Vibration Signal Forecasting on Rotating Machinery by means of Signal Decomposition and Neurofuzzy Modeling

Daniel Zurita-Millán, 1 Miguel Delgado-Prieto, 1 Juan José Saucedo-Dorantes, 2 Jesus Adolfo Carino-Corrales, 1 Roque A. Osornio-Rios, 2 Juan Antonio Ortega-Redondo, 1 and Rene de J. Romero-Troncoso 3

1 MCIA Research Center, Department of Electronic Engineering, Technical University of Catalonia (UPC), Rambla San Nebridi No. 22, Gaia Research Building, 08222 Terrassa, Spain
2 CA Mecatronica, Facultad de Ingenieria, Universidad Autonoma de Queretaro, Campus San Juan del Rio, Rio Moctezuma 249, Colonia San Cayetano, 76807 San Juan del Rio, QRO, Mexico
3 CA Telematica, DICIS, Universidad de Guanajuato, Carretera Salamanca-Valle km 3.5+1.8, Palo Blanco, 36885 Salamanca, GTO, Mexico

Correspondence should be addressed to Daniel Zurita-Millán; daniel.zurita@mcia.upc.edu

Received 6 May 2016; Revised 18 July 2016; Accepted 18 July 2016

Academic Editor: Mario Terzo

Copyright © 2016 Daniel Zurita-Millán et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vibration monitoring plays a key role in the industrial machinery reliability since it allows enhancing the performance of the machinery under supervision through the detection of failure modes. Thus, vibration monitoring schemes that give information regarding future condition, that is, prognosis approaches, are of growing interest for the scientific and industrial communities. This work proposes a vibration signal prognosis methodology, applied to a rotating electromechanical system and its associated kinematic chain. The method combines the adaptability of neurofuzzy modeling with a signal decomposition strategy to model the patterns of the vibrations signal under different fault scenarios. The model tuning is performed by means of Genetic Algorithms along with a correlation based interval selection procedure. The performance and effectiveness of the proposed method are validated experimentally with an electromechanical test bench containing a kinematic chain. The results of the study indicate the suitability of the method for vibration forecasting in complex electromechanical systems and their associated kinematic chains.

1. Introduction

Rotating machinery, like compressors, steam turbines, industrial fans, and so forth, is widely used in many industrial fields. Its reliability is an extensively investigated field, aimed at prolonging their life span and minimizing their maintenance cost. Maintenance programs try to avoid fatal breakdowns of machines and prevent production loss and human casualties. The early fault diagnosis is a challenging problem and has received more and more attention in recent years [1]. In order to reduce maintenance costs and maintain machine uptime at the highest possible level, maintenance should be carried out in a proactive way. That means a transformation of maintenance strategy from the traditional fail-and-fix practices (diagnostics) to a predict-and-prevent methodology (prognostics) [2]. The current and future health status of a machine, component, or system have to be known in order to predict and prevent an eventual occurrence of a failure, enabling the achievement of near-zero downtime performance.

Vibration measurement is an effective, nonintrusive method to monitor machine condition during start-ups, shutdowns, and normal operation [3]. Vibration analysis is used on rotating equipment to determine the operating and operation condition of the equipment. A major advantage is that vibration analysis can identify developing problems before they become too serious and cause unscheduled downtime. A majority of the mechanical faults such as defective bearings, mechanical looseness, worn or broken gears, misalignment, and unbalance could be detected by
vibration analysis. Vibration signals collected from this equipment during operation contain valuable information about the equipment condition. Great research effort has been carried out for the development of several algorithms for faults detection and diagnosis in rotating machinery based on vibration monitoring [4]. The most popular methods of rotating machine condition monitoring utilize the steady state spectral components of the vibration signals. These spectral components of the vibration are used to diagnose faults of rotating machinery and the energy of vibration signals in the frequency domain, being the most commonly used method to evaluate the severity of a fault.

Nevertheless, the vibration signal is often a complex signal which contains stationary, nonstationary, and noisy components. Therefore, the appropriate signal processing techniques have to be employed in order to compute numerical fault indicators or features [5]. Features are parameters derived from the measured data that robustly indicate the presence of faults in rotating machinery. The major feature-calculation methods are time domain methods, frequency domain methods, and time-frequency methods. Time domain methods such as peak amplitude, Root-Mean-Square amplitude, crest factor, kurtosis, and shock pulse counting have been successfully applied [6, 7]. Frequency domain methods include Fourier spectra time waveform, cepstrum analysis, and the spectral envelope technique [8, 9]. A comparative study of various feature-calculation methods in frequency domain, such as the Fast Fourier Transform (FFT) and spectral envelope, and in the time-frequency domain, such as the Short Time Fourier Transform (STFT), Wigner Ville, and wavelet analysis, is presented in [10, 11].

Clearly, the opportunity of conducting forecasting based on vibrations presents multiple benefits: (i) correct deviations before they affect the correct behavior of the system, (ii) anticipate the response towards failures, and (iii) assist the diagnosis algorithm in order to assess the future condition of the system [12, 13]. Unfortunately, little work has been done in the area with vibration forecasting and there is a lack of methodology to perform such approaches [14]. The modeling of the vibration as a physical magnitude is a complex process; furthermore, vibration in electromechanical systems is affected by several factors such as multiple emission sources, mechanical properties, operation condition. However, the modeling of numerical features that characterize the vibration is a more affordable problem, since raw data presents redundant or nonimportant information from the electromechanical components degradation point of view. This is the case when time domain features such as the Root-Mean-Square (RMS) value or kurtosis offer a more reliable view of the machine condition [15]. As a result, the temporal modeling of features extracted from the vibration signal is a matter of interest for many researchers. Therefore, the utility of a vibration's feature forecasting model relies on the early detection of deviations from the correct behavior as early as the model output is extended in the future. The optimal scenario of such approaches is the modeling of the RMS value of the vibrations during the start-up operation of the machine till it reaches the thermal stability in order to indicate the abnormal behaviors as soon as possible to minimize the damage of operating under failure conditions.

Consequently, the forecasting of such vibration feature signals can be approached from the time series feature point of view. Current modeling and forecasting techniques can be categorized into three classes, namely, model-based, data-driven, and hybrid prognostics approaches [16]. Model-based techniques describe the future evolution of the system based on the mathematical equations that describe its physical properties [17]. Data-driven techniques take advantage of the past records to learn the system behavior and conduct the predictions. Hybrid techniques are a combination of model-based and data-driven techniques [18]. In this regard, Adaptive Neurofuzzy Inference System (ANFIS) represents a hybrid data-driven forecasting method that fuses the parametric adaptability of ANN and the generalization capabilities of the fuzzy logic. ANFIS offers a reliable and robust condition predictor, since it can capture the system dynamic behavior quickly and accurately [19].

When facing vibration forecasting, the problem is how to decide for a given application the optimal forecasting horizon and the configuration of the models. In this regard, vibration signals usually present faster or at least equal dynamics in comparison with other magnitudes inherent to an electromechanical system, such as voltage, electric current, and temperature. Due to this fact, the modeling problem should be approached as a time series forecasting [20]. It is necessary to emphasize that the result of a forecasting method is critically associated with the forecasting horizon used to train and validate the models. In many applications, the forecasting horizon is usually fixed by application requirements or modeling limitations [21]. Yet, the procedure to select the best horizon for a given application is not well established in the literature. For this reason, this paper proposes performing a study relating the selected forecasting horizon and the prediction error once the application is defined. In this regard, three open issues can be found in literature: (i) the decision of the optimal prediction horizon, (ii) the configuration of the models in terms of number of inputs and past values to be considered, and (iii) the decomposition of the signal in different details in order to increase accuracy.

In this paper, the authors propose a study in order to optimize these three open issues to obtain a methodology for modeling the RMS of the vibrations from an electromechanical system and its associated kinematic chain. Therefore, the contributions of this work include a novel vibration forecasting method that utilizes the theory behind vibration signature under electromechanical failures to decompose the vibration signal in three frequency bands related to the vibration signature to enhance the forecasting accuracy and adaptability towards different dynamic contents. Then, the core of the methodology consists of the generation of three collaborative ANFIS models that are in charge of forecasting the evolution of specific spectral content of the vibration signal. The configuration of the model is optimized one step beyond the literature by analyzing cross-correlations between the signal and past instants, which are selected by using Genetic Algorithm (GA) optimization, to find the best
input configuration. Complementarily, the paper includes a systematic study of the forecasting horizon affectation of the model configuration step, in order to establish a valid method for selecting the best forecasting horizon with regard to a proposed application. Finally, the proposed method is validated experimentally by means of vibration data extracted from a complex kinematic chain working under different failure conditions.

The paper is structured as follows: Section 2 introduces the theoretical aspects of the vibration in rotating machinery and the theory behind ANFIS for time series forecasting. In Section 3, the basis of the proposed method is explained. Section 4 presents the verification of the method capabilities with experimental results of vibration data extracted from a kinematic chain. Conclusions are presented in Section 5.

2. Theoretical Background

2.1. Vibration Analysis in Rotating Machinery. The classical vibration condition monitoring schemes, focused on mechanical and electrical defects, are based on the consideration of specific vibrational frequency components [22]. The spectral analysis of the vibration signature, either in the motor case, in an external bearing house, or in any specific component, allows the evaluation of characteristics, fault frequencies related to cyclic single-point defects. For instance, in bearing failures, the fault frequencies presented in the vibration spectrum are related to the bearing design geometry and follow the well-known equations showed in [23]. Vibration analysis is also applied to diagnose the degradation of rotor bars in induction motors. The failure here is presented in a form of a partial or complete crack in the section of the bar. Such bars may break because of manufacturing defects, frequent starts at rated voltage, thermal stresses, and/or mechanical stress caused by bearing faults and metal fatigue. Furthermore, the diagnosis of the broken rotor bar is performed by calculating specific spectral indexes [24].

Nevertheless, these characteristic frequencies of the vibration modes, apart from possible electrical and mechanical noise during the acquisition, are affected by the specific structure of the considered system (i.e., the kinematic chain) and the component deterioration stage. Indeed, the intensity of the impact produced between raceways and rolling elements, under a bearing defect, is attenuated throughout the propagation from the generation point to the transducer location. Therefore, the vibration measurement is done, generally, as close as possible to the mechanical element under test. Yet, though favoring the measurement of bearing faults vibration effects, the resulting spectrum contains additional vibration modes produced by the rest of the mechanical interactions. These additional vibration frequencies mask or complicate the comprehension of the signal.

The detection of one of the characteristic fault frequencies in the resulting spectrum should be interpreted as an existence of the corresponding fault in the component. However, the absence of clear characteristic fault frequencies should not be interpreted as a completely healthy condition. In this regard, a more general approach to the monitoring of the electromechanical components can be found in the literature as a four-stage process [25]. These four stages are related to the apparition of failure effects in some of the four zones in which the vibration spectrum is divided for electromechanical failure detection. A graphical example of the frequency spectrum of the vibration signal versus the different spectral zones in which electromechanical failures occur is presented in Figure 1. The frequency band $Z_1$ is related to electromechanical failures, eventually in an advanced degradation stage or presenting high severity, which affects or is related to the rotational speed of the machine. This kind of failures appears in $Z_1$ as sidebands of the fundamental frequency. $Z_2$ is related to the regular degradation such as gears or bearings. Indeed, $Z_2$ is the frequency band in which the characteristic bearing faulty frequencies are located. $Z_3$ corresponds to a high frequency band in which early or incipient degradation occurs. The failure detection in this area is generally done by means of feature bases characterization, since the failure pattern is not as specific as the characteristic ones in the low frequency bands. $Z_4$ covers the very-high frequency part of the spectrum. This area is related to the premature degradation of any electric or mechanical component. Nevertheless, it is also very close to the electrical noise inherent in every electromechanical system. Due to this fact, the patterns in this area are very chaotic and are rarely considered for regular diagnosis purposes.

In conclusion, the first three spectral zones are of major interest from the diagnosis and the vibration monitoring point of view. The forecasting method is intended to take advantage from this spectral decomposition in order to isolate each frequency band and gain forecasting performance by modeling each band, from $Z_1$ to $Z_3$, with a dedicated ANFIS model. It should be noticed that $Z_4$ is discarded for forecasting purposes since the noisy and random patterns make the model very complex with regard to the vibration information provided by this frequency region.

2.2. Adaptive Neurofuzzy Inference System. There are different techniques to design forecasting models. Yet, the primarily used schemes are frequently based on the concept of hybrid forecasting, which means that the method integrates different techniques in order to take advantage of each one involved. In this topic, one of the most important hybrid systems is ANFIS [26]. This method fuses the parametric adaptability of artificial neural-networks and the generalization capabilities of the fuzzy logic. ANFIS based prognostic

![Figure 1: Main frequency zones in mechanical degradation.](image-url)
A vibration system offers a reliable and robust condition predictor, since it can capture the system dynamic behavior quickly and accurately [27].

The neural-fuzzy architecture can be divided into five different layers, as shown in Figure 2. For the specific architecture shown in this figure, the method aims to find a relationship between the output signal, \( z \), and a set of different input signals, \( x \) and \( y \), by means of the activation of different polynomials regarding the fuzzification of the input signals by means of the membership functions, \( A_1, A_2, B_1, \) and \( B_2 \). The output is computed regarding the pertinence of the inputs and the two if-then rules as defined in (1).

The input layer, \( L_1 \), is in charge of addressing the input signals to a node. Each node, \( g \), of each layer, \( j \), is an adaptive node with an output function, \( O_g \), defined in (2), where \( \mu_{A_j} \) and \( \mu_{B_j} \) represent the membership degree of the respective input with regard to the membership functions. In the fuzzification layer, \( L_2 \), the weight of each rule is computed by means of a fuzzy AND operation that multiplies the input signals. The outputs of this layer, \( W_g \), represent the activation variables of a fuzzy rule (3). The fuzzy rules definition layer, \( L_3 \), provides the antecedent consequence statements of fuzzy logic. Each node calculates the reason between the \( g \)th activation variable and the sum of all the activation variables (4). The outputs of this layer are called normalized activation variables, \( \overline{W}_g \). In the defuzzification layer, \( L_4 \), each node computes the contribution of the \( g \)th fuzzy rule over all the outputs (5). It should be noticed that \( p_g, q_g, \) and \( r_g \) are the polynomial coefficients that would be estimated during the learning phase of the algorithm. The outputs of each node are named consequents. The nodes of the output layer, \( L_5 \), calculate the output signal, \( z \), as the summation of all the input signals (6). Consider

\begin{align*}
\text{Rule 1:} & \quad \text{if } x \in A_1 \cup y \in B_1 \rightarrow z_1 = p_1x + q_1y + r_1 \\
\text{Rule 2:} & \quad \text{if } x \in A_2 \cup y \in B_2 \rightarrow z_2 = p_2x + q_2y + r_2 \\
O^1_g &= \mu_{A_g}(x), \quad g = 1, 2 \\
O^1_g &= \mu_{B_{g+2}}(y), \quad g = 3, 4 \\
O^2_g &= \overline{W}_g = \mu_{A_g}(x) \cdot \mu_{B_{g+2}}(y), \quad g = 1, 2
\end{align*}

\( O_g = \overline{W}_g = \frac{w_g}{w_1 + w_2}, \quad g = 1, 2 \) (4)

\( O_g = \overline{W}_g \cdot z_g = \overline{W}_g \cdot (p_gx + q_gy + r_g), \quad g = 1, 2 \) (5)

\( O_g = z = \sum_{g} \overline{W}_g \cdot z_g = \sum_{g} \overline{W}_g \cdot z_g \). (6)

The ANFIS structure allows the consideration of a data-driven modeling approach with adaptive rule changing capability and fast convergence rate and does not require extensive experience about the process to include such patterns in the fuzzy rules.

### 3. Methodology

The objective of the proposed method is to model and forecast the evolution curve of the vibration’s RMS signal during the start-up and the thermal stabilization of an electromechanical actuator. The main point is to obtain a forecasting model of the RMS capable of giving the future value with enough resolution taking into consideration the dynamics of different failures occurring to the system. After the output of the method, the presence of the failure can be anticipated by knowing the future value of the RMS and a simply statistical characterization as a diagnosis method.

Therefore, the challenge is to develop a vibration model with enough performance, in terms of forecasted signal error and forecasting horizon that considers the dynamics modes generated by different faulty scenarios of different elements of the electromechanical actuator. The general block diagram of the method is shown in Figure 3, including the training procedure and the online operation. The main limitation in the modeling of complex signals is that a unique forecasting model is not able to capture all the particular dynamics that different fault scenarios may cause. In this regard, the method uses a decomposition of the vibration signal in the three frequency bands in order to provide resolution and generalization towards multiple failure scenarios by improving the response towards their specific spectral components dynamics. Then, the previous prefiltering stage and the RMS calculation are the author’s contribution in order to increase the model resolution by limiting the amount of dynamics that each ANFIS model should face.

#### 3.1. Training Procedure.

The training procedure corresponds to the off-line learning process in order to do the following: (i) tuning of the structure of the forecasting models, (ii) analysis of the optimum inputs delays, and (iii) identification of the longest forecasting horizon.

The first step, Step 1, aims to decompose the vibration signal, \( y(t) \), with a length \( L \), in three different details that are modeled with a specific ANFIS. This decomposition is made with Finite Impulse Response (FIR) digital filters; the cutting frequencies follow theoretical considerations of mechanical fault appearance explained in Section 2.1. Then, the output signal of a certain filter \( Z_i \) is considered to be \( y_{z_i}(t) \). Three FIR filters are defined:
Prior to the model generation, the identification of the best past intervals to enhance the forecasting performance is proposed as Step 3. The optimization process gets simple as the search space is being bounded. The interval selection is made by calculating the correlation of the objective signal, $E_{Z_i}(t+p)$, with successive delays of the same signal, $E_{Z_i}(t-n)$, where $n \in [0, 10p]$. The resulting correlation coefficient is obtained by (8) and represents a measure of statistical similitude between $y(t)$ and its successive delays represented by $n$. The evolution of this coefficient as $n$ increases shows signal oscillation modes and periodicities in which the target signal and its delays are significantly correlated. Then, the interval is selected by setting a correlation threshold, $C_{th}$, and selecting those intervals with the highest accumulated correlation:

$$
C_{coeff}(n) = \frac{\text{cov}(E_{Z_i}(t-n), E_{Z_i}(t+p))}{\sigma(E_{Z_i}(t-n)) \cdot \sigma(E_{Z_i}(t+p))} \quad (8)
$$

for $n = 0, \ldots, 10p$.

It is common to find modeling situations, especially with vibration signals, in which it is necessary to deal with different training sets or operating conditions. In these situations, the search for the optimal intervals turns into an iterative process that tries to localize those intervals with a consistent correlation, and once identified, select the maximum common range that accumulates the highest correlation.

The following step, Step 4, consists in the generation and the training of the ANFIS based models to predict the energy of each filter output, being, respectively, $M_{Z_1}$, $M_{Z_2}$, and $M_{Z_3}$. The structure of the models is fixed and corresponds to (i) the current value of the energy in the $i$th filter, $E_{Z_i}(t)$, (ii) a past value of the energy in the $i$th filter delayed $n_1$ samples, $E_{Z_i}(t-n_1)$, and (iii) another past value of the RMS in the $i$th filter delayed $n_2$ samples, $E_{Z_i}(t-n_2)$.
with the ANFIS model give the predicted vibration energy $E(t + p)$ as shown in

$$E(t + p) = \sum_{i=1}^{3} \text{ANFIS}_{Z_i}(E_{Z_1}(t), E_{Z_2}(t - n_1), E_{Z_3}(t - n_2)).$$

Although each model presents the same quantity of inputs, the delayed samples $n_1$ and $n_2$ could be different for each filter depending on the optimization results. The identified correlation intervals are here introduced as boundaries for the Genetic Algorithm (GA). The cost function of the optimization procedure is defined as the Root-Mean-Square Error (RMSE) of the forecasting model that can be seen in (10). The three models are trained independently with all the available information regarding the different operating conditions and failure scenarios. Ultimately, in Step 5 the final forecasting outcome of the vibration is obtained by means of the direct combination of three filters models outputs.

In order to evaluate the performance of the models, classical statistical metrics have been used, which are RMSE defined in (10), Mean Absolute Error (MAE) in (11), and Mean Absolute Percentage Error (MAPE) in (12). These metrics have been widely used and accepted from the research community as indicated in [28] or [29]. RMSE is one of the most used performance metrics for the development of forecasting models; it is a measure of the standard deviation of the differences between predicted values and observed values. It is useful in order to analyze the global behavior of the model, but it is very sensitive to the amplitude. In this regard the MAE is used to evaluate the forecast since it is less sensitive to outliers. Finally, the MAPE helps to determine the mean deviation of each sample normalized by the amplitude, so it helps to unify the scale and can be used to compare the errors of signals with different levels of amplitudes:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{L} (E(t) - \hat{E}(t))^2}{L}}$$

(10)

$$\text{MAE} = \frac{\sum_{t=1}^{L} |E(t) - \hat{E}(t)|}{L}$$

(11)

$$\text{MAPE} = \frac{\sum_{t=1}^{L} \left| \frac{E(t) - \hat{E}(t)}{E(t)} \right|}{L} \cdot 100\%.$$  

(12)

3.2. Online Behavior. One of the primary characteristics of the proposed method is that it is prepared to work continuously online, that is, receiving new data from the machine as shown in Figure 3. Therefore, after the training procedure of the method, vibration signal is acquired periodically from the machine under study. Then, the signal is filtered with the defined three digital filters and the corresponding RMS values are calculated. As the structure of the models and the best past inputs are already defined, the trained models are used to project the three energies among the specified $p$. Finally, the output of the three models is combined giving with it the new forecasted value of the vibration energy.

4. Experimental Results

The test bench used to obtain experimental vibrations is shown in Figure 4. This test bench consists of a kinematic chain composed of a three-phase 1492 W induction motor, WEG 00236ET3E145T-W22, whose speed is controlled by a variable frequency drive (VFD), and WEG CFW08, the operating speed being fixed to 60 Hz for all experiments. A 4:1 ratio gearbox, BALDOR GCF4X01AA, is used to couple the drive motor to a DC generator, BALDOR CDP3604. The DC motor is used as a noncontrolled mechanical load that comprises around 20% of the nominal torque of the driving motor. The DAS is a proprietary low-cost design based on field programable gate array technology. The output rotational speed is obtained by using a digital encoder; the motor start-up is controlled by a relay in order to automatize the test run. A 12-bit 4-channel serial-output sampling analog-to-digital converter, ADS7841, is used in the onboard data acquisition system (DAS).

Vibration signal from the perpendicular plane of the motor axis is acquired using a triaxial accelerometer, LIS3L02AS4, mounted on a board with the signal conditioning and antialiasing filtering. Sampling frequency is set to 3 kHz for vibration acquisition. The data retrieved by the DAS is stored in a regular computer (PC).

Four scenarios have been considered, that is, the healthy condition, HC, a bearing failure (BF) condition, a half-broken rotor bar (HBBR) condition, and full-broken rotor bar (FBRR). The detail of the failures is shown in Figure 5. The failure in a 6205-2ZNR bearing has been induced by means of a hole of 1.191 mm in the outer race; the hole has been produced by a tungsten drill bit. HBBR failure is artificially produced by drilling a 6 mm hole with a depth of 3 mm that corresponds mostly to the 22% of the section of the rotor bar. Finally, FBRR is produced by a through-hole with a diameter of 6 mm and a depth of 14 mm, which corresponds to the complete section of the rotor bar.
Two different datasets with the same length have been acquired; the first one is used to train the proposed method, and the second dataset is used to validate the performance of the method. In this regard, periodical acquisitions of 30 seconds containing 90 ksamples are obtained from the accelerometer; the acquisitions are temporally spaced at 30 seconds. The average duration of the experiments for all operating conditions is configured to be 3600 seconds.

The selected duration allows obtaining the complete response of the vibrations with regard to the thermal stability of the electromechanical system. This is done in order to consider the thermal evolution of the vibrations as part of the system and facilitate the representation of the failure. The proposed approach brings the experimental validation closer to the real behavior of complex industrial machinery, in which the behavior of the vibration presents a nonconstant dynamic from the starting point till its steady state. As a result, the modeling of the vibration considering the thermal evolution represents an additional challenge to the validation of the proposed method. Furthermore, Figure 6 shows the behavior of the vibration in healthy condition, HC, with regard to the temperature of the motor from the starting point till the end of the experiment; note that the thermal steady state is reached.

As can be appreciated in Figure 6, the raw signal of the acceleration gives no important information regarding the condition of the system since it is difficult to detect underlying patterns such as the affection of the thermal stabilization, since it is modulated in the oscillatory waveform of the vibration. For this reason, the necessity of calculating a statistical feature representative of the condition of the system is justified after the preprocessing stage.

The vibration signal filtering and the corresponding windowed RMS calculation are carried out by means of FIR filters and considering the rotating speed, the sampling frequency, and, also, the theoretical background with regard to mechanical failures. As a result, the cut-off frequencies of the filters in order to isolate the three primary frequency bands, $Z_1$, $Z_2$, and $Z_3$, are shown in Table 1. The attenuation is fixed to get $-60$ dB in the cutting frequency. It should be noticed that as it was expected, the order of the filter decreases with the bands since the bandwidth of each filter increases.

Finally, the RMS feature is calculated for each filter output; the results are 3 RMS signals that summarize the information regarding the vibration of the electromechanical actuator under a concrete failure condition and are the target signals to be modeled. For this experimental application, the temporal window is configured to be $w_h = 3000$ samples, which corresponds to a window of 1 second of actuator operation, and an overlapping factor of $ov_{RMS} = 25\%$ is selected to smooth the filter response. The resulting signals to be predicted are shown in Figure 7 for all considered failures. How the faulty signal corresponding to a bearing fault is mostly located as it was expected in the second filter, $Z_2$, from 100 to 500 Hz can be appreciated. The figure justifies the decomposition of the signal since different spectral behavior

<table>
<thead>
<tr>
<th>Filter</th>
<th>Cut-off frequency 1</th>
<th>Cut-off frequency 2</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{Z1}$</td>
<td>0 Hz</td>
<td>120 Hz</td>
<td>248</td>
</tr>
<tr>
<td>$F_{Z2}$</td>
<td>120 Hz</td>
<td>350 Hz</td>
<td>136</td>
</tr>
<tr>
<td>$F_{Z3}$</td>
<td>350 Hz</td>
<td>1000 Hz</td>
<td>120</td>
</tr>
</tbody>
</table>
Figure 7: RMS of each filter output for all operating conditions: (a) HS, (b) BF, (c) FBRB, and (d) HBRB.

Figure 8: RMS of the vibration signal during the thermal stabilization period. Note that the different failures can be identified during this period, and as it was expected, the BF case shows the highest vibration.

can be observed in each filter output with the presence of failure. This means that each specific failure focuses its affection to a certain frequency band, conclusion that matches with the theoretical considerations exposed previously. The next step in the method deals with the design and the generation of the optimal dedicated ANFIS forecasting model for each filter output.

Finally, to understand the fundamentals of the method, Figure 8 shows the characteristic RMS signature of each faulty condition. It demonstrates how the calculated RMS provides the necessary information to distinguish the different failures presented in the system during the thermal stabilization of the actuator. It also justifies the suitability of modeling the RMS of the vibration signal for forecasting and diagnosis purposes. It should be remarked that these RMS are the target signals to be obtained by the models in the defined forecasting horizon \( p \), the final result of the proposed method.

4.1. Study of Forecasting Horizon Selection. The selection of the forecasting horizon is a crucial step that needs to be faced before the modeling. Most of the related literature selects the forecasting horizon simply by application requirements. Yet, the procedure to analyze the behavior of the target signal to find the optimal forecasting horizon is not established in the literature. It is proposed, then, in this work, to select the optimal horizon by analyzing the performances of the method at different horizon values.

The proposed analysis interval comprises from \( p = 1 \) sample, 1 sec., to \( p = 600 \) samples with an increase ratio of \( P_{inc} = 5 \). Once the evaluation range is established, the proposed models are trained and evaluated iteratively for all the specified horizons. The forecasted outcome of the three models is evaluated by the proposed performance metrics considering the four operating conditions. As a result, the mean performance of all operating conditions with regard to the metric used and the evaluated model is shown in Figure 9. Nevertheless, the error of the model depends on the performance of the optimization algorithm used to select the best past inputs; for this reason, a smoothed response of the error has been generated in order to estimate the average expected error with the selected horizon.

The general results show that as it was expected in time series forecasting applications in which the forecasted signal does not present a periodic pattern, the lowest error can be found in the initial area that comprises \( p = [1\text{--}50] \). Then, the error increases for all the models, especially the model of the first frequency band that exhibits a linear behavior by increasing gradually with the forecasting horizon. However, for the model of the second and third frequency bands, there is a second decay of error growth around \( p = 200 \). This area also represents a low error region that can be utilized to develop the forecasting models, and, thus, \( p = 200 \) is selected for the current application.

In specific conclusions, it should be noticed that low frequency dynamics are easier to be modeled independently of the selected forecasting horizon, since they represent
the tendency of the vibration, and thus they suffer slower changes among time. It should be observed how error increases in the higher frequency bands, since fast dynamics are more punctual and closely related to adjacent samples of the current vibration value.

4.2. Competency of the Method. The proposed method begins once the forecasting horizon is selected and the RMS of the three filter outputs is calculated. The structure of the three models, $M_{Z1}$, $M_{Z2}$, and $M_{Z3}$, is fixed by a single output, $E_{Z1}(t + p)$, and three inputs (i) $E_{Z1}(t)$, (ii) $E_{Z1}(t - n_{1i})$, and (iii) $E_{Z1}(t - n_{2i})$.

At this point, the problem derives to the selection of the best past values, $n_{1i}$ and $n_{2i}$, for each frequency band. To do so, a generic GA based optimization algorithm has been used. Yet, in order to increase the effectiveness of the optimization and reduce the associated computational cost in terms of search iterations, the method proposes a prior past index interval finding step. The correlation coefficient (8) has been calculated for each filter considering the average correlation between all the considered scenarios.

Due to the fast dynamics of the acquired vibration signal, the delayed signals of more than 20 acquisitions are considered to be nonsignificant for the application; therefore, the total number of delayed signals is set to 600. The amount of necessary correlation to consider a strong relation between the target signal and its self-delayed signals is not defined. Thus, the experimental results show that useful intervals are considered with a correlation threshold above $40\%$ ($C_{th} > 0.4$). Additionally, in order to avoid finding empty intervals in which only few indexes are found, intervals with a length lower than 5 indexes are discarded. The interval that accumulates the highest $C_{th}$ value is selected as the upper and lower values constraints of the optimization algorithm.

Figure 10 shows the average correlation coefficient for the experimental set-up.

The correlation of the target signal in $Z_1$ shows a quick drop in the correlation coefficient in the first 20 delays; this behavior is characteristic of vibration forecasting applications, such as the used test bench, since the dependence of the target signal with the correlation delays is high. The correlation coefficient decreases continuously until delay 200, where there is an area between delays 200 to 400 in which the coefficient increases moderately. This coefficient increase is due to a certain similarity that can be found in the signal. This interval can be exploited by the GA in order to find the best index for $M_{Z1}$. This fact confirms the previous selection of the forecasting horizon for $Z_1$, since the correlation increases in the same area that the selected horizon does.

The correlation of $Z_2$ presents a smoother decrease of the coefficient as the delay time increases. Moreover, it presents some oscillations as the delay increases; this resonance corresponds to signal periodicities that are characteristic of the dynamics of the signal. These periodicities can be used to enhance the forecasting capabilities of the model, since the past signals introduced are more correlated to the target signal. In this regard, the selected interval in $M_{Z2}$ to exploit these periodicities is set at $I_{Z2} = [200 \ 350]$ seconds. Note that, another time, the periodicity also matches with the selected horizon extracted from the proposed method error analysis. Additionally, notice that the correlation in the second band is the highest with regard to the others. This happens because the fundamental harmonic of the operating speed is allocated primarily in the second band, and, therefore, it presents a soft dynamic easier to be modeled.

Finally, $Z_3$ exhibits a lower correlation coefficient for the entire analyzed interval. This behavior is expected since it corresponds to the highest frequency band. In vibration
Table 2: Selected past index as a result from the constrained GA based optimization.

<table>
<thead>
<tr>
<th></th>
<th>(n_1)</th>
<th>(n_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M_{Z_1})</td>
<td>217</td>
<td>389</td>
</tr>
<tr>
<td>(M_{Z_2})</td>
<td>285</td>
<td>397</td>
</tr>
<tr>
<td>(M_{Z_3})</td>
<td>247</td>
<td>348</td>
</tr>
</tbody>
</table>

As a result, the forecasting models have been trained and validated using the validation set. The performance of the model for all operating conditions is quantitatively analyzed by means of RMSE, MAE, and MAPE coefficients, shown in Table 3.

The modeling of each frequency band vibration presents remarkable results for all operating conditions, as it achieves a MAPE lower than 5% in all operating conditions. Nevertheless, the healthy condition presents a dynamic that is more difficult to model in comparison with the others. The behavior of the vibration in healthy condition presents a steady pattern since there is an absence of a strong signature related to the failure such as the BF case. This healthy pattern should seem a priori easier to be modeled; however, as the model is being forced to learn different dynamics, it dedicates more efforts to those datasets that are similar to each other.

As a consequence, the healthy state presents a different dynamic with regard to the faulty scenarios and therefore the modeling error slightly increases, around 2% of error increase in terms of MAPE. The learning ability of the model is finite; thus, there should be a trade-off between the number of different dynamics that the model needs to consider and the performance achieved by the model versus a single set. Furthermore, regarding the models, the best performance is achieved by \(M_{Z_3}\), since it presents the lowest RMSE for all datasets. This means that the model does not present outliers and the forecasted outcome follows the target signal in its mean value. Regarding the operating condition, the bearing failure is the easiest scenario to be modeled in bands \(Z_1\) and \(Z_2\), since the fault signature presents a monotonic dynamic easily to be captured by the model. The outputs of three models are combined to obtain the final forecasting information. In this regard, the prediction performance during the validation can be seen in Figure 11.
Table 3: Error achieved by the model with the four different sets used in the experimental validation. $E_1$ corresponds to RMSE, $E_2$ to MAPE, and $E_3$ to MAE.

<table>
<thead>
<tr>
<th>$M_{Z1}$</th>
<th>HC</th>
<th>$E_1$</th>
<th>0.014</th>
<th>$M_{Z2}$</th>
<th>HC</th>
<th>$E_1$</th>
<th>0.020</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>3.65%</td>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>4.120%</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.1</td>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.127</td>
</tr>
<tr>
<td>$M_{Z1}$</td>
<td>BF</td>
<td>$E_1$</td>
<td>0.010</td>
<td>$M_{Z2}$</td>
<td>BF</td>
<td>$E_1$</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>1.699%</td>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>1.213%</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.09</td>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.097</td>
</tr>
<tr>
<td>$M_{Z1}$</td>
<td>HBRB</td>
<td>$E_1$</td>
<td>0.099</td>
<td>$M_{Z2}$</td>
<td>HBRB</td>
<td>$E_1$</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>1.625%</td>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>2.245%</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.084</td>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.116</td>
</tr>
<tr>
<td>$M_{Z1}$</td>
<td>FBRB</td>
<td>$E_1$</td>
<td>0.018</td>
<td>$M_{Z2}$</td>
<td>FBRB</td>
<td>$E_1$</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>3.397%</td>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>2.815%</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.118</td>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.105</td>
</tr>
<tr>
<td>$M_{Z1}$</td>
<td>Mean</td>
<td>$E_1$</td>
<td>0.013</td>
<td>$M_{Z2}$</td>
<td>Mean</td>
<td>$E_1$</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>2.593%</td>
<td></td>
<td>$E_2$</td>
<td>RMSE</td>
<td>2.5981%</td>
</tr>
<tr>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.0982</td>
<td></td>
<td>$E_3$</td>
<td>MAE</td>
<td>0.111</td>
</tr>
</tbody>
</table>

Figure 11: Results of the forecasting method. (a) Combined output applied to HC. (b) Combined output applied to HBRB. (c) Combined output applied to FBRB. (d) Combined output applied to BF versus the output of the single model method to the same scenario.

Table 4: Performance of the statistical error metrics applied to the combination of all model outputs. Also, the performance of the forecasting without the frequency decomposition is also shown in the table.

| Method | Single modeling | RMSE | MAPE | MAE
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td></td>
<td>0.013</td>
<td>1.69%</td>
<td>0.135</td>
</tr>
<tr>
<td>BF</td>
<td></td>
<td>0.020</td>
<td>2.27%</td>
<td>0.125</td>
</tr>
<tr>
<td>HBRB</td>
<td></td>
<td>0.022</td>
<td>2.82%</td>
<td>0.133</td>
</tr>
<tr>
<td>FBRB</td>
<td></td>
<td>0.025</td>
<td>3.44%</td>
<td>0.139</td>
</tr>
</tbody>
</table>

The results show that when the outputs of the three models are combined, the obtained signals outperform in terms of error the individual results of all models. In this regard, low error is achieved in all the considered scenarios. The method is capable of obtaining a MAPE lower than 2% in all the cases studied. Also the MAE shows the stability of the output by maintaining the absolute error near 0.125 in all cases. Additionally, as can be seen in the figure, and by terms of RMSE, the proposed method presents a smooth response that presents a marginal number of outliers.

To conclude the results, the performance of the proposed method is tested versus a forecasting approach in which the vibration signal is not decomposed. The performance achieved by the second method is named single modeling that can be also seen in Table 4.

The main limitation of the single model approach is that, despite the fact that the model is able to learn the first scenario, HC, with virtually fewer errors than the proposed
method, it is not capable of modeling additional dynamics contained in the other scenarios. For example, as it can be seen in Figure III(d), the model gets lost with the BF set, which, furthermore, was the easiest to be modeled with the method. The point here is that the amount of signal dynamic each forecasting model can face is limited, and thus, by filtering the vibration signal and facing the forecasting problem from a signal decomposition point of view, the final performance and the adaptability of the models towards different behaviors and operating condition increase.

5. Conclusions

This paper presents a novel vibration signal modeling methodology applied to a kinematic chain under different failure conditions. First, it should be pointed that the RMS values of the vibration have been proved to be a representative feature regarding the health condition of the electromechanical actuator.

Furthermore, there are three important aspects in this new method having a strong implication in the vibration forecasting field. The first one is the application of a signal decomposition strategy. It has been proved that signal decomposition allows a better characterization of the signal dynamics and improves the forecasting accuracy by reducing the amount of dynamic that each model should face. The second is the optimal model configuration by means of the proposed analysis of the prediction horizon and time delays. This analysis leads to an optimal configuration of the optimization and thus the simplification of the process assuring the convergence to a reliable and robust forecasting. The third is the calculation of the forecasting value by means of the linear combination of the models. How the combination of the three individual models concludes in a better performance when facing the original forecasting method has been seen; this fact is evidenced when the proposed method is compared with the classical approach with only one model.

For the experimental validation of the method, four different operating conditions have been considered which represent an important range of system condition possibilities. Under these experimental conditions, the proposed prognosis methodology achieves reliable results with an error lower than 2% in all the exposed scenarios. Moreover, considering the analyzed model error parameters, the prognosis methodology shows still enough dynamic range to include additional patterns.

The aim of the study has been to propose a new reliable prognosis methodology for diagnosis analysis under multiple failure conditions. The study is based on different motor faults scenario. Therefore, it should be pointed that the results obtained in this work suggest that this methodology may be also useful for any other component in the kinematic chain.

A limitation of this study is that the affectation of other sources of information, such as the temperature of the motor or the stator current, has not been considered. This additional information should be correlated with the vibration of the system and thus may help to increase the response and adaptation capabilities of the models towards new scenarios. In this regard, further research can be focused on investigating information management methods such as information compression or topology preservation to exploit and add other information available in the kinematic chain in the design of the forecasting models.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported in part by the Spanish Ministry of Education, Culture, and Sport under Grant FPU13/00589. The authors also wish to acknowledge financial support from the Generalitat de Catalunya (GRC MCI, Grant no. SGR 2014-101).

References


Submit your manuscripts at http://www.hindawi.com