THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING IN
MACHINE AND VEHICLE SYSTEMS

A simulation-driven machine learning framework for condition monitoring of rotating machinery

Yu-Hung Pai

Department of Mechanics and Maritime Sciences
Chalmers University of Technology
Gothenburg, Sweden, 2026

A simulation-driven machine learning framework for condition monitoring of rotating machinery

Yu-Hung Pai

Thesis for the degree of Licentiate of Engineering

Department of Mechanics and Maritime Sciences
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Phone: +46(0)31 772 1000

Printed by Chalmers Digitaltryck,
Gothenburg, Sweden 2026.

To the friction that sharpens my edges and propels me toward growth.

A simulation-driven machine learning framework for condition monitoring of rotating machinery

Yu-Hung Pai
Department of Mechanics and Maritime Sciences
Chalmers University of Technology

Abstract

Rotating machinery is widely used across many industrial fields and functions as the core of operations. Such machines usually work in harsh environments, which accelerates machine degradation. To ensure reliability of rotating machinery, the implementation of condition monitoring (CM) and fault diagnosis strategies is crucial for identifying anomalies at an early stage.

This thesis focuses on the development of a simulation-driven transfer learning framework, with particularly emphasis on bearing fault diagnosis across different rotational speeds. The framework defines a target domain where the actual task takes place, and a well-defined source domain where the fault-related knowledge is acquired. Through transfer learning, knowledge gained from the source domain is effectively applied to real diagnostic tasks in the target domain. To build the target domain, an open-source bearing dataset is first used. After conducting experiments on the test rig at LUT University, an experimental bearing dataset is constructed by collecting vibration signals under different operating conditions.

In parallel, multibody dynamics (MBD) models are developed for generating vibration responses that reproduce representative fault characteristics. To build the source domain, a full-scale MBD model of the test rig is developed, where the geometries of all components and faults are consistent with the physical system. The model is then validated against the measurements, demonstrating its accuracy in reproducing fault signatures. For large-scale data generation, the computational cost of the MBD model is assessed in comparison to an analytical model. This process provides valuable insights into how the system dynamics and computational efficiency evolve with increasing model fidelity.

During the diagnostic model training stage, feature extraction methods are integrated into the transfer learning framework. These methods are designed to extract fault-related signatures across different operating conditions, while preserving consistent physical information. Finally, the diagnostic model, which is trained using the simulated-fault features, successfully distinguishes bearing conditions in the measurement data. The results demonstrate the proposed framework's potential for addressing more complex scenarios in future applications.

Keywords: condition monitoring, transfer learning, multibody dynamics simulation, bearing modelling, feature extraction, fault diagnosis.

List of Publications

This thesis is based on the following publications:

[**Paper A**] Y. H. Pai, K. Shehzad, P. T. Piiroinen, H. Johansson, C. Nutakor, J. Sopanen, I. Poutiainen, S. Kumar, T. Choudhury, "A full-scale multibody dynamics framework for simulation-driven transfer learning in bearing fault diagnosis," *To be submitted for international publication.*

[**Paper B**] Y. H. Pai, P. T. Piiroinen, H. Johansson, S. Kumar "Simulation-driven diagnostic method via a multibody dynamics model for rotating machinery," *Submitted to the 12th IFToMM International Conference on Rotordynamics 2026.*

[**Paper C**] Y. H. Pai, P. T. Piiroinen, S. Kumar and H. Johansson, "Machine-Learning Based Fault Diagnosis for a Rotordynamic System Using Multibody Simulations," *Proceedings of 15th International Workshop on Structural Health Monitoring (IWSHM), 2025.*

The appended papers were prepared in collaboration with the co-authors. The author of this thesis was responsible for the main contributions to the work, including the conceptualization, development of the machine learning framework and the multibody dynamics models, implementation of the signal processing, execution of the experiments, and manuscript writing.

Preface and Acknowledgment

The present work was conducted between August 2023 and January 2026 at the Division of Dynamics, Department of Mechanics and Maritime Sciences, Chalmers University of Technology. Financial and institutional support for this work was provided by the Department of Mechanics and Maritime Sciences, Chalmers University of Technology. The experiments and faulty bearing fabrication were carried out in the Laboratory of Machine Dynamics at LUT University, Lappeenranta, Finland.

I would like to express my sincere gratitude to all those who have supported and accompanied me throughout the completion of this work.

First and foremost, I am sincerely grateful to my supervisors and examiner, Prof. Petri Piiroinen, Prof. Shivesh Kumar and Prof. Håkan Johansson, for their guidance throughout this project. Their expertise and insightful feedback have introduced me to the field of dynamics and significantly shaped this work. I also deeply appreciate their continuous encouragement and willingness to engage in thoughtful discussions whenever I sought their advice.

I would like to extend my appreciation to Prof. Jussi Sopanen for his support and guidance during my time at LUT University, Finland. I am also grateful to Ali, Khurram and other colleagues at LUT, for their assistance and support throughout the bearing measurements. I felt warmly welcomed during that research period.

I would like to appreciate all my dear colleagues in the Division of Dynamics for creating such a wonderful working environment. I greatly cherish the memories we have together and have always felt assured of your support. Also, I would like to thank all my friends at Chalmers University of Technology for the valuable knowledge exchanges and the moment we shared.

I owe my deepest gratitude to my mother for her unconditional love and unwavering support. I will never forget the sacrifices she made, which have shaped me into the person I am today. Finally, I would like to express my sincere appreciation to my partner, Tzu-Ying Chien, for her boundless support and understanding. She has been a constant source of strength and provided me with a safe harbor throughout this journey.

Gothenburg, January 2026,

Yu-Hung Pai

Yu-Hung Pai

Contents

Part I

Extended Summary

CHAPTER 1

Introduction

Rotating machinery refers to equipment that operates rotational motion to perform its function. Unlike static equipment such as pipes or vessels, rotating machinery converts input energy into mechanical motion. It plays a pivotal role in power transmission across a wide range of industrial sectors. Examples include pumps, turbines, motors, and propellers, which serve as the core components of many mechanical systems, analogous to the human heart.

With increasing demands for precision, reliability and sustainability in industries, condition monitoring (CM) has been essential for ensuring operational effectiveness and optimal performance of rotating machinery [1], [2]. To enable predictive maintenance strategies, CM identifies anomalies at an early stage during operation that may indicate the development of severe damages. For rotating machinery, fault diagnosis of rolling bearings is particularly critical, as bearing faults account for approximately 30-50% of all machine failures [3], [4]. Therefore, bearing fault diagnosis is a fundamental starting point for CM of entire rotating systems.

1.1 Bearing fault characteristics

Since ball bearings are one of the most used components in rotating machinery [5], deep-groove ball bearings with localized defects are selected as the targets in this thesis. A typical ball bearing consists of four primary components: an outer ring (OR), an inner ring (IR), rolling elements (REs) and a cage. Normally, the OR is fixed to the bearing housing, while the IR rotates synchronously along with the shaft. The cage maintains a uniform spacing between the REs and rotates at a relative speed with respect to the IR.

When REs pass over a localized defect at a constant rotational speed, periodic impulses are generated during each revolution. These impulses manifest themselves at characteristic fault frequencies. Specifically, defects on the outer race and the inner race generate vibration components at the ball-pass frequency of the outer race (BPFO) and the ball-pass frequency of the inner race (BPFI), respectively, which are the target fault frequencies in this thesis and defined as [6]:

$$\text{BPFO (Hz)} = \frac{nf_r}{2}\left(1 - \frac{d}{D}\cos(\alpha)\right), \tag{1.1}$$

$$\text{BPFI (Hz)} = \frac{nf_r}{2}\left(1 + \frac{d}{D}\cos(\alpha)\right), \tag{1.2}$$

where n is the number of rolling elements, f_r is the rotational speed of the IR and normally identical to the shaft rotational speed, d is the diameter of the REs, D is the pitch diameter of the bearing, and α is the contact angle, respectively. To visualize the parameters, Figure 1.1 shows the radial cross-section of a ball bearing.

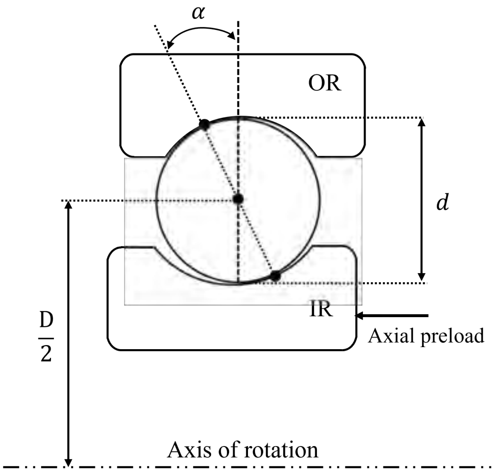


Figure 1.1: Radial ball bearing cross-section with contact angle due to axial preload.

Typically, the contact angle α of a deep-groove ball bearing is assumed to be zero. However, the introduction of an axial preload changes the contact angle, which leads to an increase in the BPFO and a decrease in the BPFI, as indicated by Equations (2.1) and (2.2). A detailed discussion of this effect is presented in **Paper A**.

1.2 Transfer learning in fault diagnosis

Achieving fault diagnosis relies heavily on effective data acquisition, since monitoring data is used to establish health indicators that extract fault-related information from measured signals [7]. Although traditional data-driven deep learning (DL) methods have achieved notable success in bearing fault diagnosis, their performance heavily relies on a substantial amount of labeled data for model training [4], [8], [9]. However, rotating machinery typically operates under healthy conditions for most of its lifespan, making the acquisition of sufficient fault data from operating real systems both challenging and impractical [10]. This imbalance in training data between health and fault datasets can adversely affect the performance of diagnostic models [11]. In addition, collecting diverse fault scenarios from physical systems is both cost-intensive and time-consuming [12]. To overcome such limitations, transfer learning (TL) methods are considered as a promising alternative to conventional DL-based training strategies [13], [14].

Unlike traditional DL approaches which train models independently within isolated domain, TL involves two datasets respectively drawn from a source domain and a target domain. The source domain typically consists of abundant labeled data (containing foundational knowledge), whereas the target domain represents a new task with limited labeled data where the acquired knowledge is applied. By leveraging knowledge acquired from the source domain, TL reduces the data preparation effort required for individual diagnostic cases. Therefore, TL is particularly well-suited for scenarios with limited labeled data [15]. It effectively mitigates data scarcity while enhancing diagnostic robustness in practical industrial applications.

1.3 Simulation-driven transfer learning based on multibody dynamics simulation

To reduce the reliance of TL on measurement data, simulation-driven methods provide a scalable approach of generating large quantities of labeled data for constructing the source domain [10], [16]. Although many analytical bearing models have been adopted in TL applications [17], [18], the target domain is constructed using data generation from a single bearing model. The limited number of degrees of freedom (DOF) inherent in analytical models constrains their ability to generate various fault data [19], especially compound faults data.

For the development of more comprehensive diagnostic models, both the diversity and the physical information of the simulation data are critical. Multibody dynamics (MBD) simulations offer higher physical fidelity and are therefore particularly well-suited for generating a wider range of fault scenarios under realistic operating conditions. Recent studies have demonstrated the effectiveness of simulation-driven approaches via MBD models in fault diagnosis, highlighting their potential for broader applications [20], [21]. However, high-fidelity MBD models typically lead to high computational cost, making large-scale data generation more challenging than using analytical models [22]. Therefore, addressing this trade-off between model fidelity and computational efficiency remains a major research challenge in TL applications.

1.4 Scope of research

This thesis presents a simulation-driven TL framework based on MBD simulations, with particularly emphasis on bearing fault diagnosis across different rotational speeds. The main objectives of this research are summarized as follows:

- Development of a simulation-driven TL framework for bearing fault diagnosis.
- Establishment of in-house bearing fault dataset.
- Development and evaluation of MBD models with manageable computational cost.
- Integration of a robust signal processing procedure for bearing fault diagnosis under varying operating conditions.

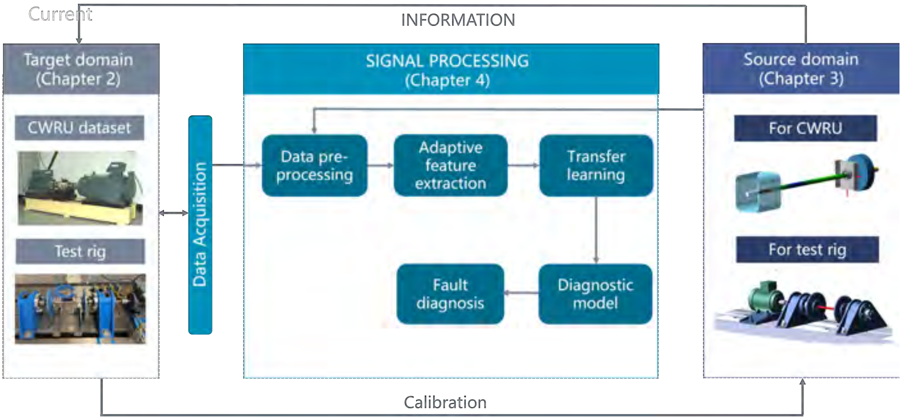


Figure 1.2: The TL framework for this thesis.

To this end, the proposed framework integrates three interconnected parts: a target domain, signal processing, and a source domain, as illustrated in Figure 1.2.

In the target domain (left), before building our bearing dataset, the open-source bearing dataset from Case Western Reserve University (CWRU) [23] is first utilized. A short introduction of the CWRU database is presented in Chapter 2.1. Following the establishment of our test rig and the fabrication of faulty bearings, the experimental measurements are carried out on the test rig presented in Chapter 2.2 and 2.3.

In the source domain (right), a single-rotor-bearing MBD model is built, and the bearing fault data are generated for diagnosing the CWRU bearing dataset. After collecting an in-house bearing dataset using the test rig, a virtual counterpart of the test rig is developed to construct a more comprehensive source-domain dataset.

In the signal processing procedure, fault-related features are extracted from both domains and used to train and fine-tune the diagnostic model through TL approaches. The adaptive feature extraction and TL are detailed in Chapter 4. Finally, the fine-tuned diagnostic model is used to diagnose the target domain dataset.

The remainder of this thesis is organized as follows. Chapter 2 describes the bearing dataset obtained from both the open-source CWRU database and the test rig. Chapter 3 introduces the development of the multibody dynamics model for the test rig. Chapter 4 introduces the TL framework and the associated signal processing procedure. Chapter 5 summarizes the main conclusions of the thesis and discusses future research directions. Summaries of the appended papers are provided in Chapter 6, followed by the three papers.

CHAPTER 2

Target domain

In this Chapter, construction of the target domain is presented. The open-source bearing dataset from the Case Western Reserve University (CWRU) database is first introduced. Following the establishment of the test rig at LUT University, the bearing vibration data are collected. The experimental arrangements and faulty bearing fabrication are also introduced here.

2.1 Case Western Reserve University bearing dataset

The CWRU bearing dataset [23] has been widely accepted as a standard reference, which provides access to the healthy and faulty bearing data [24], [25]. The test bench consists of a two-hp motor, a torque transducer and a dynamometer, as shown in Figure 2.1. The test bearings support the motor shaft. One SKF 6205-2RS deep-groove ball bearing is used for the drive-end bearing, and one SKF 6203-2RS is used for the fan-end bearing. Single-point faults were introduced to the test bearings using electro-discharge machining with different fault diameters. Vibration signals are collected by several accelerometers under healthy and single-fault bearings at both drive-end and fan-end.

In **Paper C**, a dataset comprising one healthy bearing and three bearing fault types, including an inner-race fault, an outer-race fault and a RE fault, is used. The selected fault diameters are 0.021 inches (0.53 mm) for all cases. The fault depths are 0.05 inches (1.27 mm)for both the inner race and outer race fault, and 0.15 inches (3.81 mm) for the RE fault. Each case is recorded for 10 seconds at a sampling rate of 12 kHz. For demonstration, an example of the one-second acceleration signals for all cases are shown in Figure 2.2.

2.2 Test rig arrangement

The experiments were carried out in the Laboratory of Machine Dynamics at LUT University, Lappeenranta, Finland, in April 2025. The experimental test rig, illustrated in Figure 2.3, comprised an AC motor driving a shaft, two discs,

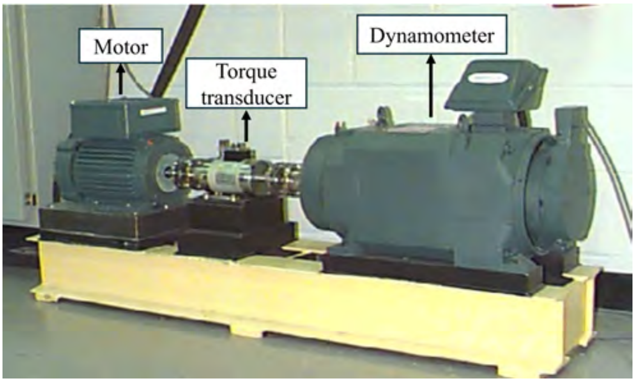


Figure 2.1: The CWRU test bench [23].

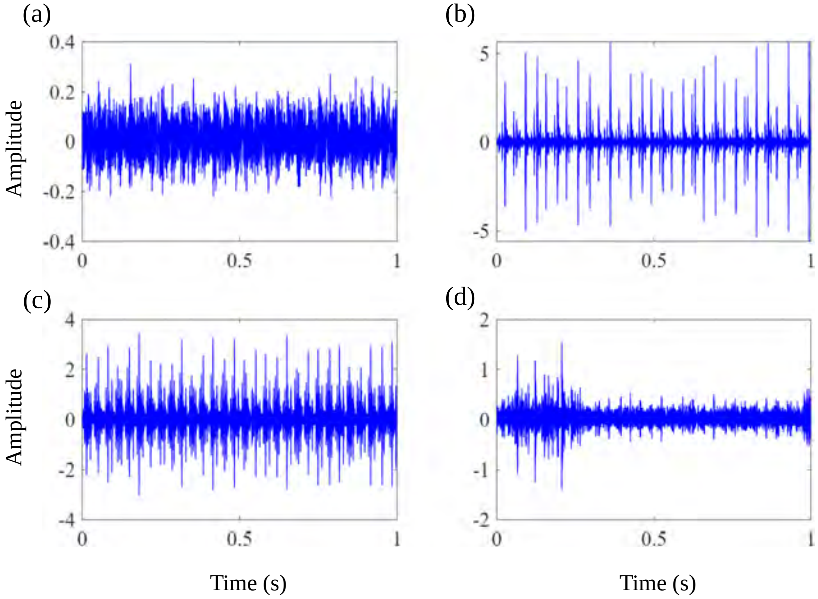


Figure 2.2: The CWRU data for four bearing conditions: (a) healthy bearing; (b) bearing with an outer-race fault; (c) bearing with an inner-race fault; (d) bearing with a RE fault.

two bearing housings and two deep-groove ball bearings. Based on the test rig design, the SKF 6007-2Z ball bearing was selected. The ball bearings were positioned at the drive end (DE) and non-drive end (NDE) of the shaft, with the NDE bearing intentionally selected as the target for fault diagnosis.

To acquire vibration responses, a PCB 356A03 tri-axial accelerometer was mounted on the NDE bearing housing, and it was connected to a DeweSoft SIRIUS® data acquisition unit operating at a 10 kHz sampling rate.

A remote optical tachometer (ROS-W) was installed adjacent to the DE of the shaft to monitor the shaft rotational speed. To investigate bearing fault diagnosis under varying rotational speed, two rotational speeds were used in this test. Although the nominal rotational speeds were set to 600 RPM and 1200 RPM, the measured rotational speeds were 588 RPM ($f_r = 9.8$ Hz) and 1188 RPM ($f_r = 19.8$ Hz). These measured rotational speeds were used as the input rotational speeds for the simulations presented in Chapter 3.

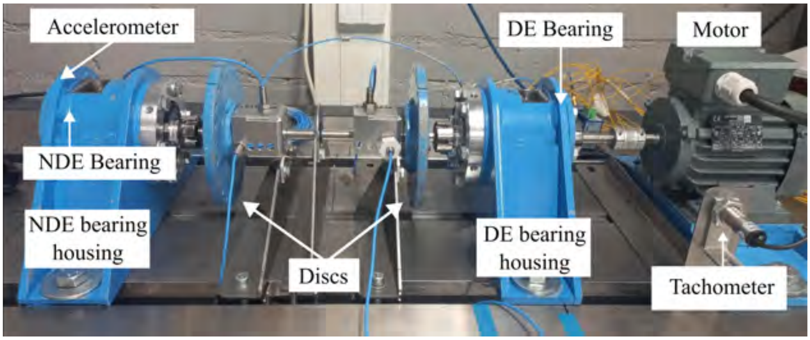


Figure 2.3: The experimental setup of the test rig.

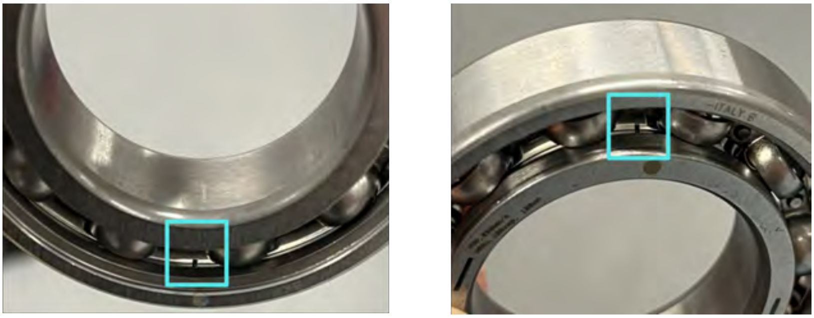


Figure 2.4: The test ball bearings with single ORF (left) and single IRF (right).

2.3 Faulty bearing fabrication

To ensure that the BPFO and BPFI can be clearly identified in the vibration response, artificial defects were designed as rectangular-shaped dents on the raceways. Two ball bearings with a single localized fault, one outer race fault (ORF) and one inner race fault (IRF), were fabricated using laser machining, as shown in Figure 2.4. The fabrication was conducted in the Laboratory of Laser Processing and Additive Manufacturing at LUT University, Lappeenranta, Finland, in April 2025. The machining process was performed under identical operating parameters and cutting angles to maintain consistency. Each defect was measured by a microscope to be approximately 2 mm in length, 0.5 mm in width and 0.2 mm in depth.

2.4 Measurement data from the test rig

The vertical vibration data of the NDE bearing housing are used in **Paper A and Paper B**. For demonstration, an example of the one-second time waveform for all three cases are shown in Figure 2.5.

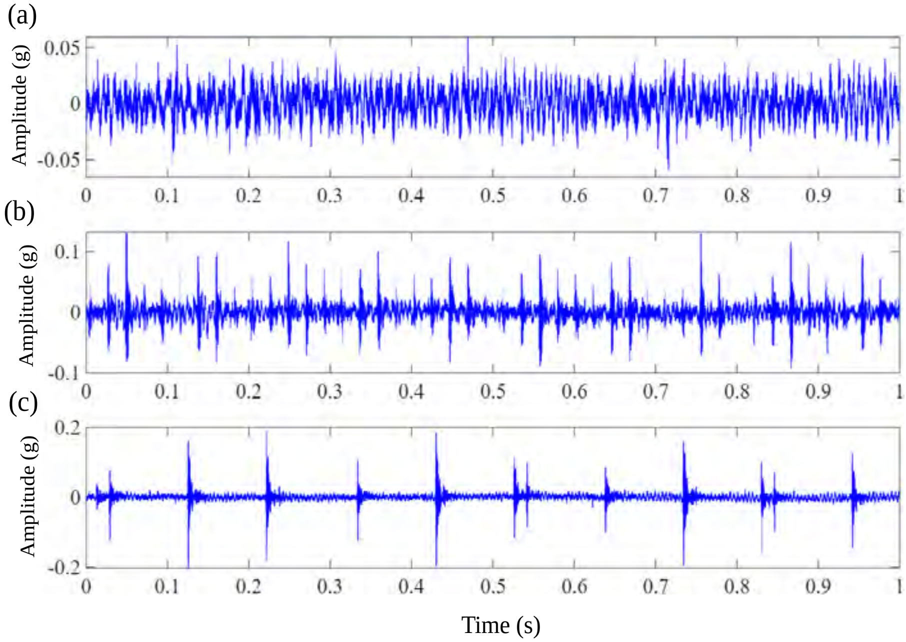


Figure 2.5: The measurement data from the test rig for three bearing conditions when $f_r = 9.8$ Hz: (a) healthy bearing; (b) faulty bearing with an ORF; (c) faulty bearing with an IRF.

CHAPTER 3

Source domain

In this Chapter, source-domain construction using multibody dynamics (MBD) simulations is presented. Since the MBD model for our test rig is more comprehensive, the single-rotor-bearing model is not presented in this chapter. A virtual counterpart of the test rig is built using the MBD software MSC Adams [26], which all dimensions consistent with those of the physical test rig. Subsequently, to preserve manageable computational cost, a high-fidelity (hi-fi) bearing model with simplified cage-pocket interaction is introduced. For comparison, an analytical bearing model is implemented within the same MBD model. The model validation and comparative analysis between two bearing models are summarized here.

3.1 Multibody dynamics model

The full MBD assembly is built using the MBD software MSC Adams [26], as depicted in Figure 3.1. In this model, the shaft is modelled as a flexible body, while all other components are considered as rigid bodies. An overview of the model topology is provided in Figure 3.2. At the drive-end (DE) part, the shaft is connected to the motor with a cylindrical joint with an axial spring-damper element, allowing rotation and displacement about axial direction (X-axis in Figure 3.1) under axial preload. A velocity-dependent driving torque is directly applied to the cylindrical joint, and the axial spring-damper element is included to control longitudinal motion and mitigate additional axial vibration.

To reduce the computational load, the DE bearing is modelled as a bushing element, since the non-drive-end (NDE) bearing is the primary objective of the fault diagnosis. The bushing representation comprises three rotational and three translational spring-damper elements. In contrast, the NDE bearing is modelled using both the hi-fi bearing model and the analytical bearing model for comparison. Two disks are rigidly fixed to the shaft between the two bearing housings and rotate about the same longitudinal axis as the shaft.

The simulated vibration responses are obtained at the NDE bearing housings, matching the location of the accelerometer in the experimental setup as shown in Figure 2.3. The interfaces between the housings and the supporting platform

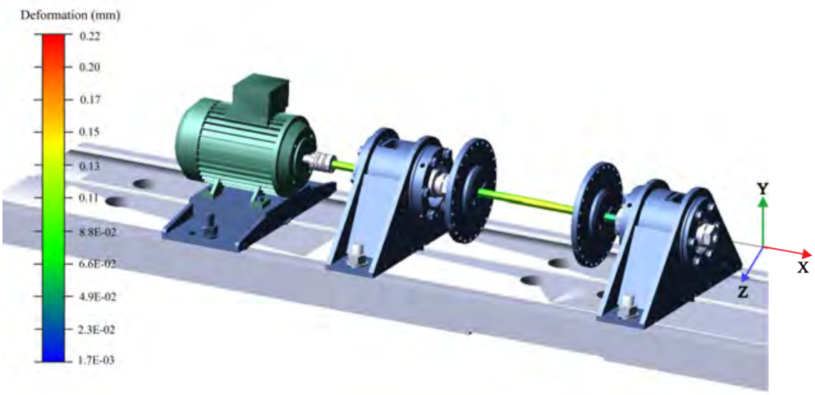
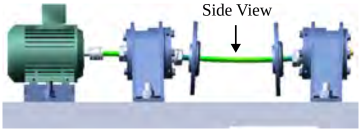
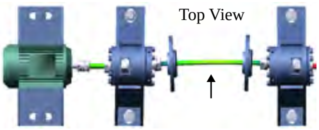


Figure 3.1: The full-scale MBD model of the test rig. The Y and Z axes are the radial direction and the X axis is the longitudinal direction in the global coordinate.

are represented by undamped bushing elements. Because the actual stiffness and damping values of these connections are not directly measurable, their parameters are calibrated to match the two dominant modal frequencies of the test rig obtained from experimental modal analysis. The calibrated results are shown in Table 3.1, ensuring that the simulated structural boundary conditions approximate the dynamic behavior of the physical system at given rotating speeds.

Table 3.1: The modal frequencies after calibration in the simulation against the experimental modal analysis.

	1^{st} shaft vertical bending mode (Hz)	1^{st} shaft lateral bending mode (Hz)
Experiments	31.25	33.75
Simulations	33.82	34.93
		

3.2 Bearing models

In the hi-fi bearing model, the 3D CAD geometries of the IR and OR are obtained from the manufacturer [27] and the complete assembly is shown in

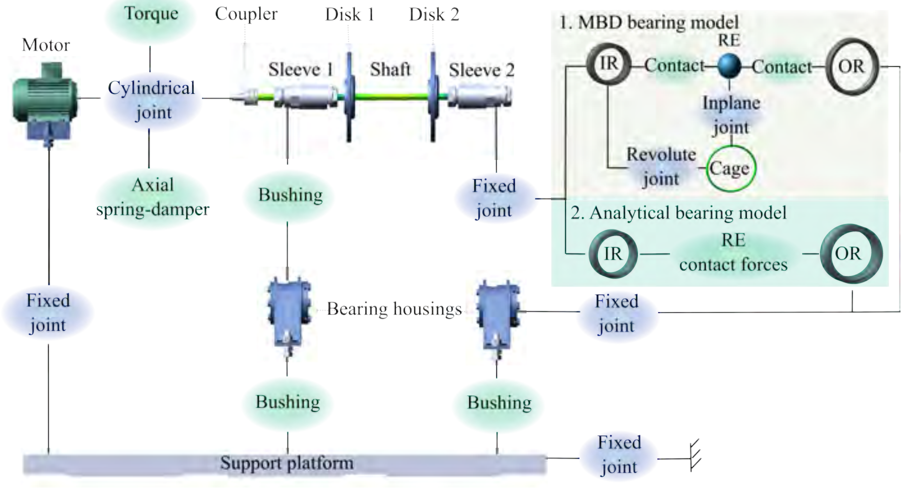


Figure 3.2: The topology of the proposed MBD model.

Figure 3.3: The complete hi-fi bearing model.

Figure 3.3. The cage is modelled as a rigid torus and connected to the IR through a revolute joint at its center of mass. Following the findings reported in [28], contact forces between the cage pockets and the REs are neglected in this study, as they significantly increase the computational time while offering limited benefit to the simulated fault characteristics.

To represent the ball-cage interaction, an in-plane joint is applied based on the work in [29]. Each RE is connected to the cage by an in-plane joint, allowing three rotational DOF about the X’, Y’ and Z’ axes and two translational DOF about the X’ and Y’ axes in the reference frame of the cage, as illustrated in Figure 3.4. Each in-plane joint is attached to the cage in its reference frame at the corresponding RE position and rotates along with the cage. Consequently, the REs are constrained to remain within the same geometric plane, thereby preserving equal angular spacing between REs while rotating.

Following the artificial faulty bearings in Figure 2.4, the faults are modelled by introducing cut-off geometries to replicate the dimensions of the actual faults, as shown in Figure 3.5.

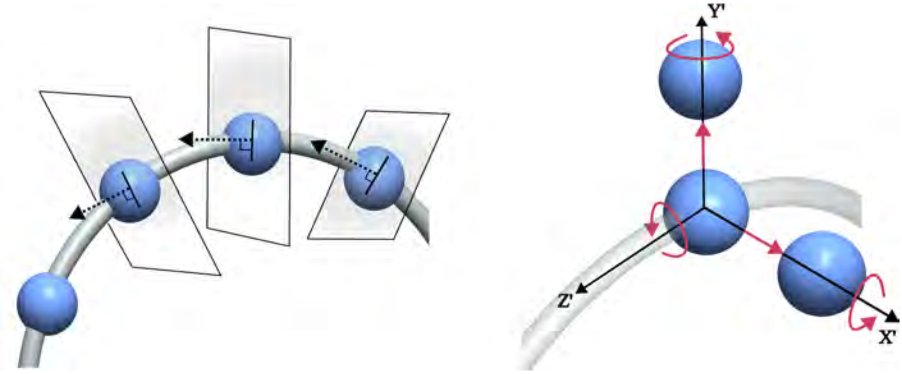


Figure 3.4: Each RE is constrained by an in-plane joint (left), which rotates along with the cage about the same axis and eliminates the DOF about the perpendicular direction (black arrow). In each in-plane joint (right), five red arrows depict five corresponding DOF .

In contrast, a 6-DOF analytical ball bearing model considering both an elastohydrodynamic film and non-Hertzian contact deformations is implemented within the same MBD model. This bearing model is detailed in [5], [30].

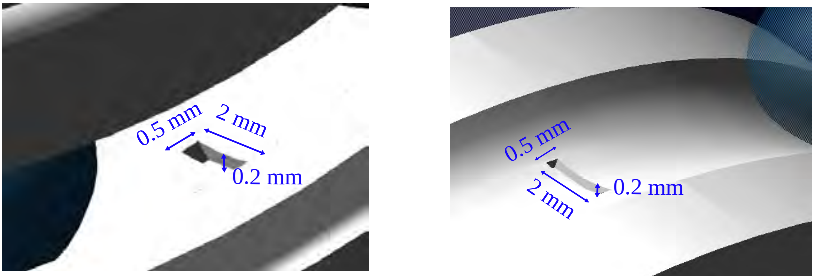


Figure 3.5: The hi-fi bearing model with single ORF (left) and IRF (right).

3.3 Comparison of the MBD models and the experimental measurements

For clarity, two cases for comparison are hereafter referred to as:

- Case 1: the DE bearing is modelled as a bushing element and the NDE bearing is modelled as a hi-fi bearing presentation.
- Case 2: the DE bearing is modelled as a bushing element and the NDE bearing is modelled as an 6-DOF analytical bearing representation.

To validate the fault characteristics, the envelope spectra of three scenarios

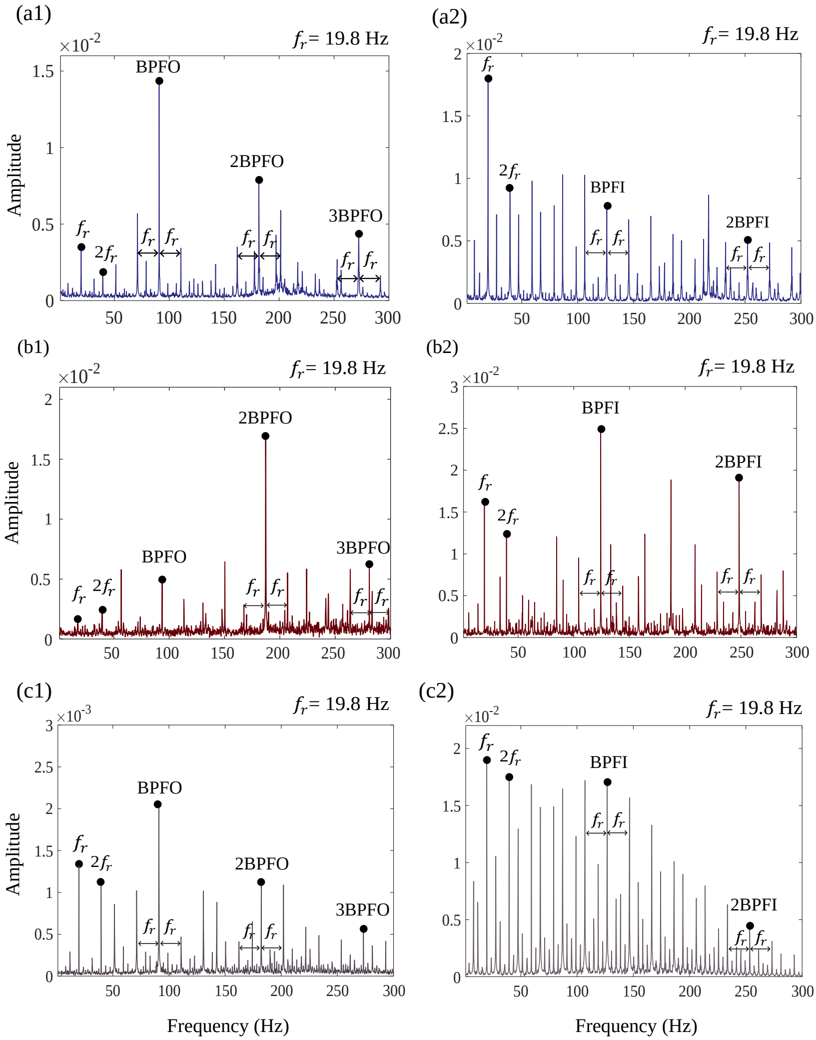


Figure 3.6: Envelope spectra of the vertical NDE housing acceleration for all cases: (a1) measurement data with single ORF; (a2) measurement data with single IRF (b1) Case 1 with single ORF (b2) Case 1 with single IRF (c1) Case 2 with single ORF (c2) Case 2 with single IRF.

are compared, namely measurement data, Case 1, and Case 2. The results for the single ORF case are presented in Figures 3.6(a1), (b1) and (c1). Both the measurement and simulated results exhibit comparable fault characteristics, with dominant frequency components corresponding to the BPFO and its harmonics. However, in Figure 3.6(b1), the amplitude of the second BPFO is higher than that of the first BPFO in Case 2. This discrepancy can be attributed to the simplified geometry of the simulated ORF, which generates two impacts when each RE rolls in and out the defect, thereby amplifying the second BPFO. Further investigations of the influence on simulated fault geometry and model calibration will be addressed in future work.

The results for the single IRF case are presented in Figures 3.6(a2), (b2) and (c2). The characteristics of BPFIs and their sidebands are clearly visible in the simulated results, as shown in Figures 3.6(b2) and (c2). These results highlight that the MBD model successfully reproduces the bearing fault characteristics that can be used for TL applications.

To evaluate the computational efficiency, Case 1 and Case 2 are simulated under identical computational conditions and compared with a model with two simplified bearings (modelled as bushing element). Table 3.2 shows an example of the computational cost of the proposed MBD framework, highlighting its suitability for large-scale fault simulation studies. The simulation time of Case 1 is only 1.13 times compared to Case 2, which is acceptable for generating massive simulated data. Moreover, the proposed MBD model is capable of accommodating more complex fault conditions, including RE faults and compound fault scenarios. More discussion and details are presented in **Paper A**.

Table 3.2: Simulation time of three scenarios. The simplified bearing is modelled as a bushing element.

Scenarios	f_r (Hz)	Simulation period (s)	Step size (s)	Elapsed time (s)
Case 1	19.8	10	10^{-6}	8251.25
Case 2	19.8	10	10^{-6}	7255.17
2 simplified bearings	19.8	10	10^{-6}	5408.75

CHAPTER 4

Simulation-driven transfer learning framework

This Chapter introduces the TL framework and the associated signal processing procedure. The goal is to diagnose three bearing conditions: outer-race fault (ORF), inner-race fault (IRF) and healthy status. The target domain is constructed by the bearing dataset measured from our test rig presented in Chapter 2.4, while the source domain is built by the simulation data generated by the MBD model in Case 2 presented in Chapter 3.3. Next, the feature extraction procedure for varying rotational speed is applied on both domains to acquire the fault-related features (knowledge for fault diagnosis). These features are then used to train and fine-tune a diagnostic model. Finally, this chapter concludes the results of the proposed simulation-driven TL framework.

4.1 Transfer learning framework

As introduced in Chapter 1.2, TL enables the diagnostic model to learn fault discrimination from the source domain and transfer this knowledge to the target domain. The proposed TL framework is shown in Figure 4.1.

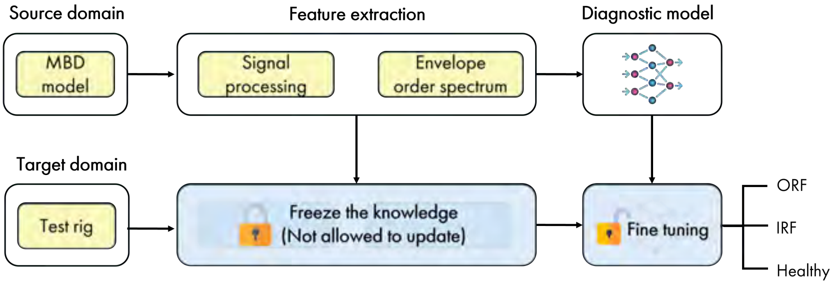


Figure 4.1: The proposed TL framework.

The source domain is composed of the bearing vibration data generated by the MBD model that incorporates a 6-DOF analytical bearing representation. Meanwhile, the target domain contains the bearing vibration data measured from the physical test rig. Initially, feature extraction techniques are applied to the source domain to capture fault-related features while preserving their physical interpretations. The extracted features are then used to construct a pre-trained diagnostic model.

In the fine-tuning stage, the parameters of the feature extraction stage are frozen, and only the parameters of the diagnostic model are fine-tuned using a small portion of the target domain dataset. This fine-tuning process allows the model to effectively adapt to real data and improves its classification performance.

4.2 Feature extraction

Feature extraction in fault diagnosis refers to the transformation of raw data into representative features that highlight fault-sensitive information. Therefore, reliable feature extraction is essential for improving fault diagnosis performance [31]. Based on the bearing fault characteristics introduced in Chapter 1.1, this research aims to design a feature extraction procedure that tracks fault frequencies across different rotational speeds. The proposed feature extraction procedure is depicted in Figure 4.2.

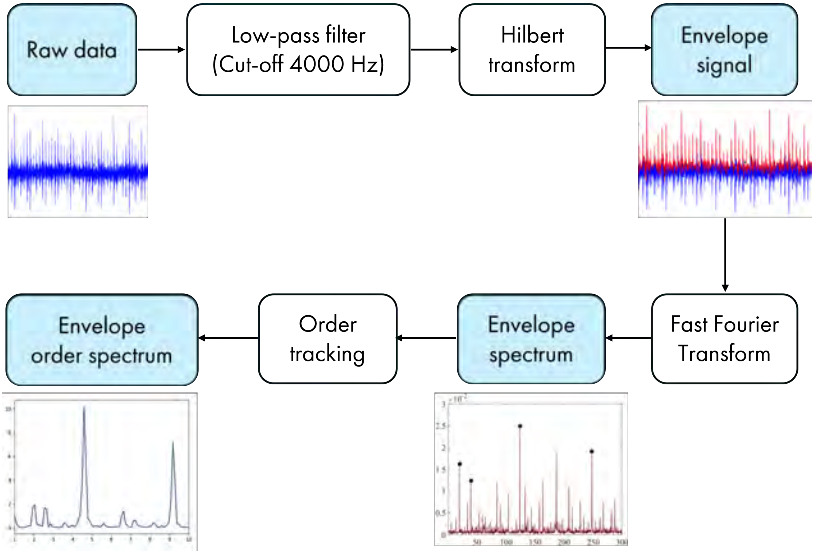


Figure 4.2: Feature extraction procedure.

Both source domain and target domain data are sampled at 10 kHz frequency, resulting in a 5000 Hz range. A low-pass filter with a 4000 Hz cut-off frequency is first applied to suppress aliasing and high-frequency noise. In the field of

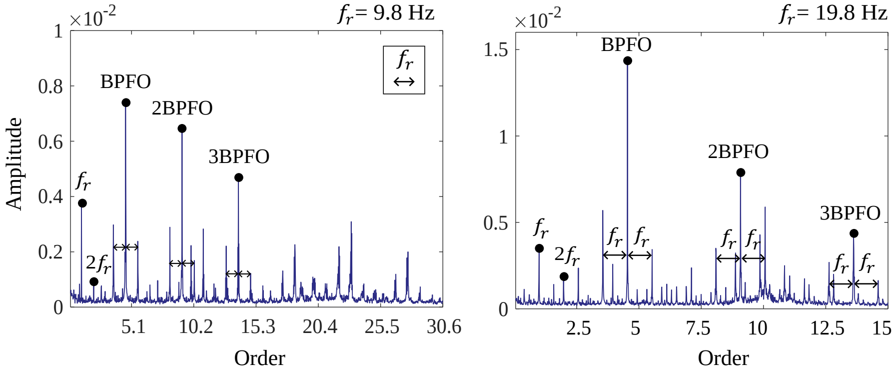


Figure 4.3: Envelope order spectra of the measurement data at two rotational speeds: when f_r = 9.8 Hz, BPFO = 4.59 order (left); when f_r = 19.8 Hz, BPFO = 4.6 order (right).

bearing fault detection, since bearing fault signals are usually modulated by low-frequency waveforms, the Fast Fourier Transform (FFT) usually falls short of identifying fault frequencies [6], [32]. Instead, envelope analysis (envelope spectrum) that uses the Hilbert transform to extract envelope signals, followed by FFT, is widely accepted as an efficient method to demodulate vibration signals and to identify their fault frequencies and harmonics [33]. The envelope spectra of the target domain and source domain are shown in Figure 3.6 and discussed in Chapter 3.3.

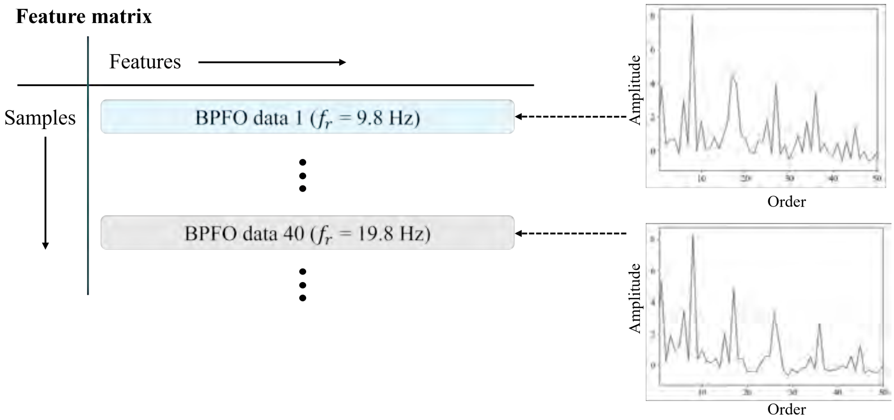


Figure 4.4: Construction of the feature matrix, where each row corresponds to a sample and each column represents a feature.

According to the bearing fault frequencies introduced in Chapter 1.1, the BPFO and BPFI vary with shaft rotational speeds. To represent the same fault type, the fault frequencies should be mapped to consistent locations in

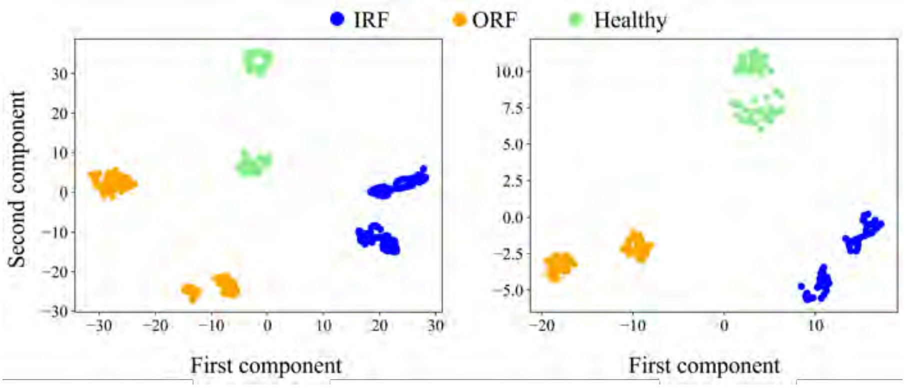


Figure 4.5: Group scatter after feature extraction: the source domain data (left) and the target domain data (right).

the feature matrix for diagnostic model training. Therefore, envelope spectra are converted into envelope order spectra using order tracking [34], where features are represented by multiples of rotational orders rather than absolute frequencies, as shown in Figure 4.3.

To construct the feature matrix, the order axis of each envelope order spectrum is resampled at 0.5-order intervals, ranging from 0.5 to 50, leading to 100 data points. The amplitude within each interval is computed using root-mean-square (RMS) value. As shown in Figure, the RMS amplitudes of the resampled envelope order spectrum are used as input features for diagnostic model training, resulting a feature matrix with a dimensionality of the numbers of samples multiplied by 100.

To verify that feature extraction process preserves the critical fault-related information, t-distributed Stochastic Neighbor Embedding (t-SNE) is applied to visualize the high-dimensional data. By projecting the data onto two t-SNE components, as shown in Fig. 4.5, the three different conditions are well separated in both the source and target domains, confirming that the envelope order spectra effectively capture the bearing fault signatures at different rotational speeds.

4.3 Diagnostic model

The diagnostic model (classifier) is implemented using a feed-forward neural network. After evaluating several network configurations, the final architecture consists of three fully-connected layers with ReLU activation function, followed by an output layer with Softmax classification [35]. Table 4.1 lists the architecture of the diagnostic model and the visualization is shown in Figure 4.6.

In the fine-tuning stage, the parameters of the fully-connected layers are frozen, and only the output layer (dense_4) of the diagnostic model is fine-tuned using a small portion of the target domain dataset. This fine-tuning process allows the model to effectively adapt to real data and improves its classification performance.

Table 4.1: Architecture of the diagnostic model.

Layer(type)	Output shape	Parameters
dense_1 (Dense)	(None, 128)	12928
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dropout_2 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 128)	16512
dropout_3 (Dropout)	(None, 128)	0
dense_4 (Dense)	(None, 3)	387

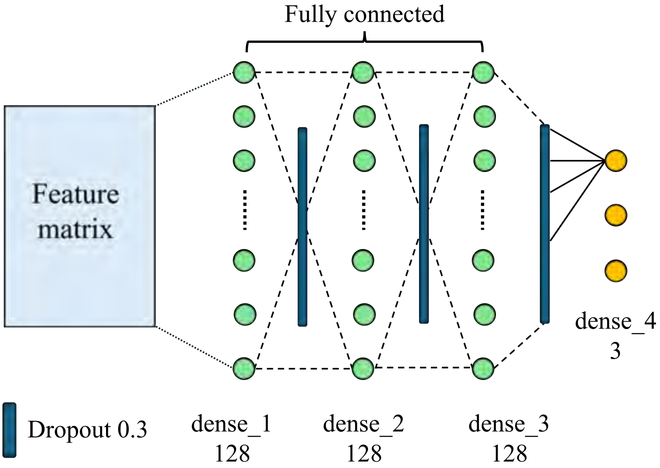


Figure 4.6: Visualization of the design model.

4.4 Diagnostic results

Both the source domain and target domain contain three types of the labeled data, namely IRF, ORF and healthy. Table 4.2 lists the dataset arrangement. The source domain comprises a total of 378 simulated samples, whereas the target domain comprises a total of 180 experimental samples.

After constructing the pre-trained diagnostic model using 90% of the simulated data for training and 10% for validation, the model is first evaluated on the experimental dataset. The diagnostic result is shown in Figure 4.7(a).

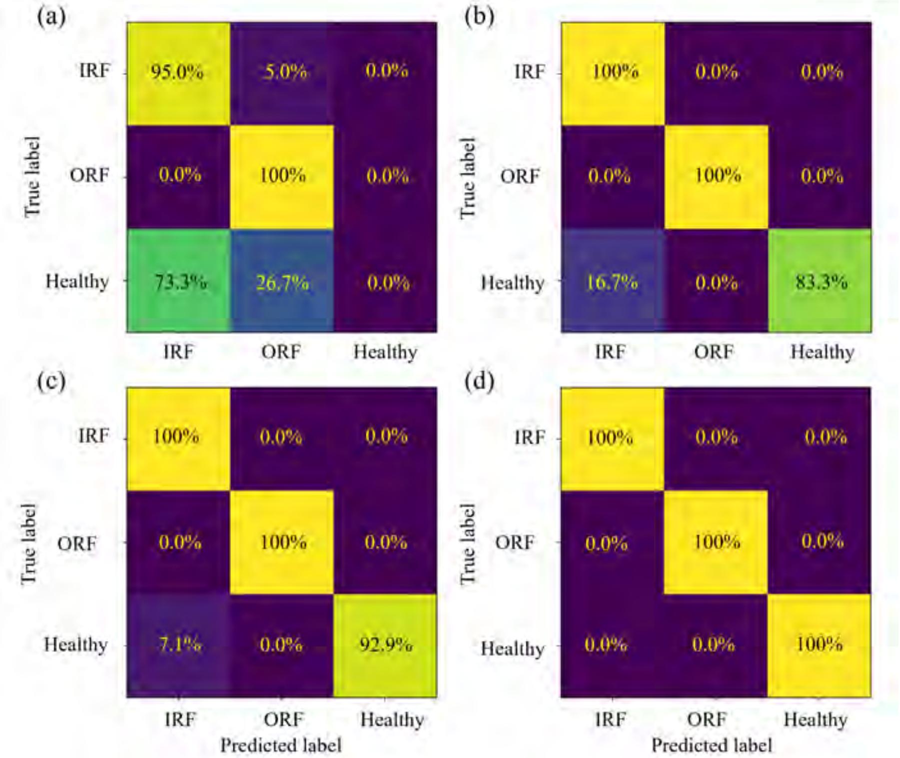


Figure 4.7: Confusion matrix of bearing fault diagnosis at different stages: (a) before fine-tuning (b) after fine-tuning by using 10% experimental data (c) after fine-tuning by using 20% experimental data (d) fine-tuned model re-evaluation on the simulation data.

The pre-trained model successfully distinguishes the IRF and ORF cases, indicating that the simulated features well represent the realistic faulty bearing behavior. However, it performs poorly on healthy data, resulting in an overall accuracy of only 65%. This is because of the discrepancy between the simulated and experimental healthy data, highlighting the necessity of applying a TL method to adapt the pre-trained model to real data.

Table 4.2: Dataset arrangement for TL framework.

Scenarios	Faulty condition	Simulated samples	Experimental samples
1	IRF	67	30
2	IRF	67	30
3	ORF	67	30
4	ORF	67	30
5	Healthy	55	30
6	Healhty	55	30

As shown in Figures 4.7(b) and (c), after fine tuning the diagnostic model using 10% of the measurement data, 83.3% of the healthy data are identified, leading to an overall classification accuracy of 94.4%. Increasing the fine-tuning dataset to 20% further improves the overall accuracy to 97.2% with 92.9% of the healthy data identified. To ensure that the model does not overfit the measurement data, the updated diagnostic model is re-evaluated on the simulated dataset after fine-tuned by 20% of the measurement data. As shown in Figure 4.7(d), the updated model achieves 100% overall accuracy on the source-domain data, demonstrating it does not overfit the fine-tuning data.

CHAPTER 5

Summary & Future work

This thesis presents a simulation-driven transfer learning framework (TL) that successfully diagnoses real bearing faults across varying rotational speeds, using data generated by multibody dynamics (MBD) models. The contributions of this work are summarized as follows:

- Model validation against the experiment data demonstrates the MBD model is capable of reproducing representative fault features.
- Comparative analysis between the hi-fi bearing and the analytical bearing representations highlights the computational efficiency of the proposed model.
- Feature extraction procedure successfully captures the fault-related signatures across different rotational speeds, highlighting its potential for addressing varying operation conditions.
- The fine-tuning process in TL significantly improves the diagnostic performance by a small portion of real data.

However, the results reveal discrepancies between the MBD simulations and the experimental measurements, indicating the need for further model calibration, particularly in the estimation of uncertain model parameters such as friction coefficient, stiffness and damping values. Moreover, the target-domain dataset in this thesis lacks sufficient diversity, which limits the comprehensive evaluation of the proposed feature extraction methods. Furthermore, the idealized bearing faults may not fully represent real fault characteristics, suggesting that more extensive case studies with realistic fault types are required to further enhance robustness of the proposed framework. Based on these findings, the directions for future work are summarized as follows:

- **MBD model calibration**
  Several estimated model parameters are currently selected based on literature review or trial-and-error approaches. Future work will focus on parameter optimization using experimental measurements to reduce discrepancies between simulated and measured responses.

- **Realistic faults and multi-fault diagnosis**
  To enrich the target domain dataset, a new test rig is currently under construction. It allows for the introduction of different levels of misalignment, unbalance and soft support conditions. Moreover, bearing faults with more realistic geometries and degradation characteristics should be investigated. Future studies will investigate various combinations of compound faults to further validate the proposed framework.

- **TL architectures and the associate feature extraction methods**
  Although the proposed TL framework performs well at the current stage, further investigation is required as the size of the target-domain dataset increases. Moreover, it is important to investigate the effect of the amount of fine-tuning data. The goal is to minimize the required fine-tuning samples, ideally enabling the diagnostic model to be fine-tuned using only health-condition data. Furthermore, more advanced architectures, such as convolution neural network (CNN), may provide improved robustness and diagnostic performance.

CHAPTER 6

Summary of appended papers

Paper A

Y. H. Pai, K. Shehzad, P. T. Piiroinen, H. Johansson, C. Nutakor, J. Sopanen, I. Poutiainen, S. Kumar, T. Choudhury, ”A full-scale multibody dynamics framework for simulation-driven transfer learning in bearing fault diagnosis,” *To be submitted for international publication.*

This paper presents the development of a full-scale MBD model that incorporates a hi-fi bearing representation. This work aims to reproduce bearing fault characteristics for data generation in TL applications, while preserving manageable computational efficiency. To do this, a test rig is built to acquire vibration signals and to construct experimental dataset. Subsequently, a full-scale MBD model is developed and validated against the experimental measurements. To further investigate the simulated fault features and computational efficiency, an analytical bearing model is implemented within the same MBD framework and used as a benchmark for comparison. The fault frequencies generated by the hi-fi bearing model are successfully reproduced, matching the experimental results and theoretical predictions. The comparative analysis highlights that the computational time of the hi-fi bearing model is only 1.13 times compared to the analytical model. In addition, the comparison provides deeper insight into the simulated fault features when the model fidelity increases. However, discrepancies remain between the simulated fault patterns and the experimental measurements, indicating that further model calibration is required in future work.

Paper B

Y. H. Pai, P. T. Piiroinen, H. Johansson, S. Kumar ”Simulation-driven diagnostic method via a multibody dynamics model for rotating machinery,” *Submitted to the 12th IFToMM International Conference on Rotordynamics 2026.*

To develop a diagnostic model for bearing fault diagnosis across different rotational speeds, this paper proposes a simulation-driven TL framework based on the MBD model that incorporates the analytical bearing representation introduced in Paper A. The source-domain dataset is fully generated using the MBD model, while the target-domain is constructed by the experimental measurements reported in **Paper A**. Subsequently, feature extraction is performed using order envelope spectra to capture fault-related features across different rotational speeds. The extracted features from the source domain are used to train a pre-trained diagnostic model. This model is then fine-tuned using a small subset of the target-domain data, enabling effective adaption to the real measurements. The proposed simulation-driven TL framework successfully classifies the experimental data, highlighting its potential for extension to more complex fault scenarios in future industrial integrations.

Paper C

Y. H. Pai, P. T. Piiroinen, S. Kumar and H. Johansson, "Machine-Learning Based Fault Diagnosis for a Rotordynamic System Using Multibody Simulations," *Proceedings of 15th International Workshop on Structural Health Monitoring (IWSHM), 2025.*

This paper presents a TL framework for bearing fault diagnosis based on a single-rotor-bearing MBD model. The source domain is constructed by the simulated data, while the target domain is formed using the open-source bearing fault dataset from Case Western Reserve University (CWRU). During feature extraction procedure, the fault-related features are extracted from envelope spectra using wavelet packet decomposition. An autoencoder neural network is subsequently applied to identify and select the most critical features. Finally, the diagnostic model is established within the TL framework and successfully classifies bearing faults in CWRU dataset. This paper lays a solid methodology foundation for subsequent studies presented in **Paper A** and **Paper B**.

Bibliography

[1] M. Vishwakarma, R. Purohit, V. Harshlata and P. Rajput, ‘Vibration analysis & condition monitoring for rotating machines: A review,’ *Materials Today: Proceedings*, vol. 4, pp. 2659–2664, 2 2017.

[2] S. K. Nithin, K. Hemanth, V. Shamanth, R. S. Mahale, P. C. Sharath and A. Patil, ‘Importance of condition monitoring in mechanical domain,’ *Materials Today: Proceedings*, vol. 54, pp. 234–239, 2022.

[3] D. T. Hoang and H. J. Kang, ‘A survey on deep learning based bearing fault diagnosis,’ *Neurocomputing*, vol. 335, pp. 327–335, 2019.

[4] S. Zhang, S. Zhang, B. Wang and T. G. Habetler, ‘Deep learning algorithms for bearing fault diagnosticsx - a comprehensive review,’ *IEEE Access*, vol. 8, pp. 29 857–29 881, 2020.

[5] J. Sopanen and A. Mikkola, ‘Dynamic model of a deep-groove ball bearing including localized and distributed defects. part 1: Theory,’ *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, vol. 217, pp. 201–211, 3 2003.

[6] S. A. McInerny and Y. Dai, ‘Basic vibration signal processing for bearing fault detection,’ *IEEE Transactions on Education*, vol. 46, pp. 149–156, 1 2003.

[7] H. Zhou et al., ‘Construction of health indicators for condition monitoring of rotating machinery: A review of the research,’ *Expert Systems with Applications*, vol. 203, p. 117 297, 2022.

[8] R. Liu, B. Yang, E. Zio and X. Chen, ‘Artificial intelligence for fault diagnosis of rotating machinery: A review,’ *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[9] Z. Chen, A. Mauricio, W. Li and K. Gryllias, ‘A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks,’ *Mechanical Systems and Signal Processing*, vol. 140, p. 106 683, 2020.

[10] W. Hou, C. Zhang, Y. Jiang, K. Cai, Y. Wang and N. Li, ‘A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data,’ *Measurement*, vol. 215, p. 112 879, 2023.

[11] K. Xu, X. Kong, Q. Wang, S. Yang, N. Huang and J. Wang, 'A bearing fault diagnosis method without fault data in new working condition combined dynamic model with deep learning,' *Advanced Engineering Informatics*, vol. 54, p. 101 795, 2022.

[12] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li and A. K. Nandi, 'Applications of machine learning to machine fault diagnosis: A review and roadmap,' *Mechanical Systems and Signal Processing*, vol. 138, p. 106 587, 2020.

[13] Z. Zhao et al., 'Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study,' *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–28, 2021.

[14] T. Han, C. Liu, R. Wu and D. Jiang, 'Deep transfer learning with limited data for machinery fault diagnosis,' *Applied Soft Computing*, vol. 103, p. 107 150, 2021.

[15] B. Yang, Y. Lei, F. Jia and S. Xing, 'An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings,' *Mechanical Systems and Signal Processing*, vol. 122, pp. 692–706, 2019.

[16] S. Tang, J. Ma, Z. Yan, Y. Zhu and B. C. Khoo, 'Deep transfer learning strategy in intelligent fault diagnosis of rotating machinery,' *Engineering Applications of Artificial Intelligence*, vol. 134, p. 108 678, 2024.

[17] C. Sobie, C. Freitas and M. Nicolai, 'Simulation-driven machine learning: Bearing fault classification,' *Mechanical Systems and Signal Processing*, vol. 99, pp. 403–419, 2018.

[18] J. Liu, H. Cao, S. Su and X. Chen, 'Simulation-driven subdomain adaptation network for bearing fault diagnosis with missing samples,' *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106 201, 2023.

[19] H. Cao, L. Niu, S. Xi and X. Chen, 'Mechanical model development of rolling bearing-rotor systems: A review,' *Mechanical Systems and Signal Processing*, vol. 102, pp. 37–58, 2018.

[20] J. Koutsoupakis, P. Seventekidis and D. Giagopoulos, 'Machine learning based condition monitoring for gear transmission systems using data generated by optimal multibody dynamics models,' *Mechanical Systems and Signal Processing*, vol. 190, p. 110 130, 2023.

[21] J. Koutsoupakis and D. Giagopoulos, 'A simulation-enhanced deep learning framework for bearing fault diagnosis using an improved bearing force model,' *Mechanical Systems and Signal Processing*, vol. 241, p. 113 519, 2025.

[22] L. Giraudo, L. G. D. Maggio, L. Giorio and C. Delprete, 'Dynamic multibody modeling of spherical roller bearings with localized defects for large-scale rotating machinery,' *Sensors 2025, Vol. 25, Page 2419*, vol. 25, p. 2419, 8 2025.

[23] *Case Western Reserve University Bearing Data Center*, `https://engineering.case.edu/bearingdatacenter`, last accessed 15 Jan 2026.

[24] D. Neupane and J. Seok, 'Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review,' *IEEE Access*, vol. 8, pp. 93 155–93 178, 2020.

[25] W. A. Smith and R. B. Randall, 'Rolling element bearing diagnostics using the case western reserve university data: A benchmark study,' *Mechanical Systems and Signal Processing*, vol. 64-65, pp. 100–131, 2015.

[26] *MSC Adams, Hexagon*, `https://hexagon.com/products/product-groups/computer-aided-engineering-software/adams`, last accessed 14 Jan 2026.

[27] *SKF 6007-2z deep-groove ball bearing*, `https://www.skf.com/group/products/rolling-bearings/ball-bearings/deep-groove-ball-bearings/productid-6007-2Z`, last accessed 15 Jan 2026.

[28] R. G. Moreno, F. Marques, E. C. Abad, J. M. Alonso, P. Flores and C. Castejon, 'Enhanced modelling of planar radial-loaded deep groove ball bearings with smooth-contact formulation,' *Multibody System Dynamics*, vol. 60, pp. 121–159, 1 2024.

[29] M. Vehviläinen et al., 'Adapting geometry-based polygonal contacts for simulating faulty rolling bearing dynamics,' *Mechanism and Machine Theory*, vol. 192, p. 105 552, 2024.

[30] J. Sopanen and A. Mikkola, 'Dynamic model of a deep-groove ball bearing including localized and distributed defects. part 2: Implementation and results,' *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, vol. 217, pp. 213–223, 3 2003.

[31] L. Wang and Y. Shao, 'Fault feature extraction of rotating machinery using a reweighted complete ensemble empirical mode decomposition with adaptive noise and demodulation analysis,' *Mechanical Systems and Signal Processing*, vol. 138, p. 106 545, 2020.

[32] Y. Cheng et al., 'An improved envelope spectrum via candidate fault frequency optimization-gram for bearing fault diagnosis,' *Journal of Sound and Vibration*, vol. 523, p. 116 746, 2022.

[33] D. Yu, J. Cheng and Y. Yang, 'Application of emd method and hilbert spectrum to the fault diagnosis of roller bearings,' *Mechanical Systems and Signal Processing*, vol. 19, pp. 259–270, 2 2005.

[34] J. Cheng, Y. Yang and D. Yu, 'The envelope order spectrum based on generalized demodulation time–frequency analysis and its application to gear fault diagnosis,' *Mechanical Systems and Signal Processing*, vol. 24, pp. 508–521, 2 2010.

[35] A. F. Agarap, *Deep learning using rectified linear units (relu)*, 2019. arXiv: `1803.08375 [cs.NE]`.

Part II

Appended Papers A-C

Paper A

A full-scale multibody dynamics framework for simulation-driven transfer learning in bearing fault diagnosis

Y. H. Pai, K. Shehzad, P. T. Piiroinen, H. Johansson, C. Nutakor,
J. Sopanen, I. Poutiainen, S. Kumar, T. Choudhury

To be submitted for international publication

A full-scale multibody dynamics framework for simulation-driven transfer learning in bearing fault diagnosis

Yu-Hung Pai[a], Khurram Shehzad[b], Petri T. Piiroinen[a], Håkan Johansson[a], Charles Nutakor[b], Jussi Sopanen[b], Ilkka Poutiainen[b], Shivesh Kumar[a,c], and Tuhin Choudhury[b]

[a]*Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden.*
[b]*LUT School of Energy Systems, Lappeenranta-Lahti University of Technology (LUT), FI-53850 Lappeenranta, Finland.*
[c]*Robotics Innovation Center, German Research Center for Artificial Intelligence (DFKI GmbH), 28359 Bremen, Germany.*

Abstract

Bearing fault diagnosis is a critical aspect of condition monitoring in rotating machinery. Conventional deep learning approaches rely on large labeled datasets from operating machines, which are often impractical and costly to obtain in industrial applications. As an alternative, simulation-driven fault diagnosis provides a scalable data generation for training diagnostic models. Therefore, developing a comprehensive and physical consistent model for data generation is essential. This study presents a full-scale multibody dynamics (MBD) model that incorporates a high-fidelity ball bearing representation for a physical test rig, designed to reproduce bearing fault characteristics with manageable computational cost for simulation-driven approaches. The MBD model is first validated against the experimental measurements from the physical test rig, demonstrating its capability to accurately replicate bearing fault signatures. A comparative analysis with an analytical bearing model further highlights the advantages of the proposed MBD model in both accuracy and computational efficiency. By capturing more detailed physical interactions and realistic operating defect responses, this study advances simulation-driven fault diagnosis and provides a promising foundation for future integration with transfer learning methods in industrial applications.

Keywords: Multibody dynamics simulation; bearing fault diagnosis; ball bearings; simulation-driven transfer learning; data generation.

---

1. Introduction

Condition monitoring systems are integral to the reliable operation of rotating machinery to ensure operational effectiveness and working performance at optimal levels [1–3]. They serve as predictive maintenance tools for identifying anomalies at an early stage during operation that can indicate the development of severe damages. For rotating machinery, fault diagnosis of rolling bearing plays a pivotal role, given that such faults are responsible for 30% to 50% of all machine failures [4–7]. In the past, the application of deep learning (DL) methods in bearing fault diagnosis has been widely studied [7–9]. Such approaches extract fault-related features automatically from measurement data, rather than rely on engineers' experience, and subsequently diagnose the current condition of machines.

Although DL methods have demonstrated significant success in bearing fault diagnosis, their performance heavily depends on a substantial amount of labeled data for model training [10, 11]. In practice, since rotating machines operate under normal conditions for most of their lifespan, it is challenging to obtain sufficient fault data from real systems in operation [12]. An imbalance in training data between health and fault datasets could adversely affect the performance of diagnostic models. In addition, collecting various types of faults from physical systems is both cost intensive and time consuming. To overcome such limitations, transfer learning (TL) methods are considered a promising alternative to conventional data generation approaches for training DLs [10, 13–15]. Applications of TL enable diagnostics models to leverage knowledge acquired from one or multiple source domains and apply them to relevant, but different, target domains [16, 17]. Such approaches facilitate effective condition monitoring, especially in the field of rotating machinery, where fault data is insufficient and thereby enhancing the adaptability and robustness of diagnostic models across varying operating conditions.

Simulation-driven methods, using dynamic bearing models, provide a straightforward way to construct a source domain for TL methods [11, 18]. The primary objective is to generate sufficient bearing fault signatures from analytical models or numerical simulations for diagnostic model training. In analytical bearing models, simplistic quasi-static or dynamic models of bearings, with localized defects have laid a solid foundation for reproducing bearing fault characteristics [19–21]. The joints between components are assumed to be ideal and formulated by kinematic constraints, thereby accelerating the generation of simulated data for TL applications. Based on these analytical bearing models, several studies have proposed simulation-driven TL frameworks for bearing fault diagnosis [22–24]. However, the limited number of degrees of freedom (DOF) inherent in analytical models constrains their ability to simulate a wider range of defective conditions, thus limiting the applicability of TL approaches. For instance, analytical models often fall short in accurately capturing dynamic behavior associated with localized defects on rolling elements. The limitation arises primarily because rolling elements are typically modeled as a single aggregated body rather than as a set of discrete components. Moreover, localized defects are frequently represented in two dimensions, which constrains the analysis and fails to account for the complex three-dimensional interactions between individual rolling element and the defect. To overcome such weaknesses and thus obtain more detailed dynamic behaviors within TL frameworks, comprehensive bearing models incorporating a higher number of DOF, accounting for more complex non-linear dynamics, have been suggested [25, 26].

Multibody dynamics (MBD) simulation has been widely accepted as a comprehensive and accurate modelling technique in mechanical engineering [27]. It is particularly well-suited for bearing models and large-scale rotor systems, as it enables a precise representation of given diametral clearances and contact mechanisms [28–32]. Therefore, MBD simulations, characterized by the flexibility to accommodate diverse system configurations, offer significant advantages in constructing source domains for TL applications. Such models allow for simulations and control of various defective scenarios. However, an increase in model complexity significantly extends computational time and makes large-scale data generation more challenging. For instance, Vehviläinen et al. [28] developed a full-rigid-body MBD model for a ball bearing, using a polygonal contact method, to investigate the responses of an outer race fault in a bearing. Although this model successfully reproduced the distinct fault frequencies of the outer race defect, the average computational time required for an 8-second simulation was approximately 10 hours. Giraudo et al. [33] developed an MBD model with all DOF considered to investigate the dynamic behavior of spherical roller bearings with localized defects, precisely reproducing fault frequencies with a maximum deviation of 0.22% from theoretical values. However, the simulation was computationally expensive, requiring one day to compute 1.26 seconds. Hence, to strike a balance between computational efficiency and accurate representation of fault characteristics, an appropriate level of simplification for bearings in MBD models is essential, particularly for TL applications.

To the best of our knowledge, comprehensive MBD models suitable for simulation-driven diagnostic model have not yet been developed and tested [25]. Although existing MBD models accurately reproduce bearing fault frequencies, their computational cost is prohibitively high to generate large-scale datasets required for diagnostic model training, even when flexible bodies are neglected. In addition, most proposed approaches only simulate bearing dynamics rather than a complete rotor system, and many are only tested on public experimental datasets [7]. As a result, fault data generated from single-bearing models may not adequately represent real rotating machinery, particularly in cases involving compound faults from other components in rotating systems.

The present study aims to advance fault diagnosis in rotating machinery by using full-scale rigid-flexible MBD simulations to construct source-domain data for TL applications. The proposed MBD framework is to highlight its computational efficiency for full-scale simulations and its modularity for various machine operating scenarios. To do this, a dedicated test rig is established to acquire signals from the defective ball bearings. Next, a full-scale MBD model is developed to reproduce bearing fault characteristics, are evaluated and validated against experimental results. Finally, a comparative analysis of the proposed high-fidelity bearing model and an analytical bearing model is performed. Both bearing models are implemented within the same test rig framework to highlight the computational efficiency on one hand and the modularity of the MBD framework on the other. The comparison not only demonstrates the accuracy of the simulated fault characteristics, but also provides valuable insights for the future development of simulation-driven approaches in bearing fault diagnosis.

The remainder of this paper is organized as follows. Section 2 introduces the methodology for bearing fault analysis and describes the target fault characteristics in this study. Section 3 illustrates the experimental arrangements, including the configuration of the experimental test rig and the fabrication of defective ball bearings. The systematic development of the proposed MBD model is detailed in Section 4, covering geometric modelling, constraints and contact mechanisms. This section also presents the proposed high-fidelity bearing model and a 6-DOF analytical bearing model. Section 5 validates the proposed MBD model against the experimental measurements, followed by a comparative analysis of two bearing models to highlight their respective advantages and limitations. Finally, Section 6 provides the concluding remarks of this study.

2. Bearing fault analysis

A typical ball bearing comprises four primary components: an outer ring (OR), an inner ring (IR), rolling elements (REs) and a cage. In industrial applications, the OR is normally fixed on the bearing housing, while the IR rotates synchronously along with the shaft. The cage maintains a uniform spacing between the REs and rotates at a relative speed with respect to the IR.

When defects occur in ball bearings, they typically generate periodic impulses that appear at the characteristic frequencies when the REs roll over a local defect at a constant rotating speed in every revolution. The fault frequencies associated with the OR and IR are related to the ball-passing frequency of outer race (BPFO) and the ball-passing frequency of inner race (BPFI), which are the target fault frequencies in this study and defined as for instance in [34]:

$$\text{BPFO (Hz)} = \frac{nf_r}{2}\left(1 - \frac{d}{D}\cos(\alpha)\right), \tag{1}$$

$$\text{BPFI (Hz)} = \frac{nf_r}{2}\left(1 + \frac{d}{D}\cos(\alpha)\right), \tag{2}$$

where n is the number of rolling elements, f_r is the rotational frequency of the IR and identical to the rotational frequency of the shaft in this study, d is the diameter of the REs, D is the pitch diameter of the bearing, and α is the contact angle, respectively.

Rolling bearing fault signals are characterized by both amplitude and frequency modulation due to the nonstationarity of dynamic behavior. These signals usually interfere with the identification of bearing fault frequencies when applying the Fourier transform to the measurement data. Instead, signal demodulation needs to be conducted before applying the Fast Fourier Transform (FFT) [34–36]. In the field of bearing fault detection, envelope analysis (envelope spectrum) that uses the Hilbert transform to extract envelope signals, followed by FFT, is widely accepted as an efficient method to demodulate vibration signals and to identify their fault frequencies and harmonics [37]. In this study, therefore, both simulated and experimental results are presented after the envelope spectrum has been applied when comparing bearing fault frequencies. Since arbitrary impulse oscillations are removed, the main frequency components in the envelope spectra from the simulated signals are expected to closely correspond to those observed in the experimental results.

3. Experimental arrangement

In this study, ball bearing fault diagnosis was conducted using the experimental test setup illustrated in Fig. 1. The bearings used in this study were standard SKF 6007-2Z deep-groove ball bearings, featuring 11 REs and a bore diameter of 35 mm. The detailed specification of the bearing type is provided in Table 1. Two ball bearings were mounted at the drive end (DE) and non-drive end (NDE) of the shaft, with the faulty bearing placed at the NDE for fault analysis. The shaft had a total length of 600 mm and a diameter of 12 mm. Two discs were mounted on the shaft, each with a diameter of 150 mm, a thickness of 10 mm, and a mass of 1.69 kg. The first disc was located at 222.5 mm from the DE, while the second disc was located 420 mm from the same reference point, leading to a distance of 297.5 mm between the two discs.

An AC electric motor (Thurm K21R 63 K 2 H, 0.18kW) was mounted at the DE and connected to the shaft via a flexible coupling. The motor rotation speed was controlled by a three-phase inverter with an operating range of 0-2765 revolutions per minute (RPM). The shaft was monitored using four eddy current sensors (SKF-CMSS 785), arranged in horizontal and vertical pairs near the locations of the two weight discs. A remote optical LED sensor (ROS-W) was installed near the motor to measure the rotational speed. Additionally, a tri-axial accelerometer (PCB 356A03) was mounted on the NDE bearing housing to capture the vibration responses. The accelerometer was connected to a data acquisition unit (DeweSoft SIRIUS®), and the sampling rate was set to 10 kHz.

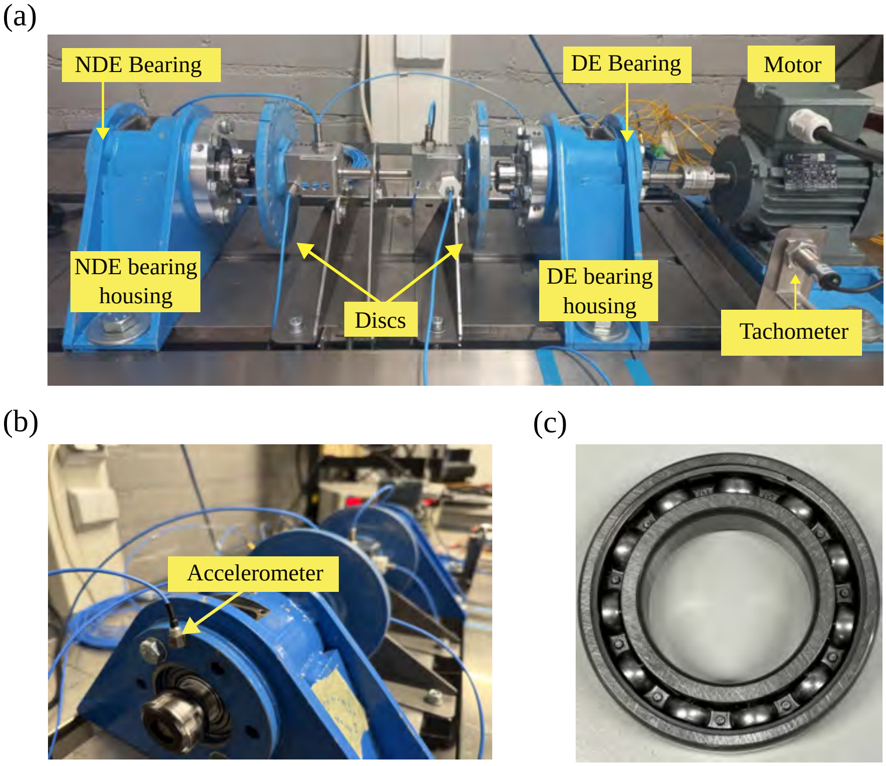


Fig. 1: The experimental setup of the test rig: (a) the arrangement of the sensors and the components (b) the location of the accelerometer (c) the standard SKF 6007-2Z deep-groove ball bearing.

Two ball bearings with a single localized fault, outer race fault (ORF) and inner race fault (IRF) were fabricated by laser machining as shown in Fig. 2. The machining process was conducted under identical operating parameters and cutting angles to ensure consistency. In both cases, the bearing faults were positioned approximately at, and oriented perpendicular to, the centerline of their respective raceways. Each fault measured approximately 2 mm in length, 0.5 mm in width and 0.2 mm in depth.

Four sets of vibration acceleration measurements were conducted under faulty bearing conditions at two distinct rotational speeds, as shown in Table 2. The nominal motor speeds were set to 600 RPM ($f_r = 10$ Hz) and 1200 RPM ($f_r = 20$ Hz). However, the actual speeds of the shaft were measured as 588 RPM ($f_r = 9.8$ Hz) and 1188 RPM ($f_r = 19.8$ Hz), respectively. These experimentally determined rotational speeds were subsequently adopted in the simulations. All acceleration signals were recorded after the shaft speed stabilized at the target rotational speeds.

Table 1: Specifications of SKF 6007-2Z deep-groove ball bearing.

Denotation	Dimension	Value (mm)
d	Ball diameter	7.94
D_o	Outside diameter	62
D_b	Bore diameter	35
W_r	Ring width	14
D	Pitch diameter	48
R_{out}	Radius of outer race way	3.99
R_{in}	Radius of inner race way	3.99
C_r	Radial internal clearance	0.03

Fig. 2: The artificial ORF (left) and IRF (right) are created by laser machining.

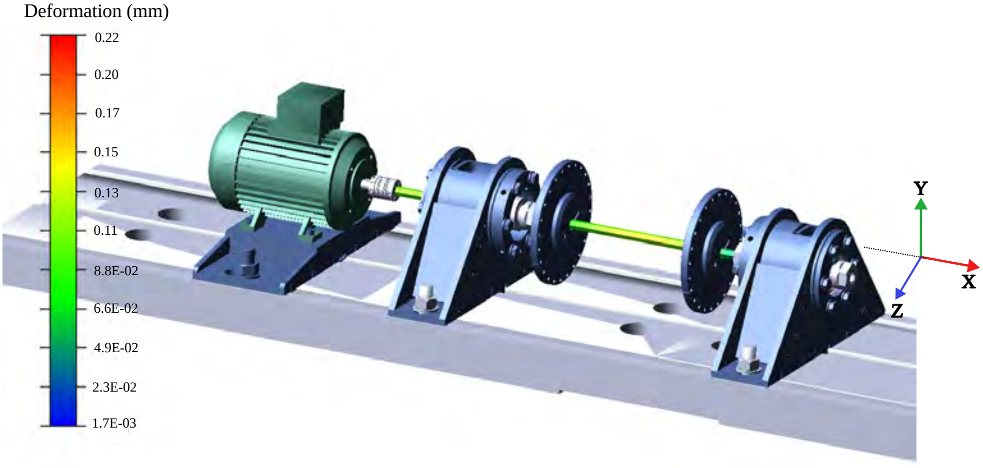


Fig. 3: The full-scale rigid-flexible MBD model. The Y and Z axes are the radial direction in the global coordinate, and the X axis is the longitudinal direction in the global coordinate.

Table 2: Experimental measurement sets.

Set	Condition	Rotational frequency f_r (Hz)
1	ORF	9.8
2	ORF	19.8
3	IRF	9.8
4	IRF	19.8

4. Development of the rigid-flexible hybrid MBD model

To replicate the experimental setup and reproduce the simulated vibration signals, a full-scale MBD model is built using the MBD analysis software, MSC Adams/View (Adams), with all dimensions consistent with those of the physical system. In this model, the shaft is modelled as a flexible body, while other components are considered to be rigid bodies. To further highlight the benefits of this MBD framework for data generation, two different bearing models are introduced, a high-fidelity (hi-fi) bearing model and an analytical bearing model. The hi-fi bearing model is designed to incorporate more realistic dynamic behavior with efficient computational cost, while the analytical bearing model serves as a benchmark for comparative analysis between two models. The development of the two bearing models is detailed in Section 4.2 and Section 4.3.

4.1. Connections and force elements

The complete assembly includes a shaft, a motor, two bearing housings, one simplified DE bearing, one hi-fi NDE bearing, two disks and a supporting platform. The MBD model is developed as shown in Fig. 3. Initially, a 3D CAD geometry of the test rig is built based on the dimensions of the physical system. This CAD file is then imported into Adams to create the full-scale MBD model.

The model topology is presented in Fig. 4. The shaft is supported by two bearings - a DE bearing and an NDE one. Since the NDE bearing is the primary objective in the experimental setup, the dynamic behavior of the DE components is not discussed further here. Therefore, the DE bearing is simplified as a bushing element, connecting to the bearing housing 1 and the sleeve 1. It comprises three rotational and three translational spring-damper elements to reduce the computational load. The rotational stiffness and damping values are set to zero in all three directions. In the longitudinal direction (X axis in Fig. 3), the translational stiffness and damping values are set to zero, thereby allowing the NDE bearing to carry the axial load. In the radial direction (Y and Z axes in Fig. 3), the translational stiffness and the damping values are set to $2 \cdot 10^4$ N/mm and $2 \cdot 10$ Ns/mm, respectively, to approximate the maximum radial clearance in the actual bearing [28]. In terms of the NDE bearing, the OR is fixed on the bearing housings, while the IR is fixed on the sleeve 2 and rotates about the same axis at the same speed as the shaft.

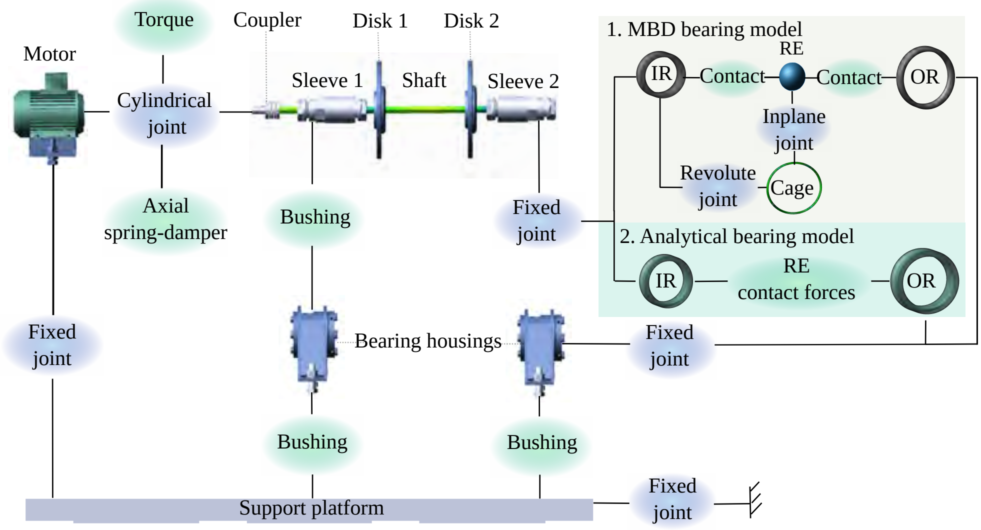


Fig. 4: The MBD model topology. Blue ovals represent joint connection and green ovals represent force elements.

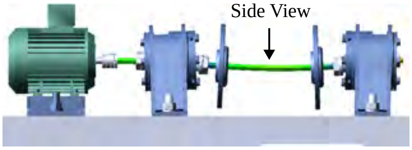
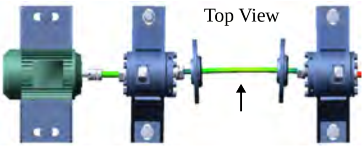
The DE of the shaft is connected to the motor with a cylindrical joint and a spring-damper element, allowing rotation about the longitudinal direction and longitudinal displacement under preload conditions. The velocity-dependent driving torque is applied on this cylindrical joint directly. The spring-damper element is used to control the shaft displacement and to reduce the additional longitudinal vibration. The stiffness and damping values of this spring-damper element are set to 400 N/mm and 20 Ns/mm, respectively.

The NDE of the shaft is supported by the NDE bearing, which is modelled using both the MBD bearing model and the analytical bearing model, as described in detail in Section 4.2 and 4.3. Two disks are fixed on the shaft between the two bearing housings and rotate about the same longitudinal axis as the shaft.

To obtain the vibration responses from the bearing housings, the interfaces between the housings and the supporting platform are modelled using undamped bushing elements. As the actual stiffness and damping properties of these connections are not directly measurable, the bushing parameters are systematically calibrated to match the two dominant modal frequencies obtained from experimental modal analysis of the test rig. The modal frequencies after calibration are presented in Table 3, where the translational stiffness values in the X, Y and Z directions are set to $1 \cdot 10^4$ N/mm, $6 \cdot 10^3$ N/mm and $9 \cdot 10^3$ N/mm respectively, and the corresponding rotational stiffness values are assigned as $5 \cdot 10^5$ $\mathrm{N \cdot mm/deg}$, $1 \cdot 10^5$ $\mathrm{N \cdot mm/deg}$ and $5 \cdot 10^4$ $\mathrm{N \cdot mm/deg}$ about the X, Y and Z axes. This calibration ensures that the simulated structural boundary conditions approximate the dynamic behavior of the physical system.

The flexible shaft is built using ViewFlex which is an finite element analysis tool based on MSC Nastran in Adams. The flexible shaft is modelled using 53 modes, with a damping ratio of 0.01 assigned to all modes below 100 Hz and 0.1 assigned to modes in the frequency range 100-1000 Hz. To solve the DAEs, the Hilber-Hughes-Taylor implicit Method (HHT-α) [38] is used in the numerical solver with an error tolerance of 10^{-7}, a maximum time step of 10^{-6} and a maximum of 50 iterations for convergence.

Table 3: The modal frequencies after calibration in the simulation against the experimental modal analysis.

	1^{st} shaft vertical bending mode (Hz)	1^{st} shaft lateral bending mode (Hz)
Experiments	31.25	33.75
Simulations	33.82	34.93
		

4.2. The high-fidelity MBD bearing model

The complete assembly of the bearing model is shown in Fig. 5. The 3D geometry of the IR and OR are obtained from the manufacturer. The REs are modelled as rigid spheres, and the cage is modelled as a rigid torus and is attached from its center point to the IR by a revolute joint. Based on the investigations in [39], the contact forces between the cage pockets

Fig. 5: The proposed hi-fi ball bearing model.

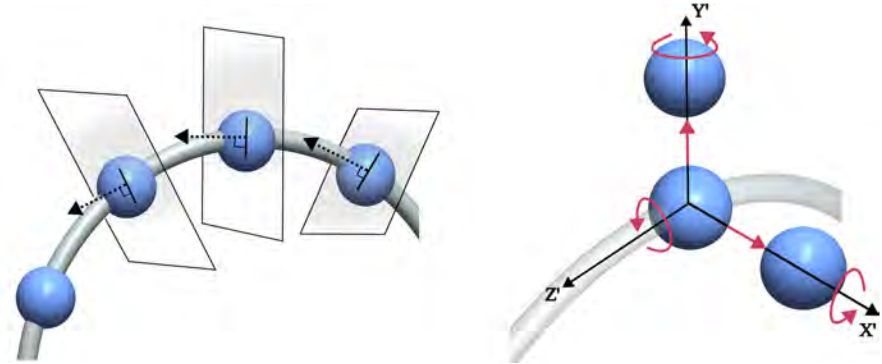


Fig. 6: Each RE is constrained by an in-plane joint (left), which rotates along with the cage about the same axis and eliminates the DOF about the perpendicular direction (black arrow). In each in-plane joint (right), five red arrows depict the corresponding DOF regarding three rotations and X'-Y' plane translations.

and the REs are neglected in this study, as they significantly increase the computational time without substantially improving the accuracy of the simulated fault responses. To simplify the cage structure, an in-plane joint is applied to represent the interaction between REs and the cage following the work in [28]. Each RE is connected to the cage by an in-plane joint, which allows for rotations about three axes and translations along two local radial directions of the cage, as shown in Fig. 6. Each in-plane joint is attached to the cage in its body-fixed reference frame (X'Y'Z' reference frame in Fig. 6) at the corresponding RE position and rotates along with the cage. In other words, the cage maintains equal angular spacing between the REs, as each RE is constrained to remain within the same plane.

The contact force settings are defined between the IR and the REs, and the OR and the REs. The IMPACT function in Adams is used to define the contact mechanism between two objects, and the contact force F_{IMPACT} is defined as:

$$F_{\text{IMPACT}} = \begin{cases} 0, & x > x_1, \\ \text{Max}\left[0, k(x_1 - x)^e - \dot{x} \cdot STEP(x, x_1 - d_{pen}, c_{\max}, x_1, 0)\right], & x \leq x_1, \end{cases} \tag{3}$$

where

$$STEP(x, x_1 - d_{\text{pen}}, c_{\max}, x_1, 0)) = \begin{cases} c_{\max}, & x \leq x_1 - d_{\text{pen}}, \\ c_{\max} - c_{\max}\lambda^2(3 - 2\lambda), & x_1 - d_{\text{pen}} < x < x_1, \\ 0, & x \geq x_1, \end{cases} \tag{4}$$

where x denotes the current distance between two objects and x_1 is the nominal distance of x. The parameters k, e, $c_{\max}$, and d_{pen} correspond to the contact stiffness, the exponent of the force-deformation characteristic, the maximum damping coefficient, and the penetration depth, respectively. The parameter λ is defined as $(x - x_1 + d_{\text{pen}})/d_{\text{pen}}$. The IMPACT function is activated when the distance between two objects is smaller than x_1, i.e, when two objects are colliding. The contact force consists of two components: a stiffness component and a damping component. The stiffness component is a function of the penetration of two objects, acting to resist further penetration, as shown in Fig. 7(a). The parameter e characterizes the non-linear stiffness behavior, representing a stiffening ($e > 1$) or softening ($e < 1$) spring response. The damping component is a function of the relative speed of two objects during collision, which opposes the relative motion of the objects. To avoid discontinuities in the damping force during contact, the damping coefficient is defined as a cubic step function ($STEP$) of penetration, as shown in Fig. 7(b). Once the penetration exceeds d_{pen}, the contact damping coefficient remains constant at its maximum value $c_{\max}$. The parameters are discussed in Section 5.1.

The contact mechanisms of the rolling bearings are detailed in [40]. For ball bearings, the RE-raceway contact force F is calculated as:

$$F = k_{\text{total}}\delta^{1.5}, \tag{5}$$

where k_{total} is the total contact stiffness and δ is the contact deflection. The value of k_{total} takes into account the contacts between REs and both raceways, which can be expressed as [40]:

$$k_{\text{total}} = \frac{1}{\left[(1/k_{\text{in}})^{2/3} + (1/k_{\text{out}})^{2/3}\right]^{3/2}}, \tag{6}$$

where k_{in} and k_{out} denote the contact stiffness between the RE and the inner and outer raceways. Accordingly, in Eq. (3), the contact stiffness k is set to k_{total} and the force-deformation exponent e is taken as 1.5.

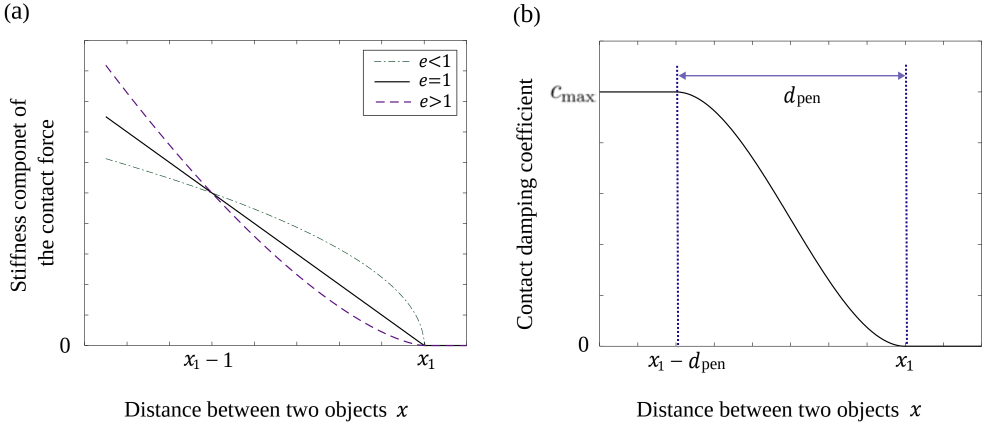


Fig. 7: The visualization of the IMPACT function in MSC Adams.

A velocity-based friction model is applied to the ball-raceway interfaces to approximate the lubrication effects [41]. This friction model presents a similar approach to smooth Coulomb friction model and it requires four values to define: the static friction coefficient μ_s, the dynamic friction coefficient μ_d, the stiction transition velocity V_s and the friction transition velocity V_d, as demonstrated in Fig. 8. The values of μ_s and μ_d are tuned to minimize slip in the ball-raceway contacts. The parameters V_s and V_d are usually set to small values to approach to smooth Coulomb friction in simulations.

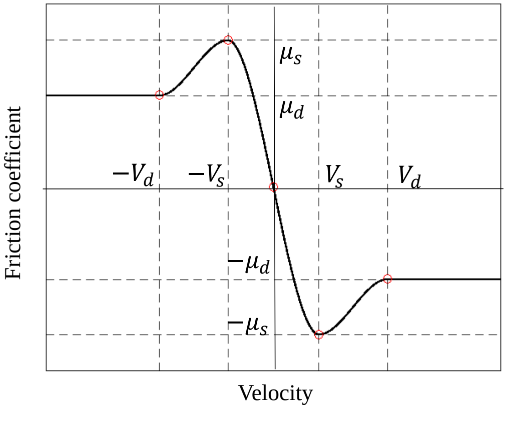


Fig. 8: Velocity-based friction model.

The axial preload plays a critical role in reducing the RE slip and promoting continuous ball-raceway contacts, especially in our case with simplified lubrication conditions. Accordingly, a constant axial load is applied at the center of mass of the IR in the simulations. In the physical test rig, the axial preload is only provided by the fixation screws; however, quantifying this actual preload is challenging. Based on SKF's documentation, the axial preload F_{pre} is suggested as:

$$F_{\text{pre}} = k_{\text{pre}} D_b, \tag{7}$$

where k_{pre} is a parameter tha is suggested to lie in a range of 0.005 to 0.01 kN/mm for small electric motors, and D_b is the bore diameter of the bearing. Given that $D_b = 35$ mm, the corresponding range of the preload force F_{pre} can be set between 175 N and 350 N. Due to the simplified cage constraints in the model, the REs are allowed to move more freely than in reality, where their motions are physically restricted by the cage pockets. Consequently, when preload is applied, the resulting relative displacement between the IR and the REs induces an initial contact angle, which subsequently affects the cage rotational frequency f_{cage} defined as [40]:

$$f_{cage} = \frac{f_r}{2}\left(1 - \frac{d}{D}\cos(\alpha)\right), \tag{8}$$

where an increase in the contact angle α leads to a corresponding increase in both the cage rotational frequency and the BPFO. In addition to the axial load, the radial loads of the simulated bearing are contributed by the gravitational force and the inertial forces.

Following the artificial defective bearings in Fig. 2, the faults are modelled by introducing cut-off geometries to replicate the dimensions of the actual faults, as shown in Fig. 9.

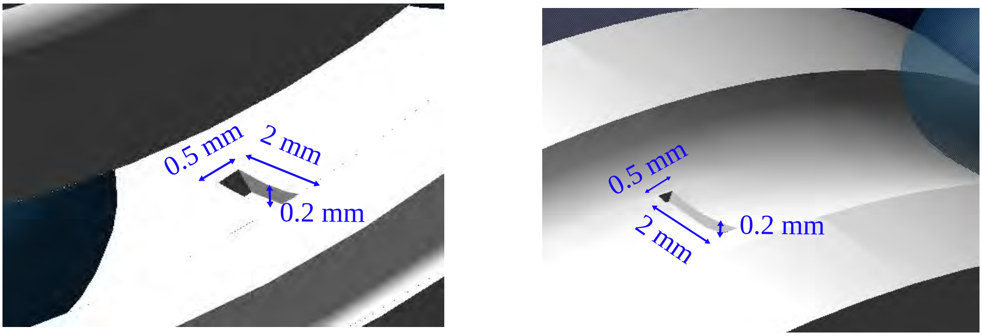


Fig. 9: Simulated faults in the hi-fi bearing model: ORF (left) and IRF (right).

4.3. The 6 DOF analytical model of ball bearing

To further validate the hi-fi bearing model and examine the corresponding fault characteristics, an analytical bearing model is introduced in this study. The analytical ball bearing model incorporates both an elastohydrodynamic film and non-Hertzian contact deformations, as detailed in [20, 42]. In Fig. 10, the 6 DOF bearing model considered three bearing force components along the X, Y and Z axes, torques about the Y and Z axes, as well as the bearing friction torque around the X axis. Moreover, it captures the effects of the fault length and depth on the additional displacement excitation induced by the localized defects. A step-function model is developed to describe the additional displacement excitation, and an example of an IRF is shown in Fig. 11, where h_{defect} represents the depth of the fault, α_1, and α_2 are the angles calculated by

$$\alpha_1 = \frac{L_{\text{defect}}}{R_{in}} \left(1 - \frac{\beta_i}{\theta_{in}}\right), \tag{9}$$

$$\alpha_2 = \frac{1}{2}\alpha_1, \tag{10}$$

where L_{defect} is the length of defect, R_{in} is the radius of inner raceway, β_i is the azimuth angle of ball i and θ_{in} is the rotation angle of the IR.

In the comparative analysis, the hi-fi bearing model is replaced with this bearing model at the same location, with the IR fixed on sleeve 2 and the OR fixed on the NDE housing. In addition, unlike the hi-fi bearing model where the DOF of REs are explicitly considered, this analytical model does not consider preload effects. Instead, the RE-raceway interactions are represented using theoretical contact forces that follow the predicted rotational frequencies.

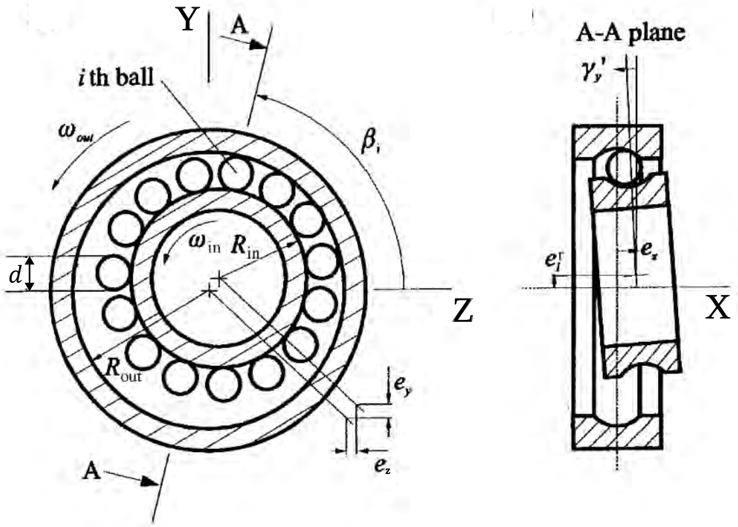


Fig. 10: Axial (left) and lateral (right) cross-section of the 6 DOF analytical model [20].

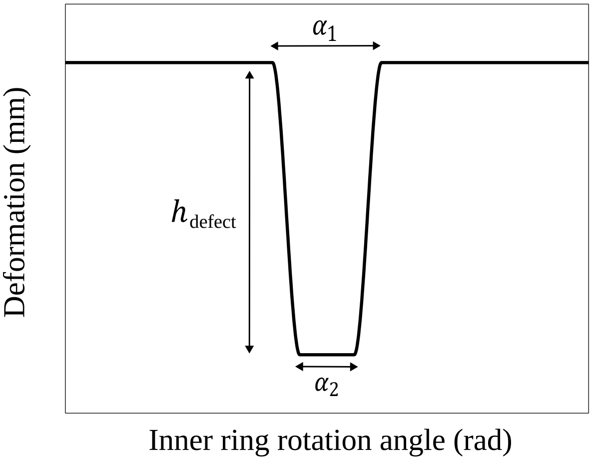


Fig. 11: Shape of the defect in IR as a function of the angle of rotation of IR [20].

5. Results

First, the contact parameters of the hi-fi bearing model, introduced in Section 4.2, are validated, and the influence of axial preload is also examined. Subsequently, the full-scale MBD model is validated against the experimental measurements by comparing the dominant bearing fault frequencies, BPFO and BPFI, using the envelope spectra described in Section 2. For both experiments and simulations, acceleration responses are obtained from the same measurement location at given driving speeds to ensure consistency. Similarly, the 6 DOF analytical bearing model is integrated into the same MBD framework to further validate the hi-fi bearing model and to compare the relative computational costs. Finally, the advantages and limitations of both bearing models are evaluated.

5.1. Validating the contact parameters of the hi-fi bearing model

Based on the model development in Section 4.2, all contact parameters have been calculated and are summarized in Table 4. According to the geometry of the SKF 6007-2Z ball bearing, the contact stiffness between the RE and the inner race, k_{in}, and that between the RE and the outer race, k_{out}, are calculated to be $8.184 \cdot 10^5$ N/mm and $8.526 \cdot 10^5$ N/mm, respectively [40]. From Eq. (6), the total contact stiffness k_{total} is calculated to be $2.953 \cdot 10^5$ N/mm, which is used for the contact stiffness k in Eq. (3). Moreover, based on the recommendations in [40], the penetration depth d_{pen} is set to 0.01 mm, and the exponent of the force-deformation characteristic e is set to 1.5 from Eq. (5). However, determining the exact value of contact damping remains challenging. Based on the sensitivity analysis performed in [43], the maximum contact damping coefficient c_{max} can be a proportion of the value of total contact stiffness k_{total}. The findings indicated that setting the c_{max} to 1/1000 of the contact stiffness k_{total} resulted in more stable vibration responses. Therefore, in this study, the maximum contact damping coefficient c_{max} is set to 1/1000 of contact stiffness, resulting in $c_{\text{max}} = 295.3$ Ns/mm for the hi-fi bearing model. Furthermore, the static and dynamic friction coefficients

($mu_s = 0.05$ and $mu_d = 0.03$) are tuned to minimize slip of the REs. The transition velocities ($V_s = 0.1$ mm/s and $V_d = 1$ mm/s) are set to maintain a consistent dynamic friction regime.

Table 4: Contact parameters used in this study.

Contact parameters	Denotation	Values
Contact stiffness	k_{total}	$2.953 \cdot 10^5$ N/mm
Contact damping coefficient	$c_{\max}$	295.3 Ns/mm
Force exponent	e	1.5
Penetration depth	d_{pen}	0.01 mm
Static friction coefficient	μ_s	0.05
Dynamic friction coefficient	μ_d	0.03
Stiction transition velocity	V_s	0.1 mm/s
Friction transition velocity	V_d	1 mm/s

Due to the simplified cage constraints, the relationship between the contact angle and the axial preload is investigated. To evaluate the influence of the axial preloads, the contact angles are calculated using the vertical and lateral contact forces in the local X'-Y' plane, as presented in Fig. 6. The cage rotational frequencies are extracted from its center of mass. The resulting contact angles of one RE under different preloads are shown in Fig. 12(a) and (b). First, the contact angle decreases as the RE enters the loaded zone, where the RE is subjected to the shaft load. This additional vertical load increases the vertical contact force and consequently reduces the contact angle. Conversely, when the RE is rolling into the unloaded zone, the contact angle increases. Second, the fluctuation in contact angle is more pronounced under lower preload, since a larger preload suppresses the REs' transverse motion and leads to a more stable contact behavior. This result is also reflected in Fig. 12(c) and (d), where the cage rotational frequency fluctuates less when $F_{\text{pre}} = 350$ N. Third, comparing Fig. 12(a) and (b), the variation of the IR rotational frequency f_r has a negligible influence on the contact angle.

Theoretical cage rotational frequencies are calculated by using the average contact angles from Fig. 12(a) and (b) as a input for the α in Eq. (8), while simulated cage rotational frequencies are extracted from Fig. 12(c) and (d). A comparison of the simulated cage frequencies and the theoretical predictions is presented in Table 5, demonstrating strong agreement and validating the accuracy of the MBD model in representing REs and cage dynamics.

Based on this investigation, a preload of 350 N is selected for all subsequent bearing fault analysis, as it ensures more stable contact dynamics and consistent cage rotational behavior. In addition, reducing contact angle fluctuations ensures that the REs are able to consistently hit both the ORF and the IRF during each revolution.

To verify that all contact parameters are appropriately defined and that fault signals can be accurately reproduced, the contact forces between each RE and the raceway are examined. As an example, the vertical contact forces during two revolutions for a single RE under the axial preload $F_{\text{pre}} = 350$ N and IR rotational frequency $f_r = 9.8$ Hz in both faulty cases, as shown in Fig. 13. In the ORF case, the vertical contact force varies as the RE orbits within the bearing, with peaks occurring when it passes over the fault located in the loaded zone. The time interval between two peaks is 0.238 s, matching the simulated cage frequency of 4.2 Hz from Table 5 which corresponds to a revolution period of 0.238 s. In the IRF case, the time interval between two peaks is 0.179 s, corresponding to a frequency of 5.6 Hz. This matches the theoretical BPFI from Eq (2), calculated using $n = 1$ and $\alpha = 28.87$ deg, confirming the accuracy of the fault signal representation.

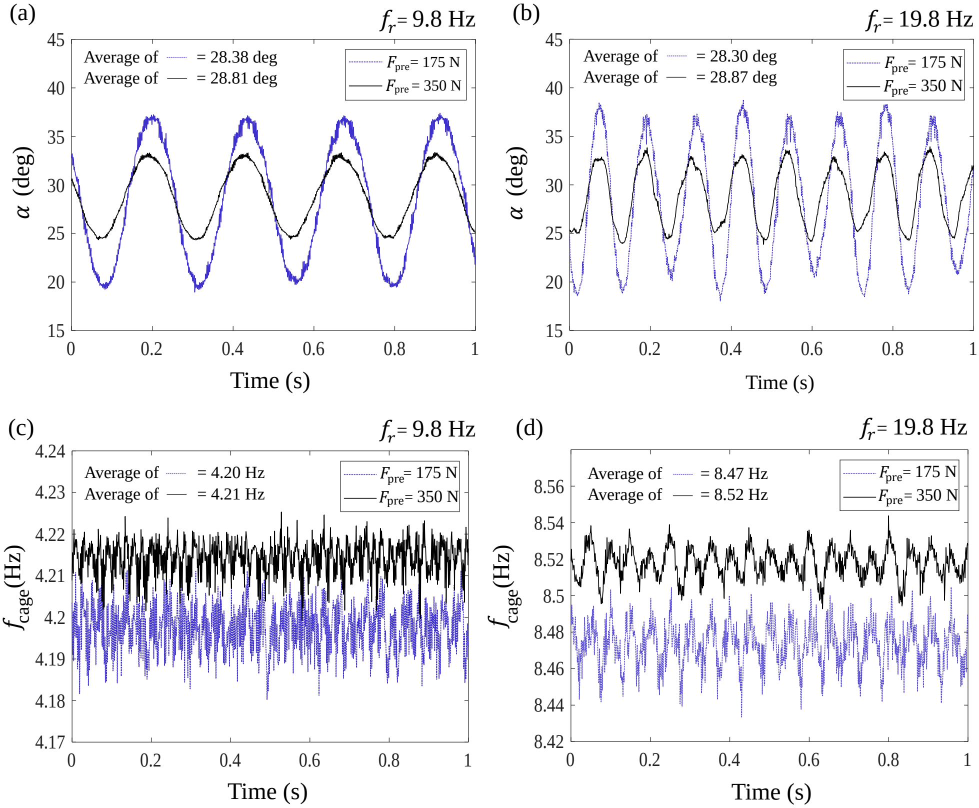


Fig. 12: Influence of two different axial preloads on the contact angles and the cage rotational frequency at different f_r: (a) the contact angles of one RE in the local X'-Y' plane under different preloads when $f_r = 9.8$ Hz (b) the contact angles of one RE in the local X'-Y' plane under different preloads when $f_r = 19.8$ Hz (c) the cage rotational frequencies under different preloads when $f_r = 9.8$ Hz (d) the cage rotational frequencies under different preloads when $f_r = 19.8$ Hz.

Table 5: Comparison between simulated and theoretical cage speed.

f_r (Hz)	F_{pre} (N)	Average contact angle α_{avg} (Deg)	Simulated cage frequency (Hz)	f_{cage} (Hz) $\alpha = \alpha_{\text{avg}}$	f_{cage} (Hz) $\alpha = 0$
9.8	175	28.38	4.20	4.20	4.10
	350	28.81	4.21	4.20	4.10
19.8	175	28.30	8.47	8.48	8.28
	350	28.87	8.52	8.49	8.28

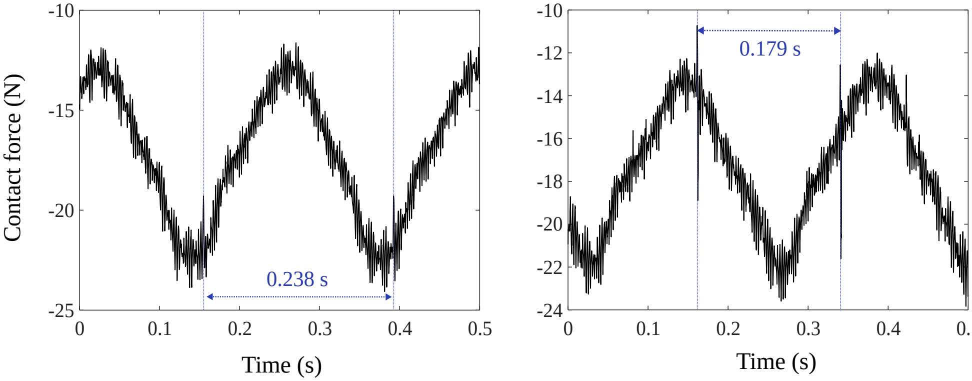


Fig. 13: Vertical RE-raceway contact forces in the local X'-Y' plane for the ORF bearing (left) and the IRF bearing (right), at F_{pre}=350 N and f_r = 9.8 Hz.

5.2. Model validation of faulty bearings using experimental data

A comparison between experimental results and the MBD faulty bearing models at 588 RPM (f_r = 9.8 Hz) and 1188 RPM (f_r = 19.8 Hz) are presented in Fig. 14 and 15. To simplify the data analysis, the envelope spectra are calculated using only the vertical acceleration (Y axis) measured at the NDE bearing housing. To highlight the key features, the rotational frequencies, BPFOs, BPFIs, and their respective harmonics are marked in the envelope spectra. Both simulated and experimental data are sampled at 10 kHz for 40 seconds during a 5-second time window corresponding to the steady rotational speed. The ORF is located in the loaded zone in both simulations and experiments.

The results for the ORF case are presented in Fig. 14. It is worth highlighting that both the experimental and simulated results demonstrate similar characteristics, with dominant frequency components corresponding to the BPFO across different rotational speeds. However, in Fig. 14 (b1) and (b2), the second BPFO amplitude exceeds the first BPFO, which contrasts with the experimental results. This discrepancy can be attributed to the simplified geometry of the simulated ORF, which produces two impacts when each RE roll in and out the defect, thereby amplifying the second BPFO. Another possible explanation is a mismatch between the damping characteristics of the model and those of the physical test rig. It is likely that the model underestimates damping effects at 2BPFO, resulting in the overestimation of

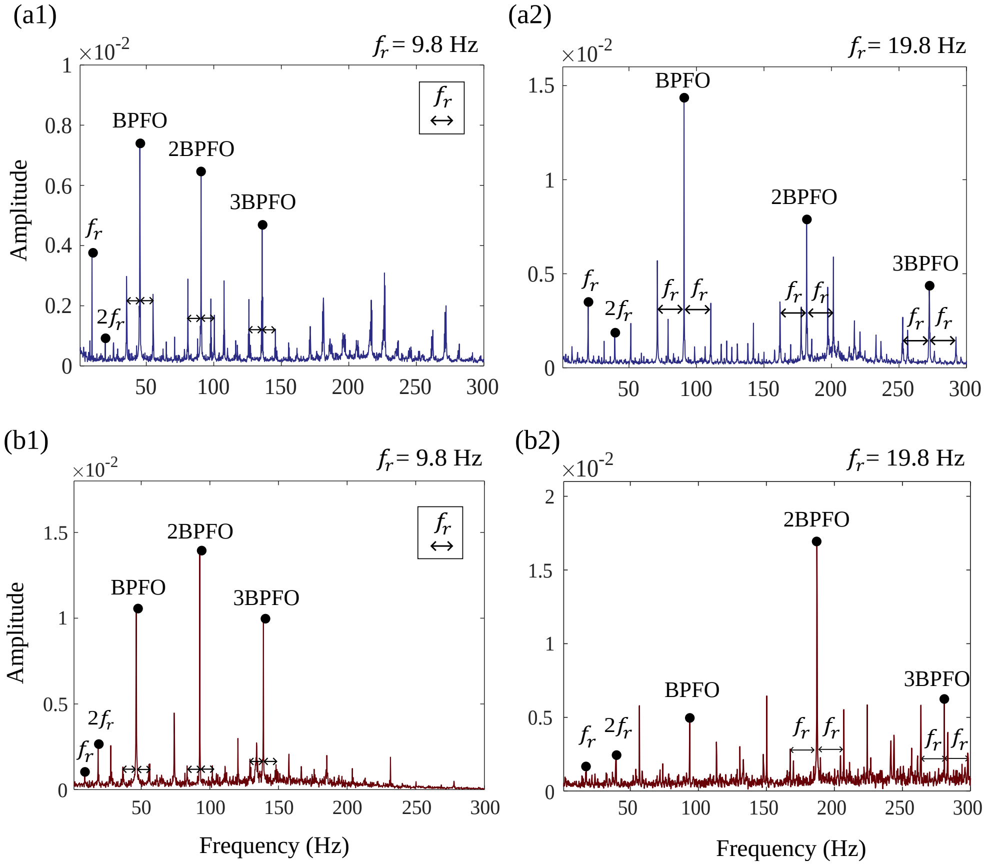


Fig. 14: Comparison between the experimental and the simulated data with the ORF: (a1) the envelope spectrum of the experimental data when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the experimental data when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the simulated data when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the simulated data when $f_r = 19.8$ Hz.

harmonic amplitudes in the simulated spectra. In contrast, the real ORF may contain small irregularities, which produces an impact profile closer to a half-sine pulse, thereby leading to a different harmonic distribution.

Additionally, the simulated amplitude in Fig. 14 (b1) and (b2) are higher than the experimental measurements in Fig. 14 (a1) and (a2). This discrepancy can be attributed to the difficulty in replicating the exact RE-raceway contact forces and the actual spring-damper properties in simulations. Although notable differences in signal amplitude can be observed in both experimental and simulated signals, they are not expected to adversely affect the performance of TL in the future, as the presence and accuracy of fault frequency components are more critical than matching matching absolute amplitude [12]. Overall, the proposed MBD models successfully reproduce the BPFO features despite these differences.

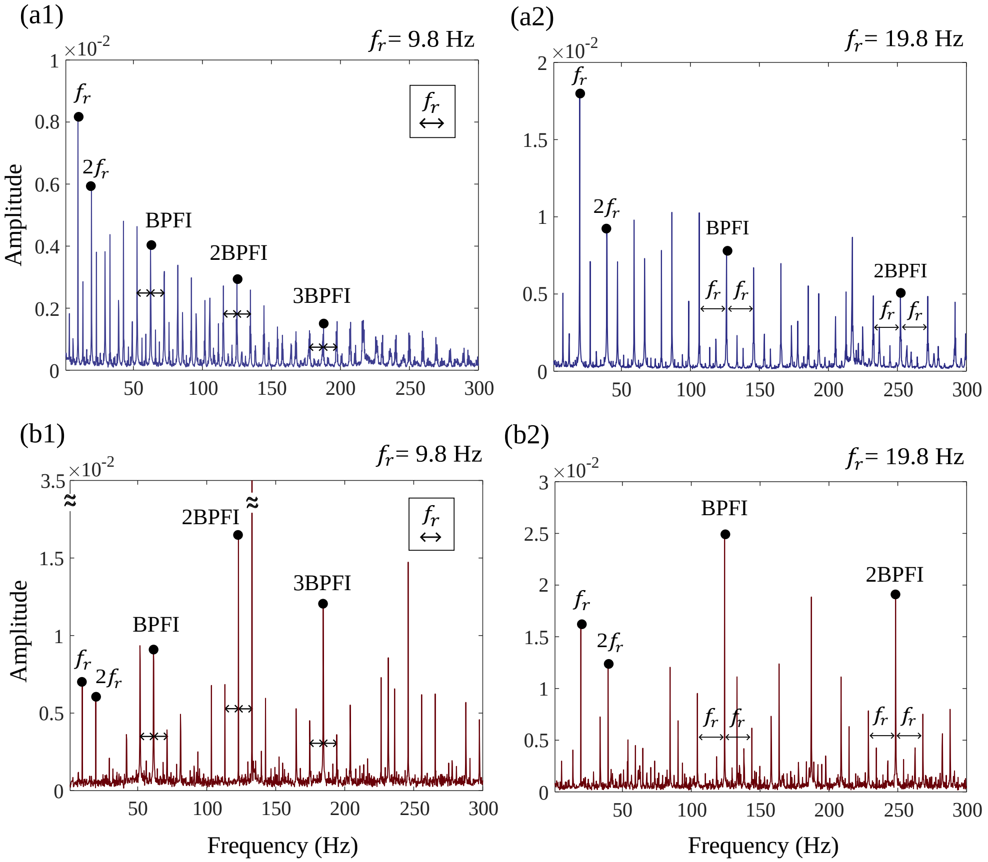


Fig. 15: Comparison between the experimental and the simulated data with the IRF: (a1) the envelope spectrum of the experimental data when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the experimental data when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the simulated data when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the simulated data when $f_r = 19.8$ Hz.

The results for the IRF case are presented in Fig. 15. In both the experimental and simulated envelope spectra, the characteristic of BPFI signature is clearly visible, together with its harmonics and sidebands. Notably, in Fig. 15 (a1) and (a2), all sidebands of the BPFI are more pronounced than those observed in the BPFO case. This behavior reflects the stronger modulation of the BPFI by the shaft rotational frequency, which is attributed to the varying contact forces between the REs and the IRF during the rotations. One sideband of the 2BPFI component in Fig. 15 (b1) exhibits a relatively higher amplitude than the others, as it coincides with the natural frequency of the housing in the vertical direction, resulting in resonance amplification. As observed in Fig. 15 (b1) and (b2), the relative amplitudes of BPFIs (BPFI–3BPFI) in the simulation do not perfectly match the experimental results, primarily due to simplifications in the defect geometry and contact modeling. Nevertheless, the

MBD model successfully reproduces the BPFI and the corresponding modulation sidebands, demonstrating its capability to capture the essential dynamic characteristics associated with IRFs.

Table 6 summarizes the characteristic frequencies obtained from the theoretical predictions (Eq. (1) and (2)), the experimental measurements, and the simulation results. The experimental results closely match the theoretical predictions, confirming that the contact angle of the deep-groove ball bearing in this study is approximately zero. Based on the discussion in Section 5.1, as expected, the simulated BPFOs are higher than the experimental results while the BPFIs are slighly lower than the experimental results, with approximately 2% error in both cases. However, when the simulated contact angle is incorporated into Eq. (1) and Eq. (2), the simulated results show strong agreement with the theoretical predictions, validating the proposed MBD model in reproducing faulty bearing behavior.

Table 6: Characteristic frequencies of experimental and simulated results at different rotating speeds.

f_r (Hz)	Cases	BPFO (Hz)	2BPFO (Hz)	3BPFO (Hz)	BPFI (Hz)	2BPFI (Hz)	3BPFI (Hz)
9.8	Theoretical values ($\alpha = 0$)	45.08	90.16	135.24	62.72	125.44	188.16
	Experiments	45.30	90.60	135.80	62.40	124.6	187.4
	Theoretical values ($\alpha = 28.3$ deg)	46.13	92.26	138.4	61.67	123.34	185.01
	MBD model	46.40	92.60	139.00	61.40	122.8	184.40
19.8	Theoretical values ($\alpha = 0$)	91.08	182.16	273.24	126.72	253.44	-
	Experiments	90.80	181.70	272.5	126.2	252.2	-
	Theoretical values ($\alpha = 28.87$ deg)	93.30	186.6	279.9	124.50	249.00	-
	MBD model	93.00	186.00	276.00	124.2	248.4	-

5.3. Comparative analysis between two different bearing models

After validating the MBD model against the experimental measurements, the next step is to investigate the differences between the hi-fi bearing model and the analytical 6 DOF bearing dynamic model, with particular emphasis on bearing fault characteristics and computational efficiency. For this comparison, a 6 DOF analytical bearing model represents the NDE bearing, while all other parameters are kept consistent with the analysis described in Section 5.2. Here, to reflect the dimensions of the actual fault, L_{defect}, which is the defect width of the real fault, and h_{defect} are set to 0.5 mm and 0.2 mm, respectively. For clarity, the MBD model incorporating a hi-fi bearing model is hereafter referred to as hi-fi BM, which has a 6 DOF dynamic bearing model as 6DOF-BM.

The comparisons between the two bearing models with an ORF at given rotational frequencies are presented in Fig. 16. The hi-fi BM exhibits fault characteristics similar to those of the 6DOF-BM, with the dominant frequency components corresponding to the BPFO. These feature patterns indicate that the detailed bearing model is capable of reproducing similar fault-related features. However, the noise level in the hi-fi BM is noticeably higher than in the 6DOF-BM. Most of this noise originates from the ball-raceway interactions, which generates additional small vibrations that are not fully damped in simulations. In contrast, the 6DOF-BM represents ball-raceway contacts as idealized impacts with an elastohydrodynamic film, resulting in a lower noise level and clearer BPFI harmonics observed in Fig. 16(a) compared to Fig. 16(b).

For the IRF case as shown in Fig. 17(a1) and (a2), the 6DOF-BM successfully reproduces the BPFI harmonics and sidebands, and their distribution is consistent with the experimental data. Likewise, the simulated fault frequency patterns in Fig. 17(b1) and (b2) still show the comparable results from the hi-fi BM. The BPFIs are more pronounced due to the applied preload, as the increased contact forces between the REs and the IRF enhance the fault-induced impacts.

Table 7 summarizes the fault frequencies extracted from both bearing models. The 6DOF-BM generates exact fault frequencies, since it relies directly on theoretical impact calculations while simplifying the complex contact mechanism between REs and raceways. In contrast, the proposed MBD model captures the fault frequencies with an offset of 2% compared to the experimental measurements.

To compare the computational time, three cases are run using a desktop equipped with an Intel i9-13900 processor, 64 GB RAM, and an NVIDIA GeForce RTX 3070 GPU. Table 8 demonstrates the efficient computational cost of the proposed MBD framework, highlighting its suitability for large-scale fault simulation studies. The simulation time of the hi-fi BM is only 1.13 times compared to the 6DOF-BM, it is acceptable for generating massive simulated data and is capable of accommodating more complex fault conditions, such as RE faults and multi-fault scenarios.

Table 7: Characteristic frequencies of both bearing models at different rotating speeds.

f_r (Hz)	Cases	BPFO (Hz)	2BPFO (Hz)	3BPFO (Hz)	BPFI (Hz)	2BPFI (Hz)	3BPFI (Hz)
9.8 Hz	Theoretical values ($\alpha = 0$ deg)	45.08	90.16	135.24	62.72	125.44	188.16
	6DOF-BM	45.00	90.20	135.20	62.80	125.40	188.20
	Theoretical values ($\alpha = 28.3$ deg)	46.13	92.26	138.4	61.67	123.34	185.01
	Hi-fi BM	46.40	92.60	139.00	61.40	123.4	186.00
19.8 Hz	Theoretical values ($\alpha = 0$ deg)	91.08	182.16	273.24	126.72	253.44	-
	6DOF-BM	91.00	182.20	273.20	126.80	253.40	-
	Theoretical values ($\alpha = 28.87$ deg)	93.30	186.6	279.9	124.50	249.00	-
	Hi-fi BM	93.00	186.00	276.00	125.2	250.4	-

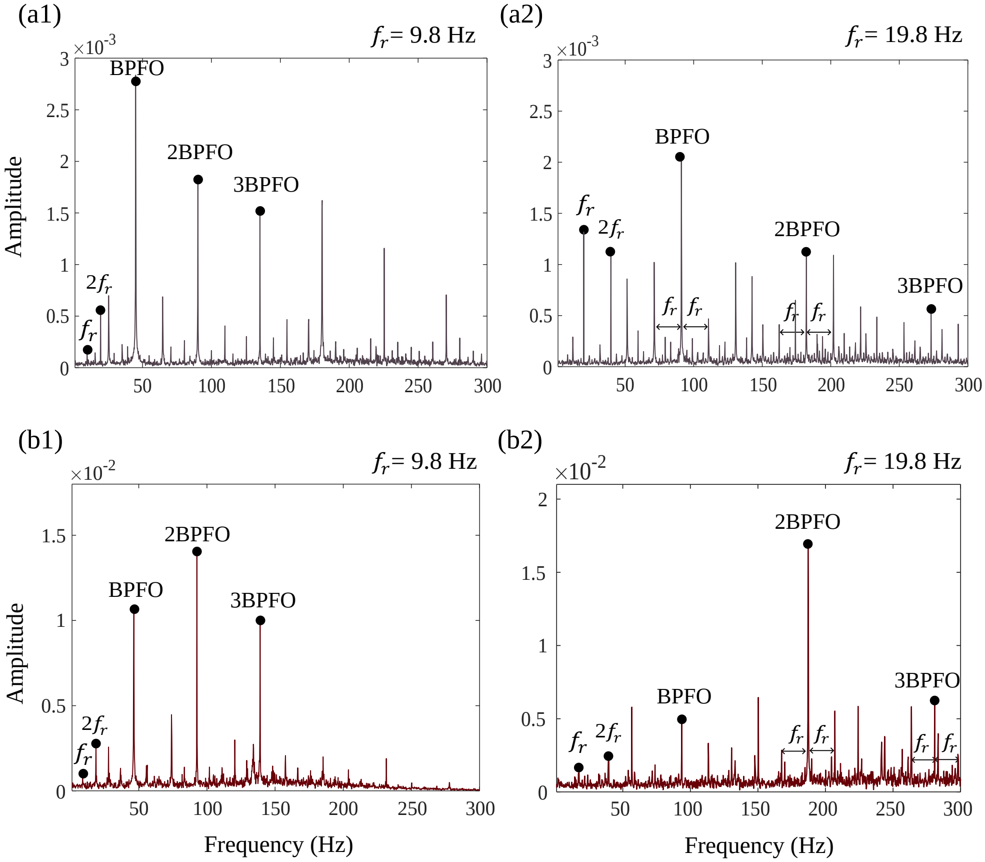


Fig. 16: Comparison between two bearing models with the ORF: (a1) the envelope spectrum of the 6DOF-BM when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the 6DOF-BM when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the hi-fi BM when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the hi-fi BM when $f_r = 19.8$ Hz.

Table 8: Simulation time of three scenarios. The simplified bearing is modelled by a bushing element.

Cases	f_r (Hz)	Duration of the simulated period (s)	Step size (s)	Elapsed time (s)
1 simplified bearing + Hi-fi BM	9.8	10	10^{-6}	6924.68
	19.8	10	10^{-6}	8251.25
1 simplified bearing + 6DOF-BM	9.8	10	10^{-6}	6110.64
	19.8	10	10^{-6}	7255.17
2 simplified bearings	9.8	10	10^{-6}	3106.29
	19.8	10	10^{-6}	5408.75

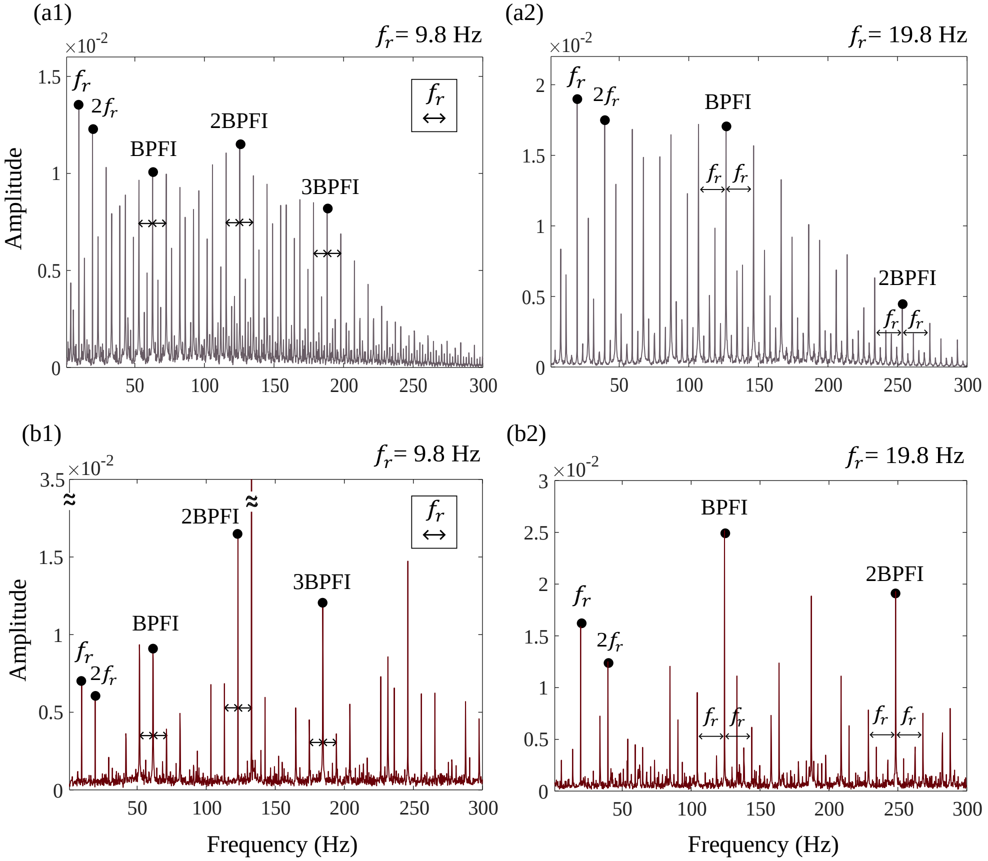


Fig. 17: Comparison between two bearing models with the IRF: (a1) the envelope spectrum of the 6DOF-BM when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the 6DOF-BM when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the hi-fi BM when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the hi-fi BM when $f_r = 19.8$ Hz.

5.4. Discussion

The proposed MBD framework integrating a hi-fi BM effectively reproduces the dominant bearing fault frequencies while maintaining a manageable computational cost. However, one of the primary challenges associated with the MBD model is accurately representing the external loading conditions on bearings and the resulting RE-raceway contact forces. First, the MBD simulations reveal that preload significantly affects the rolling trajectory of the REs. In certain cases, it may reduce the possibility of striking the bearing faults, which may also occur in real systems. Second, to simplify RE–cage interactions and improve computational efficiency, a 5-DOF in-plane joint is used. Consequently, the absence of ball-cage pocket contacts leads to an initial contact angle induced by the preload, leading to an 2% error of BPFI and BPFO in hi-fi BM. Despite this trade-off, the benefits of reduced computational cost remain significant.

Another challenge for the full-scale model is that the acceleration responses are highly sensitive to the assumed stiffness and damping properties of the structural components. This leads to the differences between experimental and simulated responses. Future work may incorporate parameter-optimization strategies to further refine the model behavior and enhance fidelity without compromising computational efficiency.

Implementation of the two different bearing models highlights the modularity of the proposed MBD framework, enabling fault data generation across different operating conditions and modeling objectives. The comparative analysis between the hi-fi BM and the 6DOF-BM demonstrates their advantages and limitations. The 6DOF-BM shows the results that closely align with the experimental setup used in this study and is particularly suitable for scenarios approximating ideal operating conditions. On the other hand, the hi-fi BM offers greater tunable configurations, allowing more detailed representation of bearing mechanics and providing valuable insights for experimental design and system-level studies. This flexibility highlights the fidelity in capturing complex dynamics, especially in cases where full-scale interactions and parameter variability play a significant role.

To summarize, in the future TL application, the hi-fi BM can be used to generate data in various bearing fault scenarios, such as RE and cage defects, while the 6DOF-BM can be utilized for data augmentation to emphasize idealized bearing fault characteristics. Collectively, these two models can complement each other in training a more comprehensive and robust diagnostic model.

6. Conclusions

This study details the development of a full-scale MBD model that incorporates a hi-fi ball bearing representation, designed to replicate a physical test rig for the purpose of generating bearing fault data. Unlike previous studies that focus solely on a single bearing, the proposed MBD model is designed to align closely with the configuration of the physical test rig and provides more simulation scenarios for developing fault diagnosis methods in future works. The MBD model is first validated against the experimental measurements, demonstrating that the MBD model successfully reproduces the fault characteristics of the BPFO, BPFI, and the corresponding harmonics at two rotational speeds. In addition, the investigation for the effects of the applied preloads demonstrates that the potentials of the proposed work in replicating more complex operating conditions, such as different contact angles or different location of the faults.

To further highlight the advantages of the proposed MBD framework, a comparative analysis is conducted between the proposed hi-fi BM and a conventional 6DOF-BM. The analytical bearing model accurately reproduces the exact fault features, while the detailed hi-fi model offers greater fidelity for simulating more complex fault scenarios. Overall, the findings demonstrate that the proposed framework not only captures the comparable fault characteristics with a balanced trade-off between accuracy and computational efficiency, but also provides a versatile foundation for simulation-driven diagnostic methods in future applications.

Future work will focus on training diagnostic models using the fault data generated by the proposed MBD framework. In addition to single-fault detection, the model is also capable of generating multi-fault signals, enabling the study of more complex fault scenarios.

Acknowledgments

This work is supported by the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Sweden, and the Academy of Finland's Centre of Excellence in High-Speed Energy Conversion Systems, Finland. The experiments were conducted in the Laboratory of Machine Dynamics at LUT University, Finland. The development of the MBD models in MSC Adams was supported by Fredrik Sjögren, senior technical specialist at MSC software, Hexagon, Mölndal, Sweden.

References

[1] K. Elbhbah, J. K. Sinha, Vibration-based condition monitoring of rotating machines using a machine composite spectrum, Journal of Sound and Vibration 332 (2013) 2831–2845. `doi:10.1016/J.JSV.2012.12.024`.

[2] M. Vishwakarma, R. Purohit, V. Harshlata, P. Rajput, Vibration analysis & condition monitoring for rotating machines: A review, Materials Today: Proceedings 4 (2017) 2659–2664. `doi:10.1016/J.MATPR.2017.02.140`.

[3] C. Malla, I. Panigrahi, Review of condition monitoring of rolling element bearing using vibration analysis and other techniques, Journal of Vibration Engineering and Technologies 7 (2019) 407–414. `doi:10.1007/S42417-019-00119-Y/TABLES/1`.

[4] R. N. Bell, D. W. McWilliams, P. O'Donnell, C. Singh, S. J. Wells, Report of large motor reliability survey of industrial and commercial installations, part i, IEEE Transactions on Industry Applications IA-21 (1985) 853–864. `doi:10.1109/TIA.1985.349532`.

[5] S. Nandi, H. A. Toliyat, X. Li, Condition monitoring and fault diagnosis of electrical motors - a review, IEEE Transactions on Energy Conversion 20 (2005) 719–729. `doi:10.1109/TEC.2005.847955`.

[6] D. T. Hoang, H. J. Kang, A survey on deep learning based bearing fault diagnosis, Neurocomputing 335 (2019) 327–335. `doi:10.1016/J.NEUCOM.2018.06.078`.

[7] S. Zhang, S. Zhang, B. Wang, T. G. Habetler, Deep learning algorithms for bearing fault diagnosticsx - a comprehensive review, IEEE Access 8 (2020) 29857–29881. `doi:10.1109/ACCESS.2020.2972859`.

[8] R. Liu, B. Yang, E. Zio, X. Chen, Artificial intelligence for fault diagnosis of rotating machinery: A review, Mechanical Systems and Signal Processing 108 (2018) 33–47. `doi:10.1016/J.YMSSP.2018.02.016`.

[9] Z. Chen, A. Mauricio, W. Li, K. Gryllias, A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks, Mechanical Systems and Signal Processing 140 (2020) 106683. `doi:10.1016/J.YMSSP.2020.106683`.

[10] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, A. K. Nandi, Applications of machine learning to machine fault diagnosis: A review and roadmap, Mechanical Systems and Signal Processing 138 (2020) 106587. `doi:https://doi.org/10.1016/j.ymssp.2019.106587`.

[11] T. Ai, Z. Liu, J. Zhang, H. Liu, Y. Jin, M. Zuo, Fully simulated-data-driven transfer-learning method for rolling-bearing-fault diagnosis, IEEE Transactions on Instrumentation and Measurement 72 (2023) 1–11. `doi:10.1109/TIM.2023.3301901`.

[12] W. Hou, C. Zhang, Y. Jiang, K. Cai, Y. Wang, N. Li, A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data, Measurement 215 (2023) 112879. `doi:https://doi.org/10.1016/j.measurement.2023.112879`.

[13] C. Li, S. Zhang, Y. Qin, E. Estupinan, A systematic review of deep transfer learning for machinery fault diagnosis, Neurocomputing 407 (2020) 121–135. `doi:10.1016/J.NEUCOM.2020.04.045`.

[14] Z. Zhao, Q. Zhang, X. Yu, C. Sun, S. Wang, R. Yan, X. Chen, Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study, IEEE Transactions on Instrumentation and Measurement 70 (2021) 1–28. `doi:10.1109/TIM.2021.3116309`.

[15] T. Han, C. Liu, R. Wu, D. Jiang, Deep transfer learning with limited data for machinery fault diagnosis, Applied Soft Computing 103 (2021) 107150. `doi:10.1016/J.ASOC.2021.107150`.

[16] B. Yang, Y. Lei, F. Jia, S. Xing, An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings, Mechanical Systems and Signal Processing 122 (2019) 692–706. `doi:10.1016/J.YMSSP.2018.12.051`.

[17] T. Han, C. Liu, W. Yang, D. Jiang, Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application, ISA Transactions 97 (2020) 269–281. `doi:10.1016/J.ISATRA.2019.08.012`.

[18] X. Liu, S. Liu, J. Xiang, R. Sun, A transfer learning strategy based on numerical simulation driving 1d cycle-gan for bearing fault diagnosis, Information Sciences 642 (2023) 119175. `doi:10.1016/J.INS.2023.119175`.

[19] S. Sassi, B. Badri, M. Thomas, A numerical model to predict damaged bearing vibrations, JVC/Journal of Vibration and Control 13 (2007) 1603–1628. `doi:10.1177/1077546307080040`.

[20] J. Sopanen, A. Mikkola, Dynamic model of a deep-groove ball bearing including localized and distributed defects. part 1: Theory, Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics 217 (2003) 201–211. `doi:10.1243/14644190360713551`.

[21] H. Cao, S. Su, X. Jing, D. Li, Vibration mechanism analysis for cylindrical roller bearings with single/multi defects and compound faults, Mechanical Systems and Signal Processing 144 (2020) 106903. `doi:10.1016/J.YMSSP.2020.106903`.

[22] C. Sobie, C. Freitas, M. Nicolai, Simulation-driven machine learning: Bearing fault classification, Mechanical Systems and Signal Processing 99 (2018) 403–419. `doi:10.1016/j.ymssp.2017.06.025`.

[23] J. Liu, H. Cao, S. Su, X. Chen, Simulation-driven subdomain adaptation network for bearing fault diagnosis with missing samples, Engineering Applications of Artificial Intelligence 123 (2023) 106201. `doi:10.1016/J.ENGAPPAI.2023.106201`.

[24] P. Aduwenye, C. Nutakor, L. Roininen, J. Sopanen, Parameter-based transfer learning for bearing fault diagnosis using small samples under variable working conditions, IEEE Access 13 (2025) 44214–44230. `doi:10.1109/ACCESS.2025.3548549`.

[25] H. Cao, L. Niu, S. Xi, X. Chen, Mechanical model development of rolling bearing-rotor systems: A review, Mechanical Systems and Signal Processing 102 (2018) 37–58. `doi:10.1016/J.YMSSP.2017.09.023`.

[26] C. Wen, X. Meng, C. Fang, J. Gu, L. Xiao, S. Jiang, Dynamic behaviors of angular contact ball bearing with a localized surface defect considering the influence of cage and oil lubrication, Mechanism and Machine Theory 162 (2021) 104352. `doi:10.1016/J.MECHMACHTHEORY.2021.104352`.

[27] A. A. Shabana, Dynamics of multibody systems, Cambridge university press, 2020.

[28] M. Vehviläinen, P. Rahkola, J. Keränen, J. Halme, J. Sopanen, O. Liukkonen, A. Holopainen, K. Tammi, A. Belahcen, Adapting geometry-based polygonal contacts for simulating faulty rolling bearing dynamics, Mechanism and Machine Theory 192 (2024) 105552. `doi:10.1016/J.MECHMACHTHEORY.2023.105552`.

[29] L. X. Xu, Y. H. Yang, Y. G. Li, C. N. Li, S. Y. Wang, Modeling and analysis of planar multibody systems containing deep groove ball bearing with clearance, Mechanism and Machine Theory 56 (2012) 69–88. `doi:10.1016/J.MECHMACHTHEORY.2012.05.009`.

[30] C. Bovet, L. Zamponi, An approach for predicting the internal behaviour of ball bearings under high moment load, Mechanism and Machine Theory 101 (2016) 1–22. `doi:10.1016/J.MECHMACHTHEORY.2016.03.002`.

[31] L. E. Stacke, D. Fritzson, P. Nordling, Beast—a rolling bearing simulation tool, Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics 213 (1999) 63–71. `doi:10.1243/1464419991544063`.

[32] Z. Qi, G. Wang, Z. Zhang, Contact analysis of deep groove ball bearings in multibody systems, Multibody System Dynamics 33 (2015) 115–141. `doi:10.1007/S11044-014-9412-0/FIGURES/24`.

[33] L. Giraudo, L. G. D. Maggio, L. Giorio, C. Delprete, Dynamic multibody modeling of spherical roller bearings with localized defects for large-scale rotating machinery, Sensors 2025, Vol. 25, Page 2419 25 (2025) 2419. `doi:10.3390/S25082419`.

[34] S. A. McInerny, Y. Dai, Basic vibration signal processing for bearing fault detection, IEEE Transactions on Education 46 (2003) 149–156. `doi:10.1109/TE.2002.808234`.

[35] Y. Cheng, S. Wang, B. Chen, G. Mei, W. Zhang, H. Peng, G. Tian, An improved envelope spectrum via candidate fault frequency optimization-gram for bearing fault diagnosis, Journal of Sound and Vibration 523 (2022) 116746. `doi:10.1016/J.JSV.2022.116746`.

[36] B. Chen, W. Zhang, J. X. Gu, D. Song, Y. Cheng, Z. Zhou, F. Gu, A. D. Ball, Product envelope spectrum optimization-gram: An enhanced envelope analysis for rolling bearing fault diagnosis, Mechanical Systems and Signal Processing 193 (2023) 110270. `doi:10.1016/J.YMSSP.2023.110270`.

[37] D. Yu, J. Cheng, Y. Yang, Application of emd method and hilbert spectrum to the fault diagnosis of roller bearings, Mechanical Systems and Signal Processing 19 (2005) 259–270. `doi:10.1016/S0888-3270(03)00099-2`.

[38] H. M. Hilber, T. J. Hughes, R. L. Taylor, Improved numerical dissipation for time integration algorithms in structural dynamics, Earthquake Engineering & Structural Dynamics 5 (1977) 283–292. `doi:10.1002/EQE.4290050306;JOURNAL:JOURNAL:10969845;WGROUP:STRING:PUBLICATION`.

[39] R. Gismeros Moreno, F. Marques, E. Corral Abad, J. Meneses Alonso, P. Flores, C. Castejon, Enhanced modelling of planar radial-loaded deep groove ball bearings with smooth-contact formulation, Multibody System Dynamics 60 (2024) 121–159. `doi:10.1007/S11044-023-09952-2/METRICS`.

[40] T. Harris, M. Kotzalas, Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition, Rolling Bearing Analysis, Fifth Edtion, CRC Press, 2006.

[41] E. Pennestrì, V. Rossi, P. Salvini, P. P. Valentini, Review and comparison of dry friction force models, Nonlinear Dynamics 83 (2016) 1785–1801. `doi:10.1007/S11071-015-2485-3/TABLES/8`.

[42] J. Sopanen, A. Mikkola, Dynamic model of a deep-groove ball bearing including localized and distributed defects. part 2: Implementation and results, Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics 217 (2003) 213–223. `doi:10.1243/14644190360713560;WGROUP:STRING:PUBLICATION`.

[43] A. Sapietová, L. Gajdoš, V. Dekýš, M. Sapieta, Analysis of the influence of input function contact parameters of the impact force process in the msc. adams, Advances in Intelligent Systems and Computing 393 (2016) 243–253. `doi:10.1007/978-3-319-23923-1_37`.

Paper B

Simulation-driven diagnostic method via a multibody dynamics model for rotating machinery

Y. H. Pai, P. T. Piiroinen, H. Johansson, S. Kumar

Submitted to the 12th IFToMM International Conference on Rotordynamics 2026

Simulation-driven diagnostic method via a multibody dynamics model for rotating machinery

Yu-Hung Pai$^{1[0000-0001-7928-1711]}$, Petri T. Piiroinen$^{1[0000-0002-3578-1877]}$, Shivesh Kumar$^{1,2[0000-0002-6254-3882]}$ and Håkan Johansson$^{1[0000-0002-7207-8486]}$

[1] Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden,
[2] Robotics Innovation Center, German Research Center for Artificial Intelligence, Bremen, Germany
`paiy@chalmers.se`*

Abstract. Fault diagnosis is a pivotal aspect of condition monitoring in rotating machinery. Although fault diagnosis methods that use deep-learning approaches have achieved significant success, they require massive, labelled datasets from operating machines, which is often impractical and expensive for industrial applications. To overcome this challenge, simulation-driven fault diagnosis provides a scalable alternative by generating simulated data for training diagnostic models. This study presents a simulation-driven diagnostic framework for a physical test rig, where fault characteristics are learned via a multibody dynamics (MBD) model capable of incorporating more complex fault conditions in rotating machinery. The MBD model is first validated against experimental measurements from the physical test rig, demonstrating its capability to accurately replicate realistic fault signatures. The simulated dataset generated by the MBD model is then used to train a diagnostic model, which is further adapted to the physical test rig through transfer-learning methods. By capturing the realistic operating defect responses, this study demonstrates the potential of simulation-driven fault diagnosis and provides a promising foundation for future industrial integration.

Keywords: Transfer learning, multibody dynamics simulation, bearing modelling, fault diagnosis.

1 Introduction

Condition monitoring (CM) for rotordynamic systems has become more feasible and increasingly desirable for industries aiming to optimize maintenance efficiency. For rotordynamic systems, fault diagnosis of rolling bearing plays a pivotal role, given that such faults are responsible for 30% to 50% of all machine failures [8]. However, a major challenge in the development of machine learning models for rotating machines is the insufficiency of fault data [2]. Since most rotating machines operate under normal conditions for most of their lifespan, it is challenging to obtain sufficient fault data from operating systems. To overcome this, simulation-driven methods provide an efficient way to construct diagnostic models.

Simulation-driven methods, combined with transfer learning (TL) techniques, enable diagnostics models to leverage knowledge acquired from one or multiple source domains and apply them to relevant target domains [13]. To generate bearing fault signals, several analytical bearing models have been developed [1, 9, 11], and successfully implemented into diagnostic model training [5, 10]. However, relying only on single-bearing model limits the diversity of training dataset, thereby constraining the development of diagnostic models. To construct more comprehensive diagnostic models, multibody dynamics (MBD) simulations offer higher physical fidelity and ability to generate a broader range of fault scenarios. Recent studies have demonstrated the effectiveness of simulation-driven approaches via MBD models in fault diagnosis, highlighting their potential for broader applications [3, 4].

This work aims to develop a simulation-driven diagnostic model that leverages MBD simulation to address the scarcity of fault data in CM of rotordynamic systems. First, experimental vibration data are obtained under both healthy and faulty bearing conditions at two rotational speeds using a dedicated test rig. Second, a full-scale MBD model, incorporating an analytical bearing representation, is developed to generate simulated data with identical localized faults, while capturing realistic operating conditions. Finally, a TL framework is established to train a diagnostic model using the simulated data and subsequently adapt it to the experimental dataset. By integrating simulated and experimental data, the TL approach bridges the gap between virtual and physical systems. The proposed framework enables the construction of a pre-trained diagnostic model using simulated data and requires only a small subset of real data for fine-tuning. This framework lays a solid foundation for future development of simulation-driven fault diagnosis.

2 Experimental arrangement

To construct the target domain for TL approach, a dedicated test rig was designed to acquire bearing vibration data, as shown in Fig. 1(a). The rig was composed of a 600 mm long and 12 mm diameter shaft, two steel discs with a diameter of 150 mm, a thickness of 10 mm, and a mass of 1.69 kg, one AC electric motor, two bearings housings, and two deep-groove ball bearings. The bearings used were standard SKF 6007-2Z deep-groove ball bearing, featuring 11 REs of 7.94 mm diameter, and a pitch diameter of 48 mm. The bearings were mounted at the drive end (DE) and the non-drive end (NDE) of the shaft, with the inner rings (IRs) rotating together with the shaft while the outer rings (ORs) were fixed to the housings.

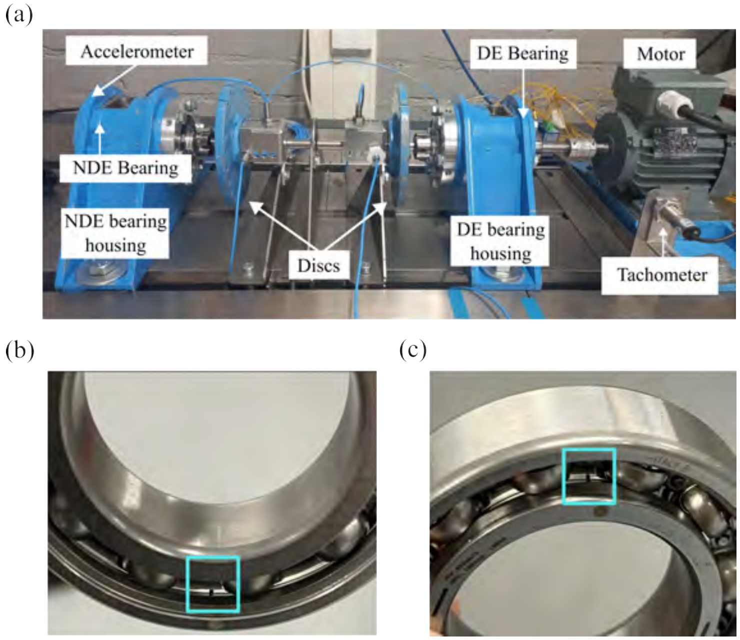


Fig. 1. The experimental setup of the test rig: (a) the arrangement of the sensors and the components (b) the tested bearing with ORF (c) the tested bearing with IRF.

Two ball bearings with a single localized fault, one with an outer race fault (ORF) and the other with inner race fault (IRF), were fabricated by laser machining as shown in Fig. 1(b) and (c). Each defect measured approximately 2 mm in length, 0.5 mm in width and 0.2 mm in depth. Vibration responses were recorded using a tri-axial accelerometer (PCB 356A03) mounted on the NDE bearing housing, where the faulty bearings were installed. The accelerometer signals were acquired through a data acquisition system (DeweSoft SIRIUS®) at a sampling rate of 10 kHz.

Six scenarios of vibration acceleration measurements were conducted with healthy and two faulty bearings at two rotational frequencies ($f_r = 9.8$ Hz and $f_r = 19.8$ Hz). The experimentally determined rotational speeds were subsequently adopted in the simulations. All acceleration signals were recorded after the shaft speed stabilized at the target rotational speeds.

3 Multibody dynamics model

To replicate the experimental setup, a full-scale MBD model is built using the MBD analysis software MSC Adams [7], with all dimensions consistent with those of the physical system. The complete assembly includes a shaft, a motor, two bearing housings, one simplified DE bearing, one analytical NDE bearing, two disks and a supporting platform. In this model, the shaft is modelled as a flexible body, while other components are considered to be rigid bodies. The Hilber-Hughes-Taylor implicit Method (HHT-α) is used in the numerical solver with an error tolerance of 10^{-7}, a maximum time step of 10^{-6}, a maximum of 50 iterations for convergence, and a sampling rate of 10 kHz.

The model topology is presented in Fig. 2. Because the diagnostic focus is on the NDE bearing, the DE bearing is simplified as a bushing element composed of three rotational and three translational spring-damper elements to reduce the computational load. The rotational stiffness and damping values are set to zero in all three directions. In the longitudinal direction (X axis), the translational stiffness and damping values are set to zero. In the radial direction (Y and Z axis), the translational stiffness and damping values are set to $2 \cdot 10^4$ N/mm and $2 \cdot 10$ Ns/mm, respectively, to approximate the maximum radial clearance in the actual bearing [12]. In contrast, the NDE bearing is modelled using an six-degree-of-freedom analytical bearing, where the IRF and ORF are considered as the additional displacement excitation based on the defect length and depth [11]. The OR is fixed to the NDE bearing housing, while the IR is fixed on sleeve 2 and rotates about the same shaft axis at the given rotational frequency.

To obtain the vibration responses from the bearing housings, the connections between the housings and the supporting platform are modelled as zero-damping bushing elements. As the actual stiffness and damping properties of these connections are not directly measurable, the bushing parameters are systematically calibrated to match the two dominant modal frequencies obtained from experimental modal analysis of the test rig. Therefore, The translational stiffness values in the X, Y and Z directions are set to $1 \cdot 10^4$ N/mm, $6 \cdot 10^3$ N/mm and $9 \cdot 10^3$ N/mm, respectively, and the corresponding rotational stiffness values are assigned as $5 \cdot 10^5$ N-mm/deg, $1 \cdot 10^5$ N-mm/deg and $5 \cdot 10^4$ N-mm/deg about the X, Y and Z axes, respectively. This calibration ensures that the simulated structural boundary conditions approximate the dynamic behavior of the physical system.

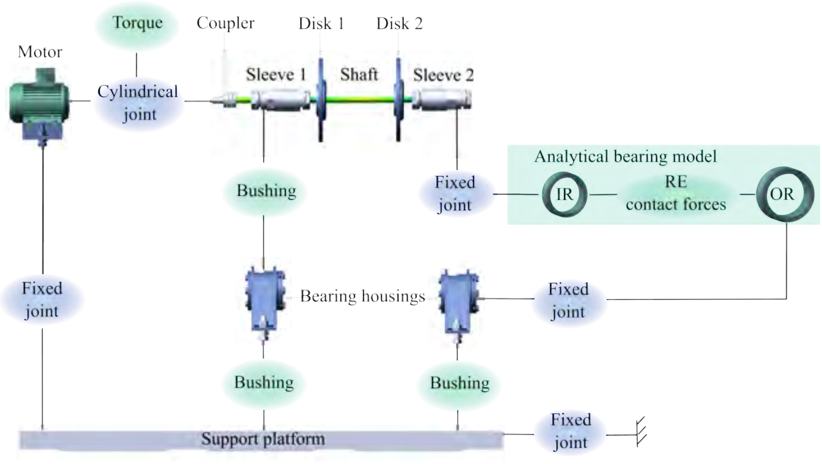


Fig. 2. Topology of the proposed MBD model.

4 Transfer learning framework

The TL framework proposed in this study is depicted in Fig. 3. The source domain is composed of bearing vibration data generated by the proposed MBD model, whereas the target domain contains experimental bearing vibration data measured from the physical test rig. Initially, feature extraction techniques are applied to the source domain to capture fault-related features while preserving their physical interpretations. The extracted features are then used to construct a pre-trained diagnostic model. The diagnostic model consists of using three fully-connected layers with ReLU activation function, followed by an output layer with Softmax classification. Subsequently, the parameters of the feature extraction stage are frozen, and only the parameters of the diagnostic model are fine-tuned using the target domain dataset. This fine-tuning process allows the model to effectively adapt to real data and improves its classification performance.

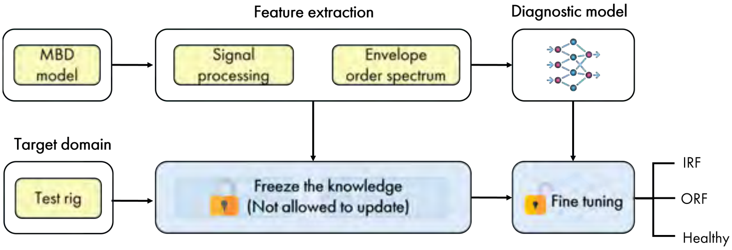


Fig. 3. The proposed TL framework.

In feature extraction procedure, a low-pass filter with a 4000 Hz cut-off frequency is first applied to suppress aliasing and high-frequency noise. Since bearing fault signals are usually modulated by low-frequency waveforms, signal demodulation needs to be conducted prior to the Fast Fourier Transform (FFT). The Hilbert transform is therefore applied to extract envelope signals, allowing clear identification of fault frequencies and their harmonics. Finally, to ensure consistency across different rotational speeds, the envelope spectra are converted into envelope order spectra using order tracking, where features are represented by multiples of rotational order rather than absolute frequencies.

5 Results

First, the simulated results for both faulty bearing conditions at two different rotational frequencies are validated against the experimental measurements. Only the vertical vibration responses are considered in this work. The corresponding envelope spectra are presented in Figs. 4 and 5. In Figs. 4(b1) and (b2), the MBD model successfully reproduces the BPFIs and their modulation sidebands at both rotational speeds, closely matching the fault characteristics observed in Figs. 4(a1) and (a2). Similarly, the simulated BPFOs in Figs. 5(b1) and (b2) exhibits strong

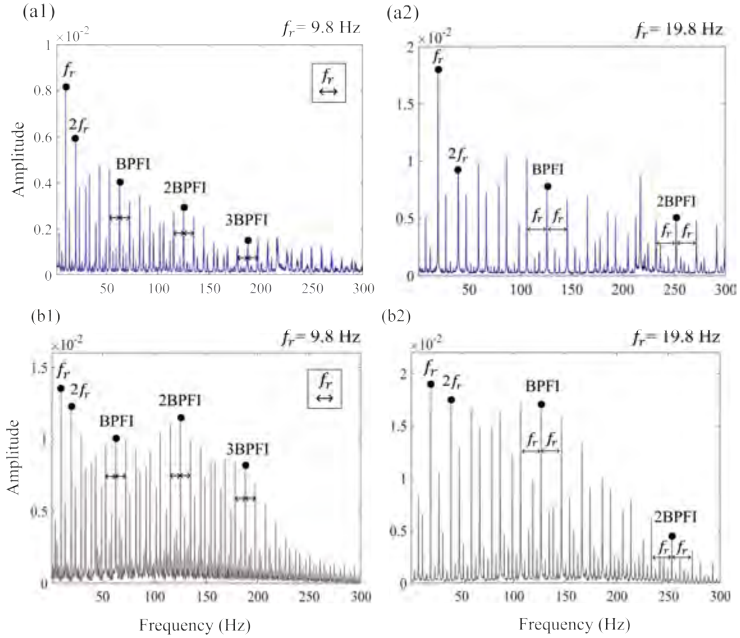


Fig. 4. Comparison between the experimental and the simulated data with the IRF: (a1) the envelope spectrum of the experimental data when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the experimental data when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the simulated data when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the simulated data when $f_r = 19.8$ Hz.

agreement with the experimental results in Figs. 5(a1) and (a2). A summary of the fault-frequency comparisons is provided in 1.

Following model validation, simulation datasets are generated according to the experimental scenarios. The arrangement of the datasets is shown in Fig. 2. Scenario 1-4 are simulated for 67 s with data sampled every 1 s, while Scenario 5 and 6 are simulated for 55 s at the same sampling interval, leading to a total of 378 simulated samples. In contrast, each experimental scenario is measured for 30 s, with data sampled every 1 s, resulting in 180 experimental samples in total.

After converting all samples into envelope order spectra, the first 50 orders are selected. To verify that feature extraction process preserves the critical fault-related information, t-distributed Stochastic Neighbor Embedding (t-SNE) is applied to visualize the high-dimensional data. By projecting the data onto two t-SNE components, as shown in Fig. 6, the three different conditions are well separated in both the source and target domains, confirming that the envelope order spectra effectively capture the bearing fault signatures at different rotational speed.

After constructing the pre-trained diagnostic model using 90% of the simulated data for training and 10% for validation, the model is first evaluated on the experimental dataset. The corresponding diagnosis result is shown in Fig. 7. The pre-

Table 1. Comparison of the fault frequencies.

f_r	Cases	BPFO (Hz)	BPFI (Hz)
9.8	Theoretical values [6]	45.08	62.72
	MBD model	45.00	62.80
	Experiments	45.30	62.40
19.8	Theoretical values [6]	91.08	126.72
	MBD model	91.00	126.80
	Experiments	90.80	126.20

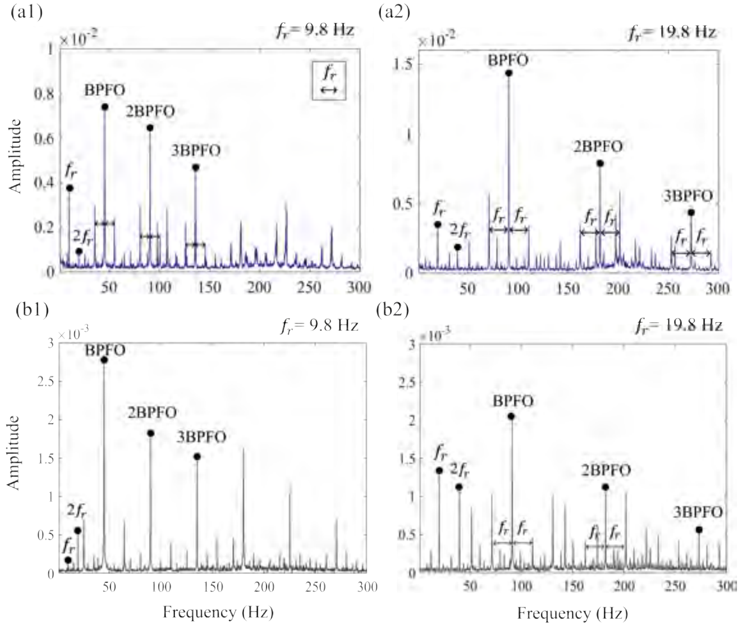


Fig. 5. Comparison between the experimental and the simulated data with the ORF: (a1) the envelope spectrum of the experimental data when $f_r = 9.8$ Hz, (a2) the envelope spectrum of the experimental data when $f_r = 19.8$ Hz, (b1) the envelope spectrum of the simulated data when $f_r = 9.8$ Hz, (b2) the envelope spectrum of the simulated data when $f_r = 19.8$ Hz.

trained model is able to distinguish between the BPFI and BPFO samples; however, it performs poorly on healthy data, achieving an overall accuracy of only 65%, as shown in Fig. 7(a). This is because of the discrepancy between the simulated and experimental healthy data, highlighting the necessity of applying a TL method to adapt the pre-trained model to real data.

As shown in Fig. 7(b) and (c), after fine tuning the diagnostic model using 10% of the experimental data, the classification accuracy increases to 94.4%. Increasing the fine tuning dataset to 20% further improves the accuracy to 97.2%. To ensure that the model does not overfit the experimental data, the updated diagnostic model

Table 2. Dataset arrangement for TL framework.

Scenarios	Faulty condition	Simulated samples	Experimental samples
1	IRF	67	30
2	IRF	67	30
3	ORF	67	30
4	ORF	67	30
5	Healthy	55	30
6	Healhty	55	30

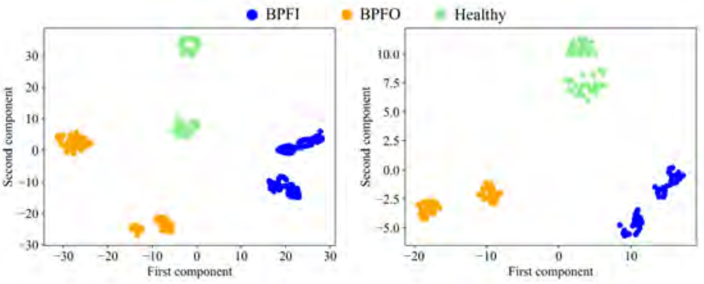


Fig. 6. Group scatter after feature extraction: the source domain data (left) and (right) the target domain data.

is re-evaluated on the simulated dataset, where it achieves 100% accuracy, as shown in 7(d).

6 Conclusions

This work proposes a simulation-driven TL framework that integrates an MBD-based bearing fault diagnosis method across different rotational speeds. The MBD model successfully reproduces the fault characteristics, showing strong agreement with the experimental measurements. The bearing fault frequencies extracted from both simulated and experimental data closely match the theoretical predictions, confirming the physical validity of the generated data. After applying the proposed TL framework, the diagnostic accuracy significantly improves from 65% to 97.2%, demonstrating its effectiveness in bridging the gap between simulated and real data. Overall, this study provides a promising direction for future development of simulation-driven bearing fault diagnosis methods.

Acknowledgments

This work is supported by the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden. The experiments were con-

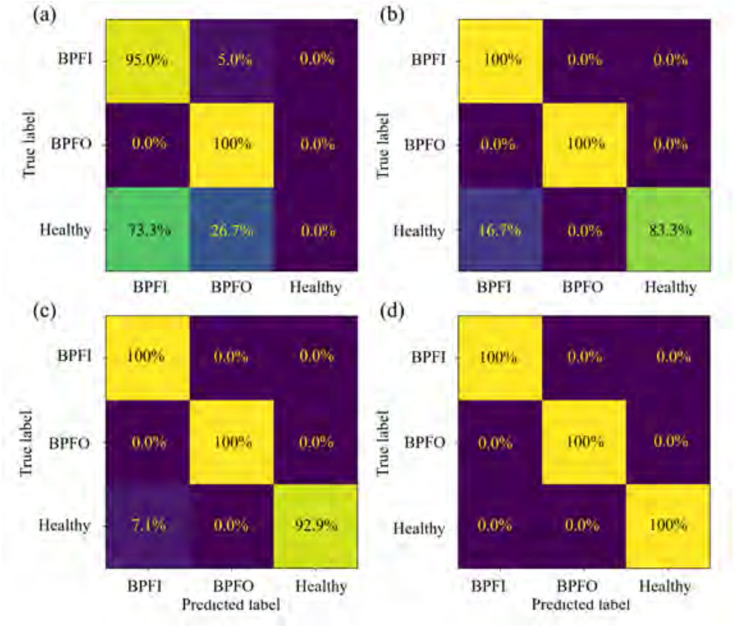


Fig. 7. Confusion matrix of fault diagnosis: (a) before fine-tuning (b) after fine-tuning by using 10% experimental data (c) after fine-tuning by using 20% experimental data (d) update model re-evaluation on the simulation data.

ducted in the Laboratory of Machine Dynamics at LUT University, Lappeenranta, Finland.

References

1. Cao H, Su S, Jing X, Li D (2020) Vibration mechanism analysis for cylindrical roller bearings with single/multi defects and compound faults. Mechanical Systems and Signal Processing 144:106,903
2. Hou W, Zhang C, Jiang Y, Cai K, Wang Y, Li N (2023) A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data. Measurement: Journal of the International Measurement Confederation 215
3. Koutsoupakis J, Giagopoulos D, Chatziparasidis I (2023) Ai-based condition monitoring on mechanical systems using multibody dynamics models. Engineering Applications of Artificial Intelligence 123:106,467
4. Koutsoupakis J, Seventekidis P, Giagopoulos D (2023) Machine learning based condition monitoring for gear transmission systems using data generated by optimal multibody dynamics models. Mechanical Systems and Signal Processing 190:110,130
5. Liu J, Cao H, Su S, Chen X (2023) Simulation-driven subdomain adaptation network for bearing fault diagnosis with missing samples. Engineering Applications of Artificial Intelligence 123:106,201
6. McInerny S, Dai Y (2003) Basic vibration signal processing for bearing fault detection. IEEE Transactions on Education 46(1):149–156
7. MSC Adams, Hexagon. https://hexagon.com/products/product-groups/computer-aided-engineering-software/adams, last accessed 14 Nov 2025

8. Nandi S, Toliyat HA, Li X (2005) Condition monitoring and fault diagnosis of electrical motors - a review. IEEE Transactions on Energy Conversion 20:719–729
9. Sassi S, Badri B, Thomas M (2007) A numerical model to predict damaged bearing vibrations. JVC/Journal of Vibration and Control 13:1603–1628
10. Sobie C, Freitas C, Nicolai M (2018) Simulation-driven machine learning: Bearing fault classification. Mechanical Systems and Signal Processing 99:403–419
11. Sopanen J, Mikkola A (2003) Dynamic model of a deep-groove ball bearing including localized and distributed defects. part 1: Theory. Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics 217:201–211
12. Vehviläinen M, Rahkola P, Keränen J, Halme J, Sopanen J, Liukkonen O, Holopainen A, Tammi K, Belahcen A (2024) Adapting geometry-based polygonal contacts for simulating faulty rolling bearing dynamics. Mechanism and Machine Theory 192:105,552
13. Yang B, Lei Y, Jia F, Xing S (2019) An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. Mechanical Systems and Signal Processing 122:692–706

Paper C

Machine-learning based fault diagnosis for a rotordynamic system using multibody simulations

Y. H. Pai, P. T. Piiroinen, S. Kumar, H. Johansson

Proceedings of 15th International Workshop on Structural Health Monitoring (IWSHM), 2025

Title: **Machine-learning based fault diagnosis for a rotordynamic system using multibody simulations**

Authors : Yu-Hung Pai
Petri T. Piiroinen
Shivesh Kumar
Håkan Johansson

ABSTRACT

Machine-learning based fault diagnosis plays an important role in condition monitoring for rotating machinery to prevent systems from catastrophic faults. It is important to note that the performance of data-driven methods relies highly on a large quantity of training fault data. Since rotating machinery operates under normal condition most of the time, collecting sufficient fault data from experiments takes a huge amount of time and expense, and under various operating conditions. To overcome the fault data insufficiency, building a virtual testbed for generating fault data is a promising way in bridging the gap between data requirement and prediction accuracy.

Many simplified dynamic models have been proposed to generate a single fault on some rotordynamic systems. These methods, however, cannot reflect complex operation conditions such as variant rotation speed or multi-faults. To better reveal vibration responses of local defects, this research aims to establish a multibody dynamics (MBD) model that can simultaneously analyze complete dynamic behavior and simulate a wider range of fault scenarios.

In this research, a simulation-driven fault diagnosis method is proposed to generate the simulation fault data. Firstly, a rigid-flexible hybrid model of a single-rotor-bearing system is established using MSC ADAMS, which is based on MBD and finite element analysis. Different fault conditions are simulated including outer race bearing faults, inner race bearing faults, and rolling element faults. After generating fault data, a time-frequency feature extraction method is developed based on Hilbert envelope and wavelet packet decomposition, extracting a large amount of features from the original signals. In addition, an autoencoder model is built to highlight the critical features, enhancing the performance of the classifier. This feature extraction is made to obtain fault-related features, which train the machine learning classifiers for discriminating the fault categories.

Yu-Hung Pai, Ph.D. student, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden. Email: paiy@chalmers.se. Petri T. Piiroinen, Associate Professor, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden. Håkan Johansson, Professor, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden. Shivesh Kumar, Assistant Professor, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Göteborg, Sweden. Robotics Innovation Center, German Research Center for Artificial Intelligence, Bremen, Germany.

To validate the simulation results, the Case Western Reserve University (CWRU) bearing dataset that has been widely accepted as a standard reference is introduced. A comparison of bearing fault frequencies between simulations and the CWRU dataset is then conducted. Meanwhile, a transfer learning method is applied using the CWRU dataset to fine tune the fault diagnosis classifier. This research lays a solid foundation for future development of a digital twin and simulation-driven transfer learning for fault diagnosis of rotating machines.

INTRODUCTION

Rotordynamic system, a key component in most industrial sectors, is prone to various defects during operations such as bearing faults, misalignment and unbalance. When a local fault grows to a critical level, it often leads to a long downtime, and in severe cases causing damage to the entire machine. Therefore, accurately diagnosing faults at an early stage is necessary for rotating machines.

Machine-learning-based fault diagnosis integrating traditional signal processing and machine learning methods serves as a predictive maintenance technique to identify the anomalies from monitoring data. However, a major obstacle in developing machine-learning models for rotating machines is the insufficiency of fault data [1]. Since most rotating machines operate under normal conditions for most of their lifespan, it is challenging to obtain sufficient fault data from physical systems. To solve the problem with missing fault samples, building a virtual counterpart of a rotating machinery plays an important role to generate various simulated fault data.

To generate fault signals, many simplified bearing models have been proposed [2] [3]. However, these models often fail to reflect the complete dynamic behavior of the rotating system and are limited in replicating certain fault types-particularly faults on rolling elements, which have been rarely discussed in the previous research. To overcome these limitations, multi-body dynamics (MBD) simulations are used in this work to reveal the realistic operating conditions and to model a wider range of fault scenarios in rotating machines. Recent work by [4] has demonstrated the effectiveness of optimal MBD simulations in condition monitoring, highlighting their potential for broader applications.

This work aims to develop a simulation-data-driven method to tackle the insufficiency of fault data in condition monitoring of rotordynamic system. First, a tunable simulation for the rotor-bearing system is developed to generate different bearing fault data. Second, several feature extraction techniques is presented to capture the fault-related features, which are used to build the fault diagnosis classifier. Third, the experimental datasets is used to validate the MBD model and fine tune the classifier. By combining simulation data and experimental data, the transfer-learning method bridges the gap between a virtual system and a real machine.

The proposed framework allows simulation data to build a fault diagnosis pre-trained classifier and only require a small portion of real data to fine tune the classifier, reinforcing the performance and increasing robustness of condition monitoring systems.

FRAMEWORK OF TRANSFER LEARNING METHOD

The typical transfer learning approach involves initially constructing a pre-trained model using data from a source domain, followed by refining this model with data from a target domain. The framework in this study is depicted in the Figure 1. Here, the source domain comprises simulated bearing fault data from the multibody dynamics (MBD) model, whereas the target domain contains experimental bearing data from the Case Western Reserve University (CWRU) dataset [5]. Initially, feature extraction techniques are applied to the source domain to effectively capture fault-related features while preserving their physical interpretations. After implementing signal processing methods and an autoencoder neural network to highlight these fault-related features, a pre-trained classifier is established based on the extracted source domain features. Subsequently, the parameters of the feature extraction process are frozen, and the classifier parameters are fine-tuned using the target domain dataset, allowing the model to adapt effectively to real data and improving its classification accuracy.

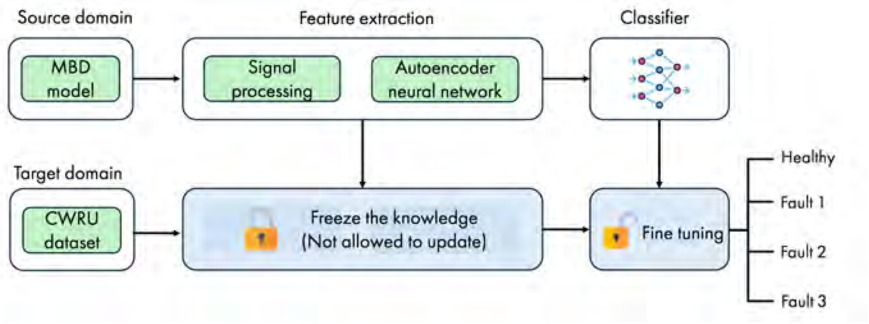


Figure 1. Framework of the proposed fault diagnosis method.

MULTIBODY DYNAMICS SIMULATION

In this work, the single-rotor-bearing model is built using the MBD analysis software, MSC ADAMS [6]. For a rigid-flexible hybrid MBD system, the equations of motion is described in the following general form [4]:

$$\begin{cases} M\ddot{q} + \Phi_q^T\lambda + F_q = Q(q) \\ \Phi(q,t) = 0 \end{cases} \tag{1}$$

where M is the system mass matrix, Φ_q is the derivative matrix of constraint equations with respect to the system generalized coordinates q, λ is the vector of Lagrangian multipliers associated with the constraints, $F(q)$ is the system elastic force vector, $Q(q)$ is the system external generalized forces, $\Phi(q,t)$ is the vector containing the system constraint equations and t is the time. The dynamic behaviors of each component can be calculated at each moment under the forces and torques applied.

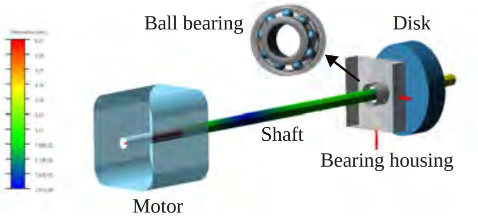


Figure 2. The single-rotor-bearing MBD model.

Figure 3. (left) Fault on the outer ring; (middle) fault on the inner ring and (right) fault on the ball.

This model comprises a motor, a shaft, a ball bearing, a bearing housing and a disk as shown in Figure 2. The shaft is modelled as a flexible body, while other components are simulated as rigid bodies. Initially, allowing the simulation to be validated by the CWRU dataset, the SKF 6205-2RS deep groove ball bearing is chosen. The bearing faults are shown in Figure 3. In this bearing model, the 3D CAD geometry of the bearing is obtained from the SKF official website [7]. Since the roller elements from the CAD file are unified as one rigid body, they are replaced with the sphere bodies in MSC ADAMS that can be tuned separately. The bearing model contains 18 contacts, 9 of which are defined between the roller elements and the inner race and 9 of which are defined between the roller elements and the outer race. The parameters of the contacts are demonstrated in [8]. All the intervals of adjacent rolling elements are confined, and the rolling elements can only rotate along with the longitudinal axis during rotation. Subsequently, the shaft is connected to the motor, the inner race of the bearing and the disk, allowing the rotation around the longitudinal axis and the radial translation confined by the bearing. The outer race of the bearing is fixed on the bearing housing, and the constrains between the housing and the ground are 3 rotational and 3 translational springs. Vertical acceleration responses are measured from the bearing housing to enable comparison with the CWRU dataset.

METHODS OF FAULT DIAGNOSIS

Based on the geometry parameters and rotating speed of the bearing, the bearing fault frequencies associated with the inner race, outer race and rolling element are ball passing frequency of inner race (BPFI), ball passing frequency of outer race (BPFO), and ball spin frequency (BSF). The theoretical fault frequencies are defined as follows:

$$\mathrm{BPFI} = \left[\frac{nf_r}{2}(1 + \frac{d}{D}\cos(\alpha))\right] \tag{2}$$

$$\text{BPFO} = \left[\frac{nf_r}{2}(1 - \frac{d}{D}\cos(\alpha))\right] \tag{3}$$

$$\text{BSF} = \left[\frac{Df_r}{d}(1 - (\frac{d}{D}\cos(\alpha))^2)\right] \tag{4}$$

where n is the number of the rolling elements, f_r is the rotating speed of the inner ring, d is the diameter of the rolling elements, D is the pitch diameter of the bearing and α is the contact angle.

The procedure for feature extraction is illustrated in Figure 4, which aims to identify and isolate fault-related characteristics within various frequency bands. Both bearing-fault and healthy data are generated from MBD model. Ensemble Empirical Mode Decomposition (EEMD) is applied for signal denoising, with the first intrinsic mode function (IMF) being selected for further analysis [9]. Subsequently, the Hilbert envelope method is used to demodulate the signals and to extract the fault frequencies and their harmonics. Even though the simulated acceleration signals do not perfectly replicate the CWRU dataset, their envelope spectrums consistently exhibit similar patterns of fault frequencies [1]. Next, wavelet packet decomposition (WPD) decomposes the envelope signals into different frequency bands, allowing computation of the energy distribution and root means square values across all bands [10]. These extracted features indicate that notably higher energy responses in frequency bands associate with fault frequencies. Finally, the extracted features are processed through an autoencoder neural network to identify critical features, which are then used to train a classifier capable of effectively distinguishing among various fault conditions.

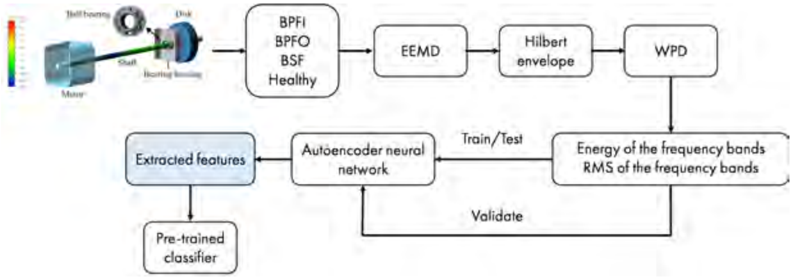


Figure 4. Procedure of feature extraction.

DISCUSSION AND COMPARISON

The rotating speed and sampling rate in the MBD model are set to 1797 rpm and 12000 Hz respectively, which are identical to the CWRU setup. Each case is simulated for 19 s, with data sampled every 0.5 s, leading to 152 $(38 \cdot 4)$ samples in total. In contrast, each case in the CWRU dataset spans 10 s and is sampled at the same 0.5 s, resulting in 80 $(20 \cdot 4)$ samples in total. Figure 5 presents a comparison of the envelope spectrums between the simulated data and the CWRU data for different bearing faults. Although the numerical model introduces additional noise, the MBD model still successfully captures the characteristic fault frequencies. TABLE I demonstrates that

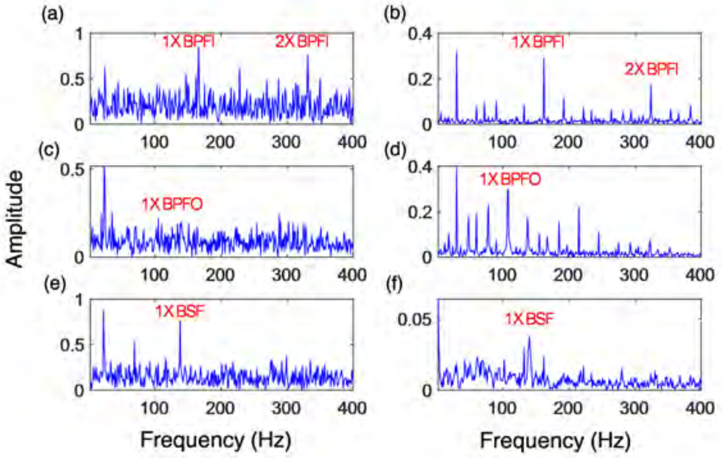


Figure 5. (left (a)(c)(e)) Envelope spectrums from simulated dataset and (right (b)(d)(f)) envelope spectrums from CWRU dataset.

TABLE I. Comparison of the fault frequencies

	MBD model (Hz)	**CWRU dataset (Hz)**	**Theoretical frequency (Hz)**
BPFI	166	162	162.18
BPFO	105	108	107.36
BSF	138	140	141.17

the errors between the MBD model and the theoretical frequencies are all within 2.3%, highlighting the accuracy of the proposed MBD model.

In the transfer learning part, the simulated dataset is divided into two datasets, 50% for training and 50% for testing to develop the initial pre-trained classifier. Similarly, the CWRU dataset is divided into two datasets, 33% for classifier fine-tuning and 67% for testing the fine-tuned classifier. Figure 6 (left) shows that the classifier cannot distinguish between BPFO and BSF and also misidentifies a part of healthy samples as fault samples, leading to a relatively low classification accuracy of 74.26 %. This result indicates that the gap between the simulated data and the experimental data is still large. However, after the fine-tuning process using the CWRU data, the results demonstrate a substantial improvement in fault diagnosis, achieving an accuracy up to 98.15%, as shown in Figure 6 (right). This significant change clearly shows the classifier successfully adapts to the experimental dataset.

CONCLUDING REMARKS

This work aims to enrich the field of study for condition monitoring using MBD simulation as the source of training data for fault diagnosis. A MBD model is first built and validated by the CWRU bearing dataset, successfully capturing the bearing fault frequencies. The flexibility of the MBD model in simulating diverse conditions and

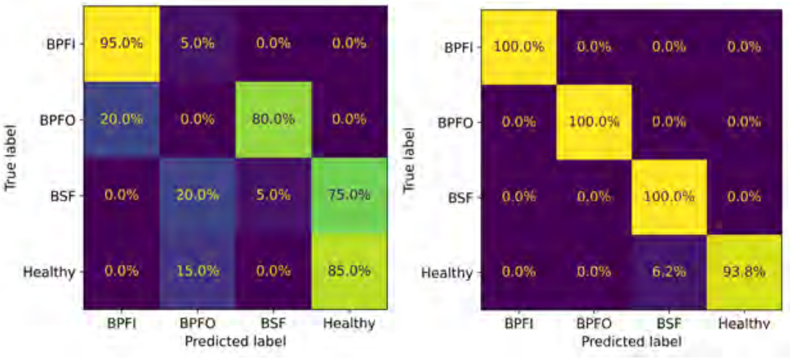


Figure 6. (left) Classification results before transfer learning and (right) classification results after transfer learning.

capturing realistic dynamic responses makes it a promising tool for advancing multi-fault diagnosis in future research.

To overcome the challenge of limited fault training data, a transfer learning strategy is presented, indicating a strong potential for enhancing model generalization. By fine-tuning the pre-trained classifier with a small portion of real data, the classification accuracy is significantly improved from 74.26% to 98.15%. This work lays a solid foundation for the integration of digital twin systems and transfer learning methods in advanced condition monitoring and structural health monitoring applications.

ACKNOWLEDGMENT

This work is carried out under the support of the Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Sweden.

REFERENCES

1. Hou, W., C. Zhang, Y. Jiang, K. Cai, Y. Wang, and N. Li. 2023. “A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data,” *Measurement: Journal of the International Measurement Confederation*, 215, ISSN 02632241, doi:10.1016/j.measurement.2023.112879.
2. Sassi, S., B. Badri, and M. Thomas. 2007. “A numerical model to predict damaged bearing vibrations,” *JVC/Journal of Vibration and Control*, 13:1603–1628, ISSN 10775463, doi: 10.1177/1077546307080040.
3. Sobie, C., C. Freitas, and M. Nicolai. 2018. “Simulation-driven machine learning: Bearing fault classification,” *Mechanical Systems and Signal Processing*, 99:403–419, ISSN 10961216, doi:10.1016/j.ymssp.2017.06.025.
4. Koutsoupakis, J., P. Seventekidis, and D. Giagopoulos. 2023. “Machine learning based condition monitoring for gear transmission systems using data generated by optimal multibody dynamics models,” *Mechanical Systems and Signal Processing*, 190:110130, ISSN 0888-3270, doi:10.1016/J.YMSSP.2023.110130.

5. Neupane, D. and J. Seok. 2020. “Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review,” *IEEE Access*, 8:93155–93178, ISSN 21693536, doi:10.1109/ACCESS.2020.2990528.
6. “MSC ADAMS, Hexagon,” https://shorturl.at/0XN9O, online; accessed 30 Apr 2025.
7. “SKF 6205 Deep groove ball bearing,” https://www.skf.com/group/products/rolling-bearings/ball-bearings/deep-groove-ball-bearings/productid-6205, online; accessed 30 Apr 2025.
8. Harris, T. and M. Kotzalas. 2006. *Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition*, Rolling Bearing Analysis, Fifth Edtion, CRC Press, ISBN 9781420006582.
9. Qin, W.-L., W.-J. Zhang, and C. Lu. 2016. “Rolling bearing fault diagnosis: A data-based method using EEMD, information entropy and one-versus-one SVM,” in *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pp. 1016–1020, doi:10.1109/WCICA.2016.7578388.
10. Liu, H., C. Liu, and Y. Huang. 2010. “Adaptive feature extraction using sparse coding for machinery fault diagnosis,” doi:10.1016/j.ymssp.2010.07.019.