Renewable Energy 71 (2014) 166–175

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Nonlinear system identification for model-based condition monitoring of wind turbines

Philip Cross*, Xiandong Ma

Engineering Department, Lancaster University, Lancaster LA1 4YR, United Kingdom

ARTICLE INFO

Article history:
Received 4 June 2013
Accepted 16 May 2014
Available online

Keywords:
Distributed generation (DG)
Wind turbine
Condition monitoring (CM)
Fault detection
Modelling and simulation
SCADA data

ABSTRACT

This paper proposes a data driven model-based condition monitoring scheme that is applied to wind turbines. The scheme is based upon a non-linear data-based modelling approach in which the model parameters vary as functions of the system variables. The model structure and parameters are identified directly from the input and output data of the process. The proposed method is demonstrated with data obtained from a simulation of a grid-connected wind turbine where it is used to detect grid and power electronic faults. The method is evaluated further with SCADA data obtained from an operational wind farm where it is employed to identify gearbox and generator faults. In contrast to artificial intelligence methods, such as artificial neural network-based models, the method employed in this paper provides a parametrically efficient representation of non-linear processes. Consequently, it is relatively straightforward to implement the proposed model-based method on-line using a field-programmable gate array.

1. Introduction

Recently, there has been increasing interest in the distributed generation (DG) of electricity due to the deregulation of utilities, environmental constraints, and concerns regarding climate change [1]. In the context of this paper, DG refers to electric power generation sources connected directly to the distribution network allowing for the integration of renewable energy resources, such as solar, CHP (Combined Heat and Power), and wind [2], with capacities ranging from less than 1 kW to over 100 MW.

The potential benefits of using DG include a more reliable electrical power supply and power loss reduction over transmission lines by generating electricity closer to the end user. Indeed, DG can provide power with little reliance on the grid. However, DG systems present several problems. Conventionally, voltage and reactive power control is based upon on the assumption that power flows from a generator to a substation, and subsequently to the feeders. The introduction of DG systems alters this power flow, causing problems with voltage regulation, equipment ratings to be exceeded, and protection schemes to be misdirected [3]. Consequently, it is vital for DG systems to operate reliably, and health condition monitoring and diagnostics schemes are essential.

Wind turbines represent a major form of distributed generation. It is common for wind turbines to be installed in remote locations on land or offshore, leading to difficulties in routine inspection and maintenance. In addition, wind turbines in these locations are often subject to harsh operating conditions. Over an operating life of 20 years, the maintenance costs for an offshore wind farm are estimated to be up to 30% of the total income [4]. Therefore, condition monitoring (CM) systems play an important role in the reliable operation of wind farms, providing information about the past and current condition of wind turbines and enabling optimal scheduling of maintenance activities, while minimising the risk of unexpected failure.

Given known input signals of a process, the corresponding output signals can be predicted using models obtained from the data generated by a monitoring system, such as SCADA (Supervisory Control and Data Acquisition) [5], and a CM scheme can be implemented by comparing actual output data with that predicted by the model. Any differences between the measurement and prediction signals could be 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. Clearly, an accurate model is essential for such a scheme, and previous research has employed a range of modelling techniques.

Mechanistic modelling techniques, for example using software such as Simulink [7], require a thorough understanding of the process, and may result in a complex or over-parameterised model

* Corresponding author. University of Leeds, Leeds LS2 9JT, United Kingdom. Tel.: +44 (0) 1524 593700.
E-mail addresses: p.cross@leeds.ac.uk (P. Cross), xiandong.ma@lancaster.ac.uk (X. Ma).

http://dx.doi.org/10.1016/j.renene.2014.05.035
0960-1481/

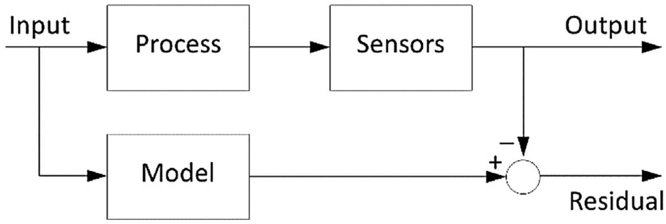

Fig. 1. Schematic diagram of the model-based condition monitoring system.

not suitable for on-line implementation as a CM system. Data-based models [8] do not require knowledge of the process or specific parameters, they are obtained directly from measured input and output signals collected either during planned experiments or by monitoring the process during normal operation. When implementing a CM scheme on-line using data-based models, it is essential to identify a low-order model. Higher-order models require increasingly more values of the past input and output signals to calculate the predicted output, resulting in an increased response time.

Many processes associated with wind turbines are non-linear. Artificial intelligence (AI) techniques are utilised by many researchers for model-based CM schemes, for example artificial neural networks (ANNs) [9] and fuzzy logic [10]. AI techniques are particularly suitable for this application being robust to noisy, incomplete and uncertain data. It should be noted that care needs to be taken to avoid 'over-fitting' of an ANN to the training data, which can lead to a poor performance when implemented on-line.

This paper proposes a non-linear data-based modelling approach to condition monitoring in which the model parameters vary as functions of the system variables, or 'state variables' [11]. The model structure and the model parameters are identified directly from the input and output data of the process. Although these state dependent parameter models cannot represent every type of non-linear behaviour, they are applicable to a wide range of processes, including those that behave chaotically. In contrast to ANNs, the data-based models employed in this paper are a parametrically efficient representation of non-linear processes, and are particularly suitable for forecasting [12], and providing the basis for automatic controller design [13]. However, the research described here represents the first occasion for which this type of model has been employed for a CM system.

The remainder of this paper is organised as follows. Section 2 describes the system identification methodology used in the research. A computer simulation of a grid-connected wind turbine is described in Section 3. Grid and power electronics faults are included in the simulation, and the non-linear model-based CM method is demonstrated by identifying these faults. In Section 4, non-linear models of an operational turbine are obtained using fault-free SCADA data, which are used to identify gearbox and generator faults in turbines on the same farm. Adaptive thresholds are derived from the model predictions, which in turn form the basis of an early warning system. A hardware implementation of the proposed approach, using an FPGA (field-programmable gate array), is presented in Section 5. Finally, conclusions to the research and future work are discussed in Section 6.

2. Model identification

A linear system can be represented by an auto-regressive with exogenous variables (ARX) model or auto-regressive moving average with exogenous variables (ARMAX) model. These model structures are usually represented by a discrete-time transfer function model, describing the relationship between the input and output signals as a ratio of polynomials. However, many processes associated with wind turbines are non-linear, and, 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 ARX models based upon ANNs have been utilised for wind turbine fault detection [9].

In this paper, non-linear processes are modelled using the following dynamic auto-regressive exogenous (DARX) model, in which the identified non-linearity is characterised by the time varying model parameters,

$$y_k = \frac{B(z^{-1},k)}{A(z^{-1},k)}u_k + \frac{1}{A(z^{-1},k)}e_k \tag{1}$$

where y_k and u_k are the kth sampled output and input variables respectively; e_k is white noise with zero mean and variance σ^2 that accounts for any random component of the observed data; $A(z^{-1},k)$ and $B(z^{-1},k)$ are appropriately defined polynomials in the backward shift operator, z^{-1},

$$A\left(z^{-1},k\right) = A\left(\chi_{1,k},z^{-1}\right) = 1 + a_1(\chi_{11,k})z^{-1} + \cdots + a_n(\chi_{1n,k})z^{-n}$$

$$\begin{aligned} B\left(z^{-1},k\right) = B\left(\chi_{1,k},z^{-1}\right) &= b_0(\chi_{21,k}) + b_1(\chi_{22,k})z^{-1} + \cdots \\ &\quad + b_m(\chi_{2m,k})z^{-m} \end{aligned} \tag{2}$$

in which the model parameters, $a_1 \ldots a_n$, $b_0 \ldots b_m$, are non-linear functions of the vectors $\chi_{1,k} = [\chi_{11} \ \cdots \ \chi_{1n}]$ and $\chi_{2,k} = [\chi_{21} \ \cdots \ \chi_{2m}]$ defined in terms of the system variables, or 'state variables', upon which the parameters are dependent. A time delay of δ samples is incorporated into the model by setting the leading δ coefficients of the $B(z^{-1},k)$ polynomial to zero.

The method of model identification employed in this paper comprises three stages [14]. In the first stage, the model structure and possible state dependent variables are identified by estimating a discrete-time linear transfer function model using any suitable method. The coefficient of determination [15] is employed here to determine the most appropriate model structure, and is defined,

$$R_T^2 = 1 - \frac{\sigma_e^2}{\sigma_y^2} \tag{3}$$

where σ_e^2 is the sample variance of the model residuals, that is, the difference between the model output and the actual output; 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 gives a good explanation of the actual output data. If the variances are similar in magnitude, R_T^2 tends to zero, indicating a poor fit.

During the second stage of model identification, stochastic time-varying parameter models are estimated using recursive Kalman filtering [16] and fixed-interval smoothing [17] algorithms. Since the variations in the parameters are functions of the state variables, the process can display severe non-linear or chaotic behaviour. Subsequently, these standard recursive estimation algorithms will not work satisfactorily. However, by sorting the data in a non-temporal order, for example, in ascending order so that the variations in the model parameters are slower and smoother, the standard identification methods can be used.

Equation (1) can be rewritten in the following vector-matrix form,