IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 10, 2018, accepted October 20, 2018, date of publication October 23, 2018,
date of current version November 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2877735

Wavelet Neural Network Based Multiobjective Interval Prediction for Short-Term Wind Speed

ZHICHAO SHI, (Student Member, IEEE), HAO LIANG, (Member, IEEE), AND VENKATA DINAVAHI, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada

Corresponding author: Zhichao Shi (zhichao1@ualberta.ca)

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). The work of Z. Shi was supported by the China Scholarship Council (CSC).

ABSTRACT As a source of clean and pollution-free renewable energy, wind power has attracted much attention and has been increasingly integrated into the existing power system. However, the uncertain and volatile wind speed makes the utilization of wind power a challenging task. Hence, it is essential to design an accurate forecast method to deal with the uncertainty caused by wind speed. This paper proposes a multiobjective interval prediction method based on wavelet neural network (WNN) for short-term wind speed forecast. This method can generate a set of Pareto optimal solutions which represents a set of prediction models that can directly construct the prediction intervals. An advanced multiobjective evolutionary algorithm, preference inspired co-evolutionary algorithm using goal vectors, is investigated to train the WNN model. Two case studies are carried out with real wind speed data of Victoria and Edmonton in Canada to justify the effectiveness of the proposed method. The numerical results also show the superiority of the proposed forecast approach compared with some benchmark methods.

INDEX TERMS Interval prediction, multiobjective optimization, neural network, wind speed.

I. INTRODUCTION

Owing to the rapid depletion of fossil fuels and rising concern of environment problem, the development of renewable energy like wind and solar has attracted significant attention in the world. With the clean and abundant advantages, wind power has been one of the most popular renewable energy resources. It is reported that approximately 6% of the electricity demand is currently supplied by wind energy in Canada and its total installed wind power reached 11,898 MW in 2016 [1]. Despite the environmental benefits, wind energy, which is chiefly determined by the variable wind speed, has the drawback of intermittency and randomness. Consequently, the high penetration of wind energy into existing power grid may pose many new issues such as the system reliability problem [2]. To alleviate the impact of uncertain wind power and ensure reliable system operation, an efficient method is to conduct accurate wind speed forecast.

Wind speed forecast has been researched for several decades and different forecast approaches have been developed in previous literature. The majority of the research focused on point prediction and the methods could mainly be classified into two groups: physical and statistical [3]. Physical models predict wind speed through solving physical equations numerically, while statistical methods implement forecast by developing statistical models. Recently, the statistical models for wind speed forecast have been widely studied, including the autoregressive integrated moving average model (ARIMA) [4], ARMA model [5], artificial neural network (ANN) [6] and support vector machine (SVM) [7]. In particular, the research on short-term wind speed prediction with ANN model or other statistical models has attracted much attention. In [8], different weather data including temperature, relative humidity and air pressure are considered to predict the wind speed by using the common backpropagation neural network. In [9], SVM model is studied to conduct short-term wind speed forecast and in this forecast method, the dataset is preprocessed by empirical model decomposition (EMD) and the parameters of SVM model are tuned by an improved cuckoo search algorithm. In addition, hybrid methods have also been studied to enhance the prediction accuracy by combining several models [10], [11]. Although the prediction methods mentioned above can achieve satisfactory results to some extent, their main problem is that the error cannot be totally eradicated and the forecast results are risky to the decision maker as the deterministic prediction results may be inaccurate. In addition, the point prediction results indicate no uncertainty level.

To overcome the drawback of point forecast, probabilistic forecast has been increasingly studied recently which can provide additional quantitative information about associated uncertainty [12]. In probabilistic forecast, interval forecast is the most visualized representation and it has been studied for wind speed [13] and wind power [14], [15]. Similarly, the mainstream methods of interval prediction consist of physical and statistic modeling methods. Physical models for wind speed forecast usually simulate the atmosphere using the complex fluid dynamics [16] and conduct interval forecast based on point forecast results [17]. Although physical models have advantages in physics process and long-term forecast, they suffer from the problems of complicate meteorological conditions and heavy computation burden. Compared with physical models, statistical models have attracted more attentions in uncertainty modeling and interval prediction such as quantile regression [18], [19], kernel density estimation [20], resampling (bootstrap) [21], ensemble method [22], lower upper bound estimation [23], and deep neural network (DNN) [24]. The forecast methods mentioned above have different strengths and weaknesses [25]. Quantile regression method can handle heterogeneity problems and is not sensitive to outliers by considering the entire distribution. However, a specific training dataset is necessary to develop the forecast model for this method, and the requirement to model each quantile increases the computational cost. For the kernel density estimation method, it is easy to construct prediction intervals based on point forecast results and a certain presumed error distribution. This is an indirect interval prediction method and the assumed error distribution is typically inaccurate. Resampling method aims to evaluate the robust properties of statistical parameters based on data resampling with replacement. It can overcome some disadvantages of quantile regression, but it is only effective for small samples problem and performs poorly when dealing with large samples. The ensemble method can improve the generalization performance of forecast engines, while the computation burden increases with the consideration of more models and parameters. The DNN model in [24] is used for short-term wind power point forecast. Then, the probabilistic forecast is conducted based on forecast error analysis. The proposed DNN model can accurately capture the dynamic information of historical data, but it is also an indirect probabilistic forecast method and several certain distributions are required for forecast error analysis.

Traditional interval prediction methods usually conduct point prediction and obtain the prediction intervals (PIs) based on certain error distribution assumption like normal distribution. However, such assumptions are usually not true for actual datasets which make the prediction intervals invalid. Recently, the direct interval prediction method, lower upper bound estimation, was proposed which makes no distributional assumption for the original data [23]. This approach directly produces the lower bound and upper bound of PIs based on a certain predictive model such as neural network (NN). Generally, the model parameters are easily adjustable and usually tuned by optimizing a comprehensive cost function which combines different PI evaluation indices. In other words, the PI construction problem is defined as a single-objective optimization problem. Considering the complicated nonlinear and non-differentiable objective function, evolutionary algorithms instead of traditional gradient based methods such as simulated annealing (SA) [23] and particle swarm optimization (PSO) [26] are adopted to solve the problem. However, PI construction is actually a multiobjective optimization problem as high quality PIs need both sufficient reliability and narrow width. Therefore, multiobjective interval prediction should be more advantageous than the single-objective prediction method. This study contributes to the multiobjective interval prediction method based on lower upper bound estimation approach.

Compared with abundant single-objective interval prediction research, there are only a few studies about multiobjective interval prediction for wind speed. Based on a simple multilayer perceptron NN model, the short-term wind speed interval prediction is performed in a multiobjective framework in [27]. The NN model was trained by a multiobjective evolutionary nondominated sorting genetic algorithm II (NSGA-II) [28]. Similarly, radial basis function (RBF) NN model is also investigated for wind speed multiobjective interval prediction [29]. In addition, SVM model is also applied to predict wind speed which is trained by the multiobjective differential evolution algorithm [30]. However, unlike the NN model, two SVMs are used to create the PIs' lower and upper bounds in this study which may increase the computational burden.

According to the discussion above, the research of multiobjective wind speed interval prediction is still not sufficient that improvement may be achieved from both the prediction model and the optimization algorithm. Different NN models have been used in the forecast tasks and good results are obtained [31], [32]. In this study, we perform the interval prediction for short-term wind speed based on NN models in a multiobjective framework. First, the wavelet NN (WNN), a kind of feedforward NN model, is proposed as the prediction model which has not been applied to the interval prediction field. WNNs are developed by combining the wavelet theory and NN models, and they have been studied for wind speed and power forecast recently [33], [34]. However, they are mainly focused on point prediction. In this work, the proposed WNN forecast model combined with the lower upper bound estimation approach is a novel interval forecast model which is different from the existing methods. Then the WNN model parameters are tuned by an advanced multiobjective evolutionary algorithm, i.e., preference-inspired coevolutionary algorithm using goal vectors (PICEA-g) [35], to find the optimal prediction model. In particular, the WNN model parameters include the connection weights of NN model and the parameters of wavelets. The construction of PIs is essentially a multiobjective optimization problem, and the reliability and interval width are considered as two objectives in this work, which are conflicting during the optimization process. With the multiobjective framework, we can optimize the two objectives (i.e., the PI evaluation indices) simultaneously and obtain a set of nondominated solutions or Pareto optimal solutions.

The primary contributions of this study can be summarized as follows:

- The WNN model is proposed for wind speed interval prediction in a multiobjective framework. Although WNN model has been studied before for point forecast tasks, it is the first time to conduct interval prediction based on WNN model in this study, i.e., this is a new interval forecast method for wind speed;
- A novel multiobjective evolutionary algorithm PICEA-g is investigated to train the NN model which considers two objectives. Considering the multiobjective essence of PI construction, the proposed multiobjective problem formulation is a more direct problem formulation compared with the indirect single-objective transformation, and this is more reasonable and practical;
- Case studies are implemented to validate the proposed prediction method based on real-world datasets. More specifically, we compare the proposed model with various single-objective and multiobjective interval prediction models based on the quality of solutions and Pareto front, and comparison results show the efficiency of the proposed approach, that is, the WNN-based multiobjective interval prediction model can achieve better forecast results.

The remainder of this paper is organized as follows. The multiobjective problem formulation for PI construction is described in Section 2. Section 3 illustrates the interval prediction methodology including the WNN model and the optimization algorithm. Numerical results and comparison based on real datasets are provided in Section 4. Finally, Section 5 concludes this work.

II. PROBLEM FORMULATION

Compared with the point forecast aimed at minimizing the forecast error, PI construction is a different prediction problem which pursues high quality PIs. In this section, we first introduce the concept of PIs and some PI evaluation indices which are the basis to assess the PI quality. Then, the problem formulation of multiobjective interval prediction is presented.

A. PI AND EVALUATION INDICES

Conventional point forecast often generates deterministic forecast values which may be highly variable. They suffer from the problem that the forecast error cannot be totally eliminated and the associated uncertainties are not indicated [36]. In comparison, PIs can describe the uncertainty level and are more informative in the decision making process. A PI copes with the uncertainty of forecasting a future value of a random variable and it is different with the confidence interval. PIs clarify more source of uncertainty and they usually cover the corresponding confidence intervals [23].

The performance of forecast methods is usually evaluated by various performance indices. In contrast with the evaluation indices for point forecast methods such as mean squared error (MSE) and mean absolute error (MAE), a different set of indices has been proposed for interval forecast. Generally, a good PI result should be more reliable and narrower. Therefore, the indices about the PI reliability and width, PI coverage probability (PICP) and PI normalized interval width (PINAW) [36], are mostly utilized to assess the performance of interval prediction methods. The formulas of these two indicators are given below:

$$\mathrm{PICP} = \frac{1}{N}\sum_{i=1}^{N}\xi_i \tag{1}$$

$$\mathrm{PINAW} = \frac{1}{R\cdot N}\sum_{i=1}^{N}(u_i - l_i) \tag{2}$$

where N is the number of prediction points, ξ_i is a binary value, R is the range of targets, l_i and u_i are lower and upper bound of a PI, respectively. For ξ_i, it equals 1 if the real value is in the PI, otherwise it is 0.

PICP is a reliability index about PIs and a higher value indicates that more true values are enveloped by the PIs. This index is usually considered to be the critical indicator of PI quality and it should be larger than a certain confidence level to attain high quality interval prediction results in general. Actually, a high PICP value can be easily achieved with sufficiently wide intervals. However, valuable information can hardly be expressed by very wide PIs. Therefore, the interval-width based index PINAW is also necessary to appraise the PIs' quality. With the same PICP level, a smaller interval width implies a better result.

As discussed above, PICP and PINAW only measure one aspect of PIs, respectively. To evaluate the PI quality with regard to both reliability and interval width, a comprehensive index, coverage width-based criterion (CWC), is designed [23]. In addition, an improved CWC index is developed to avoid the multiplication operation [37] which has the following expression:

$$\mathrm{CWC} = \begin{cases} \mathrm{PINAW} + e^{(-\eta(\mathrm{PICP}-\mu))}, & \mathrm{PICP} < \mu \\ \mathrm{PINAW}, & \mathrm{PICP} \geq \mu \end{cases} \tag{3}$$

where μ and η are preassigned hyperparameters to regulate the magnitude of CWC index. Here, μ is usually set to the nominal confidence level, and η is usually a large constant to force PIs to be valid with the exponential function. When PICP is less than the nominal confidence level, CWC will be very large and dominated by the PICP index. Otherwise, CWC is equal to PINAW that needs to be minimized. Thus, CWC is a comprehensive index to evaluate the overall quality of PIs.

In addition to the general indices introduced above, there are some other indices used in the literature, such as the average coverage error (ACE), interval score [14] and the accumulated width deviation (AWD) [25]. In PI construction process, the PI nominal confidence (PINC) is usually predefined, and the PICP index aims to approach PINC as closely as possible. In this case, ACE is defined as the difference between PICP and PINC as follows:

$$\mathrm{ACE} = \mathrm{PICP} - \mathrm{PINC}. \tag{4}$$

ACE can be utilized to assess the quality of PIs with respect to the reliability. The smaller the absolute value of ACE is, the better the quality of derived PIs is. Another index AWD can also be used for reliability evaluation of PIs. By comparing the position of the real targets and PIs, relative width deviation can be calculated, and AWD is the sum of relative width deviation as shown below:

$$\mathrm{AWD}_i = \begin{cases} \dfrac{l_i - v_i}{u_i - l_i}, & v_i < l_i \\ 0, & v_i \in [l_i, u_i] \\ \dfrac{v_i - u_i}{u_i - l_i}, & v_i > u_i \end{cases} \tag{5}$$

$$\mathrm{AWD} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AWD}_i \tag{6}$$

where v_i represents the real target. AWD index penalizes the PIs if the real targets are not enclosed, and a smaller AWD indicates higher PI quality. Note that the two basic indices PICP and PINAW are used as the objectives of the formulated multiobjective problem in this study.

To assess the overall performance of the PIs including the calibration and sharpness, we introduce another comprehensive index interval score. Denote the width of a PI as θ_i which is calculated by $\theta_i = u_i - l_i$, then the interval score S_i of a specific interval is defined as follows:

$$S_i = \begin{cases} -2\alpha\theta_i - 4(l_i - v_i), & v_i < l_i \\ -2\alpha\theta_i, & v_i \in [l_i, u_i] \\ -2\alpha\theta_i - 4(v_i - u_i), & v_i > u_i \end{cases} \tag{7}$$

where α is related to the nominal confidence level ($100(1-\alpha)\%$). Based on the interval score of each forecast point, the overall interval score can be calculated as follows:

$$\mathrm{Score} = \frac{1}{N}\sum_{i=1}^{N} S_i. \tag{8}$$

From the definition, we can find that a lower absolute value of the interval score indicates higher quality of PIs. The Score index can be used to assess the overall skill of PIs since it considers all aspects of PI evaluation [14]. Note that a lot of evaluation indices for PIs have been studied in previous literature and we employ several common indices in this work. Some other indices such as the continuous ranking probability score [38] may also be investigated for future research.

B. MULTIOBJECTIVE PROBLEM FORMULATION

According to the performance indices introduced above, the PI construction is actually an optimization problem which aims at high quality PIs. As CWC is a comprehensive evaluation index, unconstrained single-objective optimization problem based on it was first proposed as follows [23]:

$$\text{Minimize: } \mathrm{CWC}(w) \tag{9}$$

where w is the prediction model parameters to be tuned. Furthermore, taking the coverage probability as the fundamental requirement for valid PIs, constrained single-objective problem formulation was also proposed [26]. In this problem, PICP is constrained to be larger than the supposed confidence level, and the minimization of the PINAW value is the optimization objective.

Although the single-objective problem framework has been widely studied for interval prediction, the PI construction is essentially a multiobjective problem. The problem has two objectives: maximizing the reliability index and minimizing the width index, which are two conflicting objectives, i.e., boosting one objective will deteriorate the other one. Therefore, a multiobjective problem formulation is more appropriate to describe the PI construction problem and the interval prediction for wind speed is conducted in a multiobjective framework in this work. The primary multiobjective interval forecast problem can be expressed as follows:

$$\begin{aligned} \text{Objectives: } & \text{Maximize: } \mathrm{PICP}(w) \\ & \text{Minimize: } \mathrm{PINAW}(w) \\ \text{Constraints: } & 0 \le \mathrm{PICP}(w) \le 100\% \\ & \mathrm{PINAW}(w) > 0. \end{aligned} \tag{10}$$

Note that during the training process, the maximization objective can be easily transformed to the minimization of 1-PICP(w) according to the adopted training algorithm.

As a multiobjective optimization problem can be converted into a single-objective problem with some techniques such as weighted average method, the single-objective problem formulation mentioned above can be considered as such technique. However, the difference is obvious between single-objective problem and multiobjective problem. The former only optimizes one single objective and gets one optimal solution, while the latter optimizes several objectives simultaneously and obtains a set of trade-off solutions which are called Pareto optimal solutions. These solutions form the Pareto front from which the decision maker can select a most satisfactory one. Moreover, with the development of multiobjective evolutionary algorithm, multiobjective optimization problem can be solved efficiently and effectively without being transformed into a single-objective problem.

III. SOLUTION METHODOLOGY

For multiobjective interval prediction problem, a good prediction model and optimization algorithm should be designed to achieve high quality PIs. In this section, the prediction model based on WNN is first proposed, followed by the introduction of PICEA-g optimization algorithm. Then the implementation strategy of interval prediction is presented.

A. WAVELET NEURAL NETWORK BASED PREDICTION MODEL

In direct interval prediction methods based on NN models, the multilayer perceptron (MLP) model has been widely studied. In addition, RBF NN model is also reported for multiobjective interval prediction [29]. However, another feedforward NN model, WNN, has not been studied for interval prediction problem. The first WNN model was proposed to approximate arbitrary nonlinear function as an alternative of classic feedforward NN [39]. In point forecast, it was demonstrated that the WNN model outperforms the other feedforward NN models such as MLP and RBF NN models [32], [34]. Inspired by the good performance of WNN model in point forecast, it is reasonable to explore its performance in interval prediction for wind speed. In addition, considering the essence of interval prediction problem, a multiobjective optimization framework is better suited. Therefore, it is worthy to design a multiobjective prediction model based on the WNN model which is expected to have good prediction performance.

WNNs are developed by combining the wavelet theory and NN models. They belong to feedforward NNs and have been successfully used in some classification and time-series forecast problems [40]. Generally, there are two ways to combine the wavelet theory and NN models in forecasting tasks. One is using wavelet transformation to decompose the time-series data into some sub-series which are then combined with NN models to forecast future values [41]. Another method is to employ the wavelet basis function as the activation function of the hidden neurons to construct WNN model which is also studied in this work. WNNs can be classified into adaptive models where wavelet coefficients are variable and fixed grid WNNs where wavelet coefficients are fixed [42]. Adaptive WNNs have better generalization capability because of the wavelets' local properties and the adaption of wavelet shape corresponding to the training data. Consequently, we propose an adaptive WNN prediction model in this study which is shown in Fig. 1. The proposed adaptive WNN model is a more efficient structure for forecast tasks. Note that the wavelet transformation technique may also be investigated in future research.

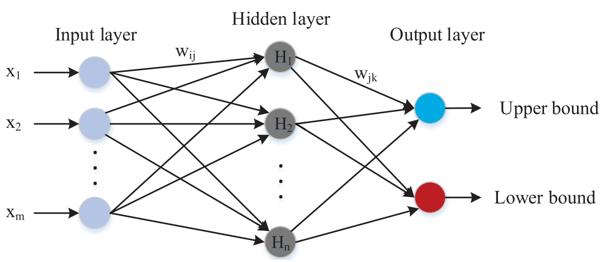

FIGURE 1. Architecture of the proposed WNN model.

As can be seen in Fig. 1, a three layer WNN model is designed where the wavelet transformation is embedded in the hidden neurons of the WNN model [43]. The output layer has two nodes which represent the upper and lower bound of a PI, respectively. According to the universal approximation theorem that a single hidden layer feedforward NN with sigmoid activation function is able to approximate any function, we can get the superposition of sigmoid wavelet [43]. Then the key problem in designing a good WNN model is to find the optimal number of hidden nodes. Among different wavelets, the Mexican hat wavelets are symmetrical and have explicit expression which can provide exact time frequency analysis. In addition, they are based on continuous wavelet transform and can be shifted and scaled smoothly over the entire domain [32]. Hence, in this work, the wavelet activation function used in the hidden units is the Mexican hat wavelet as follows:

$$\psi(x) = (1 - x^2)e^{-0.5x^2}. \tag{11}$$

Thus, the WNN model shown in Fig. 1 can be expressed as below:

$$H_j = \psi_{a_j,b_j}(\sum_{i=1}^{m} w_{ij}x_i), \quad j = 1 \cdots n \tag{12}$$

$$\psi_{a_j,b_j}(z) = \psi(\frac{z - b_j}{a_j}) \tag{13}$$

$$y_k = \sum_{j=1}^{n} w_{jk}H_j + g_k, \quad k = 1, 2 \tag{14}$$

where m and n are the number of input nodes and hidden nodes, respectively, w_{ij} and w_{jk} denote connection weights, a_j and b_j are scale (dilation) and shift (translation) parameters of wavelets, respectively, k is the number of output nodes and g represents the bias. Note that some other wavelet functions may also be used as the activation functions. But their performance needs to be further investigated in future. In this adaptive WNN model, the connection weights and wavelets parameters are all variable that need to be tuned to attain the best forecast performance.

The proposed interval prediction model is derived from the lower and upper bound estimation (LUBE) [23] method which is a direct unsupervised learning process to generate PIs. It can construct PIs simply and fast without making data distribution assumption. Compared with the supervised learning process, the proposed method only use the original data, and the lower and upper bounds are not required in the training process. Particularly, the proposed model directly generates unknown PIs which are gradually improved based on the evaluation indices. For the training set including the input and targets, the input is determined by correlation analysis method which is introduced in detail in Section IV-B. The real data points are used as the targets and the real lower upper bounds are unknown. During the training process, a set of preliminary lower and upper bounds is generated with the NN model as shown in (14) and they are compared with the real targets to calculate the evaluation indices, i.e., the optimization objectives. The PIs are adjusted iteratively based on the quality of objectives. In addition, since wavelets have shown excellent performance in nonlinear function modeling, it is expected that the proposed adaptive WNN model performs well in forecast tasks.

B. PICEA-G ALGORITHM

Various multiobjective evolutionary algorithms have been proposed such as the Pareto dominance based NSGA-II which is one of the most efficient methods by employing an elitist and diversity preservation mechanism. Recently, a new multiobjective evolutionary algorithm, PICEA-g, has been proposed and shown to perform better than other advanced methods including NSGA-II [35]. Therefore, the PICEA-g algorithm is investigated to train the the proposed WNN prediction model in this study.

It is known that preference-based methods are helpful to generate tradeoff surfaces of interest to the decision maker in objective subspaces. With the decision maker's preferences, the incomparable solutions may become comparable. As a result, the concept of co-evolving candidate populations and a set of preferences have been proposed [44] and PICEA-g algorithm is a realization of this approach. In this approach, various preference sets help generate various regions of Pareto front. It is expected to get a good representative of the whole front with many sets of preferences as the co-evolution proceeds.

The general idea of PICEA-g is summarized as follows [35]. In PICEA-g, a set of preferences, also called goal vectors, is co-evolved with the common population of potential solutions during the search process. As for fitness assignment, the potential solutions obtain fitness by satisfying some certain goal vectors in objective space, but the fitness contribution should be shared between all the solutions that meet the goals. The goal vectors' fitness is generated by satisfactory candidate solutions and higher satisfaction implies lower fitness [45]. The aim of goal vectors is to adaptively lead the potential solutions toward the Pareto front, i.e., they co-evolve with the solution population in the process.

The implementation of PICEA-g can be illustrated in an elitist framework as shown in Fig. 2. A population of potential solutions, P, and a set of goal vectors, G, are co-evolved for some certain generations. For every iteration, the genetic variation operation is conducted with the parent solution population P to produce the offspring P_v. While the new goal vectors G_v are randomly regenerated according to the predefined bounds. Then the solution population and the goal vectors are pooled respectively and sorted in terms of the fitness. Lastly, truncation selection is implemented on the sorted population to produce a fixed number of potential solutions and goal vectors as the offspring population. Note that the minimization of 1-PICP and PINAW are considered as two objectives in this work which are used to calculate the fitness during the optimization process. More details about

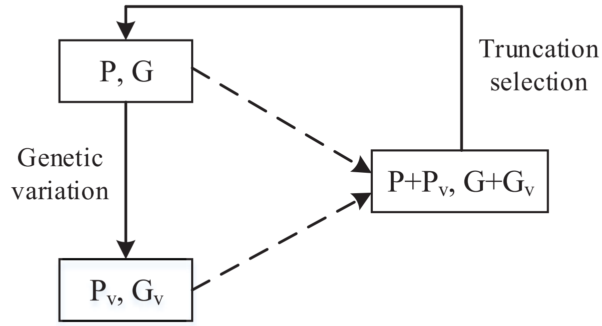

FIGURE 2. PICEA-g implementation framework.

PICEA-g algorithm including the detailed fitness function can be found in [35] and [46].

C. IMPLEMENTATION STRATEGY

Based on the proposed WNN prediction model and PICEA-g training algorithm, multiobjective interval prediction for wind speed can be implemented with real datasets. The main steps of the model implementation are summarized as follows.

Step 1: Data preprocess. Although the wind speed forecast may be influenced by many factors like the weather condition and temperature, the historical wind speed is the most relevant factor which is considered as the input in this study. The original wind speed dataset should be partitioned into training set and test set. In addition, the original data are usually normalized to speed up the model training.

Step 2: Initialize the parameters of the training algorithm. For PICEA-g algorithm, the population size, maximum number of generations and parameters of genetic operators should be specified. The population are coded with real values, i.e., real-coded chromosomes are adopted. Each individual represents one WNN model and consists of all the free parameters as follows:

$$p = [w_{ij}, w_{jk}, a_j, b_j, g_k], \quad i = 1 \cdots m,\ j = 1 \cdots n,\ k = 1, 2. \tag{15}$$

The dilation parameter a_j and translation parameter b_j of wavelet functions are randomly initialized with uniform distribution in the intervals [0.5,2] and [-3,3], respectively [34]. The weights and bias of NN model are initialized randomly in [-1,1] with uniform distribution. The real-coded genetic operators used in this study are simulated binary crossover (SBX) and polynomial mutation (PM) [28]. For SBX operator, it can be defined with the following formulas:

$$p_{idx}^{1,t+1} = 0.5 * [(1 + \beta_{idx})p_{idx}^{1,t} + (1 - \beta_{idx})p_{idx}^{2,t}] \tag{16}$$
$$p_{idx}^{2,t+1} = 0.5 * [(1 - \beta_{idx})p_{idx}^{1,t} + (1 + \beta_{idx})p_{idx}^{2,t}] \tag{17}$$

where $p_{idx}^{1,t}$ and $p_{idx}^{2,t}$ are two parent variables in generation t, $p_{idx}^{1,t+1}$ and $p_{idx}^{2,t+1}$ are two offspring variables in generation $t+1$, and the parameter β_{idx} is calculated as follows:

$$\beta_{idx} = \begin{cases} (2r)^{\frac{1}{\eta_c+1}}, & r \leq 0.5 \\ (\frac{1}{2(1-r)})^{\frac{1}{\eta_c+1}}, & r > 0.5 \end{cases} \quad (18)$$

where r is a random number in the interval [0,1] and η_c is the distribution index defined by the decision maker. The SBX operator intends to generate offspring near the parents which is helpful to inherit the valuable information. For PM operator, it can be expressed as follows:

$$p'_{idx} = \begin{cases} p_{idx} + \delta_{idx}(p_{idx} - p_{idx}^{low}), & r \leq 0.5 \\ p_{idx} + \delta_{idx}(p_{idx}^{up} - p_{idx}), & r > 0.5 \end{cases} \quad (19)$$

where p_{idx}^{low} and p_{idx}^{up} are the lower bound and upper bound of the decision variable, respectively, r is still the random number and δ_{idx} is a parameter as follows:

$$\delta_{idx} = \begin{cases} (2r)^{\frac{1}{\eta_m+1}} - 1, & r \leq 0.5 \\ 1 - (2(1-r))^{\frac{1}{\eta_m+1}}, & r > 0.5 \end{cases} \quad (20)$$

where η_m is the user-defined index parameter.

Step 3: Determine the optimal WNN structure. The parameters of WNN prediction model mainly includes the number of input nodes and hidden nodes. The input features can be determined by correlation analysis method. Specifically, the correlation analysis is implemented with the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) in this work, which will be introduced explicitly in next section. Considering the sequence of time series data, the number of hidden nodes is determined by trial and error method which has a similar idea with cross validation method [26]. The hypervolume indicator [47] is employed to assess the model performance and determine the optimal model structure. A higher hypervolume value stands for a better model.

Step 4: Model training and evaluation. After determining the parameters and optimal structure of the prediction model, the model was retrained with the training data. The termination condition is to reach the predefined maximum iteration in this study. When the training terminates, the Pareto front is attained for the test dataset which consists of the PICP and PINAW values of each individual. The hypervolume can also be calculated to evaluate the model.

Step 5: PI construction. From the multiobjective optimization method, a set of Pareto optimal solutions can be obtained. To construct high quality and satisfactory PIs, the decision makers may select the best solution according to their preferences such as the reliability requirement. This flexible selection is obviously an advantage of multiobjective interval prediction over the single-objective interval prediction method.

In wind speed forecast, the associated uncertainty can be represented by different probabilistic approaches including probability density function, moments of distribution, quantiles and intervals [21]. The most commonly used probabilistic forecast method is based on quantiles. However, we can only get one quantile in one simulation with such method and construct the PIs indirectly with a pair of these quantiles. In contrast, the proposed interval forecast method can produce a set of optimal solutions simultaneously, and the PIs are constructed directly without the estimation of quantiles. Therefore, from the decision maker's viewpoint, the proposed multiobjective interval prediction method is more efficient and concise.

IV. NUMERICAL RESULTS

To verify the effectiveness of the proposed multiobjective interval prediction model, case studies with real-world wind speed data are executed in this section. First, the datasets used as well as the parameter settings of the prediction model are depicted. Then the prediction results and comparison with other models are demonstrated. All the experiments in this work are conducted with MATLAB 2018a on a desktop computer with Intel Core TM i7-6700 CPU 3.40 GHz and 8 GB of RAM.

A. DATASETS

The wind speed data used in this study are hourly mean wind speed taken from two locations: Victoria and Edmonton in Canada [48]. The time periods of two datasets are both from 1 August 2016 to 31 July 2017 with the hour unit. However, in this time period, the Victoria dataset has 5 missing values. The missing values cannot be deleted directly to keep the wind speed distribution. As the overall data trend will not dramatically change in a very short time, the mean value of the data before and after the missing data point is used to replace the missing one in this study.

Victoria is located on Vancouver Island while Edmonton is an inland city, thus the wind speed data from these two locations are expected to have different characteristics. The descriptive statistics of the two chosen datasets are summarized in Table 1. In this study, 80% of the one year data (from August 2016 to May 2017) are used to train the prediction model, the remaining are utilized to test the model. In addition, the training set and the testing set are normalized to [-1,1], respectively. As the forecast accuracy decreases with the increase of forecast time scale, one step ahead interval forecast is conducted in this work.

TABLE 1. Descriptive statistics of the two datasets.

Location	Mean	Std.	Min.	Max.
Victoria	11.03	7.08	0	55
Edmonton	10.92	6.24	0	48

B. PARAMETER SETTINGS

Two sets of parameters need to be determined in the proposed prediction model. One is about the PICEA-g algorithm, the other one is about the WNN model. The parameters of PICEA-g algorithm used in this study are collected from the reference [35] as shown in Table 2. *Npop* is the population size of candidate solutions. The number of goal vectors is equal to the population size. *MaxGen* represents the maximum number of iterations which controls the termination of the model training. For SBX crossover operator, the recombination probability p_c is 1 and the distribution index η_c is set to 15. For PM mutation, the mutation probability p_m is related to the number of decision variables $nvar$ and the distribution index η_m is equal to 20 in this study.

TABLE 2. Parameters of PICEA-g algorithm.

Parameter	Value
Npop	40
MaxGen	1000
Crossover operator	SBX (p_c=1, η_c=15)
Mutation operator	PM ($p_m = 1/nvar$, η_m=20)

In time series forecast, correlation analysis is usually employed to identify the order of the model. In particular, the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) are often utilized to conduct correlation analysis between the forecast value and past historical data [26]. Therefore, we adopt the ACF and PACF analysis to determine the input of the WNN model, i.e., determine the input values that have maximum correlation to the forecast values. The ACF and PACF analysis method is widely used in forecast tasks such as [26] and [29]. Since the intermittent and volatile wind speed fluctuates every now and then, it shows no apparent daily and weekly trend and we can assume that it is stationary. Then the ACF and PACF analysis can be used directly without difference operation. For the Victoria dataset, the ACF and PACF are shown in Fig. 3. As can be seen from this figure, the ACF has an exponential decaying trend and the PACF is cut off at lag 3. Thus, the proper order of this time series should be 3. Considering x_t as the time series variable, the vector (x_{t-2}, x_{t-1}, x_t) is then used as the input to forecast the value x_{t+1} at next step. Likewise, the correlation analysis with ACF and PACF can also be implemented for Edmonton wind speed data. The proper time series order is also 3 and the similar analysis graph is omitted here.

The number of hidden nodes usually has an important effect on the NN model performance. To determine the optimal number of hidden nodes, two methods including cross-validation and trial and error method are often studied in the previous literature. In this study, the method proposed in [26] is adopted and combined with the hypervolume indicator to investigate the optimal number of hidden neurons which has a similar idea with cross-validation. For each selection of NN models, it is trained and validated for five times with the training and testing datasets. The hypervolume indicator is calculated for every simulation run and the average hypervolume value for each model was obtained. The hypervolume represents the proportion of the objective space calculated based

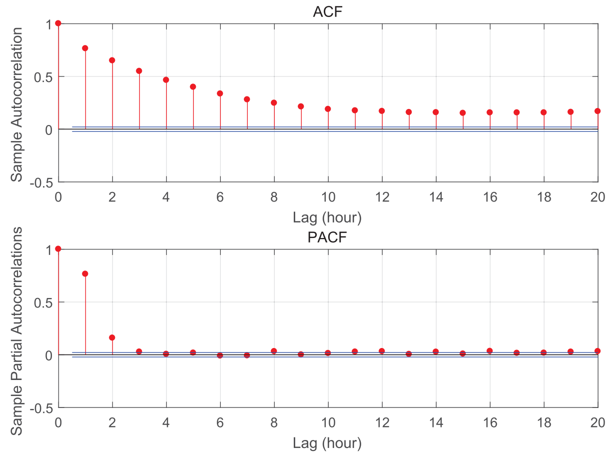

FIGURE 3. ACF and PACF analysis of Victoria wind speed.

on the obtained approximating Pareto front and a certain reference point [45]. The reference point is set to (1.2, 1.2) for the minimization problem with two objectives (1-PICP, PINAW) in this study. The hypervolume is calculated by the method developed in [49]. For the minimization problem, the model with maximum average hypervolume value is chosen to be the best model. Considering the balance of computation complexity and generalization capability, the number of hidden nodes is limited to change from 3 to 10 in this paper. The average hypervolume results for Victoria dataset are given in Fig. 4. As can be seen from this figure, the model with 8 hidden neurons has the best performance. Therefore, the optimal structure of WNN prediction model for Victoria data is 3-8-2. Similarly, the optimal number of hidden nodes of the prediction model for Edmonton dataset was determined to be 7.

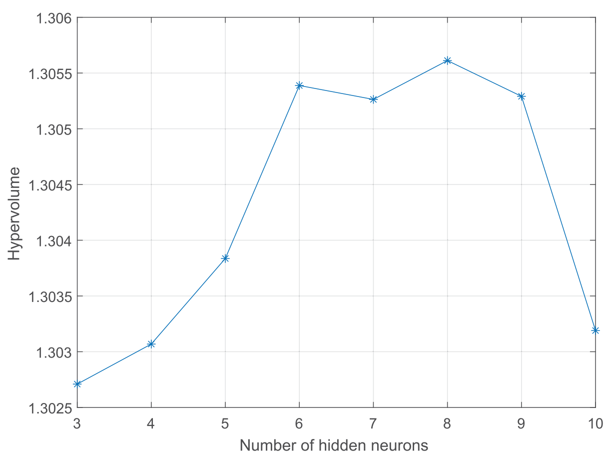

FIGURE 4. Average hypervolume results.

In single-objective interval prediction, CWC is used as the comprehensive index to evaluate the PIs' quality. In order to compare multiobjective interval prediction with single-objective interval prediction methods conveniently, CWC index is also investigated in this work. For CWC parameters, the parameter μ is specified as the nominal confidence level $1 - \alpha = 0.9$, the large constant η is set to 50 [23]. These parameters may also be variable according to the decision maker.

C. PREDICTION RESULTS

After specifying the parameters and determining the optimal model structure, multiobjective interval prediction for wind speed can be implemented. The model is first trained with the training data. After the training termination is reached, a set of Pareto optimal solutions, i.e., a set of non-dominated optimal prediction models can be obtained. Applying these models to do interval prediction with test data leads to the required Pareto front of test set.

The test results or the Pareto front with the proposed multi-objective interval prediction model for Victoria and Edmonton data are presented in Fig. 5. From this figure, we can see that good prediction results can be obtained for both datasets. Both the Pareto fronts show good convergence and diversity and have reasonable and valid objective values. Each point in the Pareto front indicates the result of a prediction model. Actually, the decision maker can choose a satisfactory prediction model among these Pareto solutions of training sets to construct PIs according to his posterior preference such as the interval prediction reliability requirement.

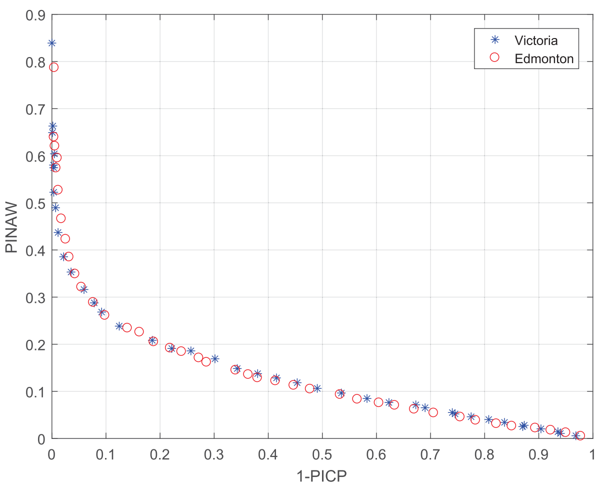

FIGURE 5. Pareto front of prediction results.

To generate high quality PIs, we consider the comprehensive index CWC and take the model with the smallest CWC value as the most appropriate model. Since the nominal confidence level is set to 0.9 in this work, choosing the model with smallest CWC value is equivalent to choose the one with smallest PINAW value among those with PICP not less than 0.9. The interval prediction results from the model with the smallest CWC value for Victoria dataset are shown in Fig. 6. Note that only the prediction results of the last week of the test dataset are shown in this figure for better visualization. In addition, as the real wind speed is impossible to be below zero, the lower bound limitation is set to zero in this study [26]. As can be seen from this figure, the PIs generated from the model are valid and narrow with PICP=90.81% and PINAW=26.83%. Both the upper bound and the lower bound vary similarly with the actual data.

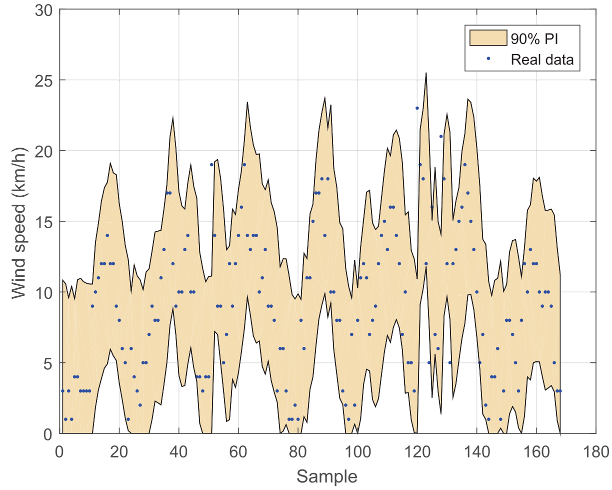

FIGURE 6. Interval prediction results for Victoria data.

Similarly, the PI construction results for Edmonton dataset can be attained as shown in Fig. 7. From this figure, it can be seen that the wind speed of Edmonton has a different fluctuation trend, but high quality PIs can still be generated by the proposed WNN prediction model. The constructed PIs are able to enclose the real targets well. For this case, PICP is 90.18% and PINAW equals 26.11%, which indicate narrow PIs on the condition that the reliability is guaranteed.

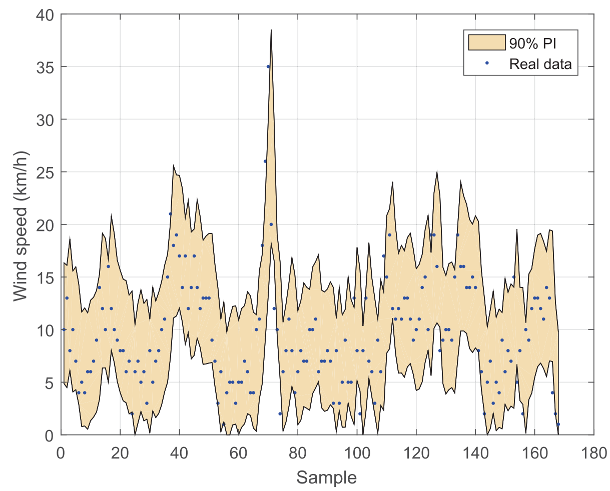

FIGURE 7. Interval prediction results for Edmonton data.

D. COMPARISON WITH OTHER MODELS

In order to substantiate the effectiveness of the proposed interval forecast approach, several benchmark models are employed to conduct interval prediction with the same

TABLE 3. Comparison results for Victoria data.

PINC	Method	PICP (%)	PINAW (%)	ACE (%)	AWD	Score	CWC
90%	WNN-PICEA-g	90.81	26.83	0.81	0.0196	-3.3890	0.2683
	NN-LUBE-PSO	96.17	36.12	6.17	0.0058	-3.6507	0.3612
	QR	85.72	29.98	-4.28	0.0535	-3.9617	8.7889
	Naive	85.80	26.39	-4.20	0.0376	-3.6267	8.4143
85%	WNN-PICEA-g	85.95	23.31	0.95	0.0310	-4.4562	0.2331
	NN-LUBE-PSO	91.26	30.54	6.26	0.0144	-5.0725	0.3054
	QR	80.24	24.86	-4.76	0.0834	-5.1405	11.0542
	Naive	81.81	23.33	-3.19	0.0548	-4.8260	5.1543
80%	WNN-PICEA-g	81.27	20.77	1.27	0.0450	-5.5404	0.2077
	NN-LUBE-PSO	87.09	26.96	7.09	0.0314	-6.3132	0.2696
	QR	74.01	21.05	-5.99	0.1202	-6.1379	20.1475
	Naive	77.71	20.94	-2.29	0.0744	-5.8641	3.3548
75%	WNN-PICEA-g	75.79	18.11	0.79	0.0679	-6.4028	0.1811
	NN-LUBE-PSO	82.92	21.65	7.92	0.0475	-6.6901	0.2165
	QR	69.50	18.46	-5.50	0.1575	-7.0067	15.8027
	Naive	73.32	18.92	-1.68	0.0970	-6.7774	2.5077
70%	WNN-PICEA-g	71.56	16.39	1.56	0.1020	-7.3752	0.1639
	NN-LUBE-PSO	77.84	20.10	7.84	0.0930	-8.0008	0.2010
	QR	64.31	16.12	-5.69	0.2053	-7.7543	17.3963
	Naive	69.21	17.14	-0.79	0.1235	-7.5989	1.6534

datasets for comparison purpose. We first compare the proposed multiobjective interval prediction method with other common single-objective interval forecast methods, then different NN models in a multiobjective framework are also compared.

The single-objective interval prediction methods considered in this study include NN based LUBE method [23], quantile regression (QR) method and Naive method. NN based LUBE method is widely studied for wind power interval prediction and the NN model is usually the MLP model. The NN based LUBE method proposed in [26] is adopted to conduct single-objective interval prediction for wind speed and the NN connection weights are tuned by particle swarm optimization (PSO) algorithm. The corresponding parameters can also be found in the reference. Note that some other feedforward NN models and optimization algorithms (e.g., genetic algorithm) can also be studied for single-objective interval forecast model which share the same principle. QR is another typical probabilistic forecast approach [50] which can be applied to interval prediction. Naive method is also a general benchmark forecast model and it works similarly with the persistence model in point forecast. Naive method forecasts future intervals based on the past historical data and it performs well for short-term forecast task. In this study, the forecast error is assumed to follow normal distribution, the last wind power value is used as the mean, and the variance is calculated based on the latest observations [14]. With the mean and variance, PIs can be constructed for the forecast horizon.

The interval prediction results from the proposed approach and the benchmark techniques for Victoria and Edmonton data are given in Table 3 and Table 4, respectively. In addition to PICP and PINAW, ACE, AWD and Score indices are also presented. The CWC value which is a comprehensive index is also listed in the results. Note that the results of the proposed WNN-PICEA-g method are generated from the most satisfactory solution of the Pareto solution set as mentioned above. For better persuasiveness, we conduct the experiments with different PINC values, i.e., PINC=85%, 80%, 75% and 70% are also studied as shown in the tables.

From Table 3, we can see that the proposed method and NN-LUBE-PSO method can generate valid PIs (PICP $\geq$ PINC) for Victoria data for all experiments. However, QR and Naive methods are not so good. Obviously, the result of the proposed method has the minimum interval width and ACE value. Although the NN-LUBE-PSO method construct PIs with a high reliability, the PIs are less informative as they are too wide, and the very high probability also leads to a slightly lower AWD value. Since CWC and Score index can measure both the coverage probability and interval width of PIs, they can be used to compare the overall performance of various forecast approaches. Therefore, the proposed multiobjective interval prediction method has the best performance in Table 3. In addition, the multiobjective interval prediction method produces a set of Pareto solutions with a simulation run which can offer more choices to the decision maker than the single-objective prediction methods. More specifically, we can select the proper solutions from the Pareto solutions according to different PINC requirements, while several simulation experiments need to be conducted with a single-objective forecast model.

Similar forecast results are also obtained for Edmonton data as shown in Table 4. The proposed WNN-PICEA-g method can still construct valid PIs with narrow width which demonstrates the stability and consistency of the method. It is still the best forecast method according to the CWC value

TABLE 4. Comparison results for Edmonton data.

PINC	Method	PICP (%)	PINAW (%)	ACE (%)	AWD	Score	CWC
90%	WNN-PICEA-g	90.18	26.11	0.18	0.0233	-3.5946	0.2611
	NN-LUBE-PSO	94.57	33.12	4.57	0.0133	-3.9449	0.3312
	QR	81.78	27.54	-8.22	0.0664	-4.2310	61.1661
	Naive	86.43	27.85	-3.57	0.0389	-3.9903	6.2351
85%	WNN-PICEA-g	87.49	23.60	2.49	0.0389	-4.9093	0.2360
	NN-LUBE-PSO	90.92	27.73	5.92	0.0237	-5.1288	0.2773
	QR	75.44	22.61	-9.56	0.1039	-5.4662	119.1755
	Naive	81.70	24.64	-3.30	0.0557	-5.2556	5.4560
80%	WNN-PICEA-g	80.18	20.19	0.18	0.0728	-5.9536	0.2019
	NN-LUBE-PSO	83.27	23.04	3.27	0.0410	-6.0864	0.2304
	QR	70.19	19.20	-9.81	0.1466	-6.4896	135.2586
	Naive	78.34	22.12	-1.66	0.0749	-6.3506	2.5200
75%	WNN-PICEA-g	76.98	19.11	1.98	0.0755	-6.9284	0.1911
	NN-LUBE-PSO	77.27	21.75	2.27	0.0715	-7.5641	0.2175
	QR	64.76	16.61	-10.24	0.1947	-7.3444	167.2513
	Naive	74.80	19.98	-0.20	0.0969	-7.2948	1.3047
70%	WNN-PICEA-g	70.42	16.14	0.42	0.1033	-7.4815	0.1614
	NN-LUBE-PSO	74.87	17.89	4.87	0.0861	-7.6088	0.1789
	QR	59.91	14.52	-10.09	0.2495	-8.0989	155.4964
	Naive	70.58	18.10	0.58	0.1226	-8.1209	0.1810

followed by the NN-LUBE-PSO method. For QR and Naive method, the PICP value cannot reach the nominal confidence level most of the time resulting a high CWC value. Furthermore, we can find that the ACE value of the proposed method is much closer to 0. In summary, the proposed multiobjective interval prediction method can construct PIs effectively and performs better than the benchmark approaches.

Since several other multiobjective interval prediction methods based on NN models have been reported in the previous literature [27], [29], we also implement multiobjective comparison between the WNN model and other NN models including MLP NN and RBF NN. The implementation strategy for MLP NN and RBF NN model is the same with that of the proposed model. The PICEA-g algorithm with the same parameters is still used as the training algorithm. To compare the performance of different models quantitatively, the hypervolume indicator is adopted to measure the obtained Pareto front.

The average hypervolume results of different NN forecast models for Victoria and Edmonton data are given in Table 5. The hypervolume values in this table are average results for the test dataset from five independent simulation runs and a larger hypervolume value means a better result for multiobjective minimization problem. As can be seen from this table, the WNN model is slightly better than the MLP NN model, but it performs much better than the RBF NN model, especially for the Edmonton dataset. Therefore, the proposed WNN forecast model performs best which has the maximum hypervolume.

In addition, NSGA-II algorithm is one of the most efficient multiobjective optimization algorithms and has been widely studied to deal with different multiobjective problems. To verify the performance of PICEA-g algorithm employed in this

TABLE 5. Multiobjective comparison results for different models.

	WNN	MLP	RBF	WNN-NSGA-II
Victoria hypervolume	1.3056	1.3053	1.2993	1.2989
Edmonton hypervolume	1.3047	1.3043	1.2899	1.2966

study, NSGA-II algorithm is used to train the WNN model for comparison. The hypervolume results obtained from the WNN model with NSGA-II algorithm are listed in the last column of Table 5, which are also average results of five individual runs. It is obvious that the hypervolume results from WNN trained by PICEA-g algorithm are better than those from WNN with NSGA-II algorithm. Thus, we can conclude that PICEA-g algorithm has good performance in the proposed multiobjective interval prediction method.

E. DISCUSSION

The proposed multiobjective interval prediction model mainly consists of WNN model and PICEA-g optimization algorithm. For PICEA-g algorithm, we maintain the widely used crossover and mutation operators, and the corresponding parameter values are collected from the reference which can be considered as optimal in this work. It is possible to study different crossover and mutation operators and corresponding parameters to further improve the performance of the forecast method in future. In addition, we study the influence of the population size *Npop*. The average training time and average training hypervolume results with different *Npop* are summarized in Table 6, which shows that the training time has a positive correlation with the *Npop* value. We can also find that all the training time for different population size

TABLE 6. Training results for different population size.

Npop	Victoria		Edmonton	
	time (s)	hypervolume	time (s)	hypervolume
30	36.90	1.3050	37.43	1.3122
40	44.73	1.3086	44.14	1.3169
50	52.36	1.3125	52.17	1.3183

TABLE 7. Hypervolume results for different feature selection methods.

Npop	Victoria		Edmonton	
	PSR method	CA method	PSR method	CA method
30	1.2874	1.3050	1.2977	1.3122
40	1.2951	1.3086	1.3048	1.3169
50	1.3017	1.3125	1.3094	1.3183

are less than one minute which shows the computational efficiency of the optimization algorithm. When the training process is finished, the testing time is less than one second. Particularly, the training time is about 45s when *Npop* is 40 which is much less than the time scale (1 hour) of the dataset. Hence, the proposed forecast model can be used to real-time wind speed forecast. The increasing hypervolume values result from the increasing evenly distributed points in the Pareto front. Therefore, the decision maker needs to select a proper population size to balance the training time and the number of Pareto solutions in practice.

In addition, various feature selection methods, such as mutual information method, recursive feature elimination, and chaotic feature selection based on phase space reconstruction, can be investigated to preprocess the input data which may potentially improve the forecast performance. The correlation analysis method is used in this work for its efficiency and simplicity. To evaluate its effectiveness, we study another feature selection method, phase space reconstruction, for comparison purpose. The phase space reconstruction technique aims to determine the delay vectors as the input. By delay embedding theorem, we need to find two parameters in terms of the embedding dimension and the time delay, which can be obtained by the mutual information method and false nearest method, respectively [51]. In this work, the embedding dimension is 8 and the time delay is 13 for Victoria dataset, and they are 8 and 15 for Edmonton dataset, respectively. Then the delay vectors or the input can be constructed. The number of the hidden nodes in the WNN model is determined with the same method as introduced before. Similarly, the average training hypervolume results based on phase space reconstruction (denoted as PSR) with various *Npop* are given in Table 7, and hypervolume results with correlation analysis (denoted as CA) method are also listed for better comparison. From this table, we can find that the correlation analysis method is effective and sufficient to determine the input for short-term wind speed forecast. For more complex forecast tasks, it is worth studying other advanced feature selection methods, which is left for future work.

Compared with the single-objective interval prediction model, we can get a Pareto front (a set of optimal solutions) from the proposed multiobjective interval prediction model. Among the nondominated optimal solutions, the decision maker can flexibly choose a proper solution according to the demand of reliability and interval width. Each solution corresponds to an interval forecast model. With the choice of a proper model, interval prediction can be implemented with new dataset. In addition, there are different ways to use the prediction intervals in reality. For instance, they can be applied to robust optimization and control problems for power systems integrated with renewable generation [51]. More specifically, in robust optimization problems with box-type uncertainty set, the prediction intervals can be directly used to describe the uncertainty without the assumption of probability distribution, i.e., only the lower and upper bounds are required in robust optimization problems. They can also be processed to get the point forecast values by some convex combination methods, such as the weighted summation method with the obtained lower and upper bounds.

V. CONCLUSION

Wind power has been increasingly integrated and used in the existing power systems due to the clean and renewable advantage. It is imperative to forecast the highly uncertain wind speed accurately which is the determinant factor of wind power. In this work, multiobjective interval prediction based on WNN model is proposed for short-term wind speed forecast. The novel multiobjective evolutionary algorithm, PICEA-g, is employed to train the WNN prediction model. Two case studies are implemented to testify the performance of the proposed model with real-world hourly wind speed data from Canada, and valid and narrow PIs are obtained. In addition, experimental results show the superiority of the proposed approach compared with other benchmark methods, the performance of PICEA-g algorithm is also verified by a comparison with the popular NSGA-II algorithm. In particular, the quality of PIs from the proposed multiobjective model is better than those from other single-objective forecast models with respect to different PINC settings including 90%, 85%, 80%, 75% and 70%. For multiobjective model comparison, the proposed model also achieves higher average hypervolume results (1.3056 for Victoria dataset and 1.3047 for Edmonton dataset) than other models. Moreover, the training time with PICEA-g algorithm for two case studies are both less than one minute which shows the feasibility of the proposed model.

As for future research, the proposed prediction model can be further improved by considering more relevant input such as the NWP data. In addition, some other advanced feature selection techniques may be adopted to determine the proper input of the model, such as recursive feature selection, mutual information based method [52] and so on. Long-term wind speed interval forecast may also be studied in the future.

REFERENCES

[1] Canadian Wind Power Association. *Wind Energy in Canada*. Accessed: Aug. 20, 2017. [Online]. Available: http://canwea.ca/wind-facts/why-wind-works/
[2] P. Hu, R. Karki, and R. Billinton, "Reliability evaluation of generating systems containing wind power and energy storage," *IET Gener., Transmiss. Distrib.*, vol. 3, no. 8, pp. 783–791, Aug. 2009.
[3] E. B. Ssekulima, M. B. Anwar, A. A. Hinai, and M. S. El Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: A review," *IET Renew. Power Gener.*, vol. 10, no. 7, pp. 885–989, Jul. 2016.
[4] E. Erdem and J. Shi, "ARMA based approaches for forecasting the tuple of wind speed and direction," *Appl. Energy*, vol. 88, no. 4, pp. 1405–1414, Apr. 2011.
[5] D. L. Faria, R. Castro, C. Philippart, and A. Gusmao, "Wavelets pre-filtering in wind speed prediction," in *Proc. Int. Conf. Power Eng., Energy Elect. Drives*, Mar. 2009, pp. 168–173.
[6] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, "Long-term wind speed forecasting and general pattern recognition using neural networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 546–553, Apr. 2014.
[7] X. Kong, X. Liu, R. Shi, and K. Y. Lee, "Wind speed prediction using reduced support vector machines with feature selection," *Neurocomputing*, vol. 169, pp. 449–456, Dec. 2015.
[8] O. Özgönenel and D. W. P. Thomas, "Short-term wind speed estimation based on weather data," *Turkish J. Elect. Eng. Comput. Sci.*, vol. 20, no. 3, pp. 335–346, 2012.
[9] C. Wang, J. Wu, J. Wang, and Z. Hu, "Short-term wind speed forecasting using the data processing approach and the support vector machine model optimized by the improved cuckoo search parameter estimation algorithm," *Math. Problems Eng.*, vol. 2016, May 2016, Art. no. 4896854, doi: 10.1155/2016/4896854.
[10] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renew. Sustain. Energy Rev.*, vol. 34, pp. 243–254, Jun. 2014.
[11] S.-W. Fei and Y. He, "Wind speed prediction using the hybrid model of wavelet decomposition and artificial bee colony algorithm-based relevance vector machine," *Int. J. Elect. Power Energy Syst.*, vol. 73, pp. 625–631, Dec. 2015.
[12] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renew. Sustain. Energy Rev.*, vol. 32, pp. 255–270, Apr. 2014.
[13] Z. Song, Y. Jiang, and Z. Zhang, "Short-term wind speed forecasting with Markov-switching model," *Appl. Energy*, vol. 130, pp. 103–112, Oct. 2014.
[14] 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.
[15] R. Ak, Y.-F. Li, V. Vitelli, and E. Zio, "Adequacy assessment of a wind-integrated system using neural network-based interval predictions of wind power generation and load," *Int. J. Elect. Power Energy Syst.*, vol. 95, pp. 213–226, Feb. 2018.
[16] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renew. Sustain. Energy Rev.*, vol. 31, pp. 762–777, Mar. 2014.
[17] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output," *Appl. Energy*, vol. 99, pp. 154–166, Nov. 2012.
[18] H. Z. Wang, G. B. Wang, G. Q. Li, J. C. Peng, and Y. T. Liu, "Deep belief network based deterministic and probabilistic wind speed forecasting approach," *Appl. Energy*, vol. 182, pp. 80–93, Nov. 2016.
[19] H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, pp. 95–108, Jan. 2006.
[20] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, Oct. 2012.
[21] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation forecasts," *IEEE Trans. Sustain. Energy*, vol. 4, no. 3, pp. 602–610, Jul. 2013.
[22] J. Zhao, Z.-H. Guo, Z.-Y. Su, Z.-Y. Zhao, X. Xiao, and F. Liu, "An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed," *Appl. Energy*, vol. 162, pp. 808–826, Jan. 2016.
[23] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 337–346, Mar. 2011.
[24] W. Wu, K. Chen, Y. Qiao, and Z. Lu, "Probabilistic short-term wind power forecasting based on deep neural networks," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Oct. 2016, pp. 1–8.
[25] J. Wang, T. Niu, H. Lu, Z. Guo, W. Yang, and P. Du, "An analysis-forecast system for uncertainty modeling of wind speed: A case study of large-scale wind farms," *Appl. Energy*, vol. 211, pp. 492–512, Feb. 2018.
[26] 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.
[27] 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.
[28] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[29] C. Zhang, H. Wei, L. Xie, Y. Shen, and K. Zhang, "Direct interval forecasting of wind speed using radial basis function neural networks in a multi-objective optimization framework," *Neurocomputing*, vol. 205, pp. 53–63, Sep. 2016.
[30] N. A. Shrivastava, K. Lohia, and B. K. Panigrahi, "A multiobjective framework for wind speed prediction interval forecasts," *Renew. Energy*, vol. 87, pp. 903–910, Mar. 2016.
[31] E. E. Elattar, "Prediction of wind power based on evolutionary optimised local general regression neural network," *IET Gener., Transmiss. Distrib.*, vol. 8, no. 5, pp. 916–923, May 2014.
[32] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network-based energy price forecasting in electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1423–1432, Aug. 2008.
[33] L. J. Ricalde, G. A. Catzin, A. Y. Alanis, and E. N. Sanchez, "Higher order wavelet neural networks with kalman learning for wind speed forecasting," in *Proc. IEEE Symp. Comput. Intell. Appl. Smart Grid (SSCI CIASG)*, Paris, France, Apr. 2011, pp. 1–6.
[34] H. Chitsaz, N. Amjady, and H. Zareipour, "Wind power forecast using wavelet neural network trained by improved clonal selection algorithm," *Energy Convers. Manage.*, vol. 89, pp. 588–598, Jan. 2015.
[35] R. Wang, R. C. Purshouse, and P. J. Fleming, "Preference-inspired coevolutionary algorithms for many-objective optimization," *IEEE Trans. Evol. Comput.*, vol. 17, no. 4, pp. 474–494, Aug. 2013.
[36] A. Khosravi and S. Nahavandi, "Combined nonparametric prediction intervals for wind power generation," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 849–856, Oct. 2013.
[37] N. A. Shrivastava, A. Khosravi, and B. K. Panigrahi, "Prediction interval estimation of electricity prices using PSO-tuned support vector machines," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 322–331, Apr. 2015.
[38] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. Roy. Statist. Soc. B (Statist. Methodol.)*, vol. 69, no. 2, pp. 243–268, 2007.
[39] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 889–898, Nov. 1992.
[40] Y. Chen, B. Yang, and J. Dong, "Time-series prediction using a local linear wavelet neural network," *Neurocomputing*, vol. 69, nos. 4–6, pp. 449–465, 2006.
[41] H. Shao, X. Deng, and F. Cui, "Short-term wind speed forecasting using the wavelet decomposition and AdaBoost technique in wind farm of East China," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 11, pp. 2585–2592, Aug. 2016.
[42] S. A. Billings and H.-L. Wei, "A new class of wavelet networks for nonlinear system identification," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 862–874, Jul. 2005.
[43] A. Mellit, M. Benghanem, and S. A. Kalogirou, "An adaptive wavelet-network model for forecasting daily total solar-radiation," *Appl. Energy*, vol. 83, no. 7, pp. 705–722, 2006.
[44] R. C. Purshouse, C. Jalbă, and P. J. Fleming, "Preference-driven co-evolutionary algorithms show promise for many-objective optimisation," in *Evolutionary Multi-Criterion Optimization—EMO* (Lecture Notes in Computer Science), vol 6576, R. H. C. Takahashi, K. Deb, E. F. Wanner, and S. Greco, Eds. Berlin, Germany: Springer, 2011.

[45] Z. Shi, R. Wang, and T. Zhang, "Picea-g using an enhanced fitness assignment method," in *Proc. IEEE Symp. Comput. Intell. Multi-Criteria Decis.-Making (MCDM)*. Orlando, FL, USA, Dec. 2014, pp. 72–77.
[46] R. Wang, R. C. Purshouse, and P. J. Fleming, "Preference-inspired co-evolutionary algorithm using adaptively generated goal vectors," in *Proc. IEEE Congr. Evol. Comput.*, Cancun, Mexico, Jun. 2013, pp. 916–923.
[47] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, Nov. 1999.
[48] *Canada Weather Stats*. Accessed: Aug. 20, 2017. [Online]. Available: https://www.weatherstats.ca/
[49] C. M. Fonseca, L. Paquete, and M. López-Ibánez, "An improved dimension-sweep algorithm for the hypervolume indicator," in *Proc. Congr. Evol. Comput.*, Palo Alto, CA, USA, Jul. 2006, pp. 1157–1163.
[50] R. Koenker and K. Hallock, "Quantile regression: An introduction," *J. Econ. Perspect.*, vol. 15, no. 4, pp. 43–56, 2001.
[51] Z. Shi, H. Liang, and V. Dinavahi, "Direct interval forecast of uncertain wind power based on recurrent neural networks," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1177–1187, Jul. 2018.
[52] N. Amjady, F. Keynia, and H. Zareipour, "Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization," *IEEE Trans. Sustain. Energy*, vol. 2, no. 3, pp. 265–276, Jul. 2011.

ZHICHAO SHI (S'17) received the B.Sc. degree from Nankai University, Tianjin, China, in 2013, and the M.Sc. degree from the National University of Defense Technology, Changsha, China, in 2015. He is currently pursuing the Ph.D. degree in electrical and computer engineering at the University of Alberta, Edmonton, AB, Canada. His current research interests include stochastic modeling in smart grid, computational intelligence and optimization, microgrid energy management, and power systems optimization.

HAO LIANG (S'09–M'14) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2013. He has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, since 2014. His current research interests include smart grid, wireless communications, and wireless networking.

VENKATA DINAVAHI (S'94–M'00–SM'08) received the B.Eng. degree in electrical engineering from the National Institute of Technology, Nagpur, India, in 1993, the M.Tech degree in electrical engineering from IIT Kanpur, India, in 1996, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2000. He is currently a Professor with the Department of Electrical and Computer engineering, University of Alberta, Edmonton, AB, Canada. His research interests include real-time simulation of power systems and power electronic systems, electromagnetic transients, device-level modeling, large-scale systems, and parallel and distributed computing.