree Q	0.1
	Number of perturbed samples β	10
NSGA-III	Maximum number of iterations	500
	Population size	100
	Lower bound of variables	-1
	Upper bound of variables	1

To meet the different needs of practical application, five different PINCs were set up in the experiment, which were 80%, 85%, 90%, 95%, and 98% respectively. The criterion for evaluating the quality of different methods is that when the PICP is as close to or even more than the PINC as possible, the interval is as small as possible. Because the simulated annealing algorithm and genetic algorithm have certain randomness, to eliminate the impact of randomness on the experimental results, we repeated the same experiment ten times, and the results were averaged.

The proposed SSMOO is compared with three other methods. The LUBE [29] combines the PICP and PINAW into a single objective and then optimize it using simulated annealing algorithm. The Multi-Objective Genetic Algorithm (MOGA) [32] is a multi-objective optimization method using the PICP and PINAW as two objectives and optimizes them using NSGA-II. NSGA-III algorithm is an improved version of NSGA-II algorithm, so we replace NSGA-II by NSGA-III for the MOGA in our experiments to facilitate comparisons between the MOGA and the proposed method. The Wavelet Neural Network – Preference Inspired Co-Evolutionary Algorithm - Goal vectors (WNN-PICEA-g) [46] is proposed based on a wavelet neural network (WNN) using the PICP and PINAW as two objectives and then PICEA-g is used to optimize the two objectives.

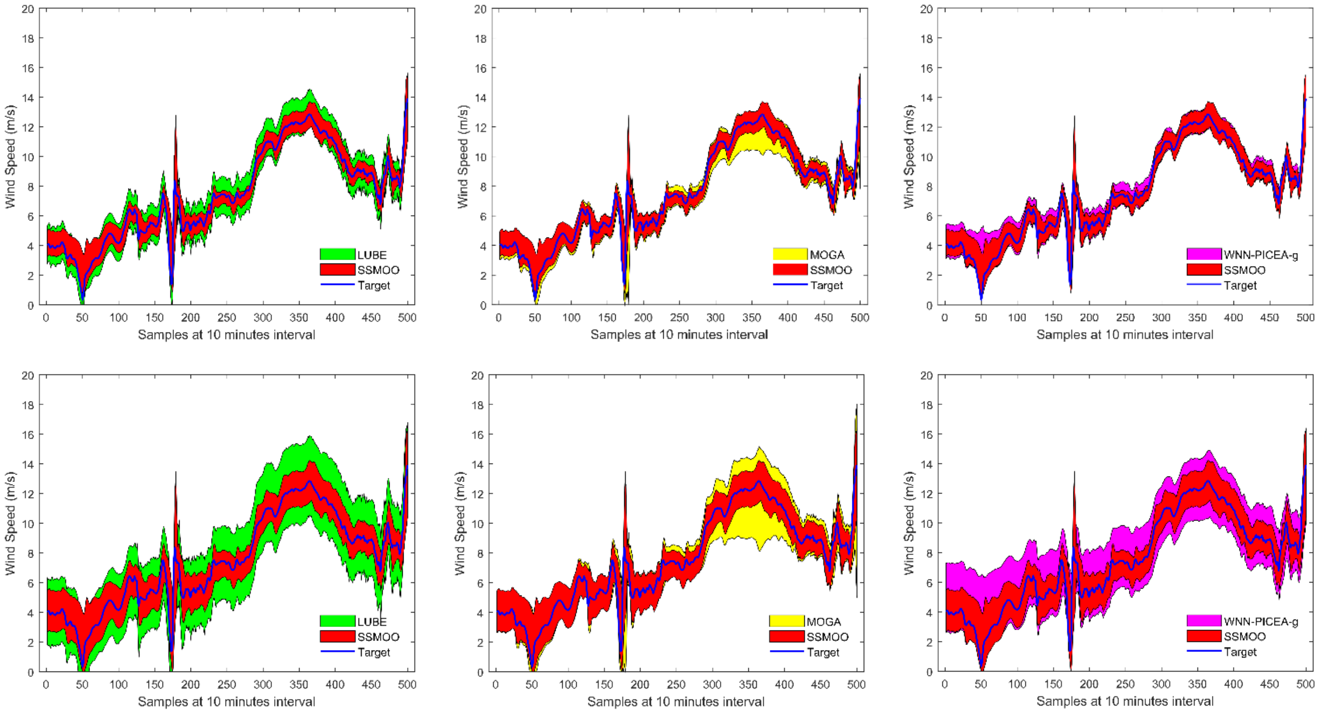

Fig. 3 PIs of the front 500 samples of the USA wind speed dataset (the PINC of the first row and second row are 80% and 95%, respectively).

Fig. 3 shows the difference between the PIs constructed by SSMOO and the three existing methods at two different PINCs on the USA wind speed dataset. The PINCs of the first and second rows are 80% and 95%, respectively. The figure shows only the first 500 samples on the testing set. Obviously, with the increase of PINC, the coverage probability of PIs constructed by the four methods also increases, but the interval width becomes larger. At 80% PINC, the intervals of the four methods are very narrow, but the target often appears near one of the boundaries; and at 95% PINC, the intervals of the four methods are obviously widened, and the target is basically perfectly included in the interval. Comparing the four different methods, we can find that the coverage probability of these methods can reach the corresponding PINCs, but in most cases the SSMOO produces the narrowest intervals. That is, the PIs constructed by SSMOO is usually included in the PIs constructed by other methods. It shows that SSMOO using stochastic sensitivity as one of the optimization objectives can effectively improve the stability and robustness of the model, and then improve the generalization ability of the model to unknown samples. Obviously, the wind speed values cannot be negative in a fixed direction. So we smooth the PIs, that is, the minimum lower bound of the PIs is set to 0.

Fig. 4 shows the solutions of the three methods on the USA wind speed testing dataset. There are 100 solutions to each method in our experiment. The solutions found by the MOGA and WNN-PICEA-g construct their corresponding Pareto fronts. Because SSMOO has three optimization objectives, it is not necessary that the solution found by the SSMOO is the Pareto optimal solution when only the PICP and PINAW are considered. It can be seen that the PINAW of SSMOO is obviously smaller than other methods in the interval of high PICP that we mainly focus on. It shows that the proposed

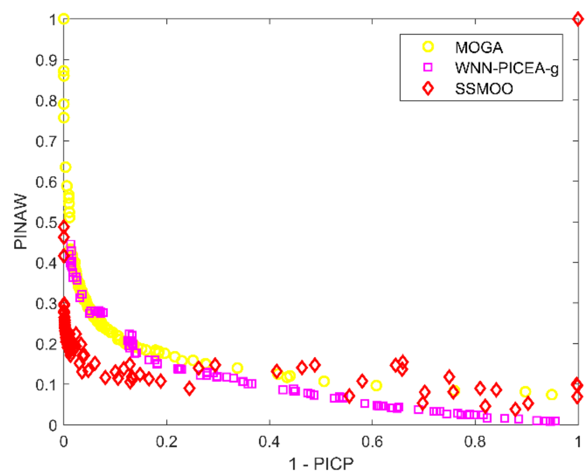

Fig. 4 Testing solutions of the USA wind speed dataset.

method can effectively improve the generalization ability to obtain higher quality PIs.

Experimental results of the four methods on the USA wind speed dataset are shown in Table III. The primary analysis shows that the single target method is the worst. At five different PINCs, the PICPs of the LUBE are not as good as the MOGA, and the PINAWs are larger than the MOGA, which shows that it is effective to transform the problem of short-term wind speed interval prediction into the problem of multi-objective optimization. It is further found that the PICPs of the WNN-PICEA-g are similar to the MOGA, but in most cases, the PINAWs are smaller than the MOGA. Finally, it can be found that the proposed method has the best performance, and the PICPs and PINAWs are optimal in almost all cases. For example, at 95% PINC, the prediction intervals constructed by the proposed method are 31.5% ((24.24-16.6) / 24.24), 23.6%, 38.8% narrower than the WNN-PICEA-g, MOGA, and LUBE, respectively. Therefore, the proposed method can construct PIs with the optimal PICP and PINAW. It shows that the way to improve the generalization ability of the model to unknown samples by using SS_PIs is effective.

Table III also contains the experimental results of the four methods on the USA wind speed dataset with respect to the stochastic sensitivity, i.e. SS_PIs. The stochastic sensitivity represents the stability of the model. In all cases, the SSMOO has the smallest SS_PIs, which shows that taking SS_PIs as one of the optimization objectives can effectively reduce the stochastic sensitivity of the model. The LUBE has the largest SS_PIs, which shows that it is not a stable algorithm and thus it yields lower PICP and higher PINAW at different PINCs. The only difference between the proposed SSMOO and the MOGA is that the MOGA does not take the SS_PIs into account and thus it yields higher SS_PIs than the SSMOO. The SSMOO yields higher PICP and lower PINAW than the MOGA in most cases, which proves that the higher quality of PIs can be obatined by reducing the SS_PIs. The SS_PIs of MOGA is far lower than that of the LUBE, which shows that the model trained with a multi-objective optimization method has better stability. In addition, it can be found that for multi-objective optimization methods, with the increase of the PINC, the SS_PIs decreases gradually. This is because when the PINC increases, the width of the PIs will become wider, so the points around the sample are more easily covered by PIs. But the single objective method LUBE does not satisfy this property, which also shows that LUBE cannot produce stable model.

The experimental results of the four methods on Spanish wind speed dataset are shown in Table IV. Because the PINAW of the LUBE is much larger than that of other methods, and in some cases, even more than twice that of other methods, the effect of the LUBE is still the worst. Further analysis shows that the proposed method is still optimal on this dataset. At five different PINCs, the PICPs of the SSMOO are almost the best. Although the WNN-PICEA-g produces the minimum PINAWs at 80% and 85% PINCs, in fact, the PICPs of the WNN-PICEA-g are far lower than those of the SSMOO. Therefore, in most cases, the SSMOO can produce the highest quality PIs.

Table IV also contains the experimental results of four methods on the Spanish wind speed dataset with respect to the stochastic sensitivity. It can be seen that in most cases, SSMOO still has the smallest SS_PIs. Although the SS_PIs of WNN-PICEA-g is the smallest when the PINC is 95% and 98%, it can be found that the quality of the PIs constructed by WNN-PICEA-g is much worse than that constructed by SSMOO. When PINC is 80%, 85% and 90%, the quality of the PIs and SS_PIs constructed by WNN-PICEA-g are not as good as SSMOO. In all cases, the quality of PIs constructed by SSMOO is better than that of MOGA and LUBE. It shows that SSMOO based on three optimization objectives can reduce the stochastic sensitivity of the model and construct the best quality PIs.The LUBE still has the largest SS_PIs and is much higher than the MOGA. It shows that the multi-objective optimization method is still better than the single objective method.

TABLE III
PICP, PINAW and Stochastic Sensitivity of Four Different Methods on Wind Speed Dataset for USA at Five Different PINCs

Performance Metrics	PICP (%) / PINAW (%) / SS_PIs				
Methods \ PINC	80%	85%	90%	95%	98%
SSMOO	**85.53** / 10.97 / **0.21**	**90.03** / **12.47** / **0.20**	**93.12** / **13.81** / **0.18**	**96.84** / **16.60** / **0.16**	98.91 / **21.00** / **0.13**
WNN-PICEA-g	85.30 / **10.28** / 0.36	88.73 / 12.63 / 0.33	92.58 / 15.48 / 0.31	96.13 / 24.24 / 0.24	97.32 / 29.03 / 0.23
MOGA	84.03 / 12.69 / 0.44	88.83 / 14.60 / 0.42	92.94 / 17.14 / 0.40	96.83 / 21.73 / 0.34	**98.99** / 28.26 / 0.27
LUBE	82.26 / 13.53 / 0.68	88.01 / 26.52 / 0.63	92.23 / 21.69 / 0.67	96.55 / 27.15 / 0.62	98.90 / 39.21 / 0.64

TABLE IV
PICP, PINAW and Stochastic Sensitivity of Four Different Methods on Wind Speed Dataset for Spain at Five Different PINCs

Performance Metrics	PICP (%) / PINAW (%) / SS_PIs				
Methods \ PINC	80%	85%	90%	95%	98%
SSMOO	**83.03** / 10.93 / **0.11**	**87.32** / 12.13 / **0.11**	**89.35** / **12.70** / **0.10**	93.60 / **15.12** / 0.09	**96.49** / **18.71** / 0.07
WNN-PICEA-g	79.58 / **9.21** / 0.16	83.36 / **10.76** / 0.13	87.23 / 12.82 / 0.11	91.40 / 16.42 / **0.08**	93.46 / 22.35 / **0.04**
MOGA	79.29 / 14.80 / 0.29	84.06 / 16.62 / 0.28	88.27 / 18.57 / 0.27	93.25 / 21.81 / 0.24	95.53 / 27.24 / 0.16
LUBE	80.06 / 35.65 / 0.51	84.67 / 26.34 / 0.49	90.49 / 25.05 / 0.46	**93.65** / 30.07 / 0.48	94.33 / 40.70 / 0.53

From the above experimental results, it can be concluded that SSMOO is the best in both the quality and the stochastic sensitivity of PIs. It shows that the performance and stability of the model can be effectively improved by adding stochastic sensitivity as the optimization objective into the optimization equation. Reducing the stochastic sensitivity can enhance the stability and robustness of the model and improve the generalization ability of unknown samples. In addition, the single objective optimization method LUBE has the worst performance among all methods, and manual combination of multiple objectives may introduce too many hyper-parameters. Therefore, the proposed method can effectively improve the quality of PI by constructing the short-term wind speed interval prediction problem as a multi-objective optimization problem and adding stochastic sensitivity as one of the optimization objectives.

In addition to the quality of PIs, we try to compare these four methods from other aspects. Compared with other methods, the LUBE has the advantage that the training speed of the model is faster, while other methods need a long time to train the model because of the use of multi-objective optimization algorithm. The advantage of a multi-objective optimization algorithm is that the final result is a set of Pareto optimal solutions, which can be selected according to actual needs in practical use. For example, after model training, no matter how much the PINC is taken, there is no need to retrain the model, only to select the appropriate solution from the set of solutions, and as long as the PINC is modified, the LUBE needs to retrain the model. Also, the LUBE needs to combine the objectives into a single objective, and needs to specify the combination mode artificially. Therefore, it involves a large number of hyper-parameters, and multi-objective optimization does not need to give each objective weight in advance, nor need to consider how to combine multiple objectives. In practical application, this is a better choice. Considering that the training data of short-term wind speed interval prediction is variable, SSMOO adds additional optimization objective SS_ PIs to improve the robustness and generalization of solutions to deal with such data environment.

V. Conclusion and Future Works

This paper presents a novel multi-objective optimization method, the Stochastic Sensitivity-based Multi-Objective Optimization Method (SSMOO), for short-term wind speed interval prediction. The proposed SSMOO not only reasonably optimizes the prediction interval coverage probability (PICP) and the prediction interval average width (PINAW) without adding too many hyper-parameters as single-objective optimization method, but also takes the proposed stochastic sensitivity of prediction intervals (SS_PIs) into account to improve the generalization ability of the model for unknown samples. To our best knowledge, the proposed SSMOO is the first work to put forward the concept of stability in short-term wind speed interval prediction and give the solution. In addition, this paper improves the traditional stochastic sensitivity, so that it can be well applied to the problem of prediction intervals.

Experiments are carried out on wind speed datasets of two countries on different continents. The results show that at different PINCs, the PICP obtained by the SSMOO can basically exceed the PINC, and the PINAW and stochastic sensitivity are smaller than other methods. It means the quality of short-term wind speed prediction intervals obtained by the SSMOO is highly better than the ones obtained by other benchmarks. The performance of single objective optimization method LUBE is the worst, which shows that it is reasonable to build short-term wind speed interval prediction problem into multi-objective optimization problem. The results show that the multi-objective optimization can be applied to the short-term wind speed interval prediction, and SS_PIs can effectively improve the stability and robustness of the model, and then improve the generalization ability of the model to unknown samples.

In the future work, we consider using dynamic optimization method instead of NSGA-III to solve the dynamic optimization problem of short-term wind speed interval prediction. In addition, we aim to further explore a residential short-term load prediction method based on the proposed SSMOO. It is also worth applying the SSMOO to solve the interval prediction problem of other renewable energy such as solar energy.

Acknowledgement

This work is in part supported by National Natural Science Foundation of China under Grants 61876066 and 61572201, in part by Guangdong Province Scienceand Technology Plan Project (Collaborative Innovation and Platform Environment Construction) 2019A050510006, in part by Brunel University London BRIEF Funding, UK and in part by the Education Department of Guangdong Province: New and Integrated Energy System Theory and Technology Research Group [Project Number 2016KCXTD022].

References

[1] Shrivastava, Nitin Anand, Kunal Lohia, and Bijaya Ketan Panigrahi. "A multiobjective framework for wind speed prediction interval forecasts." *Renewable Energy*, vol. 87, no. 2, pp. 903-910, 2017.

[2] Y. Zhang, G. Pan, Y. Zhao, Q. Li, and F. Wang, “Short-term wind speed interval prediction based on artificial intelligence methods and error probability distribution,” *Energy Conversion and Management*, vol. 224, pp. 113346, 2020.

[3] C. S. Lai and G. Locatelli, “Are energy policies for supporting low-carbon power generation killing energy storage?,” *Journal of Clearner Production*, vol. 280, pp. 124626, 2021.

[4] M. Khodayar, O. Kaynak, and M. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Trans. Ind. Inform*, vol. 13, no. 6, pp. 2770-2779, Dec. 2017.

[5] Y. Zhang, Y. Zhao, G. Pan, and J. Zhang, "Wind speed interval prediction based on Lorenz disturbance distribution" *IEEE Transactions on Sustainable Energy*, vol. 11, no. 2, pp. 807-816, April 2020.

[6] S. Chai, Z. Xu, L. L. Lai and K. P. Wong, “An overview on wind power forecasting methods,” Proceedings of the International Conference on Machine Learning and Cybernetics, IEEE, Guangzhou, China, July 2015.

[7] H.-T. Zhang, K. Chang, H. Zhang, L. L. Lai, “A novel probabilistic approach for evaluating fault ride-through capability of wind generation”, International Conference on Machine Learning and Cybernetics, Jeju, Korea, 9 -12 July 2016.

[8] E. B. Ssekulima, M. B. Anwar, A. Al Hinai, and M. S. El Moursi, “Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: A review,” *IET Renewable Power Gener*., vol. 10, no. 7, pp. 885–989, 2016.

[9] H. Yin, Z. Dong, Y. Chen, J. Ge, L. L. Lai, A. Vaccaro, and A. Meng, “An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization,” *Energy Conversion and Management,* vol. 150, pp. 108–121, 2017.

[10] J. Zhang, C. Draxl, T. Hopson, LD Monache, E. Vanvyve, and B. Hodge, “Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods,” *Applied Energy,* vol. 156, pp. 528-541, 2015.

[11] Z. Zhang, L. Ye, H. Qin, Y. Liu, C. Wang, X. Yu, X. Yin, and J. Li, “Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression,” *Applied energy*, vol. 247, pp. 270-284, 2019.

[12] J. Naik, P. Satapathy, and P. K. Dash, “Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression,” *Applied Soft Computing*, vol.70, pp. 1167-1188, 2018

[13] E. Erdem, and J. Shi, “ARMA based approaches for forecasting the tuple of wind speed and direction,” *Applied Energy*, vol. 88, no.4, pp. 1405-1414, 2011.

[14] K. Yunus, T. Thiringer, and P. Chen, "ARIMA-based frequency-decomposed modeling of wind speed time series," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2546-2556, July 2016.

[15] S. Li, P. Wang, and L. Goel, "Wind power forecasting using neural network ensembles with feature selection," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1447-1456, Oct. 2015.

[16] X. Luo et al., "Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4963-4971, Nov. 2018.
[17] H. Jahangir, M. A. Golkar, F. Alhameli, A. Mazouz, A. Ahmadian, and A. Elkamel, “Short-term wind speed forecasting framework based on stacked denoising auto-encoders with rough ANN," *Sustainable Energy Technologies and Assessments*, vol. 38, pp. 100601, 2020.
[18] N. E. Huang, Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Q. Zheng, et al., "The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis", *Proc. R. Soc. London A*, vol. 454, pp. 903-995, 1998.
[19] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method", *Adv. Adapt. Data Anal*., vol. 01, no. 01, pp. 1-41, Jan. 2009.
[20] Z. Liu., P. Jiang, L. Zhang, and X. Niu, “A combined forecasting model for time series: Application to short-term wind speed forecasting,” *Applied Energy*, vol. 259, pp. 114137, 2020.
[21] G. Ahmadi and M. Teshnehlab, "Designing and Implementation of Stable Sinusoidal Rough-Neural Identifier," *IEEE Transactions on Neural* Networks *and Learning Systems*, vol. 28, no. 8, pp. 1774-1786, Aug. 2017.
[22] X. Niu, and J. Wang, "A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting," *Applied Energy*, vol. 241, pp. 519-539, May. 2019.
[23] Z. Shi, H. Liang, and V. Dinavahi, "Direct interval forecast of uncertain wind power based on recurrent neural networks," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1177-1187, July 2018.
[24] H. M. D. Kabir, A. Khosravi, M. A. Hosen, and S. Nahavandi, "Neural network-based uncertainty quantification: A survey of methodologies and applications, " *IEEE Access*, vol. 6, pp. 36218-36234, 2018.
[25] Hong, Tao, and Shu Fan. "Probabilistic electric load forecasting: A tutorial review." *International Journal of Forecasting*, vol. 32, no.3, pp. 914-938, 2016.
[26] C. Wan, J. Wang, J. Lin, Y. Song, and Z. Y. Dong, "Nonparametric PIs of wind power via linear programming," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1074-1076, Jan. 2018.
[27] Yan, Juan, et al. "Hybrid probabilistic wind power forecasting using temporally local Gaussian process." *IEEE Transactions on Sustainable Energy*, vol. 7, no.1, pp. 87-95, 2017.
[28] Wang, Zhao, et al. "Short-term probabilistic forecasting for regional wind power using distance-weighted kernel density estimation." *IET Renewable Power Generation*, vol. 12, no.15, pp. 1725-1732, 2018.
[29] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, “Lower upper bound estimation method for construction of neural network-based PIs,” *IEEE Trans. Neural Netw*., vol. 22, no. 3, pp. 337–346, Mar. 2011.
[30] H. Quan, D. Srinivasan, and A. Khosravi, “Short-term load and wind power forecasting using neural network-based PIs,” *IEEE Trans. Neural Netw. Learn. Syst*., vol. 25, no. 2, pp. 303–315, Feb. 2014.
[31] Hu, Mengyue, et al. "A Novel Multi-Objective Optimal approach for wind power interval prediction." *Energies*, vol. 10, no.4, pp. 419-, 2017.
[32] R. Ak, V. Vitelli, and E. Zio, “An interval-valued neural network approach for uncertainty quantification in short-term wind speed prediction,” *IEEE Trans. Neural Netw. Learn. Syst*., vol. 26, no. 11, pp. 2787–2800, Nov. 2015.
[33] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, “A new fuzzy-based combined prediction interval for wind power forecasting,” *IEEE Trans. Power Syst*., vol. 31, no. 1, pp. 18–26, Jan. 2016.
[34] 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.
[35] 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.
[36] M. A. Hosen, A. Khosravi, S. Nahavandi, and D. Creighton, ‘‘Improving the quality of prediction intervals through optimal aggregation,’’ *IEEE Trans. Ind. Electron*., vol. 62, no. 7, pp. 4420–4429, Jul. 2015.
[37] G. Zhang, Y. Wu, K. P. Wong, Z. Xu, Z. Y. Dong, and H. H. C. Iu, ‘‘An advanced approach for construction of optimal wind power prediction intervals,’’ IEEE Trans. Power Syst., vol. 30, no. 5, pp. 2706–2715, Sep. 2015.
[38] Li, G. Tang, X. Xue, A. Saeed, and X. Hu, "Short-term wind speed interval prediction based on ensemble GRU model," *IEEE Transactions on Sustainable Energy*. doi: 10.1109/TSTE.2019.2926147, 2019.
[39] D. S. Yeung, J. Li, W. W. Y. Ng, and P. P. K. Chan, "MLPNN training via a multiobjective optimization of training error and stochastic sensitivity," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 5, pp. 978-992, May 2016.
[40] K Deb, H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach part I: Solving problems with box constraints", *IEEE Transactions on Evolutionary Computation*, vol. 18, no.4, pp. 577-601, 2014.
[41] H Jain, K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach part II: Handling constraints and extending to an adaptive approach", *IEEE Transactions on Evolutionary Computation*, vol. 18, no.4, pp. 602-622, 2014.
[42] A. Khosravi, S. Nahavandi, and D. Creighton, “Construction of optimal prediction intervals for load forecasting problems,” *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.
[43] Y. Jin, and B. Sendhoff, "Pareto-based multiobjective machine learning: An overview and case studies," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 3, pp. 397-415, May 2008.
[44] [Online]. Available: http://www.nrel.gov/index.html
[45] [Online]. Available: http://www.sotaventogalicia.com/en/ real-time-data/historical
[46] Z. Shi, H. Liang, and V. Dinavahi, "Wavelet neural network based multiobjective interval prediction for short-term wind speed," *IEEE Access*, vol. 6, pp. 63352-63365, 2018.