Wind Turbine Generator Condition Monitoring Using Temperature Trend Analysis

Peng Guo, David Infield, *Senior Member, IEEE*, and Xiyun Yang

***Abstract*-- Condition Monitoring can greatly reduce the maintenance cost for a wind turbine. In this paper, a new condition monitoring method based on the Nonlinear State Estimate Technique for a wind turbine generator is proposed. The technique is used to construct the normal behavior model of the electrical generator temperature. A new and improved memory matrix construction method is adopted to achieve better coverage of the generator's normal operational space. Generator incipient failure is indicated when the residuals between model estimates and the measured generator temperature become significant. Moving window averaging is used to detect statistically significant changes of the residual mean value and standard deviation in an effective manner; when these parameters exceeds predefined thresholds, an incipient failure is flagged. Examples based on data from the Supervisory Control and Data Acquisition system at a wind farm located at Zhangjiakou in northern China have been used to validate the approach and examine its sensitivity to key factors that influence the performance of the approach. It is demonstrated that the technique can identify dangerous generator over temperature before damage has occurred that results in complete shut down of the turbine.**

***Index Terms*-- condition monitoring, Nonlinear State Estimate Technique, residuals analysis, trend analysis, wind turbine, generator temperature.**

I. INTRODUCTION

CHallenging environmental factors combined with high and turbulent winds make serious demands on wind turbines and result in significant component failure rates, highlighting the importance of maintenance. Appropriate use of condition monitoring can, by detecting faults at an early stage, reduce turbine repair and maintenance costs. In [1], an overall review of Condition Monitoring (CM) methods for different wind turbine components was made. Temperatures are important and easily measured indicators of the health of many wind turbine components and are often recorded automatically by the Supervisory Control and Data Acquisition (SCADA) system. An unexpected increase in component temperature may indicate overload, poor lubrication or possibly ineffective passive or active cooling, as for example if the generator cooling system becomes partially blocked or defective for some other reason [2]. Previous work, [3], used a neural network to construct the normal operating temperature models of gearbox and generator based on SCADA data. When the residual between the model prediction and the measured value becomes very large, a potential fault is identified. Multi-agent methods can be used to combine the CM results of different components together to give a total operating condition of the wind turbine. In [4], the authors proposed a method using a Multiple Layer Perception (MLP) to build a temperature model of the gearbox. When the measured temperature value is outside the confidence range of the value predicted by the model, a fault is registered. Both of the papers just mentioned used an Artificial Neural Network (ANN) to construct the normal operating model based on SCADA data. But the ANN has demerits of requiring a time consuming training process and there can be local minima problems that may limit the improvement of model accuracy, [5]. In this paper, a temperature trend analysis method based on the Nonlinear State Estimate Technique (NSET) is proposed. At the outset, NSET is used to construct the normal operating model for the wind turbine generator temperature and then at each time step the model is used to predict the generator temperature. The time series of residuals between the real measured temperature and the estimate is smoothed using a moving average window in order to reduce the sensitivity of the method to isolated model errors, thereby improving its robustness. When the generator has a potential fault, the time evolution and the distribution of temperature residuals will be different from that for normal operation. Using the moving window approach still allows faults to be detected at an early stage, as will be demonstrated.

The paper is arranged as follows. Section two provides an introduction to turbine and generator including the air-cooled generator arrangement and available SCADA data. Section three explains how the NSET temperature model is constructed and then used to predict the generator temperature. The fourth section focuses on the moving average windowed residuals to detect early-stage faults of the wind turbine generator and examines the impact of data sampling interval. Section five presents two examples demonstrating the capability of the method for early fault detection, and also

This work was supported in part by the Fundamental Research Funds for the Chinese Central University under Grant 09MG18 and the National Natural Science Foundation of China under Grant 61074094.

Peng Guo is with the School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, P. R. China and the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, G1 1XW, UK (e-mail: huadiangp@163.com).

David Infield is with the Institute of Energy and Environment within the Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, G1 1XW, UK (e-mail: david.infield@eee.strath.ac.uk).

Xiyun Yang is with the North China Electric Power University, Beijing, 102206, P.R. China. (e-mail: yangxiyun916@sohu.com)

compares the proposed NSET modeling with an artificial neural network model. A final section provides discussion and conclusions including suggestions for further research.

II. Structure of the Air-Cooled Wind Turbine Generator and the SCADA Parameters

The wind turbines studied in this paper are located at Zhangjiakou in northern China for which SCADA data was available covering the period 04/01/2006 to 24/12/2006. The turbines, manufactured by GE (model 1.5SLE) are variable speed with rated power of 1.5MW. The asynchronous generator with slip rings (DFIG) is forced air-cooled using a closed-loop with air to air heat exchanger to discharge heat to ambient. Two Pt100 thermal resistance probes measure the stator winding temperature (for simplicity, we call it generator temperature) and the cooling air temperature.

The SCADA system at the wind farm records all wind turbine parameters every 10 seconds. Each record includes a time stamp, output power, stator current and voltage, wind speed, environment and nacelle temperature, generator stator winding and cooling air temperature amongst many others; in total 47 parameters are recorded. At the same time, the SCADA system keeps a record of wind turbine operation and fault information, such as start up, shutdown, generator over temperature, pitch system fault, etc. Each fault record includes a time stamp, state number, and fault information. For example, at 9:38, 29/04/2006, the state number was 99 indicating that a generator over temperature alarm occurred. Since 10 second resolution is unusual this data has been averaged up to both 2 and 10 minute periods prior to analysis and model development.

Analysis of a data highlighted the dramatic changes in the local climate through the year at this site reflecting the continental climate in that region. In summer, although the ambient temperature is high, wind speed is relatively low, and the generator load is low and generator failures are seldom. In winter, the wind speed is high, but the ambient temperature is very low, usually around –10 Celsius, and again generator failures are infrequent. Most generator over temperature events and failures occurred in spring and autumn, especially spring. The reason is that during springtime the ambient temperature increases significantly and the wind speed remains high. If the wind turbine shuts down due to generator fault in the windy spring period, significant energy generation will be lost due the time required to change the generator. With appropriate condition monitoring, generator faults can be detected at an early stage, and the maintenance time and cost can be reduced dramatically.

III. NSET Model Construction for Generator Temperature

A. Nonlinear State Estimate Technique (NSET) model construction

NSET is a non-parameter model construction method proposed first by Singer [6]. It is now widely used in the nuclear power plant sensor calibration, electric product lifespan prediction and software aging research [7-9]. The principle of NSET is as follows.

Let there be n variables of interest in a process or device. At time i , a single observation of the variables can be written as an observation vector:

$$\boldsymbol{X}(i)=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}^{\mathrm{T}} \tag{1}$$

Construction of a memory matrix $\boldsymbol{D}$ is the first step of NSET modeling. In a period of normal operation of the process or device, m historical observation vectors are collected covering the range of different operating conditions (such as high or low load, start up, before shut down, etc) to construct the memory matrix $\boldsymbol{D}$, denoted as:

$$\boldsymbol{D}=\begin{bmatrix}\boldsymbol{X}(1) & \boldsymbol{X}(2) & \cdots & \boldsymbol{X}(m)\end{bmatrix} = \begin{bmatrix} x_1(1) & x_1(2) & \cdots & x_1(m) \\ x_2(1) & x_2(2) & \cdots & x_2(m) \\ \vdots & \vdots & & \vdots \\ x_n(1) & x_n(2) & \cdots & x_n(m) \end{bmatrix}_{n\times m} \tag{2}$$

Each observation vector in the memory matrix represents a measured operating state of the process or device. With proper selection of the m historical observation vectors from an extended period of normal (un-faulted) operation of the process or device, the subset space spanned by the memory matrix $\boldsymbol{D}$ can be taken to represent the whole normal working space of the process or device. The construction of memory matrix $\boldsymbol{D}$ is actually the procedure of learning and memorizing the normal behavior of the process or device analogous but different from the training of ANNs.

During subsequent operation, the input to NSET at each time step is a new observation vector $\boldsymbol{X}_{\mathrm{obs}}$ and the output from NSET is a prediction $\boldsymbol{X}_{\mathrm{est}}$ for this input vector for the same moment in time. For each input vector $\boldsymbol{X}_{\mathrm{obs}}$, NSET will produce a m dimensional weight vector $\boldsymbol{W}$:

$$\boldsymbol{W}=\begin{bmatrix}w_1 & w_2 & \cdots & w_m\end{bmatrix}^{\mathrm{T}} \tag{3}$$

with:

$$\boldsymbol{X}_{\mathrm{est}}=\boldsymbol{D}\cdot\boldsymbol{W}=w_1\cdot\boldsymbol{X}(1)+w_2\cdot\boldsymbol{X}(2)+\cdots+w_m\cdot\boldsymbol{X}(m) \tag{4}$$

Equation (4) means that the estimate of NSET is a linear combination of the m historical observation vectors in the memory matrix $\boldsymbol{D}$. The residual (or model error) between the NSET estimate and the input is simply:

$$\boldsymbol{\varepsilon}=\boldsymbol{X}_{\mathrm{obs}}-\boldsymbol{X}_{\mathrm{est}} \tag{5}$$

In order to determine the weight vector $\boldsymbol{W}$ that minimizes the residuals overall the following relation is used:

$$\boldsymbol{W}=\left(\boldsymbol{D}^{\mathrm{T}}\otimes\boldsymbol{D}\right)^{-1}\cdot\left(\boldsymbol{D}^{\mathrm{T}}\otimes\boldsymbol{X}_{\mathrm{obs}}\right) \tag{6}$$

The derivation of equation (6) can be found in [7] and [10], $\otimes$ is a nonlinear operator used to replace the regular multiplying operator in matrix multiplication. It is important to note that the weight vector is recalculated at each time step, in contrast to standard multivariate regression where the weights are calculated only once, [11]. There are many optional nonlinear operators to choose from [12], and in this paper, the nonlinear operator is chosen as the Euclidean distance between the two vectors:

$$\otimes(\boldsymbol{X},\boldsymbol{Y})=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{7}$$

This norm is chosen since it is simple to evaluate but clearly reflects the difference/distance between two vectors in a manner that is easy to understand; it is the distance measure used in other application areas, see [13] and [14]. It is of course possible that alternative norms might perform better in the context of NSET, but this is a matter for future research.

The nonlinear operator has a direct physical interpretation. When two observation vectors are the same or similar, the distance between the vectors will be zero or near zero. When one vector is very different from the other, the distance between them will be great and the result of the nonlinear operator will be large. The weight vector in (6) reflects the similarities between the NSET input vector $\boldsymbol{X}_{\text{obs}}$ and the m historical observation vectors in memory matrix $\boldsymbol{D}$. If $\boldsymbol{X}_{\text{obs}}$ is similar to only one particular $\boldsymbol{X}(i)$ in the memory matrix $\boldsymbol{D}$, the corresponding weight w_i in $\boldsymbol{W}$ will be near unity and otherwise near zero.

With (4) and (6), the final estimate of the NSET model for the process or device is:

$$\boldsymbol{X}_{\text{est}}=\boldsymbol{D}\cdot\left(\boldsymbol{D}^{\text{T}}\otimes\boldsymbol{D}\right)^{-1}\cdot\left(\boldsymbol{D}^{\text{T}}\otimes\boldsymbol{X}_{\text{obs}}\right) \tag{8}$$

In the construction of memory matrix $\boldsymbol{D}$, the Euclidean distance between every two observation vectors of the m vectors should be big enough to make sure that the condition number of $\boldsymbol{D}^{\text{T}}\otimes\boldsymbol{D}$ is not too large. Otherwise, it will be very difficult to calculate the inverse matrix of $\boldsymbol{D}^{\text{T}}\otimes\boldsymbol{D}$ and the NSET model becomes an ill-conditioned problem [15].

When the process or device works normally, the input observation vector of NSET is most likely to be located in the normal working space that is represented by the memory matrix $\boldsymbol{D}$, in that it is similar to some historically measured vectors in the memory matrix. As a result, the estimate of NSET will have a very high accuracy. When problems arise with the process or device, its dynamic characteristics will change, and the new observation vector will deviate from the normal working space. In this case the linear combination of the historical vectors in the memory matrix will not provide an accurate estimate of the input and the residual will increase in magnitude.

B. The Variable Selection for the Generator Temperature NSET Model

In order to construct the NSET model of generator temperature, the variables included in the observation vector should be carefully chosen. Because we are concerned with the generator temperature, variables that have close relationship with generator temperature should be taken into account. Following a review of the 47 variables recorded by the SCADA system, the following five variables were selected to construct the observation vector.

(1) Power (P). Power has great influence on the generator temperature. When the output power is high, the stator current will be very large which will lead to the generator temperature increasing.

(2) Ambient Temperature (T). Because the local temperature that the wind turbine experiences changes greatly in the short term (from day to night for example) and in the longer term (weeks to months) due to passing weather systems and seasons it must be taken explicitly into account. At Zhangjiakou during March and April, ambient temperature changes can be as large as 30 degrees Celsius because of fast moving weather fronts.

(3) Nacelle Temperature (NT). The nacelle temperature has close relationship to the generator temperature.

(4) The temperature of the generator cooling air (GAT) directly impacts on the stator cooling.

(5) The generator stator winding temperature itself (GT).

The normal records of SCADA are divided into two groups. The first group is an extensive set (at least several thousand data points) that includes the m observation vectors selected to construct the memory matrix $\boldsymbol{D}$. The second set comprises normal records not used for selection of the memory matrix. A part of the second group, set $\boldsymbol{V}$, will be used for validation of the NSET model. After the construction of memory matrix $\boldsymbol{D}$, model estimates and residual analysis can be carried out as shown in Fig.1 to monitor the generator operating condition.

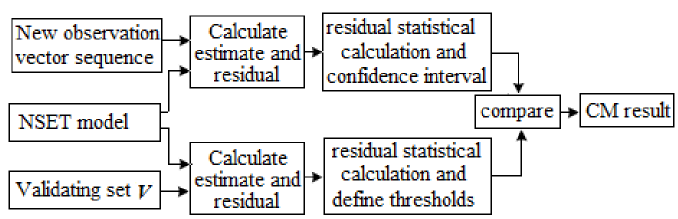

Fig.1 NEST generator temperature trend condition monitoring

C. Construction of memory matrix $\boldsymbol{D}$

Because the five variables in the observation vector have different units and their absolute values are very different, the initial values of the five variables are rescaled to the range $[0 \quad 1]$ according to their maximum and minimum values. After rescaling, each variable has same weight in the calculation of Euclidean distance, the chosen nonlinear operator.

In previous work, [7], two algorithms are used to extract observation vectors from the normal working period to construct the memory matrix. In the first algorithm, vectors are selected that correspond to the extreme normal working states. For each variable in the observation vector, the algorithm finds the minimum and maximum measurements from the normal working period. The observation vectors containing these measurements are added to the memory matrix. In the second algorithm, the remaining observation vectors from the normal working period are ordered by their Euclidean norms. For n dimensional vector $\boldsymbol{X}$, the Euclidean norm is:

$$\|\boldsymbol{X}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{9}$$

The algorithm then selects evenly spaced elements from the ordered set and adds their corresponding observation vectors to the memory matrix. This construction method is simple, but has some problems. Observation vectors can exist with Euclidean norm values that are similar or even exactly the same, but the vectors may be quite completely different, such as for the following two vectors:

$$\boldsymbol{X}_1 = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}^{\mathrm{T}} \text{ and } \boldsymbol{X}_2 = \begin{bmatrix}0 & 0 & 1 & 0 & 0\end{bmatrix}^{\mathrm{T}}$$

Selecting one of such similar Euclidean norm equivalent vectors will result in the discarding other equivalent norm vectors. Then only the working space near the selected observation vector is covered by the memory matrix, while parts of the parameter space near the discarded ones will be neglected. In order to minimize this effect a new memory matrix construction method is proposed that builds on the method used in [7] to give much improved NSET model accuracy. The new method is as follows.

Assume the historical observation vectors during the normal working period in April are $\boldsymbol{X}^N(1), \boldsymbol{X}^N(2), \cdots, \boldsymbol{X}^N(M)$ and make the matrix:

$$\boldsymbol{K} = \begin{bmatrix}\boldsymbol{X}^N(1) & \boldsymbol{X}^N(2) & \cdots & \boldsymbol{X}^N(M)\end{bmatrix}$$
$$= \begin{bmatrix} x_1^N(1) & x_1^N(2) & \cdots & x_1^N(M) \\ x_2^N(1) & x_2^N(2) & \cdots & x_2^N(M) \\ \vdots & \vdots & & \vdots \\ x_5^N(1) & x_5^N(2) & \cdots & x_5^N(M) \end{bmatrix}_{5\times M} \tag{10}$$

The number of vectors in matrix $\boldsymbol{K}$ is M. Each observation vector includes power, ambient temperature, nacelle temperature, generator cooling air temperature, generator temperature, denoted $x_1, x_2, \cdots, x_5$ respectively, and n in (1) takes the value $n = 5$.

Every observation vector in the normal working space covers the five variables and is normalized as described above. In order to ensure that memory matrix $\boldsymbol{D}$ covers the vectors with different variable values in the normal working space, for each of the five variables, the range $\begin{bmatrix}0 & 1\end{bmatrix}$ is equally divided into 100 sections and the observation vectors from matrix $\boldsymbol{K}$ selected at steps of 0.01. Observation vectors at each step of 0.01 for each variable in turn are added into memory matrix $\boldsymbol{D}$ so long as variable in question is sufficiently close to the value in the observation vector. For example, for variable x_1 (turbine power), the method of adding observation vectors to $\boldsymbol{D}$ is shown in Fig.2.

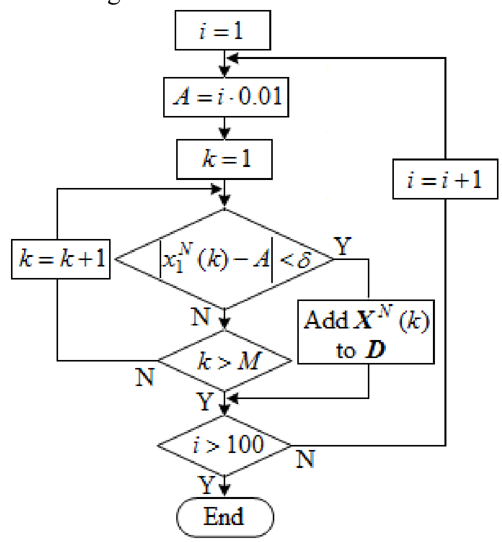

Fig.2 Construction of memory matrix $\boldsymbol{D}$ for variable x_1

In Fig.2, δ is a small positive number, taken to be 0.001 in this work. With this method, observation vectors with different variable values can be included in the memory matrix $\boldsymbol{D}$.

Memory matrix construction thus has two steps. The first step is using the method in [7] to select observation vectors from the normal working period. And the second step is using the new method outlined in this paper to add further observation vectors to $\boldsymbol{D}$. Before adding a new vector to the memory matrix, the Euclidean distances between the new potential vector and the vectors already in the memory matrix is checked. If the distance is too small, that means the memory matrix already has a vector that is quite similar to the one being considered and consequently it should not be added into the memory matrix. This check will limit the size of the memory matrix and ensure that it is well-conditioned. The validation of the NSET model will demonstrate significant improvement of model accuracy with the combined memory matrix construction method outlined above.

When the memory matrix it complete, the NSET model can be used to estimate the generator temperature for new input vectors.

IV. Generator Temperature NSET Model Residual Statistical Analysis

A. Moving window calculation of residual mean value and standard deviation statistics

When the wind turbine generator suffers from some abnormality, the new observation vector will deviate from the normal working space and the distribution and time development of the estimate residual will change significantly from the normal condition. The mean value and standard deviation will reflect a change in the distribution of the residuals. In order to detect the changes in the variables in a timely manner, a moving average calculation is used. At a certain instant, the residual sequence of generator temperature from the NSET model is:

$$\varepsilon_{\mathrm{GT}} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_N \quad \cdots] \tag{12}$$

A time window with width N is adopted to calculate the moving average or mean value and standard deviation for the N successive residuals in the window:

$$\bar{X}_\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i \ , \quad S_\varepsilon = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\varepsilon_i - \bar{X}_\varepsilon\right)^2} \tag{13}$$

The moving window is shown in Fig.3.

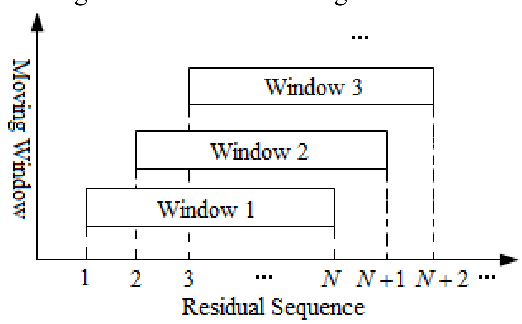

Fig.3 Residual moving average window used in residual analysis

B. Residual statistical distribution when generator has an abnormality

When the generator works normally, the NSET model provides very accurate estimates of generator temperature. The residual sequence has a mean value near zero and the standard deviation is small. When a problem occurs with the generator, the new observation vector may deviate from the normal condition and the generator temperature residual distribution will thus also change. Abnormal generator operation can be identified as follows:

(1) The residual mean value remains near zero but the standard deviation increases dramatically. In this condition, the distribution of the residuals becomes wider.

(2) The residual mean value deviates from zero with an obvious magnitude and the standard deviation remains small. In this condition, the residual systematically departs from zero.

(3) A combination of the above two situations.

In order to detect early faults of the generator, failure thresholds are needed for both the residual mean value and its standard deviation. Assume that the thresholds for them are set respectively as E_{Y} and S_{Y} where these are determined according to operator experience or determined through model validation as presented below.

The residual sequence has been obtained using observation vectors from the validation data set as input to the already identified NSET model with subsequent application of the moving average window to the time sequence of residuals. By trail and error the optimal size of the moving window is determined. The maximum absolute mean value and standard deviation allowed are, as defined above, respectively E_{V} and S_{V} but note that these will be dependent on averaging window size. Then the thresholds for generator failure detection are as follows:

$$E_{\mathrm{Y}} = \pm k_1 \cdot E_{\mathrm{V}} \ , \ S_{\mathrm{Y}} = k_2 \cdot S_{\mathrm{V}} \tag{14}$$

where k_1 and k_2 are positive coefficients and can be chosen based on operator experience. Relatively larger k_1 and k_2 will increase the robustness of CM method and reduce the false alarms. In this paper, $k_1 = 3$ and $k_2 = 2$.

As the non-parameter model is not a perfect estimator there will be some uncertainty associated with each estimate. This uncertainty should ideally be quantified when a technique for condition monitoring is presented [16]. In most applications of NSET the residual sequence has an approximately normal distribution but with mean value and variance to be determined. When calculating the mean value and standard deviation within the moving window, a confidence interval with confidence level $1-\alpha$ should be set. Assuming a normal distribution with mean value and variance unknown, the confidence intervals for the mean value and standard deviation are respectively:

$$\left(\bar{X}_\varepsilon - \frac{S_\varepsilon}{\sqrt{N}}t_{\alpha/2}(N-1), \bar{X}_\varepsilon + \frac{S_\varepsilon}{\sqrt{N}}t_{\alpha/2}(N-1)\right) \tag{15}$$

$$\left(\frac{\sqrt{N-1}\cdot S_\varepsilon}{\sqrt{\chi^2{}_{\alpha/2}(N-1)}}, \frac{\sqrt{N-1}\cdot S_\varepsilon}{\sqrt{\chi^2{}_{1-\alpha/2}(N-1)}}\right) \tag{16}$$

where, N , $\bar{X}_\varepsilon$ and S_ε are respectively the moving window width, mean value and standard deviation. $t_{\alpha/2}$ and $\chi^2{}_{\alpha/2}$ are respectively the $\alpha/2$ quantile of the t distribution and χ^2 distribution. When the confidence interval of the mean value or standard deviation exceeds the set thresholds, a failure alarm should be triggered. In our example of NSET, the operators would be notified to stop the wind turbine, if not already stopped by the control system, and check the generator condition so that serious damage can be avoided.

V. Case Studies

A. Case 1:

1) Generator Fault Description

Actual failure of major components like a generator is relatively rare and for this reason much research in this area has to make do for now with analysis of isolated incidents. This work is no exception. SCADA data of a particular turbine at the Zhangjiakou wind farm exhibited a generator failure and automatic shut down on the last day of April 2006. In the days before this event there were frequent generator over temperature alarms.

As already mentioned, the SCADA system at the Zhangjiakou wind farm is unusual in that it collects 10 second data. To reduce the size of the data sets and because 10 second resolution is not required for thermal modeling due to thermal inertia, the data was averaged up to give 2 minute data sets. These data have been used to develop an NSET model but because 10 minute SCADA data is more common, the data has also been averaged up to give more standard looking data sets and these have also been used to develop and assess the NSET approach. In fact rather little difference in model performance is found. For April, we have over twenty thousand two minute average records and over four thousand 10 minute records. According to the manufacturers fault handbook an over temperature alarm will occur and the wind turbine shuts down when the generator temperature reaches 150 Celsius over a continuous period of 60 seconds. When the temperature falls to below 135 Celsius the wind turbine will restart. According to the fault records of the wind turbine, during April 29 and 30, the over temperature alarm occurred over five times. At 9:50 on 30/04/2006, it shuts down for good due to failure of generator. The trends of the five selected variables over the five days from 10.00 am on April 25 to 10.00 am on April 30 are shown in Fig.4. The both two and ten minute averaged data sets from during this wind turbine's normal working period earlier in April are used to construct the NSET generator temperature models as described in Section Ⅱ above.

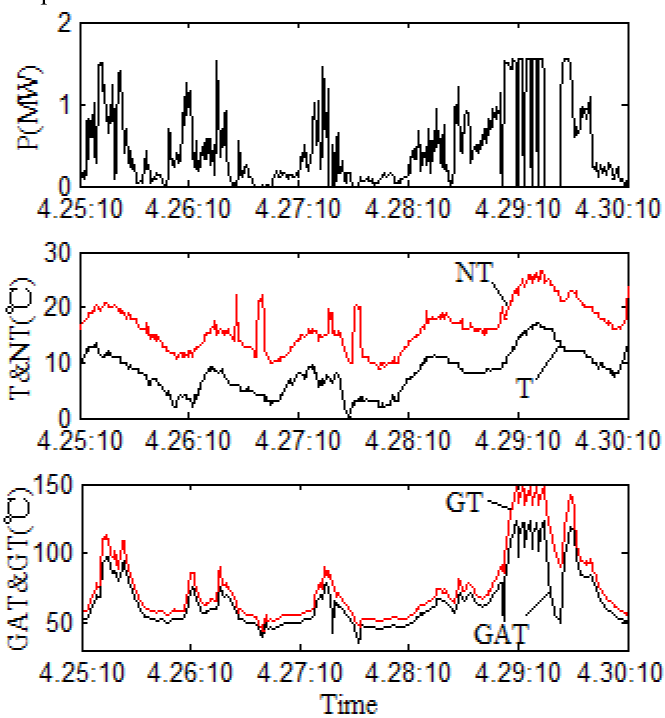

Fig.4 Trends over 5 days of the selected variables

2) The Validation of Generator Temperature NSET Model

600 successive 2 minute vectors taken from the validating data set $\boldsymbol{V}$ are used as input to the NSET model to validate its effectiveness. For simplicity, only the generator temperature element, x_5, of the vector and its residuals are studied here as this is the parameter of interest. As with data used in the NSET modeling, the generator temperature and its estimate are all normalized values. In the normal working space for April, the maximum generator stator temperature is 150 Celsius, and the minimum is 38.2 Celsius. The residual of generator temperature is:

$$\varepsilon_{GT} = x_5 - \hat{x}_5 \tag{11}$$

where, x_5 is the generator temperature within the input vector of NSET model and $\hat{x}_5$ is the corresponding model estimated temperature from the NSET output vector. Firstly, we validate the NSET model constructed using the method in [7]. The vectors with minimum and maximum Euclidean norm during normal working period are respectively:

$$\boldsymbol{X}_{\text{nmin}} = [0.049 \quad 0.133 \quad 0.080 \quad 0.028 \quad 0.048]$$

and

$$\boldsymbol{X}_{\text{nmax}} = [0.974 \quad 0.667 \quad 0.684 \quad 0.974 \quad 0.998]$$

Their Euclidean norms are respectively 0.173 and 1.951. With a Euclidean norm step set to 0.004, we select 440 vectors for memory matrix $\boldsymbol{D}$. Taking the vectors containing minimum and maximum measurements of each variable into account, the total vector number in memory matrix constructed using the method of [7] is 460, and the validation result is shown in Fig.5.

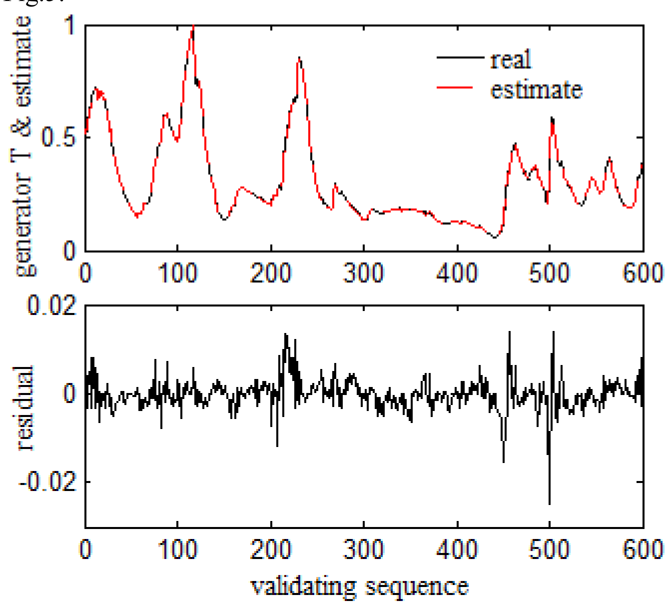

Fig.5 Validation results for the NSET model with method in [7] with 2 minute data

Next we assess the NSET model with the extended memory matrix constructed as described in Section Ⅱ. A further 401 observation vectors are added to the 460 vectors using the new algorithm. Fig.6 shows the results for the same time period and it is immediately clear that the residuals have been significantly reduced. The root Mean Square (RMS) residual is a useful but simple measure of model accuracy. The RMS for the generator temperature residual of the NSET model constructed according to [7] is 0.0034, and that for the NSET model with the extended memory matrix construction method is 0.002. It is clear that the new approach to memory matrix construction has greatly improved the model accuracy.

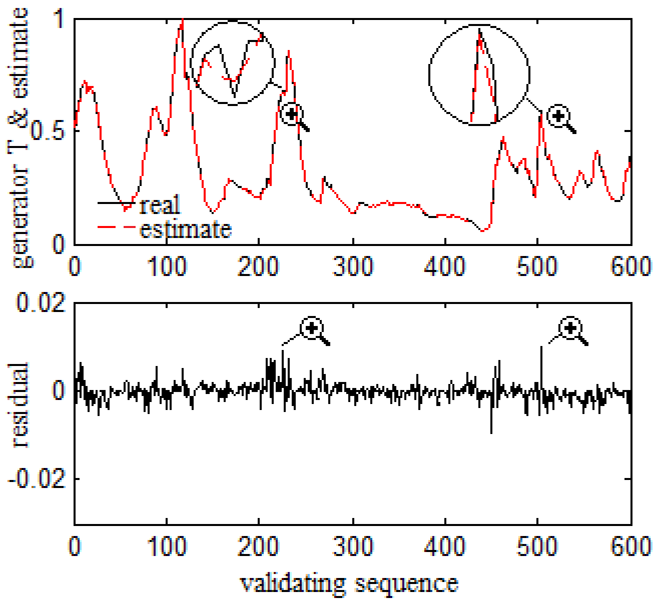

Fig.6 Validation results for the NSET model with combination method

In Fig.6, most residuals are below 0.5%. It should be noticed that in Fig.6 there are several isolated places where the residual is significantly larger, such as at points 226 and 504. This can be explained in terms of model coverage. For some certain very small working spaces, such as the areas near point 226 and 504, the NSET model's covering ability may be not as good as for other areas, and the residuals will be relatively large. And this behavior should not automatically be regarded as a generator failure indicator. The moving averaging should reduce significantly the impact of these isolated larger residuals on the proposed method of fault detection.

3) Compared With Artificial Neural Network

As mentioned in Section I, an artificial neural network (ANN) is an effective non-parametric modeling method. In order to compare with the NSET method, an ANN is developed and then used to model the normal behavior of this same wind turbine.

The ANN used here is a conventional three layer network with one input layer, one hidden layer and one output layer since there is considerable experience with such models, [17] and [18]. For equivalence to the NSET model, the network has four inputs (power, ambient temperature, nacelle temperature and generator cooling air temperature) and the network output is generator temperature. The hidden layer has twenty neurons which is typical of models of this complexity, [17].

In order to be able to benchmark against the NSET, the training data set for ANN are the historical observation vectors in the memory matrix of NSET and the validation data set for ANN is also same successive 600 vectors as used above for NSET assessment. As is common, the back propagation algorithm is used to adjust the weights and so train the network. The training process is then iterated until the total error for all training patterns satisfies a preset criteria or a set number of training cycles is completed. In order to alleviate the local minima problem and accelerate the training process, momentum term is added to the weights adjustment, [19]. For the ANN used here, the learning rate and the momentum factor are respectively 0.05 and 0.8. The training process is ended when the learning cycle reaches 10000 if not before. The network training process took about 14 minutes on a 2GHz PC. Fig.7 shows the trend of the total error for the training dataset as a function of iteration number. It can be seen that the total error reduces quickly in the early stage of the training but reaches a steady after approximately 4000 iterations.

Subsequent to training, the same successive 600 two-minute observation vector points are used to validate the ANN model. Fig.8 shows the output values predicted by the ANN and the calculated model residuals.

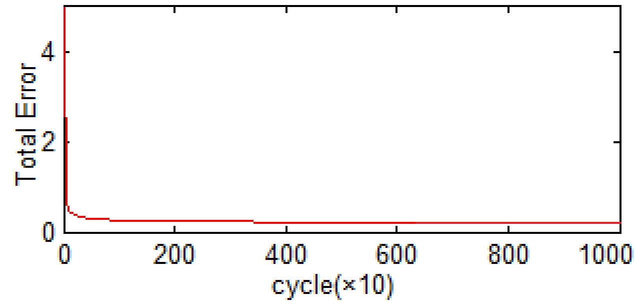

Fig.7 Trend of the total error during ANN training

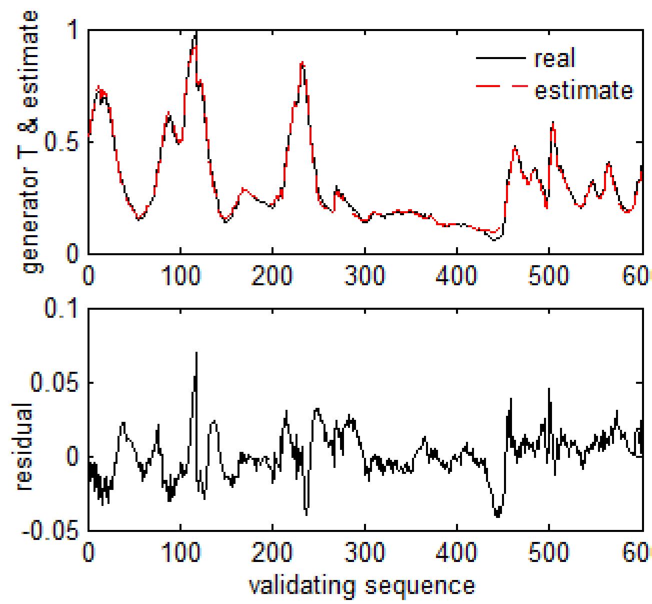

Fig.8 Validation result for artificial neural network

Comparing Fig.8 with the NSET validation result(Fig.6) it can be seen that NSET achieves considerably higher accuracy in modeling the normal behavior of the wind turbine generator temperature. And the RMS for the ANN at 0.0151 is much larger than that of the NSET.

Besides higher modeling accuracy, NSET has another benefit compared with neural network in that it can more easily adapt to new normal working condition. When the wind turbine generator's normal behavior changes due to factors such as a change in the meteorological conditions, the new normal behavior can be modeled by the selection of new (or additional) observation vectors for the memory matrix whereas the ANN will require another time consuming training process which may be inappropriate in on-line condition monitoring. Adaptive development of the NSET method is seen as a promising area of further research.

4) Moving Average Filtering and Improved Residual Analysis

First, the moving window calculation outlined in Section IV above is applied to the residual sequence of the validation set in Fig.6. The window width N should be properly selected so that the influence of the occasional isolated large residual caused by the imperfect coverage ability of memory matrix $\boldsymbol{D}$ can be minimized. At the same time, a moving window with a properly selected N must be able to detect the changes of mean value and standard deviation in a quick and effective manner. For the 2 minute data $N=150$ reflects a useful balance between these two conflicting requirements. Fig.9 shows the trends of the mean value and standard deviation for the validating set after moving average filtering.

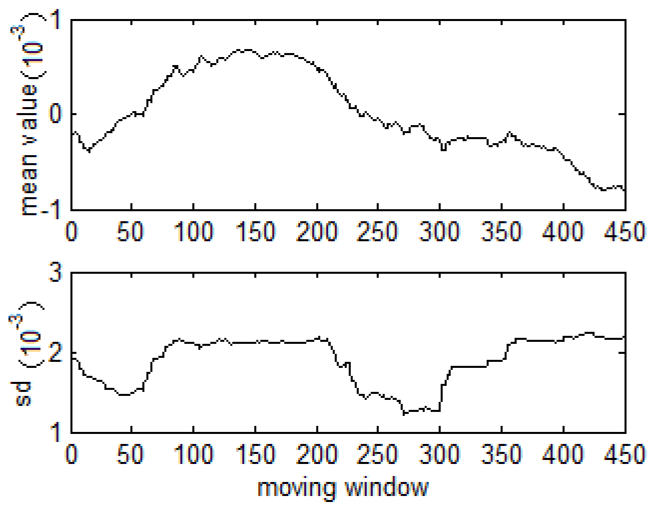

Fig.9 Statistical characteristics of validating sequence

From the trends, we estimate

$$E_{\mathrm{V}} = 0.8\times10^{-3},\ S_{\mathrm{V}} = 2.2\times10^{-3} \tag{17}$$

and the thresholds for the mean value and standard deviation are chosen respectively as:

$$E_{\mathrm{Y}} = \pm2.4\times10^{-3},\ S_{\mathrm{Y}} = 4.4\times10^{-3} \tag{18}$$

SCADA records show that the first generator over temperature alarm occurred at 9:38 29/04/2006. The 550 two-minute observation vectors prior to this alarm are used as the input to NSET model of the generator temperature. The NSET model estimate and residuals time sequences are shown in Fig.10.

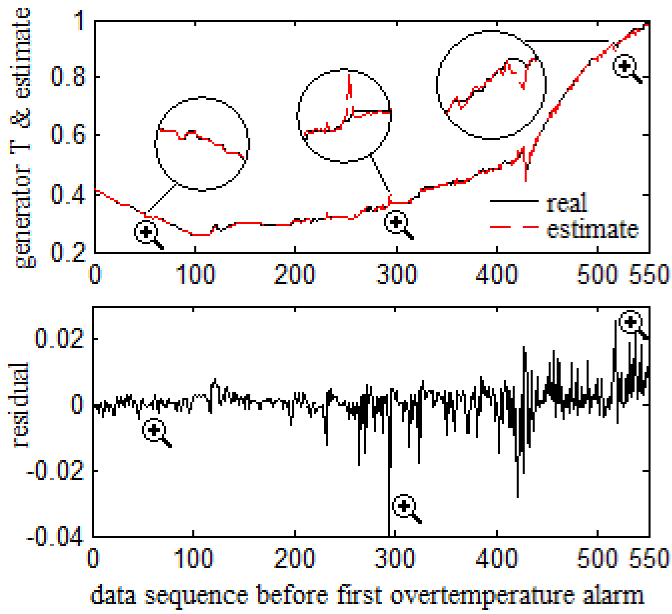

Fig.10 Estimate and residual before the first over temperature alarm

When the confidence level for the mean value and standard deviation is chosen as 95%, the trends and confidence intervals for the mean value and standard deviation of the filtered residual sequence are shown in Fig.11.

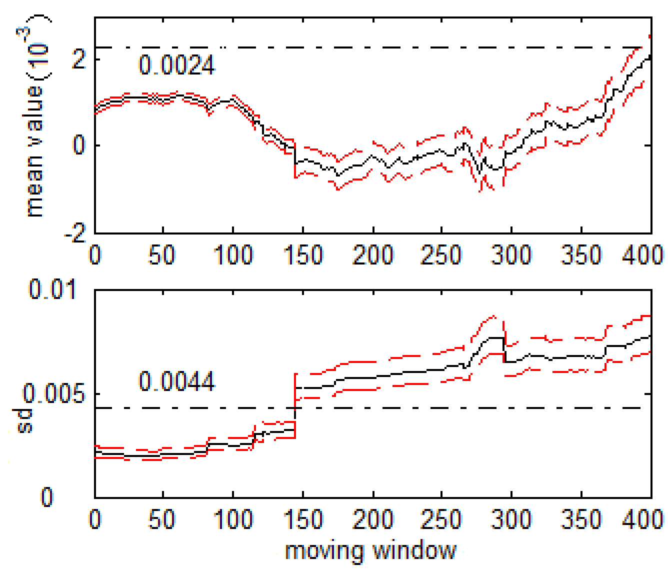

Fig.11 Statistical characteristics before the first over temperature alarm

The mean value shows no great changes and is near zero, but the standard deviation changes significantly. At time point 145, the upper confidence interval of the standard deviation exceeds the failure threshold, and an incipient generator failure has been detected. The time of incipient fault detection based on the threshold crossing of the standard deviation is eight and a half hours before the first over temperature alarm (at 9:38 29/04/2006) and over 32 hours before final shutdown and generator failure The time sequence of events is summarized in Fig.12 below.

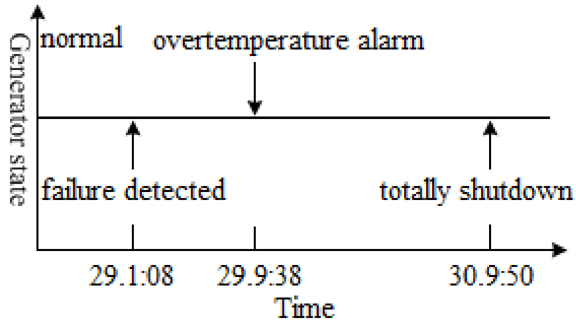

Fig.12 Generator state changes

Because 10 minute SCADA data is most common, the above analysis has been repeated with the 10 minute averaged data, both for the model identification, and the validation. It should be noted that the same time window has been used for the moving average filter so that N is reduced from 150 to 30. The final results are shown on Fig.13 and Fig.14 below and show very little difference in fault detection. :

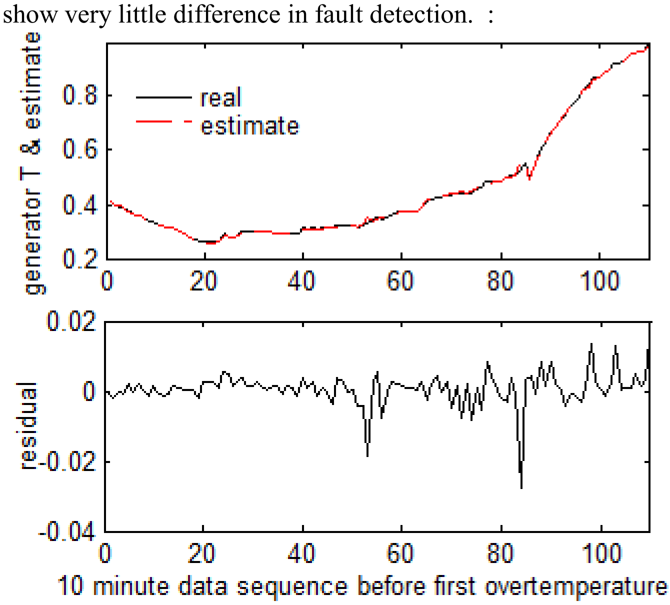

Fig.13 10 minute NSET estimate and residual

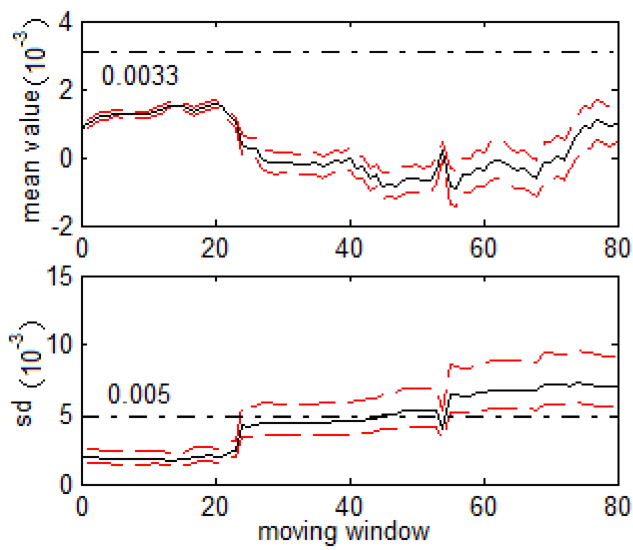

Fig.14 10 minute statistics prior to the first over temperature alarm

In Fig.14, it can be seen that, the upper limit of the standard deviation exceeds the thresholds at point 24 and the incipient generator failure was detected about nine hour before the first over temperature alarm. The 10 minute average data shows nearly the same results as for the 2 minute averaged data.

B. Case 2

In order to further validate the effectiveness of the NSET method for wind turbine generator condition monitoring, SCADA data from a second wind turbine from the same wind farm is used to further investigate the NSET method. The raw SCADA data for this wind turbine during May 2006 also have a ten second resolution, and as before have been averaged to give 2 minute data values.

For this wind turbine, 600 successive 2 minute observation vectors from 0:00 18/05/2006 are shown in Fig.15.

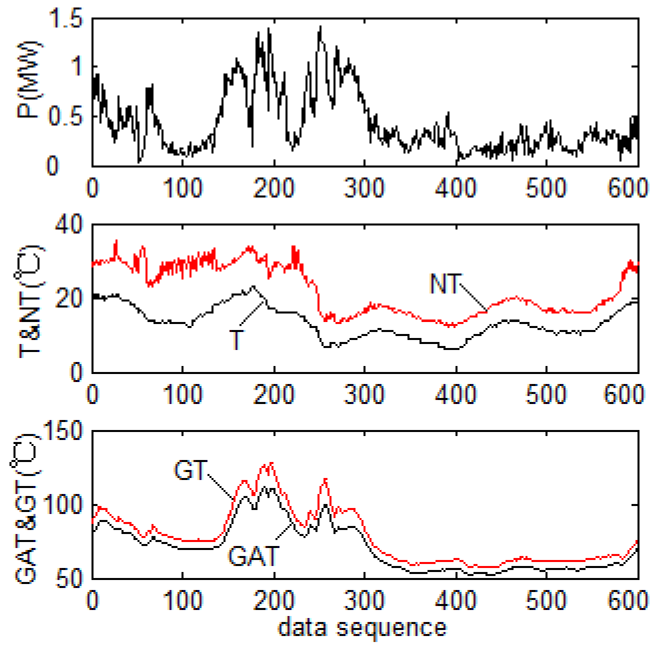

Fig.15 Trends of selected variables from 0:00 18/05/2006

Using the same method as case 1 the memory matrix was constructed for this wind turbine based on May data prior to the 18th day of the month. The total number of observation vectors in the final memory matrix is 781. Subsequently the 600 successive observation vectors shown in Fig.15 were scaled and used as validation data for the NSET model. The results are shown in Fig.16 and as before show excellent model accuracy, better indeed than in case 1.

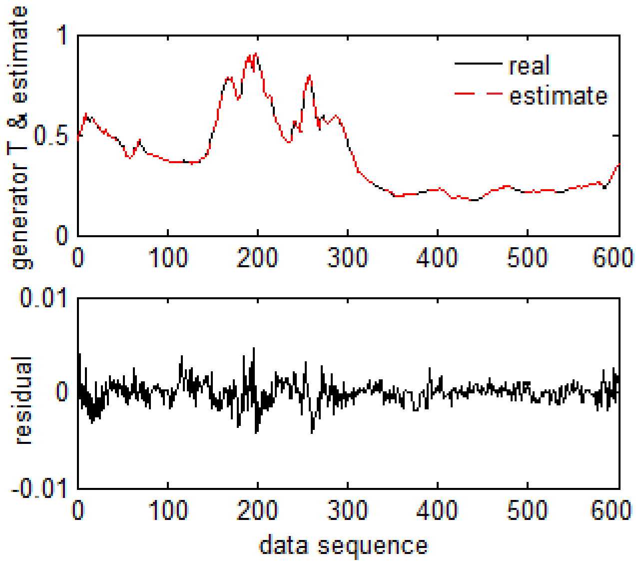

Fig.16 Validation result for wind turbine 2

Regrettably, turbine 2 showed no inclination to fail and as a result a temperature rise was manually added to the above validation sequence to test the NSET method's abnormality detection capability. A fixed temperature rise of 0.001 per time step (corresponding to about 3.3 degrees C/hour) was added from data point 451. The result is shown in Fig.17.

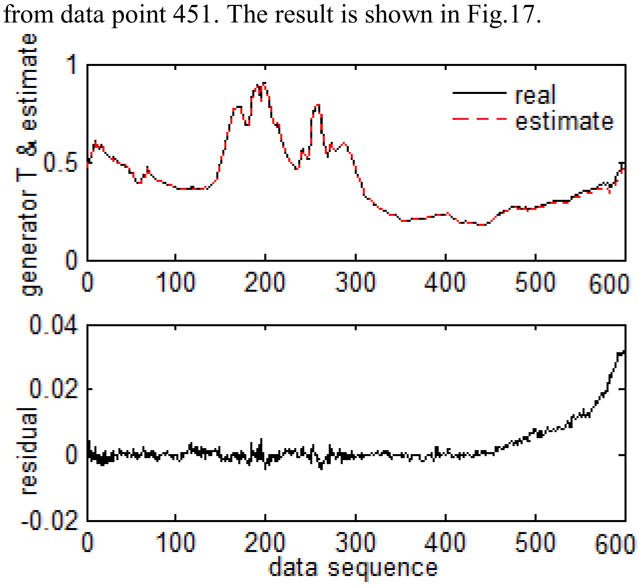

Fig.17 NSET estimate and residual after manual temperature rise added

Application of the moving average window with the width $N = 150$ as before to the validation sequence gives the maximum absolute mean value and standard deviation respectively as:

$$E_{\mathrm{V}} = 3.6 \times 10^{-4},\ S_{\mathrm{V}} = 1.7 \times 10^{-3}.$$

And the thresholds of mean value and standard deviation for case 2 are thus:

$$E_{\mathrm{Y}} = \pm 10.8 \times 10^{-4},\ S_{\mathrm{Y}} = 3.4 \times 10^{-3}.$$

Application of the moving average filter and thresholds to the data sequence with artificial trend added as in Fig.17, results in the plots and threshold crossing points as shown in Fig.18.

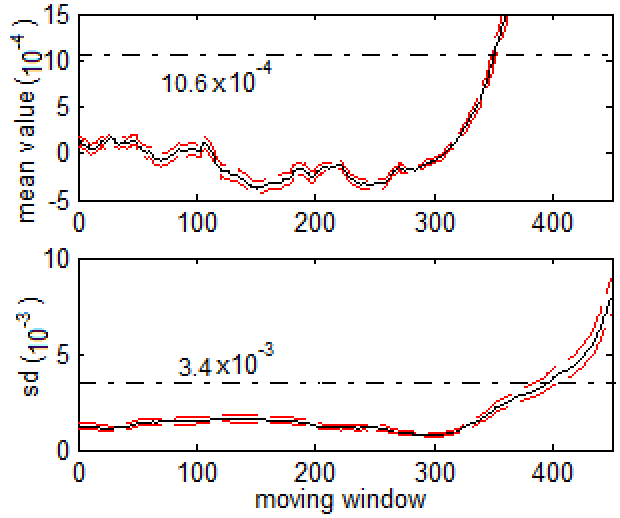

Fig.18 Statistical characteristics after temperature rise added

From Fig.18 it can be seen that the mean value first exceeds the threshold at point 349 and the artificial temperature drift has been detected. Since the manual drift was added at point 451 and the window width is 150, the drift is identified only 98 minutes after its initiation.

VI. Discussion and Conclusions

This paper uses a nonlinear state estimation technique (NSET) to construct a generator temperature model. Detail of how to construct memory matrix is provided. A new approach that improves the coverage of the normal operational space of the wind turbine by extending the memory matrix of the NSET model is presented. Compared with a neural network method, the NSET model has the advantage of not requiring a time-consuming training procedure and has been demonstrated for the case of the generator temperature to provide higher modeling accuracy. The method has the advantage of being remarkably simple computationally and conceptually (much more so than ANNs) and should thus have immediate appeal for wind industry practitioners. When the generator experiences a fault, new observation vectors will generally deviate from the normal working space and the NSET estimate of the residual distribution and its time development will change. A moving average window filter has been adopted to reveal the trends in mean value and standard deviation of the residual sequence; this is also very easy to implement computationally. Generator condition can be determined from crossing of predetermined thresholds. With the effective selection of the memory matrix from data from normal operation, the presented method can identify incipient wind turbine generator failure usefully ahead of the ultimate failure. Although the method has only been assessed for a particular turbine at a specific site the model is of a sufficiently general nature to believe that, with suitable selection of the parameters available to the user (window size, vector parameter selection, thresholds etc), the method should prove useful for wind turbine condition monitoring more widely. Future research to apply the approach to other important components of wind turbine such as the gearbox and bearings is planned.

VII. Acknowledgment

The authors gratefully acknowledge for the help of Professor Xu Daping and Lv Yuegang.

References

[1] Z.Hameed, Y.S.Hong, Y.M.Cho, S.H.Ahn, and C.K. Song, "Condition monitoring and fault detection of wind turbines and related algorithms: A review", *Renewable and Sustainable Energy Reviews*, Vol. 13, no. 1, pp. 1-39, Jan. 2009.

[2] P. Tavner, Li Ran, Jim Penman, and Howard Sedding, *Condition Monitoring of Rotating Electrical Machines*.

[3] A. Zaher, S.D.J. McArther, D.G.Infield, and Y. Patel, "Online wind turbine fault detection through automated scada data analysis", *Wind Energy*, vol. 12, no. 6, pp. 574-593, Sep. 2009.

[4] Mari Cruz Garcia, Miguel A. Sanz-Bobi, and Javier del Pico, "SIMAP: intelligent system for predictive maintenance: application to health condition monitoring of a wind turbine gearbox", *Computers in Industry*, vol. 57, no. 6, pp.552-568, Aug. 2006.

[5] M.Gori, A.Tesi, "On the problem of local minima in back-propagation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol.12, no.1, pp.76-86, Jan.1992.

[6] K.C. Gross, R.M. Singer, S.W. Wegerich, and J.P. Herzog, "Application of a model-based fault detection system to nuclear plant signals", in *proc. 9th International Conference on Intelligent Systems Application to Power System*, Seoul, Korea, Jul. 1997.

[7] F.K. Bockhorst, K.C. Gross, J.P. Herzog, and S.W. Wegerich, "MSET modeling of crystal river-3 venturi flow meters," in *Proc. International Conference on Nuclear Engineering*, San Diego, CA, 1998.

[8] S. F. Cheng and M. G. Pecht, "Multivariate state estimation technique for remaining useful life prediction of electronic products," in *Proc.AAAI Fall Symp. Artif. Intell. Prognostics*, Arlington, VA, Nov. 2007, pp. 26–32.

[9] K.J. Cassidy, K.C. Gross and A. Malekpour, "Advanced pattern recognition for detection of complex software aging phenomena in online transaction processing servers," in *Proc. Dependable Systems and Networks*, Washington, D.C., USA, Jun, 2002.

[10] R.M. Singer, K.C.Gross, J.P. Herzog, R.W. King, and S. Wegerich. "Model-based Nuclear Power Plant Monitoring and Fault Detection: Theoretical Foundation," in *proc. 9th International Conference on Intelligent Systems Application to Power System*, Seoul, Korea, Jul. 1997.

[11] Sheng Zhou, Xie Shiqian, and Pan Chengyi, *Probability Theory and Statistics*. Beijing: High Education Press, 2010.

[12] Christopher L. Black, Robert E. Uhrig, and J. Wesley Hines, "System modeling and instrument calibration verification with a nonlinear state estimate technique," in *Proc. Maintenance and Reliable Conference*, Knoxville, TN, May. 1998.

[13] Liwei Wang, Yan Zhang, and Jufu Feng, "On the Euclidean distance of images," *Trans. On pattern analysis and machine intelligence*, vol.27, no.8, pp1334-1339, 2005.

[14] Shang Zhen-hong, Fu Xiao-dong, and Yuan Mei-yu, "Line detection based on Euclidean distance," Journal of computer applications, vol. 28, no.1, pp 177-180, 2008.

[15] A.K. Cline, C.B. Moler, G.W. Stewart, and J.H. Wilkinson, "An estimate for the condition number of a matrix," *SIAM Journal on Numerical Analysis*, vol.16, no.2, pp. 368-375, Apr. 1979.

[16] J. Wesley Hines, Dustin Garvey, "Process and equipment monitoring methodologies applied to sensor calibration monitoring," *Quality and Reliability Engineering International*, vol.23, no.1, pp.123-135, Feb, 2007.

[17] Bimal K. Bose, "Neural network applications in power electronics and motor drives," Trans. On industrial electronics, vol.54, no.1, pp 14-33, 2007.

[18] Chih-fong Tsai, Jhen-wei Wu, "Using neural network ensembles for bankruptcy prediction and credit scoring," *Expert system with application*, vol.34, no.4, pp 2639-2649, 2008

[19] Nii O. Attoh-Okine, "Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance," *Advances in engineering software*, vol.30, no.4, pp 291-302, 1999,

Biographies

Peng Guo was born in Heibei, China in June 1975. He received the Ph. D. degree from the North China Electric Power University, Beijing, China, in 2004, and the B.S and M.S degrees from the North China Electric Power University, Baoding City, China, in 1998 and 2001, respectively. He is an associative professor of the School of Control and Computer Engineering, North China Electric Power University since 2007.

He is currently a Research Fellow in the Department of Electric and Electric Engineering, the University of Strathclyde. His research interests include wind energy converter condition monitoring, pitch control strategy, and wind speed prediction. He is the author or coauthor of technical papers in the above research area.

David Infield (SM'2007) was born in Paris, France in 1954 and raised and educated in England. He received a BA degree in mathematics and physics from the University of Lancaster, Lancaster, UK and a PhD degree in applied mathematics from the University of Kent, Canterbury, UK. He worked first for the Building Services Research and Information Association (BSRIA) at Bracknell, Berkshire, UK and then for the Rutherford Appleton Laboratory in Oxfordshire, UK from 1982 to 1993 researching into renewable energy supplied electricity systems. From 1993 to 2007 he was with Loughborough University, Leicestershire, UK where he established CREST, the Centre for Renewable Energy Systems Technology. He is now Professor of Renewable Energy Technologies with the Institute of Energy and Environment within the Department of Electronic and Electrical Engineering at the University of Strathclyde, Glasgow, UK. He is Editor in Chief of the IET journal Renewable Power Generation, and contributes to various IEC, CENELEC and IPPC activities.

Xiyun YANG was born in Tongliao,China, on September 1973. She received the B.E. and M.E. form Northeast Institute of Electric Power , Jilin, China, in 1994 and 1999, respectively. She received the Ph.D. degree from North China Electric Power University , Beijing, China, in 2004.

She is currently an Associate Professor in the School of Control and Computer engineering, North China Electric Power University. She has authored or coauthored more than 30 papers. Her research interests include automation, wind power system and so on.