

# Signal Based Fault Detection and Diagnosis for Rotating Electrical Machines: Issues and Solutions

Andrea Giantomassi, Francesco Ferracuti, Sabrina Iarlori,  
Gianluca Ippoliti and Sauro Longhi

**Abstract** Complex systems are found in almost all field of contemporary science and are associated with a wide variety of financial, physical, biological, information and social systems. Complex systems modelling could be addressed by signal based procedures, which are able to learn the complex system dynamics from data provided by sensors, which are installed on the system in order to monitor its physical variables. In this chapter the aim of diagnosis is to detect if the electrical machine is healthy or a change is occurring due to abnormal events and, in addition, the probable causes of the abnormal events. Diagnosis is addressed by developing machine learning procedures in order to classify the probable causes of deviations from system normal events. This chapter presents two Fault Detection and Diagnosis solutions for rotating electrical machines by signal based approaches. The first one uses a current signature analysis technique based on Kernel Density Estimation and Kullback–Liebler divergence. The second one presents a vibration signature analysis technique based on Multi-Scale Principal Component Analysis. Several simulations and experimentations on real electric motors are carried out in order to verify the effectiveness of the proposed solutions. The results show that the proposed signal based diagnosis procedures are able to detect and diagnose different electric motor faults and defects, improving the reliability of electrical machines. Fault Detection and Diagnosis algorithms could be used not only with the fault diagnosis purpose but also in a Quality Control scenario. In fact, they can be

---

A. Giantomassi (✉) · F. Ferracuti · S. Iarlori · G. Ippoliti · S. Longhi  
Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche,  
Via Brece Bianche, 60131 Ancona, Italy  
e-mail: a.giantomassi@univpm.it

F. Ferracuti  
e-mail: f.ferracuti@univpm.it

S. Iarlori  
e-mail: s.iarlori@univpm.it

G. Ippoliti  
e-mail: gianluca.ippoliti@univpm.it

S. Longhi  
e-mail: sauro.longhi@univpm.it

integrated in test benches at the end or in the middle of the production line in order to test the machines quality. When the electric motors reach the test benches, the sensors acquire measurements and the Fault Detection and Diagnosis procedures detect if the motor is healthy or faulty, in this last case further inspections can diagnose the fault.

## 1 Introduction

Mathematical process models describe the relationship between input signals  $u(k)$  and output signals  $y(k)$  and are fundamental for model-based fault detection. In many cases the process models are not known at all or some parameters are unknown. Further, the models have to be rather precise in order to express deviations as results of process faults. Therefore, process-identification methods have to be applied frequently before applying any model-based fault detection method as stated in Giantomassi (2012). But also the identification method itself may be a source to gain information on, e.g. process parameters which change under the influence of faults. First publications on fault detection with identification methods are found in Isermann (1984) and Filbert and Metzger (1982).

For dynamic processes the input signals may be the normal operating signal or may be artificially introduced for testing. A considerable advantage of identification methods is that with only one input and one output signal several parameters can be estimated, which give a detailed picture on internal process quantities. The generated features for fault detection are then impulse response values in the case of correlation methods or parameter estimates [see Isermann (2006)].

On-line process monitoring with fault detection and diagnosis can provide range of processes, as stated in Cheng et al. (2008), Giantomassi et al. (2011) and Ferracuti et al. (2010, 2011). A large number of applications have been reviewed, e.g. Isermann and Balle (1997) and Patton et al. (2000). Venkatasubramanian et al. (2000a, b, c) published an article series reviewing monitoring methods with attention in the field of chemical processes. They classified the Fault Detection and Diagnosis methods as model-based, signal-based and knowledge-based. Signal-based approaches to fault detection and isolation (FDI) in large-scale process plants are consolidated and well studied, because for these processes the development of model-based FDI methods requires considerable and eventually too high effort, and moreover because a large amount of data is collected, as stated in Chiang et al. (2000) and Isermann (2006).

Fault detection and diagnosis (FDD) in industrial applications regards two important aspects: the FDD for the production plant and for the systems that work for the plant; among these systems, induction motors are the most important electrical machineries in many industrial applications, considering that, electric

motors account about 65 % of energy use. In the field of operational efficiency, the monitoring activity of rotating electrical machines by fault detection and diagnosis is in-depth investigated: Benloucif and Balaska (2006), Ran and Penman (2008), Singh and Ahmed (2004), Taniguchi et al. (1999), Tavner (2008), Verucchi and Acosta (2008). Vibration analysis is widely accepted as a tool to detect faults of a machine since it is nondestructive, reliable and it permits continuous monitoring without stopping the machine [see Ciandrini et al. (2010), Gani and Salami (2002), Hua et al. (2009); Immovilli et al. (2010), Shuting et al. (2002); Zhaoxia et al. (2009)]. In particular analysing the vibration power spectrum it is possible to detect different faults that arise in rotating machines. In traditional machine vibration signature analysis (MVSA), the Fourier transform is used to determine the vibration power spectrum and the signature at different frequencies are identified and compared with those related to healthy motors to detect faults in the machine, as in Lachouri et al. (2008). The shortcoming of this approach is that the Fourier analysis is limited to stationary signals while vibrations are not stationary by its nature.

The use of Soft Computing methods is considered an important extension to the model-based approach Patton et al. (2000). It allows to improve residual generation in FDD when process signals show complex behaviours. Multi-scale principal component analysis (MSPCA) deals with processes that operate at different scales: events occurring at different localizations in time and frequency, stochastic processes and variables measured at different sampling rate, as reported in Bakshi (1998) and Li et al. (2000). PCA, treated in Jolliffe (2002) and Jackson (2003), decorrelates the variables by extracting a linear relationship in order to transform the multivariate space into a subspace which preserves maximum variance of the original space. Wavelets extract deterministic features and approximately decorrelate autocorrelated measurements. MSPCA combines these two techniques to extract maximum information from multivariate sensor data (Misra et al. 2002).

Rotating electrical machines are well known systems with accurate analytical models and extensive results in literature. Failure surveys, as Thomson and Fenger (2001), report that failures, in induction motors, are: stator related (38 %), rotor related (10 %), bearing related (40 %) and others (12 %). Fast and accurate diagnosis of incipient faults allows actions to protect the power system, the process leaded by the machine and the machine itself.

FDD techniques based on MVSA have received great attention in literature because by vibrations it is possible to identify directly mechanical faults regarding rotating electrical machines. In recent years, many methodologies have been developed to detect and diagnose mechanical faults of electrical machines by current measurements. In this context motor current signature analysis (MCSA) involves detection and identification of current signature patterns that are indicative of normal and abnormal motor conditions. However, the motor current is influenced by many factors such as electric supply, static and dynamic load conditions, noise, motor geometry and faults. In Chilengue et al. (2011) an artificial immune system approach is investigated for the detection and diagnosis of faults in the stator and

rotor circuits of induction machines. The proposed technique measures the stator currents to compute its representation before and after a fault condition. These patterns are used to construct a characteristic image of the machine operating condition. Moreover MCSA procedures are used to detect and diagnose not only classic motor faults (i.e. rotor eccentricity), but also gear faults (i.e. tooth spall), as presented in Feki et al. (2013). Fault Tolerant Control (FTC) as well as robust control systems have been applied in electric drive systems Ciabattoni et al. (2011a, 2011b, 2014). In Abdelmadjid et al. (2013) a FTC procedure is proposed for stator winding fault of induction motors. It consists of an algorithm which can detect an incipient fault in closed loop and switches itself between a nominal control strategy for healthy condition and a robust control for faulty condition. Samsi et al. (2009) validated a technique, called Symbolic Dynamic Filtering (SDF), for early detection of stator voltage imbalance in three-phase induction motors that involves Wavelet Transform (WT) of current signals. In Baccharini et al. (2010) a sensor-less approach has been proposed to detect one broken rotor bar in induction motors. This method is not affected by load and other asymmetries. The technique estimates stator and rotor flux and analyses the differences obtained in torque. A new saturation model that explains the experimental data is investigated in Pedra et al. (2009). The model has three different saturation effects, which have been characterized in four induction motors.

As possible solutions of the FDI problem for electrical machines, two different approaches are proposed: the first one uses vibration signals provided by accelerometer sensors placed on the machine, and the second one uses current signals provided by inverters.

In the first solution, based on current signal analysis of rotating electrical machines, different algorithms are applied for FDD: PCA is used to reduce the three-phase current space in two dimensions. Then, Kernel Density Estimation (KDE) is adopted to estimate the probability density function (PDF) of each healthy and faulty motor, which are typical features that can be used to identify each fault [see Ferracuti et al. (2013a)]. Kullback–Leibler (K–L) divergence is used as a distance between two PDF obtained by KDE. K–L allows to identify the dissimilarity between two probability distributions (that can also be multidimensional): one is related to the modelled signatures and the other one is related to the acquired data samples. The classification of each motor condition is performed by K–L divergence.

In the second approach, based on vibration analysis of rotating electrical machines, MSPCA is applied for fault detection and diagnosis (Ferracuti et al. 2013b; Lachouri et al., 2008; Misra et al. 2002). Fault identification is evaluated by calculating the contributions of each variable in the principal component subspace and in the residual space. KDE, which allows to estimate the PDF of random variables is introduced, in Odiowei and Cao (2010), to improve fault detection and isolation. The contributions PDFs are estimated by KDE, the thresholds are computed for each signal in order to improve fault detection. Faults are classified by using the contribution plots by Linear Discriminant Analysis (LDA).

The proposed data-driven algorithms for FDD based on MVSA and MCSA are tested by several simulations and experimentations in order to verify the effectiveness of the proposed methodologies.

The chapter will be organized in the following sections. In Sect. 2 the FDD algorithm based on Motor Current Signature Analysis is discussed with focus on Quality Control scenario. Experimental tests on real motors are reported in Sect. 3. The FDI algorithm based on vibration signals is described in Sect. 4. Experimental tests on real motors are reported in Sect. 5. Comments on the performances of the proposed solutions are reported in Sect. 6.

## **2 Electric Motor FDD by MCSA in Quality Control Scenario**

In industry, QC is a collection of methods that are able to improve the quality and efficiency in processes, productions and in many others industry aspects. In 1924, Walter Shewhart designed the first control chart and gave a rationale for its use in process monitoring and control (Stuart et al. 1995). The main concept of QC is the “proactiveness” that ensures the product quality, processes and signals monitoring to detect when they “go out of control”. In the last years, manufacturing industries are paying attention and efforts for the introduction of QC in the production lines. Large volumes of low-tech products involve many investigations on the efficient introduction of QC in production lines.

One of the major problems, in which these manufacturing industries are involved, is the customers satisfaction, because they usually purchase a lot of products with some unwanted defective component. In order to satisfy customers, manufacturing industries carry out some spot checks at the end of production lines. This method does not ensure the quality of products and total defective products removal. A desirable QC solution for these manufacturing industries should be minimally invasive, effective and with a low payback period. In addition, tests should be performed in a systematic way using a low-cost system based on a reduced set of sensors embedded in the test bench.

The proposed FDD system acquires sensor measurements and detects defective products. Moreover, by isolating and identifying the defective type, the FDD procedure helps to estimate in which subprocess the defect is introduced and allows to remove the defective products, improving the processes quality. The tests, performed at the end of production lines, allow to improve the quality of processes as proactive measures for the QC methodology.

## 2.1 Recalled Results

In this section authors present the algorithms used to develop the FDD procedure. They extract patterns by current signals using PCA and KDE. Then K–L divergence compares these patterns to extract the motor health index.

### 2.1.1 Principal Component Analysis

PCA is a dimensionality reduction technique that produces a lower dimensional representation in a way that preserves the correlation structure between the process variables capturing the variability in the data (Jolliffe 2002). PCA rotates the original coordinate system along the direction of maximum variance. Considering a data matrix  $\mathbf{X} \in \mathbb{R}^{N \times m}$  of  $N$  sample rows and  $m$  variable columns that are normalized to zero mean with mean values vector  $\boldsymbol{\mu}$ . The matrix  $\mathbf{X}$  can be decomposed as follows:

$$\mathbf{X} = \hat{\mathbf{X}} + \tilde{\mathbf{X}}, \quad (1)$$

where  $\hat{\mathbf{X}}$  is the projection on the Principal Component Subspace (PCS)  $S_d$ , and  $\tilde{\mathbf{X}}$ , the residual matrix, is the projection on the Residual Subspace (RS)  $S_r$  (see Misra et al. 2002). Defining the loading matrix  $\mathbf{P}$ , whose columns are the right singular vectors of  $\mathbf{X}$ , and selecting the columns of the loading matrix  $\mathbf{P} \in \mathbb{R}^{m \times d}$ , which correspond to the loading vectors associated with the first  $d$  singular values, it follows that:

$$\hat{\mathbf{X}} = \mathbf{XPP}^T \in S_d. \quad (2)$$

The residuals matrix  $\tilde{\mathbf{X}}$ , is the difference between the data matrix  $\mathbf{X}$  and its projection into the first  $d$  principal components retained in the PCA model:

$$\tilde{\mathbf{X}} = \mathbf{X}(\mathbf{I} - \mathbf{PP}^T) \in S_r, \quad (3)$$

therefore the residual matrix captures the variations in the observations space spanned by the loading vectors associated with the  $r = m - d$  smallest singular values. The projections of the observations into the lower-dimensional space are contained in the score matrix:

$$\mathbf{T} = \mathbf{XP} \in \mathbb{R}^{N \times d}. \quad (4)$$

Here, PCA is applied to the three-phase currents of induction motors in order to reduce the inputs space from the three original dimensions to two because the currents are highly correlated. Indeed for healthy motor, with three-phase without neutral connection, ideal conditions and a balanced voltage supply, the stator

currents are given by Eq. (5), where  $i_a$ ,  $i_b$  and  $i_c$  denote the three stator currents,  $I_{\max}$  their maximum value,  $f$  their frequency,  $\phi$  their phase angle and  $t$  the time. It is known that each stator current is given by the combination of the others:

$$\begin{cases} i_a(t) = I_{\max}\sin(2\pi ft - \phi) \\ i_b(t) = I_{\max}\sin(2\pi ft - 2\pi/3 - \phi) \\ i_c(t) = I_{\max}\sin(2\pi ft - 4\pi/3 - \phi). \end{cases} \quad (5)$$

The PCA transform (4), applied to the signals in Eq. (5), makes the smallest singular value equal to zero. This implies that the information of the principal component, captured by the smallest singular value is null, then the last principal component could be deleted and the original space reduced from three to two without losing information. This is justified by the fact that in Eq. (5), each stator current is perfectly correlated to the sum of the others. Adding Gaussian white noise, with standard deviation  $\sigma$ , to the stator current signals (Eq. 5), the smallest singular value will not be equal to zero, but it will depend by the ratio between  $I_{\max}$  and  $\sigma$ .

### 2.1.2 Kernel Density Estimation

Given  $N$  independent and identically distributed (i.i.d.) random vectors  $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_N]$ , where  $\mathbf{X}_i = [X_{i1}, \dots, X_{id}]$ , whose distribution function  $F(\mathbf{x}) = P[\mathbf{X} \leq \mathbf{x}]$  is absolutely continuous with unknown PDF  $f(\mathbf{x})$ . The estimated density at  $\mathbf{x}$  is given by Parzen (1962):

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{|H|^d} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{|H|^d}\right). \quad (6)$$

In the present study a two-dimensional Gaussian kernel function is used so  $d$  is 2 and a further simplification, which follows from the restriction of kernel bandwidth  $H = \{h^2I : h > 0\}$ , leads to the single bandwidth estimator so the estimated density  $\hat{f}(\mathbf{x})$  becomes:

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi h^2)^{1/2}} e^{-\frac{\|\mathbf{x} - \mathbf{X}_i\|^2}{2h^2}}. \quad (7)$$

where  $\mathbf{x} \in \mathbb{R}^d$  whose size  $n_{grid}$  is the points number in which the PDF is estimated, accordingly to Wand and Jones (1994a). It is well known that the value of the bandwidth  $h$  and the shape of the kernel function are of critical importance as stated in Mugdadi and Ahmad (2004). In many computational-intelligence methods that employ KDE, the issue is to find the appropriate bandwidth  $h$  [see for example Comaniciu (2003), Mugdadi and Ahmad (2004), Sheather (2004)]. In the present

work the Asymptotic Mean Integrated Squared Error (AMISE) with plug-in bandwidth selection procedure is used to choose automatically the bandwidth  $h$  [treated in Wand and Jones (1994b)]. In the proposed algorithm, KDE is used to model a specific pattern for each motor condition, indeed the features of the current signals are mapped in the two-dimensional principal component space, representing specific signatures of the motor conditions.

### 2.1.3 Kullback–Leibler Divergence

Given two continuous PDFs  $f_1(\mathbf{x})$  and  $f_2(\mathbf{x})$ , a measure of “divergence” or “distance” between  $f_1(\mathbf{x})$  versus  $f_2(\mathbf{x})$  is given in Kullback and Leibler (1951), as:

$$I_{1:2}(X) = \int_{\mathbb{R}^d} f_1(\mathbf{x}) \log \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} d\mathbf{x}, \quad (8)$$

and between  $f_2(\mathbf{x})$  versus  $f_1(\mathbf{x})$  is given by:

$$I_{2:1}(X) = \int_{\mathbb{R}^d} f_2(\mathbf{x}) \log \frac{f_2(\mathbf{x})}{f_1(\mathbf{x})} d\mathbf{x}. \quad (9)$$

Therefore the K–L divergence between  $f_1(\mathbf{x})$  and  $f_2(\mathbf{x})$  is:

$$\begin{aligned} J(f_1; f_2) &= I_{1:2}(X) + I_{2:1}(X) \\ &= \int_{\mathbb{R}^d} (f_1(\mathbf{x}) - f_2(\mathbf{x})) \log \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} d\mathbf{x}. \end{aligned} \quad (10)$$

The above equation is known as the symmetric K–L divergence, which represents a non negative measure between two PDFs. In the present work  $d$  is 2 and a discrete form of K–L divergence is adopted:

$$J(f_1; f_2) = \sum_{i=1}^{n_{grid}} \sum_{j=1}^d (f_1(x_{ij}) - f_2(x_{ij})) \log \frac{f_1(x_{ij})}{f_2(x_{ij})}. \quad (11)$$

The K–L divergence allows to define a fault index: if  $f_{\Omega}$  is the PDF in the PCs space estimated by KDE of the oncoming current measurements, the motor condition is that which minimizes the K–L divergence between  $f_{\Omega}$  and  $f_i$  that is the  $i$ th PDF related to each motor condition:

$$c = \arg \min_i J(f_{\Omega}; f_i), \quad (12)$$

where  $c$  is the classification output.

## 2.2 Developed Algorithm

The developed FDD procedure based on KDE consists of two stages: training and FDD monitoring. In the first, a KDE model is computed for each motor condition, in order to have one KDE model in the case of healthy motor and one for each faulty case. The training steps are summarized below:

- T1. Stator current signals for each motor condition are acquired;
- T2. Data are normalized;
- T3. PCA transform (4) is applied to stator current signals, which are projected into the two-dimensional principal component space;
- T4. The matrices  $\mathbf{P}$  and  $\boldsymbol{\mu}$  are stored;
- T5. KDE is performed on the lower-dimensional principal components space (4) using a grid of  $n_{grid}$  points and a bandwidth  $h$  for the Gaussian kernel function (7);
- T6. PDFs are estimated by KDE (7) and stored.

In diagnosis step, the models previously obtained are compared with the new data and a fault index is calculated. The diagnosis steps are summarized below:

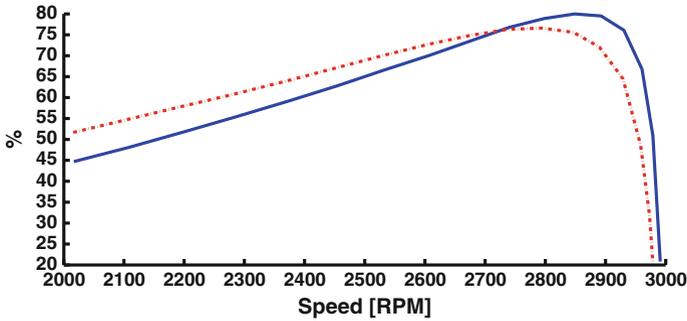
- D1. Stator current signals are acquired;
- D2. Data are normalized;
- D3. The matrices  $\mathbf{P}$  and  $\boldsymbol{\mu}$ , previously computed (T4), are applied to signals;
- D4. KDE is performed on the lower-dimensional principal component space (4) using the same points grid  $n_{grid}$  and bandwidth  $h$  used in the training step (T5);
- D5. Symmetric K–L divergence (11) is computed between the estimated PDF by KDE (7) using the acquired current signals, and those stored in the training step (one for each condition) (T6);
- D6. Diagnosis is evaluated using Eq. (12).

Faults are identified using Eq. (12) where  $f_{\Omega}$  is the PDF, estimated by KDE, in the PCs space of the oncoming current measurements and  $f_i$  is the  $i$ th PDF related to each motor condition. K–L divergence is used as an input for fault decision algorithm allowing to take decision automatically on the operating state and condition of the machine and detecting any abnormal operating condition.

The next Section introduces the FDD experimental results of induction motors in order to show the proposed method performances.

## 3 Electric Motor FDD by MCSA: Results

In order to verify the effectiveness of the proposed methodology several simulations are carried out using one benchmark and some experimentations using real asynchronous motors. The benchmark uses a Time Stepping Coupled Finite Element-State Space modelling (FEM) approach to generate current signals for induction motors as described in Bangura et al. (2003). The simulation dataset consists of



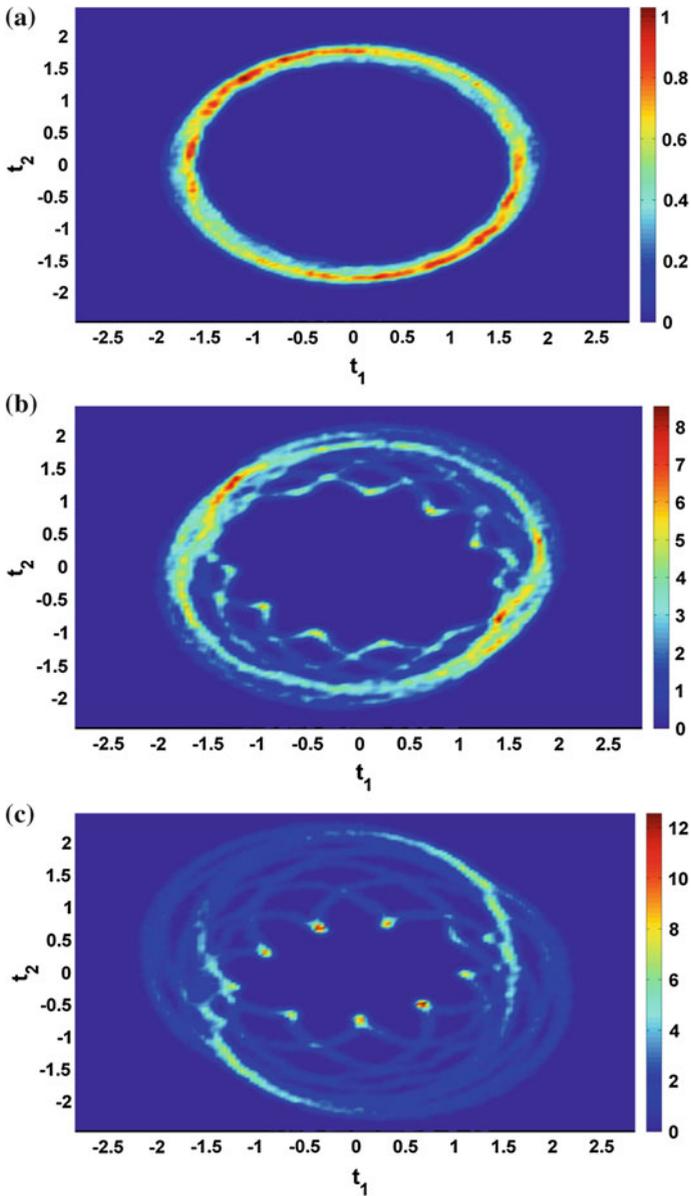
**Fig. 1** Efficiency characterization of tested induction motors. *Blue solid line* refers the healthy motor, *red dashed line* refers to motor with defective rotor

twenty-one different motor conditions, which are: one healthy condition, ten broken bars conditions and ten broken connectors conditions. Twenty time series are generated for each motor condition. Each signal consists of 1,500 samples. The dataset can be download from UCR time series data mining archive in Keogh (2013). The characteristics of the three-phase induction motors are: 208 V input voltage, 60 Hz supply frequency, 34 rotor bars, 2 poles and power 1.2 hp. The sampling rate is 33.3 kHz and the processed data, for each test, are related to 0.3 s of acquisition. White noise with standard deviation  $\sigma = 0.2$  is added to the simulated current signals. The results are the average of 200 Monte Carlo simulations where the training and testing data sets are randomly changed.

The real tests are carried out using three phase induction motors whose parameters are: 380 V input voltage, 60 Hz supply frequency, 0.75 kW power, 20 kHz sampling rate. Two different faults are tested: wrong rotor and cracked rotor. Wrong rotor refers to a non compliant rotor, in particular a single phase rotor is assembled instead of a three phase rotor. Ten motors are tested both for the healthy and faulty cases. The acquisition time is 14 s. The processed data, for each test, are related to 0.7 s of acquisition. In this case study the results are the average of 2,000 Monte Carlo simulations where the training and testing data sets are randomly changed. The motors, with a defective rotor installed, have about 3 % of efficiency drop at the operating point of 2,800 RPM, as shown in Fig. 1. So it is important to detect this defect in the energy efficiency context and QC.

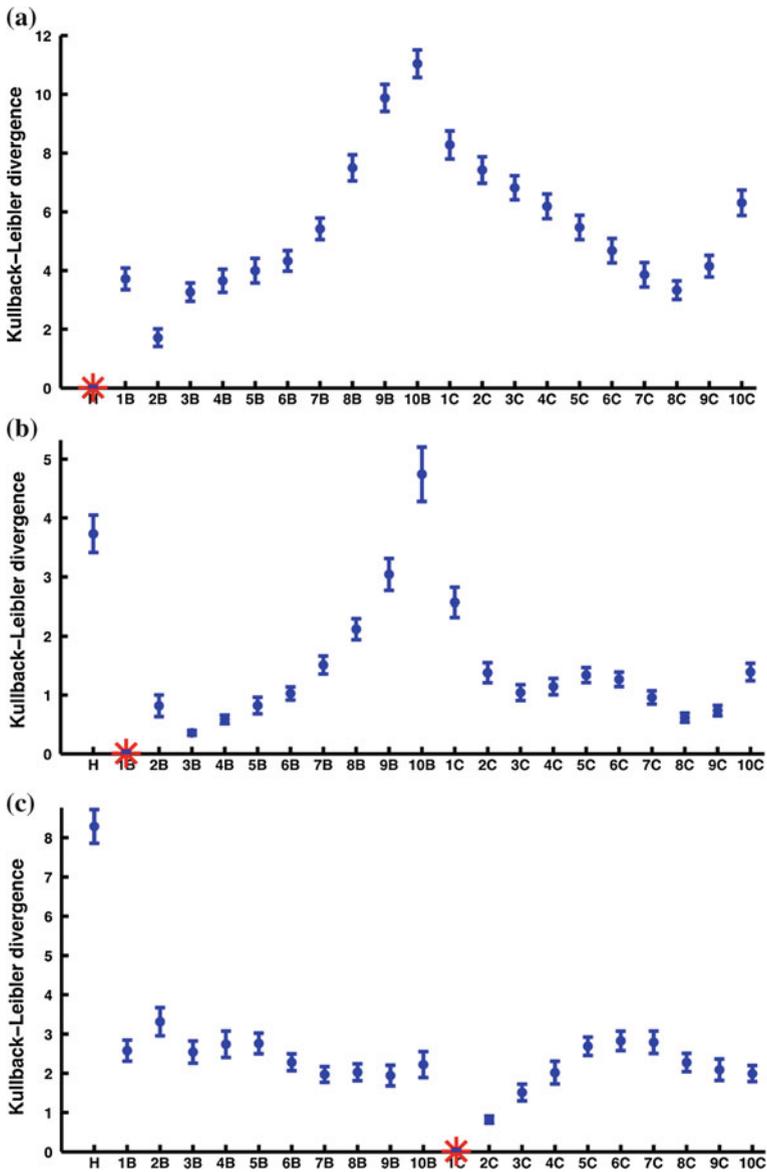
### 3.1 Results and Discussion

The proposed approach processes the three-phase stator currents in order to perform defects detection and diagnosis as described in Sect. 2.2. The following two subsections show the results related to the two cases described previously. Figures 2, 3, 4 and 5 show the simulation and experimentation results. The classification accuracy is considered as an index to evaluate the performances of the proposed

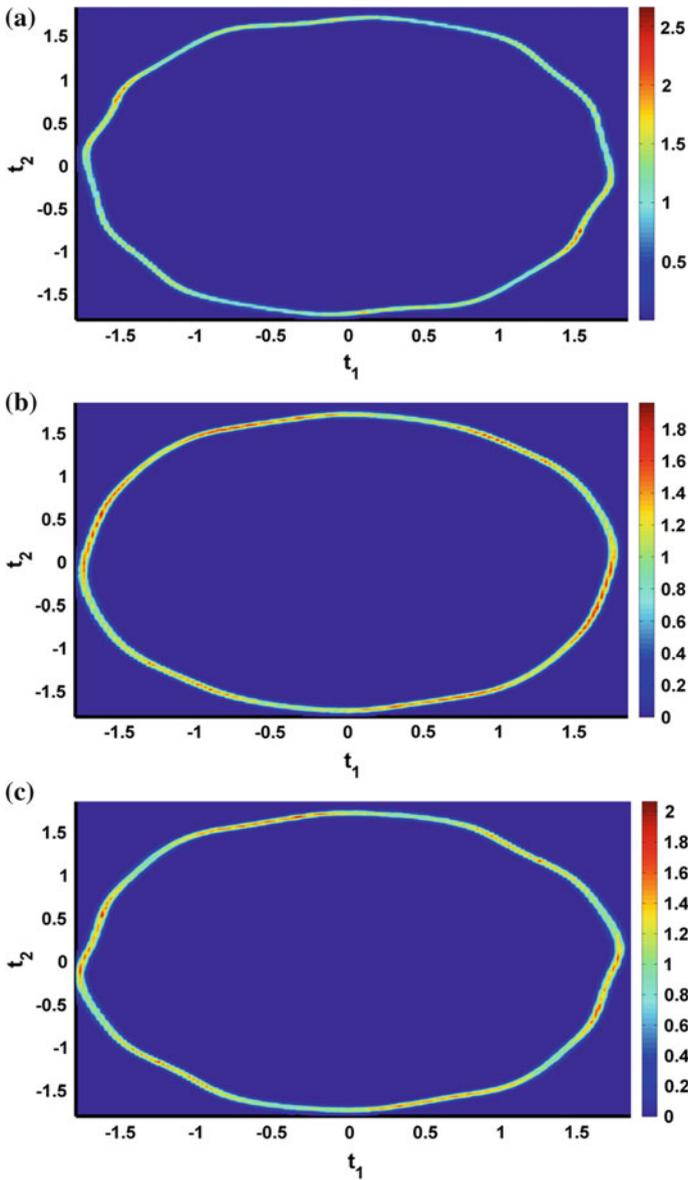


**Fig. 2** Interpolated PDFs of a finite element motor in the two-dimensional principal component space estimated by KDE. **a** Healthy motor. **b** Motor with one broken bar. **c** Motor with one broken connector

algorithm as shown in Tables 1 and 2. This index is obtained using the probability distributions of the K-L distances of each class, approximated as normal distributions and estimated by Monte Carlo trials. The simulations are carried out changing

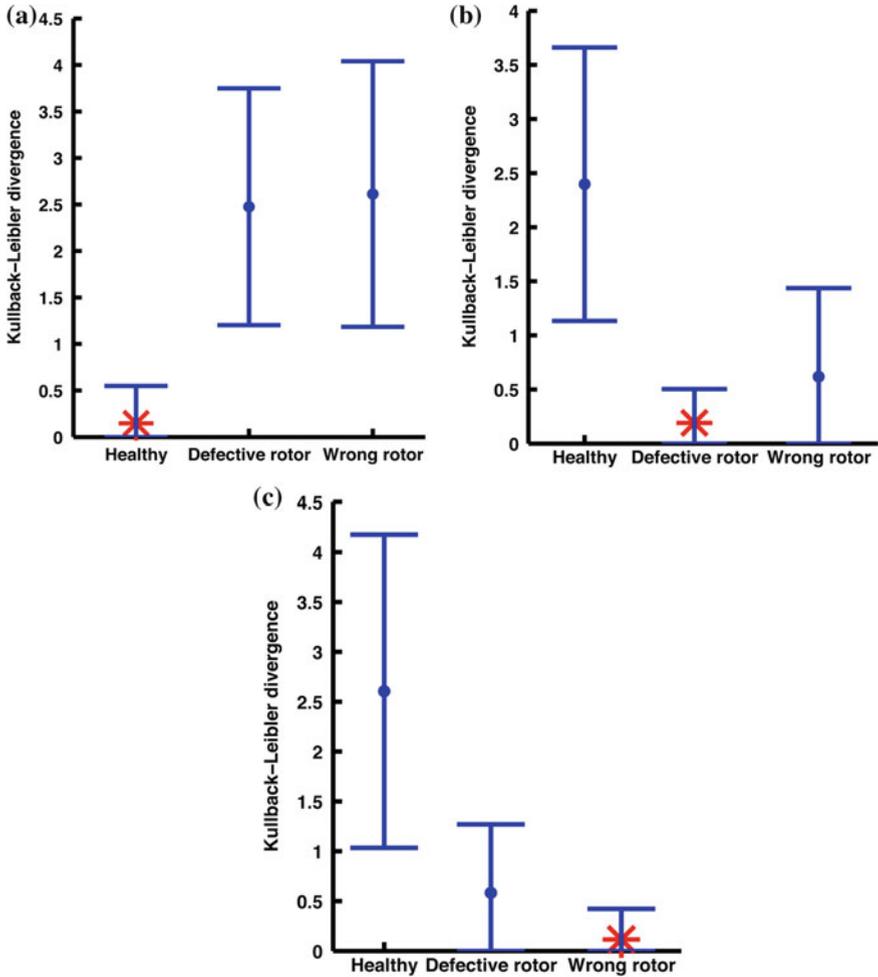


**Fig. 3** K–L divergence in the case of a finite element motor. The *blue dots* are the mean, the *blue bars* are the four times standard deviation and the *red asterisks* are the classification output. Label H means healthy motor, labels 1–10B mean broken bars with the relative number, labels 1–10C mean broken connectors with relative number. **a** Healthy motor. **b** Motor with one broken bar. **c** Motor with one broken connector



**Fig. 4** Interpolated PDFs of real motors in the two-dimensional principal component space estimated by KDE. **a** Healthy motor. **b** Motor with cracked rotor. **c** Motor with wrong rotor

$n_{grid}$ , the points number in which the PDF is estimated, and the current signals acquisition time in steady-state. Figures 3 and 5 show the K–L distances for all Monte Carlo trials. On each vertical line, the central dot is the mean and the



**Fig. 5** K–L divergence in the case of real motors. The *blue dots* are the mean, the *blue bars* are the four times standard deviation and the *red asterisks* are the classification output. **a** Healthy motor. **b** Motor with cracked rotor. **c** Motor with wrong rotor

horizontal edges are the 4 times standard deviation. The figures show the results with  $n_{grid} = 64 \times 64$  points and the acquisition time, for the benchmark and real motors, equals to 0.3 and 0.7 s respectively. This algorithm parameter setting guarantees better results for these cases taking into account the classification accuracy and the processing time. In the real motor the algorithm takes about 2.5 s for the classification output (Eq. 12): about 1 s to acquire the current signals, of which 0.25 s in transient state and 0.7 s in steady-state, and about 1.45 s to evaluate the PDF and the classification output (Eq. 12). Setting  $n_{grid} = 32 \times 32$  points, the processing time is reduced to 1.5 s but decreasing the classification accuracy as

**Table 1** Classification accuracy in the case of a finite element motor, changing  $n_{grid}$ , the points number in which the PDF is estimated, and the current signals acquisition time in steady-state

$n_{grid}$	128 × 128		64 × 64		32 × 32	
Acquisition time (s)	0.3	0.15	0.3	0.15	0.3	0.15
	%					
H	100	100	100	100	100	100
1B	100	100	100	100	100	100
2B	100	100	100	100	100	100
3B	100	100	100	100	100	99.93
4B	100	100	100	100	100	99.89
5B	100	100	100	99.96	100	99.70
6B	100	100	100	100	100	99.87
7B	100	99.98	100	99.94	100	98.19
8B	99.82	95.34	99.89	95.97	99.55	91.41
9B	100	99.74	100	99.53	100	98.56
10B	100	99.43	100	99.32	99.99	99.43
1C	100	100	100	100	100	99.99
2C	99.98	96.89	99.88	95.99	99.56	91.01
3C	99.88	91.71	99.72	95.74	98.48	93.75
4C	99.79	97.61	99.84	98.10	99.93	96.87
5C	99.98	98.96	99.99	98.49	99.96	96.37
6C	100	100	100	100	100	99.98
7C	100	100	100	100	100	99.94
8C	100	100	100	100	100	99.88
9C	100	99.99	100	100	100	99.47
10C	100	99.89	100	99.77	100	96.95
Mean	99.97	99.03	99.97	99.18	99.88	98.15

Label H means healthy motor, labels 1–10B mean broken bars with the relative number, labels 1–10C mean broken connectors with relative number

**Table 2** Classification accuracy in the case of real motors, changing  $n_{grid}$ , the points number in which the PDF is estimated, and the current signals acquisition time in steady-state

$n_{grid}$	128 × 128			64 × 64			32 × 32		
Acquisition time (s)	0.7	0.5	0.3	0.7	0.5	0.3	0.7	0.5	0.3
	%								
Healthy	100	100	100	100	100	100	100	100	100
Cracked rotor	98.82	95.08	77.08	99.00	94.74	86.54	98.29	94.02	81.45
Wrong rotor	98.97	99.47	99.49	99.18	99.36	98.56	99.85	99.41	99.21
Mean	99.26	98.18	92.19	99.39	98.03	95.03	99.38	97.81	93.55

Motor conditions are: healthy, motor with cracked rotor and motor with wrong rotor

shown in Tables 1 and 2. The tests are also performed for both cases using the asymmetric K–L divergence (Eq. 9). The results are comparable to those achieved with the symmetric K–L divergence described in the next subsections.

### 3.1.1 Broken Rotor Bars and Connectors Diagnosis

Figures 2a–c depict the patterns of a healthy motor, one broken bar and one broken connector conditions, respectively; these figures show how the PDFs, estimated by KDE in the principal component space, are used as the specific patterns for the motor conditions. The simulation results, given in Figs. 3a–c, show the faults diagnosis for broken rotor bars and connectors, setting  $n_{grid} = 64 \times 64$  and the current signals acquisition time in steady-state condition is equal to 0.3 s. Figure 3 (a) shows the K–L divergence among the PDFs, estimated by KDE, of all motor conditions (i.e. healthy, from one to ten broken rotor bars and from one to ten broken connectors) and the PDF estimated by KDE from stator current signals of healthy motor. The results show as the minimum K–L distance is exactly the healthy condition. Figure 3b shows the K–L divergence among all PDFs and the PDF estimated from stator current signals affected by one broken rotor bar. In this case the graph shows as the minimum K–L distance is exactly the broken bar condition. The last graph, Fig. 3c, shows the one broken connector diagnosis. Even in this case the K–L divergence detects and identifies the fault, that is one broken connector. By Monte Carlo simulations, all fault types are diagnosed with 100 % accuracy hence the K–L divergence figures for the other faults are not reported. Moreover the classification accuracy is 100 % with acquisition time above 0.3 s for each fault, while below 0.3 s, the classification accuracy decreases as shown in Table 1.

### 3.1.2 Real Induction Motors Diagnosis

Figures 4a–c depict the patterns of three real motors: healthy, cracked and wrong rotor; these figures show as the PDFs, estimated by KDE in the principal component space, are different and therefore can be used as specific patterns for each motor condition. Experimental results given in Figs. 5a–c show the fault diagnosis for cracked and wrong rotors, setting  $n_{grid} = 64 \times 64$  and the current signals acquisition time in steady-state is equal to 0.7 s. Figure 5a shows the K–L divergence among the PDFs, estimated by KDE, of all motor conditions (i.e. healthy, cracked and wrong rotors) and the PDF estimated by KDE from stator current signals of healthy motor. The results show as the minimum K–L distance is exactly the healthy condition. Figure 5b shows the K–L divergence among all PDFs and the PDF estimated from stator current signals where cracked rotors are diagnosed. In this case the graph shows as the minimum K–L distance is exactly the cracked rotor condition. The last graph, Fig. 5c, shows the wrong rotor diagnosis. Even in this

case the K–L divergence detects and identifies the fault. By Monte Carlo simulations, all fault types are diagnosed with accuracy reported in Table 2. It can be noticed how the classification accuracy in the case of healthy motor is always 100 %, therefore the algorithm is able to detect if motors are healthy or if there are some faults or defects. In Figs. 5b and c the blue lines of motors with cracked and wrong rotor are never overlapped to the blue lines of healthy motors so, in these tests, the algorithm never confuses the cases of healthy motors from those not healthy.

The next Section describes the well-known MSPCA algorithm for fault detection and isolation based on vibration signals.

## 4 Electric Motor FDD by MVSA

In electric motors, faults and defects are often correlated to the vibration signals, which can be processed to model the motor behaviours by patterns that represent the normal and abnormal motor conditions. Vibration analysis is widely accepted as a tool to detect faults of a rotating machine since it is reliable, not destructive and it permits continuous monitoring without stopping the machine. A brief literature review is given by: Fan and Zheng (2007), Immovilli et al. (2010), Sawalhi and Randall (2008a, b), Tran et al. (2009), Yang and Kim (2006). In particular, it is possible to detect different faults by analysing the vibration power spectrum. Most common faults are unbalance and misalignment. Unbalance may be caused by poor balancing, shaft inflection (i.e. thermal expansion) and rotor distortion by magnetic forces (a well known problem in high power electrical machines). Misalignment may be caused by misaligned couplings, misaligned bearings or crooked shaft.

In order to model the vibration signals, MSPCA is taken into account, as presented in Bakshi (1998). MSPCA deals with processes that operate at different scales, and have contributions from:

- events occurring at different localizations in time and frequency;
- stochastic processes whose energy or power spectrum changes over time and/or frequency;
- variables measured at different sampling rate or containing missing data.

MSPCA transforms the process data information at different scales by WT. The information of each different scale is captured by PCA modelling. These patterns, which represent the process conditions, can be used to identify each fault and defect.

To detect the defects, a KDE algorithm is used on the PCA residuals, and the thresholds are computed for each sensors signal. It allows to identify if, for each wavelet scale, the signals are involved in the fault or not. When Gaussian assumption is not recognized, KDE method is a robust methodology to estimate numerically the PDF, by Odiowei and Cao (2010). Fault isolation is carried out by

contribution plots, which is based on quantifying the contribution of each process variable to the single scores of the PCA. Diagnosis can be performed using the contribution plots because they represent the signatures of the rotating electrical machine conditions. The contributions are the inputs of a LDA classifier, which is a supervised machine learning algorithm used here to diagnose each motor defect. Several simulations are carried out using a benchmark provided by the Case Western Reserve University Bearing Data Center (2014).

## 4.1 Recalled Results

In this section authors present the algorithms used to develop the fault and defect diagnosis procedure. It extracts patterns by vibration signals using MSPCA and PCA contributions are used to diagnose each motor fault.

### 4.1.1 Principal Component Analysis

PCA is introduced in the Sect. 2.1.1, here an improved PCA fault detection index is described. A deviation of the new data sample  $\tilde{X}$  from the normal correlation could change the projections onto the subspaces, either  $S_d$  or  $S_r$ . Consequently, the magnitude of either  $\tilde{X}$  or  $\hat{X}$  could increase over the values obtained with normal data. The Square Prediction Error (SPE) is a statistic that measures lack of fit of a model to data. The SPE statistic is the difference, or residual, between a sample and its projection into the  $d$  components retained in the model. The description of the distribution of SPE is given in Jackson (2003):

$$SPE \equiv \|\tilde{X}\|^2 = \|X(I - PP^T)\|^2. \quad (13)$$

The process is faultless if:

$$SPE \leq \delta^2 \quad (14)$$

where  $\delta^2$  is a confidence limit for SPE. A confidence limit expression for SPE, when  $x$  follows a normal distribution, is developed in Jackson and Mudholkar (1979), Misra et al. (2002) and Rodriguez et al. (2006). The fault detectability condition is given in Dunia and Joe Qin (1998) and recalled in the following. Defining:

$$X = X^* + fE, \quad (15)$$

where the sample vector for normal operating conditions is denoted by  $X^*$ ,  $f$  represents the magnitude of the fault and  $\Xi$  is a fault direction vector. Necessary and sufficient conditions for detectability are:

- $\tilde{\Xi} = (I - PP^T)\Xi \neq 0$ , with  $\tilde{\Xi}$  the projection of  $\Xi$  on the residual subspace;
- $|\tilde{f}| = |(I - PP^T)f| > 2\delta$ , with  $\tilde{f}$  the projection of  $f$  on the residual subspace.

The drawbacks of *SPE* index for fault detection are mainly two: the first is related to the assumption of normal distribution to estimate the threshold of this index, the second is that the *SPE* is a weighted sum, with unitary coefficients, of quadratic residues  $\tilde{X}_i$ . To improve the fault detection, these two drawbacks are faced assuming that the process is faultless if, for each  $i$ :

$$\tilde{X}_i^2 \leq \delta_i \quad i = 1, \dots, m, \quad (16)$$

where  $\delta_i$  is a confidence limit for  $\tilde{X}_i^2$ . To estimate the confidence limit  $\delta_i$ , even if the normality assumption of  $\tilde{X}_i^2$  is not valid, the solution is to estimate the PDF directly from  $\tilde{X}_i^2$  through a non parametric approach. In Yu (2011a, b) and Odiwei and Cao (2010), KDE is considered because it is a well established non parametric approach to estimate the PDF of statistical signals and evaluate the control limits. Assume  $y$  is a random variable and its density function is denoted by  $p(y)$ . This means that:

$$P(y < k) = \int_{-\infty}^k p(y)dy. \quad (17)$$

Hence, by knowing  $p(y)$ , an appropriate control limit can be given for a specific confidence bound  $\alpha$ , using Eq. (17). Replacing  $p(y)$ , in Eq. (17), with the estimation of the probability density function of  $\tilde{X}_i^2$ , called  $\hat{p}(\tilde{X}_i^2)$ , the control limits will be estimated by:

$$\int_{-\infty}^{\delta_i} \hat{p}(\tilde{X}_i^2)d\tilde{X}_i^2 = \alpha. \quad (18)$$

Fault isolation and diagnosis are performed by the PCA contributions: defining the new observation vector  $x_j \in \mathbb{R}^m$ , the total contribution of the  $i$ th process variable  $X_i$  is

$$CONT_i = \sum_{j=1}^N \tilde{x}_{ij}^2 \quad i = 1, \dots, m. \quad (19)$$

### 4.1.2 Wavelet Transform

The Wavelet Transform (WT) is defined as the integral of the signal  $f(t)$  multiplied by scaled, shifted version of basic wavelet function  $\phi(t)$ , that is a real valued function whose Fourier transform satisfies the admissibility criteria stated in Li et al. (1999). Then the wavelet transformation  $c(\cdot, \cdot)$  of a signal  $f(t)$  is defined as:

$$\begin{aligned} c(a, b) &= \int_{\mathbb{R}} f(t) \frac{1}{\sqrt{a}} \phi\left(\frac{t-b}{a}\right) dt \\ a &\in \mathbb{R}^+ - \{0\} \\ b &\in \mathbb{R}, \end{aligned} \quad (20)$$

where  $a$  is the so-called scaling parameter,  $b$  is the time localization parameter. Both  $a$  and  $b$  can be continuous or discrete variables. Multiplying each coefficient by an appropriately scaled and shifted wavelet it yields the constituent wavelets of the original signal. For signals of finite energy, continuous wavelets synthesis provides the reconstruction formula:

$$f(t) = \frac{1}{K_\phi} \int_{\mathbb{R}} \int_{\mathbb{R}^+} c(a, b) \phi\left(\frac{t-b}{a}\right) \frac{da}{a^2} db \quad (21)$$

where:

$$K_\phi = \int_{-\infty}^{+\infty} \frac{|\hat{\phi}(\xi)|^2}{|\xi|} d\xi \quad (22)$$

denotes a (Wavelet specific) normalization parameter in which  $\hat{\phi}$  is the Fourier transform of  $\phi$ . Mother wavelets must satisfy the following properties:

$$\int_{-\infty}^{+\infty} |\phi(t)| dt < \infty, \quad \int_{-\infty}^{+\infty} |\phi(t)|^2 dt = 1, \quad \int_{-\infty}^{+\infty} \phi(t) dt = 0. \quad (23)$$

To avoid intractable computations when operating at every scale of the Continuous WT (CWT), scales and positions can be chosen on a power of two, i.e. dyadic scales and positions. The Discrete WT (DWT) analysis is more efficient and accurate, as reported in Li et al. (1999) and Daubechies (1988). In this scheme  $a$  and  $b$  are given by:

$$a = a_0^j, \quad b = b_0 a_0^j k, \quad (j, k) \in \mathbb{Z}^2, \quad \mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}. \quad (24)$$

The variables  $a_0$  and  $b_0$  are fixed constants that are set, as in Daubechies (1988), to:  $a_0 = 2$  and  $b_0 = 1$ . The discrete wavelet analysis can be described mathematically as:

$$\begin{aligned}
c(a, b) &= c(j, k) = \sum_{n \in \mathbb{Z}^+} f(n) \phi_{j,k}(n), \\
a &= 2^j, \quad b = 2^j k, \\
j &\in \mathbb{Z}, \quad k \in \mathbb{Z},
\end{aligned} \tag{25}$$

considering the simplified notation  $f(n) = f(n \cdot t_c)$ ,  $n \in \mathbb{Z}^+$  and  $t_c$  the sampling time, the discretization of continuous time signal  $f(t)$  is considered. The inverse transform, also called discrete synthesis, is defined as:

$$f(n) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} c(j, k) \phi_{j,k}(n). \tag{26}$$

In Mallat (1989), a signal is decomposed into various scales with different time and frequency resolutions, this algorithm is known as the multi-resolution signal decomposition. Defining:

$$\begin{aligned}
\phi_{j,k}(n) &= 2^{-j/2} \phi(2^{-j}n - k), \\
\psi_{j,k}(n) &= 2^{-j/2} \psi(2^{-j}n - k), \\
V_j &= \text{span}\{\phi_{j,k}, k \in \mathbb{Z}\}, \\
W_j &= \text{span}\{\psi_{j,k}, k \in \mathbb{Z}\},
\end{aligned} \quad (j, k) \in \mathbb{Z}^2 \tag{27}$$

the wavelet function  $\phi_{j,k}$ , is the orthonormal basis of  $V_j$  and the orthogonal wavelet  $\psi_{j,k}$ , called scaling function, is the orthonormal basis of  $W_j$ . In Daubechies (1988) is shown that:

$$\begin{aligned}
V_j &\perp W_j, \\
V_m &= W_{m+1} \oplus V_{m+1}. \quad V_m, W_m \subset \mathbf{L}^2(\mathbb{R})
\end{aligned} \tag{28}$$

Defining  $f(n) = f$  as element of  $V_0 = W_1 \oplus V_1$ ,  $f$  can be decomposed into its components along  $V_1$  and  $W_1$ :

$$f = P_1 f + Q_1 f. \tag{29}$$

with  $P_j$  the orthogonal projection onto  $V_j$  and  $Q_j$  the orthogonal projection onto  $W_j$ . Defining  $j \geq 1$  and  $f(n) = c_n^0$ , it results:

$$\begin{aligned}
f(n) &= \sum_{k \in \mathbb{Z}} c_k^1 \phi_{1,k}(n) + \sum_{k \in \mathbb{Z}} d_k^1 \psi_{1,k}(n), \\
c_k^1 &= \sum_{n \in \mathbb{Z}} h(n - 2k) c_n^0, \\
d_k^1 &= \sum_{n \in \mathbb{Z}} g(n - 2k) c_n^0, \\
h(n - 2k) &= \langle \phi_{1,k}(n), \phi_{0,n}(n) \rangle, \\
g(n - 2k) &= \langle \psi_{1,k}(n), \psi_{0,n}(n) \rangle. \\
k, n &\in \mathbb{Z}^2.
\end{aligned} \tag{30}$$

where the terms  $g$  and  $h$  are high-pass and low-pass filter coefficients derived from the bases  $\psi$  and  $\phi$ . Considering a dataset of  $N$  ( $n = 1, \dots, N$ ) samples, and introducing a vector notation,  $c_k^1$  and  $d_k^1$  can be rewrite as Daubechies (1988):

$$\begin{aligned} \mathbf{c}^1 &= \mathbf{H}\mathbf{c}^0, \\ \mathbf{d}^1 &= \mathbf{G}\mathbf{c}^0, \end{aligned} \quad (31)$$

with

$$\mathbf{H} = \begin{bmatrix} h(0) & h(1) & \cdots & h(N) \\ h(-2) & h(-1) & \cdots & h(N-2) \\ \vdots & \vdots & \cdots & \vdots \\ h(-2k) & h(1-2k) & \cdots & h(N-2k) \end{bmatrix}, \quad (32)$$

$$\mathbf{G} = \begin{bmatrix} g(0) & g(1) & \cdots & g(N) \\ g(-2) & g(-1) & \cdots & g(N-2) \\ \vdots & \vdots & \cdots & \vdots \\ g(-2k) & g(1-2k) & \cdots & g(N-2k) \end{bmatrix}. \quad (33)$$

The procedure can be iterated obtaining:

$$\begin{aligned} \mathbf{c}^j &= \mathbf{H}\mathbf{c}^{j-1}, \\ \mathbf{d}^j &= \mathbf{G}\mathbf{d}^{j-1}. \end{aligned} \quad (34)$$

Then:

$$\begin{aligned} \mathbf{c}^j &= \mathbf{H}_j\mathbf{c}^0, \\ \mathbf{d}^j &= \mathbf{G}_j\mathbf{d}^0, \end{aligned} \quad (35)$$

where  $\mathbf{H}_j$  is obtained by applying the  $\mathbf{H}$  filter  $j$  times, and  $\mathbf{G}_j$  is obtained by applying the  $\mathbf{H}$  filter  $j-1$  times and the  $\mathbf{G}$  filter once. Hence any signal may be decomposed into its contributions in different regions of the time-frequency space by projection on the corresponding wavelet basis function. The lowest frequency content of the signal is represented on a set of scaling functions. The number of wavelet and scaling function coefficients decreases dyadically at coarser scales due to the dyadic discretization of the dilation and translation parameters. The algorithms for computing the wavelet decomposition are based on representing the projection of the signal on the corresponding basis function as a filtering operation (Mallat 1989). Convolution with the filter  $\mathbf{H}$  represents projection on the scaling function, and convolution with the filter  $\mathbf{G}$  represents projection on a wavelet. Thus, the signal  $f(n)$  is decomposed at different scales, the detail scale matrices and approximation scale matrices. Defining  $L$  the decomposition levels, the approximation scale  $\mathbf{A}_L$

and the detail scales  $\mathbf{D}_j$ ,  $j = 1, \dots, L$  are the composition of  $\mathbf{c}^j$  and  $\mathbf{d}^j$  for every  $m$  variables of the data matrix  $\mathbf{X}$ :

$$\begin{aligned} \mathbf{A}_j &= [\mathbf{c}_1^j, \mathbf{c}_2^j, \dots, \mathbf{c}_m^j], \\ \mathbf{D}_j &= [\mathbf{d}_1^j, \mathbf{d}_2^j, \dots, \mathbf{d}_m^j]. \end{aligned} \quad j = 1, \dots, L \quad (36)$$

To select the wavelet decomposition level  $L$  it is considered the minimum number of decomposition levels, and used to obtain an approximation signal  $\mathbf{A}_L$  so that the upper limit of its associated frequency band is under the fundamental frequency  $f$ , as described by the following condition Antonino-Daviu et al. (2006), Bouzida et al. (2011):

$$2^{-(L+1)}f_s < f. \quad (37)$$

where  $f_s$  is the sampling frequency of the signals and  $f$  is the fundamental frequency of the machine. From this condition, the decomposition level of the approximation signal is the integer  $L$  given by:

$$L = \lfloor \log_2(f_s/f) - 1 \rfloor. \quad (38)$$

## 4.2 MSPCA Formulation

WT and PCA can be combined to extract maximum information from multivariate sensor data. MSPCA can be used as a tool for fault detection and diagnosis by means of statistical indexes. In particular, faults are detected by using Eqs. 16 and 18 and the isolation is conducted by the contribution method (Eq. 19). In this way it is possible to detect which sensor is most affected by fault (see Misra et al. 2002). Two fundamental theorems exist for the MSPCA formulation, they assess that PCA assumptions remain unchanged under the Wavelet transformation. These theorems are useful to apply MSPCA methodology, as stated in Bakshi (1998).

**Theorem 4.1** *Let  $\mathbf{W} = [\mathbf{H}'_L, \mathbf{G}'_L, \mathbf{G}'_{L-1}, \dots, \mathbf{G}'_1]'$   $\in \mathbb{R}^{N \times N}$  the orthonormal matrix representing the orthonormal wavelet transformation operator containing the filter coefficients, the principal component loadings obtained by the PCA of  $\mathbf{X}$  and  $\mathbf{WX}$  are identical, whereas the principal component scores of  $\mathbf{WX}$  are the wavelet transform of the scores of  $\mathbf{X}$ .*

**Theorem 4.2** *MSPCA reduces to conventional PCA if neither the principal components nor the wavelet coefficients at any scale are eliminated.*

The developed FDD MSPCA based procedure consists of two stages: in the first step, the faultless data are processed and a model of this data is built. MSPCA training steps are summarized below:

- T1. Data are preprocessed;
- T2. The Wavelet analysis is used, to refine the data, with a level of detail  $L$  which is chosen by Eq. (38);
- T3. Normalize mean and standard deviation of detail and approximation matrices and apply PCA to the approximation matrix  $A_L$ , of order  $L$ , and to the  $L$  detail matrices  $D_j$ , where  $j = 1, \dots, L$ ;
- T4. The PCA transformation matrix  $P$  and the signal covariance matrix  $S$  are computed for each approximation and detail matrices;
- T5. The  $\tilde{X}_i$  signals (Eq. 13) are computed, for each wavelet matrix;
- T6. The  $\delta_i$  thresholds are computed, for each detail matrix and for the approximation matrix of order  $L$ , using the KDE algorithm (Eq. 18) and a confidence bound  $\alpha$ ;

In the second step, the model previously obtained is on-line compared with the new data and a statistical index of failure is calculated. MSPCA diagnosis steps are summarized below:

- D1. The previous steps, except the threshold computation step (T6), are repeated for each new dataset, the data are standardized as in the training step (T3) and the PCA and  $\tilde{X}_i$  signals are computed using the  $P$  and  $S$  matrices, obtained in the training step;
- D2. **If** any of the  $\tilde{X}_i^2$  signals is over the thresholds  $\delta_i$ , the fault is detected and the isolation is performed by the contributions, **else** the next data set is analysed [return to (D1)];
- D3. Compute all the residual contributions, for each sensor, for all details and approximation matrices and isolate and diagnose the fault type.

The next Section introduces the FDD experimental results in order to show the MSPCA algorithm performances. Tests are carried out on real induction motors with different fault severity.

## 5 Electric Motor FDD by MVSA: Results

The diagnosis algorithm has been tested on the vibration signals provided by the Case Western Reserve University Bearing Data Center (2014). Experiments were conducted using a 2 hp Reliance Electric motor, and acceleration data was measured at locations near to and remote from the motor bearings. Motor bearings were seeded with faults using electro-discharge machining (EDM). Faults ranging from 0.007 in. to 0.040 in. of diameter were introduced separately at the inner raceway, rolling element (i.e. ball) and outer raceway. Faulty bearings were reinstalled into the test motor and vibration data was recorded for motor loads of 0–3 hp (motor speeds of 1,797–1,720 RPM). Accelerometers were placed at the 12 o'clock position at both the drive end and fan end of the motor housing. Digital data was

collected at 12,000 samples per second. Experiments were conducted for both fan and drive end bearings with outer raceway faults located at 3 o'clock (directly in the load zone), at 6 o'clock (orthogonal to the load zone), and at 12 o'clock.

## 5.1 Results and Discussion

The proposed approach described in Sect. 4.2 has been tested using a Daubechies mother wavelet of order 15, defined *db15* mother wavelet (defined kernel  $\phi$  in Sect. 4.1.2). Since the motor rotation frequency is 30 Hz and the sampling frequency is 12 kHz, applying Eq. (38), the level of detail obtained is  $L = 7$ . The dimension of principal component subspace  $d$ , chosen by the Kaiser's rule, is described in Jolliffe (2002).

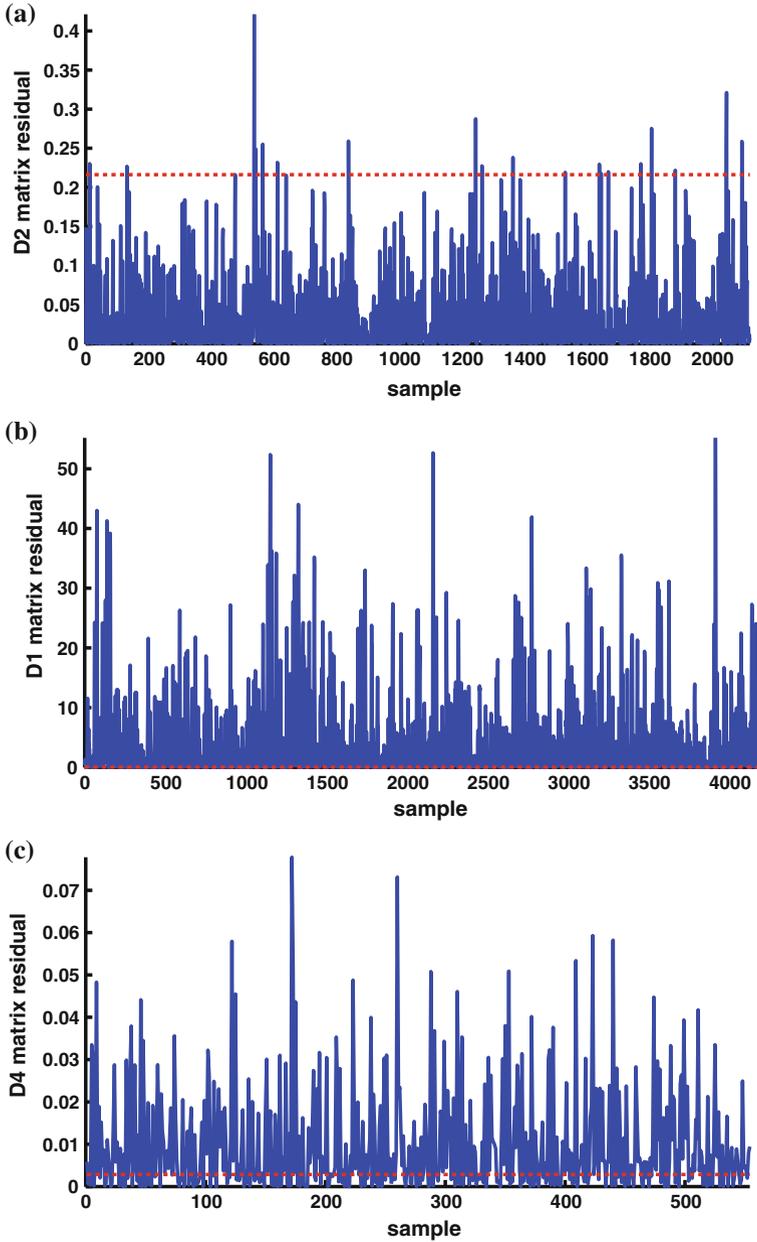
Incoming batch data samples are then fed into the MSPCA model and the PCA residual contributions are computed for the matrices  $D_j$ ,  $j = 1, \dots, L$ ,  $A_L$ . In the following, these matrices are defined *scale matrices*, and they are compared with the respective thresholds. When, at any scale, the number of residual contribution samples over the thresholds is greater than  $\alpha \cdot \gamma$ , where  $\alpha$  is the significance level used for the threshold  $\delta_i$  calculation (stated in Sect. 4.2) and  $\gamma$  is a corrective index (fixed equal to 2), a fault is detected and the motor is considered faulty.

Once a fault is detected, the isolation and diagnosis tests are performed. At this step the PCA contributions are computed for each scale matrix. Fault isolation allows to detect which sensors are involved in the fault. By using several scales for the DWT analysis, it is possible to cluster the residual contributions of each scale and define a unique signature of the motor fault, as in a MVSA approach. More in detail, the signature of each fault is given by the contributions of each variable for each scale. The results are the average of 1,000 Monte Carlo simulations where the training and testing data sets are randomly changed.

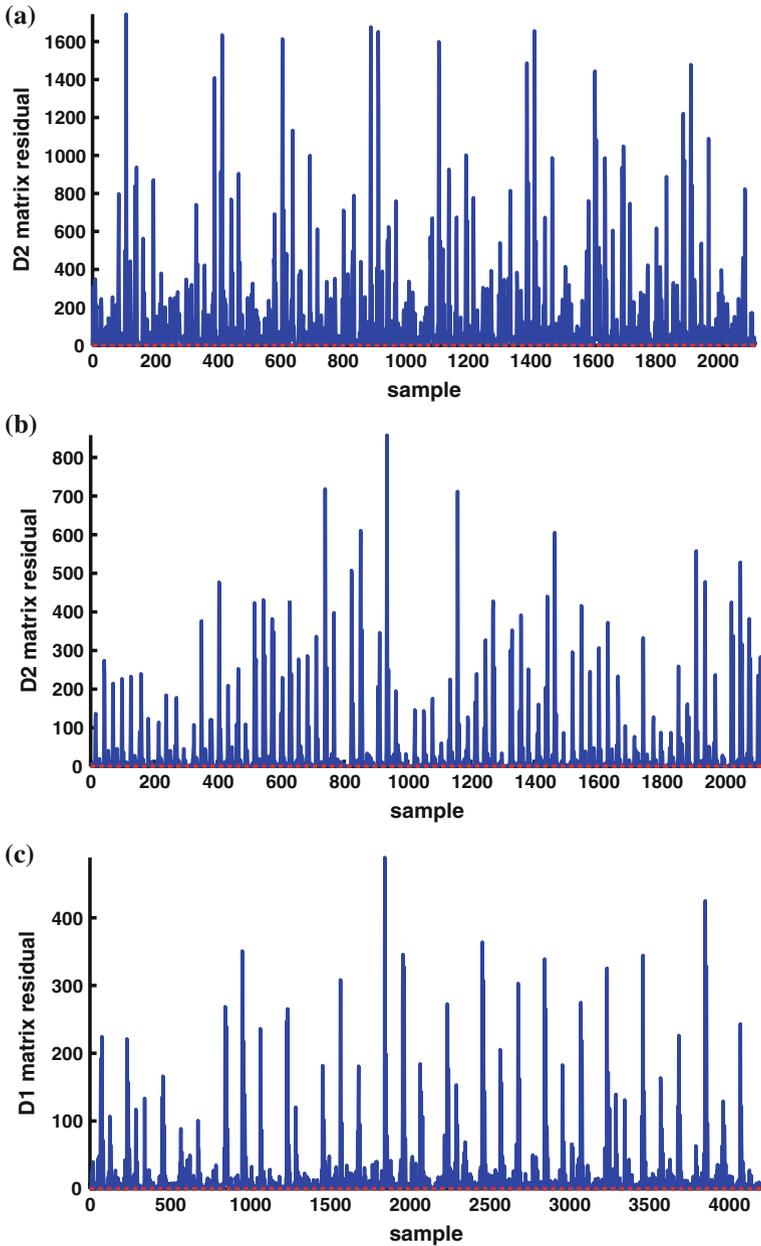
Figures 6 and 7 show the residuals of the first accelerometer (i.e. placed at the drive end) for drive end bearing faults estimated by Eq. (16). The thresholds, drawn in dashed red line, are estimated by KDE (Eq. 18). While Fig. 6a shows the residuals for healthy motor, Fig. 6b, c show the residuals of rolling element and inner raceway faults respectively at the detail scales  $D_1$  and  $D_4$ , which are, among all scales, the most affected by the faults.

Figures 7a–c show the residuals of outer raceway faults located at 3, 6, 12 o'clock respectively at the detail scales  $D_1$ ,  $D_2$  and  $D_4$ , which are, among all scales, the most affected by the faults. It can be noticed how the residuals are related to the fault type and so they can be exploited as signatures of the rotating electrical machine conditions.

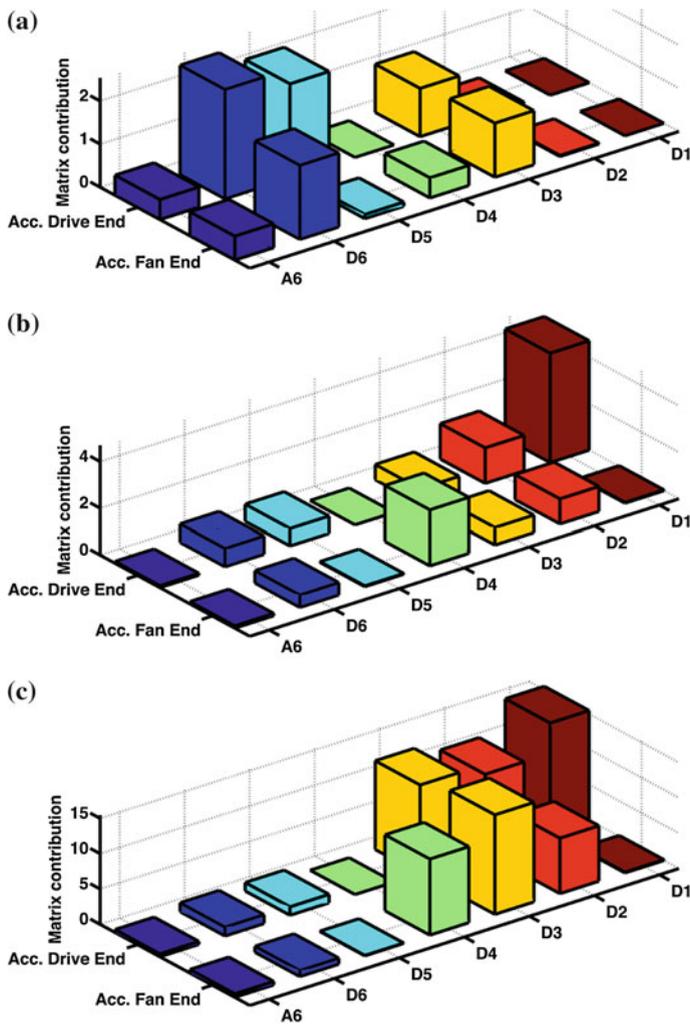
Figures 8 and 9 show the contribution plots of each accelerometer at different scales for drive end bearing fault, particularly Figs. 8a–c show the contribution plots of healthy motor, rolling element and inner raceway faults while Figs. 9a–c



**Fig. 6** Residuals of the accelerometer placed at the drive end. **a** Healthy motor at  $D_2$  detail scale. **b** Rolling element fault of drive end bearing at  $D_1$  detail scale. **c** Inner raceway fault of drive end bearing at  $D_4$  detail scale



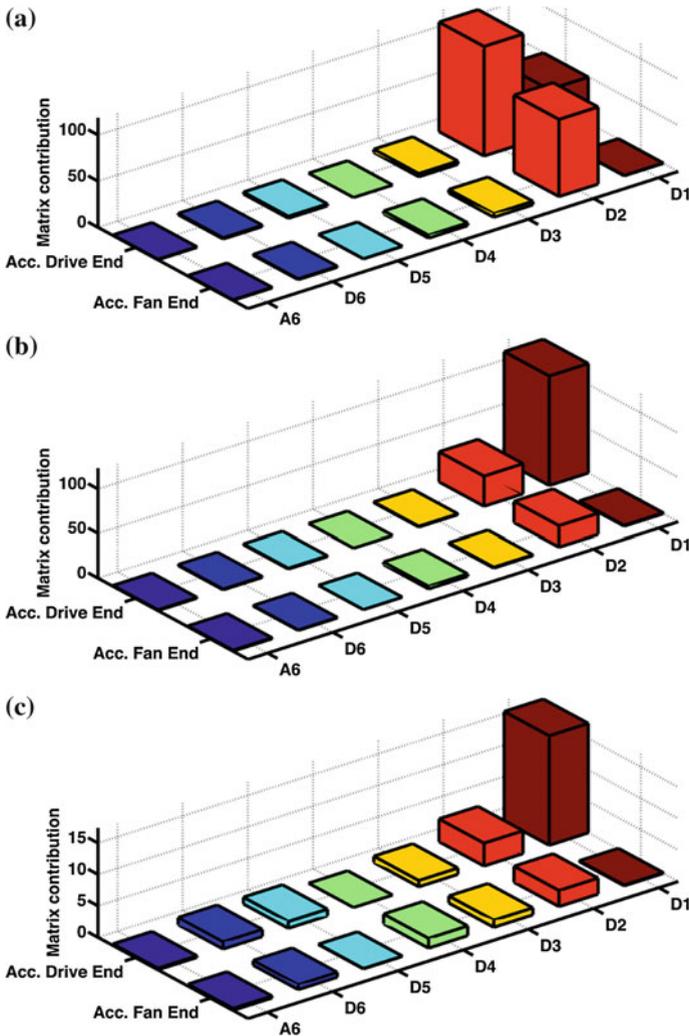
**Fig. 7** Residuals of the accelerometer placed at the drive end. **a** Outer raceway fault located at 3 o'clock of drive end bearing at  $D_1$  detail scale. **b** Outer raceway fault located at 6 o'clock of drive end bearing at  $D_2$  detail scale. **c** Outer raceway fault located at 12 o'clock of drive end bearing at  $D_2$  detail scale



**Fig. 8** Contribution plots. **a** Healthy motor. **b** Rolling element fault of drive end bearing. **c** Inner raceway fault of drive end bearing

show the contribution plots of outer raceway fault located at 3, 6, 12 o'clock respectively.

The contribution plots could be used as signatures of the electric motor conditions, so a supervised machine learning algorithm, with the PCA contributions as inputs, can be used to diagnose each motor fault. The Figs. 8 and 9 show that the identified signatures by PCA contributions are features for each fault. The accelerometers are involved in these signatures at different scales with different amplitudes. As shown by Figs. 8 and 9, all faults affect the accelerometer placed at the drive end.



**Fig. 9** Contribution plots. **a** Outer raceway fault located at 3 o'clock of drive end bearing. **b** Outer raceway fault located at 6 o'clock of drive end bearing. **c** Outer raceway fault located at 12 o'clock of drive end bearing

This points out that the contribution plots can be used to identify the sensors affected by the faults. Particularly, the fault isolation is performed by computing for each scale the average value: the sensor affected by the fault is that one with the highest average value. As shown in Figs. 8 and 9 the sensor most affected by the faults is the accelerometer placed at the drive end. The fault diagnosis is performed using the contribution plots because they are the signatures of the electric motor conditions and are features for each fault as shown in Figs. 8 and 9. Outer raceway fault located

at 6 o'clock and outer raceway fault located at 12 o'clock of drive end bearing affect the contributions at the same scales (Figs. 9a–b), but the contribution amplitudes are different so they can be used to diagnose the faults.

In order to diagnose each motor fault LDA is used. It searches a linear transformation that maximizes class separability in a reduced dimensional space. LDA is proposed in Fisher (1936) for solving binary class problems. It is further extended to multi-class cases in Rao (1948). In general, LDA aims to find a subspace that minimizes the within-class scatter and maximizes the between-class scatter simultaneously. PCA contributions are used as features input to the LDA algorithm. Tables 3, 4 and 5 show the classification accuracy. The results are the average classification accuracy of each motor conditions (i.e. healthy motor, rolling element fault, inner raceway fault, outer raceway fault located at 3, 6 and 12 o'clock) and of 4 motor loads: from 0 to 3 hp (motor speeds of 1,797–1,720 RPM). Table 3 shows the classification accuracy at different wavelet decomposition level  $L$  and acquisition time of faults occurred at drive end bearing with fault diameter of 0.007 in. The classification accuracy is over 99 % for each level  $L$  and acquisition time, so a low wavelet decomposition level and acquisition time can be chosen to diagnose effectively this fault.

Table 4 shows the classification accuracy at different wavelet decomposition level  $L$  and acquisition time of faults occurred at drive end bearing with fault diameter of 0.021 in. The classification accuracy is over 99 % at each level  $L$  and acquisition time higher 0.3 s, so a low wavelet decomposition level and acquisition

**Table 3** Average classification accuracy of drive end bearing fault with fault diameter of 0.007 in.

		Acquisition time s				
		0.1	0.2	0.3	0.5	0.7
		%				
$L$	1	99.23	99.96	99.98	100	100
	2	99.40	99.85	99.99	100	100
	3	99.55	99.95	100	99.99	100
	4	99.67	99.98	100	100	100
	5	99.47	99.95	100	100	100
	6	99.29	99.85	99.99	100	100

**Table 4** Average classification accuracy of drive end bearing fault with fault diameter of 0.021 in.

		Acquisition time s				
		0.1	0.2	0.3	0.5	0.7
		%				
$L$	1	92.66	97.73	99.19	99.91	99.96
	2	94.76	98.36	99.51	99.75	99.88
	3	95.10	98.87	99.73	99.98	99.99
	4	94.64	98.39	99.21	99.81	99.97
	5	95.00	98.54	99.50	99.79	99.96
	6	94.38	98.25	99.35	99.83	99.88

**Table 5** Average classification accuracy of fan end bearing fault with fault diameter of 0.007 in.

		Acquisition time s				
		0.1	0.2	0.3	0.5	0.7
		%				
<i>L</i>	1	73.61	82.50	89.43	93.04	94.31
	2	74.76	84.11	90.44	95.02	95.15
	3	83.08	92.54	94.82	98.49	98.60
	4	84.41	91.62	95.37	99.22	99.28
	5	84.41	90.73	95.56	98.90	98.76
	6	84.41	93.14	96.76	99.39	99.31

time of 0.3 s can be chosen to diagnose effectively this fault. Table 5 shows the classification accuracy at different wavelet decomposition level  $L$  and acquisition time of faults occurred at fan end bearing with fault diameter of 0.007 in. The classification accuracy is over 98 % for level  $L = 3$  and acquisition time higher 0.5 s, so a wavelet decomposition level of 3 and acquisition time of 0.3 s can be chosen to diagnose effectively this fault.

## 6 Summary and Conclusions

This chapter addresses the modelling and diagnosis issues of rotating electrical machines by signal based solutions. With attention to real systems, two case studies related to rotating electrical machines are discussed. The first FDD solution uses PCA in order to reduce the three-phase current space in two dimensions. The PDFs of PCA-transformed signals are estimated by KDE. PDFs are the models that can be used to identify each fault. Diagnosis has been carried out using the K–L divergence, which measures the difference between two probability distributions. This divergence is used as a distance between signatures obtained by KDE. The second FDD solution uses MSPCA, KDE and PCA contributions to identify and diagnose the faults. Several experimentations on real motors are carried out in order to verify the effectiveness of the proposed methodologies. The first solution, based on current signals, has been tested on a motor modelled by FEM and real induction motors in order to diagnose broken rotor bars, broken connector, cracked and wrong rotor. The second solution, based on vibration signals, has been tested on a real induction motors in order to diagnose bearings faults: inner raceway, rolling element (i.e. ball) and outer raceway faults with different fault severities (i.e. diameter of 0.007 and 0.021 in.). Results show that the signal based solutions are able to model the fault dynamics and diagnose the motor conditions (i.e. healthy and faulty) and identify the faults.

## References

- Abdelmadjid, G., Mohamed, B. S., Mohamed, T., Ahmed, S., & Youcef, M. (2013). An improved stator winding fault tolerance architecture for vector control of induction motor: Theory and experiment. *Electric Power Systems Research*, *104*, 129–137.
- Antonino-Daviu, J., Riera-Guasp, M., Roger-Folch, J., Martinez-Giménez, F., & Peris, A. (2006). Application and optimization of the discrete wavelet transform for the detection of broken rotor bars in induction machines. *Applied and Computational Harmonic Analysis*, *21*(2), 268–279.
- Baccarini, L. M. R., Tavares, J. P. B., de Menezes, B. R., & Caminhas, W. M. (2010). Sliding mode observer for on-line broken rotor bar detection. *Electric Power Systems Research*, *80*(9), 1089–1095.
- Bakshi, B. R. (1998). Multiscale PCA with application to multivariate statistical process monitoring. *AIChE Journal*, *44*(7), 1596–1610.
- Bangura, J., Povinelli, R., Demerdash, N. A. O., & Brown, R. (2003). Diagnostics of eccentricities and bar/end-ring connector breakages in polyphase induction motors through a combination of time-series data mining and time-stepping coupled FE-state-space techniques. *IEEE Transactions on Industry Applications*, *39*(4), 1005–1013.
- Benloucif, M., & Balaska, H. (2006). Robust fault detection for an induction machine. In *Automation Congress, 2006. WAC '06. World* (pp. 1–6). Budapest.
- Bouzida, A., Touhami, O., Ibtouen, R., Belouchrani, A., Fadel, M., & Rezzoug, A. (2011). Fault diagnosis in industrial induction machines through discrete wavelet transform. *IEEE Transactions on Industrial Electronics*, *58*(9), 4385–4395.
- Case Western Reserve University Bearing Data Center. (2014). Bearing Data Center.
- Cheng, H., Nikus, M., & Jamsa-Jounela, S. L. (2008). Evaluation of PCA methods with improved fault isolation capabilities on a paper machine simulator. *Chemometrics and Intelligent Laboratory Systems*, *92*(2), 186–199.
- Chiang, L. H., Russell, E. L., & Braatz, R. D. (2000). Fault diagnosis in chemical processes using Fischer discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, *50*(2), 243–252.
- Chilengue, Z., Dente, J., & Branco, P. C. (2011). An artificial immune system approach for fault detection in the stator and rotor circuits of induction machines. *Electric Power Systems Research*, *81*(1), 158–169.
- Ciabattini, L., Corradini, M., Grisostomi, M., Ippoliti, G., Longhi, S., & Orlando, G. (2011a). Adaptive extended Kalman filter for robust sensorless control of PMSM drives. In *Proceedings of the IEEE Conference on Decision and Control* (pp. 934–939).
- Ciabattini, L., Corradini, M., Grisostomi, M., Ippoliti, G., Longhi, S., & Orlando, G. (2014). A discrete-time vs controller based on RBF neural networks for PMSM drives. *Asian Journal of Control*, *16*(2), 396–408.
- Ciabattini, L., Grisostomi, M., Ippoliti, G., & Longhi, S. (2011b). Estimation of rotor position and speed for sensorless DSP-based PMSM drives. In *2011 19th Mediterranean Conference on Control and Automation, MED 2011* (pp. 1421–1426).
- Ciandrini, C., Gallieri, M., Giantomassi, A., Ippoliti, G., & Longhi, S. (2010). Fault detection and prognosis methods for a monitoring system of rotating electrical machines. In *Industrial Electronics (ISIE), 2010 IEEE International Symposium on* (pp. 2085–2090).
- Comaniciu, D. (2003). An algorithm for data-driven bandwidth selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(2), 281–288.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, *41*(7), 909–996.
- Dunia, R., & Joe Qin, S. (1998). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice*, *6*(4), 457–469.
- Fan, Y., & Zheng, G. (2007). Research of high-resolution vibration signal detection technique and application to mechanical fault diagnosis. *Mechanical Systems and Signal Processing*, *21*(2), 678–687.

- Feki, N., Clerc, G., & Velex, P. (2013). Gear and motor fault modeling and detection based on motor current analysis. *Electric Power Systems Research*, 95, 28–37.
- Ferracuti, F., Giantomassi, A., Iarlori, S., Ippoliti, G., & Longhi, S. (2013a). Induction motor fault detection and diagnosis using KDE and Kullback–Leibler divergence. In *Industrial Electronics Society, IECON 2013—39th Annual Conference of the IEEE* (pp. 2923–2928).
- Ferracuti, F., Giantomassi, A., Ippoliti, G., & Longhi, S. (2010). Multi-scale PCA based fault diagnosis for rotating electrical machines. In *European workshop on advanced control and diagnosis, 8th ACD* (pp. 296–301). Ferrara, Italy.
- Ferracuti, F., Giantomassi, A., & Longhi, S. (2013b). MSPCA with KDE thresholding to support QC in electrical motors production line. In *Manufacturing Modelling, Management, and Control*, 7, 1542–1547.
- Ferracuti, F., Giantomassi, A., Longhi, S., & Bergantino, N. (2011). Multi-scale PCA based fault diagnosis on a paper mill plant. In *Emerging Technologies Factory Automation (ETFA), 2011 IEEE 16th Conference on* (pp. 1–8).
- Filbert, D., & Metzger, L. (1982). Quality test of systems by parameter estimation. In *Proceedings of 9th IMEKO congress*. Berlin, Germany.
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2), 179–188.
- Gani, A., & Salami, M. (2002). Vibration faults simulation system (VFSS): A system for teaching and training on fault detection and diagnosis. In *Research and Development, 2002. SCORed 2002. Student Conference on* (pp. 15–18).
- Giantomassi, A. (2012). *Modeling estimation and identification of complex system dynamics*. LAP Lambert Academic Publishing.
- Giantomassi, A., Ferracuti, F., Benini, A., Longhi, S., Ippoliti, G., & Petrucci, A. (2011). Hidden Markov model for health estimation and prognosis of turbofan engines. In *ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 1–6).
- Hua, L., Zhanfeng, L., & Zhaowei. (2009). Time frequency distribution for vibration signal analysis with application to turbo-generator fault diagnosis. In *Control and Decision Conference, 2009. CCDC '09. Chinese* (pp. 5492–5495). Guilin.
- Immovilli, F., Bellini, A., Rubini, R., & Tassoni, C. (2010). Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Transactions on Industry Applications*, 46(4), 1350–1359.
- Isermann, R. (1984). Process fault detection on modeling and estimation methods? A survey. *Automatica*, 20(4), 387–404.
- Isermann, R. (2006). *Fault-diagnosis systems*. Berlin: Springer.
- Isermann, R., & Balle, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5), 709–719.
- Jackson, J. E. (2003). *A user's guide to principal components* (Vol. 587). New York: Wiley-Interscience.
- Jackson, J. E., & Mudholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21(3), 341–349.
- Jolliffe, I. T. (2002). *Principal component analysis*. Berlin: Springer.
- Keogh, E. (2013). UCR time series data mining archive.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22(1), 79–86.
- Lachouri, A., Baiche, K., Djeghader, R., Doghmane, N., & Ouhtati, S. (2008). Analyze and fault diagnosis by Multi-scale PCA. In *Information and communication technologies: From theory to applications, 2008. ICTTA 2008. 3rd International Conference on* (pp. 1–6).
- Li, W., Yue, H. H., Valle-Cervantes, S., & Qin, S. J. (2000). Recursive PCA for adaptive process monitoring. *Journal of Process Control*, 10(5), 471–486.
- Li, X., Dong, S., & Yuan, Z. (1999). Discrete wavelet transform for tool breakage monitoring. *International Journal of Machine Tools and Manufacture*, 99(12), 1944–1955.

- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.
- Misra, M., Yue, H., Qin, S., & Ling, C. (2002). Multivariate process monitoring and fault diagnosis by multi-scale PCA. *Computers and Chemical Engineering*, 26(9), 1281–1293.
- Mugdadi, A., & Ahmad, I. A. (2004). A bandwidth selection for kernel density estimation of functions of random variables. *Computational Statistics and Data Analysis*, 47(1), 49–62.
- Odiowei, P.-E., & Cao, Y. (2010). Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *IEEE Transactions on Industrial Informatics*, 6(1), 36–45.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3), 1065–1076.
- Patton, R., Uppal, F., & Lopez-Toribio, C. (2000). Soft computing approaches to fault diagnosis for dynamic systems: a survey. In *4th IFAC Symposium on Fault Detection supervision and Safety for Technical Processes* (pp. 298–311). Budapest, Hungary.
- Pedra, J., Candela, I., & Barrera, A. (2009). Saturation model for squirrel-cage induction motors. *Electric Power Systems Research*, 79(7), 1054–1061.
- Ran, L., & Penman, J. (2008). *Condition monitoring of rotating electrical machines* (Vol. 56). London, United Kingdom: Institution of Engineering and Technology.
- Rao, C. R. (1948). The utilization of multiple measurements in problems of biological classification. *Journal of the Royal Statistical Society—Series B*, 10(2), 159–203.
- Rodriguez, P., Belahcen, A., & Arkkio, A. (2006). Signatures of electrical faults in the force distribution and vibration pattern of induction motors. *Electric Power Applications, IEE Proceedings*, 153(4), 523–529.
- Samsi, R., Ray, A., & Mayer, J. (2009). Early detection of stator voltage imbalance in three-phase induction motors. *Electric Power Systems Research*, 79(1), 239–245.
- Sawalhi, N., & Randall, R. (2008a). Simulating gear and bearing interactions in the presence of faults: Part I. the combined gear bearing dynamic model and the simulation of localised bearing faults. *Mechanical Systems and Signal Processing*, 22(8), 1924–1951.
- Sawalhi, N., & Randall, R. (2008b). Simulating gear and bearing interactions in the presence of faults: Part II. simulation of the vibrations produced by extended bearing faults. *Mechanical Systems and Signal Processing*, 22(8), 1952–1966.
- Sheather, S. J. (2004). Density estimation. *Statistical science* (pp. 588–597).
- Shuting, W., Heming, L., & Yonggang, L. (2002). Adaptive radial basis function network and its application in turbine-generator vibration fault diagnosis. In *Power system technology, 2002. Proceedings of PowerCon 2002. International Conference on* (Vol. 3 pp. 1607–1610).
- Singh, G., & Ahmed, S. A. K. S. (2004). Vibration signal analysis using wavelet transform for isolation and identification of electrical faults in induction machine. *Electric Power Systems Research*, 68(2), 119–136.
- Stuart, M., Mullins, E., & Drew, E. (1995). Statistical quality control and improvement. *European Journal of Operational Research*, 88(2), 203–214.
- Taniguchi, S., Akhmetov, D., & Dote, Y. (1999). Fault detection of rotating machine parts using novel fuzzy neural network. In *Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on*, (Vol. 1, pp. 365–369). Tokyo.
- Tavner, P. (2008). Review of condition monitoring of rotating electrical machines. *IET Electric Power Applications*, 2(4), 215–247.
- Thomson, W., & Fenger, M. (2001). Current signature analysis to detect induction motor faults. *Industry Applications Magazine, IEEE*, 7(4), 26–34.
- Tran, V. T., Yang, B.-S., Oh, M.-S., & Tan, A. C. C. (2009). Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications*, 36(2, Part 1), 1840–1849.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2000a). A review of process fault detection and diagnosis part I: Quantitative model-based methods. *Computers and Chemical Engineering*, 27(3), 293–311.

- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2000b). A review of process fault detection and diagnosis part II: Qualitative models and search strategies. *Computers and Chemical Engineering*, 27(3), 313–326.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2000c). A review of process fault detection and diagnosis part III: Process history based methods. *Computers and Chemical Engineering*, 27(3), 327–346.
- Verucchi, C. J., & Acosta, G. G. (2008). Fault detection and diagnosis techniques in induction electrical machines. *IEEE Transactions Latin America*, 5(1), 41–49.
- Wand, M. P., & Jones, M. C. (1994a). *Kernel smoothing*. United Kingdom: Chapman and Hall CRC.
- Wand, M. P., & Jones, M. C. (1994b). Multivariate plug-in bandwidth selection. *Computational Statistics*, 9, 97–116.
- Yang, B.-S., & Kim, K. J. (2006). Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals. *Mechanical Systems and Signal Processing*, 20(2), 403–420.
- Yu, J. (2011a). Bearing performance degradation assessment using locality preserving projections. *Expert Systems with Applications*, 38(6), 7440–7450.
- Yu, J. (2011b). Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. *Mechanical Systems and Signal Processing*, 25(7), 2573–2588.
- Zhaoxia, W., Fen, L., Shujian, Y., & Bin, W. (2009). Motor fault diagnosis based on the vibration signal testing and analysis. In *Intelligent information technology application, 2009. IITA 2009. Third International Symposium on* (Vol. 2, pp. 433–436). Nanchang.