PHM Survey : Implementation of Diagnostic Methods for Monitoring Industrial Systems

Abdenour Soualhi[1], Bilal Elyousfi[1], Yasmine Hawwari[2], Kamal Medjaher[3], Guy Clerc[4], Razik Hubert[4] and François Guillet[1]

[1]*Univ Lyon, UJM-Saint Etienne, Saint-Etienne, France*
abdenour.soualhi@univ-st-etiene.fr
bilal.elyousfi@univ-st-etienne.fr
guillet@univ-st-etienne.fr

[2]*Univ d'Orlèans, Laboratoire PRISME, Orlèans, France*
y.hawwari@enim.ac.ma

[3]*ENIT - LGP, Tarbes, France*
kamal.medjaher@enit.fr

[4]*Univ Lyon, Univ Lyon-1, Laboratoire Ampère, France*
guy.clerc@univ-lyon1.fr
hubert.razik@univ-lyon1.fr

ABSTRACT

The modernization of industrial sectors involves the use of complex industrial systems and therefore requires condition based maintenance. This one aims at increasing the operational availability and reducing the life-cycle while increasing the reliability and life expectancy of industrial systems. This maintenance also called predictive maintenance is a part of an emerging philosophy called PHM 'Prognostics and Health Management'.

In this paper, the PHM will be emphasized on the existing diagnostic methods used for fault isolation and identification. This depicts an important part of the PHM as it exploits the data given by the signal processing step and its output is treated by the prognostic part. The diagnostic is mainly classified in three categories that will be highlighted in this paper.

1. INTRODUCTION

PHM was widely considered by industrials as an approach for the health management of systems. Their main duties are the detection of incipient components, failure mode identification and isolation, failure prognostics and health management (Atamuradov et al., 2018). PHM allows maintenance to be performed more efficiently by integrating PHM data (eg. the RUL "Remaining Useful Life") from:
- The health assessment derived from detection, maintenance data (resources and inventories), logistical constraints and information on the planned mission.
- The location and identification of failures resulting from diagnostic.

- The prognostic of the state.

The PHM is composed of a key milestone conditioning its success: the diagnostic. This latter is defined as a process of detection and location of faults. The diagnostic represents the information obtained by sensors in a space composed by features. These features generate a representation space allowing the identification of faults (Jardine et al., 2006). The diagnostic of faults, while the system is running, avoids unplanned outages and at the same time reduces productivity losses. Therefore, the diagnostic of faults is a research topic that is of interest for the industry and has been the subject of extensive research to develop more effective methods. The diagnostic step can be modeled as described by fig. 1 (Elghazel et al., 2015). First, the fault is detected according to an anomaly observed on the system. Then, this fault is isolated so that the source of fault is defined.

| Fault detection | Fault isolation | Fault identification |

Figure 1. Flowchart of diagnostic.

Diagnostic methods studied in (Alhelou, 2019; Gao et al., 2015; Gertler, 2017; Hwang et al., 2010; Skliros et al., 2019) can be classified into two categories: methods based on a physical model and data-driven methods. Besides, (Atoui et al., 2016; Benkouider et al., 2012; Ghosh et al., 2011; Lin et al., 2004; Maurya et al., 2007; Siswantoro et al., 2016; Zhao et al., 2013) highlighted a new family called hybrid methods in order to optimize the diagnostic performance. This new family is described in section 4.

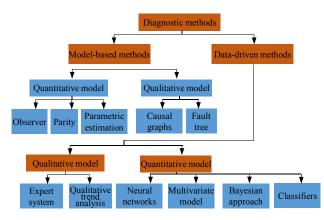The hierarchy of diagnostic approaches is shown in fig. 2.

Figure 2. General classification of diagnostic methods.

Model based methods can be classified as qualitative or quantitative. The model is typically developed on the basis of the representation of the behavior of the system. In quantitative models, this understanding is expressed in terms of mathematical relationships between the inputs and outputs of the system. In qualitative models, these relationships are expressed in terms of qualitative functions. For the database methods, only the availability of a large amount of historical data on the system is assumed. These can be processed and presented as a priori knowledge to the diagnostic. This is the feature extraction process from the historical data of the system. This initial analysis will optimize the diagnostic process. When it is possible to model the system with a lack in historical data, it is more suitable to use Model-based models (Gao et al., 2015).

2. MODEL BASED METHODS

2.1. Quantitative methods

2.1.1. Observers

The diagnostic based on observers is applied on an observable system which means that their internal states can be deduced from their outputs. Furthermore, it is necessary to identify the system parameters. The principal objective of observer-based diagnostic is the generation of a set of residuals allowing the detection and uniquely the identification of different faults (Alhelou, 2019). The differences between the measured output of the system and the values estimated by the model are used as a residue. A residue near to zero represents a system in a good health. However, a residue distant to zero indicates a discrepancy between the system output and the output predicted by the model. This difference allows detecting a fault in the system. Figure 3 shows the conceptual diagram of a residue generator based on an observer. Where y_K is the measured output, $\hat{y}_K$ is the estimated output, U_K is the system input and ε_K is the residue calculated at each discrete step time (K).
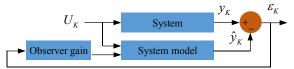

Figure 3. Residues generator based on observer.

Different observers have been developed in literature. For instance,(Eissa et al., 2019; Jlassi et al., 2017; Palak, 2017), used Luenberger observer introduced by Luenberger in 1966 (see fig.4), where the observer can be presented by the following equations:

$$\begin{cases} \hat{X}_{K+1} = A_d \hat{X}_K + B_d U_K + K_{obs} \varepsilon_K \\ \hat{Y}_K = C_d \hat{X}_K \end{cases} \tag{1}$$

With $\hat{X}$ is the estimated state of the system and A_d, B_d and C_d are respectively the state matrix, the control matrix and the measurement matrix. K_{obs} is the gain of the observer. The observer is qualified as asymptotically stable if the observer error $e(K) = \hat{X}_K - X_K$ tends to zero when K tends to infinity.
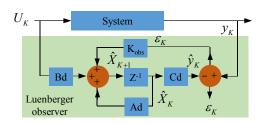

Figure 4. Structure of Luenberger observer.

Another kind of observer was introduced by Kalman in 1996 which is the Kalman filter. This observer is used for stochastic systems and consider the model uncertainty and the measurement noise by the following equation:

$$\begin{cases} \hat{X}_{K+1} = A_d \hat{X}_K + B_d U_K + W_K \\ \hat{Y}_K = C_d \hat{X}_K + V_K \end{cases} \tag{2}$$

Where W_K (the process noise) and V_K (the measurement noise) are white noises and supposed non-correlated. (Saravanakumar et al., 2014) proposed a bank of Kalman filter called dedicated observer scheme (DOS) and generalized observer scheme (GOS). The DOS and GOS detect, respectively, simultaneous and multiple sensor faults.

In DOS, each observer considers one sensor output corresponding to the fault supposed to be detected (Alrowaie et al., 2014; Dey et al., 2015). A dedicated residual set is defined as: $r_i \in \{1,2,…,M\}$ with M is the number of faults to be detected. Then, residues undergo a thresholding operation in order to ensure fault occurring and avoid false alarm : $r_i(t) > \xi_i$ where ξ_I are the thresholds for each residue r_i (Idrissi et al., 2017). For GOS, each observer considers all the sensors output except the monitored sensor.

Kalman filter (KF) is an optimal observer for linear systems and Gaussian disturbances. However, when the system is nonlinear and when the noise is still Gaussian, KF may not converge. To avoid this drawback, the extended Kalman filter (EKF) was proposed in (He et al., 2015). The EKF was used in (Yin et al., 2016) to diagnose the stator winding fault in the brushless wound field synchronous generator using the estimated rotor currents. Another alternative was proposed in (Ariola et al., 2015). It consists of developing a scheme based on an augmented stage Kalman filter (ASKF) with a bank of integrators. This model is composed by one observer and a bank of integrators where each integrator is related to the sensor fault detection. When system uncertainties and noise disturbance are absent and all sensors are normal, the integrator outputs are equal to zero. Furthermore, when a fault occurs in the *i-th* sensor, only the *i-th* integrator changes its state which is presented by y_i where $y_i=(y_1,…, y_i,….y_n)$ are the outputs of the integrator bank. Moreover, the observer gain was calculated by linearizing the system model and is kept constant.

More observers have been introduced to alleviate this drawback and can be applied to nonlinear systems and non-Gaussian distributions such as particle filtering (PF) which is adaptable to any nonlinearity and any noise distribution. In (Souibgui et al., 2011), the authors presented sequential Monte-Carlo technique implemented in the PF in order to approximate the optimal filtering with sufficient number of samples. This method, used in (Saravanakumar et al., 2014) for sensor fault diagnosis, was also used in (Chen and Saif, 2007) for actuator fault diagnosis.

2.1.2. Parity space

Parity space (PS) method was designed, first, for static systems. Then, it was extended to dynamic systems (Blesa et al., 2016). This approach consists of transforming the state-space model of the system to generate analytical redundancy relations (static or dynamic) among the inputs and outputs of the system considered as measured variables. Analytical redundancy relations are obtained by eliminating unobservable model state variables. Analytical redundancy relations are then used to construct fault indicators for diagnostic. Methods based on the parity space and observers are similar as shown in researches of (Hafaifa et al., 2015; Jiang et al., 2018; Laamami et al., 2015; Vento et al., 2015; Zogopoulos-Papaliakos and Kyriakopoulos, 2019). Figure 5 gives a conceptual scheme of the space parity for a residue generator (Ding et al., 2011).
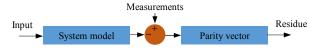

Figure 5. Residue generator by parity space.

In (Yin et al., 2014), a scheme was developed for fault detection in wind turbines. This scheme was proposed in (Ding et al., 2010) and consists of four steps. Initially, the parity space is directly designed from the measurements. Then, the parity vectors are determined by a performance index and an optimization criterion in order to make the residue generation independent of disturbances and sensitive to sensor faults. After this, *n* independent residual signals using observers are generated. Finally, the residuals signals are evaluated and compared with a threshold in order to be the input of decision logic.

2.1.3. Parameter estimation

Methods of parameter estimation were proposed by (Capolino et al., 2015; Jiang et al., 2008; Kim et al., 2013), for the detection and diagnostic of faults. It is based on the assumption that faults affect the physical parameters of the system. The occurrence of a fault can be detected by comparing estimated parameters with the real parameters of the system. Any difference between the two gives an indication about the presence of a defect. Figure 6 shows the conceptual diagram of the parametric identification for diagnostic. In this method, deviations (ε_K) from the reference model are used for faults detection and isolation, which can be also useful for fault analysis (Gertler, 2017).
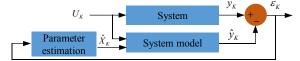

Figure 6. Schematic description of the parameter estimation.

(Ding et al., 2016) presented a parameter estimation model applied to nonlinear systems and proposed a recursive least squares (RLS) algorithm. This algorithm aims at estimating parameters recursively by minimizing the square error between the observed and the estimated outputs of the system. Wang pointed out the parameter estimation problem in block-oriented Hammerstein systems (Ding et al., 2016), which are composed of a nonlinear static block followed by a linear dynamic block with multi-input multi-output structure. (Xu et al., 2015) developed a method for parameters estimation applied to transfer function of a second order. This method is based on Newton iterative algorithm based on minimizing a defined cost function. Results showed that this method gives satisfactory results.

2.2. Qualitative methods

2.2.1. Causal models

The oriented graph allows connecting the fault causes (x_i) to their symptoms (x_j) by relations of cause and effect. The diagnostic consists of locating all possible sources of disturbance by online sensors data. The cause-effect relationships can be represented in the form of signed directed graphs (SDG). A directed graph is a graph with directed arcs between nodes (x_i, x_j). A SDG is a graph where nodes and variables are assigned by a sign (+, -, 0). These signs are obtained by comparing them to thresholds depending on

whether it is a normal value (0), too high (+) or too low (-). SDG is also represented by arcs related by signs (+,-). These signs express positive or negative relations between nodes (Yang et al., 2012). Methods based on directed graphs are attractive because little information is needed to establish a network of directed to make the diagnostic. In fig. 7, two examples of SDG are presented: in a) when x_i is getting high, x_j will decrease and in b) when x_i is getting high, x_j will not change its value suddenly but will decrease gradually and at a point of time it will increase.

Signed oriented graphs (SOG) are built from nodes and directed branches representing the paths of causality in a system without faults. Nodes of the SOG correspond to state variables, state of alarm, or the origins of defects. It is assumed that a fault affecting one of the root nodes of the SOG is the source of all disturbances. It is also assumed that the defect does not affect other pathways in the OG.

a)

$$x_i \xrightarrow{\quad - \quad} x_j \qquad x_i \xrightarrow{\quad (0,-,+) \quad} x_j$$

b)

Figure 7. Two examples of link in SDG.

Recently, (Liu et al., 2016) introduced a method based on SDG for fault detection and diagnostic applied on nuclear power plants. It consists of combining a flow chart to the experience knowledge to solve online fault deduction. A status matrix is calculated and implemented to simulate the fault propagation path and its cause. Other works for fault diagnostic based on SDG can be found in (Gao et al., 2010; Hua et al., 2011).

2.2.2. Fault Tree

A fault tree describes all combinations of events leading to a fault. It translates the behavior of a fault in a visual diagram where a set of rules, logic and symbols provides a mechanism for analyzing complex systems. Figure 8 shows the basic structure of a fault tree for the diagnostic of a fault. The upper portion (also called top event TE) of the tree is the defect to be analyzed. The lower part (called basic events BE) is connected to the upper through intermediate events (usually called nodes). This part shows how the defect is propagated in the system part. Events in a level of a fault tree are connected to the events of the next level through logic gates (AND, OR, etc.).
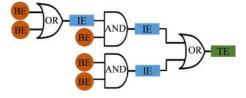

Figure 8. Basic structure of a fault tree.

Recently, (Nouri.Gharahasanlou et al., 2014) used the fault tree analysis to assess the probability of a failure occurrence for certain equipments in the Azarabadegan Khoy cement plant. The study showed that there is a probability of a failure for crushing, conveyor systems, crushing and mixing bed hall department equivalent to 73, 64 and 95 percent respectively. In the same year, a study of Purba has been published in (Purba, 2014) about the application of the fault tree analysis for nuclear power plant probabilistic safety assessment. This analysis introduces a fuzzy approach to estimate failure probabilities from insufficient historical data. These probabilities are modeled by fuzzy membership functions. This method showed very satisfactory results by applying the fault tree in the case of insufficient historical data.

Lately, Ruijters and Stoelinga established in (Ruijters and Stoelinga, 2015) the state-of-the-art in the modelization, analysis and tools of fault tree analysis. The fault tree analysis was presented in its qualitative and quantitative use and some of their extends were also presented.

3. DATA DRIVEN METHODS

Contrary to model based approach where a priori knowledge on the system is necessary, in the data-driven approach, only the analysis of the recorded data is required. The objective of this approach is to develop diagnostic methods without modeling the system. Data-driven methods can be either qualitative or quantitative.

Two of the main methods, which extract qualitative information, are expert systems and qualitative trend analysis. Methods that extract quantitative information may be statistical or non-statistical methods. Nowadays, data mining is one of the most active areas of research. The main advantage of fault diagnostic by data mining is the automatic generation of concise and accurate diagnostic models from historical data.

3.1. Qualitative methods

3.1.1. Expert system

The expert system, as shown in fig.9, is defined by Angeli (Angeli and Chatzinikolaou, 2004) as "an intelligent computer program using the knowledge base and an inference engine to solve the problems from human expertise". This knowledge base is combined to a rule base indicating the way of applying this knowledge to problems. These rules are IF-THEN statements. Many researchers worked on the application of expert systems for diagnostic problems. (Saludes et al., 2003) proposed an integrated framework including a SOM (self-organizing map) neural network subsystem and an expert system for fault detection and isolation in a hydroelectric power station. The expert system was developed in cooperation with technicians and engineers of the power station. The neural network subsystem was trained with data collected from the power station over one year. The expert system and the neural networks subsystem showed a good complementarity, where expert sys-

tem detects already known faults. Neural network subsystem was able to detect unknown faults. The advantage of an expert system is its simplicity of implementation. However, it fails to detect new faults without known signatures and needs continuous updating (Guo et al., 2019).
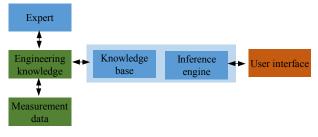

Figure 9. Architecture of an expert system.

3.1.2. Qualitative trend analysis

The qualitative trend analysis (QTA) is used to explain various important events occurred in a system. It uses information about trends present on sensor measurements via two steps: (i) the identification of trends in the measures and (ii) interpretation of trends in terms of fault scenarios. This allows fault diagnostic and prediction of future states. Many methods have been developed to represent the tendency of a system. Among them, the representation by triangular episodes and identification of primitives (episodes) by neural network and trends analysis by wavelet (Villez, 2014).

Very recently, Zhou and Ye developed a novel method of QTA called polynomial fit-based methods for qualitative trend analysis in (Zhou and Ye, 2016). This method consists of complementing the model of trend analysis developed by (Dash et al., 2004) with four polynomial fit-based trend extraction algorithms aiming at automatically estimate the parameters such as thresholds and noise level.

3.2. Quantitative methods

3.2.1. Multivariate statistical methods

Applications of multivariate statistical methods for the diagnostic of defects such as principal component analysis (PCA) have been widely reported in the literature. An overview of the use of PCA in the fault diagnostic has been given by (Ding et al., 2010; Jing and Hou, 2015; Yin et al., 2014). The PCA, as shown in fig. 10, is a method of feature reduction. It minimizes the sum of squared error (Euclidian distance) between the observations and their projections (Harmouche, 2014). It transforms a set of *n*-initial features (health indicators) into a reduced set of uncorrelated features, which explain most of the variations of the system by using tools of linear algebra (Eigen value decomposition) and statistical analysis (variance-covariance) to construct a representative model that can be used for the diagnostic of faults (Sliskovic et al., 2012). In (Pontoppidan and Larsen, 2003), a novel method for health monitoring applied on

diesel engines have been developed. This method uses Independent Component Analysis (ICA) and probabilistic anomalies detection algorithm. By considering *q* independent and non-Gaussian vibration sources and *N* sensors that give *N* output signals recorded in the same time, the ICA dissociates independent sources linearly mixed in these signals by estimating the mixing ratios of the source signals in the recorded observations (Zuo et al., 2005). Another multivariate statistical method has been developed in (Wang and Nee, 2009) and (Duda et al., 2012) is the Fisher Linear Discriminant. This method reduces the dimensionality by seeking a projection that best describes the data according to least square error in order to keep the judicious information for discrimination.
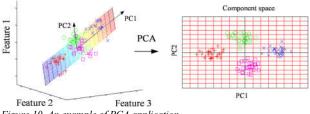

Figure 10. An example of PCA application.

3.2.2. Neural network

Neural networks (NN), used for fault diagnostic, can be classified according to the network architecture and the learning method (Liu et al., 2018). There are several papers dealing with the problem of fault diagnostic using neural networks. They can be applied to fault diagnosis using different approaches. Pattern recognition approach and residual generation decision making are the most common ones (Srivastava et al., 2014). Neural networks strength lies in the ability to combine multiple non-linearities to mimic virtually any input-output map without requiring the physical model of the system (Zappone et al., 2019).

From an architecture point of view, the NN can have several forms. In (msaaf and Belmajdoub, 2015), authors presented a comparison between the most common NN architectures used in diagnosis, namely, multi-layer (MLP) perceptions, radial basis function (RBF) networks, Kohonen chart (self-organizing map) and the random vector functional link (RVFL) network.

First, the neuron is a mathematical modeling of the natural neural as described in fig.11 where the output (y_j) is given by:

$$y_{j \atop j \in \{1,2,...,N\}} = \varphi\left(\sum_{i=1}^{n} \omega_i x_{ij} + b\right) = \varphi\left(W_j^T X_j + b\right) \quad (3)$$

Where y_j is the output (health status) corresponding to the *j-th* input vector $X_{j\,j \in \{1,2,...,N\}}$. N is the number of input vector. W denotes the vector of weight ω_i, n is the number of health indicators. $x_{ij\,i \in \{1,2,...,n\}}$ is the value of the *i-th* health indicator of the input vector X_j. b is the bias and φ is the activation function.
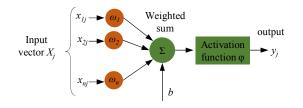
Figure 11. Simple neuron.

When the activation function is a sign function, then this form is called "perceptron" (SP), introduced by Rosenblatt in (Rosenblatt, 1958). By affecting a set of SP to successive layers, namely: input layer, hidden layers and output layer, we obtain the multilayer perceptron (MLP) (see fig. 12).
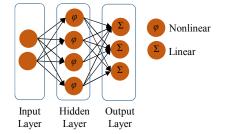

Figure 12. An example of the MLP architecture.

The computations performed by the SP with a single hidden layer of nonlinear activation functions and an output layer with linear functions can be written mathematically as follows:

$$y_{j \atop j \in \{1,2,...,N\}} = B\varphi\left(AX_j + a\right) + b \qquad (4)$$

Where X_j is an input vector of n-health indicators and y_j is the output vector. A is the matrix of weights of the first layer, a is the bias vector of the first layer. B and b are respectively, the weight matrix and the bias vector of the second layer. The function φ denotes an element wise nonlinearity.

The third architecture is the RBF developed in (msaaf and Belmajdoub, 2015) and introduced in (Broomhead, 1988). It consists of a neural network composed of typically three layers (see fig. 13). A radial basis function is used as an activation function. The output of the network is a linear combination of these functions, the inputs and neural parameters.

The most common activation function of each input is the Gaussian function:

$$RBF_i = e^{-\frac{|x_{ij} - \mu_i|}{2\sigma_i}} \qquad (5)$$

where μ_i is the center of the hidden neuron and σ_i is the standard deviation of the hidden neuron.

When applying the activation function RBF_i, it is noticeable that if the input vectors are exactly the center of the class then the function returns 1 and if the distance between them increases, the output decreases from 1 to 0.
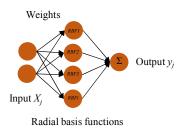

Figure 13. RBF network.

This architecture has the advantage of locally representing the space of learning towards centers and also thanks to its nonlinear approximation property. Because of their universal approximation, more compact topology and faster learning speed, RBF networks attracted considerable attention and they are widely applied in many science and engineering applications (Kurban and Beşdok, 2009).

Self-Organized Map (SOM) was used for fault detection and health monitoring of aircraft engines and bearing systems (Moshou et al., 2010). SOM or Kohonen map is an artificial neural network using low dimensional representation of the input space called a map (Cottrell et al., 2009). The theory of this architecture was presented in (Kohonen, 2001) by Kohonen. This consists of the training (setting up the map) and mapping (classifying a new input). SOM is composed of two layers, the input layer and the output layer. This layer is a two-dimensional space where each input of the vector X_j is connected to all the output neuron within weights noted $\omega_{\ell i}$. $\omega_{\ell i}$ is the weight between the i-th input (x_{ij}) of the vector X_j and the ℓ-th neuron. This architecture is used for unsupervised classification. This means that for a space of input elements, the SOM is able of grouping them in classes. (Urmos et al., 2013) presented the fourth steps of establishing SOM. In the initiation step, random weights are affected to p defined connections. Then, in the competition step, for an input $x_{ij\ i=\{1,...,n\}}$ and a weight $\omega_{\ell i}$ affected to the connection between the i-th input and ℓ-th neuron ($\ell=1,...,p$), a discriminant function for each neuron is given by:

$$d_\ell(X) = \sum_{i=1}^{n}(x_{ij} - \omega_{\ell i})^2 \qquad (6)$$

which means that the chosen neuron has the nearest weight vector to the input vector. In the cooperation step, a decreasing exponential function, depending on the distance between the chosen neuron and the other neurons, is defined. This takes into account the fact that when a neuron is chosen, their neighbors are also chosen with a topological neighboring that deceases exponentially in function of distance. In the adaptation step, as the neurons are self-organizing, the winner weight changes to resemble more to the sample vector and the neighbors weights change also but not in the same degree (see fig. 14).
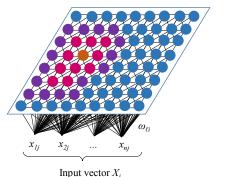
Figure 14. The winner neuron (beige), the stimulated neighboring neurons (red).

Neural networks can be used with other methods like the empirical mode decomposition. In (Ben Ali et al., 2015), it has been presented a novel method which consists of applying the EMD method to vibration signals of rolling element bearing and then selecting the most important intrinsic mode functions (IMFs) which have been used as input for the ANN. In this paper, it was proven that ANN is one of the most effective technique for automatic degradation assessment of bearing. The ANN and the support vector machine (SVM) were used for fault diagnosis in (Yang et al., 2011). This method uses vibrations signals of a gearbox, extracts the pertinent features and then, a comparison is established between applying ANN and SVM which demonstrates that SVM is more effective than ANN.

3.2.3. Classification

The diagnostic of faults can be obtained by using the classification approach (Casimir et al., 2003a, 2003b; Ondel et al., 2006; Soualhi et al., 2011). The goal is to classify measured data called observations ($X_{j\,j=\{1,..,N\}}$) into classes (Ω_1, Ω_2, ..., Ω_M) corresponding to the health status of the system (see fig. 15). Where M is the number of classes and N is the number of observations. An observation $X_j=$ [x_{1j}, x_{2j}, ..., x_{nj}] is composed by a set of n-features (faults indicators) $x_{ij\ i=\{1,..,n\}}$. These features allow constructing a representation space allowing to distinguish patterns (classes). In the case of n-features, the observation X_j is defined in an n-dimensional space.
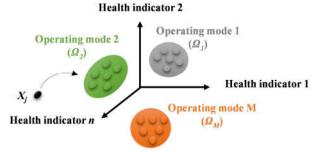

Figure 15. Classification of the observation x_i to one of the classes ($\Omega_1,\Omega_2,...,\Omega_M$).

The classification problem was approached by pattern recognition methods (*K*-means, nearest neighbor…) in (Baraldi et al., 2016; Glowacz and Glowacz, 2017), and neural networks recently (Jia et al., 2018). There is two types of classification: supervised classification and unsupervised classification.

The supervised classification deduces a function from a training set of samples. Each sample is a pair of an input (an observation X_j) and a label noted "o" (the class membership of the observation X_j). A supervised learning algorithm is then used to produce a membership function between X_j and o. Then, this membership function will be used to classify new observations. Rejection options can be implemented in order to avoid ambiguous classification.

The unsupervised classification analyses observations regardless of their labels to highlight classes with similar observations. The "clustering" is a typical example of the unsupervised classification where observations are grouped into classes based on similarity measure between observations. Supervised and unsupervised classification are two ways to solve the problems of pattern recognition. The choice of the classification type depends on the availability of the training set. If a priori knowledge of class membership is available, i.e. the health status of the corresponding system, we talk about supervised classification. Otherwise, we talk about unsupervised classification.

In (Younus and Yang, 2012), Younus and Yang presented an intelligent fault diagnosis (IFD) of rotary machinery by using features processed and extracted from infrared thermal image. The IFD uses support vector machines (SVM) and linear discriminant analysis (LDA) methods as classifiers. In the field of machine learning, SVM are supervised classification technique drawing a boundary between two classes in order to maximize the margin between them. The points employed to determine this distance are called Support Vectors (SV) (see fig. 16). A kernel can be used to classify nonlinear separable classes.

Recently, Lu and Qiao proposed in (Lu and Qiao, 2013) a novel method of SVM-based classification used for fault diagnosis of drivetrain gearboxes functioning in variable speed conditions. The experimental results of a gear crack fault demonstrate the success of this method for fault identification.
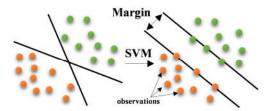

Figure 16. SVM applied on a set of observations.

In the range of the unsupervised classification, many partitioning algorithms were developed. They consist of defining gravity centers representing the centers of classes and then build classes by affecting each observation to the appropriate class. The affectation in the case of fuzzy clustering is done by a membership function indicating the degree of similarity of an observation to each class. The fuzzy *c*-means clustering algorithm which belongs to the C-means family has successfully been applied to a wide variety of problems (Leski, 2016). For this family, the gravity center of a class is the mean of all observations, weighted by their degree of belonging to the class. The assignment of an observation to a class is obtained by minimizing an objective

function: $arg \min_{\ell} \sum_{j=1}^{N} \sum_{\ell=1}^{M} \omega_{j\ell} \left\| X_j - g_\ell \right\|^2$

where each element $\omega_{j\ell}$ refers to the similarity degree of an observation X_j to the class Ω_ℓ. The algorithm returns gravity centers $\{g_1,...,g_M\}$ and a similarity matrix $W = \omega_{j\ell} \in [0,1]$, where $j=1,...,N$; $\ell=1,...,M$; N is the number of observations and M is the number of classes. In (Fontes and Pereira, 2016), this algorithm was applied and verified for the clustering of starting patterns in a gas turbine for fault detection purposes.

3.2.4. Bayesian approach

A Bayesian network (BN) represents a network of directed acyclic graph where nodes are considered as faults. This means that BN links between faults and their symptoms. The node of faults is often root node, which do not have any input arc but have prior probabilities. A descendent node has conditional probabilities that represents all possible combinations of their states and their parent nodes' states (Zhao et al., 2017). In (Zhao et al., 2015), which is the second part of the latter paper, BN-method was used for diagnosing air handling units faults and the study showed that this method is powerful even with lack on information.

4. HYBRID METHODS

In order to enhance diagnostic performance, model-based and data-driven approaches have been combined. Methods created by combining more than one data driven technique remain a data driven approach. Hybrid methods aim to enhance the diagnostic accuracy by leveraging the advantages and avoiding the limitations of their consisting methods (Skliros et al., 2019). In (Luo et al., 2010), a hybrid approach which combines model-based and data-driven methods was investigated to obtain better diagnostic performance of a vehicle's antilock braking system. First, parity equations and a nonlinear observer were combined to generate the residues. Then, statistical tests and the support vector machine were used for fault isolation. Finally, subset selection is used to accurately estimate the fault severity. In (Ghimire et al., 2011), a hybrid method was proposed to

fault detection and diagnosis for a vehicle's electronic power system. Model based methods were used for fault detection and data driven methods for fault isolation. Residues of the observed measurements to a predefined threshold were used for fault detection. Partial Least Square algorithm was used to dataset dimensionality reduction. The reduced dimensionality data was used to train machine learning algorithms (support vector machine, probabilistic neural network, and nearest neighbor) to classify different faults. Finally, SVM regression was used to estimate the fault severity.

5. CONCLUSION

This paper reviewed the most methods used in the diagnostic of faults in industrial systems. These methods reinforced by prognosis methods can lead to an effective reduction of unplanned outages and increase reliability and dependability of systems and thus reduces maintenance costs. The state of art proposed in this paper reveals that several methods can be classified in: model-based, data-driven and hybrid methods. In general, model-based approach is very reliable even with a lack in historical data but their applicability depends on detailed physical model whose coverage will necessarily be limited by their ability to integrate a physical understanding of the failure process or system degradation. In addition to this, data-driven approach does not require knowledge of the system physics but it is based on the historical data of measurements and often actual observations in order to run the method online. Finally, the hybrid approach consists on gathering methods from model-based and data driven whether from the same type or heterogeneous. These combinations aim at reducing classification error in the case of diagnostic application

Indeed, nobody can affirm that an approach is the best because the choice relies on the application and the type of inputs. However, in most of cases, the system modeling is a hard task for complex systems which makes data-driven methods either for diagnostic or prognostic more applicable in most cases. In the later researches, it is recurrent to find hybrid methods because a perfect single method does not exist.

A comparison between the different methods used for the diagnostic of defects is shown in Table 1. The purpose is to identify the strengths and weaknesses of diagnostic methods mentioned above. Research conducted by (Venkatasubramanian et al., 2003) are used to define new features to judge a diagnostic method.

Methods / Criteria	Parity space or observers	parametric estimation	expert Systems	SDG	QTA	PCA	Neural Networks	Classification
Fast detection and diagnostic	✓	X	✓	?	✓	✓	✓	✓
Identification	✓	✓	✓	X	✓	✓	✓	✓
robustness	✓	?	✓	✓	✓	✓	?	?
Identification of new faults	✓	X	X	✓	?	✓	?	✓
None required physical model	X	X	✓	✓	✓	✓	✓	✓
Identification of multiple faults	✓	✓	✓	✓	X	X	?	✓
Easy to operate	X	X	✓	✓	✓	X	X	✓
Time calculation	X	X	✓	?	✓	✓	?	?
Adaptation	X	X	X	✓	?	X	✓	✓

✓ : Favorable   X : Not favorable   ? : Depend on the situation

Table 1. Comparison of different methods used for fault diagnostic.

REFERENCES

Alhelou, H.H. (2019). Fault Detection and Isolation in Power Systems Using Unknown Input Observer. Adv. Cond. Monit. Fault Diagn. Electr. Mach. 38–58.

Alrowaie, F., Kwok, K.E., and Gopaluni, R.B. (2014). Fault Isolation based on General Observer Scheme in Stochastic Non-linear State-Space Models Using Particle Filters.

Angeli, C., and Chatzinikolaou, A. (2004). On-line Fault Detection Techniques for Technical Systems: A survey. Int. J. Comput. Sci. Appl. Vol *1*, 12–30.

Ariola, M., Mattei, M., Notaro, I., Corraro, F., and Sollazzo, A. (2015). An SFDI Observer–Based Scheme for a General Aviation Aircraft. Int. J. Appl. Math. Comput. Sci. *25*, 149–158.

Atamuradov, V., Medjaher, K., Dersin, P., Zerhouni, N., and Camci, F. (2018). A new adaptive prognostics approach based on hybrid feature selection with application to point machine monitoring. pp. 1–9.

Atoui, M.A., Verron, S., and Kobi, A. (2016). A Bayesian network dealing with measurements and residuals for system monitoring. Trans. Inst. Meas. Control *38*.

Baraldi, P., Cannarile, F., Di Maio, F., and Zio, E. (2016). Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions. Eng. Appl. Artif. Intell. *56*, 1–13.

Ben Ali, J., Fnaiech, N., Saidi, L., Chebel-Morello, B., and Fnaiech, F. (2015). Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Appl. Acoust. *89*, 16–27.

Benkouider, A.M., Kessas, R., Yahiaoui, A., Buvat, J.C., and Guella, S. (2012). A hybrid approach to faults detection and diagnosis in batch and semi-batch reactors by using EKF and neural network classifier. J. Loss Prev. Process Ind. *25*, 694–702.

Blesa, J., Puig, V., Saludes, J., and Fernández-Cantí, R.M. (2016). Set-membership parity space approach for fault detection in linear uncertain dynamic systems. Int. J. Adapt. Control Signal Process. *30*, 186–205.

Broomhead, D.S. (1988). Radial Basis Functions, Multivariable Functional Interpolation and Adaptive Networks (Royals Signals & Radar Establishment).

Capolino, G.-A., Antonino-Daviu, J.A., and Riera-Guasp, M. (2015). Modern Diagnostics Techniques for Electrical Machines, Power Electronics, and Drives. IEEE Trans. Ind. Electron. *62*, 1738–1745.

Casimir, R., Boutleux, E., Clerc, G., and Chappuis, F. (2003a). Broken bars detection in an induction motor by pattern recognition. (IEEE), pp. 313–319.

Casimir, R., Boutleux, E., and Clerc, G. (2003b). Fault diagnosis in an induction motor by pattern recognition methods. (IEEE), pp. 294–299.

Chen, W., and Saif, M. (2007). Observer-based strategies for actuator fault detection, isolation and estimation for certain class of uncertain nonlinear systems. IET Control Theory Amp Appl. 1, 1672–1680.

Cottrell, M., Gaubert, P., Eloy, C., François, D., Hallaux, G., Lacaille, J., and Verleysen, M. (2009). Fault Prediction in Aircraft Engines Using Self-Organizing Maps. In Advances in Self-Organizing Maps, (Springer, Berlin, Heidelberg), pp. 37–44.

Dash, S., Maurya, M.R., Venkatasubramanian, V., and Rengaswamy, R. (2004). A novel interval-halving framework for automated identification of process trends. AIChE J. 50, 149–162.

Dey, S., Pisu, P., and Ayalew, B. (2015). A Comparative Study of Three Fault Diagnosis Schemes for Wind Turbines. IEEE Trans. Control Syst. Technol. 23, 1853–1868.

Ding, F., Wang, X., Chen, Q., and Xiao, Y. (2016). Recursive Least Squares Parameter Estimation for a Class of Output Nonlinear Systems Based on the Model Decomposition. Circuits Syst. Signal Process. 35, 3323–3338.

Ding, S., Zhang, P., Ding, E., Naik, A., Deng, P., and Gui, W. (2010). On the application of PCA technique to fault diagnosis. Tsinghua Sci. Technol. 15, 138–144.

Ding, S.X., Zhang, P., Jeinsch, T., Ding, E.L., Engel, P., and Gui, W. (2011). A survey of the application of basic data-driven and model-based methods in process monitoring and fault diagnosis. IFAC Proc. Vol. 44, 12380–12388.

Duda, R.O., Hart, P.E., and Stork, D.G. (2012). Pattern Classification (John Wiley & Sons).

Eissa, M.A., Darwish, R.R., Bassiuny, A.M., Eissa, M.A., Darwish, R.R., and Bassiuny, A.M. (2019). New Model-Based Fault Detection Approach using Black Box Observer. J. Mechatron. Robot. 3, 42–51.

Elghazel, W., Bahi, J., Guyeux, C., Hakem, M., Medjaher, K., and Zerhouni, N. (2015). Dependability of wireless sensor networks for industrial prognostics and health management. Comput. Ind. 68, 1–15.

Fontes, C.H., and Pereira, O. (2016). Pattern recognition in multivariate time series – A case study applied to fault detection in a gas turbine. Eng. Appl. Artif. Intell. 49, 10–18.

Gao, D., Wu, C., Zhang, B., and Ma, X. (2010). Signed Directed Graph and Qualitative Trend Analysis Based Fault Diagnosis in Chemical Industry. Chin. J. Chem. Eng. 18, 265–276.

Gao, Z., Cecati, C., and Ding, S.X. (2015). A survey of fault diagnosis and fault-tolerant techniques- Part I: fault diagnosis With model-based and signal-based approaches. IEEE Trans. Ind. Electron. 62, 3757–3767.

Gertler, J. (2017). Fault Detection and Diagnosis in Engineering Systems (CRC Press).

Ghimire, R., Sankavaram, C., Ghahari, A., Pattipati, K., Ghoneim, Y., Howell, M., and Salman, M. (2011). Integrated model-based and data-driven fault detection and diagnosis approach for an automotive electric power steering system. (IEEE), pp. 70–77.

Ghosh, K., Ng, Y.S., and Srinivasan, R. (2011). Evaluation of decision fusion strategies for effective collaboration among heterogeneous fault diagnostic methods. Scopus.

Glowacz, A., and Glowacz, Z. (2017). Diagnosis of stator faults of the single-phase induction motor using acoustic signals. Appl. Acoust. 117, 20–27.

Guo, Y., Wang, J., Chen, H., Li, G., Huang, R., Yuan, Y., Ahmad, T., and Sun, S. (2019). An expert rule-based fault diagnosis strategy for variable refrigerant flow air conditioning systems. Appl. Therm. Eng. 149, 1223–1235.

Hafaifa, A., Guemana, M., and Daoudi, A. (2015). Vibrations supervision in gas turbine based on parity space approach to increasing efficiency. J. Vib. Control 21, 1622–1632.

Harmouche, J. (2014). Statistical Incipient Fault Detection and Diagnosis with Kullback-Leibler Divergence : from Theory to Applications (Supélec).

He, H., Liu, Z., and Hua, Y. (2015). Adaptive Extended Kalman Filter Based Fault Detection and Isolation for a Lithium-Ion Battery Pack. Energy Procedia 75, 1950–1955.

Hua, J., Lu, L., Ouyang, M., Li, J., and Xu, L. (2011). Proton exchange membrane fuel cell system diagnosis based on the signed directed graph method. J. Power Sources 196, 5881–5888.

Hwang, I., KIM, S., KIM, Y., and SEAH, C.E. (2010). A Survey of Fault Detection, Isolation, and Reconfiguration Methods. Surv. Fault Detect. Isol. Reconfiguration Methods 18, 636–653.

Idrissi, I., El bachtiri, R., and Chafouk, H. (2017). A Bank of Kalman Filters for Current Sensors Faults Detection and Isolation of DFIG for Wind Turbine. (IEEE), pp. 1–6.

Jardine, A.K.S., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mech. Syst. Signal Process. 20, 1483–1510.

Jia, F., Lei, Y., Lu, N., and Xing, S. (2018). Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. Mech. Syst. Signal Process. 110, 349–367.

Jiang, T., Khorasani, K., and Tafazoli, S. (2008). Parameter Estimation-Based Fault Detection, Isolation and Recovery for Nonlinear Satellite Models. IEEE Trans. Control Syst. Technol. 16, 799–808.

Jiang, Y., An, B., Huo, M., and Yin, S. (2018). Design Approach to MIMO Diagnostic Observer and its Application to Fault Detection. (IEEE), pp. 5377–5382.

Jing, C., and Hou, J. (2015). SVM and PCA based fault classification approaches for complicated industrial process. Neurocomputing 167, 636–642.

Jlassi, I., Estima, J.O., El Khil, S.K., Bellaaj, N.M., and Cardoso, A.J.M. (2017). A Robust Observer-Based Method for IGBTs and Current Sensors Fault Diagnosis in Voltage-

Source Inverters of PMSM Drives. IEEE Trans. Ind. Appl. *53*, 2894–2905.

Kim, Y.-H., Youn, Y.-W., Hwang, D.-H., Sun, J.-H., and Kang, D.-S. (2013). High-Resolution Parameter Estimation Method to Identify Broken Rotor Bar Faults in Induction Motors. IEEE Trans. Ind. Electron. *60*, 4103–4117.

Kohonen, T. (2001). Self-Organizing Maps 3rd edition.

Kurban, T., and Beşdok, E. (2009). A Comparison of RBF Neural Network Training Algorithms for Inertial Sensor Based Terrain Classification. Sensors *9*, 6312–6329.

Laamami, S., Ben, H.M., and Sbita, L. (2015). Fault detection for nonlinear systems: Parity space approach. (IEEE), pp. 1–5.

Leski, J.M. (2016). Fuzzy c-ordered-means clustering. Fuzzy Sets Syst. *286*, 114–133.

Lin, D., Wiseman, M., Banjevic, D., Jardine, A.K.S., Lin, D., Wiseman, M., Banjevic, D., and Jardine, A.K.S. (2004). An approach to signal processing and condition-based maintenance for gearboxes subject to tooth failure. Mech. Syst. Signal Process. *18*, 993.

Liu, R., Yang, B., Zio, E., and Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. Mech. Syst. Signal Process. *108*, 33–47.

Liu, Y.-J., Meng, Q.-H., Zeng, M., and Ma, S.-G. (2016). Fault diagnosis method based on probability extended SDG and fault index. In 2016 12th World Congress on Intelligent Control and Automation (WCICA), (Guilin, China: IEEE), pp. 2868–2873.

Lu, D., and Qiao, W. (2013). Adaptive feature extraction and SVM classification for real-time fault diagnosis of drivetrain gearboxes. (IEEE), pp. 3934–3940.

Luo, J., Namburu, M., Pattipati, K.R., Qiao, L., and Chigusa, S. (2010). Integrated Model-Based and Data-Driven Diagnosis of Automotive Antilock Braking Systems. IEEE Trans. Syst. Man Cybern. - Part Syst. Hum. *40*, 321–336.

Maurya, M.R., Rengaswamy, R., and Venkatasubramanian, V. (2007). A Signed Directed Graph and Qualitative Trend Analysis-Based Framework for Incipient Fault Diagnosis. Chem. Eng. Res. Des. *85*, 1407–1422.

Moshou, D., Kateris, D., Sawalhi, N., Loutridis, S., and Gravalos, I. (2010). Fault Severity Estimation in Rotating Mechanical Systems Using Feature Based Fusion and Self-Organizing Maps. In Artificial Neural Networks – ICANN 2010, (Springer, Berlin, Heidelberg), pp. 410–413.

msaaf, mohammed, and Belmajdoub, F. (2015). L'application des réseaux de neurone de type " feedforward " dans le diagnostic statique. In Xème Conférence Internationale : Conception et Production Intégrées, (Tanger, Morocco).

Nouri.Gharahasanlou, A., Mokhtarei, A., Khodayarei, A., and Ataei, M. (2014). Fault tree analysis of failure cause of crushing plant and mixing bed hall at Khoy cement factory in Iran. Case Stud. Eng. Fail. Anal. *2*, 33–38.

Ondel, O., Boutleux, E., and Clerc, G. (2006). A method to detect broken bars in induction machine using pattern recognition techniques. IEEE Trans. Ind. Appl. *42*, 916–923.

Palak, J. (2017). A Luenberger Observer-Based Fault Detection and Identification Scheme for Photovoltaic DC-DC Converters.

Pontoppidan, N.H., and Larsen, J. (2003). Unsupervised condition change detection in large diesel engines. (IEEE), pp. 565–574.

Purba, J.H. (2014). A fuzzy-based reliability approach to evaluate basic events of fault tree analysis for nuclear power plant probabilistic safety assessment. Ann. Nucl. Energy *70*, 21–29.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain. Psychol. Rev. 65–386.

Ruijters, E., and Stoelinga, M. (2015). Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. Comput. Sci. Rev. *15–16*, 29–62.

Saludes, S., Corrales, A., de Miguel, L.J., and Perán, J.R. (2003). A SOM and Expert System Based Scheme for Fault Detection and Isolation in a Hydroelectric Power Station. IFAC Proc. Vol. *36*, 999–1004.

Saravanakumar, R., Manimozhi, M., Kothari, D.P., and Tejenosh, M. (2014). Simulation of Sensor Fault Diagnosis for Wind Turbine Generators DFIG and PMSM Using Kalman Filter. Energy Procedia *54*, 494–505.

Siswantoro, J., Prabuwono, A.S., Abdullah, A., and Idrus, B. (2016). A linear model based on Kalman filter for improving neural network classification performance. Expert Syst. Appl. *49*, 112–122.

Skliros, C., Esperon Miguez, M., Fakhre, A., and Jennions, I.K. (2019). A review of model based and data driven methods targeting hardware systems diagnostics. Diagnostyka *Vol. 20, No. 1*.

Sliskovic, D., Grbic, R., and Hocenski, Ž. (2012). Multivariate statistical process monitoring.

Soualhi, A., Clerc, G., Razik, H., and Ondel, O. (2011). Detection of induction motor faults by an improved artificial ant clustering. (IEEE), pp. 3446–3451.

Souibgui, F., BenHmida, F., and Chaari, A. (2011). Particle filter approach to fault detection and isolation in nonlinear systems. (IEEE), pp. 1–6.

Srivastava, N.P., Srivastava, R.K., and Vashishtha, P. (2014). Fault Detection and Isolation (Fdi) Via Neural Networks.

Urmos, A., Farkas, M., Kóczy, L., and Nemcsics, A. (2013). Quantum structure classification by Kohonen Self-Organizing Map and by Fuzzy C-Means algorithm. In Hungarian Consortium, pp. 313–318.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S.N., and Yin, K. (2003). A review of process fault detection and diagnosis: Part III: Process history based methods. Comput. Chem. Eng. *27*, 327–346.

Vento, J., Blesa, J., Puig, V., and Sarrate, R. (2015). Set-membership parity space hybrid system diagnosis. Int. J. Syst. Sci. *46*, 790–807.

Villez, K. (2014). Qualitative trend analysis for process monitoring and supervision based on likelihood optimization: state-of-the-art and current limitations. IFAC Proc. Vol. *47*, 7140–7145.

Wang, L., and Nee, A.Y.C. (2009). Collaborative Design and Planning for Digital Manufacturing (London: Springer-Verlag).

Xu, L., Chen, L., and Xiong, W. (2015). Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. Nonlinear Dyn. *79*, 2155–2163.

Yang, F., Shah, S., and Xiao, D. (2012). Signed directed graph based modeling and its validation from process knowledge and process data. Int. J. Appl. Math. Comput. Sci. *22*, 41–53.

Yang, Z., Zhong, J., and Wong, S.F. (2011). Machine learning method with compensation distance technique for gear fault detection. In 2011 9th World Congress on Intelligent Control and Automation, pp. 632–637.

Yin, S., Ding, S.X., Xie, X., and Luo, H. (2014). A Review on Basic Data-Driven Approaches for Industrial Process Monitoring. IEEE Trans. Ind. Electron. *61*, 6418–6428.

Yin, S., Ding, S.X., and Zhou, D. (2016). Diagnosis and Prognosis for Complicated Industrial Systems—Part I. IEEE Trans. Ind. Electron. *63*, 2501–2505.

Younus, A.M.D., and Yang, B.-S. (2012). Intelligent fault diagnosis of rotating machinery using infrared thermal image. Expert Syst. Appl. *39*, 2082–2091.

Zappone, A., Di Renzo, M., and Debbah, M. (2019). Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both? ArXiv.Org.

Zhao, Y., Xiao, F., and Wang, S. (2013). An intelligent chiller fault detection and diagnosis methodology using Bayesian belief network. Energy Build. *57*, 278–288.

Zhao, Y., Wen, J., and Wang, S. (2015). Diagnostic Bayesian networks for diagnosing air handling units faults – Part II: Faults in coils and sensors. Appl. Therm. Eng. *90*, 145–157.

Zhao, Y., Wen, J., Xiao, F., Yang, X., and Wang, S. (2017). Diagnostic Bayesian networks for diagnosing air handling units faults – part I: Faults in dampers, fans, filters and sensors. Appl. Therm. Eng. *111*, 1272–1286.

Zhou, B., and Ye, H. (2016). A study of polynomial fit-based methods for qualitative trend analysis. J. Process Control *37*, 21–33.

Zogopoulos-Papaliakos, G., and Kyriakopoulos, K.J. (2019). Parity-Based Diagnosis in UAVs: Detectability and Robustness Analyses. (IEEE), pp. 6496–6502.

Zuo, M.J., Lin, J., and Fan, X. (2005). Feature separation using ICA for a one-dimensional time series and its application in fault detection. J. Sound Vib. *287*, 614–624.