Journal of Sound and Vibration 599 (2025) 118909

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Raw sensor data fusion using Johansen cointegration for condition assessment of concrete poles

Mohsen Mousavi [a,*], Ulrike Dackermann [a], Sahar Hassani [a], Mahbube Subhani [b], Amir H. Gandomi [c,d]

[a] School of Civil and Environmental Engineering, The University of New South Wales, Sydney 2052, Australia
[b] School of Engineering, Deakin University, Waurn Ponds, VIC, 3216, Australia
[c] Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia
[d] University Research and Innovation Center (EKIK), Óbuda University, 1034 Budapest, Hungary

ARTICLE INFO

Keywords:
Johansen cointegration
Variational Mode Decomposition
Raw signal fusion
Statistical features
Non-destructive testing
Normalized mutual information

ABSTRACT

This paper presents a novel approach for raw sensor data fusion using Johansen cointegration, aimed at non-destructive condition assessment of concrete poles. The proposed Johansen cointegration-based signal fusion is compared with signal averaging, a conventional method, and the Adaptive Kalman Filter (AKF), an advanced signal fusion technique. These methods are applied to data collected from concrete poles under both laboratory and real-world field conditions, using an innovative narrow-band stress wave excitation system with a center frequency of 1 kHz. Our methodology begins with fusing raw sensor data, which is subsequently decomposed into narrow-band components, known as Intrinsic Mode Functions (IMFs), using the Variational Mode Decomposition (VMD) algorithm. From these IMFs, we extract a set of non-parametric and parametric statistical features based on Instantaneous Frequency (IF) and Instantaneous Amplitude (IA) signals. The results demonstrate the superiority of Johansen cointegration over both signal averaging and AKF in scenarios involving the high nonstationarity characteristic of real-world field data. Furthermore, the findings highlight a notable similarity between AKF and signal averaging, which may reflect the dominant linear properties in the recorded signals. We also propose an index based on normalized mutual information to facilitate a fair comparison with existing fusion methods.

1. Introduction

Over the past decades, significant advancements have been made in the development of sensor technologies and analysis techniques employed for Structural Health Monitoring (SHM) and Nondestructive Testing (NDT) [1]. Numerous industries, such as the construction, manufacturing, aerospace, automotive, and energy sectors, have benefited from these advancements. The rapid progress of data science technologies has further contributed to innovative SHM techniques that can address the limitations of classical physics-based approaches. These advancements facilitated the feature extraction and fusion of data for solving classification and regression problems. However, a gap still exists in the systematic and robust fusion of data from multiple sensors.

Different hierarchies of information fusion have been employed for SHM and NDT problems including:

1. data acquisition, raw data fusion, feature extraction, decision making using machine learning results,

* Corresponding author.
E-mail address: mohsen.mousavi@unsw.edu.au (M. Mousavi).

https://doi.org/10.1016/j.jsv.2024.118909
Received 21 May 2024; Received in revised form 2 December 2024; Accepted 5 December 2024
Available online 13 December 2024

2. data acquisition, feature extraction, feature fusion, decision making using machine learning results, and
3. data acquisition, feature extraction, decisions making using several machine learning results, and decision fusion [2].

SHM systems involve the deployment of a sensory network that measures the response of a structure to natural or synthesized forces. These feedback signals contain valuable information on the interior integrity of structural elements. However, the recorded raw sensor data is often complex and polluted with noise and outliers, making it challenging to extract meaningful information. Raw data fusion, also classified as "high level information extraction" [3], aims to combine structural responses at different locations to enhance the statistical properties of the data. Therefore, this type of data fusion is an important step in the analysis of measurement signals improving the performance of the developed SHM system. Signal averaging is a common time-domain signal processing approach aimed at strengthening the signal to noise ratio (SNR) [4]. This type of raw sensor data fusion relies on averaging a set of replicate measurements, which proportionally increases the signal-to-noise ratio (SNR) relative to the square root of the number of measurements. This approach has been widely applied in various fields, including neurology [5], robotics [6,7], biomedical engineering [8], and fiber interferometers [9]. In the context of SHM, signal averaging has been employed to improve damage detection results [10]. Although signal averaging can lead to improved results, the literature lacks a comprehensive analysis of when this technique is most appropriate for SHM applications. Recent studies have begun to explore this gap by comparing signal averaging with alternative techniques [11], aiming to determine the specific conditions under which each method performs best. However, there is still a gap in determining the appropriate circumstances for the beneficial utilization of signal averaging [5]. Moreover, signal averaging is limited in its ability to handle nonstationary signals, as it lacks the capability to capture and account for the dynamic characteristics inherent in such signals. This limitation makes it less effective in scenarios where the signal properties vary over time. Nonstationary signals often exhibit time-varying statistical properties, making them challenging to analyze and fuse using traditional methods like signal averaging. One alternative approach for nonstationary signal fusion is the Adaptive Kalman Filter (AKF) [12,13]. Unlike signal averaging, the AKF is capable of dynamically adjusting its parameters based on the evolving characteristics of the signals being processed. This adaptability makes the AKF well-suited for handling the varying statistical properties of nonstationary signals over time.

After raw data fusion, the next stage in the proposed method involves feature extraction. While deep learning methods can automatically perform feature extraction in their intermediate layers, their effective application requires the availability of a large dataset. This dependency on extensive data limits their practicality in scenarios where only limited data is available. Hence, it is commendable to extract hand-crafted features when the amount of data samples per class is limited. Accordingly, feature extraction is a crucial step in most SHM problems, as it aims to derive relevant information from raw sensor data. This extracted information can then be utilized to address classification or regression problems in SHM. Various types of hand-crafted features are commonly employed in machine learning algorithms, including time-domain, frequency-domain, and time-frequency attributes. These features help capture different aspects of the structural response, enhancing the model's ability to detect patterns and anomalies. Time-domain features are directly extracted from sensor data in the time domain [14]. These features can be categorized as coefficient-based or residual-based algorithms and are set to capture the signal's behaviors in the time domain [15]. Frequency-domain features, on the other hand, include characteristics of sensor data such as frequency distribution and spectral content using Fourier transform. Some features include power spectral density, dominant frequencies, energy content in specific frequency bands, resonance frequencies, and mode shapes. Time-frequency features are used to analyze signals in the time-frequency domain and include signal decomposition algorithms such as Short-Time Fourier Transform (STFT), Wavelet Transform (WT), Empirical Mode Decomposition (EMD), and Variational Mode Decomposition (VMD) [16]. These techniques offer joint time-frequency representations of the original signal, either directly or indirectly, allowing for a more detailed analysis of nonstationary signals by capturing both time-varying and frequency-varying characteristics.

Several feature extraction strategies have been employed for SHM to capture significant attributes that can facilitate the learning of patterns in the training phase. Statistical features extracted from sensor data can provide valuable information about the distribution and variability of the statistical behavior of the data. These features include mean, standard deviation, skewness, kurtosis–higher-order moments–of the time-frequency representation of the signal such as instantaneous frequency (IF) and Instantaneous Amplitude (IA) of signal decomposition results [17]. These features are sensitive to changes in the system affecting the distribution of statistical moments in IF and IA signals. Therefore, they are useful for finding system changes caused by damage.

When a stress wave is transmitted through a medium, it interacts with the material and boundaries, resulting in several effects such as wave attenuation, dispersion, and absorption. These effects can generate higher-frequency modulations in the signal which contain valuable information about the interior of the studied specimen. Therefore, it is advantageous to analyze these modulations in isolation. To achieve this, the signal needs to be decomposed into its constituent oscillation modes. However, the choice of the decomposition algorithm is of significant importance. For instance, Fourier transformation lacks the ability to provide resolution regarding a signal's evolution over time, making it unsuitable for extracting descriptive information from nonstationary signals. On the other hand, while wavelet transformation can preserve nonstationary information through its decomposition process and provide time resolution, it is a linear time-frequency signal decomposition algorithm. Consequently, it lacks the ability to capture nonlinear effects during the decomposition process.

Huang et al. [18] introduced the Empirical Mode Decomposition (EMD) algorithm as a signal decomposition technique capable of preserving both nonstationary and nonlinear information throughout the decomposition process. Consequently, EMD has been increasingly adopted as a novel alternative to traditional algorithms in the field of SHM, offering improved accuracy in analyzing complex structural responses. This includes the introduction of Intrinsic Mode Functions (IMFs), which are extracted from the EMD process. IMFs exhibit several key properties, such as being either frequency or amplitude modulated, making them particularly

effective in representing complex, nonstationary signals by isolating underlying oscillatory modes. Moreover, EMD is an adaptive signal decomposition algorithm in a sense that the sum of its extracted IMFs can perfectly reconstruct the original signal. However, EMD has certain limitations, such as the presence of end effects and mode mixing phenomena [19]. These limitations have motivated researchers to develop new approaches that are robust against these issues.

Variational Mode Decomposition (VMD) has emerged as another adaptive nonlinear/nonstationary signal decomposition algorithm that can address the shortcomings of EMD. VMD employs a variational optimization process to update the IMFs and their corresponding center frequencies [20]. The IMFs extracted from VMD are narrow-band signals, each representing a distinct oscillation mode of the original signal. Additionally, VMD is a parametric signal decomposition algorithm, which means it requires the pre-specification of certain parameters before it can be executed. These parameters can be determined by incorporating the user's domain knowledge regarding the specific problem or by tuning them through an optimization process tailored to the problem at hand. While the former approach demands extensive domain knowledge, the latter is prone to overfitting on the training set, especially when the available data is limited. One key parameter in VMD is the number of IMFs. Choosing a large value for this parameter can result in the duplication of higher-order IMFs [20].

Two primary characteristic signals, namely the IF and IA signals, can be derived from the IMFs extracted through VMD. These IF and IA signals contain valuable information about the properties of the original signal and, by extension, the tested specimen. In their previous work [17,21,22], the authors introduced several statistical features extracted from the IF signals of the IMFs obtained through VMD. However, these statistical properties are derived based on the simplified assumption that all IMFs are stationary–similar to the assumption made by the Naive Bayes classifier, which assumes independence of all features within a given class [23]. Therefore, the objective of this paper is to address the limitations of this assumption by enhancing the stationarity of the IMFs using the Johansen cointegration method as a fusion technique.

Johansen cointegration is a concept adopted from econometrics that aims to identify a linear stationary combination of a set of nonstationary signals [24]. The fundamental assumption underlying Johansen cointegration is that the signals share the same order of nonstationarity. However, this assumption may not always hold true in certain SHM problems. Therefore, for simplicity, it has been suggested to overlook this assumption, as the primary objective in SHM problems is to facilitate damage detection [25]. In this paper, we propose to use Johansen cointegration as a signal fusion algorithm to improve the stationary quality of the fusion outcome. By applying the Johansen cointegration technique, a linear combination of a set of nonstationary, or weakly stationary, signals is sought that exhibits enhanced stationary properties compared to the original signals. This approach aims to maximize the stationary characteristics of the IMFs extracted from VMD. For comparison, we benchmark the results of the Johansen cointegration approach against conventional signal averaging and a more advanced adaptive Kalman filter (AKF) approach, both used as raw signal fusion techniques. Moreover, utilizing VMD as a feature extraction algorithm requires setting several hyperparameters. Although an optimization algorithm can be employed to identify the optimal values for these parameters, this can lead to overfitting on the training data, particularly when the available data is limited. Conversely, since the primary objective of this work is to compare different signal fusion techniques, it is crucial to ensure a fair comparison between models trained on features extracted from fused signals using various methodologies. To this end, we introduce a novel metric based on normalized mutual information (NMI) to evaluate and compare the amount of information extracted from signals across different fusion techniques. The key contributions of this work can thus be summarized as follows:

1. We propose the Johansen cointegration approach as an effective signal fusion algorithm that significantly outperforms both the conventional signal averaging and the more advanced AKF techniques in scenarios where nonstationary signals outnumber stationary ones. As a result, we demonstrate that the Johansen cointegration method is a reliable and fair technique for fusing nonstationary signals, especially when compared to alternative approaches.
2. We propose a metric based on NMI to quantify the information retained during the decomposition process. By comparing the original signal with its VMD-reconstructed counterpart, this metric evaluates the effectiveness of different fusion techniques in preserving critical signal information, enabling a fair and comprehensive comparison of these methods.
3. We explore the potential of combining various signal fusion techniques and provide a roadmap for future research directions in this field.

The remaining of the paper is organized as follows:

Section 2 briefly elaborates on the background theories, including feature extraction using the VMD (Section 2.1), a novel metric for assessing the amount of information passed through the signal decomposition to assess fairness in comparing signal fusion techniques (Section 2.2), and cointegration technique for nonstationary signal fusion (Section 2.3). The section provides insights into how the extracted features from the VMD may not capture all the nuances of the data dynamics, especially when the signals exhibit severe nonstationary properties, proposing Johansen cointegration as an effective raw signal fusion algorithm. Section 3 presents the case study of the concrete poles classification problem using nondestructive tests carried out in both laboratory and field conditions. Section 4 provides details on the adapted methodologies in this study, including the XGBoost algorithm as an effective Machine learning technique for solving the classification problem of this paper (Section 4.2), the Monte Carlo Cross Validation (MCCV) algorithm for effective assessment of the XGBoost results (Section 4.3), and the Adaptive Kalman Filter (AKF) as a more advanced alternative method for raw signal fusion for comparison (Section 4.4). Section 5 investigates the problem of this paper in full detail, starting with investigating signal statistical properties in both laboratory and in-situ field data (Section 5.1), evaluating the effect of VMD parameters on the proposed NMI metric (Section 5.2), solving the laboratory problem (Section 5.3), and the field problem (Section 5.4). Finally, Sections 6 and 7 present a road-map for future work and draw concluding remarks from the present study, respectively.

2. Background theories

2.1. Feature extraction using VMD

This section details the feature extraction process from a given signal $\mathbf{f}(t)$, using the variational mode decomposition (VMD) algorithm.

According to wave propagation theory, the probing signal undergoes modulations as it interacts with material irregularities, including density variations, boundary conditions, or internal imperfections such as damage. Consequently, a wideband signal is generated. To extract the relevant information contained in different signal bands, we propose demodulating the signal into multiple narrowband oscillation modes using the Variational Mode Decomposition (VMD) algorithm as a signal decomposition technique. We further generate Gabor's analytic signal [26] to construct the IF and IA signals that are well-defined for these narrowband components. Finally, we extract the proposed features through the procedures discussed below.

First, the signal is decomposed into its constituent oscillation modes by utilizing VMD [20]. While the theory of VMD is extensive, this section provides a concise summary of its foundational principles.

The core properties of VMD can be summarized as follows:

- VMD involves solving a variational optimization problem.
- It aims to decompose a nonlinear or nonstationary signal into a predetermined number of IMFs.
- An IMF derived from VMD is a narrowband signal that represents an oscillation mode of the original signal.

The kth IMF ($\mathbf{u}_k(t)$) obtained from VMD can be described as an amplitude- and phase-modulated signal as follows:

$$\mathbf{u}_k(t) = \mathbf{A}_k(t)\ \cos(\boldsymbol{\phi}_k(t)), \tag{1}$$

where $\mathbf{A}_k(t)$ and $\boldsymbol{\phi}_k(t)$ represent the instantaneous amplitude and phase, respectively. The IF signal can be calculated by differentiating the instantaneous phase with respect to time as $\boldsymbol{\omega}(t) = \frac{\partial \boldsymbol{\phi}(t)}{\partial t}$.

The Gabor's analytic signal ($\mathbf{u}_a(t)$) is constructed as follows:

$$\mathbf{u}_a(t) = \mathbf{u}(t) + j\hat{\mathbf{u}}(t), \tag{2}$$

where j and $\hat{\mathbf{u}}(t)$ denote the imaginary unit and Hilbert transformation [27] of the narrow-band IMF $\mathbf{u}(t)$, respectively. The IA and IF signals can then be derived as follows:

$$\mathrm{IA}(t) = \sqrt{u^2(t) + \hat{u}^2(t)}, \tag{3}$$

$$\mathrm{IF}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\tan^{-1}\left(\frac{\hat{u}(t)}{u(t)}\right)\right), \tag{4}$$

As mentioned earlier, VMD solves an optimization problem by formulating an augmented Lagrangian optimization to determine the IMFs $\mathbf{u}_k$ and their corresponding center frequencies $\boldsymbol{\omega}_k$ for $k = 1, \ldots, p$, where p denotes the prespecified number of IMFs. The objective function of the VMD can be written as follows [20]:

$$\begin{aligned}\mathcal{L}(\mathbf{u}_k, \boldsymbol{\omega}_k, \lambda) = \alpha \sum_k \left\| \partial_t \left(\delta(t) + \frac{j}{\pi t} * \mathbf{u}_k(t) \right) \times e^{-j\omega_k t} \right\|_2^2 \\ + \left\| \mathbf{f}(t) - \sum_k \mathbf{u}_k(t) \right\|_2^2 + \left\langle \lambda(t), \mathbf{f}(t) - \sum_k \mathbf{u}_k(t) \right\rangle.\end{aligned} \tag{5}$$

where in above equation:

- $\mathbf{u}_k(t)$: The mode components (or signals) extracted by the VMD algorithm from the input signal. Each mode $\mathbf{u}_k(t)$ is a band-limited signal centered around its corresponding frequency $\boldsymbol{\omega}_k$.
- $\boldsymbol{\omega}_k$: The center frequency of each mode $\mathbf{u}_k(t)$, adaptively estimated during the VMD decomposition process.
- α: A balancing parameter that controls the bandwidth constraint of the modes. Higher values of α impose stricter band limitations on the modes, effectively reducing the likelihood of capturing higher-frequency noise or other irrelevant components. As a result, this can lead to a form of denoising.
- ∂_t: The time derivative operator, which indicates temporal differentiation of the mode $\mathbf{u}_k(t)$, emphasizing changes in the mode over time.
- $\delta(t)$: The Dirac delta function, often used in conjunction with the Hilbert transform to construct the analytic signal.
- $\frac{j}{\pi t} * \mathbf{u}_k(t)$: A term representing the Hilbert transform of the mode $\mathbf{u}_k(t)$, used to determine the instantaneous frequency and phase of the mode.
- $e^{-j\omega_k t}$: A complex exponential (phasor) that shifts the mode $\mathbf{u}_k(t)$ to its center frequency ω_k, localizing the mode in the frequency domain.
- $\|\cdot\|_2^2$: The squared L_2-norm (or Euclidean norm), representing the energy of the signal in the term. As such, $\|\partial_t(\cdot)\|_2^2$ measures the energy (or bandwidth) of the time derivative of the transformed signal.

Table 1
The description of features naming.

Index	Features	Description
1	p1_IF$_i$	First quartile of IF for IMF$_i$
2	p2_IF$_i$	Second quartile of IF for IMF$_i$
3	p3_IF$_i$	Third quartile of IF for IMF$_i$
4	k_IF$_i$	Kurtosis of IF for IMF$_i$
5	v_IF$_i$	Variance of IF for IMF$_i$
6	cf_IF$_i$	Center frequency of IMF$_i$
7	sk_IF$_i$	Skewness of IF for IMF$_i$
8	p1_IA$_i$	First quartile of IA for IMF$_i$
9	p2_IA$_i$	Second quartile of IA for IMF$_i$
10	p3_IA$_i$	Third quartile of IA for IMF$_i$
11	k_IA$_i$	Kurtosis of IA for IMF$_i$
12	v_IA$_i$	Variance of IA for IMF$_i$
13	sk_IA$_i$	Skewness of IA for IMF$_i$

- **First term:**

$$\alpha \sum_k \left\| \partial_t \left(\delta(t) + \frac{j}{\pi t} * \mathbf{u}_k(t) \right) \times e^{-j\omega_k t} \right\|_2^2$$

This term penalizes the bandwidth of each mode $\mathbf{u}_k(t)$, ensuring that each mode is band-limited. The bandwidth is controlled by the parameter α, and the goal is to minimize it to ensure each mode occupies a specific frequency range.

- **Second term:**

$$\left\| \mathbf{f}(t) - \sum_k \mathbf{u}_k(t) \right\|_2^2$$

This is the reconstruction error term, which measures how well the sum of all modes $\mathbf{u}_k(t)$ reconstructs the original signal $\mathbf{f}(t)$. The goal of VMD is to minimize this error and ensure that the sum of the modes approximates the input signal as closely as possible.

- **Third term:**

$$\left\langle \lambda(t), \mathbf{f}(t) - \sum_k \mathbf{u}_k(t) \right\rangle$$

This term represents the Lagrange multiplier, enforcing the constraint that the sum of all modes $\mathbf{u}_k(t)$ should reconstruct the original signal $\mathbf{f}(t)$. $\lambda(t)$ is the Lagrange multiplier that enforces this constraint in the optimization process.

As a parametric decomposition algorithm, VMD requires the pre-specification of several parameters prior to execution [28]. For practical, real-life problems, such as the assessment and monitoring of civil infrastructure, typically only a minimal set of data from in-situ structures is available. Here, we propose a method to measure the amount of information transferred across all IMFs by adjusting various VMD parameters. The suggested approach relies on normalized mutual information, and will be explained further in the next section.

Next, the IF and IA signals corresponding to the ith IMF are constructed based on the Gabor's analytic signal. A set of statistical features are extracted from these signals, as listed below:

1. The first quartile of the IF signal.
2. The second quartile of the IF signal.
3. The third quartile of the IF signal.
4. The center frequency (ω) of the IMF.
5. The Kurtosis of the IF signal.
6. The variance of the IF signal.
7. The skewness of the IF signal.
8. The first quartile of the IA signal.
9. The second quartile of the IA signal.
10. The third quartile of the IA signal.
11. The Kurtosis of the IA signal.
12. The variance of the IA signal.
13. The skewness of the IA signal.

Notably, features 1, 2, 3, 8, 9, and 10 can be classified as nonparametric statistical features and the remaining features are parametric. The extracted features are labeled as presented in Table 1.

2.2. Normalized Mutual Information (NMI)

This section evaluates the extent to which information is retained during the decomposition process when specific values are assigned to the VMD parameters. VMD is recognized as an adaptive decomposition algorithm, which ensures that the sum of the decomposed IMFs can fully reconstruct the original signal. However, this property holds true only if denoising is not applied during the decomposition process. On the other hand, setting the denoising factor α to a relatively large value may compromise the adaptiveness of VMD by reducing the amount of information, treated as noise, that passes through the decomposition process. Here, we propose a technique based on normalized mutual information (NMI) to assess the impact of different VMD parameters on the amount of information transmitted to the decomposed IMFs.

The NMI distance between the signal $\mathbf{f}(t)$ and the sum of its extracted IMFs is defined as follows:

$$\mathrm{NMI}\left(\mathbf{f}(t), \sum_k \mathbf{u}_k(t)\right) = \frac{\mathrm{I}\left(\mathbf{f}(t), \sum_k \mathbf{u}_k(t)\right)}{\min\left(\mathrm{H}(\mathbf{f}(t)), \mathrm{H}(\sum_k \mathbf{u}_k(t))\right)} \tag{6}$$

where H(.) and I(.) denote the Shannon Entropy and mutual information, respectively. The Shannon Entropy of variable $\mathbf{X}$ and the mutual information between two variables $\mathbf{X}$ and $\mathbf{Y}$ are calculated as follows:

$$\mathrm{H}(\mathbf{X}) = -\sum_{x\in\mathbf{X}} \mathrm{P}(x)\ \log\ \mathrm{P}(x) \tag{7}$$

and

$$\mathrm{I}(\mathbf{X};\mathbf{Y}) = \sum_{y\in\mathbf{Y}}\sum_{x\in\mathbf{X}} \mathrm{P}_{(\mathbf{X},\mathbf{Y})}(x,y)\ \log\left(\frac{\mathrm{P}_{(\mathbf{X},\mathbf{Y})}(x,y)}{\mathrm{P}_{\mathbf{X}}(x)\mathrm{P}_{\mathbf{Y}}(y)}\right) \tag{8}$$

where, in above equations, $\mathrm{P}_{\mathbf{X}}(x)$, $\mathrm{P}_{\mathbf{Y}}(y)$, and $\mathrm{P}_{(\mathbf{X},\mathbf{Y})}(x,y)$ denote the marginal probability of $\mathbf{X}$ and $\mathbf{Y}$ and their joint probability, respectively. The procedures discussed in [29,30] is followed to perform the above-mentioned calculations. Here, the role of the denominator is to scale the NMI value to the range of [0, 1].

2.3. Cointegration

There is a direct correlation between the concept of Cointegration (CI) and the notation of stationary/nonstationary signals, as discussed by Doob [31]. This section provides a concise definition of the stationarity and nonstationarity in time series based on the works of Zolna et al. [32] and Dao and Staszewski [33].

The first order auto-regressive process $AR(1)$ of the signal $X(t)$ is obtained as follows:

$$X(t) = \phi\ X(t-1) + \epsilon_t \tag{9}$$

where ϵ_t represents a white Gaussian noise process which is stationary. Exploring the possible conditions on ϕ, we encounter three possible scenarios, detailed as follows:

1. $|\phi| < 1$: this case implies that the signal $X(t)$ is stationary,
2. $|\phi| > 1$: refers to a nonstationary signal, and
3. $|\phi| = 1$: corresponds to a pure random walk signal. As the variance diverges when $t \to \infty$, the signal is regarded as nonstationary.

Definition 2.1. A time series is said to be a unit root process if its characteristic equation has a root equal to one.

Definition 2.2. Nonstationarity can also arise from a deterministic trend in a time series, in which case the time series is called a trend stationary process.

The characteristic equation of a time series can be analyzed to determine the presence of unit roots. The number of unit roots n indicates the degree of nonstationarity in the series, which is expressed in terms of its integration order, denoted as $I(n)$. When the difference operator is applied to the time series, it cancels out the unit root effect and makes the series difference stationary. For instance, consider a first-order nonstationary pure random walk process, which is classified as $I(1)$. By applying the delta operator, we obtain $\Delta X(t) = X(t) - X(t-1) = \epsilon_t$, which represents a stationary white Gaussian noise process denoted by $I(0)$. Other definitions of stationary and nonstationary signals exist. For instance, statistical moments of a stationary signal are time-invariant, whereas the statistical moments of a nonstationary signal are time-variant.

The extraction of statistical moments (features) from different IMFs in this paper assumes that the signals, and hence the extracted IMFs, are stationary. While this assumption simplifies the feature extraction process, it may not necessarily hold true for all types of signals. Hence, we propose using the Johansen cointegration technique to eliminate any nonstationary trends in the recorded signals. We hypothesize that this will lead to improved classification outcomes, particularly as the proportion of nonstationary signals increases.

Note that although differentiating a nonstationary signal can potentially eliminate any nonstationary trend, it is not a recommended practice. Cross and Worden [25] presented two justifications for using cointegration in SHM rather than directly working with the difference operator:

1. Firstly, differentiating a time series can amplify the effect of high-frequency noise components. In the worst case scenario, cointegration will produce a weighted average of noise.
2. Secondly, differentiation can eliminate any trend in a signal that arises from damage.

In this paper, we propose to fuse the recorded signals using three technique: (1) by averaging signals as a traditional raw signal fusion, (2) by AKF method as a more advanced signal fusion technique theorized to be capable of fusing nonstationary signals, and (3) by using the cointegration technique as the proposed raw signal fusion algorithm suitable for fusing highly nonstationary signals. Our hypothesis is that the latter technique will be more effective when dealing with signals that exhibit more nonstationary properties. To investigate this, we apply the Johansen cointegration along with the other methods for fusing the laboratory and field signals and further employ VMD for feature extraction out of the fused signals. We finally use the extracted features for solving the classification problem of poles as either damaged or intact. In the following, we present fundamental aspects of cointegration theory.

Cointegration is a statistical technique that has found applications in several fields of engineering including condition monitoring [21,25,34] and colored noise cancellation in signals used for SHM [35]. Johansen cointegration can be regarded as a signal fusion technique that aims to obtain a linear stationary combination of a set of nonstationary signals. The following basic definitions are provided to enhance understanding of the Johansen cointegration procedures.

Definition 2.3. A time series $X(t)$ is deemed integrated of order d, denoted as $I(d)$, if $\Delta^d X(t)$ renders stationary time series. Here, d indicates the number of times the Δ operator needs to be applied to the time series to make it stationary, where $\Delta X(t) = X(t) - X(t-1)$.

Definition 2.4. A set of m signals $X_1(t), \dots, X_m(t)$ is said to be cointegrated with order d and b if they satisfy the following conditions:

1. Each signal $X_i(t)$, $i = 1, \dots, m$, is integrated of order d.
2. There exists a linear combination of the signals $X_i(t)$:

$$\Psi(t) = a_1\ X_1(t) + a_2\ X_2(t) + \cdots + a_m\ X_m(t) \tag{10}$$

such that $\Psi(t)$ is integrated of order $d - b$. In this case, the set of time series $X_1(t), \dots, X_m(t)$ is denoted as CI (d, b).

The most encountered scenario is when both d and b are equal to 1. In this scenario, the cointegration vector, denoted by $[a_1, \dots, a_m]$, can be obtained using a least squares optimization algorithm [36].

The Kwiatkowski-Phillips–Schmidt–Shin (KPSS) [37] and Augmented Dickey–Fuller (ADF) [38] tests are two widely used statistical methods in time series analysis to detect the presence of a unit root in a signal. These tests indicate how many times the difference operator, Δ, should be applied to a time series to achieve stationarity. Detailed mathematical formulations of the Johansen cointegration procedure are provided in Appendix.

3. Case study

The proposed method is demonstrated by investigating the wave propagation problem through signals recorded from testing various concrete poles, each exhibiting distinct damage conditions and concrete mixtures. The poles were tested in both controlled laboratory conditions and real-world in-situ field conditions. While in the laboratory set-up, the poles were standing freely on a Styrofoam mat, in the field, they were embedded in soil and exposed to environmental factors. Such environmental factors can compromise the accurate diagnosis of damage, as changes in the signals may be influenced by both structural degradation and external environmental changes. As highlighted in recent studies [39,40], assessing the robustness of damage detection methods while considering the impact of environmental variability is a critical aspect of SHM.

Three of the poles were manufactured using self-compacting concrete (without steel reinforcement), while the remaining two were made of generic concrete (also lacking steel reinforcement). Within the sub-set of self-compacting concrete poles, one remained undamaged, another exhibited surface void damage and the third had internal honey-comb damage. As for the generic concrete poles, one was free from damage, and the other displayed surface void damage. All tested poles had a length of 3 m and a diameter of 0.25 m.

Fig. 1 shows photos of a pole tested in the laboratory and the field. The specifications of the poles studied, including damage types and material properties, are depicted in Fig. 2 and outlined in Table 2.

To determine the health condition of the poles, they were non-destructively tested using an innovative electro-mechanical excitation system consisting of tactile transducers driven by an amplifier to generate controlled sound waves. Tactile transducers are capable of transferring controlled sound waves (vibrations) through a structure. This way, a narrow-band frequency stress wave with a center frequency of 1 kHz was generated and the wave propagation responses of the poles were captured using a network of 12 accelerometers. The accelerometers were attached in three concentric rings, each containing four accelerometers. The sensors simultaneously recorded the wave response with a sampling rate of 1 MHz and a recording time of 0.06 s. Each pole configuration was tested five times to generate multiple data sets.

The testing equipment comprised the following components: four tactile transducers (Vidsonix, model VX-GH92); sensor wedges, along with curvature adapters, for directing the wave in the longitudinal direction of the structure; a Hi-Fi amplifier to amplify and adjust the wave amplitude; a function generator to produce the desired excitation waveform and frequency; a data acquisition

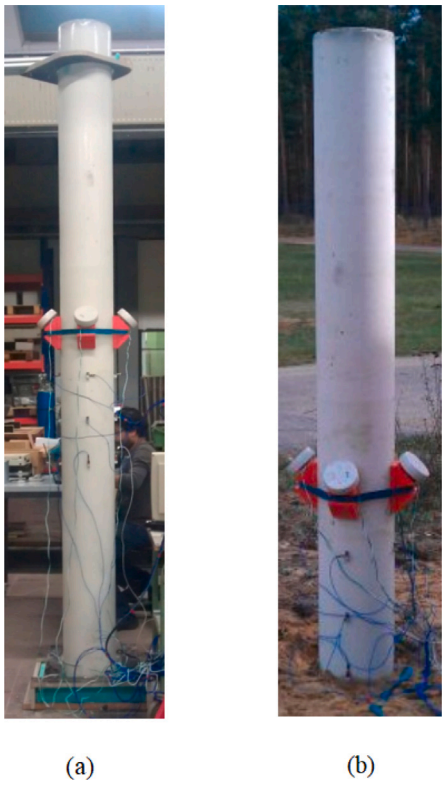


Fig. 1. Test set-up view in (a) the laboratory and (b) the field.

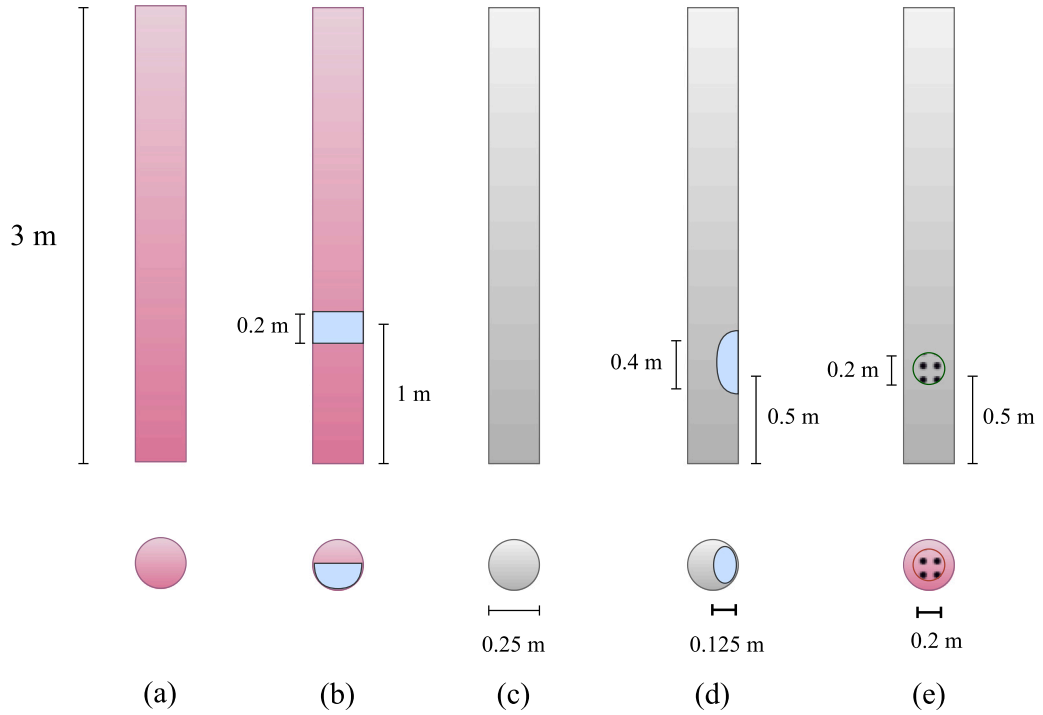


Fig. 2. Schematic view and size of the studied poles in the laboratory and field: (a) Intact Heterogeneous Concrete, (b) Damaged Heterogeneous Concrete (surface void), (c) Intact Homogeneous Concrete, (d) Damaged Homogeneous Concrete (surface void), and (e) Damaged Homogeneous Concrete (internal honeycomb).

Table 2
Types of poles and their damage scenarios tested in the laboratory and field.

Testing site	Material	Tag	Damage scenario
Laboratory	Generic concrete, heterogeneous	(a)	Intact
		(b)	Damaged, surface void
	Self-compacting concrete, homogeneous	(c)	Intact
		(d)	Damaged, surface void
		(e)	Damaged, internal honeycomb
Field	Generic concrete, heterogeneous	(a)	Intact
		(b)	Damaged, surface void
	Self-compacting concrete, homogeneous	(c)	Intact
		(d)	Damaged, surface void
		(e)	Damaged, internal honeycomb

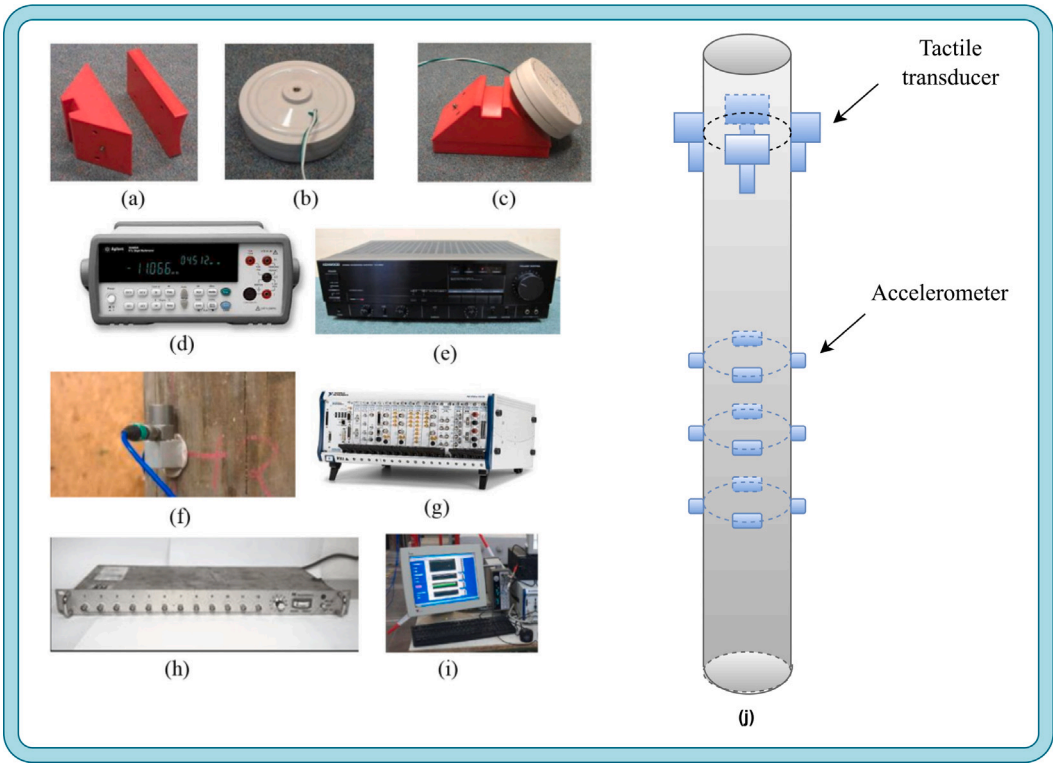


Fig. 3. Testing set-up: (a) sensor wedge, (b) tactile transducer, (c) transducer installed on sensor wedge, (d) function generator, (e) Hi-Fi amplifier, (f) accelerometer, (g) data acquisition system, (h) signal conditioner, (i) computer and (j) schematic of system setup.

system; 12 accelerometers (PCB, model 352C34) with a frequency range of 0.5 Hz to 10 kHz for measuring the wave propagation along the structure; a signal conditioner; and a computer used for data recording. Fig. 3 displays the testing equipment and the schematic of the testing setup. It is hypothesized that the sensors at different heights of the pole (different rings) will capture different wave forms. More details about the data collection and test-setups can be found in [41,42].

In this study, the test results were labeled as "Intact" or "Damaged" regardless of the type of damage and the material properties of the pole. Therefore, a binary classification problem was targeted for analyzing laboratory and field signals individually. The signals recorded from sensors placed at the same height on poles (same ring) are fused through three different methods: cointegration, signal averaging, and AKF. The results are further discussed, and conclusions are drawn in the following sections.

4. Adopted methodology

4.1. Overview

Algorithm 1, outlines the stages of the adopted methodology in this paper. The methodology begins with an assessment of the stationary properties of the raw signals using the Augmented Dickey–Fuller (ADF) test. This test is essential for determining

Algorithm 1 Signal Processing and Classification

```
Input: signals_list, significance_level, VMD parameters, number of iterations n, labels
Output: Classification Report
Stage 1: Check Stationarity of Signals using ADF Test
for each signal in signals_list do
    adf_result ← perform_adf_test(signal)
    if adf_result[p-value] < significance_level then
        Print("Signal is stationary")
    else
        Print("Signal is not stationary")
    end if
end for
Stage 2: Fuse Raw Signals
Step 2.1: Apply Johansen Cointegration
johansen_result ← johansen_cointegration_test(signals_list)
Step 2.2: Apply AKF (Adaptive Kalman Filter)
fused_signal_akf ← apply_akf(signals_list)
Step 2.3: Apply Signal Averaging
averaged_signal ← average_signals(signals_list)
Stage 3: Feature Extraction using VMD (Variational Mode Decomposition)
Step 3.1: Set VMD parameters
vmd_params ← set_vmd_parameters(alpha, tau, K, DC, init, tol)
Step 3.2: Extract Intrinsic Mode Functions (IMFs) using VMD
IMFs ← apply_vmd(fused_signal_akf, vmd_params)
Step 3.3: Construct Gabor's Analytic Signals from IMFs
for each IMF in IMFs do
    gabor_signals ← construct_gabor_analytic_signal(IMF)
end for
Step 3.4: Derive Statistical Features from IA and IF
for each gabor_signal in gabor_signals do
    IA ← compute_instantaneous_amplitude(gabor_signal)
    IF ← compute_instantaneous_frequency(gabor_signal)
    ia_features ← extract_statistical_features(IA)
    if_features ← extract_statistical_features(IF)
    features ← ia_features + if_features
end for
Stage 4: Solve Classification Problem
Step 4.1: Identify Optimal Number of Features
optimal_features ← select_optimal_features(features)
Step 4.2: Monte Carlo Cross-Validation (MCCV) for Classification
mccv_results ← monte_carlo_cross_validation(optimal_features, labels, n_iterations)
Step 4.3: Provide Classification Reports
classification_report ← generate_classification_report(mccv_results)
Print(classification_report)
```

whether the signals are stationary or nonstationary, as the presence of a unit root can significantly affect the choice of signal fusion methods. Once the stationarity of the signals is determined, three different approaches for signal fusion are applied: Johansen cointegration, signal averaging, and the Adaptive Kalman Filter (AKF). Each of these methods is selected to evaluate how they can handle the fusion of raw signals, particularly in the context of nonstationary data. Signal averaging provides a simple means of fusing data by averaging multiple signals, while AKF offers a dynamic filtering approach that accounts for noise and uncertainty.

Johansen cointegration is particularly well-suited for fusing nonstationary signals, as it tests for the existence of long-term equilibrium relationships among the signals. The logic behind comparing these three approaches is to highlight the effectiveness of Johansen cointegration when dealing with increasing instances of nonstationary signals. While signal averaging and AKF can handle certain aspects of data fusion, Johansen cointegration is effective in handling scenarios dominated by nonstationarity, as it identifies and leverages the underlying relationships among nonstationary time series. This comparison helps justify the preference for Johansen cointegration in contexts where the number of nonstationary cases rises, offering a more robust fusion method for such data.

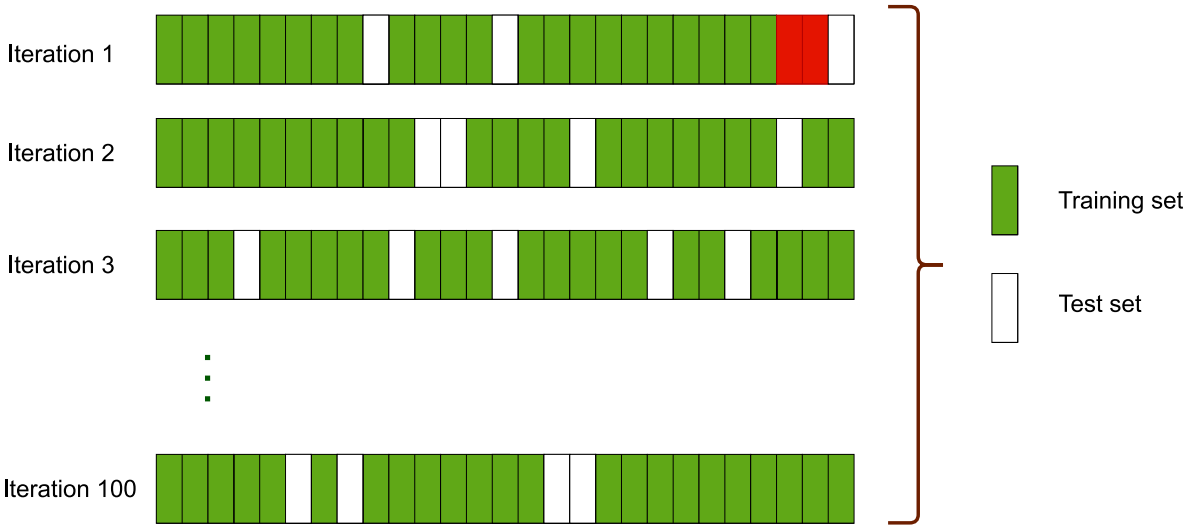


Fig. 4. Monte Carlo Cross-Validation (MCCV) scheme.

4.2. XGBoost algorithm

XGBoost (Extreme Gradient Boosting) is a sparsity-aware algorithm widely used by data scientists for the analysis of sparse data [43]. It is a scalable, distributed, and gradient-boosted decision tree algorithm, and it serves as a learning library for solving classification, regression, and ranking problems [44]. XGBoost is immune to errors stemming from the use of correlated features and the presence of multicollinearity among features. It is appealing to data scientists due to its predictive performance and efficient processing time. Therefore, the XGBoost algorithm has been explicitly employed to solve the classification problem in this paper.

4.3. Monte Carlo Cross-Validation (MCCV)

To ensure the validity of the classification results, the Monte Carlo Cross-Validation (MCCV) technique is employed to evaluate the performance of the XGBoost algorithm. Studies have demonstrated that MCCV generally yields higher accuracy and, consequently, greater reliability compared to CV (Cross-Validation), especially when dealing with limited data sets [45].

Fig. 4 displays the general scheme of the MCCV technique. Here, the data set is randomly split into training and test sets of varying size over multiple iterations. A close examination of the results of different evaluation metrics suggests that the results converge at 100 iterations. Therefore, 100 iterations of the MCCV were performed for each model. Each run involved randomly splitting the data set into training and test sets, with the test set comprising a percentage between 20% to 50% of the data set at each iteration. The mean and standard deviation of the outcomes are presented in the results section of this paper.

4.4. Adaptive Kalman Filter (AKF)

In scenarios where the noise is non-Gaussian or nonstationary, an Adaptive Kalman Filter (AKF) is likely to outperform simple signal averaging. The adaptive Kalman filter can adjust its parameters to better filter out noise while preserving the useful signal components. Therefore, we compare the performance of the cointegration technique with the AKF signal fusion technique. AKFs have been employed for sensor signal fusion in various studies including [12,13].

AKF is a variation of the Kalman filter that adapts its process variance parameter based on the difference between the predicted and measured values. This adaptation helps in handling varying uncertainties in the system dynamics. AKF operates in two main steps: Prediction Update and Measurement Update.

In the prediction step, the priori estimate and error covariance are updated based on the process variance. In the measurement step, a blending factor is calculated based on the priori error covariance and measurement variance. This factor determines the contribution of the measurement to the final estimate. The posteriori estimate and error covariance are then updated based on the measurement and blending factor. This adaptive filter adjusts the process variance based on the difference between the measurement and the posteriori estimate, thus adapting to changes in the system dynamics.

The developed AKF program fine-tunes the process variance based on the discrepancy between the fused signals and the updated estimates, enabling it to adapt to varying system dynamics. The program iteratively processes subsets of input signals and optimizes the AKF algorithm's hyperparameters to enhance fusion accuracy. Hyperparameter optimization—including process variance, measurement variance, estimated measurement variance, and damping factor—is performed using the Differential Evolution

Table 3
Summary of AKF variables and their descriptions.

Variable	Description
$\hat{x}^{k-1}$	State estimate at time step $k-1$
P^{k-1}	Error covariance at time step $k-1$
U_k^{k-1}	State transition from time step $k-1$ to k
C_k	Process noise matrix at time step k
Q_k	Process noise covariance at time step k
H_k	Measurement matrix at time step k
R_k	Measurement noise covariance at time step k
K_k	Kalman gain at time step k
$\hat{x}_{k\|k-1}^k$	Predicted state estimate at time step k
$P_{k\|k-1}^k$	Predicted error covariance at time step k
$\hat{x}^k$	Updated state estimate at time step k
P^k	Updated error covariance at time step k

algorithm [46]. This optimization aims to minimize the total squared error between the fused signals and the AKF estimates, thereby improving the accuracy of the signal fusion process.

The AKF extends the conventional Kalman filter approach by adjusting the process noise covariance matrix at time step k, represented by Q_k, based on current measurements. This adjustment is performed using the Maximum Likelihood (ML) criterion as discussed below.

The AKF algorithm is described by the following equations, which represent the overall scheme of the conventional Kalman filter [47]:

$$\hat{x}_{k|k-1}^k = U_k^{k-1}\hat{x}^{k-1} \tag{11}$$

$$P_{k|k-1}^k = U_k^{k-1}P^{k-1}(U_k^{k-1})^T + C_kQ_kC_k^T \tag{12}$$

$$K_k = P_{k|k-1}^k H_k^T(H_kP_{k|k-1}^kH_k^T + R_k)^{-1} \tag{13}$$

$$\hat{x}^k = \hat{x}_{k|k-1}^k + K_km_k \tag{14}$$

$$P^k = (I - K_kH_k)P_{k|k-1}^k \tag{15}$$

In these equations, m_k represents the measurement residual, which is calculated as:

$$m_k = z_k - H_kU_k^{k-1}\hat{x}^{k-1} \tag{16}$$

Table 3 provides a comprehensive summary of the variables used in the AKF equations and their descriptions.

Additionally, a common approach in the Russian aviation industry employs the AKF to adapt the state noise covariance matrix Q^k using the Maximum Likelihood (ML) criterion, as shown in Eq. (17):

$$C_kQ^kC_k^T = K_km_km_k^TK_k^T + P_k - U_k^{k-1}P_{k-1}U_k^{k-1T} \tag{17}$$

To simplify computations and ensure the positive semi-definiteness of Q, the term $P_k - U_k^{k-1}P_{k-1}U_k^{k-1T}$ can be neglected, leading to the following approximation:

$$C_kQ^kC_k^T = K_km_km_k^TK_k^T \tag{18}$$

This approximation reduces computational complexity, making the AKF suitable for implementation on onboard computers where swift response times and minimal computational requirements are essential.

The AKF implementation developed in this study utilizes Python for computational efficiency and flexibility. The program processes input data by iterating through subsets to apply the AKF, optimizing the hyperparameters through a combination of machine learning techniques and parallel computing. This methodology ensures robust and accurate state estimation, making the AKF a powerful tool for adaptive state estimation.

5. Results and discussions

This section presents the results of applying the proposed methodology to laboratory and field data of concrete poles for health assessment.

5.1. ADF results of laboratory and in-situ field data

Following the methodology of this paper, the Augmented Dickey–Fuller (ADF) test is first run on all signals to determine their stationarity/nonstationarity characteristics. The ADF test is widely recognized as a reliable statistical tool for detecting the presence of a unit root, as it is specifically designed to assess nonstationarity in time series data by evaluating the existence of a stochastic trend. The ADF test extends the Dickey–Fuller test by including lagged differences of the dependent variable, which helps address issues of autocorrelation in the residuals and ensures a more robust estimation [38]. Its null hypothesis is that the time series has a unit root, making it a direct and appropriate method for identifying whether a series follows a random walk or has a stochastic trend. Since the ADF test targets this specific hypothesis, it is more suitable for assessing unit roots than tests like the KPSS, which focus on testing for stationarity around a deterministic trend rather than explicitly checking for a unit root. The widespread use and validation of the ADF test in econometrics further solidifies its trustworthiness for this purpose.

The ADF test relies on several key assumptions to test for the presence of a unit root in a time series. These assumptions are important to ensure the validity and accuracy of the test results. The main assumptions of the ADF test are as follows [48]:

1. **Linearity of the Time Series:** The ADF test assumes that the time series follows a *linear autoregressive process*. This means the model fits the data based on a linear relationship between the current value of the series and its past values.
2. **Constant Variance of Residuals (Homoscedasticity):** The test assumes that the error term (residuals) has a *constant variance over time*. In other words, the variance of the errors is assumed to be homoscedastic.
3. **Uncorrelated Error Terms:** The residuals (errors) from the regression should be *uncorrelated* over time. The ADF test addresses potential autocorrelation by including lagged differences of the dependent variable in the regression, which accounts for serial correlation in the data.
4. **Stationarity of Higher-Order Lags:** The test assumes that the lagged differences included in the regression (as part of the augmentation process) are themselves stationary. This ensures that any nonstationarity is correctly attributed to the presence of a unit root in the original series rather than to higher-order effects.
5. **No Structural Breaks:** The ADF test assumes that the time series does not have any *structural breaks* (sudden changes in the mean or variance). Structural breaks, if present, can lead to incorrect conclusions about stationarity. No structural breaks are expected in the test results of this paper.
6. **Sufficient Number of Lags:** The appropriate number of lags must be included in the model to adequately capture the dynamics of the process and correct for autocorrelation. If too few lags are included, autocorrelation may remain, affecting the test's reliability. If too many lags are included, the model may overfit.

The first two assumptions of the ADF test–linearity of the time series and constant variance (homoscedasticity) of the residuals–are investigated in the following sections. These assumptions are essential to ensure the correct application of the ADF test, as non-linearity or heteroscedasticity can lead to misleading results about stationarity. The other assumptions, such as uncorrelated residuals and stationarity of higher-order lags, are implicitly satisfied by the test. The inclusion of lagged differences in the ADF model corrects for autocorrelation, while the test accounts for the possibility of higher-order nonstationarity through its augmentation process. As for structural breaks, although not directly tested, they are not expected in the nondestructive testing of concrete poles.

5.1.1. Ramsey RESET test for linearity

The Ramsey Regression Equation Specification Error Test (RESET) evaluates whether a linear regression model is correctly specified or if it suffers from omitted variables, incorrect functional forms, or nonlinearity [49]. Specifically, when testing for linearity, the null hypothesis assumes that the model does not require additional nonlinear terms.

Consider the original linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \epsilon_i \tag{19}$$

where:

- y_i is the dependent variable.
- $x_{1i}, x_{2i}, \ldots, x_{ki}$ are the independent variables.
- $\beta_0, \beta_1, \ldots, \beta_k$ are the parameters to be estimated.
- ϵ_i is the error term.

To perform the RESET test, the model is augmented with powers of the fitted values (or other transformations) up to a certain degree m:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + \cdots + \gamma_m \hat{y}_i^{m+1} + \epsilon_i \tag{20}$$

where:

- $\hat{y}_i$ represents the fitted values from the original model.
- $\gamma_1, \gamma_2, \ldots, \gamma_m$ are the coefficients for the nonlinear terms.

Table 4
Ramsey RESET test results for linearity (p-value = 0.05).

Dataset	# $p < 0.05$	Total columns	% Linearity
Laboratory Data	17	360	95.28
Field Data	14	300	95.33

The Null and Alternative Hypotheses of the RESET test are as follows:
Null Hypothesis (H_0):

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_m = 0 \tag{21}$$

The following outlines the way the results of the test can be interpreted:

- All coefficients of the additional nonlinear terms are equal to zero.
- The original linear model is correctly specified without the need for these higher-order terms.
- There is no evidence of nonlinearity or omitted variable bias in the model.

Alternative Hypothesis (H_a):

$$H_a : \text{At least one of } \gamma_1, \gamma_2, \ldots, \gamma_m \neq 0 \tag{22}$$

Interpretation:

- At least one of the nonlinear terms significantly improves the model.
- Indicates potential misspecification in the original linear model, such as nonlinearity or omitted variables.

The adopted RESET test was considered to include up to the squared term ($m = 2$). Therefore, the hypotheses become:

$$H_0 : \gamma_1 = \gamma_2 = 0 \tag{23}$$

and

$$H_a : \gamma_1 \neq 0 \quad \text{or} \quad \gamma_2 \neq 0 \tag{24}$$

By testing the null hypothesis H_0, the Ramsey RESET test helps validate the specification of the linear regression model. Failure to reject H_0 suggests that the linear model is adequate, whereas rejection indicates that the model may need to incorporate additional terms or transformations to better capture the underlying relationship in the data.
Table 4 shows results of the Ramsey RESET test run on the laboratory and field data. The table indicates that:

- **Laboratory Data:** Out of 360 columns tested, 17 columns (approximately 4.72%) rejected the null hypothesis (H_0) of correct linear specification. Consequently, 95.28% of the laboratory data columns conform to the linearity assumption.
- **Field Data:** Out of 300 columns tested, 14 columns (approximately 4.67%) rejected H_0, resulting in 95.33% conformity with the linearity assumption.

These findings suggest that the vast majority of the regression models applied to both laboratory and field datasets are appropriately specified under the linear framework. Specifically, over 95% of the models do not exhibit significant evidence of misspecification related to nonlinearity or omitted variables as per the RESET test.

5.1.2. The Goldfeld-Quandt statistical test for homoscedasticity

The Goldfeld-Quandt statistical test [50] was applied to both the laboratory and field datasets to assess the presence of heteroscedasticity. The null and alternative hypotheses for the test are as follows:
Null Hypothesis (H_0): The variance of the error terms is constant across all observations (i.e., homoscedasticity).

$$H_0 : \quad \sigma_1^2 = \sigma_2^2 \tag{25}$$

Alternative Hypothesis (H_1): The variance of the error terms is not constant across observations (i.e., heteroscedasticity exists).

$$H_a : \quad \sigma_1^2 \neq \sigma_2^2 \tag{26}$$

The results of the test with confidence level of 5% are presented in Table 5. The results indicate that more than 91.67% and 85% of the laboratory and field data conform with the homoscedasticity assumption.

The diagnostic evaluations using the Ramsey RESET and Goldfeld-Quandt tests reveal that a notable minority of the laboratory and field datasets exhibit nonlinearity and heteroscedasticity, respectively, as detailed in Tables 4 and 5. Specifically, 8.33% of the

Table 5
Goldfeld-Quandt test results for heteroscedasticity (p-value = 0.05).

Dataset	# $p < 0.05$	Total columns	% Homoscedastic
Laboratory Data	30	360	91.67
Field Data	45	300	85.00

laboratory data columns and 15.00% of the field data columns show evidence of heteroscedasticity, while approximately 4.7% of the columns in both datasets reject the null hypothesis of linearity. These specification errors suggest the presence of underlying issues that may compromise the validity of regression analyses, such as biased standard errors and misspecified functional forms. One potential source of these specification errors is nonstationarity within the datasets. Nonstationary data, characterized by properties such as time-dependent mean and variance, can lead to spurious regression results, where relationships appear significant solely due to underlying trends rather than genuine associations. To address this concern and ensure the robustness of the models, it is essential to assess the stationarity of both laboratory and field data. Consequently, conducting the ADF test is justified to investigate the nonstationary nature of the data.

Fig. 5 shows the proportion of nonstationarity in the recorded signals of all concrete poles for both laboratory and in-situ field recorded signals. It is evident from the results that the field-recorded signals exhibit a significantly higher proportion of nonstationarity across all tested poles compared to the laboratory data. Indeed, a mean increase of 57.6% in nonstationarity is found across all poles, with the highest increase of 73% (from 8% to 30%) for the intact homogeneous pole (Pole c). This underscores the significant impact of real-life testing conditions on the data and emphasizes the importance of our proposed method in mitigating the effects of increased nonstationarity. Furthermore, it is evident that the presence of damage in a pole leads to an increase in nonstationarity. This observation is expected, as damage directly affects stress wave propagation, resulting in pronounced nonstationary effects. Across all poles, an average increase in nonstationarity of 24.8% is observed when comparing intact to damaged poles. As such, for the laboratory testing, the nonstationarity increases from 17% to 25% for the damaged heterogeneous poles and from 8% to 17% for the damaged homogeneous poles. For the field results, it increases from 45% to 52%, and from 30% to 32% and 38%, for the heterogeneous and homogeneous damaged poles, respectively. Comparing the mean increases in nonstationarity attributed to the changed testing environment (57.6%) and the presence of damage (24.8%), it is evident that real-life testing conditions have a significantly larger impact on the stationarity of the signals. These findings underscore the importance of considering environmental effects during signal processing for damage assessment and highlight the necessity of the proposed method to mitigate the pronounced nonstationarity observed in field conditions. It can further be seen from the ADF test results that the signals recorded in the field show a more consistent nonstationarity across different damage scenarios. This helps to explain some of the results obtained in this study, as we compare the performance of signal averaging [51], AKF, and cointegration techniques as raw data fusion algorithms.

5.2. Evaluating the NMI for different values of VMD parameters

This section focuses on examining the influence of VMD parameters on the evaluated NMI metric between the original signal and its decomposed components. To this end, we have explored different values of VMD parameters from the options below and calculated the mean NMI value in both laboratory and field settings for each parameter combination. The following are the options for the selected values of the VMD parameters:

1. Number of IMFs (p): This parameter was selected from the options in the set {3, 5, 7}.
2. Denoising factor (α): Different values of this parameter were selected from the set {100, 500, 1000} representing minor, moderate, and severe cases of denoising, respectively.
3. Time interval (τ): This parameter determines how quickly the Lagrangian multiplier in (5) accumulates the reconstruction error. In case exact reconstruction is intended, one can set it to a small number like 0.1 to minimize the effect of α, as suggested by the authors of the VMD paper [28]. Here, we fixed τ at zero and experimented with different values of α.
4. Convergence threshold (ϵ): Different values of 10^{-5} and 10^{-7} were investigated, and their effects on the NMI metric were analyzed.
5. Center frequency initializer ($init$): This parameter can take the value of 0, 1, or 2, indicating zero initialization, uniform initialization, and random initialization, respectively. Previous work mentioned its negligible effects on the decomposition results [52].
6. DC: This is a Boolean parameter taking the value of either 0 or 1, indicating whether to fix the first IMF's center frequency at zero or not.

We seek to quantify the average value of the information passed through the decomposition process across fused signals using the proposed and adopted methods. Tables 6 and 7 show the mean and the standard deviation values of the NMI metric obtained from the application of the VMD to the entire samples of data for different combinations of the above-mentioned parameters. Our preliminary examination of the impact of various VMD parameters on the NMI results confirmed that α (denoising factor) and p (number of decompositions) are the only variables that can practically affect the NMI value. Additionally, the results demonstrate comparable values obtained by applying VMD to signals derived from both the averaging and AKF fusion methods. The results also indicate that increasing the value of α substantially decreases the amount of information transferred to the IMFs, while increasing

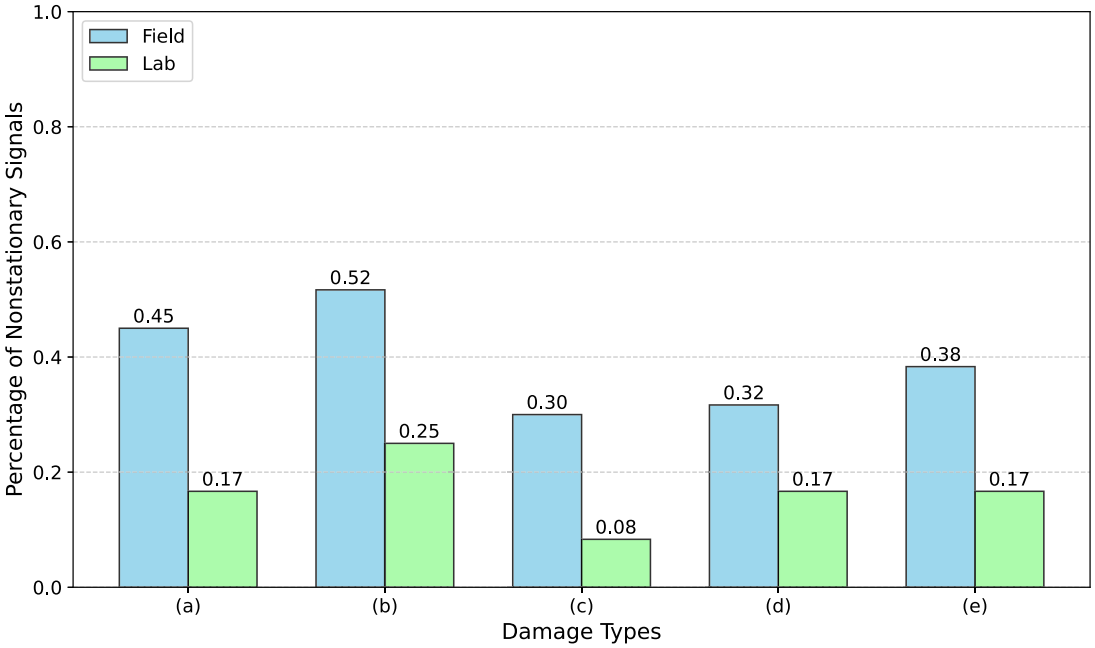


Fig. 5. Proportion of signal nonstationary of all concrete poles tested in the laboratory and field using the ADF test: (a) Intact pole – Heterogeneous concrete, (b) Damaged pole (surface void) – Heterogeneous concrete, (c) Intact pole– Homogeneous concrete, (d) Damaged pole (surface void) – Homogeneous concrete, and (e) Damaged pole (internal honeycomb) – Homogeneous concrete.

Table 6
NMI values for various combinations of hyperparameters of VMD applied to the entire fused laboratory data using various mathods ($\mu \pm \sigma$).

Averaging and AKF			
α \ p	3	5	7
100	0.68 ± 0.09	0.80 ± 0.04	0.81 ± 0.04
500	0.46 ± 0.14	0.59 ± 0.11	0.69 ± 0.08
1000	0.36 ± 0.15	0.46 ± 0.14	0.55 ± 0.13
Cointegration			
α \ p	3	5	7
100	0.62 ± 0.10	0.77 ± 0.05	0.83 ± 0.04
500	0.40 ± 0.15	0.53 ± 0.12	0.66 ± 0.08
1000	0.33 ± 0.16	0.44 ± 0.14	0.54 ± 0.12

Table 7
NMI values for various combinations of hyperparameters of VMD applied to the entire fused field data using various mathods ($\mu \pm \sigma$).

Averaging and AKF			
α \ p	3	5	7
100	0.78 ± 0.06	0.82 ± 0.05	0.86 ± 0.04
500	0.71 ± 0.08	0.73 ± 0.08	0.75 ± 0.07
1000	0.69 ± 0.08	0.71 ± 0.08	0.72 ± 0.08
Cointegration			
α \ p	3	5	7
100	0.66 ± 0.06	0.80 ± 0.04	0.85 ± 0.023
500	0.47 ± 0.10	0.55 ± 0.10	0.68 ± 0.07
1000	0.43 ± 0.09	0.48 ± 0.10	0.56 ± 0.10

the value of p amplifies it. However, there exists a practical limitation in terms of increasing the number of decompositions: it may lead to over-fitting machine learning algorithms on the training data using the extracted features [53].

Table 8
Assigned values for VMD parameters [17].

Parameters	Description	Specified values
p	Number of IMFs	5
α	Denoising factor	100
τ	Time interval	0
ϵ	Convergence threshold	10^{-5}
init	Center frequency initializer	0
DC	Boolean parameter	0

Table 8 presents suggested VMD parameters and their assigned values for the problem of this paper based on the above-mentioned results.

5.3. Laboratory experiment results

This section presents the results of applying the XGBoost algorithm to signals recorded in the laboratory. Fig. 6 illustrates examples of cointegration, averaging, and AKF signal fusion techniques derived from a set of four sensors positioned on the same ring level, along with their corresponding decomposition results. While it is challenging at this stage to identify significant differences between the decomposition outcomes of the averaged and adapted Kalman filtered signals, one clear observation is that the higher-order decomposition results effectively capture high-frequency modulations in the carrier signal (first IMF). Notably, the cointegrated signal showcases a nonlinear combination of merged signals, successfully eliminating the nonstationary effects evident in the low-frequency fluctuations of the signals. This distinction is particularly apparent When comparing in the first IMF of the cointegrated signal with those derived from the signal averaging and AKF techniques.

5.3.1. Distribution of NMI across various fusion techniques

Next, we plan to treat the NMI between the fused signals and their corresponding decompositions as a random variable. Our goal is to create a boxplot that illustrates the distribution of this variable obtained from signals resulting from all fusion techniques. Fig. 7 displays the box plots of the results. The findings show a wider range for the NMI using the averaging and AKF techniques, whereas the distribution is narrower for the cointegration technique. This observation suggests that the cointegration method concentrates the amount of information passed through the decomposition process across different signals. Consequently, a few decompositions are treated as outliers. The average value of NMIs for cointegration, averaging abd AKF techniques is calculated as 0.77, 0.79 and 0.79, respectively. This may help to explain some of the machine learning results obtained from each method in subsequent analyses. More specifically, the results indicate that the averaging and AKF techniques outperform the cointegration technique when applied to the laboratory signals. Furthermore, the averaging and AKF techniques demonstrate comparable performance.

5.3.2. Visualization of the obtained feature space

Next, the dimensionality of the feature space is reduced to facilitate the visualization of data classes and their distributions. To this end, the data is first standardized and PCA analysis is performed. Fig. 8 exhibits the screen plot of the PCA analysis result. The results indicate that over 65% of the variance in the data set can be captured by the first two principal components (PC). Therefore, plotting the first PC (PC1) versus the second PC (PC2) is expected to provide a further insight into the separability of the data set across different damage classes. Fig. 9 displays the scatter plots of PC1 versus PC2 for all the data points. The results reveal distinct clusters of data points within each class. However, while the results of the averaging and AKF techniques are comparable, determining which fusion technique achieved better separation remains challenging.

5.3.3. Exploration of feature importance

Next, each feature's importance relative to that of other features is calculated via the permutation technique. This technique evaluates the importance of features by measuring the impact of shuffling or permuting the values of each feature while keeping the relationship between the target variable and other features unchanged [54]. The core concept is that permuting the values of an important feature will significantly degrade the model's performance by disrupting the original relationship between the feature and the target variable. Each feature column is randomly permuted, and the average importance value is reported.[1]

Figs. 10(a), 10(b) and 10(c) show the relative nonzero feature importance for the cointegration, signal averaging and AKF techniques, respectively. The results show the selected 19, 14 and 14 features from the cointegration, signal averaging and AKF techniques, respectively, that have nonzero importance values. In all cases, at least three out of the five most important features are non-parametric statistical features, underscoring the critical role of non-parametric statistical features in characterizing the data obtained in the laboratory. Furthermore, the nonzero feature importance for both the signal averaging and AKF techniques is identical, with the majority of the highly important features being the same for both techniques.

The most important feature obtained from the signal averaging and AKF techniques is the third quartile of the IF corresponding to IMF_4, while the most important feature for the cointegration technique is the skewness of the IA distribution corresponding to the

[1] The number of permutations is set to default in the developed Python code.

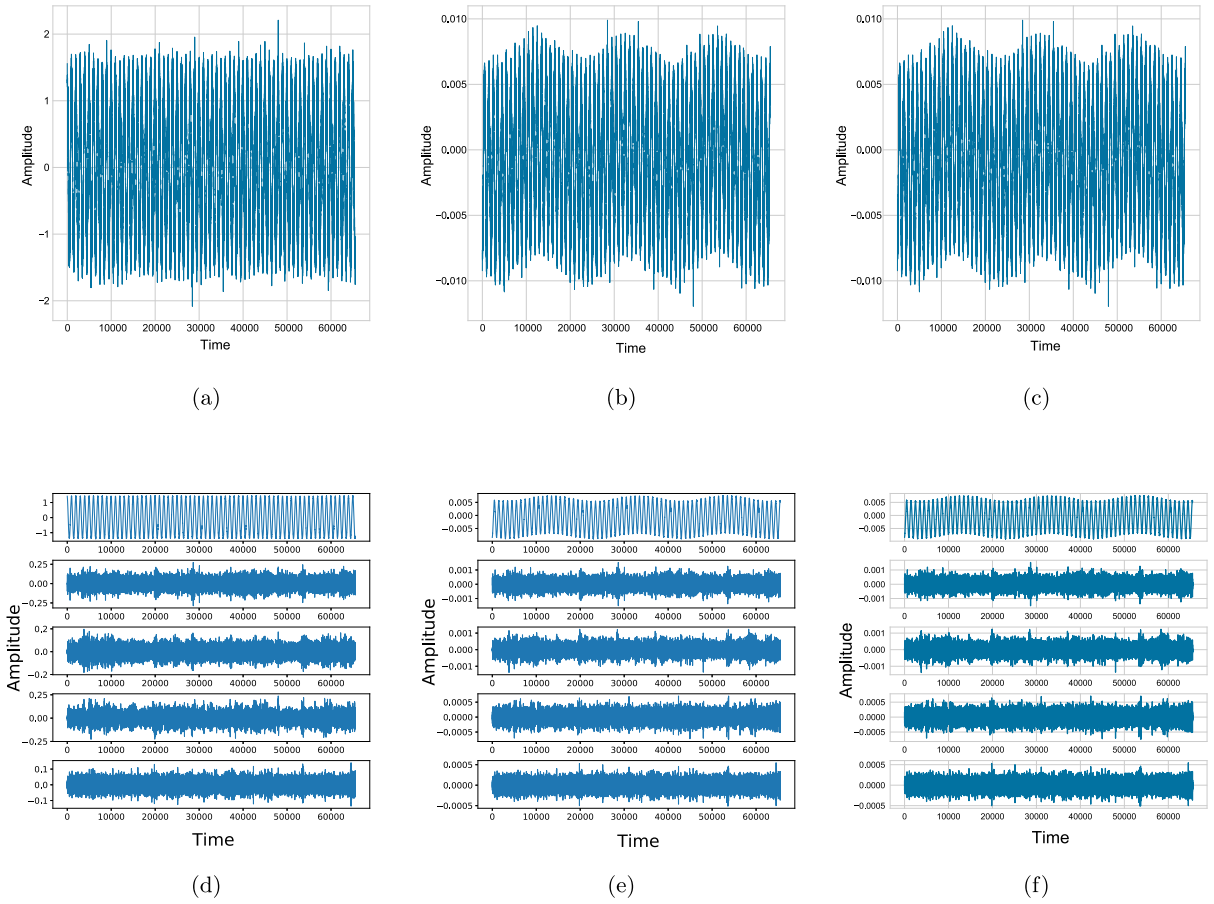


Fig. 6. Example of signal decomposition for the laboratory data (from the lowest to the highest frequency contents, arranged from top to bottom): (a) cointegrated signal, (b) averaged signal, (c) AKF, (d) decomposed cointegrated, (e) decomposed averaged, and (f) decomposed AKF.

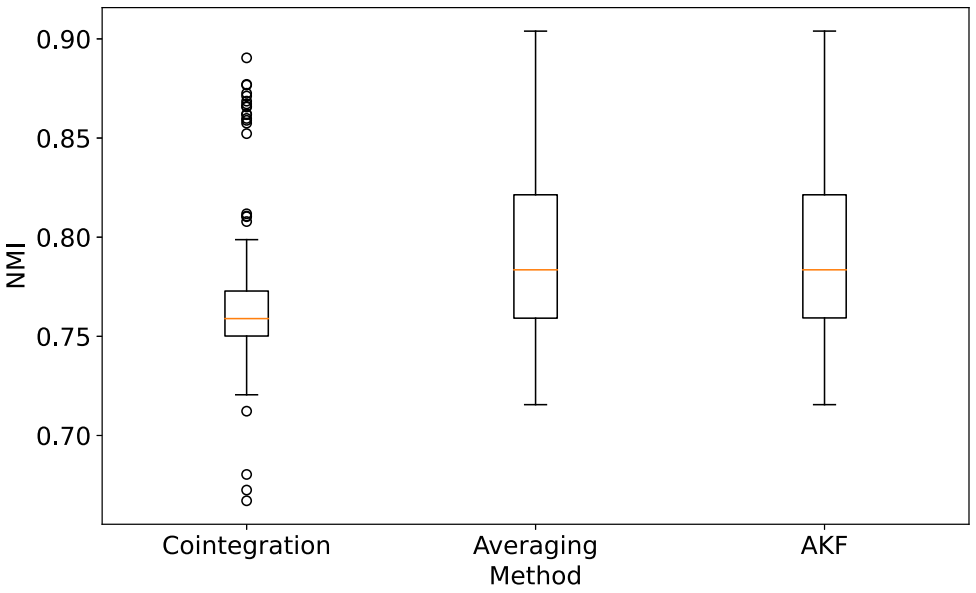


Fig. 7. NMI results of the VMD (with the selected hyperparameters) applied to all fused laboratory signals obtained from various fusion techniques.

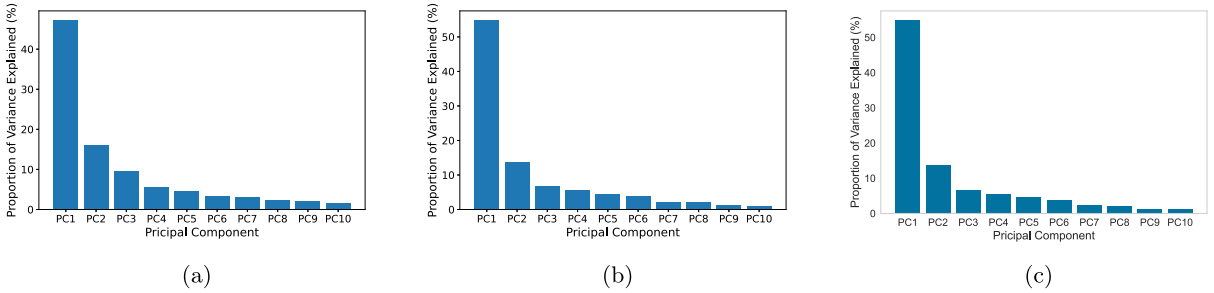


Fig. 8. Scree plots presenting the proportion of variance captured by the first 10 PCs of the laboratory features using (a) cointegration, (b) signal averaging, and (c) AKF methods.

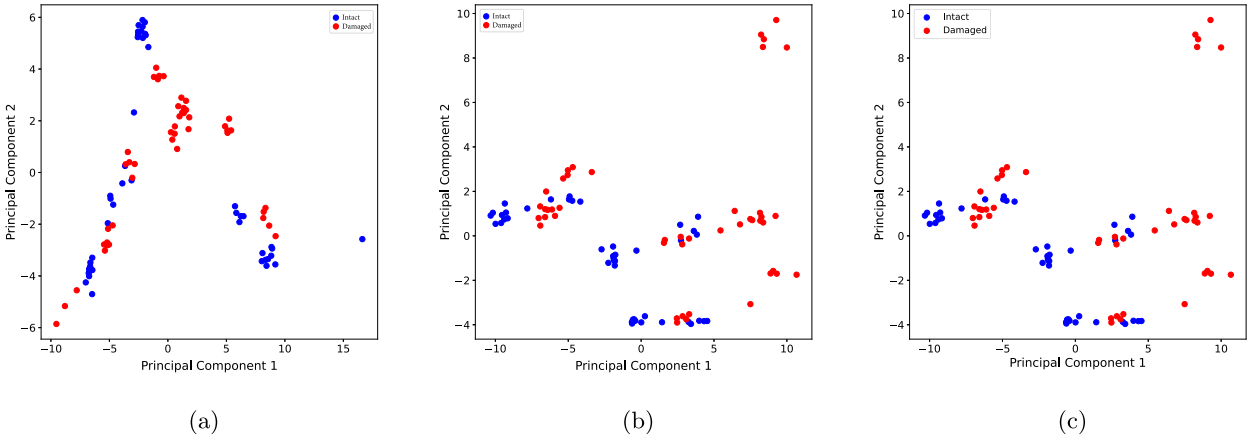


Fig. 9. Scatter plots of PC1 versus PC2 of features extracted from the laboratory signals employing (a) cointegration, (b) signal averaging, and (c) AKF methods.

IMF_3. These results underscores the importance of decomposing the fused signals into four IMFs. Furthermore, this again validates the significance of decomposing the signals as both the most important features are selected from the higher order decomposition results.

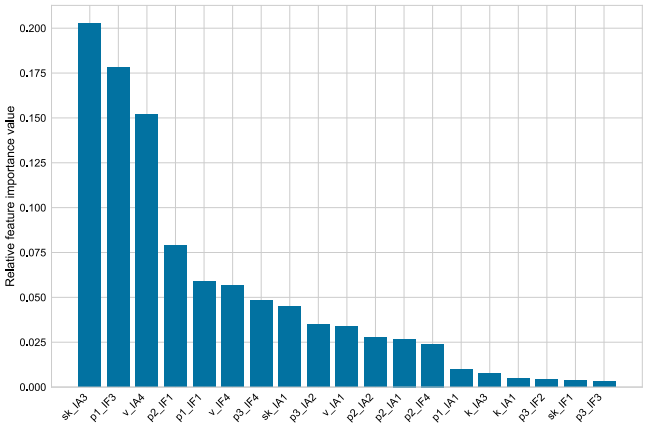
Another noteworthy observation is that no feature from IMF_5 is present in the top 15 important features of the cointegration technique. A similar result is achieved for the signal averaging technique, as no feature from IMF_5 is present in the top 12 most important features. This observation suggests that specifying the number of five IMFs in the VMD settings is sufficient for the purpose of decomposing the laboratory signals. Since over-decomposing signals using VMD primarily results in the duplication of higher-order IMFs (those with higher center frequencies) [28], it is unlikely to have a significant deteriorating effect on the overall performance of machine learning algorithms. However, performing feature selection remains crucial to avoid including highly correlated features in the models.

5.3.4. Pearson correlation matrix

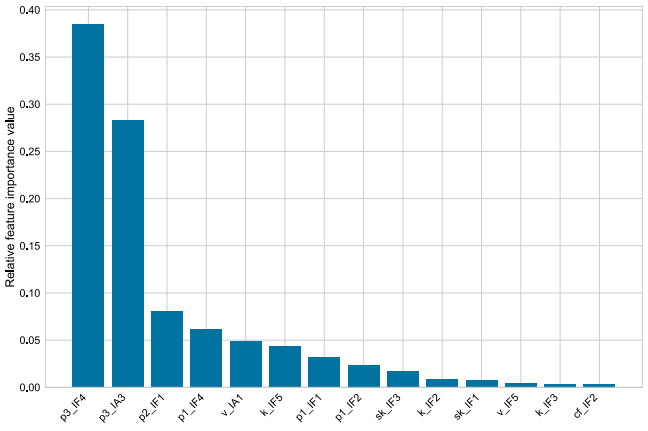
Next, the Pearson correlation matrix is calculated for the features with nonzero importance. Fig. 11 visualizes the results for all models. The results show the presence of a few highly correlated features among the features with nonzero importance. Selecting these features in a model can compromise its performance. Therefore, it is essential to employ a sifting process to discard highly correlated features from further processing.

5.3.5. Identifying the optimal number of features

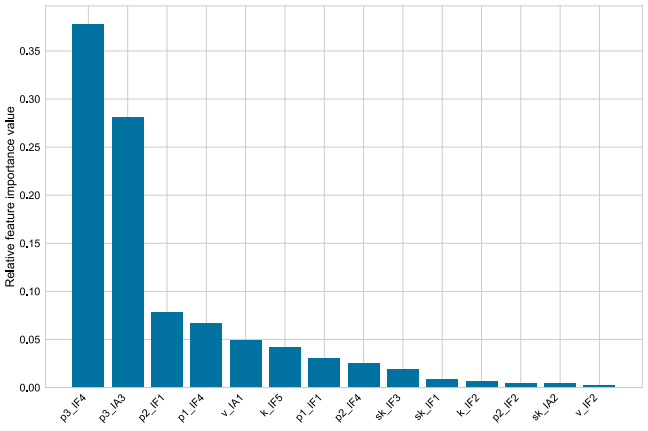
To this end, the Yellowbrick library's recursive feature elimination is employed to select features one by one based on their importance by evaluating a 5-fold cross-validation of the XGBoost models. The F1-score was set as the selection criteria. Figs. 12(a), 12(b) and 12(c) present the obtained results for the cointegration, signal averaging and AKF techniques, respectively. When running the feature-selection algorithm on models, the top 11 most important features are distinguished as optimal number of features in all three cases. Although the selected features are among those that are uncorrelated there are a few features that are highly correlated in each case. For example, regarding the cointegration model, P2IA2 is highly correlated with P3IA2 and vIA4 with 100% and 87% correlations, respectively. There are also highly correlated features among the selected features of the signal averaging technique. For example, features skIF3 is correlated with kIF2 and kIF5 with correlation factors of 90% and -83%, respectively. One reason



(a)



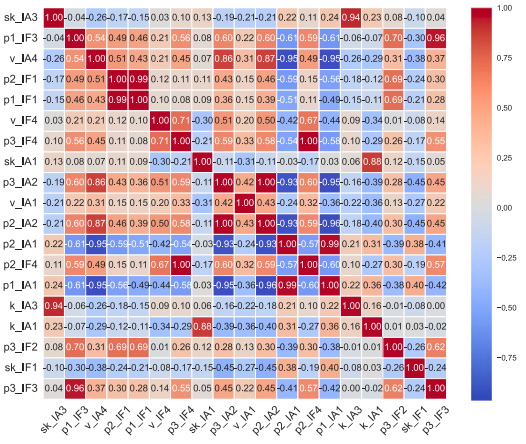
(b)



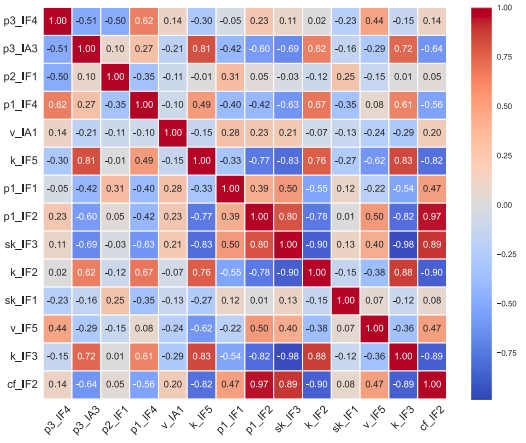
(c)

Fig. 10. Nonzero results of the permutation feature importance for the laboratory experiment using (a) cointegration, (b) signal averaging, and (c) AKF methods.

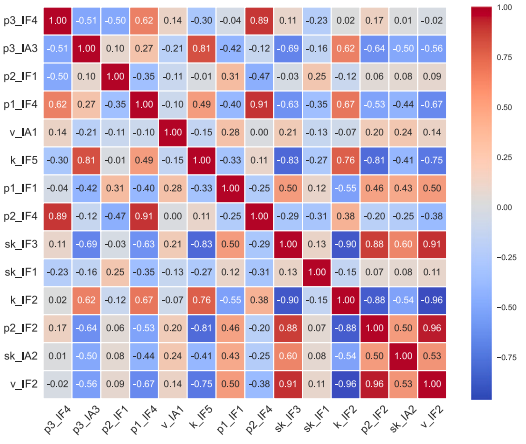
to select these features as descriptive features is likely due to the existence of collinearity among them. It is known that collinear features cannot individually predict a target. Therefore, their interaction must be considered in the machine learning models [55].



(a)



(b)



(c)

Fig. 11. Pearson correlation factors for features with nonzero importance for the laboratory experiment using (a) cointegration, (b) signal averaging, and (c) AKF methods.

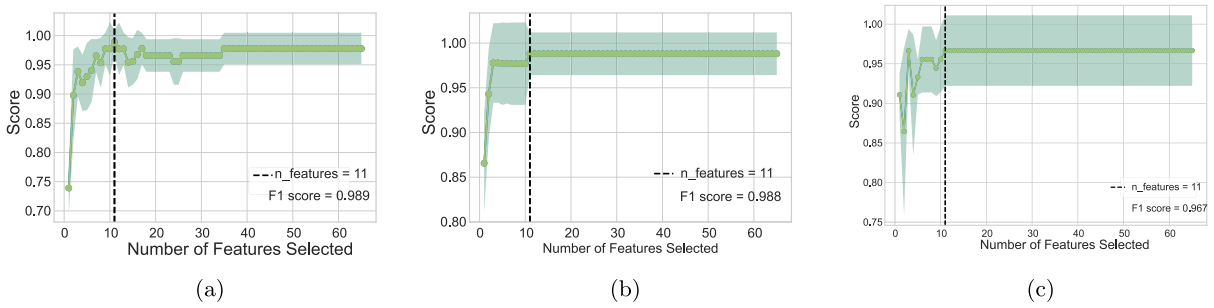


Fig. 12. F1-score results of five-fold cross-validation employing XGBoost trained on the laboratory data using recursive selection of most important features obtained from (a) cointegration, (b) signal averaging, and (c) AKF methods.

It is also evident that using features obtained from the signal averaging and AKF methods leads to significant fluctuations of the results across different folds. In contrast, the cointegration technique exhibits a smaller standard deviation compared to the signal averaging and AKF techniques.

5.3.6. Classification reports

Fig. 14 shows the mean value and standard deviation of the precision, recall, and F1-score evaluated after 100 runs of MCCV of the XGBoost models using the selected features. Different random states and train-test percentages are selected for splitting the feature matrices at each run. The parameters pertaining to the standardization of the training features were employed to standardize the test features at each run to avoid leakage. It is evident from the results that using signal averaging and AKF techniques yields superior performance in this case. Moreover, similar performance of these two techniques demonstrate the similar nature of the two techniques. Further, a smaller standard deviation is obtained across all metrics. The better results are particularly achieved for the precision and recall scores regarding the intact and damaged poles, respectively. This indicates that the signal averaging and AKF techniques can produce more effective features to reduce false positive detection. This is particularly crucial in non-destructive damage testing, where correct identification of defective poles is of utmost significance. Additionally, there is a notable increase in the recall score for the damaged class, suggesting that the signal averaging-based model excels at capturing positive instances, i.e., damaged poles. This complies with the expectation from an effective machine learning model for NDT. As such, an effective model should accurately identify as many damaged cases as possible. Furthermore, the signal averaging-based model exhibits higher F1-scores for both the intact and damaged classes, indicating its ability to strike a balance between precision and recall for both classes.

The above results suggest that both signal averaging and AKF methods are more efficient techniques for fusing the laboratory signals. To delve into reasons for this observation we refer back to the results of Fig. 5 which indicates that the majority of the laboratory signals exhibit stationary properties. Here, the majority of signals already have constant statistical features. Therefore, this may imply that fusing them through Johansen cointegration is not beneficial since Johansen cointegration typically seeks stationary trends among a set of nonstationary signals. Indeed, when this condition is not satisfied, it can lead to producing less meaningful results.

5.3.7. The Receiver Operating Characteristic (ROC) curve

The Receiver Operating Characteristic (ROC) curve can provide additional insights into the effectiveness of a model. It is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various classification thresholds. Additionally, the Area Under the Curve (AUC) provides a measure of the model's ability to correctly classify positive instances as positive (TPR) while minimizing false positive errors (FPR).

XGBoost uses a default threshold of 0.5 to classify instances as intact or damaged. A sample is classified as damaged (positive) if the obtained probability is greater than 0.5 and is classified as intact (negative) otherwise. As such, the classification decision of XGBoost is made based on the predicted probabilities. The ROC is created by varying the classification threshold and plotting the TPR versus the FPR.

A higher AUC is a good metric for evaluating the discriminatory power of a model. In other words, it provides insight into the capability of the model to distinguish between the two classes. Consequently, an ideal classifier yields an AUC of 1, whereas any value of AUC equal to or less than 0.5 represents a poor classifier. Notably, the AUC of 0.5 represents a random classifier that does not perform any better than chance. As such, a higher AUC implies that the model can learn meaningful patterns or features that are relevant to the classification task. Here, we argue that the higher value of the AUC implies the better performance of the corresponding fusion algorithm.

Fig. 13 shows the ROC curve obtained for all methods after running 100 MCCV and computing the average and standard deviation of the calculated TPR and FPR. The results demonstrate the superior performance of the signal averaging and AKF methods, as evidenced by an AUC of 0.99 and a minimal standard deviation of 0.02 compared to the cointegration method with the mean value

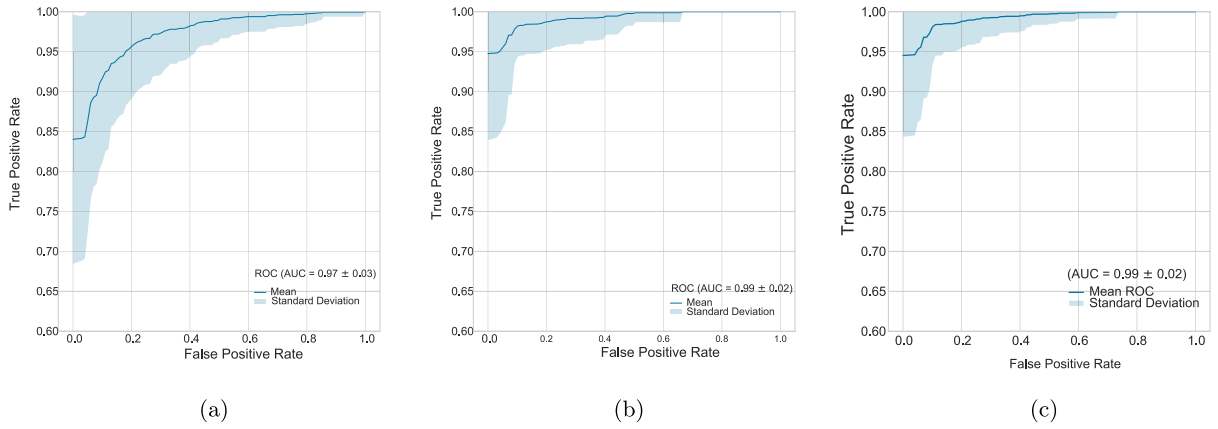


Fig. 13. ROC curves obtained for 100 runs of MCCV using the top 11 most important features of laboratory data obtained from (a) cointegration, (b) signal averaging, and (c) AKF methods.

Table 9
Comparison of the classification reports obtained from different VMD decomposition settings applied to the fused laboratory data using Johansen cointegration algorithm.

Class	Metric	Settings	
		p = 7, α = 100	p = 5, α = 100
Intact	Precision	0.98 ± 0.06	0.95 ± 0.07
	Recall	0.94 ± 0.09	0.95 ± 0.07
	F1-score	0.96 ± 0.06	0.95 ± 0.05
Damaged	Precision	0.95 ± 0.08	0.96 ± 0.07
	Recall	0.98 ± 0.06	0.95 ± 0.07
	F1-score	0.96 ± 0.06	0.95 ± 0.05

of 0.97 and standard deviation of 0.03.[2] The results also indicate a high standard deviation of TPR and FPR obtained for the signal averaging and AKF techniques compared to the cointegration method.

5.3.8. *Exploring the effect of various VMD settings*

Thus far, our analysis has been focused on fixed settings specified for VMD. We observed that both signal averaging and AKF outperform the Johansen cointegration in fusing the laboratory signals. However, upon revisiting the results presented in Table 6, it is evident that the chosen settings for VMD led to a lower NMI for the fused signals when using Johansen cointegration, in comparison to the signal averaging and AKF methods. This observation may explain why features extracted from the cointegrated signals perform less effectively than those derived from the other two techniques.

To investigate this, we adopted the settings for the VMD signal decomposition with NMI = 0.83 ± 0.04, specifically using $p = 7$ and $\alpha = 100$. This configuration produced an NMI result comparable to signal averaging and AKF. After decomposing the signals with these settings and performing recursive feature elimination, we identified the optimal number of features to be 13. We then compared the performance of this configuration with the previous settings ($p = 5$, $\alpha = 100$), where the optimal number of features was 11.

Table 9 presents the comparison of the classification reports using the optimal number of features in each case. Here, we focus on the F1-score as a metric that balances between Precision and Recall. It is evident that while increasing the number of decompositions from $p = 5$ to $p = 7$ slightly enhances normalized mutual information (NMI), it does not lead to a significant improvement in performance metrics. For instance, with $p = 7$ and $\alpha = 100$, class "Intact" achieved an F1-score of 0.96 ± 0.06. In comparison, with $p = 5$ and $\alpha = 100$, class "Intact" showed almost similar F1-score of 0.95 ± 0.05. Class "Damaged" metrics also exhibited similar results obtained for the F1-score.

These results indicate that while increasing the number of decompositions provides a more granular signal representation, as reflected in the increased NMI, it does not significantly enhance the model's ability to differentiate between classes. This may be due to further decomposition introducing features that are less relevant or redundant, thereby diminishing their value for classification. Consequently, the gains in signal representation do not translate into substantial improvements in classification performance, highlighting the need to balance decomposition depth with feature relevance for optimal outcomes.

[2] Note that the mean±std should not exceed 1. The reason this value appears to exceed 1 is due to rounding up digits.

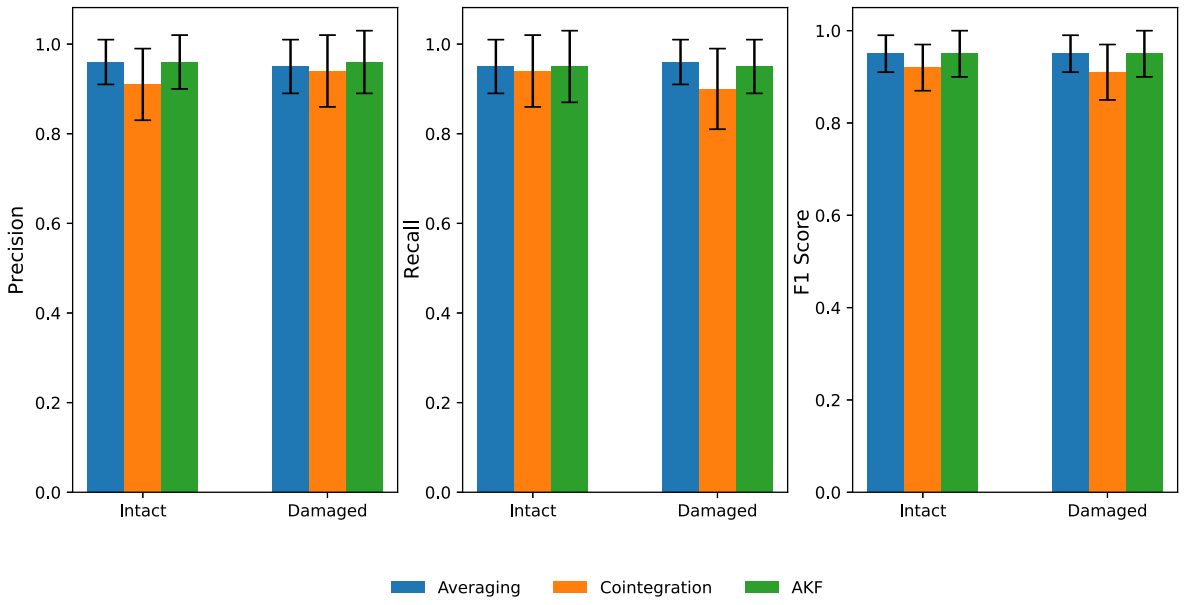


Fig. 14. Classification reports of the models trained with the top eleven important laboratory features evaluated for 100 runs of MCCV on random train-test splits of the datasets at each run.

5.4. Results of in-situ field data

In the following, we present the results of our proposed method demonstrated on the data recorded in the field. Fig. 15 depicts examples of cointegrated signals, averaged signals and AKF methods from concrete poles tested in the field along with their decomposed IMFs. A visual comparison of the decomposition results with that of Fig. 6 reveals more disparity of the features identified across different IMFs. This indicates that higher-order decompositions can capture information about modulations in the carrier signal that carry information about the interior condition of the wooden poles.

5.4.1. Distribution of NMI across various fusion techniques

Next, we calculate the NMI between the fused signals and their corresponding decompositions. The results are illustrated in the boxplots of Fig. 16. Similar to the lab results, a broader range for the NMI is obtained from the averaging and AKF techniques, while the distribution of NMIs is narrower for the cointegration technique. This observation reinforces that the cointegration technique produces signals with more concentrated information passed through the decomposition process. Moreover, fewer decompositions are treated as outliers compared to the laboratory signals. The average NMI values for the cointegration, averaging and AKF methods are calculated as 0.80, 0.82 and 0.82, respectively. Comparing with the lab outcomes, these results may suggest that machine learning results obtained from the averaging and AKF techniques should perform better than the cointegration technique in this case as well. However, the subsequent results presented in this section indicate otherwise.

5.4.2. Visualization of the obtained feature space

Fig. 17 shows the scree plots (proportion of the variance) of the features obtained from the three fusion techniques. The figure indicates that the first two PCs can capture less than 50% of the variance in the features. This result indicates that plotting the first PC versus the second PC may not provide a reasonable view of the level of separability across different classes. Fig. 18 presents the scatter plots of PC1 versus PC2 for the feature space. The results indicate the presence of distinct clusters of data points within each class. It can be seen that the features obtained from the signals obtained from averaging and AKF techniques are slightly more scattered, especially for the intact class.

5.4.3. Exploration of feature importance

Figs. 19(a), 19(b) and 19(c) display the relative importance of features using the permutation method for the cointegration, averaging and AKF models, respectively. The obtained results indicate that unlike the laboratory experiment, four out of the five most important features are parametric statistical features for both models. Furthermore, the five top features in both cases are selected from a mixture of low- and high-level IMFs, highlighting the importance of decomposing the signal into its constructive IMFs.

The most important feature for the cointegration technique is the kurtosis of the IF corresponding to IMF_3 whereas the most important feature obtained from the both signal averaging and AKF techniques is obtained from the first IMF characteristics, i.e. the first quartile of the IA and the center frequency of IF both corresponding to the first IMF signals, respectively. This underscores the

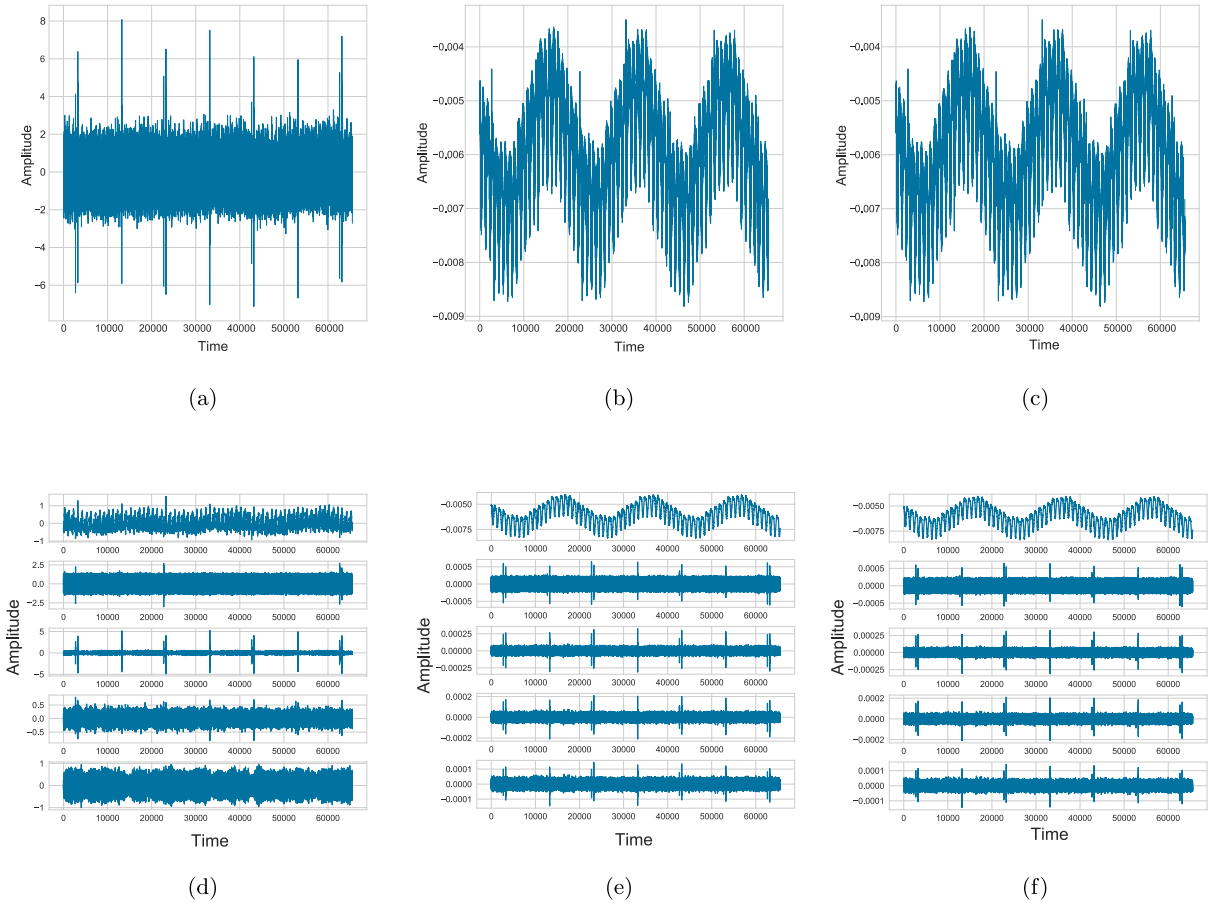


Fig. 15. Example of signal decomposition results of data captured in the field experiment: (a) cointegrated signal, (b) averaged signal, (c) AKF, (d) decomposed cointegrated signal, (e) decomposed averaged signal, and (f) decomposed AKF.

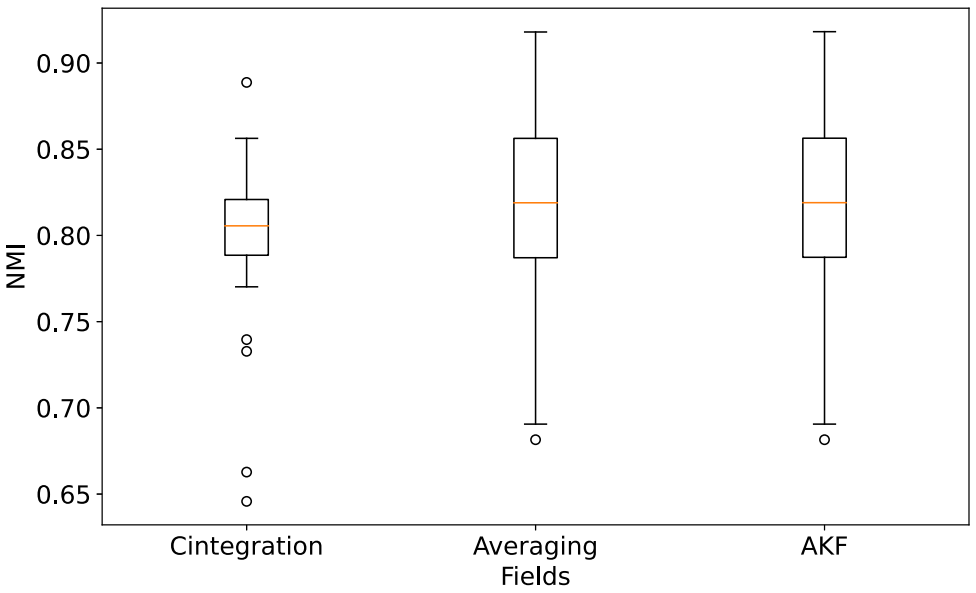


Fig. 16. NMI results of the VMD (with the selected hyperparameters) applied to all fused field signals obtained from various fusion techniques.

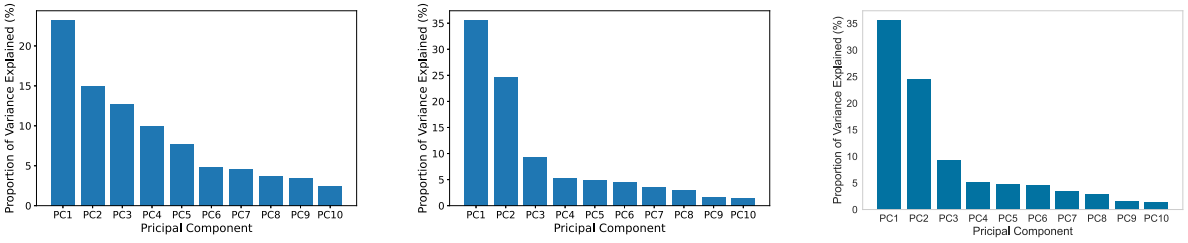


Fig. 17. Scree plots of field features after fusing signals with (a) cointegration, (b) signal averaging, and (c) AKF methods.

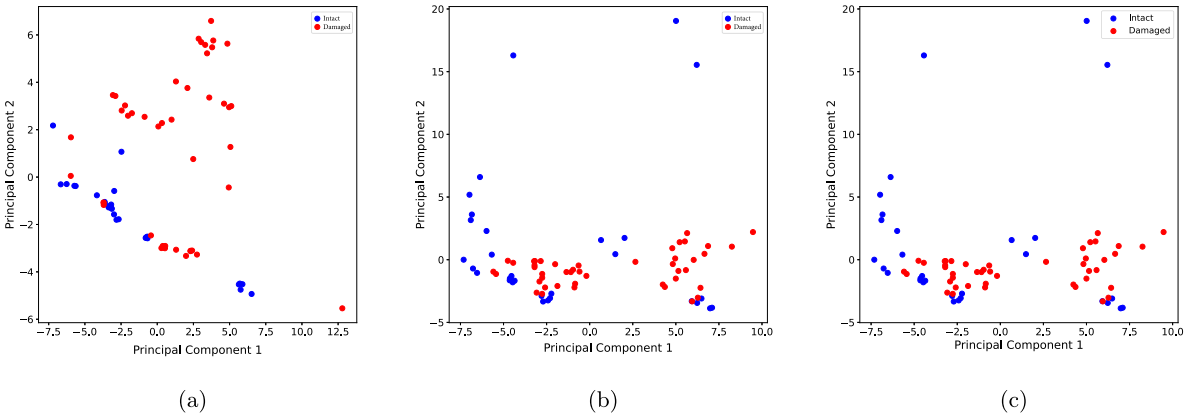


Fig. 18. Scatter plots of PC1 versus PC2 for the field data employing (a) cointegration, and (b) signal averaging and (c) AKF.

significance of considering both IA and IF for feature extraction. It appears that in the analysis mentioned, two features from the IMF_5 are considered significant for cointegration analysis, while one feature from IMF_5 is chosen for signal averaging techniques. However, no important features are selected from higher-order IMFs corresponding to the signal obtained from the AKF technique. This reasoning supports splitting the field signals into five IMFs, given that the cointegration technique provides more precise results evidenced by the upcoming discussions.

5.4.4. *Pearson correlation matrix*

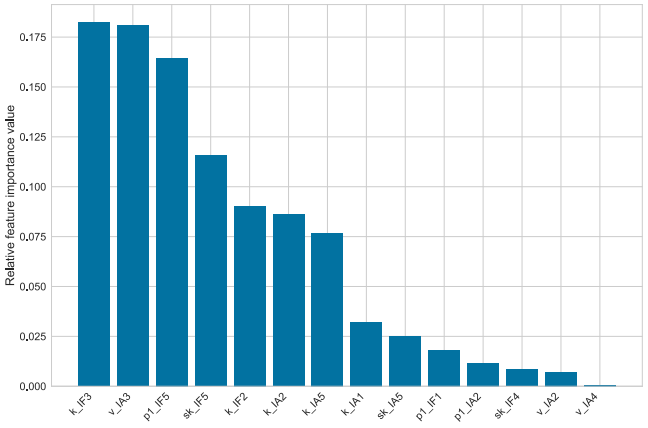
Figs. 20(a), 20(b) and 20(c) show the Pearson correlation factors for the features having non-zero importance in the cointegration, signal averaging and AKF methods, respectively. The results indicate correlation among a few features. Similar to the laboratory results, this may suggest the existence of collinearity among correlated features, where an interaction of such features needs to be incorporated in the model [55].

5.4.5. *Identifying the optimal number of features*

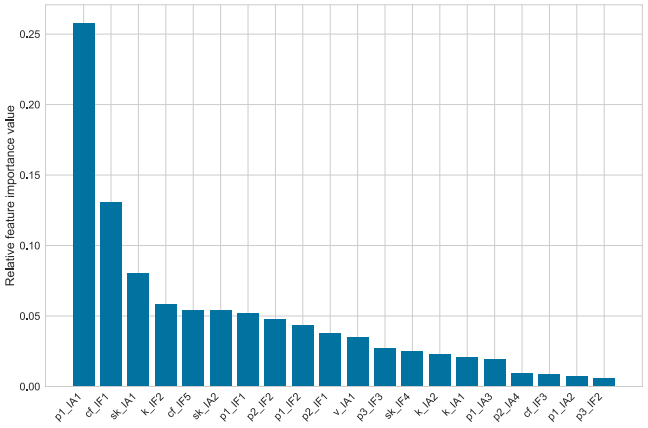
Figs. 21(a), 21(b) and 21(c) present the F1-score from 5-fold cross-validation of the XGBoost model performance on recursive selection of the most important features from the cointegration, signal averaging and AKF methods, respectively. The results indicate that the number of 7, 13 and 19 features selected respectively from the cointegration, signal averaging and AFK techniques yield the most accurate results. Once again, it is apparent from the plots that the results of the signal averaging and AKF methods are more dispersed. In contrast, the cointegration technique exhibits a smaller fluctuation of the F1 score across different folds using different number of features. This observation can further affirm the reliability of the cointegration method over the signal averaging and AKF method.

5.4.6. *Classification reports*

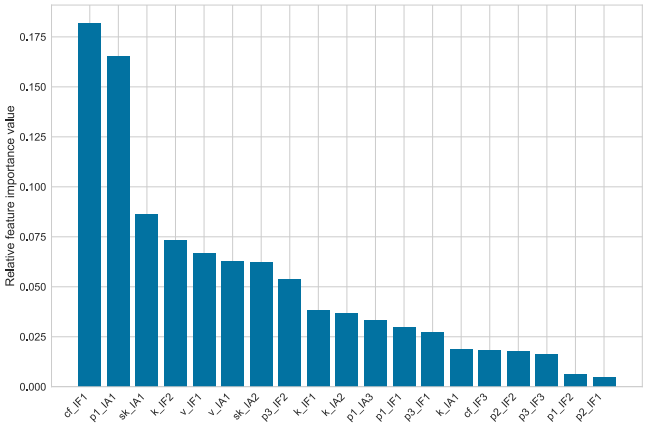
Fig. 22 shows the mean and standard deviation of various classification metrics, including precision, recall, and F1 score, evaluated after 100 runs of the MCCV of the XGBoost algorithm on all models. The datasets were randomly split into training and test sets, considering different random states at each run. The results clearly indicate the superior performance of the cointegration-based model. A significant improvement of the all metrics is evident in both classes. This result is noteworthy, as the accurate identification of defective poles is of great significance. Furthermore, using the cointegration method results in more reliable predictions, evidenced



(a)

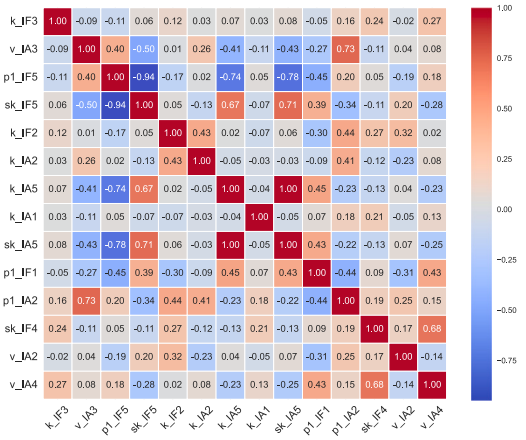


(b)

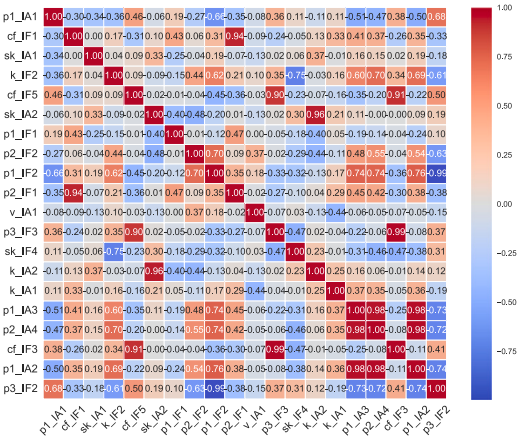


(c)

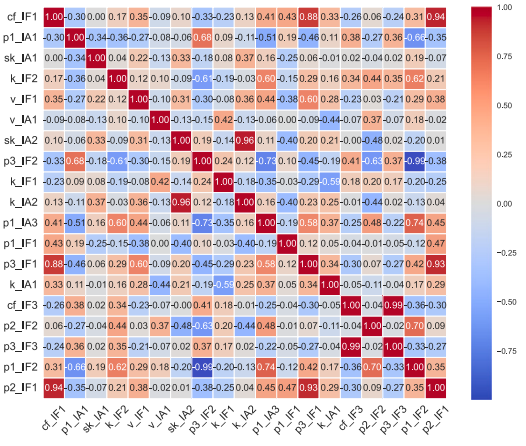
Fig. 19. Nonzero permutation feature importance results for the field data using (a) cointegration, (b) signal averaging, and (c) AKF methods.



(a)



(b)



(c)

Fig. 20. Pearson correlation factors for features with nonzero importance for the field experiment using (a) cointegration, (b) signal averaging, and (c) AKF methods.

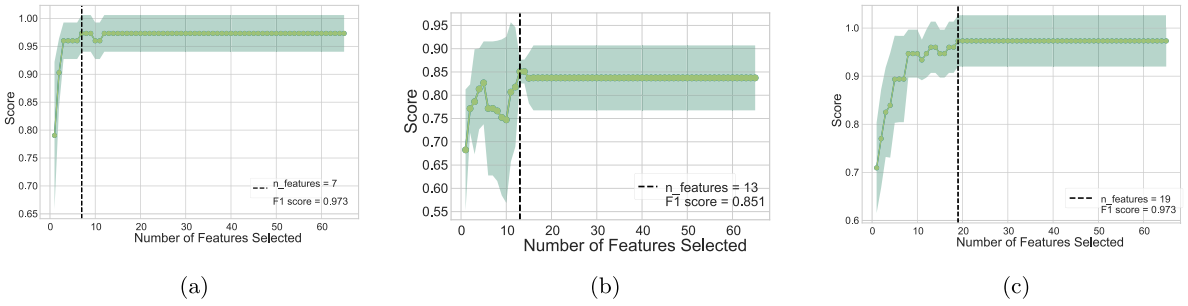


Fig. 21. Results of five-fold cross-validation F1 score for the recursive feature elimination of the field data, aimed at selecting the optimal number of features obtained from (a) cointegration, (b) signal averaging, and (d) AKF methods.

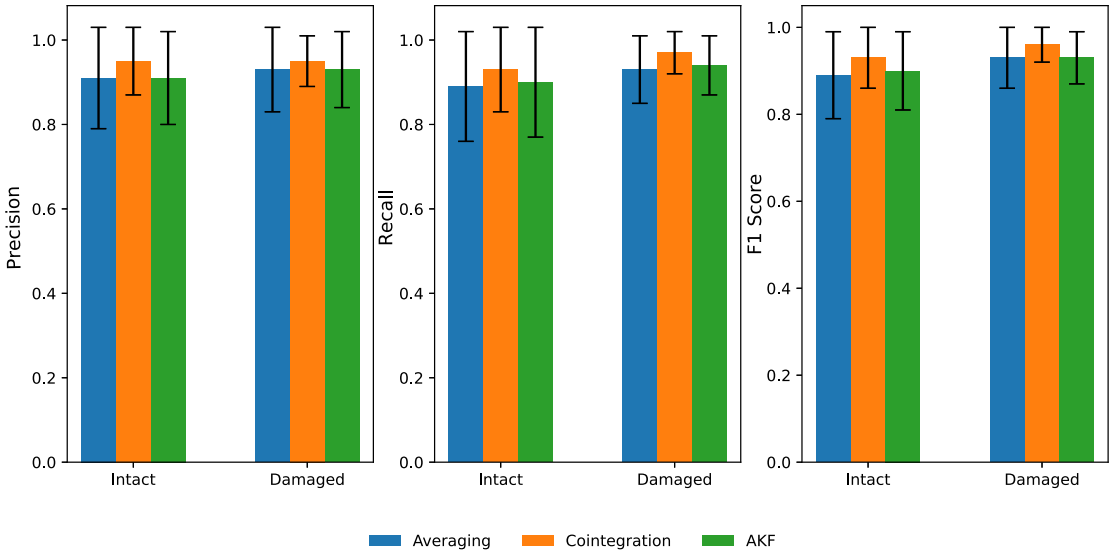


Fig. 22. Classification reports of models trained on the most important features extracted from the field data, evaluated over 100 runs of Monte Carlo Cross-Validation (MCCV) with random train-test splits for each run.

by the smaller standard deviation obtained across different folds. This result underscores the generalizability of the cointegration method compared to the signal averaging and the AKF approaches for real-life testing data recorded in the filed. Revisiting the results of the ADF test in Fig. 5 confirms that the cointegration method is a more suitable fusion technique in this case.

5.4.7. The Receiver Operating Characteristic (ROC) curve

Fig. 23 shows the ROC curves obtained for all three methods by computing the average value of the TPR and FPR after 100 MCCV runs of the XGBoost algorithm on different random splits of the field data. The results demonstrate again the superior performance of the cointegration method evidenced by an AUC of 0.99 ± 0.01 compared to the signal averaging and AKF methods with an AUC of 0.97 ± 0.04 and 0.97 ± 0.05, respectively.

Therefore, our findings show that the NMI metric does not always directly correlate with higher classification accuracy. For instance, although the mean NMI value for the field-fused data using the Johansen cointegration method is lower than that of signal averaging and AKF (see Fig. 16), the machine learning algorithm performed better on features extracted from the signals fused using Johansen cointegration. This suggests that not only the amount of information retained (as indicated by NMI) is important, but also the type and quality of the information extracted by the fusion method.

6. Future work

Thus far, we have explored the effectiveness of various sensor data fusion techniques and proposed Johansen cointegration as an effective approach, particularly when signals display significant nonstationary properties. Additionally, we introduced the NMI

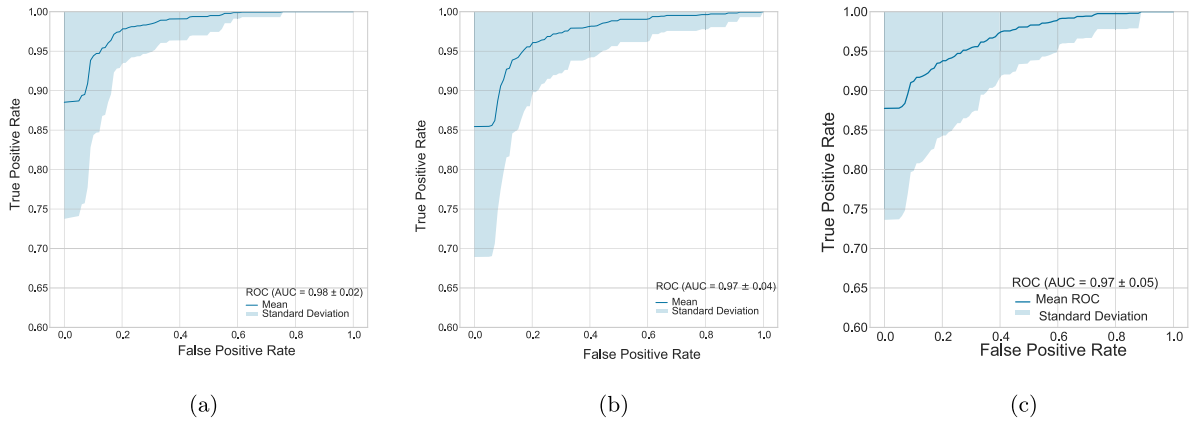


Fig. 23. ROCAUC curves obtained by averaging TPR and FPR obtained from 100 runs of MCCV of the models trained on the field data using the top most important features obtained from the (a) cointegration, (b) signal averaging, and (c) AKF methods.

Table 10
Percentage of the maximum NMI values from dynamic selection of the fusion technique for both the laboratory and field experiments.

Fusion method \ Experiment type	Laboratory (%)	Field (%)
Cointegration	33.33	36.67
Averaging	33.33	45.56
AKF	33.33	17.78

metric to evaluate fairness when comparing these techniques under specific values for VMD hyperparameters. Furthermore, we demonstrated that NMI can yield diverse values for different signals across various fusion techniques. This raises the question of whether selecting a dynamic fusion technique could enhance the final machine learning results. To illustrate this, we followed the procedures outlined below.

1. We propose to fuse signals using different signal fusion techniques.
2. We calculate the NMI metric for each fused signal obtained from different fusion strategies.
3. We select the fusion technique for each experiment that yields the maximum NMI value.
4. We construct a new feature matrix and investigate the final classification results based on it.

We have followed the above procedures for both laboratory and field data as follows. Table 10 indicates the number of incidents where using each fusion technique bore maximum NMI value for both the laboratory and field experiments. Our initial investigations suggest that combining the outcomes of various fusion methods listed in Table 10 could negatively impact the results of machine learning classification. This is likely due to the clashing information extracted from different fusion methods like cointegration, averaging, and AKF as observed in Figs. 6 and 15. A detailed analysis of these figures, along with the results presented in Figs. 10 and 11 for the laboratory experiment and Figs. 19 and 20 for the field experiment, indicates that averaging and AKF methods produce similar information. Consequently, we excluded Cointegration from the above-mentioned procedures and noticed a slight improvement of the final outcomes, implying that combining fusion techniques of the same type may improve the machine learning outcomes. Therefore, investigating a method to complement Johansen Cointegration for raw signal fusion could be a potential area for future research.

7. Conclusions

This paper presents a data-driven approach for fusing signals with pronounced nonstationary properties. To address this, a novel application of Johansen cointegration as a fusion technique for raw signals was proposed, replacing traditional raw signal averaging and more modern AKF methods. This method is particularly advantageous for scenarios where measurement signals demonstrate significant nonstationarity. The fusion based on the cointegration algorithm was shown to enhance the quality of extracted features, thereby improving classification results. The developed NMI metric confirmed that the only VMD parameters affecting the information passed through the decomposition process are the number of IMFs and the denoising factor α. Furthermore, the results highlighted a strong resemblance between the outcomes obtained from signal averaging and AKF methods. These techniques yielded similar features from the VMD algorithm, with only minor discrepancies in less significant features affecting XGBoost algorithm classification. Although AKF was preliminary developed to enhance the adaptability of the Kalman filter to nonstationary signals, it is a linear signal fusion technique in nature. Therefore, the similar results of applying signal averaging and AKF methods suggest that the signals produced by tactile transducer and further recorded at accelerometers may lack nonlinearity.

However, further investigation into the differences and similarities between AKF and signal averaging could be a subject for future research. It is worth noting that Johansen cointegration, like signal averaging and AKF, is also a linear technique in nature. The key distinction lies in its capability to handle nonstationarity in recorded signals which has been shown to be dominant compared to the AKF method.

Additionally, we have demonstrated that merging the outcomes of diverse signal fusion techniques may detrimentally affect classification results, whereas combining results from similar fusion techniques may yield only marginal enhancements. Therefore, future research could delve into a comprehensive examination of tactile transducer characteristics and the integration of various fusion strategies.

While the results of this study demonstrate the effectiveness of the proposed Johansen cointegration-based signal fusion method in field conditions, it is important to note that all damage scenarios were simulated. This approach was chosen due to the practical and ethical challenges associated with inducing real damage in critical infrastructure. Nevertheless, validating the methodology with real-world damage scenarios could further enhance the robustness and applicability of our findings. Therefore, pursuing real-world validation represents a valuable direction for future research.

Moreover, as a direction for future work, we recognize the importance of mitigating uncertainties and variability conditions that arise from environmental factors. To further enhance the robustness of our approach, integrating machine learning-aided data normalization techniques, as discussed in studies such as [56,57], could establish a solid methodological foundation. These techniques offer potential for more accurately normalizing sensor data and fusing signals across varying conditions, ultimately improving the reliability of damage detection in real-world SHM applications. Future efforts will focus on incorporating such data-driven methods to better account for environmental variability, thereby enhancing the accuracy and robustness of our proposed approach.

To improve the performance of the cointegration method applied to laboratory signals, one potential solution is to refine the treatment of outliers, which can disproportionately impact the results of the cointegration analysis. Incorporating robust statistical techniques, such as outlier detection and filtering, or leveraging robust cointegration methods that reduce the influence of outliers, may improve the identification accuracy of the cointegration method for laboratory signals. This approach represents a promising avenue for further investigation to refine our methodology.

Looking ahead, we recognize that employing a combination of fusion techniques based on the proportion of nonstationary signals within a group (e.g., four signals) could yield even better results. This adaptive approach is a promising future direction for this research, enabling dynamic selection of fusion methods depending on the characteristics of the signal set. As such, the current study represents the first phase of a broader and more comprehensive investigation into signal fusion techniques.

CRediT authorship contribution statement

Mohsen Mousavi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Ulrike Dackermann:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Data curation, Conceptualization. **Sahar Hassani:** Writing – review & editing, Writing – original draft, Visualization, Conceptualization. **Mahbube Subhani:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Funding acquisition, Data curation, Conceptualization. **Amir H. Gandomi:** Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Acknowledgments

We would like to express our sincere gratitude to Division 8.2 of the Federal Institute for Materials Research and Testing (BAM), Germany, with special appreciation to Ernst Niederleithinger and Herbert Wiggenhauser, for their invaluable support in the development of the experimental testing system and the execution of the experimental testing for this research.

Appendix. Johansen cointegration

Let $Y_t = \{X_1(t), \ldots, X_m(t)\}^T$ be a set of m signals recorded at time t. To calculate the Johansen cointegration representation of the signals the following steps are followed:

Table A.11
The model properties for different cases of $rank(\Pi) = r$.

Rank (Π) = r	Properties
r = 0	(1) The eigenvalues of Π are all equal to zero. (2) $\Pi = \mathbf{0}$. (3) The set of signals $\{X_1, \dots, X_m\}$ is not cointegrated.
r = m	(1) $\lvert\Pi\rvert \neq 0$. (2) $\{X_1, \dots, X_m\}$ have a zero order of integration ($I(0)$). (3) The association among $X_1, \dots, X_m$ can be represented using their original values and not through their differences. (4) It is unnecessary to make a reference to the error correction representation.
$0 < r < m$	(1) All the variables in $X_1, \dots, X_m$ have an order of integration of one ($I(1)$). (2) The matrix Π has r nonzero eigenvalues. (3) There exists r co-integration relationships among all the variables $X_1, \dots, X_m$.

1. First, the autoregressive (VAR) model of Y_t is established as follows:

$$Y_t = \delta D_t + \sum_{j=1}^{p} \Phi_j Y_{t-j} + u_t \tag{A.1}$$

where p is the lag order; δ is a term that captures any fixed or time-varying component not explained by the lagged values of Y_t or other independent variables included in the model; D_t indicates the vector of deterministic variables such as constant, seasonal dummy variables, or trends; Φ_j denotes an $m \times m$ coefficient matrix; and u_t is an $m \times 1$ independently and identically distributed (IID) Gaussian noise vector.

2. Next, the error correction model (ECM) is obtained by substituting $Y_t = Y_{t-1} + \Delta Y_t$, $Y_{t-1} = Y_{t-2} + \Delta Y_{t-1}$, $\dots$, $Y_{t-p} = Y_{t-p-1} + \Delta Y_{t-p}$ into (A.1) as follows:

$$\Delta Y_t = \Gamma_o D_t + \Pi Y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta Y_{t-j} + u_t \tag{A.2}$$

where, for $j = 1, \dots, p-1$, we have $\Pi = -(I - \Phi_1 - \cdots \Phi_p)$ and $\Gamma_j = -(\Phi_{j+1} + \cdots \Phi_p)$, and ΠY_{t-1} denotes error-correction term.

3. Table A.11

In such a case, it is possible to factorize Π as $\alpha\ \beta^T$, where α and β represent the adjustment and cointegration matrices, respectively. It is important to note that this factorization is not unique, and further constraints are required to obtain a unique factorization. To achieve this, Johansen proposed a maximum likelihood method, which is described as follows.

4. By substituting $Z_{0t} = \Delta Y_t$, $Z_{1t} = Y_{t-1}$, $Z_{2t} = \{\Delta Y_{t-1}, \dots, \Delta Y_{t-p-1}, D_t\}^T$, $\Psi = \{\Gamma_1, \dots, \Gamma_p, \Gamma_0\}$, and $\Pi = \alpha\ \beta^T$ into (A.2) we obtain:

$$Z_{0t} = \alpha\ \beta^T Y_{t-1} + \Psi Z_{2t} + \epsilon_t \tag{A.3}$$

5. The log-likelihood function can be formulated by assuming that $\epsilon_t \sim N(0, \Sigma)$, where Σ represents the covariance matrix, as follows:

$$lnL(\alpha, \beta, \Sigma | Y_t) = -\frac{mN}{2} log(2\pi) - \frac{N}{2} log(|\Sigma|) \tag{A.4}$$
$$-\frac{1}{2} \sum_{t=1}^{N} \left(Z_{0t} - \alpha\beta^T Z_{1t} - \Psi Z_{2t}\right)^T \Sigma^T \left(Z_{0t} - \alpha\beta^T Z_{1t} - \Psi Z_{2t}\right)$$

where N denotes the number of observations.

6. The residuals R_{0t} and R_{1t} are obtained by regressing Z_{0t} and Z_{1t}, respectively, on Z_{2t}. This results in the estimation of the unknown parameters α and β in (A.4). Therefore, the VECM of (A.2) can be written as follows:

$$S_{ij} = \frac{1}{N} \sum_{t=1}^{N} R_{it} R_{jt} \tag{A.5}$$

7. The following eigenvalue problem is solved to obtain α and β:

$$\left|\lambda_i S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0 \tag{A.6}$$

8. The cointegrating matrix β is formed using the r eigenvectors $v_1, \dots, v_r$ of (A.6) as follows:

$$\hat{\beta} = \beta_{MLE} = [v_1, \dots, v_r] \tag{A.7}$$

According to [36], the eigenvector associated with the highest eigenvalue denotes the first co-integrating vector, which can be considered as the "most stationary" one.

Data availability

The data that has been used is confidential.

References

[1] S. Hassani, U. Dackermann, A systematic review of advanced sensor technologies for non-destructive testing and structural health monitoring, Sensors 23 (4) (2023) 2204.
[2] R.-T. Wu, M.R. Jahanshahi, Data fusion approaches for structural health monitoring and system identification: Past, present, and future, Struct. Health Monit. 19 (2) (2020) 552–586.
[3] B.P.L. Lau, S.H. Marakkalage, Y. Zhou, N.U. Hassan, C. Yuen, M. Zhang, U.-X. Tan, A survey of data fusion in smart city applications, Inf. Fusion 52 (2019) 357–374.
[4] B.R. Manning, C.J. Cochrane, P.M. Lenahan, An improved adaptive signal averaging technique for noise reduction and tracking enhancements in continuous wave magnetic resonance, Rev. Sci. Instrum. 91 (3) (2020).
[5] Q. Liu, Y.-F. Chen, S.-Z. Fan, M.F. Abbod, J.-S. Shieh, Quasi-periodicities detection using phase-rectified signal averaging in EEG signals as a depth of anesthesia monitor, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (10) (2017) 1773–1784.
[6] M. Balara, D. Dupláková, D. Matisková, Application of a signal averaging device in robotics, Measurement 115 (2018) 125–132.
[7] A. Khurana, K. Nagla, Signal averaging for noise reduction in mobile robot 3D measurement system, MAPAN 33 (1) (2018) 33–41.
[8] B. Baumann, C.W. Merkle, R.A. Leitgeb, M. Augustin, A. Wartak, M. Pircher, C.K. Hitzenberger, Signal averaging improves signal-to-noise in OCT images: But which approach works best, and when, Biomed. Opt. Express 10 (11) (2019) 5755–5775.
[9] I. Chapalo, A. Petrov, D. Bozhko, M. Bisyarin, O. Kotov, Methods of signal averaging for a multimode fiber interferometer: An experimental study, in: Optical Sensors 2019, vol. 11028, SPIE, 2019, pp. 459–469.
[10] C. G. Krishnanunni, B. N. Rao, Indirect health monitoring of bridges using tikhonov regularization scheme and signal averaging technique, Struct. Control Health Monit. 28 (3) (2021) e2686.
[11] M.W. O'Brien, B.D. Schwartz, J.L. Petterson, D.S. Kimmerly, Comparison of signal-averaging and regression approaches to analyzing sympathetic transduction, Clin. Auton. Res. 32 (4) (2022) 299–302.
[12] Y. Liu, X. Fan, C. Lv, J. Wu, L. Li, D. Ding, An innovative information fusion method with adaptive Kalman filter for integrated INS/GPS navigation of autonomous vehicles, Mech. Syst. Signal Process. 100 (2018) 605–616.
[13] S.d. Pellegrini, F.C. Trigo, R. Lima, Adaptive Kalman filter-based information fusion in electrical impedance tomography for a two-phase flow, Mech. Syst. Signal Process. 150 (2021) 107326.
[14] A. Entezami, A. Entezami, Feature extraction in time domain for stationary data, in: Structural Health Monitoring by Time Series Analysis and Statistical Distance Measures, Springer, 2021, pp. 17–45.
[15] E. Figueiredo, G. Park, J. Figueiras, C. Farrar, K. Worden, Structural Health Monitoring Algorithm Comparisons Using Standard Data Sets, Technical Report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2009.
[16] T. Buckley, B. Ghosh, V. Pakrashi, A feature extraction & selection benchmark for structural health monitoring, Struct. Health Monit. 22 (3) (2023) 2082–2127.
[17] M. Mousavi, M.S. Taskhiri, A.H. Gandomi, Standing tree health assessment using contact–ultrasonic testing and machine learning, Comput. Electron. Agric. 209 (2023) 107816.
[18] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 454 (1971) (1998) 903–995.
[19] M. Mousavi, D. Holloway, J. Olivier, A.H. Gandomi, Beam damage detection using synchronisation of peaks in instantaneous frequency and amplitude of vibration data, Measurement 168 (2021) 108297.
[20] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. 62 (3) (2013) 531–544.
[21] M. Mousavi, A.H. Gandomi, Prediction error of Johansen cointegration residuals for structural health monitoring, Mech. Syst. Signal Process. 160 (2021) 107847.
[22] M. Mousavi, A.H. Gandomi, D. Holloway, A. Berry, F. Chen, Machine learning analysis of features extracted from time–frequency domain of ultrasonic testing results for wood material assessment, Constr. Build. Mater. 342 (2022) 127761.
[23] I. Rish, et al., An empirical study of the naive Bayes classifier, in: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3, 2001, pp. 41–46, 22.
[24] G.P. Dwyer, The Johansen tests for cointegration, White Pap. 516 (2015).
[25] E.J. Cross, K. Worden, Cointegration and why it works for SHM, in: J. Phys.: Conf. Ser., 382, (1) IOP Publishing, 2012, 012046.
[26] D. Gabor, Theory of communication. Part 1: The analysis of information, J. Inst. Electr. Eng. III 93 (26) (1946) 429–441.
[27] I. Muskhelishvili, Nikolaĭ, J.R.M. Radok, Singular Integral Equations: Boundary Problems of Function Theory and Their Application to Mathematical Physics, Courier Corporation, 2008.
[28] D. Zosso, Variational mode decomposition, MATLAB central file exchange, 2020, URL https://www.mathworks.com/matlabcentral/fileexchange/44765-variational-mode-decomposition. (Retrieved 27 August 2020).
[29] M.S. Khorshidi, M.R. Nikoo, M. Sadegh, Optimal and objective placement of sensors in water distribution systems using information theory, Water Res. 143 (2018) 218–228.
[30] M.S. Khorshidi, M.R. Nikoo, N. Taravatrooy, M. Sadegh, M. Al-Wardy, G.A. Al-Rawas, Pressure sensor placement in water distribution networks for leak detection using a hybrid information-entropy approach, Inform. Sci. 516 (2020) 56–71.
[31] J.L. Doob, Stochastic Processes, vol. 101, New York Wiley, 1953.
[32] K. Zolna, P.B. Dao, W.J. Staszewski, T. Barszcz, Towards homoscedastic nonlinear cointegration for structural health monitoring, Mech. Syst. Signal Process. 75 (2016) 94–108.
[33] P.B. Dao, W.J. Staszewski, Data normalisation for lamb wave–based damage detection using cointegration: A case study with single-and multiple-temperature trends, J. Intell. Mater. Syst. Struct. 25 (7) (2014) 845–857.
[34] E.J. Cross, K. Worden, Q. Chen, Cointegration: A novel approach for the removal of environmental trends in structural health monitoring data, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 467 (2133) (2011) 2712–2732.
[35] S. Hassani, M. Mousavi, U. Dackermann, Johansen cointegration of frequency response functions contaminated with nonstationary colored noise for structural damage detection, J. Sound Vib. 552 (2023) 117641.
[36] H. Shi, K. Worden, E.J. Cross, A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring, Mech. Syst. Signal Process. 103 (2018) 381–397.
[37] D. Kwiatkowski, P.C. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? J. Econometrics 54 (1–3) (1992) 159–178.

[38] D.A. Dickey, W.A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, J. Am. Stat. Assoc. 74 (366a) (1979) 427–431.
[39] A. Entezami, H. Sarmadi, B. Behkamal, Short-term damage alarming with limited vibration data in bridge structures: A fully non-parametric machine learning technique, Measurement 235 (2024) 114935.
[40] A. Entezami, H. Sarmadi, B. Behkamal, C. De Michele, On continuous health monitoring of bridges under serious environmental variability by an innovative multi-task unsupervised learning method, Struct. Infrastruct. Eng. (2023) 1–19.
[41] Y. Yu, U. Dackermann, J. Li, E. Niederleithinger, Wavelet packet energy–based damage identification of wood utility poles using support vector machine multi-classifier and evidence theory, Struct. Health Monit. 18 (1) (2019) 123–142.
[42] U. Dackermann, Y. Yu, E. Niederleithinger, J. Li, H. Wiggenhauser, Condition assessment of foundation piles and utility poles based on guided wave propagation using a network of tactile transducers and support vector machines, Sensors 17 (12) (2017) 2938.
[43] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
[44] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, et al., Xgboost: Extreme gradient boosting, 2015, pp. 1–4, R package version 0.4-2 1.
[45] G. Shan, Monte Carlo cross-validation for a study with binary outcome and limited sample size, BMC Med. Inform. Decis. Mak. 22 (1) (2022) 1–15.
[46] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, J. Global Optim. 11 (1997) 341–359.
[47] L. Zhang, W. Shaoping, M.S. Selezneva, K.A. Neusypin, A new adaptive Kalman filter for navigation systems of carrier-based aircraft, Chin. J. Aeronaut. 35 (1) (2022) 416–425.
[48] R.S. Tsay, Analysis of Financial Time Series, John Eiley and Sons, 2005.
[49] J.B. Ramsey, Tests for specification errors in classical linear least-squares regression analysis, J. R. Stat. Soc. Ser. B Stat. Methodol. 31 (2) (1969) 350–371.
[50] S.M. Goldfeld, R.E. Quandt, Some tests for homoscedasticity, J. Am. Stat. Assoc. 60 (310) (1965) 539–547.
[51] C. Brüser, J.M. Kortelainen, S. Winter, M. Tenhunen, J. Pärkkä, S. Leonhardt, Improvement of force-sensor-based heart rate estimation using multichannel data fusion, IEEE J. Biomed. Health Inf. 19 (1) (2014) 227–235.
[52] Y. Wang, R. Markert, J. Xiang, W. Zheng, Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system, Mech. Syst. Signal Process. 60 (2015) 243–251.
[53] M. Mousavi, A.H. Gandomi, Wood hole-damage detection and classification via contact ultrasonic testing, Constr. Build. Mater. 307 (2021) 124999.
[54] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[55] M. Mousavi, A.H. Gandomi, M. Abdel Wahab, B. Glisic, Monitoring onsite-temperature prediction error for condition monitoring of civil infrastructures, Struct. Control Health Monit. 29 (12) (2022) e3112.
[56] E. Figueiredo, G. Park, C.R. Farrar, K. Worden, J. Figueiras, Machine learning algorithms for damage detection under operational and environmental variability, Struct. Health Monit. 10 (6) (2011) 559–572.
[57] B. Behkamal, A. Entezami, C. De Michele, A.N. Arslan, Elimination of thermal effects from limited structural displacements based on remote sensing by machine learning techniques, Remote Sens. 15 (12) (2023) 3095.