

Article

# Entropy Indicators: An Approach for Low-Speed Bearing Diagnosis

Diego Sandoval <sup>1,2</sup>, Urko Leturiondo <sup>1,\*</sup>, Yolanda Vidal <sup>2</sup> and Francesc Pozo <sup>2</sup>

<sup>1</sup> Ikerlan Technology Research Centre, Basque Research and Technology Alliance (BRTA), Pº. J. Mª. Arizmendiarieta, 2, 20500 Arrasate-Mondragón, Spain; dasandoval@ikerlan.es

<sup>2</sup> Control, Modeling, Identification, and Applications (CoDALab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besòs (CDB), Universitat Politècnica de Catalunya (UPC), Eduard Maristany, 16, 08019 Barcelona, Spain; yolanda.vidal@upc.edu (Y.V.); francesc.pozo@upc.edu (F.P.)

\* Correspondence: uleturiondo@ikerlan.es

**Abstract:** To increase the competitiveness of wind energy, the maintenance costs of offshore floating and fixed wind turbines need to be reduced. One strategy is the enhancement of the condition monitoring techniques for pitch bearings, because their low operational speed and the high loads applied to them make their monitoring challenging. Vibration analysis has been widely used for monitoring the bearing condition with good results obtained for regular bearings, but with difficulties when the operational speed decreases. Therefore, new techniques are required to enhance the capabilities of vibration analysis for bearings under such operational conditions. This study proposes the use of indicators based on entropy for monitoring a low-speed bearing condition. The indicators used are approximate, dispersion, singular value decomposition, and spectral entropy of the permutation entropy. This approach has been tested with vibration signals acquired in a test rig with bearings under different health conditions. The results show that entropy indicators (EIs) can discriminate with higher-accuracy damaged bearings for low-speed bearings compared with the regular indicators. Furthermore, it is shown that the combination of regular and entropy-based indicators can also contribute to a more reliable diagnosis.

**Keywords:** pitch bearing; condition monitoring; entropy; low-speed bearings; vibration



**Citation:** Sandoval, D.; Leturiondo, U.; Vidal, Y.; Pozo, F. Entropy Indicators: An Approach for Low-Speed Bearing Diagnosis. *Sensors* **2021**, *21*, 849. <https://doi.org/10.3390/s21030849>

Academic Editor: Steven Chatterton  
Received: 21 December 2020  
Accepted: 23 January 2021  
Published: 27 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



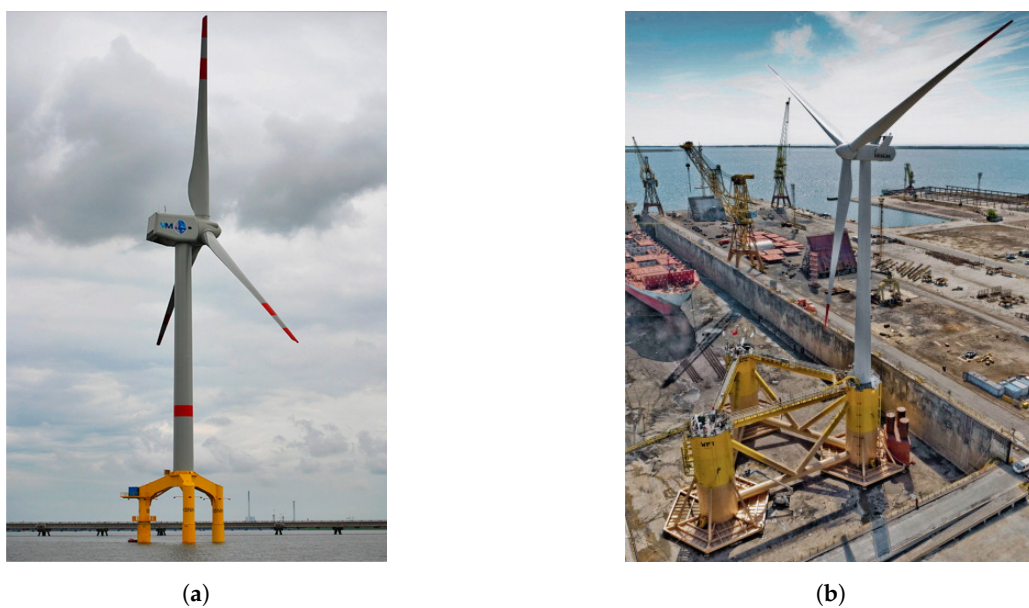
**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Of all existing sources of renewable energy, wind energy is considered to have by the European Union (EU) the largest potential to replace the fossil energy share in the next 30 years [1]. With an estimated required energy production ranging between 230 and 450 GW in Europe by 2050 [2], and with 60% of the investment in renewable energy projects in Europe focusing only on offshore wind projects [3], with the achievement of lower energy prices is a strategy to promote these projects.

Given that almost one-third of the price of 1 MWh from a wind farm is related to the payment of the operating and maintenance (O&M) fixed cost [4], new strategies are needed to reduce the impact of the mentioned cost and to consequently promote wind energy investments. One strategy is condition-based maintenance (CBM), which can alleviate the O&M fixed cost when using a preventive maintenance approach. For convenience, the definitions from standard EN 13372 [5] and EN 13306 [6] will be used. condition-based maintenance (CBM) is maintenance that is performed as governed by condition monitoring (CM) programs, which are the acquisition and processing of information and data that indicate the state of a part or machine over time. One vital aspect of CM is the diagnosis, which is known as the action taken for fault recognition, fault location, and cause identification [6]. Therefore, the diagnosis is fundamental to the achievement of CM programs.

For wind turbines (Figure 1), where several electromechanical parts are involved, bearings are critical elements as they work under varying conditions and high loads. For example, 70% of the downtime caused by the gearbox and the generator are based on a bearing malfunction [7]. Thus, the condition monitoring (CM) publications of wind turbine bearings are mainly focused on gearbox and generators [8–10]. From the bearings present on wind turbines, the pitch bearings are responsible for the pitch movement of the blade therefore, they are fundamental for energy generation control. While regular bearings have diameters that are of the order of centimeters and rotational speeds that are of the order of several thousand revolutions per minute (rpm), pitch bearings are larger (diameters of the order of meters) and operate at lower rotational speeds (10 rpm and below). As mentioned before, pitch bearings are under severe mechanical stress, with loads and moments of several hundreds of kN and thousands of kNm, respectively, [11–13]. Thus, it is particularly challenging to determine condition monitoring (CM) for this type of bearing because of the circumstances previously described.



**Figure 1.** Offshore wind turbines. (a) Wind turbine over a fixed foundation [14]. (b) Floating offshore wind turbine assembly [15].

A common approach for the diagnosis of the bearing is vibration analysis [16,17], as the characteristics in the time and frequency domains of the vibrational response of a bearing change according to its condition [18,19]. The changes mentioned can be noted in the time domain using several techniques. One regular approach found in the literature is the use of signal decomposition algorithms, such as empirical mode decomposition (EMD) [20] and its derivatives ensemble and complementary ensemble EMD [21–23], and autoregressive models [24]. Some other algorithms are inspired by EMD, such as adaptive local iterative filtering [25] and fast iterative filtering decomposition [26]. With respect to the frequency domain, the decomposition of the signal by the Fourier transform [27], the Hilbert–Huang transform (based on EMD) [28], and the continuous wavelet transform [29] can be found. In addition, the analysis of the signal by the kurtogram application [30] or spectrogram [31] are used to detect changes in the frequency characteristics. In general, the decomposition of the vibration signals for detecting and diagnosing low-speed bearing damage are not comparable to the results obtained by their application to regular bearings [32]. The main reason for this is the low energy release from the damaged bearing because of the low speed, which can be masked by background noise [33].

A different procedure for fault detection is the calculation of indicators. The kurtosis, covariance, root mean square (RMS), and skewness are indicators that are widely used in

the literature [34,35]. In addition, other indicators have been developed in the time domain, such as factors [36–38], and Hjorth parameters [39–41]. In the frequency domain, several statistical indicators have their counterparts from the time domain, which are helpful for frequency analysis [42]. Using indicators and the aforementioned techniques, the diagnosis of bearings is usually achieved [40,43,44]. However, these indicators also face the problem of low-speed bearings and their low release of energy to detect possible damage. Therefore, new types of indicators are needed to obtain more accurate information from the vibration signals of low-speed bearings.

An interesting group of indicators is especially appealing for bearing diagnosis, namely entropy-based indicators. These types of indicators have been used with promising results in different fields, such as biomedicine [45–48], economics [49,50], electronics [51], and computer science [52]. The use of entropy-based indicators in vibration analysis can be found in the literature, such as the work by Fei et al. [53], where entropy-based distances are applied for the analysis of vibration signals of rolling element bearing faults. In that work, several spectrum-based indicators are extracted from simulated faulty vibration signals for diagnosis. The research work of Gu et al. [54] uses the minimum average envelope entropy with the Teager energy operator method to diagnose incipient faults on bearings. Zhang et al. [55] decomposed vibration signals from a bearing at speeds ranging from 1720 to 1797 rpm into intrinsic mode functions, to later calculate several multiscale entropy indicators, and to finally diagnose the bearings. Qin et al. [56] used EMD and the energy entropy to select the intrinsic mode functions (IMF) component for feature extraction and subsequently bearing diagnosis. The work by Wang et al. [57] combined the wavelet packet to decompose a vibration signal, select a component according to the entropy indicator criteria, and later extract information of the selected component by using classic indicators (CIs). A similar process was followed by Vakharia et al. [58], where a multiscale entropy-based indicator was used. In general, entropy-based methods are employed for the selection of components from signal decomposition algorithms to avoid noise filtering. Overall, the use of these techniques is applied mainly for general-purpose bearings, with rotational speeds of several thousand rpm.

A small number of articles in the literature present the use of entropy-based indicators in combination with several techniques previously mentioned, such as Huo et al. [59], who used an adaptive multiscale weighted entropy and a support vector machine for the pattern recognition of damaged bearings. Yang et al. [60] worked with a multiscale symbolic dynamic entropy indicator in combination with a support tensor machine-based binary tree for health condition identification. Another example is the work by Fu et al. [61], who used a composite multiscale fine-sorted dispersion entropy, and principal component analysis for the fault diagnosis of bearings. This number decreases for articles referred to as low-speed bearing diagnosis. In this field, Caesarendra et al. [62,63] presents the application of nonlinear methods of feature extraction, which includes one entropy-based method to a bearing at 1 rpm. An et al. [64] utilized variational mode decomposition, which was later combined with a permutation entropy for fault diagnosis of a bearing at 260 rpm.

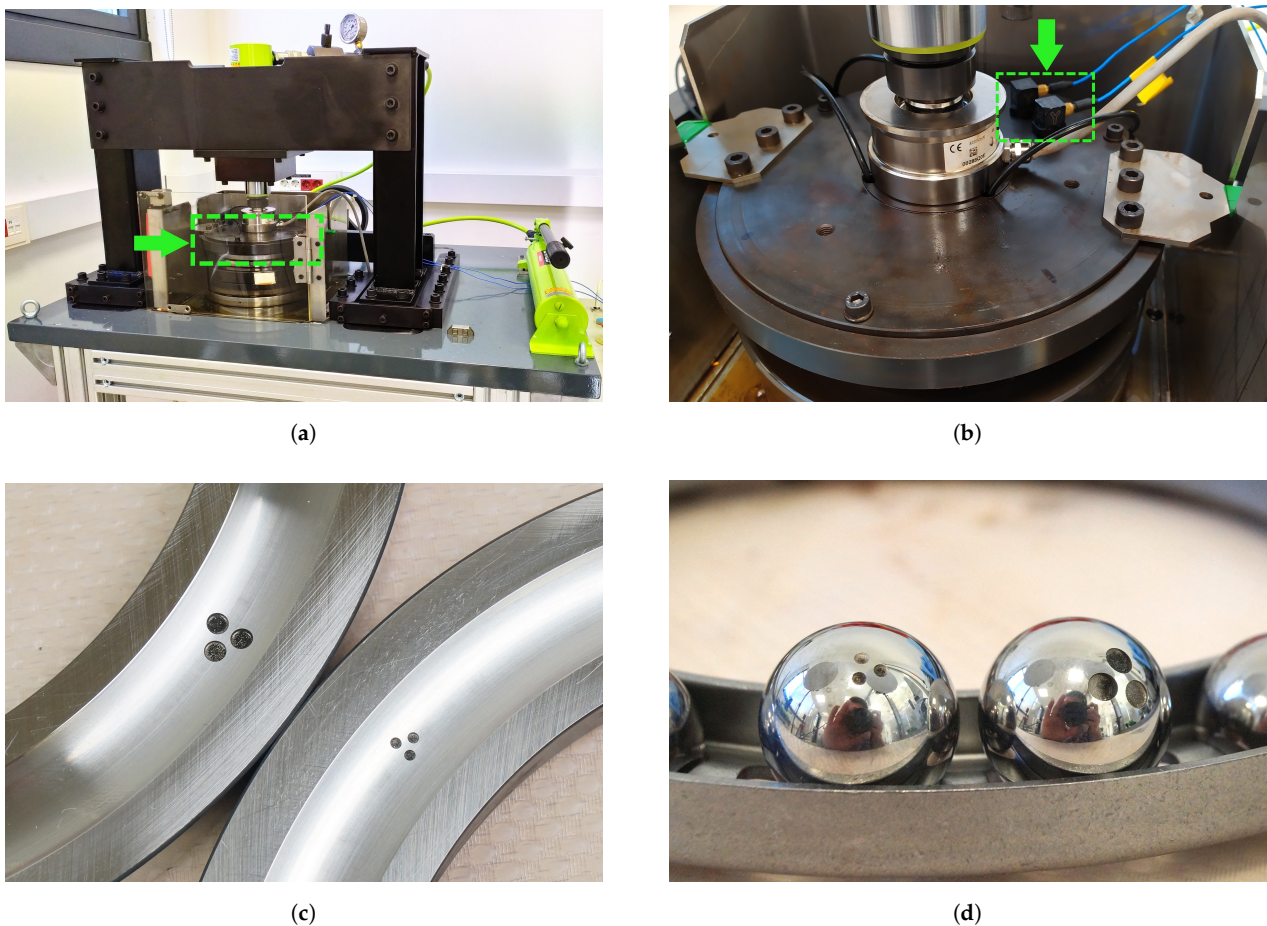
Although multiple articles use one entropy indicator for fault diagnosis, according to the published literature, the use of several entropy-based indicators for the classification of low-speed bearings is not capable of achieving damage localization. Consequently, the contribution of this study is the research on the feasibility of a set of entropy-based indicators for the classification and diagnosis of low-speed bearings with no previous signal decomposition or filtering. An advantage to highlight in the present work is the absence of signal preprocessing before the calculation of the indicators. From this perspective, a limited number of articles can claim this advantage, such as the work from Glowacz et al. [65], where the calculation of indicators based on the differences between frequency spectra of the vibration signals from different damage scenarios is proposed. From the use of several entropy-based indicators, the information extracted is compared to a set of classical indicators. To analyze how entropy-based and CIs can complement each

other, the combination of both types of indicators is studied. As the RF classifier provides excellent performance in pattern recognition [66–70], this study attempts to utilize the RF classifier to construct a classification model for fault recognition of low-speed bearings.

The overall structure of the paper has been divided into six sections, including this introductory section. Section 2 describes the experimental setup, followed by the description of the theory behind the entropy and statistical indicators. The methodology is explained in Section 4, and the results and their discussion are presented in Section 5. Finally, conclusions are given in Section 6.

## 2. Experimental Setup

Whereas open data sets are available for regular bearings in different repositories [71–73], low-speed bearing data are scarce and usually proprietary. In this work, to evaluate the proposed strategy, a dataset of vibration signals from low-speed bearings was obtained via an experimental test rig using thrust ball bearings [74] (Figure 2a). Thus, the specific bearing operating conditions were established according to the literature [22,75,76]: (i) The bearings rotated at a very low speed (between 5 and 10 rpm), and (ii) the test rig applied a high force (50 kN) to the bearing.



**Figure 2.** Test rig for low-speed bearings: (a) General view of the test rig [74]. The position of the upper support disc is highlighted with a green arrow. (b) Detail of the surface of the upper support disc. The position of the sensor is highlighted with a green arrow. (c) Damage seeded on shaft washers. (d) Damage seeded on ball bearings.

The vibration signals were acquired via a 356A16 ICP acceleration sensor from PCB Piezotronics® (highlighted by a green arrow in Figure 2b), which was placed using a wax adhesive on the surface of the upper support disc (highlighted by a green arrow in

Figure 2a). A cDAQ 9174 data acquisition system from the National Instruments® was used to acquire the raw vibration signal from the sensor at a sampling rate of 10.24 kHz.

The described experimental setup was used to obtain a dataset with healthy and damaged scenarios at different constant rotational speeds of 5, 8, and 10 rpm. The damaged scenarios include different locations and sizes thus, five FAG 51226 thrust ball bearings were used. With respect to location, damage to the shaft washer and ball bearing was introduced. Moreover, small and large damage sizes were considered for both locations. The damage used in each scenario was seeded as the shape of three circles having the same diameter and 0.3 mm deep by electro-discharge machining on the surface of the element, with special considerations for the curvature in the raceway of the shaft washer (Figure 2c) and ball bearing (Figure 2d). The circles were equally spaced by 1 mm from each other, and different diameters were used according to the damage size: 1.5 mm for small damage, and 3 mm for large damage.

In summary, the data set included time signals for different speeds (5, 8, and 10 rpm) in each of the five scenarios, which are denoted as follows: healthy scenario (HS), small raceway shaft washer damage scenario (RS), large raceway shaft washer damage scenario (RL), small ball bearing damage scenario (BS), and ball bearing damage scenario (BL).

### 3. Theoretical Background: Classic and Entropy Indicators

This section presents a brief review of the most common indicators. First, a selection of indicators found in the literature is summarized in Section 3.1. Secondly, an explanation of the proposed indicators based on entropy is presented in Section 3.2.

#### 3.1. Classic Indicators

The reviews in the literature summarize that to obtain information on the state of a low-speed bearing, several indicators can be calculated [7,16,17,77–83]. From all indicators found in the literature, a specific group is regularly used, which are referred to the remainder of this article as CIs. For the sake of clarity, a sampled vibration signal defined as  $x = \{x_1, x_2, \dots, x_N\}$ ,  $N \in \mathbb{N}$  is used for the mathematical description of CIs.

Many of the selected CIs are computed in the time domain. From all indicators, the most widely known and used is the RMS, which is calculated as:

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}. \quad (1)$$

In several cases, the value of RMS increases as the faults in a bearing are developed [34,84]. As this statement is not true for all types of bearings and faults, researchers are forced to find more reliable indicators. Complementary to RMS, the upper bound  $h_u$  and lower bound  $h_l$  of the histograms of the vibration signal are also indicators that may be sensitive to changes in the state of bearings [79,80,85]. These indicators are calculated as follows:

$$h_u = \max_{i=1, \dots, N} (x_i) + \frac{\Delta}{2}, \quad (2)$$

$$h_l = \max_{i=1, \dots, N} (x_i) - \frac{\Delta}{2}, \quad (3)$$

where

$$\Delta = \frac{\max_{i=1, \dots, N} (x_i) - \min_{i=1, \dots, N} (x_i)}{N - 1}.$$

Other indicators in the literature are specifically designed (but not limited) for bearings, such as shape factor (SF), crest factor (CF), impulse factor (IF), and margin factor

(MF) [37,86]. By definition, SF is related to the shape of the signal independent of its time extension. The calculation is performed as follows:

$$SF = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}}{\frac{1}{N} \sum_{i=1}^N |x_i|}. \quad (4)$$

In contrast to SF, CF and IF measure the impulsiveness of the signal from the bearing. Therefore, both indicators are specially used for spiky signals. Although the mathematical formulation for both indicators is similar, the main difference is the use of the RMS of the signal in the CF and the mean of the absolute values for the IF. Therefore, CF and IF are defined as follows:

$$CF = \frac{\max_{i=1, \dots, N} |x_i|}{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}}, \quad (5)$$

$$IF = \frac{\max_{i=1, \dots, N} |x_i|}{\frac{1}{N} \sum_{i=1}^N |x_i|}. \quad (6)$$

It is worth highlighting the mathematical relation among SF, CF, and IF as  $SF = \frac{IF}{CF}$ .

The last of the mentioned indicators, MF (also known as the clearance factor or clearance indicator), has been studied as its value is influenced when a damaged raceway is in contact with rolling elements [87,88]. It is calculated as:

$$MF = \frac{\max_{i=1, \dots, N} |x_i|}{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{|x_i|}\right)^2}. \quad (7)$$

In addition to previous indicators, the use of statistical moments in the time domain can be found in the literature [44,89,90]. The most widely used statistic standardized moments are the variance ( $\sigma^2$ ), the skewness ( $\gamma$ ), and the kurtosis ( $\kappa$ ). The value of  $\sigma^2$  (or second normalized moment) can be used to measure the dispersion around the mean value of the analyzed signal, and it is obtained by:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}, \quad (8)$$

where  $\mu$  is the mean value of the analyzed signal (also known as the first normalized moment). The skewness, or the third normalized moment, quantifies the number of times that a signal value is greater or smaller than the mean of the signal. It is calculated as:

$$\gamma = \frac{\sum_{i=1}^N (x_i - \mu)^3}{N\sigma^3}. \quad (9)$$

The fourth standardized moment, called kurtosis, is another value that is used for the analysis of a signal. The kurtosis measures the concentration of the signal around its mean value.

$$\kappa = \frac{\sum_{i=1}^N (x_i - \mu)^4}{N\sigma^4}. \quad (10)$$

Other statistical indicators can be obtained based on the derivatives of the signal. The first derivative of the vibration signal  $x' = \{x'_1, x'_2, \dots, x'_{N-1}\}$ ,  $N \in \mathbb{N}$  is approximated as:

$$x'_i = x_{i+1} - x_i, \quad i = 1, \dots, N - 1,$$

where  $N$  is the length of  $x$ . The second derivative  $x'' = \{x''_1, x''_2, \dots, x''_{N-2}\}$ ,  $N \in \mathbb{N}$  can also be approximated as:

$$x''_i = x'_{i+1} - x'_i, \quad i = 1, \dots, N-2.$$

These definitions are used by the Hjorth parameters [91], which are the activity ( $\mathcal{A}$ ), mobility ( $\mathcal{M}$ ), and complexity ( $\mathcal{C}$ ) [39–41]. The first activity measures the variance of the amplitude of the signal that is analyzed:

$$\mathcal{A}_x = \sigma_x^2, \quad (11)$$

where  $\sigma_x$  is the standard deviation. The mobility is the second parameter mentioned, and measures the relative average slope of the signal:

$$\mathcal{M}_x = \frac{\sqrt{\mathcal{A}_{x'}}}{\sqrt{\mathcal{A}_x}} = \frac{\sigma_{x'}}{\sigma_x}, \quad (12)$$

where  $\sigma_{x'}$  is the standard deviation of the first derivative of the vibration signal. The complexity is the ratio of mobility and its first derivative, and it indicates how close the vibration signal to a pure sine wave resembles:

$$\mathcal{C}_x = \frac{\mathcal{M}_{x'}}{\mathcal{M}_x} = \frac{\sigma_{x''}/\sigma_{x'}}{\sigma_{x'}/\sigma_x}, \quad (13)$$

where  $\sigma_{x''}$  is the standard deviation of the second derivative of the vibration signal.

The CIs used in this study is not limited to the time domain, as indicators based on the frequency domain are also considered. Some indicators proposed in the literature, such as frequency center (FC), root mean square frequency (RMSF), and root variance frequency (RVF) are used in several works [35,92]. The first two indicators can help reveal the position change of the main frequencies of the signal, and the third indicator quantifies the convergence of the power spectrum. Their formulas are as follows:

$$\text{FC} = \frac{\sum_{i=2}^{N-1} x'_i x_i}{2\pi \sum_{i=1}^N x_i^2}, \quad (14)$$

$$\text{RMSF} = \frac{1}{2\pi} \sqrt{\frac{\sum_{i=2}^{N-1} (x'_i)^2}{\sum_{i=1}^N x_i^2}}, \quad (15)$$

$$\text{RVF} = \sqrt{\frac{\sum_{i=2}^{N-1} (x'_i)^2}{4\pi^2 \sum_{i=1}^N x_i^2} - \left( \frac{\sum_{i=2}^{N-1} x'_i x_i}{2\pi \sum_{i=1}^N x_i^2} \right)^2}. \quad (16)$$

The 16 indicators described in Equations (1)–(16) are used in this study for the characterization of the vibration signals.

### 3.2. Entropy-Based Indicators

The indicators proposed in this work are based on the concept of entropy, which was first discussed for the second law of thermodynamics [93]. After this statement, the common understanding of entropy was associated as a measure of uncertainty, disorder, or lack of knowledge of the state of a system [94]. The idea of entropy for information measurement is conceived by Shannon [95], who calculates the entropy of messages in a communication system. Using this procedure, he explained that a regular event has less information on that system than an unlikely event. First, the set of probabilities  $p$  is defined as: [95]:

$$p = \{p_i\}_{i=1, \dots, N}, p_i > 0, N \in \mathbb{N},$$

where  $p_i$  is the probability that a system will be in state  $i$  from  $N$  possible states [95]. Second, the Shannon entropy  $H$  is defined as:

$$H = -K \sum_{i=1}^N p_i \log_2(p_i), \quad (17)$$

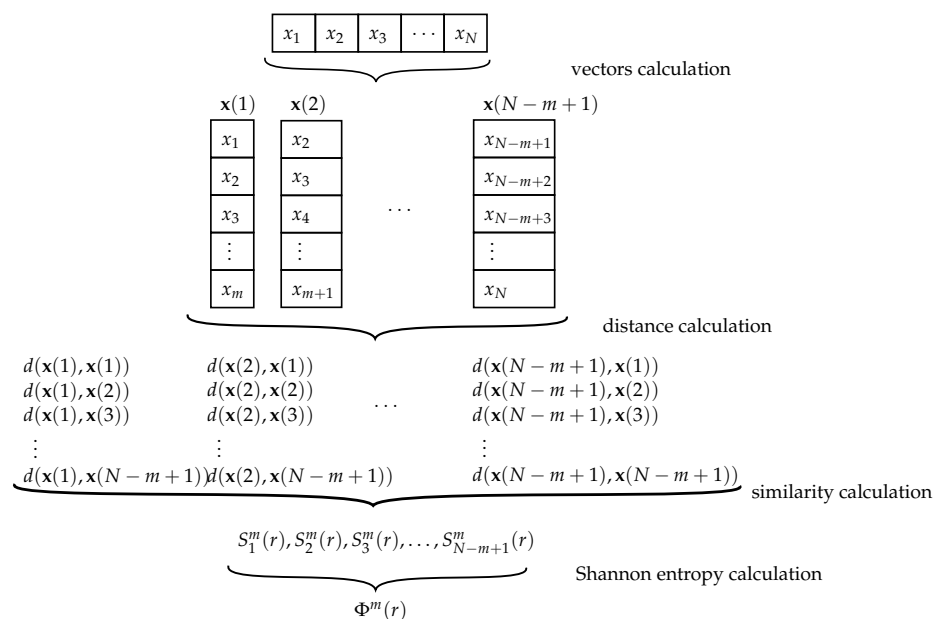
where  $K$  is a constant. Since Shannon based his work on a digital communication system, the base of the logarithm is 2. In general, the base of the logarithm depends on the application, with the natural and the decimal logarithm being the most commonly used bases [95].

The quantification of information related to an event from a system is the main basis for using entropy-based indicators (EIs) for the analysis of time signals in general and vibration signals, particularly for this work, as EIs attempts to measure the uncertainty or irregularity of the analyzed signal [96]. For this study, four EIs are used to extract information about the state of the bearing from the dataset described in Section 2: Approximate entropy, dispersion entropy, singular value decomposition entropy, and spectral entropy of the permutation entropy. To explain EIs, the general approach will be followed by a numerical example. For simplicity, the following sequence  $x$  is used in this section to provide a numerical example:

$$x = \{x_i\}_{i=1,\dots,10} = \{x_1, x_2, \dots, x_{10}\} = \{6, 6, 9, 1, 9, 8, 7, 5, 2, 4\}. \quad (18)$$

### 3.2.1. Approximate Entropy

The approximate entropy (AppEn) was proposed by Pincus [97] as a family of formulas and statistics oriented to classify complex systems. After demonstrating its capabilities for time signal analysis [98], it was used for physiological signal processing [45,46,99,100]. To date, it has been used in several fields such as climate forecasts [101,102], finance [103], image encryption authentication [89,104], and fault diagnosis [105,106]. The approximate entropy (AppEn) basically quantifies the regularity and unpredictability of the analyzed signal by comparing sliding vectors from the original signal. The steps followed for the calculation of AppEn is shown in Figure 3.



**Figure 3.** Diagram of the algorithm of approximate entropy (AppEn).



The parameters involved in the AppEn calculation are  $m$ , which is the embedding dimension, and the tolerance value  $r$ . The first step is to split the analyzed signal of length  $N$  into  $N - m + 1$  vectors  $\mathbf{x}(i)$  having  $m \in \mathbb{N}$  consecutive elements:

$$\mathbf{x}(i) = (x_i, x_{i+1}, \dots, x_{i+m-1}), i = 1, \dots, N - m + 1. \quad (19)$$

For the proposed sequence  $x$ , using  $m = 3$ , the first vector sequence is  $\mathbf{x}(1) = (6, 6, 9)$ . Subsequently, the second vector sequence is  $\mathbf{x}(2) = (6, 9, 1)$  and the third  $\mathbf{x}(3) = (9, 1, 9)$ . In total, eight  $\mathbf{x}(i)$  trios can be built. These are:

$$\begin{aligned} X &= \{\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(8)\} \\ &= \{(6, 6, 9), (6, 9, 1), (9, 1, 9), (1, 9, 8), (9, 8, 7), (8, 7, 5), (7, 5, 2), (5, 2, 4)\}. \end{aligned}$$

Secondly, a metric distance  $d$  is defined to compare two vectors  $\mathbf{x}(i)$  and  $\mathbf{x}(j)$  as the maximum absolute difference between two respective scalar values after comparing the values from all elements of the vectors [97]:

$$d(\mathbf{x}(i), \mathbf{x}(j)) = \max_{k=0, \dots, m-1} |x_{i+k} - x_{j+k}|, i, j = 1, \dots, N - m + 1. \quad (20)$$

The proposed distance is calculated for all possible pair vectors  $\mathbf{x}(i)$  and  $\mathbf{x}(j)$ . A clear consequence of this definition is that the distance from one vector to itself is zero. For the proposed example, the distance between  $\mathbf{x}(1)$  and  $\mathbf{x}(2)$  is calculated as follows:

$$\begin{aligned} d(\mathbf{x}(1), \mathbf{x}(2)) &= \max\{|x_1 - x_2|, |x_2 - x_3|, |x_3 - x_4|\} \\ &= \max\{|6 - 6|, |6 - 9|, |9 - 1|\} \\ &= 8. \end{aligned}$$

The same process is performed to compare  $\mathbf{x}(1)$  with all other vectors (itself included). The calculated values are:

$$\begin{aligned} d(\mathbf{x}(1), \mathbf{x}(1)) &= 0, & d(\mathbf{x}(1), \mathbf{x}(2)) &= 8, & d(\mathbf{x}(1), \mathbf{x}(3)) &= 5, & d(\mathbf{x}(1), \mathbf{x}(4)) &= 5, \\ d(\mathbf{x}(1), \mathbf{x}(5)) &= 3, & d(\mathbf{x}(1), \mathbf{x}(6)) &= 4, & d(\mathbf{x}(1), \mathbf{x}(7)) &= 7, & d(\mathbf{x}(1), \mathbf{x}(8)) &= 5. \end{aligned}$$

As the calculation process is performed for all pair vectors, the proposed example ends with the calculation of the distances between  $\mathbf{x}(8)$  and  $\mathbf{x}(j)$ ,  $j = 1, \dots, 8$ .

After the distances among vectors are met, the next step is the calculation of the similarity, which is unique for each vector. The correlation dimension of a vector, or similarity, is calculated as the division of the number of times that the distance between the vectors  $\mathbf{x}(i)$  and  $\mathbf{x}(j)$ ,  $j = 1, \dots, N - m + 1$  is less than or equal to the tolerance parameter  $r$ , and  $N - m + 1$ :

$$S_i^m(r) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} s_{ij}, \quad (21)$$

where

$$s_{ij} = \begin{cases} 1, & \text{if } d(\mathbf{x}(i), \mathbf{x}(j)) \leq r \\ 0, & \text{otherwise} \end{cases}.$$

From the proposed example, with the use of  $r = 5$ , the similarity values of all vectors are:

$$\begin{array}{cccc}
 S_1^3(5) = \frac{5}{8}, & S_2^3(5) = \frac{1}{4}, & S_3^3(5) = \frac{1}{4}, & S_4^3(5) = \frac{1}{8}, \\
 S_5^3(5) = \frac{3}{8}, & S_6^3(5) = \frac{5}{8}, & S_7^3(5) = \frac{1}{2}, & S_8^3(5) = \frac{1}{2}.
 \end{array}$$

The penultimate step is the Shannon entropy calculation for all  $S_i^m(r)$ , which is defined as the averaged sum of the natural logarithm of the similarities:

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{\substack{i=1 \\ S_i^m(r) > 0}}^{N-m+1} \ln(S_i^m(r)). \quad (22)$$

The AppEn is the limit of the difference between  $\Phi^m(r)$  and  $\Phi^{m+1}(r)$ , where  $\Phi^{m+1}(r)$  is the Shannon entropy from all  $S_i^{m+1}(r)$  values [97]:

$$\text{AppEn}(m, r) = \lim_{N \rightarrow +\infty} [\Phi^m(r) - \Phi^{m+1}(r)].$$

Since the analysis is performed on a time signal of finite length  $N$ , AppEn is estimated by the definition of the statistic indicator [107]:

$$\text{AppEn}(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r). \quad (23)$$

As  $m = 3$  and  $r = 5$  for the proposed example, Equation (23) is calculated by the use of  $\Phi^3(5)$  and  $\Phi^4(5)$ :

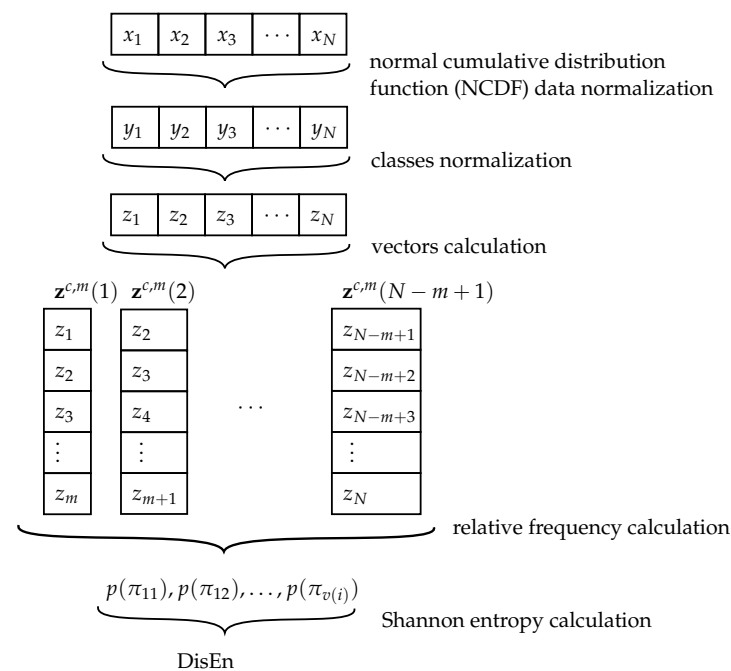
$$\text{AppEn}(3, 5, 10) = \Phi^3(5) - \Phi^4(5) \approx -1.019895 - (-1.079734) \approx 0.059839.$$

The AppEn evaluates the regularity of the signal, with the regularity defined as the tendency for the signal to have recurrent patterns throughout the signal [108]. Using this idea, the values of AppEn would be lower for regular signals and higher for complex signals. For example, a periodic vibration signal would have a low AppEn value (close to zero) because of the regularity (or periodicity) of the signal. In contrast, owing to its lower regularity, a complex vibration signal with multiple frequency components will have a higher AppEn value compared to a periodic vibration signal.

### 3.2.2. Dispersion Entropy

The next EI to be described is the dispersion entropy (DisEn), which was first proposed by Rostaghi et al. [96] as an alternative to the sample and permutation entropy [46,49,100,109]. Although this indicator is relatively new in comparison to AppEn, DisEn has gained popularity in the analysis of biomedical signals [48,110–112], economics [50], authentication processes [52], and rotary machine fault detection [113–115]. The steps followed for the calculation of DisEn is shown in Figure 4.

The calculation of DisEn starts with the parameters  $c$  and  $m$ , where  $c \in \mathbb{N}$  is the number of classes and  $m \in \mathbb{N}$  is the embedding dimension. First, the values of the analyzed signal are normalized using a normal cumulative distribution function (NCDF) to later create several embedding vectors and obtain dispersion pattern indicators. Finally, based on the Shannon entropy definition (Equation (34)), the DisEn value is obtained.



**Figure 4.** Diagram of the algorithm of dispersion entropy (DisEn).

As previously mentioned, the first step consists of the normalization of the values  $x$  into  $y = \{y_1, y_2, \dots, y_N\}$  using an NCDF to set the values between 0 and 1 [116]:

$$y_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_i} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (24)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the signal, respectively.

For the proposed example, this calculation gives the sequence  $y$  as follows:

$$y = \{y_i\}_{i=1, \dots, 10} = \{y_1, y_2, \dots, y_{10}\} = \{0.54, 0.54, 0.89, 0.03, 0.89, 0.81, 0.69, 0.39, 0.07, 0.25\}.$$

Thereafter, the values of  $y$  are classified into  $c \geq 2$  classes using the formula:

$$z_i^c = \text{round}(c \cdot y_i + 0.5). \quad (25)$$

For the numerical example proposed using the sequence  $x$ , the values of  $z_i^c$  are:

$$z^3 = \{z_i^3\} = \{2, 2, 3, 1, 3, 3, 3, 2, 1, 1\}.$$

The following step is to create the embedding vectors  $\mathbf{z}^{c,m}(i)$  with embedding dimension  $m$ , where a total of  $N - m + 1$  vectors can be calculated:

$$\mathbf{Z}^{c,m} = \{\mathbf{z}^{c,m}(1), \mathbf{z}^{c,m}(2), \dots, \mathbf{z}^{c,m}(N - m + 1)\}, \quad (26)$$

where

$$\mathbf{z}^{c,m}(i) = (z_i^c, z_{i+1}^c, \dots, z_{i+m-1}^c), \quad i = 1, \dots, N - m + 1. \quad (27)$$

Inside  $\mathbf{z}_m^c(i)$ , the order of values can follow certain patterns. These patterns are denoted as  $\pi_{v_1 v_2 \dots v_m}$ , where  $1 \leq v_i \leq c$ ,  $i = 1, \dots, m$ . The maximum number of patterns is  $c^m$ , which is the number of variations with repetition  $VR^{c,m}$ .

For the example in this work, the value of the embedding dimensions is  $m = 2$ , and the length of the sequence  $x$  is  $N = 10$ . Therefore,  $N - m + 1 = 9$  vectors can be created, which are:

$$\begin{aligned} Z^{3,2} &= \{\mathbf{z}^{3,2}(1), \mathbf{z}^{3,2}(2), \dots, \mathbf{z}^{3,2}(9)\} \\ &= \{(2, 2), (2, 3), (3, 1), (1, 3), (3, 3), (3, 3), (3, 2), (2, 1), (1, 1)\}. \end{aligned}$$

In the numerical example, the maximum number of patterns is  $3^2 = 9$ .

For the calculation of DisEn, the relative frequency of each pattern is needed. This frequency is calculated as the number of times that a pattern is present on  $Z^{c,m}$ , divided by the total number of possible dispersion patterns  $c^m$ :

$$p(\pi_{v_1 v_2 \dots v_m}) = \frac{\#\{i \in \{1, \dots, N - m + 1\} \mid \mathbf{z}^{c,m}(i) \text{ has type } \pi_{v_1 v_2 \dots v_m}\}}{N - m + 1}. \quad (28)$$

In the proposed example because one pattern is not present, eight pattern values are obtained:

$$\begin{aligned} p(\pi_{11}) &= \frac{1}{9}, & p(\pi_{13}) &= \frac{1}{9}, & p(\pi_{21}) &= \frac{1}{9}, & p(\pi_{22}) &= \frac{1}{9}, \\ p(\pi_{23}) &= \frac{1}{9}, & p(\pi_{31}) &= \frac{1}{9}, & p(\pi_{32}) &= \frac{1}{9}, & p(\pi_{33}) &= \frac{2}{9}. \end{aligned}$$

The last step is the use of the definition of the Shannon entropy with the frequencies calculated to obtain DisEn, where only the patterns  $p(\pi_{v_1 v_2 \dots v_m}) > 0$  have been considered [95]:

$$\text{DisEn}(c, m) = - \sum_{\substack{i=1 \\ p(\pi_{v(i)}) > 0}}^{c^m} p(\pi_{v(i)}) \ln(p(\pi_{v(i)})), \quad (29)$$

where  $v(i)$  is the  $i$ -th variation with the repetition of  $m$  elements from the set  $\{1, 2, \dots, c\}$ .

For the case of the numerical example, the calculation of DisEn is:

$$\text{DisEn}(3, 2) = -7 \frac{1}{9} \ln\left(\frac{1}{9}\right) + \frac{2}{9} \ln\left(\frac{2}{9}\right) \approx 2.043191.$$

This value can be normalized by dividing DisEn by  $\ln(c^m)$ . Hence, the normalized DisEn (NDisEn) is defined as:

$$\text{NDisEn} = \frac{\text{DisEn}}{\ln(c^m)}. \quad (30)$$

For the proposed example, the value of NDisEn is:

$$\text{NDisEn} = \frac{-7 \frac{1}{9} \ln\left(\frac{1}{9}\right) + \frac{2}{9} \ln\left(\frac{2}{9}\right)}{\ln(9)} \approx 0.929896.$$

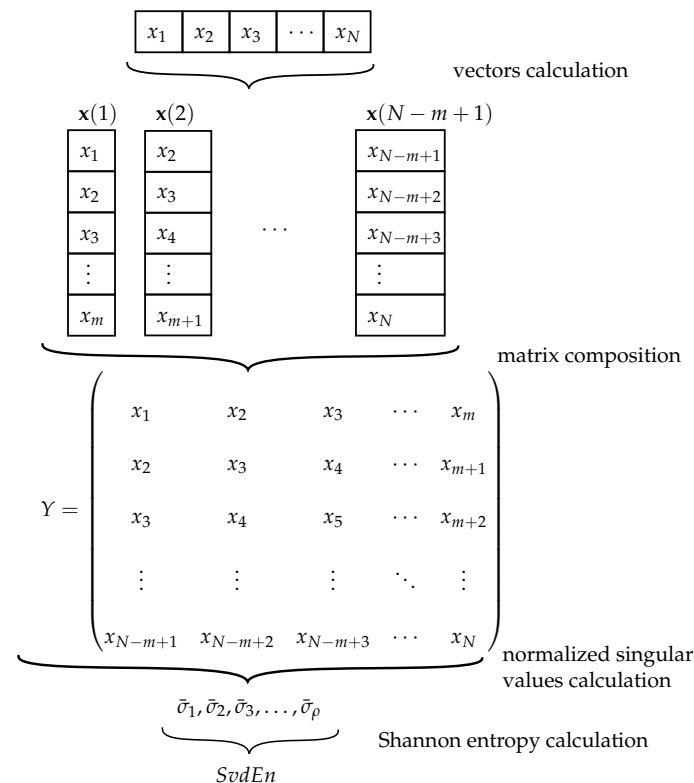
The value of DisEn is higher when all the possible dispersion patterns have equal frequency values on a signal thus, an irregular signal is present. However, the presence of fewer patterns will result from lower values of DisEn thus, a completely regular signal is obtained [96]. DisEn is relatively insensitive to noise because larger or smaller changes in the amplitude of the signal do not vary significantly the class label of the values [117].

### 3.2.3. Single Value Decomposition Entropy

The third EI used in this study is the singular value decomposition entropy (SvdEn). This entropy calculation, which was proposed by Roberts et al. [118], was an attempt

to improve the data analysis from an electroencephalogram (EEG), especially for cases where movements are thought by patients but are not performed in reality. Compared with AppEn and DisEn, SvdEn are used mostly in EEG analysis [119,120], and medical imaging [121,122]. The steps followed for the computation of SvdEn is shown in Figure 5.

The idea behind SvdEn is to indicate the number of eigenvectors needed for an adequate explanation of the analyzed signal, giving an idea of its dimensionality and complexity. Therefore, the only parameter involved in the SvdEn calculation is the embedding dimension  $m$ . To understand the procedure, the numerical calculation of SvdEn is presented with the sequence  $x$  defined in Equation (18).



**Figure 5.** Diagram of the singular value decomposition entropy (SvdEn) computation.

First, the creation of embedding vectors  $\mathbf{x}(i)$  of consecutive  $m \in \mathbb{N}$  values is needed. The procedure is similar to the vectors needed in AppEn and DisEn:

$$X = \{\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(N-m+1)\}. \quad (31)$$

For the case of the numerical example of the sequence  $x$  and considering the parameter value  $m = 5$ , the values of  $\mathbf{x}(i)$  are as follows:

$$X = \{(6, 6, 9, 1, 9), (6, 9, 1, 9, 8), (9, 1, 9, 8, 7), (1, 9, 8, 7, 5), (9, 8, 7, 5, 2), (8, 7, 5, 2, 4)\}.$$

The subsequent action is the composition of the matrix  $Y$ , which is built based on the vectors obtained from  $\mathbf{x}(i)$  in Equation (31):

$$Y = \begin{pmatrix} \mathbf{x}(1) \\ \mathbf{x}(2) \\ \vdots \\ \mathbf{x}(N-m+1) \end{pmatrix} \in \mathcal{M}_{(N-m+1) \times m}(\mathbb{R}). \quad (32)$$

For the example considered in this paper, the matrix  $Y$  is stated as:

$$Y = \begin{pmatrix} 6 & 6 & 9 & 1 & 9 \\ 6 & 9 & 1 & 9 & 8 \\ 9 & 1 & 9 & 8 & 7 \\ 1 & 9 & 8 & 7 & 5 \\ 9 & 8 & 7 & 5 & 2 \\ 8 & 7 & 5 & 2 & 4 \end{pmatrix}.$$

Once the matrix  $Y$  is built, the next step is the calculation of the  $\rho$  singular values  $\sigma = \{\sigma_1, \dots, \sigma_\rho\}$  of matrix  $Y$ , assuming that  $\rho < N - m + 1$  and  $\text{rank}(Y) = \rho$ . Once the values are obtained, the normalized singular values  $\bar{\sigma} = \{\bar{\sigma}_1, \dots, \bar{\sigma}_\rho\}$  are:

$$\bar{\sigma}_i = \frac{\sigma_i}{\sum_{j=1}^{\rho} \sigma_j}, \quad i = 1, \dots, \rho. \quad (33)$$

In the numerical example, the  $\bar{\sigma}_i$  values from  $Y$  are as follows:

$$\bar{\sigma} = \{0.537047, 0.145508, 0.116931, 0.109984, 0.0905299\}.$$

The final step, as in the rest of the previous EIs, is the use of the Shannon entropy definition on the normalized  $\bar{\sigma}$  values, and with the use of base 2 for the logarithm [118]:

$$\text{SvdEn}(\rho) = - \sum_{\substack{i=1 \\ \bar{\sigma}_i > 0}}^{\rho} \bar{\sigma}_i \cdot \log_2(\bar{\sigma}_i). \quad (34)$$

This final step can be translated to the numerical example for the proposed sequence  $x$ , with the following result:

$$\text{SvdEn}(5) = - \sum_{i=1}^5 \bar{\sigma}_i \cdot \log_2(\bar{\sigma}_i) \approx 1.912336.$$

In addition, as in the case of DisEn, the SvdEn values can be normalized, and the normalized SvdEn (NSvdEn) is defined. The step is similar to the mentioned EI, but the base of the logarithm  $e$  is changed to 2, and  $c^\rho$  is replaced by  $\rho$ , leading to:

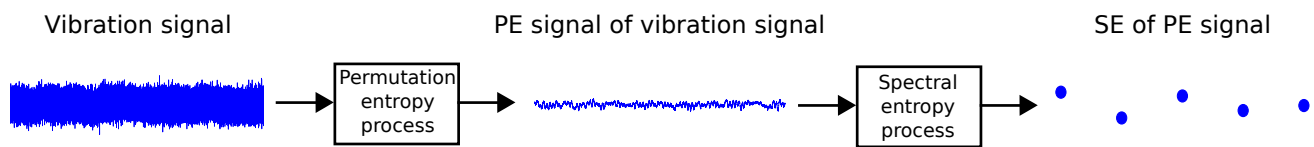
$$\text{NSvdEn} = \frac{\text{SvdEn}(\rho)}{\log_2(\rho)}. \quad (35)$$

For the proposed numerical example, the value of NSvdEn is:

$$\text{NSvdEn} = \frac{\text{SvdEn}(5)}{\log_2(5)} \approx 0.823598.$$

### 3.2.4. Spectral Entropy of the Permutation Entropy Signal

The last EI used in this study is spectral entropy of the permutation entropy signal (SepEn). This indicator is defined by Sandoval et al. [74], and it is specifically designed (but not limited to) for low-speed bearings. TheSepEn uses EIs for its process: permutation entropy (PerEn) and pectral 215 entropy (SpeEn). The steps followed for the computation of SepEn is shown in Figure 6 [74].



**Figure 6.** Diagram of the signal analysis procedure for spectral entropy of the permutation entropy signal SepEn [74].

The first step in calculating SepEn involves estimating PerEn, which starts grouping  $m$  consecutive values of the analyzed signal into  $N - m + 1$  groups [109], where  $N$  is the length of  $x$ :

$$x = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}, i = 1, \dots, N - m + 1. \quad (36)$$

For the case of the defined sequence  $x$  and considering  $m = 2$ , the groups are as follows:

$$\begin{aligned} X &= \{\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(8)\} \\ &= \{(6, 6), (6, 9), (9, 1), (1, 9), (9, 8), (8, 7), (7, 5), (5, 2), (2, 4)\}. \end{aligned}$$

Then, the permutation patterns present on the grouped values are counted. As equal consecutive values are not considered in the calculation [109], in the proposed example, only two potential permutation patterns  $\pi_j$  have been considered:  $x_i < x_{i+1}$  and  $x_i > x_{i+1}$ , which correspond to  $\pi_1$  and  $\pi_2$ , respectively. Therefore, the relative frequency for each permutation pattern was calculated as:

$$p(\pi_j) = \frac{\#\{i \in \{1, \dots, N - m + 1\} \mid \mathbf{x}(i) \text{ has type } \pi_j\}}{N - m + 1}, j = 1, \dots, m!, \quad (37)$$

obtaining  $p(\pi_1) = 3/9$  and  $p(\pi_2) = 5/9$  for the numerical example with  $x$ . Finally, the Shannon entropy is applied to the relative frequencies  $p(\pi_j)$ ,  $j = 1, \dots, m!$  to calculate the PerEn value:

$$\text{PerEn}(m) = - \sum_{\substack{j=1 \\ p(\pi_j) > 0}}^{m!} p(\pi_j) \log_2 p(\pi_j). \quad (38)$$

The value for the numerical example is:

$$\text{PerEn}(3) = -\frac{3}{9} \log_2 \left( \frac{3}{9} \right) + \frac{5}{9} \log_2 \left( \frac{5}{9} \right) \approx 0.99943.$$

As shown in Equation (38), PerEn provides a scalar value. SepEn requires the generation of a signal based on PerEn, which is calculated using a certain data window. If the data window moves forward one data set at a time on the analyzed signal, a set of values from the window slices on the entire signal can be stated, which is called the permutation signal.

The next step in the calculation of spectral entropy of the SepEn involves SpeEn, which uses the Shannon entropy of the normalized power spectral density (PSD) of a signal. The normalized PSD of a signal is defined as follows [123]:

$$\int_{f=0}^{f_s/2} P(f) df = 1, \quad (39)$$

where  $f_s$  is the sampling frequency of the analyzed signal, and  $P(f)$  is the power value of the harmonic  $f$ . The power of the analyzed signal is normalized to obtain the sum of the power of each harmonic as 1. Then, the SpeEn for an analyzed signal is:

$$\text{SpeEn} = - \sum_{f=0}^{f_s/2} P(f) \log_2 P(f), \quad (40)$$

where the base of the logarithm is 2. As the sequence  $x$  is not suitable for explaining the SpeEn, the case of two signals  $g(t)$  and  $h(t)$  is proposed. The first signal  $g(t)$  is a pure sine wave with frequency  $f_a$ , and  $h(t)$  is the sum of two sine waves with frequency  $f_b$  and  $f_c$ . While each sine wave from  $h(t)$  has a unique frequency, both the power contributions to  $h(t)$  are the same. Hence, the signals are given as follows:

$$\begin{aligned} g(t) &= \sin(f_a t), \\ h(t) &= \sin(f_b t) + \sin(f_c t). \end{aligned}$$

In the case of the signal  $g(t)$ , the power of the signal is concentrated only in one frequency  $f_a$  hence,  $P(f_a) = 1$ . Therefore, SpeEn is given as:

$$\text{SpeEn} = -1 \cdot \log_2(1) = 0.$$

The case of  $h(t)$  differs from  $g(t)$  because the power of  $h(t)$  is split between the frequencies  $f_b$  and  $f_c$ . For this reason,  $P(f_b) = P(f_c) = 1/2$ . Consequently, the SpeEn value is:

$$\text{SpeEn} = -2 \cdot \left(\frac{1}{2}\right) \log_2\left(\frac{1}{2}\right) = 1.$$

As the PSD of the signal  $g(t)$  is concentrated only in one frequency,  $g(t)$  is predictable. A different situation is stated with signal  $h(t)$ , where a greater irregularity in time is present compared with  $g(t)$ . Therefore, the SpeEn of  $h(t)$  is higher than  $g(t)$ .

The calculation process of SepEn is summarized in the following steps:

1. The permutation entropy (PerEn) is calculated for the analyzed signal [109]. Each value is obtained for a specific time window thus, a PerEn signal is achieved by sliding the time window along the analyzed signal;
2. Using the signal calculated in the previous step, for a time window equivalent to one rotation of the bearing, the SpeEn value is acquired [124].

The validity of this EI is proven by Sandoval et al. [74], who showed that the values of spectral entropy of the SepEn for healthy bearings are higher than the values obtained from vibration signals of bearings with damage.

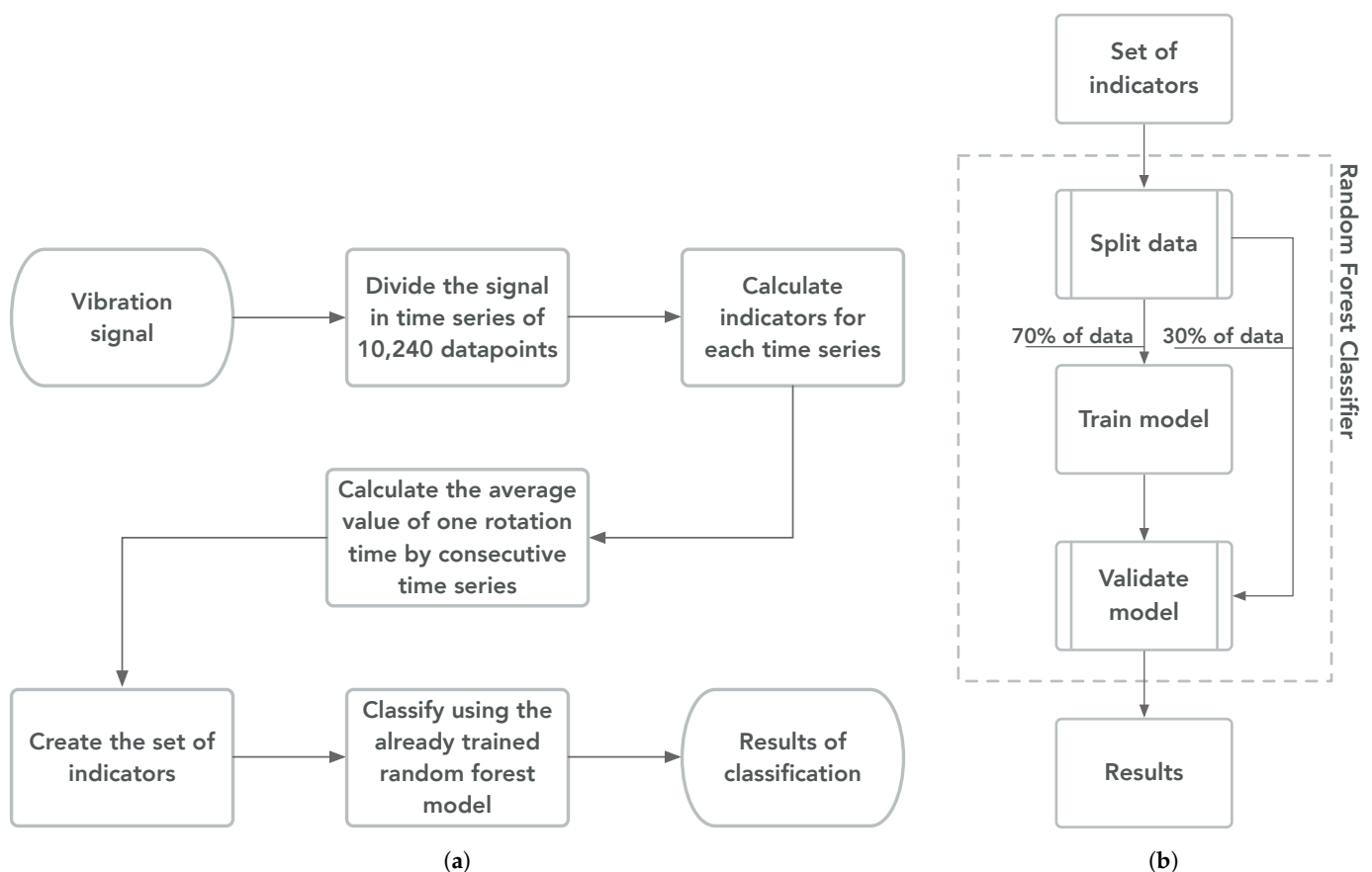
#### 4. Methodology

In this section, the proposed diagnosis strategy is first discussed. Secondly, the parameter values used to calculate the different indicators are studied. Since CIs do not depend on parameters, entropy-based indicators (EIs) is the first to be described. Finally, Section 4.3 provides the classification algorithm used in this work, which is briefly reviewed for manuscript completeness.

##### 4.1. Proposed Diagnosis Method

This paper presents a scheme in which several low-speed bearings with induced damage are classified as a feasibility study for the classification and diagnosis of low-speed bearings based on entropy-based indicators. The classification is based on the dataset presented in Section 2. First, information is extracted from the vibration signals using four EIs, which are AppEn, DisEn, SvdEn, and SpeEn. Afterward, the extracted information is used by RF the models. The steps of the proposed classification method are summarized in Figure 7.





**Figure 7.** Diagram of proposed fault diagnosis method for vibration signals of low-speed bearings. (a) Values for general method. (b) Diagram to detail the training and validation of the random forest model.

First, the vibration signals are split into a time series of 10,240 data points (one second). For each time series, the indicators were calculated to extract information from the data. Then, the analysis is performed for each entire rotation, where each rotation has a set of indicators. The value per indicator is calculated as the average of the indicators from the number of consecutive time series from the vibration signals required to represent the time of one rotation of the shaft washer, which is dependent on the rotational speed. For example, the AppEn of one rotation from a signal at 10 rpm is the average of 6 AppEn values. Finally, the indicators from each rotation are used as features to train and validate the RF model.

The presented scheme allows the possibility of switching the indicator group, hence performing a comparative study between EIs and CIs. A third group, called classic and entropy indicators (CEIs), is included to analyze the improvement of the classification by joining EIs and CIs. As the data set includes three speeds, and three indicators groups are used, nine RF models are trained.

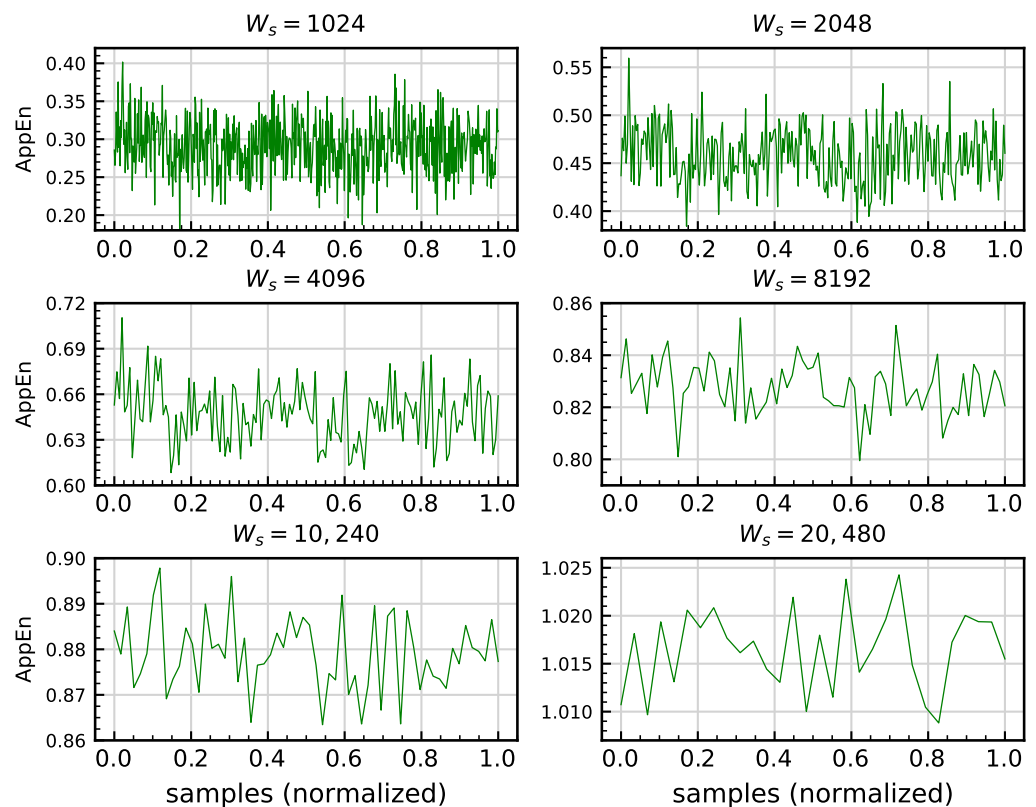
#### 4.2. Entropy-Based Indicators

The EIs is defined using multiple parameters that are selected according to the intended purpose and results. In this work, the parameter values were selected after an exhaustive study. For explanation purposes, this section uses five signals of one minute long at the same rotational speed. Each signal represents one scenario, as described in Section 2. Although the numerical results for each EI are calculated with different signal groups, the chosen group captures the general results of the explained EI.

#### 4.2.1. Approximate Entropy

As explained in Section 3.2.1, the calculation of AppEn involves two parameters:  $m$  and  $r$ , as shown in Equation (23). After consulting the literature,  $m$  values of between 3 and 10 are regularly used [42,125] thus, this range is used in this work as the initial range where the  $m$  value is sought. To analyze this parameter and its influence on the results, five signals per minute are used to compare the behavior of AppEn in different scenarios. The analysis starts by dividing each signal into a time series of  $W_s = 10,240$  samples, which corresponds to one second. This value is selected after comparing several values of  $W_s$ . For example, Figure 8 shows a comparison of the AppEn calculation for one HS signal using different  $W_s$ .

The results reveal that for values lower than  $W_s = 10,240$ , a larger amount of noise and the loss of any representativeness of the original signal can be observed in the calculated signal. The same procedure is accomplished for all EIs with the same results. Therefore,  $W_s = 10,240$ , is used in all EIs calculations.

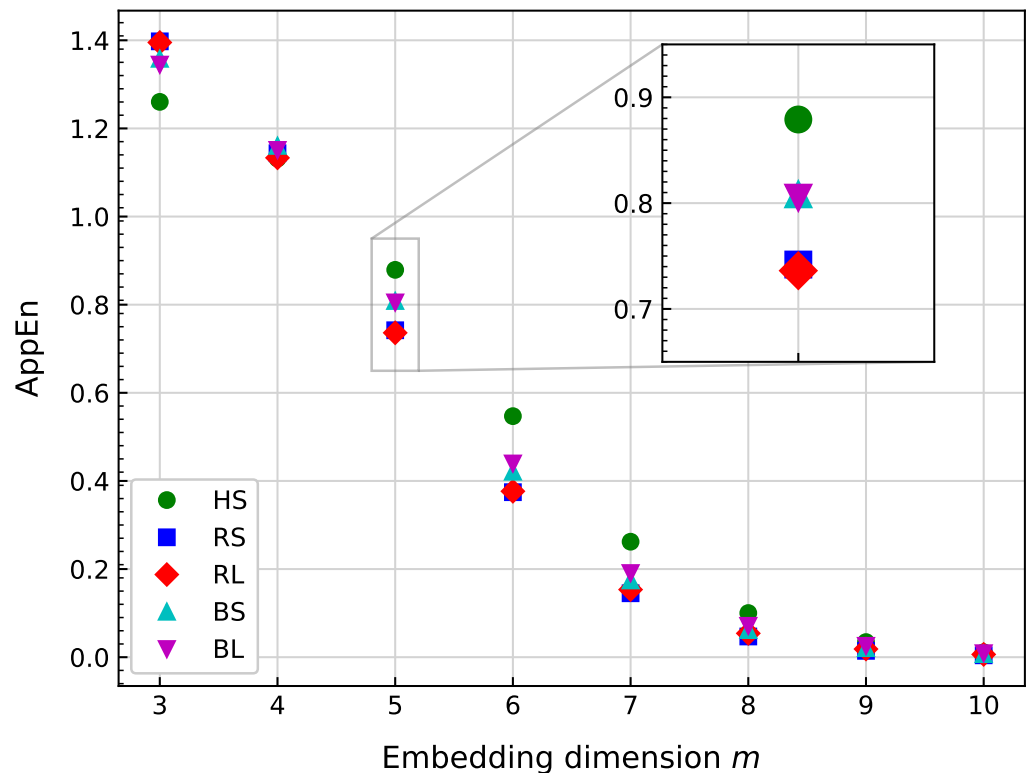


**Figure 8.** Approximate entropy (AppEn) calculation for bearing signal corresponding to healthy scenario HS with  $m = 5$ ,  $r = 0.2 \times \sigma$ , and varying  $W_s$ .

For each time series, AppEn is calculated. Finally, the average of the AppEn values for the entire signal is obtained and compared with the values from other signals. Figure 9 shows several AppEn values obtained by varying  $m$  from 3 to 10 with a fixed  $r = 0.2 \times \sigma$  (where  $\sigma$  is the standard deviation of the signal).

From Figure 9, there is no clear difference among scenarios for  $m = 4, 8, 9$ , and 10. For  $m = 3, 5, 6$ , and 7, HS can be easily discriminated from other scenarios. In addition, approximate entropy (AppEn) can differentiate scenarios by damage location for  $m = 3, 5, 6$ . In the case of the damage size, Figure 9 shows that AppEn is unable to make a difference in the results for RS and RL, as well as for BS and BL. As the damage location is the only attribute where AppEn can draw a distinction, and to maximize this difference, the value of  $m = 5$  is selected.

With respect to the  $r$  parameter, this value helps to filter the information obtained from the signal [108]. The problem with this parameter from a mathematical point of view has been detailed previously [98], and it is later calculated as a dependent value of  $\sigma$  from the analyzed signal [42,106,125]. Empirically, the nature of the vibration signals of the dataset means that no value of  $r < 0.5 \times \sigma$  affects the calculation, and  $r = 0.2 \times \sigma$  is one of the values selected normally in the literature. Therefore, in this work,  $r = 0.2 \times \sigma$  was used to obtain the results.

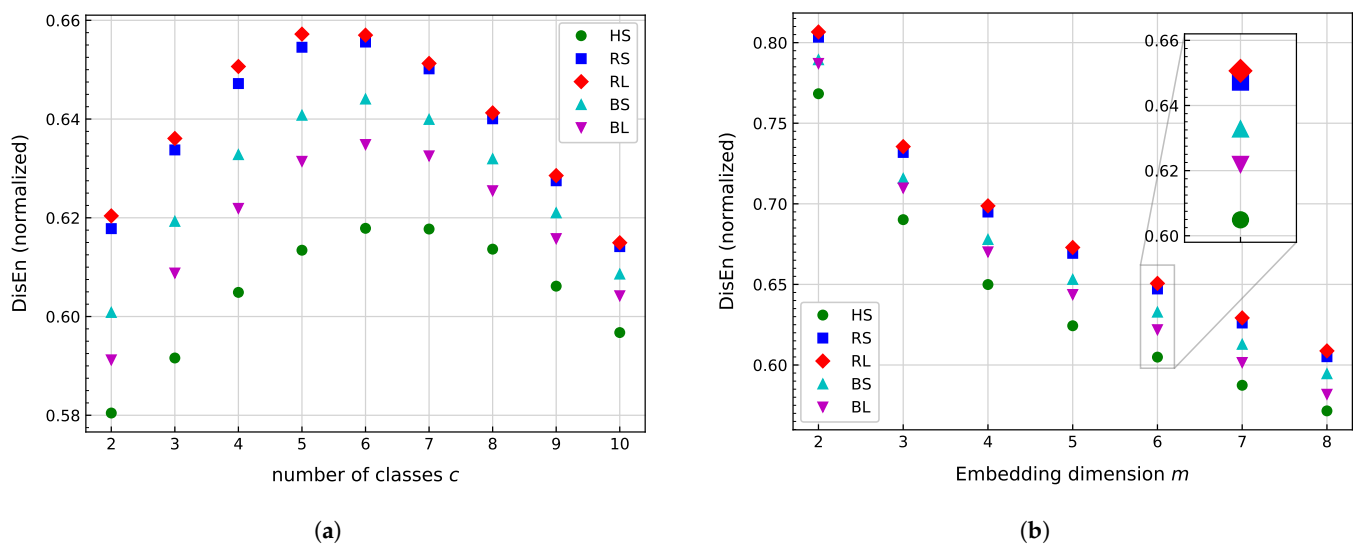


**Figure 9.** Comparison of the approximate entropy (AppEn) mean value for several bearing signals at the same rotational speed. The calculation uses  $r = 0.2 \times \sigma$ ,  $W_s = 10,240$ , and varying  $m$ .

#### 4.2.2. Dispersion Entropy

As stated in 3.2.2, the DisEn indicator depends on the parameters  $c$  and  $m$ ; see Equation (29). From the literature [126,127], the value of  $c$  is usually between 2 and 10, and  $m$  is no further than 5. Therefore, these ranges are studied for the specific application of low-speed bearing vibration signals. Consequently, the parameter values are calculated using the same methodology that was employed for the approximate entropy (AppEn): Five signals per minute are split on a time series of  $W_s = 10240$  data points. The indicator is calculated for each time series, and later, the average is obtained to contrast all signals. Figure 10 shows the results obtained with a fixed  $m = 6$  varying  $c$  and fixed  $c = 4$  varying  $m$ .

The results show that for both parameters, HS can be distinguished from other scenarios. Figure 10a illustrates that for all  $c$  values, HS has the smallest value from all signals. Based on a damage localization perspective, DisEn can be used to differentiate the cases from the ball or raceway damage scenario, but this difference decreases for  $c > 5$ . In addition, from Figure 10a, the damage size can be clearly distinguished for the ball damage scenario, yet it is somewhat less for the raceway damage scenario. To maximize the difference among scenarios,  $c = 2, 3, 4,$  and  $5$  are good candidates.



**Figure 10.** Comparison of dispersion entropy (DisEn) mean value for several bearing signals at same rotational speed, with fixed  $W_s = 10,240$ . (a) Values for fixed  $m = 6$ , varying  $c$ . (b) Values for fixed  $c = 4$ , varying  $m$ .

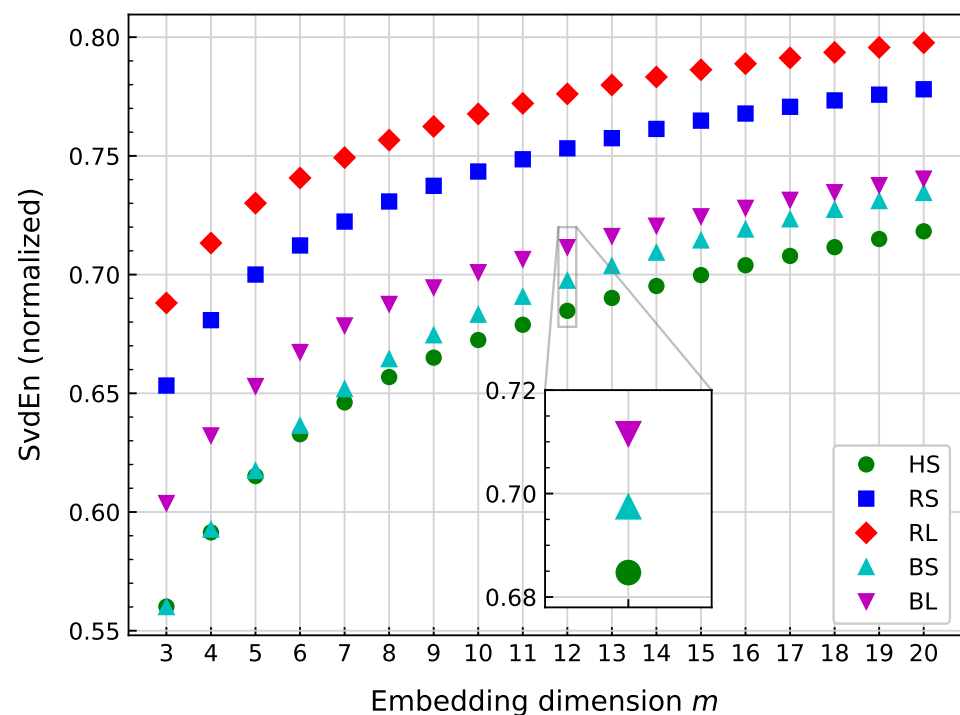
As DisEn cannot discriminate RS and RL as well as BS and BL, the major difference in value between BS and BL is noted for  $c = 4$ . In addition, the difference in value between HS and all other scenarios is higher for  $c = 4$  compared with  $c = 2, 3$ , and  $5$ . Therefore, the value selected for  $c$  was  $4$ .

Subsequently, Figure 10b shows the results of DisEn for variable  $m$ , with fixed  $c = 4$ . Similar to Figure 10a, HS has the lowest values for all cases of  $m$ . As damage scenarios can be distinguished for all values of DisEn, RS and RL are not easily distinguished. A progressive difference between BS and BL is stated as  $m$  increases. Since  $c^m$  should be less than  $W_s$ , [126], the selected value is  $m = 6$ .

#### 4.2.3. Single Value Decomposition Entropy

As stated in Section 3.2.3, SvdEn depends on one parameter,  $m$ , as shown in Equation (34). In the case of medical imaging and EEG analysis,  $m$  is usually selected between the range 5–10 [47,128]. As vibration signals are quite different from EEG signals, further analysis is performed. Figure 11 summarizes the influence of the parameter on the obtained SvdEn for all the rotational speeds, where five signals of a one minute length are analyzed, with  $W_s = 10,240$ .

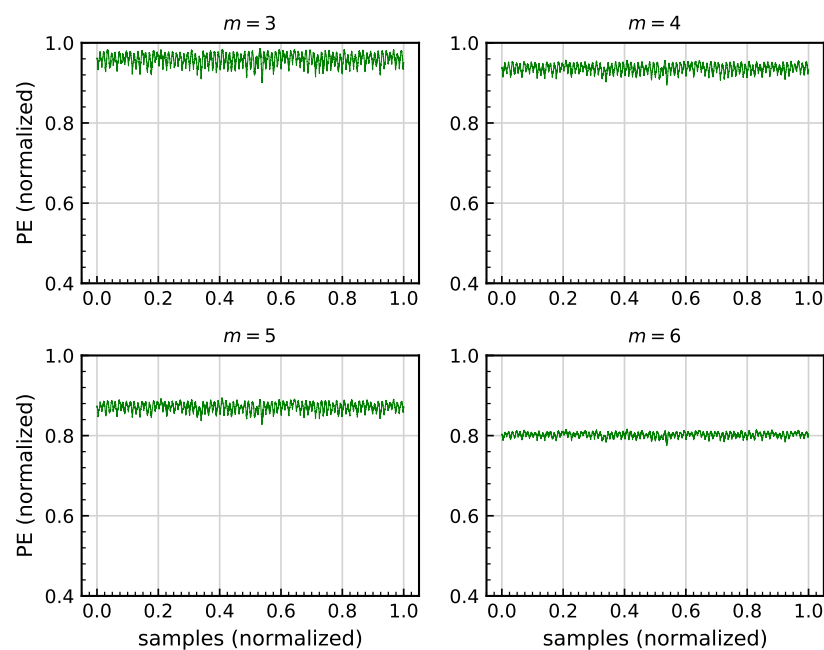
From Figure 11, the data show how SvdEn cannot establish a distinction among scenarios for values  $m < 7$  because the HS and BS values overlap. For  $m \geq 7$ , the lower values for each embedding dimension  $m$  correspond to HS. From the figure, SvdEn can supply a difference between HS and scenarios RS and RL for all values of  $m$ . In the case of BS and BL, a visible separation from HS is observed for  $m > 7$ . Another situation is observed for BS values, which are closer to HS for lower values of  $m$ . As  $m$  increases, the BS values are closer to BL than BS. For  $m = 12$ , the value difference of BS to HS and BL is a maximum compared to all other values of  $m$ . Therefore, this value is selected.



**Figure 11.** Comparison of the single value decomposition entropy (SvdEn) mean value for several bearing signals at the same rotational speed, with fixed  $W_s = 10,240$ .

#### 4.2.4. Spectral Entropy of Permutation Entropy

As stated in Section 3.2.4, spectral entropy of the permutation entropy signal (SepEn) depend on one parameter,  $m$ , as shown in Equation (36). According to previous work [74], suitable values for  $m$  can range from 3 to 6, with similar results. Since the use of a higher value only implies more computational effort, the value  $m = 3$  is used. This can be viewed in Figure 12, where 30 s of the same signal RL is used for comparison purposes.



**Figure 12.** Comparison of the permutation entropy PerEn signal calculated for 30 s of a large raceway shaft washer damage scenario RL signal, with fixed  $W_s = 10,240$ .

In contrast with the previous EIs, the values from SpeEn correspond to each rotation made by the shaft washer. To obtain the value of each rotation, the average of each  $W_s = 10,240$  is calculated [74]. This limits the use of all previous indicators to have only one value per rotation, as SpeEn does.

#### 4.3. Random Forest

The RF algorithm is used for damage classification. The first concept of RF was introduced by Ho [129] and later by Breiman [130], who gave controlled variance to the algorithm proposed by Ho. RF is an ensemble learning model that improves the capabilities of the decision tree (DT) classifier [131] by bagging ensemble learning [132]. In this respect, RF uses several DTs to decide the classification by means of decisions according to the majority of DTs. This explains the number of parameters to be determined by the RF algorithm, which starts with the parameters for the DT algorithm [131], in addition to the number of trees.

In summary, the following steps describe the operation of the RF classifier:

1. Set the parameters for RF, which include the intrinsic parameters for DT, and the number of trees to employ;
2. After the number of DT is established, a stratified sampling of the data set is set as train and test data. The proportion between both groups of data is fixed for all DT, but the elements of each data group are selected randomly [129];
3. Each DT classifies the data using a random feature from the train data, to arbitrarily classify and increase the difference between DTs. This step is fundamental because it improves the generalization error;
4. The results of all DT enable the classification of the test samples to be determined using the vote of each DT. The decision is done by majority voting.

As mentioned, the RF algorithm is an ensemble machine learning technique that has various hyper-parameters inherited from the decision tree (DT) algorithm. The hyper-parameters are those parameters that are intrinsic to the technique, whose values cannot be estimated from the data [133]. To select the value of a hyper-parameter, one recurrent process is to determine the value using experimental results [134].

For RF, the hyper-parameters inherited from DT are:

- The function criteria to measure the quality of a split, which is set as the Gini impurity;
- The maximum depth of the tree, which is set to three for this work;
- The maximum number of leafs, which is unlimited;
- The minimum number of samples required to split an internal node, which is two;
- The minimum number of samples required to be at a leaf node, which is one.

Another hyper-parameter from RF is the number of trees used for the model. Since no further changes in the results above 100 trees were observed, this value was used. Once the hyper-parameters of the RF algorithm are set, the data are split by the hold-out method for validation. A RF model is trained on 70% of the total data, and the rest of the data are used to calculate the classification accuracy score [135]. This process is repeated 1000 times for the validation of the trained models based on the indicator groups, where train and test groups are created with randomly selected data. Finally, the classification accuracy scores are averaged as representative results.

## 5. Results and Discussion

This section shows the accuracy scores obtained from the RF classifier trained models based on the dataset presented in Section 2 by using the indicators mentioned in Section 3. As the calculation is based on the number of rotations, Table 1 shows the number of rotations available for each scenario and rotational speed.

**Table 1.** Number of rotations for healthy scenario (HS), small raceway shaft washer damage scenario (RS), large raceway shaft washer damage scenario (RL), small ball bearing damage scenario (BS), and large ball bearing damage scenario (BL) at 5, 8, and 10 revolutions per minute (rpm).

	HS	RS	RL	BS	BL
10 rpm	191	195	195	40	80
8 rpm	156	156	156	156	156
5 rpm	97	97	97	97	97

As previously mentioned, three indicator groups are used in this study: CIs, entropy-based indicators (EIs), and CEIs. As the RF classifier chooses the training and testing data randomly, this process is repeated 1000 times to obtain representative results. The classification accuracy scores for the trained models are summarized in Table 2, where two columns can be found for each indicator group: The first represents the averaged accuracy scores ( $\bar{x}_{as}$ ), and the second is the standard deviation of the accuracy scores ( $\sigma_{as}$ ) of the 1000 combinations used, which can be used to enhance the description of the obtained results. Note that the first group of indicators corresponds to CIs, followed by entropy-based indicators (EIs), and finally CEIs.

**Table 2.** Average  $\bar{x}_{as}$  and standard deviation  $\sigma_{as}$  of the accuracy classification score from 1000 trained models of random forest classifier by the use of classic indicators (CIs), entropy indicators (EIs), and classic and entropy indicator (CEIs) for signals at 5, 8, and 10 rpm.

	CIs		EIs		CEIs	
	$\bar{x}_{as}$	$\sigma_{as}$	$\bar{x}_{as}$	$\sigma_{as}$	$\bar{x}_{as}$	$\sigma_{as}$
10 rpm	73%	0.027285	90%	0.017744	90%	0.018584
8 rpm	78%	0.025358	94%	0.013072	97%	0.010761
5 rpm	77%	0.033270	99%	0.003561	98%	0.014807

In general terms, the classification accuracy scores of Table 2 are above 70%. Considering groups by indicator type (columns), the CIs group has the lowest values of accuracy for the classification, from 73% to 78%. The results improved by the next group, entropy-based indicators (EIs), where all scores are above 90%. The information given by the entropy-based indicators (EIs) group allows the RF classifier to have scores up to 99%, which is the classification of data at 5 rpm. The CEIs group has an overall higher accuracy score, which is a minimum of 90% and a maximum of 98%. Although the entropy-based indicators (EIs) group has the same lower accuracy percentage value as the CEIs group, and its best value is the highest from all accuracy score values calculated (99%), at 8 rpm, it is several percentage points lower than its similarity to the CEIs group. The  $\sigma_{as}$  also reveals the variability of the results of the same groups. From this data, the lowest  $\sigma_{as}$  value is in the group of CEI models at 8 rpm, and the highest in the CI group at 5 rpm. In general, the  $\sigma_{as}$  values from the CIs trained model group are higher than for the entropy-based indicators (EIs) and CEIs groups, with differences between 30% to 90%, in comparison to  $\sigma_{as}$  values from CIs at the same rpm.

A closer view of the results from Table 2 reveals that for the same indicator group, a better classification score is obtained for lower rotational speeds. For instance, the CIs group improves by 4 percentage points from 10 to 5 rpm. The same trend occurs in the EI and CEI groups by 9 and 8 percentage points, respectively. Table 2 provides a noteworthy difference among accuracy scores from the trained models of the CIs group, and the models from the EI and CEI groups. For example, at 10 rpm, the result from CIs is 17 percentage points lower than that from the entropy-based indicators (EIs) group, with a maximal of 22 percentage points with a difference at 5 rpm. Considering groups by rows (models for

signals at same rpm), the accuracy scores at 10 rpm are the lowest for the three row group. The results for the 8 rpm group increase in value compared to the described group. In the case of the 5 rpm group, the classification score based on CIs is considerably lower than entropy-based indicators (EIs) and CIEs, with at least 20 percentage points below the corresponding values for the entropy-based indicators (EIs) and CIEs groups.

Table 2 gives a general view of the difference among classification according to several rotational speeds and types of used indicators. For a deeper analysis on each classification performance, Table 3 shows the classification accuracy score result for one particular trained model set, which is chosen randomly. To validate the trained model, a cross-validation approach of 5-fold was utilized. In addition, Table 4 shows the training time for all trained models.

The results from Table 3 show that lower accuracy scores are related to the CIs column. The classification score is improved while the rotational speed decreases for all indicator types. It is worth mentioning the particular case of perfect classification at 5 rpm, by using entropy-based indicators (EIs). This result is followed by classification at 8 and 5 rpm, with the use of CIEs. Table 4 shows the training time for each model in Table 3. In general, the CIs group consumes the highest time from all groups at the same rotational speed, followed by CEIs and EIs. A detailed explanation of the behavior from the classifiers according to the type of indicator is presented in Figure 13, where the confusion matrices for all trained models from Table 3 are provided. It is important to note that the matrices are ordered in concordance with the values from Table 3, that is, columns represent indicator groups, whereas rows represent the different operating speeds of the bearing.

**Table 3.** Accuracy score of random forest classifier by the use of classic indicators (CIs), entropy indicators (EIs), and classic and entropy indicator (CEIs) for signals at 5, 8, and 10 rpm.

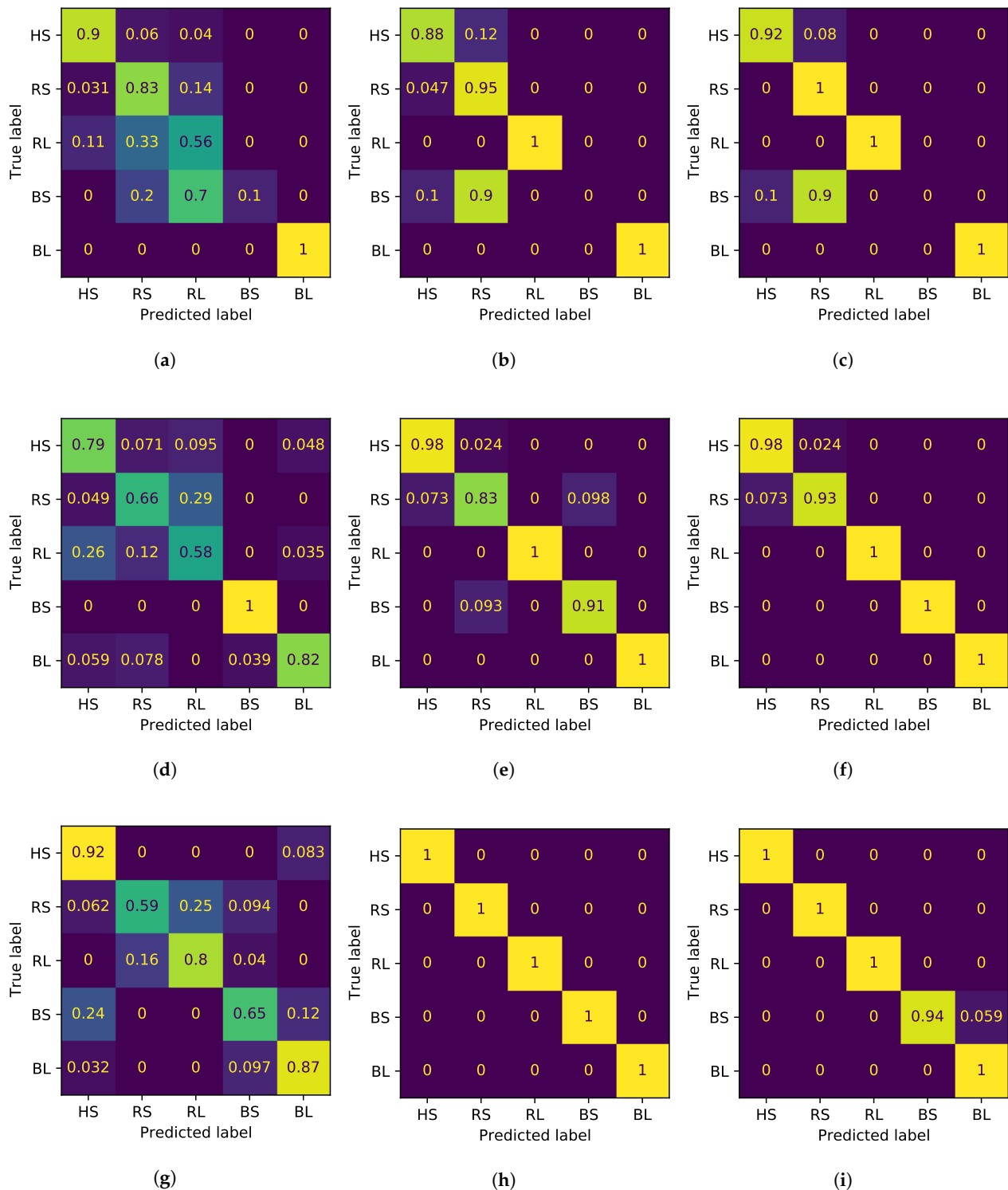
	CIs	EIs	CEIs
10 rpm	74%	90%	93%
8 rpm	76%	94%	98%
5 rpm	75%	100%	98%

**Table 4.** Training time (in seconds) of random forest classifier by the use of classic indicators (CIs), entropy indicators (EIs), and classic and entropy indicator (CEIs) for signals at 5, 8, and 10 rpm.

	CIs	EIs	CEIs
10 rpm	1.7091	1.5192	1.6991
8 rpm	1.7624	1.5625	1.7391
5 rpm	1.5992	1.4790	1.5792

The confusion matrices for the classification at 10 rpm are shown in Figure 13a for CIs, Figure 13b for entropy-based indicators (EIs), and Figure 13c for CIEs. The confusion matrix in Figure 13a shows that the classification rates for the RL and BS scenarios are the lowest for all classification scores at 10 rpm. In particular, the RL is recognized as more than half of the data, while 70% of the BS data is classified as RL. The classification from Figure 13b has the same problem as the BS, where 90% of the data is confused as RS and the rest as HS. The HS and RS scenarios are recognized above 88% of the data, but some data are misclassified. Figure 13c shows that the same problem of classification of the BS data using the entropy-based indicators (EIs) occurs using the CIEs, where 90% of the data is classified as RS and 10% as HS. In contrast, RS is always well classified and the HS rate exceeds 90% of the data.





**Figure 13.** Confusion matrices for the diagnosis of low-speed bearings using random forest classifier, varying the indicator type and the rotational speed of the data. The results were normalized over the true labels. The first row shows the results at 10 rpm, the second row at 8 rpm, and the last row at 5 rpm. The first column shows the results for classic indicators (CIs), second column for entropy indicators (EIs), and the last column shows the classic and entropy indicators (CEIs). (a) Confusion matrix for CIs and data at 10 rpm. (b) Confusion matrix for EIs and data at 10 rpm. (c) Confusion matrix for CEIs with data at 10 rpm. (d) Confusion matrix for CIs and data at 8 rpm. (e) Confusion matrix for EIs and data at 8 rpm. (f) Confusion matrix for CEIs and data at 8 rpm. (g) Confusion matrix for CIs and data at 5 rpm. (h) Confusion matrix for EIs and data at 5 rpm. (i) Confusion matrix for CEIs and data at 5 rpm.

The confusion matrices for 8 rpm are shown in Figure 13d for CIs, Figure 13e for entropy-based indicators (EIs), and Figure 13f for CIEs. Although the accuracy score for this classification is 2 percentage points above the classification for CIs at 10 rpm, the confusion matrix shows the misclassification for all scenarios, except for BS. In contrast to the results in Figure 13a, RS and RL are the scenarios with the worst classification (66% and 58%, respectively). The previous situation is improved in Figure 13e, where entropy-based indicators (EIs) has a perfect score for RL and BS. While almost 10% of the BS is misclassified, the HS and RS classifications are improved by almost 20 percentage points. Furthermore, the BL classification is improved to the maximum value. Figure 13f denotes an improvement of the classification compared with the previously described confusion matrix (where RL, BS, and BL have perfect classification). In contrast, the scores of HS and RS are above 90%, and the misclassified values from each scenario are confused with the other aforementioned scenario, respectively.

The results for classification at 5 rpm are presented in Figure 13g–i. The trained model in Figure 13g has the lowest classification rate from all models at 5 rpm. In this figure, a critical case takes place, where a significant proportion of the data from all scenarios is misclassified. For the classification of RS, 59% of the values are correctly classified, which has the lowest accuracy rate compared with other scenarios. As approximately 40% of RS data is confused, the main problem is the variety of scenarios in which the misclassified data are categorized. In contrast, only data from RL are not confused as HS. The previous situation is improved using EIs and CIEs, as shown in Figure 13h,i, where the classification is generally congruent with the original data. The only exception is the classification of the BS in Figure 13i, where less than 6% of the data is misjudged as BL.

Considering the contrast of the scenarios, HS is correctly classified in all models. The worst accuracy score occurs in the CIs-based trained model at 8 rpm, with 79% of the rotations classified. Next, RS and RL are classified at a lower accuracy by the CIs group (with lower scores at 8 rpm) compared with the EIs and CIEs groups. Moreover, several RS rotations are confused as HS, but at 5 rpm, all rotations are perfectly scored by EIs and CIEs. For RL, the lower scores are also in the CIs group. However, BS and BL are classified with higher scores by the EIs and CIEs groups at all rotational speeds compared with the EIs group. The only exception is at 10 rpm, where 90% of the rotations are classified as RS and 10% as HS.

In general, the RF trained models correctly classified over 56% of the data from each bearing scenario at all rotational speeds. The models based on CIs classified HS rotations over 79% of the rotations as HS. Based on the scores obtained from classified damaged scenarios, the CIs group had some difficulty with respect to differentiating between them (in comparison to the entropy-based indicators (EIs) and CIEs groups), as some rotations from damaged scenarios are misclassified several times as rotations from a healthy bearing. The results show that the diagnosis by entropy-based indicators (EIs) of low-speed bearing is feasible. With accuracy scores of 98%, a healthy bearing can be accurately differentiated from a damaged bearing. As the majority of classification scores from damaged scenarios are higher than the scores obtained by the CIs group, it can be stated that the information extracted using entropy-based indicators (EIs) contributes to the classification of damaged bearings with greater accuracy than entropy-based indicators (EIs). In addition, the information extracted using EIs can contribute positively to CIs to distinguishing between healthy and damaged bearings. The evidence shows that CIEs correctly classified more than 93% of the times relative to the rotations of a signal from healthy bearings, which is higher than the CIs and EIs groups. Finally, as the rotational speed decreases, the improvement of the CIs group is nearly half of the improvement from the EIs and CIEs group. Hence, a possible correlation between the addition of EIs to the trained RF model is stated. This is a positive behavior for the rotational work speed of pitch bearings, which are as low as 1 rpm.

The classification of data with four EIs is compared with the same classification process using 16 CIs, and later the indicators are combined for a third classification campaign.

The information from the data obtained by EIs enables a higher classification precision compared with the results obtained from the CIs classification. The available data, which includes several damage scenarios, show that the CIs classification is useful for determining the health of a bearing, but the declaration of the damage scenario is not as clear as with the EIs classification. The combination of both indicator types in the CIEs classification shows a general improvement in the results. Therefore, EIs can be considered as a complement for classification using CIs for better and more detailed diagnosis.

The results show the effectiveness of the entropy-based indicators for low-speed bearing diagnosis. However, several requirements are needed to implement these indicators in reality. Considering the technology readiness level of this research (level 4), the controlled laboratory environment, where no external perturbations are found, is still distant from reality. Therefore, the next step for future analysis is to validate the EIs in a relevant environment. From this perspective, the environment of pitch bearing is considerably noisier than a controlled laboratory environment. As said before, a wind turbine has several electromechanical parts, and each part contributes as an external perturbation to a vibration signal. As a result, these external perturbations can contribute negatively to the application of the EIs. Besides, the size and work regimen of the pitch bearings are needed to be considered for a final application. Another issue is the rotation of the pitch bearing, which is normally not a complete rotation. The main reason is the generation control system, which considers the random forces coming from the wind for the position of the blade against the wind. A subsequent study could bring the data set closer to reality, configuring the movement of the bearing clockwise and anticlockwise for further analysis.

The evolution of pitch bearings from a healthy state to a damaged one can last from months to years. Since the size of the data volume is unpractical for analysis, a meaningful time data acquisition per day is enough for monitoring systems. As the entropy-based indicators were calculated per second, independent of the speed, this can help its implementation in a monitoring system. One main conclusion from the results is the improvement of the information obtained from EIs at a lower rotational speed. This is particularly advantageous for its application. The focus of this work is the analysis of pitch bearings from wind turbines, but some other applications of low-speed bearings can be benefited from this method. For example, yaw bearings from wind turbines, slew bearings from cranes, or swing bearings from backhoe or excavator machinery.

## 6. Conclusions

In this paper, the use of entropy-based indicators for the classification of low-speed bearing vibration signals was proposed as an alternative to classic indicators used in the literature. To demonstrate the feasibility and benefits of entropy-based indicators (EIs), the classification of a set of bearings was proposed using vibration analysis. First, the signals from the data set were divided into time series. Then, entropy-based indicators were calculated, in addition to classic indicators used in the literature. RF models were trained based on previous results, to classify the time series. The data set employed had multiple damage scenarios at various low rotational speeds hence, the contrast among rotational speed, damage scenario types (including damage position and size), and indicators is stated.

The results showed that EIs and its use for the classification of low-speed bearing vibration signals was feasible as useful information from the signals could be obtained for diagnosis. In addition, the findings indicated that 4 EIs can be classified with an accuracy score of up to 22 percentage points above the scores obtained by 16 CIs for 5, 8, and 10 rpm. The research also demonstrated that for lower rotational speeds, the data classification score obtained using EIs could be more than 20 percentage points higher than the CIs. Finally, the results show that the use of EIs in conjunction with CIs could classify a dataset with higher accuracy compared to the classification with the indicators separately. This shows the ability of EIs to differentiate among damaged scenarios, which is an improvement in comparison to the CIs classification, which is beneficial for the diagnosis of low-speed bearings.

As the dataset used for this study has several controlled laboratory conditions, multiple challenges must be overcome before the application of EIs in systems in operational environments. For example, the forces applied to the bearing in this study were always continuous as well as the rotational speed. Thus, one of the following steps for the study consists of the design and execution of tests with variable applied forces. Moreover, because pitch bearings rotate by several degrees each time, reproducing such cyclical movement could increase the proximity to real operating conditions, in contrast to the constant rotational speed used in this study. The addition of variable rotational speeds can also help to conceptualize the boundaries of EIs. Finally, the suitability of the EIs for scenarios with rotational speeds lower than 5 rpm could be analyzed.

The damaged bearings used in this study differ from each other based on the location or size of the damage. A further study could combine the location of several damages to include scenarios where a combination of ball and raceway (or even cage) can occur. An increase in the complexity of the damage by introducing irregularities in the surfaces or testing naturally degraded bearings could also prove useful not only to improve the analysis of the current state by the use of EIs, but to also state the basis for prognosis based on EIs.

**Author Contributions:** Study conception and design were done by D.S., U.L., Y.V., and F.P. Data collection, data analysis, data interpretation, and manuscript drafting were performed by D.S. Critical revisions of the manuscript were performed by U.L., Y.V., and F.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was partially funded by the Spanish Agencia Estatal de Investigación (AEI)—Ministerio de Economía, Industria y Competitividad (MINECO), the Fondo Europeo de Desarrollo Regional (FEDER) through the research projects DPI2017-82930-C2-1-R and DPI2017-82930-C2-2-R, and by the Generalitat de Catalunya through the research project 2017 SGR 388.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflict of interest. The founders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

AppEn	approximate entropy
BL	large ball bearing damage scenario
BS	small ball bearing damage scenario
CBM	condition-based maintenance
CEI	classic and entropy indicators
CF	crest factor
CI	classic indicator
CM	condition monitoring
DisEn	dispersion entropy
DT	decision tree
EI	entropy indicator
EMD	empirical mode decomposition
EU	European Union
FC	frequency center
HS	healthy scenario
IF	impulse factor
IMF	intrinsic mode function
MF	margin factor
NCDF	normal cumulative distribution function

O&M	operating and maintenance
PerEn	permutation entropy
PSD	power spectral density
RF	random forest
RL	large raceway shaft washer damage scenario
RMS	root mean square
RMSF	root mean square frequency
rpm	revolutions per minute
RS	small raceway shaft washer damage scenario
RVF	root variance frequency
SepEn	spectral entropy of the permutation entropy signal
SF	shape factor
SpeEn	spectral entropy
SvdEn	singular value decomposition entropy

## References

- García Álvarez, M.; Moreno, B. Environmental Performance Assessment in the EU: A Challenge for the Sustainability. *J. Clean. Prod.* **2018**, *205*, 266–280. [\[CrossRef\]](#)
- Europe, W. *Our Energy, Our Future*; Technical Report; Wind Europe: Brussels, Belgium, 2019.
- Komusanac, I.; Fraile, D.; Brindley, G. *Wind Energy in Europe in 2018*; Technical Report; Wind Europe: Brussels, Belgium, 2019.
- IRENA. *Renewable Energy Cost Analysis—Wind Power*; Technical Report; International Renewable Energy Agency: Abu Dhabi, UAE, 2012.
- ISO. *Condition Monitoring and Diagnostics of Machines—Vocabulary*; Technical Report ISO 13372; ISO: Geneva, Switzerland, 2012.
- EN, C. *13306: Maintenance Terminology*; European Committee for Standardization: Brussels, Belgium, 2001.
- De Azevedo, H.; Araújo, A.; Bouchonneau, N. A Review of Wind Turbine Bearing Condition Monitoring: State of the Art and Challenges. *Renew. Sustain. Energy Rev.* **2016**, *56*, 368–379. [\[CrossRef\]](#)
- Elforjani, M.; Bechhoefer, E. Analysis of Extremely Modulated Faulty Wind Turbine Data Using Spectral Kurtosis and Signal Intensity Estimator. *Renew. Energy* **2018**, *127*, 258–268. [\[CrossRef\]](#)
- Wang, Z.; Yao, L.; Cai, Y.; Zhang, J. Mahalanobis Semi-Supervised Mapping and Beetle Antennae Search Based Support Vector Machine for Wind Turbine Rolling Bearings Fault Diagnosis. *Renew. Energy* **2020**, *155*, 1312–1327. [\[CrossRef\]](#)
- Castellani, F.; Garibaldi, L.; Daga, A.P.; Astolfi, D.; Natili, F. Diagnosis of Faulty Wind Turbine Bearings Using Tower Vibration Measurements. *Energies* **2020**, *13*, 1474. [\[CrossRef\]](#)
- Hsu, M.C.; Bazilevs, Y. Fluid–Structure Interaction Modeling of Wind Turbines: Simulating the Full Machine. *Comput. Mech.* **2012**, *50*, 821–833. [\[CrossRef\]](#)
- Kimball, R.; Goupee, A.J.; Fowler, M.J.; de Ridder, E.J.; Helder, J. Wind/Wave Basin Verification of a Performance-Matched Scale-Model Wind Turbine on a Floating Offshore Wind Turbine Platform. In Proceedings of the OMAE2014, San Francisco, CA, USA, 8–13 June 2014; [\[CrossRef\]](#)
- Manolas, D.I.; Riziotis, V.A.; Voutsinas, S.G. Assessing the Importance of Geometric Nonlinear Effects in the Prediction of Wind Turbine Blade Loads. *J. Comput. Nonlinear Dyn.* **2015**, *10*. [\[CrossRef\]](#)
- Manzano Agugliaro, F.; Sánchez-Calero, M.; Alcayde, A.; San-Antonio-Gómez, C.; Perea-Moreno, A.J.; Salmeron-Manzano, E. Wind Turbines Offshore Foundations and Connections to Grid. *Inventions* **2020**, *5*, 8. [\[CrossRef\]](#)
- Antonutti, R.; Peyrard, C.; Johanning, L.; Incecik, A.; Ingram, D. An Investigation of the Effects of Wind-Induced Inclination on Floating Wind Turbine Dynamics: Heave Plate Excursion. *Ocean Eng.* **2014**, *91*, 208–217. [\[CrossRef\]](#)
- Sandoval, D.; Leturiondo, U.; Pozo, F.; Vidal, Y.; Salgado, O. Trends in Condition Monitoring for Pitch Bearings. In Proceedings of the 16th International Conference on Condition Monitoring and Asset Management, CM 2019, Glasgow, UK, 25–27 June 2019. [\[CrossRef\]](#)
- Malla, C.; Panigrahi, I. Review of Condition Monitoring of Rolling Element Bearing Using Vibration Analysis and Other Techniques. *J. Vib. Eng. Technol.* **2019**, *7*, 407–414. [\[CrossRef\]](#)
- Boškoski, P.; Gašperin, M.; Petelin, D. Bearing Fault Prognostics Based on Signal Complexity and Gaussian Process Models. In Proceedings of the 2012 IEEE Conference on Prognostics and Health Management, Denver, CO, USA, 18–21 June 2012; pp. 1–8. [\[CrossRef\]](#)
- Jiang, K.; Xu, G.; Liang, L.; Zhao, G.; Tao, T. A Quantitative Diagnosis Method for Rolling Element Bearing Using Signal Complexity and Morphology Filtering. *J. Vibroeng.* **2012**, *14*, 1862–1875.
- Xue, W.; Dai, X.; Zhu, J.; Luo, Y.; Yang, Y. A Noise Suppression Method of Ground Penetrating Radar Based on EEMD and Permutation Entropy. *IEEE Geosci. Remote. Sens. Lett.* **2019**, *16*, 1625–1629. [\[CrossRef\]](#)
- Žvokelj, M.; Zupan, S.; Prebil, I. EEMD Based Multiscale ICA Method for Slewing Bearing Fault Detection and Diagnosis. *J. Sound Vib.* **2016**, *370*, 394–423. [\[CrossRef\]](#)
- Han, T.; Liu, Q.; Zhang, L.; Tan, A. Fault Feature Extraction of Low Speed Roller Bearing Based on Teager Energy Operator and CEEMD. *Meas. J. Int. Meas. Confed.* **2019**, *138*, 400–408. [\[CrossRef\]](#)

23. Hou, J.; Wu, Y.; Gong, H.; Ahmad, A.S.; Liu, L. A Novel Intelligent Method for Bearing Fault Diagnosis Based on EEMD Permutation Entropy and GG Clustering. *Appl. Sci.* **2020**, *10*, 386. [[CrossRef](#)]
24. Heydarzadeh, M.; Nourani, M.; Azimi, V.; Kashani-Pour, A. Vibration Based Bearing Fault Diagnosis Using Reflection Coefficients of the Autoregressive Model. In Proceedings of the American Control Conference, Milwaukee, WI, USA, 27–29 June 2018; pp. 5801–5806. [[CrossRef](#)]
25. An, X.; Zeng, H.; Yang, W.; An, X. Fault Diagnosis of a Wind Turbine Rolling Bearing Using Adaptive Local Iterative Filtering and Singular Value Decomposition. *Trans. Inst. Meas. Control* **2016**, *39*, 1643–1648. [[CrossRef](#)]
26. Xu, Y.; Fan, F.; Jiang, X. A Fast Iterative Filtering Decomposition and Symmetric Difference Analytic Energy Operator for Bearing Fault Extraction. *ISA Trans.* **2020**. [[CrossRef](#)]
27. Borghesani, P.; Pennacchi, P.; Chatterton, S.; Ricci, R. The Velocity Synchronous Discrete Fourier Transform for Order Tracking in the Field of Rotating Machinery. *Mech. Syst. Signal Process.* **2014**, *44*, 118–133. [[CrossRef](#)]
28. Yang, J.; Huang, D.; Zhou, D.; Liu, H. Optimal IMF Selection and Unknown Fault Feature Extraction for Rolling Bearings with Different Defect Modes. *Measurement* **2020**, 107660. [[CrossRef](#)]
29. Rhif, M.; Ben Abbes, A.; Farah, I.R.; Martínez, B.; Sang, Y. Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. *Appl. Sci.* **2019**, *9*, 1345. [[CrossRef](#)]
30. Gu, X.; Yang, S.; Liu, Y.; Hao, R. Rolling Element Bearing Faults Diagnosis Based on Kurtogram and Frequency Domain Correlated Kurtosis. *Meas. Sci. Technol.* **2016**, *27*, 125019. [[CrossRef](#)]
31. Pham, M.T.; Kim, J.M.; Kim, C.H. Accurate Bearing Fault Diagnosis under Variable Shaft Speed Using Convolutional Neural Networks and Vibration Spectrogram. *Appl. Sci.* **2020**, *10*, 6385. [[CrossRef](#)]
32. Pan, Y.; Hong, R.; Chen, J.; Qin, Z.; Feng, Y. Incipient Fault Detection of Wind Turbine Large-Size Slewing Bearing Based on Circular Domain. *Measurement* **2019**, *137*, 130–142. [[CrossRef](#)]
33. Caesarendra, W.; Tjahjowidodo, T.; Kosasih, B.; Tieu, A. Integrated Condition Monitoring and Prognosis Method for Incipient Defect Detection and Remaining Life Prediction of Low Speed Slew Bearings. *Machines* **2017**, *5*, 11. [[CrossRef](#)]
34. Li, N.; Lei, Y.; Lin, J.; Ding, S.X. An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings. *IEEE Trans. Ind. Electron.* **2015**, *62*, 7762–7773. [[CrossRef](#)]
35. Islam, M.R.; Kim, Y.H.; Kim, J.Y.; Kim, J.M. Detecting and Learning Unknown Fault States by Automatically Finding the Optimal Number of Clusters for Online Bearing Fault Diagnosis. *Appl. Sci.* **2019**, *9*, 2326. [[CrossRef](#)]
36. Chen, Z.; Deng, S.; Chen, X.; Li, C.; Sanchez, R.V.; Qin, H. Deep Neural Networks-Based Rolling Bearing Fault Diagnosis. *Microelectron. Reliab.* **2017**, *75*, 327–333. [[CrossRef](#)]
37. Nikula, R.P.; Karioja, K.; Pylvänäinen, M.; Leiviskä, K. Automation of Low-Speed Bearing Fault Diagnosis Based on Autocorrelation of Time Domain Features. *Mech. Syst. Signal Process.* **2020**, *138*, 106572. [[CrossRef](#)]
38. Grover, C.; Turk, N. Optimal Statistical Feature Subset Selection for Bearing Fault Detection and Severity Estimation. *Shock Vib.* **2020**, *2020*, 5742053. [[CrossRef](#)]
39. Cavalaglio, J.; Strozzi, M.; Rubini, R.; Cocconcelli, M. Analysis of NASA Bearing Dataset of the University of Cincinnati by Means of Hjorth's Parameters. In Proceedings of the International Conference on Structural Engineering Dynamics ICEDyn 2019, Viana do Castelo, Portugal, 24–26 June 2019.
40. Grover, C.; Turk, N. Rolling Element Bearing Fault Diagnosis Using Empirical Mode Decomposition and Hjorth Parameters. *Int. Conf. Comput. Intell. Data Sci.* **2020**, *167*, 1484–1494. [[CrossRef](#)]
41. Liang, T.; Lu, H. A Novel Method Based on Multi-Island Genetic Algorithm Improved Variational Mode Decomposition and Multi-Features for Fault Diagnosis of Rolling Bearing. *Entropy* **2020**, *22*, 995. [[CrossRef](#)] [[PubMed](#)]
42. Han, M.; Pan, J. A Fault Diagnosis Method Combined with LMD, Sample Entropy and Energy Ratio for Roller Bearings. *Measurement* **2015**, *76*, 7–19. [[CrossRef](#)]
43. Fu, Q.; Jing, B.; He, P.; Si, S.; Wang, Y. Fault Feature Selection and Diagnosis of Rolling Bearings Based on EEMD and Optimized Elman\_AdaBoost Algorithm. *IEEE Sens. J.* **2018**, *18*, 5024–5034. [[CrossRef](#)]
44. Tang, H.; Du, Y.; Dai, H.L. Rolling Element Bearing Diagnosis Based on Probability Box Theory. *Appl. Math. Model.* **2020**, *80*, 944–960. [[CrossRef](#)]
45. Akter, M.S.; Islam, M.R.; Iimura, Y.; Sugano, H.; Fukumori, K.; Wang, D.; Tanaka, T.; Cichocki, A. Multiband Entropy-Based Feature-Extraction Method for Automatic Identification of Epileptic Focus Based on High-Frequency Components in Interictal iEEG. *Sci. Rep.* **2020**, *10*, 7044. [[CrossRef](#)]
46. Richman, J.S.; Moorman, J.R. Physiological Time-Series Analysis Using Approximate Entropy and Sample Entropy. *Am. J. Physiol. Circ. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)]
47. Li, S.y.; Yang, M.; Li, C.c.; Cai, P. Analysis of Heart Rate Variability Based on Singular Value Decomposition Entropy. *J. Shanghai Univ.* **2008**, *12*, 433. [[CrossRef](#)]
48. Fazan, F.S.; Brognara, F.; Fazan Junior, R.; Murta Junior, L.O.; Virgilio Silva, L.E. Changes in the Complexity of Heart Rate Variability with Exercise Training Measured by Multiscale Entropy-Based Measurements. *Entropy* **2018**, *20*, 47. [[CrossRef](#)]
49. Henry, M.; Judge, G. Permutation Entropy and Information Recovery in Nonlinear Dynamic Economic Time Series. *Econometrics* **2019**, *7*, 10. [[CrossRef](#)]
50. Xiao, D.; Wang, J. Complexity Behaviours of Agent-Based Financial Dynamics by Hetero-Distance Contact Process. *Nonlinear Dyn.* **2020**, *100*, 3867–3886. [[CrossRef](#)]

51. Shang, Y.; Lu, G.; Kang, Y.; Zhou, Z.; Duan, B.; Zhang, C. A Multifault Diagnosis Method Based on Modified Sample Entropy for Lithium-Ion Battery Strings. *J. Power Sources* **2020**, *446*, 227275. [[CrossRef](#)]
52. Baldini, G.; Giuliani, R.; Steri, G.; Neisse, R. Physical Layer Authentication of Internet of Things Wireless Devices through Permutation and Dispersion Entropy. In Proceedings of the 2017 Global Internet of Things Summit (GIoTS), Geneva, Switzerland, 6–9 June 2017; pp. 1–6; [[CrossRef](#)]
53. Fei, C.W.; Choy, Y.S.; Bai, G.C.; Tang, W.Z. Multi Feature Entropy Distance Approach with Vibration and Acoustic Emission Signals for Process Feature Recognition of Rolling Element Bearing Faults. *Struct. Health Monit.* **2017**, *17*, 156–168. [[CrossRef](#)]
54. Gu, R.; Chen, J.; Hong, R.; Wang, H.; Wu, W. Incipient Fault Diagnosis of Rolling Bearings Based on Adaptive Variational Mode Decomposition and Teager Energy Operator. *Measurement* **2019**, 106941. [[CrossRef](#)]
55. Zhang, J.; Zhao, Y.; Li, X.; Liu, M. Bearing Fault Diagnosis with Kernel Sparse Representation Classification Based on Adaptive Local Iterative Filtering-Enhanced Multiscale Entropy Features. *Math. Probl. Eng.* **2019**, 2019. [[CrossRef](#)]
56. Qin, X.; Li, Q.; Dong, X.; Lv, S. The Fault Diagnosis of Rolling Bearing Based on Ensemble Empirical Mode Decomposition and Random Forest. *Shock Vib.* **2017**, 2017, e2623081. [[CrossRef](#)]
57. Wang, F.; Zhang, Y.; Zhang, B.; Su, W. Application of Wavelet Packet Sample Entropy in the Forecast of Rolling Element Bearing Fault Trend. In Proceedings of the 2011 International Conference on Multimedia and Signal Processing, Dalian, China, 14–16 October 2011; Volume 2, pp. 12–16. [[CrossRef](#)]
58. Vakharia, V.; Gupta, V.; Kankar, P. A Multiscale Permutation Entropy Based Approach to Select Wavelet for Fault Diagnosis of Ball Bearings. *J. Vib. Control* **2014**, *21*, 3123–3131. [[CrossRef](#)]
59. Huo, Z.; Zhang, Y.; Jombo, G.; Shu, L. Adaptive Multiscale Weighted Permutation Entropy for Rolling Bearing Fault Diagnosis. *IEEE Access* **2020**, *8*, 87529–87540. [[CrossRef](#)]
60. Yang, C.; Jia, M. Health Condition Identification for Rolling Bearing Based on Hierarchical Multiscale Symbolic Dynamic Entropy and Least Squares Support Tensor Machine-Based Binary Tree. *Struct. Health Monit.* **2020**, 1475921720923973. [[CrossRef](#)]
61. Fu, W.; Shao, K.; Tan, J.; Wang, K. Fault Diagnosis for Rolling Bearings Based on Composite Multiscale Fine-Sorted Dispersion Entropy and SVM With Hybrid Mutation SCA-HHO Algorithm Optimization. *IEEE Access* **2020**, *8*, 13086–13104. [[CrossRef](#)]
62. Caesarendra, W.; Kosasih, B.; Tieu, K.; Moodie, C. An Application of Nonlinear Feature Extraction - A Case Study for Low Speed Slewing Bearing Condition Monitoring and Prognosis. In Proceedings of the 2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics: Mechatronics for Human Wellbeing, AIM 2013, Wollongong, Australia, 9–12 July 2013; pp. 1713–1718. [[CrossRef](#)]
63. Caesarendra, W.; Lee, J.M.; Ha, J.M.; Choi, B.K. Slew Bearing Early Damage Detection Based on Multivariate State Estimation Technique and Sequential Probability Ratio Test. In Proceedings of the 2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), Busan, Korea, 7–11 July 2015; pp. 1161–1166. [[CrossRef](#)]
64. An, X.; Pan, L. Bearing Fault Diagnosis of a Wind Turbine Based on Variational Mode Decomposition and Permutation Entropy. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2017**, *231*, 200–206. [[CrossRef](#)]
65. Glowacz, A.; Glowacz, W.; Kozik, J.; Piech, K.; Gutten, M.; Caesarendra, W.; Liu, H.; Brumerick, F.; Irfan, M.; Khan, Z.F. Detection of Deterioration of Three-Phase Induction Motor Using Vibration Signals. *Meas. Sci. Rev.* **2019**, *19*, 241–249. [[CrossRef](#)]
66. Tian, H.; Bai, P.; Tan, Y.; Li, Z.; Peng, D.; Xiao, X.; Zhao, H.; Zhou, Y.; Liang, W.; Zhang, L. A New Method to Detect Methylation Profiles for Forensic Body Fluid Identification Combining ARMS-PCR Technique and Random Forest Model. *Forensic Sci. Int. Genet.* **2020**, *49*, 102371. [[CrossRef](#)] [[PubMed](#)]
67. Khammas, B.M. Ransomware Detection Using Random Forest Technique. *ICT Express* **2020**. [[CrossRef](#)]
68. Liu, H.; Luo, N.; Hu, C. Detection of County Economic Development Using LJ1-01 Nighttime Light Imagery: A Comparison with NPP-VIIRS Data. *Sensors* **2020**, *20*, 6633. [[CrossRef](#)]
69. Quan, Y.; Tong, Y.; Feng, W.; Dauphin, G.; Huang, W.; Xing, M. A Novel Image Fusion Method of Multi-Spectral and SAR Images for Land Cover Classification. *Remote Sens.* **2020**, *12*, 3801. [[CrossRef](#)]
70. Adams, S.M.; Feroze, H.; Nguyen, T.; Eum, S.; Cornelio, C.; Harralson, A.F. Genome Wide Epistasis Study of On-Statins Cardiovascular Events with Iterative Feature Reduction and Selection. *J. Pers. Med.* **2020**, *10*, 212. [[CrossRef](#)]
71. Smith, W.A.; Randall, R.B. Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study. *Mech. Syst. Signal Process.* **2015**, 64–65, 100–131. [[CrossRef](#)]
72. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An Experimental Platform for Bearings Accelerated Degradation Tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, PHM'12, Denver, CO, USA, 18–21 June 2012.
73. Lee, J.; Qiu, H.; Yu, G.; Services, R.T. “Bearing Data Set”, NASA Ames Prognostics Data Repository. 2007. Available online: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/publications/#bearing> (accessed on 22 October 2020)
74. Sandoval, D.; Leturiondo, U.; Pozo, F.; Vidal, Y. Low Speed Bearing Fault Diagnosis Based on Permutation and Spectral Entropy Measures. *Appl. Sci.* **2020**, *10*, 4666. [[CrossRef](#)]
75. Wang, S.; Chen, J.; Wang, H.; Zhang, D. Degradation Evaluation of Slewing Bearing Using HMM and Improved GRU. *Measurement* **2019**. [[CrossRef](#)]
76. Bao, W.; Wang, H.; Chen, J.; Zhang, B.; Ding, P.; Wu, J.; He, P. Life Prediction of Slewing Bearing Based on Isometric Mapping and Fuzzy Support Vector Regression. *Trans. Inst. Meas. Control* **2019**. [[CrossRef](#)]

77. Kim, Y.H.; Tan, A.C.C.; Kosse, V. Condition Monitoring of Low-Speed Bearings—A Review. *Aust. J. Mech. Eng.* **2008**, *6*, 61–68. [[CrossRef](#)]
78. Agahi, H.; Mahmoodzadeh, A. Decision fusion scheme for bearing defects diagnosis in induction motors. *Electr. Eng.* **2020**, *102*, 2269–2279. [[CrossRef](#)]
79. Rai, A.; Upadhyay, S.H. A Review on Signal Processing Techniques Utilized in the Fault Diagnosis of Rolling Element Bearings. *Tribol. Int.* **2016**, *96*, 289–306. [[CrossRef](#)]
80. Caesarendra, W.; Tjahjowidodo, T. A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing. *Machines* **2017**, *5*, 28. [[CrossRef](#)]
81. Wang, F.; Liu, C.; Su, W.; Xue, Z.; Li, H.; Han, Q. Condition Monitoring and Fault Diagnosis Methods for Low-Speed and Heavy-Load Slewing Bearings: A Literature Review. *J. Vibroeng.* **2017**, *19*, 3429–3444. [[CrossRef](#)]
82. Liu, C.; Wang, F. A Review of Current Condition Monitoring and Fault Diagnosis Methods for Low-Speed and Heavy-Load Slewing Bearings. In Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017, Kunming, China, 10–12 July 2017; pp. 104–109. [[CrossRef](#)]
83. Liu, Z.; Zhang, L. A Review of Failure Modes, Condition Monitoring and Fault Diagnosis Methods for Large-Scale Wind Turbine Bearings. *Measurement* **2020**, *149*, 107002. [[CrossRef](#)]
84. Tandon, N.; Choudhury, A. Review of Vibration and Acoustic Measurement Methods for the Detection of Defects in Rolling Element Bearings. *Tribol. Int.* **1999**, *32*, 469–480. [[CrossRef](#)]
85. Yang, O.S.; Widodo, A. *Introduction of Intelligent Machine Fault Diagnosis and Prognosis*; Nova Science Publishers, Incorporated: Hauppauge, NY, USA, 2010.
86. Kundu, P.; Darpe, A.K.; Kulkarni, M.S. Weibull Accelerated Failure Time Regression Model for Remaining Useful Life Prediction of Bearing Working under Multiple Operating Conditions. *Mech. Syst. Signal Process.* **2019**, *134*, 106302. [[CrossRef](#)]
87. Li, C.; Pickering, C. Robustness and Sensitivity of Non-Dimensional Amplitude Parameters for Diagnosis of Fatigue Spalling. *Cond. Monit. Diagn. Technol.* **1992**, *2*, 81–84.
88. Howard, I. *A Review of Rolling Element Bearing Vibration Detection, Diagnosis and Prognosis*; Technical Report; Defence Science and Technology Organization: Canberra, Australia, 1994.
89. Antoni, J.; Borghesani, P. A Statistical Methodology for the Design of Condition Indicators. *Mech. Syst. Signal Process.* **2019**, *114*, 290–327. [[CrossRef](#)]
90. Wang, H.; Ni, G.; Chen, J.; Qu, J. Research on Rolling Bearing State Health Monitoring and Life Prediction Based on PCA and Internet of Things with Multi-Sensor. *Measurement* **2020**, *157*, 107657. [[CrossRef](#)]
91. Hjorth, B. EEG Analysis Based on Time Domain Properties. *Electroencephalogr. Clin. Neurophysiol.* **1970**, *29*, 306–310. [[CrossRef](#)]
92. Zhang, Y.; Zuo, H.; Bai, F. Classification of Fault Location and Performance Degradation of a Roller Bearing. *Measurement* **2013**, *46*, 1178–1189. [[CrossRef](#)]
93. Clausius, R. On the Moving Force of Heat, and the Laws Regarding the Nature of Heat Itself Which Are Deducible Therefrom. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1851**, *2*, 1–21. [[CrossRef](#)]
94. Namdari, A.; Li, Z.S. A Review of Entropy Measures for Uncertainty Quantification of Stochastic Processes. *Adv. Mech. Eng.* **2019**, *11*, 1687814019857350. [[CrossRef](#)]
95. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
96. Rostaghi, M.; Azami, H. Dispersion Entropy: A Measure for Time-Series Analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [[CrossRef](#)]
97. Pincus, S.M. Approximate Entropy as a Measure of System Complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)]
98. Pincus, S. Approximate Entropy (ApEn) as a Complexity Measure. *Chaos Interdiscip. J. Nonlinear Sci.* **1995**, *5*, 110–117. [[CrossRef](#)]
99. Kurths, J.; Voss, A.; Saperin, P.; Witt, A.; Kleiner, H.J.; Wessel, N. Quantitative Analysis of Heart Rate Variability. *Chaos Interdiscip. J. Nonlinear Sci.* **1995**, *5*, 88–94. [[CrossRef](#)]
100. Lake, D.E.; Richman, J.S.; Griffin, M.P.; Moorman, J.R. Sample Entropy Analysis of Neonatal Heart Rate Variability. *Am. J. -Physiol. Integr. Comp. Physiol.* **2002**, *283*, R789–R797. [[CrossRef](#)] [[PubMed](#)]
101. Hu, J.; Liu, Y.; Sang, Y.F. Precipitation Complexity and Its Spatial Difference in the Taihu Lake Basin, China. *Entropy* **2019**, *21*, 48. [[CrossRef](#)] [[PubMed](#)]
102. Delgado-Bonal, A.; Marshak, A.; Yang, Y.; Holdaway, D. Analyzing Changes in the Complexity of Climate in the Last Four Decades Using MERRA-2 Radiation Data. *Sci. Rep.* **2020**, *10*, 922. [[CrossRef](#)] [[PubMed](#)]
103. Tapia, C.; Coulton, J.; Saydam, S. Using Entropy to Assess Dynamic Behaviour of Long-Term Copper Price. *Resour. Policy* **2020**, *66*, 101597. [[CrossRef](#)]
104. Xie, Y.; Yu, J.; Guo, S.; Ding, Q.; Wang, E. Image Encryption Scheme with Compressed Sensing Based on New Three-Dimensional Chaotic System. *Entropy* **2019**, *21*, 819. [[CrossRef](#)]
105. Zhu, K.; Li, H. A Rolling Element Bearing Fault Diagnosis Approach Based on Hierarchical Fuzzy Entropy and Support Vector Machine. *Proc. Inst. Mech. Eng. Part J. Mech. Eng. Sci.* **2015**, *230*, 2314–2322. [[CrossRef](#)]
106. Li, Y.; Wang, X.; Liu, Z.; Liang, X.; Si, S. The Entropy Algorithm and Its Variants in the Fault Diagnosis of Rotating Machinery: A Review. *IEEE Access* **2018**, *6*, 66723–66741. [[CrossRef](#)]
107. Pincus, S.M.; Goldberger, A.L. Physiological Time-Series Analysis: What Does Regularity Quantify? *Am. J. Physiol.* **1994**, *266*, H1643–H1656. [[CrossRef](#)]



108. Pincus, S.M.; Keefe, D.L. Quantification of Hormone Pulsatility via an Approximate Entropy Algorithm. *Am. J. Physiol.* **1992**, *262*, E741–E754. [[CrossRef](#)]
109. Bandt, C.; Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)]
110. Azami, H.; Abásolo, D.; Simons, S.; Escudero, J. Univariate and Multivariate Generalized Multiscale Entropy to Characterise EEG Signals in Alzheimer’s Disease. *Entropy* **2017**, *19*, 31. [[CrossRef](#)]
111. Tripathy, R.K.; Deb, S.; Dandapat, S. Analysis of Physiological Signals Using State Space Correlation Entropy. *Healthc. Technol. Lett.* **2017**, *4*, 30–33. [[CrossRef](#)] [[PubMed](#)]
112. Pham, T.D. TimeShift Multiscale Entropy Analysis of Physiological Signals. *Entropy* **2017**, *19*, 257. [[CrossRef](#)]
113. Zhou, F.; Shen, J.; Yang, X.; Liu, X.; Liu, W. Modified Hierarchical Multiscale Dispersion Entropy and Its Application to Fault Identification of Rotating Machinery. *IEEE Access* **2020**, *8*, 161361–161376. [[CrossRef](#)]
114. Yan, X.; Jia, M. Intelligent Fault Diagnosis of Rotating Machinery Using Improved Multiscale Dispersion Entropy and mRMR Feature Selection. *Knowl. Based Syst.* **2019**, *163*, 450–471. [[CrossRef](#)]
115. Yang, Y.; Zheng, H.; Li, Y.; Xu, M.; Chen, Y. A Fault Diagnosis Scheme for Rotating Machinery Using Hierarchical Symbolic Analysis and Convolutional Neural Network. *ISA Trans.* **2019**, *91*, 235–252. [[CrossRef](#)]
116. Azami, H.; Rostaghi, M.; Abásolo, D.; Escudero, J. Refined Composite Multiscale Dispersion Entropy and Its Application to Biomedical Signals. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2872–2879. [[CrossRef](#)]
117. Azami, H.; Rostaghi, M.; Fernández, A.; Escudero, J. Dispersion Entropy for the Analysis of Resting-State MEG Regularity in Alzheimer’s Disease. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 6417–6420. [[CrossRef](#)]
118. Roberts, S.J.; Penny, W.; Rezek, I. Temporal and Spatial Complexity Measures for Electroencephalogram Based Brain-Computer Interfacing. *Med Biol. Eng. Comput.* **1999**, *37*, 93–98. [[CrossRef](#)]
119. Bashashati, A.; Fatourehchi, M.; Ward, R.K.; Birch, G.E. A Survey of Signal Processing Algorithms in Brain-Computer Interfaces Based on Electrical Brain Signals. *J. Neural Eng.* **2007**, *4*, R32–R57. [[CrossRef](#)]
120. Greene, B.R.; Faul, S.; Marnane, W.P.; Lightbody, G.; Korotchikova, I.; Boylan, G.B. A Comparison of Quantitative EEG Features for Neonatal Seizure Detection. *Clin. Neurophysiol.* **2008**, *119*, 1248–1261. [[CrossRef](#)]
121. Lambert, W.; Cobus, L.A.; Frappart, T.; Fink, M.; Aubry, A. Distortion Matrix Approach for Ultrasound Imaging of Random Scattering Media. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14645–14656. [[CrossRef](#)] [[PubMed](#)]
122. Badon, A.; Barolle, V.; Irsch, K.; Boccara, A.C.; Fink, M.; Aubry, A. Distortion Matrix Concept for Deep Optical Imaging in Scattering Media. *Sci. Adv.* **2020**, *6*, eaay7170. [[CrossRef](#)] [[PubMed](#)]
123. Inouye, T.; Shinosaki, K.; Sakamoto, H.; Toi, S.; Ukai, S.; Iyama, A.; Katsuda, Y.; Hirano, M. Quantification of EEG Irregularity by Use of the Entropy of the Power Spectrum. *Electroencephalogr. Clin. Neurophysiol.* **1991**, *79*, 204–210. [[CrossRef](#)]
124. Powell, G.E.; Percival, I.C. A Spectral Entropy Method for Distinguishing Regular and Irregular Motion of Hamiltonian Systems. *J. Phys. Math. Gen.* **1979**, *12*, 2053–2071. [[CrossRef](#)]
125. Caesarendra, W.; Kosasih, P.; Tieu, A.; Moodie, C.; Choi, B.K. Condition Monitoring of Naturally Damaged Slow Speed Slewing Bearing Based on Ensemble Empirical Mode Decomposition. *J. Mech. Sci. Technol.* **2013**, *27*, 2253–2262. [[CrossRef](#)]
126. Rostaghi, M.; Ashory, M.R.; Azami, H. Application of Dispersion Entropy to Status Characterization of Rotary Machines. *J. Sound Vib.* **2019**, *438*, 291–308. [[CrossRef](#)]
127. Yan, X.; Liu, Y.; Huang, D.; Jia, M. A New Approach to Health Condition Identification of Rolling Bearing Using Hierarchical Dispersion Entropy and Improved Laplacian Score. *Struct. Health Monit.* **2020**, 1475921720948620. [[CrossRef](#)]
128. Lu, Y.; Li, Y.; Xiao, H.; Wang, Z. A Sliding Singular Spectrum Entropy Method and Its Application to Gear Fault Diagnosis. In *Advanced Intelligent Computing Theories and Applications; With Aspects of Theoretical and Methodological Issues*; Huang, D.S., Wunsch, D.C., Levine, D.S., Jo, K.H., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 669–678.
129. Tin Kam Ho. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832–844. [[CrossRef](#)]
130. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.:1010933404324. [[CrossRef](#)]
131. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 1984.
132. Breiman, L. Bagging Predictors. *Mach. Learn.* **1996**, *24*, 123–140.:1018054314350. [[CrossRef](#)]
133. Yang, L.; Shami, A. On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice. *Neurocomputing* **2020**, *415*, 295–316. [[CrossRef](#)]
134. Hutter, F.; Kotthoff, L.; Vanschoren, J. *Automated Machine Learning: Methods, Systems, Challenges*; Springer Nature: Berlin/Heidelberg, Germany, 2019.
135. Wang, Z.; Zhang, Q.; Xiong, J.; Xiao, M.; Sun, G.; He, J. Fault Diagnosis of a Rolling Bearing Using Wavelet Packet Denoising and Random Forests. *IEEE Sens. J.* **2017**, *17*, 5581–5588. [[CrossRef](#)]