View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by IntechOpen

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter 7

Condition Monitoring of Wind Turbine Structures through Univariate and Multivariate Hypothesis Testing

Francesc Pozo and Yolanda Vidal

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.78727

Abstract

This chapter presents a fault detection method through uni- and multivariate hypothesis testing for wind turbine (WT) faults. A data-driven approach is used based on supervisory control and data acquisition (SCADA) data. First, using a healthy WT data set, a model is constructed through multiway principal component analysis (MPCA). Afterward, given a WT to be diagnosed, its data are projected into the MPCA model space. Since the turbulent wind is a random process, the dynamic response of the WT can be considered as a stochastic process, and thus, the acquired SCADA measurements are treated as a random process. The objective is to determine whether the distribution of the multivariate random samples that are obtained from the WT to be diagnosed (healthy or not) is related to the distribution of the baseline. To this end, a test for the equality of population means is performed in both the univariate and the multivariate cases. Ultimately, the test results establish whether the WT is healthy or faulty. The performance of the proposed method is validated using an advanced benchmark that comprehends a 5-MW WT subject to various actuators and sensor faults of different types.

Keywords: condition monitoring, wind turbines, principal component analysis, hypothesis testing

1. Introduction

The wind energy cost depends strongly on the performance of the condition monitoring system. Advance in this area would decrease downtime periods, extend the WT lifetime, and ultimately reduce the operation and maintenance (O&M) costs, which is one of the main challenges in wind energy as stated in "20% Wind Energy by 2030" [1].

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

Usually, condition monitoring comprises different systems (vibration analysis, oil monitoring, etc. [2]) for different parts and different types of faults and makes use of expensive specific sensors that must be installed in the WT. Therefore, the advance in fault detection systems that only make use of already available data from the turbine SCADA system and comprehend different parts and different types of faults is promising (since no additional sensors or data acquisition devices are needed). The SCADA signals provide rich information on the WT performance; thus, with appropriate algorithms, they can be used effectively for condition monitoring, prognostics, and remaining useful life prediction of WTs [3]. There are some success stories about using SCADA data for condition monitoring. For example, Ruiz et al. presented a machine learning approach [4], Zaher and McArthur proposed to use the combination of abnormal detection and data-trending techniques encapsulated in a multiagent framework [5], Pozo and Vidal proposed a fault detection system based on principal component analysis [6].

In this work, following the enhanced benchmark challenge for wind turbine fault detection proposed in [7], a set of eight realistic fault scenarios are considered to develop a WT condition monitoring strategy that combines a SCADA data-driven baseline model—reference pattern obtained from the healthy wind turbine—based on MPCA in combination with uni- and multivariate hypothesis testing. Previous works using MPCA and hypothesis testing to detect structural damage [8] work under the hypothesis of guided waves. That is, the vibration (guided wave) induced to the structure is known and always the same. However, in this work, the vibration is induced by the changeful wind. The used benchmark comprehends different types of faults of a 5-MW WT given by the FAST simulator [9], which has been accepted by the scientific community and is widely used for WT-related research, e.g., [10–12].

The chapter is organized as follows. Section 2 briefly recalls the WT benchmark model. In Section 3, the condition monitoring strategy is stated. Simulation results are discussed in Section 4. Finally, conclusions are drawn in Section 5.

2. Wind turbine benchmark model

The used benchmark model is proposed in [7]. It covers a 5-MW three-bladed, variable speed WT modeled with the FAST simulator, detailed actuator and sensor models, as well as the different fault descriptions. For a complete description of the benchmark, please see reference [7]. Here, a short review is given to introduce the used notation.

The specifications of the 5-MW reference WT is documented in [13]. This model has been used as a reference by research teams throughout the world to standardize baseline on- and offshore wind turbine specifications. The wind turbine typical features are given in **Table 1**, and the assumed available SCADA data are given in **Table 2**. This work copes with the so-called full load region of operation. In order to run the simulations, turbulent wind data sets that cover this region have been generated with TurbSim [14], see **Figure 1**.

The generator-converter system can be approximated by a first-order ordinary differential equation, see [7], which is given by:

Reference wind turbine	Magnitude
Rated power	5 MW
Number of blades	3
Rotor/hub diameter	126, 3 m
Hub height	90 m
Cut-in, rated, and cut-out wind speed	3, 11.4, and 25 m/s
Rated generator speed (ω_{ng})	1173.7 rpm
Gearbox ratio	97

Table 1. WT properties.

Number	Sensor type	Symbol	Units
1	Generated electrical power	$P_{e,m}$	kW
2	Rotor speed	$\omega_{r,m}$	rad/s
3	Generator speed	$\omega_{g,m}$	rad/s
4	Generator torque	$\tau_{c,m}$	Nm
5	First pitch angle	$\beta_{1,m}$	°
6	Second pitch angle	$\beta_{2,m}$	°
7	Third pitch angle	$\beta_{3,m}$	°
8	Fore-aft acceleration at tower bottom	$a^b_{fa,m}$	m/s^2
9	Side-to-side acceleration at tower bottom	$a^b_{ss,m}$	m/s^2
10	Fore-aft acceleration at mid-tower	$a^m_{fa,m}$	m/s^2
11	Side-to-side acceleration at mid-tower	$a^m_{ss,m}$	m/s^2
12	Fore-aft acceleration at tower top	$a^t_{fa,m}$	m/s^2
13	Side-to-side acceleration at tower top	$a^t_{ss,m}$	m/s^2

These sensors are representative of the types of sensors that are available on an MW-scale commercial wind turbine.

Table 2. Assumed available measurements.

$$\dot{\tau}_r(t) + \alpha_{gc}\tau_r(t) = \alpha_{gc}\tau_c(t) \tag{1}$$

where τ_r and τ_c are the real generator torque and its reference (given by the controller), respectively. In the numerical simulations, $\alpha_{gc} = 50$, see [13]. Moreover, the power produced by the generator, $P_e(t)$, is given by (see [7]):

$$P_e(t) = \eta_g \omega_g(t)\tau_r(t) \tag{2}$$

where η_g is the efficiency of the generator and ω_g is the generator speed. In the numerical experiments, $\eta_g = 0.98$ is used, see [7].

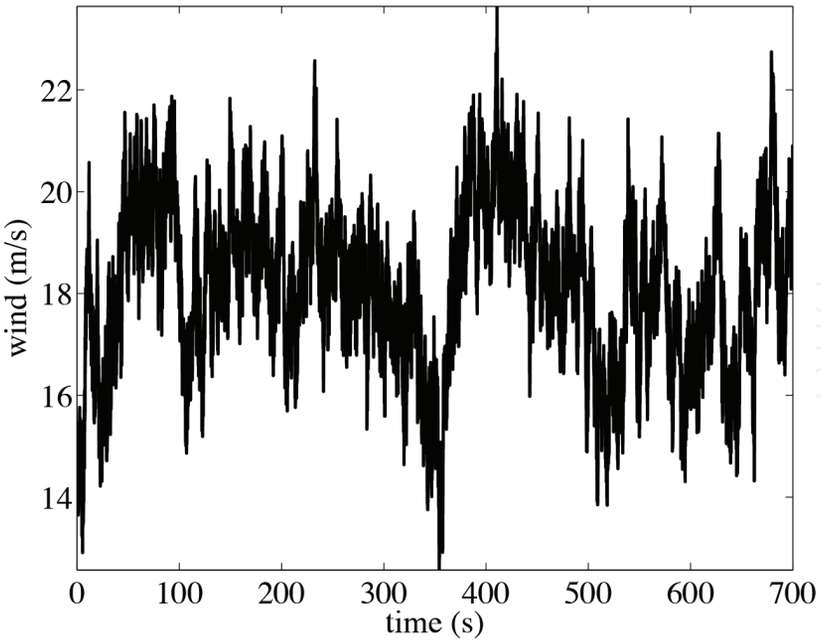

Figure 1. Wind speed signal with turbulence intensity set to 10%.

Each of the three pitch actuators is modeled as a closed loop transfer function between the pitch angle, $\beta(s)$, and its reference $\beta_r(s)$:

$$\frac{\beta(s)}{\beta_r(s)} = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2} \tag{3}$$

where ξ is the damping ratio and ω_n the natural frequency that takes the fault-free values $\xi = 0.6$ and $\omega_n = 11.11$ rad/s, see [7].

The fault detection benchmark considers different types of faults at different components (sensors and actuators), as described in **Table 3**.

Fault	Type	Description
F1	Pitch actuator	Change in dynamics: high air content in oil
F2	Pitch actuator	Change in dynamics: pump wear
F3	Pitch actuator	Change in dynamics: hydraulic leakage
F4	Torque actuator	Offset (offset value equal to 2000 Nm)
F5	Generator speed sensor	Scaling (gain factor equal to 1.2)
F6	Pitch angle sensor	Stuck (fixed value equal to 5°)
F7	Pitch angle sensor	Stuck (fixed value equal to 10°)
F8	Pitch angle sensor	Scaling (gain factor equal to 1.2)

Table 3. Fault scenarios.

3. Condition monitoring (CM) strategy

The overall CM strategy is based on a three-tier framework:

i. a multiway PCA (MPCA) model is built with the data that are collected from a healthy WT,

ii. when a new WT has to be diagnosed, the SCADA data are projected using the MPCA model created in (i), and

iii. the final decision is based on both univariate and multivariate HT.

3.1. The wind as a source for the excitation: the need for a new paradigm

In general, vibration-based structural health monitoring (SHM) is based on the fact that an alteration or difference in physical properties due to damage or structural change will motivate changes in dynamical responses that may be detected. **Figure 2** represents this paradigm in the sense that a healthy structure is excited according to a prescribed signal to build a pattern. Afterward, the structure that has to be diagnosed is affected by exactly the same signal, where the response is measured, processed, and finally compared with the previous pattern. The strategy presented in **Figure 2** is known as "guided waves in structures for SHM" [15].

In the present chapter, the field of application is wind turbines and a realistic scenario is to consider that the excitation comes from the wind turbulence. The wind turbulence cannot be controlled and it is always different. Therefore, the paradigm of guided waves in WT for SHM as in **Figure 2** cannot be considered. In this case, when the source of the excitation cannot be previously prescribed, a new paradigm is needed, as represented in **Figure 3**. The foundation of the new paradigm is that, even with a constantly different excitation, the CM strategy based on MPCA and univariate and multivariate HT will be able to disclose some hidden damage, misbehavior, or fault. To sum up, the fundamental idea behind the CM strategy is the hypothesis that a variation in the overall behavior of the WT, even with an unprescribed excitation, should be detected.

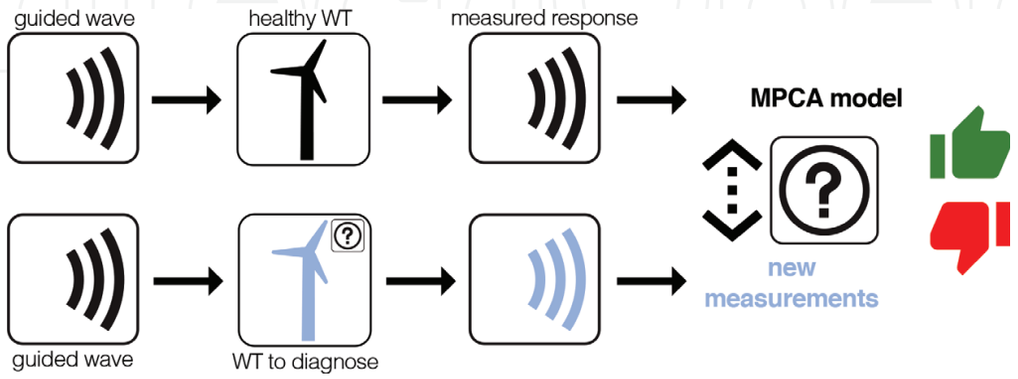

Figure 2. Vibration-based SHM is based on the fact that an alteration or difference in physical properties due to damage or structural change will motivate changes in dynamical responses that may be detected.

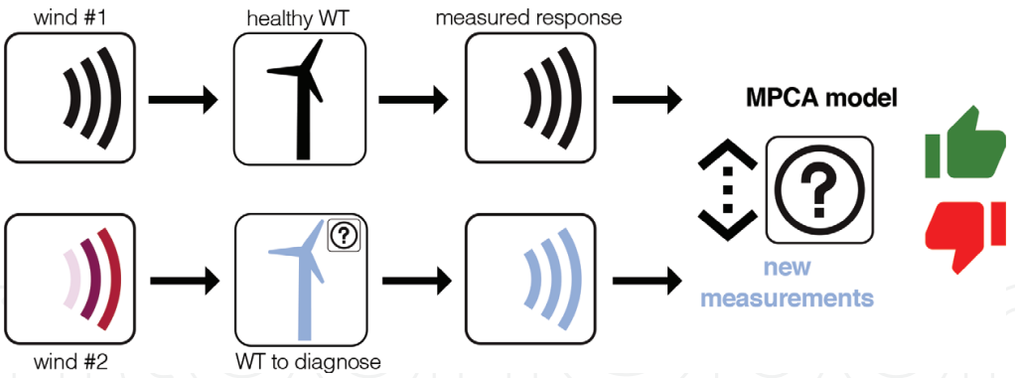

Figure 3. The key idea behind the new paradigm of the detection strategy is the assumption that a change in the behavior of the overall system, even with a different excitation, has to be detected.

However, in our application, the only available excitation of the wind turbines is the wind turbulence. Therefore, guided waves in wind turbines for SHM as in **Figure 2** cannot be considered as a realistic scenario. In spite of that, the new paradigm described in **Figure 3** is based on the fact that, even with different wind turbulence, the fault detection strategy based on PCA and statistical multivariate hypothesis testing will be able to detect some damage, fault, or misbehavior. More precisely, the key idea behind the detection strategy is the assumption that a change in the behavior of the overall system, even with a different excitation, has to be detected. Section 4 includes the simulation results of the proposed CM strategy that validates this hypothesis.

3.2. Data-driven baseline modeling based on MPCA

Multiway principal component analysis (MPCA) is a natural extension of classical principal component analysis (PCA) to manage data in multidimensional arrays [16, 17]. A conventional two-dimensional data matrix can be treated as a two-way array, where experiments and variables (or discretization instant times) form the two different ways. Frequently, this arrangement has to be extended to multiway arrays, particularly if several sensors—in different experimental trials—are gathering data at different time instants. Consequently, MPCA is equivalent to the application of standard PCA to an unfolded version of the initial multiway array.

Westerhuis et al. [18] propose six different ways of unfolding a three-way data matrix. Besides, in [18], a critical analysis of several aspects of the treatment of multiway data is provided, including how the matrix is unfolded, but also mean-centering and scaling with respect to the effects on the analysis of batch data. Ruiz et al. [19] assign one of the first six letters of the alphabet to each one of the six different ways of unfolding. In this chapter, as well as in [6, 8, 20, 21], we have considered the so-called type E. However, we will present the collected SCADA data arranged in an already unfolded matrix.

The MPCA modeling starts by measuring, from a healthy wind turbine, a sensor during $(nL-1)\Delta$ seconds, where Δ is the sampling time and $n, L \in \mathbb{N}$. The discretized measures of the sensor are a real vector

$$(x_{11}\ \ x_{12}\ \ \cdots\ \ x_{1L}\ \ x_{21}\ \ x_{22}\ \ \cdots\ \ x_{2L}\ \ \cdots\ \ x_{n1}\ \ x_{n2}\ \ \cdots\ \ x_{nL}) \in \mathbb{R}^{nL} \tag{4}$$

where the real number x_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, L$ corresponds to the measure of the sensor at time $((i-1)L + (j-1))\Delta$ seconds. These collected data can be arranged in matrix form as follows:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1L} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{iL} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nL} \end{pmatrix} \in \mathscr{M}_{n\times L}(\mathbb{R}) \tag{5}$$

where $\mathscr{M}_{n\times L}(\mathbb{R})$ is the vector space of $n \times L$ matrices over $\mathbb{R}$. It is worth noting that n is the number of rows of the matrix in Eq. (5) and L is the number of columns of the same matrix. The effect on the overall performance of the condition monitoring strategy on the choice of n and L is thoroughly analyzed on [21].

Let us assume that the SCADA data are now collected from $N \in \mathbb{N}$ sensors also during the same period of time. In this case, the collected data, for each sensor, can be organized in a matrix as in Eq. (5). Subsequently, all the collected data coming from the whole set of sensors are concatenated and disposed in a matrix $\mathbf{X} \in \mathscr{M}_{n\times(N\cdot L)}$ as follows:

$$\begin{aligned} \mathbf{X} &= \begin{pmatrix} x_{11}^1 & x_{12}^1 & \cdots & x_{1L}^1 & x_{11}^2 & \cdots & x_{1L}^2 & \cdots & x_{11}^N & \cdots & x_{1L}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1}^1 & x_{i2}^1 & \cdots & x_{iL}^1 & x_{i1}^2 & \cdots & x_{iL}^2 & \cdots & x_{i1}^N & \cdots & x_{iL}^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1}^1 & x_{n2}^1 & \cdots & x_{nL}^1 & x_{n1}^2 & \cdots & x_{nL}^2 & \cdots & x_{n1}^N & \cdots & x_{nL}^N \end{pmatrix} \\ &= \left(\underbrace{v_1|v_2|\cdots|v_L}_{\mathbf{X}^1} \underbrace{|v_{L+1}|\cdots|v_{2L}}_{\mathbf{X}^2} |\cdots| \underbrace{v_{(N-1)L+1}|\cdots|v_{N\cdot L}}_{\mathbf{X}^N} \right) \\ &= \begin{pmatrix} \mathbf{X}^1 & \mathbf{X}^2 & \cdots & \mathbf{X}^N \end{pmatrix} \in \mathscr{M}_{n\times(N\cdot L)}(\mathbb{R}) \end{aligned} \tag{6}$$

where the superindex $k = 1, \ldots, N$ of each element x_{ij}^k in the matrix represents the number of sensor. Matrix $\mathbf{X} \in \mathscr{M}_{n\times(N\cdot L)}(\mathbb{R})$—where $\mathscr{M}_{n\times(N\cdot L)}(\mathbb{R})$ is the vector space of $n \times (N \cdot L)$ matrices over $\mathbb{R}$—contains the measures from N sensors at nL discretization instants. Consequently, each row vector $x_i^T = \mathbf{X}(i,:) \in \mathbb{R}^{N\cdot L}$, $i = 1, \ldots, n$ represents the measurements from all the sensors at time instants $((i-1)L + (j-1))\Delta$ seconds, $j = 1, \ldots, L$. Equivalently, each column vector $v_j = \mathbf{X}(:,j) \in \mathbb{R}^n$, $j = 1, \ldots, N \cdot L$ represents measurements from sensor number $\left\lceil \frac{j}{L} \right\rceil$ at time instants $((i-1)L + (j-1))\Delta$ seconds, $1 = 1, \ldots, n$, where $\lceil \cdot \rceil$ is the ceiling function.

The objective of the subsequent analysis is to build the MPCA model, that is, the square orthogonal matrix $\mathbf{P} \in \mathscr{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R})$ that has to be used to transform or project the original data matrix $\mathbf{X}$ according to the following matrix-to-matrix product:

$$\mathbf{T} = \mathbf{XP} \in \mathscr{M}_{n\times(N\cdot L)}(\mathbb{R}), \tag{7}$$

where the shape of the variance-covariance matrix of matrix **T** in Eq. (7) is diagonal.

In the proposed approach in this chapter, the model defined in matrix **P** in Eq. (7) is based only on measures that come from a *healthy* wind turbine. Posteriorly, data from the current WT to diagnose will be projected using the matrix-to-matrix multiplication also defined in Eq. (7). However, a different procedure can be considered, particularly, when the goal is not just to detect a damage or a fault but to classify it. In the latter case, matrix **X** in Eq. (6) should contain measures from a WT in its healthy state but also in all the possible fault scenarios. This way, the generated model in matrix **P** in Eq. (7) contains all the possible states of the structure.

3.2.1. Centering and scaling: group scaling (GS) vs. mean-centered group scaling (MCGS)

Considering that the data stored in matrix **X** are affected by a changing wind turbulence, come from different sensors, and could have different magnitudes and scales, some kind of preprocessing step is required to rescale the data [22, 23]. According to Westerhuis et al. [18], the way this preprocessing step is carried out may affect the overall performance of the CM strategy. In the present chapter, we present two possible choices that have some common core. These two alternatives are as follows:

i. group scaling (GS) and

ii. mean-centered group scaling (MCGS).

In the former case (GS), both the arithmetic mean and the variance of all measurements of the sensor are used. More precisely, for $k = 1, 2, \ldots, N$, we define

$$\mu^k = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L} x_{ij}^k, \tag{8}$$

$$\sigma_k^2 = \frac{1}{nL}\sum_{i=1}^{n}\sum_{j=1}^{L}\left(x_{ij}^k - \mu^k\right)^2 \tag{9}$$

where μ^k and σ_k^2 are the arithmetic mean and the variance of the whole set of elements in matrix $\mathbf{X}^k$, respectively. In this case, matrix $\mathbf{X} = \left(x_{ij}^k\right)$ is centered and scaled—using GS—to define a modified matrix $\breve{\mathbf{X}} = \mathbf{X}_{\mathrm{GS}} = \left(\breve{x}_{ij}^k\right)$ as

$$\breve{x}_{ij}^k := \frac{x_{ij}^k - \mu^k}{\sqrt{\sigma_k^2}}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, L, \quad k = 1, \ldots, N. \tag{10}$$

In the latter case (MCGS), the arithmetic of all measurements of the sensor at the same column is considered in the normalization. More precisely, for $k = 1, 2, \ldots, N$, we define

$$\mu_j^k = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^k, \quad j = 1, \ldots, L, \tag{11}$$

where μ_j^k is the arithmetic mean of the measures placed at the same column. In this case, then, matrix $\mathbf{X} = \left(x_{ij}^k\right)$ is centered and scaled—using MCGS—to define a modified matrix $\breve{\mathbf{X}} = \mathbf{X}_{\text{MCGS}} = \left(\breve{x}_{ij}^k\right)$ as

$$\breve{x}_{ij}^k := \frac{x_{ij}^k - \mu_j^k}{\sqrt{\sigma_k^2}}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, L, \quad k = 1, \ldots, N. \tag{12}$$

where σ_k^2 is defined as in Eq. (9) using μ^k as in Eq. (8). It is worth noting that the only difference between the expressions in Eqs. (10) and (12) is how the elements in matrix $\mathbf{X} = \left(x_{ij}^k\right)$ are *centered*. When matrix $\mathbf{X} = \left(x_{ij}^k\right)$ is scaled and centered according to the MCGS strategy described in Eq. (12), the average value of each column vector in the scaled matrix $\breve{\mathbf{X}}$ can be calculated as

$$\frac{1}{n}\sum_{i=1}^{n} \breve{x}_{ij}^k = \frac{1}{n}\sum_{i=1}^{n} \frac{x_{ij}^k - \mu_j^k}{\sigma^k} = \frac{1}{n\sigma^k}\sum_{i=1}^{n} \left(x_{ij}^k - \mu_j^k\right) \tag{13}$$

$$= \frac{1}{n\sigma^k}\left[\left(\sum_{i=1}^{n} x_{ij}^k\right) - n\mu_j^k\right] \tag{14}$$

$$= \frac{1}{n\sigma^k}\left(n\mu_j^k - n\mu_j^k\right) = 0 \tag{15}$$

Taking advantage of the fact that the scaled matrix $\breve{\mathbf{X}}$ is a mean-centered matrix, the variance-covariance matrix can be straightforwardly computed as a matrix-to-matrix product of $\breve{\mathbf{X}}$ and its transpose, divided by $n - 1$, where n is the number of rows of matrix $\mathbf{X}$ in Eq. (6). More precisely,

$$\mathbf{C}_{\breve{\mathbf{X}}} = \frac{1}{n-1}\breve{\mathbf{X}}^T\breve{\mathbf{X}} \in \mathscr{M}_{(N\cdot L)\times(N\cdot L)}(\mathbb{R}) \tag{16}$$

Clearly, GS and MCGS are not the only ways to center and scale data. For instance, feature scaling, also known as unity-based normalization, can also be considered. In this case, data are centered with respect to the minimum value and scaled with respect to the range of the set, that is,

$$\tilde{x}_{ij}^k := \frac{x_{ij}^k - \min\left\{x_{ij}^k\right\}}{\max\left\{x_{ij}^k\right\} - \min\left\{x_{ij}^k\right\}}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, L, \quad k = 1, \ldots, N. \tag{17}$$

However, to easily compute the variance-covariance matrix in the CM strategy that we present in this chapter, the mean-centered group scaling (MCGS) is the method that we have selected

for the centering and scaling. In order to not to use the baroque notation $\breve{\mathbf{X}}$ throughout the rest of this chapter, this centered and scaled matrix is redesignated as $\mathbf{X}$, without the breve sign.

The MPCA model is described by the latent vectors

$$p_j, \quad j = 1, \ldots, N \cdot L, \tag{18}$$

also known as eigenvector or proper vectors, and the latent roots

$$\lambda_j, \quad j = 1, \ldots, N \cdot L, \tag{19}$$

also known as eigenvalues or proper values, of the variance-covariance matrix $\mathbf{C_X}$ as follows:

$$\mathbf{C_X P} = \mathbf{P}\Lambda \tag{20}$$

where

$$\mathbf{P} = \left(p_1 | p_2 | \cdots | p_{N \cdot L}\right) \in \mathscr{M}_{N \cdot L \times N \cdot L}(\mathbb{R}) \tag{21}$$

$$\Lambda = \left(\Lambda_{ij}\right) \in \mathscr{M}_{N \cdot L \times N \cdot L}(\mathbb{R}) \tag{22}$$

and

$$\Lambda_{jj} = \lambda_j, \quad j = 1, \ldots, N \cdot L \tag{23}$$

$$\Lambda_{ij} = 0, \quad i, j = 1, \ldots, N \cdot L, \quad i \neq j \tag{24}$$

The latent vectors and latent roots in Eqs. (21) and (23) are arranged in descending order with respect to the absolute values of the latent roots, that is,

$$|\lambda_i| \geq |\lambda_{i+1}|, \quad i = 1, \ldots, N \cdot L - 1 \tag{25}$$

The latent vector p_1—corresponding to the largest latent root λ_1 (in absolute value)—is called the first principal component (PC). Likewise, the latent vector p_2—corresponding to the second largest latent root λ_2 (in absolute value)—is called the second principal component. Equivalently, the latent vector $p_j, \quad j = 1, \cdots, N \cdot L$—corresponding to the latent root λ_j—is called the j−th principal component.

Matrix $\mathbf{T}$ in Eq. (7) represents the transformed or projected matrix onto the principal component space and it is also known as score matrix.

When, for the sake of dimensionality reduction, a decreased number of principal components are considered:

$$\ell < N \cdot L, \tag{26}$$

a reduced multiway PCA model is then assembled:

$$P = \left(p_1 | p_2 | \cdots | p_\ell\right) \in \mathscr{M}_{N \cdot L \times \ell}(\mathbb{R}). \tag{27}$$

3.3. HT-based condition monitoring

As said in Section 3.2, the MPCA model is based only on measures that come from a healthy wind turbine. Posteriorly, data from the current WT to diagnose—and subjected to a different wind turbulence—are gathered from as many sensors as in the modeling phase described in Section 3.2 and during a period of time, $(\nu L - 1)\Delta$ seconds, which is not necessarily equal. These new data are arranged in a new matrix $\mathbf{Y}$ in a similar way as in Eq. (6):

$$\begin{aligned}\mathbf{Y} &= \begin{pmatrix} y_{11}^1 & y_{12}^1 \cdots y_{1L}^1 y_{11}^2 \cdots y_{1L}^2 \cdots y_{11}^N \cdots y_{1L}^N \\ \vdots & \vdots \ddots \vdots \vdots \ddots \vdots \vdots \ddots \vdots \\ y_{i1}^1 & y_{i2}^1 \cdots y_{iL}^1 y_{i1}^2 \cdots y_{iL}^2 \cdots y_{i1}^N \cdots y_{iL}^N \\ \vdots & \vdots \ddots \vdots \vdots \ddots \vdots \vdots \ddots \vdots \\ y_{\nu 1}^1 & y_{\nu 2}^1 \cdots y_{\nu L}^1 y_{\nu 1}^2 \cdots y_{\nu L}^2 \cdots y_{\nu 1}^N \cdots y_{\nu L}^N \end{pmatrix} \in \mathscr{M}_{\nu \times (N \cdot L)}(\mathbb{R}) \\ &= \left(\underbrace{w_1 | w_2 | \cdots | w_L}_{\mathbf{Y}^1} | \underbrace{w_{L+1} | \cdots | w_{2L}}_{\mathbf{Y}^2} | \cdots | \underbrace{w_{(N-1)L+1} | \cdots | w_{N \cdot L}}_{\mathbf{Y}^N} \right) \\ &= \begin{pmatrix} \mathbf{Y}^1 & \mathbf{Y}^2 \cdots \mathbf{Y}^N \end{pmatrix} \in \mathscr{M}_{n \times (N \cdot L)}(\mathbb{R}) \end{aligned} \tag{28}$$

It should be noted that $\nu \in \mathbb{N}$ (the number of rows of matrix $\mathbf{Y}$) does not necessarily need to match the natural number n, which represents the number of rows of matrix $\mathbf{X}$ in Eq. (6). However, the number of columns, represented by the natural number $N \cdot L$, must agree.

The collected data in matrix $\mathbf{Y}$ in Eq. (28) are first centered and scaled to form a matrix $\breve{\mathbf{Y}} = \left(\breve{y}_{ij}^k \right)$ similar to the one in Eq. (12):

$$\breve{y}_{ij}^k := \frac{y_{ij}^k - \mu_j^k}{\sqrt{\sigma_k^2}}, \quad i = 1, \ldots, \nu, \quad j = 1, \ldots, L, \quad k = 1, \ldots, N, \tag{29}$$

where σ_k^2 and μ_j^k are the values of the variance and the arithmetic mean that have been previously calculated in Eqs. (9) and (11), respectively, with respect to $\mathbf{X}$ in Eq. (6). After the preprocessing step, that is, centering and scaling the raw data collected from the current structure to diagnose, the scores related to each row vector

$$r^i = \breve{\mathbf{Y}}(i, :) \in \mathbb{R}^{N \cdot L}, \quad i = 1, \ldots, \nu \tag{30}$$

are computed using a vector-to-matrix product:

$$t^i = r^i \cdot \hat{\mathbf{P}} \in \mathbb{R}^{\ell}, \quad i = 1, \ldots, \nu \tag{31}$$

where matrix $\hat{\mathbf{P}}$ is the reduced MPCA model in Eq. (27).

Let us consider the canonical basis

$$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_\ell\} \subset \mathbb{R}^\ell \tag{32}$$

of the $\ell-$dimensional real vector space $\mathbb{R}^\ell$.

Given a row vector r^i as in Eq. (30), the real number

$$t_1^i = t^i \cdot \mathbf{e}_1 \in \mathbb{R} \tag{33}$$

is called the first score. Likewise, the scalar

$$t_2^i = t^i \cdot \mathbf{e}_2 \in \mathbb{R} \tag{34}$$

is called the second score. In general, the scalar

$$t_j^i = t^i \cdot \mathbf{e}_j \in \mathbb{R} \tag{35}$$

is called the score associated with the principal component p_j, $j = 1, \ldots, \ell$ or, simply, score j.

In addition, an $s-$dimensional vector as can be built if more than one score is considered at the same time. Indeed,

$$\mathbf{t}_s^i = \begin{bmatrix} t_1^i & t_2^i & \cdots & t_s^i \end{bmatrix}^T \in \mathbb{R}^s, \quad s \leq \ell. \tag{36}$$

3.3.1. Scores as a random sample

As said in Section 3.1, the excitation of the WT comes from a changing turbulent wind. Somehow, this turbulent wind can be viewed as a random signal. Therefore, the response of the WT can be also viewed as a random process and so the measurements in the row vector r^i in Eq. (30). As a consequence, the vector t^i receives this random nature and it can be observed as an ℓ-dimensional random vector to construct the statistical approach in this chapter. As a motivating example, in **Figure 4**, two three-dimensional samples are represented: one is the

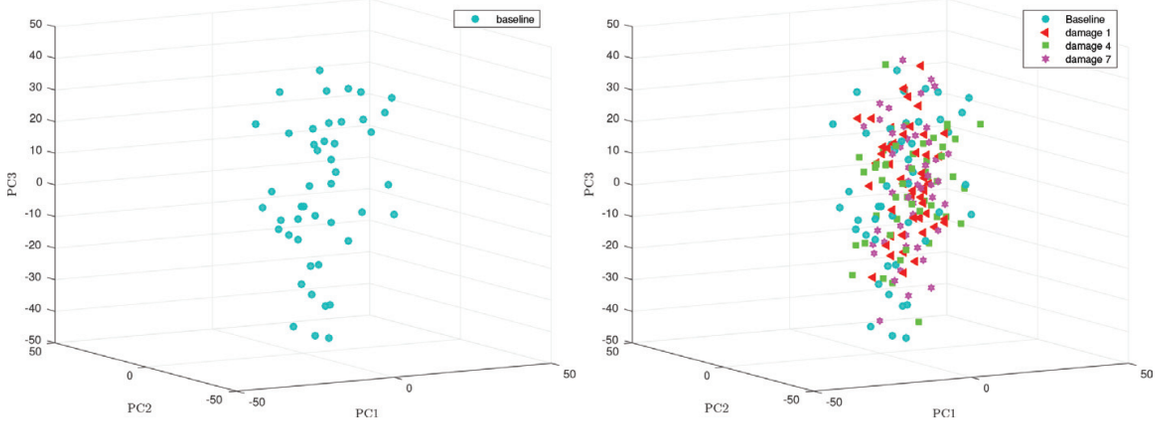

Figure 4. Baseline sample (left) and sample from the wind turbine to be diagnosed (right).

three-dimensional baseline sample (left) and the other is referred to faults 1, 4, and 7 (right). In a classic application of the PCA strategy in the field of SHM, the scores allow a separation, clustering, or visual grouping [24]. However, in this case, it can be clearly monitored in **Figure 4** (right) that a clustering, visual grouping, or separation cannot be performed. Therefore, more powerful and reliable tools are needed to be able to detect a fault in the WT.

In structural health monitoring or condition monitoring applications, the final decision on whether the structure, the actuator and/or the sensor is healthy or not should not depend on graphical approaches. One of the most common approaches to reliable indicators of damage or faults is the use of the powerful machinery of statistical hypothesis testing. The differences in this kind of strategies rely on what is the subject of the test and, of course, how the raw data collected by the sensors are arranged and preprocessed. For instance, in Zugasti et al. [25] the damage detection is based on testing for significant changes in the parameter vector of an AutoRegressive model. A comprehensive three-tier modular structural health monitoring framework is proposed by Hackell et al. [26] where the hypothesis testing is used to declare decision boundaries, control charts, and ROC curves with the ultimate goal of distinguishing between healthy and potentially damaged data on an operational wind turbine. A somehow different approach is presented by Ng et al. [27] that includes a vehicle health monitoring system where several univariate hypothesis tests are considered in parallel. Again in the field of structural health monitoring or condition monitoring of wind turbines, a recent work by Tsiapoki et al. [28] where damage and ice detection is based on data normalization, feature extraction and hypothesis testing (HT).

The use of univariate hypothesis testing as a key element for structural health monitoring or condition monitoring has been increasing in the last years as a reliable method. Variations of these univariate HT for multiple indicators include the use of univariate HT in parallel, that is, testing for each component of a parameter vector rather than testing for the whole multidimensional parameter vector. The first approach for the detection of structural changes using a multivariate hypothesis testing has been proposed by Pozo et al. [8]. One of the key results in the work [8] is that multivariate HTs allow to get better results in damage or fault detection that just univariate test. One interesting example presented in the work by Pozo et al. [8] shows that, for a given level of significance α, five independent univariate hypothesis

$$\begin{aligned} H_0: &\quad \mu_{\mathrm{c},i} = \mu_{\mathrm{h},i} \\ H_1: &\quad \mu_{\mathrm{c},i} \neq \mu_{\mathrm{h},i} \end{aligned} \tag{37}$$

where $i = 1, 2, \ldots, 5$ lead to a wrong decision while the single multivariate HT

$$\begin{aligned} H_0: &\quad \boldsymbol{\mu}_{\mathrm{c}} = \boldsymbol{\mu}_{\mathrm{h}} \\ H_1: &\quad \boldsymbol{\mu}_{\mathrm{c}} \neq \boldsymbol{\mu}_{\mathrm{h}} \end{aligned} \tag{38}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\mathrm{c}}^T &= \begin{bmatrix} \mu_{\mathrm{c},1} & \mu_{\mathrm{c},2} & \cdots & \mu_{\mathrm{c},5} \end{bmatrix} \\ \boldsymbol{\mu}_{\mathrm{h}}^T &= \begin{bmatrix} \mu_{\mathrm{h},1} & \mu_{\mathrm{h},2} & \cdots & \mu_{\mathrm{h},5} \end{bmatrix} \end{aligned} \tag{39}$$

is able to correctly classify the structure. This example shows that multivariate HT is even more reliable than univariate HT. However, these benefits come at a price, in the sense that in order to apply the multivariate HT, the statistical distribution of the data must be multinormal. Of course, it may happen that five sets of 50 samples

$$\left\{x_1^i, x_2^i, \ldots, x_{50}^i\right\} \mapsto N\left(\mu_i, \sigma_i\right), \quad i = 1, 2, \ldots, 5 \tag{40}$$

are normally distributed, while the sample vector

$$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{50}\} \not\mapsto N\left(\mu, \Sigma\right), \tag{41}$$

where

$$\mathbf{x}_j = \begin{bmatrix} x_j^1 & x_j^2 & \cdots & x_j^5 \end{bmatrix}^T, \quad j = 1, \ldots, 50 \tag{42}$$

and Σ is the variance-covariance matrix, is not multinormally distributed.

3.3.2. Univariate case: testing for the equality of means

In this section, we present how a fault is detected in the WT using univariate HT. To this end, first we have to define what we consider our *baseline*. Given a principal component $j = 1, \ldots, \ell$, the baseline sample is the set of real numbers $\left\{\tau_j^i\right\}_{i=1,\ldots,n}$ defined by

$$\tau_j^i := \left(\mathbf{X}(i,:) \cdot \widehat{\mathbf{P}}\right)(j) = \mathbf{X}(i,:) \cdot \widehat{\mathbf{P}} \cdot \mathbf{e}_j, \quad i = 1, \ldots, n, \tag{43}$$

where $\mathbf{e}_j$ is the j-th vector of the canonical basis in Eq. (32), P is the MPCA model defined in Eq. (27), and $\mathbf{X}$ is the centered and scaled matrix of the collected data from a *healthy* WT as in Eq. (6). Similarly, and given a principal component $j = 1, \ldots, \ell$, the sample of the current WT to diagnose is defined as the set of ν real numbers

$$\left\{t_j^i\right\}_{i=1,\ldots,\nu} \tag{44}$$

as defined in Eq. (35).

Before the univariate HT is applied, the following assumptions must be made:

i. the baseline sample $\left\{\tau_j^i\right\}_{i=1,\ldots,n}$ is a random sample of a random variable (RV) normally distributed with unknown mean μ_X and unknown variance σ_X^2 and

ii. the random sample $\left\{t_j^i\right\}_{i=1,\ldots,\nu}$ in Eq. (44) of the current WT to diagnose follows a normal distribution with unknown mean μ_Y and unknown variance σ_Y^2.

It is worth mentioning that the variances of these two samples are not supposed to be necessary equal.

http://dx.doi.org/10.5772/intechopen.78727

Let us define

$$\delta\mu = \mu_X - \mu_Y \tag{45}$$

as the difference between these two mean values. Since we want to know if the distribution of these two samples is related, this leads to a test of the hypothesis

$$H_0: \ \delta\mu = 0 \ \text{ versus} \tag{46}$$

$$H_1: \ \delta\mu \neq 0 \tag{47}$$

where the null hypothesis H_0 is "the sample of the WT to be diagnosed is distributed as the baseline sample" and the alternative hypothesis H_1 is "the sample of the WT to be diagnosed is not distributed as the baseline sample." In other words, if the result of the test is that H_0 is accepted, the current WT is categorized as healthy. Otherwise, if H_0 is rejected in favor of H_1, this would indicate the presence of some faults in the WT.

Given the assumptions of normality and considering that the two variances are not necessarily equal, the test for the equality of mean is based on the so-called Welch-Satterthwaite method [29], which is outlined below for the sake of completeness. If two random samples of size n and ν, respectively, are taken from two normal distributions $N(\mu_X, \sigma_X)$ and $N(\mu_Y, \sigma_Y)$ and the population variances are unknown and not necessarily equal, the random variable

$$\mathscr{WS} = \frac{\left(\overline{X} - \overline{Y}\right) + \left(\mu_X - \mu_Y\right)}{\sqrt{\left(\frac{S_X^2}{n} + \frac{S_Y^2}{\nu}\right)}} \tag{48}$$

can be approximated with a t-distribution with ρ degrees of freedom (DOF), that is

$$\mathscr{WS} \rightarrowtail t_\rho \tag{49}$$

where

$$\rho = \left\lfloor \frac{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{\nu}\right)^2}{\frac{\left(s_X^2/n\right)^2}{n-1} + \frac{\left(s_Y^2/\nu\right)^2}{\nu-1}} \right\rfloor, \tag{50}$$

S^2 is the sample variance as a random variable, s^2 is the variance of a sample, $\overline{X}, \overline{Y}$ are the sample mean as a random variable, and $\lfloor \cdot \rfloor$ is the standard floor function.

The magnitude of the test statistic using Welch-Satterthwaite method is defined as

$$t_{\text{obs}} = \frac{\overline{x} - \overline{y}}{\sqrt{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{\nu}\right)}} \tag{51}$$

where $\overline{x}, \overline{y}$ is the mean of a particular sample. The quantity t_{obs} is the fault indicator. We can then construct the following test:

$$|t_{\text{obs}}| \leq t^{\star} \quad \Rightarrow \quad \text{Accept } H_0 \tag{52}$$

$$|t_{\text{obs}}| > t^{\star} \quad \Rightarrow \quad \text{Accept } H_1 \tag{53}$$

where $t^{\star}$ is such that

$$P\left(t_{\rho} \geq t^{\star}\right) = \frac{\alpha}{2}, \tag{54}$$

where α is the level of significance for the test. To sum up,

i. H_0 is rejected if $|t_{obs}| > t^{\star}$ (the WT is classified as not healthy) and

ii. H_0 is accepted if $|t_{obs}| \leq t^{\star}$ (the WT is classified as healthy).

3.3.3. Multivariate case: testing a multivariate mean vector

In Section 3.3.2, for each principal component $j = 1, \ldots, \ell$, a test for the equality of means is performed. This means that for a single sample of the current structure to diagnose, we obtain ℓ decisions on whether the structure is healthy or not. In the present section, more than one principal component will be considered jointly thus defining a vector. Therefore, a test for the plausibility of a value for a normal population mean vector will be performed.

As in Section 3.3.2, the objective of this work is to determine whether the distribution of the multivariate random samples that are obtained from the WT to be diagnosed (healthy or not) is connected to the distribution of the baseline.

Let us define $s \in \mathbb{N}$ as the number of PCs that are considered at the same time. Before the multivariate HT is applied, the following assumptions must be made:

i. the baseline projection is a multivariate random sample (MRS) of a multivariate random variable (MRV) following a multivariate normal distribution (MVND) with known population mean vector $\boldsymbol{\mu}_{\text{h}} \in \mathbb{R}^s$ and known variance-covariance matrix $\sum \in \mathscr{M}_{s\times s}(\mathbb{R})$ and

ii. the multivariate random sample of the WT to be diagnosed also follows an MVND with unknown multivariate mean vector $\boldsymbol{\mu}_{\text{c}} \in \mathbb{R}^s$ and known variance-covariance matrix $\sum \in \mathscr{M}_{s\times s}(\mathbb{R})$.

In this case, opposite to what we have assumed in Section 3.3.2, both multivariate random variables have the same known variance-covariance matrix.

Similarly as in Section 3.3.2, the question that arises here is whether a given s-dimensional vector $\boldsymbol{\mu}_{\text{c}}$ is a reasonable value for the mean of an MVND $N_s\left(\boldsymbol{\mu}_{\text{h}}, \sum\right)$. This leads to the following test of the hypothesis

$$\begin{aligned} H_0 &: \ \boldsymbol{\mu}_{\text{c}} = \boldsymbol{\mu}_{\text{h}} \ \text{ versus} \\ H_1 &: \ \boldsymbol{\mu}_{\text{c}} \neq \boldsymbol{\mu}_{\text{h}}, \end{aligned} \tag{55}$$

where H_0 is "the MRS of the WT to be diagnosed is distributed as the baseline projection" and H_1 is "the MRS of the WT to be diagnosed is not distributed as the baseline projection." In

other words, if the result of the test is that H_0 is accepted, the current WT is categorized as healthy. Otherwise, if H_0 is rejected in favor of H_1, this would indicate the presence of some faults in the WT.

In this case, the multivariate test is based on Hotelling's T^2 statistic and it is outlined below. When an MRS of size $v \in \mathbb{N}$ is taken from an MVND $N_s(\boldsymbol{\mu}_h, \sum)$, the RV

$$T^2 = v\left(\overline{\mathbf{X}} - \boldsymbol{\mu}_h\right)^T \mathbf{S}^{-1}\left(\overline{\mathbf{X}} - \boldsymbol{\mu}_h\right) \tag{56}$$

is distributed as

$$T^2 \rightarrowtail \frac{(v-1)s}{v-s} \boldsymbol{F}_{s,\, v-s}, \tag{57}$$

where $\boldsymbol{F}_{s,\, v-s}$ denotes an RV with an F-distribution with s and $v-s$ DOF, $\overline{\mathbf{X}}$ is the sample vector mean as a MRV, and $\frac{1}{n}\mathbf{S} \in \mathscr{M}_{s\times s}(\mathbb{R})$ is the estimated variance-covariance matrix of $\overline{\mathbf{X}}$.

The value of the test statistic is defined as

$$t^2_{\text{obs}} = v\left(\overline{\mathbf{x}} - \boldsymbol{\mu}_h\right)^T \mathbf{S}^{-1}\left(\overline{\mathbf{x}} - \boldsymbol{\mu}_h\right), \tag{58}$$

and is the fault indicator. We can then construct the following test:

$$t^2_{\text{obs}} \leq \frac{(v-1)s}{v-s} \boldsymbol{F}_{s,\, v-s}(\alpha) \quad \Rightarrow \quad \text{Accept } H_0, \tag{59}$$

$$t^2_{\text{obs}} > \frac{(v-1)s}{v-s} \boldsymbol{F}_{s,\, v-s}(\alpha) \quad \Rightarrow \quad \text{Accept } H_1, \tag{60}$$

where $\boldsymbol{F}_{s,\, v-s}(\alpha)$ is the upper (100α)th percentile of the $\boldsymbol{F}_{s,\, v-s}$ distribution, that is,

$$\mathbb{P}(\boldsymbol{F}_{s,\, v-s} > \boldsymbol{F}_{s,\, v-s}(\alpha)) = \alpha, \tag{61}$$

where $\mathbb{P}$ is a probability measure and α is the level of significance for the test. To sum up,

i. H_0 is rejected if $t^2_{obs} > \frac{(v-1)s}{v-s} \boldsymbol{F}_{s,\, v-s}(\alpha)$ (the WT is classified as not healthy) and

ii. H_0 is accepted if $t^2_{obs} \leq \frac{(v-1)s}{v-s} \boldsymbol{F}_{s,\, v-s}(\alpha)$ (the WT is classified as healthy).

4. Simulation results

The results of the CM strategies presented in Sections 3.3.2 and 3.3.3 are organized into three subsections. The absolute value of samples that are correctly identified and the absolute number of false alarms and missing faults are included in Section 4.1. Sections 4.2 and 4.3 show the results, not as absolute values but as a percentage. More precisely, the sensitivity and the specificity are both comprised in Section 4.2, including the false-negative (FNR) and the

false-positive rates (FPR). Besides, the true rate of both false negatives and false positives are contained in Section 4.3.

For the validation of the CM strategies presented in Sections 3.3.2 and 3.3.3, 24 samples of $\nu = 50$ elements each have been examined, in accordance with the following organization:

• 8 samples of a faulty WT (one sample for each one of the different fault scenarios described in **Table 3**) and

• 16 samples of a healthy WT.

All samples are acquired with changing wind data sets with turbulence intensity established to 10% and computed with TurbSim [14]. These wind data have the subsequent features:

i. Kaimal turbulence model,

ii. logarithmic profile wind type,

iii. mean speed of 18.2 m/s simulated at hub height, and

iv. a roughness factor of 0.01 m.

Each sample of $\nu = 50$ elements comes from the measures collected during $(\nu \cdot L - 1)\Delta = 312.4875$ seconds. The values for these parameters are listed in **Table 4**.

We present, in Sections 4.1, 4.2, and 4.3, the results when the collected data are projected into:

i. the *first* principal component,

ii. the *second* principal component,

iii. the *third* principal component,

iv. the first *and* the second principal components, jointly,

v. the first *seven* principal components, jointly, and

vi. the first *twelve* principal components, jointly.

In the three univariate cases, (i)–(iii), we use the test for the equality of means, while in the three multivariate cases, (iv)–(vi), we use the test for the plausibility of a value for a normal population. In both cases, the chosen level of significance is $\alpha = 10\%$.

Parameter	Symbol	Magnitude
Number of rows	ν	50
Number of columns	L	500
Sampling time	Δ	0.0125
Number of sensors	N	13

Table 4. The collected measures are arranged in a $\nu \times (N \cdot L)$ matrix $\mathbf{Y}$ as in Eq. (28)

4.1. Types I and II errors

In this section, each of the 24 samples is classified as follows:

i. number of samples from the healthy WT (healthy sample), which were classified by the hypothesis test as "healthy" (accept H_0) [right decision],

ii. faulty sample classified by the test as "faulty" (accept H_1) [right decision],

iii. samples from the faulty WT (faulty sample) classified as "healthy" [wrong decision/missing fault/type II error], and

iv. healthy sample classified as "faulty" [wrong decision/false alarm/type I error].

The results displayed in **Table 6** are disposed according to the scheme in **Table 5**.

4.2. Sensitivity and specificity

As in [20, 30], two more statistical indicators are analyzed to assess the efficiency of the test. On the one hand, the *specificity* of the test is defined as the fraction of samples from the healthy structure, which are correctly classified. On the other hand, the *sensitivity*—or the power of the test—is defined as the fraction of samples from the faulty wind turbine that are correctly classified as such.

	Healthy sample (H_0)	**Faulty sample (H_1)**
Accept H_0	Correct decision	Type II error (missing fault)
Accept H_1	Type I error (false alarm)	Correct decision

Table 5. Scheme for the presentation of the results in **Table 6**

	H_0	H_1		H_0	H_1
Score 1			Scores 1–2		
Accept H_0	16	1	Accept H_0	12	0
Accept H_1	0	7	Accept H_1	4	8
Score 2			Scores 1–7		
Accept H_0	13	7	Accept H_0	13	0
Accept H_1	3	1	Accept H_1	3	8
Score 3			Scores 1–12		
Accept H_0	16	8	Accept H_0	16	0
Accept H_1	0	0	Accept H_1	0	8

Table 6. Categorization of the samples with respect to the presence or absence of a fault and the result of the test considering the first score, the second score, and the third score (left) and scores 1–2 (jointly), scores 1–7 (jointly), and scores 1–12 (jointly) (right), when the size of the samples to diagnose is $\nu = 50$ and the level of significance is $\alpha = 10\%$

The sensitivity and specificity of both the univariate HT and the multivariate case with respect to the 24 samples displayed in **Table 8** are disposed according to the scheme in **Table 7**.

4.3. Reliability of the results

Finally, the true rate of false negatives and the true rate of false positives can be used to assess the performance of the proposed CM strategy. These two measures—closely related to Bayes' theorem [31]—are described in **Table 9**. On the one hand, the true rate of false negatives is the fraction of samples from the faulty WT that have been wrongly identified as healthy. On the other hand, the true rate of false positives is the fraction of sample from the healthy WT that have been wrongly identified as faulty.

The true rate of false negatives and the true rate of false positives of both the univariate HT and the multivariate case displayed in **Table 10** are disposed according to the scheme in **Table 9**.

	Healthy sample (H_0)	**Faulty sample (H_1)**
Accept H_0	Specificity $(1-\alpha)$	False-negative rate (γ)
Accept H_1	False-positive rate (α)	Sensitivity $(1-\gamma)$

Table 7. Relationship between specificity and sensitivity.

	H_0	H_1		H_0	H_1
Score 1			Scores 1–2		
Accept H_0	1.00	0.12	Accept H_0	0.75	0.00
Accept H_1	0.00	0.88	Accept H_1	0.25	1.00
Score 2			Scores 1–7		
Accept H_0	0.81	0.88	Accept H_0	0.81	0.00
Accept H_1	0.19	0.12	Accept H_1	0.19	1.00
Score 3			Scores 1–12		
Accept H_0	1.00	1.00	Accept H_0	1.00	0.00
Accept H_1	0.00	0.00	Accept H_1	0.00	1.00

Table 8. Sensitivity and specificity of the test considering the first score, the second score, and the third score (left) and scores 1–2 (jointly), scores 1–7 (jointly), and scores 1–12 (jointly) (right), when the size of the samples to diagnose is $\nu = 50$ and the level of significance is $\alpha = 10\%$

	Healthy sample (H_0)	**Faulty sample (H_1)**
Accept H_0	$\mathbb{P}(H_0 \vert \text{accept } H_0)$	True rate of false negatives $\mathbb{P}(H_1 \vert \text{accept } H_0)$
Accept H_1	True rate of false positives $\mathbb{P}(H_0 \vert \text{accept } H_1)$	$\mathbb{P}(H_1 \vert \text{accept} H_1)$

Table 9. Relationship between the proportion of false negatives and false positives.

	H_0	H_1		H_0	H_1
Score 1			Scores 1–2		
Accept H_0	0.94	0.06	Accept H_0	1.00	0.00
Accept H_1	0.00	1.00	Accept H_1	0.33	0.67
Score 2			Scores 1–7		
Accept H_0	0.65	0.35	Accept H_0	1.00	0.00
Accept H_1	0.75	0.25	Accept H_1	0.27	0.73
Score 3			Scores 1–12		
Accept H_0	0.67	0.33	Accept H_0	1.00	0.00
Accept H_1	0.00	0.00	Accept H_1	0.00	1.00

Table 10. True rate of false negatives and true rate of false positives of the test considering the first score, the second score, and the third score (left) and scores 1–2 (jointly), scores 1–7 (jointly), and scores 1–12 (jointly) (right), when the size of the samples to diagnose is $\nu = 50$ and the level of significance is $\alpha = 10\%$

5. Concluding remarks

A multifault detection method based on MPCA through uni- and multivariate hypothesis testing has been presented in this chapter. It is noteworthy to mention the obtained performance through the study of eight realistic different faults in different components of the WT, taking into account that the proposed strategy does not need extra sensors but only uses already available data from the WT SCADA system.

The three main conclusions, which show the benefits of the multivariate statistical hypothesis testing in comparison with the univariate case, for WT condition monitoring, are the following:

1. Given a level of significance $\alpha = 10\%$, when the first 12 scores are considered jointly, an accuracy of 100% is obtained, while in all the other studied cases, misclassifications are present.
2. Multivariate analysis leads to average values of 100% for the sensitivity and 85.33% for the specificity, while for the univariate case, the average values are 33.33 and 93.67%, respectively.
3. Multivariate analysis leads to average value of the true rate of false negatives of 0% and the average value of the true rate of false positives of 20%, while for the univariate case, the average values are 24.67 and 25%, respectively.

Acknowledgements

This work has been partially funded by the Spanish Ministry of Economy and Competitiveness through the research projects DPI2014-58427-C2-1-R and DPI2017-82930-C2-1-R, and by the Generalitat de Catalunya through the research project 2017 SGR 388.

Abbreviations

The following abbreviations are used in this chapter:

DOF	degrees of freedom
CM	condition monitoring
FAST	fatigue, aerodynamics, structures, and turbulence
FD	fault detection
FNR	false-negative rate
FPR	false-positive rate
GS	group scaling
HT	hypothesis testing
MCGS	mean-centered group scaling
MPCA	multiway principal component analysis
MRS	multivariate random sample
MRV	multivariate random variable
MVND	multivariate normal distribution
O&M	operation and maintenance
PCA	principal component analysis
RV	random variable
SCADA	supervisory control and data acquisition
SHM	structural health monitoring
WT	wind turbine

Author details

Francesc Pozo* and Yolanda Vidal

*Address all correspondence to: francesc.pozo@upc.edu

Control, Modeling, Identification and Applications (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

References

[1] Lindenberg S, Smith S, O'Dell K, Demeo E, Ram B. 20% wind energy by 2030: Increasing wind energy's contribution to US Electricity Supply; Technical Report; Oak Ridge, TN, USA: U.S. Department of Energy; 2008. DOE/GO-102008-2567

[2] Tchakoua P, Wamkeue R, Ouhrouche M, Slaoui-Hasnaoui F, Tameghe TA, Ekemb G. Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges. Energies. 2014;**7**(4):2595-2630

[3] Qiao W, Lu D. A survey on wind turbine condition monitoring and fault diagnosis. Part I: Components and subsystems. IEEE Transactions on Industrial Electronics. 2015;**62**(10): 6536-6545

[4] Ruiz M, Mujica LE, Alférez S, Acho L, Tutivén C, Vidal Y, Rodellar J, Pozo F. Wind turbine fault detection and classification by means of image texture analysis. Mechanical Systems and Signal Processing. 2018;**107**:149-167

[5] Zaher A, McArthur S. A multi-agent fault detection system for wind turbine defect recognition and diagnosis. In: Power Tech. Lausanne: IEEE; 2007. pp. 22-27

[6] Pozo F, Vidal Y. Damage and fault detection of structures using principal component analysis and hypothesis testing. In: Advances in Principal Component Analysis. Springer. 2018. pp. 137-191

[7] Odgaard P, Johnson K. Wind turbine fault diagnosis and fault tolerant control—An enhanced benchmark challenge. In: Proceedings of the 2013 American Control Conference—ACC; Washington DC. USA; 2013. pp. 1-6

[8] Pozo F, Arruga I, Mujica LE, Ruiz M, Podivilova E. Detection of structural changes through principal component analysis and multivariate statistical inference. Structural Health Monitoring. 2016;**15**(2):127-142

[9] Jonkman, J. NWTC Computer-Aided Engineering Tools (FAST); Last modified 28-October-2013; NWTC Information Portal: Washington, DC, USA, 2013

[10] Vidal Y, Tutiven C, Rodellar J, Acho L. Fault diagnosis and fault-tolerant control of wind turbines via a discrete time controller with a disturbance compensator. Energies. 2015;**8**(5): 4300-4316

[11] Ochs DS, Miller RD, White WN. Simulation of electromechanical interactions of permanent-magnet direct-drive wind turbines using the fast aeroelastic simulator. IEEE Transactions on Sustainable Energy. 2014;**5**(1):2-9

[12] Beltran B, El Hachemi Benbouzid M, Ahmed-Ali T. Second-order sliding mode control of a doubly fed induction generator driven wind turbine. IEEE Transactions on Energy Conversion. 2012;**27**(2):261-269

[13] Jonkman JM, Butterfield S, Musial W, Scott G. Definition of a 5-MW reference wind turbine for offshore system development, Technical Report. Golden, Colorado: National Renewable Energy Laboratory; 2009. NREL/TP-500-38060

[14] Kelley, N.; Jonkman, B. NWTC Computer-Aided Engineering Tools (Turbsim); Last modified 30-May-2013; NWTC Information Portal: Washington, DC, USA, 2013

[15] Ostachowicz W, Kudela P, Krawczuk M, Zak A. Guided Waves in Structures for SHM: The Time-Domain Spectral Element Method. Chichester, UK: John Wiley & Sons, Ltd; 2012

[16] Chai Y, Yang H, Zhao L. Data unfolding PCA modelling and monitoring of multiphase batch processes. IFAC Proceedings Volumes. 2013;**46**(13):569-574

[17] Ruiz M, Villez K, Sin G, Colomer J, Vanrrolleghem P. Influence of scaling and unfolding in PCA based monitoring of nutrient removing batch process. In: Fault Detection, Supervision and Safety of Technical Processes 2006. Amsterdam, The Netherlands: Elsevier; 2007. pp. 114-119

[18] Westerhuis JA, Kourti T, MacGregor JF. Comparing alternative approaches for multivariate statistical analysis of batch process data. Journal of Chemometrics. 1999;**13**(3-4):397-413

[19] Ruiz M, Mujica LE, Sierra J, Pozo F, Rodellar J. Multiway principal component analysis contributions for structural damage localization. Structural Health Monitoring. 2017: 1475921717737971

[20] Pozo F, Vidal Y. Wind turbine fault detection through principal component analysis and statistical hypothesis testing. Energies. 2016;**9**(1):1-20

[21] Pozo F, Vidal Y, Serrahima JM. On real-time fault detection in wind turbines: Sensor selection algorithm and detection time reduction analysis. Energies. 2016;**9**(7):520

[22] Anaya M, Tibaduiza D, Pozo F. A bioinspired methodology based on an artificial immune system for damage detection in structural health monitoring. Shock and Vibration. 2015; **2015**:1-15

[23] Anaya M, Tibaduiza DA, Pozo F. Detection and classification of structural changes using artificial immune systems and fuzzy clustering. International Journal of Bio-Inspired Computation. 2017;**9**(1):35-52

[24] Mujica LE, Rodellar J, Fernández A, Güemes A. Q-statistic and t^2-statistic PCA-based measures for damage assessment in structures. Structural Health Monitoring. 2011;**10**(5): 539-553

[25] Zugasti E, González AG, Anduaga J, Arregui MA, Martínez F. Nullspace and autoregressive damage detection: A comparative study. Smart Materials and Structures. 2012; **21**(8):085010

[26] Häckell MW, Rolfes R, Kane MB, Lynch JP. Three-tier modular structural health monitoring framework using environmental and operational condition clustering for data

normalization: Validation on an operational wind turbine system. Proceedings of the IEEE. 2016;**104**(8):1632-1646

[27] Ng HK, Chen RH, Speyer JL. A vehicle health monitoring system evaluated experimentally on a passenger vehicle. IEEE Transactions on Control Systems Technology. 2006; **14**(5):854-870

[28] Tsiapoki S, Häckell MW, Grießmann T, Rolfes R. Damage and ice detection on wind turbine rotor blades using a three-tier modular structural health monitoring framework. Structural Health Monitoring. 2017:1475921717732730

[29] Ugarte MD, Militino AF, Arnholt A. Probability and Statistics with R. Boca Raton, FL, USA: CRC Press/Taylor & Francis Group; 2008

[30] Pozo F, Vidal Y, Salgado Ó. Wind turbine condition monitoring strategy through multiway PCA and multivariate inference. Energies. 2018;**11**(4):749

[31] DeGroot MH, Schervish MJ. Probability and Statistics. London, UK: Pearson; 2012

IntechOpen

IntechOpen

IntechOpen