Available online at www.sciencedirect.com

ScienceDirect

Journal of the Franklin Institute 355 (2018) 1913–1930

www.elsevier.com/locate/jfranklin

Improved similarity-based modeling for the classification of rotating-machine failures

Matheus A. Marins, Felipe M.L. Ribeiro*, Sergio L. Netto, Eduardo A.B. da Silva

Program of Electrical Engineering, Universidade Federal do Rio Janeiro, Rio de Janeiro, RJ 21941-972, Brazil

Received 30 November 2016; received in revised form 22 May 2017; accepted 21 July 2017
Available online 29 July 2017

Abstract

Similarity-based modeling (SBM) is a technique whereby the normal operation of a system is modeled in order to detect faults by analyzing their similarity to the normal system states. First proposed around two decades ago, SBM has been successfully used for fault detection in varied systems. In spite of this success, there is not much study performed in the literature regarding its design, that encompasses both similarity metrics and model training. This work aims at contributing with an in-depth study of SBM for fault detection considering these two design aspects. This is done in the context of proposing a novel system to identify rotating-machinery faults based on SBM, that is employed either as a standalone classifier or to generate features for a random forest classifier. New approaches for training the model and new similarity metrics are investigated. Experimental results are shown for the recently developed Machinery Fault Database (MaFaulDa) that has an extensive set of sequences and fault types, and for the Case Western Reserve University (CWRU) bearing database. Results for both databases indicate that the proposed techniques increase the generalization power of the similarity model and of the associated classifier, achieving accuracies of 98.5% on MaFaulDa and 98.9% on CWRU database.

* Corresponding author.
E-mail addresses: matheus.marins@smt.ufrj.br (M.A. Marins), felipe.ribeiro@smt.ufrj.br (F.M.L. Ribeiro), sergioln@smt.ufrj.br (S.L. Netto), eduardo@smt.ufrj.br (E.A.B. da Silva).

http://dx.doi.org/10.1016/j.jfranklin.2017.07.038
0016-0032/

1. Introduction

Maintenance of critical equipment to ensure high levels of reliability, availability, and performance is one of the major concerns on today's industrial sector [1]. Unexpected failures can lead to substantial losses, either from the maintenance procedure itself or from the resulting production halts [2]. To achieve an effective and cost-efficient procedure, new maintenance strategies are being devised based on real-time and continuous monitoring, allowing one to detect and classify operational anomalies at an early stage, thus limiting additional system degradation [2]. Applications of such techniques include, for instance, flight paths [3], natural gas and nuclear power plants [4–7], wind turbines [8], and bearing or rotating-machine faults [9–14]. Among these equipments, rotating machines are some of the most important, being a key element used in a variety of applications, including airplanes, automobiles, power turbines, oil and gas refineries, and so on [12,15].

There are many approaches for detecting faults in rotating machines. Most of them consist of extracting features from the vibration signal to assess the machine current condition, in a supervised or automatic manner. Different features are needed to extract useful information relevant to detect faults from the original sources over multiple conditions. These features can be classified considering their domain (time, spatial, time-frequency, frequency) or its computation method (e.g. transform coefficients or aggregated statistics) [16–19].

An illustrative example is the approach in Yang et al. [20]. There, a system is presented which uses an adaptive resonance theory Kohonen neural network (ART-KNN) for fault diagnosis, having as inputs features derived from the discrete Wavelet transform coefficients. Unfortunately, the fault database used is not publicly available, making its comparison with other approaches impractical.

A methodology for detecting broken and half broken bars using spectral information over a FPGA is presented in [21]. This methodology is latter extended in [22] and in [23], adding the discrete wavelet transform (DWT) coefficients as features, and combining it with discrete frequency transform coefficients. These works also treat the detection of other faults and failures. The broken bar detection problem is also approached by the authors of [24] using motor current signature analysis and mathematical morphology.

The authors of [25] focus on the feature extraction procedure proposing a novel feature extraction scheme which utilizes the generalized S transform and 2D non-negative matrix factorization to detect possible faults. Three classifiers were used to assess the system: *k*-nearest neighbors (*k*NN), naive Bayes, and support vector machines (SVM), all achieving good results. A similar approach is presented in [26] using multiscale permutation entropy for feature extraction and an SVM classifier for fault diagnosis. The work of Rauber et. al. [18] also studies the effect of the features in the system performance. It tests multiple features of different types, such as complex envelope spectrum, statistical time- and frequency-domain parameters, as well as wavelet packet analysis, together with a feature selection algorithm. A fault classification database was used as testbed, and three different classifiers (*k*NN, feedforward artificial neural networks (ANN), and SVMs) were used during the assessment, achieving good performance.

This work proposes an automatic fault detector and classifier that uses similarity-based modeling (SBM) to identify rotating-machine failures such as imbalanced load, (horizontal or vertical) shaft misalignment, and bearing defects (in rolling elements or inner/outer tracks). The similarity model can be used either as an auxiliary model to generate features for the classifier (a random forest classifier in this case) or as a standalone classifier. In this context,

new approaches for training the similarity model and new similarity metrics are investigated. Two databases were employed to evaluate the performance of the proposed techniques. The first one is the machinery fault database (MaFaulDa) [27], a relatively new, large database of problematic scenarios of rotating-machine operations [13,14]. Performance evaluation on this database included continuous monitoring of six vibration sensors, one microphone, and one tachometer [14]. The second database is the Case Western Reserve University (CWRU) bearing database [28]. This database has become a standard reference in the bearing diagnostics field [19,29] and is used as testbed for comparison between the proposed methodology against other algorithms [18,25,26,30]. Results indicate that the proposed methodology is capable of correctly diagnosing the machine operating states, achieving an accuracy of 98.5% on the MaFaulDa dataset and 98.9% on the CWRU database.

This paper is organized as follows: Section 2 presents the original SBM technique [3,5,7,9–11], devised for detecting unusual patterns in some system or machine operation. Section 3 describes the proposed modifications to the standard SBM technique that allow the detection and classification of different types of anomalous machine operations in an efficient and robust manner. Section 4 details the MaFaulDa database, used to design and evaluate the system's performance and the CWRU database, used for comparison. The methodology of performance assessment is described in Section 5. This section also describes the designed system, including the preprocessing and validation procedures. Section 6 discusses the experimental results obtained during the processes of training and selection of the best model, as well the assessment results. Comparisons to other works are also included in this section. Finally, conclusions and discussions emphasizing the main contributions of this paper are provided in Section 7.

2. Similarity-based modeling (SBM)

The SBM is a simple and yet powerful nonparametric modeling technique that puts together an ensemble of previous state vectors in a single matrix $\mathbf{D}$ to represent the normal behavior of a given system, process, or machine. The SBM then evaluates the similarity of the current state vector with all vectors within $\mathbf{D}$ to assess the normality or not of the current system operation. This technique was first proposed in [4], and since then has been used in a variety of industrial applications, such as fault diagnosis in a machinery fault simulator [9,11], modeling airplanes flight paths [3], and anomaly detection in power plants [7]. In this work the SBM technique is adapted to monitor and classify the current operation of a given rotating machine.

In the original SBM framework, a system state at time n is represented by a vector $\mathbf{x}_n = [x_n(1), x_n(2), \ldots, x_n(M)]^{\mathrm{T}}$ comprising M measurements or features from multiple sources, such as sensors or signals. Given a set of L historical data $\mathbf{x}_n$, with $n = n_1, n_2, \ldots, n_L$, corresponding to the normal behavior of a given system, one can then represent this operational mode by the $L \times M$ "memory" matrix $\mathbf{D}$ which stacks the $\mathbf{x}_n$ in a line-by-line manner, as given by Singer et al. [4]

$$\mathbf{D} = \begin{bmatrix} \mathbf{x}_{\mathbf{n_1}} & \mathbf{x}_{n_2} & \cdots & \mathbf{x}_{n_L} \end{bmatrix}^{\mathrm{T}}. \tag{1}$$

Using this matrix, one may evaluate if a new input state $\mathbf{x}$ corresponds to the system normal operation by attempting to describe this state as a linear combination of the previously selected representative states contained in $\mathbf{D}$, that is

$$\hat{\mathbf{x}}_n = \mathbf{D}^{\mathrm{T}}\mathbf{w}_n. \tag{2}$$

By defining the estimation error and estimation error function respectively as

$$\mathbf{e}_n = \mathbf{x}_n - \hat{\mathbf{x}}_n, \tag{3}$$

$$e_n = \left\| \mathbf{x}_n - \hat{\mathbf{x}}_n \right\|_2^2, \tag{4}$$

the optimal linear estimate of $\mathbf{x}_n$ becomes

$$\hat{\mathbf{x}}_n = \mathbf{D}^{\mathrm{T}} \mathbf{w}_n^o = \mathbf{D}^{\mathrm{T}} \left(\mathbf{D}\mathbf{D}^{\mathrm{T}} \right)^{-1} \mathbf{D}\mathbf{x}_n, \tag{5}$$

where $\mathbf{w}_n^o$ is the optimal estimator.

However, in [4], the authors argue that this result has numerous limitations, such as an inability to accommodate random uncertainties and non-random defects, the need of a very large number L of archetypical states that define $\mathbf{D}$, and the requirement that $\mathbf{D}\mathbf{D}^{\mathrm{T}}$ must be nonsingular.

The SBM was proposed as an alternative approach that copes with all these issues by substituting the dot product implicit in Eq. (2) by a *similarity operation* $s(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i \circ \mathbf{x}_j$. This operation provides a similarity score $0 \leq s \leq 1$ between two vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, such that non-similar vectors yield $s \approx 0$ and very similar vectors correspond to $s \approx 1$. An example of similarity function considered in the original SBM formulation is given by Wegerich et al. [9]

$$s(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{1 + \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_2}. \tag{6}$$

The incorporation of distinct similarity operations into the SBM formulation is one of the contributions of the present work, as detailed in Section 3.2.

Following the linear approach and using the similarity operation, the current state vector $\mathbf{x}_n$ in the SBM algorithm can be estimated as

$$\hat{\mathbf{x}}_n = \mathbf{D}^{\mathrm{T}} \frac{\mathbf{w}_n}{\left\| \mathbf{w}_n \right\|_1}, \tag{7}$$

with

$$\mathbf{w}_n = \left(\mathbf{D} \circ \mathbf{D}^{\mathrm{T}} \right)^{-1} \left(\mathbf{D} \circ \mathbf{x}_n \right) = \mathbf{G}^{-1} \mathbf{a}_n. \tag{8}$$

that is, $\mathbf{G} = \mathbf{D} \circ \mathbf{D}^{\mathrm{T}}$ and $\mathbf{a}_n = \mathbf{D} \circ \mathbf{x}_n$. The vector $\mathbf{a}_n$ evaluates the similarity between the current state and the representative states in matrix $\mathbf{D}$ and matrix $\mathbf{G}$ transforms the similarity vector $\mathbf{a}_n$ in a set of weights for each state in $\mathbf{D}$. When $\mathbf{G} = \mathbf{I}$, the model is called auto-associative kernel regression (AAKR) [6], a particular case of SBM, equivalent to assuming no similarity between the state samples within $\mathbf{D}$. We mean by no similarity $s(\mathbf{x}_i, \mathbf{x}_j) = 1$ if $i = j$ and zero otherwise. In this case, if $\mathbf{x}_n$ belongs to $\mathbf{D}$, one then gets $\hat{\mathbf{x}}_n = \mathbf{x}_{n_l}$ such that the vector $\mathbf{w}_n$ becomes all null with the exception of an '1' entry at its n_lth position.

A key aspect within the SBM formulation is the strategy for composing matrix $\mathbf{D}$. Using all L historical samples for the normal behavior would incur in high computational expenses and very redundant data. Choosing an inadequate vector set when opting for a smaller L leads to performance impairments. The best possible set, therefore, would have the minimal number of vectors still yielding the same performance level as the complete set. In [5], a strategy is proposed for selecting a proper reduced set of historical samples. It comprises two selection steps:

1. One chooses as representatives the samples with index in the set $J = \{j_1, j_2, \ldots, j_K\}$, $K \leq 2M$, built such that

$$j \in J \quad \text{if} \quad \exists m : x_j(m) = \min_n \left\{ x_n(m) \right\} \vee x_j(m) = \max_n \left\{ x_n(m) \right\}; \tag{9}$$

2. The other samples are sorted by their ℓ_2 norm in decreasing order and decimated by a factor of t. The ones that remain after the decimation complement the set of representative samples.

The first step inserts in $\mathbf{D}$ all vector states which present the minimum and maximum value of each vector component. The second step decimates the remaining vectors using the ℓ_2 norm as ordering criterion. However, this second step may lead to not so good choices, because vectors with similar (even identical) norm values can be completely different [8]. Also, given a small decimation factor t and the number of samples L, the number of chosen samples is $\bar{L} = K + \lfloor (L-K)/t \rfloor$, which may be not much lower than L. Additional strategies for composing the matrix $\mathbf{D}$ are proposed in Section 3.3 in an attempt to overcome these issues.

3. Proposed SBM enhancements

This section presents the proposed enhancements to the SBM formulation, which include: a generalization of the SBM framework that allows it to operate in a multiclass (more than two classes) scenario; introduction of alternative similarity operations; and the development of a new strategy to compose the matrix $\mathbf{D}$.

3.1. Multiclass SBM

SBM was originally devised to detect abnormal operating conditions, which are associated with a low similarity level between a current state vector $\mathbf{x}_n$ and its SBM estimate $\hat{\mathbf{x}}_n$ given in Eq. (7).

Such framework can be extended, however, to detect and classify several types of system operational modes by defining a distinct model-matrix $\mathbf{D}_c$ for each operational class c. In the proposed multiclass SBM formulation, given a new input state $\mathbf{x}_n$, a different estimate can be determined for each class

$$\hat{\mathbf{x}}_{n,c} = \mathbf{D}_c^{\mathrm{T}} \frac{\mathbf{w}_{n,c}}{\left\| \mathbf{w}_{n,c} \right\|_1}, \tag{10}$$

where

$$\mathbf{w}_{n,c} = \left(\mathbf{D}_c \circ \mathbf{D}_c^{\mathrm{T}} \right)^{-1} \left(\mathbf{D}_c \circ \mathbf{x}_{n,c} \right) = \mathbf{G}_c^{-1} \mathbf{a}_{n,c}. \tag{11}$$

The current state $\mathbf{x}_n$ is then associated to the class c_n^* which maximizes the similarity,

$$c_n^* = \operatorname{argmax}_c \left\{ s_{n,c} \right\} = \operatorname{argmax}_c \left\{ s\left(\mathbf{x}_n, \hat{\mathbf{x}}_{n,c} \right) \right\}. \tag{12}$$

3.2. Alternative similarity functions

A *distance metric* is a function which maps a pair of elements in a set to a non-negative real number and satisfies a set of conditions. A classic example of distance metric is the

family of the p-distance metrics [31], defined as

$$d_{ij} = \left\|\mathbf{x}_i - \mathbf{x}_j\right\|_p = \left(\sum_{m=1}^{M} \left|x_{im} - x_{jm}\right|^p\right)^{\frac{1}{p}}. \tag{13}$$

The above equation is also know as ℓ_p norm. It is typically used with either $p = 2$, the ℓ_2 norm or Euclidean distance, or $p = 1$, the ℓ_1 norm or Manhattan distance.

In this work, five distinct similarity functions are employed, which can be separated in two main families: the multiquadric set and the exponential set. The multiquadric kernels include three of the selected functions which are based on the *inverse multiquadric function*, defined as [32]

$$f(d_{ij}) = \frac{1}{\left(r^2 + d_{ij}^2\right)^{\alpha}}. \tag{14}$$

The first similarity function is a direct application of the above equation. By making $\alpha = 1/2$, $r^2 = 1/\gamma^2$, and $s_{\mathrm{IMK}}(d_{ij}) = f(d_{ij})/\gamma$, one gets

$$s_{\mathrm{IMK}}(d_{ij}) = \frac{1}{\sqrt{1 + \gamma^2 d_{ij}^2}}, \tag{15}$$

which is the *inverse multiquadric kernel* (IMK) similarity function.

The second similarity function, the *Cauchy kernel* [33], is a direct variation of the IMK function, and can be defined as

$$s_{\mathrm{CCK}}(d_{ij}) = s_{\mathrm{IMK}}^2(d_{ij}) = \frac{1}{1 + \gamma^2 d_{ij}^2}. \tag{16}$$

The third and last multiquadric function is a modified version of the original SBM function, so-called *Wegerich similarity function*, shown in Eq. (6), and defined as

$$s_{\mathrm{WSF}}(d_{ij}) = \frac{1}{1 + \gamma d_{ij}}. \tag{17}$$

The last two functions are representatives of the exponential set. The *exponential* or *Laplacian kernel* is defined as

$$s_{\mathrm{EXP}}(d_{ij}) = \mathrm{e}^{-\gamma d_{ij}}, \tag{18}$$

and the *radial basis function kernel* as

$$s_{\mathrm{RBF}}(d_{ij}) = \mathrm{e}^{-\gamma d_{ij}^2}. \tag{19}$$

Each one of these functions were employed and evaluated during the validation procedure, to select the combination of parameter γ and similarity function that would provide the system with best performance.

3.3. Improved training method

In addition to the original strategy proposed in [5] to determine the model-matrix, described in the end of Section 2, an alternative approach is proposed here in an attempt to reach the best compromise between the associated computational complexity and the resulting system fault-classification ability.

In the proposed approach, one selects the state vectors $\mathbf{x}_{n,c}$ to form a model-matrix $\mathbf{D}_c$ iteratively. In each iteration a new vector is added to $\mathbf{D}_c$ taking into account the similarity between the currently selected state vectors and the remaining vectors available for class c. More specifically, given a vector set $\mathcal{X}_c$, the iterative procedure starts by selecting its geometric median

$$\mathbf{v}_c = \text{argmin}_{\mathbf{z}\in\mathcal{X}_c} \sum_{\mathbf{x}_i\in\mathcal{X}_c} \|\mathbf{x}_i - \mathbf{z}\|_2 \tag{20}$$

as the first representative state. Since there is no closed form to compute this median, it increases in complexity as the number of samples in $\mathcal{X}_c$ increase. In order to reduce the overall complexity, in this work we approximated the vector median using the algorithm described in [34]. The subsequent states that will compose $\mathbf{D}_c$ are selected according to the following strategy: each new sample $\mathbf{x}_{n,c}$ is compared against the current selected elements in $\mathbf{D}_c$. If the similarity between the $\mathbf{x}_{n,c}$ and any element of $\mathbf{D}_c$ is below a threshold τ, this sample is selected as an element of $\mathbf{D}_c$, otherwise the sample is discarded. More formally, $\mathbf{x}_{n,c}$ is included in $\mathbf{D}_c$ if

$$\mathbf{x}_{n,c} \circ \mathbf{x}_i < \tau, \quad \forall \mathbf{x}_i \in \mathbf{D}_c. \tag{21}$$

The effects in the SBM fault detection performance caused by the improvements proposed in this section are analyzed in the remaining of this paper.

4. Databases used

Two databases were used to evaluate the contributions of this paper. The first one, named machinery fault database (MaFaulDa) [14,27] is a comprehensive database including multiple types of faults covering different severities and rotation frequencies. This database was extensively used to validate the proposed approach and to search for the models with the best set of parameters based on their performance. The second database is the Case Western Reserve University bearing database [28], the standard reference in bearing faults [19,29]. It is used to assess the proposed approach against other ones found in the literature. A brief description of each database is presented below.

4.1. MaFaulDa database

This database is composed of multivariate time-series acquired by sensors on a SpectraQuest's machinery fault simulator (MFS) alignment-balance-vibration trainer (ABVT) [35]. This equipment emulates the dynamic of motors with two shaft-supporting bearings and allows the study of multiple faults, such as imbalanced mass, axis misalignment, and bearing problems. The experimental setup used in this work is illustrated on Fig. 1.

The system was monitored by two distinct sets (one for each bearing) of three accelerometers (on the axial, radial, and tangential directions), a tachometer (for measuring the system rotation frequency), and a microphone (for capturing the sound during the system operation). During the signal acquisition procedure, a variety of faults were imposed on the MFS. These faults are described below:

- *Normal operation*: this class represents the system operating under normal condition without any fault. It includes a set of 49 distinct scenarios, each with a fixed rotating speed within the range from 737 rpm to 3686 rpm with steps of approximately 60 rpm.

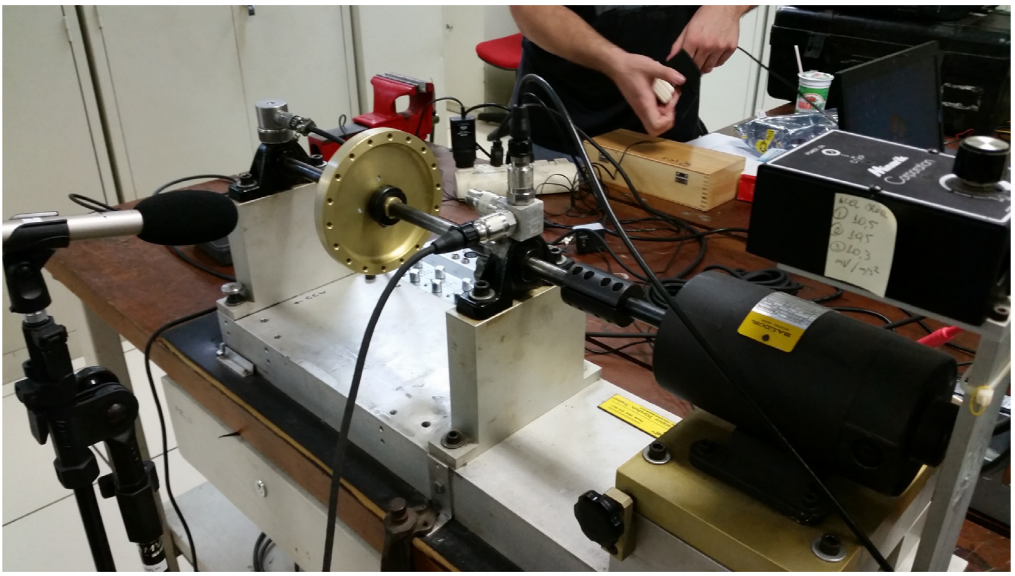

Fig. 1. Experimental setup used to produce the MaFaulDa database.

- *Imbalance*: to simulate different degrees of imbalanced operation, distinct load values of 6 g, 10 g, 15 g, 20 g, 25 g, 30 g, and 35 g were coupled to the rotor. For each load value below 30 g, the rotation frequency assumed in the same 49 values employed in the normal-operation case. For loads equal to or above 30 g, however, the resulting vibration makes impracticable for the system to achieve rotation frequencies above 3300 rpm, limiting the number of distinct rotation frequencies to only 44 in these cases. As such, the database includes a total of 333 different imbalance-operation scenarios.
- *Horizontal parallel misalignment*: this type of fault was induced into the MFS by shifting the motor shaft horizontally of 0.5 mm, 1.0 mm, 1.5 mm, and 2.0 mm. Using the same range for the rotation frequency as in the normal operation for each horizontal shift, a total of 197 different scenarios were considered for this class.
- *Vertical parallel misalignment*: this fault was induced into the MFS by shifting the motor shaft vertically of 0.51 mm, 0.63 mm, 1.27 mm, 1.4 mm, 1.78 mm, and 1.9 mm. Using the same range for the rotation frequency as in the normal operation for each vertical shift, a total of 301 different scenarios were considered for this fault class.
- *Bearing faults*: as one of the most complex elements of the machine, the rolling bearings are the most susceptible elements to fault occurrence. The ABVT manufacturer provided three defective bearings, each one with a distinct defective element (outer track, rolling elements, and inner track), that were placed one at a time in two different positions in the MFS experimental stand: between the rotor and the motor (underhang position), or in the external position, having the rotor between the bearing and the motor (overhang position). Bearing faults are practically imperceptible when there is no imbalance. So, the three masses of 6 g, 10 g, and 20 g were added to induce a detectable effect, with different rotation frequencies as before, leading to a total of 558 underhang scenarios and 513 overhang scenarios.

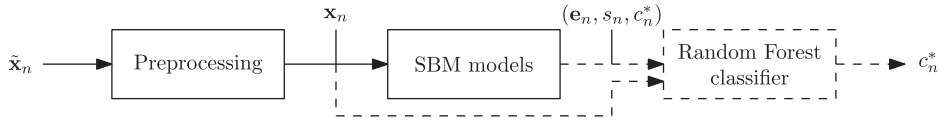

Fig. 2. Block diagram of the proposed system, composed by a preprocessing module, followed by the SBM and, possibly, by a classifier.

Considering all operating conditions described above, the MaFaulDa database comprises a total of 1951 different scenarios, each one described by 8 signals acquired at 50 kHz over a time interval of 5 s. The whole database is available for download at [27].

4.2. CWRU bearing database

The data from this database was acquired from the bearing center of the Case Western Reserve University (CWRU) [28]. It consists of 161 scenarios grouped in four categories, as described in [29]. Each scenario can be composed of acceleration signals in three directions: on the drive-end bearing, which occurs in all scenarios; on the fan-end bearing housing, which occurs in most of the scenarios; and on the motor supporting base plate, which occurs in some scenarios. The sample rates used were 12 kHz for some scenarios and 48 kHz for others. The vibration signals were obtained from different states of the bearings: normal condition, inner race fault, ball fault, and outer race fault. A more complete description of this database is found in [28].

The CWRU bearing database was selected for two main reasons. The first one is its public availability. The second one is its wide use in the literature for reporting results of automatic bearing fault detection methods, which allows comparison of the performance of proposed method against the one of other works. In this paper only scenarios containing both the fan-bearing and the drive-end signals were used, reducing the total number of valid scenarios to 153.

5. Experimental methodology

This section describes the experimental methodology employed to evaluate the modified SBM performance in detecting and classifying the ABVT's faulty scenarios within the databases described in Section 4.

The proposed system follows a modular architecture similar to the ones described in [2,36] for a condition-based maintenance system. It comprises three blocks (see Fig. 2): the *preprocessing* module converts the original data to a feature space which is more descriptive for the given application; the SBM model acts as a *state-monitoring* module, returning the similarity between the current input data and the previously modeled conditions; and the *classifier* or *diagnostic* module uses the information from previous blocks to identify the current input among the pre-specified set of classes. In such a framework the SBM can act in a stand alone manner or can be combined to a specific classifier (*random forest* [37], for instance, as employed here). In this paper, both strategies are considered.

The preprocessing block has the objective of reducing the original data to a set of more informative, relevant, and less redundant set of values. This is often important for reducing the system burden of learning and generalizing on the original data [38].

Given the distinct nature of each of the databases employed in this work, the preprocessing block should be different for each database, although its purpose is the same for both. The two preprocessing blocks are described as follows:

(i) *MaFaulDa*: three types of features were extracted from the original multivariate time-series: the *rotation frequency*, 21 additional *spectral features*, and 24 other *statistical features*.
The rotation frequency f_r was determined directly from the discrete Fourier transform (DFT) of the tachometer signal, as detailed in [13,15]. This procedure is very similar to the one presented in [21] for half-broken bars on induction motors.
The other spectral features correspond to the magnitudes of the spectrum of the signals other than the tachometer at frequencies f_r, $2f_r$, and $3f_r$.
The additional statistical features include, for each of the eight measured signals in each operational scenario, the statistical mean, the entropy, and the kurtosis. The variance feature is not employed as the signals are normalized to unit variance to reduce the effect of energy variations caused by changes in the acquisition setup.
(ii) *CWRU*: the statistical features presented in [18], together with the mean, variance and entropy were extracted from each signal, totaling 36 features.

The extracted features are then input to the subsequent stages in order to perform fault detection and classification. The two databases are treated independently for performance assessment of the proposed methods. The whole MaFaulDa database was randomly separated in two disjoint training and test sets, comprising respectively 90% and 10% of the given scenarios. The random choice of each set was constrained so that both presented the same fault proportion as the whole database. The best set of parameters was chosen using a k-fold cross-validation procedure on the training samples, with $k = 10$. Then, the performance of the best models are evaluated on the test set, producing the final results shown in Section 6.

As for the CWRU database, a process similar to cross dataset validation is applied. The best setups found for the MaFaulDa are directly used on the CWRU database. As such, this database is used to assess the generalization power of the classifiers obtained using the proposed methodology. Results have been obtained using k-folds with $k = 10$.

6. Experimental results and discussion

6.1. Experiment description

This subsection describes the experiments made during the validation procedure to select the best model for the proposed task considering all the following system variations:

- Feature types: only spectral features, only statistical features, or both families of features, as discussed in Section 5.
- Use of full SBM formulation (as given in Eq. (8)) or the AAKR scheme, which considers $\mathbf{G} = \mathbf{I}$ in this same equation.
- Choice of similarity function, as presented in Section 3.2, with distinct values of $\gamma \in \{0.01, 0.1, 0.5, 1, 10\}$ and different ℓ_p norms ($p \in \{1, 2\}$).
- Classification procedure either solely based on the stand-alone SBM or combining it to a specific classifier algorithm (e.g. random forest).

- SBM strategy for building model-matrix **D**: full matrix with all training feature vectors, original SBM training method [5] (see Section 2), named "decimation method", for decimation factors $s \in \{2, 3, 5, 7, 11\}$, or proposed threshold-based method (see Section 3.3) for threshold values $\tau \in \{0.05, 0.1, \ldots, 0.95\}$.

Clearly the full combination of the options described above leads to a prohibitively large number of possible system configurations. Therefore, in this work we significantly reduce this number by presenting the validation results following a sequential order of decisions, where each new decision seeks an improvement on the resulting performance. These decisions can be grouped in four main experiments:

- *Experiment 1* evaluates the influence of used feature types.
- *Experiment 2* compares the standard SBM model against the AAKR particularization.
- *Experiment 3* evaluates the different classification strategies, investigating whether one should use the stand-alone SBM to accomplish this function or the SBM output should be fed to a classification algorithm. In latter case, we also investigate the type of feature (similarity value or estimation error vector (Eq. 4)) that the SBM module should deliver to the subsequent classifier.
- *Experiment 4* selects the set of the remaining parameters (including procedure for building the model-matrix **D** and choice of similarity function) that produces the best SBM model overall.

6.2. Validation results

This subsection presents the validation results using cross-validation for each of the experiments. The experiments are presented in the aforementioned order, where the best configurations found in one experiment are carried out to the next one.

6.2.1. Experiment 1

This experiment assesses the influence of the feature types in the resulting validation performance of the classification system. To reduce the influence of other parameters during this evaluation, a very simple system was used, which employed the SBM as a classifier using the AAKR particularization, and the RBF similarity function with ℓ_2 norm. The kernel width γ assumed all values within the set $\{0.1, 0.5, 1.0\}$. All methods for building the SBM model-matrix **D** were assessed, with the decimation parameter fixed at $t = 5$, and the threshold parameter fixed at $\tau = 0.6$.

Table 1 presents the obtained cross-validation results for each parameter combination in Experiment 1. From this table, one can readily notice the superior performance achieved by the use of all combined (spectral and statistical) 46 features, which is carried on to all configurations considered in the subsequent experiments.

6.2.2. Experiment 2

This experiment compares the AAKR particularization with the standard SBM model by using the same set of parameters as of Experiment 1. Results from this experiment are summarized in Table 2.

From this table, one notices that the standard SBM outperformed its AAKR particularization in all configurations considered here. Given these results, the standard SBM approach was selected as the best performing option to be considered in the experiments that follow.

Table 1
Experiment 1 cross-validation accuracy(%), using the AAKR particularization as a classifier with an ℓ_2-norm RBF similarity function. f_r is the rotation frequency.

SBM model-matrix building	γ	f_r + spectral features	f_r + statistical features	All features
Full **D**	0.1	40.36 ± 3.28	68.27 ± 3.96	$\mathbf{71.61 \pm 3.72}$
	0.5	50.76 ± 3.18	76.55 ± 3.36	$\mathbf{81.00 \pm 3.07}$
	1.0	57.75 ± 3.18	77.0 ± 3.40	$\mathbf{81.91 \pm 2.77}$
Decimation [5]($t = 5$)	0.1	37.20 ± 2.10	63.64 ± 3.55	$\mathbf{67.53 \pm 5.04}$
	0.5	46.60 ± 2.48	66.19 ± 4.56	$\mathbf{72.45 \pm 4.69}$
	1.0	49.01 ± 3.21	63.97 ± 3.88	$\mathbf{69.50 \pm 4.15}$
Threshold($\tau = 0.6$)	0.1	41.02 ± 2.68	61.38 ± 2.78	$\mathbf{67.08 \pm 3.02}$
	0.5	50.36 ± 2.55	76.01 ± 3.28	$\mathbf{80.81 \pm 2.47}$
	1.0	56.70 ± 2.24	77.08 ± 3.69	$\mathbf{81.55 \pm 3.08}$

Table 2
Experiment 2 cross-validation accuracy(%), comparing the SBM and AAKR.

SBM model-matrix building	γ	SBM	AAKR
Full **D**	0.1	$\mathbf{84.80 \pm 2.81}$	71.61 ± 3.72
	0.5	$\mathbf{83.66 \pm 2.45}$	81.00 ± 3.07
	1.0	$\mathbf{82.50 \pm 2.39}$	81.91 ± 2.77
Decimation [5]($t = 5$)	0.1	$\mathbf{78.75 \pm 3.39}$	67.53 ± 5.04
	0.5	$\mathbf{72.80 \pm 3.26}$	72.45 ± 4.69
	1.0	$\mathbf{70.29 \pm 3.46}$	69.50 ± 4.15
Threshold ($\tau = 0.6$)	0.1	$\mathbf{74.00 \pm 2.28}$	67.08 ± 3.02
	0.5	$\mathbf{82.77 \pm 2.20}$	80.81 ± 2.47
	1.0	$\mathbf{83.01 \pm 2.40}$	81.55 ± 3.08

6.2.3. Experiment 3

This experiment evaluates the SBM method either as a stand-alone classifier or as an auxiliary input to an off-the-shelf random forest (RF) classifier (see Fig. 2). To this end, we have evaluated four different system configurations: (i) stand-alone SBM classifier; (ii) stand-alone RF classifier; (iii) combined SBM-RF classifier using the SBM similarities to each class as a complementary feature; (iv) combined SBM-RF classifier using the SBM estimation error vector Eq. (3) as a complementary feature. In this experiment, the SBM model used the best configurations found in the previous experiments, which include all 46 features previously considered and its standard SBM formulation.

Table 3 presents the cross-validation results for all tested SBM-based configurations. As a basis for comparison, note that in our simulations the stand-alone RF configuration achieved an accuracy score of 92.70%.

These results show that the combined SBM-RF schemes are more discriminative than the stand-alone SBM or RF models. We can see that the case where the original features are extended by the similarities to each class estimated by the SBM produces consistently good results. However, the best results were obtained by extending the original features with the SBM estimation error vector (Eq. (3)) instead of similarities.

Table 3 also indicates that, when one uses the combined SBM-RF configuration with the additional SBM estimation error features, the full model-matrix **D** is greatly outperformed by the ones built using the other two methods. Therefore, in Experiment 4, we consider only the decimation [5] and threshold methods for building **D**.

Table 3
Experiment 3 cross-validation accuracy(%), comparing different SBM-based classifier configurations. The accuracy score of a stand-alone RF configuration is 92.70%.

		Classifier		
SBM model- matrix building	γ	SBM	RF + SBM estimation error	RF + SBM similarities
Full **D**	0.1	84.80 ± 2.81	44.19 ± 4.38	**96.32 ± 1.66**
	0.5	83.66 ± 2.45	39.66 ± 3.31	**93.88 ± 1.88**
	1.0	82.50 ± 2.39	41.97 ± 3.50	**93.24 ± 1.39**
Decimation [5] ($t = 5$)	0.1	78.75 ± 3.39	**98.20 ± 1.11**	96.43 ± 0.89
	0.5	72.80 ± 3.26	**97.66 ± 0.91**	94.98 ± 1.25
	1.0	70.29 ± 3.46	**97.44 ± 0.96**	94.77 ± 1.21
Threshold ($\tau = 0.6$)	0.1	74.00 ± 2.28	94.66 ± 1.19	**96.59 ± 1.70**
	0.5	82.77 ± 2.20	93.87 ± 0.88	**94.64 ± 1.87**
	1.0	83.01 ± 2.40	85.92 ± 2.56	**94.02 ± 1.28**

Table 4
Experiment 4 cross-validation accuracy(%) for the best 10 SBM configurations. The parameter of the last column is t for the decimation method for building the model-matrix and τ for the proposed threshold method.

			Parameters		
SBM model-matrix building	Similarity function	Accuracy(%)	Norm	γ	t or τ
Decimation [5]	WSF	**98.91 ± 0.58**	1	0.01	7
	RBF	98.02 ± 0.82	2	0.1	11
	IMK	98.62 ± 0.90	2	0.1	5
	CCK	98.56 ± 0.79	1	1	7
	EXP	98.57 ± 0.93	1	0.1	11
Threshold	WSF	**98.68 ± 0.89**	1	0.01	0.9
	RBF	96.66 ± 1.17	2	0.01	0.55
	IMK	98.56 ± 1.37	2	0.1	0.8
	CCK	**98.72 ± 0.78**	2	0.1	0.5
	EXP	98.33 ± 1.14	1	0.1	0.5

6.2.4. Experiment 4

This last experiment performs a fine tuning of the SBM method by selecting the best possible procedure for building the model-matrix **D** together with the best similarity function, including all their parameters.

Table 4 presents the 10 best configurations for this experiment, where the 'Parameters' column shows the chosen parameters for each of these scenarios. The similarity functions used are WSF (Eq. (17)), RBF (Eq. (19)), IMK (Eq. (15)), CCK (Eq. (16)), and EXP (Eq. (18)).

All the 10 models discriminated in Table 4 present superior performance than the one found in previous works using the same database [13,14], attesting the SBM capability to successfully solve the fault-classification problem in rotating machines.

The analysis of the validation results in Table 4 leads us to choose three models as the best ones. They are:

1. Model A: similarity function WSF, $\gamma = 0.01$, ℓ_1 norm. Using the decimation method for building **D** [5], with $t = 7$.

Table 5
Average number of representative states for each combination of model-matrix building model and class (normal **N**, imbalanced **I**, horizontal misalignment $\mathbf{H}_M$, vertical misalignment $\mathbf{V}_M$, underhang faulty bearing $\mathbf{U}_B$, overhang faulty bearing $\mathbf{O}_B$).

Configuration	N	I	H_M	V_M	U_B	O_B
Model A	33.1	87.6	65.7	80.1	131.2	118.2
Model B	5.8	73.8	8	8	94.8	29.8
Model C	5	73.5	4.9	5	74.5	6

2. Model B: similarity function WSF, $\gamma = 0.01$, ℓ_1 norm. Using the proposed threshold method for building **D**, with $\tau = 0.9$.
3. Model C: similarity function CCK, $\gamma = 0.1$, ℓ_2 norm. Using the proposed threshold method for building **D**, with $\tau = 0.5$.

The complexity of a given SBM model is given by its number of representative states. Therefore, in order to analyze the model complexities when using the three above models, we present in Table 5 the average number of representative states, over the 10 validation folds, for each combination of selected model and class (normal **N**, imbalanced **I**, horizontal misalignment $\mathbf{H}_M$, vertical misalignment $\mathbf{V}_M$, underhang faulty bearing $\mathbf{U}_B$, overhang faulty bearing $\mathbf{O}_B$).

From this table, one can readily draw two conclusions regarding the average number of representatives in each case: first, considering each model-matrix building scheme, the table shows, as expected, that the proposed threshold method is more selective than the decimation method. This is an important result, as the proposed method requires less storage space and processing time, making it well suited for deployment in a real-time. As an example, Model B requires about 162 μs to process a sample, while Model A requires 350 μs[1]. However, in the context of this work, this difference is irrelevant, as the preprocessing phase consumes 156 ms to process a single sample. One should notice that the system is capable to work on real-time, since each sample is 5 s long. Second, analyzing the size of the model-matrices for each failure, we observe that the imbalance failure and the underhang bearing fault require larger number of states, and are thus difficult to discriminate.

6.3. Results on the testing sets

In this subsection the performance of the proposed system on the testing set is analyzed. For this study, the three models chosen in Section 6.2.4 will be considered: Model A, Model B and Model C. It is important to notice that the original scheme leads to a simpler model-matrix building stage but to a larger matrix which results in a more complex classification procedure, as observed in Table 5.

As mentioned on Section 5, the test dataset is composed by 10% of the available MaFaulDa scenarios. Using the test dataset, Model A and Model B achieved an accuracy of 98.49% and Model C achieved and accuracy of 97.47%, indicating that all three models are capable of generalizing well for other samples. The confusion matrices for the first two models are shown in Table 6a and 6b.

[1] All times were measured on an Intel(R) Core(TM) i7-4790K CPU @ 4.00 GHz machine.

Table 6
Confusion matrices in test dataset.

a) Model A

Class	N	I	H_M	V_M	U_B	O_B
N	4	0	1	0	0	0
I	0	34	0	0	0	0
$\mathbf{H_M}$	0	0	20	0	0	0
$\mathbf{V_M}$	0	0	0	31	0	0
$\mathbf{U_B}$	0	1	0	0	55	0
$\mathbf{O_B}$	0	1	0	0	0	51

b) Model B

Class	N	I	H_M	V_M	U_B	O_B
N	5	0	0	0	0	0
I	0	34	0	0	0	0
$\mathbf{H_M}$	0	0	19	1	0	0
$\mathbf{V_M}$	0	0	0	31	0	0
$\mathbf{U_B}$	0	1	0	0	54	1
$\mathbf{O_B}$	0	0	0	0	0	52

Table 7
Accuracy results for the 10-class identification problem on the MaFaulDa database.

Model	Model-matrix building method	Similarity function	γ	t or τ	p	Acc. (%)
Model A	Decimation	WSF	0.01	7	2	98.48
Model B	Threshold	WSF	0.01	0.9	1	98.48
Model C	Threshold	CCK	0.1	0.5	2	97.48

Results in these tables are consistent with some already discussed aspects of the MaFaulDa database. As stated in Section 4, there are much less scenarios when the machinery operates on normal conditions then there are faulty cases. This discrepancy makes the model-matrix building more difficult for the normal class (N). Still analyzing the confusion matrices, it is possible to observe that sometimes bearing faults are misclassified as imbalance faults. We argue that this is somewhat expected, since the system needs to be unbalanced in order for bearing faults to be observed.

Also in Section 4, when the bearing faults are described, it was mentioned that each one of the bearings (underhang and overhang) where substituted by one out of three defective bearings provided by the manufacturer. Taking this into consideration, the three best configurations, Model A, Model B, and Model C, were also used to classify the signals in 10 classes. These classes were derived by further subdividing each bearing fault in 3 classes according to the defective element (outer race, inner race, or rolling ball) employed. The results are presented in Table 7, where the good accuracy figures indicate that the proposed system is also robust when applied to more complex fault classification problems.

6.4. CWRU results and discussion

This subsection presents the results on the CWRU bearing dataset. As described in Section 5, this dataset is used for assessing the performance of the three best models selected on the MaFaulDa dataset, namely the Model A, Model B, and Model C schemes (see Section 6.2.4).

Using the same methodology as [18], each CWRU signal was divided into 15 segments, and the extended dataset was subdivided into the training and testing sets following a 9/1 ratio. The results presented in Table 8 are accuracy averages over 10 folds chosen randomly. For each fold configuration, the model-matrix **D** is computed using the data in 9 folds and the accuracy result is measured in the remaining fold. From this table, one can observe that the

Table 8
Accuracy (%) results of SBM-based classifiers on the CWRU database.

Model	Mean
Model A	98.95 ± 0.72
Model B	98.91 ± 0.75
Model C	98.91 ± 0.95

SBM-based classifier has good generalization capability for all three configurations considered here.

6.5. Comparison with previous works

As mentioned in Section 1, several other works in the literature addressed the same problem that we have addressed in this paper, that is, the automatic detection and classification of faults in rotating machines. Some of these works have used the MaFaulDa database. In [14] the faults in the MaFaulDa database have been classified using perceptron neural networks with multiple layers, considering several subsets of the features investigated here. Six classes have been considered: normal, overhang and underhang faults, imbalance, horizontal and vertical misalignment. The accuracy obtained was 95.8%, inferior to the ones obtained with the proposed use of SBM and described in Table 7, that reach, for one configuration, the average figure of 98.48%.

For the CWRU database, even though there are many works using such dataset [29], its very difficult to make a direct comparison, as most works do not present their results in a quantitative manner, but only in a qualitative manner. As such, the comparison is restricted to a small set of works. In [25] the *k*NN, naive Bayes, and SVM classifiers achieved accuracies of 98.83%, 98% and 98.97%, respectively. The SVM classifier found in [26] obtained accuracies above 98% for different rotation frequencies. The SVM and ELM classifiers using the procedure described in [30] achieved accuracies of 82.4% and 97.5%, respectively. Lastly, the *k*NN, SVM, and ANN classifiers using the feature selection method proposed in [18] obtained accuracies between 93% and 100%. From the above results and Table 8, one can conclude that the proposed SBM-based fault classifier achieves, for the CWRU database, competitive results to those. It is important to point out that, as demonstrated by the results over the MaFaulDa database, the proposed system is able to detect and classify, with high accuracy, a wide range of machine faults, including misalignment and unbalanced faults.

7. Conclusion

This paper addressed the automatic fault diagnosis in rotating machines. The use of similarity based modeling (SBM) was investigated, either as a stand-alone classification method or in combination with an off-the-shelf classifier, in this case a random forest classifier. The system is evaluated in two databases. One of them is a comprehensive database with multiple faults referred to as MaFaulDa [27]. The other is the CWRU bearing database [28], that is the current standard database for bearing fault diagnosis.

The extension of the SBM to a multiclass model is an important contribution of this paper. Other contributions include the use of new similarity metrics and the development of a novel method for building the SBM model-matrices. This work also presented an extensive

study of the evaluation of all the proposed modifications on the MaFaulDa database. These contributions achieved the goal of increasing the SBM performance in a fault classification scenario while reducing its computational complexity. The usage of SBM either as a stand-alone classifier or as a feature generator for off-the-shelf classifiers has also been investigated. Our results have shown that the use of the proposed enhancements to the SBM consistently increased the accuracy of a random forest classifier.

The proposed system showed to be robust, reaching an accuracy of around 98.5% in the MaFaulDa database, higher than previous works along the same base. For the CWRU dataset the proposed system yielded an accuracy level of 98.9%, which is as good as previous results reported in the literature. It is worth emphasizing that the proposed class of methods is able to detect and classify, with high accuracy, a wide range of faults, which indicates that the proposed approach based on SBM is worth further investigation.

Acknowledgments

This research was supported by CNPq and Petrobras. The authors would like to thank Dr. Kenneth A. Loparo and the Case Western Reserve University Bearing Data Center for providing the CWRU dataset for this study, and Dr. Wade A. Smith for his help.

References

[1] J.P. Herzog, D. Gandhi, B. Nieman, Making decisions about the best technology to implement, Technical report, Smartsignal/GE Intelligent Platforms, 2011.
[2] J.B. Coble, Merging data sources to predict remaining useful life – an automated method to identify prognostic parameters, University of Tennessee, 2010 (Ph.D. thesis).
[3] J. Mott, M. Pipke, Similarity-based modeling of aircraft flight paths, in: Proceedings of the IEEE Aerospace Conference, vol. 3, 2004.
[4] R.M. Singer, K.C. Gross, J.P. Herzog, R.W. King, S. Wegerich, Model-based nuclear power plant monitoring and fault detection: theoretical foundations, in: Proceedings of the International Conference on Intelligent Systems Applications to Power Systems, 1997.
[5] J.P. Herzog, S.W. Wegerich, K.C. Gross, F.K. Bockhorst, MSET modeling of Crystal River-3 venturi flow meters, in: Proceedings of the International Conference on Nuclear Engineering, 1998.
[6] J. Garvey, D. Garvey, R. Seibert, J.W. Hines, Validation of on-line monitoring techniques to nuclear plant data, Nucl. Eng. Technol. 39 (2) (2006) 133–142.
[7] F.A. Tobar, L. Yacher, R. Paredes, M.E. Orchard, Anomaly detection in power generation plants using similarity-based modeling and multivariate analysis, in: Proceedings of the American Control Conference, vol. 3, 2011.
[8] P. Guo, N. Bai, Wind turbine gearbox condition monitoring with AAKR and moving window statistic methods, Energies 4 (11) (2011) 2077–2093.
[9] S.W. Wegerich, A.D. Wilks, R.M. Pipke, Nonparametric modeling of vibration signal features for equipment health monitoring, in: Proceedings of the IEEE Aerospace Conference, vol. 7, 2003.
[10] S.W. Wegerich, Similarity-based modeling of time synchronous averaged vibration signals for machinery health monitoring, in: Proceedings of the IEEE Aerospace Conference, vol. 6, 2004.
[11] S.W. Wegerich, Similarity based modeling of vibration features for fault detection and identification, Sens. Rev. 25 (2) (2005) 114–122.
[12] J. Liu, W. Wang, F. Golnaraghi, An enhanced diagnostic scheme for bearing condition monitoring, IEEE Trans. Instrum. Meas. 59 (2) (2010) 309–321.
[13] A.A. de Lima, T.d.M. Prego, S.L. Netto, E.A.B. da Silva, R.H.R. Gutierrez, U.A. Monteiro, A.C.R. Troyman, F.J.d.C. Silveira, L. Vaz, On fault classification in rotating machines using Fourier domain features and neural networks, in: Proceedings of the Latin American Symposium on Circuits and Systems, 2013.
[14] D. Pestana-Viana, R. Zambrano-López, A.A. de Lima, T.d. M. Prego, S.L. Netto, E.A.B. da Silva, The influence of feature vector on the classification of mechanical faults using neural networks, in: Proceedings of the Latin American Symposium on Circuits and Systems, 2016.

[15] P. Li, F. Kong, Q. He, Y. Liu, Multiscale slope feature extraction for rotating machinery fault diagnosis using wavelet analysis, Measurement 46 (19) (2013) 497–505.
[16] R.B. Randall, J. Antoni, Rolling element bearing diagnostics – a tutorial, Mech. Syst. Signal Process. 25 (2) (2011) 485–520.
[17] J.H. Zhou, L. Wee, Z.W. Zhong, A knowledge base system for rotary equipment fault detection and diagnosis, in: Proceedings of the Eleventh International Conference on Control Automation Robotics Vision, 2010.
[18] T.W. Rauber, F. de Assis Boldt, F.M. Varejão, Heterogeneous feature models and feature selection applied to bearing fault diagnosis, Mech. Syst. Signal Process. 62 (1) (2015) 637–646.
[19] A. Boudiaf, A. Moussaoui, A. Dahane, I. Atoui, A comparative study of various methods of bearing faults diagnosis using the Case Western Reserve University Data, J. Fail. Anal. Prev. 16 (2) (2016) 271–284.
[20] B. Yang, T. Han, J. An, ART-KOHONEN neural network for fault diagnosis of rotating machinery, Mech. Syst. Signal Process. 18 (3) (2004) 645–657.
[21] J. de Jesus Rangel-Magdaleno, R. de Jesus Romero-Troncoso, R.A. Osornio-Rios, E. Cabal-Yepez, L.M. Contreras-Medina, Novel methodology for online half-broken-bar detection on induction motors, IEEE Trans. Instrum. Meas. 58 (5) (2009) 1690–1698.
[22] L.M. Contreras-Medina, R. de Jesus Romero-Troncoso, E. Cabal-Yepez, J. de Jesus Rangel-Magdaleno, J.R. Millan-Almaraz, FPGA-based multiple-channel vibration analyzer for industrial applications in induction motor failure detection, IEEE Trans. Instrum. Meas. 59 (1) (2010) 63–72.
[23] J. de Jesus Rangel-Magdaleno, R. de Jesus Romero-Troncoso, R.A. Osornio-Rios, E. Cabal-Yepez, A. Dominguez-Gonzalez, FPGA-based vibration analyzer for continuous CNC machinery monitoring with fused FFT-DWT signal processing, IEEE Trans. Instrum. Meas. 59 (12) (2010) 3184–3194.
[24] J. de Jesus Rangel-Magdaleno, H. Peregrina-Barreto, J.M. Ramirez-Cortes, P. Gomez-Gil, R. Morales-Caporal, FPGA-based broken bars detection on induction motors under different load using motor current signature analysis and mathematical morphology, IEEE Trans. Instrum. Meas. 63 (5) (2014) 1032–1040.
[25] B. Li, P. Zhang, D. Liu, S. Mi, G. Ren, H. Tian, Feature extraction for rolling element bearing fault diagnosis utilizing generalized S transform and two-dimensional non-negative matrix factorization, J. Sound Vib. 330 (10) (2011) 2388–2399.
[26] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, C.-C. Wang, Bearing fault diagnosis based on multiscale permutation entropy and support vector machine, Entropy 14 (8) (2012) 1343–1356.
[27] MaFaulDa – Machinery Fault Database, 2016, http://www02.smt.ufrj.br/~offshore/mfs/ (Accessed 22. 11. 16).
[28] K.A. Loparo, Bearings vibration data set, Case Western Reserve University, 2003, (http://csegroups.case.edu/bearingdatacenter/home) (Accessed 25. 11. 16).
[29] W.A. Smith, R.B. Randall, Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study, Mech. Syst. Signal Process. 64–65 (2015) 100–131.
[30] Y. Li, X. Wang, J. Wu, Fault diagnosis of rolling bearing based on permutation entropy and extreme learning machine, in: Proceedings of the Chinese Control and Decision Conference, 2016.
[31] G.H. Golub, C.F.V. Loan, Matrix Computations, (third, Johns Hopkins University Press, Baltimore, MD, USA, 1996.
[32] C.A. Micchelli, Interpolation of scattered data: distance matrices and conditionally positive definite functions, Constr. Approx. 2 (1) (1986) 11–22.
[33] J. Basak, A least square kernel machine with box constraints, in: Proceedings of the International Conference on Pattern Recognition, 2008.
[34] H. Cardot, P. Cénac, P.-A. Zitt, Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm, Bernoulli 19 (1) (2013) 18–43.
[35] SpectraQuest, Inc., 2016, (http://www.http://spectraquest.com/) (Accessed 27. 11. 16).
[36] G. Palem, Condition-based maintenance using sensor arrays and telematics, Int. J. Mobile Netw. Commun. Telemat. 3 (3) (2013) 19–28.
[37] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[38] Y.S. Abu-Mostafa, M. Magdon-Ismail, H.-T. Lin, Learning from Data, AMLBook, US, 2012.