Research Article

A New Wind Power Forecasting Approach Based on Conjugated Gradient Neural Network

Tian Li, Yongqian Li, Mingwei Liao, Weikang Wang, and Chujie Zeng

North China Electric Power University, Baoding 071000, China

Correspondence should be addressed to Tian Li; ncepulitian@126.com

Prediction of the output power of wind plants is of great significance for running a power system comprising large amount of wind generators. According to the prediction results, it is possible to determine the quotas of power generation in power generators and distribute resources in a scientific and reasonable way. In the past, the Grey Neural Network was widely applied in predicting wind power while it could hardly meet the engineering requirements due to the structure of ANN. The problem of slow convergence speed and large amount of iterations, especially in case of large scale data, would pose challenges to power prediction and the sensitivity of automatic control. This paper optimizes ANN model by applying conjugate gradient descent and creating Conjugated Gradient Neural Network (CGNN) in weights updating process. Experiments performed on different scale datasets have proved that the performance of CGNN improves substantially as the average iterations decreased by almost 90% without the sacrifice of prediction accuracy.

1. Introduction

In the context global crisis in energy ecology, developing clean energy has become an unavoidable path for China and even the whole world. With the development of technology and the decrease of cost, wind power has evolved as one of the rapidly developing clean energy sources. China is a leading nation in wind power development and Chinese wind market accounts for approximately 26.3% of the global total [1]. In order to distribute the workload of wind generators, it is significant to predict wind power for connecting wind plants with power grids. Currently, there are two ways to predict the output power of wind generators. One is to predict wind power output according to forecasted environment elements, such as wind speed and air humidity. The other one is to directly predict power of the next few hours or days based on historical power data. China has implemented a considerable amount of researches on wind power generation with some methods proposed to forecast the power of a wind power station, including Persistence Method [2], Kalman Filtering Method [3], ARMA [4–6], ANN [7–11], Fuzzy Logic [12], Grey Theory [13, 14], Expert System, Wavelet Decomposition

[15, 16], SVM [17, 18], Nearest Neighbor Analysis, and Spatial Correlation Models [13, 17, 19].

Kalman Filtering Method is a common method for prediction. It constructs a state space model with wind speed as a state variable. With the establishment of the model, it is possible to predict power by Kalman Filtering Algorithm. Meanwhile, the result is based on the assumption that the statistic characteristics of noise are given, but in reality it is hard to estimate the statistic characteristics. Stochastic Time Serial Method utilizes a large amount of historical data to build a prediction model with its mature and simple theory. It applies to short-term prediction but works not well in long-term prediction [16, 17, 20]. Grey Theory can resolve some uncertain problems by studying a small amount of data and information. In contrast, $GM(1, 1)$ needs only little data; then it can be used in predicting wind speed in a short time period. However, it does not work well when faced with catastrophe points [16, 20, 21]. As for Spatial Correlation Method, since a lot of things need to be considered, it needs large scale datasets. The accuracy of this method is acceptable, but the method is still on development [16, 17, 20]. The fluctuation of wind power caused by the uncertainty and intermittence

of wind can bring increasing impact on power grids. In this case, it challenges the quality and stability of power as well as the control of power generation [16, 17, 22–24]. Then it is more than necessary to develop a precise and fast wind power prediction approach.

ANN has been used in forecasting due to its ability of approaching nonlinear mapping [8–11]. Recently, some models combining ANN and other methods have been developed to increase the accuracy of prediction. Reference [25] hybridizes the Fifth Generation Mesoscale Model (MM5) with ANN to forecast wind speed in the next 48 hours. It [26] uses the Moving Window and ANN to forecast the linear and nonlinear part of the data from financial market, respectively. In addition, it [27] proposes to use a Neural Network-Markov Chain (MK) model to forecast the second-scale and hour-scale wind speed. In detail, MK is applied to forecast the long-term wind speed while ANN is applied to forecast the short-term wind speed. It [28] also combines the coral reefs optimization (CRO) and Extreme Learning Machine (ELM) approach to perform short-term wind speed forecasting. The result shows that the ELM approach has greatly reduced the computation cost. The CRO-ELM model converges only after 35 times iteration. Reference [29] optimizes the work in [28] by using CRO and the Harmony Search Operator (HS). The error of CRO-HS-ELM model is 10% less than that of CRO-ELM model. In addition to ANN, there are some alternative methods for predicting wind speed. For instance, [30] recommends using empirical mode decomposition (EMD) and ANN to predict the wind speed. The EMD model is used to divide the input data into smaller pieces which are used as the input of ANN to forecast the partial results. Finally, these partial results are integrated to get the final result of wind speed forecasting. Reference [31] puts forward a more complex model of decomposition forecasting. It uses the fast ensemble empirical model decomposition to divide the data into raw pieces. Then, it applies the Multilayer Perception Artificial Neural Network (optimized by Genetic Algorithms and Mind Evolutionary Algorithm) to forecast the partial results which are integrated to form the final forecasting results.

There are also some researches aiming at comparing the performance of ANNs used in forecast. Reference [32] compares the performance of three different types of ANNs including Feed Forward Back Propagation (FFBP), Racial Basis Function (RBF), and Adaptive Linear Element (ADA-LINE) neural networks. In hourly wind speed forecasting, the RBF network realizes the lowest forecasting error and the fastest convergence speed while BP network performs the worst in the wind speed forecast. Deepest Descent Method is usually applied to update weights in traditional BP neural network. But it leads to the problem that the time of network training is too long to satisfy the need of automatic control.

This paper incorporates the Conjugated Gradient Descent Method into BP neural network to get Conjugated Gradient Neural Network (CGNN). In this way, the convergence speed can essentially increase by almost 90%. In addition, since it can avoid falling into local maxima, it improves the accuracy and reliability of neural network. In part I of this paper, the structure and the algorithm
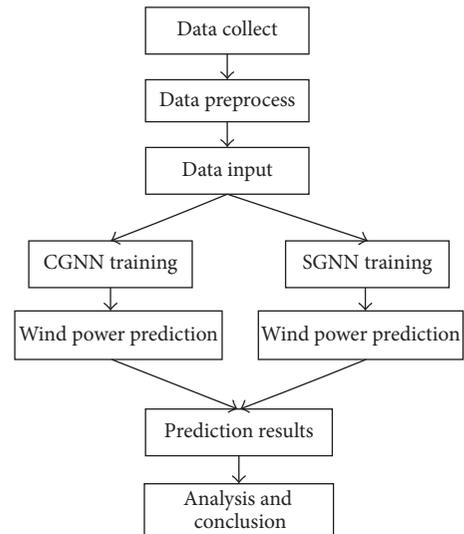
Figure 1: Flow chart of forecasting process.

procedure of wind power prediction system are introduced. As for part II, ANN and Conjugated Gradient Method are analyzed, following which the modeling procedure of wind power prediction is presented. Finally, the test data and the comparison on the results of several ANNs including CGNN, SGNN, RBF, and ELM are discussed.

2. Wind Power Prediction System Construction

In reality, there are a lot of factors that can affect output power of wind power generators while each of them has different effects. On this condition, predicting wind power by utilizing linear regression algorithm can hardly acquire a good result. One of current research trends is applying nonlinear algorithms to predict wind power. Neural network is capable of approximating nonlinear functions via updating matrices between layers. This paper uses the excellent nonlinear mapping ability of ANN to predict wind power. To optimize ANN, it is feasible to improve its inner structure and the accuracy of prediction improves greatly. The exact ANN applied is Back Propagation Neural Network with the cost function calculated as the sum of the squared errors. The optimization algorithm used in this paper is Conjugated Gradient Descent Method which can substantially improve the convergence speed and accuracy of ANN. Figure 1 shows the algorithm procedure.

3. Introduction of Algorithms

3.1. Artificial Neural Network. Neural network is able to approach nonlinear mapping in any requirement of precision and dig up unknown information from data. In terms of the structure, neural network is distributed in data storage and computing. Therefore, systems constructed by neural network possess considerable robustness and the ability of solving difficult problems. Steepest Gradient Method is a

---

(1) Take the initial point $x^{(0)} \in R^n$
(2) **if** $\nabla f(x^{(0)}) = 0$, set $x^{(0)}$ as supposed point
   else calculate the Hesse Matrix A of $f(x)$
   end if
(3) Calculate the initial conjugated vector by
   $p^{(0)} = -\nabla f(x^{(0)})$
(4) Find a proper λ to minimize
   $f(x^{(0)} + \lambda p^{(0)})$
(5) Calculate the best step by
   $\lambda_0 = \dfrac{-\nabla f\left(x^{(0)}\right)' p^{(0)}}{p^{(0)'} A p^{(0)}}$
   And $x^{(1)} = x^{(0)} + \lambda_0 p^{(0)}$
(6) **if** $\nabla f(x^{(1)}) = 0$
   set $x^{(1)}$ as the supposed point
   else if $\nabla f(x^{(1)}) \neq 0$ and $\nabla f(x^{(1)}) p^{(0)} = 0$
   search a vector $p^{(1)}$ in the sub space of
   $\nabla f(x^{(1)})$ and $p^{(0)}$, so that $p^{(1)} A p^{(0)} = 0$
   end if
(7) Set $p^{(1)} = -\nabla f(x^{(1)}) + \alpha p^{(0)}$
(8) Then $\alpha = \nabla f(x^{(1)})' A p^{(0)} / p^{(0)'} A p^{(0)}$
(9) Let $x^{(i)} = x^{(i-1)} + \lambda_{i-1} p^{(i-1)}$, calculate λ_i
(10) **if** $\nabla f(x^{(i)}) = 0$
   set $x^{(i)}$ as the supposed point
   else
   let $p^{(i)} = -\nabla f(x^{(i)}) + \alpha p^{(i-1)}$ and
   $p^{(i)} A p^{(i-1)} = 0$
   then $\alpha_{i-1} = \nabla f(x^{(i)})' A p^{(i-1)} / p^{(i-1)} A p^{(i-1)}$.
(11) Loop (3).

ALGORITHM 1: Conjugate gradient descent steps.

common optimization method to update weights in present artificial neural network system. Steepest Gradient Method converges fast in local maxima conditions. However, it mostly performs badly when faced with global maxima problems. Therefore, it is hard to get global maxima with this method. This paper introduces BP neural network based on Conjugate Gradient Method and why it applies this method instead of Steepest Gradient Method in weights updating. The reason is that it resolves the problem of slow convergence speed caused by Steepest Gradient Method and converges well in global maxima condition in the ANN system.

3.2. Conjugated Gradient Descent Method. Conjugated Gradient Descent (CGD) Method is a medium method between Steepest Gradient Descent (SGD) Method and Newton Method. With the first-order derivative, it can not only reach convergence quickly but also avoid the disadvantage of Newton Method in calculating Hesse Matrix and its inverse matrix. CGD is not only the most effective way to solve the massive equation group, but also the most valuable way to optimize the equation group. CGD requires relatively small storage space and it has the advantage of step convergence, which makes it stable. Furthermore, CGD does not require any outer parameters. It is shown in Algorithm 1.

4. Theoretical Model

4.1. Factor Selection. In reality, there are a lot of factors affecting wind power. According to [8], wind power can be represented by the following:

$$P = \frac{C_p A \rho v^3}{2}, \tag{1}$$

where P is the output power of wind turbines; C_p is rotor power coefficient; ρ is atmospheric density; A is the swept area of rotor; and v is wind speed. Among them, wind speed is the most influential factor [9].

Also, because the wind fleet usually adopts array structure, the wake effect will occur after the wind power is absorbed by the front turbines. Because of the wake effect, the output power of wind power stations will decrease at specific wind directions [8]. Therefore, wind direction is also an affecting factor.

According to (1), ρ is an important affecting factor as well and it has positive correlation with the output power. Meanwhile, atmospheric density is influenced by temperature, humidity, and atmospheric pressure. So these factors need to be concerned.

To sum up, seven factors affecting wind power can be represented as F. F is defined as follows:

$$F = \left(v_t, \cos\theta, \sin\theta, P_t, h_t, T_{\min}, T_{\max}\right), \tag{2}$$

where v_t is the average measured value of wind speed during time slot t; $\cos\theta$ and $\sin\theta$ are the sine and cosine of wind direction; P_t, h_t, $T_{\min}$, and $T_{\max}$ are the average measured value of atmospheric pressure, humidity, minimum temperature, and maximum temperature during time slot t.

4.2. Data Collection. As explained in previous parts, factors that affect the output power of wind turbines include wind speed, wind direction, atmospheric pressure, humidity, and temperature, the data of which need to be collected. This paper collects these data of a wind fleet for two months from 6 am to 12 am a day with the interval of 15 minutes. Among the collected data, there are seven affecting factors for corresponding wind power. With additional timestamp functioning as index, 8-tuple D is formed and saved in the database. D is defined as follows:

$$D = \left(v_t, \cos\theta, \sin\theta, P_t, h_t, T_{\min}, T_{\max}, ts\right), \tag{3}$$

where ts is the timestamp.

Note that the collected value might be null due to the instability of data acquisition equipment and the intermittence of wind fleet, so operators need to remove these "dirty" data.

4.3. Artificial Neural Network Model Construction. In order to judge the accuracy and convergence speed of Conjugate Gradient Neural Network precisely, this paper constructs two different Neural Network Models. One is based on Steepest Gradient Method and the other is based on Conjugate Gradient Method.
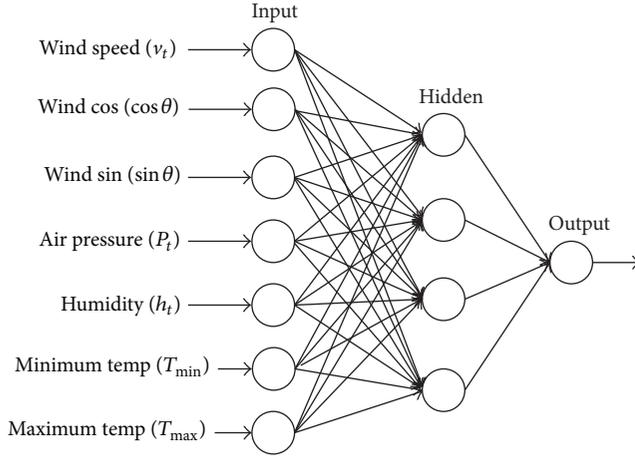
Figure 2: The structure of CGNN.

First of all, wind power data is featured with nonlinearity and randomness. Thus this paper adopts BPNN as the basic structure because of its excellent performance in nonlinear mapping, generalization, and fault tolerance. Secondly, input and output parameters should be clarified before determining the number of nodes in input and output layer. The number of input nodes is determined by affecting factors. Meanwhile, since this paper selects seven different factors, the number of input nodes will be seven. The number of output nodes is one because the output power is the only thing to predict. The data of affecting factors is used as input data and the target output is the data of corresponding wind power. The cost function is the sum of the squared errors between output value and target output value. Finally, empirical equation (4) can be used to determine the number of intermediate layers:

$$p = \sqrt{m + n} + \alpha, \tag{4}$$

where p is the number of intermediate layer nodes and m and n are the number of input nodes and the number of output nodes. Usually α is a regulatory factor, which is obtained by training the sample data with neural network. According to (4), this paper comprehensively considers the trade-off between accuracy and convergence speed and determines that p is equal to four. Finally, the CGNN model is represented as in Figure 2.

4.4. Training Strategy. In order to get rid of the influence of physical dimension and speed up convergence process, data normalization should be performed before training NN model. In this paper, Feature Scaling Method is adapted to bring all value into the range [0, 1]. Consider

$$x_i^k = \frac{x_i^k - x_{\min}^k}{x_{\max}^k - x_{\min}^k}. \tag{5}$$

In the equation, x_i^k is the value of feature k in ith tuple; $x_{\max}$ is the maximum of feature k; and $x_{\min}$ is the minimum.

After normalization is performed, training procedure can be started. The cost function is defined as follows:

$$E = \frac{1}{2} \sum_{t=1}^{Q} \sum_{k=1}^{l} \left(d_k(t) - o_k(t) \right)^2, \tag{6}$$

where Q is sample size; l is the number of output nodes; $d_k(t)$ is the target output value of node k for sample t; and $o_k(t)$ is the real output value.

Afterwards, Conjugate Gradient Method can be performed to get acceptable weights in the model.

The predicted steps are shown as follows:

(1) Standardize the data according to (5).

(2) Update the weight of ANN.

(3) Compare the target fault rate and the real fault rate. If the real one is larger than the target one, then loop (2).

(4) Forecast the relative value of wind power.

(5) Destandardize the values.

(6) Calculate the fault and then analyze it.

5. Samples and Analysis

The data for this paper comes from two wind farms located in Inner Mongolia, China, during January 2014 and March 2014. These wind farms have total installed capacity of 1,500 MW. As mentioned above, the model collects data from 6 am to 12 am with interval of 15 minutes. The data in January is used to train the network and the trained model is used for predicting data in March.

In order to make the advantages of Conjugate Gradient Method clearer and more intuitive, we compared Conjugated Gradient Neural Network (CGNN), Steepest Gradient Neural Network (SGNN), Racial Basis Function Neural Network (RBFNN), and Extreme Learning Machine (ELM) in the aspect of time and accuracy. On the other hand, to prove the robustness of the model, this paper uses data from both wind farm I and wind farm II. The test is performed on these two types of datasets separately. Due to the performance limit of data acquisition equipment and the instability of wind fleet, there is some empty data within the test sum.

To eliminate the disturbances, we removed the empty data and got 131 groups of mid-term data from wind farm I and 548 groups of long-term data from wind farm II. The details of the neural networks used in the forecasting are shown in Table 1.

5.1. Accuracy Analysis. In order to improve accuracy, we should figure out the causes of errors. As for predicting wind power, one controllable error is from model structure. Another one is the measurement error, which may have a strong impact on prediction, but it depends on the devices. So this error is not discussed in this paper. The last one is the instability of wind fleet, which means the output power may differ even in the same condition. It is represented by robustness comparison among different networks.

In this paper, the comparison on long-term power forecasting results of different networks is demonstrated in

TABLE 1: Model of different networks.

Neural network	Nodes	Data volume
CGNN	$7 \times 4 \times 1$	548×8
SGNN	$7 \times 4 \times 1$	548×8
RBFNN	$7 \times 4 \times 1$	548×8
ELM	$7 \times 4 \times 1$	548×8
CGNN	$7 \times 4 \times 1$	131×8
SGNN	$7 \times 4 \times 1$	131×8
RBFNN	$7 \times 4 \times 1$	131×8
ELM	$7 \times 4 \times 1$	131×8
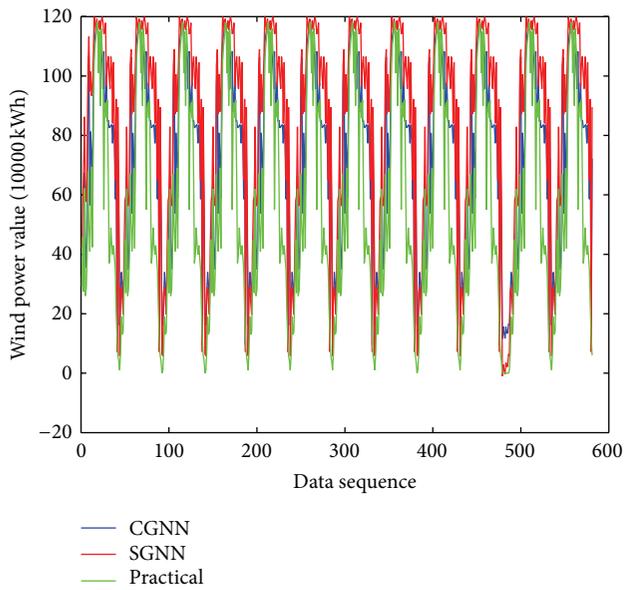
FIGURE 3: The prediction results of CGNN and SGNN with long-term data.

FIGURE 4: The prediction results of CGNN and RBF with long-term data.
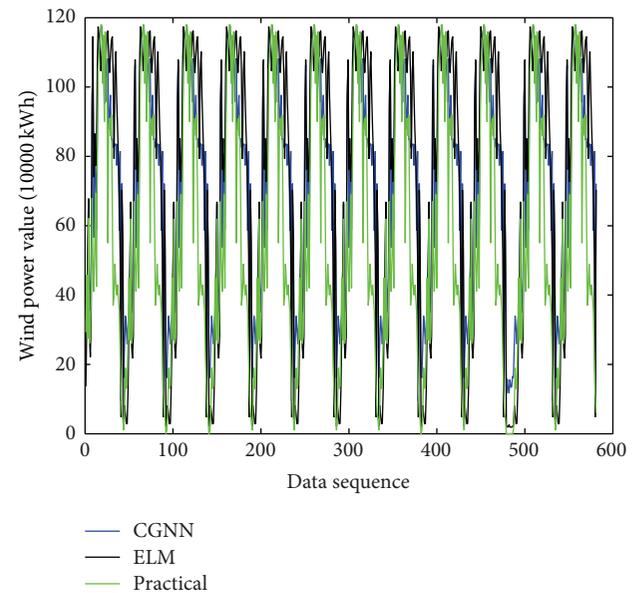
FIGURE 5: The prediction results of CGNN and ELM with long-term data.

Figures 3, 4, and 5. The comparison on mid-term forecasting results is presented in Figures 6, 7, and 8, respectively. In Figures 3 and 4, the green curve represents practical output power; the blue curve is for the prediction result of CGNN; the red curve indicates the result of SGNN; the yellow curve shows the result of RBF; the black curve refers to the result of ELM.

According to Figure 3, the blue curve (CGNN) is close to the green curve (practical) when the power is between 50 MW and 1200 MW. In contrast, the red curve (SGNN) approaches black curve only in case the power is lower than 50 MW and has rather larger error or when the power is between 700 MW and 900 MW.

Similarly, Figure 4 shows that the yellow curve (RBF) is near green curve only if the power is lower than 50 MW and has rather larger error when the power is between 700 MW and 1100 MW. Finally, Figure 5 presents the fact that the black curve (ELM) comes near green curve only in the event that the power is lower than 40 MW and has rather larger error if the power is between 40 MW and 1100 MW.

Figure 4 shows when trained with mid-term data, the blue curve is almost overlapping with the green curve except

for some minor disturbances which are probably due to the shutting down of wind generators, since the model will not identify these occurrences. Figure 4 also presents that the performance of CGNN is better than that of SGNN as a result of the less departure of predicted result on the practical data. Similarly, Figure 5 shows that the yellow curve (RBF) presents greater departure when the practical power output is 40 MW–60 MW and 70 MW–90 MW while the curve of CGNN overlaps the practical output curve. Figure 6 shows that the black curve (ELM) reveals greater departure in 0 MW–20 MW and 100 MW–120 MW.
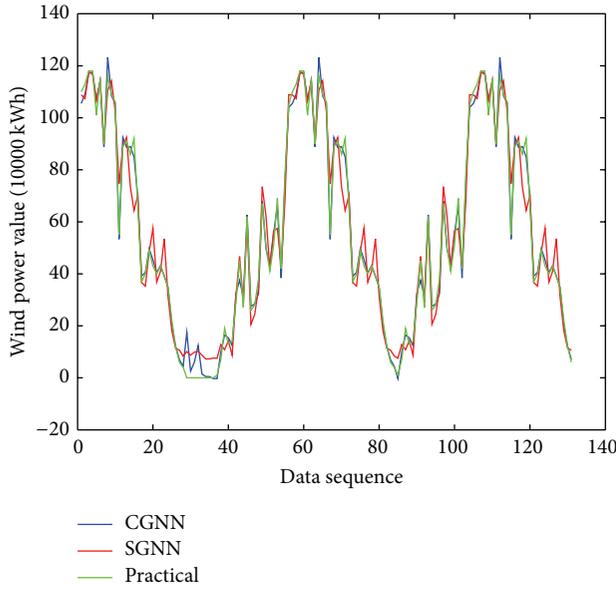
FIGURE 6: The prediction results of CGNN and SGNN with mid-term data.
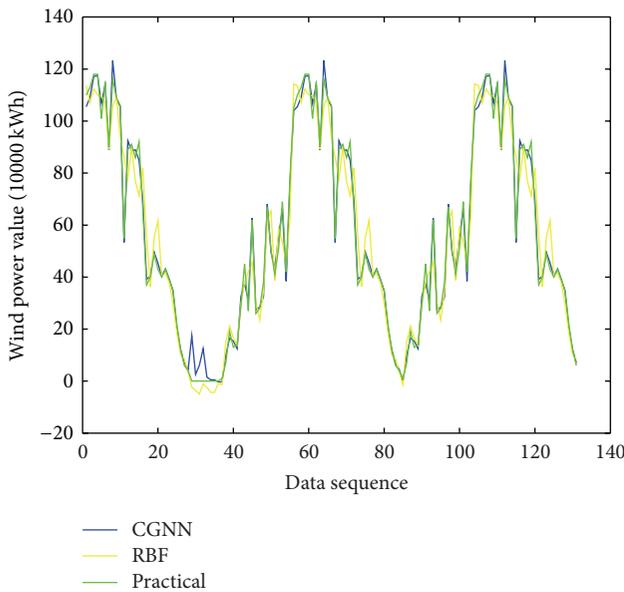
FIGURE 7: The prediction results of CGNN and RBF with mid-term data.

This paper also demonstrates the long-term and mid-term numerical prediction errors of different networks in Table 2, respectively. In the table, the maximum (Max), the minimum (Min), the average (Aver), and the sum error (Sum) of mid-term and long-term data are shown. The result indicates that the fault rate of CGNN is much less than that of SGNN.

5.2. Iteration and Convergence Analysis. In an autocontrol system, the time spent on prediction should be minimized as long as the accuracy is acceptable. In engineering situation,

TABLE 2: The faults of different networks.

Type	Term	Min (MW)	Max (MW)	Aver (MW)	Sum (MW)
CGNN	*Mid*	*0.23*	*27.62*	*2.54*	*283.95*
SGNN	Mid	0.53	17.45	5.27	590.47
RBF	Mid	0.48	31.08	7.07	791.67
ELM	Mid	0.18	46.46	7.95	889.81
CGNN	*Long*	*0.40*	*66.18*	*19.03*	*10714.14*
SGNN	Long	0.43	83.52	28.56	16084.44
RBF	Long	1.41	60.77	19.15	10786.64
ELM	Long	0.34	67.20	23.65	13319.70
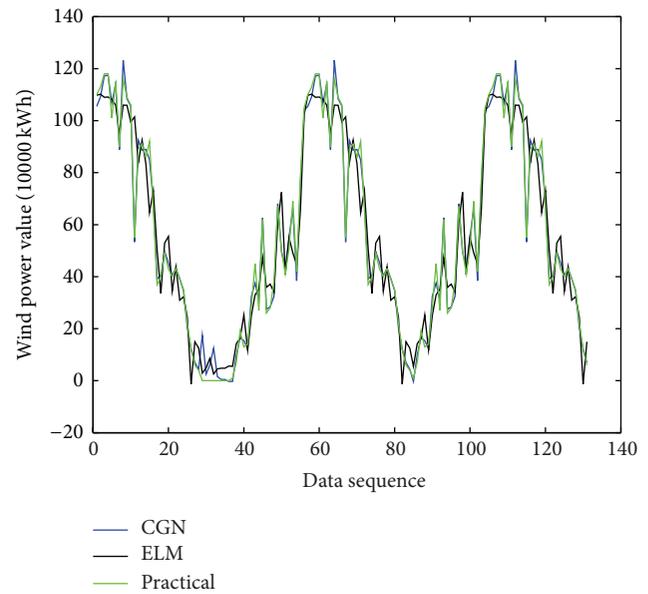
FIGURE 8: The prediction faults of CGNN and ELM with mid-term data.

ANN usually works dynamically, which means it needs to be trained every time the prediction is performed. The training process will go through many epochs until the model converges. In general, neural networks do different work in each epoch due to different structures. To measure the time consumption of the neural network, we predefined an error limit, Mean Squared Error, MSE, so that the totally consumed training time is measured from the beginning of the algorithms (input the first row of data) to the end of the algorithms (reach the predefined limit). Table 3 shows the training time consumption of four different networks in long-term and mid-term prediction when MSE equals 0.004. Table 4 shows the training time consumption of four different networks in long-term and mid-term prediction when MSE equals 0.002.

Table 3 shows that, in case MSE equals 0.004, the training time consumption of CGNN is the shortest while that of SGNN is the longest either in short-term or in long-term forecast. The training time consumption of CGNN is even less than that of ELM (29 s to 49 s and 12 s to 27 s). Table 4 indicates that even if MSE is changed to 0.002, the training time consumption of CGNN is still the shortest while that

TABLE 3: The training time consumption of four models when MSE equals 0.004.

Type	Term	MSE	Time (s)
CGNN	*Long*	*0.004*	*29*
SGNN	Long	0.004	70
RBF	Long	0.004	57
ELM	Long	0.004	49
CGNN	*Mid*	*0.004*	*12*
SGNN	Mid	0.004	43
RBF	Mid	0.004	36
ELM	Mid	0.004	27

TABLE 4: The training time consumption of four models when MSE equals 0.002.

Type	Term	MSE	Time (s)
CGNN	*Long*	*0.002*	*158*
SGNN	Long	0.002	686
RBF	Long	0.002	592
ELM	Long	0.002	361
CGNN	*Mid*	*0.002*	*68*
SGNN	Mid	0.002	387
RBF	Mid	0.002	294
ELM	Mid	0.002	176

of SGNN is the longest. The training time consumption of CGNN is still less than that of ELM (158 s to 361 s and 68 s to 176 s). Due to the structures of some models, the training time consumption does not change linearly. Given that the MSE is changed from 0.004 to 0.002, the training time consumption of other networks is nearly ten times of those with MSE equaling 0.004. If the MSE of CGNN decreases from 0.004 to 0.002, the training time consumption increases by no more than five times after Conjugate Gradient Method is adopted in neural network. So it is reasonable to draw a conclusion that Conjugate Gradient Method demonstrates irreplaceable strengths in optimizing ANN.

6. Conclusion

Prediction of wind power is of great significance for connecting wind plants with power grids. This paper optimizes neural network model with Conjugate Gradient Method and makes run-time control of wind power possible. According to the experiment result, Conjugate Gradient Method is capable of improving the convergence speed and prediction accuracy of neural networks dramatically at the same time. In mid-term prediction, CGNN performs best when fault, Max fault, Aver fault, and the Sum fault of CGNN are 0.23 MW, 27.62 MW, 2.54 MW, and 283.95 MW, respectively. When it comes to long-term prediction, CGNN performs best in case Min fault, Max fault, Aver fault, and the Sum fault are 0.40 MW, 66.18 MW, 19.03 MW, and 10714.14 MW, respectively. The results show that the model deviation decreases after Conjugated Gradient Descent (CGD) is adopted in ANN model.

Besides, the training time consumption of it is the least when compared with other three different neural networks. When MSE is set as 0.002, CGNN converges after 158 s and 68 s in long-term and mid-term prediction, respectively. In long-term prediction, the time consumption of CGNN is 76.97%, 73.31%, and 56.23% less than SGNN, RBF, and ELM separately. As for mid-term prediction, the time consumption of CGNN is 82.42%, 76.87%, and 61.36% less than SGNN, RBF, and ELM, respectively. When MSE is set as 0.004, CGNN converges after 29 s and 12 s in long-term and mid-term prediction, respectively. In long-term prediction, the time consumption of CGNN is 58.57%, 49.12%, and 40.81% less than SGNN, RBF, and ELM. In mid-term prediction, the time consumption of CGNN is 72.09%, 66.67%, and 55.56% less than SGNN, RBF, and ELM.

The results prove that CGNN performs well in either long-term or mid-term power prediction of wind output. The training time consumption of the CGNN is relatively less than those of SGNN, RBF, and ELM. Generally speaking, this paper paves the way for the coming research concerning autocontrol of wind power connected with power grid.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

References

[1] Global Wind Energy Council, Global wind report, annual market market update[EB/OL], 2011, http://www.gwec.net/index.php?id=190&L=0.

[2] G. Wu, Y. Xiao, and S. Weng, "Discussion aboutshort-term forcast of wind speed on wind farm," *Jilin Electric Power*, vol. 6, pp. 21–24, 2005.

[3] S. Han, Y. Yang, and Y. Liu, "Application study of the methods in wind speed prediction," *Journal of North China Electric Power University*, vol. 35, no. 3, pp. 57–61, 2005.

[4] E. A. Bossanyi, "Short-term wind prediction using Kalman filters," *Wind Engineering*, vol. 9, no. 1, pp. 1–8, 1985.

[5] L. Kamal and Y. Z. Jafri, "Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan," *Solar Energy*, vol. 61, no. 1, pp. 23–32, 1997.

[6] P. Guo, "Wind speed prediction research based on double direction regression moving average," *Modern Electricity*, vol. 26, no. 6, pp. 66–69, 2009.

[7] J. L. Torres, A. García, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Solar Energy*, vol. 79, no. 1, pp. 65–77, 2005.

[8] G. Fan, W. Wang, C. Liu et al., "Wind power prediction based on artificial neural network," *Proceedings of the Chinese Society for Electrical Engineering*, vol. 28, no. 34, pp. 118–123, 2008.

[9] Y. Meng, J. Lu, H. Sun, X. Pan, D. Gao, and Y. Liao, "Short-term wind power forecasting based on similar days and artificial neural network," *Power System Technology*, vol. 34, no. 12, pp. 163–167, 2010.

[10] H. Shi, J. Yang, M. Ding, and J. Wang, "A short-term wind power prediction method based on wavelet decomposition and BP neural network," *Automation of Electric Power Systems*, vol. 35, no. 16, pp. 44–48, 2011.

[11] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models," *IEEE Transactions on Energy Conversion*, vol. 21, no. 1, pp. 273–284, 2006.

[12] J.-F. Li, B.-H. Zhang, G.-L. Xie, Y. Li, and C.-X. Mao, "Grey predictor models for wind speed-wind power prediction," *Power System Protection and Control*, vol. 38, no. 19, pp. 151–160, 2010.

[13] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama, "Improved Grey predictor rolling models for wind power prediction," *IET Generation, Transmission and Distribution*, vol. 1, no. 6, pp. 928–937, 2007.

[14] Y. Du, J. Lu, Q. Li et al., "Short term wind speed forecasting of wind farm based on least square-support vector machine," *Power System Technology*, vol. 32, no. 15, pp. 62–66, 2008.

[15] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, no. 6, pp. 939–947, 2004.

[16] A. Ye, *Research on Application of Wind Power Prediction Using Gray BP Neural Network*, Lanzhou Jiaotong University, 2013.

[17] D. Pan, H. Liu, and Y. Li, "A wind speed forcasting optimization model for wind farms based on time series analysis and kalmanfilter algorithm," *Power System Technology*, vol. 32, no. 7, pp. 82–86, 2008.

[18] J. Han, "Review on prediction method of wind power," *China Electric Power*, pp. 143–144, 2013.

[19] M. C. Alexiadis, P. S. Dokopoulos, and H. S. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models," *IEEE Transactions on Energy Conversion*, vol. 14, no. 3, pp. 836–842, 1999.

[20] L. Li and L. Ye, "Short-term wind power forecasting based on an improved persistence approach," *Transactions of the Chinese Society of Agricultural Engineering*, vol. 26, no. 12, pp. 182–187, 2010.

[21] L. Cao and R. Li, "Short-term wind speed forecasting model for wind farm based on wavelet decomposition," in *Proceedings of the 3rd International Conference on Deregulation and Restructuring and Power Technologies (DRPT '08)*, pp. 2525–2529, Nanjng, China, April 2008.

[22] J. Huang, *Wind Power Forecasting Based on Wavelet Analysis and Support Vector Machine (SVM)*, Shandong University, 2012.

[23] Y. Zhang, *Optimize the Design of BP Neural Network Based on Improved Artificial Fish Swarm Algorithm*, South China University of Technology, 2012.

[24] W. Ji, *The Learning Methods Research of Evolutionary Computation Optimizing Feed forward Neural Network*, Northeast Forestry University, 2013.

[25] S. Salcedo-Sanz, A. M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and D. Paredes, "Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction," *Renewable Energy*, vol. 34, no. 6, pp. 1451–1457, 2009.

[26] M. Khashei, M. Bijari, and G. A. Raissi-Ardali, "Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs)," *Neurocomputing*, vol. 72, no. 4–6, pp. 956–967, 2009.

[27] S. A. Pourmousavi Kani and M. M. Ardehali, "Very short-term wind speed prediction: a new artificial neural network-Markov chain model," *Energy Conversion and Management*, vol. 52, no. 1, pp. 738–745, 2011.

[28] S. Salcedo-Sanz, A. Pastor-Sánchez, L. Prieto, A. Blanco-Aguilera, and R. García-Herrera, "Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization—extreme learning machine approach," *Energy Conversion and Management*, vol. 87, pp. 10–18, 2014.

[29] S. Salcedo-Sanz, A. Pastor-Sánchez, J. Del Ser, L. Prieto, and Z. W. Geem, "A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction," *Renewable Energy*, vol. 75, pp. 93–101, 2015.

[30] H. Liu, C. Chen, H.-Q. Tian, and Y.-F. Li, "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks," *Renewable Energy*, vol. 48, pp. 545–556, 2012.

[31] H. Liu, H. Tian, X. Liang, and Y. Li, "New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks," *Renewable Energy*, vol. 83, pp. 1066–1075, 2015.

[32] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Applied Energy*, vol. 87, no. 7, pp. 2313–2320, 2010.