Energy 196 (2020) 117072

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Probabilistic wind power forecasting based on spiking neural network

Huaizhi Wang [a], Wenli Xue [a], Yitao Liu [a, *], Jianchun Peng [a], Hui Jiang [b]

[a] College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen, 518060, China
[b] College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen, 518060, China

A R T I C L E I N F O

Article history:
Received 5 July 2019
Received in revised form
27 December 2019
Accepted 29 January 2020
Available online 1 February 2020

Index Terms:
Spiking neural network
Wind power forecasting
Direct probabilistic forecasting
Group search optimizer

A B S T R A C T

Accurate and reliable quantification of uncertainty in wind power forecasting is critical to the economic operation and real time control of the electric power and energy system. To this end, this paper proposes a novel direct method for probabilistic wind power forecasting based on spiking neural network. In this method, a new forecasting framework is firstly formulated to simultaneously calculate the coverage probability and sharpness with associated confidence levels. Then, group search optimizer is introduced and re-designed to optimize the parameters of the forecasting framework and directly generate the prediction intervals, so that the prediction reliability and stability are ensured. The main advantage of the proposed probabilistic forecasting method is that it does not involve any distribution assumption of the prediction errors required by most existing forecasting methods. The wind power datasets from real wind farms in Belgium and China are used in the case studies. Traditional back-propagation neural network (BPNN), support vector machine (SVM) and extreme learning machine (ELM) are selected as the benchmarking algorithms. The simulation results show that the average coverage of the proposed method is improved by 72.0%, 54.9% and 51.3% respectively, when compared to BPNN, SVM and ELM. The improvement rates of sharpness index are 43.1%, 28.1% and 21.0%, respectively. These statistical results show that the proposed method outperforms BPNN, SVM and ELM in terms of forecasting accuracy, demonstrating that this method has high practical applications in real power systems.

1. Introduction

In electrical energy system, power generation and consumption must be balanced at all times as power cannot be efficiently stored on a large scale [1]. However, with the rapid increase of the installation capacity of renewable wind power generation, the randomness, volatility and reverse load characteristics of wind power will definitely aggravate the power supply-consumption imbalance, thus bringing great challenges to the economic operation, stability and security of the electric energy system [2]. As an important tool, wind power forecasting has received much attention in recent years to ensure grid safety and improve the utilization efficiency of wind energy [3]. Wind power forecasting error and its distribution not only increase the up/down reserve capacity of power system and affect equipment maintenance plan, but also cause negative impacts on energy market transactions and charging plans of energy storage station. Therefore, wind power forecasting is of great significance for the economic and secure operation of electric energy systems [4]. Many algorithms have been reported in the literature to provide accurate wind power prediction. They are usually divided into four categories: physical method, statistical model, artificial neural network (ANN) and their hybrids [5].

Physical methods are generally proposed based on numerical weather prediction model, which is used to simulate atmospheric dynamics according to physical principles and boundary conditions [6]. Physical prediction methods usually include multiple sub-models, each of which considers temperature, pressure, topography, and obstacle to provide wind energy predictions of the measurement points [7]. By combining the prediction results of each sub-model, the final prediction results can be obtained by using the wind profile model and computational fluid dynamics. Wang et al. proposed a new sequence transfer correction algorithm for wind energy forecast based on numerical weather prediction [8]. The observed wind speed was taken as the input variable in the wind speed correction model. Extensive simulation results demonstrated that the proposed algorithm was able to improve the forecasting accuracy by 1.7–16.1% in two real wind farms. Yang

* Corresponding author.
E-mail address: liuyt@szu.edu.cn (Y. Liu).

https://doi.org/10.1016/j.energy.2020.117072
0360-5442/

et al. proposed two forecasting models based on numerical weather prediction to predict a gust of wind during a typhoon event [9], including a linear regression model and a micro-genetic algorithm-based model. In addition, a successively accumulated regression process was designed to update the parameters of the proposed forecasting models, thereby improving the prediction accuracy. These studies show that physical methods are effective in predicting atmospheric dynamic performance. However, they require a large amount of computational resources [10]. Consequently, physical methods are not suitable for short-term wind power prediction [11].

The statistical models aim to reveal the mathematical relationship in historical samples [12] that reflect the past regularity of time series data, and thus infer future wind energy varying trends. Statistical models consist of the following steps [13]: preliminary analysis of historical data, determination of the prediction model, parameter estimation, extrapolation prediction and forecasting error analysis. Common statistical models include autoregressive moving average [14], Bayesian method [15], Kalman filter [16], Markov chain [17] and grey theory [18]. A statistical interpretation method combining autoregressive moving average and grey prediction was developed for robust short-term prediction of wind power generation [19]. The weighting method was introduced to integrate multiple point forecasters. Finally, the deterministic result and prediction interval of wind power forecasting can be estimated numerically. Wu et al. proposed a new distribution estimation method based on Laplace and normal distribution to model the wind power forecast error [20]. This distribution method was then applied to assess the penalties of prediction errors on the electricity market, thereby determining the optimal size of energy storage in wind farm. However, statistical models are often formulated as linear models, limiting their abilities to solve more challenging problems with large forecasting horizons [21].

The third type of research focuses on ANN-based forecasting method. Basically, ANNs are computational models that mimic the structure and function of biological brains [22]. ANN is primarily used to estimate or approximate any nonlinear function. In addition, ANN can change its internal structure according to external information, and can effectively extract the potential features in any data. Therefore, ANN based forecasting methods always have better prediction performance than physical methods and statistical methods [23]. Support vector machine (SVM) [24], extreme learning machine (ELM) [25], and fuzzy neural network [26] are commonly used to handle the nonlinear relationships between input and output of wind power samples, and new predictions are obtained by minimizing the mapping errors. A new hybrid deep learning-based ANN model was developed for 24-h ahead wind power forecasting [27]. In this model, the convolution and pooling operations were used to extract the features in wind power. Numerical results demonstrated that the proposed forecasting model had better forecasting performance than traditional statistical methods. Ding et al. proposed a new gated recurrent unit neural network for wind speed error correction [28]. Then, the wind power curve model was applied to achieve the forecasting of wind power by using corrected numerical weather prediction.

An extensive comparison of the existing forecasting models was given in Ref. [29], showing that each single model has its own benefits and shortcomings. Therefore, the fourth type of research focuses on combining different prediction algorithms as a hybrid to take advantage of each algorithm to achieve the purpose of improving prediction performance. For example, Jiang et al. proposed a new hybrid method for wind speed forecasting based on fuzzy time series and multi-objective optimization [30]. The case study showed that the average absolute percentage error of the proposed model was less than 4%, having a high forecasting accuracy. Also, a general framework consisting of data preprocessing, data clustering and forecasting module was developed for wind speed forecasting [31]. Simulation results indicated that this framework outperformed other traditional forecasting models. A comprehensive and extensive review of the hybrid methods for wind energy forecasting can be found in Refs. [32,33] and the references therein.

Most of the current research concentrates on the development of point forecasters for wind energy prediction. However, due to the chaotic nature of weather system on earth, wind power data usually exhibits strong non-stationary and volatility [34], making the prediction error inevitable. These prediction errors will undoubtedly affect the daily operation and management of power system. Wind power probabilistic forecasting aims to quantify the uncertainty in wind power prediction and obtain the probability of multiple wind power output levels, which can help dispatchers prepare for possible scenarios in advance, thus reducing the risk of power system control and management [35]. Therefore, wind power probabilistic forecasting has received much attention in recent years. To date, wind power probabilistic prediction methods can be divided into three main categories. In the first category of research, wind speed is required to be predicted at first by using the existing point forecasting techniques. Then, the wind speed forecasting results are transformed into wind power forecasting results by using a wind turbine power curve modeling method. Wind turbine power curve shows the relationship between the output power and hub height wind speed, helping in estimating the potential wind power in a candidate site. A comparative study of power curve modeling methods and a comprehensive review on wind turbine power curve modeling techniques can be found in Refs. [36,37], respectively. Finally, when the point forecasting of wind power is obtained, the probabilistic distributions with different wind power output levels can be statistically estimated by assuming that the prediction error meets a certain distribution [38], e.g. Gaussian distribution. In the second category, a point forecaster is also required for wind power forecasting in order to obtain a set of forecasting errors. Then, nonparametric methods such as kernel density estimation and quantile regression are used to directly estimate the probabilistic information of wind power generation with multiple confidence levels [39]. However, these two types of methods actually belong to indirect probabilistic forecasting methods that either need to presuppose the probabilistic distribution of the prediction errors. They also separate the uncertainty modeling of wind power from the final forecasting performance evaluation process [40]. This is actually not conducive to improving the forecasting accuracy for wind power probabilistic prediction.

The third type of research focuses on directly estimating the wind power prediction intervals based on lower upper bound estimation (LUBE) method [41]. LUBE is a non-parametric method based on neural network, which usually constructs prediction intervals under a given confidence level in a supervised learning manner. Compared with indirect probabilistic prediction methods, the main advantage of LUBE-based direct probabilistic prediction methods is that they do not need to make any assumptions about the distribution of prediction errors. At present, feedforward neural network (FNN) and recurrent neural network (RNN) have been reported in the literature for direct wind power probabilistic prediction. FNN refers to a neural network that has no feedback in the network and the input signal propagates unidirectionally from the input layer to the output layer. FNN-based probabilistic prediction methods usually estimate the prediction interval directly by simultaneously minimizing the interval prediction width and maximizing the coverage probability. Wan et al. proposed a wind power interval prediction method based on ELM, and used particle swarm optimization to optimize the network parameters of ELM

[42]. LUBE based multi-objective optimization was formulated as a new constrained single-objective problem in Ref. [43]. Electricity loads in Singapore were used to verify the feasibility of this problem transformation. Abdollah et al. proposed a new LUBE method based on fuzzy logic to overcome the instability caused by the use of neural networks in traditional LUBE methods [44]. RNN is a type of recursive neural network that takes sequence data as input and performs recursion in the evolution direction of the sequence. RNN connects all nodes in a chain, and is very suitable for processing time series data. Shi et al. constructed a new LUBE method with a comprehensive cost function based on RNN [45]. Dragonfly algorithm was used to optimize the weights of the interval prediction model. Wang et al. proposed a wind power forecasting method based on the echo state network, and adopted the ensemble approach to mitigate the model misspecification uncertainty and data noise uncertainty [46]. A comprehensive review of probabilistic wind power forecasting can be found in Ref. [47] and the references therein.

In addition, most of the existing forecasting models generally use traditional ANNs to achieve time-series prediction of wind power. These ANNs, such as multi-layer FNN, SVM, and ELM, have certain limitations [40], such as local minimum, overfitting and high computational cost. This is mainly because traditional ANN models cannot simultaneously consider neurons and synaptic states and cannot accurately reflect the biological structure of real natural neurons [48]. In addition, traditional ANNs do not incorporate temporal sequences into their networks, which limits their ability to process spatiotemporal data. Although RNN has a memory function and can be used to process time-series data [49], RNN has a certain complexity and is computationally expensive. In recent years, the spiking neural network (SNN) has attracted great attention from academia as well as industry because it not only consider both neurons and synaptic states, but also integrate the capability for processing temporal sequences into its network [50]. The neurons in SNN are not activated in every propagation period, but only when the membrane potential reaches a predetermined value. This structure makes SNN closer to the natural neurons, so SNN can process more information at the same time and exhibits powerful calculation capability. A spike response model based SNN architecture was mooted for short-term wind speed forecasting [51]. In this model, the SpikeProp learning algorithm was improved by adding momentum items and adaptively adjusting the learning rate. Simulation results demonstrated that the proposed SNN model was effective for wind speed forecasting. Madhiarasan et al. proposed a hybrid method for long-term wind speed forecasting based on SNN and modified grey wolf optimization algorithm [52]. SNN was used to model the relationship between environmental parameters and real wind speed, and grey wolf algorithm was applied to optimize the SNN parameters. The effectiveness and efficiency of the proposed hybrid method were verified by using real-time wind speed data. A new SNN-based architecture was developed in Ref. [53] for the prediction of wind farm energy production. The main benefit of this architecture is its capability for evaluating the wake effects. The real data from a large wind power plant in Italy was used to prove the superiority of the proposed forecasting architecture. More information on wind speed prediction based on SNN can be found in Ref. [54].

So far, the above SNN-based prediction methods actually belong to point forecasters. Forecasting error assumptions or non-parametric methods are required to achieve probabilistic forecasting of wind energy. In other words, the existing SNN-based prediction methods belong to indirect prediction framework, and the corresponding forecasting accuracy is largely affected due to the adopted point forecasting framework, statistical analysis method and power curve modeling technique. Therefore, it is a pressing need to develop a direct probabilistic forecasting method based on SNN to quantify the uncertainty of wind power prediction and improve the probabilistic forecasting accuracy. Recognizing the above challenges, this paper proposes a direct wind power probabilistic prediction method based on SNN. Compared with the existing studies on similar topics, this paper has three major contributions. First, this paper proposes a direct probabilistic forecasting model based on SNN and LUBE. The forecasting model uses the resistor-capacitance (RC) circuit model as the neuron model, and the prediction intervals (PI) under different confidence levels can be directly obtained through the supervised learning algorithm. The main advantage of the proposed direct forecasting model is that it does not require any prior knowledge and prediction error distribution assumptions, and thus has better prediction accuracy than the existing indirect probabilistic forecasting methods. Second, this paper proposes a new PI optimization model based on the continuous ranking probability score (CRPS). The optimization model adjusts the model parameters of the SNN while optimizing the prediction reliability and interval sharpness to ensure that the wind power probabilistic forecasting results are consistent with the distribution of the observations. Third, the group search optimizer (GSO) is introduced in this paper to solve the proposed PI optimization model. GSO is a heuristic algorithm that has been numerically proven to exhibit global search capabilities with fast convergence and gradient-free optimization. The proposed direct probabilistic prediction method is validated in two wind farms in Belgium and China.

The rest of this paper is organized as follows. Section 2 describes how to formulate a probabilistic forecasting model based on SNN. The third Section expresses how to directly achieve optimal probabilistic forecasting by using GSO. The performance criteria of probabilistic forecasting, including PI coverage and sharpness, are introduced in Section 4. Extensive case studies are carried out in Section 5, and the superiority of the proposed method is demonstrated. Finally, the conclusion is given in the Section 6.

2. Probabilistic forecasting formulation based on spiking neural network

In general, SNN is suitable for processing time-series data due to its inherent temporal coding in neurons. Therefore, a new direct wind power probabilistic forecasting framework is developed in this paper based on SNN with associated confidence levels, as shown below.

2.1. Spiking neural network

SNN is referred to as the third-generation artificial neural network, and the modeling of its neurons is close to real neuronal cells [55]. The activation level of a neuron in SNN is generally considered to be the current state, and the input pulse will increase the current state value over a period of time and then gradually decay. In addition, SNN also considers the impact of temporal information, and so it is also suitable for processing time series data [56]. Accordingly, SNN is widely used in the fields of parallel computing, image processing, pattern recognition, etc. We now elaborate the basic concepts of SNN from three aspects: neuron modeling, network modeling and supervised training.

1) Neuron modeling: In SNN, neurons are usually abstracted into a resistor-capacitance (RC) circuit. As long as the neurons receive the input current, the membrane potential rises until the activation threshold is reached. At this moment, a pulse, also termed as a spike, will be released and the membrane potential will immediately return to the resting value. In addition, the

membrane potential will gradually decay when there has no input in the neurons. This whole process with respect to the change of membrane potential is called integral and fire (IF) model, which can be used for the simulation of neuron dynamics [57]. The mathematical representation of the IF model is described as follows:

$$I(t) - V_m/R_m = C_m(dV_m(t)\,/\,dt) \tag{1}$$

where V_m represents the membrane potential; R_m is the leakage resistance; $I(t)$ is the input current in t-th step; C_m denotes the membrane capacitance.

In (1), the external input current is obtained from a pulse signal in a presynaptic neuron. Therefore, the pulse current from the j-th neuron to the i-th neuron can be expressed as:

$$I(t) = \sum_i w_{ij} \sum_f \alpha\left(t - t_j^{(f)}\right) \tag{2}$$

where w_{ij} is the weight between the i-th neuron and j-th neuron; $t(f)\ j$ represents the f-th spike of the neurons in previous layer; α denotes a Dirac function. The modeling process of a SNN neuron in presented in Fig. 1.

2) SNN network modeling: The neuron modeling process in the previous subsection can be used to build a feedforward SNN. The input layer should be the coding layer, which is mainly used to convert the input data into a standard value in per unit. The hidden layer uses the spiking neuron model in (1)–(2), and the output layer uses traditional neural networks. In order to ensure that enough spiking neurons are activated, we set the adjacent neurons to be fully connected, while the neurons in the same layer have no inner connections. In addition, we assume that each neuron has a different response delay to ensure the stability of the feedforward SNN. The basic structure of the feedforward SNN is given in Fig. 2.
3) Supervised learning algorithm: A gradient decent algorithm is implemented in this paper to train the feedforward SNN. This algorithm was proposed by the National Center for Mathematics and Computer Science in Netherland. It is known as the Spike Prop learning algorithm. It adopts a simplified response model that limits each neuron to generate only one spike. Consequently, the discontinuity problem from neuron activation can be solved effectively [58]. In Spike Prop algorithm, the membrane potential of the j-th neuron has the following form:

$$V_j(t) = \sum_{i \in \Gamma_j} \sum_k w_{ij}^k \beta\left(t - t_i^{out} - d_{ij}^k\right) \tag{3}$$

where $wk\ ij$ is the k-th weight of the connection between the i-th neuron and j-th neuron, with $dk\ ij$ representing its delay; $V_j(t)$ is the membrane potential of the j-th neuron at t-th time step; $tout\ i$ denotes the time step when the spike of the i-th neuron is generated; Γ_j is the collection of all neurons connected to the j-th

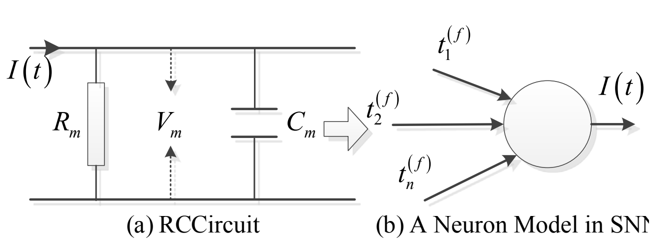

Fig. 1. The modeling process of a neuron in SNN.

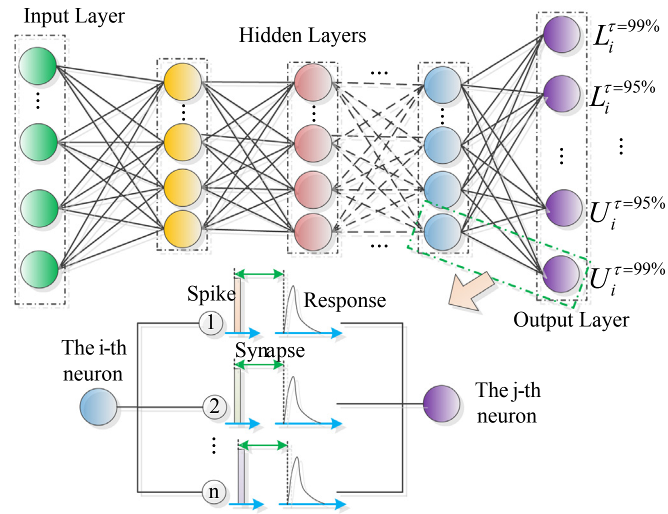

Fig. 2. The probabilistic forecasting framework based on SNN.

neuron; Finally, β denotes a kernel function that quantifies the impact of spikes on membrane potential. Obviously, the equation (3) can be rewritten as:

$$V_j(t) = \sum_{i \in \Gamma_j} \sum_k w_{ij}^k y_i^k(k) \tag{4}$$

where $yk\ i(k)$ is the consequence of the spikes on the i-th neuron.

Thereafter, an energy function is defined as a function of network errors,

$$E = \frac{1}{2} \sum_j \left(t_j^{out} - t_j^{(d)}\right)^2 \tag{5}$$

where E denotes the energy function.

According to the back-propagation algorithm, the connection weight from the hidden layer to the output layer can be updated as follows,

$$\frac{\partial E}{\partial w_{ij}^k} = \frac{\partial E}{\partial t_j^{out}} \frac{\partial t_j^{out}}{\partial V_j} \frac{\partial V_j}{\partial w_{ij}^k} \tag{6}$$

The equation (6) is very complex and it can be linearized around $tout\,j$. Consequently, V_j can be approximated as a linear function of t, as follows,

$$\frac{\partial t_j^{out}}{\partial V_j} = -\frac{1}{\sum_{i \in \Gamma_j} \sum_k w_{ij}^k \left(\partial y_j^k\left(t_j^{out}\right) / \partial t_j^{out}\right)} \tag{7}$$

Therefore, the gradient of the connection weight from the hidden layer to the output layer can be estimated as:

$$\frac{\partial E}{\partial w_{ij}^k} = y_j^k\left(t_j^{out}\right)\delta_j \tag{8}$$

where δ_j represents a gradient increment that can be updated according to:

$$\delta_j = -\frac{t_j^{out} - t_j^d}{\sum_{i \in \Gamma_j} \sum_k w_{ij}^k \left(\partial y_j^k\left(t_j^{out}\right) / \partial t_j^{out}\right)} \tag{9}$$

The connection weights between the input layer to the hidden layer can be assessed in the similar manner, and its corresponding gradient is updated based on the following equation:

$$\frac{\partial E}{\partial w_{ij}^{k}} = y_j^k\left(t_j^{out}\right)\delta_j = y_j^k\left(t_j^{out}\right)\cdot\frac{\sum_{i\in\Gamma_j}\delta_i\sum_k w_{ij}^k\left(\partial y_j^k\left(t_j^{out}\right)/\partial t_j^{out}\right)}{\sum_{h\in\Gamma_h}\sum_l w_{hi}^l\left(\partial y_h^l\left(t_i^{out}\right)/\partial t_i^{out}\right)} \tag{10}$$

where Γ_h is the set of neurons connected to the h-th neuron; δ_i is the gradient increment of the i-th neuron.

From (3)–(10), it is clear that the SNN connection weights can be updated based on the following equation:

$$\varDelta w_{ij}^k = -\eta y_j^k\left(t_j^{out}\right)\delta_j \tag{11}$$

where η represents the learning rate.

2.2. Probabilistic forecasting framework based on SNN

This subsection proposes a new direct probabilistic forecasting framework based on SNN. The idea of this framework originates from LUBE. The input of the proposed forecasting framework is wind power time-series data, and the outputs are the lower and upper limits with associated confidence levels. The proposed framework can estimate multiple pairs of limits with different nominal coverage probabilities through network optimization. The overall structure of the SNN-based forecasting framework is shown in Fig. 2, where $L\tau\ i$ and $U\tau\ i$ are the lower and upper limits, with τ representing the confidence level.

The probabilistic prediction framework attempts to directly calculate the varying intervals of future wind power with associated confidence levels. Mathematically, given a set of training samples,

$$DS = (\boldsymbol{x}_1, r_1), \ldots, (\boldsymbol{x}_i, r_i), \ldots, (\boldsymbol{x}_{N_S}, r_{N_S}) \tag{12}$$

where DS indicates a wind power dataset; $\boldsymbol{x}_i$ and r_i represent the input variables and future target to forecast; N_S means the number of samples. These samples help us calculate the lower and upper bounds such that the future target r_i is expected to be enclosed by the estimated bounds [$L\tau\ i\ U\tau\ i$], with a coverage probability [5],

$$P\left(r_i \in \left[L_i^{\tau}(x_i) \quad U_i^{\tau}(x_i)\right]\right) = 100(1-\tau)\% \tag{13}$$

where P represent the coverage probability. How to optimize the parameters of SNN is described in detail in the Section 3.

3. Optimal construction of prediction intervals based on group search optimizer

This Section discusses how to construct an optimal prediction interval to achieve probabilistic forecasting of wind power. We describe it from four aspects: optimal probabilistic forecasting structure, group search optimizer, interval optimization and advantages of the proposed probabilistic forecasting framework, as described below.

3.1. Optimal construction of prediction intervals

A new optimization model is originally proposed to optimize the prediction interval of wind power. The optimization model adopts continuous ranking probability score (CRPS) as the objective function. CRPS considers both forecasting reliability and interval sharpness (IS), and is one of the commonly-used index to evaluate probabilistic prediction performance [5]. The optimization model minimizes the objective function by adjusting the SNN weight parameters to achieve probabilistic prediction of wind power. Therefore, statistical inferences and any assumptions are no longer required and the corresponding computational effort can be saved [43]. The proposed optimization model is mathematically described as follows:

$$\min_{w_{ij}} \frac{1}{N_S}\sum\nolimits_{i=1}^{N_S} CRPS_i \tag{14}$$

$$\text{s.t.} \quad L_i^{\tau_j}(x_i) \le U_i^{\tau_j}(x_i) \tag{15}$$

$$L_i^{\tau_j}(x_i) \ge L_i^{\tau_k}(x_i) \quad and \quad U_i^{\tau_j}(x_i) \le U_i^{\tau_k}(x_i), if \quad \tau_j \le \tau_k \tag{16}$$

where $L_i^{\tau_j}(x_i)$ and $U_i^{\tau_j}(x_i)$ represent the lower and upper bounds of the i-th sample under the j-th confidence level, respectively; $CRPS_i$ stands for the CRPS index of the i-th sample and is calculated as follows:

$$CRPS_i = \int_{y=0}^{\infty} \left[CDF_i - H(y - r_i)\right]^2 dy \tag{17}$$

where CDF_i is the cumulative distribution function of the i-th sample; $H\,(y\text{-}r_i)$ represents an indicator function whose value is 0 if $r < y_i$, otherwise, its value is 1. By minimizing the objective function (14), the prediction intervals with a given confidence level can be estimated. GSO is applied in this paper to minimize (14).

3.2. Group search optimizer

The GSO algorithm is developed based on an information sharing model and a producer-scrounger model in nature [59]. It simulates a group of carnivores' foraging strategies. In short, when a group of carnivores goes out for food eat, each member of this group will have different tasks and responsibilities. The GSO algorithm imitates the division of labor of carnivores and their corresponding search strategies. GSO consists of three types of members: producer, scroungers and dispersed members with different missions and roles [60]. Producers aim to perform information integration, search the best resources and send the position of the optimum to other members. Scroungers are designed to approach the producer and evaluate the fitness value of the points on this path. Dispersed members perform a random walk strategy to prevent the algorithm from falling into local optimums. Compared with other population-based heuristic algorithms, the GSO algorithm can achieve self-learning and self-evolution. Therefore, GSO has better performance and is widely-used in economic dispatch and other fields.

In GSO, the group responsible for performing the searching strategy is called a population, and each individual in the population is called a member. All the members perform their respective search strategy by simulating the way a group of carnivores goes out for food. Considering a n-dimensional search space, the i-th member has the current position $Xk\ i \in \boldsymbol{R}^n$ at the k-th iteration, and has a forward angle $\phi k\ i = (\varphi_{i_1}^k, \ldots \varphi_{i_{n-1}}^k) \in \boldsymbol{R}^{n-1}$. It can be seen that the search direction of the i-th member is a unit vector $Dk\ i\,(\phi k\ i) = (d_{i_1}^k, \ldots, d_{i_n}^k) \in R^n$, which can be obtained by applying Cartesian coordinate transformation [61]:

$$d_{i_1}^k = \prod_{q=1}^{n-1} \cos\left(\varphi_{i_q}^k\right) \tag{18}$$

$$d_{i_j}^k = \sin\left(\varphi_{i_{j-1}}^k\right) \prod_{q=1}^{n-1} \cos\left(\varphi_{i_q}^k\right) \tag{19}$$

$$d_{i_n}^k = \sin\left(\varphi_{i_{n-1}}^k\right) \tag{20}$$

At each iteration, each member performs its own search strategy. Specifically, the producer scans at zero degree and then scan laterally by randomly sampling three points in its forward, left and right directions. Given the maximum search angle $\theta_{\max}$ and the maximum searching distance $l_{\max}$, the scanning strategy of producer is presented as follows:

$$X_{front}^k = X_p^k + \lambda_1 l_{\max} D_p^k\left(\phi^k\right) \tag{21}$$

$$X_{right}^k = X_p^k + \lambda_1 l_{\max} D_p^k\left(\phi^k + \lambda_2(\theta_{\max}/2)\right) \tag{22}$$

$$X_{left}^k = X_p^k + \lambda_1 l_{\max} D_p^k\left(\phi^k - \lambda_2(\theta_{\max}/2)\right) \tag{23}$$

where λ_1 represents a random number, and λ_2 stands for a random sequence in the range (0,1); *Xk p* is the producer's position; *Xk front*, *Xk right* and *Xk left* are the positions of the three scanning points.

The current position and the positions of the above three scanning points are compared. The producer will move to a scanning point when this point has a better fitness value than the original point. Otherwise, the producer remains at the same position with an updated search angle:

$$\phi^{k+1} = \phi^k + \lambda_2 \kappa_{\max} \tag{24}$$

where $\kappa_{\max}$ represents the maximum turning angle. After several iterations, the forward direction angle of the producer will be reset to 0° if no point with better fitness is found.

The scroungers' behavior is simply a regional replication. It is a common behavior in carnivores' foraging process. At each iteration, the scroungers move toward the producer in accordance with the established search strategy. In GSO, the search strategy of the *i*-th scrounger in the *k*-th iteration can be expressed as the following formula:

$$X_i^{k+1} = X_i^k + \lambda_3 \circ \left(X_p^k - X_i^k\right) \tag{25}$$

where ° is a vector multiplication.

If a scrounger finds a point with better fitness in its search process, the scrounger becomes the new producer, and the original producer becomes a scrounger in the next iteration.

The dispersed members are evenly distributed within the population. They perform a random walk in each iteration. If the *i*-th individual is selected as the dispersed member in the *k*-th iteration, the forward angle and the search distance are randomly selected according to (26) and (27), respectively. And this dispersed member moves to a new position, as follows:

$$\phi^{k+1} = \phi^k + \lambda_4 \kappa_{\max} \tag{26}$$

$$l_i = \lambda_5 l_{\max} \tag{27}$$

$$X_i^{k+1} = X_i^k + l_i D_i^k\left(\phi^{k+1}\right) \tag{28}$$

where λ_4 and λ_5 denote two random numbers.

The above descriptions are the main steps of the GSO. Obviously, GSO mimics the hunting behavior of carnivores. The optimization procedure is significantly different from other intelligent heuristic algorithms. It does not depend on individual search capabilities, but on the search capabilities of the entire community. This is the cornerstone of the GSO. Therefore, this paper adopts GSO to optimize the proposed prediction interval model (14)–(17).

3.3. GSO based PIs optimization

The proposed wind power probabilistic forecasting method adopts GSO to directly optimize the objective function of SNN to statistically estimate the optimal prediction intervals. The core idea of this method is to approximate the prediction intervals with the best quality through a GSO-based regression procedure while ensuring the reliability and sharpness of the PIs. The flowchart of the proposed probabilistic forecasting method is shown in Fig. 3. It can be seen from Fig. 3 that the proposed wind power probabilistic forecasting method is mainly divided into two parts. The first part is to initialize the SNN weight parameters based on historical samples. The second part is GSO.

In the first part, we create a training dataset based on historical wind power data. Then, by scaling up or down the output of the original dataset, two independent datasets are also generated, as

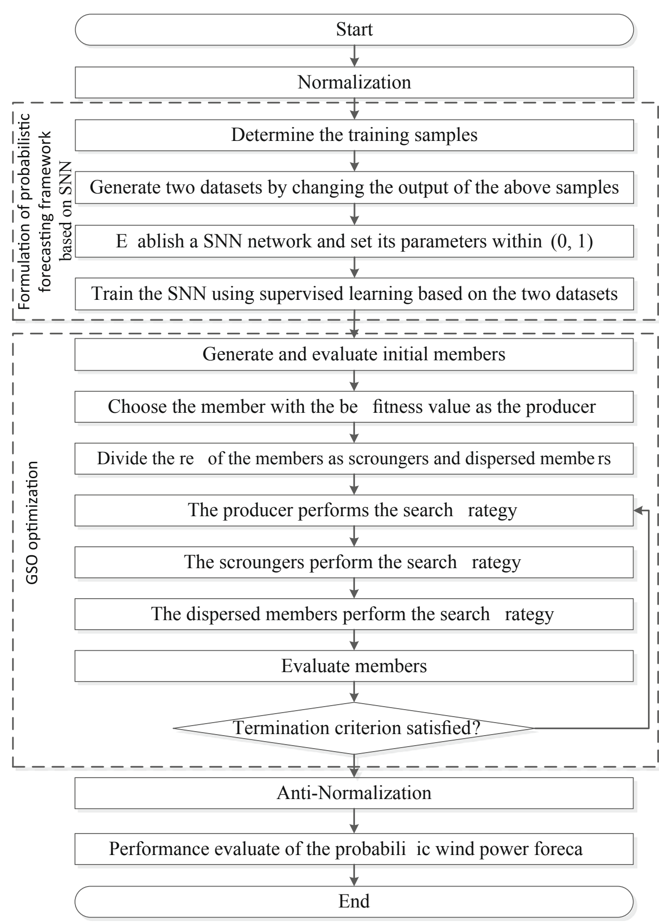

Fig. 3. The flowchart of the proposed probabilistic forecasting framework.

shown in Fig. 3. These two datasets are used to initialize the upper and lower bounds of the prediction intervals. This is because the optimized interval should contain the real output of the original wind power dataset. Afterwards, a SNN model is built. SNN parameters are randomly initialized between 0 and 1, and then well-trained by using the supervised learning method based on the gradient descent algorithm (3)–(11). Note that we use the two modified datasets to train the SNN, making it possible to directly calculate the lower and upper limits of the PIs. However, the obtained upper and lower bounds are artificially set, and are not actually optimal. Therefore, the second part is the use of GSO to optimize the lower and upper bounds of the PIs.

In the GSO part, the number of the members in a group is determined at first. Then, GSO members are initialized. Each member in the group should contain all SNN weight parameters, which are randomly-initialized around their training values. The fitness value of each member is evaluated via (14)–(16) by taking the prediction intervals generated by the SNN as the input. The member with the best fitness value is set as the producer, and the remaining 95% of the members are set as scroungers. The rest members are considered as dispersed members. In each iteration, the producer performs the search strategy in (21)–(23) based on the current angle and the largest search distance $l_{\max}$, and then updates its search angle according to (24). The scroungers independently perform a search strategy according to the location of the producer (25). And the dispersed members perform the strategy in (26)–(28). If any member's fitness value is higher than the producer's fitness value, the member will become the new producer and the original producer will become a new scrounger or dispersed member. Note that the group members need to stay within the given search space in case they exceed the boundary. When a decision variable in a member exceeds its boundary, the boundary value is taken as the new value of this decision variable. The iteration process continues until satisfactory fitness or maximum number of iterations is reached. Finally, the obtained SNN model parameters can be used to formulate the PIs. The pseudo-code of the SNN based probabilistic forecasting framework is given in Fig. 4.

3.4. Advantages of the SNN-based probabilistic forecasting framework

According to the flowchart and pseudo-code in Figs. 3–4, we can conclude that the proposed wind power probabilistic forecasting method has at least four advantages. First, the proposed method can directly obtain the wind power prediction interval. No other statistical inference mechanism in the existing point forecasting methods is required. The second advantage is that the proposed method does not require any quantile analysis and priori-hypothesis of the distribution of prediction error, and therefore has higher prediction accuracy than traditional error-distribution based prediction models. The third advantage is that the proposed method adopts SNN to map the relationship between the input and output. SNN has a better nonlinear mapping capability than traditional neural networks because SNN considers both synaptic states and temporal sequences. In addition, the proposed probabilistic forecasting method is established based on LUBE, which is a general framework known as fast convergence speed [41]. Therefore, the fourth advantage is that the proposed method has a high computational efficiency. Due to these advantages, the proposed direct probabilistic forecasting method exhibits promising prediction performance than traditional prediction methods, as demonstrated in Section 5.

4. Performance criteria

In this paper, average coverage error (ACE) and interval sharpness (IS) are used to evaluate the statistical performance of the

Step 1	Collect the original wind power dataset, and normalize it.
Step 2	Classify the normalized wind power dataset into training samples DS=(x, t)
Step 3	Generate two new training samples $DS^{\pm}$=(x, t×(1±σ%)), where σ is a random number within (0, 100).
Step 4	Construct a SNN based forecasting framework and initialize its parameters within (0, 1).
Step 5	For iteration=1:max_iteration For training_sample=1: number_of_samples The equations (3)-(5) are used to estimate the prediction errors. The equations (6)-(11) are back-propagated to update the weight parameters of SNN. end end
Step 6	Generate GSO members and initialize them.
Step 7	Evaluate these members and choose the best fitness one as the producer.
Step 8	Classify the rest GSO members into scroungers and dispersed members.
Step 9	For iteration=1:max_GSOiteration For member=1:number_of_members If the member is producer Perform the searching strategy (21)-(24). Elseif the member is scrounger Perform the searching strategy (25). Else Perform the searching strategy (26)-(28). End The member with the best fitness is chosen as the producer and the original producer becomes a scrounger or dispersed member. If any decision variable exceeds its boundary the boundary value→the decision variable. end end end
Step 10	Evaluate the prediction performance of the SNN-based framework

Fig. 4. The pseudo-code of the SNN-based forecasting framework.

proposed probabilistic forecasting method. The ACE is a measure of how well the predicted quantiles match the observed quantiles. IS index is used to evaluate the quality of the prediction interval by rewarding narrower PIs and penalizing wider PIs. The ACE and IS are calculated as follows [11]:

$$ACE = \frac{1}{N_S}\sum_{i=1}^{N_S} r_i \times 100\% - \text{PINC} \tag{29}$$

$$IS = \frac{1}{N_S}\sum_{i=1}^{N_S}\begin{cases} -2\tau\delta_i^\tau - 4\left[L_i^\tau - WS_i^\tau\right], & if\ WP_i^\tau < L_i^\tau \\ -2\tau\delta_i^\tau & if\ WP_i^\tau \in \left(L_i^\tau \quad U_i^\tau\right) \\ -2\tau\delta_i^\tau - 4\left[WS_i^\tau - U_i^\tau\right], & if\ WP_i^\tau > U_i^\tau \end{cases} \tag{30}$$

where N_S indicates the number of samples; PINC means prediction interval nominal confidence; The symbol δ_i^τ represents the width of the prediction interval, and can be computed as (U_i^α-L_i^α); WP_i^τ stands for wind power data under the τ-th confidence level; Finally, the indicator r_i is defined as:

$$r_i = \begin{cases} 0, & WP_i^a \in \left(L_i^\tau \quad U_i^\tau\right) \\ 1, & WP_i^a \notin \left(L_i^\tau \quad U_i^\tau\right) \end{cases} \tag{31}$$

The ACE and IS performance of our proposed probabilistic wind power forecasting method are presented in the following Section.

5. Case studies

In this section, two wind power datasets from real wind farms in Belgium and China are adopted to extensively and comprehensively estimate the prediction performance of SNN-based probabilistic forecasting method.

5.1. Investigations on Belgium wind farm

1) Experimental Settings: The wind power data of Belgian wind farm comes from Ref. [62], with North latitude 50°51 and East longitude 4°21. The wind power data in this dataset is collected at a 15-min interval and covers all data from January 2017 to December 2017. During this time period, the installation capacity of wind power increased from 1960.01 MW to 2621.924 MW. We divided the entire wind power dataset into 12 groups and each group covers one month of data. Then, cross validation method is used to train and test the feasibility and effectiveness of the proposed probabilistic forecasting method. A SNN structure with four-layers is designed based on trial-and-error method. It consists of input layer of 20 neurons, two hidden layers and an output layer of one neuron. The numbers of neurons in the hidden layers are 11 and 17, respectively. Subsequently, we divide the whole wind power data into multiple samples according to the established SNN structure. The outputs in all samples are increased by 1% and reduced by 1% as the target for GSO optimization, thus forming two new datasets. These two datasets are used to train the model parameters of the SNN. As soon as the SNN training process is completed, we need to determine the GSO parameters. Since GSO parameters are not sensitive to the optimization results, the settings of GSO parameters can be referred to the guidelines in Ref. [59]. The number of group members is set to 50 and the maximum number of iterations is 2000. In addition, the maximum search angle is set to $\pi/(A_{\max})^2$, where $A_{\max}$ is the maximum search area, and the maximum turning angle is set to one tenth of the maximum search angle.

In order to verify the superiority of the proposed probabilistic forecasting method, the forecasting performance is compared with other benchmarking methods, including traditional SVM, backward propagation (BP) algorithm and ELM. Both the SVM and BP use a Gaussian error hypothesis to create the prediction interval, while the ELM adopts quantile regression to create the probabilistic prediction result. All the prediction algorithms are implemented in MATLAB R2014a and was performed on a personal computer with an Intel(R) Core(TM)-i5-7200 2.5-GHz CPU and 8.00 GB RAM.

2) Forecasting Results: A series of simulations is carried out to verify the feasibility and effectiveness of the proposed method. The wind power is predicted in 15-min ahead because of two reasons. First, the original wind power data is collected at a 15-min interval. Second, wind power probabilistic forecasting results are helpful for dispatcher to perform economic dispatch and contingency analysis, which are executed every 15-min in some reginal power systems, such as China Southern Power Grid. ACE and IS are used as performance evaluation indicators. The PINC is set to 95% or 99% due to the high reliability required for power system operation. Under these two PINCs, the 15-min ahead forecasting performance of the training dataset and testing dataset are tabulated in Tables 1 and 2 in terms of ACE and IS, respectively. In addition, the prediction performances of the three benchmarking algorithms are also listed in the Table. We have drawn several daily interval forecasting results in four seasons, as presented in Figs. 5–8, where each step indicates a 15-min interval. To further demonstrate the benefits of the proposed probabilistic forecasting model, several wind power dataset with different resolutions are also adopted for simulations. The training samples are obtained by using the interval sampling method on the original wind power dataset. The absolute values of ACE and IS indices on the training and testing samples are shown in Figs. 9–11, respectively.

3) Analysis: It can be seen from Table 1 that the ACE of the proposed probabilistic forecasting method is the best among the four benchmarking algorithms. Specifically, the proposed method has an ACE of −0.3 when the PINC is 95%, which is significantly better than that of BP, SVM and ELM algorithms. When the PINC is 99%, the ACE index based on SNN and GSO is −0.15, which is also superior to the other three benchmarking algorithms. Therefore, these numerical results show that the ACE obtained from the proposed probabilistic forecasting method is closer to the corresponding nominal confidence level, indicating that the proposed method has better forecasting reliability. In addition, Table 1 also shows that the proposed method has the best IS index. The statistical results indicate that

Table 1
The ACE and IS of four methods over the training samples.

PINC	Index	Proposed	BP	SVM	ELM
95%	ACE	**−0.30**	−3.83	−2.36	−2.09
	IS	**−2.87**	−7.51	−5.45	−5.23
99%	ACE	**−0.15**	−3.02	−2.15	−1.94
	IS	**−1.38**	−3.49	−2.04	−1.92

Table 2
The ACE and IS of four methods over the testing samples.

PINC	Index	Proposed	BP	SVM	ELM
95%	ACE	**−0.65**	−4.72	−3.13	−3.00
	IS	**−4.39**	−8.67	−6.28	−5.74
99%	ACE	**−0.23**	−3.35	−3.01	−2.49
	IS	**−1.43**	−4.20	−2.94	−2.51

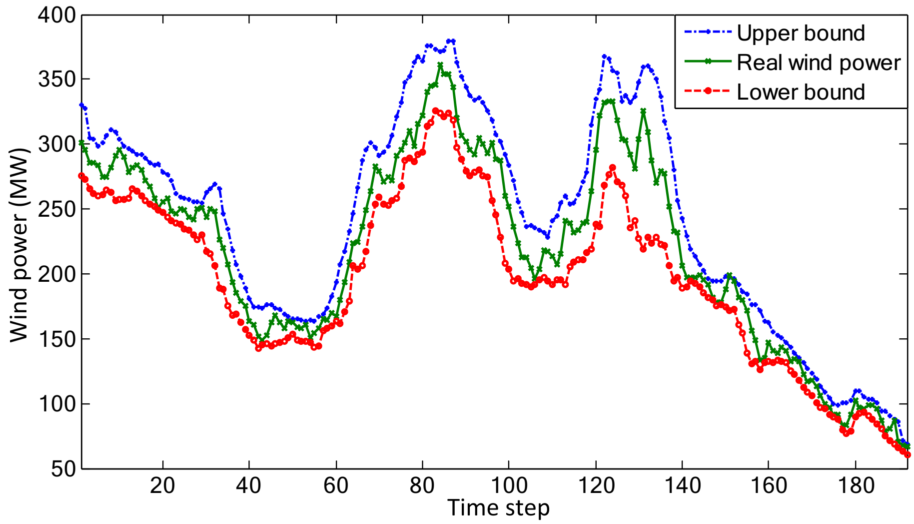

Fig. 5. The interval forecasting results of two given days in spring.

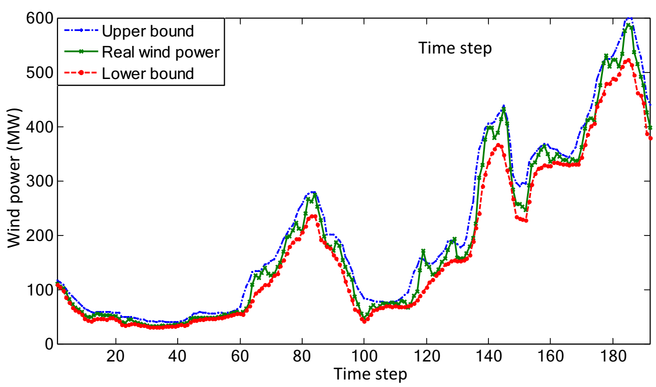

Fig. 6. The interval forecasting results of two given days in summer.

the IS indices of the proposed method are −2.87 and −1.38 when PINCs are 95% and 99%, respectively. The IS indices on these two training samples from the SNN-based forecasting model are also significantly better than that of the other three benchmarking algorithms. These statistical results of IS show

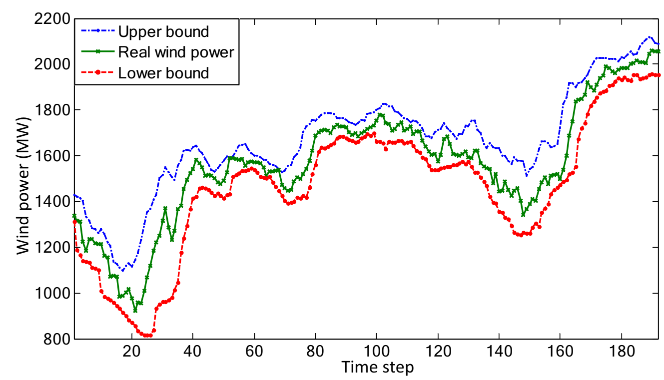

Fig. 8. The interval forecasting results of two given days in winter.

that the proposed method has better forecasting capability in terms of interval sharpness and is therefore more attractive than the other three benchmark algorithms.

As can be seen from Table 2, the proposed forecasting method

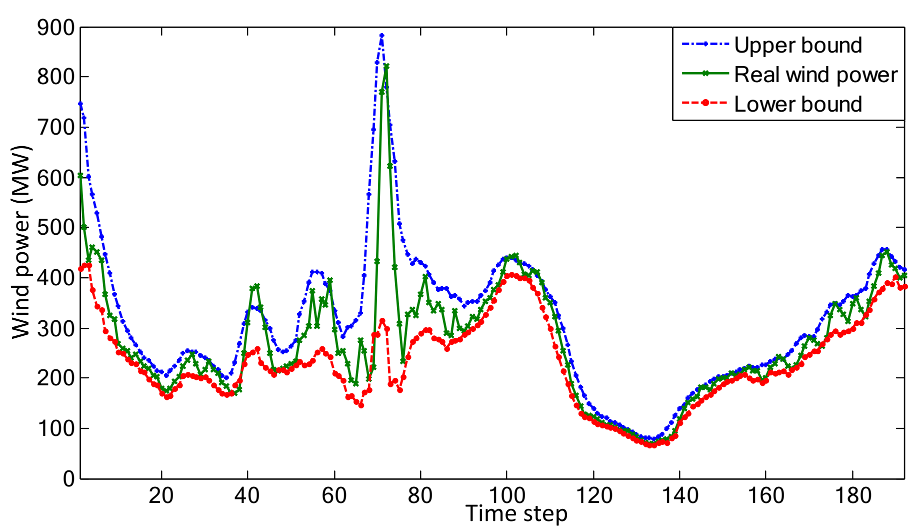

Fig. 7. The interval forecasting results of two given days in autumn.

has better ACE and IS performance on the testing samples, when compared to the three benchmark algorithms. In addition, the performance indices in Table 2 are worse than the indices in Table 1 because the indices in Table 2 are from the testing samples, while the results in Table 1 are from the training samples. Figs. 5–8 show that the real wind power curves are mostly within the lower and upper bounds of the constructed PIs. It can also be seen that the variation trends of the upper boundary line, the lower boundary line and the real wind power curve are basically the same. The high coverage of the real wind power curve and the line similarity demonstrate that the proposed probabilistic forecasting method has satisfactory performance. In addition, it is clear that the prediction interval is wider when the wind power curve varies greatly, and the prediction interval is narrower when the wind power curve is relatively smooth. For example, the upper bound curve of the prediction interval in Fig. 6 almost saturates due to the smoothness of the real wind power curve in summer. This shows that the proposed method is very flexible and can adapt to different shapes of wind power curves.

Figs. 9–11 show that the proposed method exhibits better performance in terms of ACE and IS at different data resolutions.

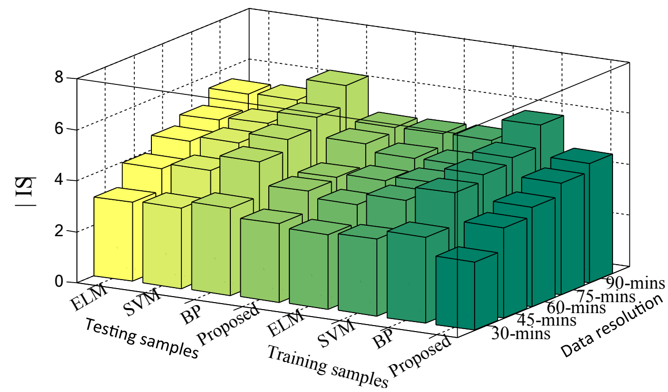

Fig. 11. IS performance on the training and testing samples.

Specifically, the ACE bias of the proposed probabilistic forecasting method is around 0.5 when the wind power data is with a 30-min resolution. The ACE bias increases as the data resolution decreases. In addition, the interval sharpness index is about 3 or so when the

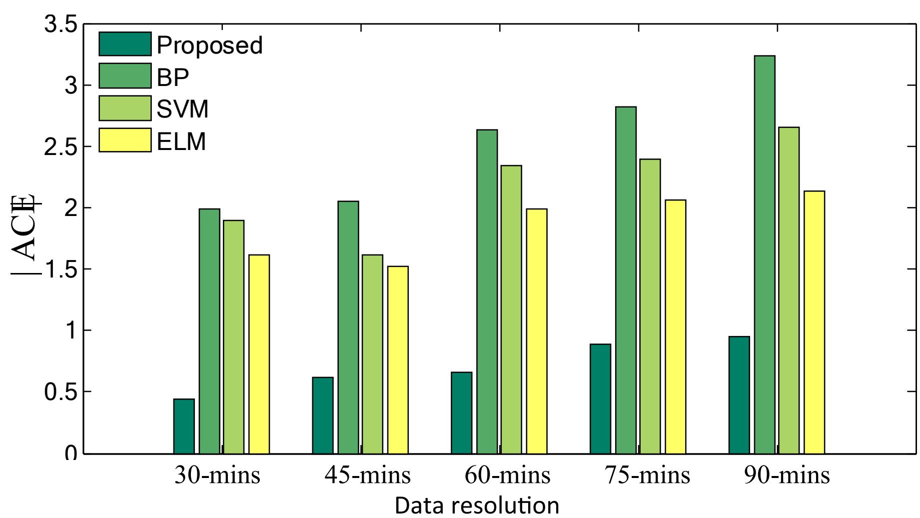

Fig. 9. ACE deviation on the training samples using different data resolutions.

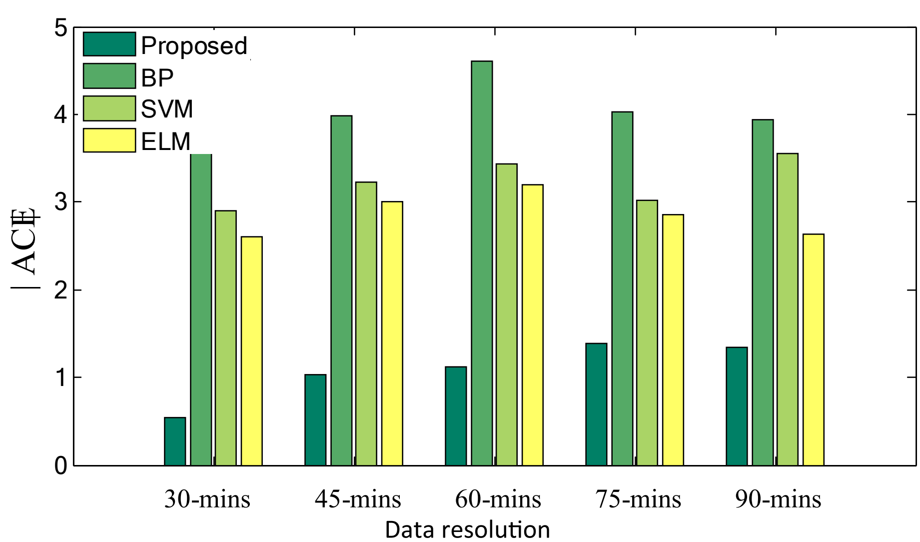

Fig. 10. ACE deviation on the testing samples using different data resolutions.

wind power is with a 30-min resolution. The value of interval sharpness is basically doubled when the wind power data is with a 90-min resolution. This is because large data resolution means that the wind power data accuracy is degraded. In Figs. 9–11, it can also be concluded that the prediction performance on the training samples is better than that of the testing samples. The proposed method performs better than the other three benchmark algorithms. These two conclusions are consistent with the conclusions from Tables 1 and 2 Moreover, it is clear that the proposed SNN-based probabilistic forecasting method can be used on several wind power datasets with different data resolutions. In other words, the proposed forecasting method can be applied on wind power dataset with any time interval. This can be explained by the fact that neural network only processes the size of the input data and cannot identify the time interval of the input data. Moreover, Tables 1 and 2 and Figs. 5–11 also show that the variation trend in ACE bias appears to be irregular and confusing.

ACE and IS are two commonly used probabilistic forecasting indices to assess the effectiveness of wind power probabilistic predictions. Based on the forecasting results obtained here, we can see that the proposed probabilistic prediction method has higher prediction reliability and interval sharpness, demonstrating that the proposed method is superior to the three benchmarking algorithms. Therefore, the SNN-based prediction method has a high potential in practical applications.

5.2. Investigations on Chinese wind farms

1) Experimental Settings: The wind power data used in this subsection is collected from Shangchuan Island wind farm (SIWF). The SIWF has a rated capacity of 48.45 MW and contains 57 wind turbines with a capacity of 0.85 MW each. The original wind power dataset covers the period from January 2013 to December 2013 with a data resolution of 15 min. SIWF is located on the southern coast of China. This area has a tropical maritime climate so that the wind power data exhibits strong fluctuations and seasonal differences. Therefore, we divide the wind power data over the whole year into four groups, each of which covers one quarter of wind power data. The SNN structure and GSO parameter settings are similar to these described in 5.1. BP, SVM and ELM are adopted as the benchmarking algorithms. In addition, the standard LUBE probabilistic forecasting method is also chosen for performance comparison because the SNN-based probabilistic prediction method is basically a variants of LUBE.
2) Forecasting Results: Cross-validation method is used here to train and test the proposed method. Tables 3–6 show the seasonal ACEs and ISs of the proposed probabilistic forecasting method with 95% PINC. In addition, different confidence levels indicate different reliabilities for wind power prediction within the estimated interval. In power system, the prediction intervals with different confidence levels indicate that the dispatchers have different preferences for risk management. Therefore, it is necessary to calculate the prediction intervals under multiple confidence levels at the same time. The ACE and IS indices are given in Table 7.
3) Analysis: Considering the testing samples in Tables 3–6, the minimum and maximum ACE values of the proposed method are −0.92 and −0.77, respectively. And the minimum and maximum values of IS are −2.71 and −2.37, respectively. The ACE indices of the BP algorithm vary a lot between −3.18 and −1.83, and the IS indices vary between −4.81 and −3.07. In addition, the SVM algorithm exhibits an ACE from −1.58 to −1.20 and an IS from −3.53 to −3.01. The forecasting performance of the ELM is better than that of BP and SVM. Moreover, the ACE index of LUBE has a minimal of −1.14 and a maximum of −0.87. And the IS of LUBE varies slightly from −2.40 to −2.93. It can be seen that the proposed forecasting method has the best prediction reliability and interval sharpness among the five benchmarking algorithms. While the BP algorithm has the worst prediction performance. Obviously, the prediction performances of spring and winter are better than that of summer and autumn. This is because SIWF is

Table 3
The ACE and IS performance in Spring.

Algorithms	Training samples		Testing samples	
	ACE	IS	ACE	IS
Proposed	**−0.43**	**−1.83**	**−0.79**	**−2.45**
BP	−1.78	−3.05	−2.02	−3.72
SVM	−1.01	−2.39	−1.33	−3.07
ELM	−0.92	−2.36	−1.20	−2.94
LUBE	−0.54	−2.01	−0.87	−2.58

Table 4
The ACE and IS performance in Summer.

Algorithms	Training samples		Testing samples	
	ACE	IS	ACE	IS
Proposed	**−0.48**	**−1.97**	**−0.92**	**−2.58**
BP	−2.13	−3.64	−2.37	−4.32
SVM	−1.10	−2.89	−1.45	−3.33
ELM	−1.00	−2.73	−1.40	−3.28
LUBE	−0.55	−2.23	−1.14	−2.79

Table 5
The ACE and IS performance in Autumn.

Algorithms	Training samples		Testing samples	
	ACE	IS	ACE	IS
Proposed	**−0.55**	**−2.20**	**−0.87**	**−2.71**
BP	−2.44	−3.65	−3.18	−4.81
SVM	−1.24	−2.86	−1.58	−3.53
ELM	−1.12	−2.37	−1.55	−3.29
LUBE	−0.71	−2.28	−0.99	−2.93

Table 6
The ACE and IS performance in Winter.

Algorithms	Training samples		Testing samples	
	ACE	IS	ACE	IS
Proposed	**−0.42**	**−1.73**	**−0.77**	**−2.37**
BP	−1.70	−3.15	−1.83	−3.17
SVM	−0.97	−2.19	−1.20	−3.01
ELM	−0.95	−1.99	−1.05	−2.47
LUBE	−0.54	−1.81	−0.89	−2.40

Table 7
The ACE and IS performance under different confidence levels.

PINC	Indices	Proposed	BP	SVM	ELM	LUBE
85%	ACE	**−2.45**	−4.63	−3.17	−3.20	−2.83
	IS	**−7.07**	−17.10	−11.58	−10.52	−8.06
90%	ACE	**−1.04**	−2.70	−1.49	−1.43	−1.14
	IS	**−4.57**	−5.99	−5.76	−5.61	−4.90
95%	ACE	**−0.84**	−2.35	−1.39	−1.30	−0.97
	IS	**−2.53**	−4.01	−3.24	−3.00	−2.68
99%	ACE	**−0.25**	−1.41	−0.95	−0.84	−0.45
	IS	**−1.31**	−2.52	−1.66	−1.57	−1.49

located in a tropical maritime climate, and its wind power data has greater volatility in summer and autumn, so it is less likely to be predicted. Tables 3–6 also show that the proposed algorithm performs well in the training samples and testing samples in all four seasons, indicating that the proposed algorithm has high stability and strong robustness.

In Table 7, the proposed algorithm has an ACE variation range of −2.45 to −0.25 and an IS variation range of −7.07 to −1.31. The minimum and maximum ACE values of the BP are −4.63 and −1.41, respectively. The minimum and maximum IS of BP are −17.10 and −2.52, respectively. In addition, the prediction performances of SVM and ELM are almost the same. Their ACEs vary between −0.84 and −3.20, and ISs vary between −1.57 and −11.58, respectively. The ACE average of LUBE is −1.35 and the IS average is −4.28. It is clear that the proposed probabilistic forecasting method has the best prediction reliability and interval sharpness, and the LUBE is the second best. These results demonstrate that the proposed method not only significantly improves the prediction performance, but also shows high stability and strong robustness, further proving the superiority and high potential of the proposed method for practical applications.

6. Discussion

The simulation results in Tables 1–7 and Figs. 3–11 show that the average of the probability coverage of the proposed method is improved by 72.0%, 54.9% and 51.3% respectively, when compared to traditional BPNN, SVM and ELM. The corresponding improvement rates of sharpness index are 43.1%, 28.1% and 21.0%, respectively. The SNN-based forecasting method performs a slightly better than LUBE. Four aspects explain why the SNN-based probabilistic forecasting method has relatively better performance. First, the proposed method is a direct probabilistic forecasting method that does not depend on any prediction error distribution hypothesis such as Gaussian distribution and inaccurate statistical inference such as quantile regression. Second, this paper adopts SNN to map the nonlinear relationship between statistical input and output of wind power samples. The neurons in SNN mimic synaptic behavior and temporal change in real neuron in nature, and thus have more accurate nonlinear mapping capability than traditional neurons. The mapping capability provides an efficient way to extract the intrinsic features and invariant structures in wind power data. Third, this paper adopts GSO to optimize the parameters of SNN model. GSO is a heuristic algorithm with powerful global search capabilities. GSO's search capabilities can help SNN model parameters find their global optimal values, thereby reducing ACE bias and sharpness. At last, this paper proposes a new optimization model based on CRPS that is a score quantifying the difference between a continuous probability distribution and a deterministic observation sample. CRPS considers both prediction reliability and interval sharpness simultaneously. Therefore, the SNN based probabilistic prediction method performs the best among the five benchmarking algorithms.

In order to evaluate the computational efficiency of the proposed method based on SNN and GSO, we collect its training time, as shown in Table 8. It can be seen that the training time of the proposed prediction model is 1.31s when there are 15,000 samples. The time it took to train 30,000 samples is 2.56s. Obviously, the probabilistic prediction model proposed in this paper does not require high computing cost. Therefore, the high computing efficiency also shows that the proposed prediction model has a high potential in practical applications.

Table 8
The training time of the SNN-based forecasting method.

Number of Samples	Training Time
15,000	1.31s
30,000	2.56s

The above numerical results and statistical analysis demonstrate that the proposed probabilistic forecasting method has higher prediction accuracy, stability and computing efficiency. More accurate probabilistic forecasting method can effectively reduce the uncertainty in wind power, and thus plays an important role in the operation schedule and real-time control of electric power and energy systems. For example, probabilistic forecasting results can be used to defend the cyber-attack of electric energy systems [34]. First, the prediction interval of each renewable generator needs to be calculated, and the variation interval of energy system state variables is then estimated based on the worst-case analysis. Finally, a cyber-attack detection method based on "OR" operation is proposed to realize the active defense of the power system against false data injection attack. For another example, the study in Ref. [5] shows that wind power probabilistic forecasting results can be used to improve the scheduling efficiency of power market operators. More specifically, wind power uncertainty can be used to adjust the day-ahead scheduling results in real time, thereby reducing the reserve capacity of the generators. This idea is validated in Illinois power system simulation model. Therefore, from these application scenarios, it is clear that wind power probabilistic forecasting method with high accuracy has high potential for practical applications.

7. Conclusions

This paper proposes a novel wind power probabilistic forecasting method based on spiking neural network and group search optimizer. Spiking neural network is used to construct the statistical relationship in historical wind power data. Group search optimizer is applied to optimize the parameters in spiking neural network to achieve the purpose of directly obtaining the prediction interval. The prediction results of the proposed method are extensively compared with four commonly used algorithms, namely back propagation algorithm, support vector machine, extreme learning machine and lower upper bound estimation. The prediction results obtained from different seasons, different prediction horizons and different wind power datasets show that the proposed probabilistic forecasting method is superior to the four benchmarking algorithms in terms of prediction reliability, prediction accuracy and interval sharpness. Therefore, the proposed method has better prediction performance and thus can accurately estimate the uncertainty exhibited in wind power. The superiority of the proposed probabilistic forecasting method is mainly attributed to the spiking structure and the group search optimizer. The former can effectively extract the high-level features and deep invariant structures in the wind power data, and the latter can reduce the prediction bias of probability coverage and interval sharpness. Therefore, the probabilistic forecasting method based on spiking neural network is attractive in future application scenarios of electric power and energy systems.

Acknowledgements

This work was jointly supported by the National Natural Science Foundation of China (51707123) and the Foundations of Shenzhen Science and Technology Committee (JCYJ20180305125407996, GJHZ20180928160212241).

References

[1] Wang H, Ruan J, Zhou B, et al. Dynamic data injection attack detection of cyber physical power systems with uncertainties. IEEE Transactions on Industrial Informatics 2019;15(10):5505–18.
[2] Lyu Xue, Jia Youwei, Xu Zhao, Østergaard Jacob. Mileage-responsive Wind Power Smoothing. arXiv:1807.05780.
[3] Jia Y, Lyu X, Lai C, et al. A retroactive approach to microgrid real-time scheduling in quest of perfect dispatch solution. J Mod Power Syst Clean Energy 2019;7:1608–18.
[4] Wang Yun, Xie Zongxia, Hu Qinghua, Xiong Shenghua. Correlation aware multi-step ahead wind speed forecasting with heteroscedastic multi-kernel learning. Energy Convers Manag 2018;163:384–406.
[5] Wang Huai-zhi, Gang-qiang Li, Wang Gui-bin, Peng Jian-chun, Jiang Hui, Yi-tao Liu. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl Energy 2017;188:56–70.
[6] Zhao Xinyu, Liu Jinfu, Yu Daren, Chang Juntao. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data. Energy Convers Manag 2018;164:560–9.
[7] Carlos González-Mingueza, Francisco Muñoz-Gutiérrez. RETRACTED: wind prediction using Weather Research Forecasting model (WRF): a case study in Peru. Energy Convers Manag 2014;81:363–73.
[8] Wang Han, Han Shuang, Liu Yongqian, Yan Jie, Li Li. Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. Appl Energy 2019;237:1–10.
[9] Yang Tsun-Hua, Chin-Cheng Tsai. Using numerical weather model outputs to forecast wind gusts during typhoons. J Wind Eng Ind Aerod 2019;188:247–59.
[10] Allen DJ, Tomlin AS, Bale CSE, Skea A, Vosper S, Gallani ML. A boundary layer scaling technique for estimating near-surface wind energy using numerical weather prediction and wind map data. Appl Energy 2017;208:1246–57.
[11] Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 2016;182:80–93.
[12] Aasim Singh SN, Abheejeet Mohapatra. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renew Energy 2019;136:758–68.
[13] Nathaniel S, Lukas G. Statistical approach for improved wind speed forecasting for wind power production. Sustainable Energy Technol Assessments 2018;27:180–91.
[14] Lydia M, Suresh Kumar S, Immanuel Selvakumar A, Edwin Prem Kumar G. Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Convers Manag 2016;112:115–24.
[15] Wang Yun, Wang Haibo, Srinivasan Dipti, Hu Qinghua. Robust functional regression for wind speed forecasting based on Sparse Bayesian learning. Renew Energy 2019;132:43–60.
[16] Yang Dazhi. On post-processing day-ahead NWP forecasts using Kalman filtering. Sol Energy 2019;182:179–81.
[17] Wang Yun, Wang Jianzhou, Xiang Wei. A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: a case study of wind farms in northwest China. Energy 2015;91:556–72.
[18] Wu Lifeng, Gao Xiaohui, Xiao Yanli, Yang Yingjie, Chen Xiangnan. Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China. Energy 2018;157:327–35.
[19] Natapol Korprasertsak, Thananchai Leephakpreeda. Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models. Energy 2019;180:387–97.
[20] Wu Junli, Zhang Buhan, Li Hang, Li Zhongcheng, Yi Chen, Miao Xiaogang. Statistical distribution for wind power forecast error and its application to determine optimal size of energy storage system. Int J Electr Power Energy Syst 2014;55:100–7.
[21] Khodayar M, Wang J, Manthouri M. Interval deep generative neural network for wind speed forecasting. IEEE Transactions on Smart Grid 2019;(99). 1-1.
[22] Liu Hui, Xi-wei Mi, Yan-fei Li. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manag 2018;156:498–514.
[23] Sun Wei, Wang Yuwei. Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network. Energy Convers Manag 2018;157:1–12.
[24] Yu Chuanjin, Li Yongle, Bao Yulong, Tang Haojun, Zhai Guanghao. A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. Energy Convers Manag 2018;178:137–45.
[25] Ali Mumtaz, Prasad Ramendra. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. Renew Sustain Energy Rev 2019;104:281–95.
[26] Amir Sharifian, Jabbari M, Sahand Ghavidel, Li Li, Zhang Jiangfeng. A new method based on Type-2 fuzzy neural network for accurate wind power forecasting under uncertain data. Renew Energy 2018;120:220–30.
[27] Hong Ying-Yi, Rioflorido P. A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. Appl Energy 2019;250:530–9.
[28] Ding Min, Zhou Hao, Xie Hua, Wu Min, Nakanishi Yosuke, Yokoyama Ryuichi. A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting. Neurocomputing 2019;365:54–61.
[29] Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 2014;34:243–54.
[30] Jiang Ping, Yang Hufang, Heng Jiani. A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. Appl Energy 2019;235:786–801.
[31] He Qingqing, Wang Jianzhou, Lu Haiyan. A hybrid system for short-term wind speed forecasting. Appl Energy 2018;226:756–71.
[32] Wang Huaizhi, Lei Zhenxing, Zhang Xian, Zhou Bin, Peng Jianchun. A review of deep learning for renewable energy forecasting. Energy Convers Manag 2019;198:111799.
[33] Zheng Qian, Yan Pei, Hamidreza Zareipour, Chen Niya. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl Energy 2019;235:939–53.
[34] Wang H, Ruan J, Wang G, et al. Deep learning based interval state estimation of AC smart grids against sparse cyber attacks. IEEE Transactions on Industrial Informatics 2018;14(11):4766–78.
[35] Kim Deockho, Jin Hur. Short-term probabilistic forecasting of wind energy resources using the enhanced ensemble method. Energy 2018;157:211–26.
[36] Ulas Eminoglu, Omer Turksoy. Power curve modeling for wind turbine systems: a comparison study. Int J Ambient Energy 2019:1–22 (just-accepted).
[37] Lydia M, Suresh S, Immanuel Selvakumar A, Edwin Prem Kumar G. A comprehensive review on wind turbine power curve modeling techniques. Renew Sustain Energy Rev 2014;30:452–60.
[38] Zhang Yao, Wang Jianxue, Wang Xifan. Review on probabilistic forecasting of wind power generation. Renew Sustain Energy Rev 2014;32:255–70.
[39] Erotokritos Xydas, Meysam Qadrdan, Charalampos Marmaras, Liana Cipcigan, Jenkins Nick, Hossein Ameli. Probabilistic wind power forecasting and its application in the scheduling of gas-fired generators. Appl Energy 2017;192:382–94.
[40] Wan C, Xu Z, Pinson P, Dong ZY, Wang KP. Optimal prediction intervals of wind power generation. IEEE Trans Power Syst 2014;29(3):1166–74.
[41] Khosravi A, Nahavandi S, Creighton D, Atiya AF. Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans Neural Network 2011;22(3):337–46.
[42] Wan C, Xu Z, Pinson P. Direct interval forecasting of wind power. IEEE Trans Power Syst 2013;28(4):4877–8.
[43] Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Transactions on Neural Networks and Learning Systems 2014;25(2):303–15.
[44] Kavousi-Fard A, Khosravi A, Nahavandi S. A new fuzzy-based combined prediction interval for wind power forecasting. IEEE Trans Power Syst 2016;31(1):18–26.
[45] Shi Z, Liang H, Dinavahi V. Direct interval forecast of uncertain wind power based on recurrent neural networks. IEEE Transactions on Sustainable Energy 2018;9(3):1177–87.
[46] Wang Huaizhi, Lei Zhenxing, Liu Yang, Peng Jianchun, Liu Jing. Echo state network based ensemble approach for wind power forecasting. Energy Convers Manag 2019;201:112188.
[47] Zhang Yao, Wang Jianxue, Wang Xifan. Review on probabilistic forecasting of wind power generation. Renew Sustain Energy Rev 2014;32:255–70.
[48] Javier Pérez, Cabrera Juan A, Castillo Juan J, Velasco Juan M. Bio-inspired spiking neural network for nonlinear systems control. Neural Network 2018;104:15–25.
[49] Koschwitz D, Frisch J, Treeck C van. Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX Recurrent Neural Network: a comparative study on district scale. Energy 2018;165(A):134–42.
[50] Bahadir Kasap, Opstal A, John van. Dynamic parallelism for synaptic updating in GPU-accelerated spiking neural network simulations. Neurocomputing 2018;302:55–65.
[51] Han Fengyang, Li Runmin, Qian Dianwei. Short-term wind speed forecasting model based on spiking neural network. In: International cConference on advanced Mechatronic systems (ICAMechS) 30 Aug.-2 Sept; 2018. p. 359–63.
[52] Madhiarasan M, Deepa SN. Long-term wind speed forecasting using spiking neural network optimized by improved modified grey wolf optimization algorithm. Int J Adv Res 2016;4:356–68.
[53] Sebastian Brusca, Giacomo Capizzi, Grazia Lo Sciuto, Gianluca Susi. A new design methodology to predict wind farm energy production by means of a spiking neural network–based system. Int J Numerical Modelling July 2017;32(4):e2267.
[54] Qian D, Ma Y. Spiking neural network-based wind speed forecasting Design & implement. Publisher: LAP Lambert Academic Publishing; 2019. May 13.
[55] Amirhossein Tavanaei, Masoud Ghodrati, Saeed Reza Kheradpisheh, Timothée Masquelier, Anthony Maida. Deep learning in spiking neural networks. Neural Network 2019;111:47–63.
[56] Kim Jaehyun, Kim Heesu, Huh Subin, Lee Jinho, Choi Kiyoung. Deep neural networks with weighted spikes. Neurocomputing 2018;311:373–86.
[57] Kulkarni Shruti R, Bipin Rajendran. Spiking neural networks for handwritten digit recognition—supervised learning and network optimization. Neural Network 2018;103:118–27.
[58] Voutsas K, Adamy J. A biologically inspired spiking neural network for sound

source lateralization. IEEE Trans Neural Network 2007;18(6):1785–99.
[59] Jiang He, Dong Yao, Xiao Ling. A multi-stage intelligent approach based on an ensemble of two-way interaction model for forecasting the global horizontal radiation of India. Energy Convers Manag 2017;137:142–54.
[60] Mehrdad Tarafdar Hagh, Saeed Teimourzadeh, Manijeh Alipour, Parinaz Aliasghary. Improved group search optimization method for solving CHPED in large scale power systems. Energy Convers Manag 2014;80:446–56.
[61] Mehrdad Tarafdar Hagh, Manijeh Alipour, Saeed Teimourzadeh. Application of HGSO to security based optimal placement and parameter setting of UPFC. Energy Convers Manag 2014;86:873–85.
[62] http://www.elia.be/en/grid-data.