lajss.org
Latin American Journal of Solids and Structures
www.lajss.org

Structural damage detection utilizing probabilistic divergence of FFT coefficient ratios and DBSCAN cluster algorithm

Guoqing Li 1[a*] https://orcid.org/0009-0003-9712-5087, Hao Wei 2[b] https://orcid.org/0009-0006-2251-7344

[a] College of Civil and Transportation Engineering, Shenzhen University, Shenzhen, China. Email: 2250471001@email.szu.edu.cn
[b] Designing Institute, Zhejiang Southeast Space Frame Company Limited, Zhejiang, China. Email: hhhw3691@gmail.com

* Corresponding author

Abstract

In structural health monitoring (SHM), uncertainties from environmental noise and modeling errors affect damage detection accuracy. This paper introduces a new concept: the Fast Fourier Transform Coefficient Ratio (FFTCR), which develops a damage indicator (DI) based on a probability model. The method involves calculating the probability density function (PDF) of the ratio and the Jensen-Shannon divergence (JSD) square root distance, which measures similarity between probability distributions. A baseline threshold is proposed using the Markov Chain Monte Carlo (MCMC) method and Monte Carlo discordancy test to handle uncertainties in statistical parameters. The robustness of the method is validated through the DBSCAN algorithm, which distinguishes damage states by applying a 1% exclusive threshold to the clustering results. The method requires no postprocessing of raw data, relying solely on response measurements. It is validated through two experiments, including laboratory ambient vibration tests and a field test on a bridge under forced vibration. Results show superior accuracy and broader applicability compared to the Mahalanobis Squared Distance (MSD) method.

Keywords
Damage detection, probability density function, uncertainty quantification, probabilistic distance, clustering

Graphical Abstract

1 INTRODUCTION

Engineering infrastructures are vulnerable to degradation due to human activities and environmental influences, which can accelerate damage accumulation and reduce the lifespan of structures. To address this, SHM systems have been widely adopted across various disciplines, including mechanical, aerospace, and civil engineering (Hu et al., 2022; Yang and Yang, 2018).

Structural damage detection techniques are generally classified into two main categories: non-destructive testing (NDT) and vibration-based methods. NDT methods are localized approaches that often fail to identify internal structural damages (He and Ng, 2017; Yan et al., 2020). In contrast, vibration-based methods analyze alterations in the overall vibration characteristics of structures, thus providing a comprehensive assessment (Avci et al., 2021). These methods have gained significant attention over recent decades due to their global approach to identifying structural integrity (Hou and Xia, 2021).

The field of structural damage detection has increasingly focused on leveraging the alterations in dynamic properties derived from ambient vibration data. One of the persistent challenges in practical applications is the insufficient sensitivity of conventional dynamic indicators like natural frequencies and modal shapes to the presence of structural damage (Capecchi et al., 2016). This has driven the need to develop novel diagnostic indicators more responsive to such damages. As the search for more effective damage detection methods continues, the Transmissibility Function (TF) has emerged as a promising alternative (Chesné and Deraemaeker, 2013). Unlike the Frequency Response Function (FRF)(Bandara et al., 2014; Liu et al., 2009), which delineates the input-output relationships within dynamic systems, TF provides a mathematical framework for understanding output-to-output relationships without requiring direct measurement of the input forces (Devriendt and Guillaume, 2007). This characteristic notably reduces the dependency on precise input measurements and alleviates the impact of variations in input on the damage detection process (Soofi and Bitaraf, 2022; Wu et al., 2021). Furthermore, TF-based approaches offer significant advantages by eliminating the need for modal identification and avoiding reliance on analytical or discretized structure models (Maia et al., 2011). These attributes make TF-based damage detection not only innovative but also practical for real-world applications. As a result, TF-based strategies have gained considerable interest for their robustness and efficacy in identifying structural damages across various engineering domains.

A significant advancement occurred when a NASA research team suggested utilizing the integral across a frequency band of the discrepancy between two TFs from intact and damaged states as a damage indicator (Zhang et al., 1999). This approach has subsequently been embraced and expanded upon in subsequent studies (Schallhorn and Rahmatalla, 2015; Zhu et al., 2010). Wu et al. (2021) introduced a novel method for detecting and localizing debonding beneath concrete pavement through transmissibility function (TF) analysis, emphasizing easy and low-cost implementation. It detailed the measurement node array layout, signal processing, and testing procedures, and used finite element analysis (FEA) and field tests to validate the effectiveness of TF analysis in debonding detection and localization. Zhang et al. (2016) introduced a novel generalized transmissibility damage indicator (GTDI) that enhanced condition monitoring of engineering structures by improving sensitivity and reducing noise.

Worden and his team (Manson et al., 2003; Worden and Manson, 2003, 2007; Worden et al., 2000) devised a range of innovative indices using the TF, integrating pattern recognition and machine learning algorithms to identify potential features for damage detection. Recently, it has been extensively documented that clustering methods are effective for novelty detection (Amarbayasgalan et al., 2018; Yu and Kang, 2023). Mei et al. (2022) introduced a data-driven structural novelty detection technology that utilized the transmissibility function (TF) and Gaussian mixture model (GMM) clustering.

TF-based damage detection methods are increasingly favoured in recent proposals. Nonetheless, existing theories largely focus on deterministic approaches. In practical settings, numerous factors introduce uncertainty (Yan et al., 2019), including the stochastic nature of vibrational responses, environmental variability, estimation errors, measurement noise and so on. These factors can cause variability in damage indicators, which may result in inaccurate interpretations of data. The cautionary characteristics resulting from multiple uncertainties might overshadow those stemming from genuine structural damage. Consequently, there is a critical demand for the development of theoretically robust methods that can effectively manage these uncertainties within the framework of TF-based damage detection. In this context, exploring the characterization of uncertainty in TF merits additional research. In this regard, Mao and Todd's research is a pioneering effort that developed quantified uncertainty models for TFs, estimated using Power Spectral Density (PSD) (Mao and Todd, 2012, 2013). Recently, there has been some progress in research on the uncertainty of TF-based damage detection (Mei et al., 2022; Mei et al., 2023; Yan et al., 2022).

In this study, to improve the precision of anomaly detection by addressing these uncertainties, a new data-driven damage detection approach is proposed that relies solely on response measurements. This method requires no postprocessing of the raw data. Inspired by the concept of the TF, the Fast Fourier Transform Coefficient Ratio (FFTCR) is introduced and detailed in Sect. 2 of this article. A novel DI based on the probability model of FFTCR is developed, involving the computation of the PDF of FFTCR and the square root distance derived from the JSD. The JSD is utilized to assess the similarity between two probability distributions and has been proposed as a novel diagnostic tool for structural damage detection due to its high sensitivity in detecting minor changes between distributions (Zhang et al., 2020).

To accommodate the uncertainties of statistical parameters, the MCMC resampling method is used to determine the statistical threshold under the intact state. The determination of this threshold can effectively distinguish the DI values of the intact state from those of the potentially damaged state. DBSCAN is a clustering algorithm that identifies clusters of varying shapes and sizes in datasets. It effectively detects outliers (noise points) and does not require the specification of the number of clusters in advance (Civera et al., 2023). In this study, the DBSCAN clustering algorithm is used to further investigate the robustness of the proposed method against uncertainties. Simultaneously, combining the results of the clustering with the threshold method also allows for the validation of structural damage outcomes. Experimental data from two case studies are employed to confirm the feasibility and effectiveness of the proposed methodology.

There are several advantages to the FFTCR-based damage detection method: (i) This proposed method effectively bypasses the need for complex operations involving imaginary numbers by ensuring that all calculations and procedures are conducted strictly within the real number domain, thereby simplifying the computational process and enhancing practical applicability. (ii) The FFTCR-based damage detection exhibits greater sensitivity to structural damages than traditional dynamic properties. (iii) FFTCR data can be directly utilized, circumventing potential errors associated with modal identification. (iv) The FFTCR can yield a richer dataset as it can be measured at multiple points within a structure. (v) Damage detection can be enhanced by utilizing numerous data points from FFTCRs across an extensive frequency range.

2 Analytical probability model of FFTCR

Given a linear system subjected to stationary excitation, time-domain response measurements are collected at two spatial locations, m and n, represented as $\mathbf{y}_m(t)$ and $\mathbf{y}_n(t)$. These measurements consist of N_L discrete data points: $\mathbf{y}_m = \{y_{m,1}, y_{m,2}, \dots, y_{m,N_L}\}$ and $\mathbf{y}_n = \{y_{n,1}, y_{n,2}, \dots, y_{n,N_L}\}$. By applying the Fast Fourier Transform (FFT), these responses are converted into their frequency domain, $\mathbf{Y}_m^k = \mathbf{Y}_m^{\mathfrak{R}_k} + i\mathbf{Y}_m^{\mathfrak{I}_k}$ and $\mathbf{Y}_n^k = \mathbf{Y}_n^{\mathfrak{R}_k} + i\mathbf{Y}_n^{\mathfrak{I}_k}$,where $\mathbf{Y}_i^{\mathfrak{R}_k}$ and $\mathbf{Y}_i^{\mathfrak{I}_k}$ $(i = m,n)$ represent the real and imaginary parts of $\mathbf{Y}_i^k$, respectively. In this study, it should be noted that the symbol 'k', whether it appears as a subscript or a superscript, represents the frequency w_k .

To simplify, the frequency domain outcomes at points m and n are represented by the real vectors $\tilde{\mathbf{Y}}_m^k$ and $\tilde{\mathbf{Y}}_n^k$, defined respectively as $\begin{bmatrix} \mathbf{Y}_m^{\mathfrak{R}_k} & \mathbf{Y}_m^{\mathfrak{I}_k} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{Y}_n^{\mathfrak{R}_k} & \mathbf{Y}_n^{\mathfrak{I}_k} \end{bmatrix}$. According to the central limit theorem, as the sample size $N_L \to \infty$, both $Y_i^{\mathfrak{R}_k}$ and $Y_i^{\mathfrak{I}_k}$ $(i = m,n)$ exhibit approximate Gaussian distributions, characterized by means θ_i , variances σ_i^2 $(i = m,n)$, and the correlation coefficient ρ_{mn} (Yuen, 2010)

The Fast Fourier Transform Coefficient Ratio (FFTCR) is defined as the ratio of the real part of one point to the real part of another, and the ratio of the imaginary part of one point to the imaginary part of another in the FFT, as illustrated in the following expression:

$$T_{mn}^{\mathfrak{R}_k} = \frac{Re(\mathbf{Y}_m^k)}{Re(\mathbf{Y}_n^k)} = \frac{Y_m^{\mathfrak{R}_k}}{Y_n^{\mathfrak{R}_k}} \tag{1}$$

$$T_{mn}^{\mathfrak{I}_k} = \frac{Im(\mathbf{Y}_m^k)}{Im(\mathbf{Y}_n^k)} = \frac{Y_m^{\mathfrak{I}_k}}{Y_n^{\mathfrak{I}_k}} \tag{2}$$

The T_{mn}^k matrix for any two points is defined as follows:

$$\mathbf{T}_{mn}^k = \begin{bmatrix} \dfrac{Y_m^{\mathfrak{R}_k}}{Y_n^{\mathfrak{R}_k}} & \dfrac{Y_m^{\mathfrak{I}_k}}{Y_n^{\mathfrak{R}_k}} \\[2ex] \dfrac{Y_m^{\mathfrak{I}_k}}{Y_n^{\mathfrak{R}_k}} & \dfrac{Y_m^{\mathfrak{I}_k}}{Y_n^{\mathfrak{I}_k}} \end{bmatrix} \tag{3}$$

Eqs. (1) and (2) represent the ratio of two correlated normal random variables, and the Probability Density Function (PDF) of this ratio can be expressed as follows (Hinckley, 1969):

$$P(t_{mn}^{\Re_k}) = \frac{b(t_{mn}^{\Re_k})d(t_{mn}^{\Re_k})}{\sqrt{(2\pi)}\sigma_m^k\sigma_n^k a^3(t_{mn}^{\Re_k})}\left[\Phi(y_1)-\Phi(y_2)\right] + \frac{\sqrt{\left(1-\left(\rho_{mn}^k\right)^2\right)}}{\pi\sigma_m^k\sigma_n^k a^2(t_{mn}^{\Re_k})}\exp\left\{-\frac{c}{2\left(1-\left(\rho_{mn}^k\right)^2\right)}\right\} \tag{4}$$

$$P(t_{mn}^{\Im_k}) = \frac{b(t_{mn}^{\Im_k})d(t_{mn}^{\Im_k})}{\sqrt{(2\pi)}\sigma_m^k\sigma_n^k a^3(t_{mn}^{\Im_k})}\left[\Phi(y_1)-\Phi(y_2)\right] + \frac{\sqrt{\left(1-\left(\rho_{mn}^k\right)^2\right)}}{\pi\sigma_m^k\sigma_n^k a^2(t_{mn}^{\Im_k})}\exp\left\{-\frac{c}{2\left(1-\left(\rho_{mn}^k\right)^2\right)}\right\} \tag{5}$$

where t_{mn}^k represents the value of the random variable T_{mn}^k evaluated at the frequency w_k ,

$$a(t_{mn}^k) = \left(\frac{(t_{mn}^k)^2}{\left(\sigma_m^k\right)^2} - \frac{2\rho_{mn}^k t_{mn}^k}{\sigma_m^k\sigma_n^k} + \frac{1}{\sigma_2^2}\right)^{\frac{1}{2}},\ b(t_{mn}^k) = \frac{\theta_m^k t_{mn}^k}{\left(\sigma_m^k\right)^2} - \frac{\rho_{mn}^k(\theta_m^k + \theta_n^k t_{mn}^k)}{\sigma_m^k\sigma_n^k} + \frac{\theta_n^k}{\left(\sigma_n^k\right)^2},\ c = \frac{\left(\theta_m^k\right)^2}{\left(\sigma_m^k\right)^2} - \frac{2\rho_{mn}^k\theta_m^k\theta_n^k}{\sigma_m^k\sigma_n^k} + \frac{\left(\theta_n^k\right)^2}{\left(\sigma_n^k\right)^2},$$

$$d(t_{mn}^k) = \exp\left\{\frac{b^2(t_{mn}^k) - ca^2(t_{mn}^k)}{2\left(1-\left(\rho_{mn}^k\right)^2\right)a^2(t_{mn}^k)}\right\},\ \Phi(y) = \int_{-\infty}^{y}\phi(u)du,\ \ \phi(u) = \frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}u^2},\ y_1 = \frac{b(t_{mn}^k)}{\sqrt{\left(1-\left(\rho_{mn}^k\right)^2\right)}a(t_{mn}^k)}, y_2 = -y_1.$$

Eqs. (4) and (5) define the PDF for T_{mn}^k . These equations will be used to compute the square root distance of the JSD under different structural states, which will aid in applying JSD to damage detection.

3 Square root distance of JSD between two different PDFs

3.1 Definition of square root distance of JSD

The JSD is a method of measuring the similarity between two probability distributions (Lin, 1991). It is a symmetrized and smoothed version of the Kullback-Leibler divergence (KLD) (Kullback and Leibler, 1951). The KLD is widely applied in various fields, such as machine learning, where it supports algorithms like variational Bayes, and in information theory, particularly for signal processing. It is utilized wherever quantifying the ‘distance’ between probability distributions is necessary.

The KLD from a distribution $p(x)$ to a distribution $q(x)$, denoted $D_{KL}(p(x)\,||\,q(x))$, is a measure of the information lost when $q(x)$ is used to approximate $p(x)$ .It is given by:

$$D_{KL}(p(x)\,||\,q(x)) = \begin{cases} \sum\limits_{x\in X} p(x)\log\left(p(x)/q(x)\right) \\ \int_{-\infty}^{+\infty} p(x)\log\left(p(x)/q(x)\right)dx \end{cases} \tag{6}$$

However, KLD is not symmetric, meaning $D_{KL}(p(x)\,||\,q(x)) \neq D_{KL}(q(x)\,||\,p(x))$ .To overcome the asymmetry and to smooth the distributions, a mixture distribution is introduced, which is the average of $p(x)$ and $q(x)$:

$$M(x) = \frac{1}{2}(p(x)+q(x)) \tag{7}$$

The JSD is then defined using the KLD and the mixture distribution M. It is the average of the KLD of $p(x)$ to $M(x)$ and the KLD of $q(x)$ to $M(x)$:

$$JSD(p(x)\,||\,q(x)) = \frac{1}{2}D_{KL}(p(x)\,||\,M(x)) + \frac{1}{2}D_{KL}(q(x)\,||\,M(x)) \tag{8}$$

Substituting Eq. (6) into the Eq. (8) yields:

$$JSD(p(x)\,||\,q(x)) = \frac{1}{2}\int_{-\infty}^{+\infty} p(x)\log\left(\frac{p(x)}{M(x)}\right)dx + \frac{1}{2}\int_{-\infty}^{+\infty} q(x)\log\left(\frac{q(x)}{M(x)}\right)dx \tag{9}$$

The JSD possesses the properties of being symmetric, $JSD(p(x)\,||\,q(x)) = JSD(q(x)\,||\,p(x))$, and always yields values between 0 and 1. Here, 0 indicates identical distributions, and 1 indicates that the distributions do not overlap at all.

It is also common to take the square root of the JSD to obtain a metric that is a true distance measure in the mathematical sense because it satisfies the triangle inequality. This is known as the Square Root of the JSD (SR_{JSD}):

$$
\begin{aligned}
SR_{JSD}(p(x)\|q(x)) &= \sqrt{JSD(p(x)\parallel q(x))} \\
&= \sqrt{\frac{1}{2}\int_{-\infty}^{+\infty} p(x)\log\left(\frac{p(x)}{M(x)}\right)dx + \frac{1}{2}\int_{-\infty}^{+\infty} q(x)\log\left(\frac{q(x)}{M(x)}\right)dx} \\
&= \sqrt{\frac{1}{2}\int_{-\infty}^{+\infty}\left(p(x)\log\left(\frac{p(x)}{M(x)}\right) + q(x)\log\left(\frac{q(x)}{M(x)}\right)\right)dx}
\end{aligned}
\tag{10}
$$

This logic underpins the use of JSD in applications where a symmetric and smooth measure of divergence between two distributions is required, such as clustering probability distributions or comparing statistical samples in machine learning. In the subsequent section, the SR_{JSD} will be employed to develop a novel damage indicator.

3.2 Square root of JSD between FFTCRs under two structural states

Damage can alter FFTCRs, and these changes can be detected by computing the SR_{JSD} between the FFTCRs under undamaged and potentially damaged state. The PDFs of the real part-based and imaginary part-based FFTCRs under intact state are presented in Eqs. (4) and (5). To avoid confusion, the PDFs of FFTCR under damaged state are expressed as follows:

$$
P^d(t_{mn}^{\Re_k}) = \frac{b(t_{mn}^{\Re_k})d(t_{mn}^{\Re_k})}{\sqrt{(2\pi)}\sigma_m^{d,k}\sigma_n^{d,k}a^3(t_{mn}^{\Re_k})}\left[\Phi(y_1)-\Phi(y_2)\right] + \frac{\sqrt{\left(1-\left(\rho_{mo}^{d,k}\right)^2\right)}}{\pi\sigma_m^{d,k}\sigma_n^{d,k}a^2(t_{mn}^{\Re_k})}\exp\left\{-\frac{c}{2\left(1-\left(\rho_{mn}^{d,k}\right)^2\right)}\right\}
\tag{11}
$$

$$
P^d(t_{mn}^{\Im_k}) = \frac{b(t_{mn}^{\Im_k})d(t_{mn}^{\Im_k})}{\sqrt{(2\pi)}\sigma_m^{d,k}\sigma_n^{d,k}a^3(t_{mn}^{\Im_k})}\left[\Phi(y_1)-\Phi(y_2)\right] + \frac{\sqrt{\left(1-\left(\rho_{mn}^{d,k}\right)^2\right)}}{\pi\sigma_m^{d,k}\sigma_n^{d,k}a^2(t_{mn}^{\Im_k})}\exp\left\{-\frac{c}{2\left(1-\left(\rho_{mn}^{d,k}\right)^2\right)}\right\}
\tag{12}
$$

where the superscript d represents the potential damage state. Substituting Eqs. (4) and (5), as well as Eqs. (11) and (12) into Eq. (10), yields expressions for the real part-based $SR_{JSD}^{\Re_k}$ and the imaginary part-based $SR_{JSD}^{\Im_k}$ between the intact state and the potentially damaged state at w_k :

$$
SR_{JSD}^{\Re_k}(p(t_{mn}^{\Re_k})\parallel p^d(t_{mn}^{\Re_k})) = \sqrt{\frac{1}{2}\int_{-\infty}^{+\infty}\left(p(t_{mn}^{\Re_k})\log\left(\frac{p(t_{mn}^{\Re_k})}{M(t_{mn}^{\Re_k})}\right) + p^d(t_{mn}^{\Re_k})\log\left(\frac{p^d(t_{mn}^{\Re_k})}{M^d(t_{mn}^{\Re_k})}\right)\right)dx}
\tag{13}
$$

$$
SR_{JSD}^{\Im_k}(p(t_{mn}^{\Im_k})\parallel p^d(t_{mn}^{\Im_k})) = \sqrt{\frac{1}{2}\int_{-\infty}^{+\infty}\left(p(t_{mn}^{\Im_k})\log\left(\frac{p(t_{mn}^{\Im_k})}{M(t_{mn}^{\Im_k})}\right) + p^d(t_{mn}^{\Im_k})\log\left(\frac{p^d(t_{mn}^{\Im_k})}{M^d(t_{mn}^{\Im_k})}\right)\right)dx}
\tag{14}
$$

where the mixture distribution $M^d(x) = \frac{1}{2}(p(x)+p^d(x))$.

The square root of JSD is often preferred over JSD itself when the properties of a metric are required. It allows for the application of geometric interpretations and techniques that rely on metric properties, which is not possible with the JSD or KLD alone. Its boundedness, symmetry, and satisfaction of the triangle inequality make it a robust tool for comparing probability distributions. The square root of JSD, being a bona fide metric, offers a robust method for quantitatively assessing and visually representing the resemblance between probability distributions. This inherent advantage is applied in actual damage monitoring, which helps to accurately quantify the probability distribution differences under different damaged states.

3.3 Damage indicator based on the square root of JSD

The square root of JSD between FFTCRs under the intact state and the possibly damaged state can be utilized to detect novelties. To augment the robustness of this analytical technique, data fusion is performed across a variety of measurements and frequency points, thereby enriching the dataset and enhancing the accuracy of the detection process. For a multi-degree-of-freedom structure, select r response measurements for analysis $\mathbf{y}(t) = \{\mathbf{y}_1(t), \mathbf{y}_2(t), \ldots, \mathbf{y}_r(t)\}$ .Under the same reference response measurements $\mathbf{y}_n(t)$, the real part-based and the imaginary part-based $\mathbf{T}_n^k$ vectors, according to previous Eqs. (1) and (2), can be obtained:

$$
\mathbf{T}_n^{\Re_k} = \left\{T_{1n}^{\Re_k}, T_{2n}^{\Re_k}, \ldots, T_{rn}^{\Re_k}\right\}_{1\times(r-1)} = \left\{\frac{Y_1^{\Re_k}}{Y_n^{\Re_k}}, \frac{Y_2^{\Re_k}}{Y_n^{\Re_k}}, \ldots, \frac{Y_r^{\Re_k}}{Y_n^{\Re_k}}\right\}_{1\times(r-1)}
\tag{15}
$$

$$\mathbf{T}_n^{\Im_k} = \left\{ T_{1n}^{\Im_k}, T_{2n}^{\Im_k}, \ldots, T_{rn}^{\Im_k} \right\}_{1\times(r-1)} = \left\{ \frac{Y_1^{\Im_k}}{Y_n^{\Im_k}}, \frac{Y_2^{\Im_k}}{Y_n^{\Im_k}}, \ldots, \frac{Y_r^{\Im_k}}{Y_n^{\Im_k}} \right\}_{1\times(r-1)} \tag{16}$$

According to Eqs. (13) and (14), the square root of JSD between the intact state and the potentially damaged state can be calculated within a specific frequency band $[w_1, w_2]$. The damage indicator (DI) is derived by initially averaging the values from the selected response measurements and subsequently averaging these results across the entire frequency band. The specific formulas are as follows:

$$DI_T^{\mathfrak{R}} = \frac{1}{n_\omega (r-1)} \sum_{k=k_1}^{k_2} \sum_{m=1, n\neq m}^{r} SR_{JSD}^{\mathfrak{R}_k} \left(p\left(t_{mn}^{\mathfrak{R}_k}\right) \| \, p^d\left(t_{mn}^{\mathfrak{R}_k}\right) \right) \tag{17}$$

$$DI_T^{\Im} = \frac{1}{n_\omega (r-1)} \sum_{k=k_1}^{k_2} \sum_{m=1, n\neq m}^{r} SR_{JSD}^{\Im_k} \left(p\left(t_{mn}^{\Im_k}\right) \| \, p^d\left(t_{mn}^{\Im_k}\right) \right) \tag{18}$$

where n_ω is the total amount of frequency points throughout the given frequency range $[w_{k1}, w_{k2}]$. It should be explicitly noted that the DI proposed here does not depend solely on measurements from two specific points. Although the FFTCR is initially defined between two arbitrary points, the final DI values are derived by averaging FFTCR-based Jensen-Shannon divergence values across multiple pairs of measurement points and frequency points, as clearly indicated in Eqs. (17) and (18). This averaging procedure significantly reduces the sensitivity of the proposed method to the specific selection of individual measurement locations, thereby enhancing its robustness and general applicability to various structural systems. In the subsequent section, the paper will elaborate on the application of the DI for identifying structural anomalies. This process involves the use of threshold-based decision criteria and clustering algorithms, which enable the differentiation between normal and anomalous structural behaviors.

4 Damage detection based on threshold technology and DBSCAN clustering

Once these DI values from different structural states are calculated, the subsequent task involves distinguishing which values indicate a damaged state and which suggest a healthy state. The use of decision-making criteria based on threshold and DBSCAN clustering algorithm has been proven effective. The threshold-based method establishes a boundary between healthy and potentially damaged states; values exceeding this threshold indicate structural potential damage. The data grouped by clustering algorithm consist of DI values. Clustering algorithm categorizes all DI values, grouping similar values to denote the same state. The robustness of the DI values against uncertainty is directly reflected in the number of noise points within the clustering results. Therefore, while the clustering algorithm is used for state discrimination, it also serves to validate the robustness of the proposed method.

4.1 Damage detection based on threshold value

It is important to highlight that the accuracy of SRJSD-based structural anomaly detection is influenced by the variability of the model parameters $\Theta_k = \left\{ \sigma_m^k, \theta_m^k, \sigma_n^k, \theta_n^k, \rho_{mn}^k \right\}, (k \in [k_1, k_2])$, which encapsulates the inherent uncertainties from observational data and model assumptions. To mitigate these uncertainties, a MCMC resampling scheme is adopted. This scheme involves repeatedly sampling and estimating the model parameters to better capture the uncertainty. By employing the MCMC method to generate samples from the distribution of FFT coefficient matrices ($\mathbf{A}_{n_w \times n_z}$ and $\mathbf{B}_{n_w \times n_z}$ are composed of real part data and imaginary part data, respectively), Monte Carlo significance testing is subsequently conducted to determine thresholds for the 1% discordancy tests from DIs.

4.1.1 MCMC resampling scheme

MCMC is a statistical method used for estimating the distribution of parameters within complex models, especially when direct sampling is challenging. This approach simulates a sequence of random samples, which are designed to converge over time to the target probability distribution. Resampling is conducted on the FFT coefficient matrices ($\mathbf{A}_{n_w \times n_z}$ and $\mathbf{B}_{n_w \times n_z}$) derived from the structural response. In these matrices, rows correspond to frequency points and columns to segments of the measurement data. The resampling adheres to the probability distribution specific to each row within matrices $\mathbf{A}_{n_w \times n_z}$ and $\mathbf{B}_{n_w \times n_z}$,ensuring that the resampled data maintains consistency with the original probability distribution at specific frequencies. During the sampling process of MCMC methods, the Metropolis-Hastings Algorithm was employed. Take the resampling of matrix $\mathbf{A}$ as an example to illustrate the implementation process of the Metropolis-Hastings Algorithm. The resampling for matrix $\mathbf{B}$ is similar. The Metropolis-Hastings Algorithm generally consists of three main steps: Step (1): Statistical parameters calculation; Step (2): Initial sample; Step (3): Proposal and Acceptance. The formula for acceptance probability is as shown in Eq. (19). The flowchart of the algorithm is shown in Figure. 1.

$$\alpha(x_A, x_{Apro}) = \min\left(1, \exp\left(-\frac{(x_{Apro} - mean(\mathbf{A}_{n_w \times n_z}(i,:)))^2}{2std(\mathbf{A}_{n_w \times n_z}(i,:))^2} + \frac{(x_A - mean(\mathbf{A}_{n_w \times n_z}(i,:)))^2}{2std(\mathbf{A}_{n_w \times n_z}(i,:))^2}\right)\right) \tag{19}$$

Figure 1 detailed explanation the implementation of the Metropolis-Hastings algorithm for resampling within a specified statistical framework, typically involving the mean and standard deviation of distributions. Each step is crafted to ensure that the sampled values closely mirror the underlying characteristics of the distribution, thereby leveraging the principles of MCMC to attain effective and representative sampling.

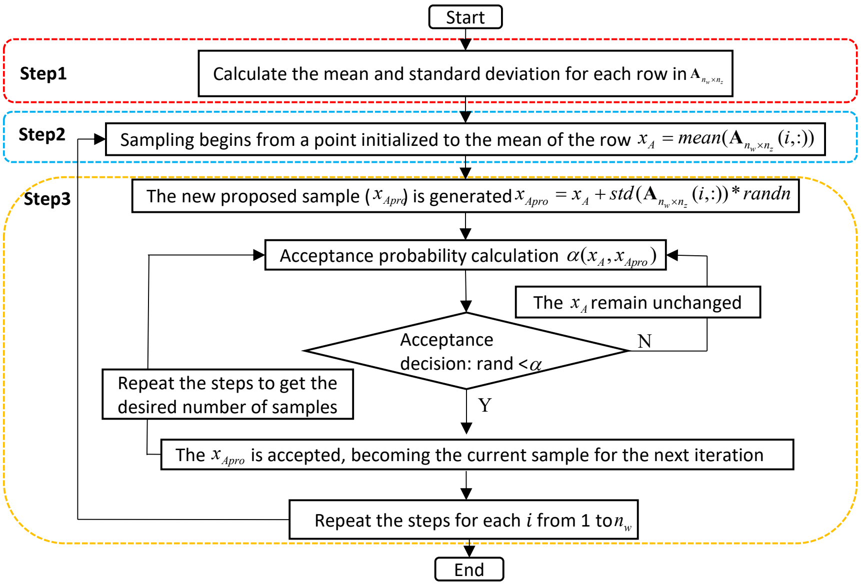


***Figure 1** The flowchart of the Metropolis-Hastings Algorithm.*

4.1.2 The determine of statistical threshold

According to the flowchart shown in Figure 1, we generate multiple sets of samples within the specific frequency band $[w_{k1}, w_{k2}]$ under intact state, each set containing the desired number of samples. Multiple sets of sampled data obtained at a specific frequency can be analyzed to derive the statistical model parameters $\Theta_k^i = \{\sigma_m^{i,k}, \theta_m^{i,k}, \sigma_n^{i,k}, \theta_n^{i,k}, \rho_{mn}^{i,k}\}$, $(k \in [k_1, k_2])$, Where i represents the i-th resampled data group. Subsequently, apply Eqs.(17) and (18) to compute the DIs for each sampled group. Finally, determine the statistical threshold using Monte Carlo discordance testing. The threshold determination method is detailed below:

(i) Firstly, the FFT coefficient matrices ($\mathbf{A}_{n_w \times n_z}$ and $\mathbf{B}_{n_w \times n_z}$) are established under undamaged conditions. Subsequently, resampling is performed using the MCMC method based on the distributions of each row in the matrices.

(ii) Calculating the statistical parameters $\Theta_k^i = \{\sigma_m^{i,k}, \theta_m^{i,k}, \sigma_n^{i,k}, \theta_n^{i,k}, \rho_{mn}^{i,k}\}, (k \in [k_1, k_2])$ from the resampling results, and then use Eqs. (4) and (5) to compute the PDFs of the FFTCRs.

(iii) The square root Jensen-Shannon distances are calculated from the PDFs of the of FFTCRs using Eqs. (13) and (14).

(iv) By using Eqs. (17) and (18), the averaged square root Jensen-Shannon distances are calculated by combining all selected measurements and frequency points, and the calculated values are then stored.

(v) Repeat steps (i) to (iv) across numerous trials to form an array sorted by the magnitude of the highest damage indices. Set the threshold for the 1% discordancy tests by locating the point in the array where damage indices exceed the top 1%.

By following the above steps, one can establish a statistically rigorous threshold that indicates damage when exceeded. This method ensures that the threshold is both precise and reliable, providing a robust indicator of structural damage.

4.2 Robustness validation based on DBSCAN clustering

The DBSCAN technique was first introduced by Ester (1996). DBSCAN clustering has also been widely applied in damage detection in recent years (Li et al., 2020; Wodecki et al., 2021). DBSCAN is a popular clustering algorithm primarily used to identify clusters in a dataset, which are dense groups of points, and to handle noise or outliers effectively. This algorithm is particularly well-suited for dealing with clusters of complex shapes and sizes, and it does not require specifying the number of clusters in advance, setting it apart from many other clustering algorithms such as K-means. DBSCAN's ability to handle complex data without predetermined cluster numbers is particularly advantageous in the context of dealing with uncertainties in structural health monitoring.

To further investigate the impact of statistical parameter uncertainties, induced by environmental noise and modeling errors, on the DI values derived from different structural states. The DBSCAN clustering algorithm is utilized to assess the robustness of the proposed method against uncertainties, based on the number of noise points in the clustering results. Furthermore, the algorithm inherently possesses the advantage of effective discrimination, thus enabling the efficient classification of DIs from various structural states. This contributes to damage detection. In this article, the dataset consisting of DIs, calculated using Eqs. (17) and (18) after MCMC resampling under different structural states, is classified using the DBSCAN algorithm. The implementation of the DBSCAN clustering algorithm in this context is based on two main formulas:

1. ε -neighborhood:

$$N_{\varepsilon}(m)=\{m\in DIs\mid \mathrm{dist}(m,o)\le\varepsilon\} \tag{20}$$

Where $N_{\varepsilon}(m)$ denotes the ε -neighborhood of point m , comprising all points located within an ε distance from m . DIs is the dataset, m and o are points in DIs ,and $\mathrm{dist}(m,o)$ is the distance between m and o .

2. Core Point Condition:

$$\overline{DI}_T=\frac{DI-\theta}{\sigma} \tag{21}$$

The main steps of the DBSCAN clustering algorithm are as follows:

(i) Initialization: Standardize the data DI as follows:

$$\overline{DI}_T=\frac{DI-\theta}{\sigma} \tag{22}$$

(ii) For each unvisited point m :

● Retrieve all points within ε distance, $N_{\varepsilon}(m)$.

● If m is a core point (i.e. $|N_{\varepsilon}(m)|\ge \mathrm{MinPts}$), Start a new cluster or expand an existing one. adding recursively all core points in $N_{\varepsilon}(m)$ and their reachable points to the cluster.

● If m is not a core point, classify m as a border point if applicable; otherwise, classify m as noise.

(iii) Iterate until all points have been processed.

By employing the DBSCAN algorithm, the DIs derived from different structural states are classified. This enables the distinction between DIs from an intact state and those indicative of a potentially damaged state. Furthermore, by applying the 1% exclusive threshold under the intact state, as described in Sect. 4.1.2, to these clustering results enables the identification of structural damage.

5 Procedures of structural damage detection

As described in Section 4, a novel damage detection framework is developed, based on the Fast Fourier Transform Coefficient Ratio (FFTCR) and the square root of the Jensen-Shannon divergence (JSD). Within this unified framework, two procedures are implemented: (i) damage identification through a threshold-based decision strategy, and (ii) robustness verification through DBSCAN clustering and threshold cross-checking. The overall process is summarized in Figure 2 and includes the following steps:

Step (1): Establish the statistical threshold for damage detection

● Extract matrices $\mathbf{A}_{n_w\times n_z}$ and $\mathbf{B}_{n_w\times n_z}$ via FFT of n_z segments of structure response data under intact state.

●Based on the results of MCMC resampling, calculate these statistical parameters $\Theta_k^i=\left\{\sigma_m^{i,k},\theta_m^{i,k},\sigma_n^{i,k},\theta_n^{i,k},\rho_{mn}^{i,k}\right\}$, $(k\in[k_1,k_2])$ for each frequency point within the selected frequency band $[w_1,w_2]$.

● Establish the damage indicator threshold as outlined in Section. 4.1.

Step (2): Damage detection using the threshold value

- Extract matrices $\mathbf{A}_{n_w \times n_z}$ and $\mathbf{B}_{n_w \times n_z}$ through FFT of n_z segments of structure response measurement under different states.
- Calculate the DIs under different states using Eqs. (17) and (18).
- Compare the DIs calculated under different states to the statistical threshold; any values exceeding this threshold indicate damage.

Step (3): Robustness validation via DBSCAN clustering

- Use the MCMC resampling method, statistical parameters Θ_k^i under the various states are determined, and the PDFs for each group of samples are calculated using Eqs. (11) and (12). Subsequently, the DI dataset for each structural states is computed using Eqs. (17) and (18).
- The DBSCAN algorithm is used to classify the DI dataset calculated from various states. The distribution of noise points and the clustering consistency are analyzed to evaluate the robustness of the proposed damage detection method.

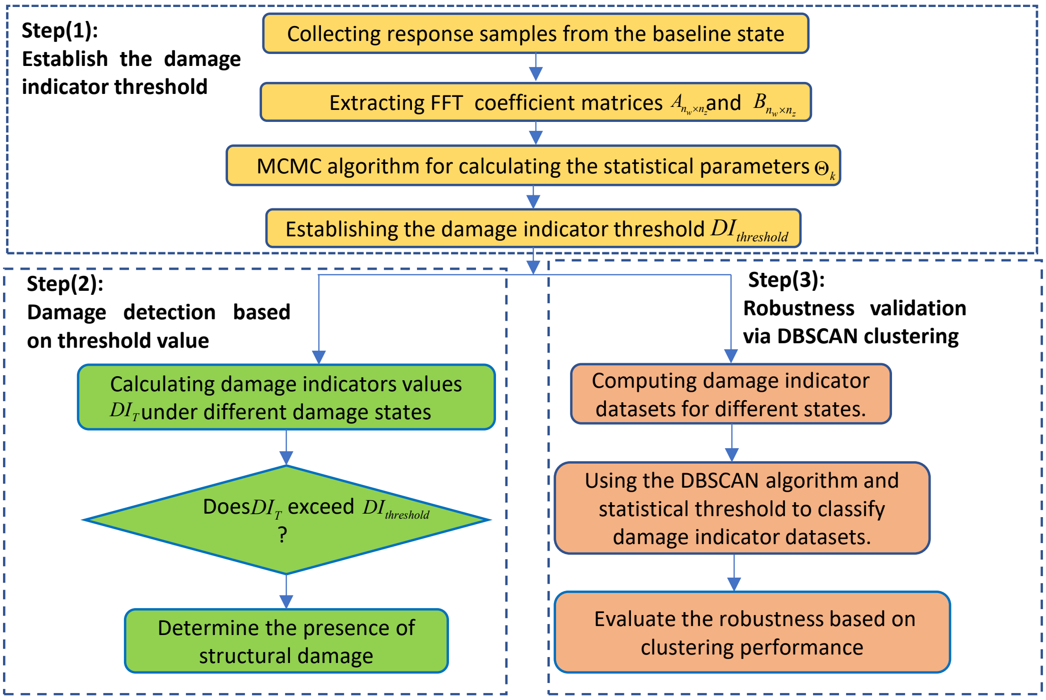


***Figure 2** Flowchart of the proposed damage identification method in this paper.*

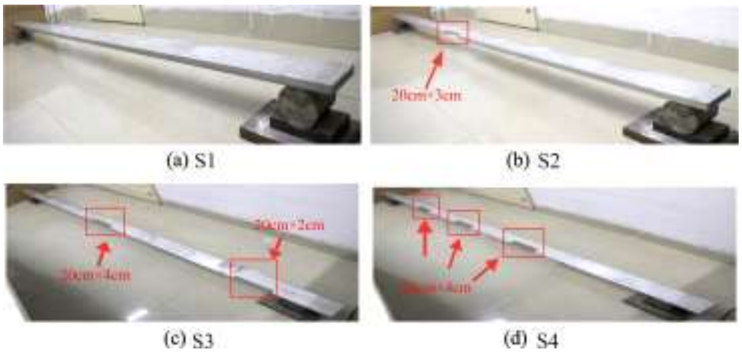
6 Case studies

In this section, we introduce two realistic cases to verify the effectiveness and accuracy of the proposed method. The first case involves four simply supported beams in a laboratory setting, while the second examined a Z24 bridge experiencing progressive damage on site. In both experiments, various damage conditions representative of practical applications were simulated.

6.1 Damage detection of simply supported beams

6.1.1 Damage detection based on threshold value

Four simply supported beams are shown in Figure 3, one of which is undamaged, while the other three are sequentially damaged at one, two and three levels, respectively. Each beam, uniformly fabricated from aluminum within the same production facility, measures 300 cm in length, 10 cm in width, and 3 cm in thickness. Figure 3 illustrates that the number of cuts quantifies the structural damage in this study. The damaged states are designated as S1, S2, S3, and S4. S1 serves as the reference undamaged structure, while S2, S3, and S4 contain one, two, and three cuts, respectively. The damage scenarios are introduced by manually cutting through-thickness cuts into the steel beam specimens, as shown in Figure 3. Although these artificial cut-outs simplify actual bridge damage mechanisms, such as fatigue cracking or corrosion, they offer a controllable and repeatable means of altering the dynamic behavior of the structure. This allows for a clear and consistent evaluation of the proposed method's sensitivity to structural changes under well-defined experimental conditions. Eight sensors were uniformly distributed and installed on each beam. The specifications of the accelerometer are as follows: a sampling frequency of 200 Hz; a sensitivity of 2000 millivolts per gravity (mV/g); and a dynamic range of ±2g. Figure 4 shows the data acquisition system. It should be noted that data collection during this experiment was conducted under ambient vibration. At this sampling frequency, data from the four beams were continuously acquired over 3 hours. The collected data were divided into 400 segments, and an FFT was applied to each segment, yielding 400 pairs of FFT coefficients (each pair containing a real and an imaginary part) for each frequency point. This process produced two matrices: the real matrix $\mathbf{A}_{n_w \times 400}$ and the imaginary matrix $\mathbf{B}_{n_w \times 400}$.



(a) S1 (b) S2

(c) S3 (d) S4

Figure 3 *Four simply supported beams under different damaged states.*

Figure 4 *Data acquisition system.*

In this article, the analysis steps and process of the proposed method are illustrated using the real part-based $t_{13}^{\Re_k}$ and the imaginary part-based $t_{13}^{\Im_k}$. Figures 5-7 compare the PDFs of $t_{13}^{\Re_k}$ and $t_{13}^{\Im_k}$ for the intact state S1 and the damaged scenarios Si (i=2,3,4) at w_k =15 Hz . The corresponding JSD is also shown. Significant changes in the PDFs of $t_{13}^{\Re_k}$ and $t_{13}^{\Im_k}$ are observed between the intact state S1 and the damaged states S2 and S3, as depicted in Figures. 5a, 5c, 6a and 6c. These deviations between the PDFs permit the assessment of structural state. However, Figure. 7a shows that the shift between the PDFs under the damaged state S4 is relatively minor, complicating the assessment of structural damage. This indicates that accurately assessing the structural integrity under all states based on a single frequency point is challenging.

In order to further investigate the impact of different frequency points on PDF of $t_{13}^{\Re_k}$, a comparative graph of multiple frequency points is made. The comparison of PDFs of $t_{13}^{\Re_k}$ before and after damage at different frequencies under different damage states is shown in Figure 8. The solid lines of the same color in the figure denote the undamaged state S1 at the same frequency point, while the three red dashed lines at each frequency indicate the damage state, namely S2, S3, and S4. Comparative analysis reveals significant relative shifts in the PDFs of $t_{13}^{\Re_k}$ at frequencies of 10Hz and 20Hz before and after damage, enabling accurate damage detection in the structure. However, at a frequency of 15Hz under damage state S3, and at 25Hz and 30Hz under damage states S2 and S4, the changes between the PDFs of $t_{13}^{\Re_k}$ are less discernible, posing challenges to damage identification. This observation highlights that reliance on single-frequency analysis may overlook potential structural damage. Therefore, an analytical method integrating multi-frequency points is required to evaluate the structural state, thereby enabling a more reliable damage identification. The method proposed in this paper effectively addresses this challenge by calculating the SR_{JSD} across multiple frequency points. This process involves calculating the area under the JSD curve, as illustrated in Figures 5-7 b and d. Even in scenarios where traditional probabilistic methods are inadequate, the method enables accurate assessments and significantly improves the detection of structural damage

The following sections first introduce the characteristics of SR_{JSD} as it varies with frequency, and then validate the effectiveness and advancement of the DI derived from SR_{JSD} in damage identification.

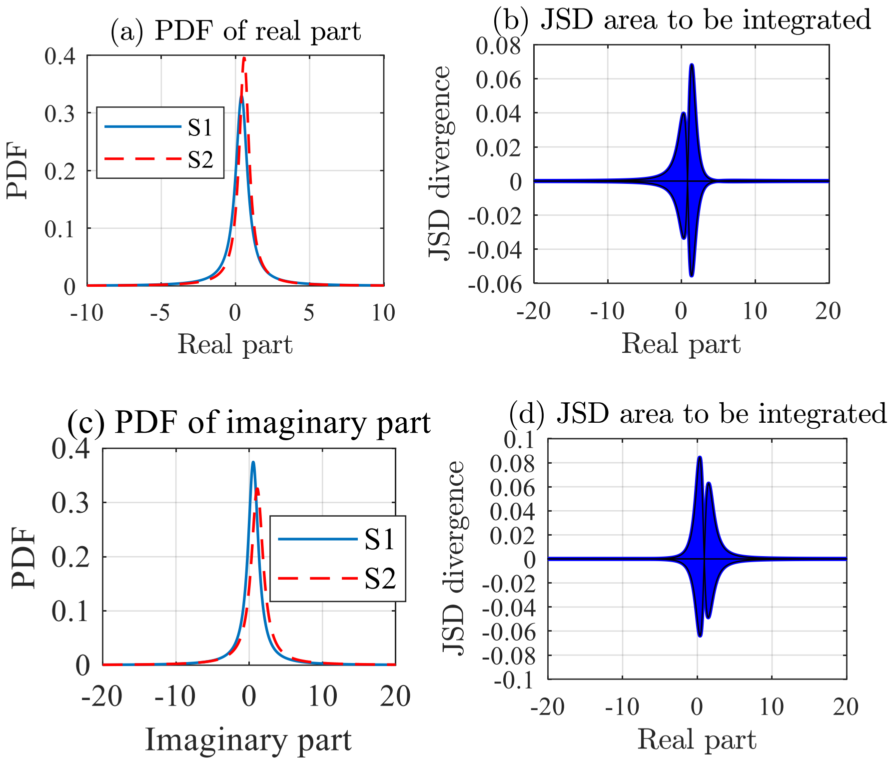


Figure 5 *The PDFs and JSDs for real part-based* $t_{13}^{\Re_{15}}$ *and imaginary part-based* $t_{13}^{\Im_{15}}$ *under two states of S1 and S2.*

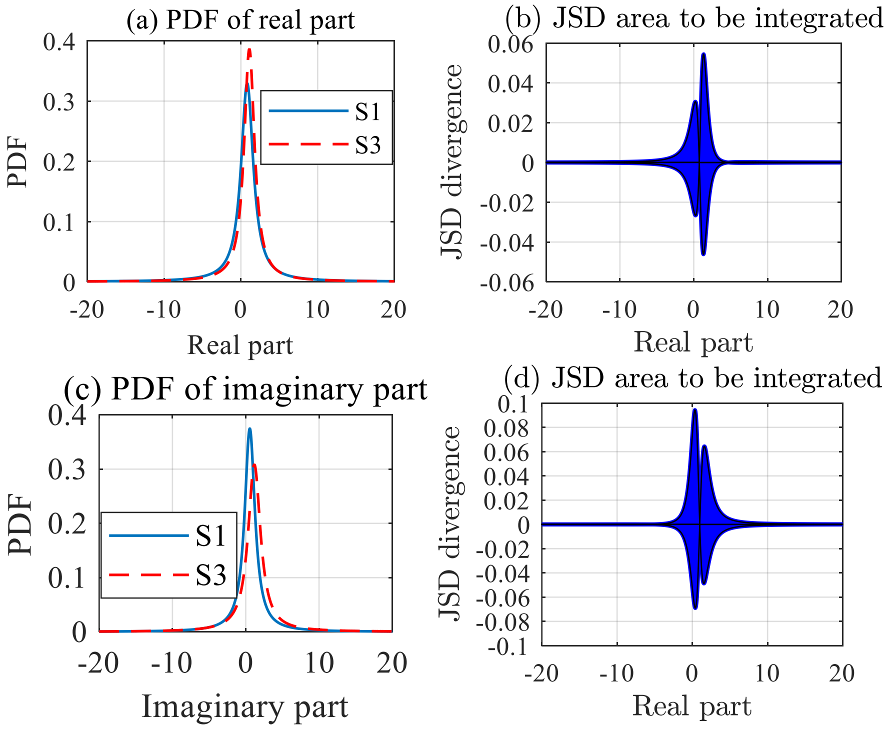


Figure 6 *The PDFs and JSDs for real part-based* $t_{13}^{\Re_{15}}$ *and imaginary part-based* $t_{13}^{\Im_{15}}$ *under two states of S1 and S3.*

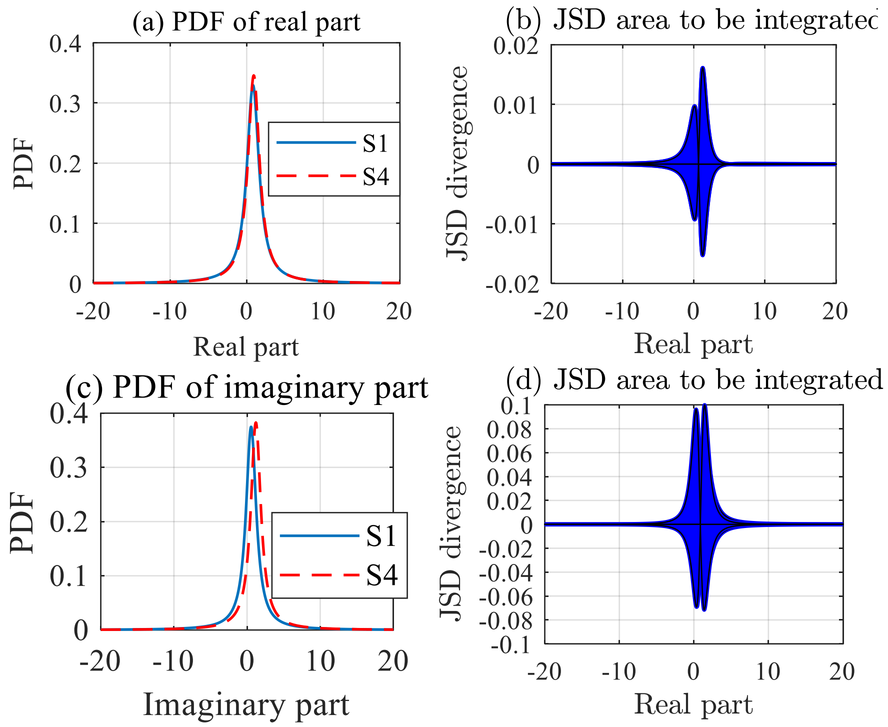


Figure 7 *The PDFs and JSDs for real part-based* $t_{13}^{\Re_{15}}$ *and imaginary part-based* $t_{13}^{\Im_{15}}$ *under two states of S1 and S4.*

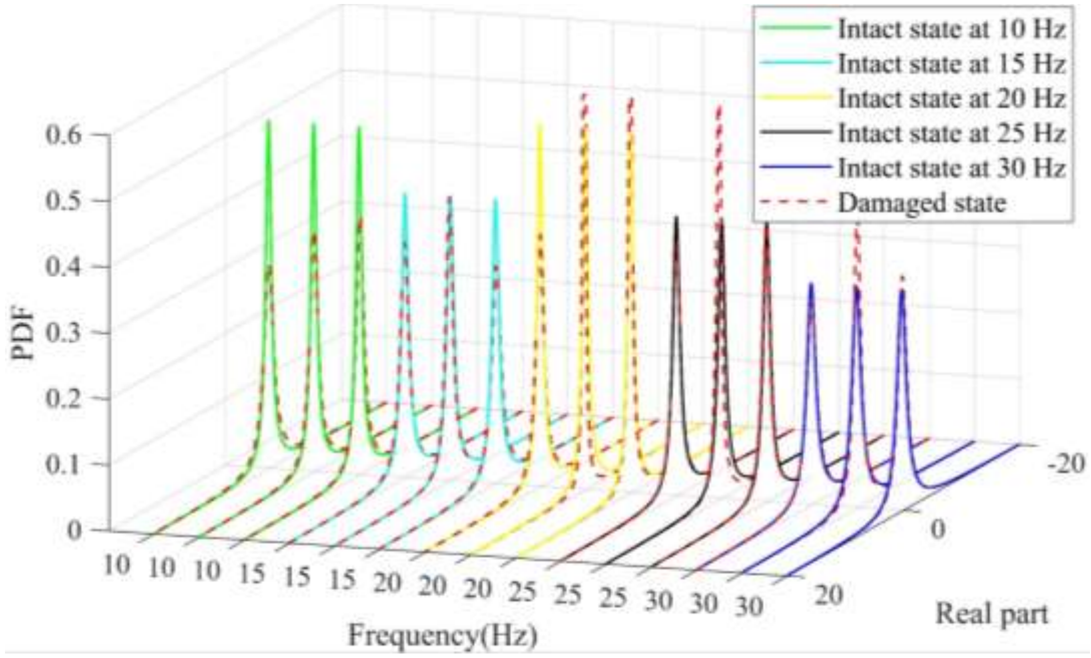
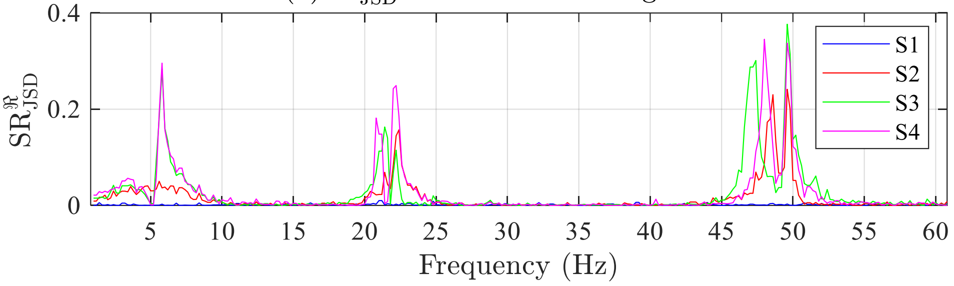


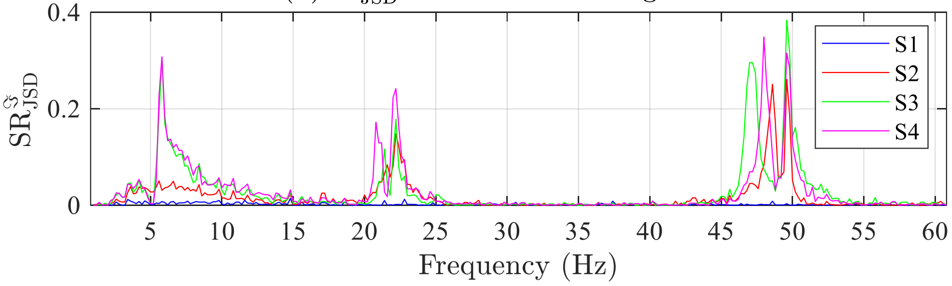
Figure 8 *Comparison of PDFs of* $t_{13}^{\Re_k}$ *before and after damage at different frequencies under different damage states.*

According to Eqs. (13) and (14), the real part-based $SR_{JSD}^{\Re}$ and imaginary part-based $SR_{JSD}^{\Im}$ between the potentially damaged states (S2, S3 and S4) and the healthy state (S1) can be obtained within the selected frequency band, as shown in Figures 9a and b. The establishment of baseline S1 was achieved using two sets of response data in the intact state. Figure 9c shows variations of the PSD across frequencies under various structural states. From these figures, it is evident that the peaks of the SR_{JSD} and the PSD plot occur within comparable frequency ranges, and two important conclusions can be drawn: (1) Variations in the SR_{JSD} across frequencies are minimal under undamaged state, with the degree of change close to zero. However, the changes in the SR_{JSD} increase significantly under potentially damaged states, particularly in the resonance frequency region. This increase is attributable to the enhanced magnitudes of structural responses observed at these frequencies. Therefore the presented method exhibits low sensitivity to the undamaged state but high sensitivity to the damaged state, effectively demonstrating its applicability for structural damage detection. (2) $SR_{JSD}^{\Re_k}$ and $SR_{JSD}^{\Im_k}$ show consistent trends across damaged states, indicating their potential for structural health assessment. These favorable conclusions establish a theoretical foundation for the DI_T discussed in this article regarding its application in damage identification and prove its feasibility for detecting damages.

(a)$\mathrm{SR}_{\mathrm{JSD}}^{\Re}$ for different damage states



(b)$\mathrm{SR}_{\mathrm{JSD}}^{\Im}$ for different damage states



(C)PSD for different damage states

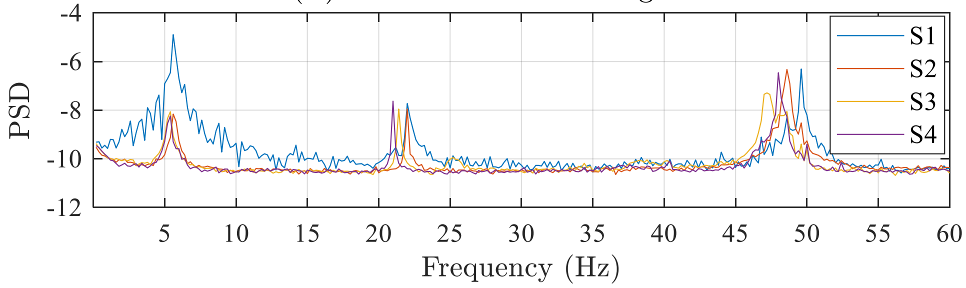


Figure 9 *The variation of* SR_{JSD} *and PSD along with* w_k *in various damage scenarios.*

The establishment of the DI_T proposed in this paper generally involves two main steps: Initially, the first channel is used as the reference degree of freedom to construct the vectors $\mathbf{T}_1^{\Re_k} = \left\{t_{21}^{\Re_k}, t_{31}^{\Re_k}, \ldots, t_{81}^{\Re_k}\right\}_{1\times 7}$ and $\mathbf{T}_1^{\Im_k} = \left\{t_{21}^{\Im_k}, t_{31}^{\Im_k}, \ldots, t_{81}^{\Im_k}\right\}_{1\times 7}$. The chosen frequency range for this analysis is between 1Hz and 60Hz, which encompasses the resonant frequencies most relevant to the structural behaviors being assessed. Subsequently, the averaged damage indicators $DI_T^{\Re}$ and $DI_T^{\Im}$ are calculated according to Eqs. (17) and (18). These damage indicators are vital for assessing the structural integrity and identifying potential damages by quantifying deviations from the baseline conditions. These damage indicators under different states are presented in Figure 10, where Figure 10(a) displays the results based on the real part, and Figure. 10(b) shows the results based on the imaginary part. The red dashed line in the figure indicates the DI threshold, which is determined based on the MCMC resampling method described in Sect. 4.1.2. This threshold is a critical benchmark for distinguishing between normal and abnormal structural states. The blue bars in the figure represent the DI values under various structural damage scenarios. It is evident that all DI values exceed the red threshold line, indicating that states (S2, S3, S4) are damage states. This demonstrates that the method proposed in this paper can accurately detect structural damage across different states and exhibits high robustness. Moreover, its damage identification accuracy surpasses that of the method based solely on the PDF changes before and after damage, as shown in Figure 8.

To further demonstrate the performance of the proposed method in damage detection, this study employs the MSD of the magnitude transfer functions (TFs) for comparison, following the methodology described in Reference (Worden et al., 2000). MSD is a statistical metric utilized to quantify the distance between a point and a distribution, expressed as follows:

$$\mathrm{MSD}_{\zeta} = \left(\{x_{\zeta}\} - \{\bar{\mu}_{\zeta}\}\right)^{T} [\Sigma]^{-1} \left(\{x_{\zeta}\} - \{\bar{\mu}_{\zeta}\}\right) \tag{23}$$

Where $\{x_\zeta\}$ is the potential outlier datum, $\{\bar{\mu}_\zeta\}$ is the mean vector of the sample observations, $[\Sigma]$ is the sample covariance matrix. The determination of threshold value and method for generating the testing dataset for damage detection are both adopted from the literature (Worden et al., 2000). In both the undamaged state (S1) and the damaged states (S2, S3, S4), the magnitudes of the transmissibility functions (TFs) for channels 1 and 3 were replicated 1000 times each within the frequency band [1Hz, 60Hz]. This process resulted in a test dataset comprising 4000 observations. The damage detection results, as illustrated in Figure 11, reveal that despite the post-processing of the raw TFs, the MSD failed to identify most observations as outliers under damage states (S2, S3, S4). In contrast, the method proposed in this study successfully identified structural damage states, as demonstrated in Figure 10, thereby underscoring the superior accuracy of the proposed approach.

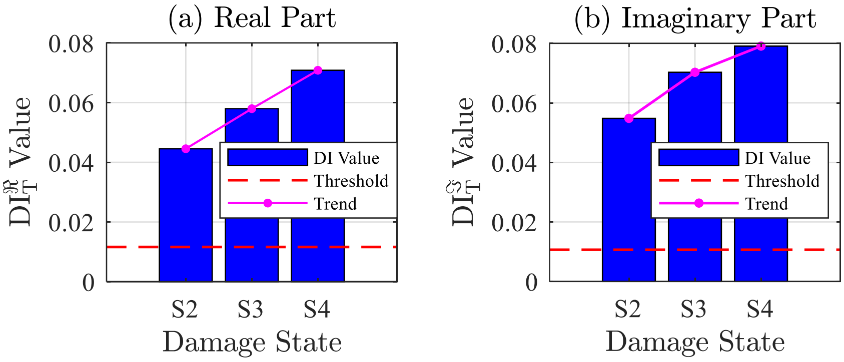


Figure 10 *The real part-based $DI_T^{\Re}$ and imaginary part-based $DI_T^{\Im}$ for S2, S3, and S4.*

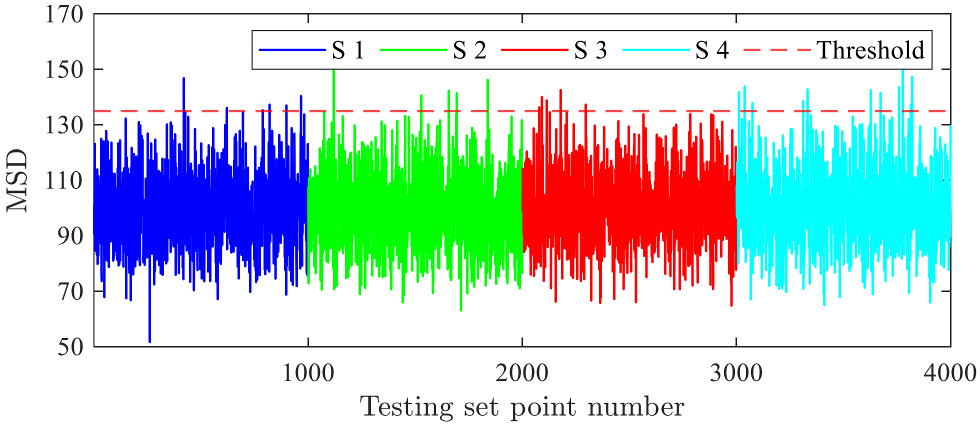


Figure 11 *Outlier detection using MSD under various structural states.*

6.1.2 Research on the Uncertainty of Statistical Parameters

To rigorously test the impact of environmental noise and modeling errors on the proposed approach, the clustering results of the DBSCAN algorithm are used to demonstrate the performance of the proposed method in adapting to uncertainties associated with statistical parameters Θ_k . DI samples are calculated using the MCMC resampling scheme under various damage states, as detailed in Sect. 4.1.1. For each state, 120 DI samples are calculated, resulting in a total of 480 DI samples across the four states. These samples are then clustered using the DBSCAN algorithm, as outlined in Sect. 4.2. The parameters for the DBSCAN algorithm are set as follows: ε = 0.15 and MinPts = 4. The clustering results, displayed in Figure. 12, illustrate how the DBSCAN algorithm segregates the calculated DI values under different structural states into distinct clusters, each represented by a unique color corresponding to a damaged state. Figure. 12a shows the clustering results of the DIs based on the real part, while Figure. 12b shows the clustering results of the DIs based on the imaginary part. By establishing a 1% exclusive threshold for DIs under the intact state (S1), it is observed that the DI values for the other states (S2, S3, S4) exceed this threshold. Therefore, states S2, S3, and S4 can be identified as damage states. Although a few noise points are present, they do not compromise the overall classification of structural states. These results confirm that the proposed method exhibits high noise resistance and robustness, ensuring the reliability of the structural damage identification.

To evaluate the classification performance of DBSCAN, a confusion matrix is employed. The confusion matrix is a machine-learning visualization tool that uses columns to represent expected class outcomes and rows to show actual class results. The study employs the False Positive Rate (FPR) and False Negative Rate (FNR) to quantitatively assess the performance of the proposed methods:

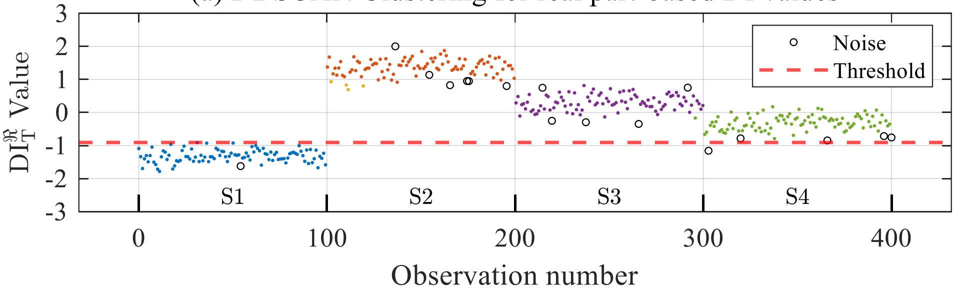
$$FPR = \frac{FP}{TN + FP} \times 100\% \tag{24}$$

$$FNR = \frac{FN}{FN + TP} \times 100\% \tag{25}$$

Where True Positives (TP) represent numbers correctly identified as 'damaged,' and False Positives (FP) represent numbers incorrectly identified as such. True Negatives (TN) denote numbers correctly identified as 'intact,' and False Negatives (FN) denote numbers incorrectly identified as such.

The confusion matrices of the DBSCAN clustering for DI values are shown in Figure 13. In Figure 13a, the confusion matrix reveals a FPR of 2.00% and a FNR of 1.67%, with 295 true positives and 98 true negatives out of total samples. Conversely, Figure. 13b exhibits a FPR of 3.00% and a FNR of 1.00%, identifying 297 true positives and 97 true negatives. Matrix (b) shows a higher accuracy in classifying the damaged state (99%) compared to Matrix (a) (98.33%), while Matrix (a) exhibits a slightly higher accuracy in classifying the intact state (98%) compared to Matrix (b) (97%). Both matrices have a small number of misclassifications, indicating a high level of accuracy in the classification of 'intact' versus 'damaged' categories. It has been demonstrated that the distribution of DI values remains relatively concentrated across various individual structural states. This consistency is maintained even in the presence of environmental noise, indicating that external factors do not significantly influence the dispersion of these values. Such robustness further underscores the effectiveness of DI values as reliable indicators for classifying damage states.

(a) DBSCAN Clustering for real part-based DI values



(b) DBSCAN Clustering for imaginary part-based DI values

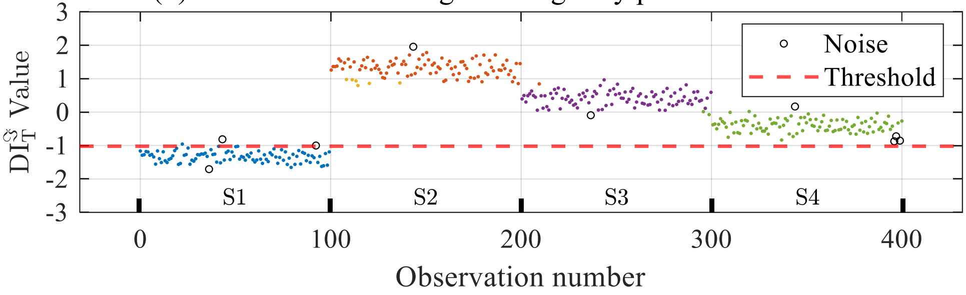


***Figure 12** The clustering results for the real and imaginary parts of DI values.*

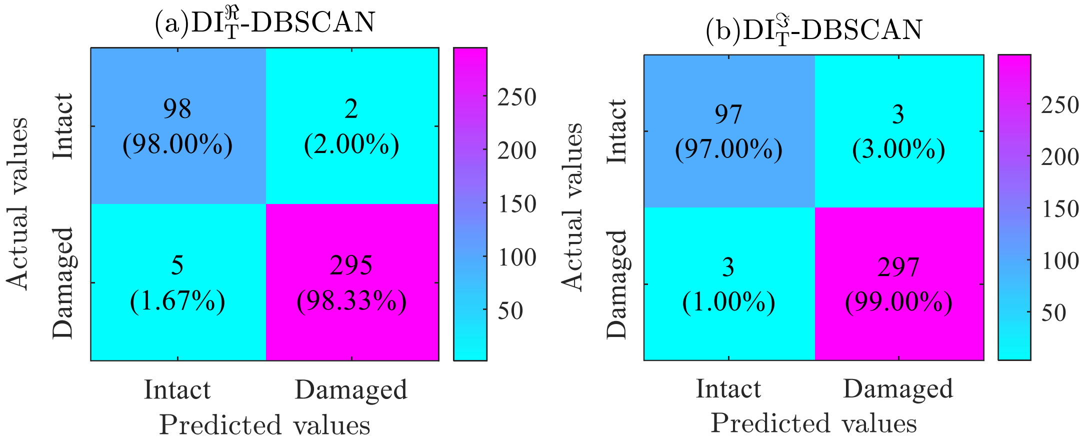


***Figure 13** The confusion matrices of the DBSCAN clustering for DI values.*

6.2 Structural damage detection for the Z24 bridge

In this section, the efficacy and accuracy of the proposed damage detection method are further validated using the famous Z24 bridge (Peeters and De Roeck, 2001; Reynders and De Roeck, 2010), located in the Canton of Bern, Switzerland, which connects Koppigen and Utzendorf. This bridge, a three-span prestressed concrete structure, features a main span of 30 meters flanked by two 14-meter side spans. The geometric layout is illustrated in Figure 14, and Figure 15 displays the bridge's cross-section and the arrangement of sensors used for the vibration test. A subset of sensors numbered between 199 and 243 from those installed on the bridge's surface, was selected for this case study. The complete sensor layout diagram is provided in Ref. (Maeck and De Roeck, 2003a). In this study, vertical acceleration response data, recorded from progressive damage tests conducted from 4 August to 9 September 1998, are utilized for damage detection. These tests were performed continuously just before the demolition of the bridge. An overview of all progressive damage tests is presented in Table 1. It should be noted that the damage scenarios in the Z24 bridge dataset were intentionally introduced during the decommissioning phase to emulate real-world degradation mechanisms. These scenarios include the removal of structural components (e.g., diaphragms, bearings), as well as the weakening of joints and supports, which result in significant alterations to the load path and boundary conditions. While these damages are artificially imposed, they represent plausible deterioration patterns observed in operational bridges, such as localized stiffness loss and fatigue-induced weakening. According to Peeters and De Roeck (2001), the severity of these damage stages ranges from minor to severe, thereby offering a broad spectrum of structural conditions suitable for evaluating the performance of damage detection methods.

Unlike Case 1, which utilizes acceleration response data from ambient operational vibration tests on four simply supported beams, forced vibration testing is employed for damage detection in the Z24 bridge case. This choice allows validation of the proposed method's effectiveness and applicability under ambient operational and forced vibration conditions. Forced excitation of the bridge deck was achieved using two vertical shakers.

Figure 14 *A full picture of the Z24 bridge, as well as a closer look at the bridge deck and one of the piers (Maeck and De Roeck, 2003a).*

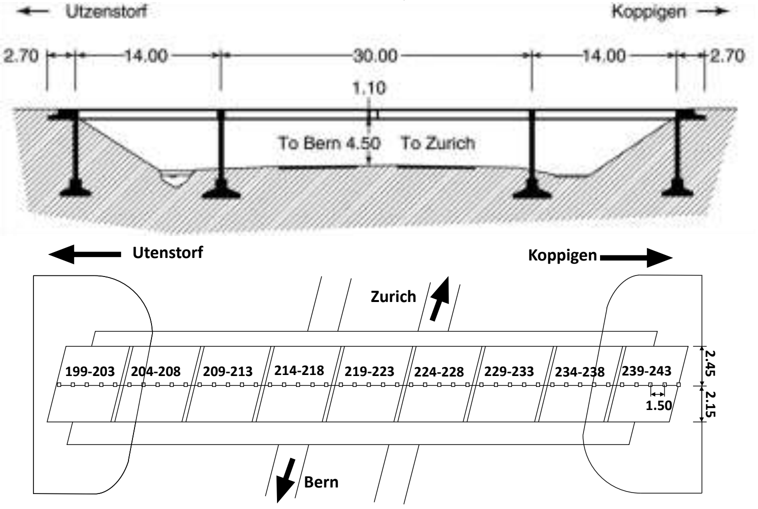


Figure 15 *Cross section of the bridge and layout diagram of sensors on the bridge.*

Table 1 Progressive damage scenarios of Z24 bridge (Maeck and De Roeck, 2003b).

No	Scenario	Description/simulation of real damage cause
State 1	First reference measurement	• Healthy structure
State 2	Second reference measurement	• After installation of lowering System
State 3	Lowering of pier, 20 mm	• Settlement of subsoil, erosion

State 4	Lowering of pier, 40 mm	
State 5	Lowering of pier, 80 mm	
State 6	Lowering of pier, 95 mm	
State 7	Tilt of foundation	• Settlement of subsoil, erosion
State 8	Third reference measurement	• After lifting of the bridge to its initial position
State 9	Spalling of concrete, 24m2	• Vehicle impact, carbonization, and subsequent corrosion of reinforcement
State 10	Spalling of concrete, 12m2	
State 11	Landslide at abutment	• Heavy rainfall, erosion
State 12	Failure of concrete hinge	• Chloride attack, corrosion
State 13	Failure of anchor heads I	• Corrosion, overstress
State 14	Failure of anchor heads II	
State 15	Rupture of tendons I	• Erroneous or forgotten injection of tendon tubes, chloride influence
State 16	Rupture of tendons II	
State 17	Rupture of tendons III	

For each applied damage scenario, 65536 samples per channel were collected at a sampling rate of 100 Hz. To facilitate analysis, these 65,536 sample points were divided into 50 segments, with each segment encompassing approximately 13.1 seconds of sampling time. An FFT was applied to each segment, yielding 50 pairs of FFT coefficients (each pair containing a real and an imaginary part) for each frequency point. This resulted in two matrices: the real matrix $\mathbf{A}_{n_w \times 50}$ and the imaginary matrix $\mathbf{B}_{n_w \times 50}$. In this case study, the frequency range chosen for detecting structural anomalies spanned from 0.1 Hz to 45 Hz. Using data from the First reference measurement(S1), the damage indicator threshold was established following the method described in Sect. 4.1.2. This threshold, derived from the baseline state S1, was compared with the DI values from other states (S2-S17) to detect anomalies. Exceeding this threshold indicates potential structural damage. The averaged DI values for the damage scenarios (S2-S17) listed in Table 1 are displayed in Figure. 16. Figure 16a presents the DI values based on the real part, while Figure 16b shows the DI values based on the imaginary part. These figures clearly show that the DI values for states S2 to S17, whether based on the real or imaginary part, exceed the threshold established from the intact State (S1). Therefore, the damage states S2 to S17 can be accurately identified.

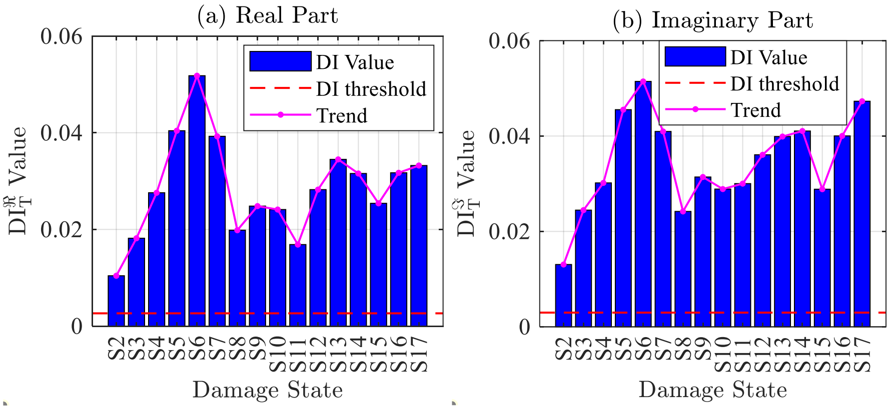
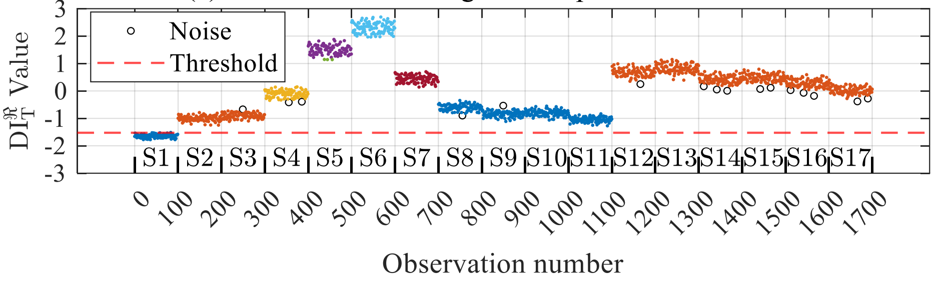


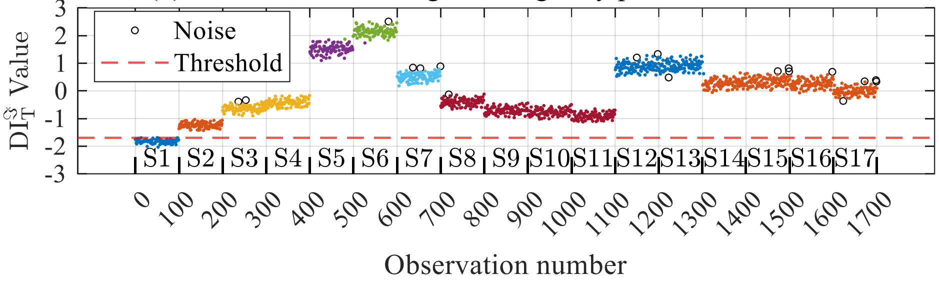
Figure 16 *The real part-based* $DI_T^{\Re}$ *and imaginary part-based* $DI_T^{\Im}$ *for S2-S17.*

Then, 100 DI samples for each scenario are calculated using MCMC to assess the proposed method's capacity to accommodate the uncertainties associated with statistical parameters caused by environmental noise. The parameters for the DBSCAN algorithm are set as follows: ε = 0.15 and MinPts = 4. This clustering technique analyzes 1700 samples across 17 different states. A 1% exclusive threshold of DIs, derived from the intact state (S1), is applied to the clustering analysis results, as illustrated in Figure 17. It is evident that the DBSCAN technique segregates the samples into distinct clusters (different colors indicate different clusters), with each cluster corresponding to its specific structural state. Regardless of whether the DI values are derived from the real or imaginary parts, those for clusters corresponding to structural states S2 to S17 surpass the threshold established under the intact state, indicating that these states S2 to S17 are potential damage states, and these damage detection result is the same as shown in Figure 16. Additionally, the clustering results reveal that the number of noise points within the results is minimal, which does not significantly influence the clustering groups. This effectively demonstrates the proposed method's robustness against the uncertainties associated with statistical parameters.

Figure 18 displays two confusion matrices from the DBSCAN clustering analysis for DI values. Each matrix compares actual values (rows) against predicted values (columns) to visualize the classification results. Both matrices demonstrate similar high performance, with a correct classification rate of 100% for the damaged state (1600 samples correctly classified as damaged in each matrix). The classification accuracy for the intact state is slightly lower at 98% (98 samples correctly classified as intact), with a misclassification rate of 2% (2 samples incorrectly classified as damaged). There are no false negatives in either matrix, but a small number of false positives are present. These high classification rates reflect the robustness and reliability of the DI measures. It has been shown that the distribution of DI values is relatively uniform across different structural states and remains stable despite the presence of environmental noise. This uniformity and stability emphasize the effectiveness of DI values in classifying damage states.



(a) DBSCAN Clustering for real part-based DI values



(b) DBSCAN Clustering for imaginary part-based DI values

Figure 17 *The clustering results for the real and imaginary parts of DI values.*

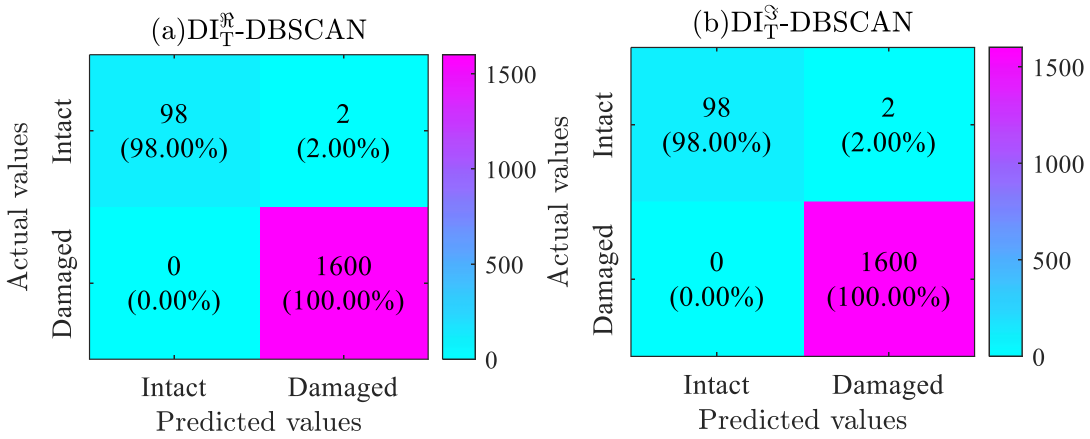


Figure 18 *The confusion matrices of the DBSCAN clustering for DI values.*

7 CONCLUSIONS

In this paper, a novel damage detection method based on the probabilistic divergence of FFT coefficient ratios (FFT-CRs) is proposed. The method begins by computing FFTCRs from the real and imaginary parts of multi-segment frequency response data measured between arbitrary pairs of points. To effectively quantify and characterize the uncertainty of statistical parameters, the PDF of these FFTCRs is introduced. The square root distance of the JSD is utilized to measure the similarity between two probability distributions, namely the intact state and the possibly damaged state. Subsequently, a new damage index is derived by averaging the JSD square root distances across different measurement points and selected frequency points. The establishment of a threshold provides a basis for assessing structural damage. Finally, the robustness of the proposed method is effectively evaluated using DBSCAN clustering for the classification of DIs under different states.

The proposed method is validated through two experiments: one involving ambient vibration tests on four simply supported beams in a laboratory, and the other a field test on the Z24 Bridge under forced vibration. It can be concluded that the sensitivity of the PDF of the FFT coefficient ratio to damage varies at different frequency points. Thus, determining sensitive frequency points prior to structural damage assessment is crucial. The methodology proposed in this paper integrates multiple frequency points, obviating the need for selecting specific frequencies before assessing structural damage. This highlights the practicality and advanced nature of the approach. Research has also shown that this method has higher accuracy compared to traditional methods that solely rely on deviations in the PDFs before and after damage to determine structural damage. The DBSCAN algorithm is used to evaluate the robustness of the proposed method against uncertainty. The results show that DBSCAN accurately detects damaged states, thereby demonstrating the stability and dependability of the proposed method. Additionally, applying a 1% exclusive threshold for DIs obtained from the intact state to the clustering results can also accurately identify structural damage states.

In summary, this paper introduces a novel approach to damage detection using an innovative algorithm that utilizes statistical and data-driven techniques that rely solely on response data. This method analyzes raw data directly without any postprocessing to identify potential damage, and it is capable of adapting to various systems and environments by learning from actual data instead of depending on potentially inaccurate models. However, the success of the algorithm critically depends on the quality of the data and computational capacity. Another challenge is the quantification of structural damage and the localization of localized damages, which require further in-depth research in the future.

ACKNOWLEDGMENT

The authors highly appreciate the KU Leuven Structural Mechanics Section for sharing the field test data of Z24 Bridge.

Author's Contribuitions: Conceptualization, Methodology, and Original Draft Writing: Guoqing Li; Supervision and Editing: Hao Wei.

Editor: Rogério José Marczak

References

Amarbayasgalan, T., Jargalsaikhan, B., Ryu, K. H. (2018). Unsupervised Novelty Detection Using Deep Autoencoders with Density Based Clustering. Applied Sciences-Basel, v. 8, n. 9, p. 18. doi: 10.3390/app8091468

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. Mechanical Systems and Signal Processing, v. 147, n. p. 45. doi: 10.1016/j.ymssp.2020.107077

Bandara, R. P., Chan, T. H. T., Thambiratnam, D. P. (2014). Frequency response function based damage identification using principal component analysis and pattern recognition technique. Engineering Structures, v. 66, n. p. 116-128. doi: 10.1016/j.engstruct.2014.01.044

Capecchi, D., Ciambella, J., Pau, A., Vestroni, F. (2016). Damage identification in a parabolic arch by means of natural frequencies, modal shapes and curvatures. Meccanica, v. 51, n. 11, p. 2847-2859. doi: 10.1007/s11012-016-0510-3

Chesné, S., Deraemaeker, A. (2013). Damage localization using transmissibility functions: A critical review. Mechanical Systems and Signal Processing, v. 38, n. 2, p. 569-584. doi: 10.1016/j.ymssp.2013.01.020

Civera, M., Sibille, L., Fragonara, L. Z., Ceravolo, R. (2023). A DBSCAN-based automated operational modal analysis algorithm for bridge monitoring. Measurement, v. 208, n. p. 15. doi: 10.1016/j.measurement.2023.112451

Devriendt, C., Guillaume, P. (2007). The use of transmissibility measurements in output-only modal analysis. Mechanical Systems and Signal Processing, v. 21, n. 7, p. 2689-2696. doi: 10.1016/j.ymssp.2007.02.008

Ester, M. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Proc.int.conf.knowledg Discovery & Data Mining, v. n. p. doi:

He, S., Ng, C. T. (2017). Guided wave-based identification of multiple cracks in beams using a Bayesian approach. Mechanical Systems and Signal Processing, v. 84, n. p. 324-345. doi: 10.1016/j.ymssp.2016.07.013

Hinckley, D. (1969). On the ratio of two correlated normal random variables. Biometrika, v. 56, n. 3, p. 635-639. doi:

Hou, R. R., Xia, Y. (2021). Review on the new development of vibration-based damage identification for civil engineering structures: 2010-2019. Journal of Sound and Vibration, v. 491, n. p. 32. doi: 10.1016/j.jsv.2020.115741

Hu, J. Y., Zhang, S. S., Chen, E., Li, W. (2022). A review on corrosion detection and protection of existing reinforced concrete (RC) structures. Construction and Building Materials, v. 325, n. p. doi: 10.1016/j.conbuildmat.2022.126718

Kullback, S., Leibler, R. A. (1951). On information and sufficiency. Annals of Mathematical Statistics, v. 22, n. 1, p. 79-86. doi: 10.1214/aoms/1177729694

Li, S. P., Qin, N., Huang, D. R., Huang, D. Q., Ke, L. Y. (2020). Damage Localization of Stacker's Track Based on EEMD-EMD and DBSCAN Cluster Algorithms. Ieee Transactions on Instrumentation and Measurement, v. 69, n. 5, p. 1981-1992. doi: 10.1109/tim.2019.2919375

Lin, J. H. (1991). DIVERGENCE MEASURES BASED ON THE SHANNON ENTROPY. Ieee Transactions on Information Theory, v. 37, n. 1, p. 145-151. doi: 10.1109/18.61115

Liu, X., Lieven, N. A. J., Escamilla-Ambrosio, P. J. (2009). Frequency response function shape-based methods for structural damage localisation. Mechanical Systems and Signal Processing, v. 23, n. 4, p. 1243-1259. doi: 10.1016/j.ymssp.2008.10.002

Maeck, J., De Roeck, G. (2003a). Damage assessment using vibration analysis on the Z24-bridge. Mechanical Systems and Signal Processing, v. 17, n. 1, p. 133-142. doi: 10.1006/mssp.2002.1550

Maeck, J., De Roeck, G. (2003b). Description of Z24 benchmark. Mechanical Systems and Signal Processing, v. 17, n. 1, p. 127-131. doi: 10.1006/mssp.2002.1548

Maia, N. M. M., Almeida, R. A. B., Urgueira, A. P. V., Sampaio, R. P. C. (2011). Damage detection and quantification using transmissibility. Mechanical Systems and Signal Processing, v. 25, n. 7, p. 2475-2483. doi: 10.1016/j.ymssp.2011.04.002

Manson, G., Worden, K., Allman, D. (2003). Experimental validation of a structural health monitoring methodology. Part II. Novelty detection on a Gnat aircraft. Journal of Sound and Vibration, v. 259, n. 2, p. 345-363. doi: 10.1006/jsvi.2002.5167

Mao, Z., Todd, M. (2012). A model for quantifying uncertainty in the estimation of noise-contaminated measurements of transmissibility. Mechanical Systems and Signal Processing, v. 28, n. p. 470-481. doi: 10.1016/j.ymssp.2011.10.002

Mao, Z., Todd, M. (2013). Statistical modeling of frequency response function estimation for uncertainty quantification. Mechanical Systems and Signal Processing, v. 38, n. 2, p. 333-345. doi: 10.1016/j.ymssp.2013.01.021

Mei, L. F., Yan, W. J., Yuen, K. V., Beer, M. (2022). Structural novelty detection with Laplace asymptotic expansion of the Bhattacharyya distance of transmissibility and Bayesian resampling scheme. Journal of Sound and Vibration, v. 540, n. p. 19. doi: 10.1016/j.jsv.2022.117277

Mei, L. F., Yan, W. J., Yuen, K. V., Ren, W. X., Beer, M. (2023). Transmissibility-based damage detection with hierarchical clustering enhanced by multivariate probabilistic distance accommodating uncertainty and correlation. Mechanical Systems and Signal Processing, v. 203, n. p. 23. doi: 10.1016/j.ymssp.2023.110702

Peeters, B., De Roeck, G. (2001). One-year monitoring of the Z24-Bridge: environmental effects versus damage events. Earthquake Engineering & Structural Dynamics, v. 30, n. 2, p. 149-171. doi: 10.1002/1096-9845(200102)30:2<149::Aid-eqe1>3.0.Co;2-z

Reynders, E., De Roeck, G. (2010). A local flexibility method for vibration-based damage localization and quantification. Journal of Sound and Vibration, v. 329, n. 12, p. 2367-2383. doi: 10.1016/j.jsv.2009.04.026

Schallhorn, C., Rahmatalla, S. (2015). Crack detection and health monitoring of highway steel-girder bridges. Structural Health Monitoring-an International Journal, v. 14, n. 3, p. 281-299. doi: 10.1177/1475921714568404

Soofi, Y. J., Bitaraf, M. (2022). Output-only entropy-based damage detection using transmissibility function. Journal of Civil Structural Health Monitoring, v. 12, n. 1, p. 191-205. doi: 10.1007/s13349-021-00535-3

Wodecki, J., Michalak, A., Zimroz, R. (2021). Local damage detection based on vibration data analysis in the presence of Gaussian and heavy-tailed impulsive noise. Measurement, v. 169, n. p. doi: 10.1016/j.measurement.2020.108400

Worden, K., Manson, G. (2003). Experimental validation of a structural health monitoring methodology: Part I. Novelty detection on a laboratory structure. Journal of Sound and Vibration, v. 259, n. 2, p. 323-343. doi: 10.1006/jsvi.2002.5168

Worden, K., Manson, G. (2007). The application of machine learning to structural health monitoring. Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, v. 365, n. 1851, p. 515-537. doi: 10.1098/rsta.2006.1938

Worden, K., Manson, G., Fieller, N. R. J. (2000). Damage detection using outlier analysis. Journal of Sound and Vibration, v. 229, n. 3, p. 647-667. doi: 10.1006/jsvi.1999.2514

Wu, D. F., Zeng, M. Y., Zhao, H. D., Wang, Y., Du, Y. C. (2021). Detection and localization of debonding beneath concrete pavement using transmissibility function analysis. Mechanical Systems and Signal Processing, v. 159, n. p. 18. doi: 10.1016/j.ymssp.2021.107802

Yan, W. J., Chronopoulos, D., Papadimitriou, C., Cantero-Chinchilla, S., Zhu, G. S. (2020). Bayesian inference for damage identification based on analytical probabilistic model of scattering coefficient estimators and ultrafast wave scattering simulation scheme. Journal of Sound and Vibration, v. 468, n. p. 23. doi: 10.1016/j.jsv.2019.115083

Yan, W. J., Chronopoulos, D., Yuen, K. V., Zhu, Y. C. (2022). Structural anomaly detection based on probabilistic distance measures of transmissibility function and statistical threshold selection scheme. Mechanical Systems and Signal Processing, v. 162, n. p. 22. doi: 10.1016/j.ymssp.2021.108009

Yan, W. J., Zhao, M. Y., Sun, Q., Ren, W. X. (2019). Transmissibility-based system identification for structural health Monitoring: Fundamentals, approaches, and applications. Mechanical Systems and Signal Processing, v. 117, n. p. 453-482. doi: 10.1016/j.ymssp.2018.06.053

Yang, Y. B., Yang, J. P. (2018). State-of-the-Art Review on Modal Identification and Damage Detection of Bridges by Moving Test Vehicles. International Journal of Structural Stability and Dynamics, v. 18, n. 2, p. doi: 10.1142/s0219455418500256

Yu, J. H., Kang, J. H. (2023). Clustering ensemble-based novelty score for outlier detection. Engineering Applications of Artificial Intelligence, v. 121, n. p. 21. doi: 10.1016/j.engappai.2023.106164

Yuen, K.-V. (2010). Bayesian methods for structural dynamics and civil engineering: John Wiley & Sons.

Zhang, H., Schulz, M. J., Ferguson, F., Pai, P. F. (1999). Structural health monitoring using transmittance functions. Mechanical Systems and Signal Processing, v. 13, n. 5, p. 765-787. doi: 10.1006/mssp.1999.1228

Zhang, L., Lang, Z. Q., Papaelias, M. (2016). Generalized Transmissibility Damage Indicator With Application to Wind Turbine Component Condition Monitoring. Ieee Transactions on Industrial Electronics, v. 63, n. 10, p. 6347-6359. doi: 10.1109/tie.2016.2580519

Zhang, X. X., Delpha, C., Diallo, D. (2020). Incipient fault detection and estimation based on Jensen-Shannon divergence in a data-driven approach. Signal Processing, v. 169, n. p. 12. doi: 10.1016/j.sigpro.2019.107410

Zhu, D. P., Yi, X. H., Wang, Y., Lee, K. M., Guo, J. J. (2010). A mobile sensing system for structural health monitoring: design and validation. Smart Materials and Structures, v. 19, n. 5, p. 11. doi: 10.1088/0964-1726/19/5/055011