

Article

Operator-Centred Visualization of Rolling-Element Bearing Faults: A Comparison of the Zhao–Atlas–Marks Distribution and CEEMDAN, with a Non-Specialist Readability Assessment of the ZAMD-Based Framework

Christos Tsiafis ^{1,*}, Constantine David ¹ and Apostolos Korlos ²

¹ Laboratory of Manufacturing Technology and Production Systems, Mechanical Engineering Department, International Hellenic University, Serres Campus, 62124 Serres, Greece; david@ihu.gr

² Advanced Materials and Manufacturing Technologies Laboratory, Department of Industrial Engineering and Management, School of Engineering, International Hellenic University, 57400 Thessaloniki, Greece; apkorlos@ihu.gr

* Correspondence: tsiafis@gmail.com; Tel.: +30-6977416966

Abstract

Rolling-element bearings remain a leading cause of unplanned downtime in industrial machinery, while vibration-based condition monitoring has matured, the post-2018 literature has converged on machine-learning classifiers whose interpretability layer remains restricted to expert analysts. This paper presents an operator-centred visualization framework supported by two time-frequency methods: the Zhao–Atlas–Marks Distribution (ZAMD), a Cohen’s-class representation with a cross-term-suppressing cone kernel, and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), evaluated through its Hilbert spectral analysis output. Both methods produce two-dimensional time-frequency artefacts with a similar visual structure—impact-related energy bursts that recur at the characteristic fault frequencies—and are presented in side-by-side form for each fault class. A four-stage framework wraps either method with the characteristic fault frequencies (supplied as a comparison reference) and colour-coded, healthy baseline-referenced scaling. The framework is demonstrated on a laboratory bearing rig (KOYO 6302, 600 RPM) across inner-race, outer-race, and ball-spin fault classes. A preliminary readability assessment of annotated ZAMD-generated artefacts, with twelve non-specialist participants from a brewing and packaging industrial context, recorded 89.8% aggregate classification accuracy (194 of 216 trials) at a mean response time of 15.4 s. Because no label-free or alternative-format control conditions were included, this result characterises the annotated artefact as a whole and does not isolate the contribution of the time-frequency representation from that of the annotation layer; it is established for the ZAMD engine only. The two methods are compared as visualization engines—qualitatively, through the structure of their side-by-side time-frequency artefacts, and quantitatively, through computational cost—whereas the non-specialist readability assessment characterises the ZAMD-based framework specifically. CEEMDAN is positioned as a candidate alternative engine whose time-frequency output is shown to be structurally similar but whose operator readability has not been tested with human participants and is identified as future work.

Keywords: rolling-element bearing; condition monitoring; Zhao–Atlas–Marks distribution; CEEMDAN; Hilbert spectral analysis; time-frequency analysis; vibration signal processing; operator interpretability; explainable diagnostics



Academic Editors: Kai Wang, Jingmang Xu and Yao Qian

Received: 9 June 2026

Revised: 9 July 2026

Accepted: 9 July 2026

Published: 13 July 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 Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Rolling-element bearings and rotating components are ubiquitous in industrial machinery and remain among the most frequent root causes of unplanned production downtime. Industry surveys consistently attribute 40–50% of rotating-machinery failures to bearing degradation, with secondary damage to shafts, seals, and adjacent components inflating both direct repair costs and indirect production losses [1,2]. The economic and safety motivation for early, reliable detection is therefore well established, and vibration-based condition monitoring has emerged as the dominant non-invasive diagnostic paradigm.

The detection challenge is fundamentally one of signal interpretation. Bearing faults generate periodic impulse trains at characteristic kinematic frequencies—ball-pass frequency outer race (BPFO), ball-pass frequency inner race (BPFI), ball spin frequency (BSF), and fundamental train frequency (FTF)—superimposed on broadband noise and modulated by shaft rotation. Under varying load and speed, these signatures are inherently non-stationary, and conventional Fourier-domain methods are insufficient to localise the fault evidence in time as well as in frequency. Four decades of research have produced a rich family of time-frequency methods: the Wigner–Ville Distribution (WVD), the Zhao–Atlas–Marks Distribution (ZAMD), Empirical Mode Decomposition (EMD), Ensemble EMD (EEMD), Complete Ensemble EMD with Adaptive Noise (CEEMDAN), and Variational Mode Decomposition (VMD), among others [3–7]; recent surveys review these non-parametric time-frequency, adaptive-decomposition, and deconvolution techniques in a structural-health-monitoring context [8].

A clear trend characterises the post-2018 literature: bearing diagnostics have converged on supervised machine-learning classifiers—Support Vector Machines, Convolutional Neural Networks, Long Short-Term Memory networks, and Transformer architectures—that consume mode decompositions (predominantly VMD and CEEMDAN) as feature vectors and deliver high classification accuracy on standardised benchmarks [9–12]. The resulting systems achieve laboratory accuracies above 99%, yet their interpretability layer remains restricted to expert analysts. Established industrial CBM tools continue to present spectral evidence to trained vibration specialists, but the underlying interpretive skill—sideband recognition, demodulation band selection, and kurtogram reading—is not uniformly distributed across the wider maintenance workforce. The growing literature on explainable artificial intelligence (XAI) for industrial deployment has begun to address this through post hoc attribution methods (Grad-CAM, attention visualization), but these methods explain the classifier’s decision, not the signal itself [13,14]; even recent large-language-model-based interpretation layers operate on the classifier’s output rather than on the raw signal [15].

Two operator profiles must be distinguished. By vibration analyst we refer to a specialist with formal signal-processing training who selects demodulation bands, recognises modulation sidebands, and interprets FFT, envelope, cepstrum, and kurtogram representations to reach a diagnosis—the established condition-monitoring expert. By operator we refer to front-line maintenance personnel—maintenance technicians, line operators, and shift staff—whose responsibilities include equipment surveillance but who are not assumed to possess this signal-processing literacy; for this profile, a diagnostic representation is useful only if it reduces the task to a direct perceptual judgement—recognising a repeating impact-burst pattern and comparing its rate against a small set of reference frequencies—rather than a quantitative spectral interpretation. The framework proposed in this paper is directed at the operator profile and is intended to complement, not replace, the analyst-facing workflow.

The present paper addresses the complementary question of whether a more accessible visual representation of the signal can extend interpretability to a broader operator profile. The two time-frequency methods evaluated here have been previously developed and

applied by the authors to bearing fault detection: the ZAMD, introduced by Zhao, Atlas, and Marks in 1990 [3] and shown by the authors to be robust under impulsive noise [16–18]; and CEEMDAN, introduced by Torres et al. in 2011 [7] and demonstrated by the lead author to overcome the mode-mixing limitation of the Empirical Mode Decomposition for early-stage bearing fault detection [19]. Although the two methods belong to different theoretical families (kernel-based time-frequency distribution versus adaptive signal decomposition followed by Hilbert spectral analysis), their final diagnostic artefacts—two-dimensional time-frequency maps showing impact-related energy bursts that recur at the characteristic fault frequencies—share the same perceptual structure. This observation motivates a unified operator-facing visualization framework in which either method can serve as the analytical engine.

This paper makes three contributions. First, it presents a four-stage industrial visualization framework that wraps either ZAMD or the CEEMDAN-based Hilbert spectrum with the contextual cues required for non-specialist interpretation. Second, it demonstrates the framework on a laboratory bearing test rig (KOYO 6302) covering healthy, inner-race, outer-race, and ball-spin fault classes, with side-by-side ZAMD and CEEMDAN artefacts for each condition. Third, it reports a readability assessment with twelve non-specialist participants from a brewing and packaging industrial context, conducted with ZAMD-generated artefacts, which establishes the operator readability of the ZAMD-based framework. CEEMDAN is presented as a candidate alternative engine: its time-frequency output is compared with that of ZAMD qualitatively (visual structure) and quantitatively (computational cost), but its operator readability is not tested here and is identified as future work.

The remainder of the paper is organised as follows. Section 2 reviews the development of time-frequency methods for bearing diagnostics. Section 3 presents the ZAMD and CEEMDAN formulations and the visualization framework. Section 4 describes the experimental setup. Section 5 presents results. Section 6 discusses method comparison, limitations, and complementarity. Section 7 concludes.

2. Background and Related Work

2.1. Characteristic Bearing Fault Signatures

A localised defect on a bearing component generates a periodic impulse each time a rolling element traverses it. The repetition rate depends on the bearing kinematics: BPFO and BPFI scale with shaft speed and the bearing geometry (pitch diameter, ball diameter, number of rolling elements, contact angle), while BSF describes the rotation of the rolling element about its own axis. These frequencies are widely tabulated and computed analytically [1,20]. In the early stage of fault development, the impulse amplitude is small, and the resulting spectral signature is buried in broadband background vibration. The detection task is therefore to extract a weak, non-stationary, periodic impulse train from a noisy multi-component signal—a task to which time-frequency methods are inherently suited.

2.2. From Wigner–Ville to Cohen’s Class: The ZAMD

The WVD provides a high-resolution joint time-frequency representation but suffers from cross-term interference: when a signal contains multiple components, spurious “ghost” terms appear at their cross-frequencies, obscuring the genuine signal evidence [4]. Cohen’s class of distributions addresses this by applying a smoothing kernel to the ambiguity function of the signal; different kernels yield different distributions with different trade-offs between cross-term suppression and resolution loss. The ZAMD employs a cone-shaped kernel—non-zero only within a cone-shaped region of the time-lag plane—that selectively suppresses off-diagonal cross-terms while leaving the auto-term ridges along the time-frequency diagonal largely intact [3]. The result is a time-frequency map in which genuine

signal components remain sharply localised and cross-term artefacts are heavily attenuated. For impulsive bearing fault signals, this property is particularly valuable.

2.3. From EMD to CEEMDAN

EMD [5] introduced an adaptive, data-driven approach: rather than projecting the signal onto a predefined basis, it iteratively extracts Intrinsic Mode Functions (IMFs)—narrowband oscillatory components—through a sifting procedure based on local extrema. EMD requires no prior parameters and adapts to signal content, but its principal limitation is mode mixing: under signal intermittency or noise, a single IMF may contain oscillations of disparate frequencies, obscuring the diagnostic signature. EEMD addressed this by averaging IMF sets from multiple EMD realisations of the signal with added white noise [6]. CEEMDAN [7] further refined the approach by adding adaptive noise at each decomposition stage, eliminating reconstruction error while preserving the noise-cancellation benefit. Applied to bearing diagnostics through Hilbert spectral analysis (HSA) of the resulting IMFs, CEEMDAN produces a time-frequency representation of energy distribution that, although derived through a fundamentally different computational path than ZAMD, presents the same diagnostic content in the same visual format. The lead author has applied CEEMDAN to rolling-bearing wear monitoring with success on early-stage faults [19].

VMD [21] is non-recursive and solves a constrained variational problem for K band-limited modes and their centre frequencies simultaneously. It is robust to noise and avoids the iterative drawbacks of EMD-family methods but requires preset parameters K (mode number) and α (bandwidth penalty). VMD has become the dominant method in post-2018 literature, but is not evaluated here; a parallel benchmarking study by the same authors compares ZAMD, CEEMDAN, and VMD on the CWRU public dataset, focused on detection accuracy rather than operator readability.

2.4. The Machine-Learning Turn and Operator-Facing Interpretability

Since approximately 2018, the bearing-diagnostics literature has been overwhelmingly oriented toward end-to-end ML pipelines [9–12,22], a direction consolidated in recent reviews of AI-driven predictive maintenance for Industry 4.0 [23]. The dominant architecture consumes IMFs from CEEMDAN or modes from VMD as features for a downstream classifier (SVM, Random Forest, CNN, Transformer), with reported accuracies on the CWRU dataset routinely exceeding 99% [10,11,24]. This trend has favoured methods that produce one-dimensional features that map cleanly into ML input layers; two-dimensional time-frequency representations (STFT, CWT, ZAMD, and the Hilbert spectrum derived from CEEMDAN) survive primarily as inputs to CNN-based visual classifiers [25], rather than as artefacts intended for direct human inspection. In parallel, established industrial CBM platforms continue to present spectral evidence to trained vibration analysts; however, the interpretability of these established formats presumes signal-processing literacy that is not universally distributed across the maintenance workforce. The complementary question—whether a more accessible representation of the signal can extend interpretability to a wider operator profile—is the question this paper addresses, irrespective of which of the two evaluated methods is used as the analytical engine.

2.5. Positioning of This Paper

This paper takes the position that both ZAMD and the CEEMDAN-based Hilbert spectrum are well suited to broader-audience diagnostics precisely because of the property—two-dimensional time-frequency output—that placed both outside the dominant ML-feature pipeline. With appropriate contextual support (the characteristic fault frequencies supplied as a comparison reference, a healthy-baseline panel, and consistent baseline-referenced colour scaling), either method's output can be interpreted by personnel

with limited signal-processing background. The resulting framework is positioned as complementary to ML-based diagnostics, providing an additional explainability layer for the wider maintenance workforce.

3. Methodology and Visualization Framework

3.1. ZAMD Formulation

Given a vibration signal $x(t)$, the ZAMD is computed as a Cohen's-class generalised time-frequency representation with a cone-shaped kernel. Following the formulation in the lead author's doctoral thesis and Zhao, Atlas and Marks [3], the ZAMD is expressed as

$$\text{ZAMD}_x(t, v) = \int_{-\infty}^{+\infty} \left[h(\tau) \int_{t-\tau/2}^{t+\tau/2} x\left(s + \frac{\tau}{2}\right) x^*\left(s - \frac{\tau}{2}\right) ds \right] e^{-j2\pi v \tau} d\tau, \quad (1)$$

where $h(\tau)$ is a real lag window and the time integral over s implements the cone-shaped kernel that suppresses off-diagonal cross-terms in the time-lag plane. The output $\text{ZAMD}_x(t, v)$ is a real-valued two-dimensional map over the time-frequency plane, representing the local energy distribution of the signal. For impulsive bearing fault signals, this map resolves each impact as a sharply localised broadband burst, with successive bursts recurring at the characteristic fault frequency. The kernel follows the standard Cohen's-class cone-shaped formulation—non-zero only within the cone-shaped region of the time-lag plane, bounded by the cone lines $|\tau| = 2|t|$ —which selectively suppresses off-diagonal cross-terms while preserving the auto-term ridges along the time-frequency diagonal, as adopted by the authors in earlier ZAMD applications to rolling-element bearings under impulse noise [16] and in their conference work on operator-readable spectra [17]. The lag window $h(\tau)$ that sets the degree of cross-term suppression is applied with the settings reported in Section 4.

3.2. CEEMDAN-Based Hilbert Spectral Analysis

CEEMDAN decomposes $x(t)$ iteratively into a set of Intrinsic Mode Functions (IMFs) $\{c_1(t), c_2(t), \dots, c_n(t)\}$ plus a monotonic residue $r(t)$. At each decomposition stage, adaptive white noise $\varepsilon \cdot w_i(t)$ —scaled by the standard deviation of the residue at that stage—is added before the EMD sifting, and ensemble averaging is performed over I realisations. The result, following Torres et al. [7], reconstructs the original signal as

$$x(t) = \sum_{k=1}^n c_k(t) + r(t). \quad (2)$$

Once IMFs are extracted, each is passed through the Hilbert transform to compute instantaneous amplitude $a_k(t)$ and instantaneous frequency $f_k(t)$. The Hilbert spectrum $H(f, t)$ is then constructed by summing the instantaneous amplitude contributions of all IMFs across the time-frequency plane:

$$H(f, t) = \sum_{k=1}^n a_k(t) \cdot \delta[f - f_k(t)]. \quad (3)$$

The Hilbert spectrum $H(f, t)$ is, like the ZAMD, a two-dimensional time-frequency map showing the local energy distribution of the signal. For bearing fault signals, this map exhibits impact-related energy bursts in the structural-resonance bands, recurring at the characteristic fault frequencies, visually similar in structure to those produced by ZAMD. The principal CEEMDAN parameters are the ensemble size I and the noise amplitude ratio ε . In the present work, following the lead author's calibration in [19], $I = 100$ and $\varepsilon = 0.2$ are used. The iterative sifting terminates on the Rilling stopping

criterion with thresholds $(\theta_1, \theta_2, \alpha) = (0.05, 0.5, 0.05)$ [16]. The random seed for the added-noise realisations was not fixed; reproducibility of the decomposition is instead ensured by the large ensemble size ($I = 100$), which averages out the specific noise realisation so that run-to-run variation in the aggregated Hilbert spectrum is negligible. All resulting IMFs are retained and aggregated into the Hilbert spectrum, consistent with Equation (3); no kurtosis-based preselection of IMFs is applied.

3.3. The Four-Stage Visualization Framework

Direct application of either ZAMD or the CEEMDAN-based Hilbert spectrum produces a raw time-frequency map, but a raw map is not yet a shop-floor diagnostic artefact. The proposed framework wraps either method with four sequential stages that transform the raw map into an operator-readable visualization.

- Stage 1—Signal acquisition and preprocessing. Vibration acceleration is acquired at 10 kHz (or higher, as the application permits) from a sensor mounted on the bearing housing. The signal is segmented into non-overlapping analysis windows of 1 s (10,000 samples at the 10 kHz sampling rate), long enough to capture several fault-impact cycles at the operating speed, and each segment is zero-mean normalised. No band-pass filtering or envelope demodulation is applied at this stage: the framework deliberately exposes the full-bandwidth content so that the high-frequency structural-resonance bands excited by the impacts, and the temporal repetition of the impact bursts are both visible in the raw signal evidence.
- Stage 2—Time-frequency computation. For each segment, either the ZAMD (Section 3.1) or the CEEMDAN-based Hilbert spectrum (Section 3.2) is computed, producing a two-dimensional energy map over the analysis bandwidth. The choice between the two methods is a configuration decision: ZAMD offers deterministic, non-iterative computation; CEEMDAN offers adaptive decomposition with established robustness on noisy industrial signals. The downstream stages of the framework—and the operator artefact they produce—are identical.
- Stage 3—Contextual reference. Bearing kinematic parameters (pitch diameter, ball diameter, number of rolling elements, contact angle) and the shaft speed are combined to compute the characteristic fault frequencies BPFI, BPFO, BSE, and FTF and their low-order harmonics. Because a localised defect strikes the housing impulsively, the diagnostic evidence in the raw signal is an impact-related burst pattern in the high-frequency structural-resonance bands whose repetition rate equals the characteristic fault frequency. The computed frequencies (and harmonics) are therefore supplied to the operator as the reference set against which the observed burst repetition—together with any amplitude modulation or sidebands—is compared. A healthy-baseline map at the same operating condition is provided in a side-by-side panel.
- Stage 4—Colour scaling and differentiation from the healthy baseline. The energy map is rendered on a fixed colour scale that is applied identically to the current map and to the healthy-baseline map, so that the two side-by-side panels are directly comparable, and it is the differentiation of the current map from the healthy baseline that produces the diagnostic colour contrast. Each energy map $M(t, f)$ —the ZAMD magnitude or the CEEMDAN Hilbert-spectrum energy—is first shifted to non-negative values and then normalised by a common reference peak P (the maximum magnitude observed across the healthy-baseline and fault maps of the dataset),

$$\hat{M}(t, v) = \frac{M(t, v) - \min_{t,v} M(t, v)}{P}, \quad \hat{M}(t, v) \in [0, 1], \quad (4)$$

and displayed through a fixed colour map over the unit interval. Because the healthy-baseline panel is rendered with the identical normalisation and colour map, the impact-burst energy of a faulty bearing appears warmer (towards red) on the current map than on the baseline, whereas the baseline-consistent background remains in the cooler part of the scale. The colour therefore expresses the current map's departure from the healthy baseline as a qualitative status cue, without requiring the operator to read absolute energy values.

3.4. The Resulting Artefact

The output of the four-stage framework is a single composite figure presented to the operator: a centre panel showing the time-frequency map (ZAMD or CEEMDAN Hilbert spectrum) for the current signal with the colour scale applied; a right panel showing the healthy-baseline map at the same operating condition; the characteristic fault frequencies and their harmonics supplied alongside as the comparison reference; and, as a proposed at-a-glance summary, a bottom panel of three indicator lights (BPFI, BPFO, and BSF) coloured green/orange/red.

The diagnostic cue that the operator is asked to identify is defined as follows. A localised bearing fault strikes the housing impulsively, and each impact excites the structural resonance of the housing; in the time-frequency map, each impact therefore appears as a short, broadband energy burst concentrated in the resonance bands (here of the order of 1–4 kHz), and successive bursts recur at the characteristic fault frequency. The operator's task is to recognise this impact-burst pattern and to judge whether its repetition rate—together with any amplitude modulation or sidebands—is consistent with BPFI, BPFO, or BSF (or their harmonics) provided as reference, and whether its colour indicates elevation above the healthy baseline. This is an interpretive comparison rather than a deterministic reading: the burst pattern is affected by noise, overlapping harmonics, fault severity, and the method-specific artefacts of the underlying representation—which is precisely why a non-specialist readability assessment is meaningful rather than trivial.

Because the cognitive task imposed on the viewer—recognise the impact-burst pattern and matching its repetition to a characteristic frequency—is structurally the same regardless of whether the underlying time-frequency map was produced by ZAMD or by the CEEMDAN Hilbert spectrum, the same readability characteristics may be expected for either method; this is a hypothesis motivated by the qualitative visual comparison reported in Section 5, and it is tested in the present study only for the ZAMD engine.

4. Experimental Setup

4.1. Laboratory Bearing Test Rig

The experimental validation uses the bearing test rig developed by the lead author during the doctoral research at the University of Kragujevac [16]. The rig consists of a variable-speed electrical motor driving a shaft supported by two simple ball bearings; the test bearing is mounted in a known position, and an accelerometer is placed in immediate proximity to the bearing housing. The acquired analog signal is digitised through an A/D converter and stored on a personal computer for offline processing. Particular attention has been paid to the mechanical mounting of the rig: shaft alignment was verified using a coordinate measuring machine, and the rubber-mounted base was characterised for vibration isolation in order to minimise external interference and ensure that the recorded signals reflect bearing-internal phenomena [16,18]. The mechanical assembly, the physical realisation of the rig, and the instrumented acquisition setup are shown in Figures 1–3.

The experimental validation uses the authors' own laboratory rig rather than a public benchmark such as the Case Western Reserve University (CWRU) dataset [26]. The princi-

pal value of CWRU is as a common benchmark for comparing classifier detection accuracy across studies, and such comparisons are already extensively reported in the literature, with accuracies above 99% (Section 2). The present study does not address detection accuracy but the human readability of visual artefacts—a question orthogonal to classification benchmarking, for which cross-study accuracy comparability is not the relevant criterion. We therefore use the rig, bearing dataset, and ZAMD and CEEMDAN implementations established in the lead author's peer-reviewed doctoral work [16,19], which provide a self-consistent and fully documented experimental basis under a single controlled operating condition. CWRU is, in turn, employed by the present authors for detection-accuracy benchmarking in a parallel study of ZAMD, CEEMDAN, and VMD (Section 2); the two studies are complementary, addressing operator readability and detection accuracy, respectively.

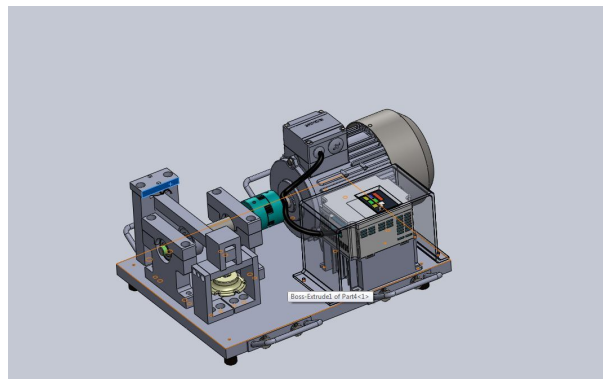


Figure 1. Bearing test rig—mechanical assembly (CAD model).

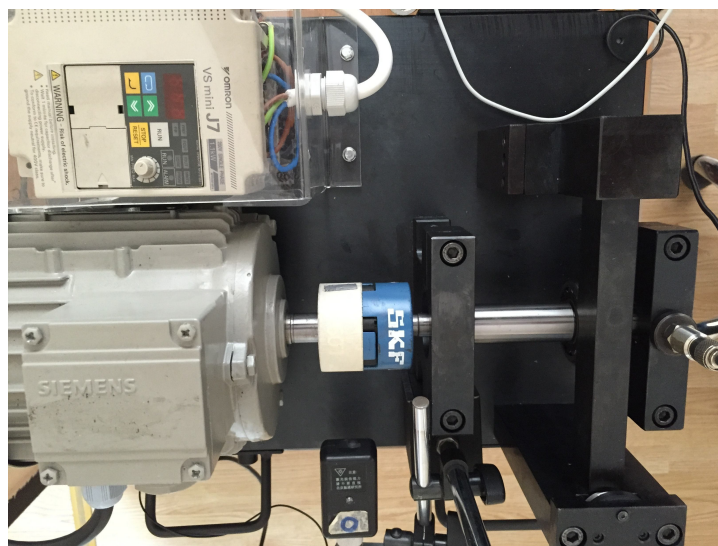


Figure 2. Physical realisation of the bearing test rig: variable-speed electric motor and inverter drive (left), flexible coupling, shaft, and the two bearing housings mounting the support and test bearings on the base frame (right).

The test bearing is a deep-groove ball bearing KOYO 6302 with the following kinematic parameters: pitch diameter $PD = 28.5$ mm, ball diameter $BD = 7.938$ mm, number of rolling elements $N = 7$, and contact angle $\theta = 0^\circ$. The faulty specimens were factory pre-damaged KOYO 6302 (C4 clearance) bearings supplied by the manufacturer (JTEKT/Koyo, Osaka, Japan), each carrying a localised defect on the inner race, the outer race, or a rolling element; together with a healthy reference bearing of identical type, these provide three discrete fault classes and a healthy baseline. Each specimen was accompanied by a manufacturer noise-characterisation certificate (JTEKT/Koyo Noise Conversion Table),

which grades the defect by its standardised vibration signature—acceleration, velocity, decibel and Anderon values across low-, mid- and high-frequency bands—rather than by geometric dimensions; this certificate is provided as Supplementary Table S1. The use of manufacturer-prepared and independently characterised specimens ensures controlled and consistent defect conditions across the dataset.

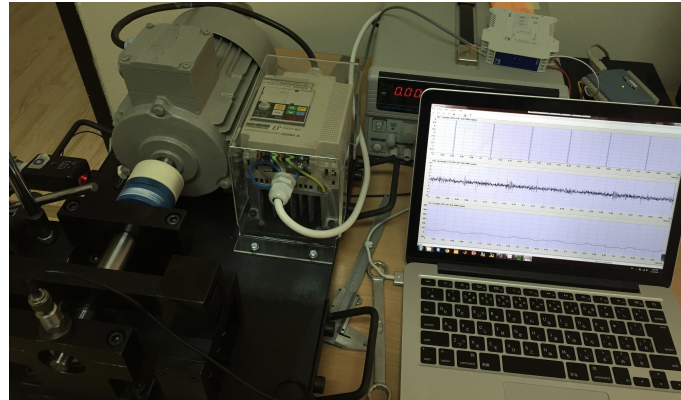


Figure 3. Instrumented acquisition setup: motor, accelerometer, A/D converter, and PC.

4.2. Instrumentation

The shaft was driven by a variable-speed three-phase induction motor (Siemens, Munich, Germany) whose rotational speed was regulated by a frequency inverter, allowing operation at the nominal test speed of 600 RPM ($f_r = 10$ Hz). Vibration was measured with a piezoelectric accelerometer (Brüel & Kjær 4507 B; Brüel & Kjær, Nærum, Denmark, sensitivity 100 mV/g) mounted on the bearing housing in immediate proximity to the test bearing. The accelerometer output was digitised by a PMD-1608FS data-acquisition module (Measurement Computing Corporation, Middleboro, MA, USA; 16-bit resolution) at a sampling rate of 10 kHz and stored on a host computer, with all subsequent processing performed in MATLAB R2018b (The MathWorks, Natick, MA, USA). The rig was seated on a high-stiffness metal base plate and isolated from the supporting bench by AVM-2055 natural-rubber/styrene-butadiene (NR/SBR) urethane mounting pads, whose elastic and damping properties (damping coefficient $D \approx 0.04$) were characterised and validated by combined finite-element (ANSYS Workbench 15.1, Academic Research licence; ANSYS Inc., Canonsburg, PA, USA) and FFT modal analysis [16]. Shaft alignment and the dimensional tolerances of the bearing housings were verified with a coordinate measuring machine (PH10MQ probing head, TP200 probe; Renishaw plc, Wotton-under-Edge, UK), confirming a 42 mm (H7) housing bore and a shaft-alignment concentricity of 0.062 mm.

4.3. Acquisition Protocol

Measurements were acquired under fixed operating conditions: shaft rotational speed 600 RPM (corresponding to a shaft frequency $f_r = 10$ Hz), no external radial load, and a sampling rate of 10,000 Hz. Each recording lasted 15 s; for every fault class and the healthy baseline, ten consecutive recordings were collected. Bearing characteristic fault frequencies were computed analytically from the KOYO 6302 geometry and verified against the SKF online frequency calculator, with agreement within 0.5% (Table 1).

Table 1. Bearing characteristic fault frequencies for KOYO 6302 at 600 RPM ($f_r = 10$ Hz), no load.

Fault Location	Symbol	Formula	Value (Hz)
Inner race	BPI	$(N/2)(1 + (BD/PD) \cos \theta) \cdot f_r$	44.75
Outer race	BPFO	$(N/2)(1 - (BD/PD) \cos \theta) \cdot f_r$	25.25
Ball spin	BSF	$(PD/(2BD))(1 - ((BD/PD) \cos \theta)^2) \cdot f_r$	16.56
Train (cage)	FTF	$(1/2)(1 - (BD/PD) \cos \theta) \cdot f_r$	3.61

4.4. Implementation

Both the ZAMD and the CEEMDAN-based Hilbert spectrum computations are implemented in MATLAB. The ZAMD uses the standard Cohen's-class cone-shaped kernel with Gaussian smoothing windows—a length-9 Gaussian time-smoothing window and a length-27 Gaussian lag window $h(\tau)$ —applied without further problem-specific tuning; the CEEMDAN ensemble size is $I = 100$ with noise amplitude ratio $\varepsilon = 0.2$. Sampling was at $f_s = 10,000$ Hz. Each 15 s recording is divided into non-overlapping 1 s analysis windows of 10,000 samples, giving fifteen windows per recording (150 per class over the ten recordings and 600 across the four classes); each window is zero-mean normalised before the time-frequency computation, and the resulting energy map is then colour-scaled by the common reference-peak normalisation of Stage 4 (Section 3). Every map in Figures 4–7 is one such 1 s window—which is why the time axis of these figures spans 0–1 s—selected as a representative window of its class; the fault specimen shown in each is identified by its manufacturer-certificate row in Supplementary Table S1 (inner-race = row 4, outer-race = row 5, ball = row 8, healthy = row 9). The 2048-sample (0.2048 s) window used for the computational-performance comparison (Section 5) is a separate, deliberately small fixed reference workload, chosen only so that the ZAMD and CEEMDAN timings are measured on an identical small unit; it is not the analysis window of Figures 4–7, and, because processing time scales approximately linearly with window length, the reported per-window timings scale accordingly for the 1 s analysis window. Because the diagnostic evidence in the housing signal is the pattern of impact bursts in the high-frequency structural-resonance bands rather than a spectral line at the (low) characteristic frequency, the fault frequency is recovered from the temporal repetition of the bursts rather than from fine spectral resolution: at 600 RPM, BSF (≈ 16.56 Hz), BPFO (≈ 25.25 Hz), and BPFI (≈ 44.75 Hz) correspond to inter-burst intervals of the order of 20–60 ms, which are well resolved along the time axis and are read directly from the map. The framework therefore does not depend on resolving the characteristic frequencies as narrow low-frequency spectral lines. The fundamental train frequency FTF (≈ 3.61 Hz) is not among the fault classes considered. The same acquisition and processing settings are used for both engines to keep the comparison controlled. All processing was performed on a standard desktop workstation (Intel Core i7-class CPU, 32 GB RAM) running Windows 10 and MATLAB R2018b. Computation times for both methods are reported in Section 5.

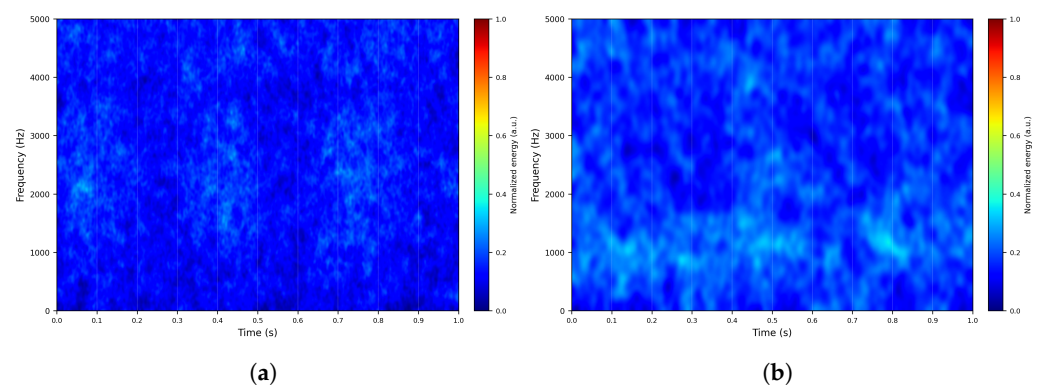


Figure 4. Healthy baseline (certificate specimen 9, Supplementary S1). Neither representation shows impact bursts; the housing energy is diffuse and low-level, without a periodic burst structure. No inter-burst-interval markers are drawn, as no periodic burst pattern exists to verify. (a) ZAMD—healthy. (b) CEEMDAN Hilbert spectrum—healthy.

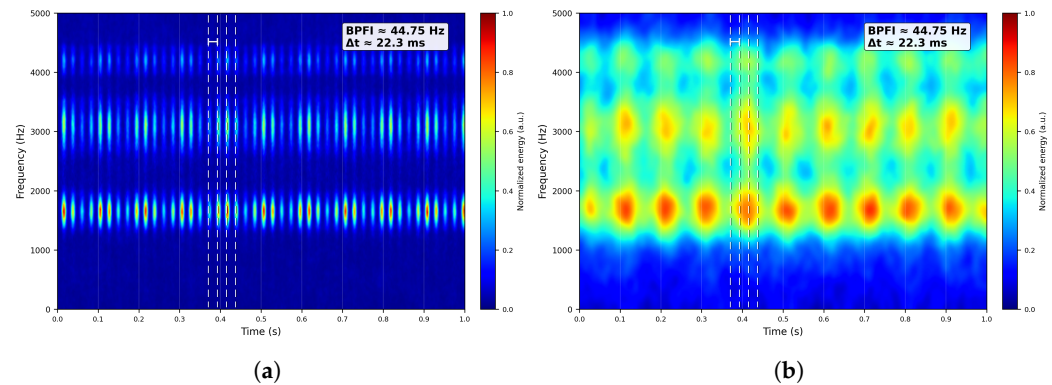


Figure 5. Inner-race fault (certificate specimen 4, Supplementary S1). Impact bursts in the resonance bands recur at BPFI (≈ 44.75 Hz, Table 1) and are amplitude-modulated as the defect traverses the load zone, giving a less stationary, sideband-rich pattern; the ZAMD map (left) resolves the bursts more sharply. The dashed vertical markers indicate the expected inter-burst interval $\Delta t = 1/\text{BPFI} \approx 22.3$ ms, allowing the burst repetition to be verified directly against the characteristic frequency; observed spacings agree with the computed value to within $\sim 1\text{--}2\%$, consistent with rolling-element slip. (a) ZAMD—inner-race fault. (b) CEEMDAN Hilbert spectrum—inner-race fault.

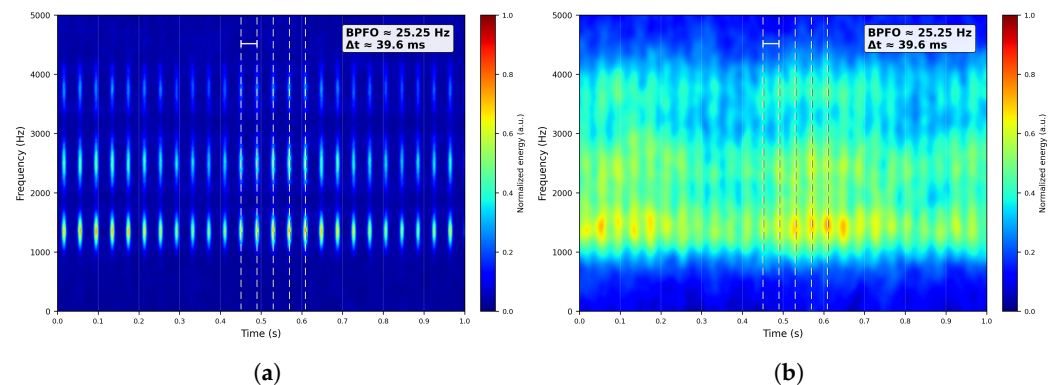


Figure 6. Outer-race fault (certificate specimen 5, Supplementary S1). Stable, evenly spaced impact bursts in the structural-resonance bands recur at BPFO (≈ 25.25 Hz, Table 1); the fixed defect position relative to the sensor gives the most regular and repeatable pattern of the three fault classes. The dashed vertical markers indicate the expected inter-burst interval $\Delta t = 1/\text{BPFO} \approx 39.6$ ms; the burst spacing matches the markers across the record to within $\sim 1\text{--}2\%$, consistent with rolling-element slip. (a) ZAMD—outer-race fault. (b) CEEMDAN Hilbert spectrum—outer-race fault.

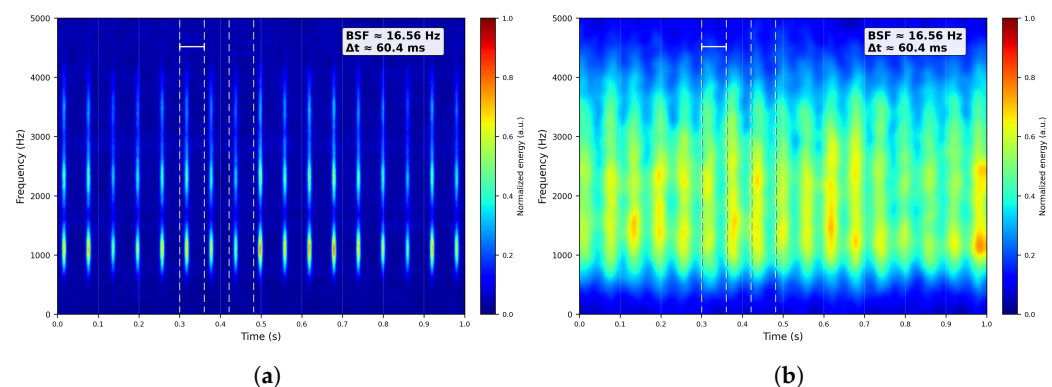


Figure 7. Ball (rolling-element) fault (certificate specimen 8, Supplementary S1). The impact bursts near BSF (≈ 16.56 Hz, Table 1) are more diffuse and less regular than for the inner- and outer-race cases, consistent with the continuously changing ball-contact geometry and the lower BSF excitation under the unloaded condition; this is the most difficult class to read. The dashed vertical markers indicate

the expected inter-burst interval $\Delta t = 1/\text{BSF} \approx 60.4$ ms; the dominant burst spacing follows the markers, with the larger irregularity expected for this class. (a) ZAMD—ball fault. (b) CEEMDAN Hilbert spectrum—ball fault.

5. Results

5.1. Side-by-Side Comparison: ZAMD and CEEMDAN

Figures 4–7 present, for each condition of the KOYO 6302 bearing at 600 RPM, the time-frequency maps produced by the two methods side-by-side: the ZAMD map on the left and the CEEMDAN-based Hilbert spectrum on the right. Both maps are displayed on the same physical frequency axis (0–5 kHz) and time axis. In both panels the ordinate is physical frequency in hertz: for the ZAMD it is the frequency variable of the Cohen’s-class time-frequency distribution, and for the CEEMDAN engine it is the frequency coordinate of the Hilbert spectrum $H(f, t)$ (Equation (3)); although the Hilbert spectrum is assembled from the instantaneous-frequency estimates of the individual IMFs, its output is itself a frequency–time energy map, which is what allows the two artefacts to be compared side-by-side.

The maps show the vibration energy measured at the bearing housing. A localised defect strikes the housing impulsively and excites its structural resonances (here concentrated between roughly 1 and 4 kHz), so each impact appears as a short, broadband vertical energy burst, and successive bursts recur at the characteristic fault frequency. The diagnostic information is therefore carried by the repetition rate of the bursts rather than by a stationary band at the characteristic frequency itself. For the outer-race case, the bursts are stable and evenly spaced at BPFO (≈ 25.25 Hz); for the inner-race case, they recur at BPFI (≈ 44.75 Hz) but are amplitude-modulated as the defect passes through the load zone, producing a less stationary, sideband-rich pattern; and for the ball case, the bursts are more diffuse and less regular near BSF (≈ 16.56 Hz), consistent with the continuously changing contact geometry. The healthy baseline shows no such impact bursts. The structural appearance is closely similar across the two methods, but the ZAMD map renders the bursts more sharply against a more strongly suppressed background (its cone kernel attenuates cross-terms), while the CEEMDAN Hilbert spectrum presents the same bursts with a more diffuse texture (a consequence of residual mode mixing). Recognising this burst pattern and comparing its repetition against the characteristic frequencies of Table 1 is the cognitive task imposed on the non-specialist observer, and it is the same for both methods. To allow the reader to verify this correspondence independently, Figures 5–7 include dashed vertical markers spaced at the expected inter-burst interval of each fault class ($\Delta t = 1/f$: ≈ 22.3 ms for BPFI, ≈ 39.6 ms for BPFO, ≈ 60.4 ms for BSF); the observed burst spacing follows these markers to within ~ 1 – 2% , consistent with rolling-element slip. These markers are an annotation added for the reader’s verification; they were not part of the stimuli shown to participants in the readability assessment (Section 5.3).

Across the four conditions, the side-by-side comparison shows that ZAMD and CEEMDAN, despite their fundamentally different computational paths, produce diagnostic visualizations of closely similar structure. This qualitative visual similarity motivates the hypothesis that the two engines would be comparably readable to non-specialists; that hypothesis is; however, tested in the present study only for ZAMD (Section 5.3), and its confirmation for CEEMDAN is left to future work.

5.2. Composite Operator Artefact

Within the four-stage framework, the central time-frequency map (from either method) is presented to the operator alongside the healthy-baseline reference panel; the proposed full design additionally envisages a three-light status summary (BPFI, BPFO, and BSF)

as described in Section 3.4. In that envisaged deployment the composite would support single-glance interpretation: a maintenance technician walking past a monitoring station would see the indicator lights and, if any is non-green, consult the central time-frequency map for the visual signature. The indicator-light panel is a proposed design element only it was not implemented as a live classifier and was not shown to participants in the readability assessment reported below, which used the time-frequency map and the healthy-baseline panel alone (Section 5.3). The signature itself—an impact-burst pattern whose repetition rate matches a characteristic fault frequency—is assessed against the supplied reference frequencies, regardless of which of the two time-frequency methods underlies the map.

5.3. Non-Specialist Readability Assessment

To assess whether the framework genuinely supports interpretation by personnel without time-frequency-analysis training, a structured visual classification task was carried out with twelve participants drawn from outside the signal-processing community. Participants included maintenance technicians, line operators, and engineering students recruited from a brewing and packaging industrial context, providing realistic operator-profile diversity beyond a purely academic sample. None of the participants had prior formal training in time-frequency analysis, envelope spectra, or condition-monitoring software.

Protocol. Each participant received a brief (5-min) standardised orientation: “Look for repeated impact bursts in the map; judge how often they repeat and check whether that rate matches one of the listed fault frequencies (BPFO, BPFI, or BSF).” They then classified eighteen stimulus images—four to five per class, balanced across healthy, inner-race-fault, outer-race-fault, and ball-fault cases drawn from the KOYO 6302 dataset—presented in randomised order. Each stimulus comprised a single ZAMD time-frequency map shown alongside the healthy-baseline panel at the same operating condition; the characteristic fault frequencies and the burst pattern expected for each fault type were conveyed in the orientation. The proposed indicator-light summary (Section 3.4) was not part of the stimuli and was not shown to participants. The stimulus pool was formed from the same non-overlapping 1 s windows used throughout: each class contributes 150 candidate windows (fifteen from each of its ten 15 s recordings), 600 across the four classes, from which the eighteen stimuli presented to each participant (four to five per class) were drawn. The stimulus order was counterbalanced across participants to control for learning effects, and class labels were not revealed during the session. This yields 216 classification trials in total (12×18). Across participants, class assignment was balanced so that each of the four classes accumulated 54 trials in total (each participant classified four or five stimuli of a given class, drawn from a shared pool, counterbalanced so that the per-class totals were equal); the per-class counts $n = 54$ reported in Table 2 are therefore aggregate totals across all twelve participants, not per-participant counts. Classification accuracy was scored for every trial and retained as aggregate and per-class counts; a subjective readability score (1 = unintelligible, 5 = immediately clear) was collected once per participant at the end of the session. Response time was not logged per trial: the total elapsed classification time was measured for each participant’s session with a handheld stopwatch, and the mean response time reported below is this total elapsed time divided by the number of trials—a coarse session-level aggregate rather than a per-trial measurement. The readability outcomes are reported as descriptive pilot statistics rather than as inferential estimates. The 216 trials are clustered by participant and by stimulus and are therefore not mutually independent, so a confidence interval computed under an independence assumption would understate the true uncertainty; a clustering-aware analysis (participant-level bootstrap or a mixed-effects logistic model) and a full confusion matrix require per-trial records, which were not retained in this anonymous pilot (only aggregate outcomes were kept). The individual

stimulus images were likewise not archived. Both analyses, and a fixed version-controlled stimulus set, are planned for the confirmatory study identified in Section 6.4.

Method scope. The artefacts used in the readability assessment were generated using the ZAMD as the time-frequency engine; accordingly, the readability outcomes reported in Table 2 are established for the ZAMD-based framework specifically. The side-by-side comparison of Figures 4–7 shows that the CEEMDAN-based Hilbert spectrum presents the same cognitive task to the observer—recognising the impact-burst pattern and comparing its repetition against the characteristic fault frequencies—which motivates the expectation that CEEMDAN would be comparably readable; this remains, however, an untested hypothesis. A confirmatory readability study with CEEMDAN-generated artefacts is identified as future work (Section 6.4).

Table 2. Non-specialist classification of the annotated ZAMD artefacts (time-frequency map with healthy-baseline panel)—aggregate and per-class. Readability for the CEEMDAN engine is not tested here and is left to future work (see Section 6.4).

Outcome	Value
Correct classifications	194/216
Aggregate accuracy	89.8% (194/216 trials)
Mean response time	15.4 s per artefact
Subjective readability	4.4/5
Healthy ($n = 54$)	94%
Inner-race fault ($n = 54$)	91%
Outer-race fault ($n = 54$)	90%
Ball-spin fault ($n = 54$)	84%

Interpretation. The aggregate accuracy provides preliminary descriptive evidence that the framework’s composite artefact is interpretable by non-specialist personnel, with a mean response time well below those typically associated with detailed spectral inspection. As an anonymous pilot without retained per-trial records, a confusion matrix, or a clustering-aware analysis, this result is descriptive and hypothesis-generating rather than a validated performance claim; the suitability of the framework for practical fault triage is thus indicated but not established and remains to be confirmed by the instrumented study of Section 6.4. It is important to be transparent about the source of this performance: a substantial fraction of the readability gain is attributable to the contextual reference (the supplied BPFI/BPFO/BSF characteristic frequencies and harmonics, together with the healthy-baseline panel) rather than to the underlying time-frequency decomposition alone. This reference encodes engineering knowledge that, in conventional FFT or envelope-spectrum workflows, the analyst must hold in working memory. In this sense, the framework is best understood as the combination of (i) a two-dimensional time-frequency representation (either ZAMD or CEEMDAN Hilbert spectrum) and (ii) a knowledge-injecting annotation layer; both elements are essential to the result. We emphasise that the present design does not isolate the contribution of the representation from that of the annotation: because no label-free control condition and no comparison against conventional formats (FFT, envelope spectrum, and STFT/CWT) were included, the reported accuracy characterises the annotated framework as a whole and cannot be attributed to the time-frequency representation alone. The standardised orientation given to participants (“look for repeated impact bursts and judge whether their repetition matches one of the listed fault frequencies”) supplies the reference knowledge that a specialist would otherwise recall from memory, but the operator must still interpret the burst pattern against it rather than perform an independent spectral diagnosis; this is consistent with the framework’s stated purpose of extending interpretability to non-specialists, but it also means the assessment measures

the readability of the composite artefact, not the standalone diagnostic power of ZAMD or CEEMDAN. Control conditions that remove the annotation layer and a head-to-head comparison with conventional annotated formats are required to separate these contributions and are identified as future work (Section 6.4).

The per-class breakdown identifies ball faults as the weakest class (84%). This is consistent with the underlying physics: the ball defect produces the most diffuse and least regular impact-burst pattern because the contact geometry changes continuously, and under the unloaded operating condition employed here, the BSF excitation is comparatively low. More generally, the fact that the aggregate accuracy is of the order of 85–90% rather than near 100% reflects the genuine difficulty of the task. Fault localisation from these maps is not a deterministic “one line, one fault” reading but a probabilistic interpretation—the operator recognises energy concentration, modulation, and harmonics and compares them against BPFI, BPFO, and BSF to reach the most likely diagnosis. The burst pattern is affected by broadband noise, overlapping harmonics and sidebands, weak or low-severity faults, residual mode mixing in the CEEMDAN representation and residual cross-terms in the ZAMD representation, and per-bearing variability in defect severity. An accuracy in this range is therefore the expected, realistic outcome and is precisely what makes an operator readability study meaningful: were the localisation visually obvious and deterministic, neither the study nor the framework would be necessary. The 84% ball figure is reported without correction as a transparent indication of where the visualization is least reliable.

A further, practically important distinction is between fault detection and fault classification. The accuracy figures above concern the harder task of identifying the fault type. Fault detection—distinguishing a faulty bearing from a healthy one—is markedly more reliable because any of the three fault classes produces a conspicuous impact-burst pattern that is entirely absent from the healthy baseline (Figure 4). For the shop-floor triage role that the framework is designed to support, this detection capability is the primary value: a non-specialist operator can recognise that a bearing is faulty and escalate to a specialist or schedule a bearing replacement at a planned maintenance window, even without identifying the exact fault location. In this framing, exact fault-type classification is a secondary, best-effort output, whereas the reliable, high-confidence signal is simply the presence or absence of the impact-burst signature.

5.4. Computational Performance

The two methods differ in computational profile. Computation times were benchmarked on a standard 2048-sample window (0.2048 s) at 256×256 time-frequency resolution—a fixed reference workload, distinct from the 1 s analysis window of Figures 4–7, chosen so that the two methods are timed on an identical small unit; processing time scales approximately linearly with window length for both methods. ZAMD requires approximately 380 ms per such window on the desktop workstation described in Section 4.4. CEEMDAN, with an ensemble size of $I = 100$, requires approximately 2.4 s per window of the same size; the additional cost is dominated by the ensemble averaging and per-stage adaptive-noise sifting. Both timings were measured for a single 2048-sample analysis window, excluding figure rendering and file input/output, under identical experimental conditions—the same input window, the same 256×256 output grid, the same signal, and the same workstation (Section 4.4; Intel Core i7-class CPU, 32 GB RAM, Windows 10, MATLAB R2018b)—with each method operating at the settings used throughout this study (the ZAMD cone kernel of Section 3; CEEMDAN ensemble size $I = 100$ and noise amplitude ratio $\varepsilon = 0.2$). Because the two methods belong to different computational families, their method-specific parameters cannot be matched to one another; the reported $\approx 6\times$ ratio therefore compares each method at its recommended configuration rather than under a

shared parameter set and is specific to $I = 100$ —CEEMDAN’s cost scales approximately linearly with the ensemble size, so a smaller I would narrow the gap at some cost to noise robustness. Both methods are therefore comfortably suited to the periodic inspection workflows typical of industrial CBM deployments, where maps are computed at scheduled intervals rather than continuously, but neither is appropriate for high-frequency streaming applications without specialised acceleration.

5.5. Positioning Against Conventional Visualization

The framework—whether configured with ZAMD or CEEMDAN—is intended to complement, not replace, conventional analyses. Established formats such as FFT spectra, envelope spectra, and wavelet scalograms remain the diagnostic standard for trained vibration analysts and are not displaced by the present framework. The motivation for the framework is access: those conventional formats rely on familiarity with signal-processing concepts (sideband modulation, demodulation-band selection, impulse-train recognition) that are not uniformly distributed across the wider maintenance workforce. The framework proposed here offers a representation in which the diagnostic content is presented in a more directly perceptual form, supported by the supplied reference frequencies and healthy baseline, and is therefore intended to extend interpretability to a broader operator profile, alongside—not in place of—established analyst-facing workflows.

6. Discussion

6.1. When to Use Which Method

Within the four-stage framework, ZAMD is the engine validated for operator readability in this study; the CEEMDAN-based Hilbert spectrum is a candidate alternative whose visual output is qualitatively similar (Section 5.1) but whose operator readability has not yet been tested. Should that readability be confirmed, the choice between the two would become a configuration decision rather than a diagnostic-quality decision, driven by the computational and parametric considerations set out below. ZAMD offers deterministic, non-iterative computation with a fixed kernel (no ensemble averaging and no convergence loop) and approximately six times lower computational cost than CEEMDAN at the settings used here. CEEMDAN offers adaptive, data-driven decomposition that performs strongly under signal intermittency and non-stationary noise—conditions in which kernel-based distributions can struggle. In an industrial deployment with relatively clean signals and a need for fast, predictable computation, ZAMD is the natural default; in environments with higher noise variability or atypical impulse patterns, CEEMDAN provides a robust alternative without requiring any change to the downstream framework stages.

6.2. Complementarity with Machine Learning

The framework is not proposed as a replacement for ML-based bearing diagnostics. In practice, the two approaches are complementary. An industrial CBM system can deploy a high-accuracy ML classifier as the primary alarm-generation layer and the framework (in either ZAMD or CEEMDAN configuration) as the explanation layer activated when an alarm is raised. In the taxonomy of explainable AI, the framework belongs to the “intrinsically interpretable model” tier—the signal evidence itself is the explanation, not a post hoc attribution over a classifier’s internal state.

6.3. Limitations

The framework has clear boundaries that should be stated explicitly.

- Variable-speed and variable-load operation. The reference frequencies are computed analytically from a fixed, known shaft speed, so the framework as demonstrated

assumes quasi-stationary operation—the condition of the constant 600 RPM laboratory rig. Two robustness considerations arise under dynamic field conditions. First, speed variation: because the characteristic frequencies scale linearly with shaft speed (Table 1), a relative speed fluctuation of $\pm p\%$ shifts the burst repetition rate by the same $\pm p\%$. The comparison against a reference frequency therefore remains valid as long as this drift is small relative to the spacing between adjacent characteristic frequencies; for the KOYO 6302 frequencies (BPFI 44.75, BPFO 25.25, BSF 16.56 Hz), this tolerates a few percent of speed variation, but not the large excursions of start-up/shut-down or genuinely variable-speed duty. Second, load variation: bearing characteristic frequencies are kinematic and depend on geometry and speed rather than on load, so load principally modulates the burst amplitude—which the colour scaling is designed to capture—with only minor slip-induced shifts of the rolling element and cage frequencies (BSF, FTF) of a few percent. For operation outside this quasi-stationary regime, order-tracking pre-processing (angular resampling to a constant number of samples per revolution) converts the time axis to an angular axis and stabilises the burst pattern at fixed orders; this is identified as the primary extension in Future Work (Section 6.4) and applies to both ZAMD and CEEMDAN.

- Unloaded operating condition. All measurements were taken under no external radial load, whereas industrial bearings normally operate under load; the present data therefore do not reproduce the load-zone amplitude modulation of a loaded drive. The unloaded condition was chosen to isolate clean, unambiguous bearing-internal fault signatures for a first readability demonstration, and it is the conservative case for this purpose: the fault impulse strength is proportional to the rolling-element contact force, so fault energy at the characteristic frequencies is generally weaker under no load than under load [1]. The readability outcomes reported here—including the 84% floor for the low-excitation ball-spin class—therefore represent a lower bound, and loading, which strengthens precisely the impact bursts the operator is trained to recognise, would be expected to maintain or improve framework readability. A validation campaign across representative load levels is identified as future work.
- Coupled rotating components. When the analysis bandwidth contains a gearbox, additional shafts, or multiple bearings, mesh frequencies and harmonics can overlap with bearing characteristic frequencies. The impact-burst pattern then carries ambiguous attribution in both methods.
- Low signal-to-noise ratio. The cone kernel substantially suppresses cross-terms for ZAMD, and ensemble noise averaging mitigates noise-induced mode mixing for CEEMDAN; nonetheless, at very low SNR, residual artefacts re-emerge for both methods and the contextual reference can no longer fully compensate.
- Multi-fault scenarios. All conditions reported here are single-fault. Compound faults can produce overlapping or modulated burst patterns that the present framework is not designed to disambiguate.
- Readability scope—the principal empirical limitation. The readability result of Section 5.3 was obtained exclusively with ZAMD-generated artefacts; the CEEMDAN-based Hilbert spectrum, although applied experimentally to the same bearing recordings (Figures 4–7), has not yet been subjected to a human-subject readability test. Any expectation that the two engines are comparably readable for non-specialist interpretation therefore rests, at present, on a logical and perceptual argument (Section 5.1) supported by the qualitative side-by-side comparison, and not on independent readability data for CEEMDAN. We regard this as the most important limitation of the present study: accordingly, the readability claim of this paper is confined to the ZAMD engine, and no readability result is claimed for CEEMDAN. Comparable readability

of the CEEMDAN engine is a well-motivated but unconfirmed hypothesis, whose confirmation is left to the study identified in Future Work.

- Statistical scope of the pilot. The readability figures are descriptive pilot statistics, not inferential estimates. The trials are clustered by participant and by stimulus, so an independence-based confidence interval would understate the uncertainty; and, because only aggregate outcomes were retained in this anonymous pilot, a clustering-aware reanalysis and a full confusion matrix cannot be produced post hoc. These are deferred to the instrumented confirmatory study of Section 6.4.
- Stimulus provenance and the indicator-light element. Each pilot stimulus comprised a ZAMD time-frequency map shown alongside the healthy-baseline panel, with the characteristic frequencies and expected per-fault burst patterns conveyed in the orientation. The three-light status summary of Section 3.4 is a proposed design element that was neither implemented as a live classifier nor shown to participants, and no readability result is attributed to it. Moreover, the individual stimulus images were not archived (only aggregate outcomes were retained), so the exact images presented cannot be reproduced post hoc; a fixed, version-controlled stimulus set is part of the instrumented confirmatory study of Section 6.4.
- Readability is not classification accuracy. The pilot result quantifies how interpretable the artefact is to non-specialists; it does not quantify field detection performance under operational conditions and is not a substitute for a calibrated alarm-generation classifier.

6.4. Future Work

Five concrete extensions are identified: (i) a confirmatory human-subject readability assessment using CEEMDAN-generated artefacts under the identical protocol of Section 5.3, explicitly designed to test the perceptual-equivalence premise and to establish the feasibility of the CEEMDAN-based Hilbert spectrum as an operator-facing artefact, and logging per-trial outcomes so that a clustering-aware analysis (participant-level bootstrap or a mixed-effects logistic model) and a full confusion matrix can be reported—this is planned as the immediate next step of this work; (ii) a controlled readability study incorporating label-free control conditions and a head-to-head comparison against conventional annotated formats (FFT, envelope spectrum, STFT/CWT), designed to separate the contribution of the time-frequency representation from that of the annotation layer; (iii) order-tracking integration for variable-speed conditions, applicable to both methods; (iv) integration with prognostic remaining-useful-life estimation so that the visualization conveys fault evolution as well as current state; (v) a controlled deployment study in a production environment, measuring the framework's effect on maintenance decision speed and accuracy under realistic shop-floor conditions, including representative bearing load levels.

7. Conclusions

This paper has presented a four-stage operator-centred visualization framework that can be configured with either the Zhao–Atlas–Marks Distribution or the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (via Hilbert spectral analysis) as its time-frequency engine. Side-by-side demonstration on a laboratory KOYO 6302 rig across healthy, inner-race, outer-race, and ball-spin fault conditions showed that the two methods produce diagnostic artefacts of closely similar visual structure, despite their fundamentally different computational paths.

A preliminary readability assessment of annotated ZAMD-generated artefacts, with twelve non-specialist participants from a brewing and packaging industrial context, recorded an aggregate classification accuracy of 89.8% (194 of 216 trials) at a mean response time of 15.4 s, with per-class performance between 84% (ball-spin fault) and 94%

(healthy baseline). Because no label-free or alternative-format control conditions were included, this result characterises the annotated artefact as a whole and does not isolate the contribution of the time-frequency representation from that of the annotation layer. It is established for the ZAMD-based framework; because the assessment used ZAMD-generated artefacts only, its extension to the CEEMDAN engine remains an untested hypothesis, and a confirmatory study using CEEMDAN-generated stimuli is identified as future work.

ZAMD is therefore the demonstrated engine of the framework, with CEEMDAN positioned as a candidate alternative; a qualitative visual and quantitative computational comparison of the two is provided to inform future engine selection, subject to confirmation of CEEMDAN's operator readability. ZAMD offers deterministic, non-iterative computation and lower computational cost; CEEMDAN offers adaptive decomposition robust to non-stationary noise. Both are intended as complements to machine-learning bearing diagnostics rather than replacements, providing the intrinsic-interpretability layer that purely classifier-based systems lack.

The broader observation is that the explainability requirement of industrial condition monitoring can be addressed not only by post hoc explanation of ML decisions, but by re-deploying classical signal-processing methods that produce intrinsically interpretable diagnostic artefacts—and that, when wrapped with appropriate contextual annotation, the choice among such methods can reduce to a matter of engineering preference rather than diagnostic outcome, provided that each candidate engine's operator readability has been independently established, as demonstrated here for the ZAMD-based framework and identified as future work for CEEMDAN.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/eng7070342/s1>, Table S1. Manufacturer defect-characterisation certificate (JTEKT/Koyo Noise Conversion Table)—supplied with the pre-damaged KOYO 6302 test bearings, characterising each fault specimen (inner-race, outer-race, and ball defects, plus healthy references) by its standardised acceleration, velocity, decibel, and Anderon values across low-, mid-, and high-frequency bands. Figure S1. Photograph of the JTEKT/Koyo “Noise Conversion Table” supplied with the pre-damaged KOYO 6302 test bearings.

Author Contributions: Conceptualisation, C.T.; methodology, C.T.; software, C.T.; validation, C.T. and C.D.; formal analysis, C.T.; investigation, C.T.; data curation, C.T.; writing—original draft preparation, C.T.; writing—review and editing, C.T., C.D. and A.K.; visualization, C.T.; supervision, C.D. and A.K.; project administration, C.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. The readability assessment was an informal pilot evaluation conducted with adult volunteers and did not collect personally identifying information, biometric data, sensitive personal data, or any data permitting re-identification of individual participants. Only aggregate classification outcomes and aggregate response-time statistics were retained. Under these conditions, formal institutional review board approval was not required.

Informed Consent Statement: Informed consent was obtained from all participants involved in the readability assessment. Participants were informed of the study's purpose, the anonymous nature of the data collected, and their right to withdraw at any time without consequence.

Data Availability Statement: The vibration data underlying this study are derived from the laboratory bearing rig described in Section 4 and are available from the corresponding author on reasonable request. Aggregate readability-assessment outcomes are reported in full in Section 5.3, Table 2; individual participant records are not shared to preserve anonymity.

Acknowledgments: The authors thank the participants of the readability assessment for their time and the International Hellenic University (IHU) for laboratory access. The lead author acknowledges the doctoral work conducted at the University of Kragujevac that underpins the experimental rig, bearing dataset, and the original ZAMD and CEEMDAN implementations used in this study.

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

Abbreviations

The following abbreviations are used in this manuscript:

BPFI	Ball-Pass Frequency Inner race
BPFO	Ball-Pass Frequency Outer race
BSF	Ball Spin Frequency
CBM	Condition-Based Maintenance
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
CWRU	Case Western Reserve University
CWT	Continuous Wavelet Transform
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
FFT	Fast Fourier Transform
FTF	Fundamental Train Frequency
HHT	Hilbert–Huang Transform
HSA	Hilbert Spectral Analysis
IMF	Intrinsic Mode Function
ML	Machine Learning
RUL	Remaining Useful Life
SNR	Signal-to-Noise Ratio
STFT	Short-Time Fourier Transform
VMD	Variational Mode Decomposition
WVD	Wigner–Ville Distribution
XAI	Explainable Artificial Intelligence
ZAMD	Zhao–Atlas–Marks Distribution

References

1. Randall, R.B.; Antoni, J. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [\[CrossRef\]](#)
2. Tandon, N.; Choudhury, A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribol. Int.* **1999**, *32*, 469–480. [\[CrossRef\]](#)
3. Zhao, Y.; Atlas, L.E.; Marks, R.J. The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1084–1091. [\[CrossRef\]](#)
4. Cohen, L. *Time-Frequency Analysis*; Prentice Hall: Englewood Cliffs, NJ, USA, 1995.
5. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.-C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **1998**, *454*, 903–995. [\[CrossRef\]](#)
6. Wu, Z.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [\[CrossRef\]](#)
7. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the IEEE ICASSP, Prague, Czech Republic, 22–27 May 2011; pp. 4144–4147. [\[CrossRef\]](#)
8. Zhou, Y.; Ma, Z.; Fu, L. A review of key signal processing techniques for structural health monitoring: Highlighting non-parametric time-frequency analysis, adaptive decomposition, and deconvolution. *Algorithms* **2025**, *18*, 318. [\[CrossRef\]](#)
9. Wang, X.; Li, J.; Jing, Z.; Li, H.; Xing, Z.; Yang, Z.; Cao, L.; Zhou, X. Fault diagnosis method of rolling bearing based on SSA-VMD and RCMDE. *Sci. Rep.* **2024**, *14*, 30637. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Wang, S.; Wang, C.; Lian, Y.; Luo, B. Research on bearing fault diagnosis based on VMD-RCMWPE feature extraction and WOA-SVM-optimised multidataset fusion. *Sensors* **2025**, *25*, 5139. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Wang, L.; Li, H.; Xi, T.; Wei, S. Fault feature extraction method for rolling bearings based on CEEMDAN and VMD. *Sensors* **2023**, *23*, 9441. [\[CrossRef\]](#) [\[PubMed\]](#)

12. Li, Y.; Sheng, J.; Liu, J.; Liu, S. Fault diagnosis and classification of rolling bearings using ICEEMDAN–CNN–BiLSTM and acoustic emission. *Sensors* **2026**, *26*, 507. [[CrossRef](#)] [[PubMed](#)]
13. Sanakkayala, D.C.; Varadarajan, V.; Kumar, N.; Karan; Soni, G.; Kamat, P.; Kumar, S.; Patil, S.; Kotecha, K. Explainable AI for bearing fault prognosis using deep learning techniques. *Micromachines* **2022**, *13*, 1471. [[CrossRef](#)] [[PubMed](#)]
14. Zhang, Q.; Wei, X.; Wang, Y.; Hou, C. Convolutional neural network with attention mechanism and visual vibration signal analysis for bearing fault diagnosis. *Sensors* **2024**, *24*, 1831. [[CrossRef](#)] [[PubMed](#)]
15. Bao, P.; Yi, W.; Zhu, Y.; Shen, Y.; Peng, H. A spectral interpretable bearing fault diagnosis framework powered by large language models. *Sensors* **2025**, *25*, 3822. [[CrossRef](#)] [[PubMed](#)]
16. Tsiafis, C.; Zaharis, Z.; Xanthopoulou, M.; Skeberis, C.; Tsiafis, I.; Todorovic, P.; Xenos, T. Detection of non-linear signal distortions due to external impulse stimulations in rolling bearing experimental device. *J. Balk. Tribol. Assoc.* **2015**, *21*, 233–245.
17. Tsiafis, C.; Giagloglou, E.; Skeberis, C.; Tsiafis, I.; Todorovic, P.; Xenos, T. Avoiding human error through easy readable spectra provided by Zhao–Atlas–Marks distribution for monitoring rolling-element bearings. In Proceedings of the 49th ESReDA Seminar, Brussels, Belgium, 29–30 October 2015.
18. Tsiafis, C.; Zaharis, Z.D.; Xanthopoulou, M.; Skeberis, C.; Tsiafis, I.; Todorovic, P.; Xenos, T. A novel methodology to detect non-linear signal distortions due to external impulse stimulations in ball bearing fault detection setup. In Proceedings of the BalkanTrib'14, Sinaia, Romania, 30 October–1 November 2014; pp. 245–253.
19. Tsiafis, C. An effective method for monitoring rolling bearings wear status via the application of noise-assisted Hilbert–Huang transform. *J. Balk. Tribol. Assoc.* **2021**, *27*, 826–841.
20. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64–65*, 100–131. [[CrossRef](#)]
21. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
22. Hendriks, J.; Dumond, P.; Knox, D.A. Towards better benchmarking using the CWRU bearing fault dataset. *Mech. Syst. Signal Process.* **2022**, *169*, 108732. [[CrossRef](#)]
23. Mishra, R.; Singh, S.K.; Kumar, R. AI-driven predictive maintenance for industry 4.0: A systematic review of models, methods, and challenges. *Int. J. Adv. Manuf. Technol.* **2026**, *143*, 4861–4876. [[CrossRef](#)]
24. Rosa, R.K.; Braga, D.; Silva, D. Benchmarking deep learning models for bearing fault diagnosis using the CWRU dataset: A multi-label approach. *arXiv* **2024**, arXiv:2407.14625. [[CrossRef](#)]
25. Pham, M.; Kim, J.-M.; Kim, C.H. A comparative study of time-frequency representations for bearing and rotating fault diagnosis using Vision Transformer. *Machines* **2025**, *13*, 737. [[CrossRef](#)]
26. Loparo, K.A. Bearing Data Center. Case Western Reserve University. Available online: <https://engineering.case.edu/bearingdatacenter> (accessed on 24 May 2026).

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.