Neural Computing and Applications (2024) 36:10279–10296
https://doi.org/10.1007/s00521-024-09604-2

ORIGINAL ARTICLE

Fault detection of wind turbine system based on data-driven methods: a comparative study

Lamiaa M. Elshenawy[1,2] · Ahmed A. Gafar[1] · Hamdi A. Awad[1] · Mahmoud S. AbouOmar[1]

Received: 7 August 2023 / Accepted: 9 February 2024 / Published online: 14 March 2024

Abstract
Fault detection plays a crucial role in ensuring the safety, availability, and reliability of modern industrial processes. This study focuses on data-driven fault detection methods, which have gained significant attention across various industrial sectors due to the rapid development of industrial automation technologies and the availability of extensive datasets. The objectives of this paper are to comprehensively review and present the theoretical foundations of widely used data-driven fault detection approaches. Specifically, these approaches are applied to fault detection in wind turbine systems, with performance evaluation conducted using multiple statistical measures. The data utilized in this study were collected from a simulated benchmark of a wind turbine system. The data-driven methods are tested under the assumption that the wind turbine operates in a steady-state region. Additionally, a comparative study is conducted to identify and discuss the primary challenges associated with the practical application of these methods in real-world scenarios. Simulation results show the effectiveness and efficacy of data-driven approaches concerning the sensitivity and robustness of wind turbine sensor faults as applied in practical industrial environments.

Keywords Fault detection · Statistical process monitoring · Data-driven methods · Wind turbine system.

1 Introduction

Fault detection and process monitoring have been active areas of the research in the control community over the last several decades [1–3]. Due to a large amount of stored data in industrial databases, the data-driven process monitoring approaches have attracted more attention because of their simple design methods and low requirements on the underlying mechanisms [4, 5]. To effectively use the data-driven approaches for process monitoring of large-scale industrial systems with a huge amount of data, it is necessary to perform pre-processing on the collected data to extract information, followed by dimensionality reduction and the selection of the variables that explains a significant part of the observed process variation.

The multivariate statistical process monitoring methods are considered the most popular among data-driven techniques where they directly use the input–output measurements for process monitoring purposes. The basic multivariate statistical process monitoring methods including principal component analysis (PCA) [6, 7], partial least squares (PLS) [8], total projection to latent structures (TPLS) [9], modified partial least squares (MPLS) [10], orthogonal projection to latent structures (O-PLS) [11], modified orthogonal projection to latent structures (MOPLS) [12], expectation-maximization partial robust M-regression (EMPRM) [13], and total principal component regression (TPCR) [14]. All these methods

✉ Ahmed A. Gafar
ahmed.abdallah@el-eng.menofia.edu.eg

Lamiaa M. Elshenawy
lamiaa.elshenawy@el-eng.menofia.edu.eg;
lamiaa.elshenawy@rst.edu.eg

Hamdi A. Awad
hamdy.awad@el-eng.menofia.edu.eg

Mahmoud S. AbouOmar
mahmoud-samy@el-eng.menofia.edu.eg

[1] Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

[2] Faculty of Computers and Artificial Intelligence, AlRyada University for Science and Technology, Sadat City 32897, Egypt

have been successfully applied to many large industrial processes, e.g., chemical plants, water treatment processes, power grids, and cyber-physical systems [15–17]. It should be noted that these monitoring methods are efficiently used with linear, single-mode, and time-invariant processes [18].

Today, wind turbine systems have been widely used to convert wind energy into electricity as a renewable energy source [19–21]. Some wind turbines can produce power up to 4.8 MW [22]. However, wind turbines can be subjected to several faults whether they are sensor faults, actuator faults, and system faults. For a wind turbine, the sensor faults include pitch position sensor faults, rotor speed sensor faults, and generator speed sensor faults. On the other hand, the actuator faults are due to converter coupling faults and pitch system faults. Furthermore, system faults can be found in the wind turbine drive train. It is worth mentioning that unexpected failures of wind turbine parts are the major cause of increased repair costs [23, 24].

Fault detection and diagnosis in different fields have attracted more attention in the literature. The research papers can be classified according to the used fault detection techniques into statistical techniques [25], machine learning techniques [26], deep learning techniques [27–29], vibration analysis techniques [30], and Hybrid techniques [31]. Furthermore, there are two sources of data in these researches that can be either simulated data [32] or SCADA data [33]. Moreover, deep learning methods have been applied in many fields, and among these fields is the field of fault detection and diagnosis. Therefore, its advantages are their ability to improve the accuracy of predictions and decision-making, the ability to learn from unstructured data, such as images, text, and audio, and the ability to automatically learn features from data, eliminating the need for manual feature engineering, but their limitations are their high computational cost and their dependence on the quality and quantity of data used for training.

Recently, data-driven techniques have been applied to a wide range of applications, e.g fault classification [34], signal processing [35], image processing and pattern recognition [36], modeling [37], real-time fatigue life prediction of structures [38], and fuzzy sewage treatment processes [39].

Although significant efforts have been made in the process monitoring of wind turbine systems, to the best knowledge of the authors, there is no systematic comparative studies of data-driven fault detection strategies are available in the literature. Therefore, this is the main motivation behind this study. Due to the nature of this study, the amount of mathematical equations behind the different methodologies have been minimized, as there is a wealth of knowledge in the literature on data-driven methodologies. In addition to that, a detailed description of the benchmark of wind turbines including the models, variables, and faults is presented to be as a reference to other research studies.

The remainder of this study is presented in the following structure. Section 2 gives an overview of the basic data-driven fault detection methods and their variants. Section 3 provides a detailed description of the wind turbine system. The presented fault detection methods are applied to a simulated benchmark of wind turbines, and the comparative results are presented in Sect. 4. Finally, the conclusions are provided in Sect. 5.

2 Overview of data-driven techniques

In order to make it easier to compare different data-driven techniques, we categorize them along a standard fault detection and diagnosis (FDD) work-flow for gathering and analyzing measurements from manufacturing process. Additionally, the basic multivariate statistical process monitoring methods including: The principal component analysis (PCA) model is a statistical procedure in which a set of correlated variables is converted into a set of linearly independent variables that are called principal components. Moreover, PCA is regarded as one of the dimensional reduction techniques in which the number of principal components is always lower than the number of original variables. As well known, the fault detection issue in the scope of the statistics is formulated by two hypotheses, the null hypothesis, which represents the fault-free case, and the alternative hypothesis, which represents the faulty-case [40]. Assume that J_{index} and $J_{\text{th,index}}$ are the fault detection index and its corresponding threshold, respectively. Then, the fault detection logic is as follows:

$$\begin{cases} J_{\text{index}} \leq J_{\text{th,index}} & \text{null hypothesis (fault-free case),} \\ J_{\text{index}} > J_{\text{th,index}} & \text{alternative hypothesis (fault case).} \end{cases} \tag{1}$$

Partial least square (PLS) is a popular input–output technique that is used for modeling, regression, fault detection, and classification purposes. Although the success of the standard PLS model as a monitoring tool for quality-related process problems, it suffers from some problems such as it requires many latent variables that include variations orthogonal to the output Y but are not useful for predicting the output Y [41]. Moreover, residual subspace usually has quite large variations that are not proper to be monitored. [9] proposed the TPLS algorithm to treat the standard PLS model-associated problems. TPLS model works well for characterizing the observed X and is appropriate for monitoring various parts of X. In spite of this, TPLS is not able to totally eliminate impacts. Therefore, as fault amplitude

increases, the amount of undesired variability will rise to a point where the post-processing techniques are useless.

MPLS is another variant of PLS in which the orthogonal decomposition on regression variable space is applied to eliminate the useless variations for output prediction. The related computation cost of the modified strategy eliminates the shortcomings of the traditional PLS algorithm and is significantly simpler than that of the standard technique. Orthogonal projection to latent structures (OPLS) has been used to reduce the number of latent variables while maintaining good prediction accuracy. The OPLS approach is a combination between PLS and a pre-processing technique that is used for removing components that orthogonal to the output Y from the input X [42].

MOPLS methodology is a combination of both OPLS and MPLS algorithms defined in the previous subsections. One of its main advantages is its lower complexity since it reduces the number of latent variables. The EMPRM algorithm has a smaller predicted offset and a more accurate predicted output than the PLS algorithm. The expectation-maximization (EM) phase consists of two steps: the expectation step and the maximization one. In the former, the missing elements are filled with the expected values. In the latter, the expected values are updated using the data in which missing elements are included have been filled in. TPCR model is used to solve the quality-related fault detection issue for linear systems [14]. However, TPCR is inappropriate for automobile applications but is suitable for statistical processes. In the following subsections, the essential mathematical background for the discussed data-driven techniques is presented.

2.1 Principal component analysis

The first step in building a PCA model for the fault detection process is to collect the normal dataset $X = [x_1\, x_2 \ldots x_N]^{\mathrm{T}} \in \Re^{N\times M}$, with M is the process variables and N is the number of measurements. The second step is to normalize the data matrix X with data mean and variance to ensure that all variables have equal weight and to prevent a set of variables from dominating the fault detection process [43]. The third step is to build the PCA model which can be done by calculating the sample covariance matrix $S = \frac{1}{N-1}X^{\mathrm{T}}X$, and then applying the singular value decomposition (SVD) method to calculate the loading vectors which are considered the new coordinates of the data. Generally, PCA model decomposes the original data measurement space into the principal component subspace (PCS), S_p = span(P) and the residual subspace (RS), S_{r} = span$(\widetilde{P})$ [18, 41].

$$X = \widehat{X} + \widetilde{X} = TP^{\mathrm{T}} + \widetilde{T}\widetilde{P} \tag{2}$$

where $T \in \Re^{N\times l}$ and $\widetilde{T} \in \Re^{N\times M-l}$ are the score matrices in the PCS and RS, respectively, and $l \ll M$ is the number of the principal components that can be determined by using different methods [44]. In general, The PCS contains the data that show the most variation, while the RS typically includes data with little variation, which is primarily noise. For an online sample vector, $x \in \Re^{M}$, the principal and residual components are obtained by projecting x on both two subspaces, PCS and RS according to the following:

$$x = \widehat{x} + \widetilde{x} = PP^{\mathrm{T}}x + \widetilde{P}\widetilde{P}^{\mathrm{T}}x = PP^{\mathrm{T}}x + \left(I - PP^{\mathrm{T}}\right)x \tag{3}$$

where $I \in \Re^{M\times M}$ is an identity matrix. Typically, there are two indices are used to monitor normal the variability in PCS and RS, i.e., Hotelling's T^2 and the squared prediction error SPE. Hotelling's T^2 captures the variations in the PCS. On the other hand, the SPE index measures the variations in RS. Therefore, the two monitoring indices with their thresholds for a given significant level [45, 46] can be calculated as Table 1:

2.2 Partial least squares

Generally, the PLS partitions the input space into a principal subspace S_{p} and a residual subspace S_{r}. Given a measurement data matrix (regression variables) $X = [x_1\, x_2 \ldots x_N]^{\mathrm{T}} \in \Re^{\mathrm{NxM}}$ and $Y = [y_1\, y_2 \ldots y_N]^{\mathrm{T}} \in \Re^{N\times q}$ consisting of N samples of q product quality variables (outputs), the PLS projects both X and Y onto a low dimensional subspace defined by a l latent variables. Therefore, both X and Y become:

$$X = \mathrm{TP}^{\mathrm{T}} + \widetilde{X} = \mathrm{XRP}^{\mathrm{T}} + \widetilde{X} \tag{4}$$

$$Y = \mathrm{TQ}^{\mathrm{T}} + \widetilde{Y} = \mathrm{XRQ}^{\mathrm{T}} + \widetilde{Y} \tag{5}$$

where $T = [t_1 \ldots t_l] \epsilon\, \Re^{N\times l}$is the score matrix. $P \in \Re^{M\times l}$ and $Q \in \Re^{q\times l}$ are the loading vectors of X and Y, respectively. $P^{\mathrm{T}}R = R^{\mathrm{T}}P = I_l$, $R \in \Re^{M\times l}$. The PLS is

Table 1 PCA monitoring statistics and thresholds

Statistics	Calculation	Threshold
T^2	$x^{\mathrm{T}}P\,\Lambda^{-1}P^{\mathrm{T}}x$	$\frac{l\left(N^2-1\right)}{N(N-l)}F_\alpha(l, N-l)$
SPE	$\|\widetilde{x}\|^2 = \|(I - PP^{\mathrm{T}})x\|^2$	$\theta_1\left(\frac{c_\alpha h_0\sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right)^{\frac{1}{h_0}}$

where $\Lambda = \mathrm{diag}(\lambda_1\ \lambda_2\ \ldots,\ \lambda_l)$ are the leading eigenvalues of S. $F_\alpha(l, N-l)$ is an F distribution with l and $N-l$ degrees of freedom. c_α is the confidence interval that corresponds to the $(1-\alpha)$ percentile of the normal distribution. $\theta_i = \sum_{j=l+1}^{m}\left(\lambda_j\right)^2$, $i = 1, 2, 3$, $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$

implemented using the nonlinear iterative partial least squares algorithm (NIPALS) to calculate P, T, Q, and R matrices [47, 48]. Similarly to the PCA model, there are two indices used for fault detection extracted from PLS model, i.e., Hotelling's T^2_{PLS} statistic that is used for monitoring the variations in the S_{p} related to the quality output data Y. On the other side, the residual subspace S_{r} is monitored by the SPE_{PLS} statistic which represents the variations unrelated to Y. Therefore, the two monitoring indices with their thresholds can be calculated as Table 2:

2.3 Total projection to latent structures (TPLS)

The main idea of the TPLS model is to decompose the input data space, X, into four parts instead of two parts as in the standard PLS:

$$\begin{aligned} X &= T_y P_y^{\text{T}} + T_{\text{o}} P_{\text{o}}^{\text{T}} + T_{\text{r}} P_{\text{r}}^{\text{T}} + \widetilde{X}_{\text{r}} \\ Y &= \widehat{Y} + \widetilde{Y} = TQ^{\text{T}} + \widetilde{Y} \end{aligned} \tag{6}$$

where T_y is a score matrix that is directly correlated with Y in the original T, and T_{o} is orthogonal to Y in the original T. Furthermore, T_{r} is the main part in $\widetilde{X}$, and $\widetilde{X}_{\text{r}} = \widetilde{X}\left(I - P_{\text{r}} P_{\text{r}}^{\text{T}}\right)$ is the residual part in $\widetilde{X}$ that represents the noise. P_y, P_{o}, P_{r} are the loading vectors that span the corresponding subspaces. It is clear that the TPLS model is able to monitor different parts of X. Therefore, there are four monitoring statistics, i.e., T_y^2, T_{r}^2, T_{o}^2, and SPE_{r} with their thresholds that can be calculated according to the following [49] as Table 3:

More details about the TPLS model are given in the research work of [9]. It should be noted that both T_y^2 and SPE_{r} are used to detect the faults related to Y. On the contrary, T_{o} and T_{r} are used together to detect the faults that are not related to Y. It should be noted that SPE_{r} is more sensitive to incipient faults compared with SPE in the standard PLS [9].

2.4 Modified partial least squares (MPLS)

The following desired relation can be calculated:

Table 2 PLS monitoring statistics and thresholds

Statistics	Calculation	Threshold
T^2_{PLS}	$x^{\text{T}} R \left(\frac{T^{\text{T}} T}{N-1}\right)^{-1} R^{\text{T}} x$	$\frac{l(N^2-1)}{N(N-l)} F_\alpha(l, N-l)$
SPE_{PLS}	$\|\widetilde{x}\|^2 = \left\|\left(I_{\text{MxM}} - \text{PR}^{\text{T}}\right)x\right\|^2$	$g\chi^2_{h,\alpha}$

where $g\chi^2_{h,\alpha}$ is the χ^2 distribution under a significance level α with scaling factors $g = \gamma/2v$ and $h = 2v^2/\gamma$. v and γ are the sample mean and the variance of the SPE statistic [46, 49].

$$Y = \widehat{Y} + \widetilde{Y} = XM + \widetilde{Y} \tag{7}$$

where M is the matrix of the regression coefficient and contains correlation information between X and Y. Moreover, $\widehat{Y}$ and $\widetilde{Y}$ are the subspaces that are correlated and uncorrelated with X, respectively. The coefficient matrix, M can be easily calculated as:

$$M = \left(X^{\text{T}} X\right)^{-1} X^{\text{T}} Y \tag{8}$$

By applying the SVD technique, the MPLS model decomposes the original data measurement space into the principal component subspace (PCS), $S_p = \text{span}(P_M)$ and the residual subspace (RS), $S_{\text{r}} = \text{span}(\widetilde{P}_M)$ [50].

$$MM^{\text{T}} = \begin{bmatrix} P_M & \widetilde{P}_M \end{bmatrix} \begin{bmatrix} \Lambda_M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P_M^{\text{T}} \\ \widetilde{P}_M^{\text{T}} \end{bmatrix} \tag{9}$$

where $P_M \in \Re^{M \times q}$, $\widetilde{P}_M \in \Re^{M \times (M-q)}$ and $\Lambda_M \in \Re^{q \times q}$

Accordingly, the fault detection process can be achieved using two monitoring statistics, T_M^2 and SPE_M. It is worth mentioning that T_M^2 statistic is used for monitoring $\widehat{X}$ which enables detecting faults that are related to Y. On the other hand, the SPE_M statistic is used for monitoring $\widetilde{X}$, thus it can detect faults that are unrelated to Y. Therefore, the monitoring statistics, i.e., T_M^2, SPE_M with their thresholds can be calculated as Table 4

More details about the MPLS algorithm are given in the research work of [50].

2.5 Orthogonal projection to latent structures (OPLS)

The original measurement data matrix X is converted into a filtered matrix X_{opls} which in turn is decomposed by the OPLS algorithm into two subspaces $\widehat{X}_{\text{opls}}$ and $\widetilde{X}_{\text{opls}}$.

$$X_{\text{opls}} = \widehat{X}_{\text{opls}} + \widetilde{X}_{\text{opls}} \tag{10}$$

where $\widehat{X}_{\text{opls}}$ is highly correlated with Y and $\widetilde{X}_{\text{opls}}$ is uncorrelated with Y. Similarly, the aforementioned subspaces can be monitored by two indices, i.e., T^2_{opls} and SPE_{opls}. Therefore, the two monitoring statistics, i.e., T^2_{opls}, SPE_{opls} with their thresholds can be calculated as Table 5:

2.6 Modified orthogonal projection to latent structures (MOPLS)

In this technique, the process data matrix X is decomposed into orthogonal data, $X_\perp$, and filtered data X_{opls}. First, the PCA model is used for Monitoring $X_\perp$ that are uncorrelated with the quality variables. This can be done by performing SVD on $(1/N-1)X_\perp^{\text{T}} X_\perp$.

Table 3 TPLS monitoring statistics and thresholds

Statistics	Calculation	Threshold
T_y^2	$t_y^{\mathrm{T}}\Lambda_y^{-1}t_y$	$\frac{(N+1)}{N}F_\alpha(1,N-1)$
T_{o}^2	$t_{\mathrm{o}}^{\mathrm{T}}\Lambda_{\mathrm{o}}^{-1}t_{\mathrm{o}}$	$\frac{(l-1)(N^2-1)}{N(N-l+1)}\ F_\alpha(l-1,N-l+1)$
T_{r}^2	$t_{\mathrm{r}}^{\mathrm{T}}\Lambda_{\mathrm{r}}^{-1}t_{\mathrm{r}}$	$\frac{l_{\mathrm{r}}(N^2-1)}{N(N-l_{\mathrm{r}})}\ F_\alpha(l_{\mathrm{r}},N-l_{\mathrm{r}})$
$\mathrm{SPE}_{\mathrm{r}}$	$\|\widetilde{x}_{\mathrm{r}}\|^2=\left\|\left(I_{\mathrm{MxM}}-P_{\mathrm{r}}P_{\mathrm{r}}^{\mathrm{T}}\right)\left(I_{\mathrm{MxM}}-PR^{\mathrm{T}}\right)x\right\|^2$	$g_{\mathrm{r}}\chi^2_{h_{\mathrm{r}},\alpha}$

where t_k are the score vectors for an online sample data vector,x, which are calculated as: $t_y = QR^{\mathrm{T}}x$, $t_{\mathrm{o}} = P_{\mathrm{o}}^{\mathrm{T}}(P - P_yQ)R^{\mathrm{T}}x$, and $t_{\mathrm{r}} = P_{\mathrm{r}}^{\mathrm{T}}(I_{\mathrm{MxM}} - PR^{\mathrm{T}})x$. Here, l_y is the number of principal components related to Y; l_{r} is the number of principal components unrelated to Y. $\Lambda_y = \frac{1}{N-1}t_y^{\mathrm{T}}t_y$ is the variance of t_y which is estimated by the training samples. Moreover, Λ_{o} and Λ_{r} are the covariance matrices of t_{o} and t_{r}, respectively. $g_{\mathrm{r}}\chi^2_{h_{\mathrm{r}},\alpha}$ is the critical value of the χ^2 distribution under a significance level α with scaling factors $g_{\mathrm{r}} = \gamma_{\mathrm{r}}/2v_{\mathrm{r}}$ and $h_{\mathrm{r}} = 2v_{\mathrm{r}}^2/\gamma_{\mathrm{r}}$, where v_{r} and γ_{r} are the sample mean and variance of the $\mathrm{SPE}_{\mathrm{r}}$ statistic [46, 49].

Table 4 MPLS monitoring statistics and thresholds

Statistics	Calculation	Threshold
T_M^2	$x^{\mathrm{T}}P_M\left(\frac{P_M^{\mathrm{T}}X^{\mathrm{T}}XP_M}{N-1}\right)^{-1}P_M^{\mathrm{T}}x$	$\frac{q(N^2-1)}{N(N-q)}\ F_\alpha(q,N-q)$
SPE_M	$x^{\mathrm{T}}\widetilde{P}_M\left(\frac{\widetilde{P}_M^{\mathrm{T}}X^{\mathrm{T}}X\widetilde{P}_M}{N-1}\right)^{-1}\widetilde{P}_M^{\mathrm{T}}x$	$\frac{(M-q)(N^2-1)}{N(N-M+q)}\ F_\alpha(M-q,N-M+q)$

$$\left(\frac{1}{N-1}\right)X_\perp^{\mathrm{T}}X_\perp = \begin{bmatrix}\Gamma_{\mathrm{pc}} & \Gamma_{\mathrm{res}}\end{bmatrix}\begin{bmatrix}\Lambda_{\mathrm{pc}} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\Gamma_{\mathrm{pc}}^{\mathrm{T}}\\ \Gamma_{\mathrm{res}}^{\mathrm{T}}\end{bmatrix} \tag{11}$$

where $\Gamma_{\mathrm{pc}} \in \Re^{\mathrm{Mxl_{pc}}}$, $\Gamma_{\mathrm{res}} \in \Re^{\mathrm{Mx(M-l_{pc})}}$, $\Lambda_{\mathrm{pc}} \in \Re^{l_{\mathrm{pc}}xl_{\mathrm{pc}}}$, and l_{pc} is the number of principal components. For a new data sample $x \in \Re^M$, this subspace can be monitored using two statistics that are $T_\perp^2$ and $\mathrm{SPE}_\perp$. Second, the MPLS model is used to monitor X_{opls} which in turn decomposed into $\widehat{X}_{\mathrm{opls}}$ and $\widetilde{X}_{\mathrm{opls}}$. The monitoring results of $\widehat{X}_{\mathrm{opls}}$ and $\widetilde{X}_{\mathrm{opls}}$ reveal quality-related and quality-unrelated faults, respectively. The principal and residual components are obtained by projecting x_{opls} on both two subspaces, PCS and RS according to the following:

$$x_{\mathrm{opls}} = \widehat{x}_{\mathrm{opls}} + \widetilde{x}_{\mathrm{opls}} \tag{12}$$

Table 5 OPLS monitoring statistics and thresholds

Statistics	Calculation	Threshold
T_{opls}^2	$\widehat{x}_{\mathrm{opls}}^{\mathrm{T}}\left(\frac{\widehat{X}_{\mathrm{opls}}^{\mathrm{T}}\widehat{X}_{\mathrm{opls}}}{N-1}\right)^{-1}\widehat{x}_{\mathrm{opls}}$	$\frac{M(N^2-1)}{N(N-M)}\ F_\alpha(M,N-M)$
$\mathrm{SPE}_{\mathrm{opls}}$	$\widetilde{x}_{\mathrm{opls}}^{\mathrm{T}}\widetilde{x}_{\mathrm{opls}}$	$g_{\mathrm{opls}}\chi^2_{h_{\mathrm{opls}},\alpha}$

where $\widehat{x}_{\mathrm{opls}}$ is the projection of x_{opls} in the PCS that are associated with the quality variables. $\widetilde{x}_{\mathrm{opls}}$ is the projection of x_{opls} in the RS that is independent of quality variables. g_{opls} and h_{opls} are scaling factors defined as in [51].

Therefore, there are four monitoring statistics, i.e., $T_\perp^2$, $\mathrm{SPE}_\perp$, $T^2_{\widehat{x}_{\mathrm{opls}}}$, and $T^2_{\widetilde{x}_{\mathrm{opls}}}$ with their thresholds that can be calculated according to the following [12] as Table 6:

where $\theta_i = \sum_{j=l_{\mathrm{pc}}+1}^{m}(\lambda_j)^2$, $i=1,2,3$, $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$. λ_j is the diagonal elements of Λ_{pc}. Here, $t_{\widehat{x}_{\mathrm{opls}}}$ and $t_{\widetilde{x}_{\mathrm{opls}}}$ are considered the score vectors as defined in [12].

2.7 Expectation–maximization partial least squares (EMPRM)

Algebraically, the final regression coefficient vector M_{EM} is obtained from the last PLS step in the final iteration of EMPRM algorithm that is summarized in [52].

Similarly, by applying the singular value decomposition (SVD) method, the EMPRM model decomposes the original data measurement space X and Y into the principal component subspace (PCS), S_p = span(P_{EM}) and the residual subspace (RS), S_{r} = span($\widetilde{P}_{\mathrm{EM}}$).

$$M_{\mathrm{EM}}M_{\mathrm{EM}}^{\mathrm{T}} = \begin{bmatrix}P_{\mathrm{EM}} & \widetilde{P}_{\mathrm{EM}}\end{bmatrix}\begin{bmatrix}\Lambda_{\mathrm{EM}} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}P_{\mathrm{EM}}^{\mathrm{T}}\\ \widetilde{P}_{\mathrm{EM}}^{\mathrm{T}}\end{bmatrix} \tag{13}$$

where $P_{\mathrm{EM}} \in \Re^{M\times q}$, $\widetilde{P}_{\mathrm{EM}} \in \Re^{M\times(M-q)}$ and $\Lambda_{\mathrm{EM}} \in \Re^{q\times q}$

The fault detection process can be achieved using two monitoring statistics, T_{EM}^2 and $\mathrm{SPE}_{\mathrm{EM}}$ for monitoring $\widehat{X}$, $\widetilde{X}$, respectively. It is worth mentioning that T_{EM}^2 statistic which enables detecting faults related to Y. The other statistics $\mathrm{SPE}_{\mathrm{EM}}$ is used for monitoring $\widetilde{X}$ that detects unrelated

Table 6 MOPLS monitoring statistics and thresholds

Statistics	Calculation	Threshold
$T_{\perp}^2$	$x_{\perp}^{\mathrm{T}}\Gamma_{\mathrm{pc}}\Lambda_{\mathrm{pc}}^{-1}\Gamma_{\mathrm{pc}}^{\mathrm{T}}x_{\perp}$	$\frac{l_{\mathrm{pc}}\left(N^2-1\right)}{N\left(N-l_{\mathrm{pc}}\right)}F_{\alpha}(l_{\mathrm{pc}},N-l_{\mathrm{pc}})$
$\mathrm{SPE}_{\perp}$	$x_{\perp}^{\mathrm{T}}\Gamma_{\mathrm{res}}\Gamma_{\mathrm{res}}^{\mathrm{T}}x_{\perp}$	$\theta_1\left(\frac{c_{\alpha}\sqrt{2\theta_2h_0^2}}{\theta_1}+1+\frac{\theta_2h_0(h_0-1)}{\theta_1^2}\right)^{\frac{1}{h_0}}$
$T_{\underset{\sim}{x}_{\mathrm{opls}}}^2$	$t_{\underset{\sim}{x}_{\mathrm{opls}}}^{\mathrm{T}}\left(\frac{T_{\underset{\sim}{x}_{\mathrm{opls}}}^{\mathrm{T}}T_{\underset{\sim}{x}_{\mathrm{opls}}}}{N-1}\right)^{-1}t_{\underset{\sim}{x}_{\mathrm{opls}}}$	$\frac{q(N^2-1)}{N(N-q)}\ F_{\alpha}(q,N-q)$
$T_{\underset{\sim}{x}_{\mathrm{opls}}}^2$	$t_{\underset{\sim}{x}_{\mathrm{opls}}}^{\mathrm{T}}\left(\frac{T_{\underset{\sim}{x}_{\mathrm{opls}}}^{\mathrm{T}}T_{\underset{\sim}{x}_{\mathrm{opls}}}}{N-1}\right)^{-1}t_{\underset{\sim}{x}_{\mathrm{opls}}}$	$\frac{(M-q)(N^2-1)}{N(N-M+q)}\ F_{\alpha}(M-q,N-M+q)$

faults to Y. Therefore, the two monitoring statistics with their thresholds can be calculated as Table 7.

More details about the EMPRM model are given in the research work of [52].

2.8 Total principal component regression (TPCR)

The basic idea of this algorithm is that the original process matrix X is first projected onto the score matrix T by PCA. T_{c} is a score matrix that is extracted from T by the least squares regression between T and Y. After that, X_{c} is reconstructed from X by T_{c} such that X_{c} is highly correlated with Y, leaving the remaining part X_{u} is uncorrelated with Y. To build the TPCR model, the PCA is performed on $\widehat{Y}$ to get its score matrix T_{c} and load matrix Q_{c} such that:

$$T_{\mathrm{c}}=\widehat{Y}Q_{\mathrm{c}}=\ TQ^{\mathrm{T}}Q_{\mathrm{c}} \tag{14}$$

where $\widehat{Y}$ is the online prediction of Y, $Q^{\mathrm{T}}=(T^{\mathrm{T}}T)^{-1}T^{\mathrm{T}}Y$, and T is re-projected to T_{c} by $Q^{\mathrm{T}}Q_{\mathrm{c}}$. Next, we reconstruct X_{c} from T_{c} by the following steps:

$$P_{\mathrm{c}}^{\mathrm{T}}=\left(T_{\mathrm{c}}^{\mathrm{T}}T_{\mathrm{c}}\right)^{-1}T_{\mathrm{c}}^{\mathrm{T}}\widehat{X} \tag{15}$$

Thus:

$$X_{\mathrm{c}}=T_{\mathrm{c}}P_{\mathrm{c}}^{\mathrm{T}}=T_{\mathrm{c}}\left(T_{\mathrm{c}}^{\mathrm{T}}T_{\mathrm{c}}\right)^{-1}T_{\mathrm{c}}^{\mathrm{T}}TP^{\mathrm{T}} \tag{16}$$

And

$$X_{\mathrm{u}}=X-X_{\mathrm{c}} \tag{17}$$

By performing PCA on X_{u}, we get its score matrix T_{u} and load matrix P_{u} with only the eigenvectors corresponding to zero and extremely small eigenvalues. For each online sample x, the correlated score vector $t_{\mathrm{c}}^{\mathrm{T}}$ is calculated as follows:

$$t_{\mathrm{c}}^{\mathrm{T}}=t^{\mathrm{T}}Q^{\mathrm{T}}Q_{\mathrm{c}}=x^{\mathrm{T}}PQ^{\mathrm{T}}Q_{\mathrm{c}} \tag{18}$$

Similarly, the score vector of the uncorrelated part is calculated as follows:

$$t_{\mathrm{u}}^{\mathrm{T}}=x_{\mathrm{u}}^{\mathrm{T}}P_{\mathrm{u}}=\left(x^{\mathrm{T}}-t_{\mathrm{c}}^{\mathrm{T}}P_{\mathrm{c}}^{\mathrm{T}}\right)P_{\mathrm{u}}=\left(x^{\mathrm{T}}-x^{\mathrm{T}}PQ^{\mathrm{T}}Q_{\mathrm{c}}P_{\mathrm{c}}^{\mathrm{T}}\right)P_{\mathrm{u}} \tag{19}$$

Therefore, the two monitoring statistics, i.e., T_{c}^2, T_{u}^2 of X_{c} and X_{u}, respectively, with their thresholds can be calculated as Table 8.

To summarize the main components of the fault detection approaches given in this paper, the offline and online phases are described in the following flowchart, see Fig. 1. The main terminologies and abbreviations of the algorithms are summarized in table 15.

3 Wind turbine system

It is well known that wind turbines convert the kinetic energy of wind into electrical energy. The main components are the tower, the rotor and hub (including three blades), the nacelle, and the generator as shown in Fig. 2. Wind turbines may generate several GW yearly that are used around the world. Therefore, they have attracted many investments in the field of renewable energy sources. It is worth mentioning that the parts of wind turbines may have malfunctions that should be detected using fault detection schemes. As mentioned in the introduction section, there are two sources of the wind turbine systems data including the SCADA and simulated data. In the paper, we collect the data from a benchmark of a wind turbine system that is popularly used for evaluating the controller and process monitoring schemes [53–55]. The wind turbine model consists of a horizontal-axis three-blade turbine with full converter coupling and is connected to the generator via a gearbox, see Fig. 2. The conversion from wind energy to mechanical energy can be controlled using the aerodynamics of the wind turbine. Using the generator coupled to a converter coupling, mechanical energy is converted to electrical energy. The drive train (Gearbox) between the rotor and the generator increases the generator's speed.

Table 7 EMPRM monitoring statistics and thresholds

Statistics	Calculation	Threshold
T^2_{EM}	$x^{\mathrm{T}} P_{\mathrm{EM}} \left(\frac{P^{\mathrm{T}}_{\mathrm{EM}} X^{\mathrm{T}} X P_{\mathrm{EM}}}{N-1} \right)^{-1} P^{\mathrm{T}}_{\mathrm{EM}} x$	$\frac{q(N^2-1)}{N(N-q)} \; F_{\alpha}(q, N-q)$
$\mathrm{SPE}_{\mathrm{EM}}$	$x^{\mathrm{T}} \widetilde{P}_{\mathrm{EM}} \left(\frac{\widetilde{P}^{\mathrm{T}}_{\mathrm{EM}} X^{\mathrm{T}} X \widetilde{P}_{\mathrm{EM}}}{N-1} \right)^{-1} \widetilde{P}^{\mathrm{T}}_{\mathrm{EM}} x$	$\frac{(M-q)(N^2-1)}{N(N-M+q)} \; F_{\alpha}(M-q, N-M+q)$

Table 9 summarizes the main signals exchanged among the subsystems.

The wind turbine model consists of different parts, i.e., controller, drive train, generator/converter, wind, blade and pitch subsystems as shown in Fig. 3. The main variables of the wind turbine system are summarized in table 16.

Moreover, because the wind turbine system is a multi-mode system, the controller has to work in four operating zones, which are determined according to the average wind speed within a certain time window as depicted in Fig. 4 [56]. It is clearly shown from this figure that the turbine is at standstill at Region I and Region II represents the power optimization with partial load. Whenever, Region III and Region IV describe the constant power generation and high wind speed, respectively. In this paper, the wind turbine is assumed to work at Region III which represents the steady-state operation of the wind turbine system. The turbine is controlled at wind speed between 0 and v_{rated} m/s in order to achieve the optimal power generation. The speed ratio is given by:

$$\mu = \frac{\omega_{\mathrm{r}}.\mathrm{r}}{\upsilon_{\mathrm{w}}} \tag{20}$$

where r is the radius of the blades.

The benchmark model implemented in Simulink is available at the URL address: https://www.mathworks.com/matlabcentral/fileexchange/35130-award-winning-fdi-solution-in-wind-turbines.

The utilization of wind energy for the wind power generation system is a subject of research interest and in the recent years, the focus is on the cost-effective use of wind energy with the aim of providing electricity of high quality and reliability. In the past twenty years, wind turbine sizes have evolved from 20-kW to 5-megawatts, while even more powerful wind turbines are being developed. Therefore, in order to prevent major component failures, fault detection algorithms enable early alarms of mechanical and electrical faults. Side effects on other components can be significantly reduced.

Furthermore, many faults can be detected even when the faulty component is still working. Thus, necessary repairs can be planned in time and do not have to be carried out immediately. Therefore, this is important because wind power generation system is inaccessible because they are located on extremely high towers, typically 20 m or higher [57]. This is also particularly important for wind turbine installations, where adverse weather conditions (storms, high tides, etc.) can prevent repair action for several weeks.

Therefore, maintenance costs and downtime of wind power generation system can be significantly reduced [58]. Therefore, due to the importance of fault detection and diagnosis in wind power generation system (blades, drive train, and generator), this paper is presented to be as a reference to other research studies as shown in Fig. 5 [25, 57]. For illustration, this figure shows the main wind turbine components that are concerned by the above benchmark model.

In summary, the sensors are mainly used for the blade load reduction based on the individual pitch control strategy, especially in offshore wind turbines. The lifetime of the sensors in wind turbine systems is usually not very long. There are several factors that lead to higher failure rates. The strain in the blades can be very high, which affects the gauges themselves and the bonding. Harsh environmental factors such as lightning, salt spray, moisture, corrosion can directly affect the bonding and wiring of the sensors. Maintenance personnel can easily damage the sensors [25].

Table 8 TPCR monitoring statistics and thresholds

Statistics	Calculation	Threshold
T^2_{c}	$t^{\mathrm{T}}_{\mathrm{c}} \left(\frac{T_{\mathrm{c}}{}^{\mathrm{T}} T_{\mathrm{c}}}{N-1} \right)^{-1} t_{\mathrm{c}}$	$\frac{(N+1)}{N} F_{\alpha}(1,\ N-1)$
T^2_{u}	$t^{\mathrm{T}}_{\mathrm{u}} \left(\frac{T_{\mathrm{u}}{}^{\mathrm{T}} T_{\mathrm{u}}}{N-1} \right)^{-1} t_{\mathrm{u}}$	$\frac{(l-1)(N^2-1)}{N(N-l+1)} \; F_{\alpha}(l-1,\ N-l+1)$

3.1 Wind model

The combined wind model is [59]:

$$\upsilon_{\mathrm{w}}(t) = \upsilon_{\mathrm{m}}(t) + \upsilon_{\mathrm{s}}(t) + \upsilon_{\mathrm{ws}}(t) + \upsilon_{\mathrm{ts}}(t) \tag{21}$$

where $\upsilon_{\mathrm{m}}(t)$ is the mean wind, $\upsilon_{\mathrm{s}}(t)$ is the stochastic part, $\upsilon_{\mathrm{ws}}(t)$ is the wind shear, and $\upsilon_{\mathrm{ts}}(t)$ is the tower shadow. The wind shear model is:

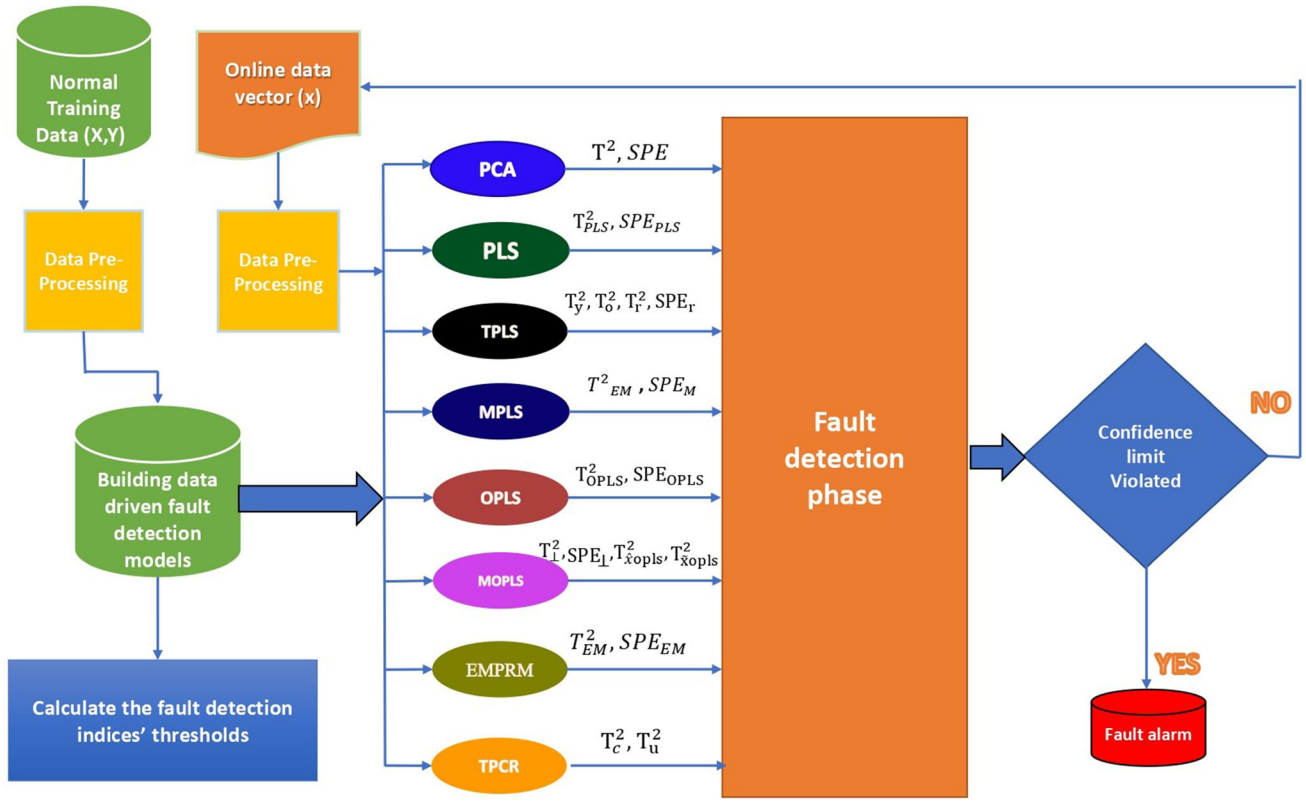

Fig. 1 Flowchart of the fault detection schemes

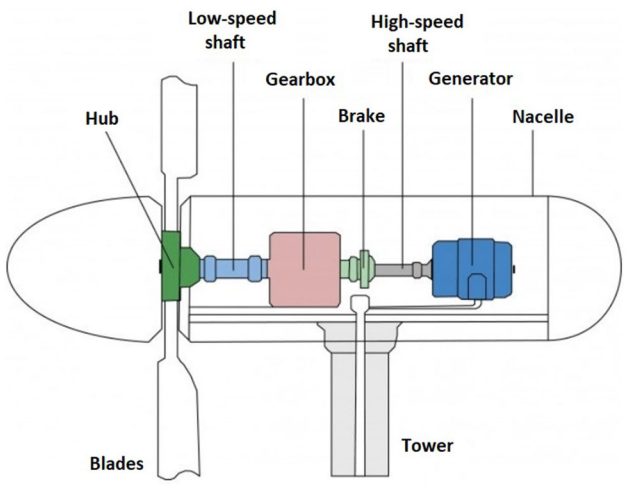

Fig. 2 Wind turbine model

$$\upsilon_{\mathrm{ws},i}(t) = \frac{2\upsilon_{\mathrm{m}}(t)}{3r^2}\left(\frac{r^3\sigma}{3H}\Xi + \frac{r^4}{4}\sigma\frac{\sigma-1}{2\,H^2}\Xi^2\right) + \frac{2\upsilon_{\mathrm{m}}(t)}{3r^2}\left(\frac{r^5}{5}\frac{(\sigma^2-\sigma)(\sigma-2)}{6\,H^3}\Xi^3\right) \tag{22}$$

where $\Xi = \cos(\vartheta_{r*}(t))$, and ϑ_{r*} is the angular position of the three blades, $\vartheta_{r1}(t) = \vartheta_{\mathrm{r}}(t)$, $\vartheta_{r2}(t) = \vartheta_{\mathrm{r}}(t) + (2/3)\pi$, and $\vartheta_{r3} = \vartheta_{\mathrm{r}}(t) + (4/3)\pi$. σ and H are two aerodynamic parameters. Furthermore, the tower shadow model is:

$$\upsilon_{\mathrm{ts},i}(t) = \frac{m\overline{\vartheta}_{\mathrm{r},i}(t)}{3r^2}(\psi + \upsilon) \tag{23}$$

where

$$\psi = 2a^2\frac{r^2 - r_0^2}{\left(r^2 + r_0^2\right)\sin\left(\overline{\vartheta}_{\mathrm{r},i}(t)\right)^2 + k^2)} \tag{24}$$

$$\upsilon = 2a^2k^2\frac{(r_0^2 - r^2)(r_0^2\sin(\overline{\vartheta}_{\mathrm{r},i}(t))^2 + k^2)}{r^2\sin\left(\overline{\vartheta}_{\mathrm{r},i}(t)\right)^2 + k^2} \tag{25}$$

$$m = 1 + \frac{\sigma(\sigma-1)\ r_0^2}{8\ H^2} \tag{26}$$

$$(\overline{\vartheta}_{\mathrm{r},i}(t) = \vartheta_{\mathrm{r}}(t) + \frac{(i-1)2\pi}{3} - \mathrm{floor}\left(\frac{\vartheta_{\mathrm{r}}(t) + \frac{(i-1)2\pi}{3}}{2\pi}\right)2\pi \tag{27}$$

r_0 is the blade hub radius and k is an aerodynamic parameter.

Table 9 Variables of wind turbine system

Variable	Notation	Description
v_1	v_{hub}	Wind speed at hub height
v_2	$\omega_{\text{r},m1}$	Measured rotational speed of the rotor 1
v_3	$\omega_{\text{r},m2}$	Measured rotational speed of the rotor 2
v_4	$\omega_{g,m1}$	Measured rotational speed of the generator 1
v_5	$\omega_{g,m2}$	Measured rotational speed of the generator 2
v_6	τ_{g}	Generator torque
v_7	P_{g}	Power produced by the generator
v_8	$\beta_{1,m1}$	Pitch 1 position for blade 1
v_9	$\beta_{1,m2}$	Pitch 2 position for blade 1
v_{10}	$\beta_{2,m1}$	Pitch 1 position for blade 2
v_{11}	$\beta_{2,m2}$	Pitch 2 position for blade 2
v_{12}	$\beta_{3,m1}$	Pitch 1 position for blade 3
v_{13}	$\beta_{3,m2}$	Pitch 2 position for blade 3
v_{14}	$\tau_{g,r}$	Torque reference to the generator
v_{15}	ω_{r}	Rotational speed of the rotor of the drive train
v_{16}	ω_{g}	Rotational speed of the generator of the drive train

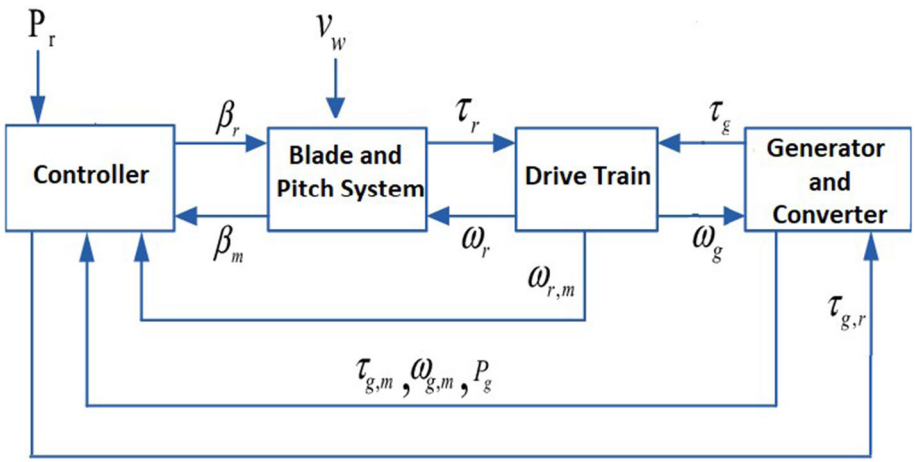

Fig. 3 Overview of the benchmark model

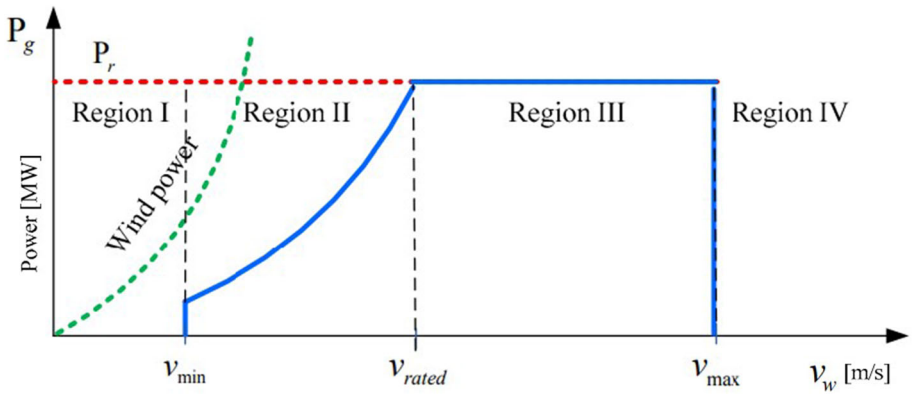

Fig. 4 Regions of power operation

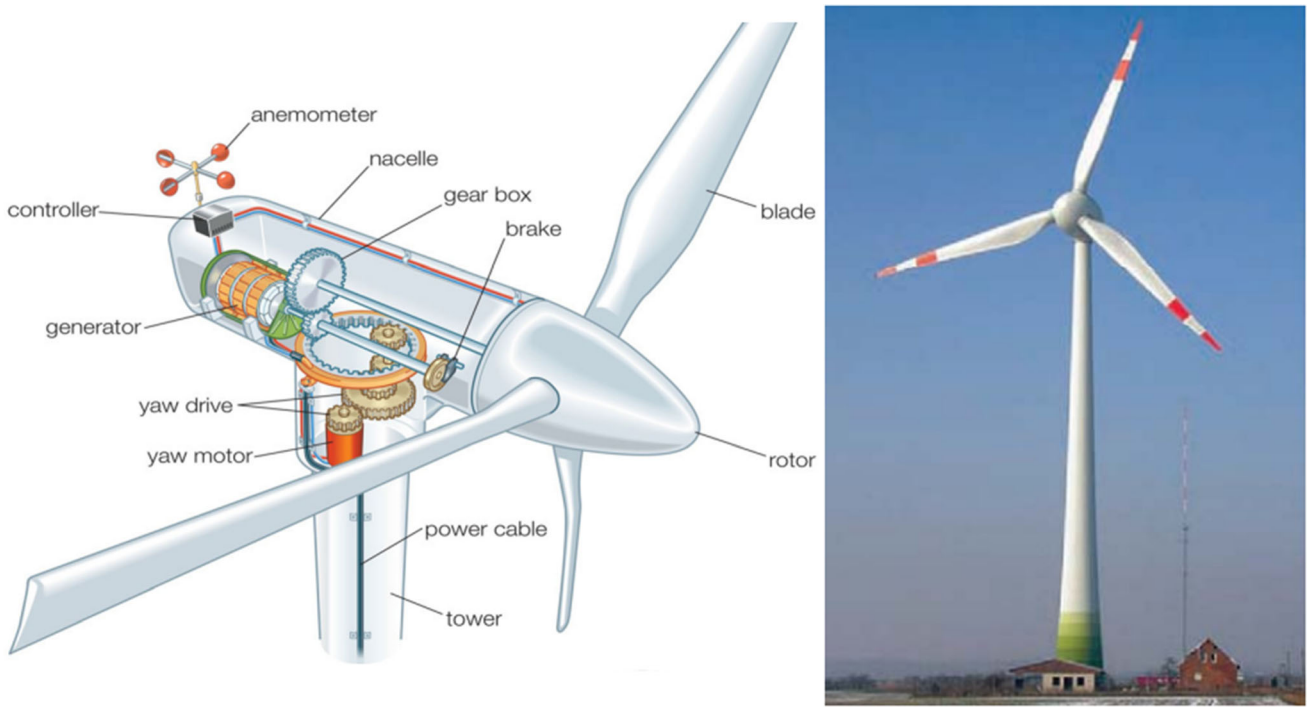

Fig. 5 Wind turbine configuration

3.2 Blade and pitch model

This model combines the aerodynamic and the pitch models. The aerodynamic subsystem describes the forces that an air flow develops on a wind turbine and transforms the three-dimensional wind field into concentrated forces. As can be seen in the block diagram of Fig. 3, the inputs to this subsystem are the wind speed v_{w}, the pitch angle β, and the rotation speed of the rotor ω_{r}. The output of this subsystem is the aerodynamic torque τ_{r}. The torque equation of this subsystem is [56]:

$$\tau_{\mathrm{r}}(t) = \frac{\rho \pi r^3 C_q(\gamma(t), \beta(t)) \upsilon_{\mathrm{w}}(t)^2}{2} \tag{28}$$

where $C_q(\gamma(t), \beta(t))$ is a map of the torque coefficients that represent a function of the speed ratio with the lead angle and ρ is the air density. A simple representative is used to model the three blades to obtain their pitch angle value. This assumption suppose that the torque of each blade is one-third of the torque that is given by the three blades. Therefore, the torque equation of this subsystem is:

$$\tau_{\mathrm{r}}(t) = \sum_{i=1}^{3} \frac{\rho \pi r^3 C_q(\gamma(t), \beta_i(t)) \upsilon_{w,i}(t)^2}{6} \tag{29}$$

where β_i is the pitch position.

On the other hand, the pitch subsystem is an actuator that generally rotates all the blades or a part of them. Therefore, the model of the hydraulic pitch system is considered as a closed-loop transfer function between the measured pitch angle (β_{m}) and its reference (β_{r}). In principle, this subsystem can be modeled by a second-order transfer function [60] as:

$$\frac{\beta_{\mathrm{m}}(s)}{\beta_{\mathrm{r}}(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{30}$$

where β_{r} is the input to the closed-loop transfer function; β_{m} is the output of the transfer function; $\zeta$is the damping factor, and ω_n is the natural frequency.

3.3 Drive train model

A two-mass model of the drive train is used in this benchmark model. The torque from the rotor is transferred to the generator through the drive train. From the low-speed rotor side to the high-speed generator side, the rotational speed is increased using a gearbox. A two-mass drive train model can be represented by:

$$J_{\mathrm{r}}\dot{\omega}_{\mathrm{r}}(t) = \tau_{\mathrm{r}}(t) - K_{\mathrm{dt}}\vartheta_{\Delta}(t) - (B_{\mathrm{dt}} + B_{\mathrm{r}})\omega_{\mathrm{r}}(t) + \frac{B_{\mathrm{dt}}}{N_{\mathrm{g}}}\omega_{\mathrm{g}}(t) \tag{31}$$

$$J_{\mathrm{g}}\dot{\omega}_{\mathrm{g}}(t) = \frac{\eta_{\mathrm{dt}}K_{\mathrm{dt}}}{N_{\mathrm{g}}}\vartheta_{\Delta}(t) + \frac{\eta_{\mathrm{dt}}B_{\mathrm{dt}}}{N_{\mathrm{g}}}\omega_{\mathrm{r}}(t) - \left(\frac{\eta_{\mathrm{dt}}B_{\mathrm{dt}}}{N_{\mathrm{g}}^2} + B_{\mathrm{g}}\right)\omega_{\mathrm{g}}(t) - \tau_{\mathrm{g}}(t) \tag{32}$$

$$\dot{\vartheta}_{\Delta}(t) = \omega_{\mathrm{r}}(t) - \frac{1}{N_{\mathrm{g}}}\omega_{\mathrm{g}}(t) \tag{33}$$

Table 10 The benchmark model parameters

Subsystem	Parameter name	Value and unit	Description
Wind model	σ	0.1	Aerodynamic parameter
	H	81 m	Blade hub height
	r_o	1.5 m	Radius of blade hub
Blade and pitch model	r	57.5 m	Radius of blades
	ρ	1.225 kg/m^3	Air density
	ζ	0.6	Damping factor
	ω_n	11.11 rad/s	Natural frequency
Drive train model	B_{dt}	775.49 Nms/rad	Torsion damping coefficient of drive train
	B_r	7.11 Nms/rad	Viscous friction of low-speed shaft
	B_g	45.6 Nms/rad	Viscous friction of high-speed shaft
	N_g	95	The gear ratio
	K_{dt}	2.7×10^9 Nm/rad	Torsion stiffness of drive train
	η_{dt}	0.97	Efficiency of drive train
	η_{dt2}	0.92	Lower drive train efficiency
	J_g	390 kg m^2	Moment of inertia of high-speed shaft
	J_r	55×10^6kg m^2	Moment of inertia of low-speed shaft
Generator and converter model	σ_{gc}	50 rad/s	Generator and converter model parameter
	η_{gc}	0.98	Generator and converter efficiency
Controller model	K_{opt}	1.2171	Optimal value of K
	K_i	1	Controller gains of PI
	K_p	4	Controller gains of PI
	ω_{nom}	162 rad/s	Nominal generator speed
	ω_Δ	15 rad/s	Small offset
Sensor model	v_ω	1.5 m/s	Mean value of wind speed
	S_ω	0.5 m/s	Variance of wind speed
	v_{wr}	0 rad/s	Mean value of rotor speed
	S_{wr}	0.025 rad/s	Variance of rotor speed
	v_{wg}	0 rad/s	Mean value of generator speed
	S_{wg}	0.05 rad/s	Variance of generator speed
	v_{τ_g}	0 Nm	Mean value of generator torque
	S_{τ_g}	90 Nm	Variance of generator torque
	v_{Pg}	0 W	Mean value of generator power
	S_{Pg}	1000 W	Variance of generator power
	v_β	0°	Mean value of pitch angle
	S_β	0.2°	Variance of pitch angle

where J_r, J_g are the moment of inertia of the low-speed and high-speed shafts, respectively, K_{dt} is the torsion stiffness, η_{dt} is the efficiency, B_{dt} is the torsion damping coefficient, B_r, B_g are the viscous friction of the low-speed and high-speed shafts, respectively, N_g is the gear ratio, and $\vartheta_\Delta(t)$ is the torsion angle.

3.4 Generator and converter model

The frequency range used in this model is much slower than the electrical system in the wind turbine system. A first-order transfer function can be used to represent the generator and converter dynamics at the wind turbine system level.

$$\frac{\tau_g(s)}{\tau_{g,r}(s)} = \frac{\sigma_{gc}}{s + \sigma_{gc}} \tag{34}$$

Table 11 Sensor faults in the wind turbine system

Fault number	Related variables	Fault type
F1, F3, F5, F7, F9, F11	$\beta_{1,m1}$, $\beta_{1,m2}$, $\beta_{2,m1}$, $\beta_{2,m2}$,$\beta_{3,m1}$, $\beta_{3,m2}$	Fixed value
F2, F4, F6, F8, F10, F12	$\beta_{1,m1}$, $\beta_{1,m2}$, $\beta_{2,m1}$, $\beta_{2,m2}$,$\beta_{3,m1}$, $\beta_{3,m2}$	Gain factor
F13, F15	$\omega_{r,m1}$, $\omega_{r,m2}$	Fixed value
F14, F16	$\omega_{r,m1}$, $\omega_{r,m2}$	Gain factor
F17, F19	$\omega_{g,m1}$, $\omega_{g,m2}$	Fixed value
F18, F20	$\omega_{g,m1}$, $\omega_{g,m2}$	Gain factor

where σ_{gc} is the model parameter of generator and converter.

The generator output power is:

$$P_g(t) = \eta_g \omega_g(t) \tau_g(t) \tag{35}$$

where η_g is the generator efficiency. Moreover, the controller is implemented in discrete-time form, with a sampling frequency of 100 Hz. The controller changes from mode 1 to mode 2 in the following case:

$$P_g(k) \geq P_r(k) \vee \omega_g(k) \geq \omega_{nom} \tag{36}$$

Here, ω_{nom} is the nominal speed of the generator. The controller changes from mode 2 to mode 1 in the following case:

$$\omega_g(k) < \omega_{nom} - \omega_\Delta \tag{37}$$

ω_Δ is a small offset subtracted from the nominal speed of the generator. At mode 1, the optimal value of γ is represented by γ_{opt}. This optimal value is realized when the pitch reference to zero ($\beta_r(k) = 0$), and the reference torque to the converter $\tau_{g,r}$ is:

$$\tau_{g,r}(k) = K_{opt} \left(\frac{\omega_g(k)}{N_g} \right)^2 \tag{38}$$

$$K_{opt} = \frac{1}{2} \rho * A r^3 \frac{C_{Pmax}}{\gamma_{opt}^3} \tag{39}$$

$A = \pi r^2$ is the area swept by the wind turbine blades, and K_{opt} is the optimal value of k, C_{Pmax} is the maximum value of the power coefficient. On the other hand, at mode 2, the major control actions are handled by the pitch system using a PI controller trying to keep $\omega_g(k)$ at ω_{nom}

$$\beta_r(k) = \beta_r(k-1) + K_P e(k) + (K_i T_s - K_P) e(k-1) \tag{40}$$

$e(k) = \omega_g(k) - \omega_{nom}$, and the controller gains are K_P and K_i. In this case, the converter reference is:

$$\tau_{g,r}(k) = \frac{P_r(k)}{\eta_{gc} \omega_g(k)} \tag{41}$$

where η_{gc} is the efficiency of the generator and converter subsystems. Moreover, a stochastic noise component is added to the actual variable value to model each sensor. The parameters used in the benchmark model are listed in Table 10.

It is worth mentioning that many fault types are occurring in wind turbine systems including sensor faults, actuator faults, and process faults. This paper is dedicated to studying the efficiency of the presented process monitoring methodologies for sensor faults detection. The sensor faults can be in the pitch position measurements, e.g., $\beta_{1,m1}$ $\beta_{1,m2}$ $\beta_{2,m1}$ $\beta_{2,m2}$ $\beta_{3,m1}$,$\beta_{3,m2}$; in the rotor speed measurements,$\omega_{r,m1}$ and$\omega_{r,m2}$; in the generator speed measurements $\omega_{g,m1}$ and $\omega_{g,m2}$. The details of these faults are listed in Table 11.

4 Simulation results and discussion

In this simulation study for a wind turbine system, the measurement data are collected from 16 variables at normal operation. The dataset includes 10^5 samples, and the fault scenarios are presented at a sample time $k = 5 \times 10^4$ to the end of the simulation. Twenty different fault scenarios are given in Table 11 that occur in the wind turbine system. The rotational speed of the rotor in the drive train is chosen as the quality variable to construct Y matrix (output variable) and the other 15 process variables are chosen as input variables. Furthermore, the different fault scenarios are monitored using the presented data-driven fault detection approaches in this paper and evaluated by using two indices, i.e., fault detection rate (FDR) and false alarm rate (FAR) [9, 43, 61].

Table 12 Design parameters

PCA	PLS	TPLS	MPLS	OPLS	MOPLS	EMPRM
l = 6	LV = 12	LV = 6	LV = 12	LV = 1	LV = 1	LV = 12

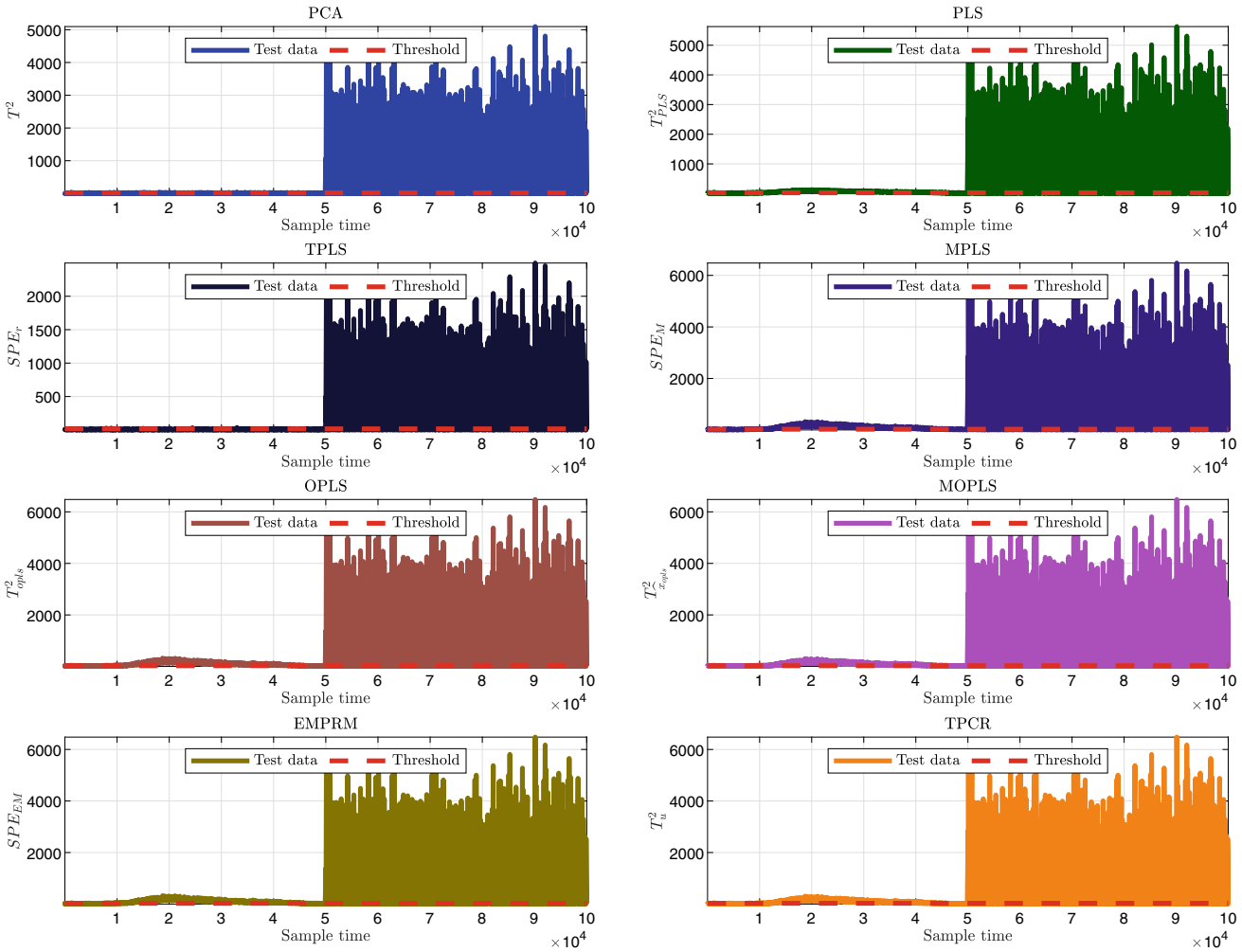

Fig. 6 Process monitoring using different algorithms in case of (F4)

$$\text{FDR} = \frac{\text{No. of samples } (J > J_{\text{th}} | f \neq 0)}{\text{total samples } (f \neq 0)} \times 100\%$$

$$\text{FAR} = \frac{\text{No. of samples } (J > J_{\text{th}} | f = 0)}{\text{total samples } (f = 0)} \times 100\%$$

The number of principal components (PCs) and the number of LVs are selected by using the cumulative percent variance method (CPV) and the cross-validation method, respectively [9, 44, 62]. The design parameters are summarized in Table 12.

Two fault scenarios are given in details, i.e., F14 and F19. The first fault occurs as a gain factor in the pitch position sensor. Figure 6 shows the fault detection charts of the described monitoring models in this paper. It is clearly shown that the monitoring models achieved FDR with average about 80%. Furthermore, PCA, TPLS recorded the lowest FAR, but the other methods have the highest degree of FAR compared with the PCA and TPLS.

Besides, F19 represents a fixed value in the sensor that measures the generator speed. The fault detection results of the presented methods of this fault are shown in Fig. 7. It is clearly shown that all the tested algorithms showed satisfactory fault detection performance with FDRs of approximately 99.99% except for PCA algorithm that failed to detect fault successfully as shown in Fig. 7a. As well as, all monitoring methodologies exhibit acceptable FARs.

To evaluate the fault detection approaches according to all possible sensor faults, Tables 13 and 14 give the FDRs and FARS. As shown in Table 13, most fault detection methods successfully detected all faults with high FDRs. On the other side, the PCA and TPLS could not detect the faults (F13, F14, F15, F16, F19) and F9, respectively, in which the boldface denotes the lowest FDRs.

It is clearly shown from Table 14 that the FARs of both PCA and TPLS had the lowest FAR rates for all fault scenarios, which means they are the most robust fault detection methods. As well as other algorithms in some fault scenarios have high FARs which are denoted by the boldface.

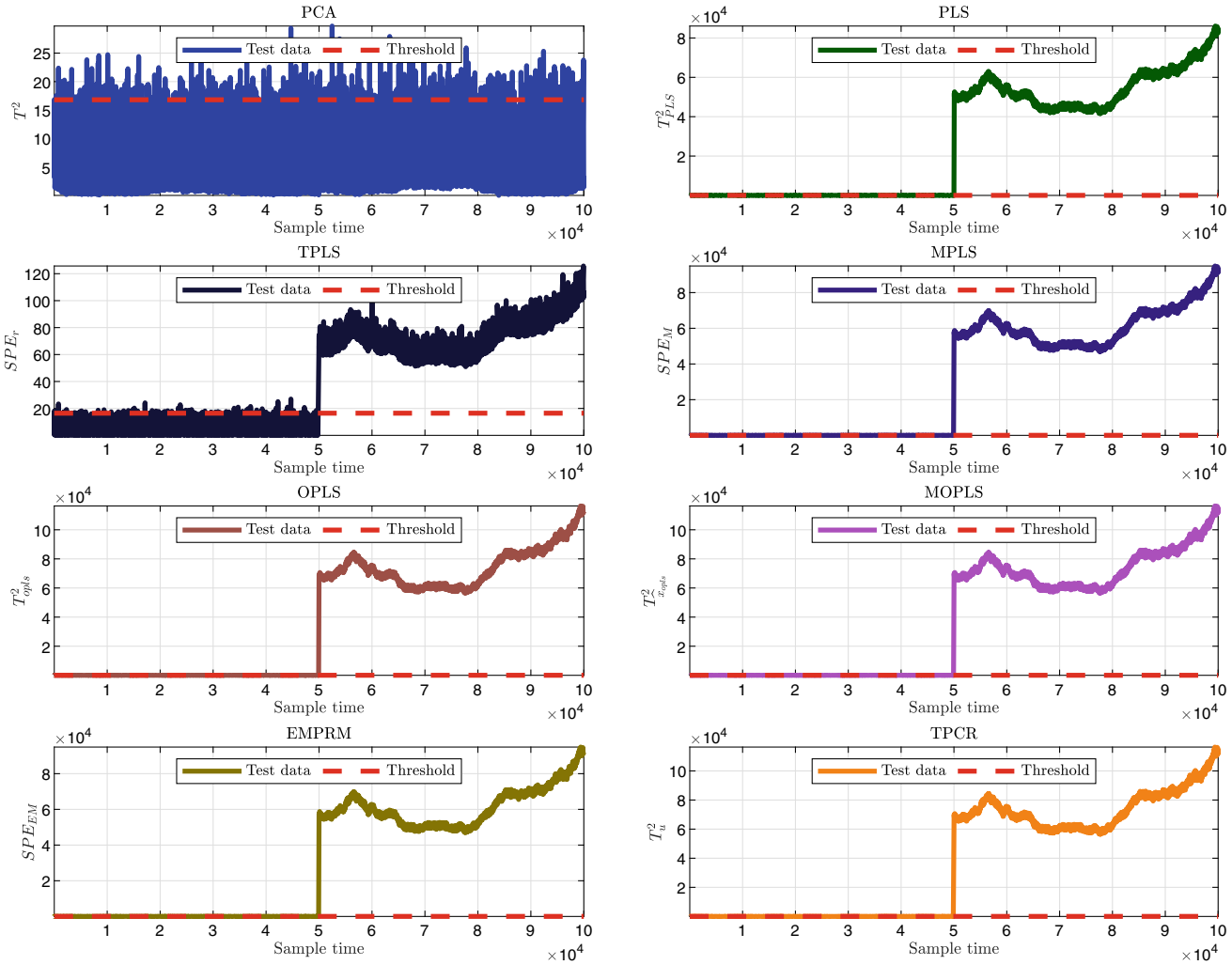

Fig. 7 Process monitoring using different algorithms in case of (F19)

The detectability of the described fault detection approaches in this paper is clearly shown in Fig. 8 that represents the average value of FDR's of all possible sensor fault cases. The MPLS achieved the highest average FDR comparable to other methodologies. As well as, From the point of view of robustness, Fig. 9 introduces a comparison of the fault detection techniques in terms of FARs. It is well proven that both TPLS and PCA are the most robust fault detection techniques.

In summary, the data-driven methods are tested under the assumption that the wind turbine operates in a steady-state region. Simulation results demonstrate that PLS (Partial Least Squares) and its variants exhibit the highest sensitivity to wind turbine sensor faults. Additionally, most of the fault detection methodologies, including MPLS, EMPRM, MOPLS, TPCR, OPLS, PLS, and TPLS, successfully detected all faults with a high Fault Detection Rate (FDR) with average about 90.57%, 88.55%, 88.52%, 88.12%, 88.09%, 86.55%, and 80.95%, respectively. On the other hand, PCA exhibited the lowest FDR compared to the other methods with average about 68.86%. Additionally, this is because PCA cannot establish the correlation between quality and process variables. However, PCA still outperformed PLS and its variants in terms of robustness to these faults, which directly relates to the False Alarm Rate (FAR), i.e, PCA with average FAR about 1.02% is less than TPCR, MOPLS, PLS, OPLS, EMPRM, and MPLS with average FAR about 14.17%, 14.70%, 17.93%, 19.18%, 19.74%, and 22.96%, respectively, except TPLS has the lowest average FAR of about 0.2%.

5 Conclusions

In this paper, we present a comprehensive review and evaluation of the most commonly used multivariate statistical techniques for fault detection, specifically focusing on their application in wind turbines. The primary

Table 13 FDRs (%) of the wind turbine system

Fault	PCA	PLS	TPLS	MPLS	OPLS	MOPLS	EMPRM	TPCR
F1	99.99	99.99	99.97	99.99	99.99	99.99	99.99	99.99
F2	63.59	59.22	39.42	66.75	60.20	61.19	62.01	60.74
F3	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F4	83.71	83.15	76.66	85.34	83.51	83.46	83.46	82.86
F5	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F6	76.46	72.80	68.20	82.91	79.29	79.78	80.25	79.61
F7	99.99	99.99	75.64	99.99	99.99	99.99	99.99	99.99
F8	72.39	65.89	46.96	73.11	67.76	68.81	69.30	68.35
F9	67.80	54.58	**5.76**	76.12	59.80	62.97	60.22	57.64
F10	58.43	46.15	58.92	61.23	55.48	56.88	56.45	55.74
F11	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F12	76.66	63.73	73.19	77.89	72.99	73.80	74.59	73.62
F13	**30.79**	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F14	**14.42**	99.65	99.23	99.70	99.60	99.62	99.63	99.62
F15	**42.95**	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F16	**1.03**	85.84	75.12	88.39	83.25	84.08	85.20	84.27
F17	94.03	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F18	94.33	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F19	**0.61**	99.99	99.99	99.99	99.99	99.99	99.99	99.99
F20	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99

Table 14 FARs (%) of the wind turbine system

Fault	PCA	PLS	TPLS	MPLS	OPLS	MOPLS	EMPRM	TPCR
F1	0.65	5.44	0.18	12.26	7.11	2.53	7.99	2.12
F2	1.82	**44.41**	0.21	**48.82**	**45.43**	**41.66**	**45.93**	**41.28**
F3	1.00	22.12	0.25	**30.04**	**24.42**	15.80	**25.24**	14.86
F4	3.84	**66.30**	0.10	**71.56**	**67.68**	**62.33**	**68.36**	**61.81**
F5	0.94	18.88	0.23	**25.32**	**20.79**	14.00	**21.61**	13.16
F6	0.79	18.34	0.19	**26.76**	**20.53**	13.84	**21.57**	13.04
F7	0.93	**23.77**	0.19	**28.24**	**24.78**	**21.04**	**25.22**	**20.55**
F8	0.61	0.60	0.22	2.23	0.61	0.38	0.63	0.27
F9	0.74	1.38	0.20	5.07	1.72	0.56	1.86	0.38
F10	0.92	**29.39**	0.19	**37.78**	32.50	19.31	**33.78**	17.81
F11	0.61	0.58	0.22	2.19	0.71	0.48	0.66	0.3040
F12	0.56	0.51	0.19	2.05	0.53	0.50	0.57	0.37
F13	0.93	**23.77**	0.19	**28.24**	**24.78**	**21.04**	**25.22**	**20.55**
F14	1.82	**44.41**	0.21	**48.82**	**45.43**	**41.66**	**45.93**	**41.28**
F15	0.74	1.38	0.20	5.07	1.72	0.56	1.86	0.38
F16	0.79	18.34	0.19	**26.76**	**20.53**	13.84	**21.57**	13.04
F17	0.68	8.63	0.19	16.42	10.89	4.22	11.87	3.64
F18	0.92	**29.39**	0.19	**37.78**	**32.50**	19.31	**33.78**	17.81
F19	0.54	0.42	0.20	1.75	0.45	0.50	0.52	0.36
F20	0.56	0.51	0.19	2.05	0.53	0.50	0.57	0.37

objective of this study is to conduct a comparative analysis of data-driven fault detection strategies within the context of wind turbine applications. The data-driven methods are tested under the assumption that the wind turbine operates in a steady-state region. Simulation results demonstrate that PLS (Partial Least Squares) and its variants exhibit the highest sensitivity to wind turbine sensor faults. Additionally, most of the fault detection methodologies, including MPLS, EMPRM, MOPLS, TPCR, OPLS, PLS, and TPLS, successfully detected all faults with a high Fault

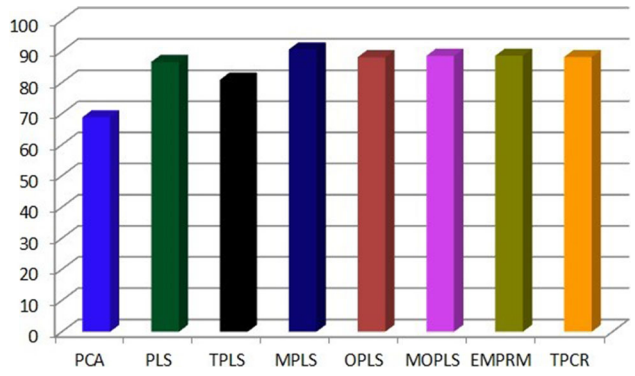

Fig. 8 Average FDR of the presented algorithms

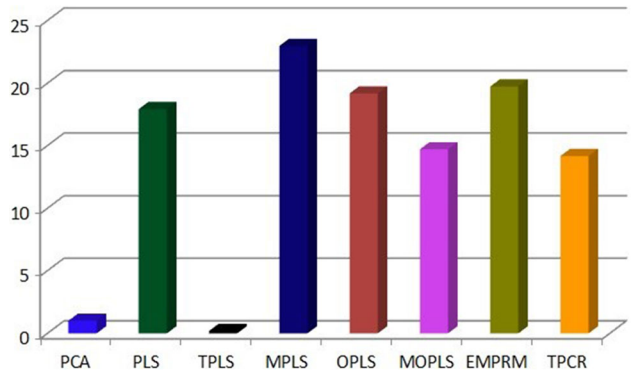

Fig. 9 Average FAR of the presented algorithms

Detection Rate (FDR) averaging around 90%. On the other hand, PCA exhibited the lowest FDR compared to the other methods. However, PCA still outperformed PLS and its variants in terms of robustness to these faults, which directly relates to the False Alarm Rate (FAR). It should be noted that the discussed fault detection methods in this paper are efficient in dealing with single-mode, time-invariant systems. Real wind turbines, however, are multi-mode systems with high nonlinearity. Therefore, future research efforts should focus on the development of process monitoring methods capable of effectively handling these challenges.

Acronyms and abbreviations

See Tables 15 and 16.

Table 15 Abbreviations of algorithms

Abbreviation	Definition
FDD	Fault detection and diagnosis
PCA	Principal component analysis
PLS	Partial least squares
TPLS	Total projection to latent structures
MPLS	Modified partial least squares
OPLS	Orthogonal projection to latent structures
MOPLS	Modified orthogonal projection to latent structures
EMPRM	Expectation–maximization partial robust M-regression
TPCR	Total principal component regression
SVD	Singular value decomposition
PCS	Principal component subspace
RS	Residual subspace
SPE	Squared prediction error
l	No. of latent variables
NIPLS	Nonlinear Iterative Partial Least Squares
EM	Expectation-Maximization
FDR	Fault Detection Rate
FAR	False Alarm Rate
CPV	Cumulative Percent Variance

Table 16 Acronyms of wind turbine system

Acronym	Definition
τ_w	The wind torque affected on the turbine blades
τ_r	The rotor torque
β_r	The reference pitch position
β_m	The measured pitch position
$\tau_{w,m}$	The measured wind torque due to wind speed
$\omega_{r,m}$	The measured rotational speed of the rotor
$\omega_{g,m}$	The measured rotational speed of the generator
$\tau_{g,m}$	The measured generator torque
P_r	The torque reference to the the power reference

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability The authors confirm that the data supporting the findings of this study are available. Raw data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Yin S, Ding SX, Xie X, Luo H (2014) A review on basic data-driven approaches for industrial process monitoring. IEEE Trans Ind Electron 61:6418–6428
2. Elshenawy LM, Yin S, Naik AS, Ding SX (2010) Efficient recursive principal component analysis algorithms for process monitoring. Ind Eng Chem Res 49:252–259
3. Wang D et al (2023) A correlation-graph-CNN method for fault diagnosis of wind turbine based on state tracking and data driving model. Sustain Energy Technol Assess 56:102995
4. Ding SX (2014) Data-driven design of fault diagnosis and fault-tolerant control systems. Springer, Berlin
5. Ding J, Modares H, Chai T, Lewis FL (2016) Data-based multiobjective plant-wide performance optimization of industrial processes under dynamic environments. IEEE Trans Ind Inf 12:454–465
6. Zhang C, Gao X, Xu T, Li Y, Pang Y (2018) Fault detection and diagnosis strategy based on a weighted and combined index in the residual subspace associated with pca. J Chemom 32:e2981
7. Elshenawy LM, Mahmoud TA, Chakour C (2020) Simultaneous fault detection and diagnosis using adaptive principal component analysis and multivariate contribution analysis. Ind Eng Chem Res 59:20798–20815
8. Harrou F, Nounou MN, Nounou HN, Madakyaru M (2015) PLS-based EWMA fault detection strategy for process monitoring. J Loss Prev Process Ind 36:108–119
9. Zhou D, Li G, Qin SJ (2010) Total projection to latent structures for process monitoring. AIChE J 56:168–178
10. He S, Wang Y, Liu C (2018) Modified partial least square for diagnosing key-performance-indicator-related faults. Can J Chem Eng 96:444–454
11. Zhou J, Ren Y, Wang J (2018) Quality-relevant fault monitoring based on locally linear embedding orthogonal projection to latent structure. Ind Eng Chem Res 58:1262–1272
12. Yin S, Wang G, Gao H (2015) Data-driven process monitoring based on modified orthogonal projections to latent structures. IEEE Trans Control Syst Technol 24:1480–1487
13. Zheng J, Song Z, Ge Z (2016) Probabilistic learning of partial least squares regression model: theory and industrial applications. Chemom Intell Lab Syst 158:80–90
14. Wang G, Jiao J (2018) Quality-related fault detection and diagnosis based on total principal component regression model. IEEE Access 6:10341–10347
15. Chen H, Jiang B, Lu N (2018) An improved incipient fault detection method based on Kullback–Leibler divergence. ISA Trans 79:127–136
16. Zhai L, Zhai J, Xie Y (2022) Fault detection and isolation of industrial fermentation process based on semi-supervised convex nonnegative matrix factorizations. J Chem Eng Jpn 55:358–364
17. Han H-G, Wang C-Y, Sun H-Y, Qiao J-F (2022) Data-based robust model predictive control for wastewater treatment process. J Process Control 118:115–125
18. Elshenawy LM, Chakour C, Mahmoud TA (2022) Fault detection and diagnosis strategy based on k-nearest neighbors and fuzzy c-means clustering algorithm for industrial processes. J Franklin Inst 359:7115–7139
19. Kandukuri ST, Klausen A, Karimi HR, Robbersmyr KG (2016) A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management. Renew Sustain Energy Rev 53:697–708
20. Dao PB (2022) Condition monitoring and fault diagnosis of wind turbines based on structural break detection in SCADA data. Renewable Energy 185:641–654
21. Yin S, Wang G, Karimi HR (2014) Data-driven design of robust fault detection system for wind turbines. Mechatronics 24:298–306
22. Odgaard PF, Stoustrup J, Kinnaert M (2009) Fault tolerant control of wind turbines-a benchmark model. IFAC Proc Vol 42:155–160
23. Qiao W, Lu D (2015) A survey on wind turbine condition monitoring and fault diagnosis-part I: Components and subsystems. IEEE Trans Ind Electron 62:6536–6545
24. Wang J, Yang Y, Li N (2023) Randomization-based neural networks for image-based wind turbine fault diagnosis. Eng Appl Artif Intell 121:106028
25. Chakour C, Hamza A, Elshenawy LM (2021) Adaptive CIPCA-based fault diagnosis scheme for uncertain time-varying processes. Neural Comput Appl 33:15413–15432
26. Leoni L, De Carlo F, Abaei MM, BahooToroody A, Tucci M (2023) Failure diagnosis of a compressor subjected to surge events: a data-driven framework. Reliab Eng Syst Saf 233:109107
27. Chen G, Pei Q, Kamruzzaman M (2020) Remote sensing image quality evaluation based on deep support value learning networks. Signal Process: Image Commun 83:115783
28. Liang P, Wang B, Jiang G, Li N, Zhang L (2023) Unsupervised fault diagnosis of wind turbine bearing via a deep residual deformable convolution network based on subdomain adaptation under time-varying speeds. Eng Appl Artif Intell 118:105656
29. Liu N, Xu Y, Tian Y, Ma H, Wen S (2019) Background classification method based on deep learning for intelligent automotive radar target detection. Futur Gener Comput Syst 94:524–535
30. Teng W et al (2021) Vibration analysis for fault detection of wind turbine drivetrains—a comprehensive investigation. Sensors 21:1686
31. Wen X, Xu Z (2021) Wind turbine fault diagnosis based on Relieff-PCA and DNN. Expert Syst Appl 178:115016
32. Laouti N, Sheibat-Othman N, Othman S (2011) Support vector machines for fault detection in wind turbines. IFAC Proc Vol 44:7067–7072
33. Dao PB (2021) A CUSUM-based approach for condition monitoring and fault diagnosis of wind turbines. Energies 14:3236
34. Yang H, Meng C, Wang C (2020) Data-driven feature extraction for analog circuit fault diagnosis using 1-d convolutional neural network. IEEE Access 8:18305–18315
35. Ying Y et al (2013) Toward data-driven structural health monitoring: application of machine learning and signal processing to damage detection. J Comput Civ Eng 27:667–680
36. Dai X, Gao Z (2013) From model, signal to knowledge: a data-driven perspective of fault detection and diagnosis. IEEE Trans Ind Inf 9:2226–2238
37. Chen H, Chai Z, Dogru O, Jiang B, Huang B (2021) Data-driven designs of fault detection systems via neural network-aided learning. IEEE Trans Neural Netw Learn Syst 33:5694–5705

38. Feng S, Han X, Ma Z, Królczyk G, Li Z (2020) Data-driven algorithm for real-time fatigue life prediction of structures with stochastic parameters. Comput Methods Appl Mech Eng 372:113373
39. Zeng W et al (2023) Data-driven management for fuzzy sewage treatment processes using hybrid neural computing. Neural Comput Appl 35:23781–23794
40. Ding SX (2013) Basic Ideas, Major Issues and Tools in the Observer-Based FDI Framework. Springer, Berlin, pp 13–19
41. Qin SJ (2012) Survey on data-driven industrial process monitoring and diagnosis. Annu Rev Control 36:220–234
42. Trygg J, Wold S (2002) Orthogonal projections to latent structures (O-PLS). J Chemom: J Chemom Soc 16:119–128
43. Chiang LH, Russell EL, Braatz RD (2000) Fault Detection and Diagnosis in Industrial Systems. Springer, Berlin
44. Valle S, Li W, Qin SJ (1999) Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. Ind Eng Chem Res 38:4389–4401
45. Jackson JE, Mudholkar GS (1979) Control procedures for residuals associated with principal component analysis. Technometrics 21:341–349
46. Tracy ND, Young JC, Mason RL (1992) Multivariate control charts for individual observations. J Qual Technol 24:88–95
47. Höskuldsson A (1988) PLS regression methods. J Chemom 2:211–228
48. Dayal BS, MacGregor JF (1997) Improved PLS algorithms. J Chemom: J Chemom Soc 11:73–85
49. Nomikos P, MacGregor JF (1995) Multivariate SPC charts for monitoring batch processes. Technometrics 37:41–59
50. Yin S, Ding SX, Zhang P, Hagahni A, Naik A (2011) Study on modifications of PLS approach for process monitoring. IFAC Proc Vol 44:12389–12394
51. Wang G, Jiao J, Yin S (2017) Quality-related fault detection approaches based on data preprocessing. IFAC-PapersOnLine 50:15740–15747
52. Yin S, Wang G, Yang X (2014) Robust PLS approach for KPI-related prediction and diagnosis against outliers and missing data. Int J Syst Sci 45:1375–1382
53. Bianchi FD, De Battista H, Mantz RJ (2007) Wind turbine control systems: principles, modelling and gain scheduling design, vol 19. Springer, Berlin
54. Munteanu I, Bratcu AI, CeangĂ E, Cutululis N-A (2008) Optimal control of wind energy systems: towards a global approach, vol 22. Springer, Berlin
55. Burton T, Jenkins N, Sharpe D, Bossanyi E (2011) Wind energy handbook. Wiley, New York
56. Johnson KE, Pao LY, Balas MJ, Fingersh LJ (2006) Control of variable-speed wind turbines: standard and adaptive techniques for maximizing energy capture. IEEE Control Syst Mag 26:70–81
57. Amirat Y, Benbouzid MEH, Al-Ahmar E, Bensaker B, Turri S (2009) A brief status on condition monitoring and fault diagnosis in wind energy conversion systems. Renew Sustain Energy Rev 13:2629–2636
58. Amirat Y, Benbouzid ME, Bensaker B, Wamkeue R (2007) Condition monitoring and ault diagnosis in wind energy conversion systems: a review. In: IEEE. vol 2, pp 1434–1439
59. Dolan DS, Lehn PW (2006) Simulation model of wind turbine 3p torque oscillations due to wind shear and tower shadow. IEEE Trans Energy Convers 21:717–724
60. Manring ND, Fales RC (2019) Hydraulic control systems. Wiley, New York
61. Lee J-M, Qin SJ, Lee I-B (2006) Fault detection and diagnosis based on modified independent component analysis. AIChE J 52:3501–3514
62. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130