Fault early warning of pitch system of wind turbine based on GA-BP neural network model

Sihan Chen, Yongguang Ma, and Liangyu Ma

North China Electric Power University, Baoding, 071000, China

Abstract. A fault early warning method based on genetic algorithm to optimize the BP neural network for the wind turbine pitch system is proposed. According to the parameters monitored by SCADA system, using correlation analysis to screen out the parameters of the pitch system with strong power correlation. The BP neural network optimized by genetic algorithm is used to establish the model of the pitch system under normal working conditions. The verification results show that the input parameters of the pitch system model determined by the correlation coefficient are more reasonable, and the accuracy of the pitch system model established by the genetic algorithm-optimized BP neural network is higher than that of the unoptimized model. Based on the above model, a sliding window model is established, and the early warning threshold is determined through the statistics of the residuals of the sliding window to realize the fault early warning of the pitch system of the wind turbine. The example shows that the method can give early warning in the event of failure, and verifies the effectiveness of the method.

1 Introduction

With the development of the wind power industry in recent years, the installed capacity of wind turbines in China has also increased year by year. The pitch system is one of the important control systems. It includes pitch motor, pitch control cabinet, pitch driver, etc. These are fault-prone components. Due to the blindness and lag of the wind farm's fault handling of the pitch system, as the frequency of system failures increases, it has a great impact on the normal operating time of the wind turbine.

At present, the fault warning for wind turbines mainly focuses on the research of gearbox oil temperature and oil pressure and generator temperature. Reference [1] established a neural network model of gearbox bearing temperature based on principal component analysis to realize early warning of gearbox failure. Reference [2] used a weighted principal component analysis method to establish a generator temperature model under normal operating conditions, and accurately analyzed generator temperature faults by analyzing statistics and squared prediction error trends. Reference [3] based on nonlinear state estimation of the wind turbine pitch control system model can diagnose major and minor faults. Although the above literature has research on the pitch system, it only diagnoses the fault and does not propose a warning method.

This article takes 1.5MW double-fed wind turbine as research object. Aiming at the data collected by SCADA system, a genetic algorithm optimized BP (GA-BP) neural network model of the pitch system is established. Based on this, the fault early warning threshold is determined based on the sliding window model, and the fault early

warning of the pitch system is realized.

2 Characteristic parameters of pitch system

A total of 22 parameters of the pitch system monitored by a wind farm SCADA system include blade angle, pitch motor current, pitch battery voltage, etc. In order to filter out the characteristic parameters of the pitch system with a large power correlation coefficient, Pearson correlation coefficient was introduced for correlation analysis [4]. The calculation formula is:

$$r_{xy} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (1)$$

Where, X_i and Y_i ($i = 1 \cdots n$) are two parameters for calculating the correlation, $\bar{X}$ and $\bar{Y}$ are the average values of the sample parameters. The correlation coefficient $r \in [-1,1]$, the greater the absolute value of r, the stronger the correlation. Here, selecting the characteristic parameter with a correlation coefficient $|r| \geq 0.6$ indicates a strong correlation with power. The selected characteristic parameters are shown in Table 1.

Table 1. Correlation analysis results

Characteristic parameters	Correlation coefficient
Wind wheel speed	0.8447
1 # Paddle motor temperature	0.7670
2 # Paddle motor temperature	0.7672
3 # Paddle motor temperature	0.7752
1 # Paddle blade angle	0.6601
2 # Paddle blade angle	0.6605
3# Paddle blade angle	0.6604

The operating conditions of the pitch system in different wind speed ranges are different. Below the rated wind speed, the speed is tracked to control the wind speed to optimize the power; above the rated wind speed, the torque controller and the pitch controller are used to make the power and torque relatively stable, so the wind speed and generator torque are included in the model parameters. Therefore, the input parameters of the pitch system model are finally selected: wind turbine speed, 1 # 2 # 3 # paddle motor temperature, 1 # 2 # 3 # paddle blade angle, wind speed and generator torque; the output parameter is power.

3 Model of pitch system

3.1 BP neural network model

BP neural network is a kind of multilayer feedforward neural network, which is composed of input layer, hidden layer and output layer. It mainly includes two processes of forward signal propagation and reverse error propagation [5]. Forward signal propagation refers to the process of passing input parameters through the input layer to the hidden layer and finally outputting through the output layer. If the output reaches the expected value this time, a forward signal propagation process is completed; if the output fails to reach the expected value and there is a large error from the actual value, the neural network will propagate the error back and modify the connection weight and threshold to minimize the error until the network requirements are met.

According to the selected pitch system characteristic parameters, it is determined that the number of input layers of the BP neural network is 9, the number of output layers is 1, and the number of hidden layers is 5. The excitation function of the hidden layer is selected as "tansig", the excitation function of the output layer is selected as "purelin", and the training function of the neural network is selected as "trainlm". Establishing BP neural network model, the structure is shown in Figure 1.
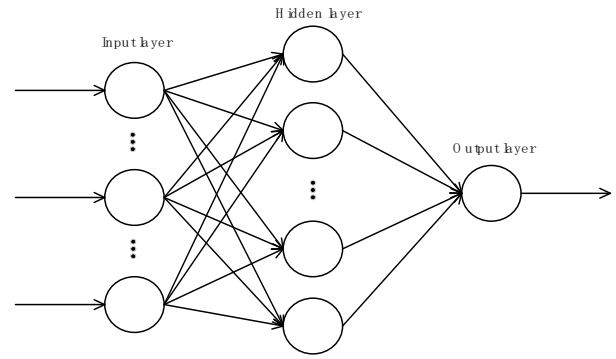

Fig.1 BP neural network model

3.2 GA-BP neural network model

BP neural network is easy to fall into the local minimum, resulting in training failure. Using genetic algorithm to optimize BP neural network can solve its problems. The genetic algorithm searches for the global optimal solution by simulating the evolution process in nature, which is mainly divided into three steps of selection, crossover and mutation [6].

（1）Selection

According to the principle of "survival of the fittest", individuals with high fitness and high selection rate are selected from the group, and individuals with low selection rate are eliminated.

$$p_i = f_i / \sum_{j=1}^{n} f_i \qquad (2)$$

Where, n represents the group size, i represents the individual, f_i represents the fitness of the individual i, p_i represents the probability that individual i is selected.

（2）Crossover

Replacing and reorganizing part of the structure of the two parent individuals to generate new individuals. The real-number crossover method is usually used. The crossover operation formula for the kth chromosome a_k and the lth chromosome a_l at the j position is as follows:

$$a_{kj} = a_{kj}(1 - b) + a_{ij}b$$
$$a_{ij} = a_{ij}(1 - b) + a_{kj}b \qquad (3)$$

Where, b is a random number between [0,1].

（3）Mutation

First determine whether all individuals are mutated. Then, the individuals who need to be mutated are randomly selected to mutate. The local search ability of the mutation operator accelerates convergence to the optimal solution. Select the jth gene a_{ij} of the ith individual for mutation. The mutation formula is as follows:

$$a_{ij} = a_{ij} + \left(a_{max} - a_{ij}\right) \times f(g), \quad r > 0.5$$
$$a_{ij} = a_{ij} - \left(a_{ij} - a_{min}\right) \times f(g), \quad r \leq 0.5$$
$$f(g) = r'(1 - \frac{g}{G_{max}})^2 \qquad (4)$$

Where, a_{max} is the upper bound of the gene a_{ij};

a_{min} is the lower bound of the gene a_{ij}; g is the current iteration number; G_{max} is the maximum evolution number; r' is a random number between [0,1].

Genetic algorithm-based BP neural network is to find the optimal initial connection weights and thresholds of BP neural network through genetic algorithm, so that the BP neural network after determining the optimal initial parameters can better predict the output after training [7]. Set the initial parameters of the optimization process with a population size of 80, the number of iterations, a cross probability of 0.4, and a mutation probability of 0.001.

3.3 Model simulation analysis

This article collects data of a wind turbine SCADA system for a continuous month in April. The rated power of the unit is 1.5MW, the cut-in wind speed is 3m/s, and the cut-out wind speed is 25m/s. Preprocess the data before training the neural network, remove fault data, and screen out normal operating data with power greater than 0 and wind speed between 3-25m/s. The pitch system models of BP neural network and GA-BP neural network were established respectively, and the accuracy of the models was compared. The prediction results of the BP neural network and GA-BP neural network models are shown in Figs.2 and 3, and the comparison results of the two models is shown in Fig.4.
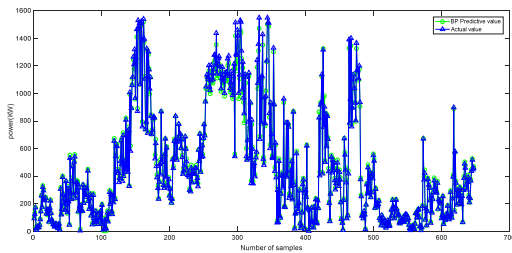

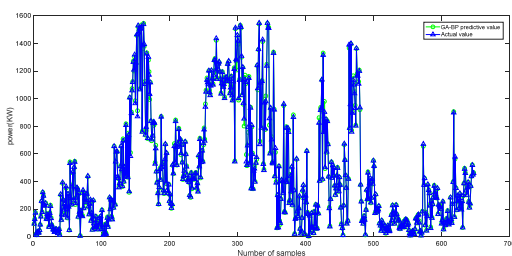**Fig.2.** BP neural network model prediction results

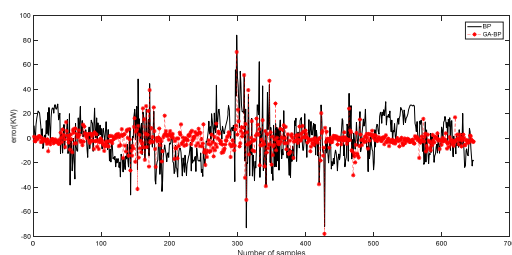**Fig.3.** GA-BP neural network model prediction results

Fig.4 BP and GA-BP model prediction errors

From Fig.4 it is found that the error range of the

prediction model of the pitch system established by the BP neural network is between [-73,84]. The error range of the prediction model of the pitch system established by GA-BP neural network is between [-70,70]. Although the error fluctuation range of the two models is not much different, the fluctuation range of the GA-BP neural network model is more stable.

The mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) are used to calculate and analyze the training and prediction errors of the two models. The results are shown in Table 2. The above error calculation formulas are:

$$MAE = \frac{1}{m} \sum_{i=1}^{m} |f(x_i) - y_i| \qquad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(f(x_i) - y_i)^2}{m}} \qquad (6)$$

$$MAPE = \sum_{i=1}^{m} \left| \frac{f(x_i) - y_i}{h(x_i)} \right| \times \frac{100}{m} \qquad (7)$$

Where, m represents the total number of samples; $f(x_i)$ represents the actual value of the parameter; y_i represents the calculated value of the neural network.

Table2. Evaluation index of neural network model

Evaluation index	BP	GA-BP
MAE	14.256	5.552
RMSE	17.887	9.459
MAPE	12.271	2.453

From the Table2, the three evaluation indexes of the GA-BP neural network model are smaller than the BP neural network model. It is fully proved that the prediction effect of GA-BP neural network is better. Therefore, the GA-BP neural network model is selected as the fault early warning model of the wind turbine's pitch system.

4 Fault warning of pitch system based on sliding window

4.1 Sliding window model

The GA-BP neural network model is based on the model established under normal working conditions, so the model effect is better during normal operation. However, when an abnormal situation occurs in the equipment, the prediction of the model will deviate from the normal state, resulting in a large deviation [8]. In order to better monitor the running state of the wind turbine pitch system on-line, this paper establishes a sliding window model and implements fault early warning through the sliding window residual statistics method.

The model established by this method takes time as the unit, divides the real-time updated data stream into the same time window, and recognizes the state by updating the sliding of the window, which can avoid interference from external factors and false alarms caused by extreme operating points [9]. This method can continuously reflect the change of the residual distribution characteristics, and is suitable for online analysis. In a window of width h, calculate its running status index:

$$C = \sqrt{\frac{\sum_{i=1}^{h}(f(x_i) - y_i)^2}{\sum_{i=1}^{h} f(x_i)^2}} \qquad (8)$$

Where, $f(x_i) - y_i$ indicates the difference between the actual value and the predicted value.

The time interval of the data collected in this article is 1 minute, and the window width $h = 10$ is chosen to ensure the accuracy and reliability of data analysis. The sliding window model is shown in Fig.5.
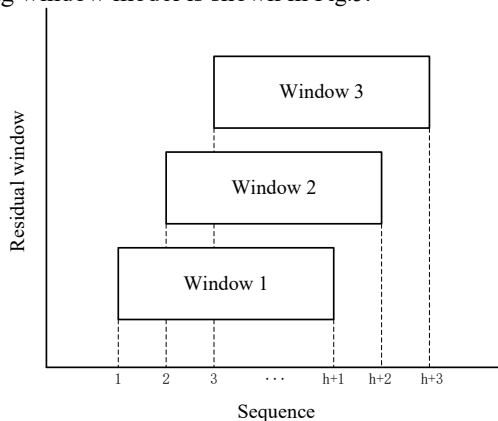

Fig.5. Sliding window residual statistic

4.2 Operating state threshold

The operating conditions of the wind turbine's pitch system are variable and randomly switched, so the operating index is also a random quantity, but according to experience, it is known that the random operating index obeys a certain statistical distribution law [10]. This paper determines the upper limit of the operating index of the pitch system according to the statistical formula.

$$C_{th} = \overline{C_{std}} + \frac{S_{std}}{\sqrt{N_{std}}} t_\alpha (N_{std} - 1) + 3S_{std} \qquad (9)$$

Where, $\overline{C_{std}}$ is the average value of the normal operation indicator; S_{std} is the standard deviation of the normal operation indicator; and N_{std} is the total amount of the window during normal operation. t_α is the distribution of t when $\alpha = 0.0005$, and the calculated value is $C_{th} = 0.052$, so the fault early warning threshold of the pitch system is 0.052.

4.3 Failure warning example

In order to verify the effectiveness of the above method, the normal operating data of different months are selected for verification. It is known that the failure at 8:06 on May 13th is that the angle difference between the two blades of the pitch change is 2 degrees. The experiment selects the operation data from 0:00 on May 12th to 8:05 on May 13th for verification. It can be seen from Fig.6 that the residual value of the 938th sliding window sequence exceeds the upper threshold, and the corresponding time of this window is 20:04 on May 12th. This proves the effectiveness of the early warning model, and the model shows an early warning 12 hours in advance.
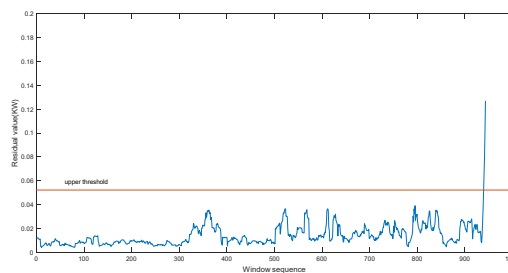

Fig.6 Fault warning of pitch system based on sliding window

5 Conclusions

(1) Combined with the data of SCADA system, the method of Pearson correlation analysis is introduced to screen the input and output parameters of the pitch system model. The selected parameters have strong reliability.

(2) According to the input and output parameters, the BP neural network model and GA-BP neural network model of the pitch system are established respectively. The results show that the predicted output of the GA-BP neural network fits the actual output better and the model accuracy is higher.

(3) By establishing a fault early warning model for the pitch system of a wind turbine, it is possible to provide early warning before a fault occurs, which proves the validity and reliability of the model and is of great significance for the operation and maintenance of the pitch system.

Acknowledgements

References

1. Tao L, Gang L, Ruiqi C, Wenhuan L. Fault gear warning of fan gearbox based on bearing temperature[J]. Renewable Energy, 2018,36(12): 1877-1882.

2. Xuan L, Jianping S. Research on Early WarningMethod for Temperature Rise of Wind Turbine[J]. Electric Power Science and Engineering,2016,32(6):38-43.

3. Shi Y, Zhongyuan Y, Kaifeng M, Chuang L, Fault Identification of Pitch Control System for Wind Turbines Based on Nonlinear State Estimation [J]. Transactions of China Electrical Engineering, 2014,34(S1):160-165.

4. Zhongshan H, Ling T, Dong X. PCA and SPC-Dynamic Neural Network for Prediction of Oil Temperature Trend of Gearboxes in Wind Turbi-nes[J]. Journal of Tsinghua University (Science and Technology),2018,58(6):539-546.

5. Xiaotang X, Tingyang L.A fuzzy control model based on BP neural network arithmetic for opti-mal

control of smart city facilities[J]. Personal and Ubiquitous Computing,2019,23(3-4).

6. Qian X, Wenhua L, Zhigang L. Prediction of wind speed and wind power based on improvedwavelet BP neural network [J]. Power System Protection and Control,2014, 42(15):80-86.

7. Qing Z, Wenya X, Tao T. Optimization design of cassava straw cutting based on neural network genetic algorithm[J]. Journal of Chongqing University of Science and Technology (Natural Science Edition),2016,18(2):114-116.

8. Hui L, Chao Y, Xuewei L, Haiting J, Xing Q, Yaojun C. Mining of State Characteristic Param-eters and Abnormal Recognition of Electric Fan Pitch System[J]. Chinese Journal of Electrical Engineering, 2014,34(12):1922-1930.

9. Jiayu D, Chang X, Lichao G, Jie Y, Shuai X, Yuntao L. Research on fault prediction of wind turbine based on bearing temperature model [J]. Renewable Energy, 2018,36(02):276-282.

10. Fan Z, Deshun L, Juchuan D, Chao W, Xiangbing S. A Method for Identifying Wind Turbine Operating State Based on SCADA Parameter Relationship [J]. Journal of Mechanical Engineering, 2019,55(4):1-9.