

## Article

# AI-Based Health Monitoring for Class I Induction Motors in Data-Scarce Environments: From Synthetic Baseline Generation to Industrial Implementation

Duter Struwig <sup>1</sup> , Jan-Hendrik Kruger <sup>1</sup> , Henri Marais <sup>2,\*</sup>  and Abrie Steyn <sup>1</sup><sup>1</sup> School of Mechanical Engineering, North-West University, Potchefstroom 2531, South Africa<sup>2</sup> School of Electrical, Electronic and Computer Engineering, North-West University, Potchefstroom 2531, South Africa

\* Correspondence: henri.marais@nwu.ac.za

## Abstract

Condition-based maintenance strategies using AI-driven health monitoring have emerged as valuable tools for industrial reliability, yet their implementation remains challenging in industries with limited operational data. Class I induction motors ( $\leq 15$  kW), which power critical equipment in industries such as grain handling facilities, represent a significant portion of industrial assets but lack established healthy vibration baselines for effective monitoring. A fundamental challenge exists in deploying AI-based health monitoring systems when no historical performance data is available, creating a ‘cold-start’ problem that prevents industries from adopting predictive maintenance strategies without costly pilot programs or prolonged data collection periods. This study developed a data-driven health monitoring framework for Class I induction motors that eliminates the dependency on long-term historical trends. Through extensive experimental testing of 98 configurations on new motors, a correlation between vibration amplitude at rotational frequency and motor power rating was established, enabling the creation of a synthetic signal generation algorithm. A robust Health Index (HI) model with integrated diagnostic capabilities was developed using the JPCCED-HI framework, trained on both experimental and synthetically generated healthy vibration data to detect degradation and diagnose common failure modes. The regression analysis revealed a statistically significant relationship between motor power rating and healthy vibration signatures, enabling synthetic generation of baseline data for any Class I motor within the rated range. When implemented at an operational grain silo facility, the HI model successfully detected faulty behavior and accurately diagnosed probable failure modes in equipment with no prior monitoring history, demonstrating that maintenance decisions could be made based on condition data rather than reactive responses to failures. This framework enables immediate deployment of AI-based condition monitoring in industries lacking historical data, eliminating a major barrier to adopting predictive maintenance strategies. The synthetic data generation approach provides a cost-effective solution to the data scarcity problem identified as a critical challenge in industrial AI applications, while the successful industrial implementation validates the feasibility of this approach for small-to-medium industrial facilities.



Academic Editors: Rui Araújo and Alexandru Stanciu

Received: 29 December 2025

Revised: 12 January 2026

Accepted: 13 January 2026

Published: 16 January 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.**Keywords:** condition monitoring; health index; JPCCED-HI framework; rotating machines; predictive maintenance; data-scarce environments

## 1. Introduction

Industrial maintenance strategies directly impact operational profitability, with maintenance costs that can comprise up to 70% of total production costs [1]. Despite McKinsey's predictions that digital maintenance technologies could increase asset availability by 5% to 15% and reduce maintenance costs by 18% to 25% [2], the adoption of condition-based maintenance (CBM) remains limited—particularly for Class I induction motors ( $\leq 15$  kW as defined by ISO 20816 [3]). These motors constitute a substantial portion of industrial assets, powering critical equipment across manufacturing sectors, yet the implementation of systematic health monitoring faces persistent barriers.

The financial case for transitioning from run-to-failure approaches is compelling. Mobley [4] demonstrated that repairs under run-to-failure paradigms typically cost three times more than equivalent preventive maintenance interventions, while predictive maintenance implementations have achieved 25–30% reductions in maintenance costs along with 35–45% reductions in downtime [5,6]. However, Anomughanran [7] found that approximately 55% of industrial plants continue to operate under corrective maintenance paradigms, suggesting that technical or economic barriers prevent widespread adoption of CBM, despite its demonstrated benefits.

A fundamental implementation barrier exists: effective condition-based maintenance requires baseline data that represent healthy equipment operation [8], yet many industrial facilities lack such data—particularly those with legacy equipment. Standards, including ISO 20816 [3], provide zone boundaries for acceptable vibration severity, but do not establish healthy baselines specific to individual motor configurations. This creates a “cold-start problem”: industries cannot deploy AI-based health monitoring without historical performance data, forcing costly pilot programmes or prolonged data collection periods that undermine the economic feasibility of CBM implementation.

The challenge is compounded by the fundamental uncertainty about the age of the equipment as a failure predictor. For example, Bengtsson [1] found that only 11% of the components of the aircraft in their study exhibited clear ageing characteristics, with 89% randomly failing without significant ageing patterns—a phenomenon that became even more pronounced with increased complexity of the equipment. This undermines time-based maintenance strategies while simultaneously highlighting the data requirements that condition-based alternatives demand. For facilities with equipment of unknown initial health status, this circularity prevents CBM adoption: without baseline data, health cannot be assessed; without health assessment capability, baseline establishment remains impossible.

Class I induction motors present a particularly acute instance of this problem. These motors power conveyors, elevators, fans, and processing equipment in various industrial applications, yet there is no established correlation between motor specifications and expected healthy vibration characteristics. Practitioners face a stark choice: assume that current operational equipment is healthy (and potentially contaminate all subsequent analyses) or undertake expensive measurement campaigns on new equipment for each motor size and coupling configuration encountered.

This study addresses the cold-start problem through the development of a data-driven health monitoring framework that eliminates the dependency on historical operational data. The approach combines extensive experimental characterisation of healthy motor vibration signatures, regression analysis establishing correlations between motor power rating and vibration amplitude, synthetic baseline generation algorithms, and robust Health Index modelling with integrated diagnostic capabilities. Implementation in an operational facility with no prior monitoring history demonstrated that maintenance decisions can be based

on condition data rather than reactive responses to failures, even when the initial health status of the equipment is unknown.

The remainder of this paper is structured as follows: Section 2 provides a critical review of health index construction methods, data scarcity challenges in industrial settings, and gaps in existing motor health monitoring approaches—particularly the JPCCED-HI framework’s component-level validation limitations. Section 3 details the experimental methodology, regression analysis procedures, synthetic signal generation algorithms, and construction of the Health Index model with diagnostic capabilities. Section 4 presents regression results establishing the power–vibration correlation, synthetic signal validation, model verification performance, and industrial implementation outcomes across ten motors. It discusses the validation of diagnostic results through physical inspection, practical implications for the resolution of cold-start problems, and study limitations. Section 5 concludes with primary contributions and directions for future research.

## 2. Literature Review

### 2.1. Health Index Construction Methods

The Health Index (HI) has emerged as a critical tool for extracting degradation information from condition-indicating parameters to evaluate asset health [9]. Zhou et al. [9] categorise HI construction methods into four primary approaches: feature-based methods employing signal processing techniques in the time, frequency, and time-frequency domains; multi-feature fusion methods combining several features to improve prediction accuracy; distance-based methods measuring similarity between healthy and faulty signals using Euclidean or Mahalanobis calculations; and deep learning-based methods utilising convolutional neural networks (CNNs), recurrent neural networks (RNNs), and autoencoders to automatically learn representations from datasets.

Feature-based approaches have been extensively developed for rotating machinery applications. Widodo and Yang [10] introduced principal component analysis (PCA) to fuse peak value, kurtosis, and entropy into unified health metrics. Zhao et al. [11] used S-transform and Gaussian pyramid methods to obtain time-frequency representations on multiple scales, using PCA-based indicators for bearing health assessment. Hong et al. [12] applied empirical mode decomposition with self-organising maps to analyse vibration signals and calculate confidence values for the bearing health states. These methods demonstrate the value of multi-domain feature extraction, but typically require extensive historical data for baseline establishment.

Deep learning approaches have gained considerable attention for their ability to automatically extract features and construct HI models. Convolutional neural networks [3,13], recurrent neural networks [14,15], unsupervised deep belief networks [16,17], and convolutional autoencoder networks [18,19] have been applied to the construction of HIs. However, these methods present significant practical limitations. Song et al. [20] identified three fundamental problems with conventional HI models: differences in operating conditions between training and testing data reduce model performance; data are required throughout the entire equipment life cycle under different operating conditions, making acquisition expensive; and complete datasets combined with complex models make real-time implementation difficult.

Methods based on reconstruction errors offer an alternative approach. Qiu et al. [21] introduced wavelet filter-based weak signature enhancement combined with self-organising map (SOM) methods to assess performance degradation. Huang et al. [22] used minimum quantisation error indicators derived from SOM trained in vibration features. Yan et al. [23] presented optimised stacked denoising autoencoders based on multidomain indicators

for a robust representation of features. Although these approaches reduce dependence on labelled fault data, they still require substantial healthy-state datasets for model training.

Distance-based and similarity-based methods provide yet another avenue. Guo et al. [24] proposed similarity features related to the time-frequency range, addressing varying statistical feature ranges. Wang et al. [25] estimated degradation by calculating the deviation of vibration signal statistics from known healthy states. Li et al. [26] calculated Mahalanobis Distance from time and time-frequency domain characteristics, obtaining monotonically increasing HIs through cumulative sum approaches. Schwartz et al. [27] proposed automatic unsupervised Kernel PCA approaches for similarity-based HI creation. These methods effectively quantify deviation from baseline conditions, but require that baseline conditions be accurately established—a requirement that proves problematic in data-scarce industrial environments.

A persistent challenge across all HI construction categories is the establishment of upper and lower bounds. Even for methods requiring only healthy-state training data, challenges persist in establishing appropriate limits without faulty reference data [20]. The lack of standardised HI bounds complicates intuitive health assessment and threshold setting for maintenance decisions [1].

## 2.2. The Data Scarcity Challenge in Industrial Settings

The successful implementation of condition-based maintenance (CBM) depends on the availability of reliable data that accurately represent asset degradation and allow for the detection of condition changes [8]. Nevertheless, many industries have been slow to adopt advanced maintenance technologies due to persistent data-related challenges. Jardine, Lin, and Banjevic [28] identify key barriers, including inadequate or inappropriate data collection and storage, poor communication between developers and practitioners, insufficient validation methods, and implementation difficulties arising from evolving designs, technologies, and business policies.

Infrastructure constraints further impede CBM adoption, particularly the absence of facilities capable of supporting monitoring equipment, which significantly increases instrumentation and installation costs [29]. Many industrial plants were not designed with future condition monitoring in mind, necessitating substantial investment in specialised hardware and software for data acquisition, storage, and analysis. These costs are difficult to quantify and compare with traditional maintenance approaches, especially when financial benefits are realised only after extended periods of operation [8].

Organisational and human factors present additional barriers. Effective implementation of CBM requires coordination between organisational levels, yet management and maintenance personnel often resist changes that increase short-term workload while offering delayed benefits [1]. Common obstacles include resistance to change, limited teamwork, weak leadership, and insufficient trust. Mobley [6] reports that CBM programmes often fail within 24 months when senior management does not sustain investment in training and consulting support.

Baseline data availability is a critical requirement for vibration-based condition monitoring. A meaningful comparative analysis depends on baseline measurements that represent healthy operating conditions, established at installation or shortly after maintenance [6]. Accurate baseline use requires data resetting after maintenance, correct machine train identification, and defined operating envelopes for different process conditions [6]. Mais [30] similarly emphasises the need for healthy baseline datasets for each machine and operating state. However, such baseline data are often unavailable for legacy equipment, substantially limiting the effectiveness of CBM. Recent comprehensive reviews [31–33] highlight the continued dominance of data-driven approaches in industrial condition monitoring, yet

consistently identify data availability as a critical barrier to implementation—particularly for facilities lacking historical operational records.

### 2.3. The Cold-Start Problem

There is a persistent gap in both the literature and industrial practice regarding the initiation of condition monitoring programmes. Many organisations implement condition-based maintenance (CBM) without establishing the initial health state of monitored assets, implicitly assuming healthy operating conditions. This assumption compromises trend analysis, as baselines may be derived from data representing unknown or degraded states. Although historical maintenance records may be available, they are often incomplete and insufficient for rigorous analysis.

This issue is particularly pronounced for Class I induction motors ( $\leq 15$  kW). The ISO 20816 standard series [3] employs the same machine classification framework and defines general vibration evaluation zones but does not specify healthy baseline values for this motor class, noting only that RMS vibration velocity typically lies toward the lower end of the prescribed ranges. In the absence of standardised baselines, practitioners must assume asset health or perform costly measurements on new equipment.

Time-based methods provide limited mitigation. Bengtsson [1] showed that only 11% of aircraft components exhibited clear ageing characteristics, while manufacturing studies found ageing patterns in only 30% of components, with predictability decreasing as system complexity increased [34]. These findings demonstrate that the age of the equipment is an unreliable indicator of health, which explains the ineffectiveness of time-based maintenance strategies.

Although CBM was introduced to address these shortcomings [1], its effectiveness depends on historical and baseline data that are often unavailable in industrial settings, particularly for legacy equipment. This data scarcity challenge is not unique to rotating machinery; analogous issues have been identified across multiple industrial domains, including railway point machine monitoring, where the absence of comprehensive fault data similarly impedes the deployment of intelligent condition monitoring systems [35].

### 2.4. Induction Motor Health Monitoring

Induction motors are fundamental for industrial operations, accounting for approximately 45–50% of global industrial energy consumption [36]. Their widespread use is driven by robust construction, although failures can result in substantial operational disruption. Motor failures arise from a combination of design limitations, manufacturing tolerances, installation errors, adverse operating environments, demanding load conditions, and inappropriate maintenance practices [36–38].

In normal operation, electromagnetic and mechanical forces interact to produce stable performance with low vibration levels [37]. Faults disturb this balance, leading to increased vibration and eventual failure. Common mechanical faults include unbalance, misalignment, bearing defects, and looseness, while electrical faults include stator winding defects, broken rotor bars, and eccentricity-related phenomena [36,37,39].

Vibration monitoring is the most widely adopted condition monitoring technique for rotating machinery, including induction motors, due to its high sensitivity to a broad range of mechanical and electrical faults [6]. Frequency-domain analysis enables the identification of faults through characteristic signatures: unbalance at the  $1\times$  rotational frequency, misalignment at the  $1\times$  to  $3\times$  harmonics, bearing defects at characteristic fault frequencies, and electrical faults through side bands around the supply frequency harmonics [30,40].

Despite its effectiveness, vibration-based monitoring depends on reliable healthy baselines for a meaningful comparison. Although ISO 20816 [3] defines general vibration severity zones, it does not provide equipment-specific baseline values. For Class I motors, the standards simply note that vibration levels tend toward the lower range, however, offering no predictive link between motor specifications and expected healthy vibration behaviour. This lack of guidance complicates the establishment of a baseline for previously unmonitored equipment.

### 2.5. The JPCCED-HI Model: Advances and Remaining Gaps

The Joint Principal Component Cumulative Empirical Distribution Health Index (JPCCED-HI) model [20] addresses data scarcity challenges for rotating machinery by requiring only healthy-state data for baseline establishment and providing a bounded health index (0–1) for intuitive interpretation. However, a critical examination reveals significant limitations for a broader industrial application.

#### 2.5.1. Component-Level vs. System-Level Validation

Song et al. [20] validated the model exclusively with isolated components: a two-stage parallel shaft gearbox (40 rpm, controlled conditions) and NSK 6804DD bearings from the IEEE PHM 2012 Prognostic Challenge. Neither case addressed complete machine systems comprising driver and driven equipment.

#### 2.5.2. Absence of Motor-Machine Coupling Effects

Industrial induction motors rarely operate in isolation—they drive machinery through mechanical couplings (chain-sprocket, V-belt, direct) that introduce additional vibration sources and frequency components. The JPCCED-HI model was not validated to distinguish motor health degradation from coupling-induced vibration changes.

#### 2.5.3. Absence of Size to Vibration Amplitude Correlation

Neither validation addressed the relationship between the power rating of the equipment and the expected healthy vibration amplitude. For Class I induction motors ( $\leq 15$  kW per ISO 20816 [41]), no established correlation exists in the literature, preventing the generation of a healthy synthetic baseline and forcing the reliance on physical measurements.

#### 2.5.4. Laboratory vs. Industrial Operating Conditions

Typically, CBM research is performed under laboratory conditions [20,42] where machines are operated under controlled conditions: constant speeds, consistent loading, direct sensor mounting, and continuous operation until natural or induced failure. Industrial environments present different challenges including intermittent operation, variable loading, constrained sensor placement, environmental noise from adjacent equipment, and extended periods between measurements.

#### 2.5.5. Detection Without Diagnosis

The JPCCED-HI model provides health detection, but not diagnostic functionality. Therefore, when HIs indicate degradation, maintenance personnel lack information on which failure modes are developing—information that is critical to guide corrective actions and procurement of spare parts.

### 2.6. Research Gap and Contribution

The literature reveals a circular challenge: condition-based maintenance requires healthy baseline data, yet many facilities lack such data. The JPCCED-HI framework [20]

addressed the requirements of healthy-state training, but was validated only with isolated components under controlled conditions.

For Class I induction motors HI frameworks, four critical gaps therefore remain:

**No power–vibration correlation.** ISO Standards [3] provide zone boundaries, but not predictive equations that relate motor specifications to expected vibration characteristics.

**No system-level validation.** Previous research focused only on isolated components, not motors coupled to driven equipment.

**No synthetic baseline capability.** Predictive models require extensive operational data for each motor, perpetuating data scarcity.

**No industrial validation.** Laboratory studies do not address the cold-start problem in facilities with no monitoring history.

The present study addressed these gaps through experimental investigation of the power–vibration correlation for Class I motors, synthetic baseline generation, and industrial implementation and validation, with fault confirmation through operational measurements and physical inspection.

### 3. Methodology

#### 3.1. Experimental Setup

##### 3.1.1. Layout of Test Bench

The reconfigurable test bench shown in Figure 1 was used for the experimental aspects.

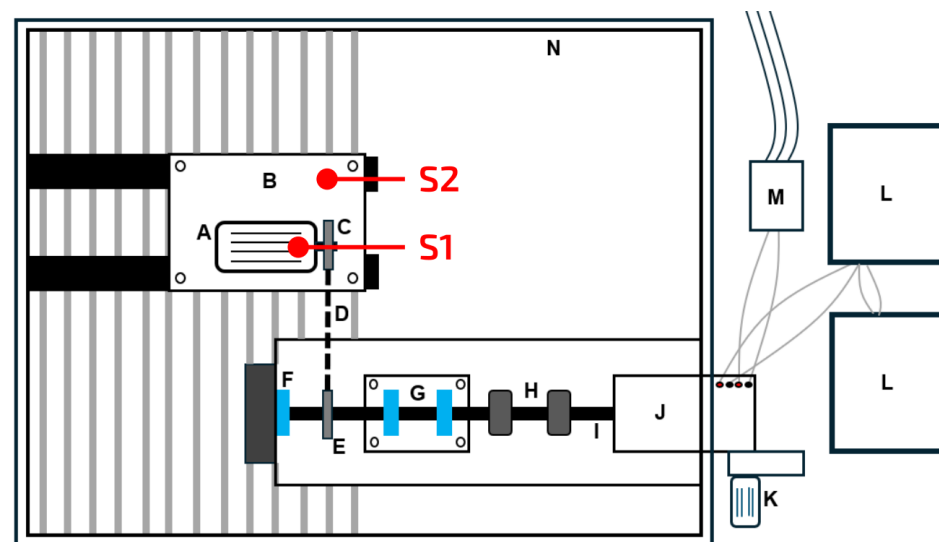


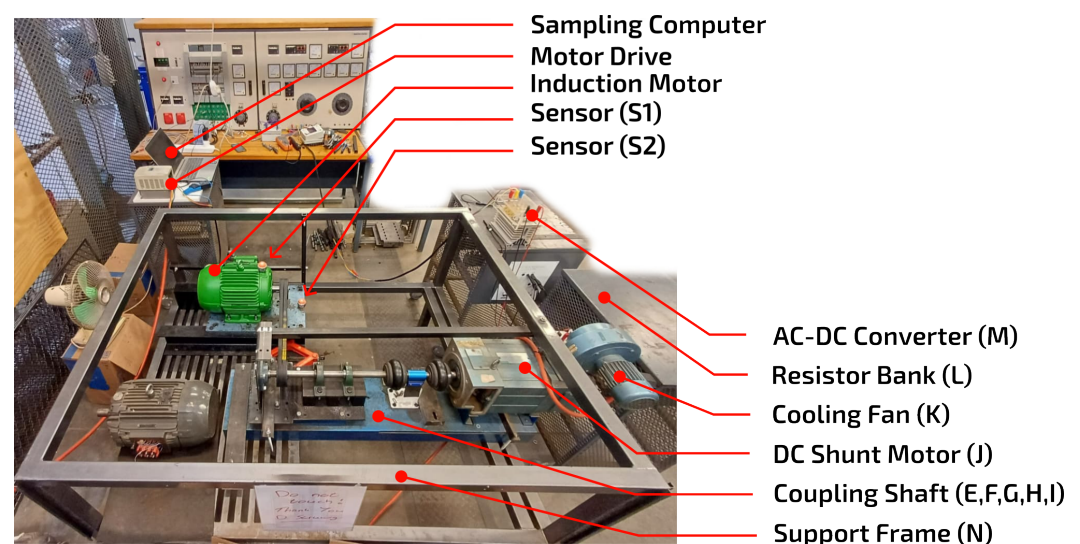
Figure 1. Experimental test bench layout.

1. **Induction motor (A):** The mounting position for all induction motors used during data collection.
2. **Mounting plate (B):** The mounting plate, which serves as the base onto which the different induction motors are mounted, ensuring rigid mounting.
3. **Driving end (C):** The sprocket/pulley mounted on the driving end (DE) of the induction motor, transferring torque to the chain/V-belt.
4. **Power transmission element—chain/V-belt (D):** The connection between the driving and driven side of the setup; used to transfer torque from the induction motor shaft to the loaded shaft.
5. **Driven end (E):** The sprocket/pulley mounted on the shaft connected to the load, transferring the torque from the sprocket/pulley onto the loaded shaft.

6. **End bearing and support structure (F):** An end bearing mounted onto the end of the loaded shaft, supporting the shaft to reduce the moment at the point of load, thus preventing bending due to the load point on the shaft.
7. **Support bearing blocks (G):** Two support bearing blocks through which the shaft is fitted, providing support between the driven side and the load.
8. **Flexible couplings (H):** Two flexible couplings with a torque transducer mounted between them (the transducer was not used in the current study).
9. **Shaft (I):** The final shaft connected to the load.
10. **Direct current (DC) shunt motor (J):** The DC shunt motor that acts as a generator is used to convert the mechanical energy created by the induction motors into electrical energy, which then is “dumped” in the resistor banks. This results in a mechanical load on the shaft, providing a means to operate the induction motors at full load condition (FLC). The shunt motor is excited with a voltage from the high-current bench.
11. **Three-phase fan (K):** A three-phase fan that is powered during the loaded experimental runs to ensure adequate cooling of the DC shunt motor.
12. **Resistor banks (L):** Two resistor banks are used to “dump” the electrical energy generated by the DC shunt motor during the loaded experimental tests.
13. **AC-DC converter (M):** This component takes the alternating current (AC) input from the high-current bench and converts it to a direct current (DC), which is then connected to the DC shunt motor, inducing a load onto the shaft.
14. **Steel frame (N):** Serves as the support structure, ensuring rigid mounting for all components mounted onto the structure.

To control the speed of the induction machine, a WEG CFW08 variable speed drive (VSD) was used. The actual motor speed was verified with a digital tachometer and Erbesd Instruments EPH-V11E Triaxial wireless vibration sensors were placed at locations S1 and S2 in Figure 1.

A typical configuration of the test bench is shown in Figure 2, where the active motor in green can clearly be seen, along with the mechanical coupling and load.



**Figure 2.** Photo of the experimental test bench layout.

### 3.1.2. Experimental Operating Conditions

A selection of induction motors was made to represent a sufficiently broad range of induction motors within Class I (see Table 1). All induction motors used during data collection were new W22 IE3 Premium Efficiency, four pole WEG induction motors, without any prior operation.

**Table 1.** Power ratings and frame sizes of induction motors for experimental data collection.

| Power Rating (kW) | Frame Size (IEC) |
|-------------------|------------------|
| 0.55              | 80               |
| 0.75              | 80               |
| 1.5               | 90 L             |
| 2.2               | 100 L            |
| 3                 | 100 L            |
| 4                 | 112 M            |
| 5.5               | 132 S            |

From a cursory investigation of common industrial processing plants, particularly those in the agri-processing space, it was determined that the most common types of coupling between the motor and the driven equipment were chain and belt-based with limited direct coupling. For this reason, in this study the following configurations were investigated:

**Stand-alone.** The induction motor is mounted onto the baseplate, but no other coupling connections are made; this configuration was used primarily for preliminary analysis purposes, see Figure 3a.

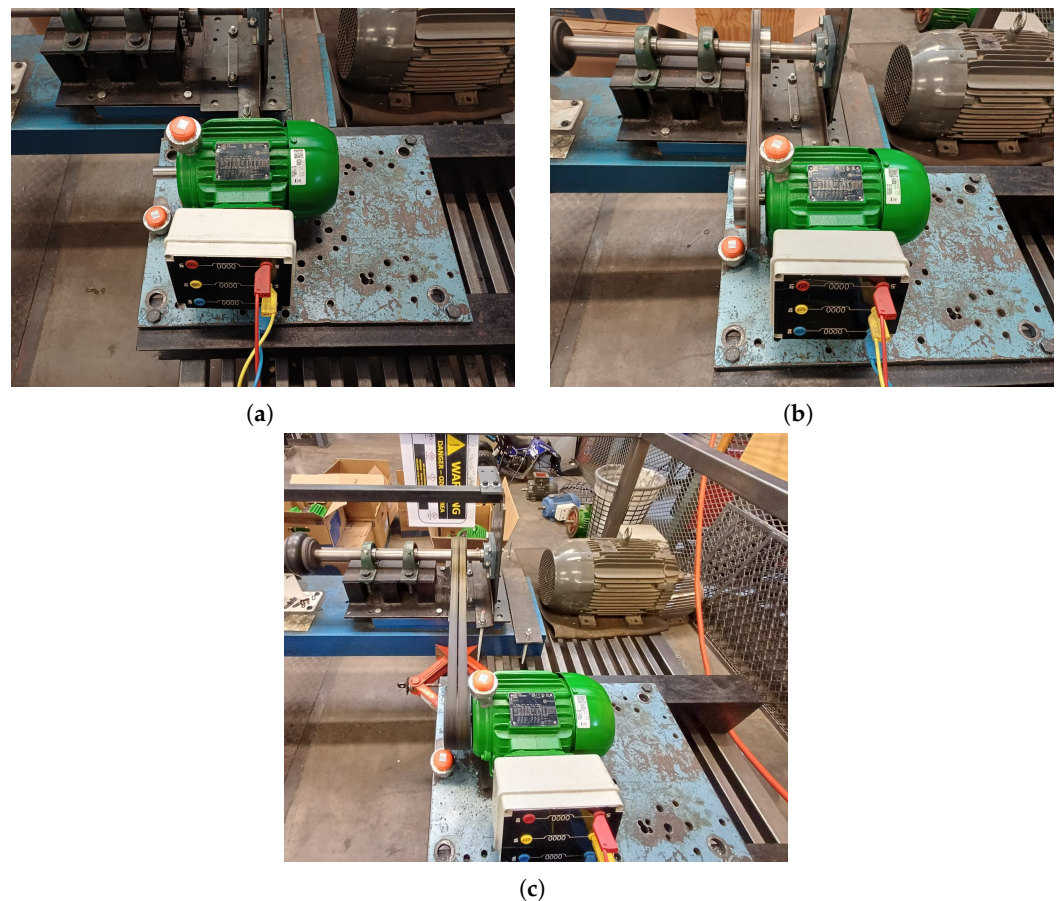
**Chain-sprocket condition (Unloaded).** The induction motor is coupled with the shaft connected to the load using a chain-sprocket coupling, but the load is not excited with a voltage, and thus the motor only spins the shaft with no additional load, except for the friction of the components, see Figure 3b.

**Chain-sprocket condition (Loaded).** The induction motor is coupled with the shaft connected to the load using a chain-sprocket coupling, and the load is excited to a voltage, which results in the induction motor operating at or close to the specified FLC as stated on its nameplate.

**V-belt pulley condition (Unloaded).** The induction motor is coupled to the shaft connected to the load using a V-belt pulley coupling, but the load is not excited with a voltage, and therefore the motor only spins the shaft with no additional load, except for the friction of the components; see Figure 3c.

**V-belt pulley condition (Loaded).** The induction motor is coupled with the shaft connected to the load using a V-belt pulley coupling, and the load is excited to a voltage, which results in the induction motor operating at or close to the specified FLC as stated on its nameplate.

In the industrial settings explored, most applications operated the motors at a fixed speed close to the rated speed of the motor, and for this reason a 1:1 ratio was used for all coupling configurations. In addition to the different coupling configurations, data was also collected at different operating speeds. For chain-sprocket coupling configurations, the percentages of synchronous speed were 25%, 50%, 75%, and 100%. For the V-belt pulley configuration, measurements were made only at a speed setting of 100%. For the stand-alone and chain-sprocket coupling conditions, 30 vibration measurements were made for each individual configuration. For the V-belt pulley coupling configuration, only 10 measurements were taken for each motor size.



**Figure 3.** Different configurations on which experimental measurements were made. (a). Stand-alone configuration. (b). Chain-sprocket coupling. (c). V-belt pulley coupling.

A summary of the experimental configurations investigated is provided in Table 2.

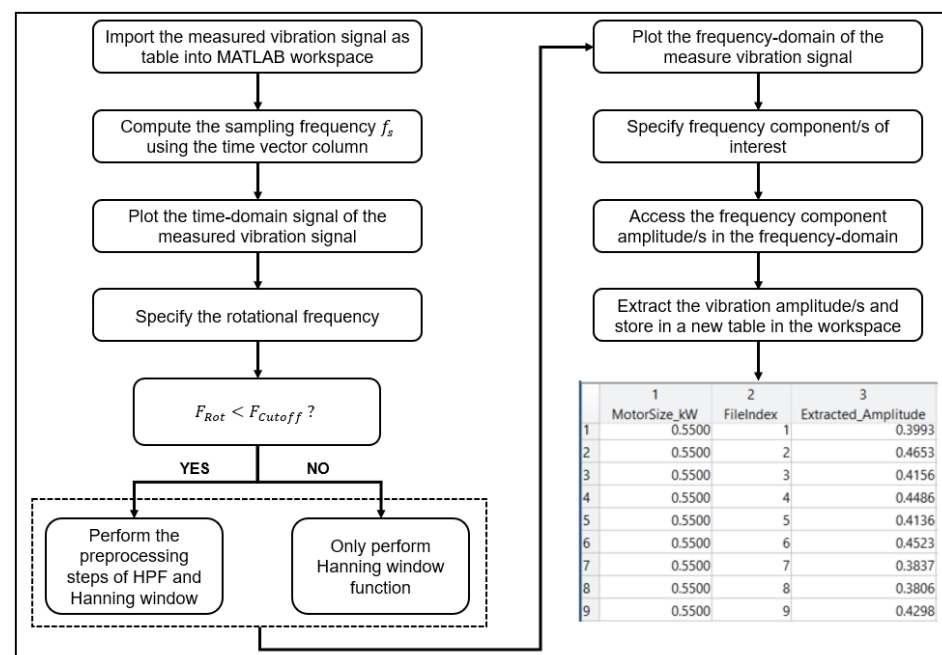
**Table 2.** Experimental configurations and number of signals measured.

| Configuration             | Sub Configurations | Number of Measurements | Total Signals |
|---------------------------|--------------------|------------------------|---------------|
| Stand-alone               | 28                 | 30                     | 840           |
| Chain-sprocket (unloaded) | 28                 | 30                     | 840           |
| Chain-sprocket (loaded)   | 28                 | 30                     | 840           |
| V-belt pulley (unloaded)  | 7                  | 10                     | 70            |
| V-belt pulley (loaded)    | 7                  | 10                     | 70            |

Vibration measurements were acquired using the EPH-V11E triaxial wireless sensors at a sampling frequency of 12,800 Hz with a signal duration of 6 s per measurement, yielding 76,800 samples per signal. Frequency-domain analysis employed the Fast Fourier Transform (FFT) with a Hanning window to reduce spectral leakage, providing a frequency resolution of approximately 0.17 Hz. A 10 Hz high-pass filter was applied to eliminate low-frequency structural effects. For bearing fault assessment, the RMS amplitude in the 200 Hz to 6 kHz band was evaluated against a  $3\sigma$  threshold derived from healthy baseline data.

### 3.2. Power–Vibration Correlation procedure

Using the procedure described in Figure 4, the fundamental frequencies for each motor in the sample set were extracted.



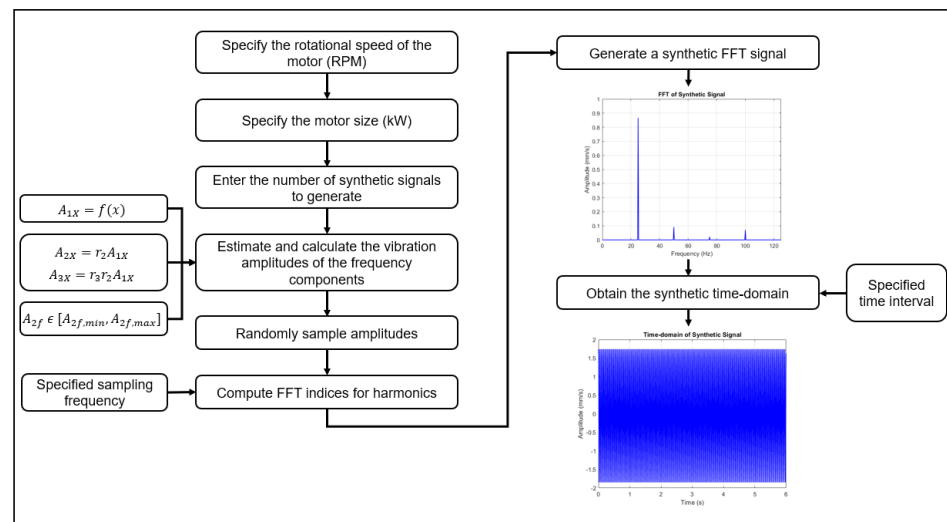
**Figure 4.** General signal analysis procedure.

These vibration amplitudes and the associated motor power were used in conjunction with the MATLAB CurveFitter application to perform curve fitting. Based on the percentage error calculated between the fitted curve and the actual values, the optimal curve was determined and the associated parameters for the curve and the 95% confidence interval extracted. Regression analysis was performed in MATLAB for the 100% operating speed condition, informed by the observation that a typical induction motor installed in industry operates at (or close to) the synchronous speed.

### 3.3. Synthetic Signal Generation

Since it would be difficult and expensive to construct an extensive database of healthy vibration measurements on a large number of Class I induction motors under operational conditions in an industrial setting (data scarcity problem), a method is required to synthetically generate these signals for training the HI models. This enables the generation of unlimited healthy vibration signals and essentially addresses the information gap between experimental, theoretical, and industrial settings.

Figure 5 outlines the steps to generate healthy, statistically representative vibration spectra by means of the following steps:



**Figure 5.** Steps for the synthetic signal generation methodology.

- Step 1:** Input specification: The motor rotational speed (RPM) and power rating (kW) are specified as input parameters, corresponding to the target motor for which baseline signals are required.
- Step 2:** Fundamental amplitude estimation: The expected healthy vibration amplitude at the rotational frequency ( $1\times$ ) is calculated using the regression Equation (13). The 95% confidence bounds, derived from 2660 measurements across 98 configurations encompassing multiple coupling types and loading conditions, define the permissible amplitude range that captures real-world operational variability.
- Step 3:** Harmonic amplitude calculation: Amplitudes for the  $2\times$  and  $3\times$  harmonics are computed as proportions of the fundamental amplitude, constrained by established vibration analysis practice [30]:  $A_{2x} = r_2 \cdot A_{1x}$  where  $r_2 \in [0, 0.5]$ , and  $A_{3x} = r_3 \cdot A_{2x}$  where  $r_3 \in [0, 0.5]$ . Supply frequency components (50 Hz and 100 Hz) are included at amplitudes consistent with healthy motor electromagnetic behaviour.
- Step 4:** Stochastic sampling: For each synthetic signal, amplitude values are randomly sampled within the established bounds using uniform distributions. This stochastic process ensures that each generated signal exhibits unique characteristics while remaining statistically consistent with the experimentally observed healthy motor population.
- Step 5:** Frequency-domain construction: FFT indices corresponding to each frequency component are computed based on the specified sampling frequency. The frequency-domain representation is populated with the sampled amplitudes at the appropriate indices.
- Step 6:** Time-domain transformation: The synthetic frequency-domain signal is transformed via inverse FFT to obtain the time-domain signal of specified duration.

This procedure enables the generation of an unlimited number of healthy vibration signals that are both reproducible (given identical input parameters and random seed) and representative of real-world variability, subject to the following assumptions:

- Dominant frequency:** For a healthy induction motor operating under direct-on-line conditions at or near synchronous speed, the dominant frequency in the frequency domain will be that of the rotational frequency of the motor, known as the fundamental frequency. This frequency peak is due to the normal mechanical unbalance in the motor and the rotation-related forces, as well as the magnetic field interactions. Thus, a dominant, stable, and relatively low amplitude peak will be observed at the rota-

tional frequency of all healthy induction motors. The exact amplitude at this frequency for each motor size is to be established during the experimental data analysis.

2. **Harmonics:** For a healthy induction motor, some harmonics of the rotational frequency will be present in the spectrum, which will most often only be the  $2\times$  harmonic, but the  $3\times$  harmonic might also appear under certain conditions. These are also due to mechanical rotation of motor components and should be lower than the fundamental frequency, with [30] specifying that the healthy amplitude of the  $2\times$  harmonic should typically not be greater than 50% of the amplitude at the fundamental frequency. The specification of the  $3\times$  harmonic follows the same logic as the  $2\times$  harmonic, with the assumption that the amplitude should be restricted to less than 50% of the  $2\times$  harmonic amplitude. Specifying  $2\times$  and  $3\times$  as functions of the  $1\times$  amplitude ensures that the amplitudes are not scaled unrealistically and also ensures that a wide enough range of amplitudes is included.
3. **Supply frequency:** The frequency of the electrical supply (50 Hz) and its  $2\times$  harmonic will always be present in the vibration spectrum of a healthy induction motor, irrespective of the type of loading and coupling. The reason why these components will be present in the healthy spectrum is because of the alternating magnetic forces in the stator field and the magnetic pull forces in the motor.

The assumptions underpinning synthetic signal generation were consistently validated across the 2660 experimental measurements encompassing multiple coupling configurations and loading conditions. However, applications involving variable frequency drives, significant partial-load operation, or coupling arrangements beyond those tested may require re-evaluation of the dominant frequency assumption and corresponding adjustment of the synthetic signal parameters.

### 3.4. Health Index Model Training

The development of the HI model is similar to the detailed description of the model developed in [20], with the main difference being the signal features used to build the model. The key features from both the time and frequency domains were manually extracted after the FFT was computed, applying the Hanning window to reduce spectral leakage. These features were then stored in a feature vector and thus each synthetic signal became a feature vector.

The final list of 18 extracted features is a combination of 15 features in the time domain and 3 features in the frequency domain and is listed in Tables 3 and 4.

**Table 3.** HI Time-domain modelling features.

| Time-Domain Signal Features  |                   |                         |                       |                  |
|------------------------------|-------------------|-------------------------|-----------------------|------------------|
| (1) Mean                     | (2) Median        | (3) Standard Deviation  | (4) Variance          | (5) Peak-to-Peak |
| (6) Mean Absolute Difference | (7) RMS           | (8) Mean Absolute Value | (9) Kurtosis          | (10) Skewness    |
| (11) Crest Factor            | (12) Shape Factor | (13) Impulse Factor     | (14) Clearance Factor | (15) Peak Value  |

**Table 4.** HI Frequency-domain modelling features.

| Frequency-Domain Signal Features |                      |                       |
|----------------------------------|----------------------|-----------------------|
| (16) Spectral Centroid           | (17) Spectral Spread | (18) Spectral Entropy |

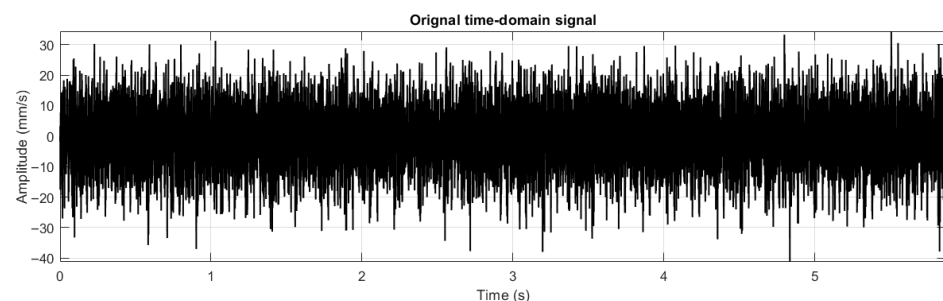
The other parameters of interest are as follows:

1. QBS —The number of healthy baseline samples to be used for training, selected as 30% of the number of healthy synthetic signals supplied to the training algorithm.
2.  $\delta$ —The log-scaling floor of the ECDF, to avoid  $\log(0)$  when scoring, and also to provide the lower limit for health degradation. Kept at the default of 0.01.
3. Variance threshold—The percentage of variance the PCA should retain, selected as 97% for this study.

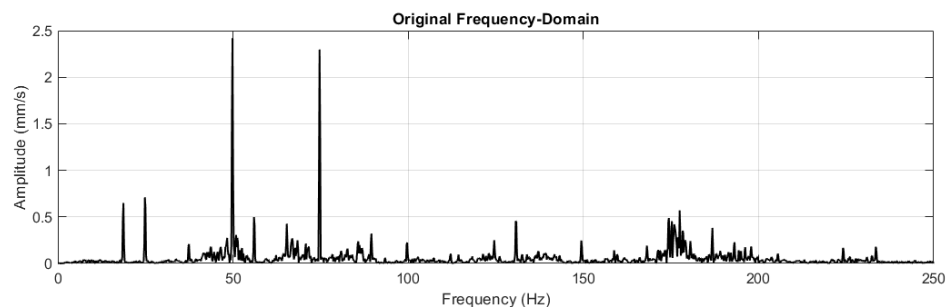
### 3.5. Fault Detection

The first step in any CBM strategy is to determine whether the monitored asset shows signs of deterioration—commonly termed fault detection in the FDI literature. An appropriate error analysis of the vibration signal of the asset can be used for this.

Due to the process of sampling and processing digital signals, the captured vibration signals contain redundant information. The original time domain of an example signal is shown in Figure 6 and the frequency domain in Figure 7, and they show the inherent “noise” in these signals.

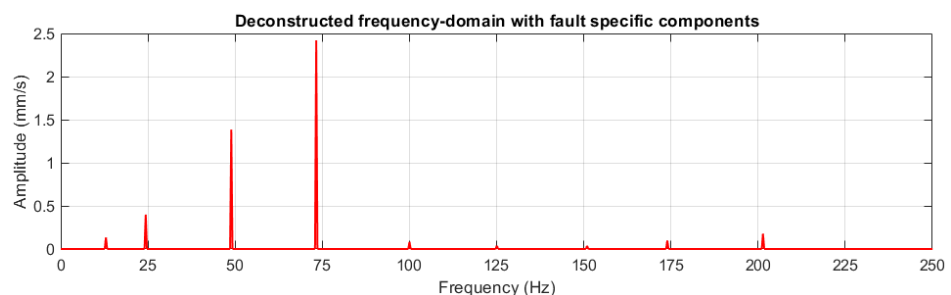


**Figure 6.** Original time-domain signal of the example signal to be evaluated.



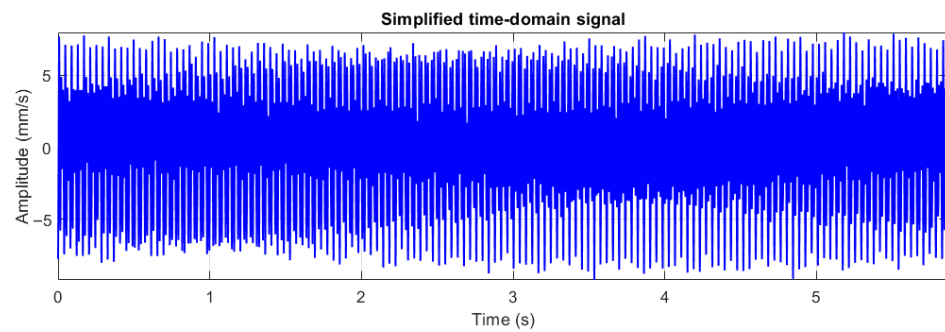
**Figure 7.** Original frequency domain of the example signal to be evaluated.

Through the application of band-pass filters in the frequency domain, it is possible to extract only the components of interest before calculating the HI, as shown in Figure 8.



**Figure 8.** Frequency domain of the example signal with fault-frequency components and their amplitudes.

The “clean” version of the frequency domain of the signal is then transformed to obtain a clean and filtered time-domain signal for the HI model, as shown in Figure 9.



**Figure 9.** New time-domain signal only containing the specified fault-frequency components from the example signal.

The filtered time-domain signal will then be evaluated by the HI model, where a non-parametric percentile-based threshold strategy is employed to classify the Health Index of the new signal against the trained healthy signals. Thresholds are required to establish the degradation state of the signal evaluated. Thresholds are therefore specified at three levels of deviation, which are at the 95th, 99th, and 99.9th percentile of the healthy HI distribution.

$$\text{Status} = \begin{cases} \text{Potential fault} & \text{if } \text{HI}_{\text{New}} > T_{999} \\ \text{Warning} & \text{if } T_{99} < \text{HI}_{\text{New}} \leq T_{999} \\ \text{Shift noted} & \text{if } T_{95} < \text{HI}_{\text{New}} \leq T_{99} \\ \text{Healthy} & \text{Otherwise} \end{cases} \quad (1)$$

These thresholds are mathematically defined in Equation (1) [20] (where  $T$  denotes “Threshold”), and indicate the following:

1. Shift noted —A shift is noted when the HI of the new signal lies outside of the 95th percentile of the computed HIs of the healthy baseline data;
2. Warning—A warning is issued when the HI of the new signal lies outside of the 99th percentile of the computed HIs of the healthy baseline data;
3. Potential fault detected—When the HI of the new signal is an outlier, thus, the computed HI lies outside of the 99.9th percentile of the computed HIs of the healthy baseline data, it is classified as a deviation in the signal and thus a potential fault in the induction motor.

### 3.6. Fault Isolation

Where the HI algorithm returns a non-healthy result, the underlying cause of this needs to be determined in order to provide industrial relevance from a maintenance management perspective. The diagnostic algorithm was developed to evaluate the harmonic amplitude relationships and the RMS energy to classify mechanical and electrical faults in induction motors. The inclusion of relationships, together with the absolute amplitude for diagnosis, is inspired by the work of [30]. A rule-based framework defines the logical conditions for the common failure mode using both amplitude evaluation through z-score comparison and amplitude ratios relative to the fundamental frequency.

To enhance diagnostic exclusivity, a probabilistic weighting model is applied that penalises overlapping or underdetermined modes and rewards uniquely satisfied conditions.

The requirements for each failure mode can be constructed from the following aspects.

1. **Elevated:** Requires that the amplitude be elevated compared to the average  $1 \times$  frequency amplitude for a the healthy component;
2. **Ratio:** The specified ratio of the fault-frequency component to the healthy  $1 \times$  amplitude;

3. **Dominance:** The amplitude of the specified component must be dominant (more than 2 times the amplitude of the average of the frequency components  $1\times$ ,  $2\times$ ,  $3\times$  and  $4\times$ ).

Table 5 specifies the fault-frequency conditions evaluated by the diagnostic algorithm. For each failure mode, all listed conditions must be satisfied simultaneously (logical AND). The *Elevated components* column specifies which harmonic frequencies must exceed their respective healthy thresholds. The *Ratio* column defines amplitude ratio requirements relative to the fundamental frequency ( $A_{1\times}$ ), expressed as  $A_{n\times}/A_{1\times}$ . The *Dominant components* column identifies which frequency component must exhibit the highest relative amplitude among the first four harmonics. A failure mode is considered a candidate diagnosis only when all specified conditions in its row are met.

The final list of requirements for each failure mode is provided in Table 5.

**Table 5.** Fault-frequency conditions for the diagnostic algorithm.

| Failure Mode         | Elevated Components                  | Ratios                            | Dominant Components |
|----------------------|--------------------------------------|-----------------------------------|---------------------|
| Unbalance            | $1\times$                            | $\frac{2\times}{1\times} < 0.15$  | $1\times$           |
| Misalignment         | $1\times, 2\times, 3\times$          | $\frac{2\times}{1\times} > 0.5$   | $2\times$           |
| Bent shaft           | $1\times, 2\times$                   | $\frac{2\times}{1\times} > 1$     | $2\times$           |
| Mechanical looseness | $1\times, 2\times, 3\times, 4\times$ | $\frac{0.5\times}{1\times} > 0.1$ | $3\times, 4\times$  |
|                      |                                      | $\frac{3\times}{1\times} > 0.2$   |                     |
|                      |                                      | $\frac{4\times}{1\times} > 0.2$   |                     |
|                      |                                      | $\frac{5\times}{1\times} > 0.2$   |                     |
|                      |                                      | $\frac{6\times}{1\times} > 0.2$   |                     |
|                      |                                      | $\frac{7\times}{1\times} > 0.2$   |                     |
| Structural looseness | $2\times, 3\times$                   | $\frac{8\times}{1\times} > 0.2$   | $3\times$           |
|                      |                                      | $\frac{2\times}{1\times} > 1$     |                     |
| Electrical fault     | $4\times$                            | $\frac{3\times}{1\times} > 1$     | $4\times$           |
|                      |                                      | $\frac{4\times}{1\times} > 0.5$   |                     |

When the algorithm detects that any specific failure mode is not dominant, it defaults to the failure mode “Investigate coupling health”. This occurs when the failure mode cannot be identified with confidence, specifying that the observed deviation might also be due to the influence of the system coupled to the induction motor.

### 3.7. Fault Classification

Suppose that a sample consists of nine harmonic components, as per Equation (2), and thresholds are required to determine which aspects are elevated relative to the healthy baseline.

$$A = [A_{0.5\times}, A_{1\times}, A_{2\times}, A_{3\times}, A_{4\times}, A_{5\times}, A_{6\times}, A_{7\times}, A_{8\times}] \quad (2)$$

An inter-harmonic ratio [30] to the  $1\times$  harmonic is calculated using Equation (3), where a small floor value of  $10^{-12}$  is used for the  $1\times$  amplitude to ensure numerical stability.

$$r_n = \frac{\text{Amplitude}(n\times)}{\max(|\text{Amplitude}(1\times)|, 10^{-12})} \quad (3)$$

For the  $1\times, 2\times, 3\times, 4\times$  harmonics, a percentile-based threshold is determined by

$$T_h = P_{ph}(A_{T,h}) \quad (4)$$

where  $P_{ph}(\cdot)$  denotes the  $p_h^{th}$  percentile of the training amplitudes at harmonic  $h$ , and the percentile is calculated as

$$p_h = 95 + 4.9 \times R_h \quad (5)$$

with  $R_h$  the relaxation factor defined as

$$R_h = [0.0, 0.2, 0.6, 1.0] \text{ for } 1\times, 2\times, 3\times, 4\times \quad (6)$$

The  $R_h$  progressively increases the percentile threshold for higher harmonics since higher harmonics will have much lower healthy amplitudes. For the sub-harmonic and higher harmonics ( $0.5\times$  and  $5\times$  to  $8\times$ ), elevation is determined instead by the amplitude ratios relative to the  $1\times$  component, with the ratios specified in Table 5. The use of percentiles ensures that 95% to 99% of healthy amplitudes fall below the threshold. By increasing the scale for higher harmonics, the algorithm prevents false positives caused by minor natural fluctuations. After calculating these thresholds, the algorithm compares the test amplitudes Equation (2) against the thresholds to determine which components are elevated.

$$E_h = \begin{cases} 1, & \text{if } A_h > T_h \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

An important note to make is the identification of a potential bearing failure. Since detailed diagnosis requires bearing type information, which is expected to be unknown, the average RMS vibration amplitude in the 200 Hz to 6 kHz band is used. The possibility of bearing failure is indicated if the experimental values fall outside of ( $3\sigma$ ) of the healthy baseline data.

### 3.8. Classification Rule Engine

The dominance ratio for each harmonic  $A_{h\times}$  is defined as the ratio of its amplitude to the mean amplitude of all other key harmonics:

$$D_{h\times} = \frac{A_{h\times}}{\frac{1}{n-1} \sum_{\substack{k=1 \\ k \neq h}}^n A_{k\times}} + \varepsilon \quad (8)$$

where:

- $D_{h\times}$  is the dominance ratio for harmonic  $h$ ;
- $n = 4$  for the first four harmonics;
- $\varepsilon$  is a small constant (approaching zero) to prevent division by zero.

For higher harmonics, the mean of the first four harmonics is used as a stable reference baseline. This introduces a relative weighting system to evaluate the prominence of a specific harmonic. Unlike percentile-based elevation, which identified harmonics exceeding normal limits, dominance essentially quantified the fault signature's intensity.

$$\text{Base Score} = \left( \frac{\text{Number of conditions met}}{\text{Total required conditions}} \right) \times 100 \quad (9)$$

Only using Equation (9) can lead to false positives, particularly when fault-frequency conditions for diagnosis overlap. To address this, additional factors are introduced such that

$$\text{Final Score} = \text{Base Score} \times P_{\text{ex}} \times B \quad (10)$$

where  $P_{\text{ex}}$  and  $B$  are discussed in the next section. The resulting final scores are then ranked so that the highest-scoring failure mode will be set to 100%.

### 3.8.1. Exclusivity Penalisation

Addresses situations where a fault defined by a simple singular rule, such as unbalance, is active, but many other unrelated conditions are also true. Applied to failure modes that require a small number of conditions ( $\leq 2$ ) to be met, the penalty factor reduces the score based on the fraction of other active conditions as follows:

$$P_{\text{ex}} = \max\left(1 - \gamma \cdot \frac{N_{\text{other}}}{N_{\text{total}}}, P_{\text{min}}\right) \quad (11)$$

where

- $\gamma = 0.8$  is the penalty strength;
- $N_{\text{other}}$  is the number of true conditions not required by the mode;
- $N_{\text{total}}$  is the number of true conditions;
- $P_{\text{min}} = 0.2$  is a floor to prevent the score from completely being eliminated.

### 3.8.2. Specificity Boost

The mechanism prioritises failure modes triggered by unique conditions, where uniqueness is defined as conditions exclusive to a specific failure mode. The boost factor increases scores proportionally to the fraction of uniquely satisfied conditions for each failure mode, calculated using Equation (12).

$$B = 1 + \alpha \cdot \left(\frac{N_{\text{unique met}}}{N_{\text{required}}}\right) \quad (12)$$

where

- $\alpha = 0.4$  is the boost strength and can be tuned;
- $N_{\text{unique met}}$  is the count of conditions met that are unique to that mode;
- $N_{\text{required}}$  is the total number of conditions required by the mode.

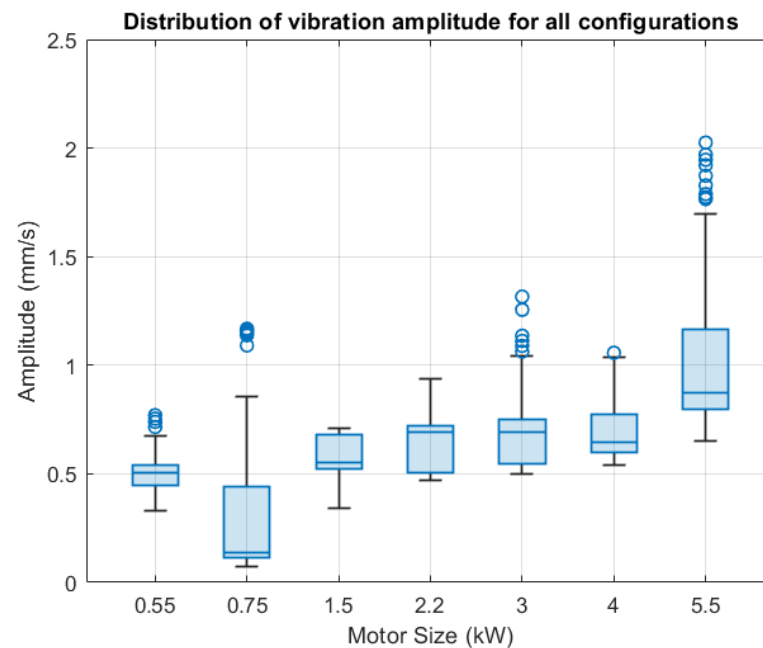
### 3.8.3. Superset Suppression

The final refinement suppresses low-scoring faults whose evidence is entirely subsumed by higher-scoring faults. If the conditions satisfied for fault B constitute a complete subset of those for fault A and fault A scores significantly higher, fault B is suppressed by a tunable factor of 0.5.

## 4. Results

### 4.1. Power–Vibration Correlation

Figure 10 shows the vibration amplitudes extracted at the rotational frequency component for all vibration signals of each motor size. Vibration amplitudes generally increased with motor size (also observed in previous analyses), with larger motors exhibiting greater variability and occasional outliers. Therefore, it is important that during the estimation of the amplitude, the final regression equation accounts for this increased spread of the vibration amplitude with the increased motor size.



**Figure 10.** Vibration amplitudes for the rotational frequency for all configurations and motor sizes.

A selection of regression fit types was evaluated to determine the most appropriate goodness of fit across the entire range of Class I induction machines. The results of these studies and the resultant average accuracy (predicted vs actual) are provided in Table 6.

**Table 6.** Accuracy for the regression fits under evaluation.

| Regression Fit    | 0.55 kW | 0.75 kW | 1.5 kW | 2.2 kW | 3 kW   | 4 kW   | 5.5 kW | Average |
|-------------------|---------|---------|--------|--------|--------|--------|--------|---------|
| Linear polynomial | 84.87%  | 67.16%  | 96.93% | 88.24% | 98.80% | 74.63% | 87.93% | 85.51%  |
| Exponential       | 89.83%  | 61.97%  | 95.91% | 85.33% | 97.23% | 77.76% | 84.39% | 84.63%  |
| 1-term power      | 75.35%  | 73.54%  | 98.13% | 93.89% | 94.00% | 73.30% | 93.72% | 85.99%  |
| 2-term power      | 87.35%  | 65.04%  | 95.87% | 86.63% | 99.50% | 75.53% | 86.39% | 85.19%  |
| Logarithmic       | 73.60%  | 70.24%  | 92.31% | 97.73% | 92.30% | 75.10% | 99.38% | 85.81%  |

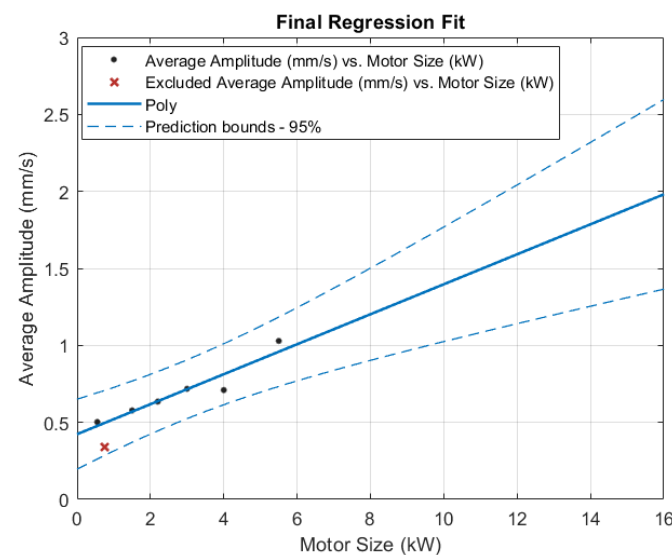
The final regression equation was obtained by excluding the data point of the 0.75 kW motor and recalculating the coefficients for the mean and the upper and lower confidence bounds. The removal of the 0.75 kW motor data point resulted in an R-Square value increasing to 0.903 from 0.878, and a reduction in the SSE from 0.0338 to 0.0162, indicating that the quality of fit was further improved. It should be noted that the 0.75 kW motor was excluded based on improvement of fit and not a formal statistical outlier determination.

Table 7 summarises the final coefficients of the regression equation (Equation (13)), with these coefficients used in synthetic signal generation. Figure 11 indicates the final regression fit in the MATLAB R2024a Curve Fitter application. The wider confidence interval ensures that predictions for healthy amplitude ranges consider the greater spread observed during analysis.

$$f(x) = p_1x + p_2 \quad (13)$$

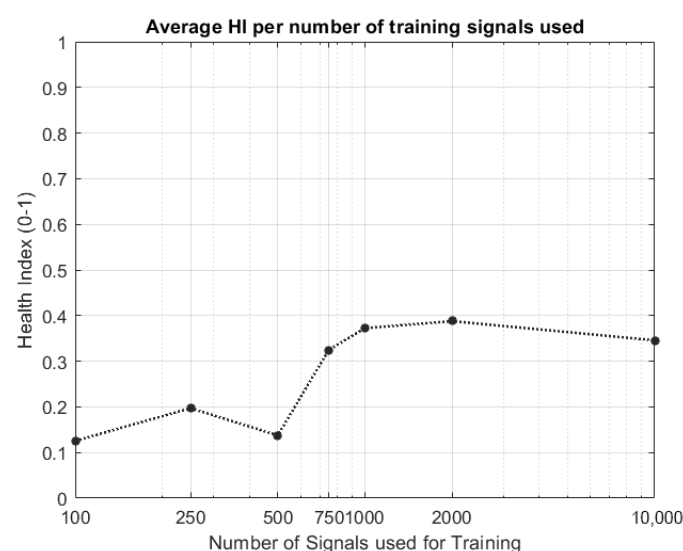
**Table 7.** Final regression equation coefficients.

|                        | p1     | p2     |
|------------------------|--------|--------|
| Upper confidence bound | 0.1416 | 0.5668 |
| Coefficient            | 0.0973 | 0.4236 |
| Lower confidence bound | 0.0530 | 0.2804 |

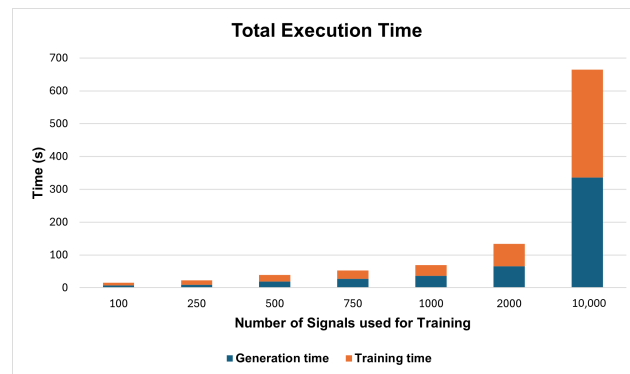
**Figure 11.** Final regression fit for estimating  $1\times$  vibration amplitudes of healthy Class I motors.

#### 4.2. HI Model Training

The prediction accuracy of the HI model is dependent on the QBS parameter and the number of healthy synthetic signals supplied to the model for training. Sweeping the number of healthy synthetic signals from 100 to 10,000 signals indicates an optimal point at the 2000 signal mark, beyond which overfitting effects become evident; see Figure 12.

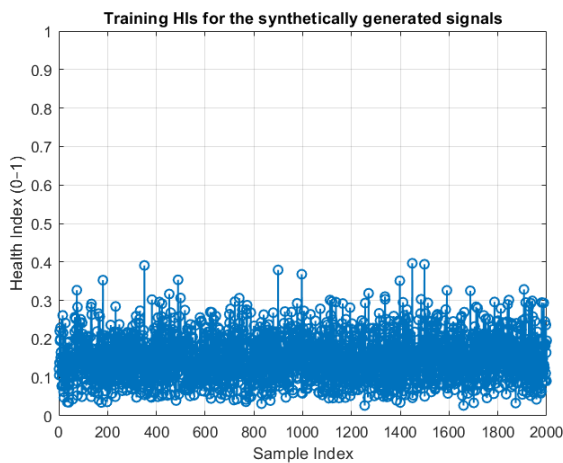
**Figure 12.** Average HI for different number of training signals used.

Additionally, the healthy synthetic signals used for training require generation, and a comparison of the generation vs training time (Figure 13) indicates exponential growth at the 10,000 signal point. Therefore, based on Figures 12 and 13, the number of healthy synthetic signals supplied was set at 2000, where the QBS is 30% of this.

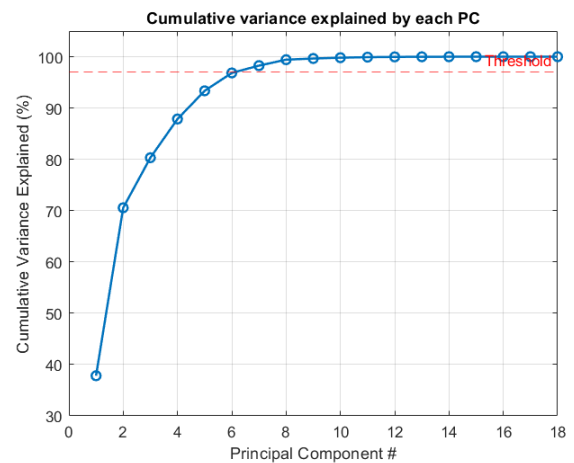


**Figure 13.** Computational cost (time) for different number of training signals used.

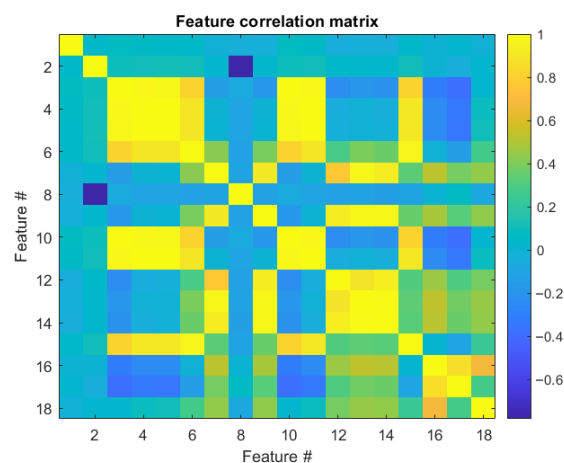
After model training using a 3 kW motor as an illustrative example, the following results were obtained: Figure 14a provides the HI of the trained signals and Figure 14b verifies that the PCA aligns with the selected parameter. Figure 14c further illustrates the correlation matrix of the selected features. The colour scale ranges from  $-1$  (dark blue), indicating a strong negative correlation, to  $+1$  (yellow), indicating a strong positive correlation, between the respective features.



(a)



(b)



(c)

**Figure 14.** General results after the HI model training algorithm was activated. (a) Health Index of the synthetically generated training signals. (b) Cumulative variance explained by the principal components. (c) Correlation matrix between selected features.

Several features exhibit moderate to strong positive correlations, indicating that some extracted features capture similar signal characteristics, resulting in a degree of redundancy within the feature set. However, there are also some regions with lower or near-zero correlations, which suggests that many features still provide unique and complementary diagnostic information about the signals. However, the training algorithm was implemented so that highly correlated features are removed from the HI computation before training and therefore do not cause redundancy during training.

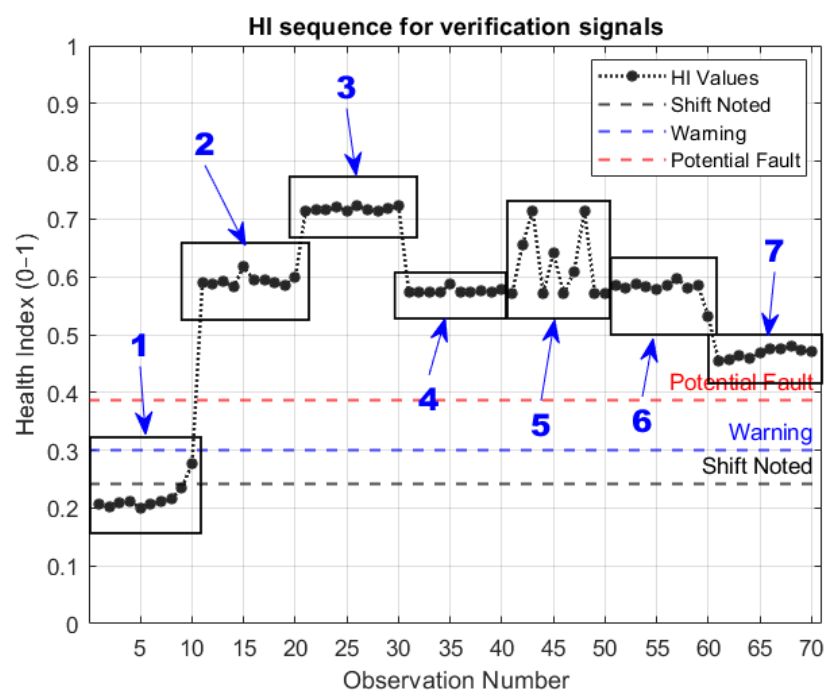
#### 4.3. Detection and Diagnostic Accuracy

In order to evaluate the model performance, known failure modes must be excited and submitted to the HI model for verification purposes. The applicable failure modes are indicated in Table 8.

**Table 8.** Vibration amplitudes specified for each failure mode—used for model verification.

|                        | 0.5× | 1×   | 2×   | 3×   | 4×   | 5×   | 6×   | 7×   | 8×   |
|------------------------|------|------|------|------|------|------|------|------|------|
| 1—Healthy              | 0    | 0.60 | 0.15 | 0.05 | 0.25 | 0    | 0    | 0    | 0    |
| 2—Unbalance            | 0    | 5.00 | 0.15 | 0.05 | 0.25 | 0    | 0    | 0    | 0    |
| 3—Misalignment         | 0    | 5.00 | 3.50 | 1.75 | 0.25 | 0    | 0    | 0    | 0    |
| 4—Bent shaft           | 0    | 5.00 | 7.50 | 0.10 | 0.25 | 0    | 0    | 0    | 0    |
| 5—Mechanical looseness | 1.25 | 5.00 | 2.00 | 1.50 | 1.25 | 1.25 | 1.20 | 1.15 | 1.05 |
| 6—Structural looseness | 0    | 0.60 | 2.40 | 3.00 | 0.25 | 0    | 0    | 0    | 0    |
| 7—Electrical fault     | 0    | 0.60 | 0.15 | 0.05 | 2.50 | 0    | 0    | 0    | 0    |

For each failure mode, ten HI values were computed, corresponding to the noise level added, which are 40, 30, 25, 20, 15, 10, 5, 0, −5, and −10 dB, respectively. The text arrows with the numbers (1–7) in Figure 15 indicate the calculated HI values for each failure mode. These numbers correspond to the numbers assigned to each failure mode, as presented in Tables 8 and 9.



**Figure 15.** Computed HI values for all verification signals and failure conditions.

To show the performance of diagnostic accuracy, a confusion matrix is created. This confusion matrix, presented in Table 9, illustrates the classification performance of the diagnostic algorithm under the different conditions evaluated. The diagonal elements represent the correctly classified samples, while the off-diagonal values indicate misclassified cases. The values within the cells show how many of the ten signals at each noise level evaluated were accurately diagnosed for each failure mode, whereas the numbers in the heading rows correspond to those of each failure mode. The accuracy of the diagnostics for each failure mode is also provided.

**Table 9.** Confusion matrix of diagnostic accuracy.

|                        | 1  | 2  | 3  | 4 | 5  | 6  | 7  | Accuracy |
|------------------------|----|----|----|---|----|----|----|----------|
| 1—Healthy              | 10 |    |    |   |    |    |    | 100%     |
| 2—Unbalance            |    | 10 |    |   |    |    |    | 100%     |
| 3—Misalignment         |    |    | 10 |   |    |    |    | 100%     |
| 4—Bent shaft           |    |    |    | 8 | 2  |    |    | 80%      |
| 5—Mechanical looseness |    |    |    |   | 10 |    |    | 100%     |
| 6—Structural looseness |    |    |    |   |    | 10 |    | 100%     |
| 7—Electrical fault     |    |    |    |   |    |    | 10 | 100%     |

In general, the model achieved high diagnostic precision, and almost all conditions were diagnosed with 100% accuracy across all SNRs. A minor misclassification is observed for a bent shaft failure mode, where the evaluated signals at noise levels  $-5$  and  $-10$  dB were incorrectly diagnosed as mechanical looseness. These misclassifications are likely due to the overlap of diagnostic fault-frequency components or the model not being able to accurately extract the amplitudes at the specific frequency components due to the spectrum noise. Despite this, the overall performance demonstrates that the model accurately diagnoses the most common health states of the induction motors, confirming its strong diagnostic capability.

The developed HI model demonstrates stable, repeatable, and interpretable performance in all verification stages. The model produced consistent HI results under repeated evaluation of identical signals, confirming the computational robustness. Applying vibration signals of known health and failure modes, the model accurately detected healthy, transitional, and degraded states, achieving high detection accuracy. In evaluating diagnostic performance, the model also consistently diagnosed the different failure modes, even when additional noise was added to the signals to be evaluated, verifying their diagnostic accuracy and performance. In general, the results prove that the developed HI model provides a reliable quantitative indicator of machine health.

#### 4.4. Industrial Validation

Industrial validation of the proposed methodology was conducted from 18 August 2025 to 8 September 2025 in an operational grain silo plant, with 10 critical assets being inspected 30–50 times during this period. A summary of the implementation is provided in Table 10. Physical inspections validated algorithmic results, with the majority of motors classified as healthy by the HI model, and four unhealthy diagnoses were made.

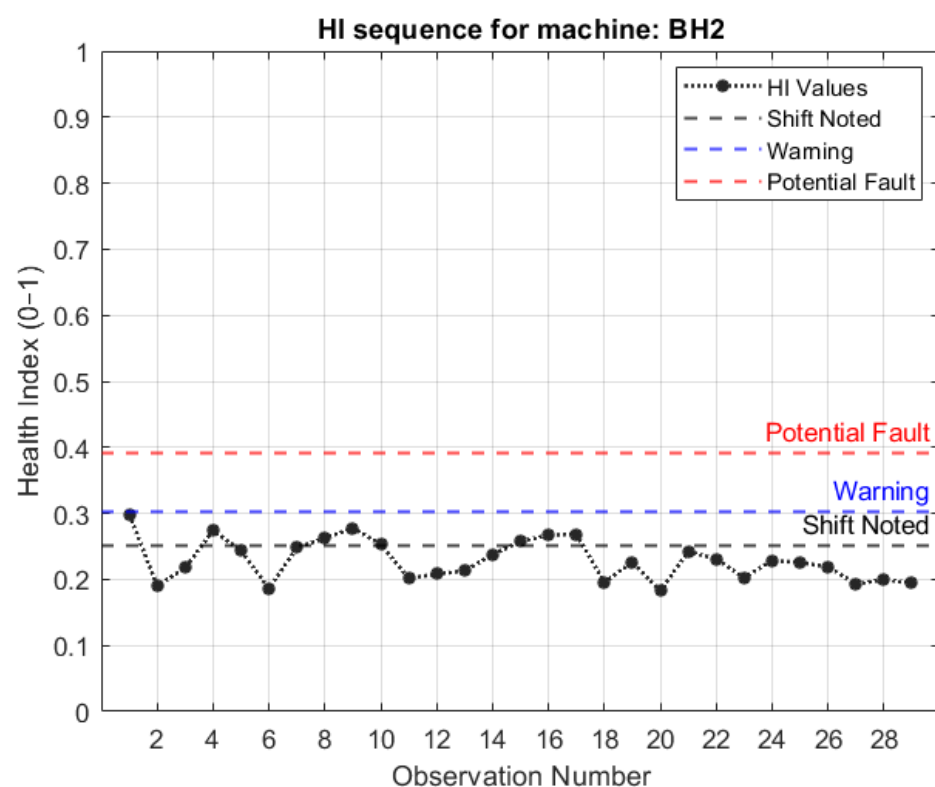
The detailed validation for the faulty assets, BH7, VB1, and VB9, is described below. The situation of VB2 is also described, which was flagged as healthy, but with a shift in condition. Only in cases where the algorithm provided an unhealthy result is the full diagnostic algorithm output included for transparency.

**Table 10.** Summary of the results for each asset on which the HI model was implemented.

| Equipment ID | Within Healthy Range | Shift Noted | Warning | Potential Fault | Diagnostic Result    |
|--------------|----------------------|-------------|---------|-----------------|----------------------|
| BH2          | 21                   | 8           | 0       | 0               | Healthy              |
| BH7          | 0                    | 0           | 1       | 46              | Electrical fault     |
| KLL5         | 22                   | 17          | 1       | 0               | Healthy              |
| KLL6         | 34                   | 6           | 0       | 0               | Healthy              |
| KT1          | 0                    | 0           | 30      | 5               | Healthy              |
| KT2          | 0                    | 0           | 22      | 15              | Healthy              |
| VB1          | 0                    | 0           | 0       | 50              | Structural looseness |
| VB2          | 0                    | 40          | 0       | 0               | Healthy              |
| VB9          | 0                    | 0           | 0       | 40              | Unbalance            |
| SW2          | 0                    | 0           | 28      | 5               | Healthy              |

#### Bucket Elevators (BH2, BH7)

The HI sequence of BH2 (see Figure 16) shows that the induction motor generally has a healthy condition, with intermittent periods of elevated risk. Most HI values fluctuate approximately between 0.2 and 0.3, all of which are below the “Warning” threshold of the baseline feature statistics of the trained model. Operation in the “Shift noted” region suggests only a minor deviation from the healthy baseline. The trend shows no significant sustained deterioration or acute spikes in the higher regions, confirming that BH2 is operating in a healthy condition throughout the measurement sequence.

**Figure 16.** Calculated Health Index trend of BH2.

Observing the HI trend of BH7 (see Figure 17), it is seen that the calculated HI values fall in the range of 0.45 and 0.55, somewhat above the failure threshold of the healthy trained model. This indicates that a potential failure mode is present in the induction motor

throughout the observation period. The immediate and consistent crossing of the failure threshold suggests that a fault was present in the induction motor prior to the start of the monitoring sequence. The trend currently does not show an accelerating sharp trajectory; however, the high sustained HI value suggests investigation and intervention to prevent further degradation.

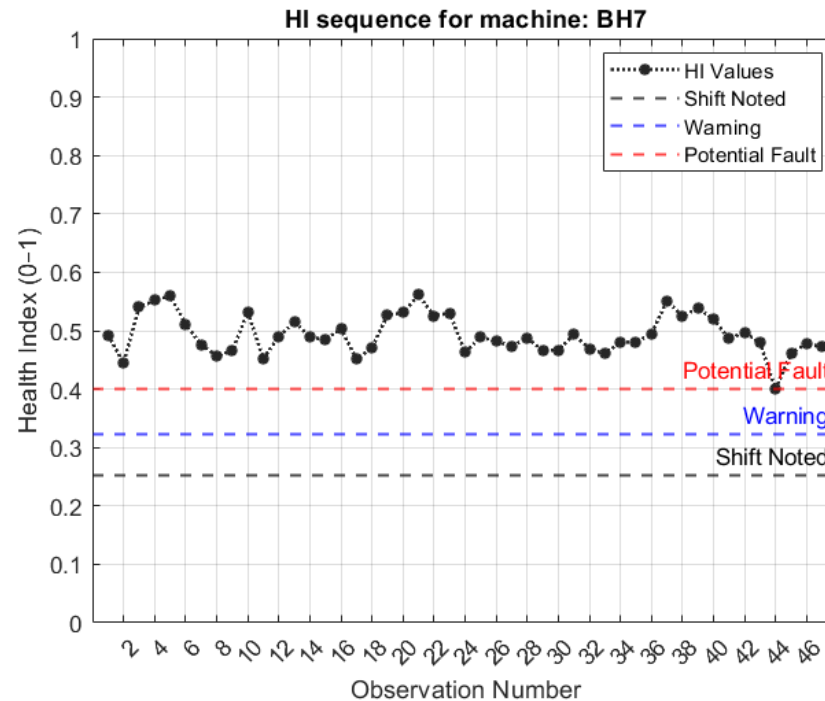


Figure 17. Calculated Health Index trend of BH7.

Figure 18 provides the output result of the diagnostic algorithm for BH7, indicating the potential failure mode as an electrical fault. Figure 19 shows the frequency-domain of the vibration signal of the last observation on BH7. Inspecting the amplitude at the rotational frequency ( $\pm 25$  Hz) in the frequency-domain, it is seen that the amplitude is within the healthy range for a 15 kW induction motor. However, it is also evident that the amplitudes of the frequency components  $2\times$  and  $3\times$  are elevated, compared to the specified healthy baseline.

```

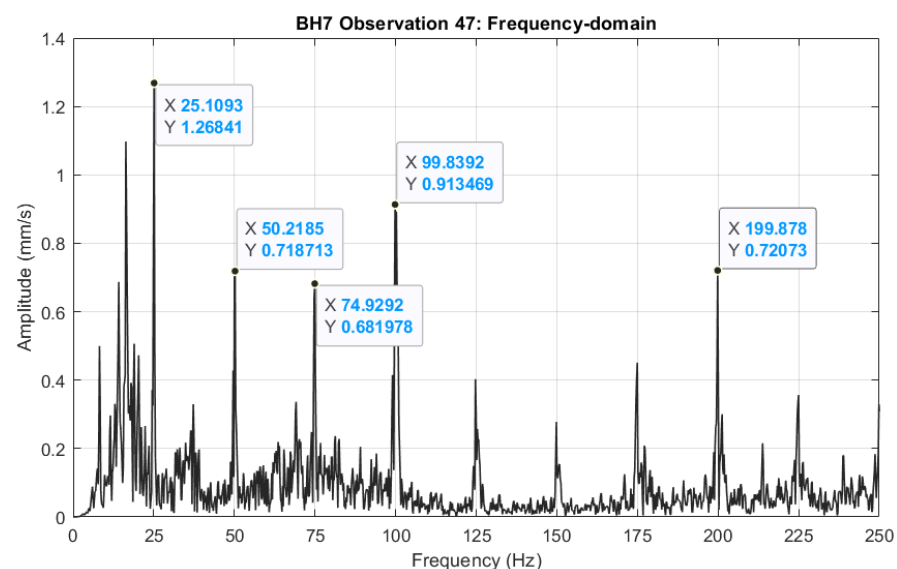
-----
Failure Mode Probabilities:
Unbalance           :    0.0% (met 0/3, penalty 1.00, uniq 0)
Misalignment        :   85.8% (met 3/5, penalty 1.00, uniq 1)
Bent Shaft          :    8.3% (met 1/4, penalty 1.00, uniq 0)
Mechanical Looseness :   80.0% (met 7/13, penalty 1.00, uniq 4)
Structural Looseness :   13.2% (met 2/5, penalty 1.00, uniq 0)
Bearing Failure     :    0.0% (met 0/1, penalty 0.20, uniq 0)
Electrical fault     :  100.0% (met 2/3, penalty 1.00, uniq 1)
Investigate coupling health :  36.1% (met 1/1, penalty 0.27, uniq 0)
-----
Most probable failure mode: Electrical fault (100.0%)
-----
RMS Frequency Range (from machine history): 0.0089
Specified RMS: 0.0100
Threshold (3x specified): 0.0300
Bearing RMS within acceptable limits.
-----

```

Figure 18. Diagnostic algorithm result for BH7.

A significant amplitude is also observed at 100 Hz, which is the  $2\times$  supply frequency component. When these observations are compared with the specifications for the diagnosis of failure modes, the result of the diagnostic algorithm in Figure 18, which is an electrical fault, is validated. This is mainly due to the spike in the amplitude at the 100 Hz component. The physical inspection of the motor was inconclusive in identifying the diagnosed failure mode, because the induction motor could not be opened to locate any possible electrical faults. The reason for this is because the motor was in operation and could not be disabled and opened for inspection, as this would have had a negative effect on the plant operations for that specific time period. Therefore, only theoretical validation of this diagnostic result was possible.

One could also argue that the spectrum in Figure 19 indicates a misalignment failure mode (elevated  $1\times$ ,  $2\times$  and  $3\times$ ); however, the requirement for elevated  $1\times$  is not met. The diagnostic results in Figure 18 show that this failure mode is the second most likely. Another failure mode which is also somewhat expected from a spectrum such as Figure 19 is mechanical looseness, which is characterised by multiple harmonic components of the rotational frequency. This is true for the spectrum in Figure 19; however, the specificity requirement for the additional component of  $0.5\times$  is not met. Thus, the most probable diagnosis is not mechanical looseness, and it is therefore specified as only the third most probable cause of failure in Figure 18.



**Figure 19.** Frequency domain of observation 47 measured on BH7.

#### 4.5. Conveyor Belts (VB1, VB2, VB9)

##### 4.5.1. Conveyor VB1

The calculated HI of VB1 reveals a state of severe and persistent degradation throughout all 50 observations, as seen in Figure 20. The HI values are consistently clustered around 0.58, and the minimal fluctuation indicates that the fault is stable, but severe, with the only deviation being some minor dips at some of the observations. Continuous high-level HI, well into the “Potential fault” region, strongly suggests that a major fully developed fault is and was present in the motor. The HI thus suggests that the machine is operating in an unhealthy state, with the diagnostic algorithm indicating structural looseness as the most probable cause of failure (see Figure 21). The VB1 result demonstrates the consistency of the model, where most of the 50 observations made had a very similar HI value, indicating the robustness and repeatability of the model.

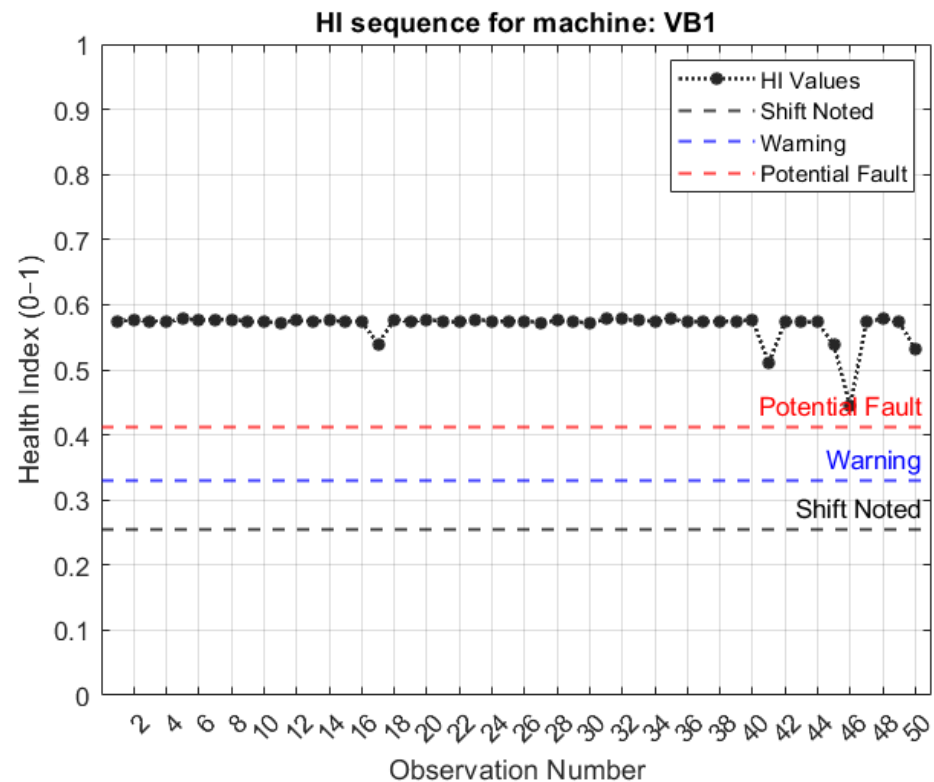


Figure 20. Calculated Health Index trend of VB1.

Investigating the results of the measurement of the last observation on VB1 Figure 22, a significant peak is observed at the  $3\times$  frequency component. The diagnosis is structural looseness, which is characterised by elevated  $2\times$  and  $3\times$ , with the unique requirement of having a dominant  $3\times$  amplitude, more than twice the amplitude of  $1\times$ . This is true for the spectrum in Figure 22, where the amplitude is almost 15 times that of the amplitude at the rotational frequency. However, the  $2\times$  component is not prominent in the frequency domain. This is the reason why “Inspect coupling health” also has an increased probability percentage, as seen in Figure 21. In conclusion, the result of structural looseness from the diagnostic algorithm is valid due to the dominant  $3\times$  component.

```

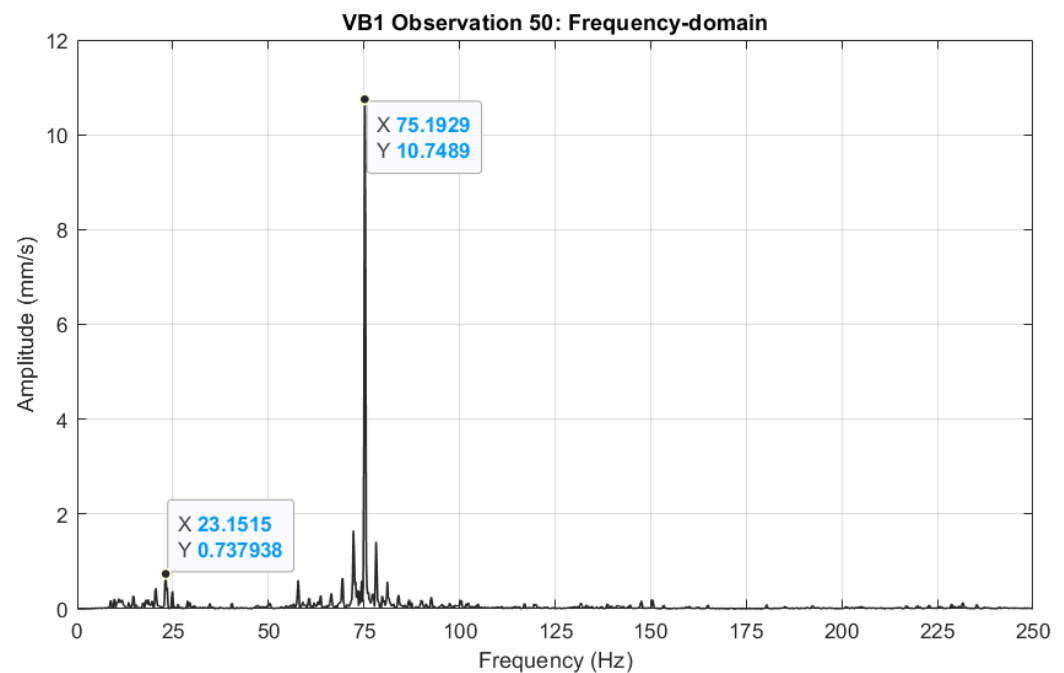
-----
Failure Mode Probabilities:
Unbalance           :    0.0% (met 0/3, penalty 1.00, uniq 0)
Misalignment        :    7.7% (met 1/5, penalty 1.00, uniq 0)
Bent Shaft          :    0.0% (met 0/4, penalty 1.00, uniq 0)
Mechanical Looseness :   64.8% (met 5/13, penalty 1.00, uniq 3)
Structural Looseness :  100.0% (met 3/5, penalty 1.00, uniq 1)
Bearing Failure     :    0.0% (met 0/1, penalty 0.20, uniq 0)
Electrical fault     :    0.0% (met 0/3, penalty 1.00, uniq 0)
Investigate coupling health :  48.5% (met 1/1, penalty 0.31, uniq 0)
-----

Most probable failure mode: Structural Looseness (100.0%)
-----

RMS Frequency Range (from machine history): 0.0026
Specified RMS: 0.0100
Threshold (3x specified): 0.0300
Bearing RMS within acceptable limits.
-----

```

Figure 21. Diagnostic algorithm result for VB1.



**Figure 22.** Frequency domain of observation 50 measured on VB1.

Physical inspection of this motor provided a positive outcome when it was found that the safety structure surrounding the VB1 V-belt pulley coupling was loose. This structural looseness was captured by the vibration sensor, thus providing the result seen in Figure 22. The result thus validates the diagnostic result, showing an accurate diagnostic performance of the diagnostic algorithm, even during the implementation phase.

#### 4.5.2. Conveyor VB2

The HI trend of VB2 in Figure 23 shows a slightly elevated condition in all 40 observations, with all HI values falling within the “Shift noted” region. Minimal fluctuation of the HI values is observed and is tightly clustered slightly above 0.3. This smooth result of the calculated HI reiterates the repeatability of the model. The vibration measurements for each observation number exhibited similar feature statistics, demonstrating the model’s ability to accurately compute these statistics for comparable vibration signals. The slightly elevated HI, in the upper “Shift noted” region, indicates that the motor has some slight deviation from normal operation, which should be noted. Therefore, the evaluation of the HI trend of VB2 concludes that the motor is operating in a stable, slightly elevated but healthy condition.

#### 4.5.3. Conveyor VB9

The HI trend of VB9 in Figure 24 indicates an extremely severe and persistent state of degradation throughout all 40 observations. The HI fluctuates quite significantly between observations, which is again attributed to signal characteristics because of the severe deviation observed. The trend indicates that the induction motor has been in an advanced state of failure throughout the monitoring period, with a significant increase in HI at observation 11, indicating an increase in the severity of the fault. VB9 is operating at an unacceptable level of risk, requiring immediate attention and serious maintenance intervention. Figure 25 shows the result of the diagnostic algorithm, with unbalance being diagnosed as the most probable cause of failure, and some other possible failure modes also indicated.

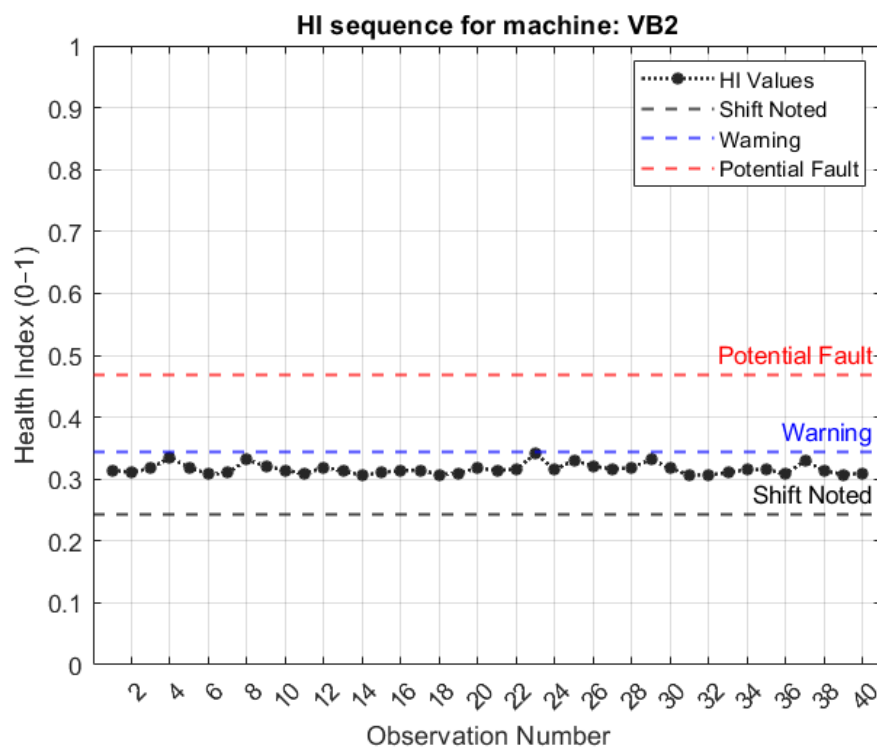


Figure 23. Calculated Health Index trend of VB2.

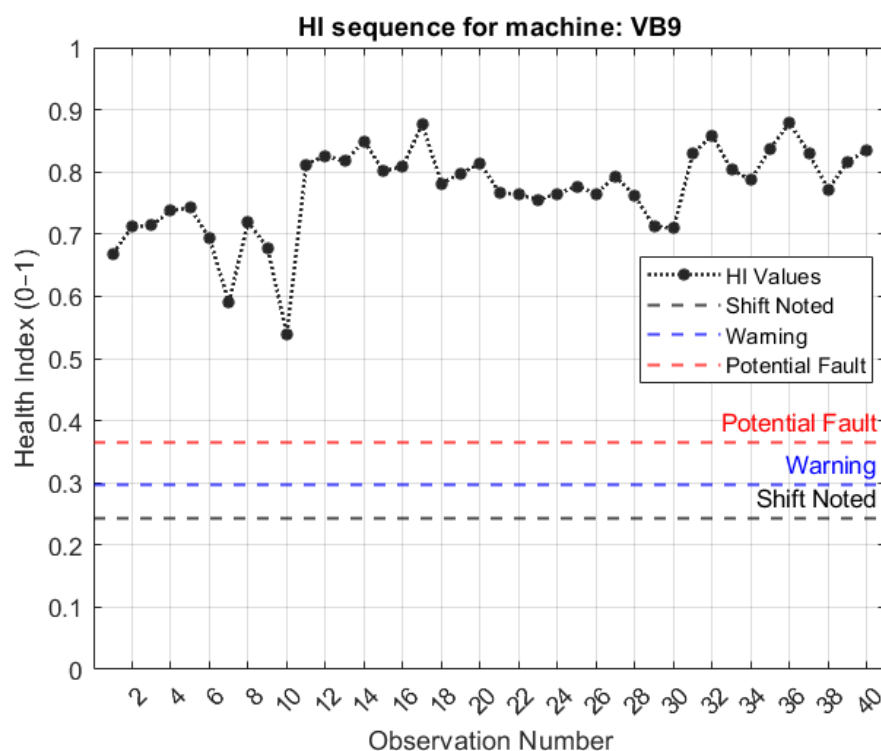


Figure 24. Calculated Health Index trend of VB9.

Observing the spectrum in Figure 26, it is evident that there is a significant peak at the rotational frequency. The amplitude is around 13 mm/s, which is significantly higher than the healthy range [3] for an 11 kW motor such as VB9. The requirement for identification of unbalance is observed by an elevated  $1\times$  amplitude evident in Figure 26, and therefore the diagnostic result can theoretically be validated as true.

```

-----
Failure Mode Probabilities:
Unbalance           : 100.0% (met 3/3, penalty 1.00, uniq 2)
Misalignment        : 47.4% (met 3/5, penalty 1.00, uniq 0)
Bent Shaft          : 39.5% (met 2/4, penalty 1.00, uniq 0)
Mechanical Looseness : 24.3% (met 4/13, penalty 1.00, uniq 0)
Structural Looseness : 15.8% (met 2/5, penalty 1.00, uniq 0)
Bearing Failure      : 0.0% (met 0/1, penalty 0.20, uniq 0)
Electrical fault     : 26.3% (met 1/3, penalty 1.00, uniq 0)
Investigate coupling health : 0.0% (met 0/0, penalty 0.20, uniq 0)
-----

Most probable failure mode: Unbalance (100.0%)
-----

RMS Frequency Range (from machine history): 0.0050
Specified RMS: 0.0100
Threshold (3x specified): 0.0300
Bearing RMS within acceptable limits.
-----

```

Figure 25. Diagnostic algorithm result for VB9.

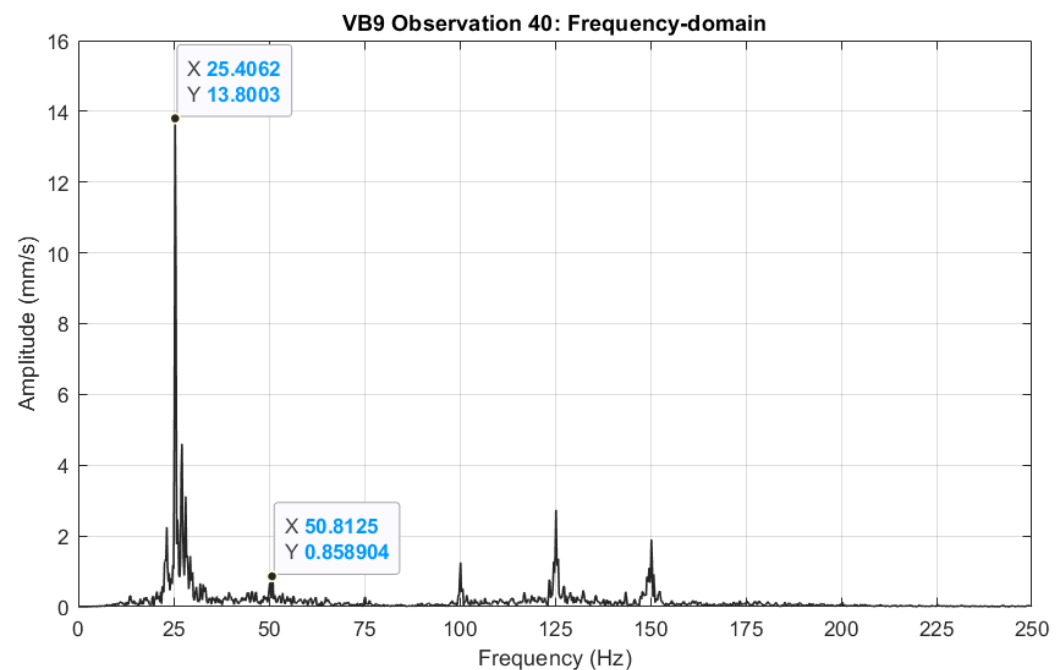


Figure 26. Frequency domain of observation 40 measured on VB9.

The results of an analysis of historical maintenance records indicated that the bearings of VB9 had to be replaced in the past, which could indicate that not enough attention was paid to the problem of unbalance failure. During the current inspection, it was found that the operating environment of this motor is harsh, with a lot of dust accumulating in the motor cooling fan, which could also cause possible unbalance. This inspection indicates that there are numerous factors that can possibly contribute to the unbalance in the rotor, thereby validating the results of the diagnostic algorithm in practice.

## 5. Conclusions

This study addressed critical gaps in the health index methodologies for industrial induction motor monitoring, specifically data scarcity and unknown initial health status, when characterising legacy equipment in operational facilities.

- An empirical correlation between motor power rating and healthy vibration amplitude at rotational frequency for Class I induction motors was developed, ex-

pressed as  $f(x) = 0.0973x + 0.4236$  ( $R^2 = 0.903$ ), based on 2660 measurements across 98 configurations, establishing predictive relationships absent in existing standards [3];

- A synthetic baseline generation algorithm was implemented that produces statistically representative healthy vibration spectra without physical measurements, addressing the cold-start problem in industrial condition monitoring [6,34].
- Validation of the JPCCED-HI framework [20] was achieved for complete motor-machine systems with industrial coupling configurations, extending beyond previously published validation cases using isolated components, to confirm its applicability under operational industrial conditions;
- A rule-based diagnostic algorithm was created that incorporates amplitude ratios, dominance weighting, and probabilistic scoring, achieving a verification accuracy of 97.1% and a successful identification of industrial faults, validated through spectrum analysis and physical inspection.

### 5.1. Practical Implications

The framework addresses the cold-start problem in industrial condition monitoring, whereby facilities initiate condition-based maintenance without knowing the health status of the equipment [6,34]. Through experimental characterisation and regression analysis, synthetic baselines are established for Class I induction motors independent of historical operational data from monitored equipment. The regression equation predicts healthy vibration amplitude ranges for Class I motors based solely on power rating, eliminating extensive baseline measurements prior to the implementation of condition monitoring. For legacy equipment of unknown condition, the synthetic baseline generation algorithm enables the immediate initiation of health monitoring.

The results of the implementation demonstrate the effective identification of fault development before machine failure. The diagnosis of the case study identified three motors that were operational, but the facility personnel were unaware of developing faults. Historical maintenance records revealed symptom-based interventions rather than root cause mitigation; VB9 exhibited 22 maintenance activities over four years, predominantly bearing replacements, while underlying unbalance remained unaddressed. The diagnostic algorithm provides actionable fault mode identification, enabling pre-emptive corrective action planning and spare parts procurement, thereby reducing maintenance costs and downtime. Rule-based probabilistic weighting ensures interpretable outputs validated against established vibration analysis practices [30,39].

Computational efficiency supports practical deployment. Training on 2000 synthetic signals optimally balances accuracy and computational cost. The JPCCED-HI model requires 4.24 ms inference time [20], enabling near real-time health assessment suitable for plant control system integration.

### 5.2. Limitations

The applicability of the framework is bounded by two primary data constraints: (i) the extent of motor power ratings that can be experimentally characterised (0.55–5.5 kW), which limits validated predictions to this range, with extrapolation beyond requiring further verification; and (ii) the single-axis vertical measurement orientation, which may reduce diagnostic sensitivity for fault types with dominant axial or horizontal signatures.

The regression equation is derived from seven motor sizes (0.55 kW to 5.5 kW), with extrapolation across the Class I range ( $\leq 15$  kW). Although demonstrating strong statistical validity ( $R^2 = 0.903$ ), measurements on 7.5 kW to 15 kW motors would strengthen extrapolated prediction confidence.

The experimental motors consisted exclusively of WEG W22 IE3 Premium Efficiency four-pole machines. Therefore, the influences of manufacturer, efficiency class, and pole count vibration–power relationships remain uncharacterised in the general sense. Manufacturer effects are presumed negligible given consistent fundamental mechanical vibration generation principles but still require empirical validation.

The coupling configurations were limited to chain-sprocket and V-belt pulley arrangements typical of grain handling applications. Direct couplings, gearboxes, and alternative transmission configurations were not tested. The applicability of the regression equation to these configurations remains uncertain, although inclusion of loaded and unloaded conditions across multiple coupling types suggests a reasonable generalisability.

Vibration measurements were acquired exclusively in the vertical direction at drive-end bearing locations. Axial and horizontal measurements, and non-drive-end measurements, were excluded. This limits the diagnostic capability for failure modes that manifest primarily in unmeasured directions, such as axial misalignment. Multi-axis measurement would enhance detection sensitivity and diagnostic discrimination.

The diagnostic algorithm was verified using synthetically generated fault signals rather than measurements from motors with confirmed physical faults. Although industrial implementation diagnoses were validated by physical inspection, systematic validation against comprehensive known-fault datasets would provide rigorous diagnostic accuracy confirmation across all failure modes.

Industrial implementation occurred at a single grain silo facility over three weeks. Longer-term monitoring would be required to observe fault progression and validate HI trending for remaining useful life prediction. Multi-facility implementation would confirm framework transferability across operational environments.

### 5.3. Future Work

Extension of the experimental database to include direct couplings and gearbox configurations would broaden the applicability of the regression equation beyond the investigated chain-sprocket and V-belt arrangements. Measurements on 7.5 kW to 15 kW motors would provide direct validation of extrapolated predictions for larger Class I motors.

Multi-axis vibration measurement (vertical, horizontal, axial) at drive-end and non-drive-end bearing locations would enhance diagnostic discrimination, particularly for failure modes such as misalignment exhibiting directional dependencies. This would enable a comprehensive validation of the diagnostic algorithm against expanded fault signatures.

Investigation of the applicability of the framework to motors beyond tested specifications by including different manufacturers, efficiency classes, and pole configurations would establish the generalisability of the power–vibration correlation. Extension to Class II and Class III motors would significantly expand industrial relevance.

Longer-term monitoring studies are required to validate HI trending for remaining useful life prediction and to observe fault progression from incipient detection to failure. Integration with the plant control system for automated monitoring and alert generation represents a practical development path towards the implementation of full condition-based maintenance.

Evaluation of machine learning approaches for diagnostic classification, trained on expanded confirmed-fault datasets, may improve diagnostic accuracy and enable fault severity quantification beyond current binary healthy/faulty classification.

Comparative evaluation against established HI construction methods and diagnostic frameworks under controlled conditions where baseline data availability permits parallel implementation would quantify the trade-offs between the proposed approach and conventional, data-driven, or hybrid alternatives.

Real-time industrial deployment presents additional considerations warranting investigation. Computational cost analysis (Figure 13) indicates that synthetic signal generation and HI model training are one-time processes completed offline; subsequent health assessment requires only FFT computation and feature extraction, which are computationally inexpensive and suitable for edge deployment. Sensor placement optimisation—including the trade-off between single-axis drive-end measurements (as employed in this study) and multi-axis, multi-location configurations—would inform cost–benefit decisions for industrial practitioners balancing diagnostic capability against instrumentation investment.

Integration with predictive maintenance frameworks requires (i) longitudinal monitoring to capture degradation trajectories; (ii) correlation of HI trends with remaining useful life through run-to-failure datasets; and (iii) prognostic models linking HI rate-of-change to maintenance scheduling decisions.

**Author Contributions:** Conceptualisation, J.-H.K. and H.M.; methodology, software, validation, formal analysis, investigation, resources, data curation, D.S.; writing—original draft preparation, H.M.; writing—review and editing, J.-H.K., D.S., and A.S.; visualisation, D.S.; supervision, J.-H.K., H.M., and A.S.; project administration, J.-H.K.; funding acquisition, J.-H.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Senwes Grainlink, Klerksdorp, South Africa.

**Data Availability Statement:** The supporting data for this study is available on request from the authors.

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

## Nomenclature

### Roman Symbols

|                           |                                                                  |
|---------------------------|------------------------------------------------------------------|
| $A$                       | Amplitude vector of harmonic components                          |
| $A_{n \times}$            | Vibration amplitude at $n \times$ harmonic frequency (mm/s)      |
| $A_{T,h}$                 | Training amplitudes at harmonic $h$                              |
| $B$                       | Specificity boost factor                                         |
| $D_{h \times}$            | Dominance ratio for harmonic $h$                                 |
| $E_h$                     | Elevation indicator (binary)                                     |
| $f(x)$                    | Regression function for power–vibration correlation              |
| $h$                       | Harmonic index                                                   |
| $n$                       | Number of harmonics considered                                   |
| $N_{\text{other}}$        | Number of true conditions not required by fault mode             |
| $N_{\text{required}}$     | Total conditions required by fault mode                          |
| $N_{\text{total}}$        | Total number of true conditions                                  |
| $N_{\text{unique met}}$   | Count of uniquely satisfied conditions                           |
| $p_1, p_2$                | Linear regression coefficients                                   |
| $p_h$                     | Percentile level for harmonic $h$                                |
| $P_{\text{ex}}$           | Exclusivity penalty factor                                       |
| $P_{\text{min}}$          | Minimum penalty floor (0.2)                                      |
| $r_n$                     | Inter-harmonic amplitude ratio ( $A_{n \times} / A_{1 \times}$ ) |
| $R^2$                     | Coefficient of determination                                     |
| $R_h$                     | Relaxation factor for harmonic $h$                               |
| $T_h$                     | Threshold at harmonic $h$                                        |
| $T_{95}, T_{99}, T_{999}$ | Health Index percentile thresholds                               |
| $x$                       | Motor power rating (kW)                                          |

## Greek Symbols

|               |                                        |
|---------------|----------------------------------------|
| $\alpha$      | Specificity boost strength (0.4)       |
| $\gamma$      | Exclusivity penalty strength (0.8)     |
| $\delta$      | Log-scaling floor for ECDF (0.01)      |
| $\varepsilon$ | Small constant for numerical stability |
| $\sigma$      | Standard deviation                     |

## Abbreviations

|      |                                            |
|------|--------------------------------------------|
| CBM  | Condition-based maintenance                |
| DE   | Drive end                                  |
| ECDF | Empirical cumulative distribution function |
| FFT  | Fast Fourier transform                     |
| HI   | Health Index                               |
| PCA  | Principal component analysis               |
| QBS  | Quantile-based sample size                 |
| RMS  | Root mean square                           |
| RUL  | Remaining useful life                      |
| SSE  | Sum of squared errors                      |

## References

1. Bengtsson, M. On Condition Based Maintenance and Its Implementation in Industrial Settings. Ph.D. Thesis, Mälardalen University, Västerås, Sweden, 2007. [\[CrossRef\]](#)
2. Mołęda, M.; Małysiak-Mrozek, B.; Ding, W.; Sunderam, V.; Mrozek, D. From Corrective to Predictive Maintenance—A Review of Maintenance Approaches for the Power Industry. *Sensors* **2023**, *23*, 5970. [\[CrossRef\]](#) [\[PubMed\]](#)
3. ISO 20816-1:2016; Mechanical Vibration—Measurement and Evaluation of Machine Vibration—Part 1: General Guidelines. ISO: Geneva, Switzerland, 2016.
4. Mobley, R.K. *Maintenance Fundamentals*, 2nd ed.; Elsevier: Burlington, MA, USA, 2004.
5. Poór, P.; Ženíšek, D.; Basl, J. Historical Overview of Maintenance Management Strategies: Development from Breakdown Maintenance to Predictive Maintenance in Accordance with Four Industrial Revolutions. In Proceedings of the International Conference on Industrial Engineering and Operations Management, Pilsen, Czech Republic, 23–26 July 2019.
6. Mobley, R.K. *An Introduction to Predictive Maintenance*; Elsevier: Amsterdam, The Netherlands, 2002.
7. Anomughanran, O.G. Condition-Based Monitoring and Predictive Maintenance for Plant Equipment. Master's Thesis, Engineering Institute of Technology, Lagos, Nigeria, 2018.
8. Teixeira, H.N.; Lopes, I.; Braga, A.C. Condition-based maintenance implementation: A literature review. *Procedia Manuf.* **2020**, *51*, 228–235. [\[CrossRef\]](#)
9. Zhou, J.; Yang, J.; Qin, Y. A systematic overview of health indicator construction methods for rotating machinery. *Eng. Appl. Artif. Intell.* **2024**, *138*, 109356. [\[CrossRef\]](#)
10. Widodo, A.; Yang, B.-S.; Han, T. Combination of Independent Component Analysis and Support Vector Machines for Intelligent Faults Diagnosis of Induction Motors. *Expert Syst. Appl.* **2007**, *32*, 299–312. [\[CrossRef\]](#)
11. Bousdekis, A.; Magoutas, B.; Apostolou, D.; Mentzas, G. A proactive decision making framework for condition-based maintenance. *Ind. Manag. Data Syst.* **2015**, *115*, 1225–1250. [\[CrossRef\]](#)
12. Hong, S.; Zhou, Z.; Zio, E.; Wang, W. An adaptive method for health trend prediction of rotating bearings. *Digit. Signal Process.* **2014**, *35*, 117–123. [\[CrossRef\]](#)
13. Rastegari, A.; Bengtsson, M. Implementation of Condition Based Maintenance in Manufacturing Industry—A Pilot Case Study. In Proceedings of the 2014 International Conference on Prognostics and Health Management, Cheney, WA, USA, 22–25 June 2014. [\[CrossRef\]](#)
14. Irfan, M.; Saad, N.M.; Ibrahim, R.; Chai, M.; Asirvadam, V.S. Analysis of distributed faults in inner and outer race of bearing via Park vector analysis method. *Neural Comput. Appl.* **2019**, *31*, 683–691. [\[CrossRef\]](#)
15. Touret, T.; Changenet, C.; Ville, F.; Lalmi, M.; Becquerelle, S. On the use of temperature for online condition monitoring of geared systems—A review. *Mech. Syst. Signal Process.* **2018**, *101*, 197–210. [\[CrossRef\]](#)
16. WEG. *Basic Parts of a Three Phase (3 –  $\Phi$ ) Squirrel Cage Induction Motor*; WEG: Jaraguá do Sul, Brazil, 2025.

17. Sarma, N.; Tuohy, P.; Djurovic, S. Condition monitoring of rotating electrical machines. In *Encyclopedia of Electrical and Electronic Power Engineering*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 143–154. [\[CrossRef\]](#)
18. PRÜFTECHNIK. *Vibration Handbook: Fundamentals, Transducers, Measurement Methods*; PRÜFTECHNIK: Singapore, 1992.
19. Rao, S. *Mechanical Vibrations*, 6th ed.; Pearson Education: London, UK, 2018.
20. Song, H.; Sun, Y.; Gao, H.; Guo, L.; Wu, T. A novel HI construction method based on healthy-state data training for rotating machinery components. *Struct. Health Monit.* **2024**, *24*, 3849–3867. . [\[CrossRef\]](#)
21. Qiu, H.; Lee, J.; Lin, J.; Yu, G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J. Sound Vib.* **2006**, *289*, 1066–1090. [\[CrossRef\]](#)
22. Huang, Y.; Yu, J.; Qiu, M.; Chen, Z. Bearing Condition Recognition and Degradation Assessment Based on Neighborhood Preserving Embedding. *Shock Vib.* **2014**, *2014*, 781583. [\[CrossRef\]](#)
23. Yan, X.; Liu, Y.; Jia, M. Health condition identification for rolling bearing using a multi-domain indicator-based optimized stacked denoising autoencoder. *Struct. Health Monit.* **2019**, *19*, 1602–1626. [\[CrossRef\]](#)
24. Guo, L.; Li, N.; Jia, F.; Lei, Y.; Lin, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* **2017**, *240*, 98–109. [\[CrossRef\]](#)
25. Li, Q.; Yan, C.; Chen, G.; Wang, H.; Li, H.; Wu, L. Remaining Useful Life prediction of rolling bearings based on risk assessment and degradation state coefficient. *ISA Trans.* **2022**, *129*, 413–428. [\[CrossRef\]](#)
26. Li, Q.; Yan, C.; Wang, W.; Babiker, A.; Wu, L. Health indicator construction based on MD-CUMSUM with multi-domain features selection for rolling element bearing fault diagnosis. *IEEE Access* **2019**, *7*, 138528–138540.
27. Schwartz, S.; Montero Jiménez, J.J.; Vingerhoeds, R.; Salaün, M. An unsupervised approach for health index building and for similarity-based remaining useful life estimation. *Measurement* **2022**, *199*, 111510. [\[CrossRef\]](#)
28. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [\[CrossRef\]](#)
29. van Blerk, W. A Web-Based Multilevel Framework for Condition Monitoring of Industrial Equipment. Master's Thesis, North-West University, Potchefstroom, South Africa, 2019.
30. Mais, J. *Spectrum Analysis: The Key Features of Analyzing Spectra*; SKF USA Inc.: Philadelphia, PA, USA, 2002.
31. Dadfarnia, M.; Sharp, M.E.; Herrmann, J.W. Comprehensive evaluations of condition monitoring-based technologies in industrial maintenance: A systematic review. *J. Manuf. Syst.* **2025**, *82*, 449–477. [\[CrossRef\]](#)
32. Sharma, R.; Parey, A.; Tandon, N. Review of rotating machinery elements condition monitoring using acoustic emission signal. *Expert Syst. Appl.* **2024**, *255*, 124536. [\[CrossRef\]](#)
33. Qi, R.; Zhang, J.; Spencer, K. A Review on Data-Driven Condition Monitoring of Industrial Equipment. *Algorithms* **2023**, *16*, 9. [\[CrossRef\]](#)
34. Bengtsson, M. Condition Based Maintenance Systems—An Investigation of Technical Constituents and Organizational Aspects. Master's Thesis, Mälardalen University, Västerås, Sweden, 2004.
35. Hu, X.; Tang, T.; Tan, L.; Zhang, H. Fault Detection for Point Machines: A Review, Challenges, and Perspectives. *Actuators* **2023**, *12*, 391. [\[CrossRef\]](#)
36. Covary, T.; Reddy, L.; Meintjes, S. *Cost Benefit Analysis to Introduce Minimum Energy Performance Standards (MEPS) for Electric Motors in South Africa*; Report prepared for CLASP for submission to the Department of Mineral Resources and Energy (DMRE) and the South African National Energy Development Institute (SANEDI) 1; South African National Energy Development Institute (SANEDI): Johannesburg, South Africa, 2021.
37. Singh, G.K.; Al Kazzaz, S.A.S. Induction machine drive condition monitoring and diagnostic research—A survey. *Electr. Power Syst. Res.* **2003**, *64*, 145–158. [\[CrossRef\]](#)
38. Collacott, R. *Mechanical Fault Diagnosis and Condition Monitoring*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
39. Mehala, N. Condition Monitoring and Fault Diagnosis of Induction Motor Using Motor Current Signature Analysis. Ph.D. Thesis, National Institute of Technology, Kurukshetra, India, 2010.
40. Mehala, N. Current signature analysis for condition monitoring of motors. *Int. J. Electron. Comput. Sci. Eng.* **2012**, *1*, 1629–1633.
41. Stols, A. Implementing a Remote Condition Monitoring System for South African Gold Mines. Ph.D. Thesis, North-West University, Potchefstroom, South Africa, 2019.
42. Ayankoso, S.; Dutta, A.; He, Y.; Gu, F.; Ball, A.; Pal, S.K. Performance of vibration and current signals in the fault diagnosis of induction motors using deep learning and machine learning techniques. *Struct. Health Monit.* **2024**, *25*, 196–212. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.