

Research Article

Damage Detection in Bridge Structures through Compressed Sensing of Crowdsourced Smartphone Data

Mohammad Talebi-Kalaleh  and Qipei Mei 

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada

Correspondence should be addressed to Qipei Mei; qipei@ualberta.ca

Received 20 November 2023; Revised 16 March 2024; Accepted 28 March 2024; Published 4 April 2024

Academic Editor: Zoran Rakicevic

Copyright © 2024 Mohammad Talebi-Kalaleh and Qipei Mei. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traditional bridge health monitoring methods that necessitate sensor installation are not only costly but also time-consuming. In contrast, utilizing smartphone data collected from vehicles as they traverse bridges offers an efficient and cost-effective alternative. This paper introduces a cutting-edge damage detection framework for indirect monitoring of bridge structures, leveraging a substantial volume of acceleration data collected from smartphones in vehicles passing over the bridge. Our innovative approach addresses the challenge of collecting and transmitting high-frequency data while preserving smartphone battery life and data plans through the integration of compressed sensing (CS) into the crowdsensing-based monitoring framework. CS employs random sampling and signal recovery from a significantly reduced number of samples compared to the requirements of the Nyquist-Shannon sampling theorem. In the proposed framework, acceleration signals from vehicles are initially acquired using smartphone sensors, undergo compression, and are then transmitted for signal reconstruction. Subsequently, feature extraction and dimensionality reduction are performed using Mel-frequency cepstral coefficients and principal component analysis. Damage indexes are computed based on the dissimilarity between probability distribution functions utilizing the Wasserstein distance metric. The efficacy of the proposed methodology in bridge monitoring has been substantiated through the utilization of numerical models and a lab-scale bridge. Furthermore, the feasibility of implementing the framework in a real-world application has been investigated, leveraging the smartphone data from 102 vehicle trips on the Golden Gate Bridge. The results demonstrate that damage detection using the reconstructed signals obtained through compressed sensing achieves comparable performance to that obtained with the original data sampled at the Nyquist measurement sampling rate. However, it is observed that to retain severity information within the signals for accurate damage severity identification, the compression level should be limited to 20%. These findings affirm that compressed sensing significantly reduces the data collection requirements for crowdsensing-based monitoring applications, without compromising the accuracy of damage detection while preserving essential damage-sensitive information within the dataset.

1. Introduction

The implementation of intelligent solutions for condition monitoring of infrastructure represents a pivotal stride toward the realization of smart cities. Regular monitoring of structures not only prevents potential loss of lives and wealth due to sudden collapses but also enables proactive maintenance scheduling, leading to reduced maintenance and life-cycle costs. Bridges, as a critical component of the public transportation system, are deteriorating due to a variety of factors. This type of degradation significantly impairs the

performance and service life of bridges [1–3]. As a result, it is vital to create effective and efficient ways for the early detection of deterioration and damage [2].

The deteriorating state of infrastructure in the United States and Canada, particularly with regard to bridges constructed during the 1950s and 1960s, has raised significant concerns [4]. Presently, nearly 9% of bridges in the United States are classified as structurally deficient, as reported by the Federal Highway Administration (FHWA) [5]. Similarly, the Canadian Infrastructural Report Card for 2019 estimates that around 40% of Canadian bridges fall into

the “fair,” “poor,” or “very poor” condition categories [6]. This situation poses an elevated risk of potential failure for these bridges, highlighting the crucial need for vigilant monitoring and maintenance practices to ensure their structural stability and long-term performance.

Traditional bridge health monitoring techniques typically involve the installation of sensors directly on the structure, with subsequent analysis of the collected data to detect damage [7–11]. Recent research on fixed-sensor-based damage detection has emphasized Bayesian methods. For example, Zhang et al. [12] proposed a novel damage detection method based on a fundamental Bayesian two-stage model and sparse regularization, demonstrating improved performance due to the consideration of uncertainty. Wang et al. [13] introduced a probabilistic data-driven damage detection method using Sparse Bayesian Learning (SBL) and validated its capabilities with field monitoring data from a cable-stayed bridge. In another work [14], the authors applied an improved sparse Bayesian learning (iSBL) scheme for high-precision data modeling, producing accurate probabilistic predictions in both time and frequency domains. Wang and Wu [15] presented an improved explicit connectivity Bayesian networks (ECBNs) for system reliability assessment, especially suitable for systems with rare damage field data, obtaining promising results. Despite their accuracy and automation, these approaches involve substantial costs, time commitments, and potential bridge closures due to sensor installations. Consequently, many short- and medium-span bridges forego structural health monitoring (SHM) systems due to practical, economic, and installation constraints. To address these challenges, researchers have explored indirect health monitoring methods based on the concept of vehicle-bridge interaction.

Previous research has extensively explored the concept of indirect bridge condition monitoring through vehicle-assisted measurements. Yang et al. [16] were among the pioneers in this field, where they investigated the equation of motion integrating the dynamic properties of both the vehicle and bridge. They represented the acceleration of the passing vehicle as a function of the bridge’s dynamic characteristics. Subsequent studies by various researchers worldwide have further contributed to indirect health monitoring [17–20]. For instance, Bu et al. [21] proposed an approach based on dynamic response sensitivity analysis using acceleration measurements on the vehicle to detect bridge damage. Their method involved an iterative procedure using 3-parameter and 5-parameter vehicle models, exposing damage in terms of flexural stiffness reduction. Numerical analyses demonstrated the effectiveness of their approach in the presence of measurement noise and road surface roughness. Talebi-Kalaleh and Mei [22] introduced an innovative method for bridge modal analysis using vehicle-mounted accelerometers. They mapped the crossing vehicle contact-point responses to some virtual sensing nodes on the bridge, applying an inverse problem solution with cubic spline functions to accurately identify the mode shapes of the bridge. They further improved accuracy with an ARX-based signal prediction approach. Zhang et al. [23] proposed an approach to extract the mode shape square

(MOSS) of a bridge from a passing vehicle. By exciting the bridge using tapping devices while the vehicle passes, they defined the damage feature as the difference in MOSS between intact and damaged bridges. Their numerical analysis and experiments demonstrated successful damage localization even in the presence of high-level noise. Matarazzo and Pakzad [24] proposed structural identification using the expectation maximization (STRIDE) method to address the challenge of missing observations in modal analysis. They successfully applied this technique to data collected from the Golden Gate Bridge using mobile sensing, accurately identifying 19 modes. In addition, Matarazzo and Pakzad [25] introduced a dynamic sensor network (DSN) approach to efficiently store measurement data from a large number of mobile sensing nodes. They also presented a truncated physical model for data processing, demonstrating the efficiency of DSN through examples of high-resolution mobile sensing and big data processing. Despite notable advancements in this field, several challenges persist. These include the limited interaction time between vehicles and bridges, which can hinder the acquisition of sufficient information to detect potential damages and assess their severity. In addition, sensitivity to ambient noise remains a concern, and the requirement for knowledge regarding vehicle configurations and contact-point responses of the vehicles remains a critical consideration.

The rapid advancement of Internet of Things (IoT) technologies has facilitated the effortless installation of sensors on smart devices such as electric vehicles and cellphones, allowing the collection of large-scale real-time data for SHM applications [26, 27]. Mei and Gul [28, 29] introduced a crowdsensing-based methodology for damage detection in bridges using data collected by a large number of passing-by vehicles. Their innovative framework was validated through both numerical simulations and laboratory experiments, allowing them to establish a correlation between damage severity and feature magnitude. Matarazzo et al. [26] conducted a real-field experiment to assess the feasibility of crowdsensing-based bridge monitoring techniques. Leveraging smartphone data from everyday vehicle trips on both the Golden Gate Bridge and a concrete bridge, which included controlled field experiments and uncontrolled Uber rides, they could accurately determine the critical modal properties of the bridges, such as natural frequencies using the smartphone data from different vehicle trips.

However, continuous data collection and transmission with high sampling frequencies pose challenges, particularly in terms of public participation due to potential smartphone battery drain and data plan consumption [29]. This paper introduces a novel crowdsensing-based framework that leverages mobile sensors, including smartphones or embedded sensors in smart vehicles. The framework employs compressed sensing (CS) for data collection and indirect monitoring of bridges, enhancing efficiency by recovering signals from fewer samples. By aggregating acceleration data from numerous vehicles passing over a bridge, the framework extracts features through Mel-frequency cepstral analysis. It then estimates probability distributions and

performs a comparative analysis between baseline and unknown cases for damage detection, enabling continuous and simultaneous bridge monitoring. Numerical analyses and laboratory experiments validate the methodology, demonstrating that damage detection using reconstructed signals remains comparable to results obtained from original data, even with significant compression levels. The effectiveness of the proposed method is further demonstrated through its implementation and validation in a real-world scenario. Smartphone data from 102 vehicles crossing the Golden Gate Bridge is utilized, providing a practical and tangible application of the methodology.

The paper is structured as follows: Section 1 offers an introduction and reviews related work in the realm of bridge condition monitoring. Section 2 provides a detailed exposition of the proposed methodology, while Sections 3 and 4 delve into numerical investigations and laboratory experiments, respectively. The feasibility and real-world application of the proposed method are explored in Section 5. The paper concludes with Section 6, providing a summary of the key findings and discussing potential avenues for future research.

2. Methodology

This study employs a novel methodology that utilizes compressed data obtained from crowdsourced acceleration responses of a large number of passing vehicles, utilizing the smartphones of the drivers to detect and quantify damage on bridges. The framework overview is illustrated in Figure 1.

In this methodology, acceleration signals from each passing vehicle are collected using a random sampling technique to reduce the number of measurement data points while retaining the necessary information and minimizing Internet data usage. Subsequently, the compressed data is transmitted to a processing center via the Internet, where the CS theory is employed to reconstruct the original high-frequency signals from the compressed data.

To extract engineering features from the reconstructed signals, Mel-frequency cepstral (MFC) analysis is employed. In addition, principal component analysis (PCA) is used to reduce the dimensionality of the extracted MFCCs while preserving the most informative components. Probability distribution functions (PDFs) of the extracted features are computed separately for the baseline case (representing the intact bridge) and the unknown cases (indicating a potentially damaged bridge). The damage index is computed by quantifying the dissimilarity between the probability distributions of the baseline case and an unknown case using the Wasserstein distance metric. The index is then normalized with respect to the baseline validation case to define a damage measure. A value greater than one on this index indicates the presence of damage or anomalies in the bridge structure.

2.1. Compressed Sensing. CS is a revolutionary technique in signal processing that enables the efficient acquisition and reconstruction of sparse or compressible signals from highly

incomplete random sets of measurements [30]. It challenges the traditional Nyquist–Shannon sampling theorem by leveraging the prior knowledge that most real-world signals are sparse or can be represented sparsely in a certain domain. In SHM, CS is mainly used in data loss applications. Bao et al. [31] presented a literature review on the emerging use of CS technology in SHM data management. They highlighted that CS offers a new sampling theory for reducing data acquisition by reconstructing sparse or compressible signals from incomplete measurements. The authors discussed various applications of CS in SHM, including acceleration data, lost data recovery, acoustic emission data, moving load distribution identification, and structure damage identification. Their investigation demonstrated the promising potential of CS in SHM. Almasri et al. [32] explored the use of discrete cosine transform (DCT) as a data compression technique for SHM data. They integrated a time-frequency blind source separation technique with DCT-based compression to assess its accuracy in modal identification. Their results, validated through numerical and experimental studies, demonstrated that DCT can effectively compress vibration data containing damage signatures and low energy modes. CS has also proven invaluable in handling missing values within the spatio-temporal response matrix of bridge structures in scenarios involving mobile sensors instead of fixed ones. For instance, Jana and Nagarajaiah [33] introduced a formation control framework that harnesses data from multi-agent mobile sensors to estimate the dense full-field vibration response matrix of the structure, utilizing the compressed sensing algorithm in the spatial domain. Their proposed method successfully obtained highly accurate responses, demonstrating the efficacy of compressed sensing in filling in missing data within the spatial domain.

A sparse signal contains mostly zero or close to zero elements. For instance, in an image, many pixel values may be zero or have minimal intensity. CS aims to exploit this sparsity during the sampling process to acquire the signal more efficiently [34]. The acquisition process in CS involves projecting the signal onto a lower-dimensional subspace using a sensing matrix. This matrix, typically random or pseudo-random, captures linear combinations of the original signal. By reducing the signal's dimensionality while preserving essential information, CS achieves more efficient acquisition. Consider a sparse signal \mathbf{x} of length N , which can be represented as a vector in some basis or transform domain (e.g., Fourier, wavelet, etc.).

$$\mathbf{x} = \mathbf{A}_N^{-1} \mathbf{X}, \quad (1)$$

where \mathbf{X} ($N \times 1$) is the discrete cosine transform coefficients of the original signal in the frequency domain, and \mathbf{A}_N^{-1} is the inverse operator of the discrete cosine transform (I-DCT) ($N \times N$) that depends only on the time step size and the number of samples. The main reason for using DCT in CS is that it tends to concentrate the signal energy in a smaller number of coefficients compared to the discrete Fourier transform (DFT). This property makes it particularly suitable for data compression applications

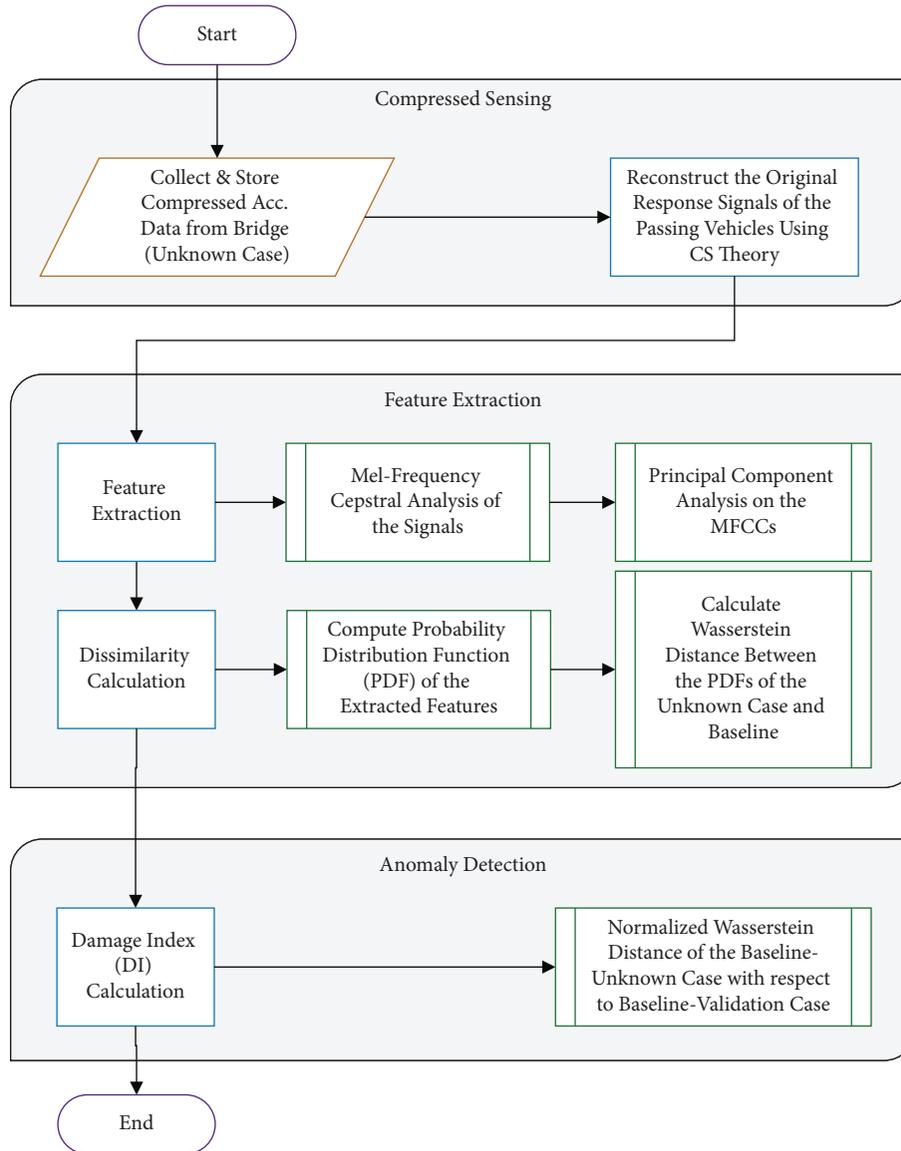


FIGURE 1: Proposed framework for damage detection based on CS.

where retaining the most significant information is important. DCT has several advantages over other signal processing tools [35]: (1) it has the ability to concentrate energy in the lower frequencies, and (2) it can reduce the blocking artifact's effect that results from the boundaries between subimages as they become visible. Owing to the

above two properties, the DCT provides a very good compromise between information-packing ability and computational complexity.

The commonly used discrete cosine transform is DCT Type-II, which is an orthonormal transform, and its inverse operator is as follows [35]:

$$\mathbf{A}_N^{-1} = \sqrt{\frac{2}{N}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \cos\left(\frac{\pi}{N}\right) & \dots & \cos\left(\frac{(N-1)\pi}{N}\right) \\ \frac{1}{\sqrt{2}} & \cos\left(\frac{2\pi}{N}\right) & \dots & \cos\left(\frac{2(N-1)\pi}{N}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{2}} & \cos\left(\frac{(N-1)\pi}{N}\right) & \dots & \cos\left(\frac{(N-1)(N-1)\pi}{N}\right) \end{bmatrix}. \quad (2)$$

The compressed measurement of \mathbf{x} , denoted as \mathbf{y} ($M \times 1$), is obtained through a linear transformation of the original signal using a randomly generated sensing matrix, Ψ , with dimensions $M \times N$, where $M \ll N$. This matrix comprises entries of zeros and ones (a logical matrix), adhering to a Gaussian distribution [36]. It is crucial to emphasize that the number of data points to be measured (M) is contingent on the desired compression level. In addition, matrix Ψ must be generated before initiating the measurement process, and it needs to be provided as input to the computer responsible for reconstructing the original signal from the compressed data.

$$\mathbf{y} = \Psi \mathbf{x}. \quad (3)$$

Hence, the goal is to determine the unknown discrete cosine transform of the original signal (\mathbf{X}) based on the given compressed measurement (\mathbf{y}) by minimizing the reconstruction error using the following equation:

$$\min_{\mathbf{X}} \left[\|\Psi \mathbf{A}_N^{-1} \mathbf{X} - \mathbf{y}\|^2 + \gamma \|\mathbf{X}\|_1 \right]. \quad (4)$$

In the equation above, γ represents the regularization parameter, while $\|\mathbf{X}\|_1$ denotes the Manhattan norm, which quantifies the sum of absolute values for the vector's elements. By using any optimization package, we can find an optimal estimate of \mathbf{X} (i.e. $\hat{\mathbf{X}}$). Then, we can substitute it into equation (1) to obtain the reconstructed signal with the same size as the original signal ($\mathbf{x}_{\text{recon}}$).

The conventional algorithm for the CS method to reconstruct original signals with a higher sampling rate using the compressed measurement is as follows (Algorithm 1):

2.2. Feature Extraction Using Mel-Frequency Cepstral Analysis. The application of Mel-frequency cepstral (MFC) analysis for feature extraction in indirect health monitoring was initially introduced in the domain of bridge health monitoring by the authors in 2019 [28, 29]. Cepstral analysis was chosen in the proposed method due to its capability to extract information across a wide range of frequencies, rather than focusing solely on peaks. Although conventional cepstral analysis techniques assign equal weights to different frequency ranges, MFC analysis assigns higher weights to lower frequencies, making it more suitable for bridge monitoring. The design of Mel-frequency cepstral coefficients (MFCCs) was inspired by the human auditory system's response to auditory stimuli. Similar to the human perception of sound, where the perceptual difference between 100 Hz and 200 Hz is much more significant than between 10,000 Hz and 10,100 Hz, despite their equal linear distances, the natural frequency of vibration for bridges exhibits a similar characteristic. The discrepancy in the lower frequency range, which encompasses the majority of significant modes, tends to be more substantial compared to the difference in higher frequency ranges [28].

The procedure for calculating MFCCs from signal \mathbf{x} can be summarized as follows:

- (1) Perform the Fourier transform on the acceleration data to convert the signal from the time domain to the frequency domain, resulting in the signal's power spectrum (\mathbf{X}).
- (2) Apply a set of triangular filters ($H_i(k)$) to the power spectrum (n filter in total). Each triangular filter is defined as follows:

$$H_i(k) = \begin{cases} \frac{k - f_{i-1}}{f_i - f_{i-1}}, & \text{if } f_{i-1} \leq k < f_i, \\ \frac{f_{i+1} - k}{f_{i+1} - f_i}, & \text{if } f_i \leq k < f_{i+1}, \\ 0, & \text{otherwise,} \end{cases} \quad ; i = 2, 3, \dots, n-1, \quad (5)$$

$$f_i = M^{-1} \left(M(f_1) + (i-1) \times \frac{M(f_n) - M(f_1)}{n-1} \right),$$

Input: Compressed measurement of \mathbf{y} and sensing matrix Ψ
Output: Reconstructed signal with higher sampling rate ($\mathbf{x}_{\text{recon.}}$)
Step 1: Determine the number of rows (M) and the number of columns (N) of matrix Ψ .
Step 2: Determine the inverse discrete cosine transform operator matrix \mathbf{A}_N^{-1} .
Step 3: Solve the optimization problem:
 $\hat{\mathbf{X}} = \operatorname{argmin}_{\mathbf{X}} [\|\Psi \mathbf{A}_N^{-1} \mathbf{X} - \mathbf{y}\|^2 + \gamma \|\mathbf{X}\|_1]$
Step 4: Reconstruct the original signal using the estimated DCT of that:
 $\mathbf{x}_{\text{recon.}} = \mathbf{A}_N^{-1} \hat{\mathbf{X}}$

ALGORITHM 1: Compressed sensing algorithm.

where $H_i(k)$ and f_i are the magnitude response and center Hertz-scale frequency of the i -th triangular filter, respectively. k is the frequency index, and n is the total number of the selected frequencies.

- (3) Calculate the Mel-scale and Hertz-scale frequencies using the following mappings:

$$m_i = M(f_i) = 5 \ln \left(1 + \frac{f_i}{5} \right), \quad (6)$$

$$f = M^{-1}(m_i) = 5 \left(e^{\frac{m_i}{5}} - 1 \right).$$

Here, m_i denotes the corresponding evenly spaced Mel-scale frequency [28].

4. Compute the logarithms of the powers (sum of the products of triangular filters and power spectrum) at each Mel-scale frequency using the following equation (7):

$$\log E(i) = \log \left(\sum_k H_i(k) \times |X(k)|^2 \right); \quad i = 2, 3, \dots, n-1. \quad (7)$$

- (5) Extract the MFCCs by applying the discrete cosine transform (DCT) to the logged powers:

$$\text{MFCC}_j = \sum_{i=2}^{n-1} \log E(i) \cos \left[\frac{\pi}{n-2} \left(i - \frac{3}{2} \right) j \right]; \quad j = 1, 2, \dots, n. \quad (8)$$

The MFC analysis process is depicted in Figure 2, providing an overview of the involved steps. For the purpose of damage detection in bridges, a specific number of MFCCs can be employed [28]. In this study, we have selected 20 and 30 MFCCs for numerical and experimental investigations, respectively. The n -selected MFCCs are extracted from each of the recorded acceleration signals (m signals in total), which represent observations collected from different passing vehicles at various time windows within the same system, encompassing either a damaged case or the baseline. These chosen MFCCs form the feature matrix \mathbf{U} , resulting in dimensions of m by n .

2.3. Anomaly Detection Using Extracted Features. In the context of anomaly detection, the dissimilarity between the probability density functions (PDFs) of a feature in two different states can be used to define anomalies. Figure 3 presents an illustrative example of anomaly detection using bivariate features with a normal distribution.

After extracting the feature matrices through MFC analysis of the crowdsensed acceleration data, n -dimensional PDFs are computed based on these matrices for both the k -th unknown case (\mathbf{Q}_k) and the baseline case (\mathbf{P}). These PDFs provide valuable insights into the statistical characteristics of the observed data and serve as the basis for quantifying anomaly detection and the identification of structural damage. Figure 4 illustrates a sample pairwise scatterplot of the first 3 MFCCs, demonstrating the statistical pattern of the feature matrices for the baseline and a damaged case, highlighting their applicability for anomaly detection. An underlying assumption in this paper is that the extracted feature matrix should exhibit stable distributions when the bridge condition remains unchanged.

PCA is utilized to analyze the extracted feature matrices and effectively reduce their dimensionality while identifying the most significant components. The distance between two multivariate probability PDFs using the Wasserstein distance, also known as the Earth Mover's distance [37], is a fundamental concept in probability theory for quantifying dissimilarity between probability measures. The p -Wasserstein distance, where $p \geq 1$, is defined as follows:

For $p \in [1, \infty)$ and Borel probability measures \mathbf{P} and \mathbf{Q}_k on \mathbb{R}^n with finite p -moments, their p -Wasserstein distance [37] is given by

$$W_p(\mathbf{P}, \mathbf{Q}_k) = \left(\inf_{\pi \in \Gamma(\mathbf{P}, \mathbf{Q}_k)} \int_{\mathbb{R}^n \times \mathbb{R}^n} |\mathbf{X} - \mathbf{Y}|^p, d\pi \right)^{1/p}. \quad (9)$$

Here, samples $X = (X_1, \dots, X_n)$ are drawn from the probability measure \mathbf{P} , and samples $Y = (Y_1, \dots, Y_n)$ are drawn from the probability measure \mathbf{Q}_k . Both \mathbf{P} and \mathbf{Q}_k are probability measures (functions) defined on R^n , where n represents the dimensionality of the data. The symbol π represents a joint probability measure on $R^n \times R^n$. It is an element of the set $\Gamma(\mathbf{P}, \mathbf{Q}_k)$, which consists of all possible joint probability measures whose marginals are \mathbf{P} and \mathbf{Q}_k . In other words, for any subset $A \subset R^n$, we have $\pi(A \times R^n) =$

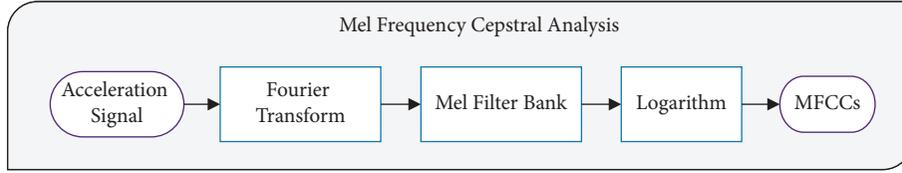


FIGURE 2: Overview of the MFCC analysis process [28].

$\mathbf{P}(A)$ and $\pi(R^n \times A) = \mathbf{Q}_k(A)$. The p-Wasserstein distance, denoted as $W_p(\mathbf{P}, \mathbf{Q}_k)$, measures the minimum cost of transforming the distribution represented by \mathbf{P} into the distribution represented by \mathbf{Q}_k . This cost is computed by minimizing the integral expression over all feasible joint probability measures π between the samples \mathbf{X} and \mathbf{Y} . Smaller values of the metric indicate a closer resemblance, while larger values indicate a greater degree of dissimilarity. It is worth noting that the Wasserstein distance is always nonnegative and finite. To establish a normal range for the damage index, this measure can be computed using the validation data for the baseline structure initially. Then, for an unknown case, the damage index (DI) is defined by normalizing the measured value with respect to the calculated baseline-validation value:

$$DI = \frac{W_p(\mathbf{P}, \mathbf{Q}_{\text{unknown}})}{W_p(\mathbf{P}, \mathbf{Q}_{\text{validation}})} \quad (10)$$

If the DI is approximately 1, it indicates no damage, while values greater than 1 suggest the presence of anomalies or damage in the structure.

3. Numerical Investigations

3.1. Model Setup. To validate the proposed method, a set of numerical data is generated using the finite element software Abaqus. In this analysis, an undamped, simply supported bridge is modeled under a moving mass. The bridge span is set at 25 meters, similar to the bridge described in Yang et al.'s work [16]. The bridge, constructed from reinforced concrete with a density of 2400 kg/m^3 and an elastic modulus of 27.5 GPa , has a cross-sectional area of 2.0 m^2 and a moment of inertia of 0.12 m^4 . It is discretized into 16 elements, and its first three natural frequencies are determined as 2.08 Hz, 8.33 Hz, and 18.75 Hz. The vehicle-bridge interaction is represented as a one-axle moving spring-mass system, as illustrated in Figure 5. The parameters for the base vehicle, including the spring constant, mass, and speed, are defined following the specifications outlined by Yang et al. [16].

To simulate real-world scenarios involving the passage of diverse vehicles across the bridge, various parameters of the spring-mass model were adjusted. These parameters included the mass, spring constant, and speed of the vehicle. The mass of the vehicle was selected from a range of 960, 1200, 1440, 1680, 1920, 2160, or 2400 kg. The spring constant could be set to 200, 250, 300, 350, 400, 450, or 500 kN/m. The vehicle speeds were chosen from a list of 28.8, 36, 43.2, 50.4, 57.6, 64.8, and 72 km/h. By combining all of the possible values for these three parameters, a total of $7 \times 7 \times 7 = 343$ simulations were performed, each representing a unique set of vehicle

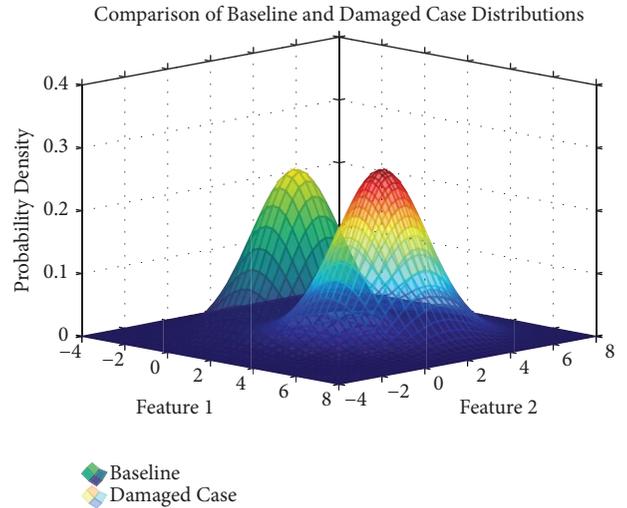


FIGURE 3: The intuition of anomaly detection using a bivariate feature.

configurations. In addition, each vehicle acceleration response was replicated by adding artificial random Gaussian noise at 5 different times, each with a magnitude of 5%. Thus, for each damaged case of the bridge, there were a total of $343 \times 5 = 1715$ recorded acceleration data from the passing vehicle. It is important to note that the 1715 recorded data were separately generated for each damaged case and baseline scenario. Furthermore, by incorporating various configurations of mass and frequency, the study encompasses a comprehensive range of vehicle-to-bridge mass ratios, spanning from 0.8% to 2%. The natural frequencies of the vehicles also cover a broad spectrum, ranging from 9.13 to 22.82 Hz. This inclusive approach serves to thoroughly test the robustness of the framework under diverse conditions.

Considering the fact that different types of vehicles, albeit with a similar distribution, would cross the bridge at different times, only 50% of the 1715 data entries were randomly sampled for each damage case, while the remaining 50% of each case was not used. For the intact case, the 1715 data entries were randomly sampled twice, with 50% being used as the baseline and the other 50% for validation purposes. The validation data was crucial for normalization in anomaly detection. The details of the vehicle simulations and data sampling can be found in Table 1.

The extensive numerical data allowed for the thorough evaluation of the proposed damage assessment algorithm under various damage scenarios and vehicle configurations. The inclusion of artificial Gaussian noise and random sampling ensured the reliability and robustness of the

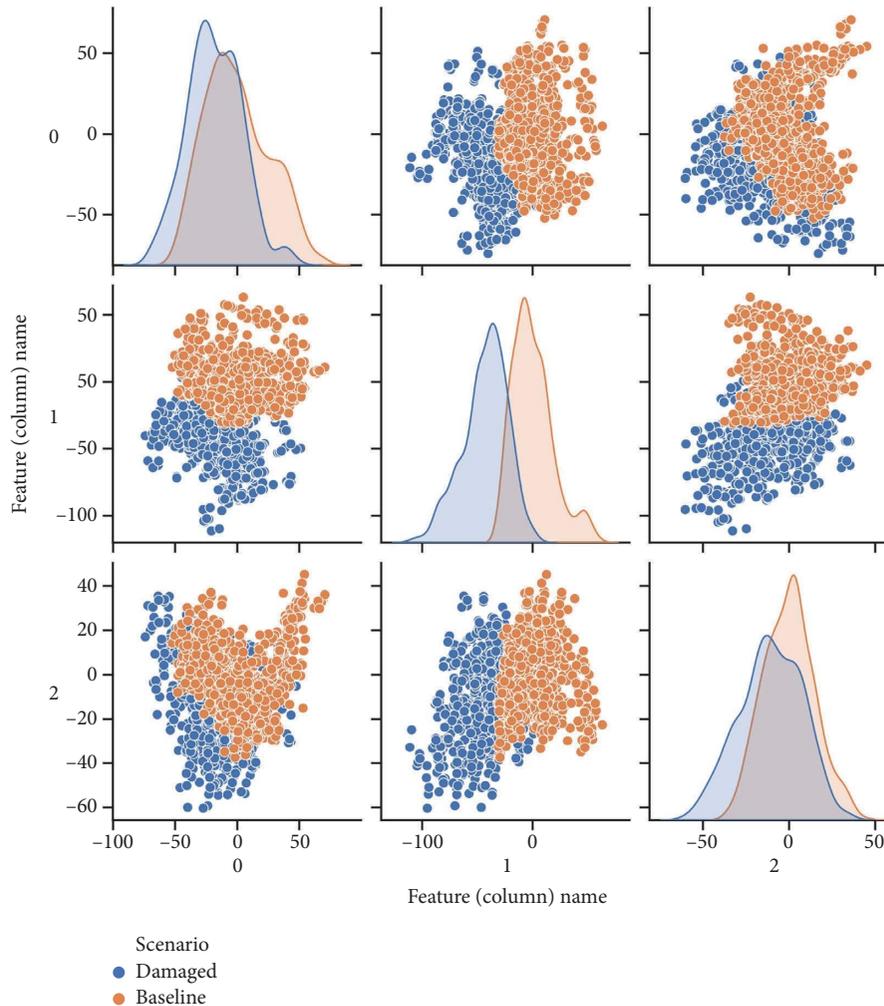


FIGURE 4: Pairwise scatterplot of the first 3 columns of the feature matrices for a sample damaged case and the baseline.

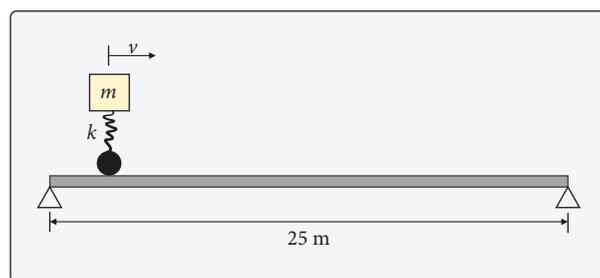


FIGURE 5: Schematic of the simulated vehicle-bridge model.

validation process. It is important to note that the original data obtained from a step-by-step dynamic analysis in Abaqus has a sampling rate of 100 Hz. However, to ensure a fair comparison between the results of the compressed data and the original data, the data generated by the software is uniformly downsampled to match the Nyquist–Shannon sampling theorem, which states that the sampling rate must be at least twice the frequency to be captured in the signal. In

this case, the downsampling rate is set to twice the frequency of the third mode, which is approximately 40 Hz (2×18.75).

It is worth noting that, in the numerical simulations, road surface roughness is deliberately excluded to assess the performance of the method under ideal conditions. However, the impact of road roughness is inherently considered in the experimental investigations, providing a more realistic evaluation of the method's effectiveness in actual scenarios later on.

TABLE 1: Vehicle simulation parameters.

Simulation	Mass (kg)	Spring constant (kN/m)	Speed (km/h)	Repetitions
1	960	200	28.8	5
2	960	200	36.0	5
3	960	200	43.2	5
...
341	2400	500	57.6	5
342	2400	500	64.8	5
343	2400	500	72.0	5

3.2. *Damage Cases.* To validate the proposed anomaly detection framework, five specific damage scenarios were taken into account, in addition to a baseline case representing the intact structure. The baseline case, denoted as DC0, represents an undamaged bridge. In DC1a and DC1b, stiffness reductions of 15% and 30% were, respectively, applied to the mid-span of the bridge. Similarly, in DC2a and DC2b, stiffness reductions of 15% and 30% were respectively applied to the quarter-span of the bridge. It is important to note that these damage cases were simulated by reducing the elastic modulus of the corresponding structural elements. In DC3, the support conditions at both ends were changed from hinged to fixed. Figure 6 illustrates the visual representation of these damaged cases.

As evident in Table 2, the modal analysis of the damaged bridge highlights that quarter-span damage scenarios exhibit a slightly higher severity compared to mid-span damage cases, as evidenced by their higher impact on the fundamental natural frequency. Moreover, the case involving a change in support conditions is considerably more severe than the others, as it induces a substantial alteration in the bridge's fundamental natural frequency. This observation will be further discussed in Section 3.4.

3.3. *Compressed Sensing Scenarios.* In practical applications of compressed sensing for data collection, specific random data points from the response signal are recorded without the need for downsampling. In essence, only the measurement matrix and the compressed measurements need to be stored. However, in this study, the original signals recorded at a higher frequency are artificially downsampled to mimic CS behavior in real-world applications. Three distinct compression levels are considered (CL=20%, 50%, and 80%). For instance, a compression level of 80% indicates that only 20% of the original data, sampled at the Nyquist frequency of 40 Hz, is randomly downsampled, retaining only 20% of the initial data for use as input in the subsequent anomaly detection phase. The numerical validation is carried out using Python, with the complete methodology implemented across various compression levels, and each compression level is executed 30 times to have a robust analysis. The procedural steps for the algorithm are as follows (Algorithm 2):

3.4. *Interpretation of Results.* This subsection presents the damage indexes calculated using the proposed methodology using the data generated from the numerical models for each

of the damage cases (DC1, DC2, and DC3) and across different compression levels (0%, 20%, 50%, and 80%). Furthermore, the section commences by presenting the results of signal reconstruction, emphasizing the effectiveness of CS in extracting crucial vibration information from signals.

3.4.1. *Signal Reconstruction Using Compressed Sensing.* The performance of CS in reconstructing the original signals using the compressed ones is conducted on an acceleration signal recorded from a random vehicle crossing the intact bridge. This signal corresponds to a vehicle with a mass of 1200 kg, a spring constant of 500 kN/m, and a traveling speed of 36 km/h. Figure 7 displays the reconstruction results obtained from the original signal sampled at 40 Hz (Nyquist frequency) for varying compression levels of 20%, 50%, and 80%. It can be observed that even with only 20% of the data points (equivalent to an 80% compression level), the reconstruction of the original acceleration signal remains highly accurate. It is evident that a higher percentage of data points, corresponding to the lowest compression level, results in superior signal reconstruction. The R^2 values, computed to assess the fit of reconstructed data against the original data are 0.996, 0.989, and 0.957 for compression levels of 20%, 50%, and 80%, respectively. In summary, using 50% of the data points for reconstruction yields a satisfactory level of accuracy in comparison to this signal. These results demonstrate the potential of CS to accurately capture essential information from acceleration signals sampled at the Nyquist frequency with much fewer data points.

3.4.2. *Damage Detection Results for Mid-Span Cases.* Figure 8 illustrates the damage index calculated for mid-span damage cases across different compression levels. To ensure result consistency, the damage index was computed for 30 independent runs for each damage case. As depicted in the figure, for case DC1b, the results from all runs, along with their means, clearly indicate the presence of damage in the bridge. On the other hand, detecting the presence of damage in the lighter damage scenario (DC1a) becomes slightly challenging at higher compression levels (50% and 80%) as it is mixed with the validation case results. Nevertheless, the mean results derived from the 30 runs still effectively indicate the presence of damage and capture the underlying pattern of information within the compressed data. It is worth noting that damage detection based on the original signals distinctly highlights the damage with higher indexes.

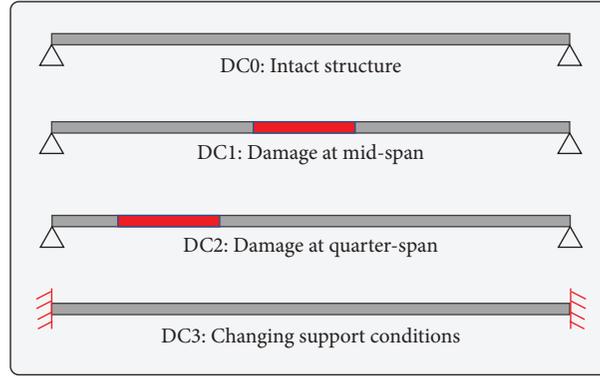


FIGURE 6: Illustration of the considered damage scenarios.

TABLE 2: Damage cases considered in the numerical modeling and corresponding fundamental natural frequencies of the bridge.

Damage case	Location	Stiffness reduction	Bridge fundamental frequency (Hz)
DC1a	Mid-span	15%	2.06
DC1b	Mid-span	30%	2.03
DC2a	1/4 span	15%	2.04
DC2b	1/4 span	30%	1.98
DC3	Boundary condition change	—	4.72

Input: Acceleration signals for the baseline and an unknown case

Output: Damage index for the unknown case

Data Compression: Acquire and compress the acceleration signals of vehicles passing over the bridge, considering a specified compression level for both cases.

Reconstruction of Original Signals: Reconstruct all original signals using the CS theory.

for each run (Run number < 30) do

Signal Selection: Randomly select 50% of the baseline signals as a baseline set. Use the remaining signals as a validation set. Similarly, select only 50% of the recorded signals for the unknown case.

Feature Extraction: Extract feature matrices using MFCC analysis from the selected signals in both cases separately.

PCA: Apply PCA to reduce dimensionality and obtain principal components of the feature matrices separately.

Probability Distribution Functions: Calculate the multivariate probability distribution of the projected features for both cases.

Dissimilarity Measure between the PDFs: Calculate the Wasserstein distance between the probability distribution of the baseline and the unknown case.

end

Average Wasserstein distance: Calculate the average Wasserstein distance over the 30 runs.

Damage Index: Normalize the average Wasserstein distance value with respect to the average from the validation set.

Output: Output the normalized damage index for the unknown case.

ALGORITHM 2: Numerical investigation procedure.

However, as anticipated, increasing the compression level, which results in the loss of more information from the vibration signals, leads to a decrease in the damage index for both cases, gradually approaching the validation case (intact structure).

In summary, for the 30% damage level (DC1b), the higher damage indexes obtained compared to DC1a suggest that the severity of damage can be discerned even from the compressed data. In addition, it is evident that with higher compression levels, the variance of damage indexes calculated from different runs diminishes. This trend can be attributed to the intuition that with more information, the anomaly detection task becomes

somewhat more consistent. Finally, the presented figures consistently demonstrate the reproducibility of the method, as the results align across multiple runs. This consistency further underscores the reliability and robustness of the proposed method, which adeptly extracts damage features even from a mere 20% of the original data in the mid-span damage case.

3.4.3. Damage Detection Results for Quarter-Span Cases. Similar observations can be concluded from Figure 9, representing the results obtained for quarter-span damage cases across different compression levels. According to the

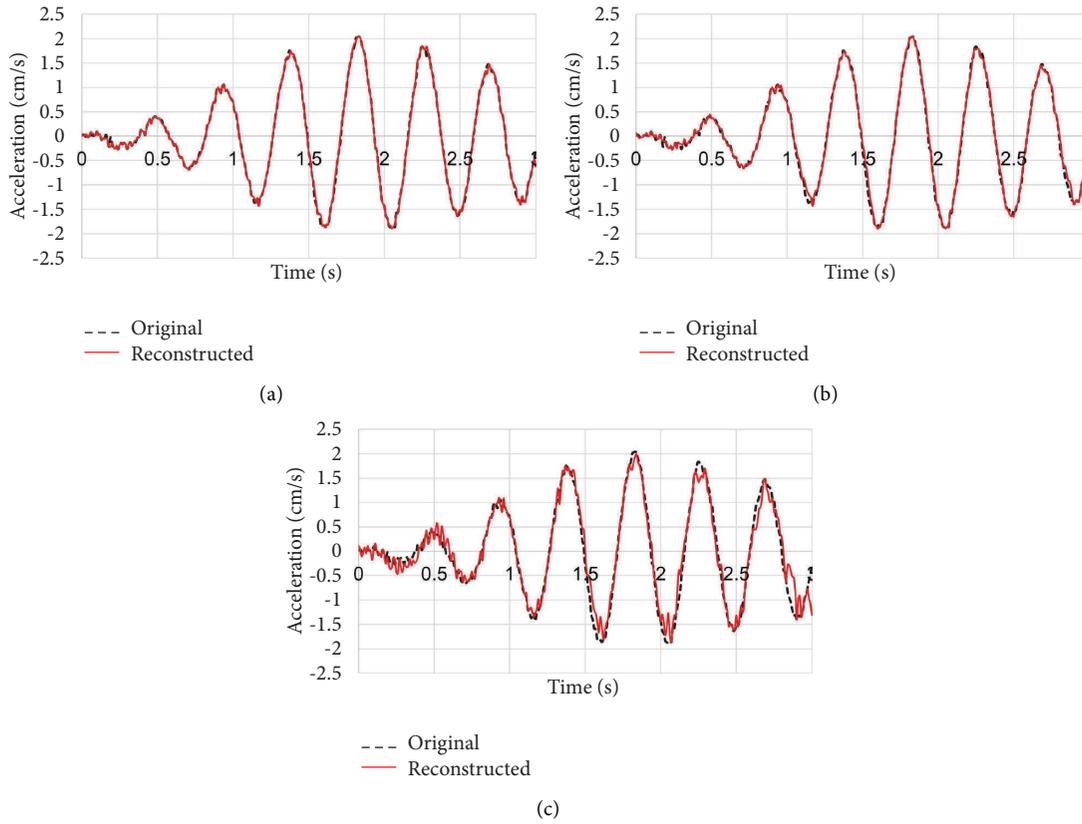


FIGURE 7: Comparison between the original signal (sampled at 40 Hz) and the reconstructed signal for different compression levels, recorded by a randomly selected vehicle while crossing the intact bridge (numerical). (a) Compression level = 20%. (b) Compression level = 50%. (c) Compression level = 80%.

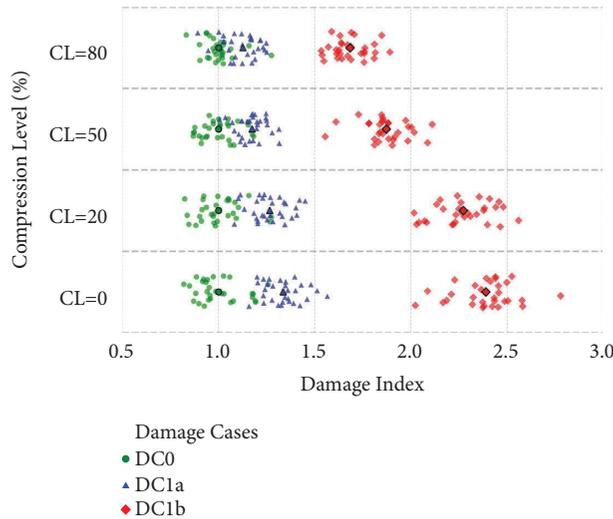


FIGURE 8: Damage index for mid-span damage cases (DC1a and DC1b) at different compression levels (numerical data).

figure, the proposed methodology effectively detects damage at the quarter-span location. Even at higher compression levels (e.g., CL = 80%), where the distribution of the damage indexes starts to overlap with the baseline validation case for the 15% damage level, the average values still indicate the presence of damage.

The ability to accurately detect damage in this case, even under compressed conditions, demonstrates the robustness of the proposed method and its potential applicability in damage detection scenarios without losing important information while saving memory for data collection. The severity of the damage can be identified well, similar to the

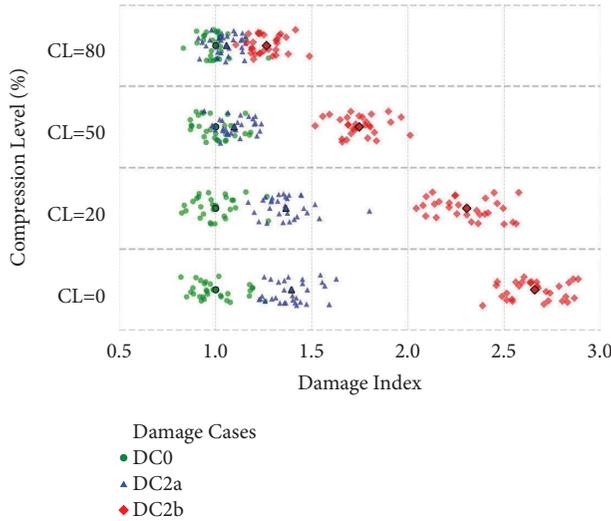


FIGURE 9: Damage index for quarter-span damage cases (DC2a and DC2b) at different compression levels (numerical data).

pattern of the original data, even when using only 20% of the data. This observation becomes more apparent in the 30% damage intensity case. In conclusion, storing only 80% of the original signals (sampled at Nyquist frequency) does not compromise the damage detection results for this case either, and we can still retain enough information in the signals for damage detection.

3.4.4. Damage Detection Results for the End Supports Change Case. Figure 10 displays the results of the damage indexes calculated for the baseline-validation case and the DC3 damage case, where the support conditions were changed from hinge to fixed. The damage index is significantly higher for the DC3 case compared to the previous cases (about 5 times greater), indicating a substantial change in the structure's stiffness. It can be observed that the damage index results for the DC3 case consistently show the same trend across 30 runs, which underscores the reliability of the proposed method in detecting such structural changes. In contrast to the previous cases, this consistency is maintained even at higher compression levels (e.g., CL=80%). The pattern of information loss, resulting in smaller values for the damage indexes with increasing compression levels similar to those of the previous cases, is more evident in this case.

3.4.5. Comparison of the Damage Detection Results. This section offers a comprehensive comparison, providing an overview of the distribution and variability of the damage indexes across different compression levels and damage scenarios. Figure 11(a) presents the mean damage index results for mid-span and quarter-span damage cases at different compression levels, along with error bars representing one standard deviation around the means. In the absence of signal compression (CL=0%), the figure indicates that structural damage at quarter-span locations manifests with slightly greater severity than observed in mid-

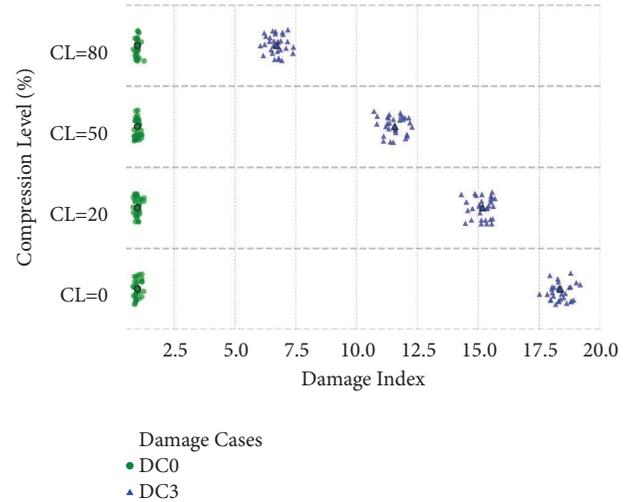


FIGURE 10: Damage index for the change in the end supports case (DC3) at different compression levels (numerical data).

span damage cases. This observed pattern aligns with the behavior evident in the fundamental natural frequency of the bridge for all the damage scenarios, as detailed in Table 2. However, as the compression level increases and important damage-related information is lost in the compressed signals, distinguishing between mid-span and quarter-span damage becomes progressively more challenging. This trend reverses after reaching a compression level between 20% and 50%. Despite this, even at higher compression levels (e.g., 50% or 80% of the original data), the presence of damage remains easily identifiable due to the sufficient difference from the baseline-validation case. This highlights the effectiveness of the CS-based anomaly detection method in detecting damage presence in all scenarios, albeit contradicting the Nyquist–Shannon theorem. It is important to note that the proposed method aims to identify the existence and severity of damage from the compressed signals rather than focusing on damage localization. The reduced variance in the damage index values for highly compressed data (CL=80%), compared to the original data, is attributed to the utilization of a smaller dataset in the CS approach.

Figure 11(b) illustrates the line plot with error bars for the damage index resulting from changes in boundary conditions. Increasing the compression level leads to a linear decrease in the damage indexes, primarily due to the loss of damage-sensitive information as the sampling data points decrease below what's required by the Nyquist–Shannon theorem. Similar patterns and trends are observed in the boundary condition change case compared to other damage scenarios. However, the damage indexes for the boundary condition change case are considerably higher than those for the other damage cases. This substantial difference can be attributed to the significant alteration in the vibrational characteristics of the structures due to the boundary condition change.

By comparing the figures for all damage cases, it can be concluded that the optimal compression level depends on the severity of the damage. For severe damage cases, higher

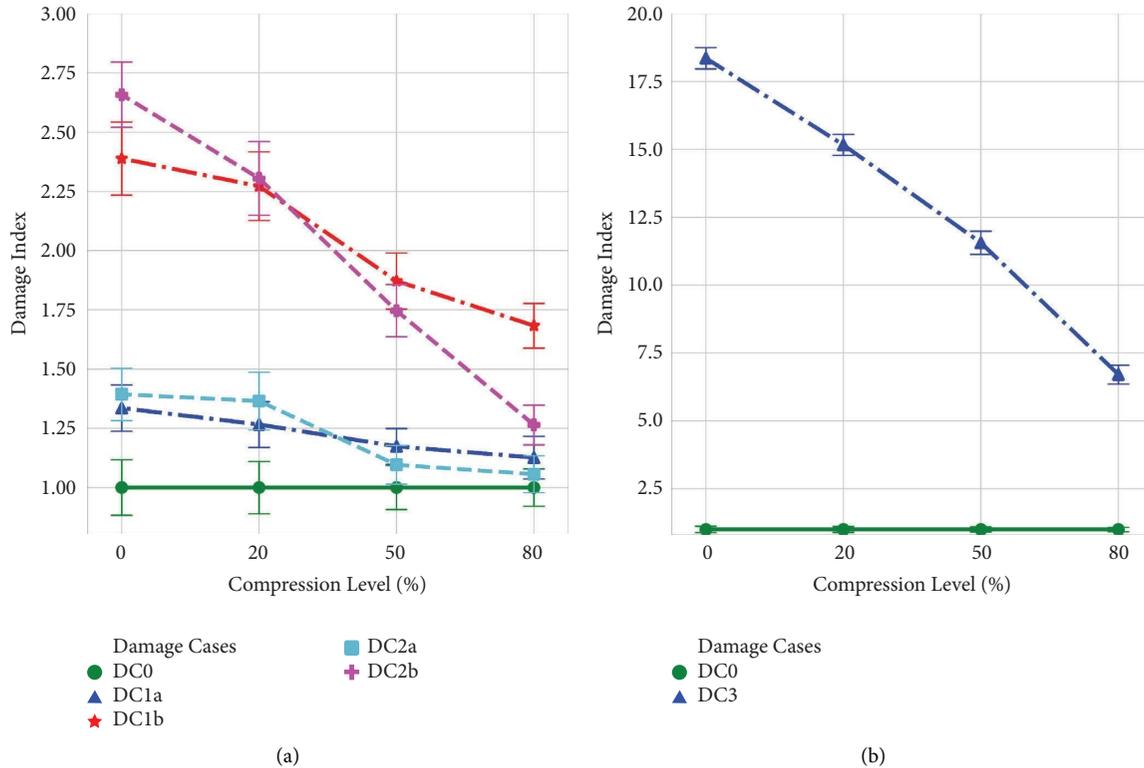


FIGURE 11: Variations of the damage detection results for all damage scenarios at different compression levels (numerical data). (a) Damage at mid-span and quarter-span. (b) Change in end supports (DC3).

compression levels (less information) can be utilized. Moreover, in the context of the numerical model explored in this section, a 20% compression level can still provide all the necessary information for damage identification and quantification, making it a practical choice for SHM applications based on moving sensors.

4. Laboratory Experiments

4.1. Experiment Setup. To validate the proposed method using real sensor-recorded signals, a laboratory-scale, simply supported bridge (depicted in Figure 12) was employed. The bridge deck was constructed using W44 hot-rolled steel, with a modulus of elasticity of 200 GPa. This bridge had a span of 2 meters, a width of 330 mm, and a thickness of 6.35 mm. To assess the effectiveness of the proposed method in damage detection, artificial damages were introduced to the experimental bridge model, following a similar approach to the numerical investigation. Figure 13 provides an overview of the five damage scenarios implemented in the experiments.

The damage scenarios involve various stiffness reductions. DC1a and DC1b represent reductions at the mid-span, while DC2a and DC2b represent reductions at the 1/4 span. In addition, DC3 involves a change in the boundary conditions at both ends of the bridge. Table 3 illustrates the specific dimensions of the cuts for each damage case. To simulate the stiffness reductions, precise cuts are made in the steel bridge. For example, DC1a includes a 24.8 mm × 250 mm cut centered at the mid-span on each side. Similarly,

DC2a involves a comparable cut positioned at a distance of 0.5 m from one end. To achieve a 30% stiffness reduction, DC1b and DC2b incorporate 49.5 mm × 250 mm cuts on each side of the bridge. It is important to note that steel flat bars of the same size as the cut area are loosely attached to the bridge using hot glue to compensate for the mass reduction caused by the cuts. For DC3, each end of the bridge is mounted on a short I-beam using four bolts to implement the boundary condition change.

Emphasizing the inherent distinctions between the lab-scale bridge model and the numerical model is crucial, as they serve different purposes within the scope of this study. Modal analysis of the experimental bridge, as indicated by the fundamental natural frequencies in Table 3, reveals that the damage in the mid-span exhibits greater severity than that in the quarter-span, attributed to a slightly closer positioning of the quarter-span cut towards the mid-span in the experimental bridge model. This distinction is vital to bear in mind when analyzing the results of experimental damage detection and drawing comparisons with outcomes derived from the numerical model. The model vehicle used in the experiments consists of two aluminum plates. Two G-Link-200 wireless accelerometers are mounted on the sides of the top plate, and a Galaxy S5 smartphone is positioned at the center of the top plate. The wireless accelerometers have a sampling frequency of 128 Hz, while the smartphone has a sampling frequency of 100 Hz during the test. The collected data from the two wireless accelerometers is averaged. Various parameters of the model vehicle are considered to

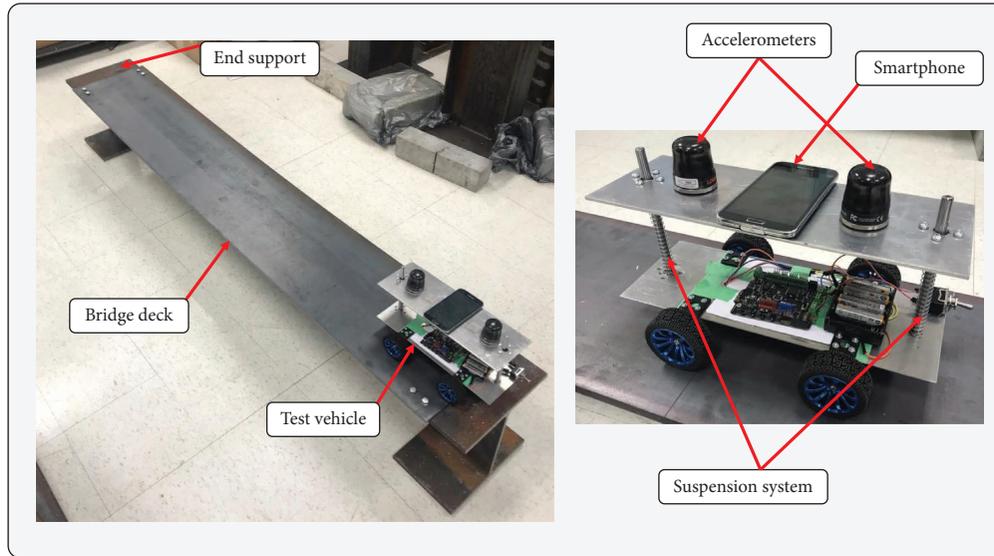


FIGURE 12: Test setup.



FIGURE 13: Created artificial damage in the bridge.

TABLE 3: Damage scenarios considered in the laboratory experiments and their corresponding fundamental natural frequencies of the bridge.

Damage case	Location	Stiffness reduction	Cut dimensions	Bridge fundamental frequency (Hz)
DC1a	Mid-span	15%	24.8 mm × 250 mm	3.62
DC1b	Mid-span	30%	49.5 mm × 250 mm	3.52
DC2a	1/4 span	15%	24.8 mm × 250 mm	3.66
DC2b	1/4 span	30%	49.5 mm × 250 mm	3.61
DC3	Boundary condition change	—	—	8.40

replicate real-world scenarios. The springs in the model vehicle are replaceable, allowing for different spring constants to be used. The spring constant is changed among five different values: 155, 288, 425, 615, and 726 N/m. In addition, the weight of the model vehicle can be adjusted by placing additional masses on the top plate. It is changed among five different levels: 0.898, 0.988, 1.084, 1.170, and 1.270 kg. Furthermore, the speed of the model vehicle can be controlled by programming the Arduino board of the robot, and it varies among three different values: 0.25, 0.33, and 0.40 m/s. The inclusion of a diverse range of values for both the vehicles' mass and spring constant contributes to a broad spectrum of vehicle-to-bridge mass ratios, spanning from 2.73% to 3.86%. The natural fundamental frequencies of these vehicles vary from 11.05 to 28.43 Hz. To ensure statistical significance and reliability, each test is repeated three

times for each model vehicle configuration. Combining all the parameter changes and repeated tests, there are a total of 225 tests conducted for each bridge state. This extensive range of vehicles serves as a reliable evaluation of the framework's robustness to diverse mechanical properties, demonstrating its insensitivity to variations in vehicle characteristics. Table 4 presents the different combinations of vehicle parameters and their corresponding test numbers.

Before conducting the main tests, an impact test is performed to identify the first three modal frequencies of the bridge, which are identified as 3.71, 14.9, and 33.4 Hz. Following the numerical data approach, the initial data collected from both the wireless accelerometer and smartphone sensors is uniformly downsampled to the Nyquist frequency of 70 Hz, approximately twice the third natural frequency of the experimental model. To ensure diversity in

TABLE 4: Vehicle parameters and test combinations.

Test	Spring constant (N/m)	Weight (kg)	Speed (m/s)
1	155	0.898	0.25
2	155	0.898	0.33
3	155	0.898	0.40
...
75	726	1.270	0.40

the vehicle selection for different damage cases, 50% of the data entries from the vehicle pool are randomly sampled. Similar to the numerical analysis, 50% of the tests on the intact bridge are reserved as the baseline case, while the remaining 50% are used for validation. For the other damage cases, only 50% of the tests are selected for damage detection. This approach reflects real-world scenarios where the exact same set of vehicles passing across the bridge at different times is unlikely. However, since the sets of vehicles are sampled from the same pool of configurations, they are expected to follow similar distributions. The sampling process introduces randomness. Therefore, to ensure robustness and verify the method's effectiveness, the method is implemented 30 times with different samples included.

To assess the CS-based anomaly detection method's performance, various compression levels (0% = original data, 20%, 50%, and 80%) are utilized to randomly down-sample the initial data obtained from both smartphones and accelerometers. This downsampling replicates compressed sensing applications in practice. The compressed datasets serve as inputs for the anomaly detection framework.

4.2. Experimental Results. The experimental results are presented in two main sections to validate the method's effectiveness at different compression levels. These comparisons utilize both wireless accelerometers and smartphone sensors. The results for each individual run, as well as the mean of 30 runs, are provided in the plots. It is important to note that the smartphone data underwent initial pre-processing and cleaning procedures before analysis.

4.2.1. Damage Detection Results. By comparing the results obtained from wireless sensors for mid-span and quarter-span damage scenarios, as depicted in Figures 14 and 15, it becomes evident that in the experimental model, mid-span damage exhibits a higher degree of severity compared to quarter-span damage. This difference is particularly pronounced at the 30 percent damage level. Moreover, similar to the results obtained from the numerical model, higher damage levels consistently yield higher damage index values in the experimental setup, demonstrating the consistency and efficiency of the proposed damage metric.

The impact of compression ratios on the real data can be visualized through Figures 14(a) and 15(a), along with Figures 14(b) and 15(b). As these figures are observed, a discernible trend emerges in which an increase in the compression ratio leads to a gradual loss of structural information within the signals. Consequently, lower damage index values are obtained, confirming the inverse relationship

between CR and data quality. Notably, this trend emphasizes the importance of the optimal compression range between 50% and 80%, where damage-sensitive information is preserved without compromising data fidelity.

The comparison between smartphone data and wireless sensor data, as presented in Figures 14 to 16, reveals that smartphone data exhibits sparser patterns and yields lower damage indices when compared to wireless sensor data. This underscores the fact that wireless sensors can capture finer structural details. Despite this disparity, smartphone data remains a viable and cost-effective option for damage detection tasks, eliminating the need for costly commercial sensors.

Examining the impact of boundary conditions, as depicted in Figure 16(a), particularly the transition from a hinge to a fixed boundary condition, provides insightful observations. Unlike the results of the numerical model, this transition does not significantly elevate the observed damage index for this scenario when compared to the results obtained from the simply supported bridge model with damage at mid-span or quarter-span. This observation can be attributed to the nonideal hinge boundary condition inherent in the simply supported experimental model and the incomplete transition to the ideal fixed support. Consequently, the effect of transitioning from hinge to fixed in the lab experiments appears to resemble a damage level slightly exceeding 30 percent within the span of the simply supported bridge.

Furthermore, the relationship between compression levels and damage severity can be observed in Figures 14 and 15. These figures illustrate that for both mid-span and quarter-span damage scenarios, compression levels of up to 20% effectively capture and convey information regarding damage severity. It is noteworthy that the 30% damage scenario consistently yields higher damage index values compared to the 15% damage scenario, although some overlap is discernible in the case of quarter-span damage scenarios acquired via smartphone data.

Finally, an increase in the compression level results in a reduction in variance among damage index values across the 30 runs, as observed. This trend, consistent with our numerical modeling observations, can be attributed to the fact that higher compression levels may reduce the number of damage-sensitive features within the signals, ultimately decreasing sparsity and variance in the signal distributions.

In conclusion, the experimental findings underscore the effectiveness of compressing signals sampled at Nyquist frequencies up to 20% in retaining critical information related to damage severity. This approach not only preserves data integrity but also enhances data collection and transmission efficiency, offering valuable insights for structural health monitoring and damage detection applications.

4.2.2. Comparison of Damage Detection Results and Their Variance. In this section, the distribution and variability of the damage indexes across various compression levels and damage scenarios are investigated, as obtained from both wireless and smartphone sensors. Figure 17 depicts the

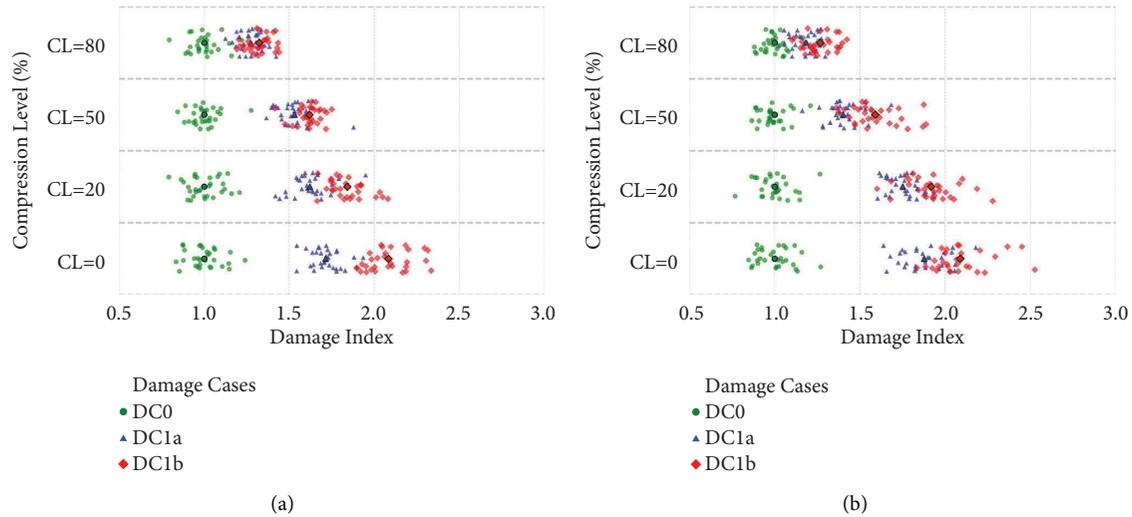


FIGURE 14: Damage detection using wireless accelerometer and smartphone sensor data for mid-span damage cases (DC1a and DC1b). (a) Wireless accelerometer. (b) Smartphone sensor.

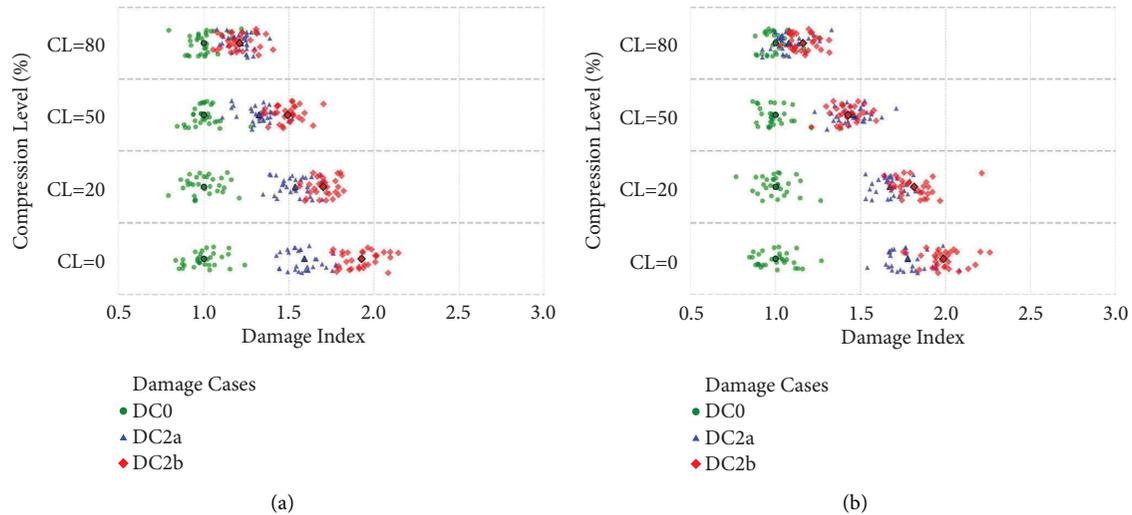


FIGURE 15: Damage detection using wireless accelerometer and smartphone sensor data for the quarter-span damage cases (DC2a and DC2b). (a) Wireless accelerometer. (b) Smartphone sensor.

results derived from wireless and smartphone sensors data for the 15% and 30% damage scenarios at both mid-span and quarter-span locations (DC1 and DC2), as well as the boundary condition change case (DC3). Based on the signals recorded by the wireless accelerometer sensor, Figure 17(a) presents the mean damage indexes for the mid-span and quarter-span damage cases at different compression levels. It is evident that the mean damage indexes for mid-span damage scenarios (DC1a and DC1b) exhibit higher values than their quarter-span cases (DC2a and DC2b) up to a compression level between 20% and 50%. This observation reflects the heightened damage severity in mid-span damage cases.

In addition, as previously mentioned, the greater severity of damage is discernible from the higher damage indexes calculated for the 30% damage level for both mid-span and

quarter-span cases. This suggests that, based solely on the mean of the 30 runs, the optimal compression level can be chosen between 20% and 50%. However, considering the variation in results, it is advisable to opt for a 20% compression level to retain the information regarding damage severity within the signals. It is worth noting that, while the proposed method may miss out on severity information in signals at an 80% compression level, it can still identify the presence of damage in all damage scenarios. This implies that even 20% of the data points required by the Nyquist–Shannon theorem are sufficient for damage detection using the proposed methodology.

Regarding the variability of results, a consistent pattern emerges where an increase in the compression level leads to lower variance among the damage indexes obtained from the 30 independent runs. This observation aligns with the

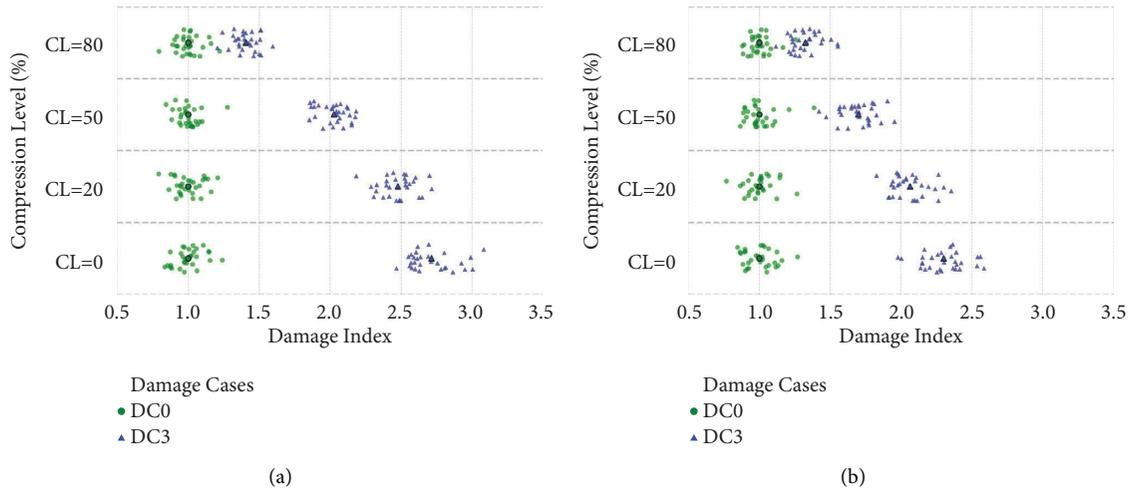


FIGURE 16: Damage detection using wireless accelerometer and smartphone sensor data for the boundary condition change case (DC3). (a) Wireless accelerometer. (b) Smartphone sensor.

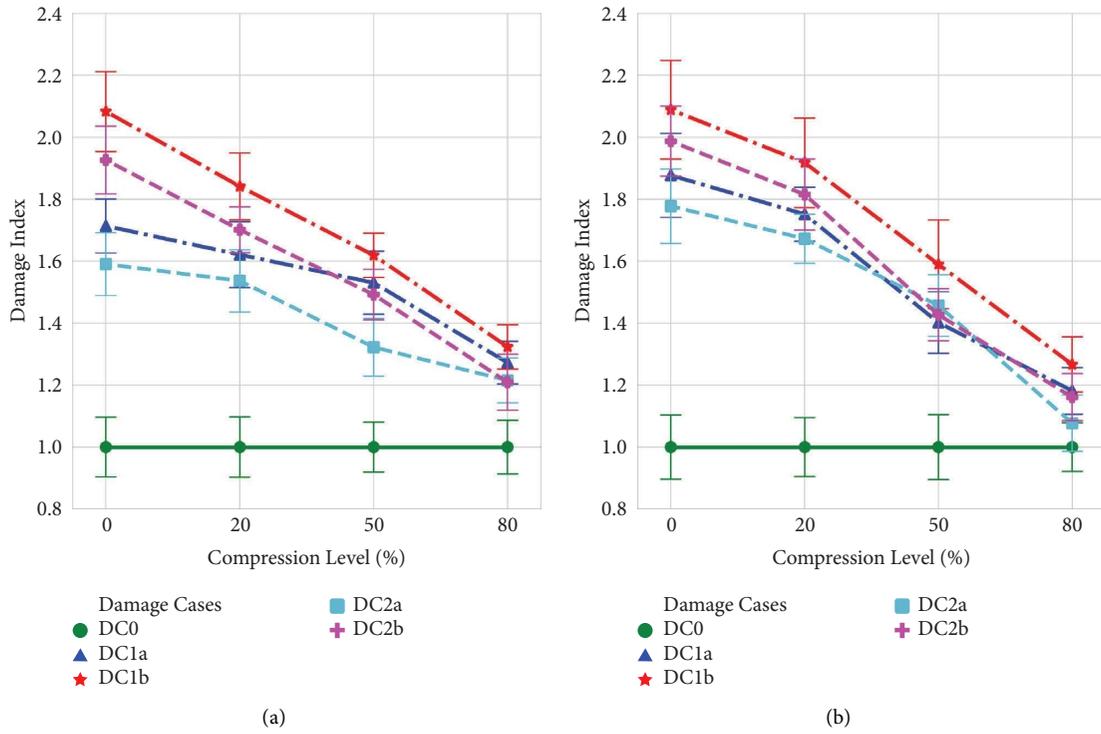


FIGURE 17: Continued.

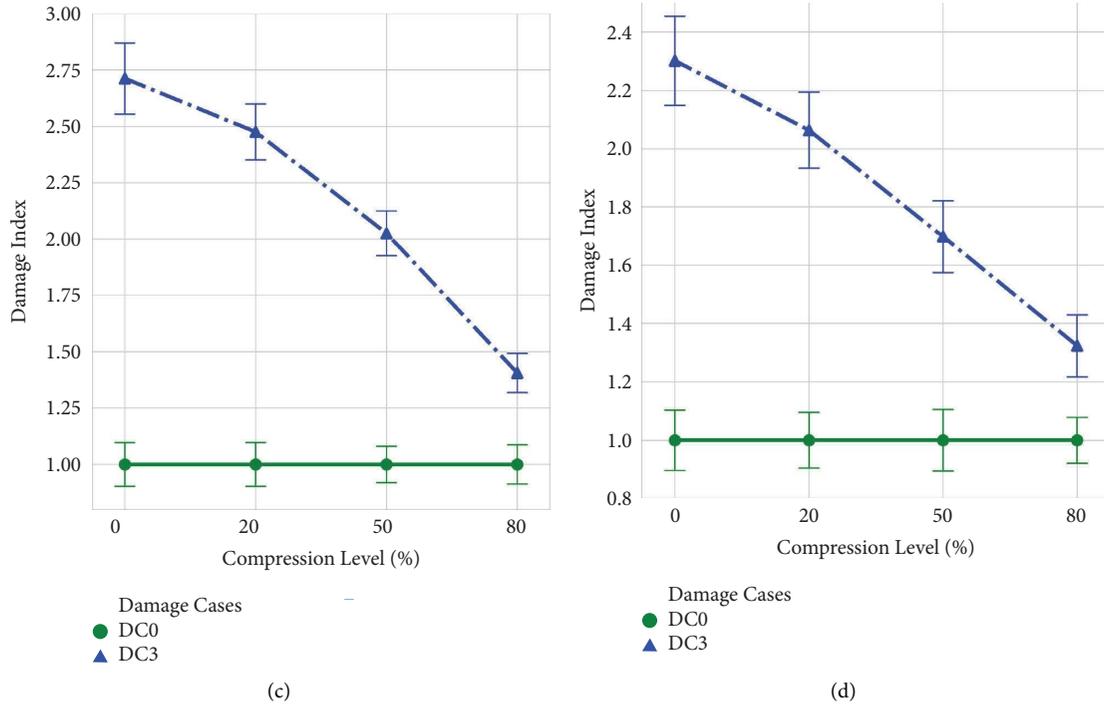


FIGURE 17: Variations of the damage detection results for all damage scenarios at different compression levels. (a) and (c) represent results from the commercial accelerometer, while (b) and (d) are from the smartphone accelerometer. (a) Damage at mid-span and quarter span (accelerometer). (b) Damage in mid-span and quarter-span (smartphone). (c) Change in supports (accelerometer). (d) Change in end support (smartphone).

numerical findings and can be attributed to the fact that higher compression levels may reduce the number of damage-sensitive signal features, ultimately decreasing sparsity and variance within the signal distributions. Comparable patterns and trends are observed in the boundary condition change case (Figure 17(c)) compared to the other damage scenarios. Despite the higher damage indexes for this case, increasing the compression level tends to decrease the damage index, yet damage remains clearly identifiable even at a compression level of 80% for this case.

To assess the capabilities of smartphone sensors in damage detection, which is a primary focus of this study, Figures 17(b) and 17(d) present comparative results based on smartphone sensors. Similar to the wireless sensor results, mid-span damage cases appear more severe than quarter-span damage cases when considering only the mean results. However, to retain information about damage severity within the signals, the optimal compression level can be chosen between 20% and 50%. It should be noted that the damage indexes are smaller, and the variability of results based on smartphone sensors is higher than those based on wireless sensors, indicating that wireless accelerometers can extract more sensitive vibration information from the bridge than smartphone sensors. However, this may depend on the type of sensor used in the smartphone.

5. Case Study: The Golden Gate Bridge

To validate and assess the feasibility of implementing the compressed sensing framework in real-life applications, we

employed a publicly accessible dataset of smartphone accelerations. This dataset was obtained from 102 vehicle trips crossing the Golden Gate Bridge. The Golden Gate Bridge, located in California, USA, is a long-span suspension bridge with a main span length of 1280 meters (see Figure 18 for more details). The natural frequencies of its first four modes are 0.106, 0.132, 0.170, and 0.216 Hz, respectively [39].

5.1. Introduction of the Dataset. The dataset employed in the case study originates from a controlled field test conducted by Matarazzo et al. [26]. They conducted 102 trips across the bridge, recording data using iPhone 5 smartphones equipped with the Sensor Play App. Data collection occurred during morning and afternoon rush-hour periods over five consecutive days (June 18–22, 2017). Each acceleration signal was resampled to 100 Hz. Two sedan-style vehicles were used, a Nissan Sentra for the first fifty trips, and a Ford Focus for the remaining fifty-two, with speeds defined at 32, 40, 48, 56, and 64 km/h.

5.2. Implementation of the Framework. The dataset utilized for this study focuses solely on the intact bridge since it is impractical to induce damage to real bridges. Consequently, the investigation of the field data centers on implementing compressed sensing in crowdsensing-based bridge monitoring applications. In addition, we examine the sensitivity of the damage index at different compression levels, exclusively studying its application to intact bridges. Out of the



FIGURE 18: Scenic overview of the golden gate bridge (adopted from [38]).

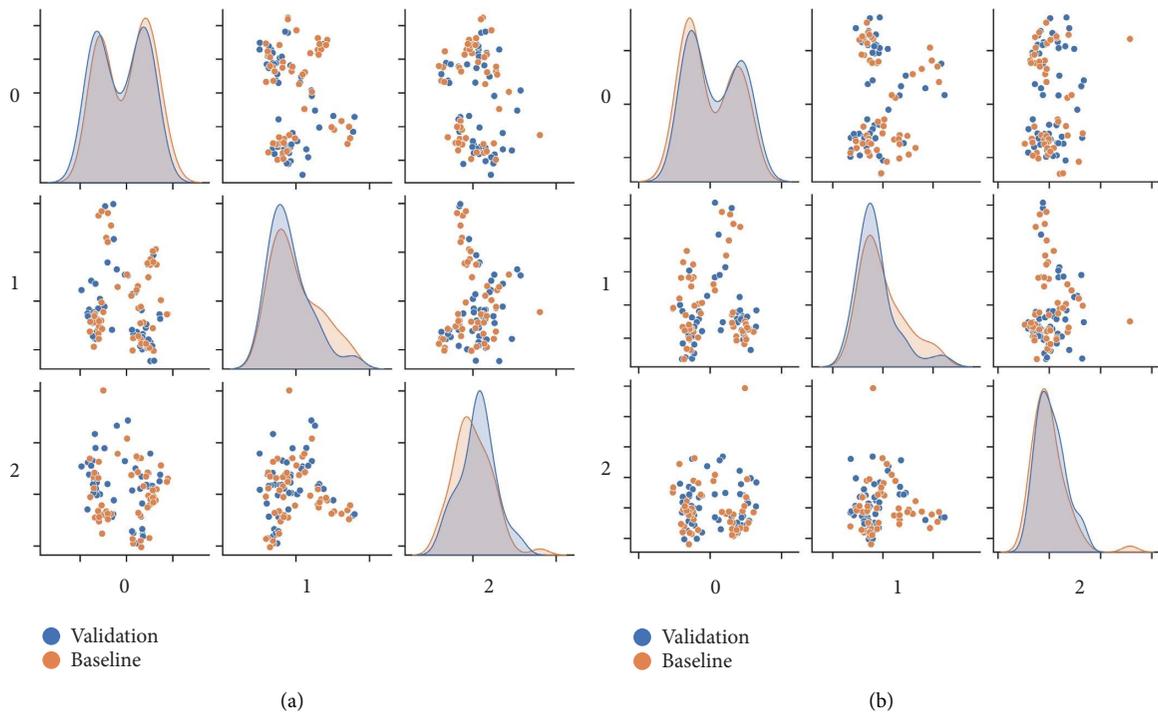


FIGURE 19: Continued.

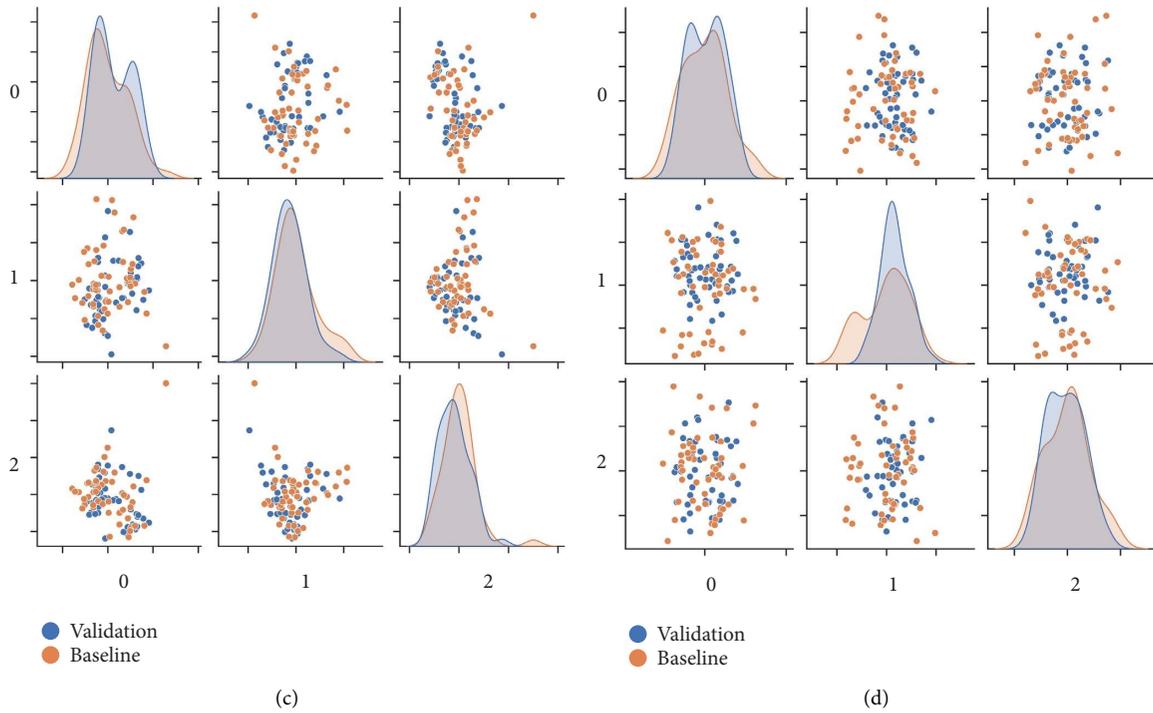


FIGURE 19: Comparison between the first three extracted features (MFCCs) for the baseline and validation sets at different compression levels. (a) Compression level = 0%. (b) Compression level = 20%. (c) Compression level = 50%. (d) Compression level = 80%.

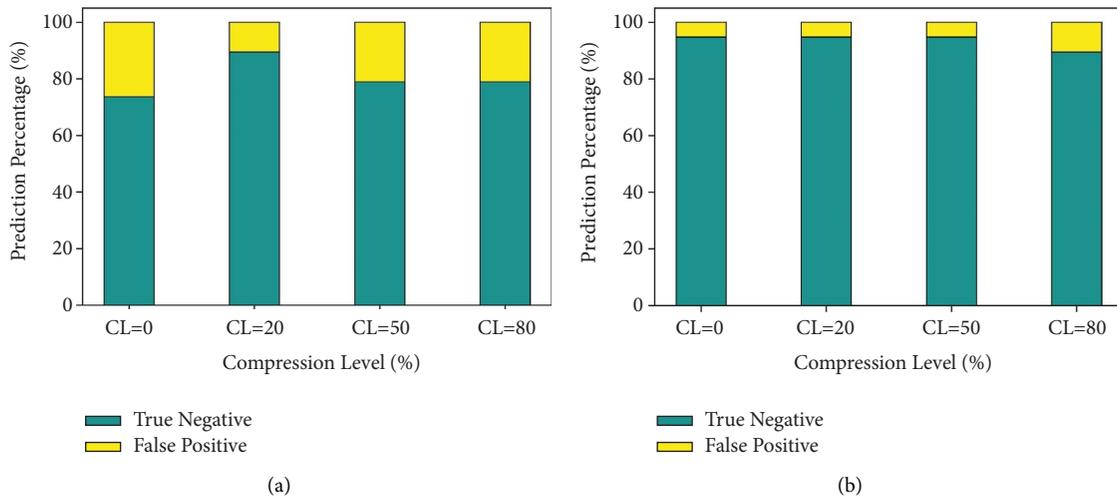


FIGURE 20: Exploration of true negative and false positive results in damage detection under two distinct threshold settings. (a) Threshold = 1.05. (b) Threshold = 1.10.

102 acceleration datasets obtained from two different vehicles at various speeds, 50% of the signals are allocated for training the framework (fitting the baseline). The remaining data serves as an unknown case (validation set) for damage index calculations. Given the intact nature of the bridge throughout all trips, a damage index very close to one is anticipated. To assess the framework’s robustness to the training set, the random selection of the training set among the 102 trips is repeated 30 times.

5.3. *Interpretation of Results.* Figure 19 illustrates the statistical distribution of the first three features (MFCCs) for the training set (baseline) and the unknown set (validation). A notable change in the probability density function (PDF) of both baseline and validation cases is observed when transitioning from CL = 0% to CL = 80%. However, a clear consistency between the two sets prevails, with low dissimilarity indicating an absence of damage within the bridge. The validation case closely follows the baseline pattern.

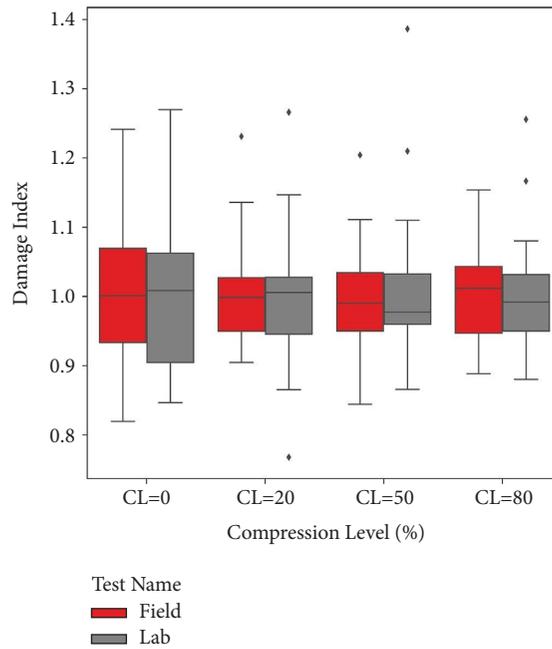


FIGURE 21: Comparison of damage indices computed for intact bridges in both laboratory and field experiments under varying compression levels.

The damage index proposed in this study quantifies damage intensity. In a real-world implementation, to trigger an alarm for the existence of damage in bridges, one can set an intact/damaged threshold. This threshold, while challenging to determine, can be optimized by comparing the baseline set and validation sets using data collected during a specific period when the bridge was intact. Figure 20 demonstrates the impact of two different threshold values (1.05 and 1.10) on true and false alarms. Increasing the threshold from 1.05 to 1.10 reduces false alarms (false positive predictions) from around 20% to less than 5%, highlighting the importance of a sufficiently small threshold.

As previously mentioned, introducing synthetic damage to real bridges is unfeasible. Therefore, evaluating the damage detection capabilities of the proposed framework poses a challenge, prompting a laboratory experiment in this research project. To comprehend the implementation challenges of the proposed compressed-sensing-based damage detection, a comparison between the results of the intact (validation) case from laboratory experiments (using smartphone data) and the field test is beneficial. Figure 21 reveals a close match in the variation of the validation case between both experiments. Similar to the results from the lab experiments, an increase in compression level augments result variation. However, as highlighted in sections 3 and 4, higher compression levels may lead to overlooking damage-sensitive features and cause interference in pinpointing the damage location.

In conclusion, the field experiments successfully evaluated the feasibility of implementing the proposed damage detection framework. The study demonstrated the consistency of the introduced damage index in real-world applications, accounting for environmental and operational effects, as well as different vehicle speeds and properties.

6. Conclusion

This paper introduces a novel methodology for damage detection on bridges, leveraging the use of smartphones to collect acceleration responses from passing vehicles. The collected data is then subjected to data compression and transmission, followed by signal reconstruction using CS theory. Subsequently, MFC analysis is employed to extract pertinent engineering features from the recovered signals, further reduced through PCA. PDFs are computed for both baseline and unknown cases, enabling the calculation of a damage index based on their dissimilarity measured through Wasserstein distance. This integrated approach offers an efficient and cost-effective solution for crowdsensing-based bridge health monitoring.

To validate the methodology, comprehensive numerical investigations were conducted on a simply supported bridge model using finite element simulations in Abaqus. Various damage scenarios, including alterations to the elastic modulus and boundary conditions of structural elements, were considered, all while introducing a controlled 5% artificial noise level into the simulation results. The performance of the proposed method was evaluated by calculating damage indexes across different damage cases and compression levels. The numerical results convincingly demonstrated the method's efficacy in identifying structural damage. Remarkably, even at a compression level of 80% compared to the original data sampled at the Nyquist-Shannon frequency, the method consistently detected the existence of damage, highlighting its robustness. The derived damage indexes provided valuable insights into the presence and severity of the damage, exhibiting consistency and reproducibility across multiple runs.

It is noteworthy that the proposed framework has insensitivity to both speed and mechanical properties of the passing vehicle, which conventionally contribute to additional peaks in the Fourier transform of the vehicle's response, referred to as the driving frequency and fundamental frequency of the vehicle. This robustness is closely tied to the framework's incorporation of PCA following the extraction of the frequency domain features (MFCCs) from the responses of all passing vehicles in the dataset. Through this process, the framework adeptly filters out the driving frequencies and natural frequencies of the vehicles, underscoring its efficacy in mitigating the impact of speed-related variations.

To thoroughly validate the efficacy of the proposed method, experimental investigations were conducted on a simply supported bridge model. Artificial damage scenarios, such as stiffness reductions and changes in boundary conditions, were introduced to simulate practical situations. Data collected from both wireless accelerometers and smartphone sensors during these tests confirmed the method's effectiveness in damage detection. While accurately assessing the severity of damage poses challenges in certain scenarios due to factors like ambient noise and the limitations of smartphone sensors, the method reliably detected damage even at the 80% compression level.

To assess the method's real-world applicability and implementation, the framework was tested using actual smartphone data collected from the intact state of the Golden Gate Bridge. The results demonstrated that the proposed compressed sensing-based method can be seamlessly integrated into real-life crowdsensing-based monitoring applications. The introduced damage index on the validation set yielded a value close to one, indicating the intact state of the bridge. Furthermore, the observed variation in damage indexes across different compression levels mirrored the patterns observed in laboratory applications, underscoring the reliability of the proposed framework.

However, it is crucial to acknowledge the inherent limitations of this methodology. Challenges stemming from traditional indirect health monitoring problems, such as road profile roughness, limited vehicle-bridge interaction (VBI) time, and environmental effects, may impact the framework's performance. These factors introduce additional uncertainties and variations in the collected acceleration data, potentially affecting the accuracy of damage detection. In addition, accurately distinguishing the severity of damage can be challenging in some cases due to factors such as ambient noise and smartphone sensor limitations.

Future research endeavors will be dedicated to comprehending and addressing the aforementioned challenges, aiming to enhance the reliability of the proposed methodology. A systematic investigation of the road roughness should be further conducted. In practice, to reduce the influence of environmental and operational effects, the framework can be complemented with other monitoring data sources, such as weather information or temperature. The integration of such additional data can offer a more comprehensive understanding of the bridge's condition, accounting for potential external factors that may impact

damage detection. Moreover, a promising avenue for future investigation involves inducing controlled damage on bridges slated for demolition. Before and following the introduction of damage, extensive vehicle trips should be done on both the intact and damaged bridges. This dataset holds significant promise for evaluating the efficacy of the proposed methodology in detecting structural damage under real-world conditions.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Grant (RGPIN-2022-04160) and Alliance Grant (ALLRP 576826-22). We gratefully acknowledge Matarazzo et al. for providing the source data used in this study. Their contributions facilitated our research and enabled the validation of our findings.

References

- [1] C.-W. Kim, K.-C. Chang, S. Kitauchi, and P. J. McGetrick, "A field experiment on a steel girder-truss bridge for damage detection utilizing vehicle-induced vibrations," *Structural Health Monitoring*, vol. 15, no. 2, pp. 174-192, 2016.
- [2] Q. Mei and M. Gül, "Novel sensor clustering-based approach for simultaneous detection of stiffness and mass changes using output-only data," *Journal of Structural Engineering*, vol. 141, no. 10, 2015.
- [3] M. Gul and F. N. Catbas, "Structural health monitoring and damage assessment using a novel time series analysis methodology with sensor clustering," *Journal of Sound and Vibration*, vol. 330, no. 6, pp. 1196-1210, 2011.
- [4] Parsons Brinckerhoff Engineering and Industrial Heritage, "Nchrp project 25-25, task 15: a context for common historic bridge types," *Privileged document, The National Cooperative Highway Research Program, Transportation Research Council, National Research Council, Ottawa, Canada*, 2005.
- [5] Federal Highway Administration, "Deficient bridges by highway system," 2016, <https://www.fhwa.dot.gov/bridge/nbi/no10/defbr16.cfm>.
- [6] Canadian Construction Association, Canadian Public Works Association, Canadian Society for Civil Engineering, and Federation of Canadian Municipalities, "Informing the future: assessing the health of our communities' infrastructure. canadian infrastructural rep. card 2019," Technical report, Canadian Infrastructure Report Card (CIRC), Ottawa, Canada, 2019.
- [7] R. Niyirora, W. Ji, E. Masengesho, J. Munyaneza, F. Niyonyungu, and R. Nyirandayisabye, "Intelligent damage diagnosis in bridges using vibration-based monitoring approaches and machine learning: a systematic review," *Results in Engineering*, vol. 16, Article ID 100761, 2022.

- [8] Z. Deng, M. Huang, N. Wan, and J. Zhang, "The current development of structural health monitoring for bridges: a review," *Buildings*, vol. 13, no. 6, p. 1360, 2023.
- [9] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, and S. Nagarajaiah, "Review of bridge structural health monitoring aided by big data and artificial intelligence: from condition assessment to damage detection," *Journal of Structural Engineering*, vol. 146, no. 5, Article ID 04020073, 2020.
- [10] M. Flah, I. Nunez, W. Ben Chaabene, and M. L. Nehdi, "Machine learning algorithms in civil structural health monitoring: a systematic review," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2621–2643, 2021.
- [11] X. Zheng, D.-H. Yang, T.-H. Yi, and H.-N. Li, "Development of bridge influence line identification methods based on direct measurement data: a comprehensive review and comparison," *Engineering Structures*, vol. 198, Article ID 109539, 2019.
- [12] F.-L. Zhang, D.-K. Gu, X. Li, X.-W. Ye, and H. Peng, "Structural damage detection based on fundamental bayesian two-stage model considering the modal parameters uncertainty," *Structural Health Monitoring*, vol. 22, no. 4, pp. 2305–2324, 2023.
- [13] Q.-A. Wang, Y. Dai, Z.-G. Ma et al., "Towards probabilistic data-driven damage detection in shm using sparse bayesian learning scheme," *Structural Control and Health Monitoring*, vol. 29, no. 11, 2022.
- [14] Q.-A. Wang, Y. Dai, Z.-G. Ma et al., "Towards high-precision data modeling of shm measurements using an improved sparse bayesian learning scheme with strong generalization ability," *Structural Health Monitoring*, vol. 23, no. 1, pp. 588–604, 2024.
- [15] Q. Wang and Z. Wu, "Structural system reliability analysis based on improved explicit connectivity bns," *KSCE Journal of Civil Engineering*, vol. 22, no. 3, pp. 916–927, 2018.
- [16] Y. B. Yang, C. W. Lin, and J. D. Yau, "Extracting bridge frequencies from the dynamic response of a passing vehicle," *Journal of Sound and Vibration*, vol. 272, no. 3-5, pp. 471–493, 2004.
- [17] A. Malekjafarian, F. Golpayegani, C. Moloney, and S. Clarke, "A machine learning approach to bridge-damage detection using responses measured on a passing vehicle," *Sensors*, vol. 19, no. 18, p. 4035, 2019.
- [18] H. Shokravi, H. Shokravi, N. Bakhary, M. Heidarzaei, S. S. Rahimian Kolor, and M. Petru, "Vehicle-assisted techniques for health monitoring of bridges," *Sensors*, vol. 20, no. 12, p. 3460, 2020.
- [19] A. Malekjafarian, R. Corbally, and W. Gong, "A review of mobile sensing of bridges using moving vehicles: progress to date, challenges and future trends," *Structures*, vol. 44, pp. 1466–1489, 2022.
- [20] K. Gkoumas, M. C. Galassi, D. L. Allaix et al., "Indirect Structural Health Monitoring (iSHM) of Transport Infrastructure in the Digital Age," *MITICA (Monitoring Transport Infrastructures with Connected and Automated Vehicles) Workshop Report*, Publications Office of the European Union, Luxembourg, 2023.
- [21] J. Q. Bu, S. S. Law, and X. Q. Zhu, "Innovative bridge condition assessment from dynamic response of a passing vehicle," *Journal of Engineering Mechanics*, vol. 132, no. 12, pp. 1372–1379, 2006.
- [22] M. Talebi-Kalaleh and Q. Mei, "A mobile sensing framework for bridge modal identification through an inverse problem solution procedure and moving-window time series models," *Sensors*, vol. 23, no. 11, p. 5154, 2023.
- [23] Y. Zhang, L. Wang, and Z. Xiang, "Damage detection by mode shape squares extracted from a passing vehicle," *Journal of Sound and Vibration*, vol. 331, no. 2, pp. 291–307, 2012.
- [24] T. J. Matarazzo and S. N. Pakzad, "Structural identification for mobile sensing with missing observations," *Journal of Engineering Mechanics*, vol. 142, no. 5, Article ID 04016021, 2016.
- [25] T. J. Matarazzo and S. N. Pakzad, "Truncated physical model for dynamic sensor networks with applications in high-resolution mobile sensing and bigdata," *Journal of Engineering Mechanics*, vol. 142, no. 5, Article ID 04016019, 2016.
- [26] T. J. Matarazzo, D. Kondor, S. Milardo et al., "Crowdsourcing bridge dynamic monitoring with smartphone vehicle trips," *Communications Engineer*, vol. 1, no. 1, p. 29, 2022.
- [27] J. Liu, S. Chen, M. Bergés et al., "Diagnosis algorithms for indirect structural health monitoring of a bridge model via dimensionality reduction," *Mechanical Systems and Signal Processing*, vol. 136, Article ID 106454, 2020.
- [28] Q. Mei and M. Gül, "A crowdsourcing-based methodology using smartphones for bridge health monitoring," *Structural Health Monitoring*, vol. 18, no. 5-6, pp. 1602–1619, 2019.
- [29] Q. Mei, "Application of compressed sensing on crowdsensing-based indirect bridge condition monitoring," in *Proceedings of the Canadian Society of Civil Engineering Annual Conference 2022*, R. Gupta, M. Sun, S. Brzev et al., Eds., pp. 123–132, Springer International Publishing, berlin, Germany, 2023.
- [30] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [31] Y. Bao, H. Li, and J. Ou, "Emerging data technology in structural health monitoring: compressive sensing technology," *Journal of Civil Structural Health Monitoring*, vol. 4, no. 2, pp. 77–90, 2014.
- [32] N. Almasri, A. Sadhu, and S. ray chaudhuri, "Toward compressed sensing of structural monitoring data using discrete cosine transform," *Journal of Computing in Civil Engineering*, vol. 34, no. 1, Article ID 04019041, 2020.
- [33] D. Jana and S. Nagarajaiah, "Full-field vibration response estimation from sparse multi-agent automatic mobile sensors using formation control algorithm," *Sensors*, vol. 23, no. 18, p. 7848, 2023.
- [34] R. E. Carrillo, A. B. Ramirez, G. R. Arce, K. E. Barner, and B. M. Sadler, "Robust compressive sensing of sparse signals: a review," *EURASIP Journal on Applied Signal Processing*, vol. 2016, no. 1, p. 108, 2016.
- [35] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, vol. 23, no. 1, pp. 90–93, 1974.
- [36] H. Bai, A. Wang, and M. Zhang, "Compressive sensing for dct image," in *2010 International Conference on Computational Aspects of Social Networks*, pp. 378–381, Taiyuan, China, September 2010.
- [37] C. Villani, *Optimal Transport: Old and New*, vol. 338, Springer-Verlag, Berlin, Germany, 2009.
- [38] K. Hetter, "10 secrets of the golden gate bridge," 2018, <https://edition.cnn.com/travel/article/golden-gate-bridge-secrets-california/index.html>.
- [39] S. N. Pakzad and G. L. Fenves, "Statistical analysis of vibration modes of a suspension bridge using spatially dense wireless sensor network," *Journal of Structural Engineering*, vol. 135, no. 7, pp. 863–872, 2009.