International Journal of Automation and Computing 12(1), February 2015, 25-34
DOI: 10.1007/s11633-014-0863-9

Model-based and Fuzzy Logic Approaches to Condition Monitoring of Operational Wind Turbines

Philip Cross Xiandong Ma

Engineering Department, Lancaster University, Lancaster LA14YR, UK

Abstract: It is common for wind turbines to be installed in remote locations on land or offshore, leading to difficulties in routine inspection and maintenance. Further, wind turbines in these locations are often subject to harsh operating conditions. These challenges mean there is a requirement for a high degree of maintenance. The data generated by monitoring systems can be used to obtain models of wind turbines operating under different conditions, and hence predict output signals based on known inputs. A model-based condition monitoring system can be implemented by comparing output data obtained from operational turbines with those predicted by the models, so as to detect changes that could be due to the presence of faults. This paper discusses several techniques for model-based condition monitoring systems: linear models, artificial neural networks, and state dependent parameter "pseudo" transfer functions. The models are identified using supervisory control and data acquisition (SCADA) data acquired from an operational wind firm. It is found that the multiple-input single-output state dependent parameter method outperforms both multivariate linear and artificial neural network-based approaches. Subsequently, state dependent parameter models are used to develop adaptive thresholds for critical output signals. In order to provide an early warning of a developing fault, it is necessary to interpret the amount by which the threshold is exceeded, together with the period of time over which this occurs. In this regard, a fuzzy logic-based inference system is proposed and demonstrated to be practically feasible.

Keywords: Condition monitoring, wind turbines, artificial neural network, state dependent parameter model, fuzzy logic.

1 Introduction

Wind farms are being constructed increasingly offshore to take advantage of stronger and more reliable winds, and reduced visual and noise impacts[1]. However, routine inspection and maintenance is more difficult. Over an operating life of 20 years, maintenance costs for an onshore wind farm are estimated to be 10–15 % of the total income[2]. And for an offshore wind turbine over the same period, these costs are estimated to be up to 30 % of the total income[3]. Consequently, condition monitoring (CM) systems play an important role in the operation of wind farms, providing information about the past and current conditions of the turbines and enabling optimal scheduling of maintenance activities, while minimising the risk of unexpected failure[4]. The data generated by these monitoring systems can be used to obtain models of a process operating under different conditions, e.g., these data can be acquired from a supervisory control and data acquisition (SCADA) system[5]. SCADA systems are used to monitor and control devices in many industrial applications, such as communications, water and waste control, energy, oil and gas refining, and transportation. The system comprises an interface to sensors, local control switchboxes, actuators, a communication network to transfer data to the SCADA central host, and the associated software.

The output signals of a process can be predicted from known inputs using models obtained from the data generated by these monitoring systems, and a CM system can be implemented by comparing actual output data with those predicted by the model for given input signals. Differences between the output signals are caused by changes in the process, possibly due to the occurrence of faults[6]. The model-based method is illustrated in Fig. 1, in which the residual signal can reveal potential component failures.

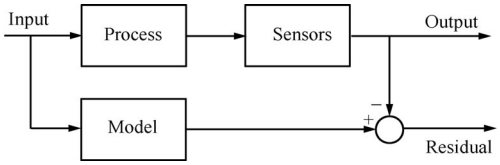

Fig. 1 Model-based condition monitoring system

Clearly, an accurate model is essential for such a system, and previous research in this area has employed a range of techniques, based upon both mechanistic models derived from physical laws[7] and data-based models obtained from measured input and output signals[8]. The former require a thorough understanding of the process, and may result in a complex or over-parameterised model not suitable for online implementation as a CM system due to an excessive response time. In contrast, data-based models do not require knowledge of the process or specific parameters. Models are identified from input and output signals collected either during planned experiments or by monitoring the process

Regular paper
Special Issue on Recent Advance in Automation and Computing
Manuscript received December 24, 2013; accepted September 24, 2014
This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) (No. EP/I037326/1).
Recommended by Associate Editor Yi Cao

during normal operation. However, similar to mechanistic models, it is essential to identify a low-order model for on-line implementation.

Many processes associated with wind turbines are non-linear, and techniques such as artificial neural networks (ANNs)[9] and fuzzy logic[10] have been utilised by many researchers for model-based CM systems. These techniques are particularly suitable for this application, being robust to noisy, incomplete, and uncertain data.

This paper extends the previous research into model-based methods for wind turbine condition monitoring[11], and discusses linear, ANN-based, and state dependent parameter (SDP)[12] modelling techniques. These methods are used to model the relationships between critical measured parameters identified with SCADA data acquired from an operational wind farm. It is found that multivariate SDP models outperform both linear and ANN-based models. Further, in contrast to ANN-based models, SDP models are a parametrically efficient representation of non-linear processes, and are particularly suitable for forecasting[13] and providing the basis for automatic controller design[14]. Subsequently, in this paper, multivariate SDP models are used to identify adaptive thresholds for selected critical output signals. In turn, these thresholds form the basis of a CM system incorporating fuzzy logic that provides an early warning of component failure and an indication of its severity.

2 Wind farm data

Many commercial wind farms employ a SCADA system to provide online information regarding various measurements and signals. These systems allow data to be acquired without the need to install additional instrumentation. Therefore, SCADA data have been employed widely by researchers as the basis for CM systems[8, 9].

For this research, the SCADA data have been acquired from a wind farm comprising 26 turbines and representing 16 months′ operation. These data have been collected by the SCADA system at a sample rate of two seconds, then processed and stored at ten minute intervals. They consist of 128 readings for various temperatures and pressures, power output, vibration, wind speed, and digital control signals. For the work presented here, the data selected represent ten minute averages for each parameter. Examples of the SCADA data are illustrated in Fig. 2, in which the wind speed and the corresponding active power output for three months′ operation for a single wind turbine are shown.

Gaps in SCADA data exist due to occasions when the turbine is inactive during periods of low and high wind speeds. Additional gaps occur due to the occurrence of scheduled maintenance and faults. Prior to model identification, it is necessary to remove these gaps. Fig. 3 shows the data after processing. After removal of the gaps in the data, a power curve for the turbine can be plotted, showing the relationship between wind speed and active power output, as illustrated in Fig. 4.

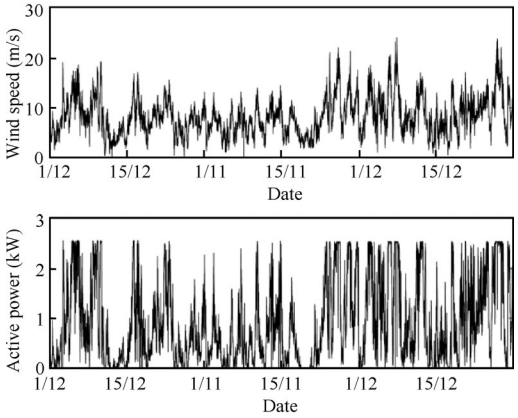

Fig. 2 Example of SCADA data

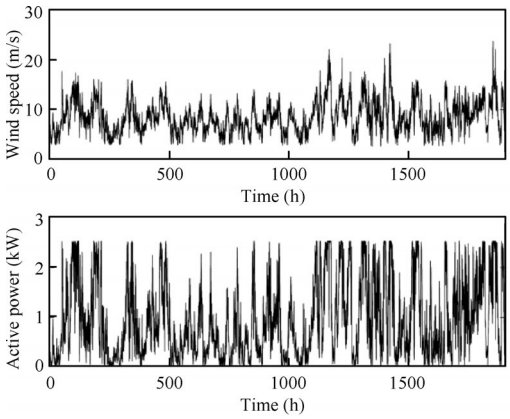

Fig. 3 Example of processed data

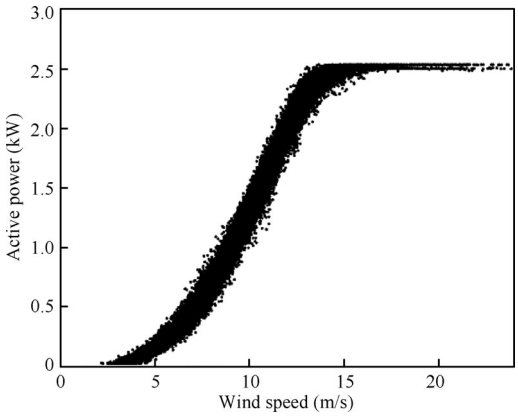

Fig. 4 Example of a power curve for a wind turbine showing the active power output as a function of wind speed

3 System identification

A model of a dynamic system can be obtained from the measured input and output data collected from either planned experiments or by monitoring the system during normal operation. These data can be utilised to identify the appropriate model structure and parameters using system identification, or to form the training set for an ANN.

3.1 Linear models

A dynamic linear system can be represented as a time invariant transfer function model. In addition, the small perturbation behavior about an equilibrium point of many non-linear systems can be approximated by a linear transfer function model. Considering a single-input single-output process sampled at a uniform interval Δt, beginning at time $t = 0$, the resulting discrete-time system can be represented by the following difference equation, describing the output variable at sample k in terms of a combination of its past values and the current and past values of the input variable:

$$\begin{aligned} y_k = \; & -a_1 y_{k-1} - a_2 y_{k-2} - \cdots - a_n y_{k-n} + b_0 u_k + \\ & b_1 u_{k-1} + \cdots + b_m u_{k-m} \end{aligned} \tag{1}$$

where y_k and u_k are the k-th sampled output and input variables, respectively. By defining the "backward shift operator" as $z^{-i} y = y_{k-i}$, where i is a delay in number of samples, the difference equation can be represented by the discrete transfer function model as

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_k \tag{2}$$

in which $A(z^{-1})$ and $B(z^{-1})$ are appropriately defined polynomials in the backward shift operator:

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + \cdots + a_n z^{-n} \\ B(z^{-1}) &= b_0 + b_1 z^{-1} + \cdots + b_m z^{-m}. \end{aligned} \tag{3}$$

A time delay of δ samples can be incorporated into the model by setting the leading δ coefficients of the $B(z^{-1})$ polynomial to zero.

This research employs the simplified refined instrumental variable (SRIV) algorithm for discrete-time transfer function model identification[15]. This algorithm employs an iterative procedure in which each step allows for a linear-like computation of recursive estimates. After the initial estimate of the model polynomials, an estimate of the noise signal is obtained by least squares. Pre-filtered input and "noisy" output signals, together with a pre-filtered estimate of the noise-free system output, are derived from the estimates obtained in the previous step. These are then used to obtain estimates of the model polynomials and the process is repeated with these updated estimates. Although the above describes the identification of a single-input single-output system, it is straightforward to extend the methodology to multiple-input single-output systems.

3.2 Artificial neural networks

Many processes associated with wind turbines are non-linear. Although linear models can predict the output of many non-linear processes over a small operating range, this may be inadequate for model-based CM systems. In this regard, non-linear models based upon artificial neural networks have been utilised by many researchers[9]. An ANN is a mathematical model inspired by the structure of biological neural networks, consisting of an interconnected group of "neurons" arranged in a series of layers. These layers comprise input and output layers, and one or more "hidden" layers.

ANNs are "trained" to perform a particular task by changing the weights associated with the connections between the neurons[16]. During the training process, the weightings are adjusted until the output data from the ANN match the target output data using a supervised learning technique. In this research, a back-propagation algorithm is employed, in which the input data are mapped to the target output data by minimising the error between these outputs and the predicted outputs using a damped least squares method. The network is validated using the input data not used for training. These data are entered into the trained network to see how well the corresponding output values are predicted and the weightings between neurons are changed if necessary. Finally, the network is tested with a further set of input data not used for either training or validation.

ANNs are categorised as static networks and dynamic networks. The former have no feedback elements and contain no delays. In contrast, the output of a dynamic ANN depends on the current input to the network and the previous inputs, outputs or states of the network. A dynamic ANN can be used to model the relationship between input and output data for a non-linear system. Such a system can be described by a non-linear autoregressive function with exogenous inputs, in which the output signal is regressed on its delayed values together with the current and delayed values of the input signal[17], i.e.,

$$y_k = f(y_{k-1},\, y_{k-2}, \cdots,\, y_{k-n},\, u_k,\, u_{k-1}, \cdots,\, u_{k-m}) \tag{4}$$

where y_k and u_k are defined as in (1). The function f can be approximated using a dynamic ANN, in which the connections between the neurons form a directed cycle, and feedback connections enclose several layers of the network, creating an internal state for the network. It has been shown that one hidden layer of neurons is sufficient to approximate any continuous function[18], and the number of neurons in the hidden layer is selected by considering the training time with respect to model fit.

3.3 State dependent parameter models

Since an ANN is purely "black box", it is difficult to interpret the identified models in a physically meaningful way. Therefore, modelling techniques have been developed to combine both linear and non-linear components, e.g., hybrid linear and non-linear autoregressive moving average[19] and quasi-auto-regressive moving average with exogenous variables[20] methods. In this paper, non-linear systems are modelled using a quasi-linear state dependent parameter model structure, in which parameters vary as functions of the state variables. Although the following describes a single-input single-output system, the methodology can be extended to multivariate systems.

A single-input single-output SDP system can be repre-

sented by the "pseudo" transfer function as

$$y_k = \frac{B_k(z^{-1})}{A_k(z^{-1})} u_k \tag{5}$$

where y_k and u_k are defined as in (1); and $A_k(z^{-1})$ and $B_k(z^{-1})$ are appropriately defined state dependent polynomials in the backward shift operator as

$$\begin{aligned} A_k(z^{-1}) &= 1 + a_{1,k}z^{-1} + \cdots + a_{n,k}z^{-n} \\ B_k(z^{-1}) &= b_{0,k} + b_{1,k}z^{-1} + \cdots + b_{m,k}z^{-m}. \end{aligned} \tag{6}$$

The state dependent coefficients, $a_{i,k}$ and $b_{j,k}$ ($i = 1, 2, \cdots, n$; $j = 0, 1, \cdots, m$), are derived from the input and output signals, but can be functions of measured state variable. Although SDP models cannot represent every type of non-linear behavior, they are applicable to a wide range of systems. The coefficients, $a_{i,k}$ and $b_{j,k}$ are direct functions of the state of the system. Subsequently, the parameters can change rapidly, and they will vary chaotically in the case of a chaotic system.

The SDP model structure and parameters are identified directly from the input and output data of the process, and the method of model identification employed here comprises three stages[21]. In the first stage, the most appropriate model structure and the state variables are identified by estimating a discrete-time linear transfer function model using any suitable method, such as the SRIV algorithm described above. The progression of the state vector associated with the state dependent coefficients can be described by a generalised random walk. This is a statistical mechanism allowing the change in the parameters to be estimated.

During the second stage of model identification, a stochastic time-varying parameter (TVP) model is estimated using a forward pass recursive Kalman filter[22] and a backward pass fixed-interval smoothing (FIS) algorithm[23]. The backward pass FIS is required because the parameter estimates generated from the forward pass filtering are lagged. Since the variations in the SDPs are functions of the state variables, the process may display severely nonlinear or chaotic behavior. Subsequently, these standard recursive estimation algorithms will not work satisfactorily. However, by sorting the data in a non-temporal order, e.g., in ascending order so that the variations in the SDPs are slower and smoother, the standard methods can be used. This recursive process continues until the forward pass estimates of the TVPs no longer change significantly. In order to apply this algorithm, it is necessary to specify the elements of the covariance matrix associated with the TVP model and the parameters of the random walk model. As these are not known a priori, they must be estimated using numerical optimisation. A straightforward maximum likelihood approach is not possible due to the non-stationary nature of the model, and therefore it must be used together with prediction error decomposition[24].

For the final identification stage, the resulting non-parametric estimates of each parameter are unsorted and plotted as a function of the variable upon which it depends, providing an indication of the location of any significant non-linearities. The non-parametrical estimates are then parameterised in terms of their associated dependent variables. In this paper, this is achieved by defining the parametric model as a polynomial function of the state variables, and the associated coefficients are obtained using standard numerical optimisation to minimise the error between the polynomial function and the non-parametric graph. The resulting model provides a parametrically efficient representation of the system, suitable for on-line implementation. It is necessary only to store the current input signal and the delayed state variables, together with the arithmetic expressions required to calculate the dependent parameters.

3.4 Coefficient of determination

The coefficient of determination[25] is employed in this paper as a measure of how well a model explains the actual output data, and is defined as

$$R_T^2 = 1 - \frac{\sigma_e^2}{\sigma_y^2} \tag{7}$$

where σ_e^2 is the sample variance of the model residuals, and σ_y^2 is the variance of the actual output. If the variance of the model residuals is low compared with the variance of the actual output, R_T^2 tends to unity, indicating the model provides a good fit. If the variances are similar in magnitude, R_T^2 tends to zero, indicating a poor fit. And if the model fit is very poor, the variance of the model residuals will be high compared with the variance of the actual output, generating a negative value for R_T^2.

4 System identification results

Models describing the relationship between wind speed and active power output, the temperature of the gearbox bearing, and the temperature of the generator winding have been identified from SCADA data obtained from a fault-free turbine. These relationships have been selected since they provide an early indication of generator, bearing, and gearbox faults.

4.1 Linear model results

After removal of baseline data, the SRIV algorithm is employed to identify the most appropriate model structure and associated parameters for the linear transfer function. For the relationship between wind speed and active power output, the best performing model is a second-order transfer function with one time delay:

$$y_k = \frac{b_1 z^{-1}}{a_1 z^{-1} + a_2 z^{-2}} u_k \tag{8}$$

in which y_k and u_k are defined as in (1), a_1, b_1, and b_2 are the model parameters.

Fig. 5 shows the fit of this model to a section of the validation data. It can be seen that the model fit deteriorates at the upper and lower extents of this relationship. The

residual between the model prediction and the actual active power output is also shown in Fig. 5.

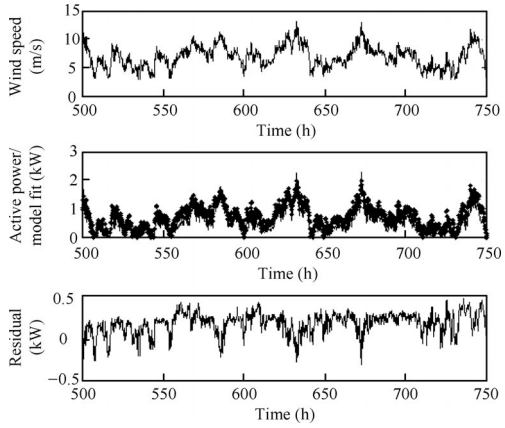

Fig. 5 Linear model response compared to SCADA data. Upper plot: wind speed – model input. Middle plot: active power (thin trace) and model output (dots). Lower plot: residual signal between SCADA data and model prediction.

4.2 Artificial neural network results

In order to identify ANN-based models, it has been assumed that the output signal is dependent upon the delayed output and input signals and the current value of the input signal. The number of neurons in the hidden layer is selected by considering training time with respect to model fit.

Fig. 6 shows the ANN-based model fit to a section of the validation data set for the relationship between wind speed and active power output. The residual between the model prediction and the actual active power output is shown in the lower plot of Fig. 6. It can be seen that the model is able to estimate successfully the output signal.

4.3 State dependent parameter model results

It has been assumed that the time varying model parameters for each relationship are dependent upon the delayed output and input signals, together with the current value of the input signal. The structure for each SDP model has been obtained from the equivalent linear transfer function model.

The parameterisation of the relationships between the state variables and the associated dependent parameters in the identified models has been achieved by defining the parametric models as polynomial functions of the state variables, and the associated coefficients are obtained using standard numerical optimisation.

The following "pseudo" transfer function has been obtained for the relationship between wind speed and active power output.

$$y_k = \frac{b_{1,k} z^{-1}}{a_{1,k} z^{-1} + a_{2,k} z^{-2}} u_k. \tag{9}$$

The state dependent parameter $b_{1,k}$ is defined by a polynomial function of the input signal u_k, and the state dependent parameters $a_{1,k}$ and $a_{2,k}$ are defined by polynomial functions of the output signal y_k.

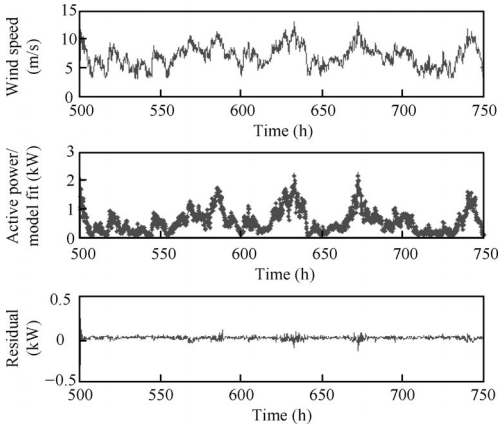

Fig. 6 ANN model response compared to SCADA data. Upper plot: wind speed – model input. Middle plot: active power (thin trace) and model output (dots). Lower plot: residual signal between SCADA data and model prediction.

Fig. 7 shows the fit of the SDP model to a section of the validation data. The residual between the model prediction and the actual active power output is shown in the lower plot in Fig. 7. It is clear that the SDP model is better able to predict the whole range of the output signal in comparison to the linear model. However, several peaks in the signal are over estimated.

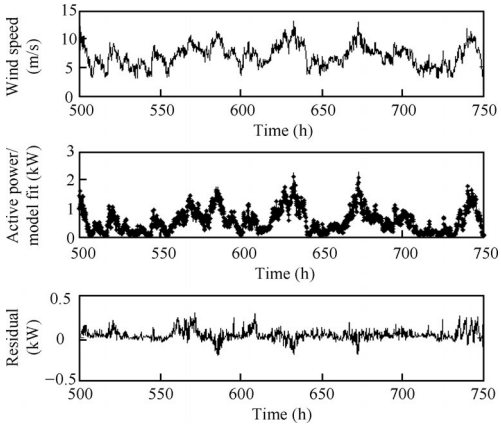

Fig. 7 SDP model response compared to SCADA data. Upper plot: wind speed – model input. Middle plot: active power (thin trace) and model output (dots). Lower plot: residual signal between SCADA data and model prediction.

4.4 Summary of single input model results

The linear transfer function models relating wind speed to the gearbox bearing temperature, the generator winding temperature and active power output obtained using the SRIV method are second-order, first-order and second-order, respectively, each with one numerator parameter and one time delay. The structures of the SDP models are the same as the linear models with the parametric models de-

scribed by second-order polynomial functions of the associated state variables. The ANN consists of a hidden layer of ten neurons, each represented by a sigmoidal transfer function and an output layer of one neuron represented by a linear transfer function. The coefficients of determination for these models are summarised in Table 1.

It is clear that the ANN-based models outperform both the linear and SDP methods for single-input single-output models.

Table 1 R_T^2 values for single-input models

Model output	Linear	SDP	ANN
Gearbox bearing temperature	0.657	0.992	0.999
Generator winding temperature	0.580	0.908	0.989
Active power output	0.786	0.961	0.995

4.5 Multivariate model results

The temperatures of the gearbox bearing and generator winding depend not only upon the wind speed, but also upon the power demand from the grid. Consequently, this subsection considers the temperature of the gearbox bearing as a function of wind speed and active power output, together with the temperature of the generator winding as a function of wind speed and active power.

As with the single-input single-output relationships obtained in the previous subsection, the data are preprocessed by removing the gaps that occur when the turbine is inactive. The multivariate versions of the identification methods are then used to obtain models for these relationships.

For the linear model, the temperature of the gearbox bearing as a function of wind speed and active power output is represented by two second-order transfer functions, each with one numerator parameter and one time delay and the same denominator parameters. The liner model for the generator winding temperature as a function of wind speed and active power output consists of two first-order transfer functions, each with one numerator parameter and one time delay and the same denominator parameters.

The structures of the SDP models are the same as the linear models with the parametric models again defined as second-order polynomial functions. As with the single input models, the ANN consists of a hidden layer of ten neurons represented by a sigmoidal transfer function as no improvement in model fit is obtained by increasing the number of neurons on the hidden layer. The output layer consists of one neuron, represented by a linear transfer function. The corresponding coefficients of determination are summarised in Table 2.

Table 2 R_T^2 values for multivariate models

Model output	Linear	SDP	ANN
Gearbox bearing temperature	0.710	0.997	0.992
Generator winding temperature	0.833	0.983	0.977

It can be seen that the fit of the linear and SDP models improves with the addition of the second input variable. In contrast, the fit of the ANN-based models appears to decrease. The relatively poor performance of the multiple-input ANN-based models is probably caused by "over fitting" of the models during the training period, due to a high correlation between the input parameters.

Fig. 8 illustrates the relationship between the SDP model prediction for bearing temperature as a function of wind speed and active power output. The relationship between the generator winding temperature, wind speed and active power output as obtained from the SDP model is shown in Fig. 9.

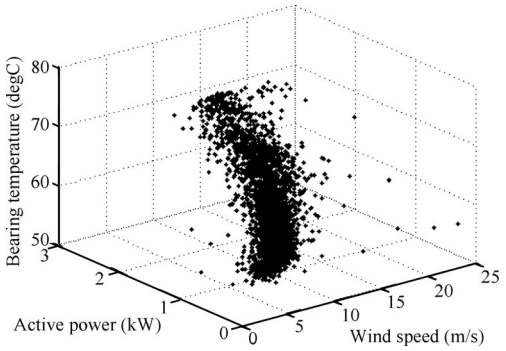

Fig. 8. Plot of gearbox bearing temperature as a function of wind speed and active power output, obtained from SDP model

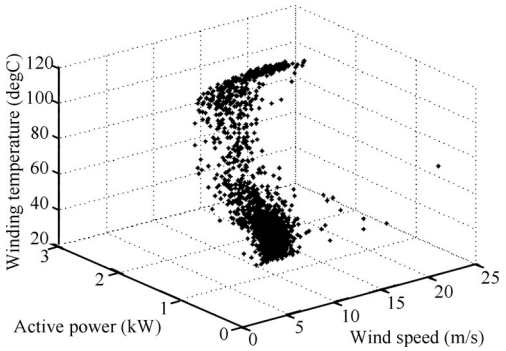

Fig. 9 Plot of winding temperature as a function of wind speed and active power output, obtained from SDP model

5 Fault detection using fuzzy logic

Power curves obtained from two turbines operating on the same wind farm are shown in Fig. 10. The upper subplot illustrates the power curve for a fault-free turbine. The lower subplot shows the power curve for the second turbine that operated with a reduced power output following a gearbox fault at some points.

In order to investigate whether the identified SDP model is able to detect changes due to the onset of the fault, the SCADA data leading up to the occurrence of the fault are acquired. The fault is identified from the event logs for the turbine and the corresponding data are shown in Fig. 11, with the fault occurring at 2900 minutes and the turbine being shut down immediately. The gearbox bearing temperature signal obtained from the SCADA data has been compared with the corresponding SDP model prediction obtained using active power output and wind speed as the

input signals, as shown in the lower plot of Fig. 11.

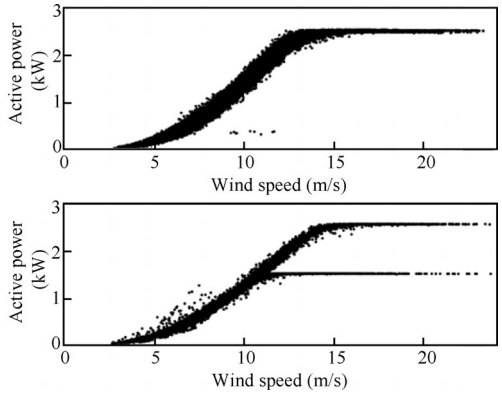

Fig. 10 Examples of power curves for turbines. Upper plot: curve for fault-free turbine. Lower plot: power curve for faulty turbine.

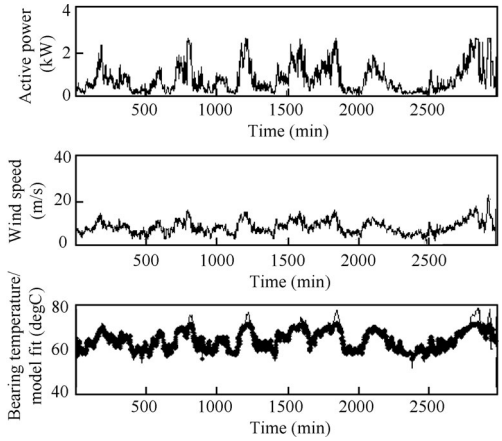

Fig. 11 Detection of gearbox bearing fault using SDP models. Upper plot: active power – model input 1. Middle plot: wind speed – model input 2. Lower plot: Gearbox bearing temperature from SCADA data (thin trace) and model output (dots).

Fig. 12 illustrates the residual signal, i.e., the difference between the SCADA data and model prediction. It can be seen from the plot that there are three large spikes in the residual signal, with deviations from the model prediction by more than 4 °C at 800, 1200, and 1800 minutes. Immediately before the fault occurring, there are two more large deviations of approximately 7 °C from the model prediction. It can be concluded that a warning system activated by the spikes in the model residual can be used to provide an early indication of the onset of the bearing fault.

5.1 Adaptive thresholds

The data illustrated in Fig. 11 are presented in the form of a three-dimensional graph in Fig. 13. A three-dimensional surface, defined by the upper limit of the bearing temperature as predicted by the SDP model, is also shown. This has been obtained using a threshold function written in Matlab together with a curve-fitting toolbox. The surface represents an adaptive threshold level for the bearing temperature. This multivariate method provides a more sensitive threshold in comparison to a conventional approach, which would derive the threshold for the bearing temperature based purely on a single input parameter. By monitoring the residual signal between the SCADA data and the adaptive threshold, it can be determined if there is potential for a fault to develop. The period of time over which the threshold is exceeded is also monitored in order to avoid false alarms caused by small fluctuations in the bearing temperature during normal operation, but not predicted by the model. This "trend analysis" is undertaken using a fuzzy logic-based inference system.

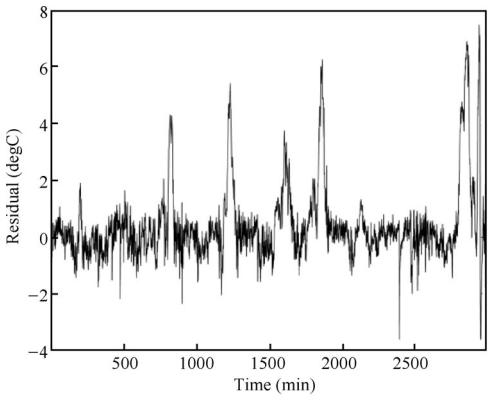

Fig. 12 Residual signal between SCADA data and model prediction

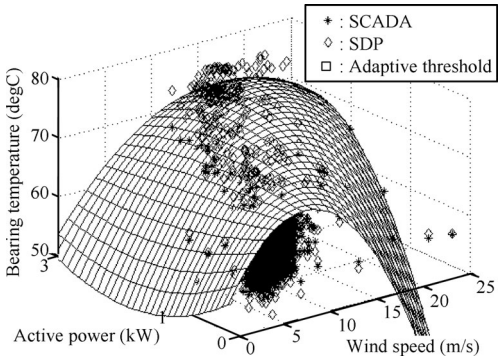

Fig. 13 Plot of gearbox bearing temperature as a function of wind speed and active power output

5.2 Fuzzy inference of thresholds

Fuzzy logic is a form of probabilistic logic that maps an input space to an output space using a set of if-then statements[26]. A membership function (MF) defines how each point in the input space is mapped to a fuzzy membership value that lies between 0 and 1. This is referred to as "fuzzification".

The MFs used for the fuzzification of the first input, representing the residual signal between the SCADA data and the adaptive threshold, are shown in Fig. 14. Each of the three MFs refers to the amount by which the threshold is exceeded: MF 1 corresponds to situations when the threshold is exceeded by a small amount (less than 3 °C); MF 2 corresponds to when the threshold is exceeded by a medium amount (between 1.5 °C and 6.5 °C); and MF 3 corresponds

to the threshold being exceeded by a large amount (more than 5 °C).

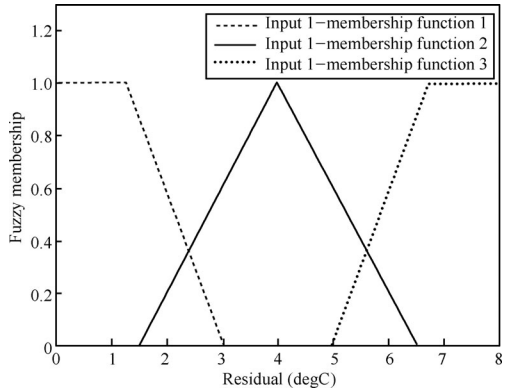

Fig. 14 Membership functions for input 1: residual signal between SCADA data and adaptive threshold

Fig. 15 shows the MFs used for the fuzzification of the second input, representing the period of time over which the threshold temperature is exceeded. In a similar fashion to the method employed for the fuzzification of the residual signal, each MF refers to the length of time the threshold is exceeded: MF 1 corresponds to the threshold being exceeded for a short period (less than 100 minutes); MF 2 corresponds to the threshold which is exceeded for a medium period (between 50 and 200 minutes); and MF 3 corresponds to the occasion when the threshold is exceeded for a long period of time (more than 150 minutes).

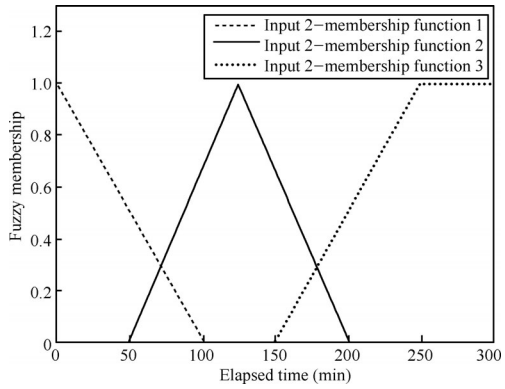

Fig. 15 Membership functions for input 2: period of time over which threshold temperature is exceeded

In order to define suitable levels of fault severity (i.e. no fault, low severity, medium severity, and critical), four MFs have been chosen for the fuzzy output signal. This output signal is obtained using the following Mamdani fuzzy rules[27]:

if (input 1 is MF 1) and (input 2 is MF 1) then (output is MF 1)

if (input 1 is MF 1) and (input 2 is MF 2) then (output is MF 1)

if (input 1 is MF 1) and (input 2 is MF 3) then (output is MF 2)

if (input 1 is MF 2) and (input 2 is MF 1) then (output is MF 1)

if (input 1 is MF 2) and (input 2 is MF 2) then (output is MF 2)

if (input 1 is MF 2) and (input 2 is MF 3) then (output is MF 3)

if (input 1 is MF 3) and (input 2 is MF 1) then (output is MF 2)

if (input 1 is MF 3) and (input 2 is MF 2) then (output is MF 3)

if (input 1 is MF 3) and (input 2 is MF 3) then (output is MF 4).

The MFs for the fuzzy output are shown in Fig. 16. An aggregate output fuzzy set is obtained by combining the outputs of the rules using a fuzzy "and" function.

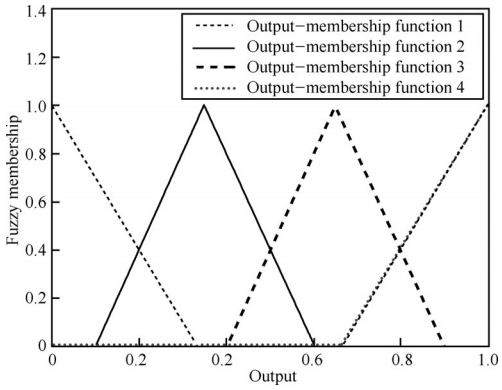

Fig. 16. Membership functions for fuzzy output

Finally, the aggregate output fuzzy set is "defuzzified" to obtain a single output value. Here, the "last-of-maximum" method is employed for the defuzzification process[28]. This method finds the last point of the maximum value described by the aggregate output fuzzy set.

Considering the residual signal shown in Fig. 12, the signals corresponding to the amount the threshold temperature is exceeded and the period of time over which the threshold temperature is exceeded are illustrated in Fig. 17. This figure also shows the defuzzified output obtained using the inference system.

Clearly, it is necessary for the output of the fuzzy inference system to correspond to distinct warning levels. In this case, a "traffic light" system is chosen to represent the severity of a fault, i.e., level 0 represents no fault, level 1 represents a low fault severity, level 2 represents medium severity, and level 3 represents a critical situation. The warning level is obtained from the defuzzified output using the following rules:

if (output $\leq$ 0.25) then (warning is level 0)
else if (0.25 $<$ output $\leq$ 0.5) then (warning is level 1)
else if (0.5 $<$ output $\leq$ 0.75) then (warning is level 2)
else (warning is level 3).

The corresponding warning output generated by the fuzzy inference system is shown in Fig. 18. It can be seen that a critical warning is generated 20 hours (i.e., 1200 minutes) prior to the fault. Indeed, the conditions for a critical warning are met twice during the period investigated. Fur-

ther, a level 1 warning is triggered four times and a level 2 warning is triggered once.

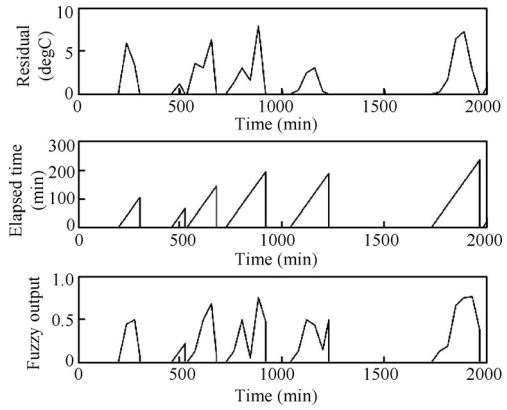

Fig. 17 Fuzzy inference system input and output signals. Upper plot: amount by which threshold is exceeded. Middle plot: elapsed time over which threshold is exceeded. Lower plot: defuzzified output signal.

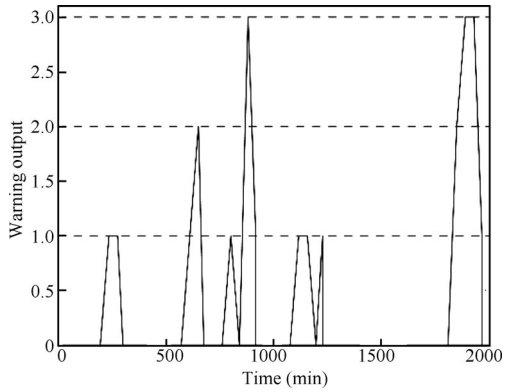

Fig. 18 "Traffic light" warning output signal

6 Conclusions

This paper has presented several methods for obtaining models for a condition monitoring system, and these have been demonstrated with SCADA data acquired from operational wind turbines.

It has been shown that the single-input models based upon artificial neural networks outperform both linear models and non-linear state dependent parameter models for the relationships between the input and output signals considered.

However, the temperatures of the gearbox bearing and generator winding depend not only upon the wind speed, but also upon the power demand from the grid. Consequently, multiple-input systems have been investigated. The goodness-of-fit of both linear models and state dependent parameter models is improved by increasing the number of input parameters. However, the corresponding multiple-input artificial neural network-based models perform relatively poorly in comparison. This is probably caused by "over fitting" of the models, due to a high correlation between the input parameters.

A multivariate adaptive threshold for gearbox bearing temperature has been obtained from the associated SDP model. In order to provide early warning of a developing fault, it is necessary to interpret the amount by which the adaptive threshold is exceeded together with the period over which this occurs. In this regard, a fuzzy logic method has been proposed. Clearly, the fuzzification of the input and output signals, the fuzzy rules, the defuzzification method, and the rules defining the warning levels may be adjusted in order to control the sensitivity of the system and hence tailor the response of the warning output.

In contrast to artificial neural network-based models, the SDP method employed in this paper provides a parametrically efficient representation of non-linear processes. When implemented on-line, it will be necessary only to store the current input signal and the delayed state variables, together with the arithmetic expressions required to calculate the dependent parameters. Other data-driven condition monitoring methodologies, such as multivariate statistical process control[29], rely on a large amount of historical data to be stored as charts. However, a condition monitoring scheme based upon the state dependent parameter methodology can be realised with a relatively low-powered, low-cost processor, and integrated into an existing monitoring system. In this regard, the current research is focusing upon a hardware implementation of the SDP model-based condition monitoring scheme employing a field programmable gate array (FPGA).

Acknowlegements

Permission to use the SCADA data obtained from Wind Prospect Ltd. is gratefully acknowledged.

References

[1] B. Snyder, M. J. Kaiser. A comparison of offshore wind power development in Europe and the US: Patterns and drivers of development. *Applied Energy*, vol. 86, no. 10, pp. 1845–1856, 2009.

[2] C. A. Walford. Wind Turbine Reliability: Understanding and Minimizing Wind Turbine Operation and Maintenance Costs, Technical Report SAND 2006-1100, Sandia National Laboratories, USA, 2006.

[3] G. J. W. van Bussel, C. Schöntag. Operation and maintenance aspects of large offshore wind farms. In *Proceedings of European Wind Energy Conference*, Dublin, Ireland, 1997.

[4] D. McMillan, G. Ault. Condition monitoring benefit for onshore wind turbines sensitivity to operational parameters. *IET Renewable Power Generation*, vol. 2, no. 1, pp. 60–72, 2009.

[5] D. Barr, P. M. Fonash. Supervisory Control and Data Acquisition (SCADA) Systems, Technical Information Bulletin NCS, Communication Technologies, Inc., Chantilly, Virginia, USA, 2004.

[6] D. Dvorak, B. Kuipers. Model-based monitoring of dynamic systems. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, San Francisco, CA, USA, pp. 1238–1243, 1989.

[7] M. Entezami, S. Hillmansen, P. Weston, M. P. Papaelias. Fault detection and diagnosis within a wind turbine mechanical braking system using condition monitoring. *Renewable Energy*, vol. 47, pp. 175–182, 2012.

[8] W. G. Garlick, R. Dixon, S. J. Watson. A model-based approach to wind turbine condition monitoring using SCADA data. In *Proceedings of the 20th International Conference on Systems Engineering*, Coventry University, Coventry, UK, 2009.

[9] M. Schlechtingen, I. F. Santos. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1849–1875, 2011.

[10] T. Üstüntaş, A. D. Şahin. Wind turbine power curve estimation based on cluster center fuzzy logic modelling. *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, pp. 611–620, 2008.

[11] P. Cross, X. D. Ma. Model-based condition monitoring for wind turbines. In *Proceedings of 19th International Conference on Automation and Computing*, IEEE, London, UK, pp. 1–7, 2013.

[12] M. B. Priestley. State-dependent models: A general approach to non-linear time series analysis. *Journal of Time Series Analysis*, vol. 1, no. 1, pp. 47–71, 1980.

[13] N. -V. Truong, L. Wang, P. K. C. Wong. Modelling and short-term forecasting of daily peak power demand in Victoria using two-dimensional wavelet based SDP models. *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 9, pp. 511–518, 2008.

[14] J. Gu, Z. H. Feng, X. Ma, J. F. Ni. Proportional-integral-plus control of robotic excavator arm utilising state-dependent parameter model. *Applied Mechanics and Materials*, vol. 48–49, pp. 1323–1327, 2011.

[15] P. C. Young. *Recursive Estimation and Time Series Analysis*, Berlin, Germany: Springer-Verlag, 1984.

[16] S. Haykin. *Neural Networks: A Comprehensive Foundation*, Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.

[17] I. J. Leontaritis, S. A. Billings. Input-output parametric models for non-linear systems Part I: Deterministic non-linear systems. *International Journal of Control*, vol. 41, no. 2, pp. 303–328, 1985.

[18] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.

[19] Y. Zheng, Z. Lin, D. B. H. Tay. State-dependent vector hybrid linear and nonlinear ARMA modeling: Theory. *Circuits Systems Signal Processing*, vol. 20, no. 5, pp. 551–574, 2001.

[20] J. Hu, K. Kumamaru, K. Inoue, K. Hirasawa. A hybrid quasi-ARMAX modeling scheme for identification of non-linear systems. *Transactions of the Society of Instrument and Control Engineers*, vol. 34, no. 8, pp. 977–985, 1998.

[21] P. C. Young, P. McKenna, J. Bruun. Identification of non-linear stochastic systems by state dependent parameter estimation. *International Journal of Control*, vol. 74, no. 18, pp. 1837–1857, 2001.

[22] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[23] G. J. Bierman. Fixed interval smoothing with discrete measurements. *International Journal of Control*, vol. 18, no. 1, pp. 65–75, 1973.

[24] F. Schweppe. Evaluation of likelihood function for Gaussian signals. *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 61–70, 1965.

[25] R. G. D. Steel, J. H. Torrie. *Principles and Procedures of Statistics*, New York, USA: McGraw-Hill, 1960.

[26] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 1, pp. 28–44, 1973.

[27] E. H. Mamdani, S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-machine Studies*, vol. 7, no. 1, pp. 1–13. 1975.

[28] W. Van Leekwijck, E. E. Kerre. Defuzzification: Criteria and classification. *Fuzzy Sets and Systems*, vol. 108, no. 2, pp. 159–178. 1999.

[29] J. F. MacGregor, T. Kourti. Statistical process control of multivariate processes. *Control Engineering Practice*, vol. 3, no. 3, pp. 403–414. 1995.

Philip Cross received the B. Sc. degree in geophysical sciences from Lancaster University, UK in 1991, and the M. Sc. degree in information technology from the University of Teesside, UK in 1992. Following several years in industry, he returned to Lancaster University where he obtained the M. Sc. degree in informatics and the Ph. D. degree in engineering in 2008 and 2012 respectively. He worked as a research associate in the Engineering Department at Lancaster University from 2012 to 2013 and is currently working as a research associate at the University of Leeds, UK.

His research interests include data-based modelling, feedback control, system identification, artificial intelligence, condition monitoring, and renewable energy.

E-mail: p.cross@leeds.ac.uk (Corresponding author)

ORCID iD: 0000-0002-1655-832X

Xiandong Ma is a senior lecturer in the Engineering Department at Lancaster University, UK. He received the B. Eng. degree in electrical engineering in 1986, the M. Sc. degree in power systems and automation in 1989, and a Ph. D. degree in partial discharge based high voltage condition monitoring in 2003. His employment experience includes the Nanjing Automation Research Institute (now the State Grid Electric Power Research Institute) in China for 10 years as a senior research engineer, Lancaster University and then the University of Manchester for 5 years as a postdoctoral research associate, Alstom Power Ltd UK for 18 months as a generator condition monitoring engineer before taking up his lectureship at Lancaster University in December 2008. He has published over 80 refereed journal and international conference papers. He is a chartered engineer and a member of IET.

His research interests include intelligent condition monitoring and fault diagnosis of distributed generation systems with wind turbines with emphasis on analytical and experimental investigation, advanced signal processing, data mining and instrumentation, and electromagnetic non-destructive testing (NDT) and imaging.

E-mail: xiandong.ma@lancaster.ac.uk