

Article

Rolling-Element Bearing Fault Diagnosis Using Advanced Machine Learning-Based Observer

Farzin Piltan ¹, Alexander E. Prosvirin ¹, Inkyu Jeong ², Kichang Im ² and Jong-Myon Kim ^{1,*}

¹ Department of Electrical, Electronics and Computer Engineering, University of Ulsan, Ulsan 44610, Korea; piltanfarzin@gmail.com (F.P.); a.prosvirin@hotmail.com (A.E.P.)

² ICT Convergence Safety Center, University of Ulsan, Ulsan 44610, Korea; inkyujeong@outlook.com (I.J.); kichang@ulsan.ac.kr (K.I.)

* Correspondence: jmkim07@ulsan.ac.kr; Tel.: +82-52-259-2217

Received: 8 November 2019; Accepted: 6 December 2019; Published: 10 December 2019



Abstract: Rotating machines represent a class of nonlinear, uncertain, and multiple-degrees-of-freedom systems that are used in various applications. The complexity of the system's dynamic behavior and uncertainty result in substantial challenges for fault estimation, detection, and identification in rotating machines. To address the aforementioned challenges, this paper proposes a novel technique for fault diagnosis of a rolling-element bearing (REB), founded on a machine-learning-based advanced fuzzy sliding mode observer. First, an ARX-Laguerre algorithm is presented to model the bearing in the presence of noise and uncertainty. In addition, a fuzzy algorithm is applied to the ARX-Laguerre technique to increase the system's modeling accuracy. Next, the conventional sliding mode observer is applied to resolve the problems of fault estimation in a complex system with a high degree of uncertainty, such as rotating machinery. To address the problem of chattering that is inherent in the conventional sliding mode observer, the higher-order super-twisting (advanced) technique is introduced in this study. In addition, the fuzzy method is applied to the advanced sliding mode observer to improve the accuracy of fault estimation in uncertain conditions. As a result, the advanced fuzzy sliding mode observer adaptively improves the reliability, robustness, and estimation accuracy of rolling-element bearing fault estimation. Then, the residual signal delivered by the proposed methodology is split in the windows and each window is characterized by a numerical parameter. Finally, a machine learning technique, called a decision tree, adaptively derives the threshold values that are used for problems of fault detection and fault identification in this study. The effectiveness of the proposed algorithm is validated using a publicly available vibration dataset of Case Western Reserve University. The experimental results show that the machine learning-based advanced fuzzy sliding mode observation methodology significantly improves the reliability and accuracy of the fault estimation, detection, and identification of rolling element bearing faults under variable crack sizes and load conditions.

Keywords: rotating machine; bearing; decision trees; sliding mode observer; fuzzy technique; advanced sliding mode observer; fault detection; fault estimation; fault identification; machine learning

1. Introduction

Rolling element bearings (REBs) have been extensively used in several industries, such as the automotive, steam and gas turbines, and power generation industries, to improve their efficiency by reducing friction [1,2]. The complexities of the required tasks, with time-varying and nonlinear parameters in rolling bearings, make their fault estimation, detection, and identification highly challenging. The fault estimation, detection, and identification (further referred to as FEDI) are intransitive to prevent the bearing's destruction. Here, the fault estimation technique is used to estimate the signal (fault) to obtain

the valuable differentiation between various conditions of bearing, the fault detection algorithm is used to detect normal and abnormal conditions, and the fault identification technique is used to identify the specific types of faults in the REBs. Various types of failures have been representing in bearings, which are divided into four foremost groups, i.e., inner race faults, outer race faults, ball or rolling-element faults, and cage faults that in this research called IF, OF, BF, and CF, respectively. To analyze the faults conditions in a bearing, various REB condition monitoring techniques such as vibration, motor current signature analysis (MCSA), and acoustic emission (AE) measurements have been reported [3]. This research exploits the vibration measurements since these signals are suitable for FEDI.

Numerous procedures have been recently presented for FEDI in various systems, which can be divided into four groups: (a) model-based (MB) techniques, (b) signal-based (SB) approaches, (c) knowledge-based (KB) procedures, and (d) hybrid methods [1,4]. Some recently published representative examples of model-based techniques have been reported in [1,5–7]. Model-based techniques for fault diagnosis identify the faults by modeling the system's dynamics using mathematical modeling or system identification techniques using a small dataset. Apart from the various advantages of the MB method such as reliability and robustness, accuracy is the main drawback of this technique [1]. Signal-based fault diagnosis approaches extract fault features and differentiate the health conditions of the system by applying various signal processing techniques to the acquired signals. Some recently published representative examples of signal-based approaches can be found in [8–17]. When using the traditional fault diagnosis frameworks (i.e., knowledge-based methods), it is crucial to select an appropriate signal processing technique that is useful for extracting the representative fault features that are used as inputs to the decision-making approaches. Recently, various types of advanced signal processing approaches were applied for rotating machinery fault diagnosis. These techniques include the conventional wavelet transform [9], wavelet-based kurtogram [10], empirical mode decomposition [11], ensemble empirical mode decomposition [12] and their modifications [13–15], as well as relatively new methods for detecting and extracting the repetitive transients caused by mechanical faults, such as spectral l2/l1 norm [16] and spectral Gini index [17]. These methods are essential for feature engineering; however, in the proposed methodology, the signal processing step is replaced by the observation technique (advanced fuzzy sliding mode observer) which is used to provide new insights and demonstrate the applicability of the control theory field for solving the problems of mechanical fault diagnosis. Knowledge-based approaches were based on the ideas of transferring the industrial knowledge and expertise in fault diagnosis from humans to the machine by creating algorithms that performed fault identification using decision tables and rules. However, recently the trend in KB-based approaches has been shifted toward artificial intelligence techniques that are used for automatically extracting valuable features and making decisions about the fault conditions of the system [8,18,19]. To address the issues of model-based approaches and knowledge-based techniques, hybrid fault diagnosis can be introduced for bearing fault diagnosis. In the hybrid approach, various algorithms from different groups can be used in parallel to improve the performance of the fault diagnosis method. In this research, a robust hybrid fault diagnosis algorithm is presented using a machine learning-based advanced fuzzy sliding mode observer for the REB.

The main challenge in designing the procedure of sliding mode observer is system modeling. The mathematical-based system modeling and system identification are the main frameworks for modeling the complex (e.g., bearing) systems [5,6]. In the mathematical-based system modeling, the Lagrange technique can be used for modeling the REBs [6]. Apart from the reliability and accuracy of physical system modeling, this technique has drawbacks in uncertain and noisy conditions. To address this issue, the system identification techniques are widely applied as indicated in [5,20,21]. To estimate the system using system identification technique, various orthonormal procedures such as Auto Regressive with eXogenous input (ARX), orthonormal function bases (OFB), and generalized orthonormal bases (GOB) methods have been used [20,21]. Apart from several advantages of orthonormal techniques compared to the classical algorithms, these techniques have two important drawbacks. The first one is related to the challenge of finding the optimal orthonormal values and the second problem is related to

the restrictions in decoupled systems [6,7,20,21]. To address these issues, ensure an efficient complexity reduction and reduce the system's estimation order, the ARX-Laguerre technique has been applied in [20,21]. Apart from the advantage of complexity reduction in ARX-Laguerre, it has a challenge related to estimation accuracy in nonlinear systems. To increase the accuracy in the ARX-Laguerre technique, the T-S fuzzy ARX-Laguerre (FAL) technique is presented in the literature [5]. In our work, this technique is used to estimate the vibration signals of REBs.

One of the main model-based techniques that can be used for FEDI is the observation technique [1]. The observation techniques can be categorized into two main groups: (a) The linear observers (e.g., proportional-integral (PI) and proportional multi integral (PMI)) and (b) nonlinear observers (e.g., feedback linearization observer (FLO), sliding mode observer (SMO), and fuzzy observer (FO)). To reduce the observation error, the linear observer only uses the feedback term. Robustness and reliability in uncertain conditions are the main drawbacks of linear observers [21,22]. Regarding the nonlinear observers, they can be designed based on driving the dynamic system's behavior in parallel with the linear observer to increase the robustness and reliability of the linear observer for fault estimation [6,7]. In this research nonlinear observer is used for FEDI. Despite the improvement of the accuracy obtained from using the FLO, robustness is the main issue of this technique [7]. The extended FLO and SMO have been presented by researchers to improve the robustness of observers [5–7]. The SMO is a robust technique for FEDI for nonlinear and complex systems (e.g., rolling bearings) which operate in uncertain and noisy conditions. The nonlinear switching function in the SMO is defined to converge the output estimation error toward zero. This technique can perform FEDI according to adaptive updates of the observer parameters, which can significantly improve the performance of FEDI in nonlinear systems [23,24]. Though the SMO increases the robustness, this scheme unfortunately suffers from the chattering phenomenon and reduced fault estimation accuracy in the presence of uncertainties and unknown conditions. The chattering phenomenon (high-frequency oscillation) is one of the significant disadvantages of SMO. The main effect of this challenge is the increase of some serious mechanical obstacles such as heats the mechanical components and saturation. To decrease the chattering, the higher-order SMO (HSMO) is presented and reported in [25,26]. To increase the performance of the HSMO, different techniques, such as the quasi-continuous (QC) algorithm [27], suboptimal (SO) method [28], and twisting technique (TW) [29], have been introduced. The main challenge of the QC-HSMO, SO-HSMO, and TW-HSMO approaches is the first-order derivative of the sliding variable. To address this issue, a higher-order super-twisting (advanced) SMO (ASMO) for measurable and unmeasurable state observer has been reported [6]. Apart from the stability, robustness, and chattering attenuation in the ASMO, this method suffers from a somewhat reduced fault estimation accuracy. Therefore, in this research, the fuzzy technique is applied to the ASMO to increase the estimation accuracy and design AFSMO.

Once the observer is designed, the decision regarding the REB condition should be made. There are various conventional techniques that can be applied for the decision-making, such as decision tables, rule-based reasoning, and case-based reasoning [30]; however, recently the solutions provided by artificial intelligence (AI) are frequently applied to resolve the problems of fault diagnosis. Machine learning (ML) is one of the fields of AI that introduces some of the most popular techniques for decision-making, such as support vector machines (SVMs) [31] and artificial neural networks (ANNs) [32,33]. To make a decision on the particular faulty condition, the SVM first attempts to find an optimal hyperplane that best separates the feature parameters corresponding to data instances of different classes. Then, when the new data sample appears, the SVM determines on which side of the hyperplane the sample lays and assigns a class corresponding to the location. When an ANN is used in fault identification, the network learns the optimal weights of its neurons during the back-propagation procedure to minimize the loss function and meet the target values given a set of input attributes (i.e., feature set) corresponding to different faulty conditions. These days, the trend in AI applications have shifted toward the deep learning (DL) approach, which focuses on learning data representations to achieve the target goals. This shift is primarily enabled by the significant increase

of computational capabilities of modern computer systems. The most popular DL-based solutions used for solving different problems in condition monitoring are convolutional neural networks [34] (fault diagnosis), autoencoders [19] (fault diagnosis, feature extraction, and data augmentation), generative adversarial networks [35] (data augmentation), and recurrent neural networks [36] (fault prediction). The principles of DL-based solutions are similar to those of the ANN; they adapt the weights of neurons and tune the hyperparameters to meet the requirements according to the task. However, unlike ANNs, the deep networks are characterized by a large number of hidden layers and nodes, which necessitate the application of huge datasets to achieve a good generalization by these networks. Also, the computational time of DL-based methods significantly increases in comparison with the conventional ML-based approaches. In this work, we employ an ML-based classification technique called a decision tree (DT) [37–39] to complete the proposed fault diagnosis methodology and implement the decision-making procedure for REB fault detection and identification. During training, the conventional DT [40] algorithm relatively fastly learns and derives the logical set of easily interpretable rules that can be used for decision-making regarding the REB faulty conditions, while providing insights into the quality of the fault estimation procedure performed by the advanced fuzzy SMO (AFSMO) in the previous step.

Figure 1 illustrates the complete block diagram of the proposed algorithm for FEDI of the bearing. This figure indicates that this algorithm has three main parts. In the first step, the system is modeled using the fuzzy ARX-Laguerre (FAL) technique. This part itself has three steps: (i) modeling the bearing based on the filtered ARX method, (ii) modifying the performance of the filtered ARX technique based on an orthonormal function and designing the ARX-Laguerre method, and (iii) improving the accuracy and performance of ARX-Laguerre bearing modeling based on the fuzzy ARX-Laguerre technique. In the second step, the AFSMO is designed for accurate fault estimation and improved performance of the decision component. The second step has three sub-blocks: (i) run the SMO, (ii) reduce the chattering and increase the robustness, in which the SMO is improved based on the advanced technique and the designed ASMO, and (iii) increase the fault estimation accuracy using the fuzzy algorithm and apply it to the ASMO. Apart from the advantages of the ASMO regarding robustness and reliability, it suffers from a suboptimal fault estimation accuracy. To address this issue, a fuzzy algorithm is used in parallel with the advanced SMO for the bearings. In the third step, the faults are detected and identified based on the DT machine learning algorithm. The decision-making for fault detection and identification of REBs contain three parts applied in sequence. These parts are the (i) residual generator, (ii) window characterization, and (iii) deriving the logical decision rules for fault detection and identification using DTs. In the residual generator, the residual signals are calculated in various (normal and abnormal) conditions. Next, these residual signals are cut into windows of equal size and the amplitude-dependent feature parameter is extracted to quantitatively characterize these obtained windows. Finally, the fault detection and identification are accomplished by the logical decision rules delivered by the DT technique. Specifically, the decision about the fault condition is accomplished by comparing the value of the extracted feature parameter from the window of the residual signal with the learned threshold provided by the DT classification algorithm. This paper has the following contributions:

- I. Robust technique for modeling the vibration signals in bearings based on the fuzzy ARX-Laguerre approach is proposed.
- II. The estimation accuracy of the higher-order sliding mode observer for vibration signals has been improved by the T-S fuzzy algorithm.
- III. The performance of fault detection and identification by the proposed hybrid observer is improved by the decision tree technique and hence, new machine learning-based hybrid observer is introduced in this paper.

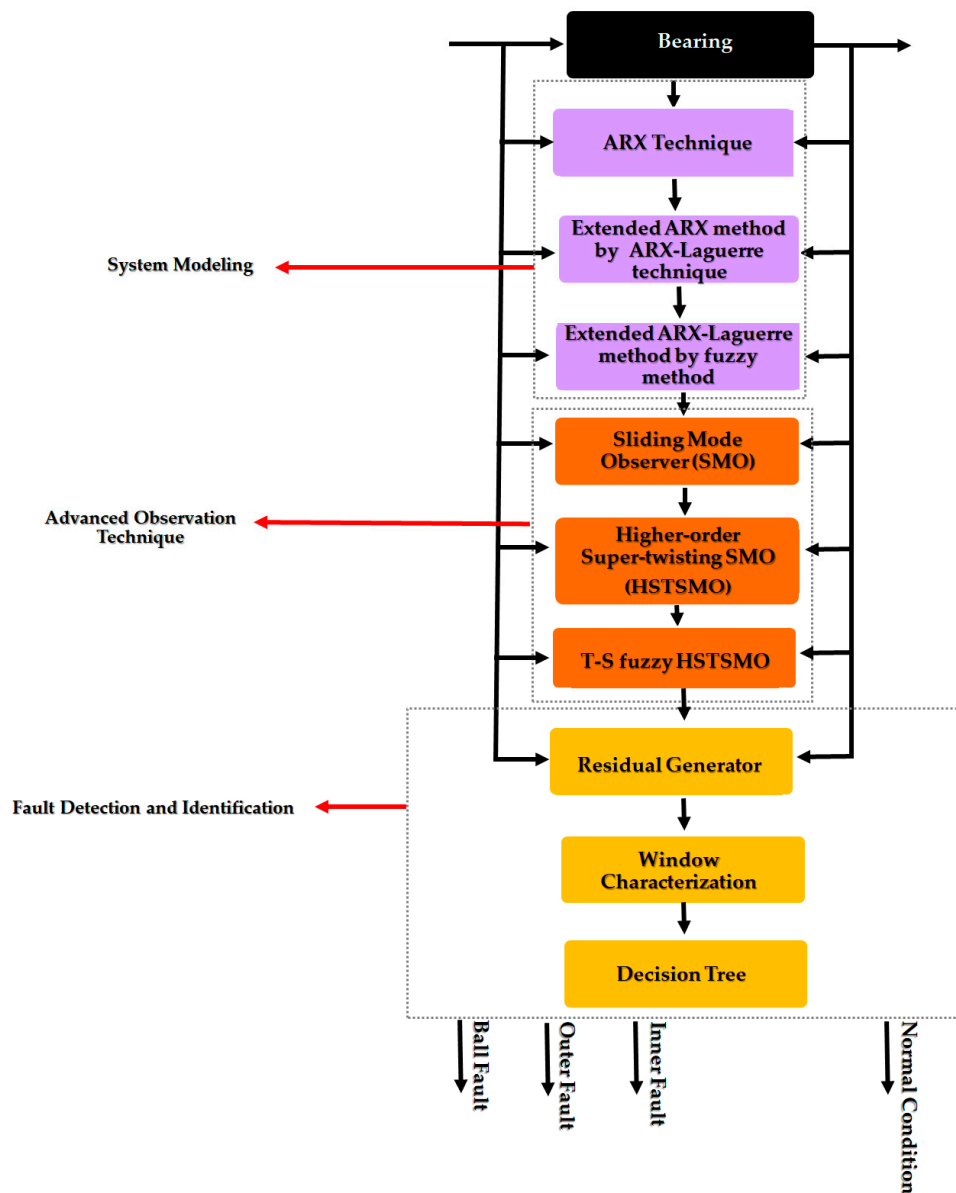


Figure 1. Proposed algorithm for FEDI (fault estimation, detection, and identification) in rolling-element bearing (REB).

The rest of this research paper is organized as follows: Section 2 provides insights into the Case Western Reserve University (CWRU) benchmark dataset used in this paper. In Section 3, the bearing is modeled based on the FAL procedure. Section 4 includes two main steps. In the first step, the AFSMO is utilized for fault estimation. In the second step, the decision tree algorithm is used for the fault detection and identification. In Section 5, fault detection, estimation, and identification results for the bearing are analyzed. Finally, the conclusions are provided in the last section.

2. Dataset

In this research, the Case Western Reserve University (CWRU) benchmark bearing dataset [2,6] is used to experimentally evaluate the effectiveness of machine learning-based AFSMO that is diagrammed in Figure 1. Figure 2 presents the experimental data acquisition system used to extract the signals in various conditions. To collect the vibration signals in normal and abnormal conditions, the vibration sensor (6205-2RS JEM SKF) is used. This sensor collects normal and single point faults that are seeded on the drive-end bearings at various bearing locations as the ball fault (BF), outer fault (OF), and inner

fault (IF), respectively. The sampling rate of the record data is 12 kHz under four different motor torques loaded under different rotation velocities from 1730 rpm to 1790 rpm. Table 1 illustrates the data description of the CWRU benchmark bearing dataset. Four different datasets are defined in this table which are categorized by motor torque loaded from 0 to 3 hp. Moreover, the vibration signals in healthy and faulty (BF, OF, IF) conditions have three different crack sizes (i.e., 0.007, 0.014, and 0.021 inches in diameter).

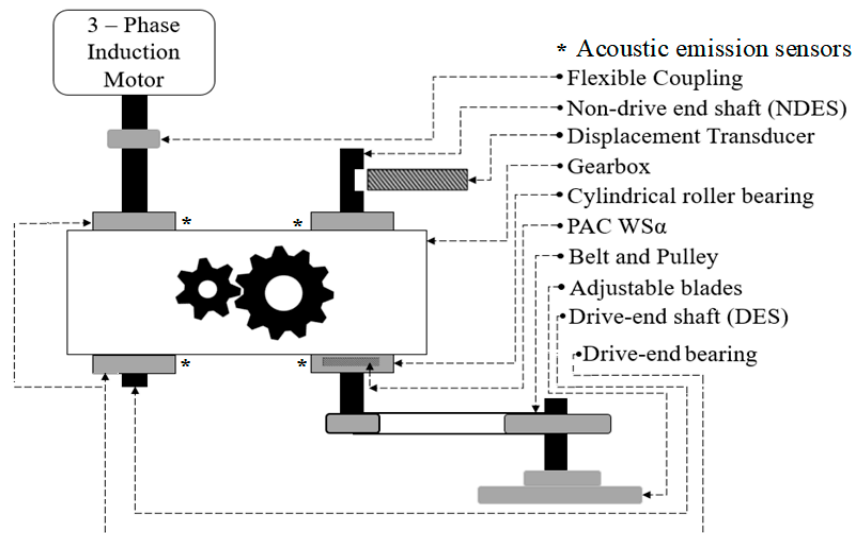


Figure 2. The system simulator’s block [22].

Table 1. The detailed description of the Case Western Reserve University (CWRU) datasets in normal and abnormal conditions.

Dataset	Conditions	Load (hp)	Crack Sizes (in)
Dataset 1	Normal	0	0.007, 0.014, and 0.021
	BF	0	
	IF	0	
	OF	0	
Dataset 2	Normal	1	0.007, 0.014, and 0.021
	BF	1	
	IF	1	
	OF	1	
Dataset 3	Normal	2	0.007, 0.014, and 0.021
	BF	2	
	IF	2	
	OF	2	
Dataset 4	Normal	3	0.007, 0.014, and 0.021
	BF	3	
	IF	3	
	OF	3	

3. Rolling-Element-Bearing Modeling

In this research, the hybrid technique is proposed for FEDI. The model-based approach is the core of the proposed method. First, the Lagrangian formulation based on potential energy, kinetic energy, and generalized forces can be expressed as the following equation [41].

$$\frac{d}{dt} \left(\frac{\partial K}{\partial \dot{q}_i} \right) - \frac{\partial K}{\partial q_i} + \frac{\partial P}{\partial q_i} = Q_i, i = 1, 2, 3, 4, \dots \tag{1}$$

Here, K, P, Q_i , and q_i are kinetic energy, potential energy, generalized force, and generalized coordinate, respectively. The energy equation can be obtained by the derivative of the Equation (1) and is expressed as follows [6,41]:

$$F_{(q)} = (A + \Delta A)(q)[\ddot{q}] + (B + \Delta B)(q)[\dot{q}] + (C + \Delta C)(q)[q] + \varphi, \tag{2}$$

where $F_{(q)}, A(q), B(q), C(q), \varphi$, and $(\Delta A, \Delta B, \Delta C)$ are the force vector, mass vector, time-variant stiffness matrix, time-variant damping matrix, fault (IF, OF, BF) vectors, and unknown modeling parameters for mass, stiffness, and damping matrix, respectively. If $\Delta = \Delta A(q)[\ddot{q}] + \Delta B(q)[\dot{q}] + \Delta C(q)[q]$, the dynamic equation of the bearings can be represented as follows:

$$F_{(q)} = A(q)[\ddot{q}] + B(q)[\dot{q}] + C(q)[q] + \Delta + \varphi, \tag{3}$$

If $\omega(q, \dot{q}) = B(q)[\dot{q}] + C(q)[q]$ and $\psi(q, \dot{q}) = A^{-1}(q)(\Delta + \varphi)$ represent the uncertainties and faults, the bearing dynamic equation is rewritten as follows:

$$[\ddot{q}] = A(q)^{-1} \{ F_{(q)} - \omega(q, \dot{q}) \} - \psi(q, \dot{q}). \tag{4}$$

Apart from several advantages of mathematical-based system modeling, in most complicated systems, such as bearing systems, the precise mathematical formulation of energy and force in the bearing is nonlinear and complicated. In addition, mathematical modeling is not accurate in the presence of uncertainty. Moreover, the dynamic behavior of the bearing in theoretical and practical applications may be different, which causes challenges in system modeling for FEDI. Therefore, the fuzzy ARX-Laguerre (FAL) technique is represented for REB modeling. This system modeling technique has three main steps. In the first step, the ARX system modeling is defined. To improve the robustness and reliability of the system modeling, the ARX-Laguerre technique is represented in the second step. In addition, to improve the system’s modeling accuracy for the ARX-Laguerre technique, the fuzzy technique is represented. The mathematical formulation for an ARX system model is represented as [20–22]:

$$f_o(k) = \sum_{i=1}^{\delta_x} \varepsilon_x(i) f_o(k-i) + \sum_{i=1}^{\delta_y} \varepsilon_y(i) f_i(k-i), \tag{5}$$

where $f_o(k), (\varepsilon_x(i), \varepsilon_y(i)), f_i(k)$, and (δ_x, δ_y) are the output, model parameters, input, and order of the system, respectively. To represent the model parameters $(\varepsilon_x, \varepsilon_y)$, the following equation is presented.

$$\begin{cases} \varepsilon_x(i) = \sum_{n=0}^{\infty} \omega_{n,x} J_n^x(i, \gamma_x) \\ \varepsilon_y(i) = \sum_{n=0}^{\infty} \omega_{n,y} J_n^y(i, \gamma_y) \end{cases} \tag{6}$$

Here, $(\omega_{n,x}, \omega_{n,y}), (\gamma_x, \gamma_y)$, and (J_n^x, J_n^y) are the coefficients of the Fourier decomposition, orthonormal basis, and orthonormal function, respectively. In the next step, ARX-Laguerre system modeling is used to obtain robust system modeling. The input and output orthonormal functions are represented in Equation (7).

$$\begin{cases} O_{n,f_o}(k, \gamma_x) = \sum_{i=1}^{\infty} J_n^x(i, \gamma_x) f_o(k-i) \\ O_{n,f_i}(k, \gamma_y) = \sum_{i=1}^{\infty} J_n^y(i, \gamma_y) f_i(k-i) \end{cases} \tag{7}$$

Here, O_{n,f_o} and O_{n,f_i} are the filtered orthonormal functions for the output and input, respectively. The ARX orthonormal function is represented as the following equation.

$$f_o(k) = \sum_{n=0}^{\delta_x-1} \omega_{n,x} O_{n,f_o}(k, \gamma_x) + \sum_{n=0}^{\delta_y-1} \omega_{n,y} O_{n,f_i}(k, \gamma_y), \tag{8}$$

In addition, the Laguerre technique is defined as Equation (9).

$$\begin{cases} L_n^x = \frac{\sqrt{1-\gamma_x^2}}{Z-\gamma_x} \left(\frac{1-\gamma_x z}{Z-\gamma_x} \right)^n \\ L_n^y = \frac{\sqrt{1-\gamma_y^2}}{Z-\gamma_y} \left(\frac{1-\gamma_y z}{Z-\gamma_y} \right)^n \end{cases} \tag{9}$$

Here, L_n^x and L_n^y are the input and output Laguerre functions, respectively. Based on Equations (7) and (9), the input and output orthonormal functions are modified to be the following equation.

$$\begin{cases} \bar{O}_{n,f_o}(k, \gamma_x) = \sum_{j=1}^{\infty} L_n^x(k) * f_o(k) \\ \bar{O}_{n,f_i}(k, \gamma_y) = \sum_{j=1}^{\infty} L_n^y(k) * f_i(k) \end{cases} \tag{10}$$

Here, \bar{O}_{n,f_o} and \bar{O}_{n,f_i} are the input and output modified orthonormal functions, respectively. Thus, the ARX-Laguerre system modeling and estimation is represented as the following [20].

$$\begin{aligned} f_o(k) &= \sum_{n=0}^{i_x-1} \omega_{n,x} \bar{O}_{n,f_o}(k, \gamma_x) W(k) + \sum_{n=0}^{i_y-1} \omega_{n,y} \bar{O}_{n,f_i}(k, \gamma_y) I(k) \\ f_o(k) &= \Psi(z^{-1})W(z) + \Lambda(z^{-1})I(z) \end{aligned} \tag{11}$$

Here, $(\Psi(z^{-1}), \Lambda(z^{-1}))$ and $(W(k), I(k))$ are the polynomial variables and filtering signals, respectively and defined as the following equations.

$$\begin{aligned} \Psi(z^{-1}) &= \sum_0^{n_x-1} \omega_{n,x} \bar{O}_{n,f_o}(k) \\ \Lambda(z^{-1}) &= \sum_0^{n_y-1} \omega_{n,y} \bar{O}_{n,f_i}(k) \end{aligned} \tag{12}$$

Therefore, the state-space ARX-Laguerre technique is represented as Equation (13).

$$\begin{cases} J_s(k+1) = [\sigma_s J_s(k) + \sigma_o J_o(k) + \sigma_i J_i(k)] + \varphi(k) \\ J_o(k+1) = (\lambda)^T J_s(k) \end{cases} \tag{13}$$

Here, $J_s(k), (\sigma_s, \sigma_o, \sigma_i), J_o(k), J_i(k), \varphi(k)$, and $(\lambda)^T$ are the system's state, modeling coefficients, measured output, input, uncertainties and fault, and Fourier coefficients, respectively. Here, the state system modeling coefficient σ_s is represented by the following equation:

$$\sigma_s = \begin{bmatrix} \sigma_{s_o} & O_{N_x, N_y} \\ O_{N_y, N_x} & \sigma_{s_i} \end{bmatrix} \tag{14}$$

where O_{N_x, N_y} and O_{N_y, N_x} are null matrices and σ_{s_o} and σ_{s_i} are represented in Equations (15) and (16), respectively.

$$\sigma_{s_o} = \begin{bmatrix} \gamma_x & \cdots & 0 \\ \vdots & \ddots & \vdots \\ (-\gamma_x)^{N_x-1}(1-\gamma_x^2) & \cdots & \gamma_x \end{bmatrix}, \tag{15}$$

$$\sigma_{s_i} = \begin{bmatrix} \gamma_y & \cdots & 0 \\ \vdots & \ddots & \vdots \\ (-\gamma_y)^{N_y-1}(1-\gamma_y^2) & \cdots & \gamma_y \end{bmatrix}, \tag{16}$$

The output system modeling coefficient σ_o is represented in Equation (17).

$$\sigma_o = \sqrt{1-\gamma_x^2} \begin{bmatrix} 1 \\ -\gamma_x \\ -\gamma_x^2 \\ \vdots \\ (-\gamma_x)^{N_x-1} \end{bmatrix}, \tag{17}$$

In addition, the input system modeling coefficient σ_i can be represented in Equation (18).

$$\sigma_i = \sqrt{1-\gamma_y^2} \begin{bmatrix} 1 \\ -\gamma_y \\ -\gamma_y^2 \\ \vdots \\ (-\gamma_y)^{N_y-1} \end{bmatrix}, \tag{18}$$

To improve the ARX-Laguerre system modeling accuracy, the fuzzy technique is recommended in this research. The fuzzy ARX-Laguerre (FAL) system modeling is defined as follows.

$$\begin{cases} J_s(k+1) = [\sigma_s J_s(k) + \sigma_o J_o(k) + \sigma_i J_i(k) + \sigma_f J_f(k)] + \varphi(k) \\ J_o(k+1) = (\lambda)^T J_s(k) \end{cases}, \tag{19}$$

Here, σ_f and J_f are the fuzzy coefficient and fuzzy function for system estimation, respectively. The fuzzy if-then rule in this research is defined by the following rule.

$$\text{If } \langle \text{first input} \rangle \text{ is } \langle \text{linguistic variable for first input} \rangle \text{ and } \langle \text{second input} \rangle \text{ is } \langle \text{linguistic variable for second input} \rangle \text{ then } \langle \text{output} \rangle \text{ is } \langle \text{linguistic variable for output} \rangle \tag{20}$$

The membership functions of a fuzzy set for the system estimation of error (e) in the interval of $[-0.3, 0.3]$ are the Gaussian and the linguistic variables, which are defined by negative high (NH), negative medium (NM), negative low (NL), zero (Z), positive low (PL), positive medium (PM), and positive high (PH). The fuzzy membership functions for the system estimation of the change of error (\dot{e}) in the interval of $[-0.1, 0.1]$ are the Gaussian and the linguistic variables, which are defined by NH, NM, NL, Z, PL, PM, and PH. In addition, the fuzzy linguistic variables for J_f in the interval of $[-0.08, 0.08]$ are the Gaussian and the fuzzy sets, which are defined by NH, NM, NL, Z, PL, PM, and PH. Table 2 illustrates the fuzzy rule table for the FAL system estimation technique. According to this table, the error has seven linguistic variables, the change of error has seven linguistic variables and the fuzzy system estimation has seven linguistic variables. Therefore, the fuzzy technique to improve system modeling has 49 rule-bases. Based on these 49 rule-bases, the fuzzy technique has improved the accuracy of system modeling to achieve the minimum estimation error.

Table 2. The fuzzy rule table for system modeling based on the fuzzy ARX-Laguerre (FAL) technique.

		Change of Error (\dot{e})						
		NH	NM	NL	Z	PL	PM	PH
Error (e)	NH	PH	PH	PH	PH	PM	PL	Z
	NM	PH	PH	PH	PM	PL	Z	NL
	NL	PH	PH	PM	PL	Z	PL	PM
	Z	PH	PM	PL	Z	NL	NM	NH
	PL	PM	PL	Z	NL	NM	NH	NH
	PM	PL	Z	NL	NM	NH	NH	NH
	PH	Z	NL	NM	NH	NH	NH	NH

Figures 3 and 4 show the estimation accuracy and errors of REB modeling for the normal and faulty conditions based on the proposed FAL technique, ARX-Laguerre system modeling technique, and ARX system estimation method. Based on these figures, the system modeling accuracy in the proposed fuzzy ARX-Laguerre system modeling is higher than the ARX-Laguerre technique and ARX system modeling techniques. Regarding Figure 3, the error rate in the proposed fuzzy ARX-Laguerre technique is close to zero.

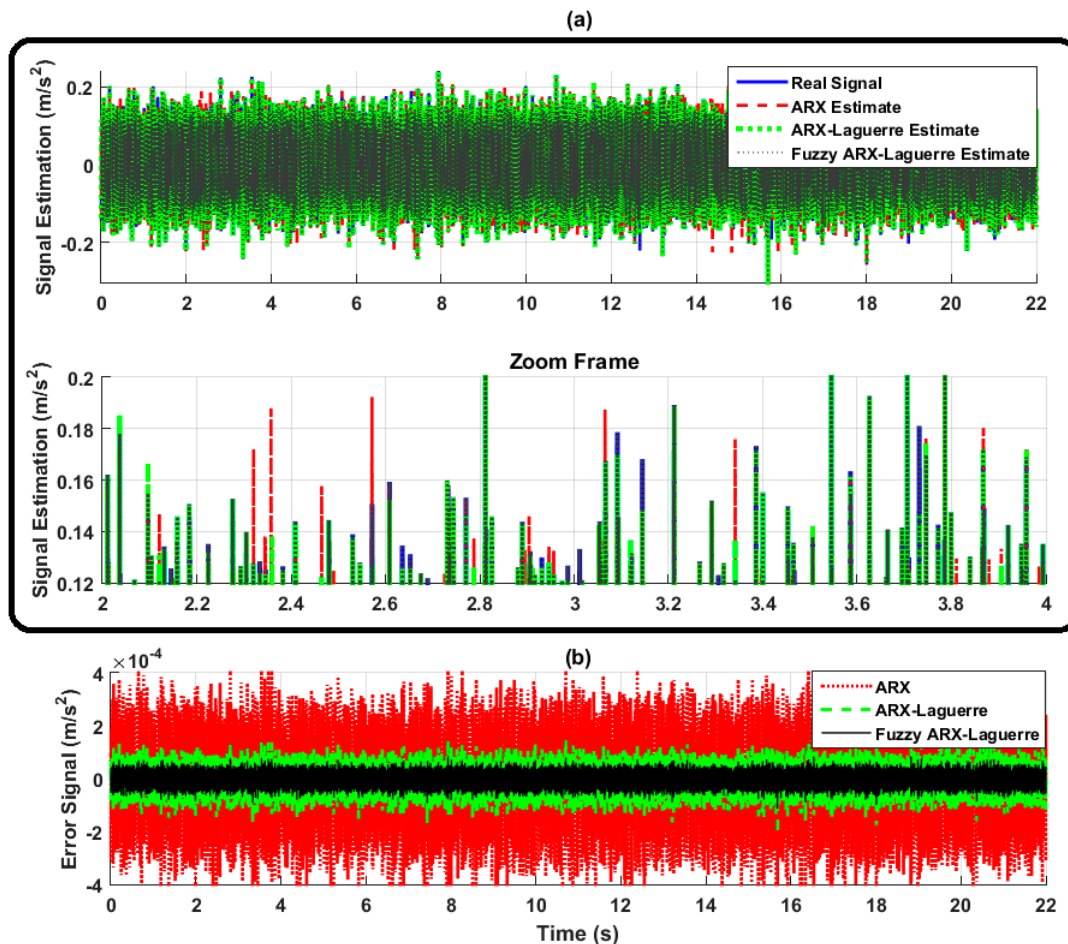


Figure 3. The estimation accuracies in the normal condition: (a) the real and estimated signal, and (b) the estimation signal’s error.

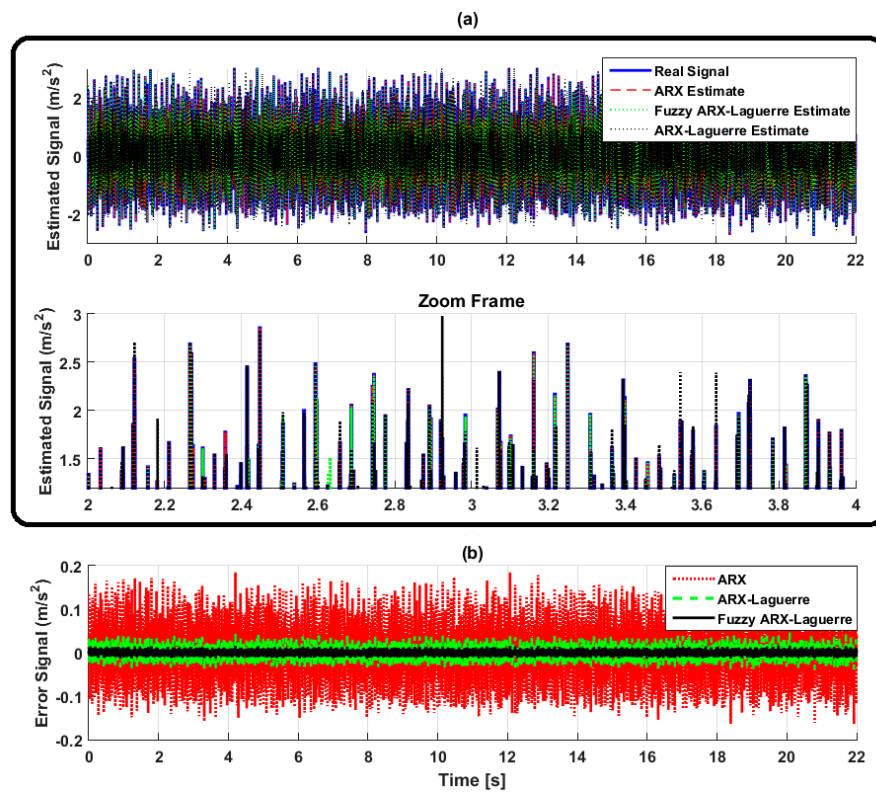


Figure 4. The estimation accuracies in the abnormal condition. (a) Real and estimated signals by three techniques, and (b) the estimation signal’s error by three techniques.

4. Proposed Algorithm for Fault Diagnosis in Rolling-Element Bearing

According to Figure 1, the fuzzy ARX-Laguerre method was used to model the bearing in normal and abnormal conditions. The next step focuses on the designed advanced fuzzy sliding mode observation technique that comprises three main blocks: (a) apply SMO for fault estimation, (b) improve the robustness and chattering attenuation based on the high-order super-twisting (advance) SMO (ASMO), and (c) improve the accuracy of the fault estimation technique, i.e., the T-S fuzzy algorithm applied to the ASMO.

4.1. Advance Sliding Mode Observer for Fault Estimation

Based on Equation (19) for FAL modeling of the REB, the SMO is proposed based on Equation (21).

$$\begin{cases} \hat{J}_{s-SMO}(k+1) = [\sigma_s \hat{J}_{s-SMO}(k) + \sigma_o \hat{J}_{o-SMO}(k) + \sigma_i J_i(k)] + \hat{\phi}_{SMO}(k) + \sigma_p (J_o(k) - \hat{J}_{o-SMO}(k)) + \sigma_{SM} \cdot \text{sgn}(J_o(k) - \hat{J}_{o-SMO}(k)) \\ \hat{J}_{o-SMC}(k+1) = (\lambda)^T \hat{J}_{s-SMO}(k) \end{cases}, \quad (21)$$

In addition, the equation to define fault estimation based on the SMO is:

$$\hat{\phi}_{SMO}(k+1) = \hat{\phi}_{SMO}(k) + \sigma_p (J_o(k) - \hat{J}_{o-SMO}(k)) + \sigma_{SM} \cdot \text{sgn}(J_o(k) - \hat{J}_{o-SMO}(k)). \quad (22)$$

where $\hat{J}_{s-SMO}(k)$, $\hat{J}_{o-SMO}(k)$, $\hat{\phi}_{SMO}(k)$, and $(\sigma_s, \sigma_o, \sigma_i, \sigma_p, \sigma_{SM})$ are the sliding mode system state estimation, sliding mode output estimation, sliding mode fault estimation, and coefficients, respectively. The SMO is a robust and stable technique for fault diagnosis in bearings, however, this technique suffers from the chattering phenomenon. To reduce the effect of the chattering, the following function is introduced:

$$Z = \sigma_x \|J_o(k) - \hat{J}_{o-SMO}(k)\|^{0.5} \text{sgn}(J_o(k) - \hat{J}_{o-SMO}(k)), \quad (23)$$

where Z and σ_x are new the observation definition and coefficient, respectively. If the unknown conditions are estimated $(\Delta + \varphi)$, the SMO error goes to zero in a finite time.

$$\begin{cases} \sigma_x \|J_o(k) - \hat{J}_{o-SMO}(k)\|^{0.5} \operatorname{sgn}(J_o(k) - \hat{J}_{o-SMO}(k)) - \dot{\hat{\delta}} \\ \dot{\hat{\delta}} = \sigma_{x_0} \times \operatorname{sgn}(J_o(k) - \hat{J}_o(k)) \end{cases}, \quad (24)$$

Here, σ_{x_0} and $\dot{\hat{\delta}}$ are the respective coefficient and the variable for the super-twisting variable. Therefore, based on Equation (23), this technique can improve the performance of fault estimation in the unknown condition. Therefore, to reduce the chattering phenomenon, the ASMO state-space and the fault estimation equation can be represented in Equations (25) and (26).

$$\begin{cases} \hat{J}_{s-HSMO}(k+1) = [\sigma_s \hat{J}_{s-HSMO}(k) + \sigma_o \hat{J}_{o-HSMO}(k) + \sigma_i J_i(k)] + \hat{\varphi}_{HSMO}(k) + \\ \sigma_p (J_o(k) - \hat{J}_{o-HSMO}(k)) + \sigma_{HSM} \cdot \operatorname{sgn} \left| (J_o(k) - \hat{J}_{o-HSMO}(k)) \right|^{\frac{2}{3}} \\ \hat{J}_{o-HSMO}(k+1) = (\lambda)^T \hat{J}_{s-HSMO}(k) \end{cases}, \quad (25)$$

$$\begin{cases} \hat{\varphi}_{HSMO}(k+1) = \hat{\varphi}_{HSMO}(k) + \sigma_p (J_o(k) - \hat{J}_{o-HSMO}(k)) + \sigma_{HSM} \times \operatorname{sgn} \\ (J_o(k) - \hat{J}_{o-HSMO}(k)) + \sigma_x \|J_o(k) - \hat{J}_{o-HSMO}(k)\|^{0.5} \operatorname{sgn}(J_o(k) - \hat{J}_{o-HSMO}(k)) - \dot{\hat{\delta}} \\ \dot{\hat{\delta}} = \sigma_{x_0} \times \operatorname{sgn}(J_o(k) - \hat{J}_o(k)) \end{cases}. \quad (26)$$

Here, $\hat{J}_{s-HSMO}(k)$, $\hat{J}_{o-HSMO}(k)$, $\hat{\varphi}_{HSMO}(k)$, and (σ_x, σ_{HSM}) are the advanced sliding mode system state estimation, advanced sliding mode output estimation, advanced sliding mode fault estimation, and coefficients, respectively. To find the sliding gain, the estimation error is calculated based on the following equation.

$$\begin{cases} \tilde{J}_{s-HSMO}(k+1) = [\sigma_s \tilde{J}_{s-HSMO}(k) + \sigma_o \tilde{J}_{o-HSMO}(k)] + \tilde{\varphi}_{HSMO}(k) + \\ \sigma_p (J_o(k) - \tilde{J}_{o-HSMO}(k)) + \sigma_{HSM} \cdot \operatorname{sgn} \left| (J_o(k) - \tilde{J}_{o-HSMO}(k)) \right|^{\frac{2}{3}} \\ \tilde{J}_{o-HSMO}(k+1) = (\lambda)^T \tilde{J}_{s-HSMO}(k) \\ \tilde{J}_{s-HSMO}(k) = J_s(k) - \hat{J}_{s-HSMO}(k) \\ \tilde{J}_{o-HSMO}(k) = J_o(k) - \hat{J}_{o-HSMO}(k) \end{cases}. \quad (27)$$

Here, $\tilde{J}_{s-HSMO}(k)$, $\tilde{J}_{o-HSMO}(k)$, and $\tilde{\varphi}_{HSMO}(k)$ are the state-estimation error, measured output estimation error, and fault estimation error based on the advanced SMO, respectively. Therefore, to guarantee the stability and robustness, the sliding surface slope coefficients are represented by the following equations [6].

$$\begin{cases} \sigma_{x_0} = 1.1 \times \Gamma \\ \sigma_{HSM} = 1.9 \times \sqrt[3]{\Gamma} \\ \sigma_x = 1.5 \times \sqrt{\Gamma} \end{cases}. \quad (28)$$

Here, Γ is an estimation error bounded and based on Equation (26), it can be determined by the nonlinear part of the estimation error performance. Though there is an improvement in the robustness and the attenuation of the chattering based on the ASMO for FEDI, this fails to improve the fault estimation accuracy. Therefore, the T-S fuzzy algorithm is used to improve the ASMO. The T-S fuzzy technique is represented based on the following equation [42].

$$\alpha_f(k) = \begin{cases} \text{if } r_x(k) \text{ is } TH_n \text{ then } \alpha_f(k+1) = \alpha_f(k) + \zeta_n (J_o(k) - \hat{J}_{o-fHSMO}(k)) \\ \text{if } r_x(k) \text{ is } TH_f \text{ then } \alpha_f(k+1) = \alpha_f(k) + \zeta_f (J_o(k) - \hat{J}_{o-fHSMO}(k)) \end{cases}. \quad (29)$$

Here, $r_x(k)$, α_f , TH_n , TH_f , and (c_n, c_f) are the residual signal, fuzzy estimation function, threshold value for the normal condition, threshold value for the faulty condition, and normal and abnormal coefficients for the T-S fuzzy observer, respectively. Therefore, the proposed advanced fuzzy SMO and the fault estimation are represented in Equations (30) and (31).

$$\begin{cases} \hat{J}_{s-fHSMO}(k+1) = [\sigma_s \hat{J}_{s-fHSMO}(k) + \sigma_o \hat{J}_{o-fHSMO}(k) + \sigma_i J_i(k)] + \hat{\phi}_{fHSMO}(k) + \\ \sigma_p (J_o(k) - \hat{J}_{o-fHSMO}(k)) + \sigma_{fHSM} \cdot \text{sgn} \left| (J_o(k) - \hat{J}_{o-fHSMO}(k)) \right|^{\frac{2}{3}} + \alpha_f(k) \\ \hat{J}_{o-fHSMO}(k+1) = (\lambda)^T \hat{J}_{s-fHSMO}(k) \end{cases} \quad (30)$$

$$\begin{cases} \hat{\phi}_{fHSMO}(k+1) = \hat{\phi}_{fHSMO}(k) + \sigma_p (J_o(k) - \hat{J}_{o-fHSMO}(k)) + \sigma_{fHSM} \times \text{sgn} \\ (J_o(k) - \hat{J}_{o-fHSMO}(k)) + \sigma_x \|J_o(k) - \hat{J}_{o-fHSMO}(k)\|^{0.5} \text{sgn}(J_o(k) - \hat{J}_{o-fHSMO}(k)) - \hat{\delta} \\ \dot{\hat{\delta}} = \sigma_{x0} \times \text{sgn}(J_o(k) - \hat{J}_{o-fHSMO}(k)) \end{cases} \quad (31)$$

Here, $\hat{J}_{s-fHSMO}$, $\hat{J}_{o-fHSMO}$, $\hat{\phi}_{fHSMO}$, and σ_{fHSM} are the advanced fuzzy sliding mode system state estimation, advanced fuzzy sliding mode output estimation, advanced fuzzy sliding mode fault estimation, and coefficients, respectively. After estimating the output based on the proposed AFSMO, the residual signal can be calculated as follows:

$$r_x(k) = J_o(k) - \hat{J}_{o-fHSMO}(k). \quad (32)$$

Figures 5–7 illustrate the residual signal in normal and abnormal conditions based on the SMO, ASMO, and AFSMO (the proposed method). Based on these figures, the difference between various states of signals in the AFSMO is more clear than those of the ASMO and SMO. In the next part, the procedure of fault detection and identification is performed and described using the decision tree machine learning technique.

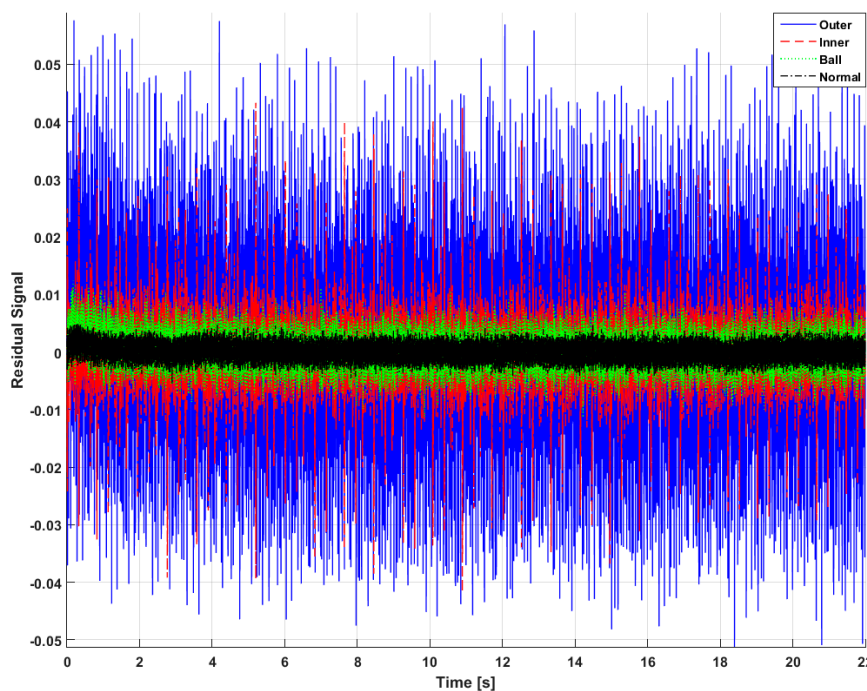


Figure 5. Residual signal based on sliding mode observer (SMO).

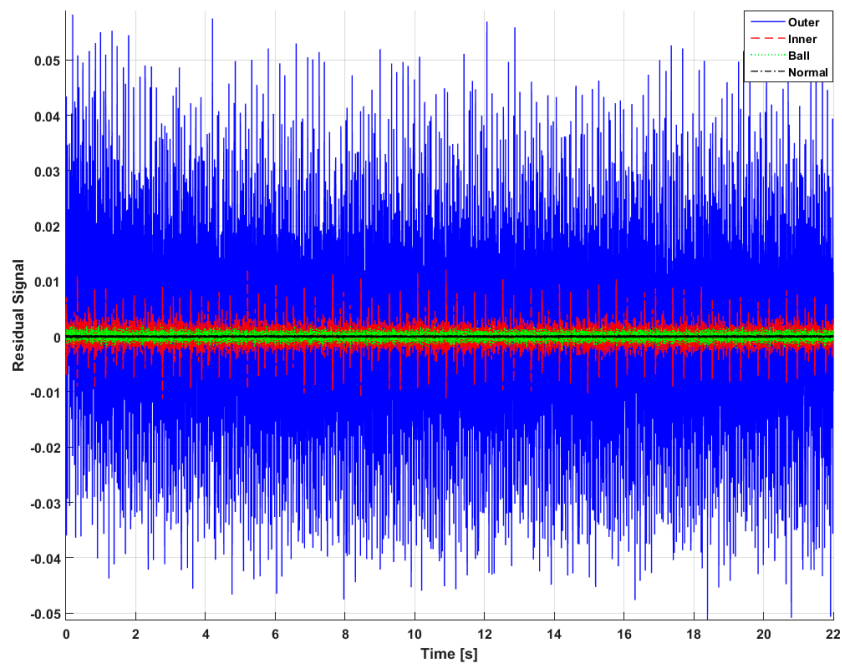


Figure 6. Residual signal based on advanced SMO (ASMO).

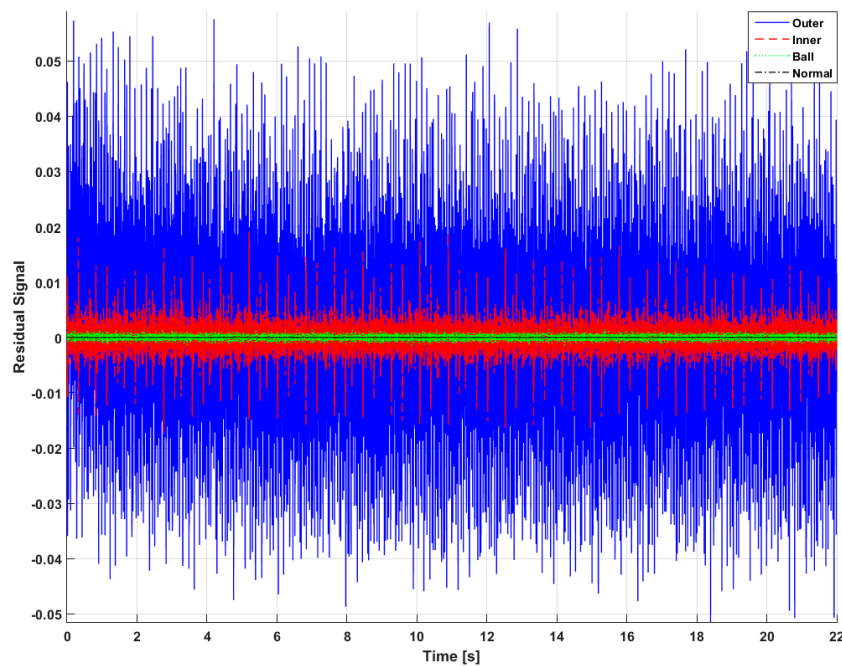


Figure 7. Residual signal based on advanced fuzzy SMO (AFSMO).

4.2. Fault Detection and Fault Identification Based on Decision Tree Technique

4.2.1. Residual Signal Characterization

Once the residual signal is obtained based on Equation (32), this signal can be successfully used to perform the fault detection and diagnosis process. To apply the DT approach for the problem of fault diagnosis, the specific numerical feature parameter sensitive to the changing signal conditions is essential. Because of the fact that the residual signal itself represents an error between the original signal and that estimated by the proposed observer, this signal cannot be quantified by most of the

conventional feature parameters that are used in fault diagnosis. However, the parameters that are sensitive to the amplitude can be applied to characterize the residual (i.e., the error) signal.

To characterize the residual signal by a numerical parameter, first, we split the time sequence into windows of equal size [22]. Then, the feature parameter called the energy of the signal is extracted from these windows to deliver a numerical value that can be used to characterize “the amount of error” in the particular window. The formulation of the energy feature parameter is provided below:

$$E = \sum_{i=1}^N r_{x_i}^2. \tag{33}$$

Here, r_{x_i} is the i th sample of the residual signal and N is the total number of samples. When the different types of mechanical faults appear in REBs, the amplitudes of the vibration signals change drastically. This signal behavior strongly affects the residues as well. Thus, the use of the energy as a numerical feature for characterizing the windows of the residual signals is a reasonable choice because the value of the energy is closely related to the amplitude of the sequence being investigated. Moreover, since the residue represents the error between the two signals, the values of its samples can be both positive and negative. The application of the energy feature allows us to consider the error values located on both sides from zero because of the squared term presented in Equation (33). Figure 8 shows the residual signal characterization based on the energy for normal and faulty conditions. As can be seen from this figure, the extracted feature parameters corresponding to different system’s conditions are well separable in 1D feature space which means that the proposed methodology can be efficiently applied for purposes of REB fault diagnosis.

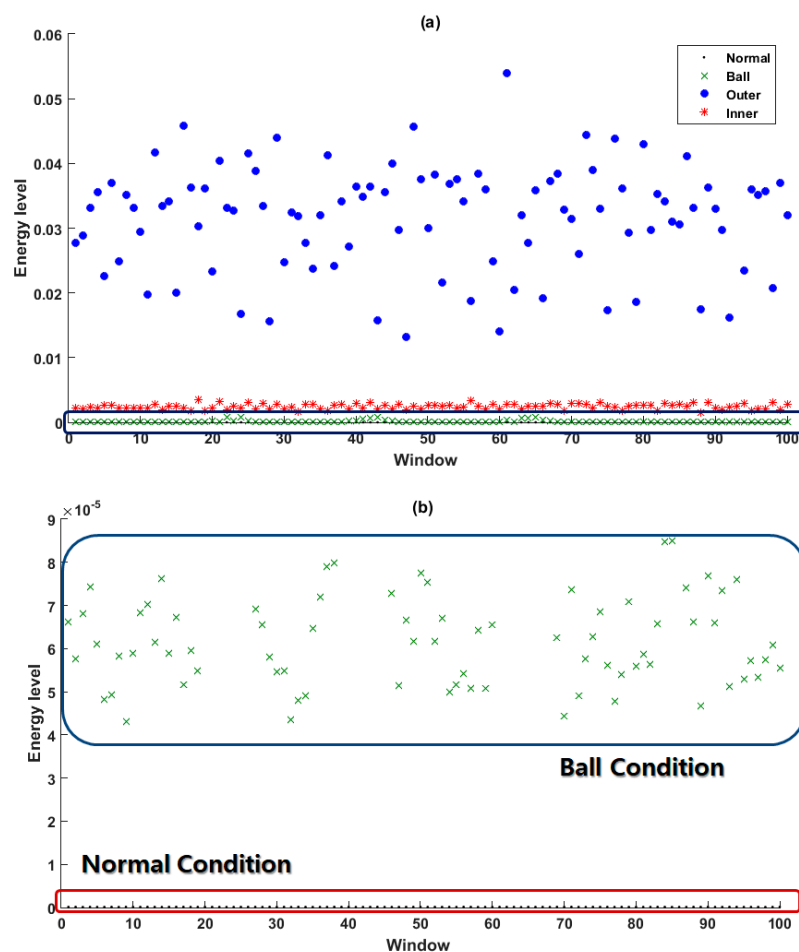


Figure 8. The energy of the residual signal based on the AFSMO: (a) original, (b) zoom frame.

4.2.2. Decision Tree-Based Fault Diagnosis

Once the windows of residual signals have been characterized using the energy feature parameter, these feature parameters are fed into the classifier, which is made using the DT. DT [40], as well as random forests (i.e., the ensemble of DTs) [43,44], is a type of machine learning technique that can perform both the classification and regression tasks while their decisions are fairly intuitive and easy to interpret. Because of their properties, decision trees are considered as white-box models; they automatically derive general classification rules that work well for the data given in the training set based on the specific attributes (i.e., feature parameters). Moreover, these obtained decision rules can be efficiently used manually for classifying different samples if necessary.

During training of the DT to perform fault detection and identification, the heuristic criteria that is based on the error and is called the Gini Diversity Index (GDI) [45], also known in the literature as the Gini Impurity, is applied in this paper. The main advantage of employing the GDI criterion is that the DTs trained with this metric tends to isolate the most frequent data class in its branch of the tree which is useful for evaluating the quality of features in training set and their separability.

The GDI is defined as follows:

$$GDI_i = 1 - \sum_{k=1}^n P_{i,k}^2. \tag{34}$$

where GDI_i is the Gini score for the i th leaf of the DT, n is the total number of classes (i.e., bearing faults) presented in data, and $p_{i,k}$ is the ratio of the instances belonging to class k among all instances in the i th leaf.

The training procedure of the DT can be summarized as follows. Given a specific feature parameter and GDI criterion, the training goal is to find a split for the features that induces a binary partition of the set of data samples with a minimum GDI criterion, where the weights of this criterion are given by the number of data samples that lie in each of the two branches [46]. The leaf node is said to be “pure” when all training samples it applies to belong to the same class. Finally, after the training of the DT with the energy values extracted from the windows of residual signals is completed, we obtain a set of rules for differentiating various bearing conditions. This set of rules is based on the threshold values automatically learned by the DT to minimize the weighted sum of Gini scores of the DT leaves. After completing the training process, the new data sample can be simply classified by tracing out a route from the root to one of the leaves of the tree while comparing the energy value with the learned threshold values for each of the leaves. In this paper, the number of assigned leaves is similar to the number of signal classes presented in the dataset, which is equal to four. Equations (35) and (36) are used to formulate the respective fault detection and fault diagnosis based on the DT algorithm.

$$\begin{cases} \text{if } E < \zeta_n \rightarrow \text{Normal} \\ \text{if } E \geq \zeta_n \rightarrow \text{Faulty} \end{cases} \tag{35}$$

$$\begin{cases} \text{if } E \geq \zeta_n \& E < \zeta_b \& E < \zeta_i \rightarrow \text{Ball Fault} \\ \text{if } E \geq \zeta_n \& E \geq \zeta_b \& E < \zeta_i \rightarrow \text{Inner Fault} \\ \text{if } E \geq \zeta_n \& E \geq \zeta_b \& E \geq \zeta_i \rightarrow \text{Outer Fault} \end{cases} \tag{36}$$

Here, ζ_n , ζ_b , and ζ_i are the normal condition threshold, ball fault condition threshold, and inner fault condition threshold, respectively. The proposed DT-based AFSMO for bearing FEDI is summarized in Algorithm 1, and an example of the DT trained using the energy values extracted from the residual signal windows for 0.021 inches crack size condition under 3 hp torque load is presented in Figure 9.

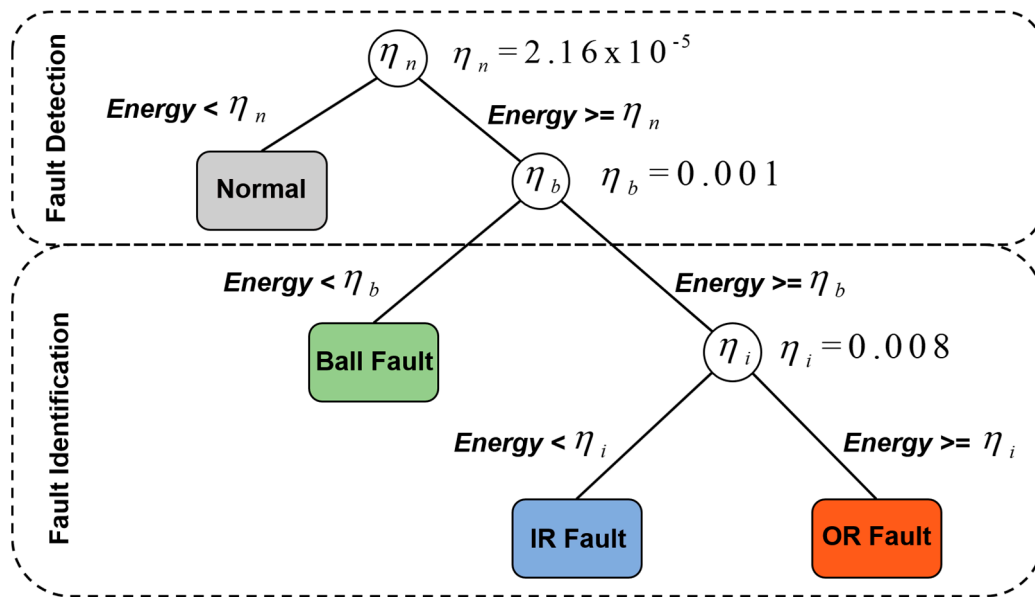


Figure 9. The decision tree trained for detecting and diagnosing various bearing faults using the energy values of the residual signal windows for 0.021 inches crack size condition under 3 hp torque load.

Algorithm 1. The AFSMO for FEDI in the bearing.

- 1: Perform system modeling based on the fuzzy ARX-Laguerre technique (19)
 - 2: Run the SMO (21, 22)
 - 3: Run the ASMO to reduce the chattering phenomenon (25, 26)
 - 4: Run the AFSMO to increase the fault estimation (30, 31)
 - 5: Run the residual generator (32)
 - 6: Run the residual signal characterization by energy (33)
 - 7: Run the learning process of decision trees (34)
 - 8: Apply the classification rules delivered by decision trees for fault detection and identification (35, 36)
-

5. Experimental Results

In this section, the fault diagnosis capabilities of the proposed technique are validated against two state-of-the-art approaches that have reported results for the same publicly available dataset. Specifically, the first employed method is the ASMO and the second approach used for the comparison is the SMO. The final fault diagnosis performance is expressed in terms of the average classification accuracy (ACA) as follows:

$$ACA = \frac{\sum_{m=1}^4 TP_m}{N_{samples}} \times 100\% \tag{37}$$

where T_p is the number of true positive predictions (i.e., the number of data samples of the specific class m correctly identified as ones belonging to the class m), and $N_{samples}$ is the total number of samples available in the particular dataset. To provide a comprehensive performance evaluation, the two case-studies are considered in this result sections: evaluating fault diagnosis accuracies on crack-variant datasets provided by CWRU and the created custom load-variant datasets.

5.1. Crack-Variant Datasets

In this subsection, we investigate the fault identification capabilities of the proposed methodology on four crack-variant datasets, where the torque load remained fixed. In other words, each of the four datasets corresponding to the particular load levels contains three subsets, where each of the subsets is formed using the data instances collected under the specific crack size. In this experiment, each subset

comprises of $N_f \times N_w$ data samples, where N_f is the number of faulty conditions available in the CWRU benchmark dataset and N_w is the number of windows (equal to 100, Section 4.2) cut from the residual signals delivered by the AFSMO in the previous step. The fault diagnosis results obtained for the REB datasets with various crack diameters under fixed load conditions are tabulated in Tables 3–6.

Table 3. The accuracy of the FEDI when the torque load is 0 hp.

Algorithms	AFSMO (Proposed Method)			ASMO			SMO		
Crack Diameters (in)	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Normal	100%	100%	100%	100%	100%	100%	100%	100%	100%
BF	100%	100%	100%	93%	94%	94%	89%	91%	92%
OF	84%	100%	99%	81%	92%	92%	80%	88%	90%
IF	96%	100%	99%	94%	90%	96%	88%	91%	94%
ACA	95%	100%	99.5%	91.25%	94%	95.5%	89.25%	92.5%	94%

Table 4. The accuracy of the FEDI when the torque load is 1 hp.

Algorithms	AFSMO (Proposed Method)			ASMO			SMO		
Crack Diameters (in)	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Normal	100%	100%	100%	100%	100%	100%	98%	98%	98%
BF	100%	99%	100%	96%	90%	98%	90%	88%	92%
OF	100%	100%	100%	94%	94%	95%	90%	90%	90%
IF	100%	92%	100%	95%	98%	98%	88%	90%	89%
ACA	100%	97.75%	100%	96.25%	95.5%	97.75%	91.5%	91.5%	92.25%

Table 5. The accuracy of the FEDI when the torque load is 2 hp.

Algorithms	AFSMO (Proposed Method)			ASMO			SMO		
Crack Diameters (in)	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Normal	100%	100%	100%	100%	100%	100%	100%	100%	100%
BF	100%	95%	100%	91%	90%	88%	88%	90%	87%
OF	100%	100%	99%	90%	84%	85%	88%	84%	82%
IF	100%	100%	100%	98%	100%	100%	89%	90%	90%
ACA	100%	98.75%	99.75%	94.75%	93.5%	93.25%	91.25%	91%	89.75%

Table 6. The accuracy of the FEDI when the torque load is 3 hp.

Algorithms	AFSMO (Proposed Method)			ASMO			SMO		
Crack Diameters (in)	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Normal	100%	100%	100%	100%	100%	100%	100%	100%	100%
BF	100%	100%	100%	92%	86%	88%	85%	85%	85%
OF	100%	100%	100%	86%	87%	90%	80%	87%	89%
IF	100%	100%	100%	94%	96%	96%	87%	93%	90%
ACA	100%	100%	100%	93%	92.25%	93.5%	88%	91.25%	91%

The experimental results demonstrated in Tables 3–6 show that the proposed methodology based on the AFSMO, and DT for fault identification of the REBs clearly outperforms the state-of-the-art techniques used for comparison in this paper for all the available fault types with various fault severity degrees (i.e., crack sizes). Despite the advantages of the proposed methodology, the significant drop in classification accuracy can be observed for OR Fault under a 0 hp load with a crack of size 0.007 inches—around 84% (Table 3). This performance drop can be explained as follows. During the analysis of this behavior, it was discovered that the CWRU signals acquired under the conditions mentioned above for the OR Fault and IR Fault overlap significantly. Since the proposed and referenced methodologies create the model of the system being investigated based on the actual data collected from the testbed, they tend to transfer some of the uncertainty conditions presented in the original data into the created model. Because of this overlapping behavior appearing in the originally collected

data, the residual signals provided by the proposed methodology and energy values extracted to characterize the windows of these residuals also overlap each other. However, the proposed technique allowed for the improvement of fault identification for the BF and IF compared to the counterparts used in this paper. Thus, even with the significant accuracy decrease for the OF condition, it is clear that the proposed methodology demonstrates better classification performance in terms of the ACA.

5.2. Load-Variant Datasets

In this subsection, based on Table 1 we re-configured the available CWRU dataset and created three custom load-variant datasets to validate the robustness of the proposed fault identification technique under changing load conditions. In more detail, each of the three new datasets corresponded to three different sizes of the crack (i.e., 0.007, 0.014, and 0.021 inches in diameter), which contains $N_f \times N_{load} \times N_w$ data instances, where N_f is the number of fault conditions presented in the available CWRU dataset, N_{load} is the number of load levels, and N_w is the number of windows cut from the residual signals delivered by the AFSMO. The results obtained during this experiment are presented in Table 7.

Table 7. The accuracy of the FEDI for the custom load-variant datasets.

Algorithms	AFSMO (Proposed Method)			ASMO			SMO		
Crack Diameters (in)	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021
Normal	100%	100%	100%	100%	100%	100%	99%	99%	100%
BF	100%	99%	100%	95%	95%	95%	89%	88%	91%
OF	75%	99%	95%	81%	90%	92%	75%	79%	84%
IF	100%	99%	93%	93%	96%	91%	90%	91%	94%
ACA	93.8%	99.1%	97%	92.3%	95.25%	94.5%	88.3%	89.3%	92.3%

The results presented in this table allow us to conclude that the proposed method is highly robust to the changing experimental conditions, such as torque loads and rotating speeds. The proposed approach outperformed its counterparts with the lowest average classification accuracy of 93.8%, achieved for the crack with a diameter of 0.007 inches under different load level conditions in this case study.

6. Conclusions

This paper proposed a decision tree-based advanced fuzzy SMO (AFSMO) to perform reliable FEDI in a rotating machine. Rolling element bearing (REB) presents a nonlinear system that can now be analyzed in uncertain conditions. To improve the system modeling accuracy, the fuzzy ARX-Laguerre technique is presented in the first step. In the second step, the SMO is designed for the FEDI. To reduce the chattering and increase the estimation accuracy of the SMO, in the third step, the AFSMO is considered. To increase the fault estimation accuracy, the T-S fuzzy algorithm is applied to the AFSMO in the fourth step. To perform fault detection and identification in the presence of uncertainties, decision trees govern the AFSMO to find the exact solution for fault diagnosis under various crack types and motor speed conditions. The effectiveness of the proposed algorithm is validated using a publicly available vibration dataset of CWRU. The proposed decision tree-based AFSMO outperformed the ASMO and SMO in terms of classification accuracy. As a result, the proposed method improved the average fault identification performance by about 4.5%, 5.3%, and 4.8% compared with the ASMO for the crack sizes of 0.007, 0.014, and 0.021 inches, respectively. In addition, the AFSMO improved the average performance of fault identification by about 8.8%, 7.6%, and 8.1% compared with the SMO for the crack sizes of 0.007, 0.014, and 0.021 inches, respectively. Moreover, the proposed technique demonstrated its robustness while performing the tasks of REB fault detection and identification under changing operating conditions, such as variable load levels and variable rotating speeds. Based on the results we can conclude that the proposed methodology is highly efficient in diagnosing bearing faults. However, to ensure the robustness and efficacy in the industrial environment, it is crucial to reduce the

error of the system modeling. To address this issue, we will focus on developing an adaptive algorithm in our future work. Moreover, we will focus on exploring and evaluating the capabilities of applying the proposed methodology in conjunction with various vibration health indicators for a task of REB fault prognosis on run-to-failure experimental data.

Author Contributions: All authors contributed equally to the conception of the idea, the design of experiments, the analysis and interpretation of results, and the writing and improvement of the manuscript.

Funding: It was supported in part by the Ministry of SMEs and Startups (MSS), Korea, under the “Regional Specialized Industry Development Program (R&D, P0002964)” supervised by the Korea Institute for Advancement of Technology (KIAT). This work was also partially supported a study on the “Leaders in INdustry-university Cooperation +” Project, supported by the Ministry of Education and National Research Foundation of Korea.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, Z.; Cecati, C.; Ding, S.X. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [[CrossRef](#)]
2. Adams, M.L. Analysis of Rolling Element Bearing Faults in Rotating Machinery: Experiments, Modeling, Fault Detection and Diagnosis. Ph.D. Thesis, Case Western Reserve University, Cleveland, OH, USA, 2001.
3. Mehrjou, M.R.; Marium, N.; Hamiruce, M.; Misron, N. Rotor fault condition monitoring techniques for squirrel-cage induction machine—A review. *Mech. Syst. Signal Process.* **2011**, *25*, 2827–2848. [[CrossRef](#)]
4. Cecati, C. A survey of fault diagnosis and fault-tolerant techniques—Part II: Fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3768–3774.
5. Piltan, F.; Kim, C.-H.; Kim, J. Adaptive Fuzzy-Based Fault Tolerant Control of a Continuum Robotic System for Maxillary Sinus Surgery. *Appl. Sci.* **2019**, *9*, 2490. [[CrossRef](#)]
6. Piltan, F.; Kim, J.-M. Bearing fault diagnosis by a robust higher-order super-twisting sliding mode observer. *Sensors* **2018**, *18*, 1128. [[CrossRef](#)] [[PubMed](#)]
7. Farzin, P.; Kim, J.-M. Bearing Fault Diagnosis Using an Extended Variable Structure Feedback Linearization Observer. *Sensors* **2018**, *18*, 4359. [[CrossRef](#)]
8. Guo, X.; Chen, L.; Shen, C. Hierarchical Adaptive Deep Convolution Neural Network and Its Application to Bearing Fault Diagnosis. *Measurement* **2016**, *93*, 490–502. [[CrossRef](#)]
9. Lou, X.; Loparo, K.A. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech. Syst. Signal Process.* **2004**, *18*, 1077–1095. [[CrossRef](#)]
10. Antoni, J. Fast computation of the kurtogram for the detection of transient faults. *Mech. Syst. Signal Process.* **2007**, *21*, 108–124. [[CrossRef](#)]
11. Cheng, J.; Yu, D.; Tang, J.; Yang, Y. Local rub-impact fault diagnosis of the rotor systems based on EMD. *Mech. Mach. Theory* **2009**, *44*, 784–791. [[CrossRef](#)]
12. Prosvirin, A.; Islam, M.; Kim, J.; Kim, J.-M. Rub-Impact Fault Diagnosis Using an Effective IMF Selection Technique in Ensemble Empirical Mode Decomposition and Hybrid Feature Models. *Sensors* **2018**, *18*, 2040. [[CrossRef](#)] [[PubMed](#)]
13. Lei, Y.; Zuo, M.J. Fault diagnosis of rotating machinery using an improved HHT based on EEMD and sensitive IMFs. *Meas. Sci. Technol.* **2009**, *20*, 125701. [[CrossRef](#)]
14. Bai, L.; Han, Z.; Li, Y.; Ning, S. A Hybrid De-Noising Algorithm for the Gear Transmission System Based on CEEMDAN-PE-TFPPF. *Entropy* **2018**, *20*, 361. [[CrossRef](#)]
15. Nguyen, P.H.; Kim, J.-M. Multifault Diagnosis of Rolling Element Bearings Using a Wavelet Kurtogram and Vector Median-Based Feature Analysis. *Shock Vib.* **2015**, *2015*, 1–14. [[CrossRef](#)]
16. Wang, D. Spectral L2/L1 norm: A new perspective for spectral kurtosis for characterizing non-stationary signals. *Mech. Syst. Signal Process.* **2018**, *104*, 290–293. [[CrossRef](#)]
17. Wang, D. Some further thoughts about spectral kurtosis, spectral L2/L1 norm, spectral smoothness index and spectral Gini index for characterizing repetitive transients. *Mech. Syst. Signal Process.* **2018**, *108*, 360–368. [[CrossRef](#)]
18. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [[CrossRef](#)]

19. Lu, C.; Wang, Z.-Y.; Qin, W.-L.; Ma, J. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Process.* **2017**, *130*, 377–388. [[CrossRef](#)]
20. Bouzrara, K.; Garna, T.; Ragot, J.; Messaoud, H. Decomposition of an ARX model on Laguerre orthonormal bases. *ISA Trans.* **2012**, *51*, 848–860. [[CrossRef](#)]
21. Najeh, T.; Njima, C.B.; Garna, T.; Ragot, J. Input fault detection and estimation using PI observer based on the ARX-Laguerre model. *Int. J. Adv. Manuf. Technol.* **2017**, *90*, 1317–1336. [[CrossRef](#)]
22. Piltan, F.; Kim, J.-M. Nonlinear Extended-state ARX-Laguerre PI Observer Fault Diagnosis of Bearings. *Appl. Sci.* **2019**, *9*, 888. [[CrossRef](#)]
23. Chu, Z.; Meng, F.; Zhu, D.; Luo, C. Fault reconstruction using a terminal sliding mode observer for a class of second-order MIMO uncertain nonlinear systems. *ISA Trans.* **2019**. [[CrossRef](#)] [[PubMed](#)]
24. Yang, H.; Yin, S. Reduced order sliding mode observer-based fault estimation for Markov jump systems. *IEEE Trans. Autom. Control* **2019**, *64*, 4733–4740. [[CrossRef](#)]
25. Guezmil, A.; Berriri, H.; Pusca, R.; Sakly, A.; Romary, R.; Mimouni, M.F. High order sliding mode observer-based backstepping fault-tolerant control for induction motor. *Asian J. Control* **2019**, *21*, 33–42. [[CrossRef](#)]
26. Fridman, L.; Davila, J.; Levant, A. High-order sliding-mode observation for linear systems with unknown inputs. *Nonlinear Anal. Hybrid. Syst.* **2011**, *5*, 189–205. [[CrossRef](#)]
27. Van, M.; Kang, H.-J.; Shin, K.-S. Backstepping quasi-continuous high-order sliding mode control for a Takagi–Sugeno fuzzy system with an application for a two-link robot control. *Proc. Inst. Mech. Eng. Part. C J. Mech. Eng. Sci.* **2014**, *228*, 1488–1500. [[CrossRef](#)]
28. Ferrara, A.; Rubagotti, M. A sub-optimal second order sliding mode controller for systems with saturating actuators. *IEEE Trans. Autom. Control* **2009**, *54*, 1082–1087. [[CrossRef](#)]
29. Bartolini, G.; Pisano, A.; Punta, E.; Usai, E. A survey of applications of second-order sliding mode control to mechanical systems. *Int. J. Control.* **2003**, *76*, 875–892. [[CrossRef](#)]
30. Adhikari, Y.R. Inference and Decision Making Methods in Fault Diagnosis System of Industrial Processes. *IFAC Proc. Vol.* **2004**, *37*, 193–198. [[CrossRef](#)]
31. Huo, Z.; Zhang, Y.; Shu, L.; Gallimore, M. A New Bearing Fault Diagnosis Method Based on Fine-to-Coarse Multiscale Permutation Entropy, Laplacian Score and SVM. *IEEE Access* **2019**, *7*, 17050–17066. [[CrossRef](#)]
32. Darwish, H.A.; Taalab, A.-M.I.; Kawady, T.A. Development and implementation of an ANN-based fault diagnosis scheme for generator winding protection. *IEEE Trans. Power Deliv.* **2001**, *16*, 208–214. [[CrossRef](#)]
33. Iannace, G.; Ciaburro, G.; Trematerra, A. Fault Diagnosis for UAV Blades Using Artificial Neural Network. *Robotics* **2019**, *8*, 59. [[CrossRef](#)]
34. Prosvirin, A.; Kim, J.; Kim, J.-M. Bearing Fault Diagnosis Based on Convolutional Neural Networks with Kurtogram Representation of Acoustic Emission Signals. In *Advances in Computer Science and Ubiquitous Computing*; Park, J.J., Loia, V., Yi, G., Sung, Y., Eds.; Springer: Singapore, Singapore, 2018; Volume 474, pp. 21–26.
35. Mao, W.; Liu, Y.; Ding, L.; Li, Y. Imbalanced Fault Diagnosis of Rolling Bearing Based on Generative Adversarial Network: A Comparative Study. *IEEE Access* **2019**, *7*, 9515–9530. [[CrossRef](#)]
36. De Bruin, T.; Verbert, K.; Babuska, R. Railway Track Circuit Fault Diagnosis Using Recurrent Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 523–533. [[CrossRef](#)] [[PubMed](#)]
37. Dietterich, T.G. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Mach. Learn.* **2000**, *40*, 139–157. [[CrossRef](#)]
38. Zhang, Q.; Yang, Y.; Ma, H.; Wu, Y.N. Interpreting CNNs via Decision Trees. *arXiv* **2018**, arXiv:1802.00121. Available online: <https://arxiv.org/pdf/1802.00121.pdf> (accessed on 25 February 2019).
39. King Saud University; Saudi, A.; Al-Barrak, M.A.; Al-Razgan, M. King Saud University, Saudi Arabia Predicting Students Final GPA Using Decision Trees: A Case Study. *Int. J. Inf. Educ. Technol.* **2016**, *6*, 528–533.
40. Quinlan, J.R. Decision trees and decision-making. *IEEE Trans. Syst. Man Cybern.* **1990**, *20*, 339–346. [[CrossRef](#)]
41. Mishra, C.; Samantaray, A.K.; Chakraborty, G. Ball bearing defect models: A study of simulated and experimental fault signatures. *J. Sound Vib.* **2017**, *400*, 86–112. [[CrossRef](#)]
42. Li, L.; Ding, S.X.; Qiu, J.; Yang, Y.; Xu, D. Fuzzy observer-based fault detection design approach for nonlinear processes. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *47*, 1941–1952. [[CrossRef](#)]
43. Iannace, G.; Ciaburro, G.; Trematerra, A. Heating, Ventilation, and Air Conditioning (HVAC) Noise Detection in Open-Plan Offices Using Recursive Partitioning. *Buildings* **2018**, *8*, 169. [[CrossRef](#)]

44. Iannace, G.; Ciaburro, G.; Trematerra, A. Wind Turbine Noise Prediction Using Random Forest Regression. *Machines* **2019**, *7*, 69. [[CrossRef](#)]
45. Sen, P.K. Gini diversity index, Hamming distance, and curse of dimensionality. *Int. J. Stat.* **2005**, *LXIII*, 329–349.
46. Laber, E.S.; Molinaro, M. Binary Partitions with Approximate Minimum Impurity. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).