Floating offshore wind turbine fault diagnosis via regularized dynamic canonical correlation and fisher discriminant analysis

Wu, Ping; Liu, Yichao; Ferrari, Riccardo M.G.; van Wingerden, Jan Willem

DOI
10.1049/rpg2.12319

Publication date
2021

Document Version
Final published version

Published in
IET Renewable Power Generation

Received: 25 June 2021 | Revised: 15 September 2021 | Accepted: 16 October 2021

IET Renewable Power Generation

DOI: 10.1049/rpg2.12319

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Floating offshore wind turbine fault diagnosis via regularized dynamic canonical correlation and fisher discriminant analysis

Ping Wu[1] | **Yichao Liu**[2] | **Riccardo M.G. Ferrari**[2] | **Jan-Willem van Wingerden**[2]

[1] School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou, China

[2] Faculty of Mechanical Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands

Correspondence
Ping Wu, School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China.
Email: pingwu@zstu.edu.cn
Riccardo M.G. Ferrari, Faculty of Mechanical Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands.
Email: R.Ferrari@tudelft.nl

Funding information
European Union via a Marie Sklodowska-Curie Action, Grant/Award Numbers: Project EDOWE, 835901; Fundamental Research Funds of Zhejiang Sci-Tech University, Grant/Award Number: 2021Q024; Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China, Grant/Award Number: ICT2021B45; Zhejiang Province Public Welfare Technology Application Research Project, Grant/Award Numbers: LGF19F030004, LGG21F030015

Abstract

Over the past decades, Floating Offshore Wind Turbine (FOWT) has gained increasing attention in wind engineering due to the rapidly growing energy demands. However, difficulties in turbine maintenance will increase due to the harsh operational conditions. Fault diagnosis techniques play a crucial role to enhance the reliability of FOWTs and reduce the cost of offshore wind energy. In this paper, a novel data-driven fault diagnosis method using regularized dynamic canonical correlation analysis (RDCCA) and Fisher discriminant analysis (FDA) is proposed for FOWTs. Specifically, to overcome the collinearity problem that exists in measured process data, dynamic canonical correlation analysis with a regularization scheme, is developed to exploit the relationship between input and output signals. Then, the residual signals are generated from the established RDCCA model for fault detection. To further classify the fault type, an FDA model is trained from the residual signals of different training faulty data sets. Simulations on a FOWT baseline model based on the widely used National Renewable Energy Laboratory FAST simulator are carried out to demonstrate the feasibility and efficacy of the proposed fault detection and classification method. Results have shown many salient features of the proposed method with potential applications in FOWTs.

1 | INTRODUCTION

Nowadays, wind energy has become one of the most important renewable sources for fulfilling the world's energy demand [1]. Compared to onshore installed wind turbines, offshore wind turbines have several advantages such as more reliable sources of wind energy, more power, less environmental impact etc. Therefore, offshore wind turbines have a large potential to grow in the future wind industry [2]. Among many types of offshore wind turbines, floating offshore wind turbine (FOWT) can easily access deeper water areas to harness the best wind resources. FOWT plays a leading role to economically harvest wind energy over the deeper sea [3, 4].

However, FOWT would frequently suffer from various faults and failures in sensors, actuators, and components, due to the harsh working environments with storing wind, rising water level. Recently, some studies showed that the ratio of the operation and maintenance (O & M) costs to the total cost of energy can reach 23% for offshore wind turbines [5]. Therefore, condition monitoring (CM) is of vital importance for floating offshore wind power systems to ensure stability and extend long service life [6]. With efficient fault detection and classification technique for online condition monitoring of wind turbines, the downtime can be reduced and severe component damage can be avoided. Furthermore, maintenance strategies can be optimized to avoid unnecessary activities, thus making fault detection and classification techniques an enabling technology for reducing O&M costs and maximizing the profit of wind energy projects [7].

In recent years, there has been an increasing amount of literature on wind turbine fault detection and classification [8]. Wind turbine fault detection and classification methods can be

roughly categorized into two classes: model-based and data-driven approaches. For model-based approaches, Kalman filter and estimators, parity equations, and observer-based fault detection techniques have been reported in the literature for wind turbines or their subsystems [9–13]. The advantage of model-based methods is that they can offer the faults insight. However, model-based approaches rely on an accurate mathematical model which is often difficult to obtain for wind turbines or their subsystems due to the complicated structure.

Different from model-based approaches, data-driven methods only require the available system I/O data for the task of fault detection and classification. Due to the model-free characteristic, the data-driven methods are more suitable and practical for wind turbines or their subsystems [14–17]. With the rapid evolution of sensor and communication technologies, it is facilitating the development of data-driven wind turbine fault diagnosis. Recently, a considerable amount of literature has been published on the data-driven wind turbine fault detection and classification methods. Bessa et al. considered the detection of the occurrence of faults as a change point detection problem in time-series data. A wind turbine fault detection method based on the Gibbs sampling algorithm was developed. Moreover, a Fuzzy/Bayesian network scheme was applied for fault isolation [18]. Pashazadeh et al. developed a wind turbine fault detection and classification scheme through a fused classifier with decision tree and k-nearest neighbor (*knn*) in parallel [19]. Based on supervisory control and data acquisition (SCADA) data, Song et al. presented a wind turbine health states monitoring method using a Bayesian framework to identify abnormal turbine states [20]. Li et al. applied long short-term memory networks (LSTM) to generate residual signals for the purpose of wind turbine fault detection and employed random forest for decision-making [21]. Jiang et al. applied denoising autoencoder (DAE) to build a multivariate reconstruction model on raw time-series data from multiple sensors. The reconstruction error of the DAE was analyzed to detect faults [22].

Compared to the above-mentioned machine learning-based methods, multivariate analysis (MVA) methods have also been widely used in wind turbine fault detection and classification. For MVA based fault detection method, a fault is detected by comparing the constructed statistic and its corresponding threshold [23]. The main advantages of MVA are its simplicity and efficiency. Among these MVA methods, principal component analysis (PCA) and Fisher discriminant analysis (FDA) have been commonly employed for wind turbine fault detection and classification. In PCA, an orthogonal transformation is learned to convert a set of correlated variables into uncorrelated variables. As a representative classification method, the FDA determines a set of projection vectors that minimize the scatter within each class while maximizing the scatter between the classes [24]. Krüger et al. adopted PCA to investigate the correlations of the measured process data for the detection of process abnormalities that happened in wind turbine gearbox and then used the FDA to classify the fault types [25]. Pozo et al. used PCA and statistical hypothesis testing to develop a fault detection scheme for wind turbine [26]. Wang et al. studied the PCA based variable selection method and used the established PCA model for developing wind turbine fault detection and identification algorithms [27]. Nevertheless, these systems are mostly designed for onshore wind turbines. The fault detection and classification of FOWTs is rarely studied.

Different from PCA-based fault detection methods, canonical correlation analysis (CCA) determines the input–output relationship of the system by maximizing the correlations between input and output signals [28, 29]. CCA-based fault detection method is more efficient while the input–output relationship is explicitly established, compared to PCA-based method [30, 31]. Since the measured process data often exhibits collinearity problem in practice [29], a regularized scheme was employed. To capture the system dynamics, CCA can be extended to dynamic CCA (DCCA) or canonical variate analysis (CVA) by taking past input and output data into consideration [32, 33]. CCA and its variants have been successfully applied for fault detection and diagnosis in a variety of industrial processes [34–38]. Furthermore, a combined canonical variate analysis and Fisher discriminant analysis was developed for fault classification for chemical processes, where the state space vector is extracted by canonical variate analysis and fed into Fisher discriminant analysis model [39]. However, to the authors' knowledge, CCA is rarely investigated for system-wide wind turbine fault diagnosis.

Inspired by the idea of combined CVA and FDA [39], we develop a new fault diagnosis method using regularized dynamic canonical correlation analysis and Fisher discriminant analysis for FOWT. First, the relationship between input and output signals is described by presenting the regularized dynamic canonical correlation analysis (RDCCA) where a regularization scheme is integrated into dynamic canonical correlation analysis. Then, the residual generator is established through the RDCCA model. The system dynamics is captured and the collinearity problem in measured process data is addressed. To further classify the fault types, FDA is combined into the framework of RDCCA. The residual vectors derived from the RDCCA model are used to build an FDA model for fault classification. The main contributions of this paper are two-fold as follows,

- The fault detection and classification problem for FOWTs is investigated. Although numerous studies of wind turbine fault detection and classification systems have been reported in the literature, little attention has been paid to FOWTs. The critical faults of the power and control system and of the FOWT drive train are considered in this paper.
- A new fault detection and classification method using RDCCA and FDA is proposed for FOWTs. The RDCCA is developed to design the fault detection of FOWTs by taking the dynamics and correlations in multivariate sensor time-series data into account. And then the fault types are classified by the integration of RDCCA and FDA methods.

The rest of this paper is organized as follows. In Section 2, a brief review is dedicated to CCA. Section 3 presents the proposed RDCCA-FDA approach based on fault detection and classification scheme. The FOWT fault simulation benchmark using the National Renewable Energy Laboratory (NREL) FAST (Fatigue, Aerodynamics, Structures and Turbulence) v8

simulator is introduced in Section 4. Performance evaluation is carried out on the simulated benchmark model in Section 5. Finally, conclusions are drawn in Section 6.

2 | BRIEF REVIEW OF CCA

As a representative multivariate analysis method, CCA has been widely employed to analyze the relationship between two variables. By maximizing the correlations between the projections of high-dimensional variables, the relationship between input and output signals can be described in an explicit way. These projections are also called canonical varieties in CCA. CCA has been proven as a promising tool for developing fault diagnosis systems [30, 32].

Let $\mathbf{u} \in \mathcal{R}^m$ and $\mathbf{y} \in \mathcal{R}^l$ be the normalized input and output vectors. The aim of CCA is to find weight vectors $\mathbf{w}_u$ and $\mathbf{w}_y$ to maximize the correlations between the projections $\mathbf{w}_u^T\mathbf{u}$ and $\mathbf{w}_y^T\mathbf{y}$. Thus, the optimization problem of CCA can be formulated as follows,

$$\begin{aligned} \max_{\mathbf{w}_u,\mathbf{w}_y} \quad & \frac{\mathbf{w}_u^T \Sigma_{uy} \mathbf{w}_y}{(\mathbf{w}_u^T \Sigma_u \mathbf{w}_u)^{1/2} (\mathbf{w}_y^T \Sigma_y \mathbf{w}_y)^{1/2}} \\ \text{s.t.} \quad & \mathbf{w}_u^T \Sigma_u \mathbf{w}_u = 1, \\ & \mathbf{w}_y^T \Sigma_y \mathbf{w}_y = 1, \end{aligned} \tag{1}$$

where Σ_{uy} is the covariance matrix between $\mathbf{u}$ and $\mathbf{y}$. Σ_u and Σ_y are the variance matrices of $\mathbf{u}$ and $\mathbf{y}$, respectively. Generally, the optimization problem Equation (1) can be easily solved through singular vector decomposition (SVD) or eigenvalue decomposition.

Assuming that N samples $\{\mathbf{u}_i, \mathbf{y}_i\}_{i=1,\dots,N}$ are collected, the input and output matrices of $\mathbf{U}$ and $\mathbf{Y}$ can be obtained as,

$$\mathbf{U} = \left[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_N\right],$$
$$\mathbf{Y} = \left[\mathbf{y}_1\ \mathbf{y}_2\ \cdots\ \mathbf{y}_N\right].$$

Then, the co-variance matrices $\Sigma_{\mathbf{u}}, \Sigma_{\mathbf{y}}$ and cross-variance matrix $\Sigma_{\mathbf{u},\mathbf{y}}$ can be estimated, respectively, as

$$\Sigma_{\mathbf{u}} = \frac{1}{N-1}\mathbf{U}\mathbf{U}^T,$$
$$\Sigma_{\mathbf{y}} = \frac{1}{N-1}\mathbf{Y}\mathbf{Y}^T,$$
$$\Sigma_{\mathbf{u},\mathbf{y}} = \frac{1}{N-1}\mathbf{U}\mathbf{Y}^T.$$

Perform SVD on the following matrix,

$$\Sigma_{\mathbf{u}}^{-1/2}\Sigma_{\mathbf{u},\mathbf{y}}\Sigma_{\mathbf{y}}^{-1/2} = \mathbf{Y}_c\Delta_c\Psi_c^T. \tag{2}$$

Then, the weight matrices are obtained,

$$\mathbf{W}_u = \Sigma_{\mathbf{u}}^{-1/2}\mathbf{Y}_c, \tag{3}$$

$$\mathbf{W}_y = \Sigma_{\mathbf{y}}^{-1/2}\Psi_c. \tag{4}$$

However, the components extracted by the CCA often suffer from ill-conditioned problem due to the sensitivity to noise in the data. To deal with this problem, CCA is usually extended to regularized CCA (RCCA) [29]. In RCCA, the regularization terms are added as follows,

$$\begin{aligned} \max_{\mathbf{w}_u,\mathbf{w}_y} \quad & \frac{\mathbf{w}_u^T \Sigma_{uy} \mathbf{w}_y}{(\mathbf{w}_u^T \Sigma_u \mathbf{w}_u)^{1/2} (\mathbf{w}_y^T \Sigma_y \mathbf{w}_y)^{1/2}} \\ \text{s.t.} \quad & \mathbf{w}_u^T (\Sigma_u + \alpha_1\mathbf{I})\mathbf{w}_u = 1, \\ & \mathbf{w}_y^T (\Sigma_y + \alpha_2\mathbf{I})\mathbf{w}_y = 1, \end{aligned} \tag{5}$$

where α_1 and α_2 are the the regularization parameters to control non-zero eigenvalues to avoid the ill-conditioned problem. The optimization problem Equation (5) can be solved through SVD in a similar manner as CCA.

Remark 1. The regularization terms are added into the objective function Equation (2) to ensure non-zero eigenvalues. While α_1 and α_2 are set to 0, RCCA will reduce to CCA.

3 | PROPOSED METHOD

The input and output signals are usually extended to the stacked vectors to capture the system dynamics. To do so, the stacked past data vector $\mathbf{z}_p \in \mathcal{R}^{(l+m)q}$, future input vector $\mathbf{u}_f \in \mathcal{R}^{mq}$ and future output vector $\mathbf{y}_f \in \mathcal{R}^{lf}$ are formed in a time-lagged way. For example, $\mathbf{z}_p$ and $\mathbf{u}_f$ are formed as follows,

$$\mathbf{z}_p(k) = \begin{bmatrix} \mathbf{u}(k-1) \\ \vdots \\ \mathbf{u}(k-q) \\ \mathbf{y}(k-1) \\ \vdots \\ \mathbf{y}(k-q) \end{bmatrix} \quad \mathbf{u}_f(k) = \begin{bmatrix} \mathbf{u}(k) \\ \mathbf{u}(k+1) \\ \vdots \\ \mathbf{u}(k+q-1) \end{bmatrix}.$$

$\mathbf{y}_f$ is also defined as $\mathbf{u}_f$. q is the number of time lags in past and future data vectors $\mathbf{z}_p(k)$, $\mathbf{u}_f(k)$ and $\mathbf{y}_f(k)$.

According to the realization theory [40], the future output vector $\mathbf{y}_f$ can be predicted through the past inputs and outputs $\mathbf{z}_p$, and the future inputs $\mathbf{u}_f$ by analyzing their correlations. For the purpose of establishing the relationship between $\mathbf{y}_f$ and $\mathbf{z}_{pf} = \left[\mathbf{z}_p^\top\ \mathbf{u}_f^\top\right]^\top$, the optimization problem of DCCA can be cast as,

$$\max_{\mathbf{w}_{dz},\mathbf{w}_{dy}} \quad \frac{\mathbf{w}_{dz}^T \Sigma_{z_{pf},y_f} \mathbf{w}_{dy}}{(\mathbf{w}_{dz}^T \Sigma_{z_{pf}} \mathbf{w}_{dz})^{1/2} (\mathbf{w}_{dy}^T \Sigma_{y_f} \mathbf{w}_{dy})^{1/2}}$$

$$\text{s.t.} \quad \mathbf{w}_{d\mathbf{z}}^T \Sigma_{\mathbf{z}_{pf}} \mathbf{w}_{d\mathbf{z}} = 1, \tag{6}$$
$$\mathbf{w}_{d\mathbf{y}}^T \Sigma_{\mathbf{y}_f} \mathbf{w}_{d\mathbf{y}} = 1,$$

where $\Sigma_{\mathbf{z}_{pf},\mathbf{y}_f}$ is the covariance matrix between $\mathbf{z}_{pf}$ and $\mathbf{y}_f$. $\Sigma_{\mathbf{z}_{pf}}$ and $\Sigma_{\mathbf{y}_f}$ are the variance matrices of $\mathbf{z}_{pf}$ and $\mathbf{y}_f$, respectively. $\mathbf{w}_{d\mathbf{z}}$ and $\mathbf{w}_{d\mathbf{y}}$ are the weight vectors to be found.

Collinearity is a common statistical issue in industrial data sets, particularly while there exists feedback control in dynamic systems. A simple way to address this collinearity issue is using a regularization scheme. In this study, we also adopt the regularization scheme to deal with the measured process data collected from the wind turbine system. Specifically, two regularization terms are added to convert $\Sigma_{\mathbf{z}_{pf}}$ and $\Sigma_{\mathbf{y}_f}$ to $\Sigma_{\mathbf{z}_{pf}} + \alpha_1 \mathbf{I}$ and $\Sigma_{\mathbf{y}_f} + \alpha_2 \mathbf{I}$, respectively. Here, α_1 and α_2 are the regularization parameters. The regularization parameters α_1 and α_2 are often determined by cross-validation. Therefore, the optimization problem of RDCCA can be cast as,

$$\max_{\mathbf{w}_{d\mathbf{z}},\mathbf{w}_{d\mathbf{y}}} \quad \frac{\mathbf{w}_{d\mathbf{z}}^T \Sigma_{\mathbf{z}_{pf},\mathbf{y}_f} \mathbf{w}_{d\mathbf{y}}}{(\mathbf{w}_{d\mathbf{z}}^T \Sigma_{\mathbf{z}_{pf}} \mathbf{w}_{d\mathbf{z}})^{1/2} (\mathbf{w}_{d\mathbf{y}}^T \Sigma_{\mathbf{y}_f} \mathbf{w}_{d\mathbf{y}})^{1/2}}$$
$$\text{s.t.} \quad \mathbf{w}_{d\mathbf{z}}^T (\Sigma_{\mathbf{z}_{pf}} + \alpha_1 \mathbf{I}) \mathbf{w}_{d\mathbf{z}} = 1, \tag{7}$$
$$\mathbf{w}_{d\mathbf{y}}^T (\Sigma_{\mathbf{y}_f} + \alpha_2 \mathbf{I}) \mathbf{w}_{d\mathbf{y}} = 1.$$

Assuming that N samples are collected, the Hankel matrices of $\mathbf{z}_{pf}$ and $\mathbf{y}_f$ can be created as,

$$\mathbf{Z}_{pf} = \left[\mathbf{z}_{pf}(q+1) \cdots \mathbf{z}_{pf}(M)\right],$$
$$\mathbf{Y}_f = \left[\mathbf{y}_f(q+1) \cdots \mathbf{y}_f(M)\right],$$

where $M = N - 2q + 1$. Then, the co-variance matrices $\Sigma_{\mathbf{z}_{pf}}, \Sigma_{\mathbf{y}_f}$ and cross-variance matrix $\Sigma_{\mathbf{z}_{pf},\mathbf{y}_f}$ can be estimated, respectively, as

$$\Sigma_{\mathbf{z}_{pf}} = \frac{1}{N-1} \mathbf{Z}_{pf} \mathbf{Z}_{pf}^T, \tag{8}$$

$$\Sigma_{\mathbf{y}_f} = \frac{1}{N-1} \mathbf{Y}_f \mathbf{Y}_f^T, \tag{9}$$

$$\Sigma_{\mathbf{z}_{pf},\mathbf{y}_f} = \frac{1}{N-1} \mathbf{Z}_{pf} \mathbf{Y}_f^T. \tag{10}$$

Based on the estimated covariance matrices and cross-variance matrix, the optimization problem Equation (7) can be readily solved through SVD. According to linear algebra theory, the SVD operation is performed as follows,

$$(\Sigma_{\mathbf{z}_{pf}} + \alpha_1 \mathbf{I})^{-1/2} \Sigma_{\mathbf{z}_{pf},\mathbf{y}_f} (\Sigma_{\mathbf{y}_f} + \alpha_2 \mathbf{I})^{-1/2} = \Upsilon \Delta \Psi^T, \tag{11}$$

to derive the solution of Equation (7). In Equation (11), Υ contains the left singular vectors and Ψ contains the right singular vectors. $\Delta = \text{diag}(\lambda_1, \ldots, \lambda_{lq})$ is a diagonal matrix which consists of descended order singular values. For the purpose of dimensionality reduction, only n largest singular values are retained. Usually, $n \leq lq$. Thus, $\Delta_n = \text{diag}(\lambda_1, \ldots, \lambda_n)$. From the result of SVD, the weight matrices $\mathbf{W}_{d\mathbf{z}}$ and $\mathbf{W}_{d\mathbf{y}}$ can be calculated,

$$\mathbf{W}_{d\mathbf{z}} = \Sigma_{\mathbf{z}_{pf}}^{-1/2} \Upsilon(:, 1:n), \tag{12}$$

$$\mathbf{W}_{d\mathbf{y}} = \Sigma_{\mathbf{y}_f}^{-1/2} \Psi(:, 1:n). \tag{13}$$

Remark 2. In the proposed RDCCA method, the regularization terms are imposed where $\mathbf{w}_{d\mathbf{z}}^T(\Sigma_{\mathbf{z}_{pf}} + \alpha_1 \mathbf{I})\mathbf{w}_{d\mathbf{z}} = 1$ and $\mathbf{w}_{d\mathbf{y}}^T(\Sigma_{\mathbf{y}_f} + \alpha_2 \mathbf{I})\mathbf{w}_{d\mathbf{y}} = 1$. While α_1 and α_2 are set to 0, RDCCA will reduce to DCCA.

Based on the established RDCCA model, the residual signals are generated from $\mathbf{z}_{pf}(k)$ and $\mathbf{y}_f(k)$ at time instant k as

$$\mathbf{r}(k) = \mathbf{W}_{d\mathbf{y}}^T \mathbf{y}_f(k) - \Delta_n \mathbf{W}_{d\mathbf{z}}^T \mathbf{z}_{pf}(k) \tag{14}$$

For fault detection, the Hotelling's statistic $T_r^2(k)$ is calculated from $\mathbf{r}(k)$,

$$T_r^2(k) = \mathbf{r}(k)^T \Sigma_r^{-1} \mathbf{r}(k) \tag{15}$$

Here, Σ_r is the co-variance matrix of $\mathbf{r}$. Under the assumption that measured process data are sampled from Gaussian distributions, the threshold J_{th} can be calculated from χ^2 distribution corresponding to a certain significance level α,

$$J_{th} = \chi_\alpha^2(n). \tag{16}$$

Using Hotelling's statistic $T_r^2(k)$ and threshold J_{th}, the fault decision logic can be concluded,

$$\begin{cases} T_r^2(k) > J_{th} \Rightarrow \text{faulty}. \\ T_r^2(k) \leq J_{th} \Rightarrow \text{fault-free}. \end{cases} \tag{17}$$

The procedure of the RDCCA-based fault detection method can be summarized as follows,

- **Offline Training**
  - Normalize the collected samples $\mathbf{U}$ and $\mathbf{Y}$ under the fault-free scenario.
  - Compute the covariance and variance matrices.
  - Perform SVD through Equation (11) with determined parameters to extract the weight matrices.
  - Obtain the residuals from Equation (14) and calculate the corresponding Hotelling's statistics from Equation (15).
  - Calculate the threshold from Equation (16) corresponding to a certain significance level.
- **Online Monitoring**
  - Collect and normalize the online sample with the mean and variance of normal samples.

- Obtain the online residuals and Hotelling's statistic from Equations (14) and (15) with calculated weight matrices in the offline training phase.
- Determine the fault-free or faulty sample by using the logic Equation (17).

To further classify the fault type after the detection of fault occurrence, FDA is integrated into RDCCA in this study. FDA is a typical pattern classification method [41]. In this paper, we conduct the FDA method using the residual signals derived from RDCCA. The basic idea of FDA is to determine a set of projection vectors that optimize the Fisher criterion. Assuming that N_{all} observations of residuals $\mathbf{r}(k)$ are obtained from all c faulty classes, the training data matrix for FDA is then arranged as $\mathbf{R} \in \mathcal{R}^{N_{all} \times n}$. The dispersion of one class can be represented by using the within-scatter matrix. As an example, for the jth class $\mathbf{R}_j$, the scatter matrix $\mathbf{S}_j$ is defined as,

$$\mathbf{S}_j = \sum_{\mathbf{r}(k) \in \mathbf{R}_j} (\mathbf{r}(k) - \mathbf{r}_{m,j})(\mathbf{r}(k) - \mathbf{r}_{m,j})^T, \tag{18}$$

where $\mathbf{r}_{m,j}$ is the mean vector of residuals of $\mathbf{R}_j$. The within-class-scatter matrix is given by summing all $\mathbf{S}_j$,

$$\mathbf{S}_w = \sum_{j=1}^{c} \mathbf{S}_j. \tag{19}$$

The dispersion of all classes is indicated by the total-scatter matrix. The total-scatter matrix $\mathbf{S}_t$ is defined as,

$$\mathbf{S}_t = \sum_{\mathbf{r}(k) \in \mathbf{R}} (\mathbf{r}(k) - \mathbf{r}_m)(\mathbf{r}(k) - \mathbf{r}_m)^T, \tag{20}$$

where $\mathbf{r}_m$ is the mean vector of $\mathbf{R}$. The between-class-scatter matrix $\mathbf{S}_b$ is calculated as,

$$\mathbf{S}_b = \sum_{j=1}^{c} n_j (\mathbf{r}_m - \mathbf{r}_{m,j})(\mathbf{r}_m - \mathbf{r}_{m,j})^T = \mathbf{S}_t - \mathbf{S}_w. \tag{21}$$

Here, n_j is the number of observations of j^{th} class. The aim of FDA is to find the optimal discriminant directions $\mathbf{w}_i$ (for $i = 1, \dots, b$ with $b \leq c - 1$) to maximize the separability between different classes while minimizing the within-class scatter. Thus, the Fisher criterion is to be maximized as follows,

$$\max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}. \tag{22}$$

The solution of the optimization problem Equation (22) is equivalent to solve a generalized eigenvalue problem,

$$\mathbf{S}_b \mathbf{w}_i = \mu_i \mathbf{S}_w \mathbf{w}_i, \tag{23}$$

where μ_i is the generalized eigenvalue. The corresponding eigenvector is the Fisher discriminant direction $\mathbf{w}_i$. From the solution of generalized eigenvalue problem Equation (23), the optimal discriminant directions can be defined as $\mathbf{W}_b = [\mathbf{w}_1\ \mathbf{w}_2\ \cdots\ \mathbf{w}_b]$, where $\mathbf{w}_1$ to $\mathbf{w}_b$ are the first b eigenvectors sorted in the descending order of the eigenvalues. Then, the projection of $\mathbf{r}(k)$ onto the discriminant subspace can be represented as,

$$\mathbf{d}(k) = \mathbf{W}_b^T \mathbf{r}(k). \tag{24}$$

To classify the class of a test residual vector $\mathbf{r}(t)$, the Fisher discriminant function is usually employed,

$$\begin{aligned} g_j(\mathbf{r}(t)) = &-\frac{1}{2}(\mathbf{d}(t) - \mathbf{d}_{m,j})^T \left(\frac{1}{n_j - 1} \mathbf{W}_b^T \mathbf{S}_j \mathbf{W}_b \right)^{-1} \\ &(\mathbf{d}(t) - \mathbf{d}_{m,j}), \\ &- \frac{1}{2} \ln \left[\det \left(\frac{1}{n_j - 1} \mathbf{W}_b^T \mathbf{S}_j \mathbf{W}_b \right) \right], \end{aligned} \tag{25}$$

where $\mathbf{d}_{m,j} = \mathbf{W}_b^T \mathbf{r}_{m,j}$. Then, the class of the test $\mathbf{r}(t)$ is determined to be one of the faulty classes by observing the Fisher discriminant function,

$$C(\mathbf{r}(t)) = arg \max_{1 \leq j \leq c} g_j(\mathbf{r}(t)) \tag{26}$$

The procedure of the RDCCA-FDA based fault classification method can be summarized as follows,

- **Offline Training**
  - Collect and normalize the faulty samples.
  - Obtain the residuals from the collected faulty samples as Equation (14).
  - Solve the generalized eigenvalue problem Equation (23) to derive the optimal discriminant directions $\mathbf{W}_b$.
- **Online Classification**
  - Collect and normalize the online faulty sample.
  - Calculate the Fisher discriminant function Equation (25).
  - The faulty sample is classified as one of the faulty types where the corresponding Fisher discriminant function has the maximum value.

In summary, the flowchart of the proposed RDCCA-FDA based fault detection and classification approach is illustrated in Figure 1.

4 | CASE STUDY

In this section, a 10MW three-bladed variable speed reference wind turbine with the Triple-Spar floating platform is simulated to verify the performance of the proposed RDCCA-FDA method. The simulation is based on the high-fidelity benchmark model developed by the Technical University of Denmark (DTU)[42] and Stuttgart Wind Energy (SWE) institute [13, 43]. The benchmark model of the operational FOWT is simulated

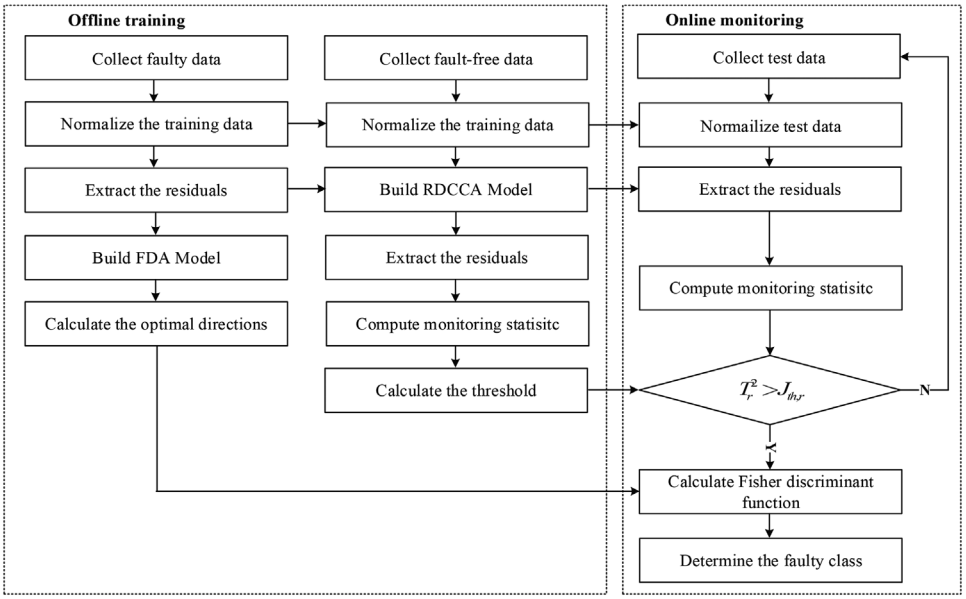

FIGURE 1 The RDCCA-FDA based fault detection and classification approach

TABLE 1 Specifications of the DTU 10MW FOWT (SWL: Sea Water Level)

Parameter	Value
Turbine system	
Rating	10 MW
Rotor orientation, configuration	Upwind, 3 blades
Pitch control	Variable speed, collective pitch
Drive-train	Medium speed, multiple stage gearbox
Rotor, hub diameter	178.3 m, 5.6 m
Hub height	119 m
Cut-in, rated, cut-out wind speed	4 m/s, 11.4 m/s, 25 m/s
Cut-in, rated rotor speed	6 rpm, 9.6 rpm
Rated tip speed	90 m/s
Floating platform	
Total height	66 m
Distance from the tower center-line	26 m
Draft	56 m
Single column diameter	15 m
Column elevation above SWL	10 m
Elevation of tower base above SWL	25 m
Water displacement	29497.7 m^3

TABLE 2 Wind and wave conditions in two LCs

Load Case	Wind condition	U_m (m/s)	H_s (m)	T_p (s)
LC1	Turbulent wind	12	2.6	7.42
LC2	Turbulent wind	20	2.66	7.42

by the physics-based FAST v8 simulator. The benchmark model specifications and parameters are listed in Table 1.

In the FOWT benchmark model, the waves are simulated by using the JONSWAP wave spectrum where the significant wave height and peak period are estimated according to the conditional probabilistic distribution of the wind speeds over the North Sea. The wind model is using the IEC61400-3 design regulation, where a turbulent wind field is modeled by mixing a mean wind and a fluctuating component. The turbulent wind is generated by using Turbsim [44] according to the Kaimal turbulence model including the turbulence intensity with IEC Class C. Similar to [13], to simulate the realistic wind fields, the turbulence intensity is a function of the wind speed at the hub height. In this study, two load cases (LCs) are employed. The parameters constant wind speed U_m, significant wave height H_s and peak-spectral period T_p are listed in Table 2.

The baseline control system consists of two individual controllers: a blade pitch and a torque controller. The blade pitch controller is designed as a proportional-integral (PI) controller, and the torque controller is based on a nonlinear state feedback function. Details of these controllers can be found in [45]. The block diagram of the closed-loop system is shown in Figure 2.

Similar to the widely used wind turbine fault detection and diagnosis benchmark model proposed by [46], a number of realistic faults including the actuator, sensor and component faults, are simulated. Table 3 lists the descriptions of these faults. Among these faults, faults 1–4 and 7-8 are sensor faults including scaling, offset and stuck; fault 5-6 are actuator faults such as stuck and offset. Fault 9, instead, is a component fault to simulate the cracking, debonding/delamination and fiber breakage of blades. The simulation of fault 9 is usually implemented as a reduction of the blade stiffness.

Variables including the measurements **y** and controller outputs **u** are selected from the closed-loop system (see Table 4) to build the monitoring model. All these variables are easily

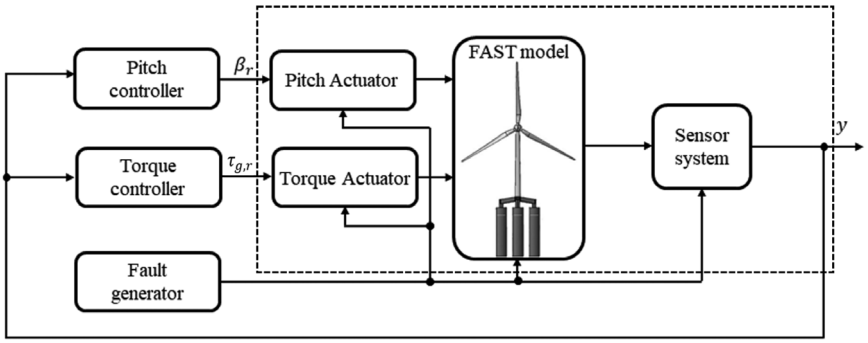

FIGURE 2 Block diagram of the closed loop system

TABLE 3 Descriptions of fault scenarios

Fault No.	Fault type	Description	Amplitude	Time
Sensor				
f_1	Scaling	Generator speed	0.95	185–210 s
f_2	Scaling	Generator power	1.1	240–265 s
f_3	Offset	Blade root bending moment	10^4KN·m	295–320 s
f_4	Scaling	Rotor speed	1.1	350–375 s
f_7	Stuck	Pitch sensor	0.2 deg	570–595 s
f_8	Scaling	Torque sensor	0.9	645–670 s
Actuator				
f_5	Stuck	Pitch actuator	0.2	420–445 s
f_6	Offset	Torque actuator	+20 KN·m	495–520 s
Component				
f_9	Scaling	Rotor blade sudden fault	0.2	700–1000 s

TABLE 4 Selected variables for fault detection

	Variable	Symbol	Units
u			
	Pitch angle reference	β_r	deg
	Generator torque reference	$\tau_{g,r}$	Nm
y			
	Generator power	$P_{g,m}$	W
	Generator angular speed	$\omega_{g,m}$	rad/s
	Pitch angle of i^{th} blade	$\beta_{i,m}$	deg
	Blade root moment of i^{th} blade	$M_{B,i,m}$	Nm
	Rotor angular speed	$\omega_{r,m}$	rad/s
	Generator torque	$\tau_{r,m}$	Nm

acquired by standard commercial SCADA systems as found in modern wind turbines. Additionally, a band-limited Gaussian white noise is added to the measurements. The specification of the band-limited Gaussian white noise can be found in [13].

The sampling interval for the selected variables is 0.01s. In each LC, the simulation duration is 1000s and the start time of these faults is listed in Table 3. For each LC, three data sets are generated: the first is produced under normal operating conditions for training fault detection model; the second is generated under 9 faulty scenarios to build the FDA model; the last data set is collected for validation. The wind power under normal operating condition is plotted in Figure 3 for LC1. The plot of wind power in LC2 is similar to LC1. Since the wind turbine start-up phase starts from 0 to 86s, all the data sets are collected from 86s for each load case. Thus, the training data contains 91400 samples which are collected under the fault-free condition.

5 | RESULTS AND ANALYSIS

For CCA-based methods, the input variables are the pitch angle reference β_r and the generator torque reference $\tau_{g,r}$, while the remaining ones are outputs. For PCA-based methods, all the variables are considered as observations. In this study, $q = 30$ is predefined for DCCA and RDCCA methods. The system order n is determined as 20 using the method in [32]. Furthermore, for simplicity, the regularization coefficients of RDCCA are set equal where $\alpha_1 = \alpha_2 = \alpha$. An appropriate α is selected where a lower false alarm rate (FAR) is obtained through grid search. The definition of FAR will be introduced later. Specifically, in the offline training phase, we use different α to train the RDCCA model. Then, several FARs are calculated. Figure 4 plots the FARs corresponding to different α for LC2. For RDCCA, the regularization coefficients are set as 10^{-5}.

Similarly, the regularization coefficients of RCCA are set as 10^{-2}. The threshold is determined according to Equation (16). In order to verify the fault detection performance, the existing PCA and dynamic PCA (DPCA) based wind turbine fault detection methods [25, 27] are employed. In terms of the reconstruction error, PCA and DPCA use the squared prediction error (SPE) statistic (Q statistic) for residual evaluation and threshold calculation. The number of principal components is determined such that the cumulative variance contribution rates exceed 99%. For all compared methods, the significance level is set as 0.01.

5.1 | Fault detection results

Figures 5 and 6 depict the monitoring charts for LC1 and LC2, respectively. For PCA, it can be found that T^2 and Q statistics

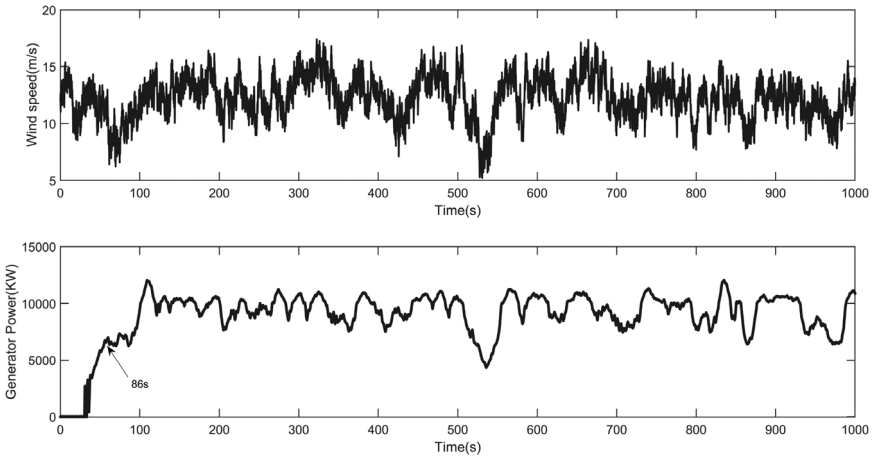

FIGURE 3 Wind power under normal operation condition in the case of wind speed of 12 m/s

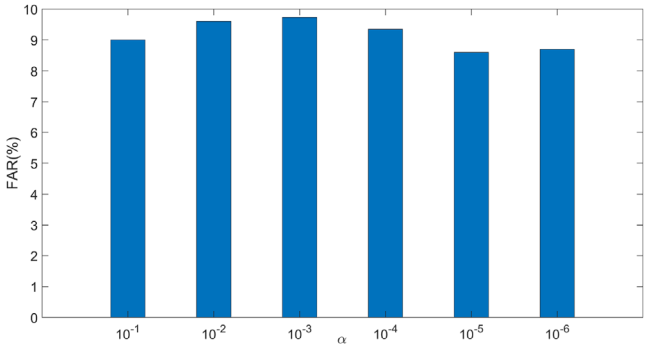

FIGURE 4 FAR with different α by RDCCA for LC2

can detect most faults, as shown in Figures 5a and 6a. As displayed in Figures 5b and 6b, the fault detection performance of DPCA Q statistic is improved, since the system dynamics is taken into account. For CCA-based method, T_r^2 statistic can provide similar performance as DPCA in LC1. Due to the introduction of regularization terms, false alarm conditions are reduced by RCCA as shown in Figures 5c and 6c. Compared to PCA, DPCA and CCA, DCCA can offer better performance as shown in Figures 5e and 6e. It can be concluded that the exploration of input–output relation and system dynamics can be beneficial for fault detection. Though more faulty samples can be detected, there are still many false alarms for DCCA-based method. As plotted in Figures 5f and 6f, RDCCA can provide similar fault detection performance as DCCA while reducing the false alarm by utilizing a regularization scheme. However, it is noticed that further improvements to detect Faults 3 and 9 are required, particularly for Fault 3. Despite the relatively poor fault detection performance for faults 3 and 9, RDCCA and DCCA have a better detection performance on these faults than other methods.

To achieve a quantitative performance evaluation, two indices are adopted. The first one is the fault detection rate (FDR). FDR is calculated as the ratio of successfully detected faulty samples to the total faulty samples. Another index is the false alarm rate (FAR), which is the ratio of the number of false alarms to the total number of fault-free samples. The definitions of FAR and FDR are as follows,

$$\text{FDR}(\%) = \frac{n_f}{n_t} \times 100,$$

where n_f denotes the number of faulty samples detected correctly by monitoring statistics and n_t is the total number of faulty samples.

$$\text{FAR}(\%) = \frac{n_{fi}}{n_{tf}} \times 100,$$

where n_{fi} denotes the number of fault-free samples incorrectly detected as faulty by monitoring statistics and n_{tf} is the total number of fault-free samples.

A good fault detection method should yield a high FDR and a low FAR. Table 5 lists the FDRs and FARs for all the fault scenarios in LC1 and LC2. Taking the results in LC1 as an example, the average FDR of DPCA Q statistic is 78.41% while for PCA Q statistic is 67.22% under LC1. On the other hand, it can be seen that the fault detection performance is improved by exploiting the relationship between input and output. For example, the average FDRs of static CCA T_r^2 statistic is 76.63%. DCCA method takes both the advantages of dynamic characteristics and input–output relationship. Thus, the average FDRs of DCCA T_r^2 statistic is 86.06%. However, the FARs of CCA and DCCA methods are also relatively high, that is, 3.38% and 6.25% for CCA and DCCA, respectively. The introduction of the regularization scheme can be beneficial to reduce false alarm rates. It can be seen that the FAR of RCCA is 2.20%. By taking both advantages of DCCA and regularization scheme, the RDCCA method can achieve the best fault detection performance: the FAR of RDCCA is 3.90% while its average FDR is 87.83%. For LC2, similar conclusions as in LC1 can be drawn.

5.2 | Fault classification results

After a fault alarm is triggered, the next crucial step is to identify the fault type for providing the operator with more information. In our study, the parameter b in Equation (19) is selected as 4. To

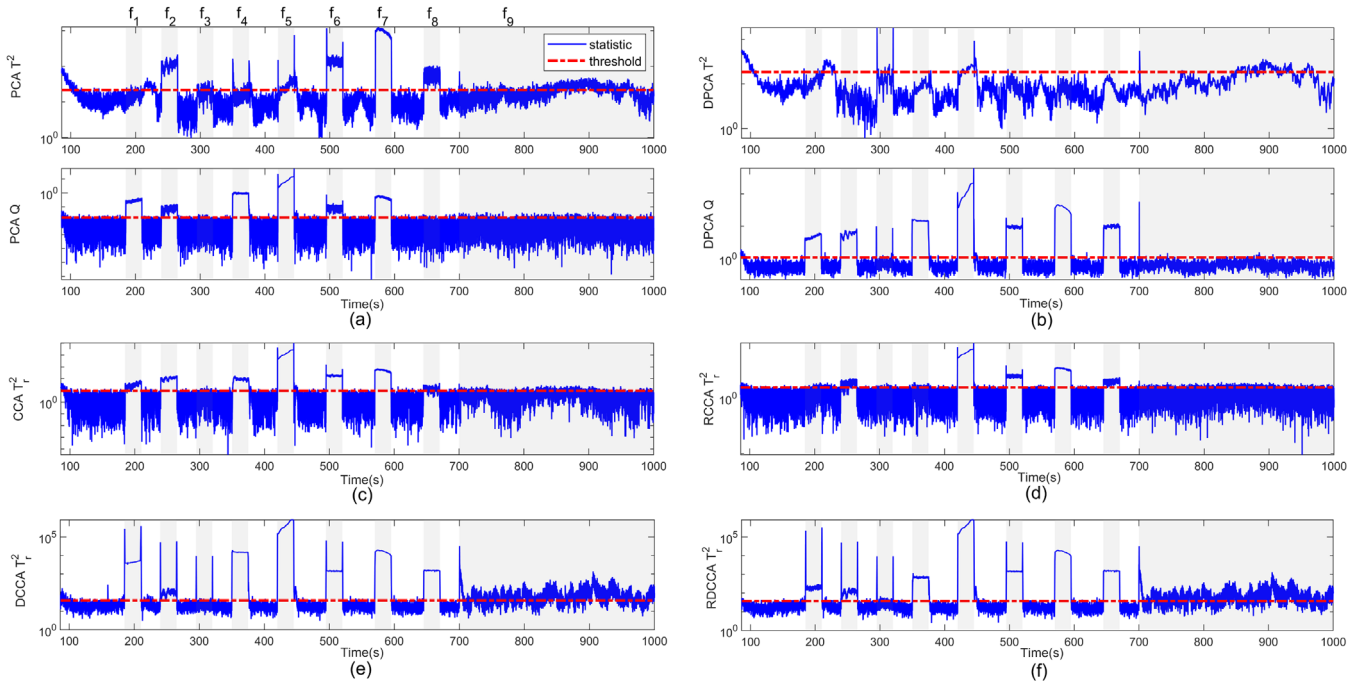

FIGURE 5 Fault detection result: LC1 (a) PCA (b) DPCA (c) CCA (d) RCCA (e) DCCA (f) RDCCA

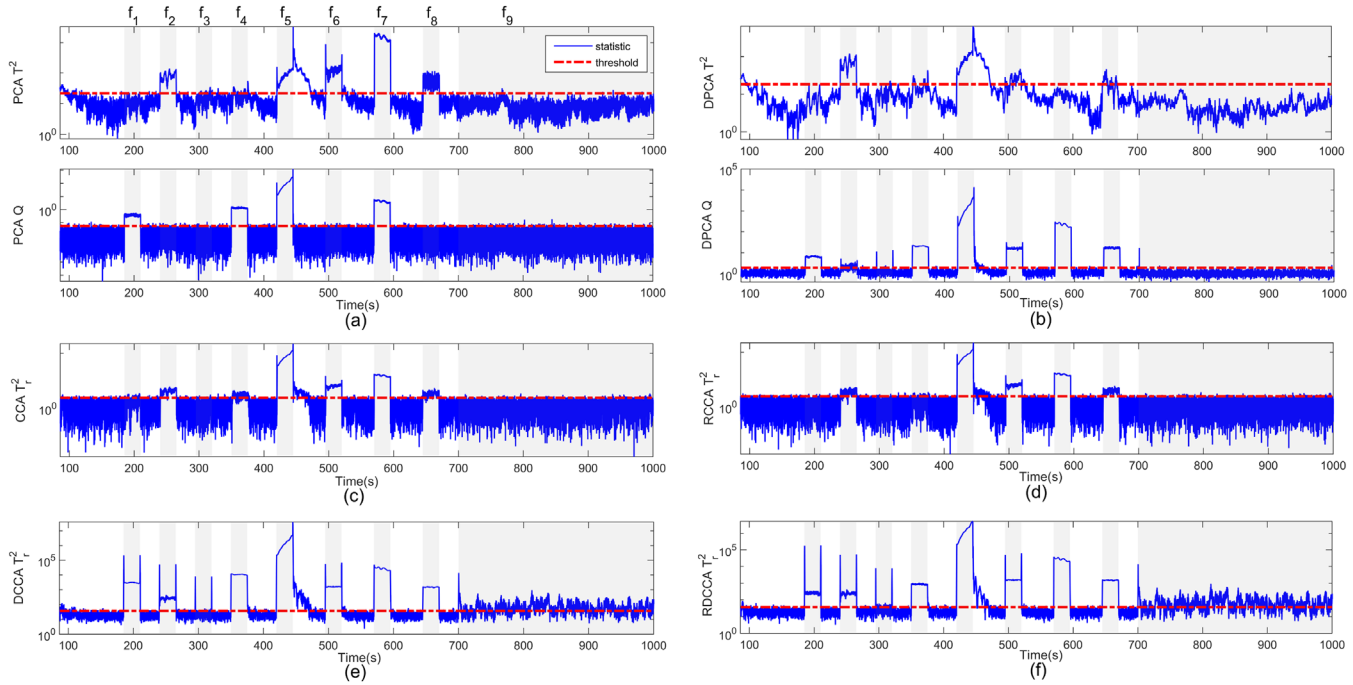

FIGURE 6 Fault detection result: LC2 (a) PCA (b) DPCA (c) CCA (d) RCCA (e) DCCA (f) RDCCA

visualize the discriminate performance of RDCCA-FDA, taking LC2 as an example, the first three FDA loading vectors are plotted in Figure 7. It can be found that the distance between different clusters of faults is large. Specifically, the clusters of Fault 1, Fault 4, Fault 5, Fault 6, Fault 7 and Fault 8 are far away from each other. Thus, these faults are easier to be classified. On the other side, the clusters of Fault 2, Fault 3 and Fault 9 are near each other. Error classification may occur as classifying Fault 2, Fault 3 and Fault 9. A similar situation exists in LC1.

Figures 8 and 9 display the confusion matrices in LC1 and LC2, respectively. A confusion matrix is designed to plot the true labels and predicted labels. Taking the results in LC1 as an

TABLE 5 Comparison of FDR (%) and FAR(%) with different methods

		PCA [27]		DPCA [25]		CCA	RCCA	DCCA	RDCCA
	Fault No.	T^2	Q	T^2	Q	T_r^2	T_r^2	T_r^2	T_r^2
FDR: LC1	1	1.64	99.96	1.48	99.88	99.28	8.28	100	99.96
	2	99.92	99.92	0.00	99.84	99.92	95.60	99.76	99.88
	3	24.00	1.80	19.88	5.32	4.20	2.60	5.52	15.44
	4	8.12	99.92	0.00	99.88	96.64	10.48	99.84	99.80
	5	60.68	99.84	65.60	99.84	99.84	99.84	99.88	99.88
	6	99.84	99.84	0.56	99.84	99.84	99.84	99.92	99.88
	7	99.92	99.92	0.00	99.88	99.92	99.92	99.92	99.92
	8	99.88	1.00	0.00	99.88	78.80	99.88	99.76	99.76
	9	32.73	2.76	26.99	1.32	11.25	6.17	70.00	75.98
	Average	58.53	67.22	12.72	78.41	76.63	58.70	86.06	87.83
FAR		9.73	1.80	9.82	3.12	3.38	2.20	6.25	3.90
FDR:LC2	1	0.24	99.96	0.04	99.88	13.56	1.48	99.96	99.96
	2	99.92	1.16	99.68	56.00	99.60	96.28	99.92	99.92
	3	5.12	0.88	5.84	1.96	1.44	1.00	6.28	17.68
	4	15.96	99.92	24.64	99.88	84.72	7.24	99.80	99.68
	5	99.76	99.84	99.44	99.84	99.84	99.84	99.88	99.88
	6	99.84	1.36	69.16	99.84	99.84	99.84	99.92	99.88
	7	99.92	99.92	0.00	99.88	99.92	99.92	99.92	99.92
	8	99.88	0.60	51.64	99.88	95.76	99.60	99.84	99.88
	9	0.08	0.69	0.00	0.35	0.91	0.93	65.07	65.77
	Average	57.86	44.92	38.94	73.06	66.18	56.24	85.62	86.95
FAR		10.13	1.00	10.44	3.19	6.85	5.83	11.43	8.60

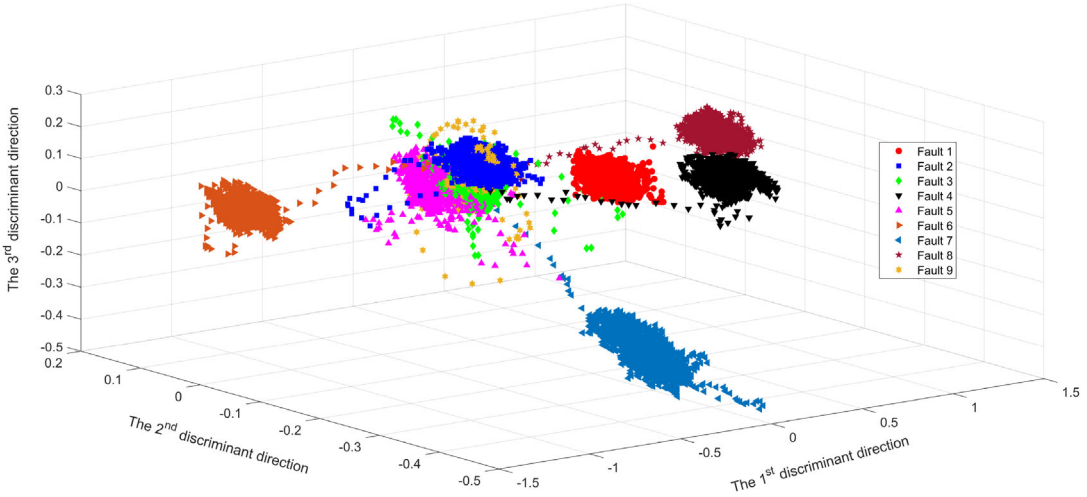

FIGURE 7 Projected training samples in Fisher discriminant subspace in the LC2: RDCCA-FDA

example, as shown in Figure 8, the rows correspond to the predicted class and the columns correspond to the true class. The values of the diagonal cells are the number of samples that are correctly classified. The off-diagonal cells correspond to incorrectly classified observations. For example, in the second row, the number of correctly classified samples is 2338 as the number of samples that belong to fault 2 is 2440. Thus, the percentage of samples belonging to fault 3 that are correctly classified is

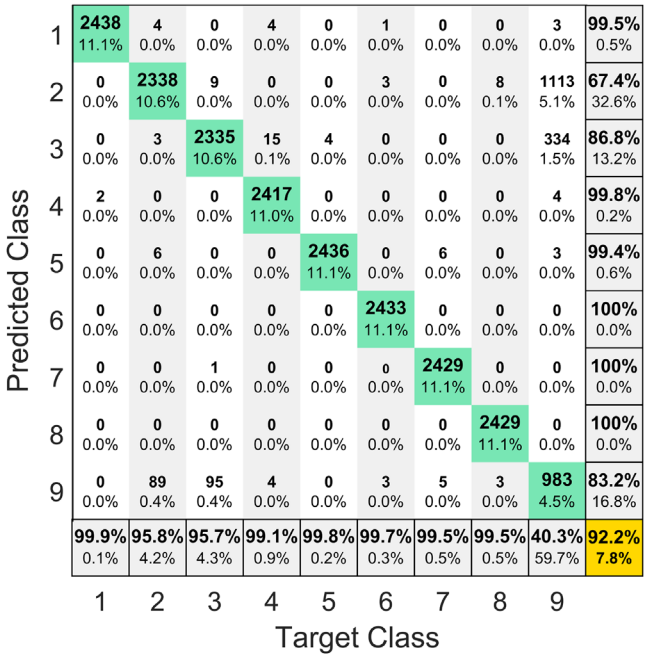

FIGURE 8 Fault classification result for the RDCCA-FDA method in the LC1

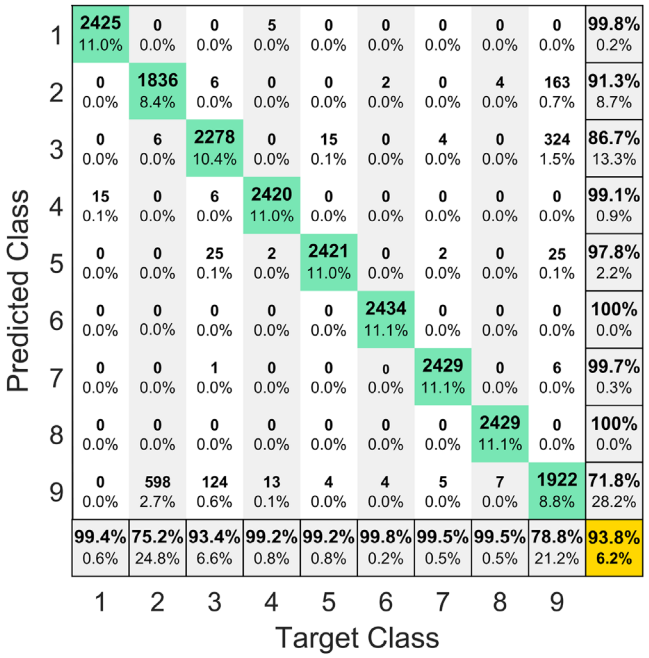

FIGURE 9 Fault classification result for the RDCCA-FDA method in the LC2

95.8%, and that are incorrectly classified is 4.2% as shown in the row at the bottom of Figure 8. On the other hand, the cluster of Fault 9 is closed to the one of Fault 2. 1113 samples which should be classified as Fault 9 are misclassified as Fault 3. The column on the far right of the plot shows the percentages of all the examples predicted to belong to each class that are correctly and incorrectly classified. These metrics are often called the Positive Predictive Value (PPV) and the False Discovery Rate (FaDR), respectively [47]. For example, in the column on the far right of the second row of Figure 8, the PPV and the FaDR are 67.4% and 32.6%, since some samples belonging to Fault 9 are misclassified as Fault 2. The overall accuracy is 92.2%, which is shown in the cell of the bottom right of Figure 8. For the LC2, the results are similar to the LC1 as shown in Figure 9. The overall accuracy is 93.8%. Nevertheless, RDCCA-FDA can provide superior fault detection and classification performances among comparable methods such as PCA and CCA.

To further verify the performance of RDCCA-FDA method, the residuals obtained from RCCA model are used instead of $\mathbf{r}_k$ of RDCCA model for comparison. Then, a new RCCA-FDA model is first trained. Similar to the RDCCA-FDA method, the projected training samples in Fisher discriminant subspace in the LC2 are plotted in Figure 10. Since the input $\mathbf{u}$ is only a two-dimensional vector, the residuals obtained from RCCA model are two-dimensional vectors. Thus, the dimensionality of Fisher discriminant subspace is set as 2.

It can be observed that the projected training samples are closed to each other. Therefore, the faulty samples are difficult to be classified to correct types. The confusion matrices of RDCCA-FDA in LC1 and LC2 are plotted in Figures 11 and 12. Through the data in Figures 11 and 12, the results of RCCA-FDA are not satisfactory where the overall accuracy is only 25.6% for LC1 and 20.3% for LC2. However, by considering the dynamics in the proposed RDCCA-FDA method, the performance can be highly enhanced.

6 | CONCLUSIONS

In this paper, a novel data-driven FOWT fault diagnosis approach is developed. The proposed RDCCA-FDA method utilizes regularized dynamic canonical correlation analysis to construct the residual generator for fault detection and feeds the residuals into the Fisher discriminant analysis model to classify the fault types. The performance of the developed method is validated on a FOWT baseline model based on the widely used NREL FAST. The simulation results show that the proposed RDCCA-FDA method can provide superior fault detection performance, by comparison with other relevant methods. Although there are misclassification errors of faults 2 and 9, RDCCA-FDA has achieved satisfying fault classification performance. More effort to correctly classify fault 2 and fault 9 should be made in future work. In addition, the reduction of false alarm rate should be studied in future work. Furthermore, the application of the proposed RDCCA-FDA method to real-world SCADA data collected from FOWTs should be investigated in future work.

ACKNOWLEDGEMENTS

This work was supported by the Zhejiang Province Public Welfare Technology Application Research Project (Grant Nos. LGF19F030004 and LGG21F030015), the Fundamental Research Funds of Zhejiang Sci-Tech University

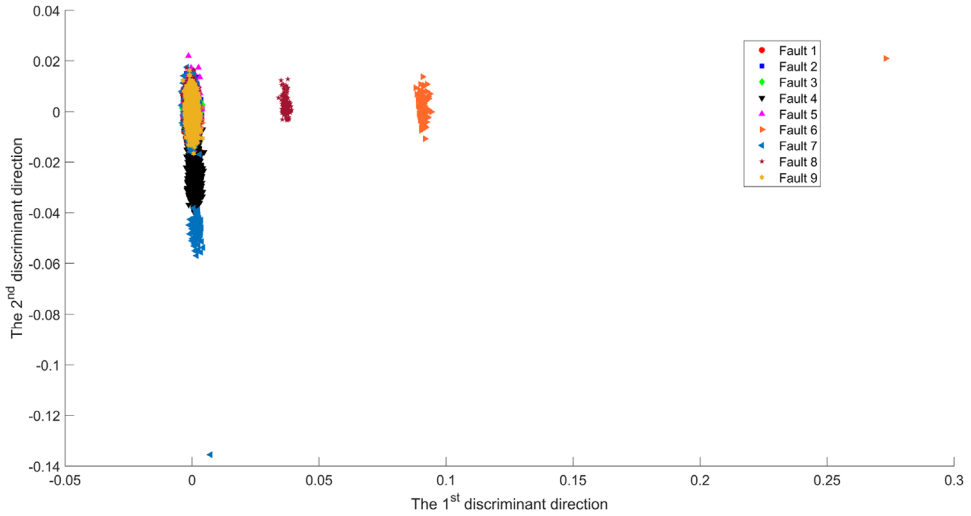

FIGURE 10 Projected training samples in Fisher discriminant subspace in the LC2: RCCA-FDA

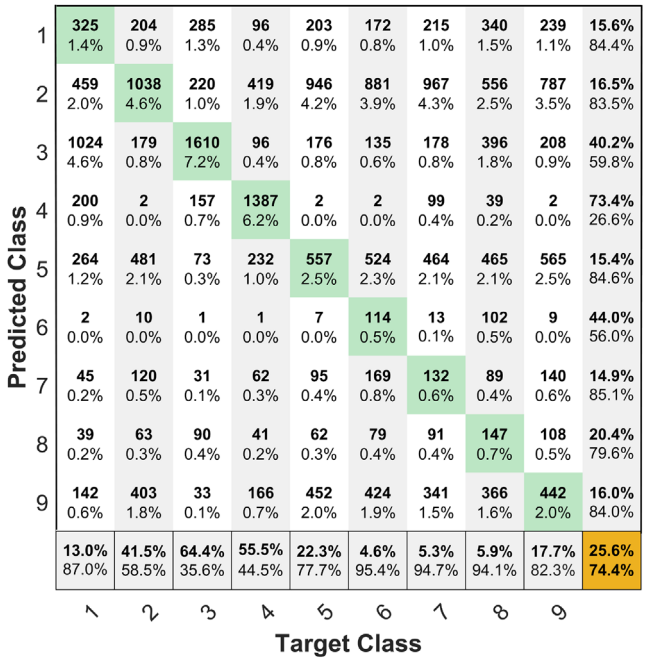

FIGURE 11 Fault classification result for the RCCA-FDA method in the LC1

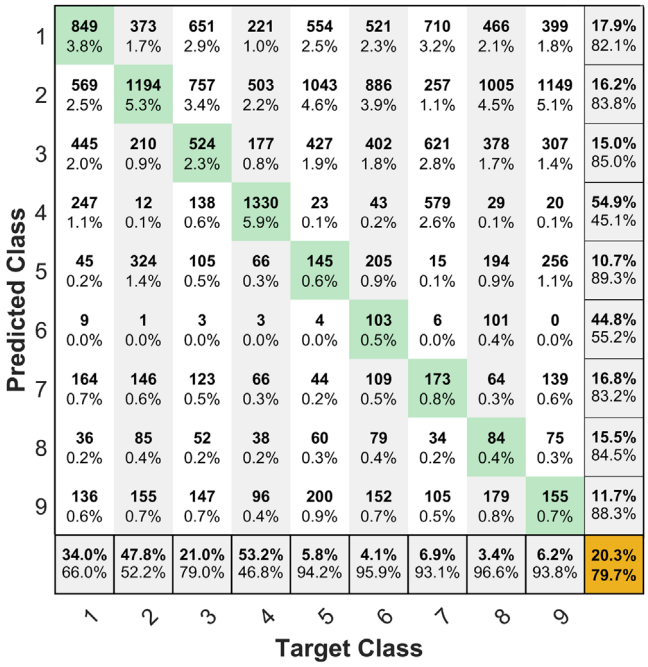

FIGURE 12 Fault classification result for the RCCA-FDA method in the LC2

(No. 2021Q024), and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (No. ICT2021B45), and the European Union via a Marie Sklodowska-Curie Action (Project EDOWE, Grant 835901).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Ping Wu https://orcid.org/0000-0002-2729-9669

REFERENCES

1. Ellabban, O., Abu Rub, H., Blaabjerg, F.: Renewable energy resources: Current status, future prospects and their enabling technology. Renew. Sustain. Energy Rev. 39, 748–764 (2014)
2. deCastro, M., Salvador, S., Gómez Gesteira, M., Costoya, X., Carvalho, D., Sanz Larruga, F.J., et al.: Europe, china and the united states: Three different approaches to the development of offshore wind energy. Renew. Sustain. Energy Rev. 109, 55–70 (2019)
3. Cruz, J., Atcheson, M.: Floating Offshore Wind Energy: The Next Generation of Wind Energy. Springer International Publishing, Switzerland (2016)
4. Bento, N., Fontes, M.: Emergence of floating offshore wind energy: Technology and industry. Renew. Sustain. Energy Rev. 99, 66–82 (2019)

5. de Novaes Pires Leite, G., Araújo, A.M., Rosas, P.A.C.: Prognostic techniques applied to maintenance of wind turbines: a concise and specific review. Renew. Sustain.e Energy Rev. 81, 1917–1925 (2018)
6. Márquez, F.P.G., Tobias, A.M., Pérez, J.M.P., Papaelias, M.: Condition monitoring of wind turbines: Techniques and methods. Renew. Energy 46, 169–178 (2012)
7. Hameed, Z., Hong, Y.S., Cho, Y.M., Ahn, S.H., Song, C.K.: Condition monitoring and fault detection of wind turbines and related algorithms: A review. Renew. Sustain. Energy Rev. 13(1), 1–39 (2009)
8. Qiao, W., Lu, D.: A survey on wind turbine condition monitoring and fault diagnosis-part i: Components and subsystems. IEEE Trans. Ind. Electron. 62(10), 6536–6545 (2015)
9. Dey, S., Pisu, P., Ayalew, B.: A comparative study of three fault diagnosis schemes for wind turbines. IEEE Trans. Control Syst. Technol. 23(5), 1853–1868 (2015)
10. Badihi, H., Zhang, Y., Hong, H.: Wind turbine fault diagnosis and fault-tolerant torque load control against actuator faults. IEEE Trans. Control Syst. Technol. 23(4), 1351–1372 (2015)
11. Cho, S., Gao, Z., Moan, T.: Model-based fault detection, fault isolation and fault-tolerant control of a blade pitch system in floating wind turbines. Renew. Energy 120, 306–321 (2018)
12. Simani, S., Farsoni, S., Castaldi, P.: Fault diagnosis of a wind turbine benchmark via identified fuzzy models. IEEE Trans. Ind. Electron. 62(6), 3775–3782 (2015)
13. Liu, Y., Ferrari, R., Wu, P., Jiang, X., Li, S., van Wingerden, J.W.: Fault diagnosis of the 10 mw floating offshore wind turbine benchmark: A mixed model and signal-based approach. Renew. Energy 164, 391–406 (2021)
14. Yin, S., Wang, G., Karimi, H.R.: Data-driven design of robust fault detection system for wind turbines. Mechatronics 24(4), 298–306 (2014)
15. Kusiak, A., Verma, A.: A data-driven approach for monitoring blade pitch faults in wind turbines. IEEE Trans. Sustain. Energy 2(1), 87–96 (2011)
16. Simani, S., Castaldi, P., Tilli, A.: Data–driven approach for wind turbine actuator and sensor fault detection and isolation. IFAC Proc. Volumes 44(1), 8301–8306 (2011)
17. Dong, J., Verhaegen, M.: Data driven fault detection and isolation of a wind turbine benchmark. IFAC Proc. Volumes 44(1), 7086–7091 (2011)
18. de Bessa, I.V., Palhares, R.M., DAngelo, M.F.S.V., Filho, J.E.C.: Data-driven fault detection and isolation scheme for a wind turbine benchmark. Renew. Energy 87, 634–645 (2016)
19. Pashazadeh, V., Salmasi, F.R., Araabi, B.N.: Data driven sensor and actuator fault detection and isolation in wind turbine using classifier fusion. Renew. Energy 116, 99–106 (2018)
20. Song, Z., Zhang, Z., Jiang, Y., Zhu, J.: Wind turbine health state monitoring based on a bayesian data-driven approach. Renew. Energy 125, 172–181 (2018)
21. Li, M., Yu, D., Chen, Z., Xiahou, K., Ji, T., Wu, Q.H.: A data-driven residual-based method for fault diagnosis and isolation in wind turbines. IEEE Trans. Sustain. Energy 10(2), 895–904 (2019)
22. Jiang, G., Xie, P., He, H., Yan, J.: Wind turbine fault detection using a denoising autoencoder with temporal information. IEEE/ASME Trans. Mechatron. 23(1), 89–100 (2018)
23. Ding, S.X.: Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems. Springer-Verlag, London (2014)
24. Cho, S., Jiang, J.: Optimal fault classification using fisher discriminant analysis in the parity space for applications to npps. IEEE Trans. Nucl. Sci. 65(3), 856–865 (2018)
25. Krüger, M., Ding, S.X., Haghani, A., Engel, P., Jeinsch, T.: A data-driven approach for fault diagnosis in gearbox of wind energy conversion system. In: 2013 Conference on Control and Fault-Tolerant Systems (SysTol), pp. 359–364. IEEE, Piscataway (2013)
26. Pozo, F., Vidal, Y.: Wind turbine fault detection through principal component analysis and statistical hypothesis testing. Energies 9(1), 3 (2016)
27. Wang, Y., Ma, X., Qian, P.: Wind turbine fault detection and identification through pca-based optimal variable selection. IEEE Trans. Sustain. Energy 9(4), 1627–1635 (2018)
28. Chen, Z., Ding, S.X., Zhang, K., Li, Z., Hu, Z.: Canonical correlation analysis-based fault detection methods with application to alumina evaporation process. Control Eng. Pract. 46, 51–58 (2016)
29. Zhu, Q., Qin, S.J.: Supervised diagnosis of quality and process faults with canonical correlation analysis. Ind. Eng. Chem. Res. 58(26), 11213–11223 (2019)
30. Jiang, Q., Ding, S.X., Wang, Y., Yan, X.: Data-driven distributed local fault detection for large-scale processes based on the ga-regularized canonical correlation analysis. IEEE Trans. Ind. Electron. 64(10), 8148–8157 (2017)
31. Chen, Z., Cao, Y., Ding, S.X., Zhang, K., Koenings, T., Peng, T., et al.: A distributed canonical correlation analysis-based fault detection method for plant-wide process monitoring. IEEE Trans. Ind. Inf. 15(5), 2710–2720 (2019)
32. Odiowei, P.P., Cao, Y.: Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. IEEE Trans. Ind. Inf. 6(1), 36–45 (2010)
33. Pilario, K.E.S., Cao, Y.: Canonical variate dissimilarity analysis for process incipient fault detection. IEEE Trans. Ind. Inf. 14(12), 5308–5315 (2018)
34. Liu, Y., Liu, B., Zhao, X., Xie, M.: A mixture of variational canonical correlation analysis for nonlinear and quality-relevant process monitoring. IEEE Trans. Ind. Electron. 65(8), 6478–6486 (2018)
35. Liu, Q., Zhu, Q., Qin, S.J., Chai, T.: Dynamic concurrent kernel cca for strip-thickness relevant fault diagnosis of continuous annealing processes. J. Process. Control 67, 12–22 (2018)
36. Bhowmik, B., Tripura, T., Hazra, B., Pakrashi, V.: Robust linear and nonlinear structural damage detection using recursive canonical correlation analysis. Mech. Syst. Sig Process. 136, 106499 (2020)
37. Chen, Z., Li, X., Yang, C., Peng, T., Yang, C., Karimi, H.R., et al.: A data-driven ground fault detection and isolation method for main circuit in railway electrical traction system. ISA Trans. 87, 264–271 (2019)
38. Li, X., Duan, F., Loukopoulos, P., Bennett, I., Mba, D.: Canonical variable analysis and long short-term memory for fault diagnosis and performance estimation of a centrifugal compressor. Control Eng. Pract. 72, 177–191 (2018)
39. Jiang, B., Zhu, X., Huang, D., Paulson, J.A., Braatz, R.D.: A combined canonical variate analysis and fisher discriminant analysis (cva–fda) approach for fault diagnosis. Comput. Chem. Eng. 77, 1–9 (2015)
40. Katayama, T.: Subspace Methods For System Identification. Springer, Heidelberg (2005)
41. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.R.: Fisher discriminant analysis with kernels. In: Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468), pp. 41–48. IEEE, Piscataway (1999)
42. Bak, C., Zahle, F., Bitsche, R., Kim, T., Yde, A., Henriksen, L.C., et al.: The dtu 10-mw reference wind turbine. DTU Wind Energy (2013)
43. Lemmer, F., Amann, F., Raach, S., Schlipf, D.: Definition of the swetriplespar floating platform for the dtu 10mw reference wind turbine. Stuttgart Wind Energy (SWE) (2016)
44. Jonkman, B.J., Buhl, M.L.: Turbsim User's Guide. National Renewable Energy Laboratory, Golden (2006)
45. Fontanella, A., Bayati, I., Belloli, M.: Control of floating offshore wind turbines: Reduced-order modeling and real-time implementation for wind tunnel tests. In: ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering, pp. V010T09A081–V010T09A081. ASME, New York (2018)
46. Odgaard, P.F., Stoustrup, J., Kinnaert, M.: Fault-tolerant control of wind turbines: A benchmark model. IEEE Trans. Control Syst. Technol. 21(4), 1168–1182 (2013)
47. Kohavi, R., Provost, F.: Glossary of terms. Mach. Learn. 30, 271–274 (1998)

How to cite this article: Wu, P., Liu, Y., Ferrari, R.M.G., van Wingerden, J.-W.: Floating offshore wind turbine fault diagnosis via regularized dynamic canonical correlation and fisher discriminant analysis. IET Renew. Power Gener. 15, 4006–4018 (2021). https://doi.org/10.1049/rpg2.12319