Research Article

Wind Power Prediction Investigation

[1]Yuanlong Liu, [2]Yuanbiao Zhang and [2]Ziyue Chen
[1]Department of Electrical and Information Engineering,
[2]Department of Packaging Engineering, Mathematical Modeling Innovative
Practice Base, Jinan University, Zhuhai 519070, China

Abstract: Daily and real-time forecast data of wind power is predicted in this study using three methods, which are Kalman filter model, GARCH model and time-series-based BP neural network model. Then, owing to evaluation to the calculation of accuracy and qualification rate, the best method, the time-series-based BP neural network model, was selected for its highest accuracy. Moreover, the prediction error influence due to convergence of wind turbine is on consideration according to the evaluation. Finally, suggestions of improving the prediction accuracy were put forward based on the discussion of accuracy-obstacle factors.

Keywords: BP neutral network, Kalman filter, time sequence prediction, wind-power prediction

INTRODUCTION

Recently, the main source of wind power generation is the close-to-ground wind, whose volatility, intermittency and low energy density leads to the instability of wind power. The large instability would bring side effects on power balance and frequency regulation of network when a large scale wind farm connects to the public network. Electric power dispatching department could make a schedule to ensure power balance and operation safety, if the wind power could be predicted. As a result, the precise prediction for power in wind farms is a pivotal problem remains settled.

According to the mathematical model, wind power prediction method can be divided into continuous prediction method, Kalman filter model, the ARMA model and intelligence method, etc. Continuous prediction method (Brown *et al.*, 1984) is the most simple prediction model, it assumes that wind speed is equal to the sliding average of the values of the last few wind prediction speed (Lexiadis *et al*., 1998), but its prediction results are not stable and precise. There are improved method, like Kalman filter (Bossanyi, 1985), ARMA model (Boone, 2005) and time sequence method combined with Kalman filter. There are also some intelligence methods, such as artificial neural network method (Kariniotakis *et al*., 1996), etc.

In the centralized exploitation mode, the main mode existed of wind power generation, the power of each wind turbine gathers in the wind farm before connecting to the network. The gather of several wind turbines might change the wave characteristics, influencing the prediction error further. The relationship of prediction error between the power of one wind turbine and the total power of several wind turbines remain description, which lead to much more researches needed for the influence that the gathering of wind turbines has on wind power prediction.

WIND POWER PREDICTION

Due to the different demand in operation mode by electric power dispatching department, prediction of wind power could be divided into two types, daily prediction and real-time prediction. In daily predictions, the 96 time points for wind power of the next 24 h could be forecasted while 16 time points for the following 4 h. The interval time of two adjacent time points in both predictions is 15 min.

There is a wind farm containing 58 wind turbines, whose rated power output is 850 kW. In this study, wind power of 4 wind turbines (P_A, P_B, P_C, P_D) and the total power (P_{58}) were predicted by both daily prediction and real-time prediction according to the data from May 10th to June 6th in 2006.

Real-time prediction of wind power based on Kalman filter: For prediction with Kalman filter, prediction recursion equation ought to be deducted using orthogonality theorem by mathematical induction. The following Eq. (1) is the final prediction recursion equation (Bossanyi, 1985):

Corresponding Author: Yuanlong Liu, Department of Electrical and Information Engineering, Jinan University, Zhuhai 519070, China

$$
\begin{cases}
\hat{X}(k+1|k+1) = \Phi(k+1,k)\hat{X}(k|k) + K(k+1) \\
\quad [Z(k+1) - H(k+1)\Phi(k+1,k)\hat{X}(k|k)] \\
K(k+1) = P(k+1,k)H^T(k+1)[H(k+1) \\
\quad P(k+1|k)H^T(k+1) + R(k+1)]^{-1} \\
P(k+1|k) = \Phi(k+1,k)P(k|k)\Phi^T(k+1,k) \\
\quad + \Gamma(k+1,k)Q(k)\Gamma^T(k+1,k) \\
P(k+1|k+1) = [I - k(k+1)H(k+1)] \\
\quad P(k+1|k)
\end{cases}
$$
(1)

In the equation, $\hat{X}(k+1|k+1)$ means the state estimation. $K(k+1)$ means the Kalman gain matrix. $P(k+1|k+1)$ means error covariance matrix predicted by Kalman filter. All the variables above are based on the time point, $k+1$. $P(k+1,k)$ is the single-step prediction-error covariance matrix from the time point k to $k+1$. $Q(k)$ is covariance matrix relating to $w(k)$. $R(k)$ means covariance matrix relating to $v(k)$. I is the unit matrix. The equations in (1) are optimal filter estimation equation, optimal gain matrix equation, single-step prediction-error covariance matrix equation and error covariance matrix equation predicted by Kalman filter.

Accurate state equation and measuring equation should be deducted as well. To reach the aim, instable differential ARMA model is used in the study, as is shown in (2):

$$X(k+1) = a_1 X(k) + a_2 X(k-1) + a_3 X(k-2) + a_4 X(k-3)$$ (2)

In the equation, a_1, a_2, a_3, a_4 are the coefficient of correspondence.

Four delays are included in (2), which have to be rewrite as a matrix in order to use it in the Kalman filter. Assuming $X_1(k) = X(k)$, $X_2(k) = X(k-1)$, $X_3(k) = X(k-2)$, $X_4(k) = X(k-3)$, equation could be get as following (3):

$$
\begin{bmatrix} X_1(k+1) \\ X_2(k+1) \\ X_3(k+1) \\ X_4(k+1) \end{bmatrix} = \begin{bmatrix} a_1\ a_2\ a_3\ a_4 \\ 1\ \ 0\ \ 0\ \ 0 \\ 0\ \ 1\ \ 0\ \ 0 \\ 0\ \ 0\ \ 1\ \ 0 \end{bmatrix} \begin{bmatrix} X_1(k) \\ X_2(k) \\ X_3(k) \\ X_4(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} w(k+1)
$$
(3)

Obviously, the expression, $Z(k+1) = X(k+1) + v(k+1)$, could be put forward. In the expression, $v(k+1)$ is the addition noise which is assumed as white noise for the convenience of modeling? The equation for measuring is expressed as the following (4):

$$
Z(k+1) = [1,0,0,0]\begin{bmatrix} X_1(k+1) \\ X_2(k+1) \\ X_3(k+1) \\ X_4(k+1) \end{bmatrix} + v(k+1)
$$
(4)
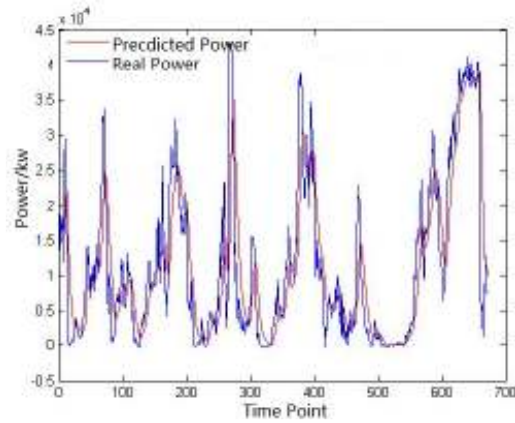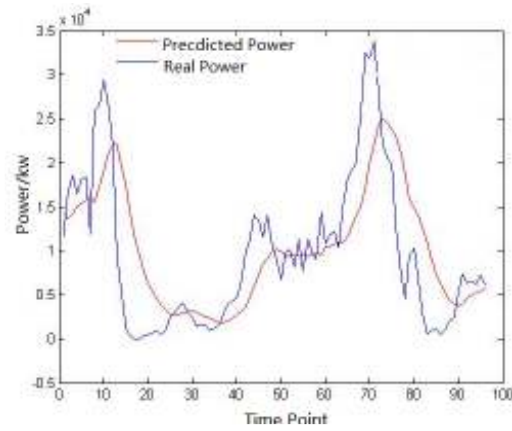

Fig. 1: Predicted and active power graph of two period of time

The occurrence of $\Phi(k+1,k)$, $\Gamma(k+1,k)$ and $H(k+1)$ could be figured out according to the Eq. (5) and the Eq. (6). After defying the initial state $X(0|0)$ and $P(0|0)$, recursive equation could be used for iterative prediction with Eq. (1). Considering convergence rate and engineering habits, the initial value is defined as $X(0|0) = [0]$ and $P(0|0) = 10l$. MATLAB is used here for the predictive value of wind power with $R(k) = [1]$ ($k = 1, 2, 3\ldots$) and $Q(k) = [1]$ ($k = 1, 2, 3\ldots$) since it is easy to realize stepwise predict.

Take P_{58} as an example, the predicted power and the active power of two period of time, from 0:00 to 23:45 in 31st May and from 0:00 to 23:45 in 6th June is shown in Fig. 1.

As is shown in Fig. 1, the real-time prediction of wind power based on Kalman filter enjoys a good effect, which basically reflects the fluctuation tendency of P_{58} though it cannot predict the fluctuation of large amplitude.

Prediction based on GARCH model:
Construction of the GARCH model (Zhou *et al.*, 2011): GARCH model enjoys the advantages in describing the volatility of time series, especially under

the condition that, q, the order of the time series is too big for ARCH (q) model to describe. GARCH model are often used to overcome low efficiency of iterative parameter, if the sample size is limited. The definition of GARCH model is described as (5):

$$\begin{cases} \gamma_t = \chi_t^T + \varepsilon_t \\ \varepsilon_t = \sigma_t \upsilon_t \qquad \upsilon_t \in NID(0,1) \\ \sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \ldots a_q \varepsilon_{t-q}^2 + \delta_1 \sigma_{t-1}^2 + \ldots + \delta_p \sigma_{t-p}^2 \end{cases} \qquad (5)$$

Relationship of the parameters in (5) applies to:

$$\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} \delta_j < 1, a_0 > 0, a_i \geq 0, \delta_j \geq 0, i = 1, 2, \cdots, q, j = 1, \cdots, p$$

Adding the auto-regression module to the expression, δ_t^2, in ARCH model, conditional variance in the GARCH model is not only related with the square of the prophase disturbance term but the prophase numerical value of the conditional variance. In fact, low order model are usually used taking place of (5), the expression of GARCH (p, q), as (6), a most frequently used model, GARCH (1, 1):

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + \delta_1 \sigma_{t-1}^2 = a_1 \varepsilon_{t-1}^2 + \delta_1 \sigma_{t-1}^2 + \delta_1 (a_0 + a_1 \varepsilon_{t-2}^2 + \delta_1 (\cdots))$$
$$= \frac{a_0}{1 - \delta_1} + a_1 (\sigma_{t-1}^2 + \delta_1 \sigma_{t-2}^2 + \delta_1 \sigma_{t-3}^2 + \cdots) \qquad (6)$$

σ_t^2, in the GARCH model does not have the characteristic of q-order censoring, which means the unlimited memory to the square of the former error term. The change makes the model much more suitable for the feature of time series in wind-power prediction. The biggest feature of the GARCH model is to describe a complex ARCH model with few parameters, lowering the number of parameters, which bring convenience for parameter estimation and parametric test.

Power prediction: Wind-power data of 31st June are predicted based on the data of wind turbine A from 28th to 30th June since the least sample that GARCH request is 200 as usual.

Stationary analysis: Figure 2 is the unit root test result of the data using EVIEWS 5.1. As is shown in the figure, the t-Statistic of the time series is larger than each ethical value of three significance levels. Thus, the raw time series are non-stationary. Based on the condition, first order difference of the raw data is calculated and the stationary test of the new data meet

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.470509	0.1238
Test critical values:	1% level	-3.453234	
	5% level	-2.871510	
	10% level	-2.572154	

Fig. 2: t -statistic of the data series

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	187.9585	34.42464	5.459999	0.0000
AR(1)	0.829172	0.033299	24.90119	0.0000
R-squared	0.685107	Mean dependent var		180.6096
Adjusted R-squared	0.684002	S.D. dependent var		177.0023
S.E. of regression	99.49952	Akaike info criterion		12.04513
Sum squared resid	2821544.	Schwarz criterion		12.07063
Log likelihood	-1726.476	F-statistic		620.0693
Durbin-Watson stat	2.457524	Prob(F-statistic)		0.000000
Inverted AR Roots	.83			

Fig. 3: Related parameter estimation and statistic value of AR (1)

F-statistic	31.33198	Prob. F(1,284)	0.000000
Obs*R-squared	28.41750	Prob. Chi-Square(1)	0.000000

Fig. 4: Residual tested by Lagrangian multiplier method

the request. As a result, the non-stationary of the raw data is weak.

Establishment of the ARCH model: According to the establishing steps of the ARCH model, the basic form of the time series is figured out. Then, residual is tested by Lagrangian multiplier method to figure out whether ARCH effect exists in the time series. If the ARCH effects do exist, related AR (k) and GARCH (p, q) model would be established based on correlation effect and the general form of GARCH model.

According to the analysis above, raw data could be fitted by AR (k) and GARCH (p, q). Auto-correlation module of time series is described in AR (k). As auto-correlation module of the time series is usually less than 5, 5 models, which are expressed as AR (1), AR (2), AR (3), AR (4) and AR (5), are established. If ARCH effect exists in AR (k), build GARCH (1, 1) and GARCH (2, 2) model for its residual. Taking AR (1) as an example, related parameter estimation and statistic value is calculated and expressed in Fig. 3 by EVIEWS 5.1.

Testing the residual of AR (1) with Lagrangian multiplier method, the following results are figured out as is shown in Fig. 4.

In Fig. 4 $LM = TR^2 = 28.4175 > x_{0.05}^2 = 3.841$. Since theLM -related probability is smaller than 0.05, the value is in the right side of critical value, which is the rejection region of null hypothesis. As a result, ARCH exists in three error sequence of the model.

	Coefficient	Std. Error	z-Statistic	Prob.
C	126.2052	44.05677	2.864603	0.0042
AR(1)	0.844319	0.044574	18.94212	0.0000
	Variance Equation			
C	17902.89	1433.439	12.48946	0.0000
RESID(-1)^2	0.098717	0.029696	3.324281	0.0009
GARCH(-1)	-0.840434	0.050467	-16.65318	0.0000
R-squared	0.681835	Mean dependent var		180.6096
Adjusted R-squared	0.677322	S.D. dependent var		177.0023
S.E. of regression	100.5458	Akaike info criterion		11.97413
Sum squared resid	2850865	Schwarz criterion		12.03789
Log likelihood	-1713.288	F-statistic		151.0829
Durbin-Watson stat	2.470618	Prob(F-statistic)		0.000000

Fig. 5: Related parameter estimation and statistic value of AR (1)-ARCH (1, 1)

Table 1: Same-model testing results

AR(k)-GARCH					
(p, q)	(1, 1, 1)	(1, 2, 1)	(2, 1, 1)	(2, 2, 1)	(2, 1, 2)
AIC	11.974	11.743	11.623	11.821	11.921
SIC	12.038	11.921	11.745	11.825	12.127

According to the GARCH (1, 1) model of residue based on the AR (1) -mean-value equation, the result could be figured out as Fig. 5.

Thus, the mean-value equation could be figured out as (7):

$$X_t = 126.2052 + 0.844319 X_{t-1} + \varepsilon_t \qquad (7)$$

The GARCH equation is:

$$\sigma_t^2 = 17902.89 + 0.09877\varepsilon_{t-1}^2 - 0.840434\sigma_{t-1}^2$$

Therefore, different LM are established for testing different AR (k)-GARCH (p, q) model.

Model selection: According to the judgment and testing results, 5 models suitable for time series of wind power could be established. Then, different models would be selected based on several criterions. As a result, AIC information criterion and SIC information criterion are chosen for reflecting goodness of fit of time series. As is shown in Table 1, AR (2)-GARCH (1, 1) enjoys the best imitative effect.

Model prediction: Predicting the wind power of 31st June produced by Wind Turbine A by AR (2)-GARCH (1, 1), the results are constructed as Fig. 6. What could be seen from Fig. 6 is cluster effect of the wind-power random perturbation? That means large fluctuation range are often followed with the range of large function and it is the same as that of small fluctuation range, which means the feature of ARCH model. Other results could be figured out with the same principle.
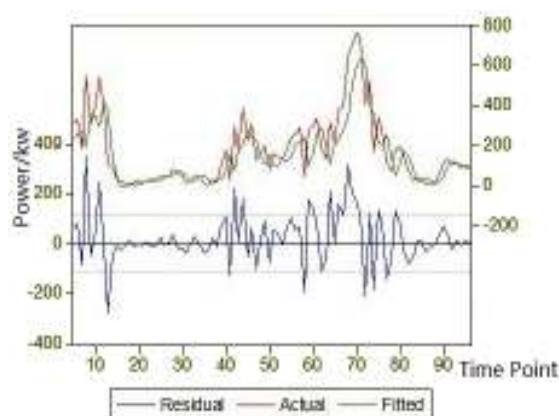

Fig. 6: Wind-power prediction result of 31st June

Real-time prediction by time-series-based BP neural network:
Establishment of the BP-neural-network prediction model (Zhang and Ghao, 2003):
Deciding network layer: BP neural network is constituted by input layer, output layer and several hidden layer. Each layer is made by several nodes, each of which represents a neuron. Nodes between layers are connected by full-associated connecting mode; no connection is made among nodes within the same layer though. Thus, the number of layer of BP neural network depends on the number of the hidden layers.

Adding the hidden layers cannot meet the request of real-time prediction, since it will increase the calculating time. As a solution to the problem, one-hidden-layer BP neural network are chosen to predict the wind power.

Deciding the neuronal number in each layer: There is a close connection between input node, output node and sample, which related with their application field. According to the request of the prediction, wind farm are required to forecast the wind power data every 15 min of the next 4 h. The neural number of input layer is 16, which means 16 wind power data should be conducted as input variable. Meanwhile, the neural number of the output layer is 1, representing the 17th wind power data. That means the times series, $x_1, x_2, \dots x_n$, satisfy with Eq. (8) as following:

$$x_{i+16} = f(x_i) \, x_i = (x_i, x_{i+1}, x_{i+2}, \cdots, x_{i+15}) \qquad (8)$$

Two points for attention for deciding the neural number of hidden layers is mentioned. The too small neural number will lead to the hardship of recognizing samples, which makes it training obstacle and even lower the network tolerance:
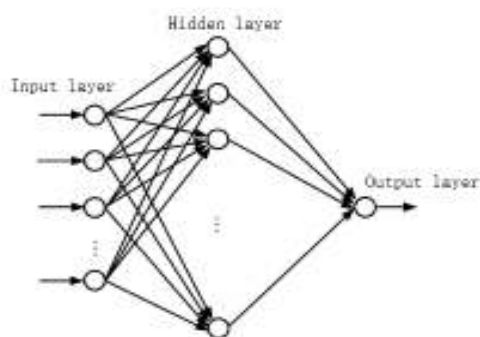
Fig. 7: BP neural network

$$i = \sqrt{n + m} + k \qquad (9)$$

i = The neural number of the hidden layer
n = The neural number of the input layer
m = The neural of the output layer

Thus, the neural number of the hidden layer is 8; the model can be described as Fig. 7.

Training BP neural network by sample vector: Usually, the weight of BP neural network would not change once decided. In this model, wind power is fluctuated since the wind power depends on wind speed, whose feature is volatility, intermittency and low energy density. Thus, weights are adjusted with the following steps:

Step 1: After training the network with the training sample data, x_1, x_2, x_3……x_{16}, x_{17}, could be predicted.
Step 2: Take x_{17} as the input sample of the series, which means the training sample of the second group is x_2, x_3……. Using the data to retraining the network, x_{18} is predicted.
Step 3: After the repeated iteration that training the network with the data, x_t, x_{t+1} …… x_{t+15} and predict x_{t+16}, the weights of the network tend to be higher precision, which better reflects the change of wind power.
Step 4: Preprocessing data: Number the data from 10th to 30th in May in order of time step. Then, normalizing the number to real number in the interval, [0, 1], since the variable range is large. The normalization is expressed as (10):

$$X_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \qquad (10)$$

where,
X_i : The normalized wind power
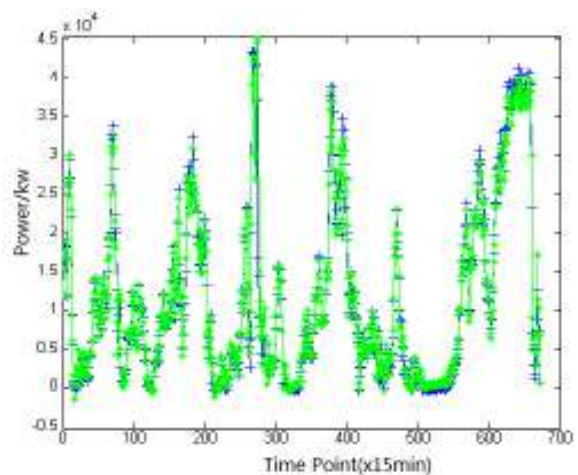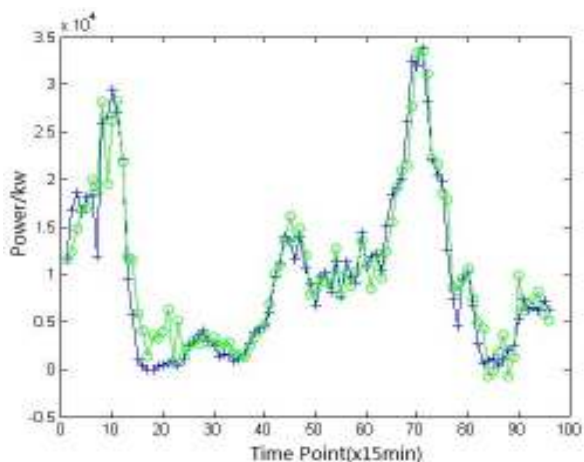x_i : The original wind power

Fig. 8: Predicted and real data of two period of time

x_{max} : The maximum of the original wind power
x_{min} : The minimum of the original wind power

Step 5: Input the preprocessed number from 10th to 30th May as training sample. The desired error is reached with the iteration number of 392.

Real-time prediction for wind power with BP neural network: Inputting the historic data to the trained network, 16 wind power data of the following 4 h could be predicted. Then, the wind power data of 31st May and that from 31st May to 6th June are calculated. Predicted data of wind power from 0:00 to 23:45 and from 0:00 31st May to 23:45 6th June could be calculated with MATLAB, which is shown in Fig. 8.

Method evaluation: Accuracy and qualification are established for comparison of the three methods in this study according to relevant standard.

Table 2: Accuracy and qualification of the three methods

		Kalman filter model		GARCH model		BP neural model	
		One day (%)	Seven day (%)	One day (%)	Seven day (%)	One day (%)	Seven day (%)
P_A	Accuracy	88.03	88.97	87.87	92.36	95.28	95.86
	Qualification	83.33	88.24	92.71	94.20	98.96	99.85
P_B	Accuracy	85.89	88.05	90.26	90.96	93.58	93.75
	Qualification	80.21	87.50	85.42	92.41	96.88	95.98
P_C	Accuracy	84.90	88.40	85.48	91.49	94.42	95.54
	Qualification	79.17	88.40	82.29	92.11	97.92	98.66
P_D	Accuracy	87.93	89.05	91.52	92.11	93.65	94.65
	Qualification	84.38	89.73	94.79	93.30	100	99.40
P_4	Accuracy	88.33	89.81	91.92	93.97	96.49	96.92
	Qualification	87.50	91.52	97.92	97.02	99.70	97.44
P_{58}	Accuracy	90.63	91.97	94.85	96.13	98.23	98.31
	Qualification	92.71	95.09	100	98.96	100	99.70

Accuracy: Define the difference between 1 and root-mean-square error as the accuracy r_1 as (11):

$$r_1 = (1 - e_{RMSE}) \times 100\% \qquad (11)$$

e_{RMSE} is the root-mean-square error.

Qualification: Define qualification as (12):

$$\begin{cases} r_2 = \dfrac{1}{n}\sum_{i=1}^{n} B_i \times 100\% \\[2mm] \dfrac{|e_i|}{P} \leq 25\%, \quad B_i = 1 \\[2mm] \dfrac{|e_i|}{P} > 25\%, \quad B_i = 0 \end{cases} \qquad (12)$$

e_i = Prediction error
P = The rated power of the wind farm

Substituting real power data and real power data calculated by the three methods to (11) and (12), accuracy and qualification within one day (from 0:00 to 23:45 31st May) and seven days (from 0:00 31st May to 23:45 6th June) could be figured out, which is shown in Table 2.

According to the comparison, it is obvious that the time-series-based BP neural network model and GARCH prediction model is better than Kalman filter model. Moreover, the time-series-based BP neural network enjoys more advantages than GARCH prediction model. As a result, the time-series-based BP neural network model is selected for prediction.

PREDICTION ERROR EXPECTATION BY GATHERING

Based on the analysis above, the law we hope to get is that the gathering of wind turbines obviously decreases the prediction error.

Table 3: Root-mean-square error of different method

Time step/min	Kalman filter	GARCH	BP neural network
15	0.167	0.154	0.072
30	0.174	0.165	0.082
60	0.187	0.177	0.090

Combining with the facts, wind power are affected by wind speed, wind direction, temperature, sir pressure and so on. The wind power would suffer fluctuation owing to the natural factors above. The fluctuation range enlarges as the range of these factors increase, which are easy to be ignored in the prediction, leading directly to the predicting error. Meanwhile, as wind turbines are built in the same area, the influence of natural factors on wind turbines are similar. After gathering, accumulation of wind-power fluctuation is scaled up with more wind turbines. As a result, the fluctuations that ignored in the single-turbine prediction would be predicted efficiently, lower its predicting error.

At present, wind power is centralized developed. Wind power of each turbine gathers before connecting to the power network by wind farms. It is meaningful for ensuring the power balance and running safety to predict wind power of wind farms accurately. According to the results analyzed, suggestions are made that more wind turbines connect to the power network.

MAIN FACTORS PREVENTING THE RAISE OF THE REAL-TIME PREDICTION

Number of historic data: Sensitivity analysis is made for analyzing the error by changing the number of historic data. As different numbers of data make contributions to different predicting errors, lacking historic data prevents the improvement of the wind-power-prediction accuracy.

Method selection: Different method contributes to different error. If features of the high-accuracy method are find out, accuracy could be improved.

Selection of time step: Predicting with different time steps, make the sensitivity analysis and error analysis. Then the root-mean-square error of predicted data in 31st June is in Table 3. Root-mean-square error of every model grows when the time step enlarges. Thus, the time step has effect on accuracy of prediction.

CONCLUSION

Change law of the wind power is well captured by the neural network model, time-series-based GARCH model and Kalman filter model, which all enjoy quantities of advantages in prediction. Among them, Kalman filter keep renewing the weight of state estimation during recursion, ensuring the results close to the real condition. GARCH model not only well considered the sequential characteristic of Root-mean-square error but explain the volatility of time series. The benefit that the neural network has is its characteristics of nonlinear complex signal processing, which greatly approximating continuous curve.

Influenced by similar natural factors, wind powers of turbines in the same area fluctuate much more violently after gathering together, which makes it easier for prediction, leading a higher accuracy.

Due to the existence of systematic errors, accuracy of wind power cannot rise without limitation. However, a more accurate prediction of wind power is of great possibility. Thus, the suggestions are given in the following part based on the factors above.

Firstly, number of the data should be increased and the time step should be narrowed. Secondly, the prediction would be much better developed if meteorological information and physical model are added. In addition, it is tend to build multi-predicting model for forecasting wind power more accurately.

REFERENCES

Boone, A., 2005. Simulation of Short-Term Wind Speed Forecast Errors Using a Multi-Variate ARMA (1, 1) Time-Series Model. Royal Institute of Technology, Sweden, pp: 70-88.

Bossanyi, E.A., 1985. Short-term wind prediction using Kalman filters. Wind Eng., 9: 1-8.

Brown, B.G., R.W. Katz and A.H. Mtlrpby, 1984. Time series models to simulate and forecast wind speed and wind power. J. Clim. Appl. Meteorol., 23(8): 1184-1195.

Kariniotakis, G.N., G.S. Stavrakakis and E.F. Nogaret, 1996. Wind power forecasting using advanced neural networks models. IEEE T. Energy Conver., 11(4): 762-767.

Lexiadis, M.A., P.S. DokoPoulos, S.H. Samanoglou, I.M. Manousaridis, 1998. Short-term forecasting of wind speed and related electrical power. Sol. Energy, 63(1): 61-68.

Zhang, Y. and Z.W. Gao, 2003. The Load Forecasting of Electric Energy Based on Time Sequence and Neural Network. J. Harbin Univ. Sci. Tech., pp: 30-36.

Zhou, H., J.X. Fang and M. Huang, 2011. Application research of wind power forecasting model GARCH. Power Syst. Protect. Control, 39(5): 108-119.