

TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT
BỘ MÔN TỰ ĐỘNG HOÁ

BÁO CÁO HỌC THUẬT
Học kỳ II, năm học 2025-2026

**DESCRIPTION OF THE DATASET AND SIGNAL
PREPROCESSING METHODS FOR BEARING FAULT
DIAGNOSIS UNDER TIME-VARYING SPEED CONDITIONS**

Người thực hiện: TS. Phạm Thị Thanh Loan

Hà Nội, 5/2026

MỤC LỤC

I. INTRODUCTION

1.1. The role and importance of electric motors in modern industry

Electric motors are one of the most crucial and common devices in modern industrial, transportation, and energy systems. They are widely used in production lines, pump-fan-compressor systems, conveyors, industrial robots, electric vehicles, as well as in mining and renewable energy systems. According to many industrial statistics, electric motors consume 60–70% of the total electricity in manufacturing plants, demonstrating their central role in energy efficiency and operational reliability.

In the context of Industry 4.0 and digital transformation, the requirements for electric motors are not limited to high efficiency, but also include continuous operation, safety, low maintenance, and real-time monitoring of their condition. Therefore, traditional time-based maintenance strategies are increasingly revealing their limitations, especially for highly automated systems and harsh operating conditions.

1.2. Basic Structure of Electric Motors and Components Prone to Failure

Structurally, a typical rotating electric motor consists of main components such as: stator, rotor, shaft, bearings, housing, and cooling system. Specifically:

- The stator creates the rotating magnetic field and is generally less susceptible to mechanical damage during normal operation.
- The rotor and shaft are subjected to direct mechanical loads and can experience faults such as misalignment, imbalance, or fracture under harsh operating conditions.
- Rolling bearings are crucial intermediate components that allow the shaft to rotate with low friction while simultaneously bearing radial and/or axial loads.

In practice, bearings are the most prone to failure in electric motors, accounting for a large proportion of the causes of malfunctions and unplanned shutdowns. The cause stems from the fact that the bearings have to work continuously under the simultaneous influence of mechanical load, vibration, high temperature, assembly misalignment, and suboptimal lubrication conditions.

1.3. Common Types of Faults in Electric Motors, Focusing on Bearing Faults

Faults in electric motors can be classified into three main groups: electrical faults, mechanical faults, and mixed electrical-mechanical faults. Among these, mechanical faults, especially rolling bearing faults, are considered the most common cause of reduced motor efficiency and lifespan.

For rolling bearings, typical failure types include:

- Inner race fault: Cracks, pitting, or wear appear on the inner ring surface, often causing characteristic impact pulses with a frequency directly related to the shaft rotation speed.
- Outer race fault: Occurs on the fixed outer ring, leading to distinctly periodic vibration pulses.
- Ball/Roller fault: Involves rolling elements, often more difficult to detect due to their complex kinematic characteristics.
- Combination fault: multiple types of failure coexist simultaneously, reflecting severe bearing degradation.

These faults often manifest as non-stationary vibration signals, especially when the motor's rotational speed changes over time, making analysis and diagnosis significantly difficult.

1.4. The Importance of Engine and Bearing Fault Diagnosis

Engine fault diagnosis, especially bearing fault diagnosis, is crucial to ensuring the reliability, safety, and economic efficiency of industrial systems. Early detection of potential failures allows for:

- Minimizing the risk of serious breakdowns and sudden machine shutdowns.
- Optimizing condition-based maintenance strategies.
- Extending equipment lifespan and reducing operating and maintenance costs.
- Ensuring safety for people and systems, especially in heavy industries and mining.

In practice, many diagnostic methods based on vibration signal and frequency spectrum analysis have been developed and successfully applied under constant rotational

speed conditions. However, when rotational speed varies over time, the spectral characteristics of bearing faults will shift and distort, significantly reducing the effectiveness of traditional diagnostic methods.

1.5. Research Motivation: Fault Diagnosis Under Variable Speed Conditions

Stemming from the challenges mentioned above, research into methods for diagnosing bearing faults under time-varying rotational speed has become a crucial and highly relevant research direction. To evaluate, compare, and develop advanced diagnostic algorithms, the existence of high-quality experimental datasets collected under non-stationary conditions is particularly important.

In this context, the Ottawa 2018–2019 bearing vibration dataset published by the University of Ottawa has become a valuable scientific reference data source, serving research on signal analysis, data preprocessing, and fault diagnosis based on machine learning and deep learning under time-varying speed conditions.

1.6. Overview of Current Bearing Fault Diagnosis Approaches

For decades, bearing fault diagnosis has attracted significant attention from the research and industrial communities due to the high failure rate and the serious impact of bearings on the operational reliability of electric motors. Bearing fault diagnosis methods can be classified according to the signal analysis domain, including: time domain, frequency domain, time-frequency domain, and angle domain. Each approach has its own advantages and limitations, especially when considering conditions of varying rotational speed.

1.6.1. Time-domain diagnostic methods

Time-domain methods are the most basic and intuitive approach, based on the direct analysis of vibration signals over time. Commonly used statistical indicators include mean, standard deviation, RMS value, crest factor, kurtosis, and skewness.

The main advantages of time-domain methods are their simplicity, ease of implementation, and low computational cost, making them suitable for online monitoring systems. In particular, indicators such as kurtosis and crest factor are sensitive to short-term shock impulses, which often occur in the early stages of bearing failure.

However, under conditions where the rotational speed changes over time, the vibration signal becomes non-stationary, leading to significant variations in statistical indicators. Consequently, it becomes difficult to distinguish between time-domain characteristics resulting from changes due to failure and changes due to operating conditions, significantly reducing the reliability of the method

1.6.2. Frequency-domain methods

Frequency domain analysis is the most common and widely applied approach in bearing fault diagnosis. This method relies on the Fourier transform (FFT) to convert vibration signals from the time domain to the frequency domain, thereby exploiting characteristic frequencies of bearing faults such as BPFI (Ball Pass Frequency Inner Race), BPFO (Ball Pass Frequency Outer Race), BSF (Ball Spin Frequency), and FTF (Fundamental Train Frequency).

Under constant rotational speed conditions, the characteristic frequency components of bearing faults appear at stable positions on the spectrum, allowing for highly accurate fault detection and classification. This forms the basis of many current industrial vibration monitoring systems.

However, when rotational speed varies, these characteristic frequencies shift over time, leading to "spectral smearing" and blurring of fault spectral peaks. Therefore, traditional frequency analysis methods prove ineffective under non-stationary conditions, especially when the rate of change is rapid or non-linear.

1.6.3. Time–frequency methods

To overcome the limitations of frequency-domain methods under non-stationary conditions, time-frequency analysis methods have been developed, allowing simultaneous representation of information in both time and frequency domains. Typical techniques include: Short-Term Fourier Transform (STFT), Wavelet Transform (WT), Wavelet Packet Transform (WPT), and Hilbert–Huang Transform (HHT).

A key advantage of these methods is their ability to track the variation of the frequency component over time, thereby improving the detection of bearing faults under varying speeds. Wavelet transforms and HHT are particularly effective in detecting short bursts and non-linear signals.

Nevertheless, time-frequency methods often have high computational complexity, are heavily dependent on parameter selection (window, wavelet function, number of decay levels), and are difficult to implement in real-time monitoring systems. Furthermore, time-frequency representations are still affected by domain shift when the rate changes sharply between training and test conditions.

1.6.4. Angle-domain methods

Angular domain analysis is considered one of the most effective approaches for processing vibration signals under varying rotational speeds. The core idea of this method is to convert the signal from the time domain to the angular domain of the shaft's rotation through angle-domain resampling based on the encoder signal or speed estimate.

In the angular domain, the mechanical phenomena associated with each shaft rotation appear at fixed positions, independent of changes in speed over time. This makes the characteristic frequencies of bearing faults relatively invariant, significantly improving fault detection and classification under non-stop conditions.

However, the angular domain method requires highly accurate speed or encoder signals, increasing the complexity of data acquisition and preprocessing. In practice, speed measurement errors, encoder noise, or signal loss can significantly affect the quality of the angular domain signal

1.6.5. Trends in Combining Preprocessing and Machine Learning in Bearing Fault Diagnosis

Recently, bearing fault diagnosis methods have increasingly tended to combine signal preprocessing techniques (such as angular domain normalization, SGA, STFT, wavelet) with machine learning and deep learning models. The goal is to extract stable features, reduce the influence of domain shift, and improve the generalizability of the model under different operating conditions.

In this context, preprocessing algorithms such as the Signal Generalization Algorithm (SGA) play a crucial role in normalizing vibration signals collected at different speeds, facilitating the training and evaluation of diagnostic models based on data.

Bảng 1. So sánh các hướng tiếp cận chẩn đoán lỗi ổ bi lăn

Method	Analysis domain	Typical features	Advantages	Main limitations	Suitable for variable speed
Time-domain	Time	RMS, kurtosis, crest factor	Simple, fast computation, easy for online implementation	Sensitive to noise and operating conditions; difficult to distinguish fault types	Low
Frequency-domain	Frequency	BPFI, BPFO, BSF, harmonics	Intuitive and effective under steady-speed conditions	Spectral smearing under variable speed; domain shift	Low
Time–Frequency	Time–frequency	STFT, WT, HHT, WPT	Tracks non-stationary signals and is effective for impulse detection	Complex, parameter-dependent, high computational cost	Medium
Angle-domain	Angular domain	Order spectrum, angular FFT	Speed-invariant features; highly suitable for non-stationary conditions	Requires accurate encoder/speed information; complex preprocessing	High
Hybrid + ML/DL	Multi-domain	Automatically learned features	High accuracy and good generalization capability	Data-dependent; requires appropriate preprocessing	Very high

1.7. Summary of the Report's Research and Guidance

From the overview above, it can be seen that, although many methods for diagnosing bearing faults have been proposed, the problem of diagnosis under variable rotational speed conditions still presents many challenges. Traditional time-domain and frequency-domain methods prove ineffective in non-stationary environments, while time-frequency and angular domain methods, despite their clear advantages, still have limitations in terms of complexity and data requirements.

Therefore, this report focuses on exploiting the Ottawa 2018–2019 dataset and the SGA preprocessing method to study the possibility of reducing domain shift and improving

the efficiency of bearing fault diagnosis under variable speed conditions, laying the foundation for further research and practical applications.

II. GENERAL INFORMATION ABOUT THE OTTAWA DATABASE

2.1. Overview of the Ottawa Bearing Fault Dataset

This study employs the bearing vibration dataset released by the University of Ottawa (Canada) in 2018–2019, which has been widely recognized as a benchmark dataset for fault diagnosis under time-varying rotational speed conditions. Unlike many classical bearing datasets collected under constant speed, the Ottawa dataset explicitly considers non-stationary operating conditions, making it particularly suitable for evaluating advanced signal processing and fault diagnosis algorithms.

The dataset is publicly available through the Mendeley Data repository and consists of two main versions: Version 1 (V1), released in 2018, and Version 2 (V2), released in 2019 with additional fault cases and experimental repetitions.

2.2. Experimental Test Rig and Measurement Setup

The data were collected using a SpectraQuest Machinery Fault Simulator (MFS-PK5M), which is a modular test rig designed to emulate common mechanical faults in rotating machinery. The test rig is driven by an electric motor controlled via an AC drive, allowing flexible programming of rotational speed profiles.

A rolling bearing (ER16K) is mounted on the test shaft, where different fault types are artificially introduced. The vibration signals are measured using an ICP® piezoelectric accelerometer mounted directly on the bearing housing, ensuring high sensitivity to localized defect-induced impacts. In addition, an incremental encoder is employed to record shaft rotational speed and angular position.

Each data record contains two synchronized channels:

- Channel 1: Vibration acceleration signal.
- Channel 2: Encoder signal representing shaft speed/angle.

This dual-channel configuration enables both time-domain analysis and angle-domain signal processing.

2.3. Bearing Specifications and Fault Characteristics

The tested bearing is a deep-groove ball bearing ER16K. Its key geometric parameters include pitch diameter, ball diameter, and the number of rolling elements, which directly determine the theoretical fault characteristic frequencies.

Based on bearing kinematics, the characteristic frequencies associated with inner race faults (BPFI) and outer race faults (BPFO) can be expressed as multiples of the shaft rotational frequency. These frequencies serve as critical references for frequency-domain and order-domain fault diagnosis.

Importantly, under time-varying speed conditions, these characteristic frequencies are no longer fixed, which significantly increases the difficulty of fault detection and motivates the use of advanced preprocessing techniques.

2.4. Data Acquisition and Sampling Parameters

All vibration signals are sampled at a high sampling frequency of 200 kHz to ensure accurate capture of high-frequency impact components generated by bearing defects. Each data file corresponds to a 10-second recording, resulting in two time-synchronized vectors stored in MATLAB (.mat) format.

The high sampling rate, combined with the availability of encoder data, makes the dataset particularly suitable for investigating:

- Non-stationary signal characteristics,
- Speed-dependent spectral shifting,
- Angle-domain resampling and normalization techniques.

2.5. Fault Categories and Speed Profiles

The dataset includes multiple bearing health conditions, namely:

- Healthy condition,
- Inner race fault,
- Outer race fault,
- Ball fault,
- Combination fault.

To simulate realistic operating scenarios, each fault condition is tested under four distinct speed profiles: increasing speed, decreasing speed, increasing–decreasing speed, and decreasing–increasing speed. Each experiment is repeated multiple times to enhance statistical reliability.

This structured design enables systematic evaluation of fault diagnosis methods across both fault types and speed variation patterns.

2.6. Data Organization and File Naming Convention

Each data file follows a consistent naming convention that encodes both the bearing condition and the speed profile. This facilitates automated data loading, labeling, and batch processing in machine learning pipelines.

The standardized structure of the dataset makes it well suited for supervised learning, domain adaptation, and cross-condition generalization studies.

2.7. Challenges and Research Opportunities

Despite its high quality, the Ottawa dataset presents significant challenges due to its non-stationary nature. The continuous variation of rotational speed causes time-varying spectral characteristics and domain mismatch between different operating conditions.

These challenges make the dataset an ideal testbed for evaluating preprocessing strategies such as Signal Generalization Algorithm (SGA), angle-domain resampling, and hybrid time–frequency representations, particularly when combined with machine learning-based fault diagnosis models.

III. EXPERIMENTAL SETUP

The experimental setup utilizes the SpectraQuest Machinery Fault Simulator (MFS-PK5M). The motor speed is controlled by an AC drive. The sensor and measurement system includes:

- One ICP Model 623C01 accelerometer mounted on the bearing housing to measure vibration.
- One Incremental Encoder Model EPC 775 to measure shaft speed/phase.
- NI USB-6212 BNC data acquisition module; each data set consists of two channels:
 - + Channel_1: vibration signal (acceleration).
 - + Channel_2: speed/encoder signal.

Data is stored as a .mat file

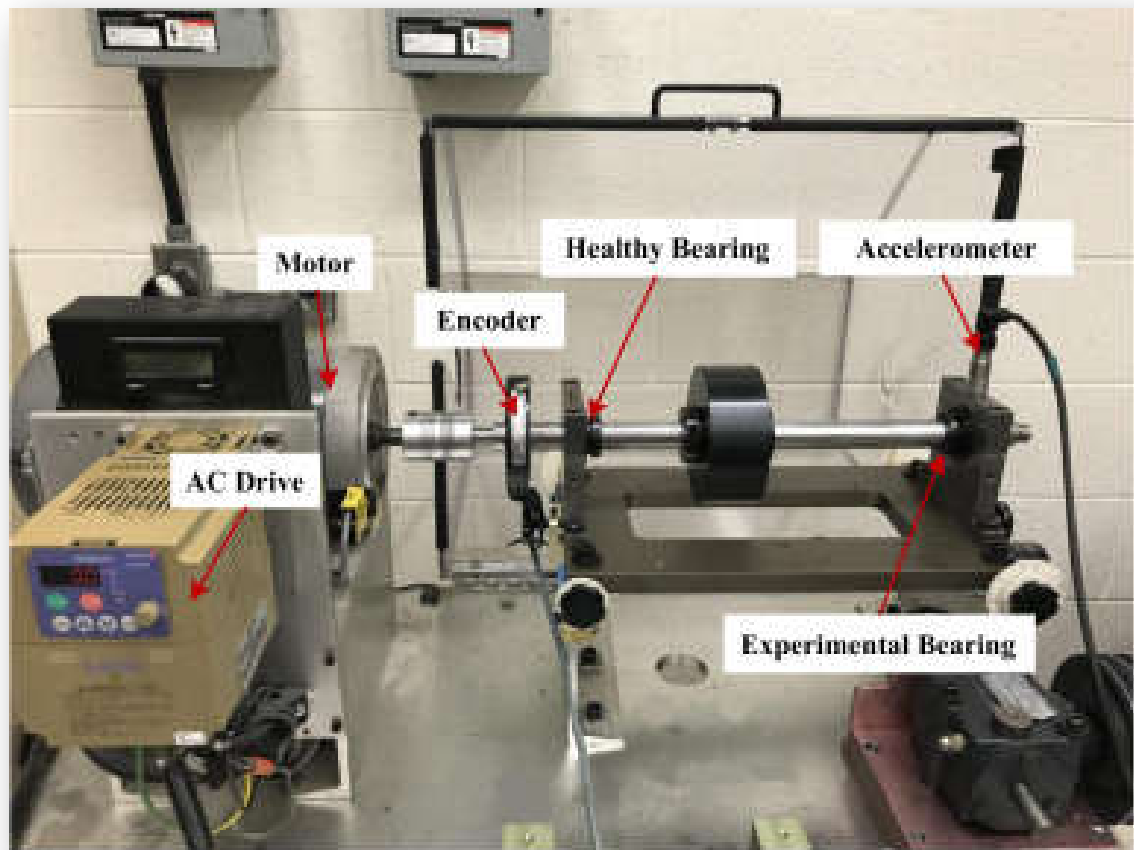


Figure 1. Data collection rig

KEY TECHNICAL INFORMATION

3.1. Specifications of the ICP Model 623C01 accelerometer sensor

Attribute	Value	Gia tốc kế
Sensing principle	Ceramic shear, industrial type, high frequency	
Sensitivity	100 mV/g $\pm 5\%$ ≈ 10.2 mV/(m/s ²)	
Frequency range (± 3 dB)	0.8–15 kHz, equivalent to 48–900,000 cpm	
Measurement range	± 50 g, equivalent to ± 490 m/s ²	
Weight	~ 51 g	
		Figure 2. Accelerometer

3.2. ER16K ball bearing specifications

Parameter	Value	Vòng bi cầu ER16K
Pitch diameter	38.52 mm	
Ball diameter	7.94 mm	
Number of balls	9	
Bearing fault characteristic frequency — inner race (BPFI)	$5.43 \cdot fr$	
Bearing fault characteristic frequency — outer race (BPFO)	$3.57 \cdot fr$	
		Figure 3. Ball bearing

3.3. Data collection parameters

Item	Description
Data acquisition module	NI Data Acquisition boards, NI USB-6212 BNC
Channel 1	Accelerometer for vibration measurement
Channel 2	Encoder for shaft speed measurement

Dataset structure	Each dataset is sampled from 2 channels
Data storage format	Data collected from the 2 channels are stored in .mat files
Channel_1	Vibration data

3.4. Data acquisition hardware

Data acquisition module: NI Data Acquisition boards (NI USB-6212 BNC)	 Figure 4. NI Data Acquisition boards (NI USB-6212 BNC)
Channel 1: Accelerometer measures vibration data.	
Channel 2: Rotary axis data measurement encoder	
Each sampled dataset consists of 2 channels.	
Data collected from the two channels is stored in a .mat file.	
+ "Channel_1": Vibration data + "Channel_2": Rotation speed data	

DATABASE STRUCTURE DESCRIPTION

3.5. Collected datasets

Total number of datasets: 110. The dataset design is categorized into two groups of factors

(1) Mechanical errors

- R-U: Rotor Unbalance
- R-M: Rotor Misalignment
- B-R: Bowed Rotor
- K-A: Broken Rotor Bars
- F-B: Faulty Bearings.

(2) Electrical error

- S-W: Stator Winding Fault.

- V-U: Voltage Unbalance and Single Phasing.

Table 1. Differentiating types of errors

Condition Type	Abbreviation	Fault Category
Healthy Motor	H-H	Healthy Motor
Rotor Unbalance	R-U	Mechanical Fault
Rotor Misalignment	R-M	Mechanical Fault
Stator Winding Fault	S-W	Electrical Fault
Voltage Unbalance and Single Phasing	V-U	Electrical Fault
Bowed Rotor	B-R	Mechanical Fault
Broken Rotor Bars	K-A	Mechanical Fault
Faulty Bearings	F-B	Mechanical Fault

Table 2. Detailed description of error types

Number	Meaning (Speed Condition)	Description of Operating Frequency	Unloaded/Loaded
1	Constant Speed	15 Hz	Fixed operating frequency around 15 Hz (Example: 15.10 Hz)
2	Constant Speed	30 Hz	Fixed operating frequency around 30 Hz (Example: 29.98 Hz)
3	Constant Speed	45 Hz	Fixed operating frequency around 45 Hz (Example: 44.91 Hz)
4	Constant Speed	60 Hz	Fixed operating frequency around 60 Hz (Example: 59.80 Hz)
5	Increasing Speed	15 Hz → 45 Hz	H-H-5-0 (Healthy, Unloaded), V-U-5-1 (Voltage Fault, Loaded), B-R-5-1 (Bowed Rotor – Mechanical Fault, Loaded)
6	Increasing Speed	30 Hz → 60 Hz	H-H-6-0 (Healthy, Unloaded),

			R-M-6-1 (Rotor Misalignment, Loaded), increasing from approximately 30 Hz to 60 Hz
7	Decreasing Speed	45 Hz → 15 Hz	B-R-7-0 (Bowed Rotor, Unloaded), R-U-7-1 (Rotor Unbalance, Loaded), decreasing from 44.92 Hz to 15.75 Hz
8	Decreasing Speed	60 Hz → 30 Hz	H-H-8-0 (Healthy, Unloaded), decreasing from 59.79 Hz to 30.57 Hz V-U-8-1 (Voltage Fault, Loaded), decreasing from 59.90 Hz to 29.79 Hz

The dataset, collected from bearings operating under conditions of varying rotational speed over time, more accurately reflects the actual operating modes of industrial electromechanical systems. Unlike most previously published datasets, which were collected under conditions of constant rotational speed, this dataset allows for the study and analysis of bearing vibration behavior in non-stationary environments. Therefore, the data can be used to investigate the variation of frequency characteristics corresponding to different bearing states as rotational speed changes. Simultaneously, the dataset also serves as an important reference platform for evaluating the effectiveness of newly developed methods for fault diagnosis or bearing condition monitoring under conditions of varying speed over time

3.6. Evaluate the data

1. Kurtosis methods

$$K = \frac{\frac{1}{N} \sum_{i=0}^N (X_i - \mu)^4}{\sigma^4}$$

- K : Sharpness
- N : Number of data points
- X_i : Amplitude value of the frequency spectrum at frequency i (i -th signal value)

- The mean value of the signal is calculated by $\frac{1}{N} \cdot \sum_{i=1}^N X_i$
- σ : Standard Deviation.

In this method, kurtosis (denoted by K) is used as a statistical indicator to assess the degree of occurrence of transient components in the frequency spectrum of a vibration signal. With a dataset of N spectral points, where X_i represents the amplitude of the frequency spectrum at the i th frequency, the average value of the signal is determined by the arithmetic mean of the spectral amplitudes. The standard deviation σ reflects the degree of dispersion of the amplitude values around the mean. Based on this, kurtosis is calculated to measure the degree of "sharpness" of the spectral distribution, thereby highlighting the existence of abnormal spectral peaks. In bearing fault diagnosis, this index is particularly sensitive to shock impulses generated by local defects on the inner ring, outer ring, or rollers, and is therefore often used for early detection of failure under conditions of nonlinearity and high noise in the signal.

2. Root Mean Square method

$$RMS = \sqrt{\frac{1}{N} \sum_{i=0}^N |X_i|^2}$$

- RMS : Mean Square
- N : Number of data points
- X_i : The amplitude value of the frequency spectrum at i

SCIENTIFIC VALUE AND APPLICATION POTENTIAL

The Ottawa dataset is particularly valuable because it was collected under conditions of time-varying rotational speed—unlike many classical datasets that only consider steady speed. Therefore, the dataset is suitable for:

- Studying the spectral/frequency characteristics of bearing failures under non-stationary conditions.
- Evaluating the effectiveness of diagnostic, fault detection, and condition monitoring methods when speed changes.

- Studying the domain shift problem caused by peak-to-peak spectral shift with speed, and techniques for normalizing/synchronizing the domain characteristics.

DATA EVALUATION

Some guiding observations from the signal analysis:

- The 10–25 Hz range often shows shaft-related components.
- The 54–135 Hz range often shows inner loop faults.
- The 37–89 Hz range often shows outer loop faults.

In addition, consider the 1X, 2X, 3X,... harmonics of the characteristic frequencies..

PRE-PROCESSING METHOD: FFT

The Fast Fourier Transform (FFT) is a popular signal preprocessing and analysis tool in the diagnosis of electric motor faults, particularly for oscillation signals, stator current, audio signals, or leakage flux. Essentially, the FFT allows the conversion of time-domain signals to frequency-domain signals, thereby clarifying the energy distribution of the signal at characteristic frequency components. In motor fault diagnosis, faults such as rotor imbalance, shaft misalignment, rotor bending, rotor conductor breakage, bearing failure, or stator winding failure often produce abnormal frequency components or alter the amplitude of spectral components related to rotational frequency, power supply frequency, and characteristic harmonic families. Therefore, instead of simply analyzing the waveform directly in the time domain, the signal after FFT transformation can provide clearer information about the nature of oscillations and the fault-causing mechanisms of the system.

In the data preprocessing process, the measured signal is first normalized to the same sample length and sampling frequency to ensure consistency across datasets. Then, the signal can be subjected to DC component removal, noise filtering, or multiplication with a window function such as Hanning/Hamming to reduce spectral leakage when performing Fourier transform on a finite signal segment. The resulting FFT generates an amplitude spectrum as a function of frequency, where each spectral point represents the contribution of a specific frequency component in the original signal. For engine fault diagnosis, this amplitude spectrum can be used directly as input to machine learning/deep learning models, or used to extract quantitative features such as amplitude at rotational frequency,

amplitude of higher-order harmonics, spectral energy in each frequency band, dominant frequency, spectral deviation, and energy ratio between frequency bands. These characteristics reflect changes in the engine's dynamic state when a fault occurs, thereby improving the ability to distinguish between a healthy state and various types of faults.

In this study, the Fast Fourier Transform (FFT) is used as a preprocessing step to convert the measured signal from the time domain to the frequency domain, thereby highlighting characteristic spectral components related to the operating state and fault modes of the motor. Assuming the discrete signal obtained from the sensor or measurement system is $x[n]$, with $n=0,1,\dots,N-1$, the Discrete Fourier Transform (DFT) of the signal is determined by:

$$X[k] = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N-1$$

where $X[k]$ is the complex spectral coefficient at frequency index k , N is the number of samples in a signal segment, and j is the imaginary unit. The FFT is an efficient computational algorithm for the DFT, reducing the computational complexity from $O(N^2)$ to $O(N\log N)$, thus suitable for motor signal processing problems with a large number of samples or requiring repetitive processing on multiple data segments.

The physical frequency corresponding to spectral index k is determined according to:

$$f_k = \frac{k f_s}{N}$$

where f_s is the sampling frequency of the signal. The one-sided amplitude spectrum is often used to represent the contribution of each frequency component in a real signal and can be approximated as:

$$A[k] = \frac{2}{N} |X[k]|, \quad k = 1, 2, \dots, 2N-1$$

The one-dimensional and Nyquist components are usually not multiplied by a factor of 2. Based on the amplitude spectrum $A[k]$, frequency domain features can be constructed, such as amplitude at characteristic frequencies, spectral energy in a frequency band, or the ratio of energy between different frequency bands. The spectral energy in the frequency band from f_a to f_b can be expressed as:

$$E_{[f_a, f_b]} = \sum_{k=k_a}^{k_b} |X[k]|^2$$

where k_a and k_b are the spectral indices corresponding to the lower and upper limits of the frequency band under investigation, respectively.

For motor fault diagnosis, defects such as rotor imbalance, shaft misalignment, rotor warping, bearing failures, rotor conductor breakage, or stator winding failures often alter the spectral energy distribution of the signal. These changes can manifest as the appearance of abnormal spectral peaks, increased amplitude at rotational frequencies and harmonic families, or energy shifts in frequency bands related to the fault mechanism. Therefore, the use of FFT allows for the extraction of characteristic information that is difficult to directly represent in time-domain waveforms. In the implementation process, the original signal is normalized to the same sample length, the DC component is removed, and it can be multiplied by the window function $w[n]$ to reduce spectral leakage:

$$x_w[n] = x[n]w[n]$$

Subsequently, the signal $x_w[n]$ is fed into the FFT transform to obtain a frequency spectrum for feature extraction or directly as input to the classification model. Thanks to its ability to clearly describe the frequency structure of the signal, the FFT enhances the distinction between healthy and faulty states, and provides a clear physical basis for interpreting diagnostic results.

IV. DATA VISUALIZATION

Data visualization is an essential preliminary step in bearing fault diagnosis, especially when the vibration signals are collected under time-varying speed conditions. By observing the raw acceleration signals and their transformed frequency-domain representations, it is possible to preliminarily identify differences in amplitude variation, transient impulses, modulation patterns, and speed-dependent spectral changes among healthy and faulty conditions. These visual observations provide an intuitive understanding of the non-stationary characteristics of the dataset and help reveal how different fault types affect the vibration response under acceleration, deceleration, unloaded, and loaded operating modes. Therefore, visualization not only supports qualitative assessment of the collected data but also provides an important basis for selecting suitable preprocessing methods, extracting diagnostic features, and designing machine-learning or deep-learning models for subsequent fault classification.

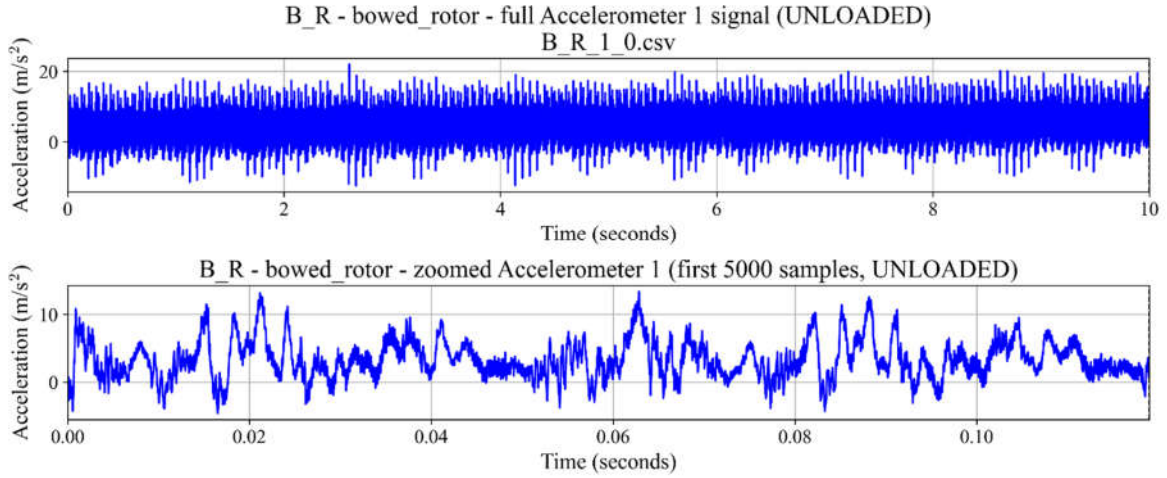


Figure 1. Raw data of bowed rotor fault_unloaded

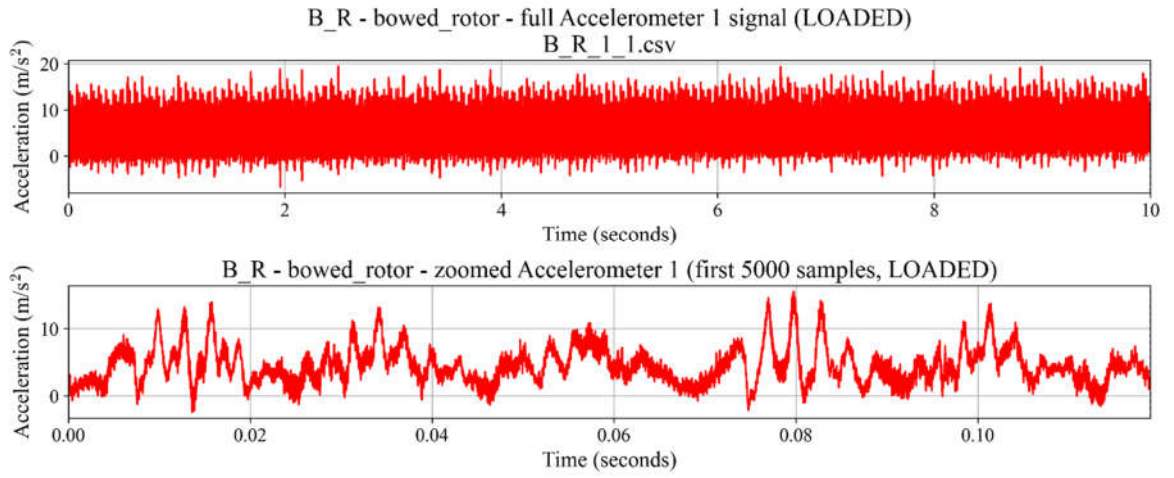


Figure 2. Raw data of bowed rotor fault_loaded

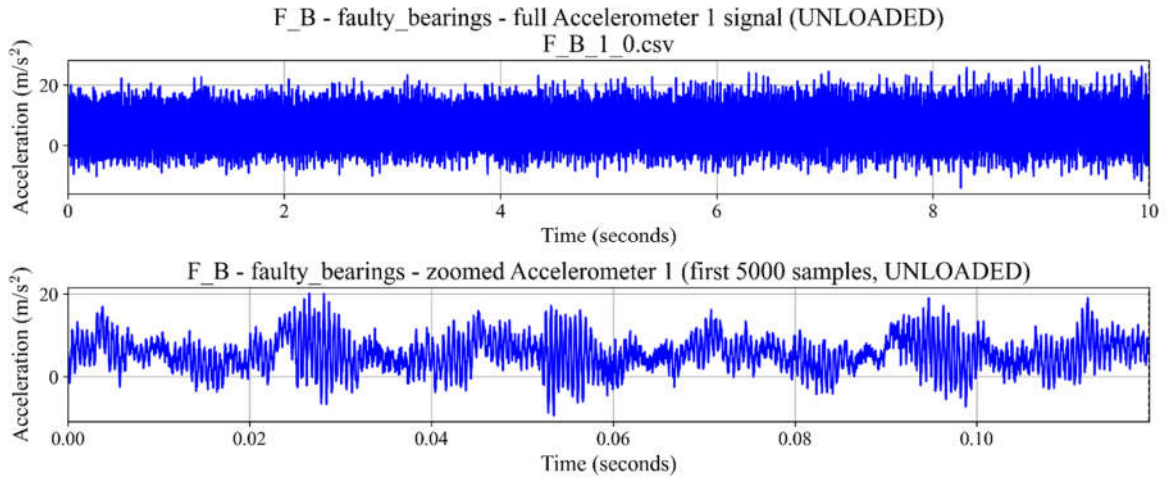


Figure 3. Raw data of bearing failure, unloaded

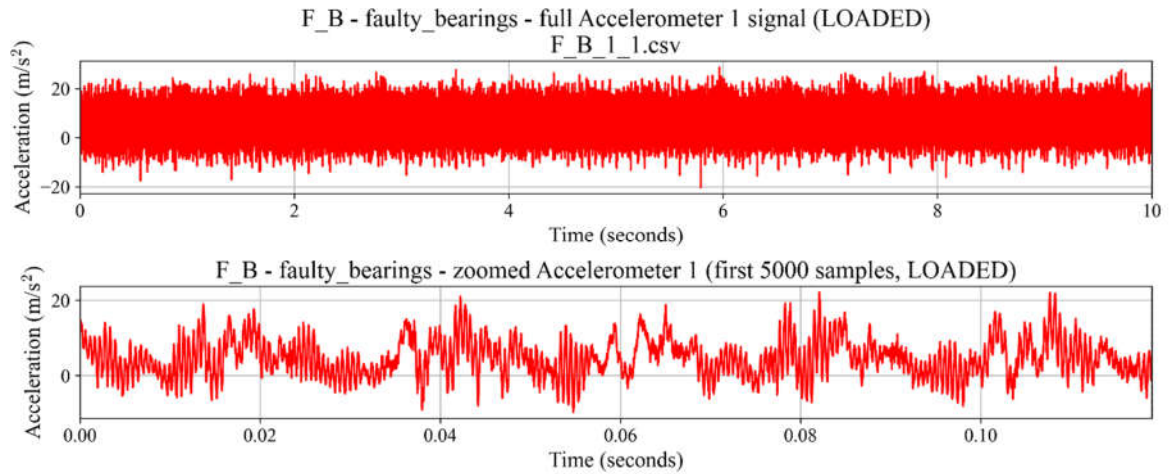


Figure 4. Raw data of bearing failure, loaded

4.1. Visualizing raw data during deceleration

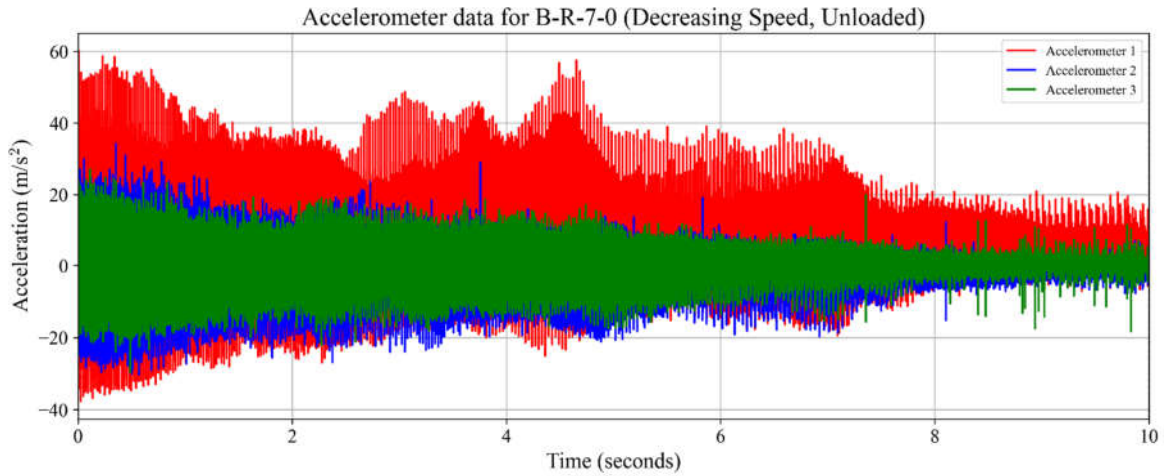


Figure 5. Bowed Rotor 44.92 Hz to 15.75 Hz

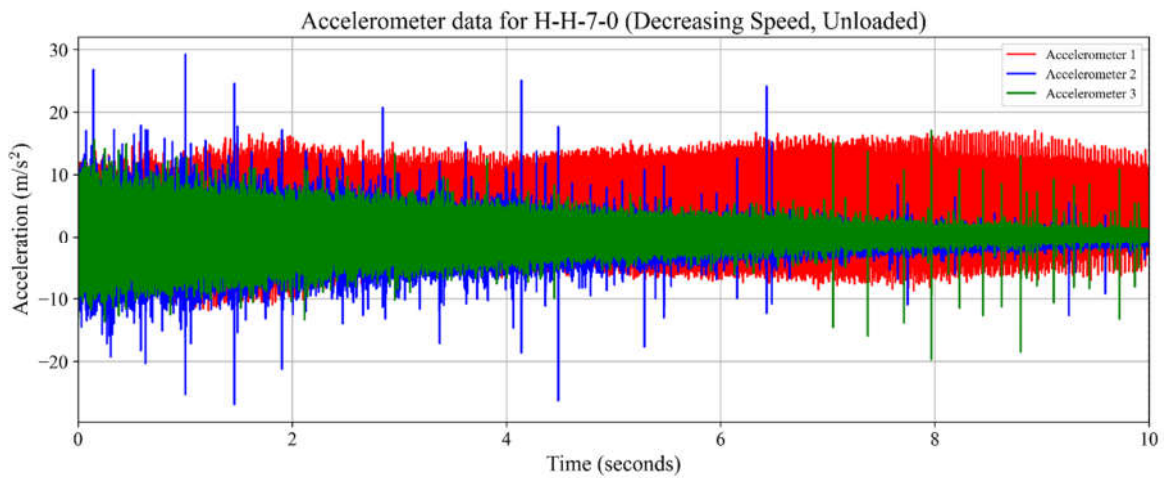


Figure 6. Healthy. 44.89 Hz $\rightarrow$ 15.12 Hz

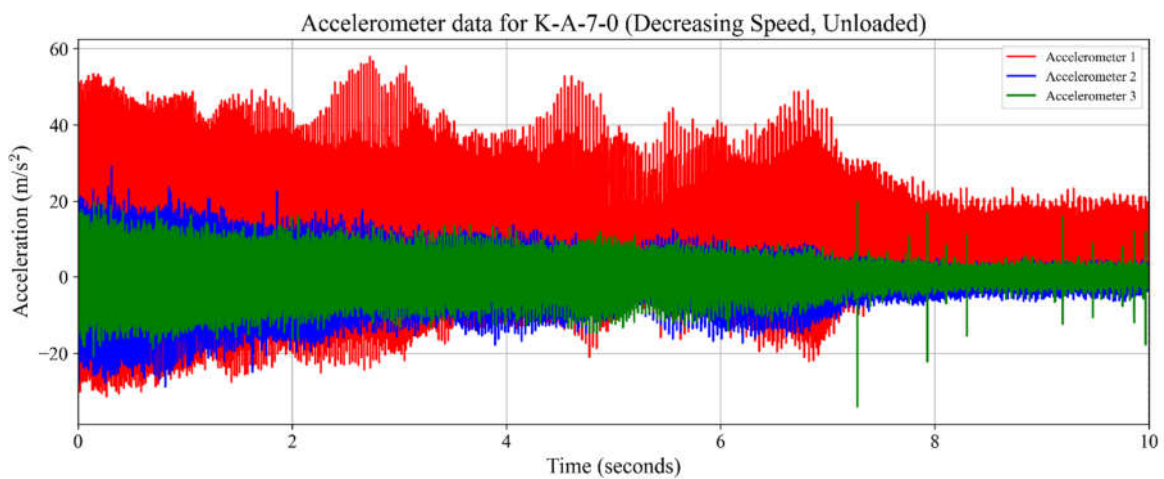


Figure 7. Broken Rotor Bars 44.96 Hz $\rightarrow$ 16.81 Hz

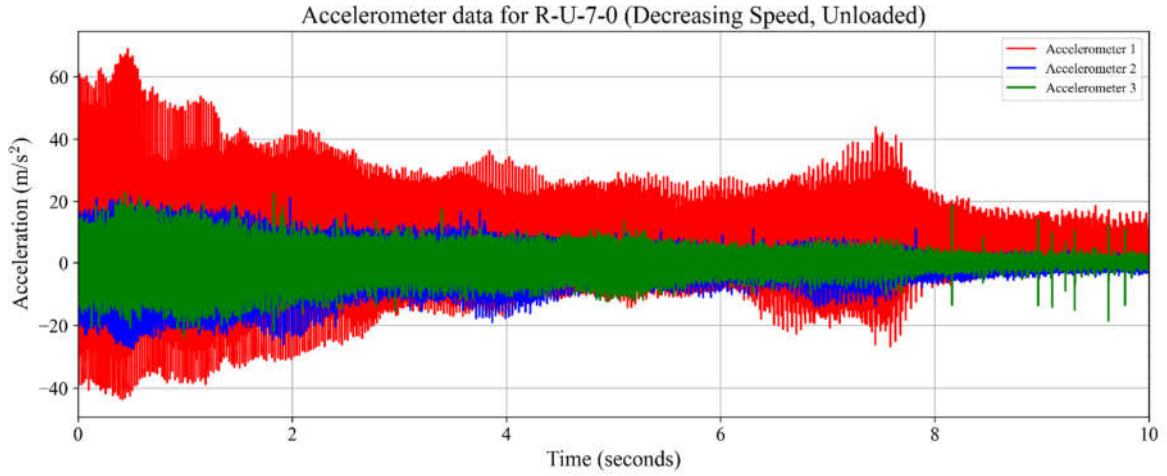


Figure 8. Rotor Unbalance 44.97 Hz → 18.82 Hz

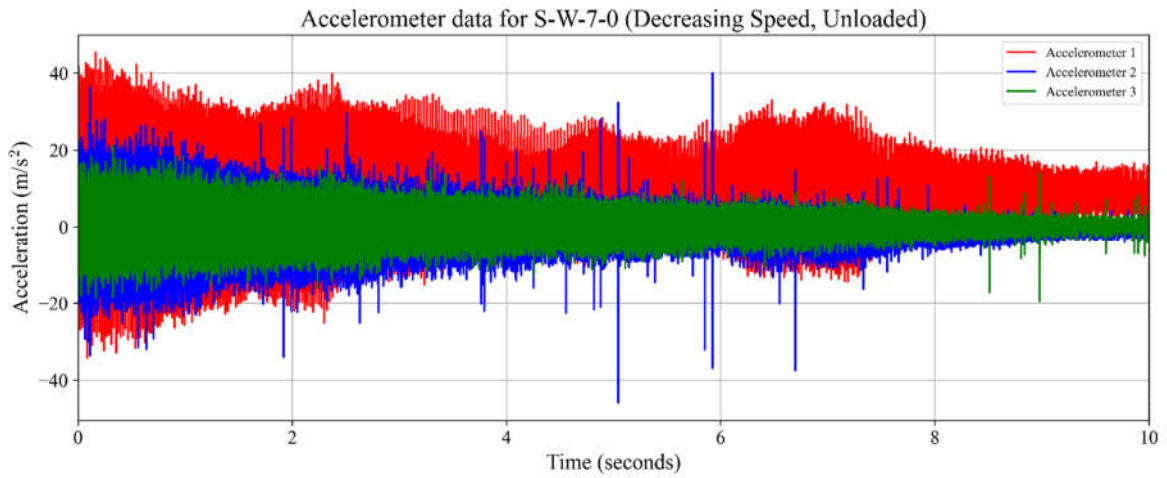


Figure 9. Stator Winding Fault. 44.92 Hz → 15.65 Hz

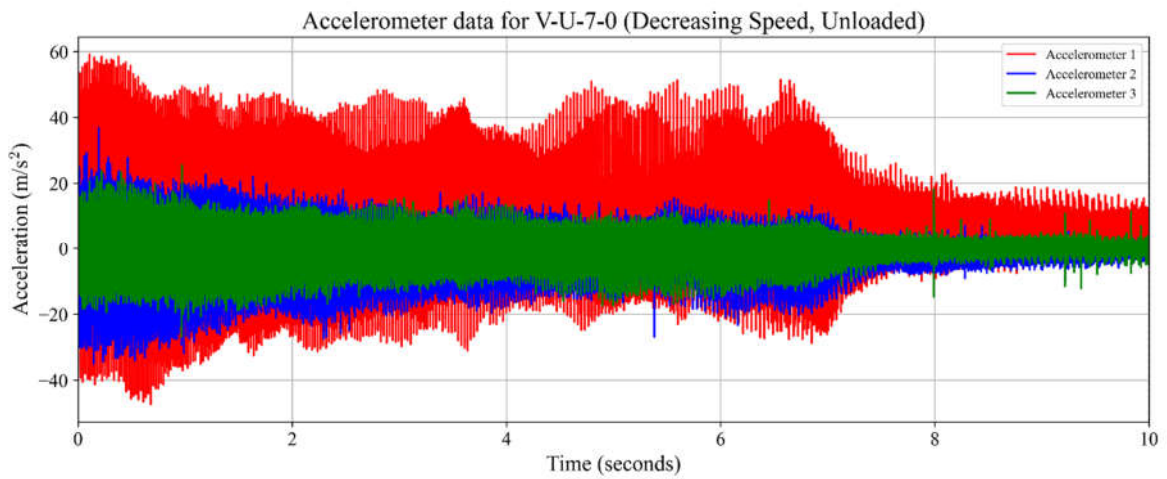


Figure 10. Voltage Unbalance and Single Phasing 44.91 Hz → 15.32 Hz

4.2. Visualizing raw data during acceleration

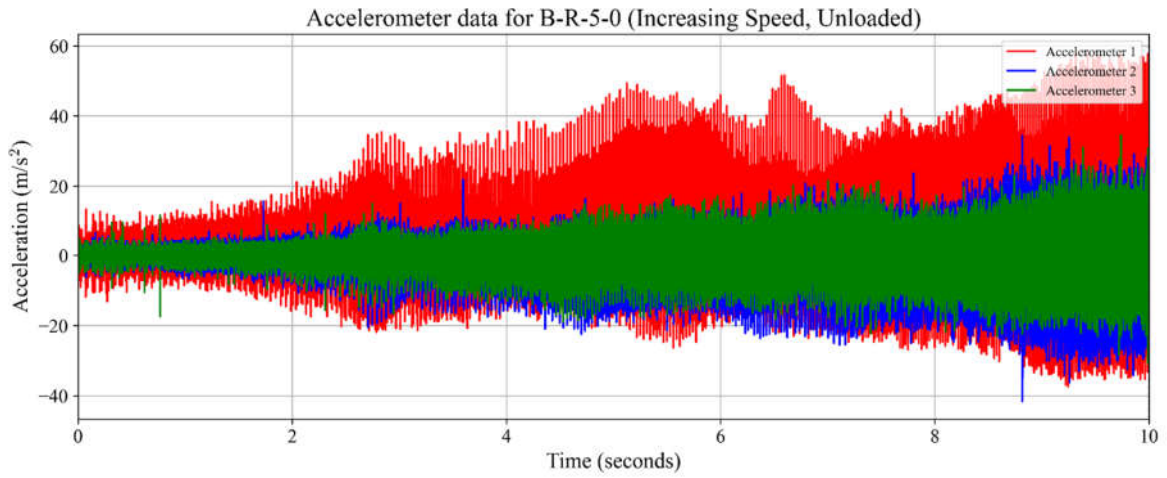


Figure 11. Bowed Rotor 15.81 Hz → 44.81 Hz

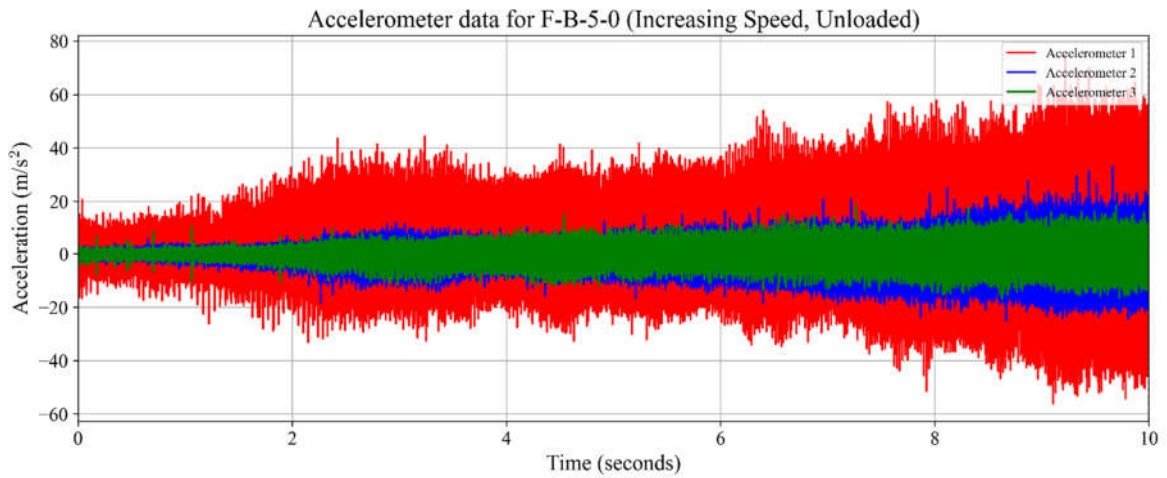


Figure 12. Faulty Bearings 15.22 Hz → 44.96 Hz

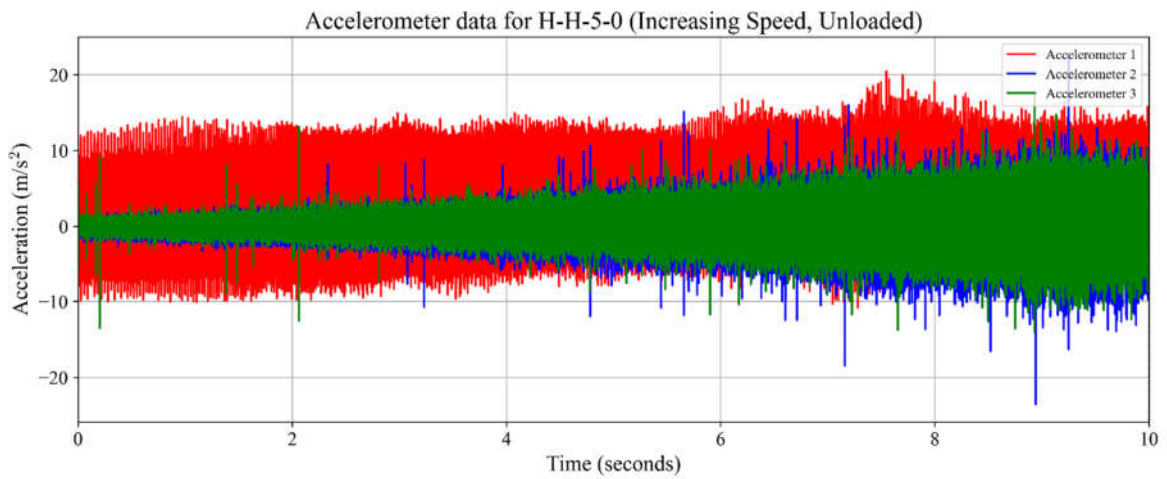


Figure 13. Healthy 14.99 Hz → 44.86 Hz

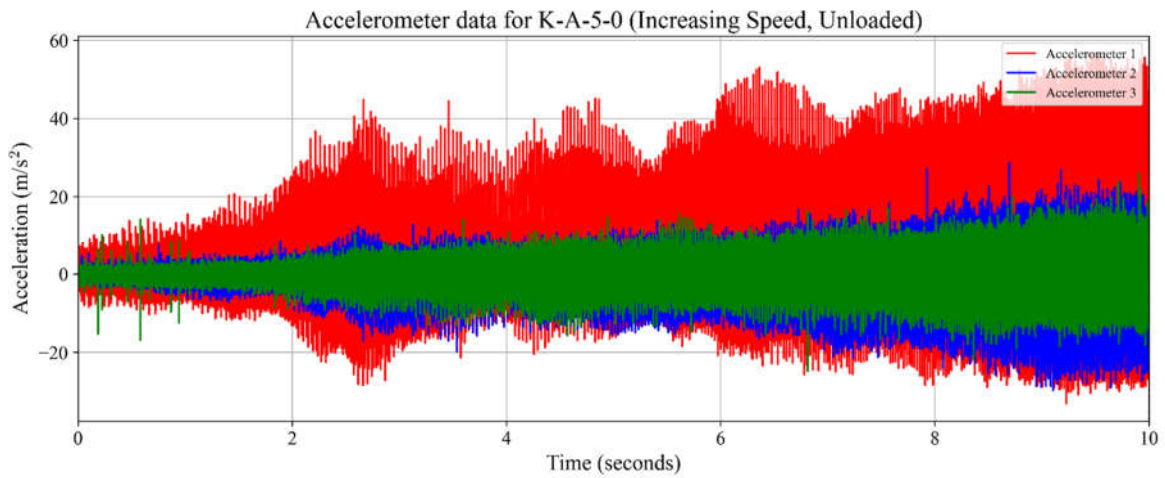


Figure 14. Broken Rotor Bars 15.13 Hz $\rightarrow$ 44.87 Hz

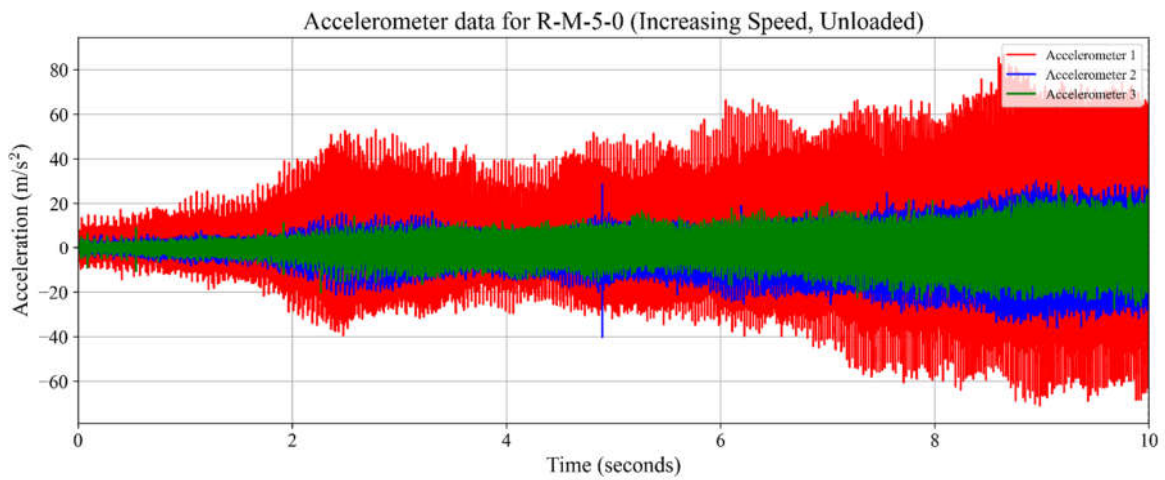


Figure 14. Rotor Misalignment 15.82 Hz $\rightarrow$ 44.91 Hz

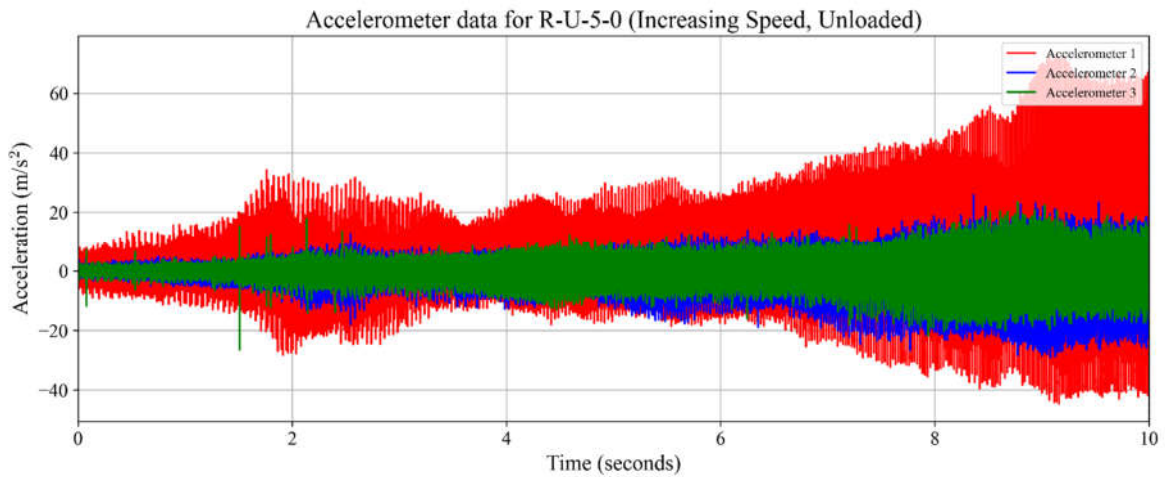


Figure 15. Rotor Unbalance 15.72 Hz $\rightarrow$ 44.97 Hz

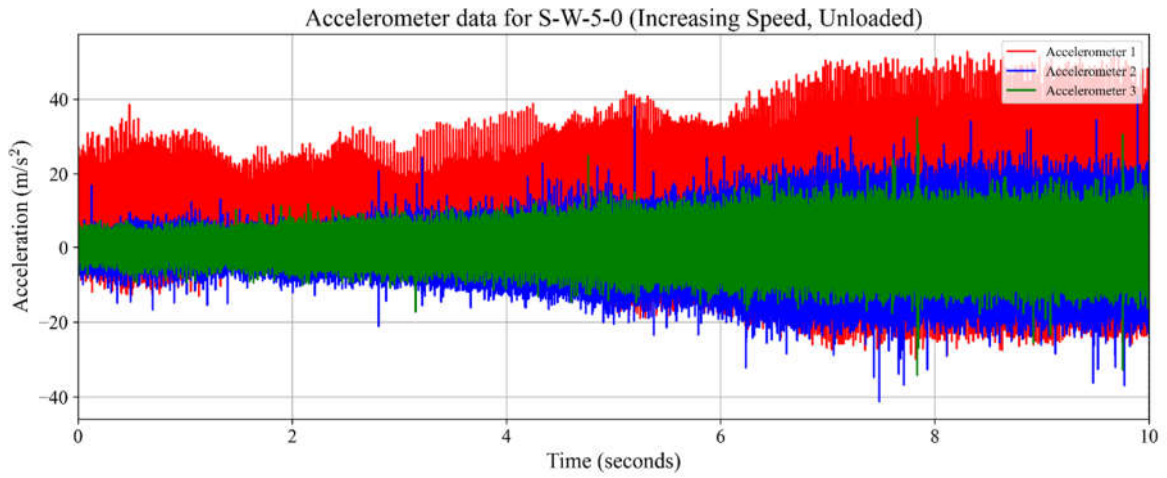


Figure 16. Stator Winding Fault 15.41 Hz $\rightarrow$ 44.96 Hz

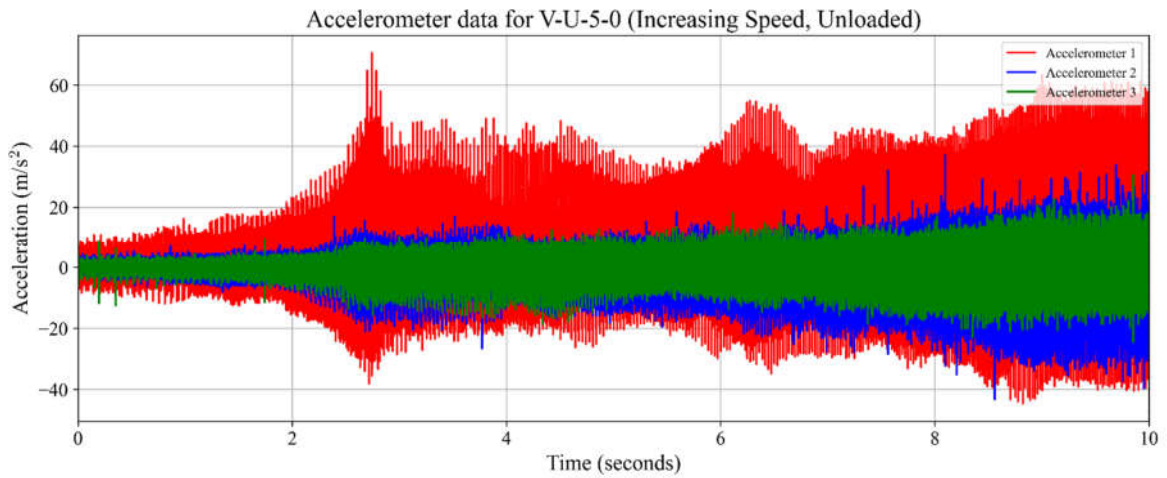


Figure 17. Voltage Unbalance and Single Phasing 15.14 Hz $\rightarrow$ 44.96 Hz

4.3. Visualizing raw data during deceleration, loaded

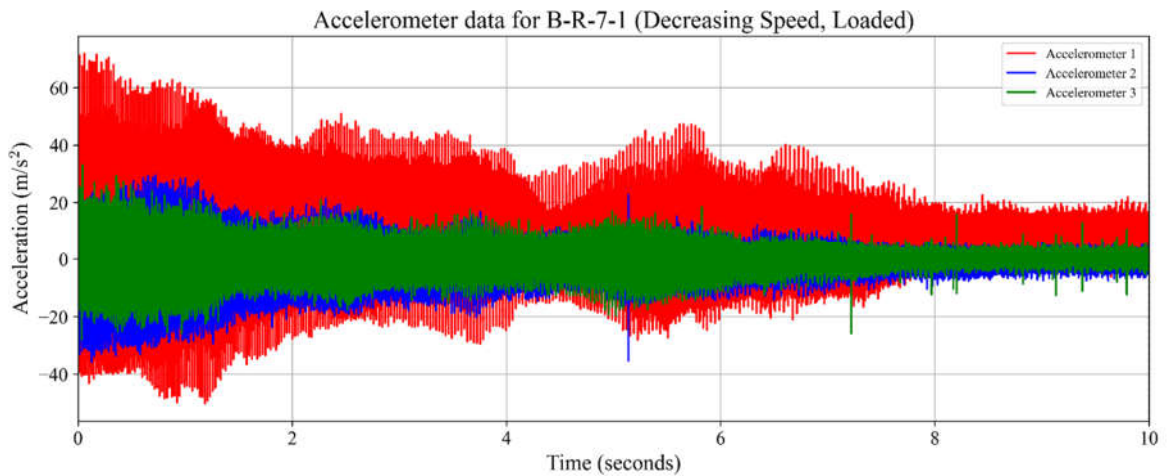


Figure 18. Bowed Rotor (44.92 Hz $\rightarrow$ 15.75 Hz)

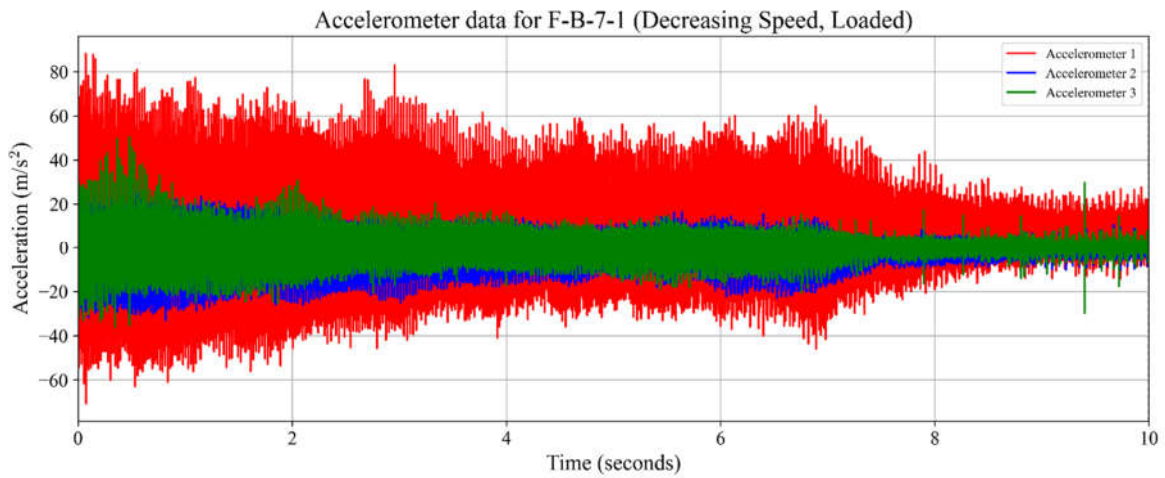


Figure 19. Faulty Bearings (Mechanical Fault). 44.90 Hz $\rightarrow$ 14.44 Hz

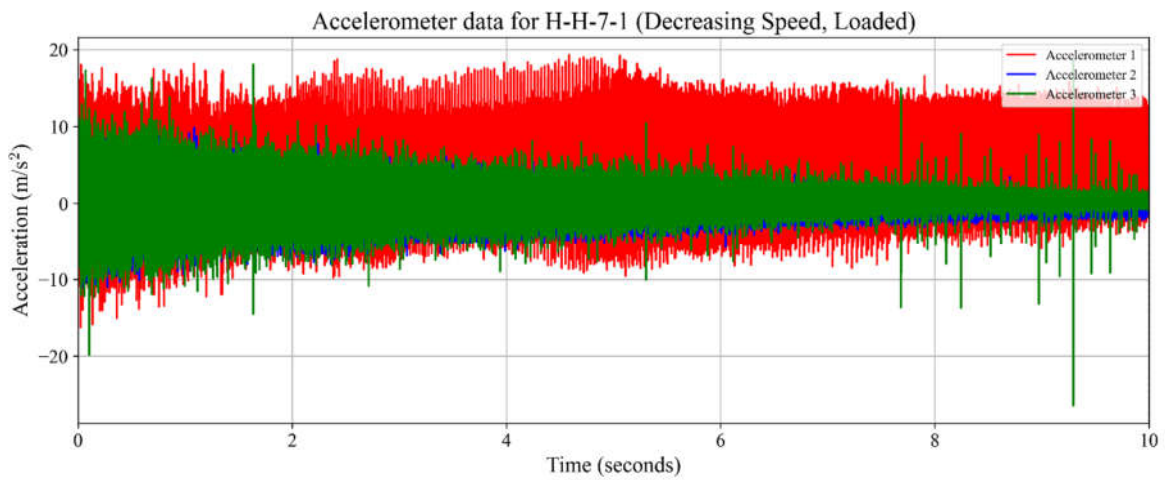


Figure 20. Healthy 44.89 Hz $\rightarrow$ 15.12 Hz

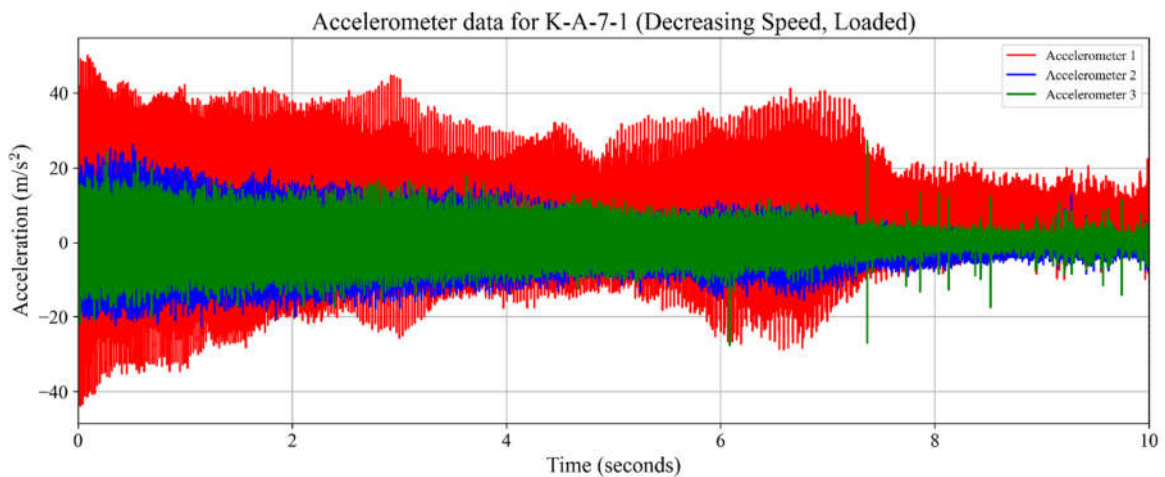


Figure 21. Broken Rotor Bars 44.96 Hz $\rightarrow$ 16.81 Hz

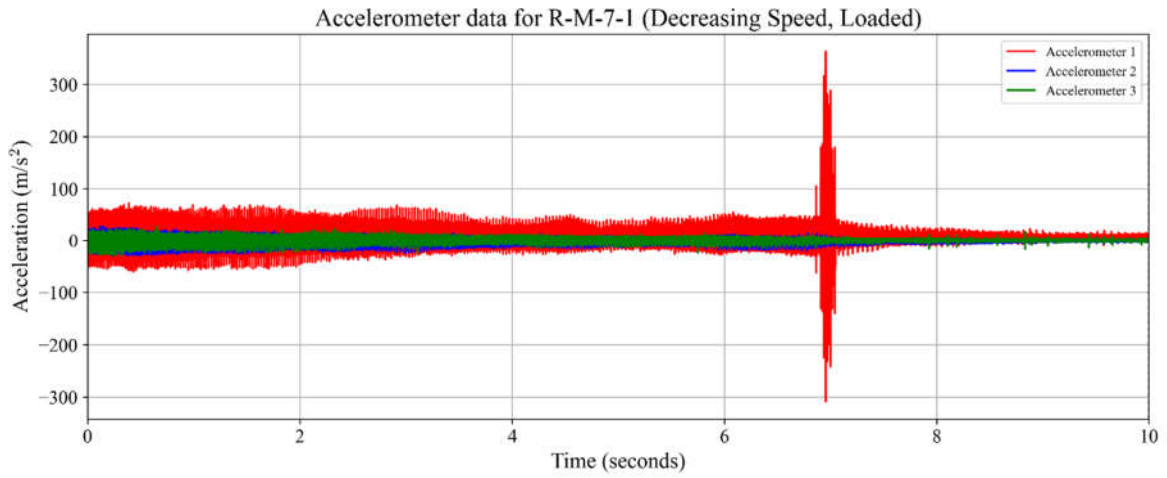


Figure 22. Rotor Misalignment. 44.97 Hz $\rightarrow$ 18.82 Hz

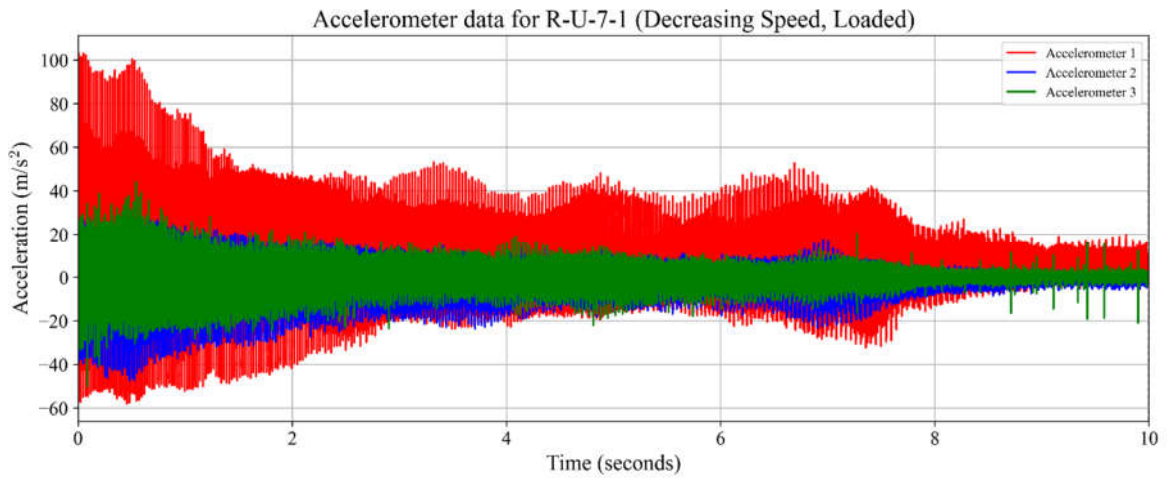


Figure 23. Rotor Unbalance Lỗi Cơ học (Mechanical Faults). 44.97 Hz $\rightarrow$ 18.82 Hz

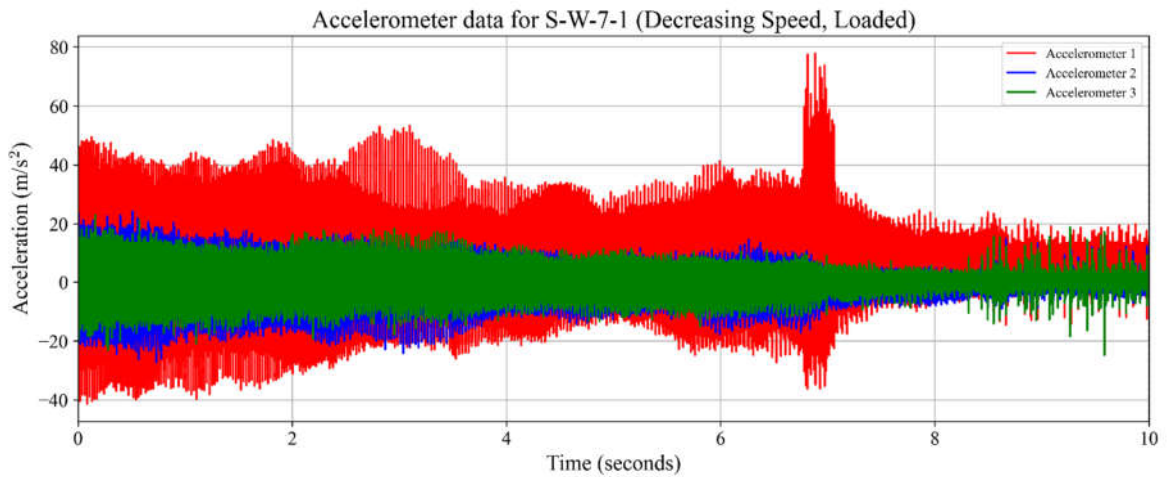


Figure 24. Stator Winding Fault. 44.92 Hz $\rightarrow$ 15.65 Hz

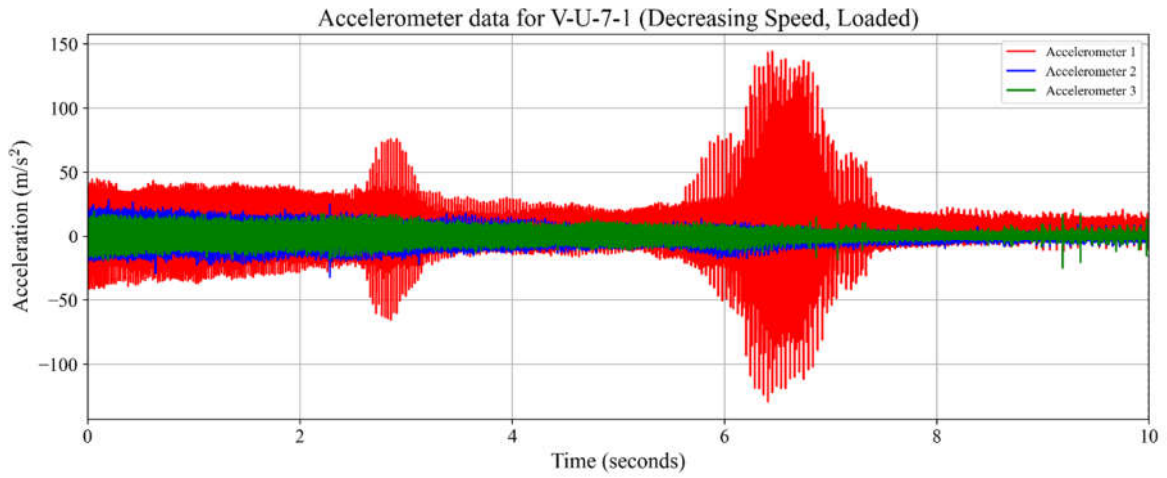


Figure 25. Electrical Fault. 44.91 Hz $\rightarrow$ 15.32 Hz

4.4. Visualizing raw data during Acceleration, loaded

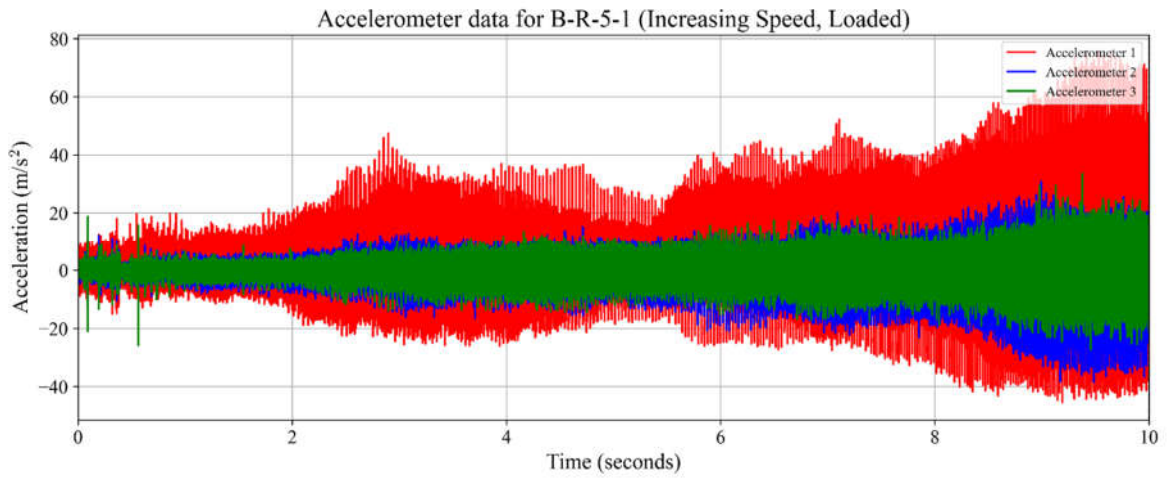


Figure 26. Bowed Rotor 15.81 Hz $\rightarrow$ 44.81 Hz

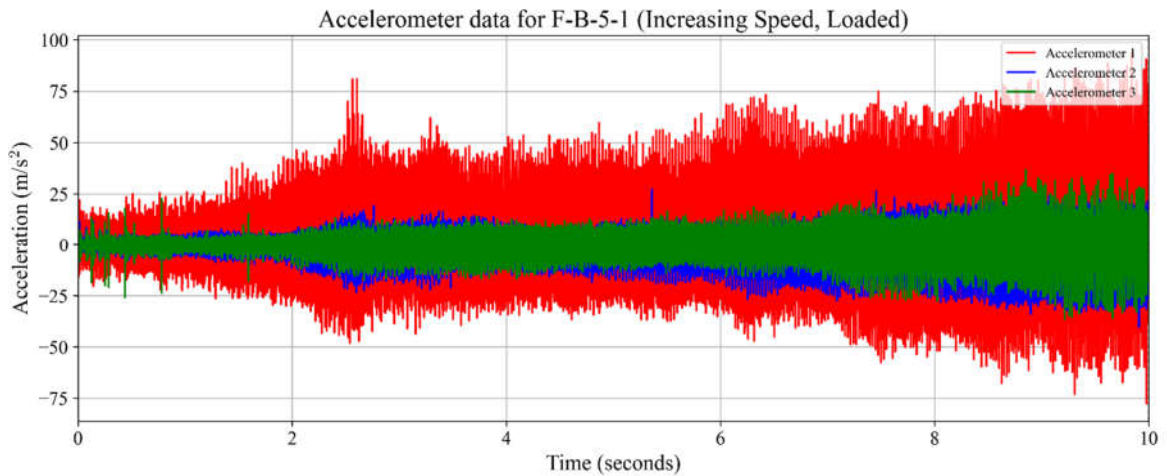


Figure 27. Faulty Bearings 15.22 Hz $\rightarrow$ 44.96 Hz

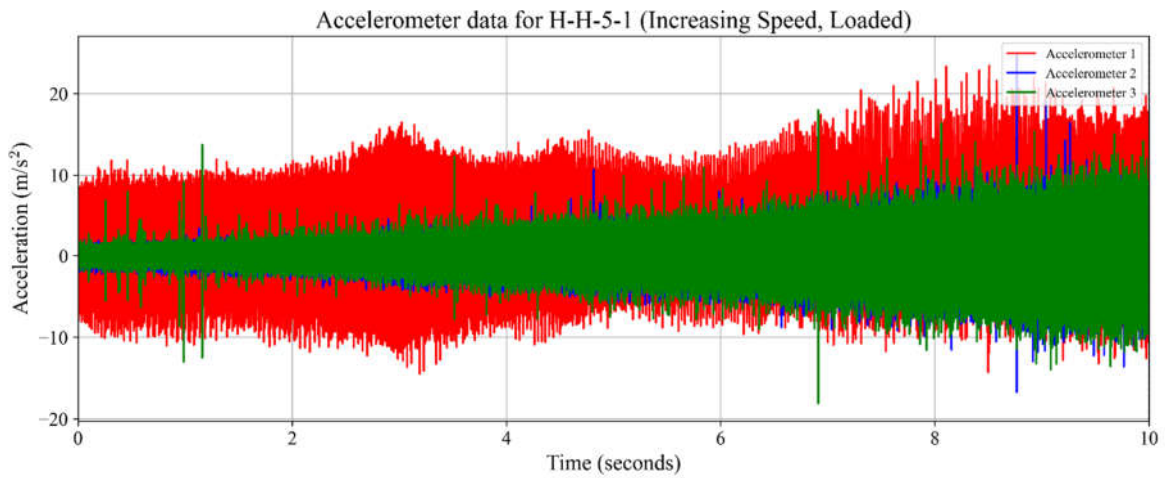


Figure 28. Healthy. 14.99 Hz $\rightarrow$ 44.86 Hz

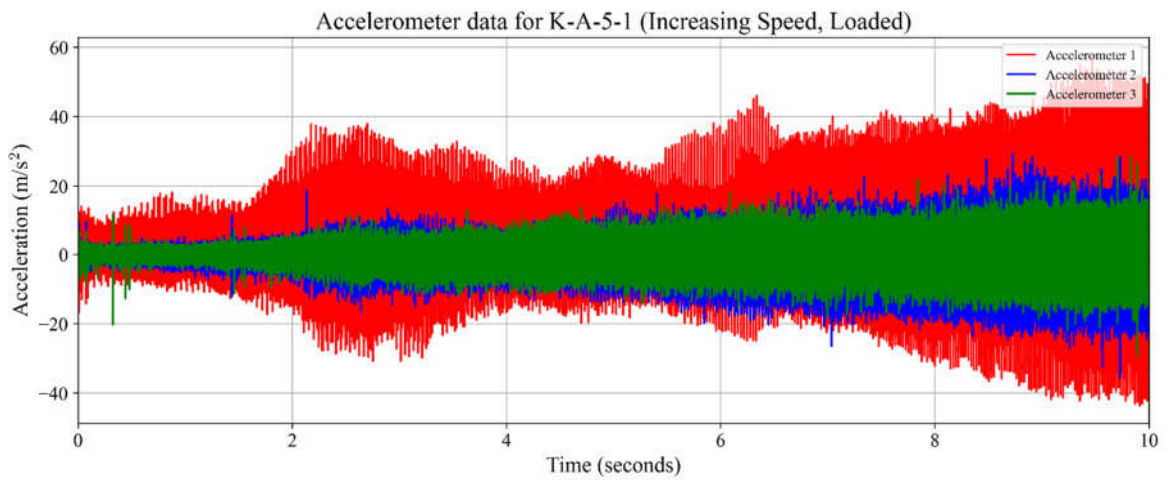


Figure 29. Broken Rotor Bars 15.13 Hz $\rightarrow$ 44.87 Hz

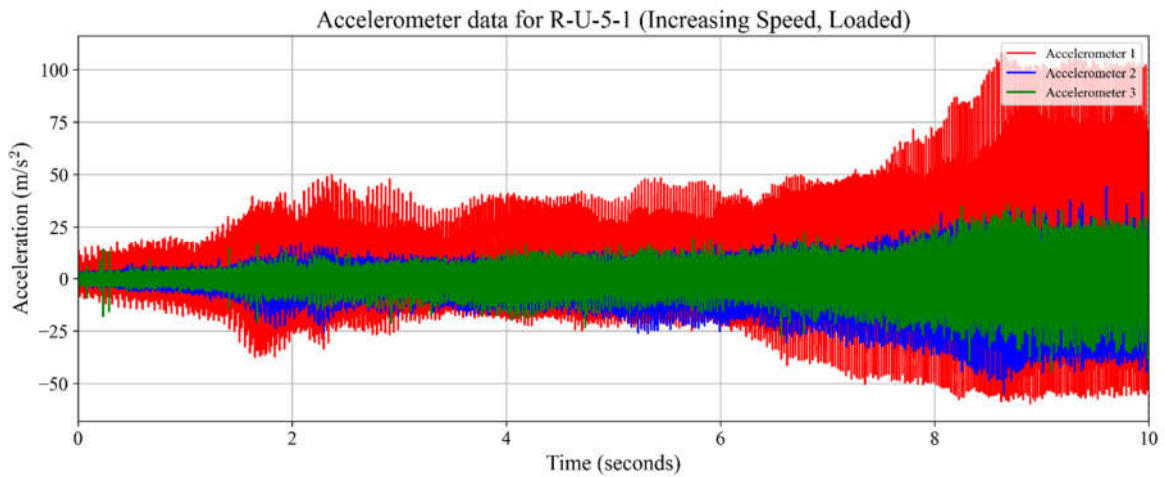


Figure 30. Rotor Unbalance 15.72 Hz $\rightarrow$ 44.97 Hz

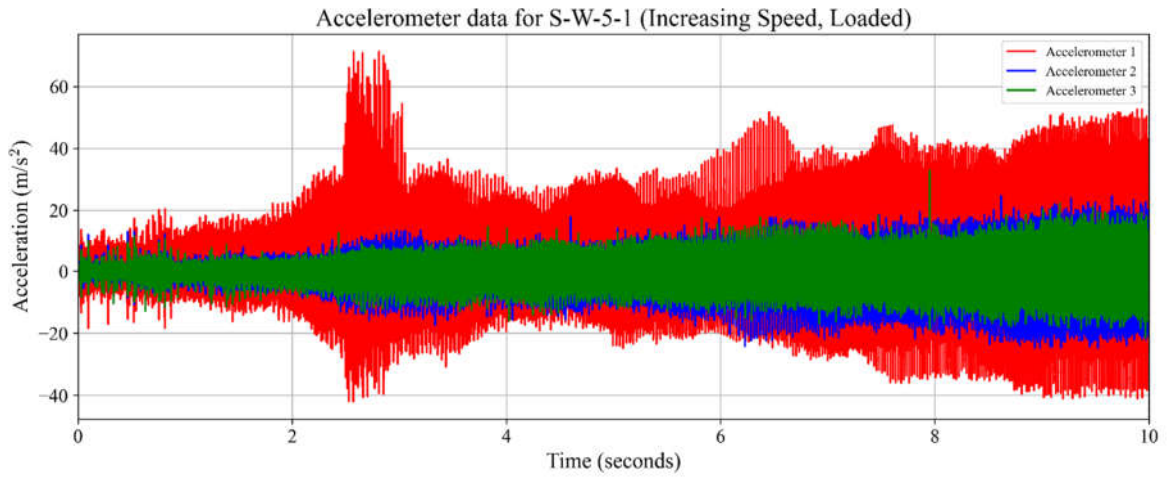


Figure 31. Stator Winding Fault 15.41 Hz $\rightarrow$ 44.96 Hz

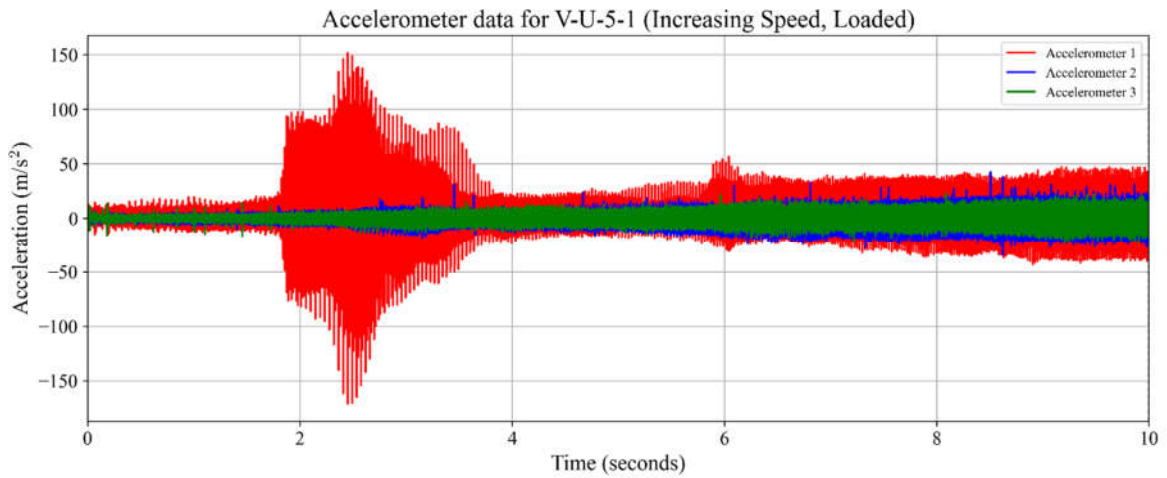


Figure 32. Voltage Unbalance and Single Phasing 15.14 Hz $\rightarrow$ 44.96 Hz

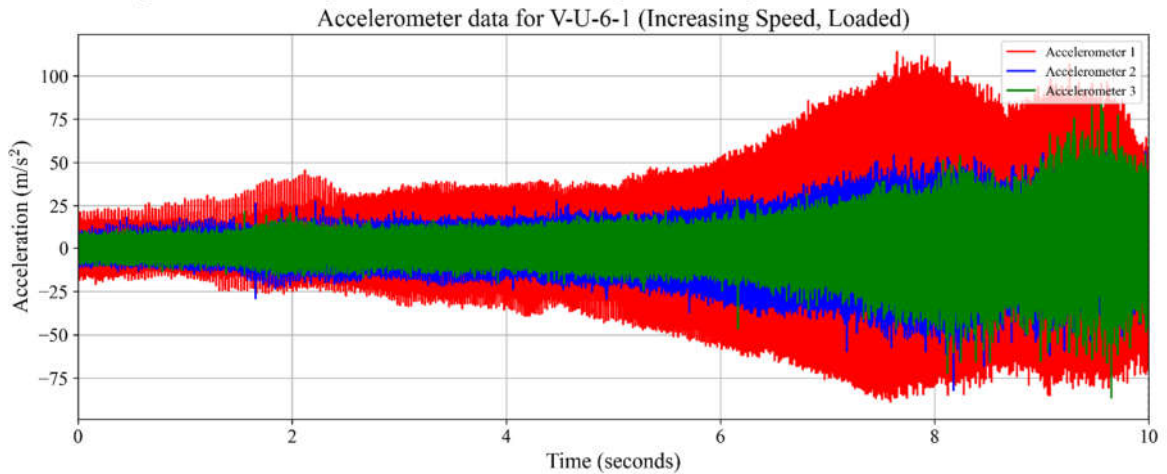


Figure 33. Voltage Unbalance 30.08 Hz $\rightarrow$ 59.88 Hz

4.5. Data visualization with FFT transformations

4.5.1. Deceleration, unloaded

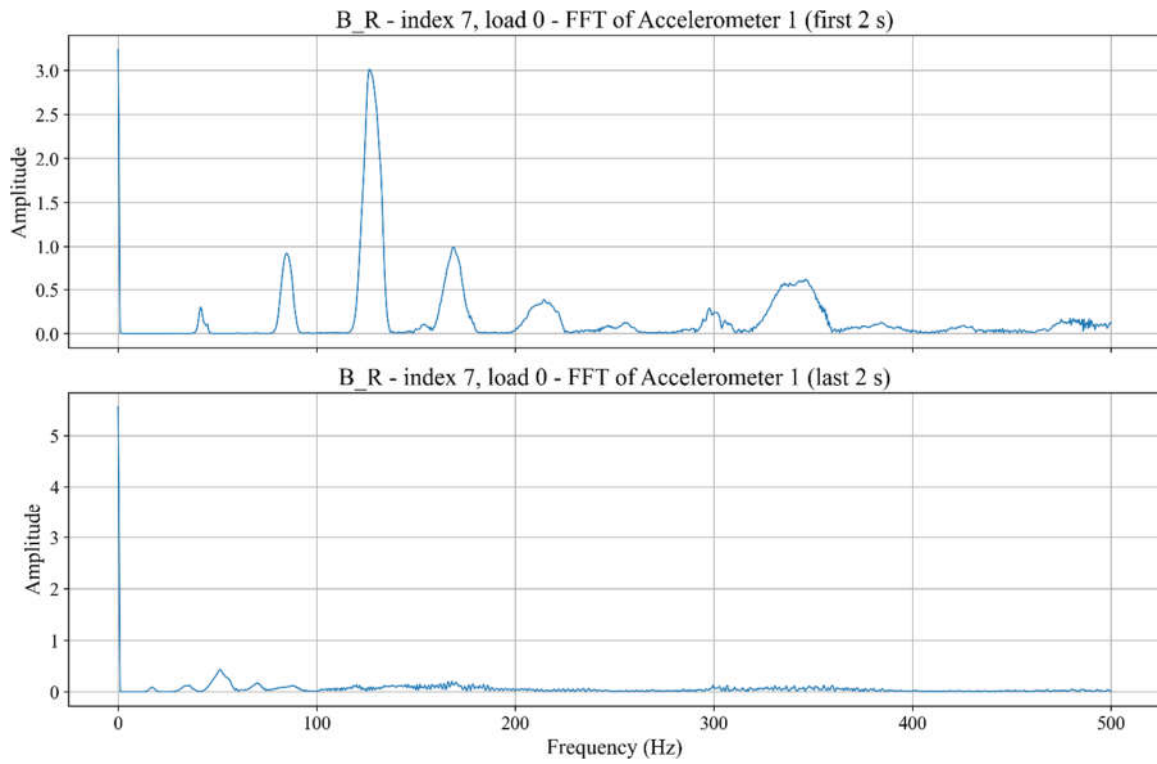


Figure 34. Bowed Rotor 44.92 Hz $\rightarrow$ 15.75 Hz

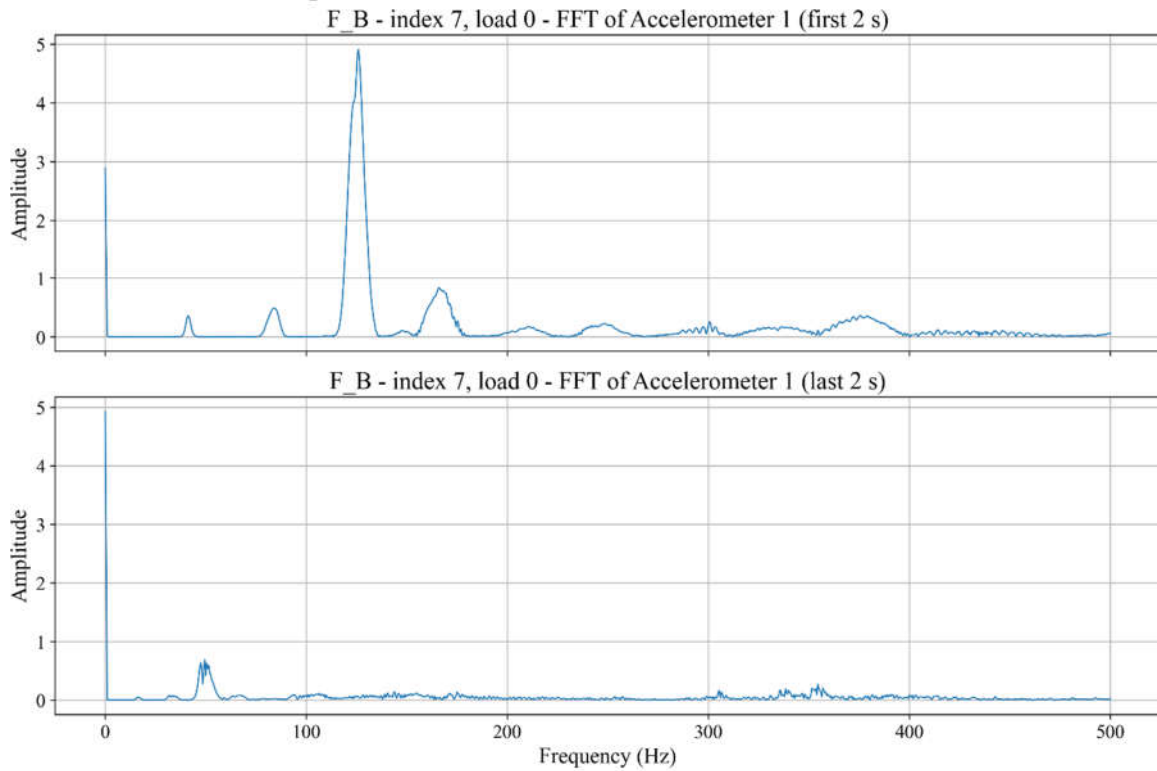


Figure 35. Faulty Bearings (Mechanical Fault). 44.90 Hz $\rightarrow$ 14.44 Hz

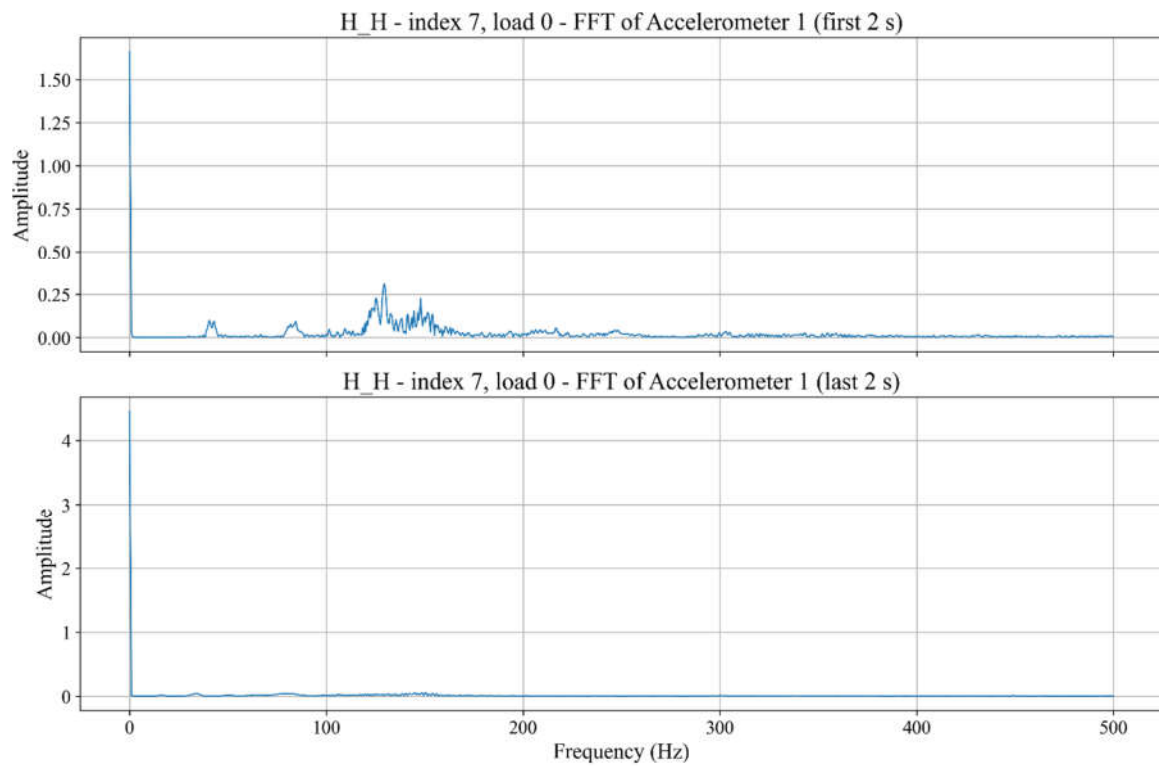


Figure 36. Healthy 44.89 Hz $\rightarrow$ 15.12 Hz

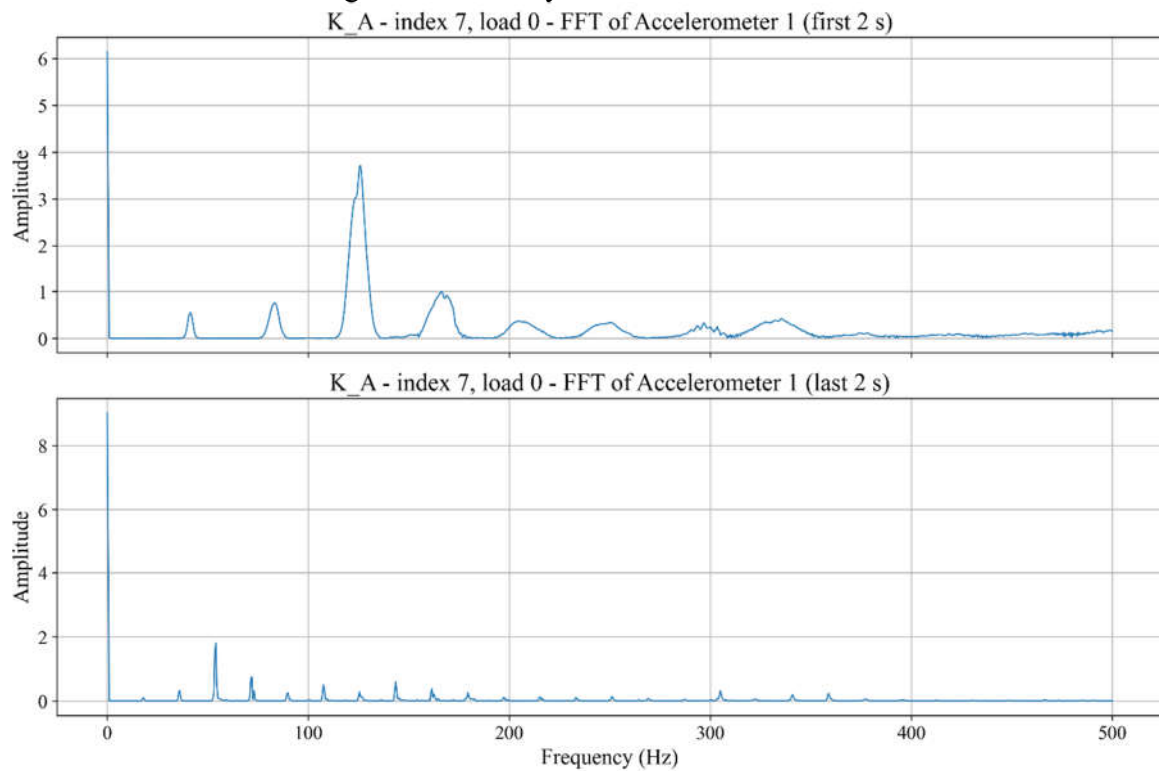


Figure 37. Broken Rotor Bars 45 Hz $\rightarrow$ 15 Hz

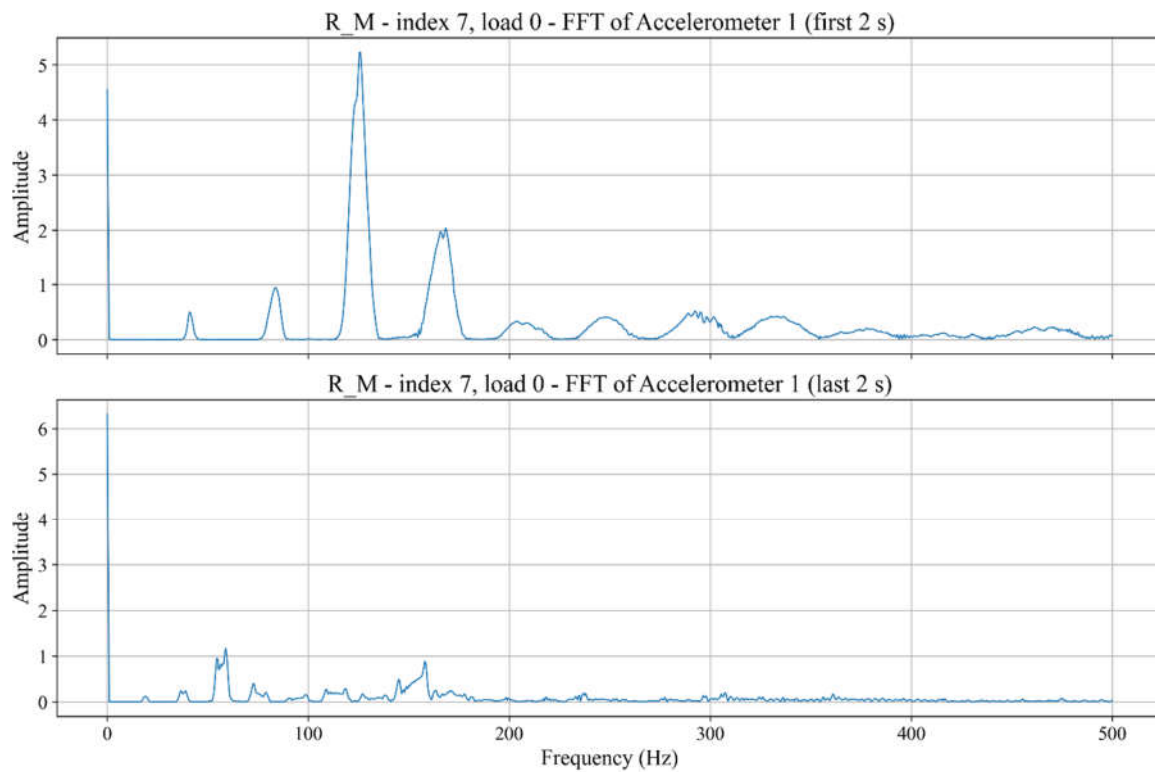


Figure 38. Rotor Misalignment 44.97 Hz $\rightarrow$ 18.82 Hz

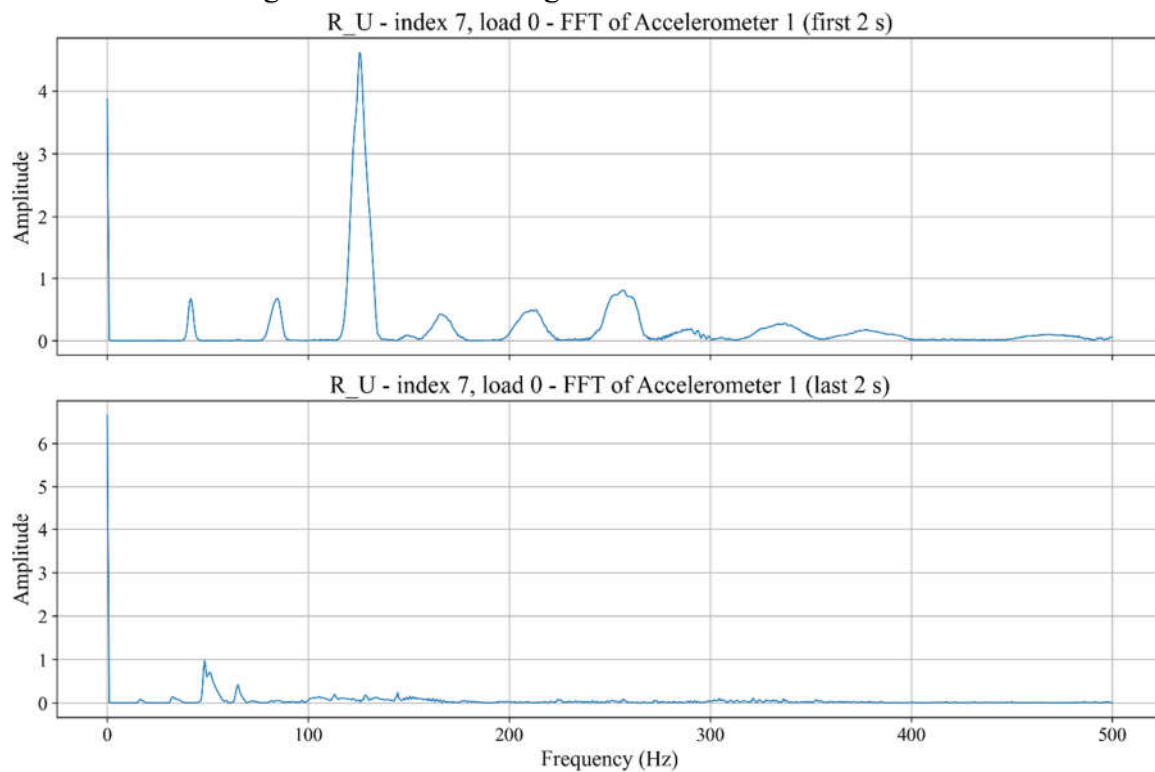


Figure 39. Mechanical Faults. 44.97 Hz $\rightarrow$ 18.82 Hz

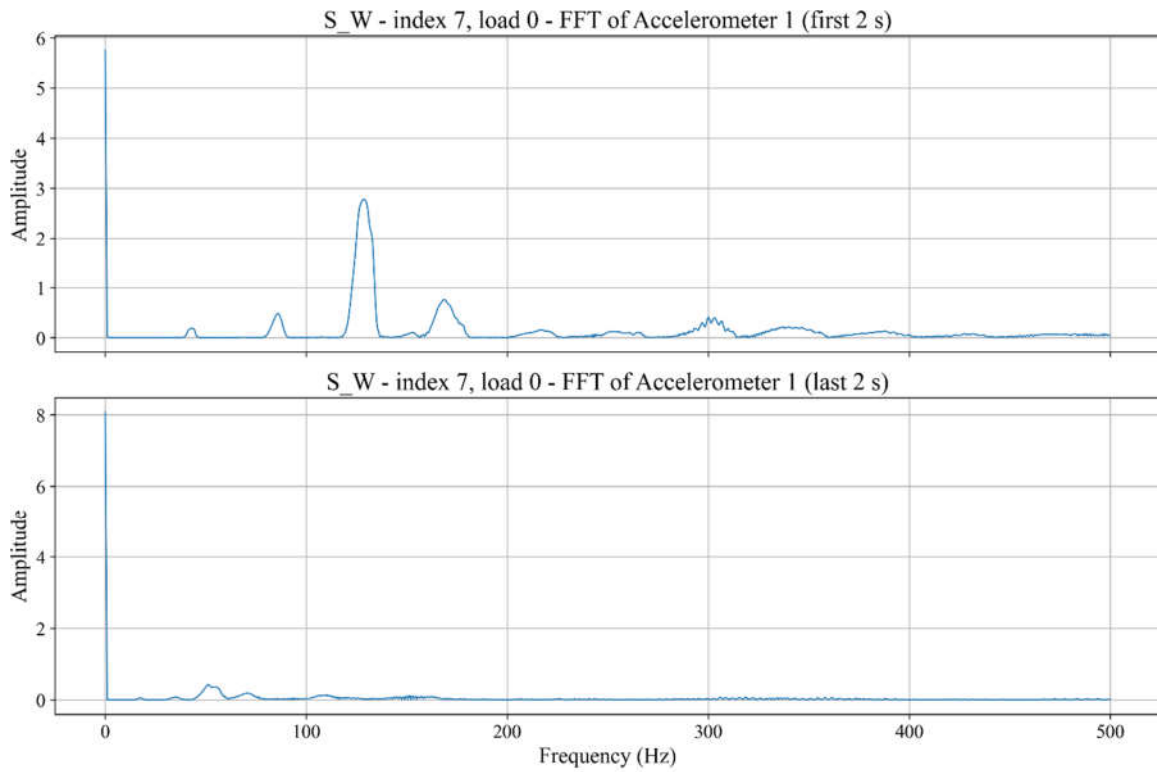


Figure 40. Stator Winding Fault 44.92 Hz → 15.65 Hz

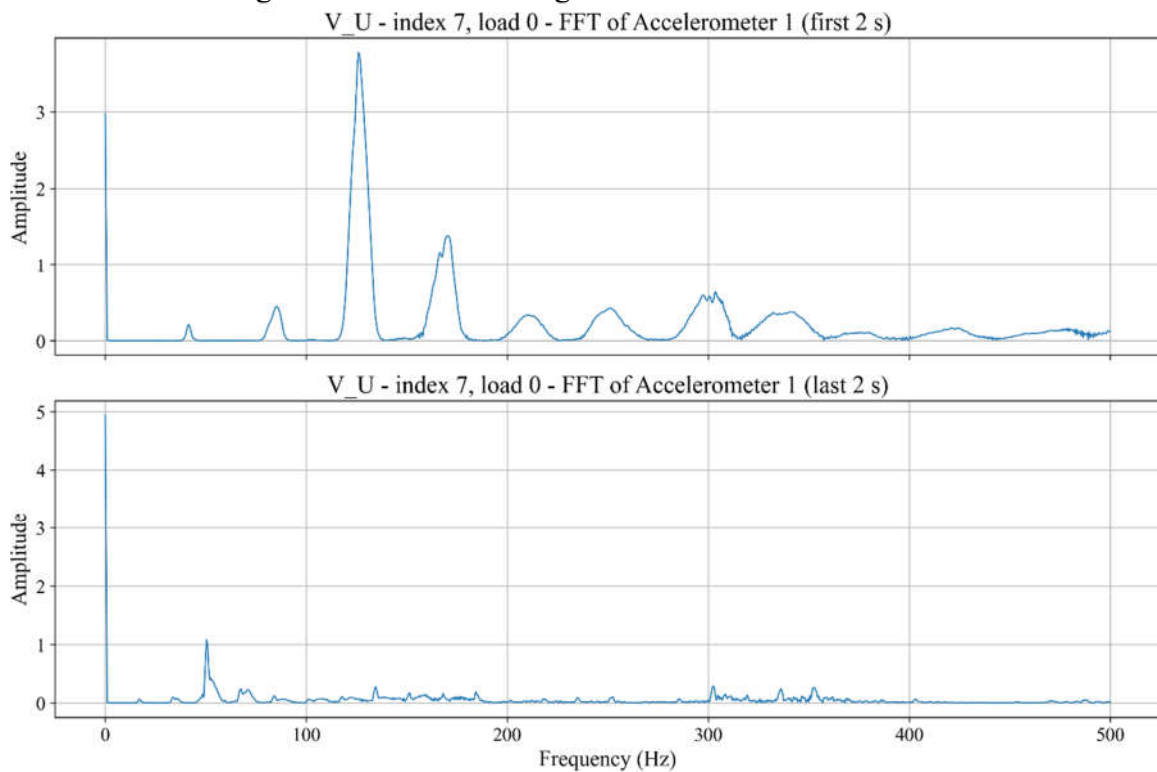


Figure 41. Electrical Fault 44.91 Hz → 15.32 Hz

4.5.2. Acceleration, unloaded

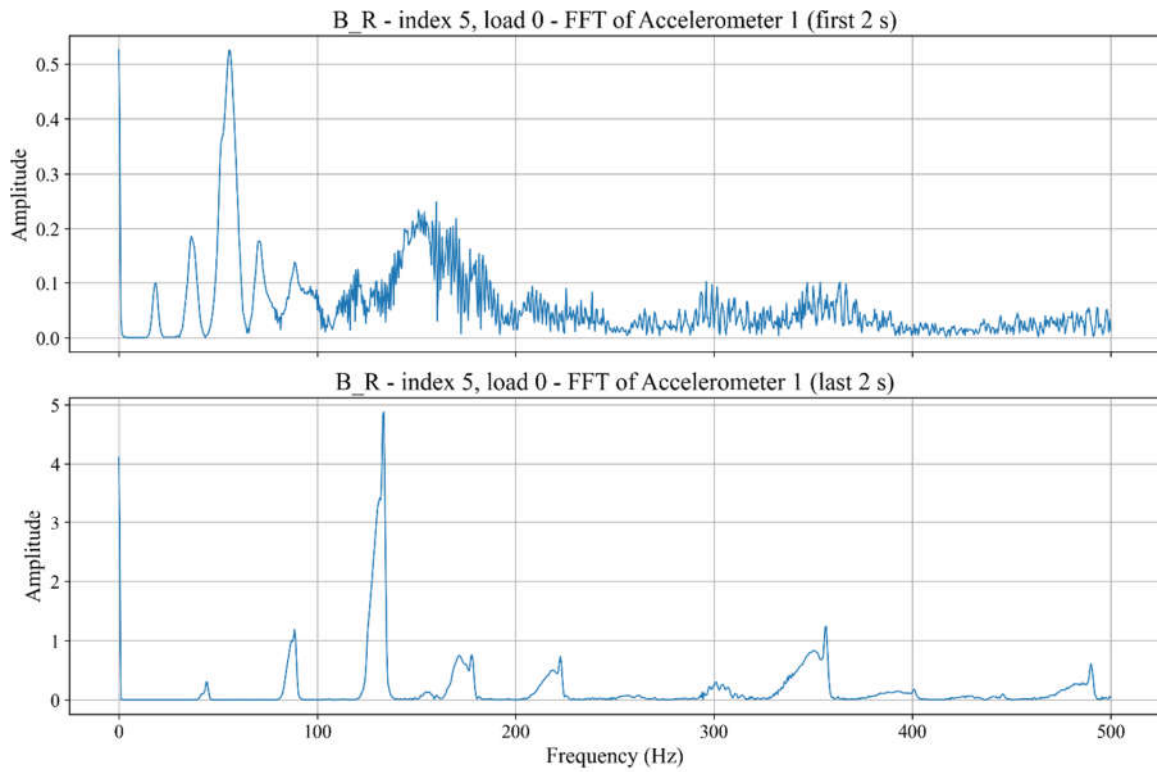


Figure 42. Bowed Rotor 15.81 Hz → 44.81 Hz

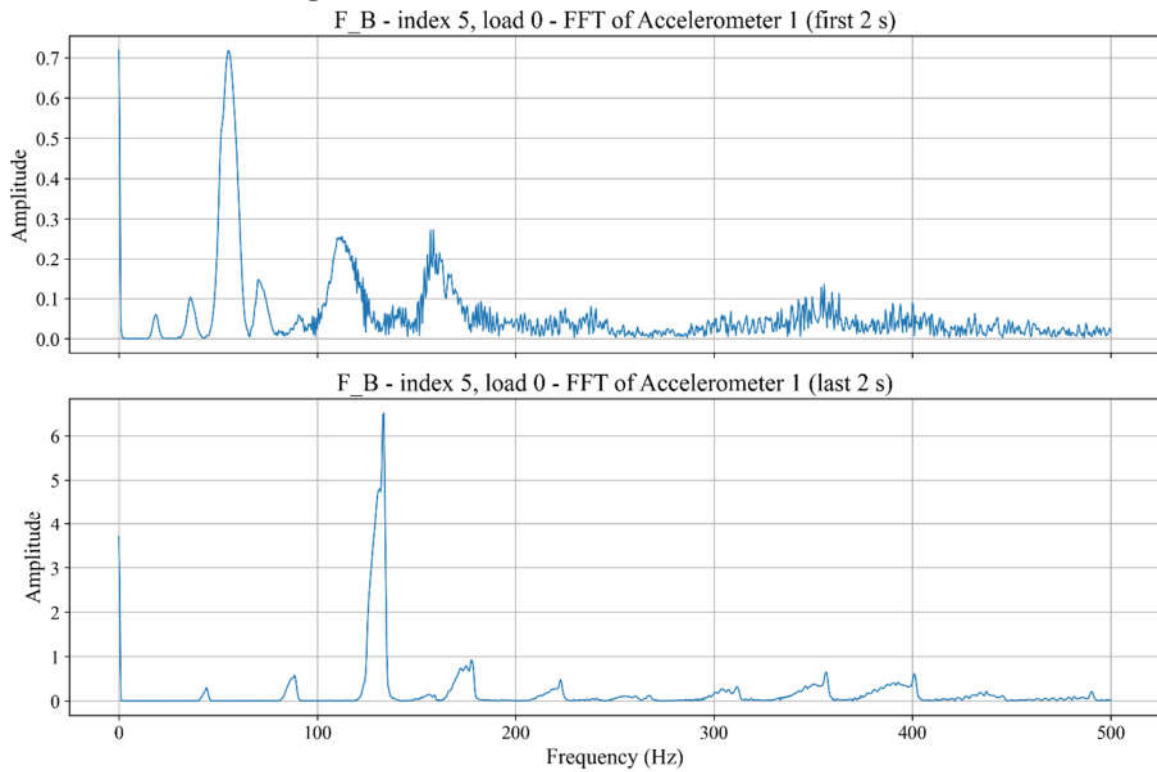


Figure 43. Faulty Bearings 15.22 Hz → 44.96 Hz

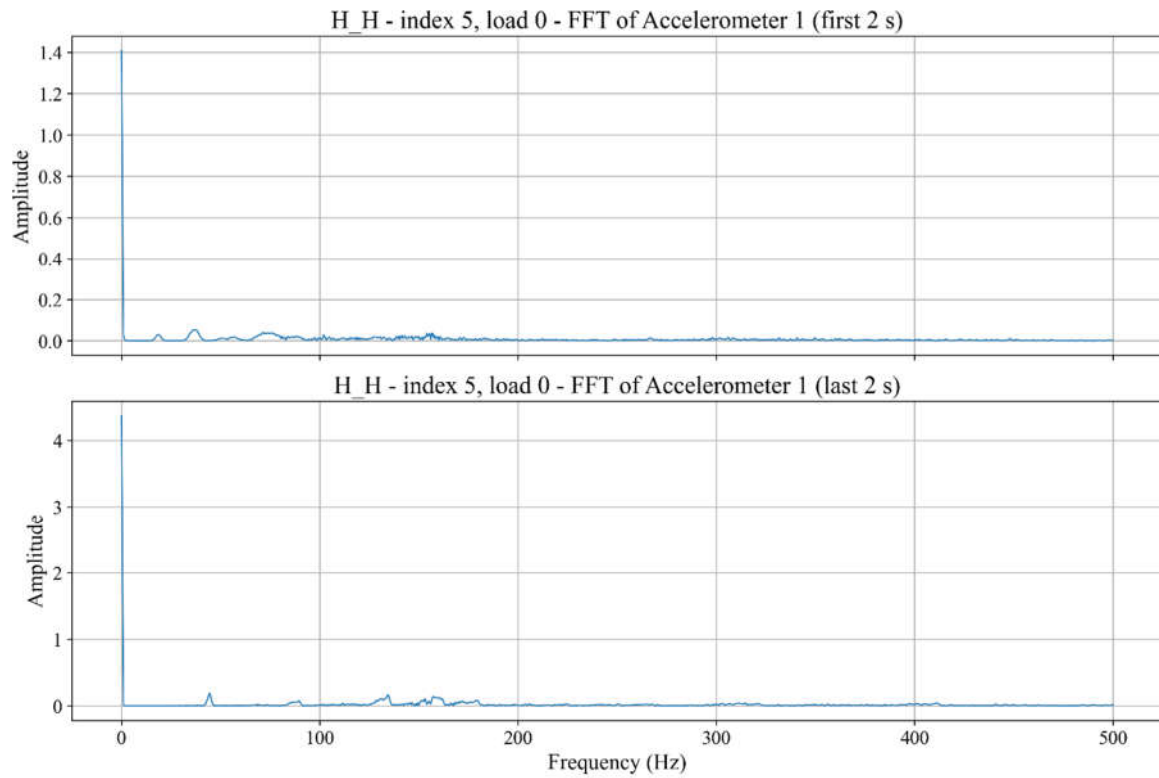


Figure 44. Healthy 14.99 Hz $\rightarrow$ 44.86 Hz

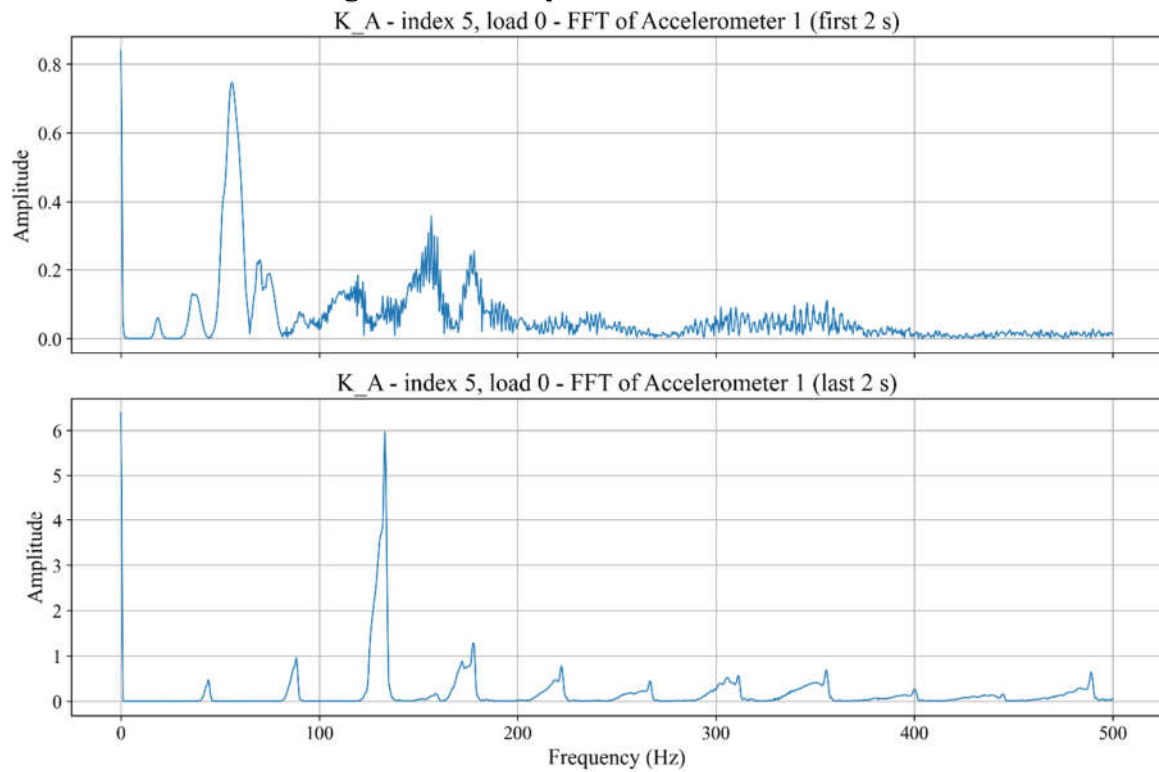


Figure 45. Broken Rotor Bars 15.13 Hz $\rightarrow$ 44.87 Hz

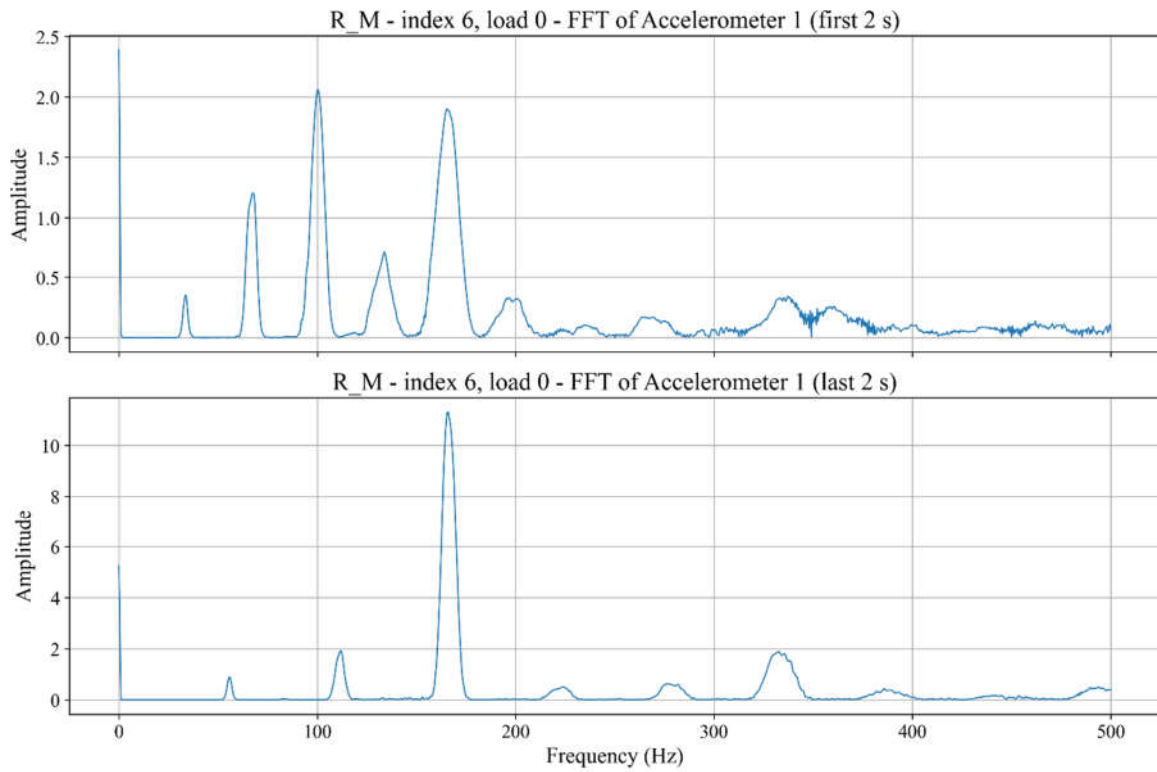


Figure 46. Rotor Misalignment 30.33 Hz → 59.96 Hz

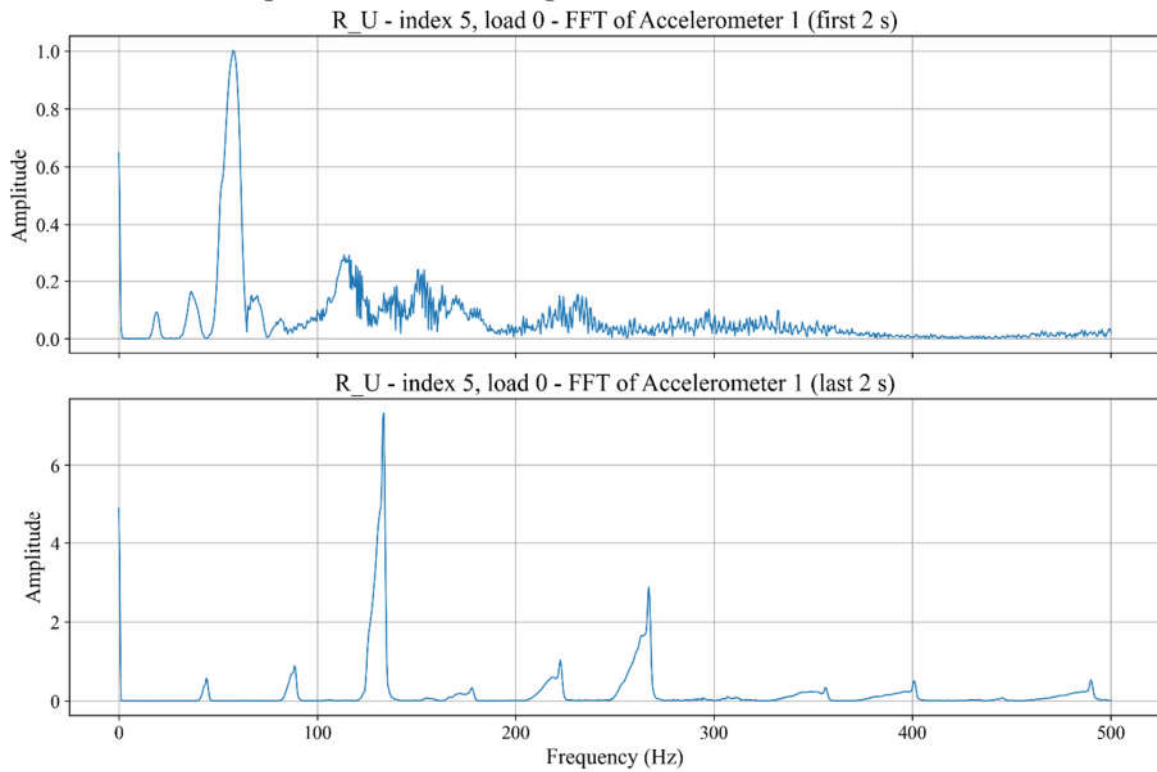


Figure 47. Rotor Unbalance 15.72 Hz → 44.97 Hz

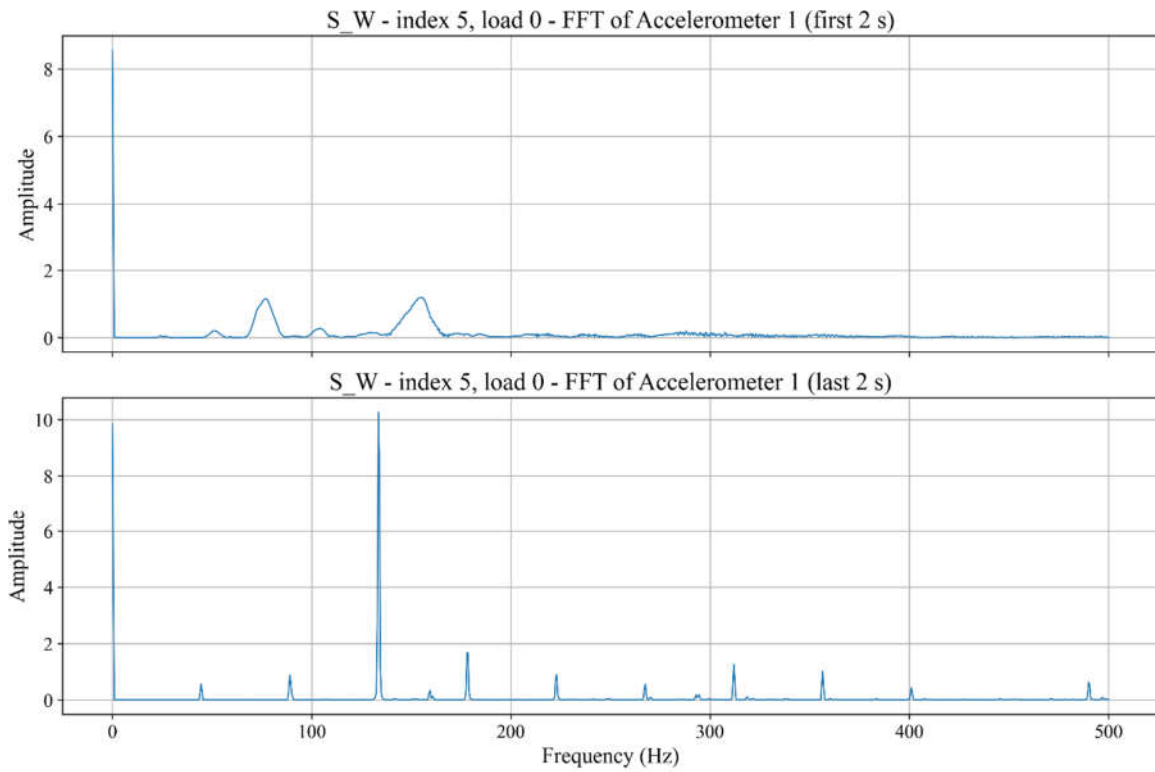


Figure 48. Stator Winding Fault 15.33 Hz → 44.92 Hz

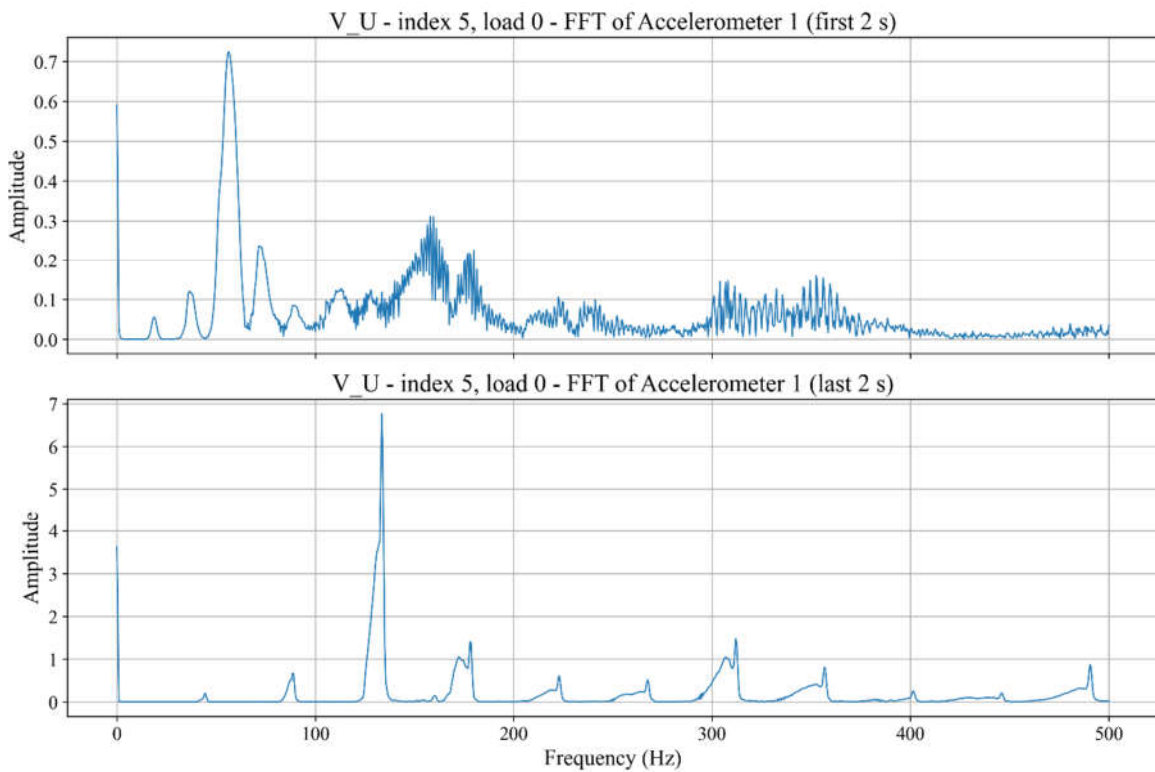


Figure 49. Voltage imbalance. 15.14 Hz → 44.96 Hz

4.5.3. Deceleration, loaded

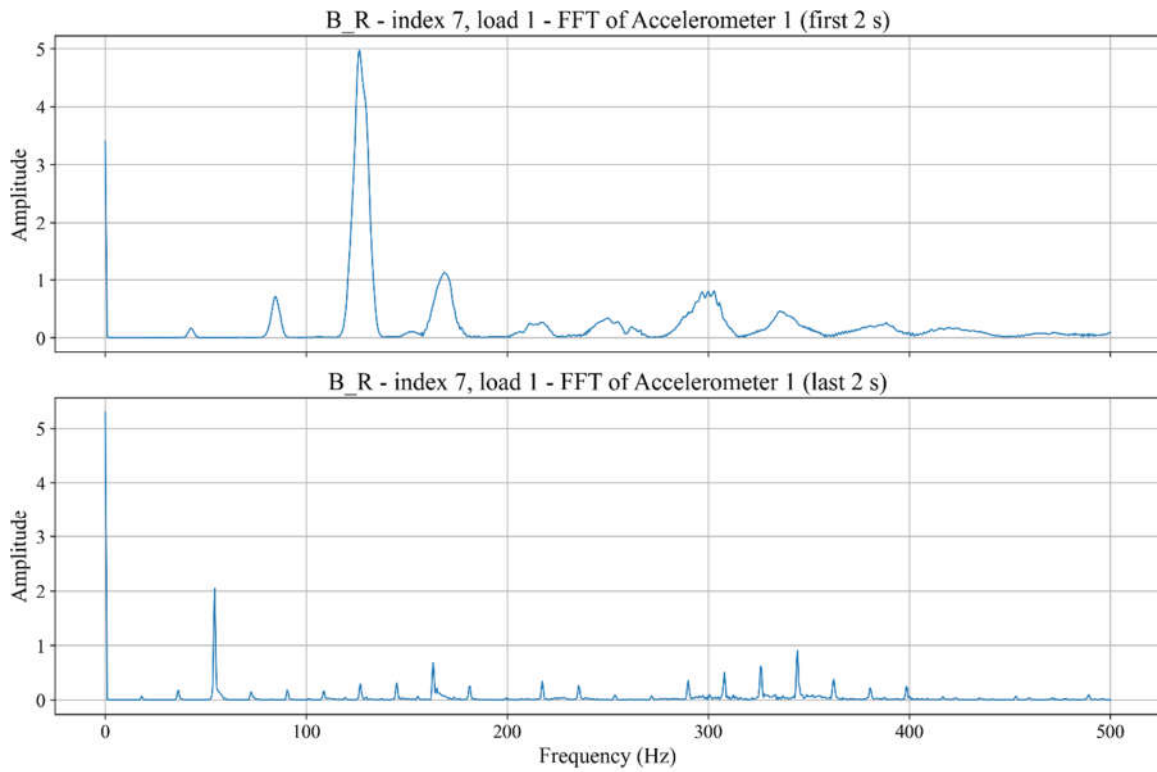


Figure 50. Bowed Rotor 44.92 Hz $\rightarrow$ 15.75 Hz

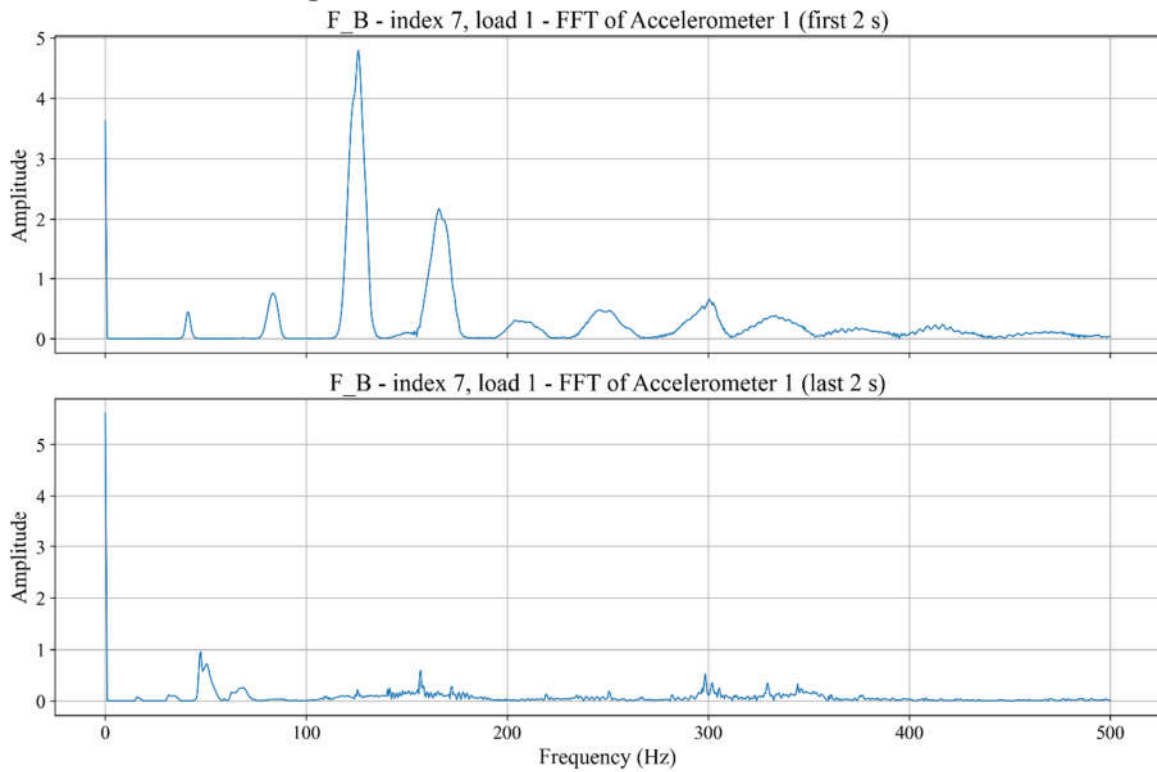


Figure 51. Faulty Bearings (Mechanical Fault). 44.90 Hz $\rightarrow$ 14.44 Hz

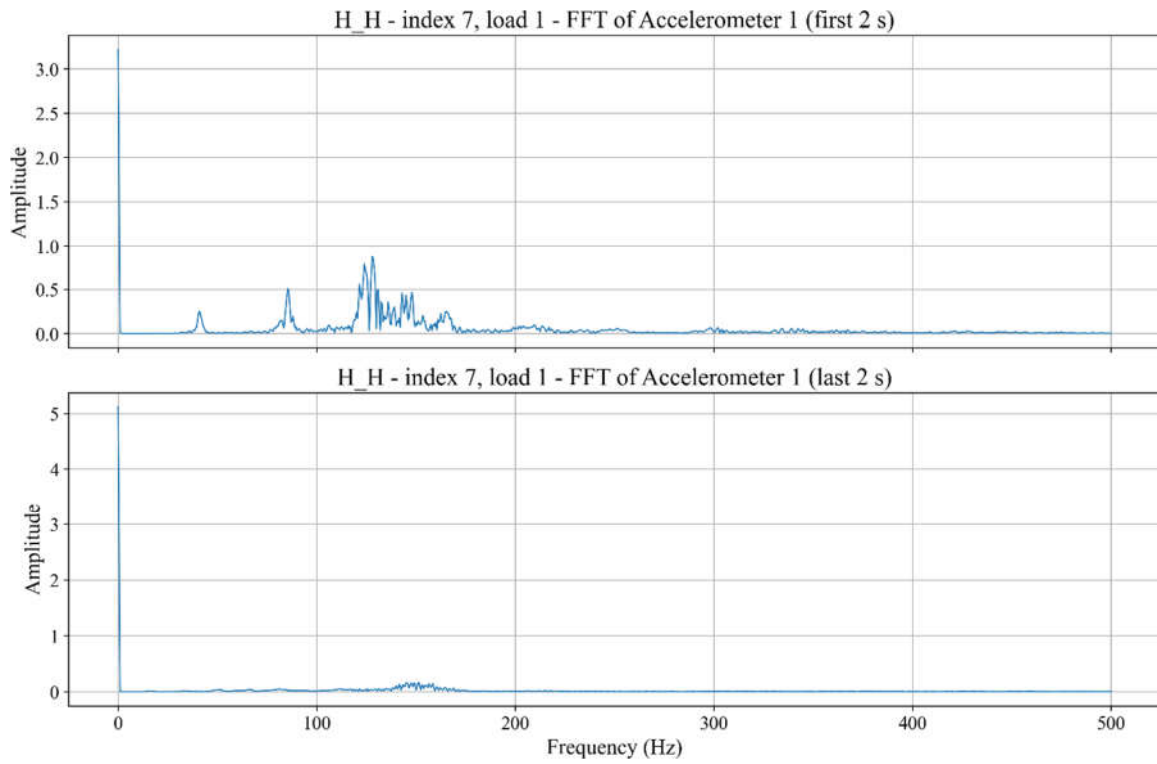


Figure 52. Healthy 44.89 Hz $\rightarrow$ 15.12 Hz

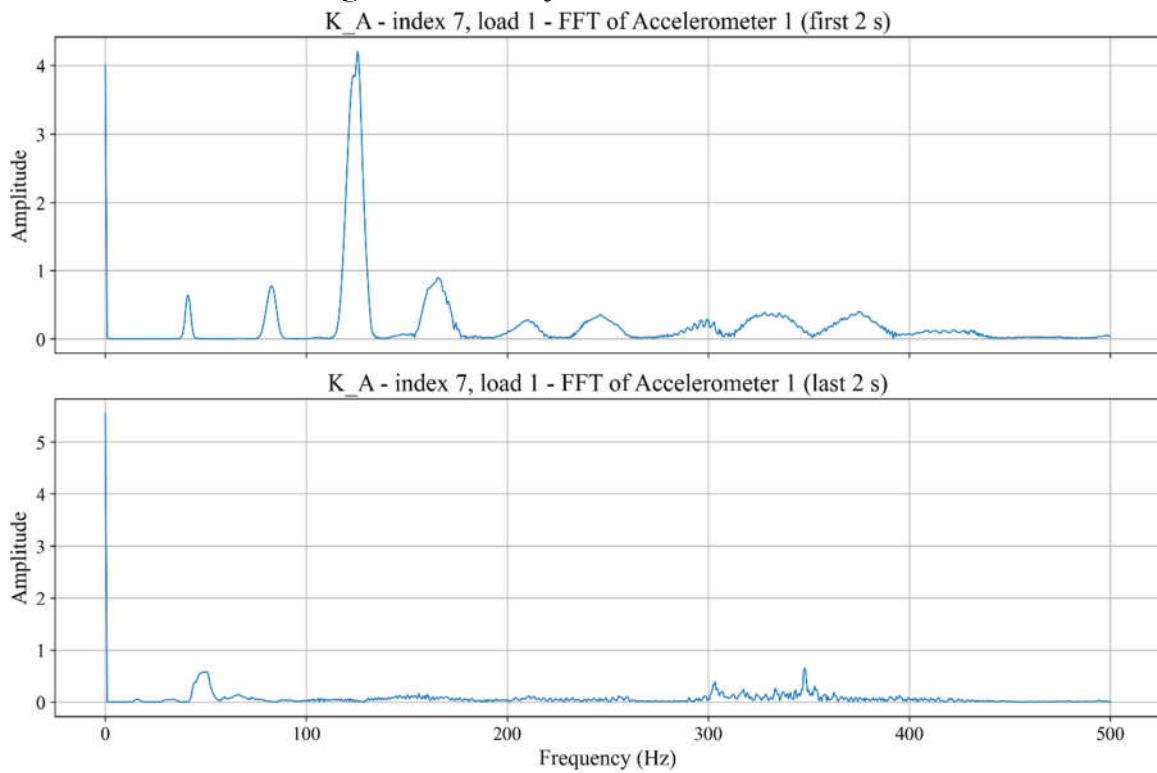


Figure 53. Broken Rotor Bars 44.96 Hz $\rightarrow$ 16.81 Hz

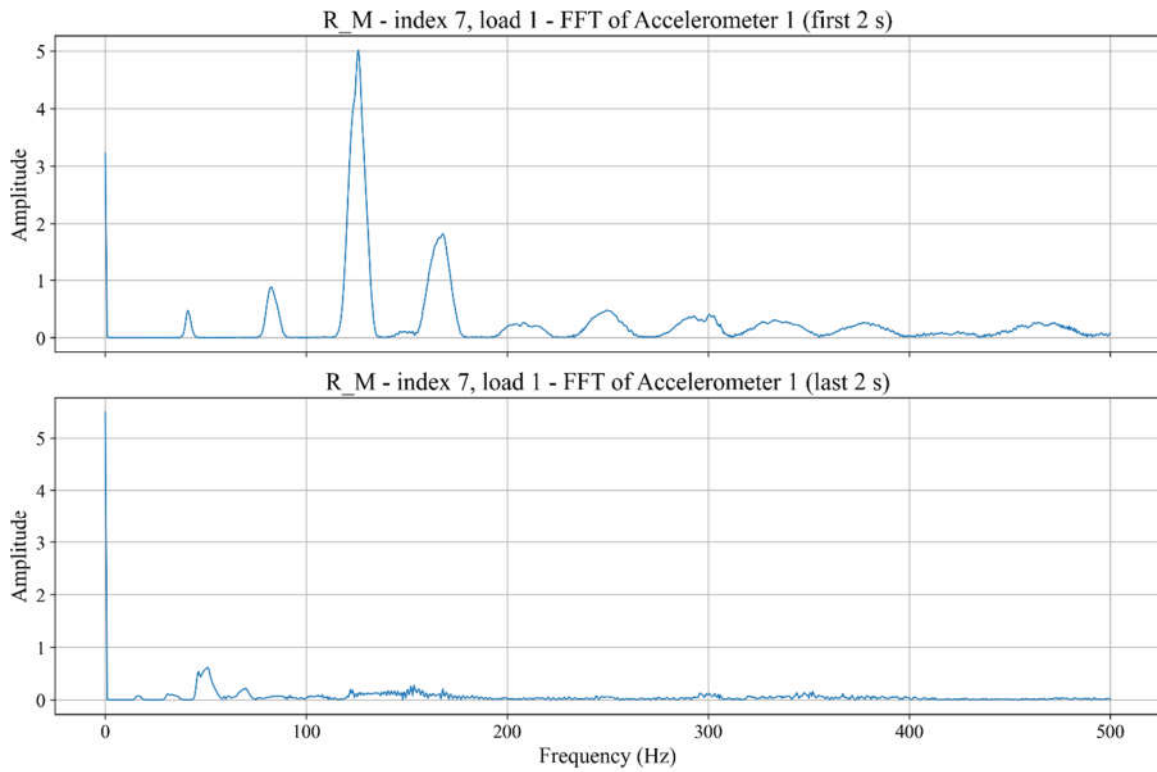


Figure 54. Rotor Misalignment 44.97 Hz - 18.82 Hz

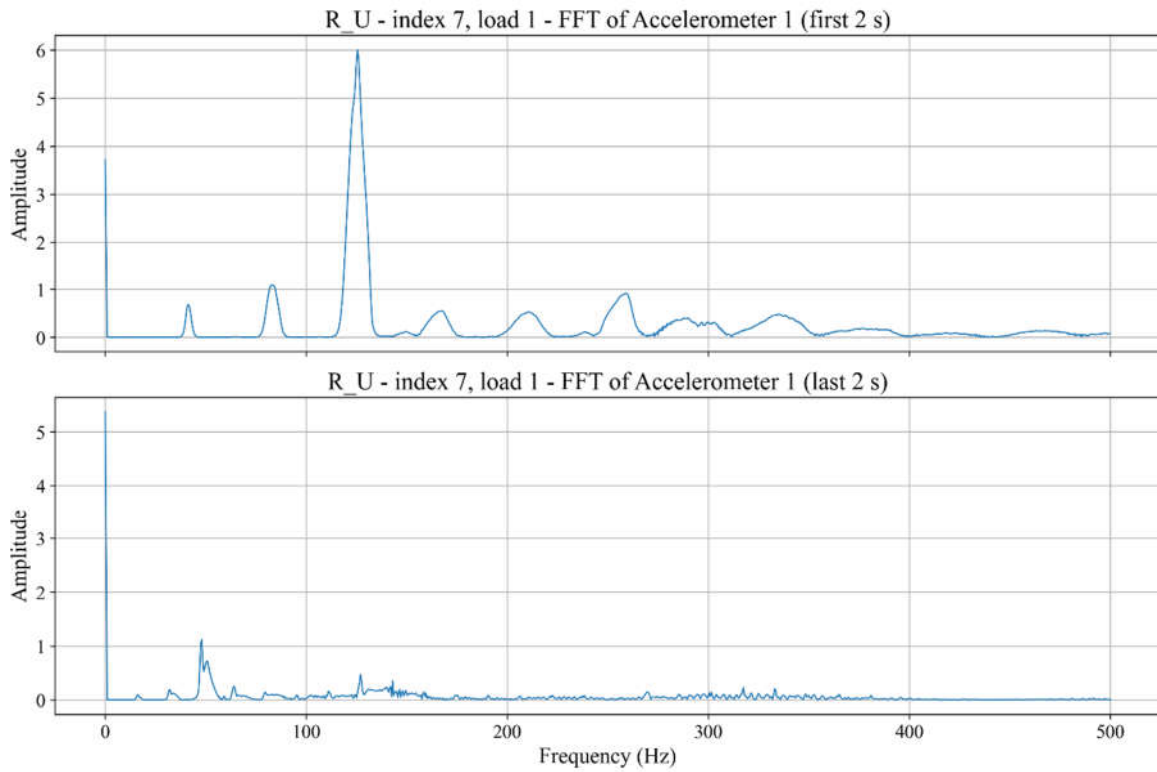


Figure 55. Mechanical Faults 44.97 Hz → 18.82 Hz

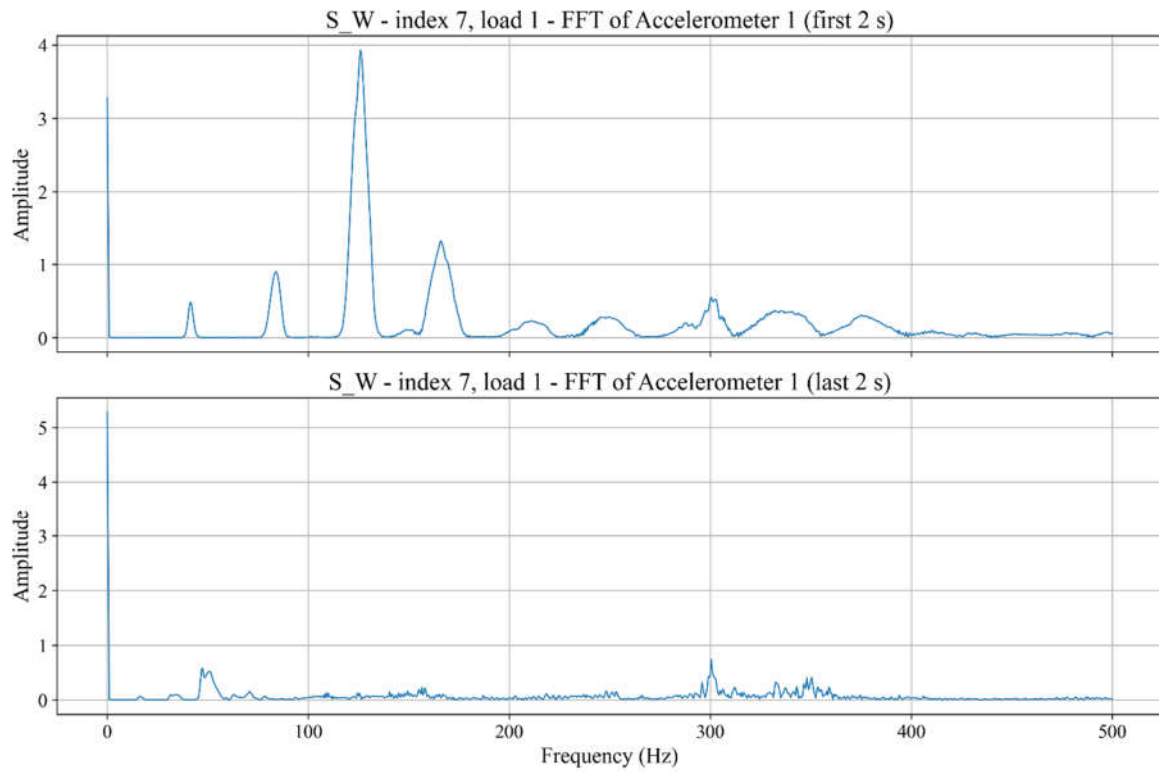


Figure 56. Stator Winding Fault 44.92 Hz $\rightarrow$ 15.65 Hz

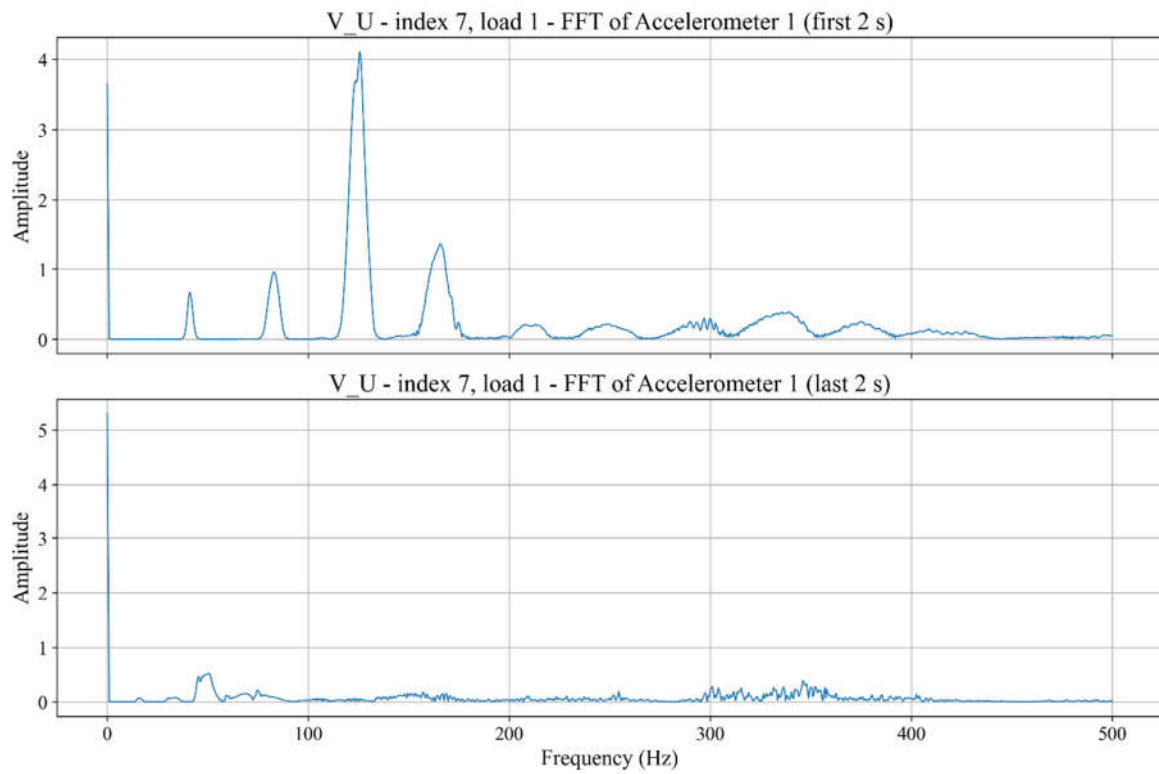


Figure 57. Electrical Fault 44.91 Hz $\rightarrow$ 15.32 Hz

4.5.4. Acceleration, loaded

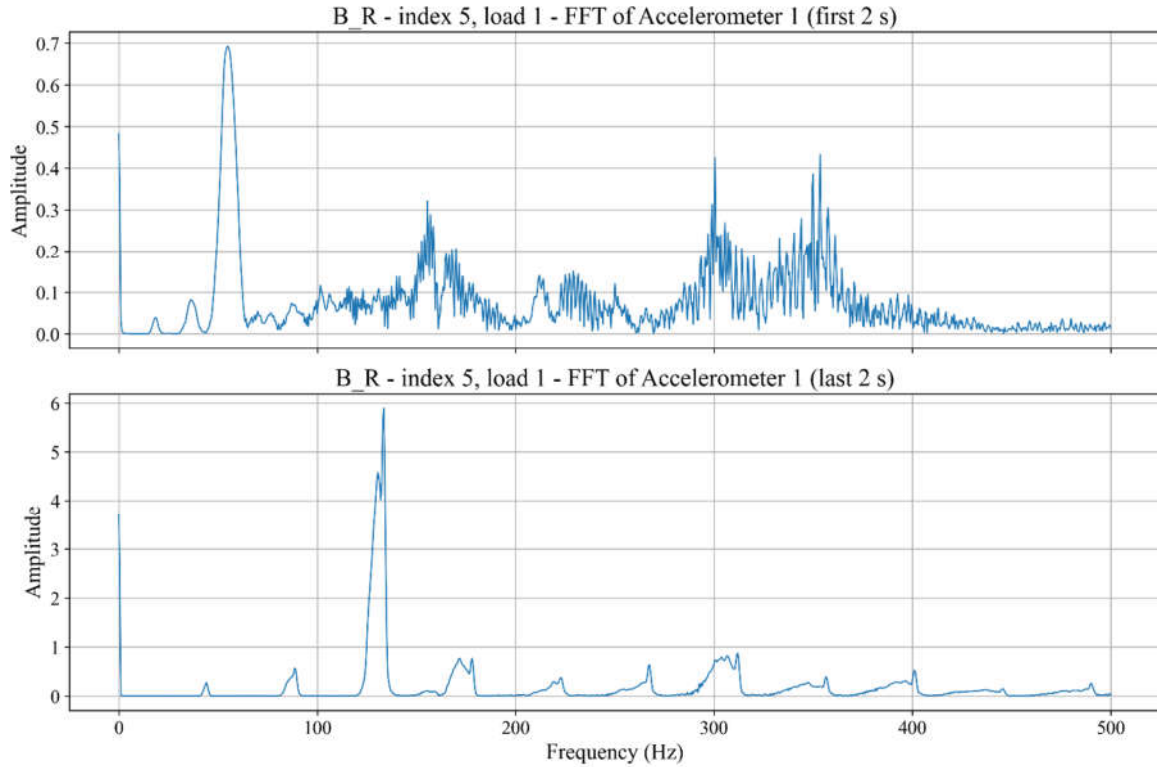


Figure 58. Bowed Rotor 15.81 Hz → 44.81 Hz

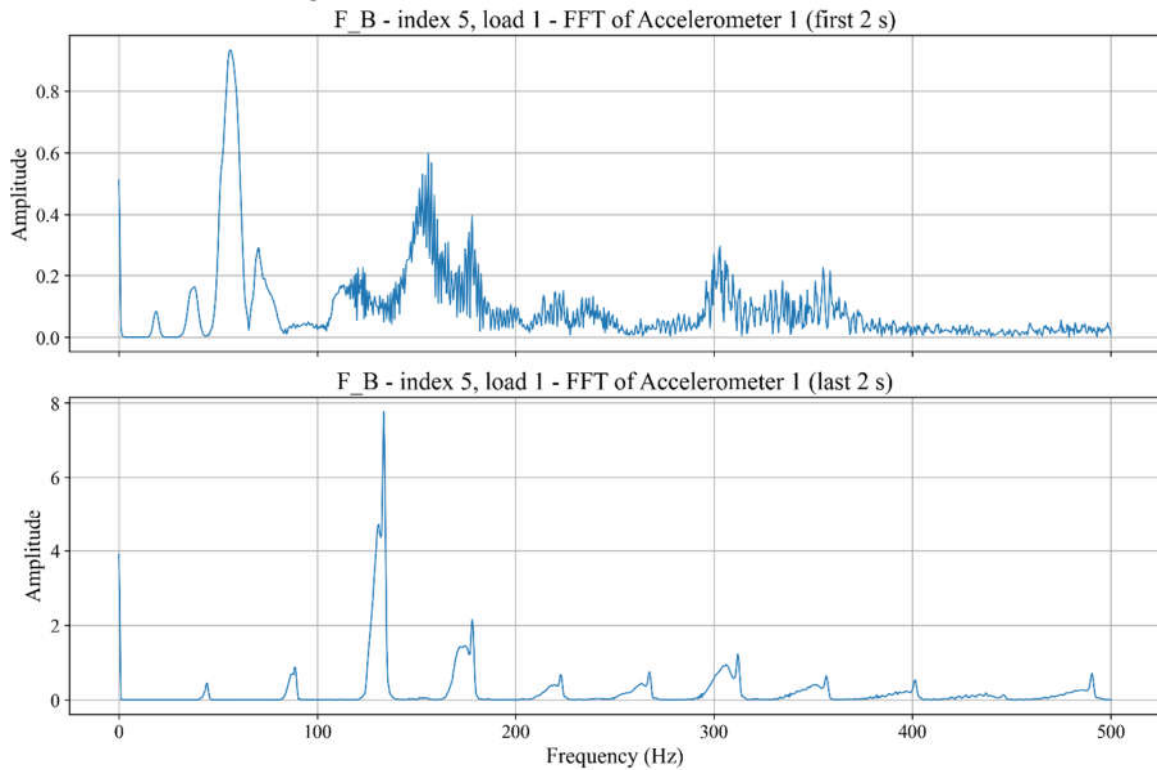


Figure 59. Faulty Bearings 15.22 Hz → 44.96 Hz

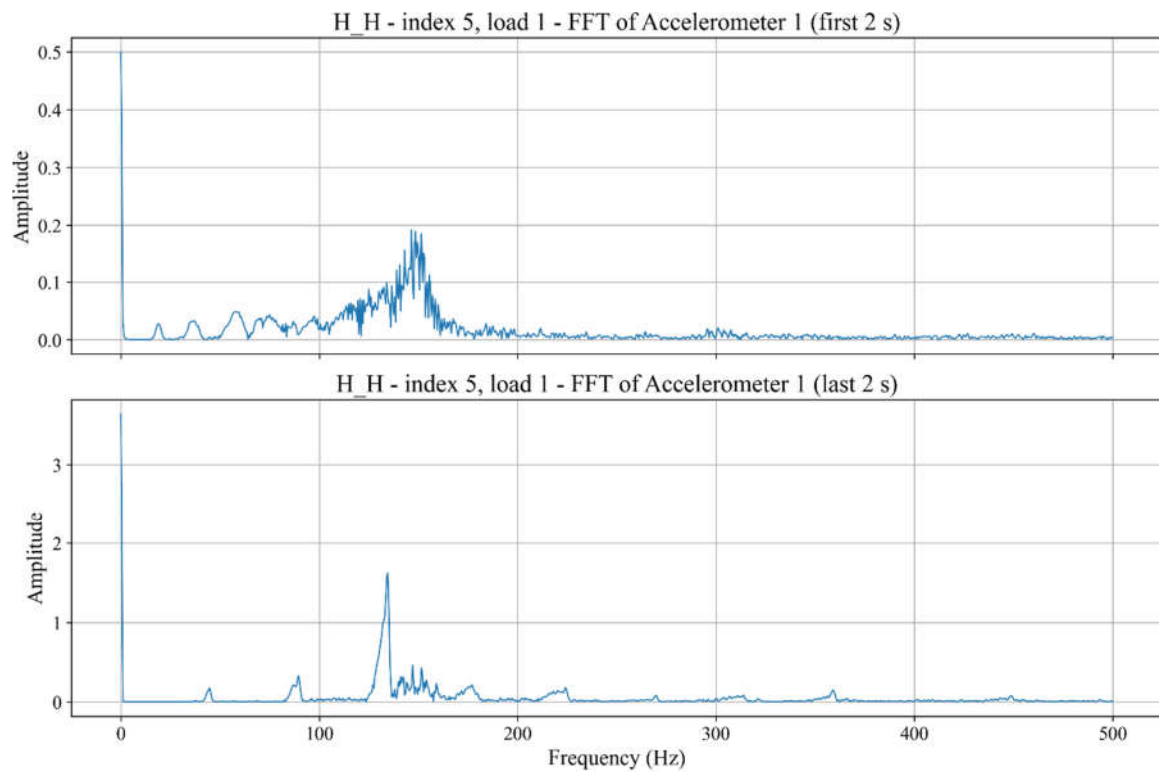


Figure 60. Healthy 14.99 Hz $\rightarrow$ 44.86 Hz

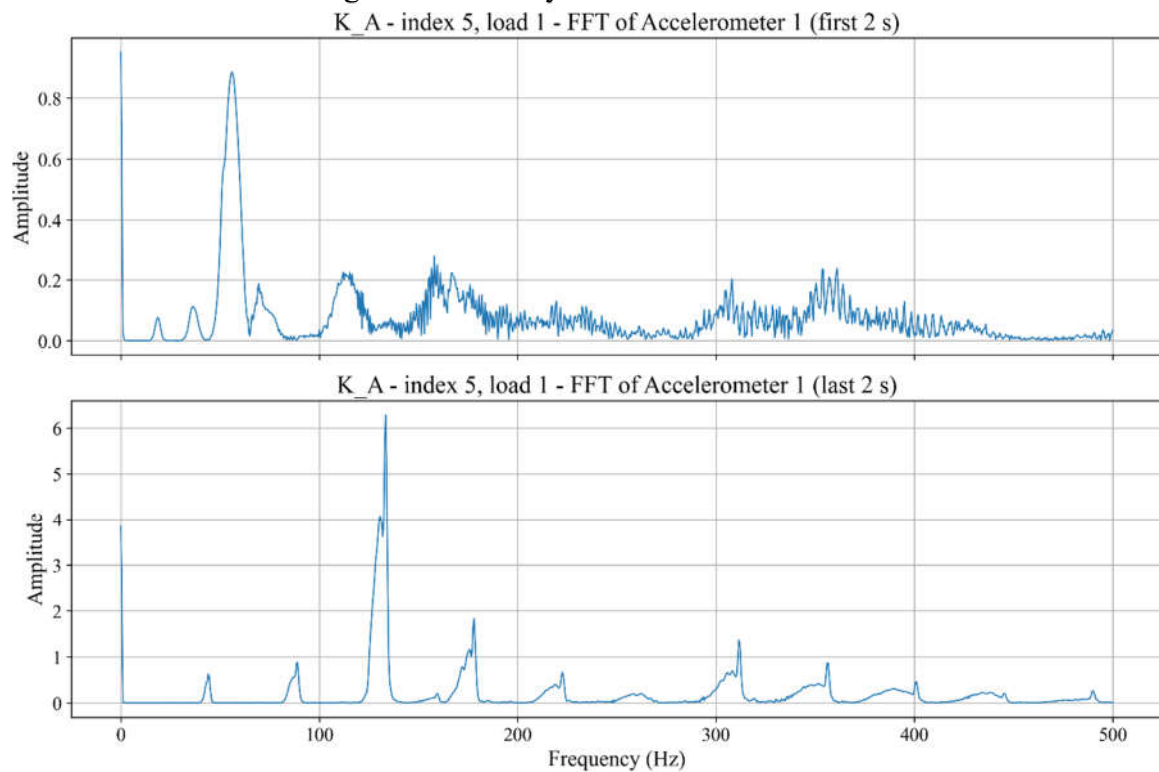


Figure 61. Broken Rotor Bars 15.13 Hz $\rightarrow$ 44.87 Hz

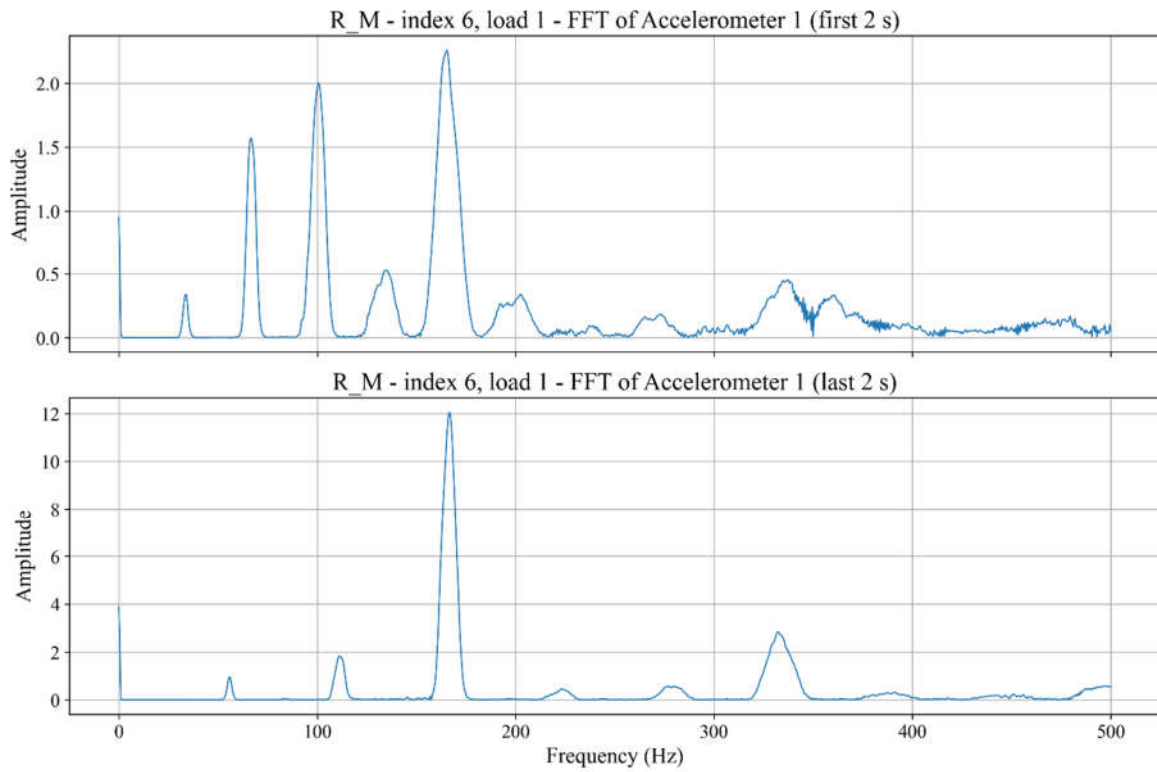


Figure 62. Rotor Misalignment 30.33 Hz $\rightarrow$ 59.96 Hz

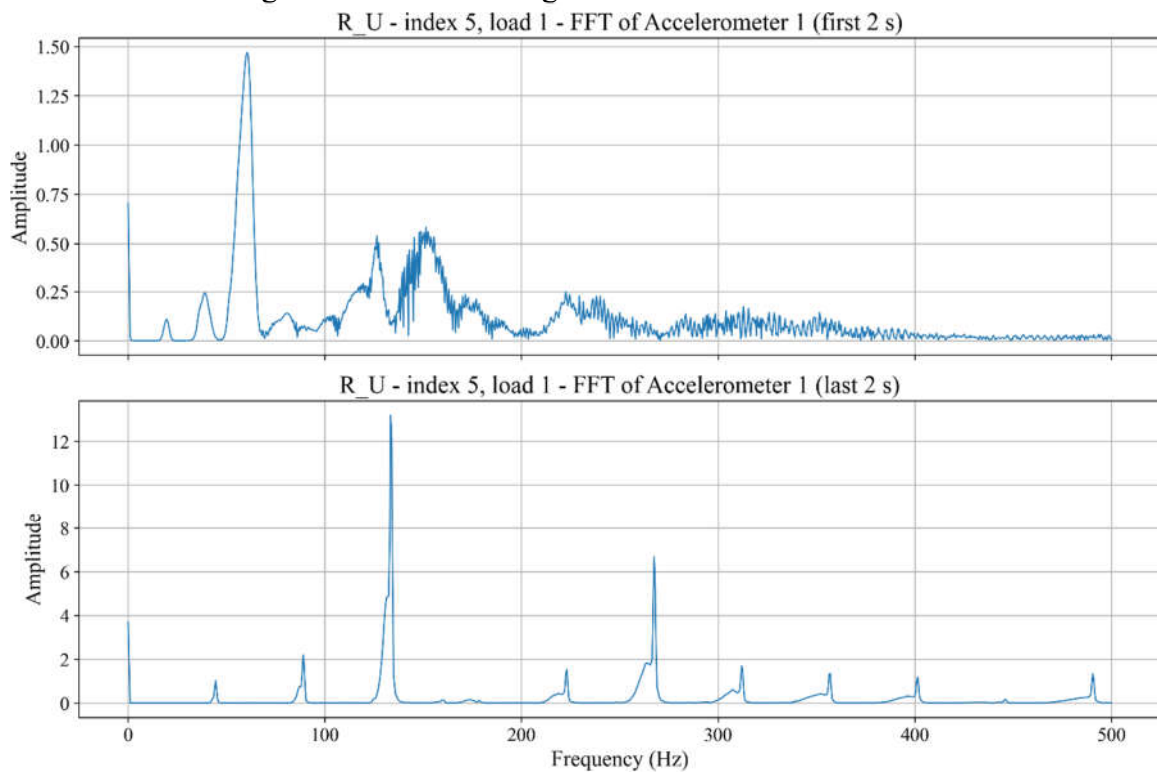


Figure 63. Rotor Unbalance 15.72 Hz $\rightarrow$ 44.97 Hz

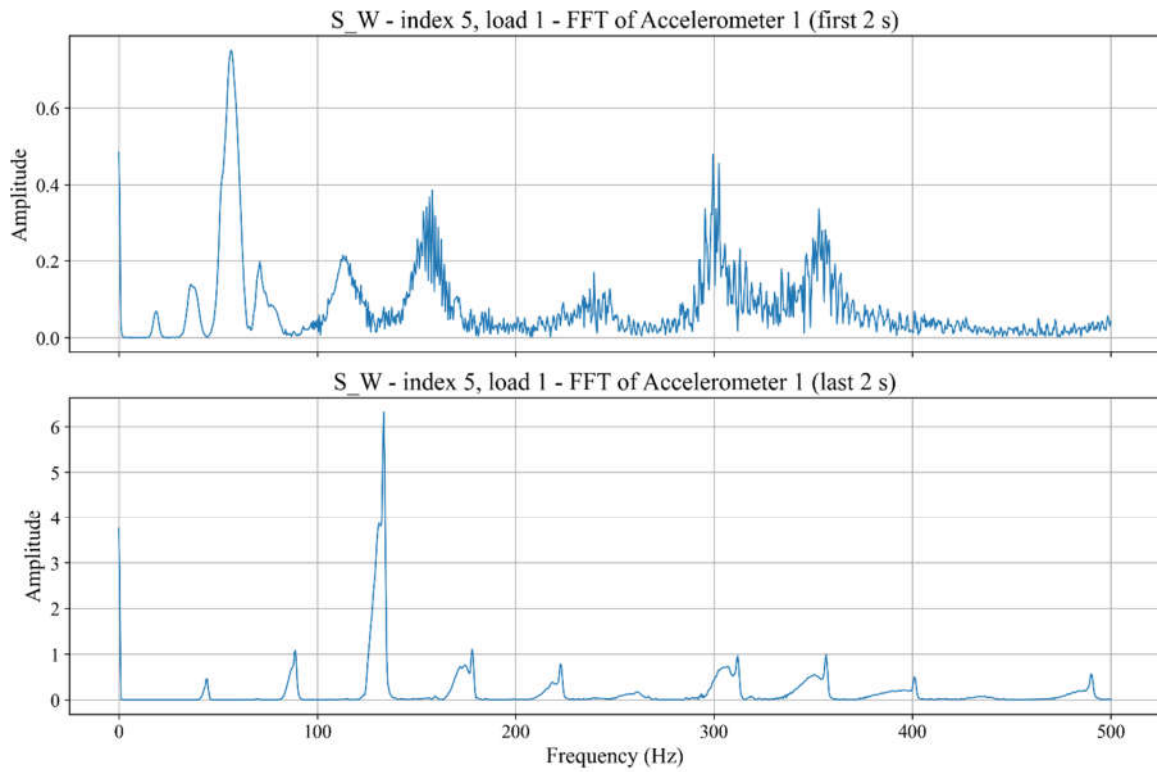


Figure 64. Stator Winding Fault 15.33 Hz → 44.92 Hz

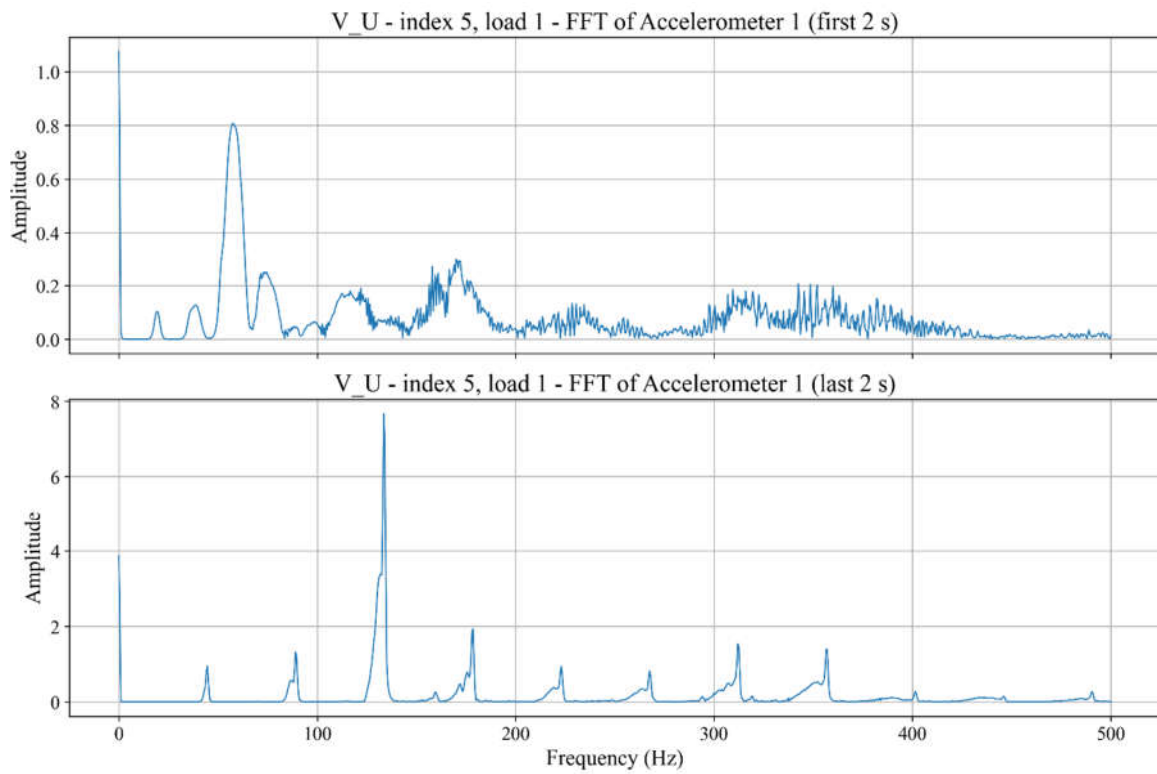


Figure 65. Voltage Imbalance and Single-Phase Fault 15.14 Hz → 44.96 Hz

V. CONCLUSION

The Ottawa 2018–2019 dataset provides a valuable experimental foundation for bearing fault diagnosis under variable speed conditions—a non-stationary problem that often significantly degrades the effectiveness of traditional spectrum-based methods. The SGA preprocessing method is introduced as a key approach to signal normalization, domain shift reduction, and speed-dependent fault characteristic stabilization. Based on this report, further development directions could include: designing preprocessing pipelines (SGA/angle-domain resampling), multi-domain feature extraction (time-frequency), and building machine learning/deep learning-based diagnostic models for quantitative evaluation.

REFERENCES

- [1] Rotational Speed Conditions,” Mendeley Data, V1, 2018. doi:10.17632/v43hmbwxpm.1.
- [2] H. Huang and N. Baddour, “Bearing Vibration Data under Time-varying Rotational Speed Conditions,” Mendeley Data, V2, 2019. doi:10.17632/v43hmbwxpm.1.
- [3] P. Van Trang and N. Thanh Lich, “Fault classification of the rolling bearing based on minimax entropy domain adaption augmented with signal generation algorithm,” Tạp Chí Khoa Học Giao Thông Vận Tải, vol. 75, no. 1, pp. 1110–1124, 2024. <https://doi.org/10.47869/tcsj.75.1.1>.