Data-driven fault identification of ageing wind turbine based on NARX

Data-driven fault identification of ageing wind turbine based on NARX

Yue Liu[1], and Long Zhang[2] *

[1] Department of Electrical & Electronic Engineering, University of Manchester, UK yue.liu@manchester.ac.uk

[2] Department of Electrical & Electronic Engineering, University of Manchester, UK long.zhang@manchester.ac.uk

Abstract. As the existing wind turbine is approaching the designed service life, it is of great significance to check the ageing condition in advance. In this study, a system identification model, nonlinear autoregressive network with exogenous inputs (NARX), was used to analyze the ageing condition of wind turbines. This data-driven approach uses the input and output data of the system directly without the need for specific mathematical models. Simulated experimental data for four different ageing conditions are used for system identification. By comparing the NARX model parameters under different conditions, the fault conditions of the system can be found and the degree of ageing can be detected.

Keywords: NARX, System Identification, Ageing Assessment, Artificial Intelligence.

1 Introduction

In recent years, the installed capacity of wind power has increased rapidly, and the proportion of wind power in the UK's power supply has also increased year by year [1]. In 2020, wind power accounted for 24.8% of the power generation in the UK power system [2], and it is estimated that by 2030, the proportion of power supply from offshore wind power alone will reach one-third of the total power generation. To reduce energy costs and improve turbine reliability, many studies are devoted to developing better condition diagnostic techniques [4]. Relevant research can help improve wind turbine design and help formulate corresponding policies, thereby accelerating the development of the wind power industry. With the continuous development of wind power, the service life of more and more wind turbines is gradually approaching the design life [3]. Wind turbines are designed to last around 20 years, ageing affects wind turbine performance over time, losing approximately 1.6% of output per year [5]. Because wind turbines are subjected to extreme loads and harsh operating environments throughout their service life, which cause accelerated machine aging, wind turbines are theoretically unlikely to last longer than they were designed for [6]. Therefore, it is very important to evaluate the ageing degree of wind turbines that have been used for several years. This will facilitate the life management of wind turbines in wind farms, allowing them to operate until their design year, saving maintenance costs and improving economic returns.

Wind turbine ageing assessment and fault diagnosis methods can be divided into three categories: signal processing method, model-based method and data-based method. Vibration signal analysis is one of the most commonly used methods for fault detection. By using fast Fourier transform, wavelet transform and other methods to analyze vibration signals, fault conditions can be found through the characteristic frequencies in the frequency spectrum [8]. In addition, there are other methods for fault diagnosis of wind turbines, such as torque [9] and current [10], which collect data from other different sensors. However, this may increase the cost of power generation, because corresponding sensors need to be installed on each wind turbine, and the working status of the sensors should also be considered. The model- based method is to conduct system analysis by establishing a mathematical model of the system under test and comparing it with the real model. But this approach requires knowledge of a detailed physical model of the system and then digitizing the system. The wind turbine system has a complex structure and is a very typical nonlinear system, so it is very difficult to digitize the entire dynamic system [14]. Similar to the model- driven approach, the data-driven approach also requires a system model. However, the data-driven method does not need to know the specific physical model of the system, and only needs to use the input and output signals of the system to build a dynamic model of the system. By analyzing the changes of the model and comparing the characteristic patterns in different situations, the fault diagnosis can be realized.

At present, there are few research on the ageing detection of wind turbines, and most of the research focus on fault diagnosis and reliability analysis. This is mainly since ageing is different from damage to components, and once a component fails, the entire system becomes completely unusable. Ageing occurs relatively slowly, and there may not be a specific component in the entire system that fails. Although it may have a certain impact on the output power, the overall system is still running normally. Therefore, it also brings difficulties to the research on ageing. The slow and lengthy ageing process makes it difficult to collect data for all ageing periods of a wind turbine. Moreover, even two identical wind turbines in the same wind farm can have completely different components and severity of ageing. This also requires the simultaneous analysis of multiple component conditions to assess the performance of the wind turbine. In addition, different components may need to use different sampling rates to best capture its dynamics. Existing datasets basically collect data for all components at a uniform sampling rate, which may not be the most suitable for some components [7]. The main contributions of this paper are as follows. The nonlinear modelling is the main contribution while considering different factors, such as different degrees of ageing, sampling rates, and lengths of time for specific components. Therefore, the effect of different sampling rates on the results of ageing experiments at the same part can be observed.

2 DFIG wind turbines

Doubly-fed induction generator (DFIG) wind turbine systems are widely used due to their low installation cost and both the stator and rotor can supply power to the grid [16]. The adjustable speed generator and blade pitch control improve system efficiency while reducing mechanical stress and other issues [15]. While making the system more complex and costly, and large infiltrations can cause problems in the power system, more energy extraction can offset this negative impact [13], [16].

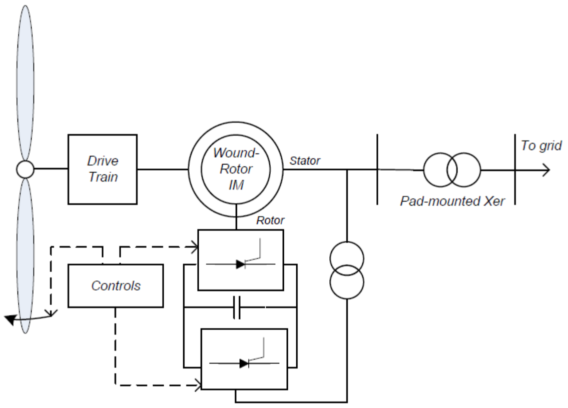


Fig. 1. DFIG wind turbine schematic [13].

3Data-Driven Method

System identification is widely used in building systems because it does not need to consider the exact physical model of the system, but only needs to consider the relationship between the input signal and the output signal. The models available for fitting can be divided into two categories: linear and non-linear.

3.1 NARX Model

Wind turbine system is a typical nonlinear system. NARX can be used to describe nonlinear dynamic systems, is suitable for time series forecasting, and is applied to solve nonlinear series forecasting problems in various fields. The mathematical formula of multiple input multiple output (MIMO) NARX can be expressed as:

$$\begin{aligned} y(t) = \quad & f\big(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u) \\ & e(t-1), \dots, e(t-n_e)\big) + e(t) \end{aligned} \tag{1}$$

Where,

$$y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix}, u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_r(t) \end{bmatrix}, e(t) = \begin{bmatrix} e_1(t) \\ \vdots \\ e_m(t) \end{bmatrix} \quad (2)$$

represent the output, input and noise vector of the system at time t; n_y, n_u and n_e represents the delay order of output, input and noise, respectively. $f(\cdot)$means non-linear function which is generally complex and difficult to obtained. Thus a common method is to use a simpler approximation formula. In this paper we use the method of neural networks.

$$\begin{aligned} y(t) = f\big(y(t-1), \dots, y\big(t-n_y\big), \\ u(t-1), \dots, u(t-n_u)\big) + e(t) \end{aligned} \quad (3)$$

Eq. (3) is a special case of (1), because it only considers the independent noise sequence. The neural network model structure of NARX can be shown as Fig.2, one or more layers could be designed between the input and output, n_l is the number of hidden layers.

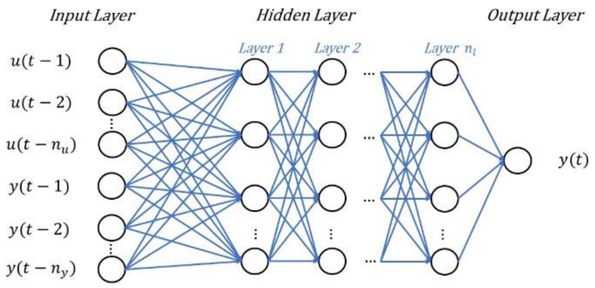


Fig. 2. The structure of multiple input NARX

Neural networks is statistical-based learning approach, and NARX is able to predict the output based on the current input, taking into account the impact of previous inputs on the system. The multi-layer structure of NARX is designed with hidden layer, input layer and output layer. Changing the number of hidden layers and neurons will affect the stability of neural network, which requires more attempts in experiments. The operation process of each neuron in the system is shown in Fig. 3, where x_i is the neuron input and y is neuron output. The functional relationship between input and output is given by threshold parameter μ, weights w_i and activation function $a(\cdot)$:

$$y = a(\sum w_i x_i + \mu) \quad (4)$$

If all activation functions are linear, then the output of the neuron is the weighted sum of the inputs.

$$y = \sum w_i x_i \tag{5}$$

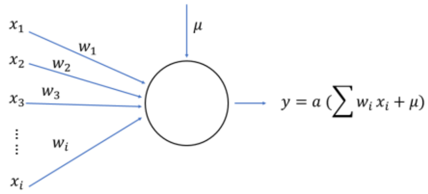


Fig. 3. The structure of a neuron

According to the universal approximation theorem, artificial neural networks have the ability to approximate arbitrary functions [11]. Fewer hidden layers make the computation faster, so if we use a neural network with one hidden layer to model the non-linear system described by Eq. (3). Define $n = mn_y + rn_u$

$$x(t) = [x_1(t) \quad \dots \quad x_n(t)]^{\mathrm{T}} = \begin{bmatrix} y^{\mathrm{T}}(t-1) \\ \cdots \\ y^{\mathrm{T}}(t-n_y)u^{\mathrm{T}}(t-1) \\ \cdots \\ u^{\mathrm{T}}(t-n_u) \end{bmatrix} \tag{6}$$

At the same time, it is stipulated that the number of hidden neurons is n_h, threshold of i-th hidden neuron is $\mu_i^{(h)}$. The weight from $x_j(t)$ to i-th hidden neuron is $w_{ij}^{(h)}$, output of i-th hidden neuron is $o_{hi}(t)$ and the weight from i-th neuron to k-th output neuron is $w_{ki}^{(o)}$.

Define $\Theta = [\theta_1 \quad \dots \quad \theta_{n_\theta}]^{\mathrm{T}}$ for all weights and thresholds. The model can be shown as

$$\hat{y}(t,\Theta) = \hat{f}(x(t);\Theta) = \left[\hat{f}_1(x(t);\Theta) \quad \dots \quad \hat{f}_m(x(t);\Theta)\right]^{\mathrm{T}} \tag{7}$$

And

$$\begin{aligned} \hat{y}_k(t,\Theta) \quad &= \hat{f}_k(x(t);\Theta) = \sum_{i=1}^{n_h} w_{ki}^{(o)} o_{hi}(t) \\ &= \sum_{i=1}^{n_h} w_{ki}^{(o)} a\left(\sum_{j=1}^{n} w_{ij}^{(h)} x_j(t) + \mu_i^{(h)}\right),\ 1 \leqslant k \leqslant m \end{aligned} \tag{8}$$

Neural networks need to choose different activation functions for training according to the purpose of use. In this paper the activation function $a(\cdot)$ is chosen as

$$a(z) = \frac{1}{1 + \exp(-z)} \tag{9}$$

and the derivative function is

$$a'(z) = a(z)[1 - a(z)] \tag{10}$$

The gradient [12] of $\hat{y}(t, \Theta)$, for $1 \leqslant i \leqslant n_\theta$ and $1 \leqslant j \leqslant$ m is

$$\Psi_{ij}(t, \Theta) = \frac{d\hat{y}_j(t, \Theta)}{d\theta_i}$$

$$= \begin{cases} o_{hk}(t) & \text{if } \theta_i = w_{jk}^{(0)}, 1 \leqslant k \leqslant n_h \\ w_{jk}^{(0)} o_{hk}(t)\big(1 - o_{hk}(t)\big) & \text{if } \theta_i = \mu_k^{(h)}, 1 \leqslant k \leqslant n_h \\ w_{jk}^{(0)} o_{hk}(t)\big(1 - o_{hk}(t)\big) x_l(t) & \text{if } \theta_i = w_{kj}^{(h)}, 1 \leqslant k \leqslant n_h, \\ & 1 \leqslant l \leqslant n \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

The mean squared error between the network output and the target can be written as

$$E_D = mse = \frac{1}{N_D} \sum_{i=1}^{N_D} \varepsilon(t, \Theta)^2 = \frac{1}{N_D} \sum_{i=1}^{N_D} (y(t) - \hat{y}(t, \Theta))^2 \tag{12}$$

where N_D represent the number of output data. The least-squares solution of model is obtained by minimizing E_D. But in some times the result will fall into one of local minima. In order to make the result be global minimum, regularization is required. Bayesian regularization minimizes the mean squared error by updating the weights as network changes. The method used in this paper is to determine the ageing severity of wind turbines by comparing the weight changes between different models, so more accurate weights are very important for research. A diagonal matrix Λ is introduced, when it has constant diagonal elements λ, cost function S(w) will be minimized with respect to the weights:

$$S(\Theta) = E_D + \lambda \sum_{j=1}^{N_P} \left(w_{kj}^{(h)}\right)^2, \text{where } 0 \leq \lambda \leq 1 \tag{13}$$

where N_p is the number of parameters (number of neurons).

$$\Delta\theta_i = -\alpha \frac{\partial S(\Theta)}{\partial \theta_i} = -\alpha \left(\frac{2}{N_D} \sum_{i=1}^{N_D} \varepsilon(t, \Theta) \Psi_{ij}(t, \Theta) + 2\lambda\delta \right) \tag{14}$$

Only when $\theta_i = w_{kj}^{(h)}, \delta = 1$, otherwise the value of δ is 0. Then the parameters of the model can then be updated as

$$\theta_i^{\text{new}} = \theta_i^{\text{old}} + \Delta\theta_i \tag{15}$$

4 Experiment

The experiment is mainly to test whether the ageing degree of the system can be judged by the NARX model. First, a healthy wind turbine model needs to be built, and then the different ageing conditions are simulated by changing the corresponding parameters to collect data. This is because the actual data is difficult to record all the ageing conditions completely, and the damaged parts are also uncontrollable. Therefore, data collection needs to be assisted by simulation experiments. Afterwards, the collected data is used for system identification, which is used for ageing degree analysis.

4.1 Reference mode

The experimental object of this experiment is a 2MW wind turbine. The establishment of health model is by modifying the parameters of the wind farm model in MATLAB. By comparing the actual data and the simulation data, it can be seen that the simulation data is basically consistent with the actual data trend. Therefore, the data from simulation model can be used instead of the actual data.

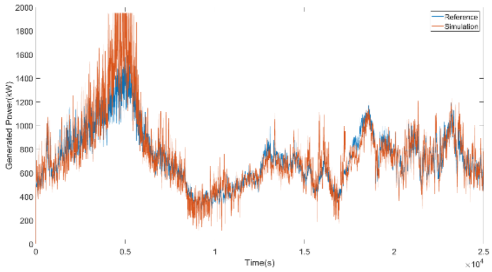


Fig. 4. The validation of health wind turbine model.

4.2 Ageing cases setting

In order to analyses whether this method can be used in different parts of wind turbine, four cases were selected for simulation. For Case 1, the changes of mechanical power are used to simulate the ageing of rotor-side converter control system. For Case 2, the change of magnetizing inductance are used to simulate the generator coils ageing. For both cases 3 and 4 are used to simulate the blade part. The changes of change rate and controller gain are used to simulate the ageing of blade bearing and pitch control system, respectively.

To describe the gradual ageing of wind turbine, 4 to 5 different stages are designed for each case. From the reality, with ageing, the performance of wind turbine will gradually decrease. Corresponding to the model, that is, the parameters will gradually change. Therefore, only the parameter reduction cases are considered during the experiment. In order to minimize the influence of noise on the experimental results, each experiment is repeated 5 times to obtain the average results. Other data that will be used in the experiment are shown in the Table 1. The input signal is wind speed, and output is generated power.

Table 1. Experimental setup parameters

Name	
Mechanical power	0.895(pu)
Constant wind speed	10.5m/s
Square wind speed	7 to 11m/s
Sampling rate	1s

The input signal used by case 1 and 2 is constant wind speed. Since case 3 and 4 involve pitch control, they use gust with a period of 120s as input to repeatedly collect the situation when blades rotate.

4.3 Validation

In order to verify whether the model obtained by system identification can replace the original model, it is necessary to make a validation. It can be seen from Fig.5 that the results obtained by the data-driven model are basically consistent with those of the original model when the same wind speed is input. This shows that the data-driven model obtained by system identification can replace the original model.

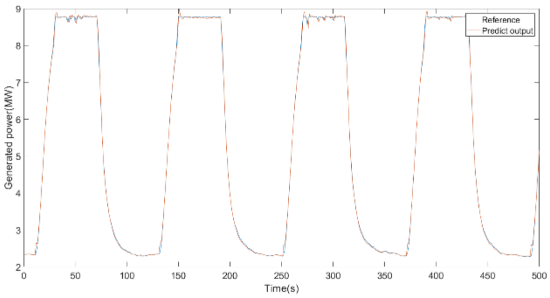


Fig. 5. The validation of pontificated model.

4.4 Different model comparison

By using different models for system identification, the mean square error (MSE) of the models obtained is shown in the Table 2. It can be seen that for all cases, the MSE of NARX is the smallest. This also shows that compared with the other two models, the predicted value calculated by NARX is closer to the real value.

Table 2. Comparison of different model

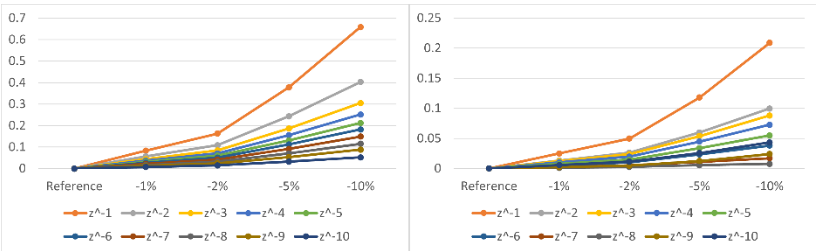
	ARX	ARMAX	NARX
Case 1	0.001267	0.001665	0.001013
Case 2	0.001267	0.001055	0.001013
Case 3	0.004635	0.004709	0.003292
Case 4	0.004608	0.00469	0.003364

5 Results

Data collection was performed by the model after reducing parameters by 1%, 2%, 5%, and 10%. Different order of delay and the layer number of the neural network will cause different results. Using the acquired data for system identification, the results obtained are shown below.

5.1 Case 1: Mechanical power

First tried the case where input and output delay order are 10 and the hidden layer is 1. The model obtained at this time is the simplest structure of the model. After system identification and data processing, the results shown in the Fig. 6 can be obtained. The results shown in the figure are the changes in model parameters. It can be seen that as the ageing degree increased, the variation of parameters also increased.

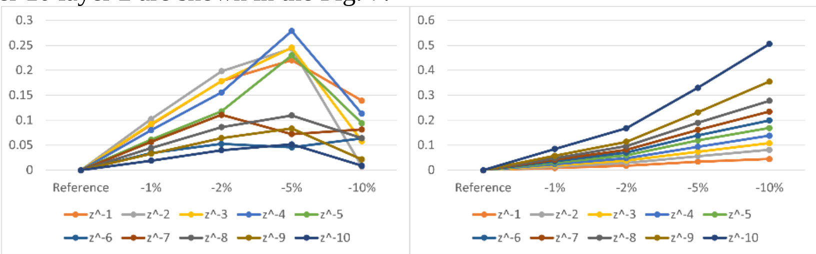


(a) input parameter (b)output parameter

Fig. 6. The variations of parameters with the mechanical power changes (10 orders, 1 layer).

The change in parameters is proportional to the degree of ageing. That means, if there is a problem with the mechanical power part, the ageing degree can be roughly determined by comparing the parameters of model. Although the changing trends of all parameters are consistent, it can be clearly seen that the change rate of each parameter is different. In theory, even if the components of other parts are also ageing, the change rate of parameters will be different. Therefore, it should be possible to identify the ageing site by the parameter change rate.

In addition, in order to observe the results of experiments with different delay orders and hidden layers, several other experiments were performed. The experimental results of order 10 layer 2 are shown in the Fig. 7.



(a) neuron 1 (b) neuron 2

Fig. 7. The variations of input parameters with the mechanical power changes (10 orders, 2 layers).

It can be seen that due to the increase of hidden layer, the number of input parameters has changed from 10 to 20. While all parameter variables in neuron 2 are still proportional, the parameter changes in neuron 1 do not always follow the same trends. With the increase of hidden layer, the change of parameters does not become more regular, but becomes more chaotic, 5 layers result shown in Fig. 8. This may be because, although the wind turbine is a nonlinear system, it can still be approximated as a linear model under some specific conditions. In this case, the simple model can have better accuracy than the complex model.

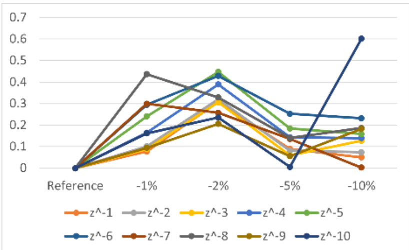
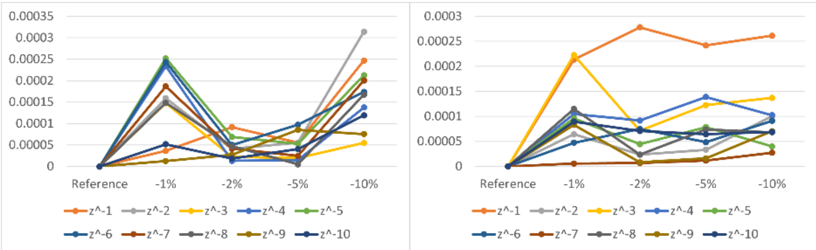


Fig. 8. The variations of input parameters with the mechanical power changes (10 orders, 5 layers).

5.2 Magnetizing inductance changes

First, as in case 1, consider the results in a linear system. The results are show in Fig. 9. However, unlike case 1, in the one-layer model, the change trend of the parameters is not consistent with the change trend of the magnetizing inductance. This may be due to the fact that magnetizing inductance also affects the reactive power, but this was not taken into account during system identification. But the same as case1, as the hidden layer increases, the parameter changes become more chaotic.

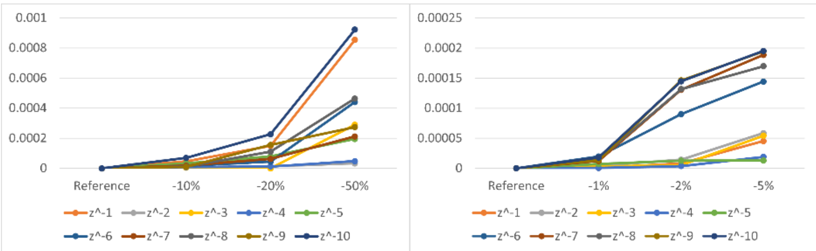


(a) neuron 1 (b) neuron 2

Fig. 9. The variations of input parameters with the magnetizing inductance changes.

5.3 Pitch angle changes

Since this case involves pitch control, in order to make the model identified by the system more accurate, the multiple-input single-output (MISO) NARX model is used for system identification. Take wind speed and pitch angle as input and generated power as output. Also consider the data analysis results under only one layer at first, and the results are shown in the Fig. 10. The change of parameters is consistent with the change trend of pitch angle.

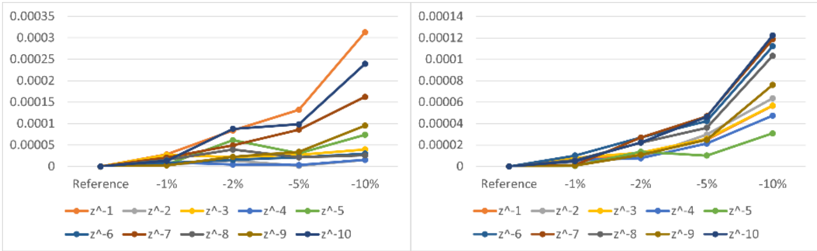


(a) input parameter (b)output parameter

Fig. 10. The variations of parameters with the pitch angle changes.

5.4 Pitch controller gain changes

This experiment also uses the MISO model, and the results are shown in the Fig. 11. In the case of linear models, it can be seen that although most of the parameter changes are proportional. But there are also individual parameter changes that are inconsistent with changes in controller gain, such as the fifth delay.



(a) input parameter (b)output parameter

Fig. 11. The variations of parameters with the pitch controller gain changes.

6 Conclusion

This paper uses a data-driven approach to assess wind turbine system ageing. For the simulation results, all ageing conditions can be detected when the model has only one hidden layer. However, in case 2, only the ageing of the system can be found, and the severity of the ageing cannot be confirmed. In other cases, the degree of ageing can be reflected in changes in model parameters. In these cases, the degree of ageing can be confirmed while the fault is detected, and corresponding warnings or maintenance advice can be issued. But all experiments perform poorly when the hidden layer is greater than 1. This may be because, under certain conditions, the local linear model is more accurate than the nonlinear model.

Under certain conditions, the parameters of the data-driven model can still reflect the ageing of wind power system. This work validates them under these operating conditions. At the same time, this method could theoretically be used for other kinds of wind turbines. Future work will also consider how to analyse the ageing combination of different components.

References

1. Li J, He D, He D X, et al. Overview of worldwide wind power industry[J]. Strategies of Sustainable Development in China's Wind Power Industry, 2020: 29-60.
2. Lux B, Pfluger B. A supply curve of electricity-based hydrogen in a decarbonized European energy system in 2050[J]. Applied Energy, 2020, 269: 115011.
3. Byrne R, Astolfi D, Castellani F, et al. A study of wind turbine performance decline with age through operation data analysis[J]. Energies, 2020, 13(8): 2086.
4. Dai J, Yang W, Cao J, et al. Ageing assessment of a wind turbine over time by interpreting wind farm SCADA data [J]. Renewable energy, 2018, 116: 199-208.
5. Staffell I, Green R. How does wind farm performance decline with age?[J]. Renewable energy, 2014, 66: 775-786.

6. Carroll J, McDonald A, McMillan D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines [J]. Wind Energy, 2016, 19(6): 1107-1119.
7. Yang W, Tavner P J, Crabtree C J, et al. Wind turbine condition monitoring: technical and commercial challenges [J]. Wind Energy, 2014, 17(5): 673-693.
8. Tran M Q, Li Y C, Lan C Y, et al. Wind Farm Fault Detection by Monitoring Wind Speed in the Wake Region[J]. Energies, 2020, 13(24): 6559.
9. Caselitz P, Giebhardt J, Krüger T, et al. Development of a fault detection system for wind energy converters[J]. 1996.
10. Bennouna O, Héraud N, Camblong H, et al. Diagnosis of the doubly-fed induction generator of a wind turbine[J]. Wind Engineering, 2005, 29(5): 431-447.
11. Lu L, Jin P, Karniadakis G E. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators[J]. arXiv preprint arXiv:1910.03193, 2019.
12. Chen S, Billings S A, Grant P M. Non-linear system identification using neural networks [J]. International journal of control, 1990, 51(6): 1191-1214.
13. Singh M, Muljadi E, Jonkman J, et al. Simulation for wind turbine generators--with FAST and MATLAB-simulink modules[R]. National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014.
14. Herp J, Ramezani M H, Bach-Andersen M, et al. Bayesian state prediction of wind turbine bearing failure[J]. Renewable Energy, 2018, 116: 164-172.
15. Muller S, Deicke M, De Doncker R W. Doubly fed induction generator systems for wind turbines[J]. IEEE Industry applications magazine, 2002, 8(3): 26-33.
16. Ngamroo I. Review of DFIG wind turbine impact on power system dynamic performances[J]. IEEJ Transactions on electrical and electronic engineering, 2017, 12(3): 301-311.