



Ensemble learning-based structural health monitoring of a bridge using an interferometric radar system

Ali Yaghoubzadehfard¹ · Elisa Lumantarna¹ · Nilupa Herath¹ · Massoud Sofi¹ · Mehmet Rad²

Received: 5 November 2023 / Accepted: 28 February 2024 / Published online: 30 April 2024
© Crown 2024

Abstract

Due to the increase in population, urbanisation, transportation development, and the aging of existing bridges, there is a growing need for new and rapid structural health monitoring (SHM) of bridges. To address this challenge, a method that stands out is the use of an interferometric radar system-based device, specifically Image by Interferometric Survey-Frequency for structures (IBIS-FS). Known for its portability and non-intrusive operation, IBIS-FS does not require direct contact with the bridge. This study utilised IBIS-FS to capture a pedestrian bridge's natural frequencies and mode shapes. The data obtained were found to be consistent with results from finite element models, demonstrating the reliability of IBIS-FS in capturing modal parameters. Building upon this foundation, the study then explores the application of advanced ensemble-based machine-learning techniques. By leveraging the data acquired from IBIS-FS, algorithms such as Random Forest, Gradient-boosted Decision Trees (GBDT), and Extreme Gradient Boosting (XGBoost) are used for bridge damage detection. These machine-learning (ML) techniques are suited to analyse the incomplete modal parameters of bridges, as captured by IBIS-FS. The study focuses on using these algorithms to interpret the changes in modal parameters, specifically identifying damage as a reduction in the stiffness of elements. This approach allows for a comprehensive analysis, where the modal parameters, including mode shapes and natural frequencies altered by varying noise levels, are fed as input to the models. It was observed that all three ML methods, with Random Forest in particular, can effectively identify the location and severity of damage, demonstrating an efficient training process. The robustness of GBDT and XGBoost in handling complex data sets also shows great promise for their application in bridge damage detection. Collectively, these results underscore the potential of combining advanced ML techniques like Random Forest, GBDT, and XGBoost with the data acquired from IBIS-FS.

Keywords Structural health monitoring (SHM) · Damage detection · Interferometric radar system · Modal analysis · Random forest · Gradient boosting decision tree · XGBoost

Abbreviations

BFD	Basic frequency domain	GBDT	Gradient-boosted decision trees
CMIF	Complex mode indication factor	GPS	Global position system
DAQ	Data acquisition	IBIS-FS	Image by Interferometric Survey-Frequency for structures
Dof	Degree of freedom	ML	Machine-learning
FDD	Frequency domain decomposition	RMSE	Root-mean-square error
FEM	Finite element model	OMA	Operational modal analysis
FN	False negative	PCC	Pearson correlation coefficient
FP	False positive	PP	Peak picking
FRF	Frequency response function	PSD	Power spectral density
		RF	Random forest
		SDoF	Single degree of freedom
		SF-CW	Stepped frequency-continuous wave
		SHM	Structural health monitoring
		SNR	Signal-to-noise ratio
		SVD	Singular value decomposition
		XGBoost	Extreme gradient boosting

✉ Massoud Sofi
massoud@unimelb.edu.au

¹ Department of Infrastructure Engineering, University of Melbourne, Parkville, VIC 3010, Australia

² Department of Computer Science, University of New South Wales, Sydney, NSW 2052, Australia

R^2	Coefficient of determination
$G_{yy}(\omega)$	One-sided output power spectrum
$G_{p_r p_r}(\omega)$	Auto-spectral density function
N_m	Number of modes
$p_r(t)$	Modal coordinate
R_{yy}	Output correlation function
S_{yy}	Two-sided output power spectrum
Q_r	Holds the information about the modal scale factor
[U]	Matrix of left singular vector
[y(t)]	Measured output
γ_r	Operation reference vector
λ_r	Rth eigenvalue/continuous time pole
ϕ_r	Rth unscaled mode shape
[Σ]	State covariance matrix
H	Hermitian adjoint
*	Complex conjugate

1 Introduction

An increasing number of structural systems require assessment for their safety and serviceability in Australia and around the world. There is a heavy reliance on the use of visual inspections during maintenance. However, it has a rudimentary capability in detecting damage (e.g., damage detection of reinforcement embedded in the concrete or inaccessible elements is challenging). For the past few decades, there have been extensive studies on using a structure's dynamic parameters (e.g., natural frequencies, mode shapes, and damping) to characterise damage and health assessment [1]. This further led to the development of a system called Structural Health Monitoring (SHM) which aims to monitor, inspect, and test the performance of critical structures to aid engineers in improving their safety and maintenance. SHM is a combination of two major components. They are smart sensing technology and damage detection algorithms. The smart sensing technology is used to record chemical or physical structural characteristics, and the damage detection algorithm translates the results from the smart sensors into damage characteristics and the health of the structure [2].

Generally, sensing systems used to capture structural response can be divided into non-contact and contact techniques. Accelerometers and strain gauges are conventional contact sensors commonly employed for SHM. The main limitation of accelerometers or strain gauges is the time-consuming setup process. The accelerometers need to be installed at each point where vibration occurs, and all of them need to be connected to a central computer. Moreover, it is difficult to attach these sensors to inaccessible elements of structures. GPS and laser-based systems are non-contact techniques. In recent years, GPS has been used in SHM of

bridges [3–8]. The main drawback revealed by real-world cases is the difficulty in achieving GPS positioning accuracy [9]. Also, since GPS has low frequency and due to multipath and other errors, it is difficult to obtain accurate results for a short-span bridge. A more detailed study on the usage of GPS in SHM can be found in Ref. [10]. The laser-based systems still have difficulty in measuring structures with non-optically collaborative surfaces. The influence of speckle effects and poor signal-to-noise ratio (SNR) on the performance of laser-based systems can also be considered the main limitation [11]. An experimental investigation of laser-based systems for damage detection can be found in Ref. [12].

Image by Interferometric Survey-Frequency for structures (IBIS-FS) is an alternative non-contact technique that has been used recently. This technique can be used for both static and dynamic monitoring of structures by capturing the deflection and displacement with better accuracy (0.01 mm) compared to other non-contact techniques such as GPS (accuracy of 10 mm) and image-based techniques (accuracy of 1.0–4.0 mm) [13]. Mayer et al. [14] measured the vertical and torsional displacement of a mid-span bridge under regular automobile and train traffic loads using GPS and IBIS-FS. Another study that compared GPS and IBIS-FS was conducted by Sofi et al. [15]. They employed both instruments to obtain modal parameters of two high-rise buildings (77-story and 50-story buildings), and both cases demonstrated better accuracy of IBIS-FS compared to GPS. Furthermore, IBIS-FS is designed based on interferometric radar techniques with the significant advantage of remotely monitoring multiple points simultaneously. Few tests have verified the in-situ performance and accuracy of IBIS-FS on some full-scale bridges. In 2008, Pieraccini et al. [16] conducted an experimental modal analysis on a cable-stayed pedestrian bridge to compare displacement measurements obtained from radar and accelerometers. Results agreed with a deviation of 0.122 mm [17, 18]. Interferometric radar was used to obtain the modal parameters of a bridge in Refs. [17] and [18]. Comparison of results from IBIS-FS with conventional accelerometers shows remarkable consistency. Sárácin et al. measured the horizontal displacement of a ten-story building and the vertical displacement of cable-stayed [19] and a metal truss bridge [20]. They demonstrated the measuring principle and the optimal choice of observation point when using IBIS-FS. Silva et al. [21] used IBIS-FS to obtain modal parameters of the telecommunication mast, and Quirkel and Barrias [22] used IBIS-FS to validate the FE model of a railway bridge under real-time tram load. Recently, IBIS-FS has been used to monitor masonry structures [23] and obtain tensile force in tie-rods [24]. Castellano et al. [25] used two synchronised IBIS-FS to measure three-dimensional displacement vectors and mode shapes of a masonry bell tower. All studies emphasise ease of installation, low

cost, and significant potential usage of IBIS-FS compared to other (contact or non-contact) sensing techniques in SHM.

As previously mentioned, the second component of SHM is the damage detection procedure. Modal-based damage detection has been used widely in the last decades. It is based on the principle that modal parameters (e.g., frequencies, mode shapes, and damping) depend on mass and stiffness [26]. Therefore, any damage to the structure that changes the stiffness and/or mass would change the modal parameters of the structure [27]. Depending on the location and severity of the damage, this change would be either negligible or severe. In recent decades, pattern recognition has been used to evaluate the relationship between these changes and modal parameters to detect damage in bridges [28]. Pattern recognition is usually applied by machine-learning (ML) algorithms. A notable advancement in ML for SHM is the application of ensemble learning techniques, which have garnered attention for their ability to enhance prediction accuracy by combining multiple models, known as weak learners, into a more robust strong learner. Ensemble learning is particularly favored in engineering prediction problems due to its robustness and improved performance over individual models. Among the various ensemble learning methods, Random Forest (RF), ExtraTree, AdaBoost, Gradient-boosted decision trees (GBDT), Extreme Gradient Boosting (XGBoost), and LightGBM have become widely used. RF and ExtraTree, for instance, employ bagging, a technique that generates diverse training subsets through random sampling, trains weak learners on these subsets, and then integrates their outputs. On the other hand, AdaBoost, GBDT, XGBoost, and LightGBM use boosting, which sequentially trains weak learners, each time focusing more on the samples that previous learners misclassified. By combining these learners in a weighted manner, boosting effectively enhances the overall predictive performance of the model.

Although a considerable amount of literature has been published on ML-based damage detection, there are only a few studies on ML-based damage detection using ensemble learning-based models. In 2013, Zhou et al. [29] proposed an RF classifier integrated with a data fusion technique to detect multiple damages on an eight-story steel shear frame benchmark. A year after, Zhou et al. [30] proposed a two-stage RF-based feature elimination method for damage detection to increase the accuracy of outputs and reduce the detection time cost. More recently, Lu et al. [31] implemented ML-based regression methods, including Support Vector Regressor (SVR), Artificial Neural Network, Regression Tree, and RF to predict the natural frequencies of the Tamar bridge considering environmental factors. Results show SVR and RF outperform other methods. Chencho et al. [32] used RF to locate and quantify single and multiple damages in a simple beam and scaled steel structure using noisy acceleration data obtained

from the numerical and experimental model. For historical structures, Chaiyasarn et al. [33] used the integrated Convolutional Neural Network method with RF and Support Vector Machine classifiers to detect cracks. Huang et al. [34] used the cross-correlation function and wavelet packet decomposition to extract features from raw data obtained from the ASCE benchmark to demonstrate the effectiveness of ensemble-based ML methods. The results indicate that RF and XGBoost outperform SVM. Furthermore, RF shows slightly better performance compared to XGBoost. In 2021, Garg et al. [35] used XGBoost, Decision Tree (DT), RF, K-Nearest Neighbors (KNN), and Logistic regression models for the condition assessment of bridges. They used four features for classification and results indicate DT and RF have the highest accuracy among other methods. In another study, Ref. [36] compared ensemble learning-based models such as XGBoost, GBDT, Adaboost, LightGBM, and Extratrees in bridge deck monitoring. In their case study, XGBoost outperforms others, while GBDT has similar results to RF with slightly better than RF. A comprehensive study on ML-based damage detection can be found in Refs. [28, 37–39].

The structure of this study is divided into eight sections. Initially, the paper provides an introduction and reviews recent studies in the IBIS-FS and the ensemble-learning-based damage detection method. It will then go on to the methodology section to describe the conduction of the study. Section 3 describes the field measurement setups and looks at how experimental data are obtained from the bridge. Section 4 concerns the operational modal analysis techniques used for this study. Section 5 presents the finite element modeling of the structure and describes how data are generated to train the damage detection method. The procedures for damage detection are laid out in Sect. 6. Lastly, Sect. 7 displays the results from the operational modal analysis, finite element modal analysis, and the damage detection process.

2 Methodology

In this study, an IBIS-FS was used for field measurement of the pedestrian bridge located in Melbourne, Australia. After pre-processing data, operational modal analysis (OMA) techniques, including the frequency domain decomposition (FDD) and the peak-picking (PP) methods, were used to obtain the modal parameters of the bridge. Then a finite element model (FEM) of the bridge was developed using SAP2000 (SAP2000 2015). The modal results from the finite element analysis were then compared with experimental results. For damage detection, data obtained from the FEM of the bridge in SAP2000 was used to train damage detection algorithms to detect pre-defined damage scenarios in the bridge with different conditions (noise-free, contaminated

by noise, and incomplete modal parameters). SAP2000 API Python was used to build a data frame. The flowchart of the study is presented in Fig. 1.

3 Field measurements

3.1 Image by Interferometric Survey-Frequency for structures (IBIS-FS)

IBIS-FS (Image by Interferometric Survey-Frequency for Structures) is an innovative interferometric radar system designed by the IDS GeoRadar company for static and dynamic monitoring of structural displacement or deformations. It can be used to remotely measure the displacement of multiple points simultaneously at different locations on a structure with an accuracy of 0.01 mm. The resonant frequencies and mode shapes can then be identified from the static deflections and vibrations [14]. IBIS-FS comprises PC, Tripod, Radar head, and Antennas (Fig. 2a). The sensor unit is a coherent radar used for electromagnetic signal transmission and reception. Electromagnetic signals generated by the radar will be reflected from each discontinuity of the structure back to the sensor [13]. The received electromagnetic signals can then be analysed to compute the displacement time histories of the selected measurement points along the structure [14]. The operational characteristics of IBIS-FS are summarised in Table 1.

The IBIS-FS radar has implemented two radar techniques, Stepped Frequency-Continuous Wave (SF-CW) and differential interferometry technique, to simultaneously measure the displacement of a series of points at different positions. SF-CW is a technique employed to distinguish the positions of different targets along the radar's line of sight. By linearly

increasing the frequency of a burst of N single tones in discrete steps, large bandwidth, and range resolution of 0.5 m can be achieved [13]. This means two targets can still be detected individually if the distance between them is greater than 0.5 m. A 1-D range profile of the radar echoes with range resolution can be plotted by computing the magnitude of each bin of obtained vector samples (Fig. 2b). The differential interferometry technique can then be implemented to evaluate the displacement of each target by comparing phase information of reflected electromagnetic waves received by the sensor at different times [41].

3.2 Pedestrian bridge description

The pedestrian bridge was constructed in 1976. This bridge is located at the University of Melbourne, Australia, connecting two university buildings (the David Caro building and the Earth Science building) over Swanston Street. The width of the bridge is 2.8 m, and the bridge's center span is approximately 22 m, with the highest point about 5.3 m above the ground. The bridge was constructed from pre-cast I profile steel core concrete beams supporting an in-situ concrete deck. The beams cantilever out from two piers, supporting a gently arched middle span [18].

3.3 Field measurement setup and loading condition

The measurements were conducted from the side of Swanston Street at 11:00 pm to minimise the impact of traffic on the IBIS-FS data collection. The radar beam emitted from the sensor module was configured to cover the middle span between the two bridge joints, as shown in Fig. 2 (c). To achieve this, geometry parameters (e.g., structural length, the exact location of the equipment relative to the bridge,

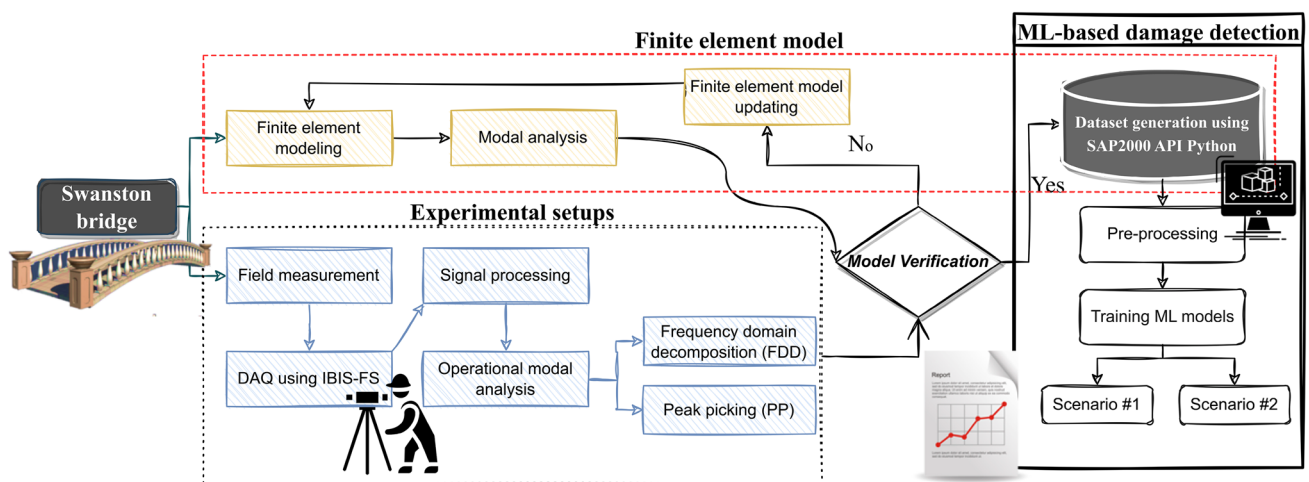


Fig. 1 Flowchart of the methodology

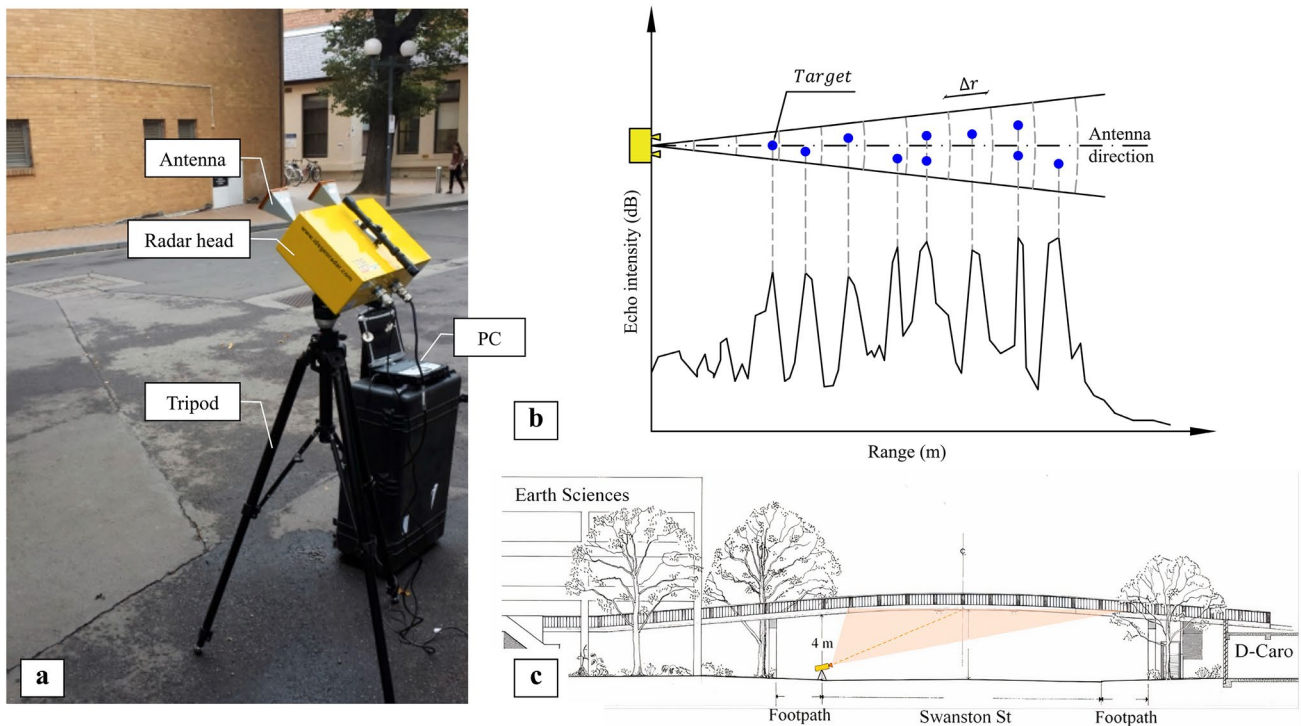


Fig. 2 a IBIS-FS schematic setup, b Idealised range profile, and c Elevation view of the pedestrian bridge from the north and the IBIS-FS setup [18]

Table 1 IBIS-FS operational characteristics

Technical specifications	Value
Sampling frequency	200 Hz
Maximum operational distance	1000 m
Range resolution	0.05 m
Operational weather condition	All
Imaging capabilities	One dimension
Displacement sensitivity	0.01

Table 2 Geometry configurations of IBIS-FS on the pedestrian bridge [18]

IBIS-FS geometry configuration	Pedestrian bridge
Structural length	21.2 m
Relative horizontal location	0 m
Vertical distance to the bridge	4 m
Vertical tilt	20

and vertical tilt angle of the sensor module) as summarised in Table 2 were inputted into the software. Even though the measurements were taken late at night, there were still several trams, cars, and pedestrian traffic. The tram line vibration and movement of trams, cars, and pedestrians can induce interference in the measurements. Furthermore, the electric lines of the trams and other surrounding structures rather than the bridge may be detected by IBIS-FS. These factors require consideration when undertaking the measurements to ensure the accuracy of results.

For this study, the IBIS-FS was located on the footpath near one of the bridge joints. The radar beam range is limited to the bridge’s middle span to avoid capturing other interferences and noise to obtain good-quality results. The bridge was excited dynamically by pedestrians jumping up and down at

the mid-span of the bridge. The speed of jumping was set to be three times per second as an initial analysis of the bridge shows a fundamental frequency of 3.2 Hz. The frequency of the manual excitation was set to excite the first mode of vibration. For each case, the measurements have been performed repeatedly (3–4 times) to improve the reliability and accuracy of the results. In this way, the radial displacement history of each measurement point on the bridge, as well as the natural frequencies of the bridge, can be captured and recorded by the IBIS-FS device.

4 The operational modal analysis

The experimental modal analysis aims to identify modal parameters from measurements of the controlled applied force and the vibration responses obtained from a structure. Heavy and expensive devices must be used to apply controlled force to a structure to receive measurable responses. To overcome this challenge, the study focuses on the opportunity that the operational modal analysis (OMA) provides. OMA techniques estimate the modal parameters of a structure only from the measurement. Artificial excitation is not needed in these techniques, and vibration responses are obtained under ambient forces and operational loads (e.g., traffic and wind) [42]. In this study, peak-picking and frequency domain decomposition methods are used to obtain modal parameters.

4.1 The peak-picking method

The peak-picking (PP) method is the simple and fast way for an output-only system identification system, commonly referred to as the Basic Frequency Domain (BFD) [42]. The PP technique can better approximate the natural frequency of systems when a structure has minimal damping and its modes are well differentiated within a particular bandwidth. Otherwise, it may yield erroneous results by identifying operational deflection shapes instead of individual mode shapes. Nevertheless, the simplicity and computational efficiency of the method, make it a valuable tool for analysts, particularly during field tests, to quickly assess the effectiveness of measurements and obtain preliminary insights into dynamic identification results. Furthermore, the PP technique can only detect unscaled mode shapes since the input is unknown [42].

As the term “peak-picking” indicates, in this method, the modes are detected by picking the peaks in the power spectrum energy (PSD) plots [42]. For the theoretical background of this method, consider partial fraction expansion of pole-residue form of two-sided PSD of the frequency response function (FRF) matrix (refer to Ref. [42] for more details):

$$[S_{yy}(\omega)] = \sum_{r=1}^{N_m} \frac{\{\phi_r\}\{\gamma_r\}^T}{i\omega - \lambda_r} + \frac{\{\phi_r\}^*\{\gamma_r\}^H}{i\omega - \lambda_r^*} + \frac{\{\gamma_r\}\{\phi_r\}^T}{-i\omega - \lambda_r} + \frac{\{\gamma_r\}^*\{\phi_r\}^H}{-i\omega - \lambda_r^*}, \quad (1)$$

where N_m is the number of modes, $\{\phi_r\}$ is the mode shape, the λ_r keeps information about natural frequencies and damping ratio, and $\{\gamma_r\}$ is the operational reference vector associated with the r th mode corresponding to the modal

participation vector in the pole-residue form of the FRF matrix.

In the PP technique it is assumed that one dominant mode around a resonance. Therefore, Eq. (1) can be approximated by the contribution of the dominant mode only (say r^{th} mode). Thus, the structural response is approximately equal to the modal response [42].

$$\{y(t)\} \approx \{\phi_r\}p_r(t), \quad (2)$$

where $\{\phi_r\}$ is the modal matrix and $p_r(t)$ is the modal coordinate related to the r^{th} mode. From Eq. (2), the correlation functions of the responses can be computed:

$$[R_{yy}(\tau)] = E[\{y(t+\tau)\}\{y(t)\}^T] = R_{p_r p_r}(\tau)\{\phi_r\}\{\phi_r\}^T, \quad (3)$$

where

$$R_{p_r p_r}(\tau) = E[\{p_r(t+\tau)\}p_r(t)]. \quad (4)$$

Equation (4) is the modal auto-correlation function. Taking the Fourier transform of Eq. (3), the PSD matrix can be obtained:

$$[G_{yy}(\omega)] = G_{p_r p_r}(\omega)\{\phi_r\}\{\phi_r\}^H, \quad (5)$$

where $G_{p_r p_r}(\omega)$ is the auto spectral density function of the modal coordinate. The PSD matrix of measure outputs must be computed to obtain modal parameters. Modes are represented by the trace of the PSD matrix at each discrete frequency and associated unscaled mode shape is one of the columns of the PSD matrix.

4.2 The frequency domain decomposition method

The frequency domain decomposition (FDD) technique has overcome the limitation of the PP method concerning the identification of closely spaced modes and removing subjective task in picking the peaks. This method was originally applied to FRFs and known as Complex Mode Indicator Function (CMIF) to highlight its ability to detect multiple roots and, therefore, the possibility to count the number of dominant modes at a given frequency [42]. Considering the single value decomposition (SVD) of the Hermitian and positive definite PSD matrix at a specific frequency ω leads to the following matrix (refer Ref. [42] for more details):

$$[G_{yy}(\omega)] = [U]\left[\sum\right][U]^H, \quad (6)$$

where $[U]$ is the unitary matrix including singular vectors and $[\sum]$ is the diagonal matrix including singular values. To obtain modal parameters, after computing PSD of measured outputs, the SVD of the PSD at each frequency calculated.

Singular values are modes and the corresponding singular vectors are mode shapes.

Assumptions and requirements specific to the FDD method are duly addressed. Due to visual inspection, the structure is linear and there is not any significant damage or change in the boundary condition of the structure. Regarding frequency resolution, a critical aspect of the methodology involves selecting a resolution of 0.025, balancing the need to capture fine details in the frequency domain with avoiding excessive computational load. This choice is informed by duration of displacement time series and a sampling rate of 200 Hz, resulting in a Nyquist frequency of 100 Hz. The time series data are segmented into non-overlapping segments, each representing a meaningful portion of the structural response. The selection of segment length considers dynamic changes while maintaining adequate frequency resolution. In terms of window function selection, the study opts for the rectangular window. This decision is based on the analysis objectives. The rectangular window is chosen for its computational efficiency and minimal distortion to the signal, aligning well with the study's focus on a linear and undamaged structure. While other window functions may offer improved frequency resolution, they often introduce more spectral leakage, which is deemed unnecessary for the precision required in representing frequencies in the context of a linear and undamaged structure.

5 The finite element model of the bridge

The IBIS-FS results obtained from dynamic excitations were verified using a numerical model of the pedestrian bridge. Using the SAP2000 a model of the pedestrian bridge was developed and subjected to a dynamic frequency analysis to calculate the structure's natural frequencies and mode shapes.

The calculation of the modal parameters using SAP2000 is highly dependent on the geometry, stiffness, and mass

of the structure. The dimensions of the bridge were measured on-site and verified with the architectural drawings. These parameters formed the basis of the structural model. However, some distinct features need to be appropriately integrated into the model. The first features are the joints that exist between the two piers. It was noted that a section of the beam cantilever from both piers seats and supports the bridge's middle span (Fig. 3). Both joints were simplified and represented as a hinge joint within the model. Another feature is the shallow arc of the bridge, as seen in Fig. 3, which was incorporated into the model using height measurements below different points of the bridge. The main factor influencing the mass and stiffness of the bridge is the material and composition of the structure. The bridge consists of pre-cast concrete beams that support the in-situ reinforced concrete deck [18] and the steel box handrails (Fig. 3). In this study, the finite element (FE) modeling approach adopts the assumption that the structure operates within a linear and serviceability state, aligning with the study's primary objective of analysing the structural response within elastic material limits. This simplification facilitates the modeling process, although it may overlook potential non-linearities that could arise under extreme loading conditions or in the presence of damage. The assumption of a perfect bond between steel and concrete is made, acknowledging that the modeling might not encompass all the intricacies of the composite beam section.

The model from SAP 2000 is presented in Fig. 4. The FEM of the bridge in SAP2000 was used as a baseline model for damage detection and sensitivity analysis. Damage was then applied to elements by reducing the stiffness of the elements (e.g., 10% damage in element e_i is defined as 10% reduction in elastic moduli (E_i) of e_i). After defining damage for each element, modal analysis was performed in SAP2000 using SAP2000 API for Python to generate and collect data.

The FE model of the bridge consists of 126 elements, including main beams and handrails; generating such a data

Fig. 3 Elevation from North of the bridge showing left pier and cantilever joint

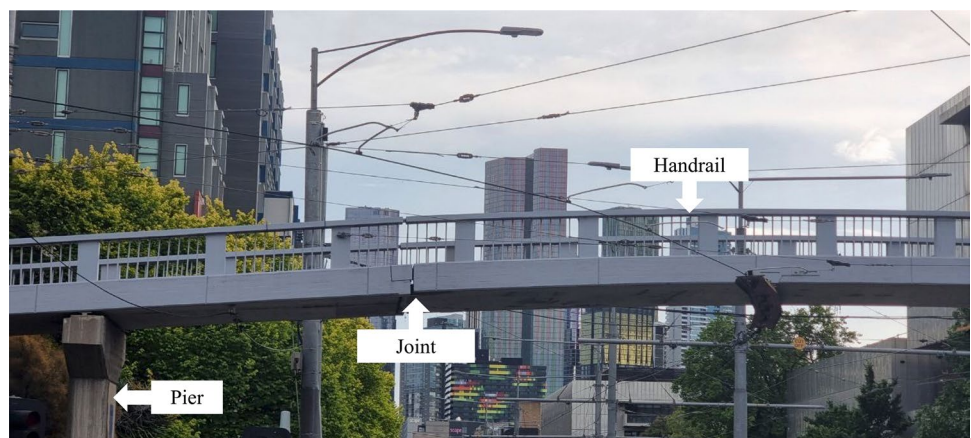
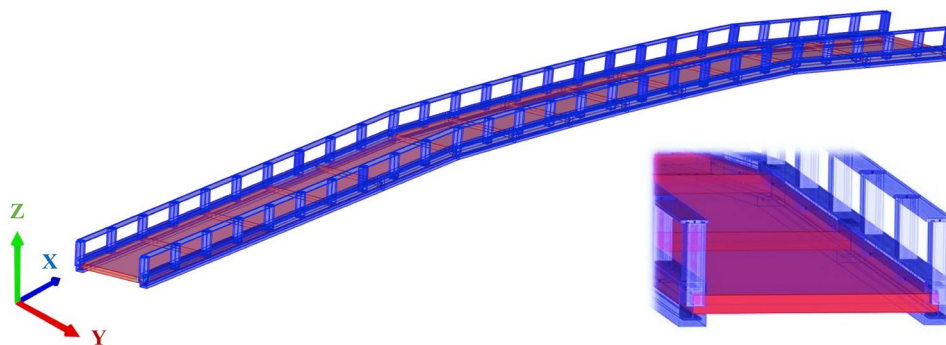


Fig. 4 FEM of the bridge

frame with many classification classes is time-consuming. Furthermore, from an ML perspective, the increase in the number of classification classes tends to reduce the performance of the model. Therefore, sensitivity analysis is conducted to reduce the number of elements for damage detection. For this purpose, the sensitivity of the first five natural frequencies to the change in stiffness of each element is evaluated. Due to the symmetry, only 63 elements are used for the sensitivity analysis.

In Fig. 5, the results of the sensitivity analysis are presented. According to the results, natural frequencies are more sensitive to the change in the stiffness of main beams since it ranges between -19% and 14.8% . In contrast, vertical and horizontal elements of handrails have a neglectable effect on natural frequencies compared to main beams. The sensitivity of the first five natural frequencies to horizontal elements ranges between -0.16% and 0.04 and for vertical elements ranges -1.2% and 0.2 . Furthermore, significant damage in a handrail can be detected via visual inspection. Therefore, in this study, handrails are not used for damage detection and main composite beams were selected (20 elements) for data generation. The damage production procedure is shown in Fig. 6.

6 Machine-learning-based damage detection

In the study presented, damage detection is approached through a two-part process: damage localisation and assessment of damage extent. As depicted in Fig. 7, two machine-learning problems must be addressed: classification to locate the damage and regression to predict the severity of the damage. Initially, data are processed using a classification algorithm to identify the location of the damage. Subsequently, a regression algorithm evaluates the extent of the damage. For both tasks, ensemble learning-based ML algorithms are employed. These ensemble models combine various individual models, commonly weak learners, such as decision trees, to form a more accurate and robust predictive model. The two popular methods of ensemble learning are

boosting and bagging. With the bagging method, individual models are trained in parallel and independently, while with the boosting method, models are trained sequentially, as shown in Figs. 8 and 9, respectively. This study specifically employs RF, GBDT, and XGBoost for both classification and regression in detecting damage. While RF and GBDT are implemented using the scikit-learn package in Python, XGBoost necessitates its own library namely as XGBoost's Python Package.

6.1 Random forest

A Random Forest (RF) is a supervised ML technique that is used to solve regression and classification problems. It uses ensemble learning, a technique that combines many classifiers to improve the performance of a model and reduce the variance of the results. As the term “forest” indicates, RF consists of many decision trees. Every decision tree consists of decision nodes, leaf nodes, and a root node. In a Random Forest, each decision tree is trained using a different data set that is substituted from the actual data set. This approach is referred to as bagging or bootstrapping. Bagging, also known as bootstrap aggregation, is a common ensemble learning method which integrates predictions from various models even when the predictions are not correlated. Each tree draws a collection of features at random from the entire feature set in the training data set, making them uncorrelated [32]. The leaf node of each tree is the final output produced by that specific decision tree. The final output of the RF system is chosen by the majority of the decision trees, as shown in Fig. 8.

6.2 Gradient boosting decision tree

A Gradient Boosting Decision Tree (GBDT) is a supervised learning technique have been employed for both regression and classification tasks. It constructs a predictive model in the form of an ensemble of decision trees, which are combined to yield a more accurate and robust final model. In the GBDT framework, each tree is built sequentially, with subsequent trees focusing on addressing the errors or residuals

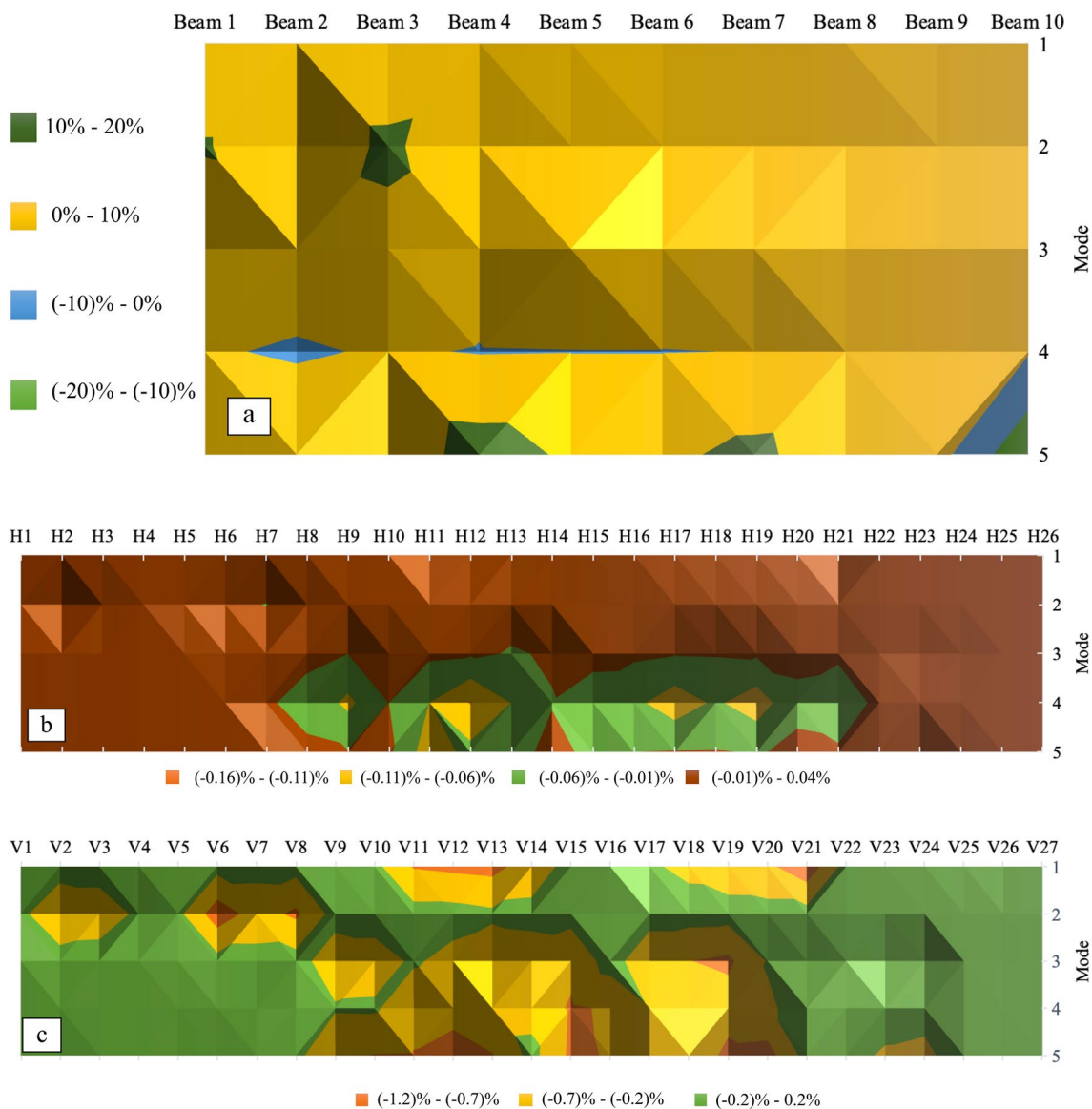


Fig. 5 Sensitivity of natural frequencies to the stiffness of (a) main beams, (b) horizontal elements of handrails, and (c) vertical elements of handrails

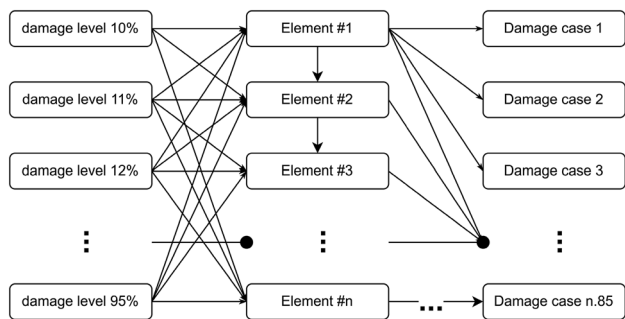


Fig. 6 Damage cases production procedure

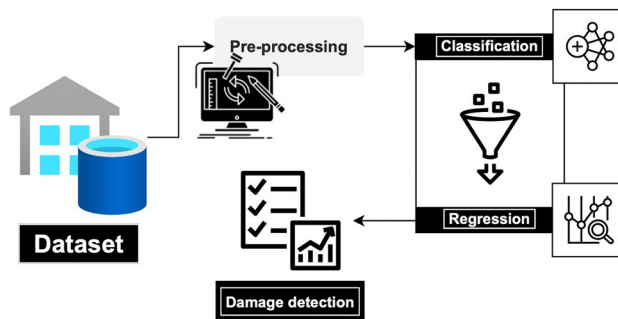


Fig. 7 Damage detection procedure

Fig. 8 Classification with Random Forest with bagging learning technique

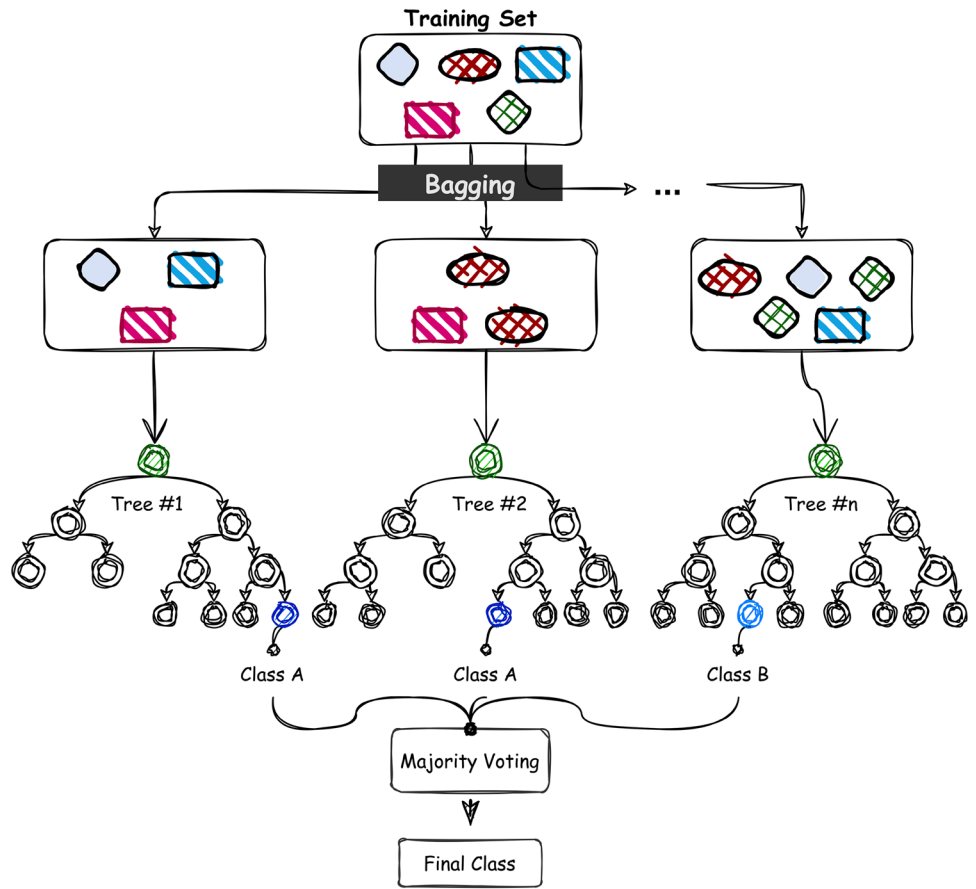
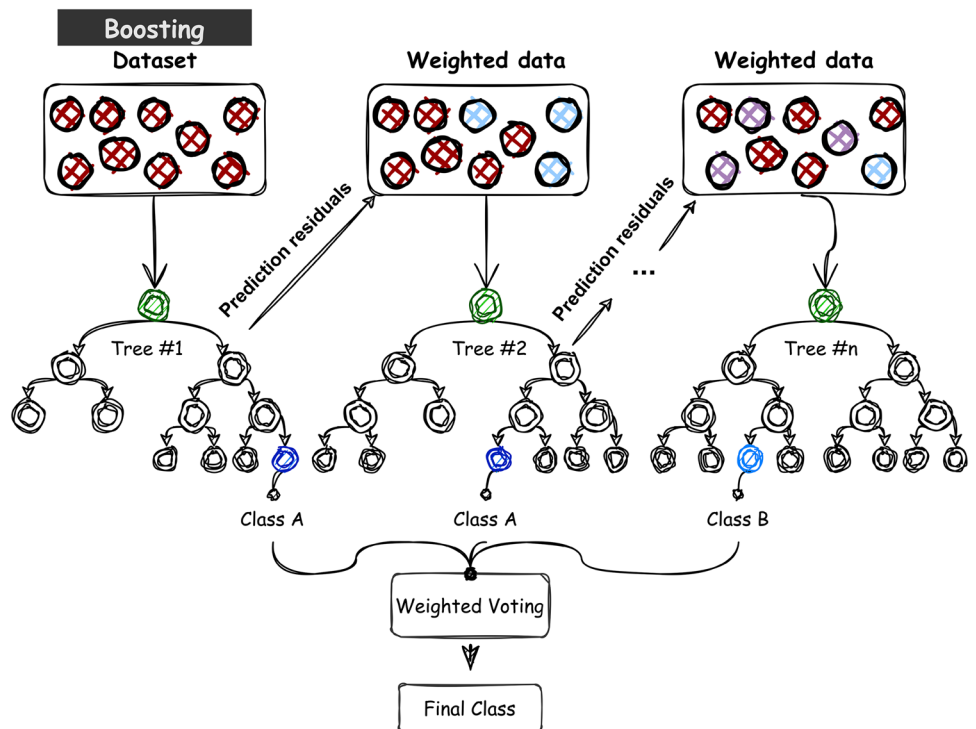


Fig. 9 Classification with GBDT with boosting learning technique



identified by their predecessors. This method, known as boosting, effectively reduces the bias and variance of the collective model. Specifically, GBDT iteratively minimises the objective function by applying gradient descent to the residuals (the discrepancies between predicted and true values). The optimisation process within GBDT involves navigating the gradient of the loss function to enhance model predictions progressively. The construction and refinement of the model via GBDT are shown in Fig. 9.

6.3 Extreme gradient boosting

Extreme Gradient Boosting (XGBoost) is an advanced ensemble ML algorithm renowned for its efficacy in both regression and classification challenges. As a supervised learning method, XGBoost refines its model by constructing a series of decision trees in a sequential manner, where each subsequent tree is crafted to correct the residuals, thereby enhancing the model's precision incrementally. This process embodies the principle of boosting; however, XGBoost distinguishes itself through the introduction of regularisation terms in its objective function, which mitigate overfitting more effectively than traditional GBDT. XGBoost augments the traditional GBDT approach by optimising its cost function with both L1 and L2 regularisation terms, a technique not typically applied in standard GBDT implementations. Additionally, XGBoost employs a novel tree-building algorithm that utilises second-order gradients, allowing for more accurate split decisions in the trees. The optimisation of the XGBoost model employs both gradient descent and a custom regularisation strategy to converge to a robust solution. This optimisation ensures efficient handling of sparse data and integrates mechanisms for parallel processing, furthering its computational advantages.

6.4 Grid search cross-validation

In machine learning, it is standard practice to divide data sets into training, validation, and test sets. However, in scenarios where the availability of sampled data is limited, the validation set is sometimes either omitted or may not be adequately representative. Furthermore, memorising or overfitting is a common problem in ML algorithms. As previously stated, the RF algorithm determines the result based on the decision trees' predictions. The average or mean output of each individual tree is obtained when the regression problem is solved with RF. This procedure prevents overfitting. Similarly, XGBoost, a powerful and efficient gradient boosting framework, also demands hyperparameter tuning to manage its complexity and to avoid overfitting. It is because XGBoost builds trees sequentially, and without proper optimisation, it can quickly fit to noise in the training data. Consequently,

Table 3 Hyperparameters considered for the RF model

Hyperparameters	Value range
n_estimators	50, 100, 150, 200
max_depth	5, 10, 15, 20
max_features	"none", "sqrt", "log2"

Table 4 Hyperparameters considered for the GBDT and XGBoost models

GBDT		XGboost	
Hyperparameters	Value range	Hyperparameters	Value range
n_estimators	100, 200, 300	n_estimators	200, 400, 500
max_depth	3, 5, 7, 9	max_depth	3, 5, 7, 9
learning_rate	0.01, 0.1, 0.2	learning_rate	0.01, 0.1, 0.2
min_samples_split	2, 5, 10	min_child_weight	1, 3, 5
min_samples_leaf	1, 2, 4	gamma	0, 0.1, 0.2

choosing optimal hyperparameters based on the training set becomes crucial for the ML model's performance [43].

The selection of optimal hyperparameters based on the training set is essential for the robust performance of ML models, particularly when it comes to preventing overfitting. ML models often involve numerous hyperparameters, and manually adjusting these parameters in isolation can be a repetitive and time-consuming task. To address this issue, the current study utilises grid search cross-validation to efficiently identify the best hyperparameters for the RF, GBDT, and XGBoost models. The specific hyperparameters under consideration for each model are outlined in Table 3 for RF and Table 4 for GBDT and XGBoost. By automating the search for the optimal hyperparameter combinations, the study enhances the models' predictive accuracy and generalisability, ultimately aiding in the effective detection of damage severity and location.

Initially, the study explores n hyperparameters with individually assigned m values. Next, the grid search method systematically enumerates all hyperparameter combinations, resulting in mn combinations. For each combination, the hyperparameters are used to train models with the K-fold cross-validation method with $k=5$. Therefore, the training data set is divided into five distinct subgroups at random. Four of these subgroups are then designated as the training set, with the fifth subgroup serving as the test set. This process is repeated five times, with each subgroup being rotated in as the test set in successive iterations. For each model, incorporating various hyperparameter combinations, the metric, I_m is calculated to assess its performance.

The choice of I_m for regression is root-mean-square error (RMSE). When the minimum RMSE value is determined, the set of hyperparameters associated with this value represents

the optimal combination. For the classification, the choice of I_m depends on the specific requirements and characteristics of the task. If missing any damaged element could lead to catastrophic outcomes, recall might be the most important metric. If false alarms are very costly and need to be minimised, precision would be more important. In this study, I_m for classification is precision. Therefore, when the highest precision value is achieved, the set of hyperparameters linked to this maximum value is recognised as the optimal combination. These steps lead to the determination of the optimal hyperparameters for each model using the training data set. Once established, the model is then applied to the test data set to accurately identify the severity and location of the damage.

6.5 Model evaluation metrics for machine learning

A common tool to evaluate the performance of classification is a confusion matrix, which separates the method results into four categories: True Negative (TN), False Negative (FN), False Positive (FP), and True Positive (TP). For damage detection analogy, the meaning of each term is as follows:

- TP: model correctly detects damaged elements,
- FP: model detects damage in element, but there is not,
- TN: model correctly detects intact elements,
- FN: model considers there is no damage in the element, but there is.

FP and FN are more important from the SHM perspective. FP increases the number of elements that will examine in further investigations (e.g., local damage detection), which is time-consuming and costly. Also, FN is the most important and dangerous since the undergo damage is not detected in the structure. The main goal of classification methods is to reduce the value of FP and FN while increasing the value of TP and TN. Accuracy, Precision, Recall, and F_1 scores were generated from the confusion matrix to help to assess performance. These scores are defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (7)$$

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{Actual\ results}, \quad (8)$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{Predicted\ results}, \quad (9)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

For performance evaluation, the precision and recall parameters are more appropriate since they show how well

the model can locate the damaged element in the data set. Both parameters may be taken into account simultaneously with equal weights using the F_1 . Ideally, the higher the F_1 parameter, the better the damage indicator is in detecting damage in the data set [44]. Furthermore, it is useful for comparing two models, especially in damage detection when both FP and FN are important. However, the choice between models depends on the specific requirements of the task and computation time.

The most undemanding metrics for regression models are RMSE, coefficient of determination (R^2), and Pearson Correlation Coefficient (PCC). If all predictions matched the expected values, the RMSE values tend to be zero. The R^2 indicates the strength of the model and evaluate the scatter of the data around the fitted regression line. It ranges from 0 to 1, the lower R^2 represents larger differences between the actual data and the predicted values (e.g., an $R^2 = 0.4$ indicates that 40% of the variance in the outcome data can be explained by the model). Regression models with higher R^2 is well-fitted models. The PCC measures the correlation between variables and ranges from -1 to 1 . It is square root of coefficient of determination. The sign of the PCC shows the direction of the relationship between variables. PCC values further way from zero represent strong linear relationship between variables (e.g., $PCC = -0.9$ indicates that there is a strong negative relationship between variables).

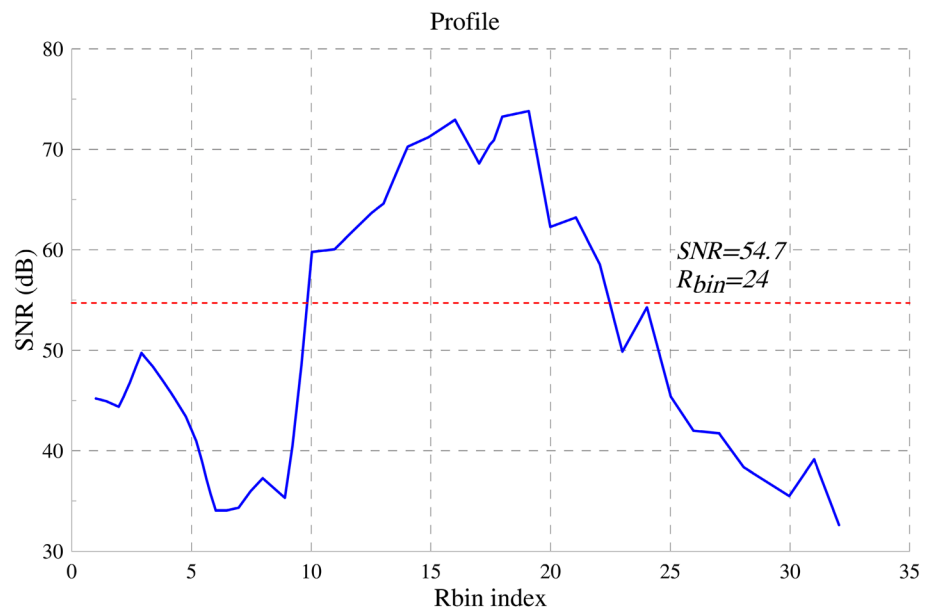
7 Results and comparison

7.1 Modal analysis

7.1.1 Operational modal analysis

The IBIS-FS performs the simultaneous monitoring of several different time series, which correspond to different reflections in the bridge. These different time series are called Range Bins (R_{bins}). It also provides a profile with the Signal-to-Noise Ratio (SNR) for each R_{bin} to select the ones with the highest SNRs corresponding to the range of the bridge. Figure 10 shows the SNR profile of the R_{bins} captured within the mid-span of the bridge [45]. The x-axis shows the R_{bin} index. The numbers correspond to the structural distance within the span relative to the starting position of IBIS-FS. The y-axis shows the SNR of a specific R_{bin} . The R_{bins} selected for analysis within each experimental measurement data were selected based on the profile shown in Fig. 10. As a criterion, only those R_{bins} with a SNR value of 55 or more, as shown by the red dashed line, were selected to ensure that the displacement accuracy captured by IBIS was within 0.01 mm. Initially, the R_{bins} with sharp peaks and troughs above the 55 dB SNR threshold were selected,

Fig. 10 Diagram showing the SNR of each Robin within the whole mid-span of the bridge



but then they were later refined with further investigation of their quality using the polar graphs.

The following polar graphs in Fig. 11 were obtained from the selected R_{bins} . The radial axis denotes the amplitude of the signal received, while the angular axis denotes the phase angle. The polar graphs in Fig. 11 indicate the quality of the data recorded. An expected good result would be for a R_{bin} selection to follow the same radial coordinate, or amplitude, for a wide range of the angular coordinates. This is best shown by R_{bin} 16 in Fig. 11a. However, R_{bin} 10 in Fig. 11a, d indicates poor data as there is only a small angular range along the amplitude coordinate. Deviations in Fig. 11 of the amplitude along the angular coordinate are noise in the data and can create unfavourable results. The difference in the quality of data between data sets 35 and 22 (Fig. 11a, d) could have been the result of vehicle, tram, or pedestrian movement during the data collection.

The following frequency displacement graphs obtained from the IBIS-FS raw data sets are shown in Figs. 12, 13, 14, 15. The frequency location of the peaks on these graphs represents a mode and can be compared with the frequencies of the modes from the FE model. The frequency displacement graphs show that the IBIS-FS radar can reasonably detect a structure's modal frequencies. The natural frequency determined by data set 13 was 3.68 Hz, while all other data sets gave a natural frequency of 3.76 Hz. This is between 0.53% and 2.65% difference from the natural frequency found with the FE model. Explicit modes could also be found at approximately 8 Hz for data sets 13, 35, and 44 and at approximately 13.4 Hz for data sets 35 and 44. These correspond to mode shapes 4 and 8 as determined by FEM and only vary between 0.25% and 1.11%. It can be noted

that the better-quality data in data set 35 produced peaks at more modal frequencies than the lower-quality data set 22.

The other non-torsional mode, mode 3 as determined by FEM, should have appeared at approximately 7 Hz; however, due to its proximity to mode shape 4 there was some interference, and the peak-picking method was not able to detect the mode accurately.

7.1.2 Numerical modal analysis

Using measurements obtained on the bridge and careful investigation of the structural plans, a finite element model of the pedestrian bridge was constructed in SAP2000. As indicated in the structural plans, the characteristic strength of concrete used in all structural elements is Grade 25, and all structural steel components are mild steel for this model. The dynamic frequency analysis was performed for the first 20 dynamics modes within a tolerance of 0.001 Hz. The resulting natural frequencies and natural period are summarised in Table 5, and some of the corresponding mode shapes identified are given in Fig. 16. Comparisons between the experimental and numerical model (see Table 5) indicated that the developed SAP2000 models can simulate the FEM with good accuracy and can be used as baseline FEM for damage detection procedures.

The following Figs. 17–19 were obtained by plotting the displacement data sets extracted from numerical modal analysis and IBIS-FS onto the same graph for each mode. The data sets obtained from FEM are plotted as smooth solid lines and are used as the reference mode shape in ideal conditions. All the data sets obtained from IBIS-FS are plotted as scattered points so that the accuracy of IBIS-FS can be

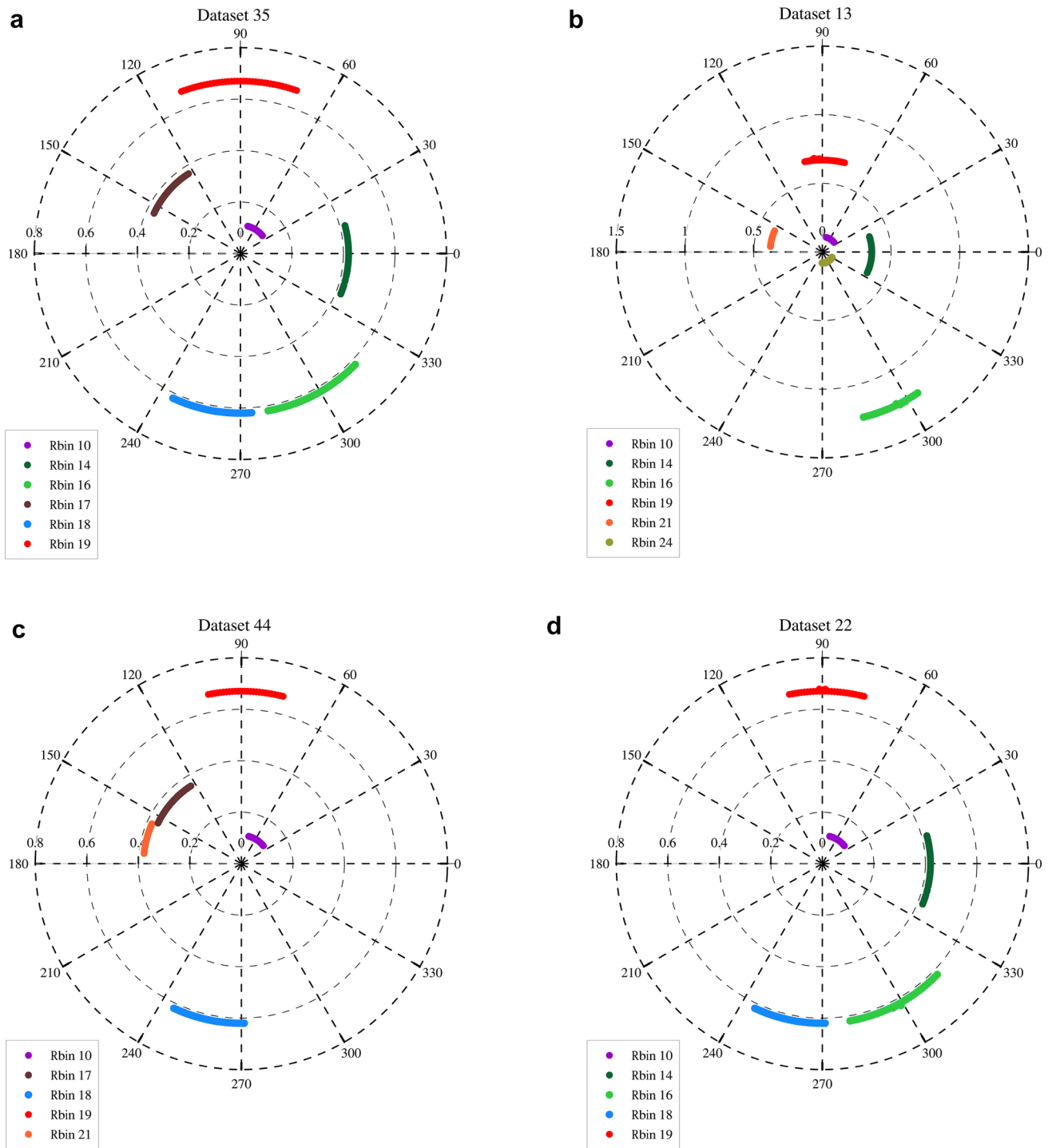


Fig. 11 Polar graphs of (a) data set 35, (b) data set 13, (c) data set 44, and (d) data set 22

validated by observing whether the data points are well scattered around the finite element mode shapes.

The mode shape plots for both FEM and IBIS-FS result illustrated in Figs. 17–19 show quite good data quality. Most of the IBIS-FS data points are close to the mode shape obtained from FEM, which validates the accuracy

of IBIS-FS measurements. However, several data points do not match well with the FEM mode shape. One reason is that the data quality is not quite good at these points, which can be directly revealed by the polar graphs, e.g., data set 35 does not match up with mode shape #4 as shown in Fig. 18, and it can be seen from its polar graph

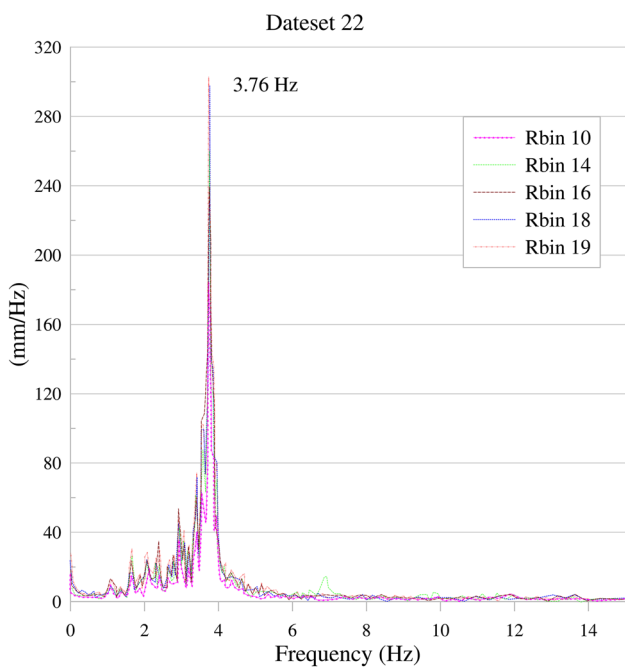


Fig. 12 Frequency displacement graph from data set 22 of IBIS-FS

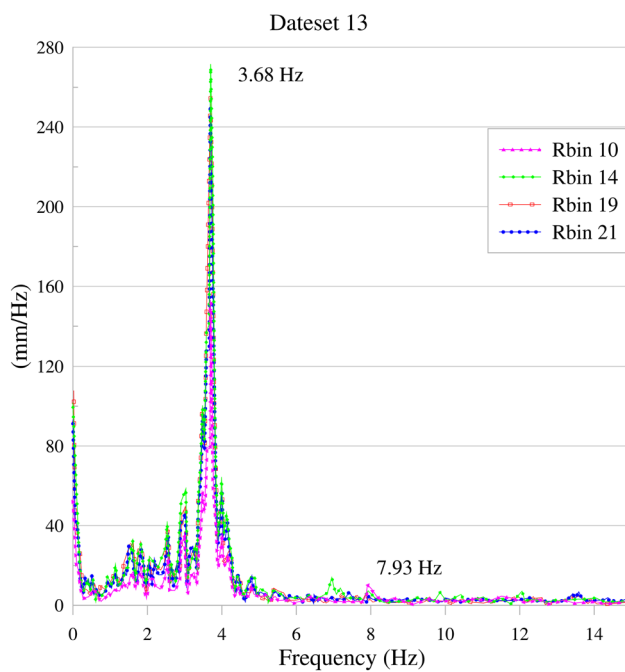


Fig. 14 Frequency displacement graph from data set 13 of IBIS-FS

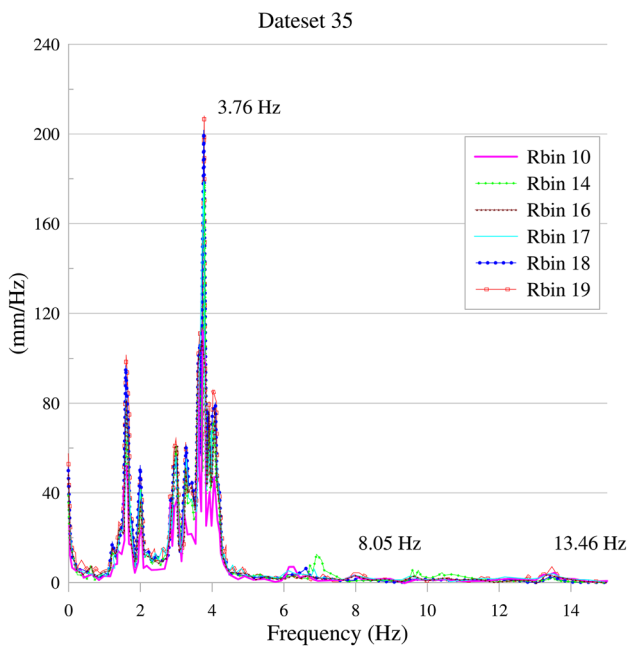


Fig. 13 Frequency displacement graph from data set 35 of IBIS-FS

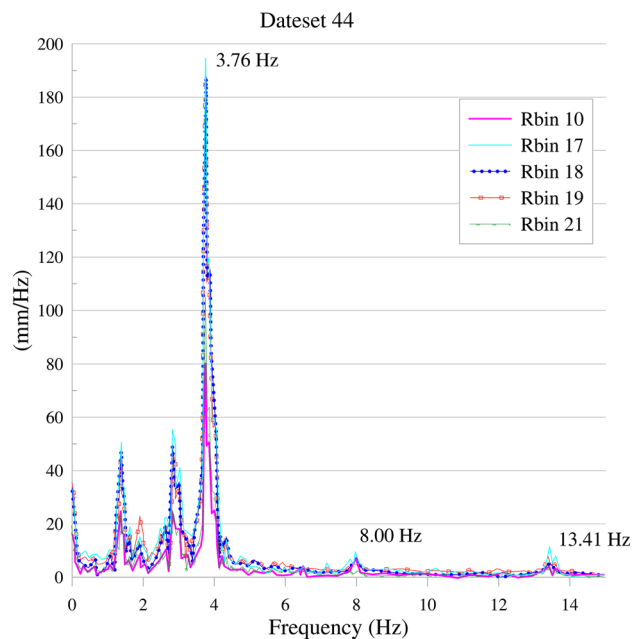


Fig. 15 Frequency displacement graph from data set 44 of IBIS-FS

that there are two lines not perfectly shaped and are a bit wide (Fig. 11a). Despite the observed scatters, the results are generally acceptable. The accuracy of the IBIS-FS in identifying the mode shapes can be improved by applying a better method (e.g., stochastic subspace identification

(SSI) methods or natural excitation techniques (NExT) to determine mode shapes from IBIS-FS results.

Table 5 Non-torsional modes obtained from finite element model and operational modal analysis

Mode	SAP2000 (Hz)	OMA PP (Hz)	OMA FDD (Hz)
#1	3.84	3.76	3.76
#4	7.63	8.05	7.91
#8	13.91	13.46	13.27

7.2 Damage detection

The term damage detection refers to identifying the damage locations and estimating the damage severity. As previously mentioned, RF, GBDT, and XGBoost with optimised

hyperparameters were considered for damage detection. The training data frame consists of 1700 raw vectors. Each raw vector contains the following data points: (a) scenario #1, which includes the mode shapes at all nodes in the structural model (151 nodes and 6 degrees of freedom). (b) scenario #2, which only includes 81 nodal values (referred to as scenario with incomplete mode shapes). For each scenario, a different number of frequencies (out of the 20 frequencies included in the numerical modal analysis) with varying levels of noise were considered. The data were contaminated by Gaussian distribution with zero mean and 1 for the standard deviation of the normal distribution. 70% of the data were used to train the models, and 30% was used for testing. As mentioned, the confusion matrix was used to evaluate

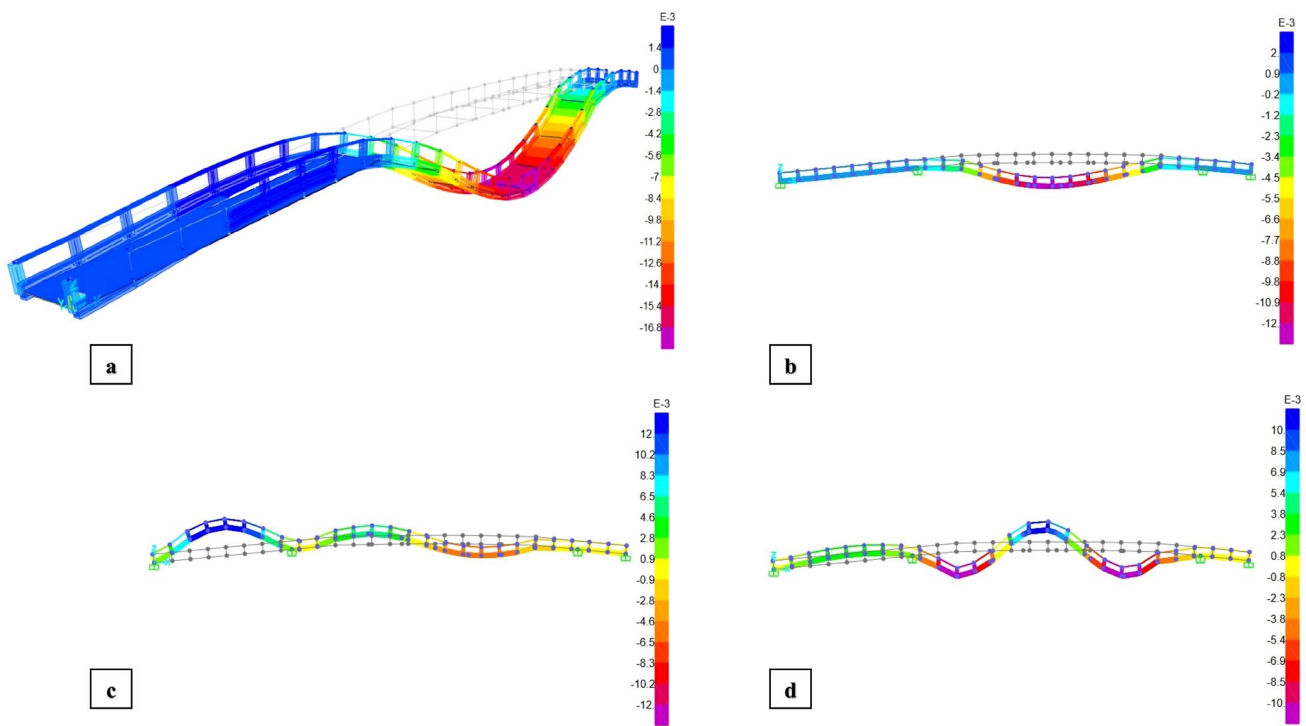


Fig. 16 Non-torsional mode shapes obtained from SAP2000: (a) Extruded view of first mode shape, (b) 2D view of first mode shape, (c) 2D view mode shape #4, and (d) 2D view of mode shape #8

Fig. 17 Mode shape #1 obtained from numerical modal analysis and IBIS-FS

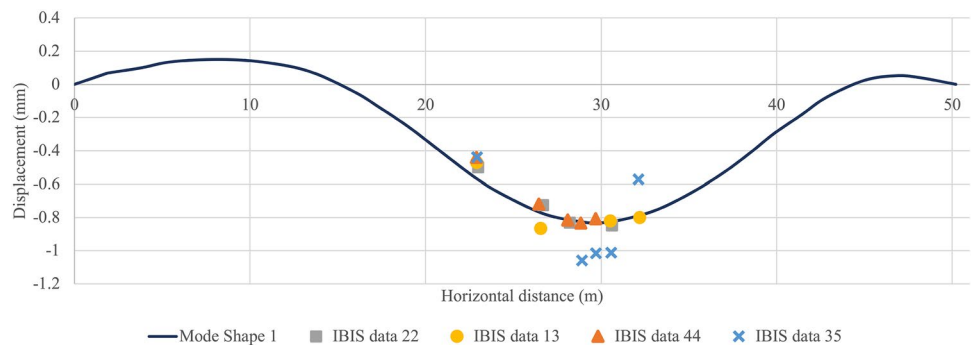
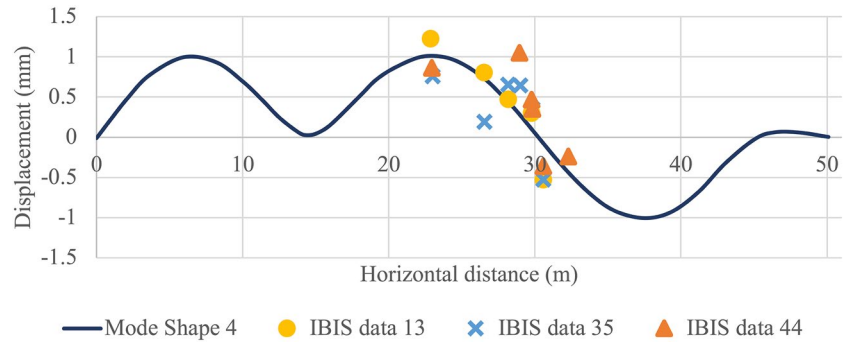


Fig. 18 Mode shape #4 obtained from numerical modal analysis and IBIS-FS



classification, while the Pearson correlation coefficient, R^2 , and RMSE are used as regression metrics. In regression tasks, the coefficient of determination, commonly known as R^2 , serves as a pivotal metric. R^2 quantifies the proportion of the variance in the dependent variable that is predictable from the independent variable(s). In the context of damage detection, R^2 facilitates the evaluation of the model's ability to accurately predict the severity or extent of structural damage, providing insights into the reliability of the regression model. Precision, recall, and F1 score are fundamental metrics for classification tasks. Precision quantifies the accuracy of positive predictions, ensuring that identified damage instances are indeed indicative of true damaged element. Recall, on the other hand, measures the model's ability to correctly identify all instances of actual damaged elements within the structure. The F1 score, being the harmonic mean of precision and recall, serves as a balanced metric in situations where class imbalances exist. It provides

a comprehensive assessment of a model's performance, particularly when both false positives and false negatives must be minimised.

Table 6 shows the damage detection results using RF and XGBoost in scenario #1. RF and XGBoost maintain high accuracy, often close to or above 0.90, indicating strong performance in correctly classifying damaged and undamaged cases. Both models exhibit very high accuracy, demonstrating that they perform well even as noise increases and with fewer modes. RF consistently demonstrates high precision, indicating a strong ability to correctly identify true damage cases. XGBoost matches RF in precision, showing its effectiveness in identifying damaged cases without many false positives. This low rate of false positives is crucial, as it indicates that the model seldom misclassifies undamaged cases as damaged, ensuring reliability and reducing the likelihood of unnecessary interventions or repairs based on incorrect damage assessments. RF has strong recall, which shows a slight decrease with fewer modes, indicating a high ability to

Table 6 Damage detection using RF and XGBoost with **complete** mode shapes (scenario #1)

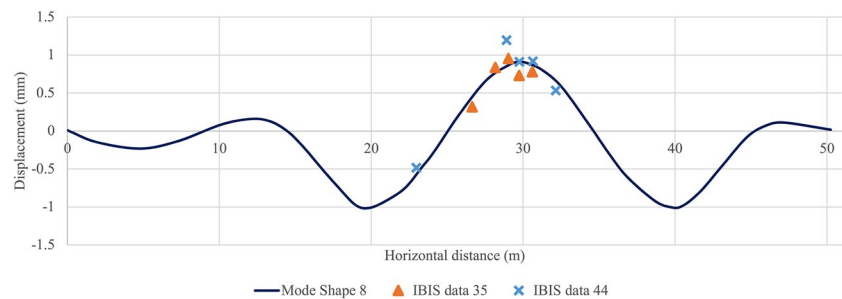
Model	Number of modes	Noise level	Acc	Prec	Rec	F_1	RMSE	R2	PCC
RF	20 modes	Noise-free	0.99	0.98	0.99	0.98	0.01	0.99	0.96
		5%	0.99	0.97	0.99	0.98	0.02	0.94	0.98
		10%	0.96	0.96	0.98	0.97	0.01	0.94	0.97
	10 modes	Noise-free	0.92	0.96	0.88	0.92	0.01	0.91	0.96
		5%	0.91	0.93	0.88	0.90	0.01	0.90	0.95
		10%	0.89	0.92	0.87	0.91	0.03	0.90	0.93
	3 modes	Noise-free	0.89	0.91	0.86	0.89	0.01	0.88	0.93
		5%	0.88	0.87	0.86	0.86	0.02	0.88	0.91
		10%	0.88	0.86	0.85	0.86	0.02	0.88	0.89
XGBoost	20 modes	Noise-free	0.99	0.99	0.97	0.98	0.00	0.97	0.98
		5%	0.97	0.99	0.97	0.98	0.03	0.94	0.96
		10%	0.97	0.93	0.95	0.94	0.03	0.93	0.97
	10 modes	Noise-free	0.95	0.93	0.96	0.94	0.00	0.97	0.97
		5%	0.90	0.92	0.88	0.90	0.03	0.91	0.94
		10%	0.91	0.85	0.84	0.84	0.05	0.89	0.90
	3 modes	Noise-free	0.89	0.90	0.90	0.90	0.01	0.91	0.86
		5%	0.91	0.85	0.85	0.85	0.03	0.88	0.87
		10%	0.88	0.86	0.82	0.84	0.04	0.86	0.87

Table 7 Damage detection using RF and XGBoost with incomplete mode shapes (scenario #2)

Model	Number of modes	Noise level	Acc	Prec	Rec	F_1	RMSE	R2	PCC
RF	20 modes	Noise-free	0.95	0.94	0.96	0.95	0.04	0.97	0.96
		5%	0.92	0.92	0.93	0.92	0.04	0.93	0.91
		10%	0.94	0.96	0.91	0.93	0.04	0.93	0.88
	10 modes	Noise-free	0.92	0.95	0.90	0.92	0.04	0.90	0.91
		5%	0.88	0.92	0.87	0.93	0.05	0.86	0.86
		10%	0.81	0.83	0.87	0.85	0.05	0.86	0.84
	3 modes	Noise-free	0.83	0.88	0.86	0.87	0.08	0.81	0.80
		5%	0.80	0.81	0.86	0.83	0.08	0.77	0.72
		10%	0.73	0.78	0.72	0.75	0.12	0.81	0.78
XGBoost	20 modes	Noise-free	0.97	0.97	0.97	0.97	0.00	0.97	0.98
		5%	0.95	0.90	0.92	0.91	0.01	0.91	0.91
		10%	0.91	0.90	0.88	0.89	0.10	0.91	0.88
	10 modes	Noise-free	0.95	0.93	0.93	0.93	0.01	0.97	0.97
		5%	0.92	0.92	0.90	0.91	0.02	0.91	0.92
		10%	0.86	0.81	0.87	0.84	0.14	0.87	0.83
	3 modes	Noise-free	0.89	0.90	0.89	0.89	0.02	0.91	0.86
		5%	0.89	0.83	0.79	0.81	0.08	0.87	0.88
		10%	0.75	0.81	0.72	0.76	0.13	0.81	0.77

Table 8 Damage detection using RF, GBDT, and XGboost in IBIS-FS simulated environment

Number of selected nodes	Number of selected Dof	Noise level	Model	Acc	Prec	Rec	F_1	RMSE	R2	PCC
7	1 (z dir.)	10%	RF	0.751	0.727	0.701	0.713	0.194	0.73	0.76
			XGBoost	0.726	0.752	0.681	0.714	0.230	0.68	0.70
			GBDT	0.647	0.688	0.641	0.663	0.257	0.68	0.64

Fig. 19 Mode shape #8 obtained from numerical modal analysis and IBIS-FS

detect true positives. XGBoost's recall is comparable to RF, suggesting it can effectively identify actual cases of damage with only a few elements going undetected. The F_1 score of both models remains high, showing balanced performance between precision and recall. RF and XGBoost show very low RMSE, suggesting accurate predictions in the severity of damage. RF presents high R^2 values, demonstrating its capability to account for the variance in damage severity. XGBoost's R^2 values are comparable, showing its capacity to predict the severity of damage effectively.

In Table 7, the RF and XGBoost results with incomplete mode shapes are presented. In this case, the model values at

70 nodes from the structure were deleted. From the results presented in Table 8, RF's accuracy decreases compared to scenario #1 but remains robust, indicating resilience in less ideal conditions. XGBoost experiences a slight drop in accuracy, especially with the highest noise level and the fewest modes, suggesting a bit more sensitivity to data completeness. RF still maintains high precision, though it shows some reduction, suggesting reliability in identifying true damage. XGBoost sees some drop in precision but remains relatively high, suggesting it may produce more false positives compared to RF in this scenario. While still effective, the increased rate of false positives with XGBoost indicates

a higher likelihood of incorrectly identifying undamaged cases as damaged. This could lead to unnecessary interventions, highlighting the need for careful consideration in choosing the appropriate model based on the acceptable balance between detecting all true damage cases and avoiding misclassification of undamaged elements. RF's recall dips in scenario #2, which may indicate challenges in detecting all true positives under more complex conditions. XGBoost shows a more noticeable decrease in recall, especially with fewer modes and more noise, indicating potential difficulties in identifying all actual damaged cases. RF's F1 score is lower in scenario #2 but still reflects a good balance between precision and recall. XGBoost's F1 score decreases slightly more than RF's, suggesting that it may not handle incomplete data as effectively as RF. RF sees an increase in RMSE but still indicates good performance, suggesting that predictions of damage severity are reasonably accurate. XGBoost's RMSE is slightly higher in this scenario, indicating a small decrease in the accuracy of severity predictions compared to RF. RF and XGBoost show a slight reduction in R^2 values but they remain strong, reflecting their ability to explain the variance in damage severity and suggesting that they retain much of their predictive power even with incomplete mode shapes.

In summary, both RF and XGBoost demonstrate strong performance in scenarios with complete and incomplete mode shapes. However, RF appears to be slightly more robust when dealing with incomplete data. While XGBoost shows a reduction in some metrics in Scenario #2, it still performs adequately and may offer advantages in specific regression tasks. Both algorithms exhibit adaptability to challenging data conditions, underscoring their utility in damage detection applications for the bridge under investigation in this study.

Regarding computing time, RF and XGBoost display different performance characteristics. To facilitate a comprehensive comparison, scenario #2 with three modes is utilised. For detecting the location of damage, RF takes 103 s, which is quite efficient given the task's complexity. RF involves constructing multiple decision trees and making predictions by averaging the results, which can be computationally intensive. However, it can perform classification relatively quickly due to the parallel processing of the trees. In contrast, XGBoost takes significantly longer for classification, requiring 647 s. XGBoost is renowned for its performance and accuracy, often surpassing other algorithms, but it can be more computationally demanding. It employs gradient boosting, an iterative process where new models are created to predict the residuals or errors of previous models, and then combined to form the final prediction. This sequential model building is inherently less parallelisable than RF's tree building, leading to longer computation times. For estimating the severity of damage, RF requires

a considerably longer time, at 388 s. Regression tasks in RF might need a more complex split criterion and the computation of more statistics at each node, compared to classification, which could explain the extended duration. On the other hand, XGBoost requires only 60 s, significantly outpacing RF. XGBoost is highly optimised for performance and can converge to a solution more quickly in regression tasks, despite its iterative nature.

To investigate if IBIS-FS can be used for damage detection, only three non-torsional modes and the associated vertical displacement of seven nodes from the mid-span of the bridge were used. The results of the RF process are shown in Table 8. Only seven nodes were selected as it is unlikely that IBIS-FS can measure the mode shapes at all nodes (with six DoFs) of the structure. This is due to the complexity of the structure, the limited number of IBIS-FS and the difficulties in synchronising IBIS-FS. Therefore, only translational DoF (the vertical z-direction) at the selected nodes: #23, #26, #28, #29, #30, #31, and #32 (see Figs. 17, 18, 19) were used for damage detection. The RF process was used in this case with spatially incomplete data and 10% noise-polluted mode shapes associated with modes #1, #4, and #8.

As shown in Table 8, among the assessed models, the RF model excels, securing the top position in accuracy, demonstrating a commendable balance between precision and recall, and indicating fewer prediction errors with its low RMSE and high R^2 score. XGBoost, though not far behind, surpasses RF in precision but falls short in recall and overall prediction consistency, as reflected by its RMSE and R^2 . On the other hand, GBDT lags in all key metrics, indicating a lower classification efficacy and the highest prediction errors, making RF the most reliable for damage detection in this study. The results indicate the possibility of damage detection using IBIS-FS with the RF method.

8 Conclusion

This study explores the capability of the IBIS-FS radar system in capturing the mode shapes of the pedestrian bridge at the University of Melbourne. By applying periodic loads at mid-span, close to the bridge's natural frequency, the IBIS-FS was used to measure the displacement at this key location. Operational modal analysis methods, such as peak picking (PP) and frequency domain decomposition (FDD), facilitated the acquisition of modal parameters, with a focus on non-torsional natural frequencies and associated mode shapes. These mode shapes, as determined through IBIS-FS, were subsequently compared to those from a finite element (FE) model, constructed based on on-site measurements and architectural drawings. While the experimental and numerical results exhibited varying levels of accuracy, they collectively affirmed the effectiveness of IBIS-FS in identifying

mode shapes. Radar interferometry, characterised by its rapid installation and ability for remote displacement and vibration measurements, is particularly advantageous in situations where installing an accelerometer is impractical due to accessibility constraints. Its capacity for remote sensing not only guarantees operator safety but also avoids invasive equipment installation on structures. Regarding the financial aspect, the IBIS-FS system represents a significant investment compared to traditional accelerometers, encompassing the necessary hardware and software. Furthermore, the system requires only a moderate training period for operation. A few hours are typically enough to acquaint an operator with its functionalities, ensuring accurate data collection and system optimisation. Additionally, the radar interferometer measures displacements in a direct manner, providing crucial data for evaluating damage and observing structural behavior after events, including the remaining displacements caused by earthquakes. However, radar interferometry, despite its benefits over traditional sensors, necessitates specific conditions for accurate data acquisition:

A clear line of sight without obstructions between the instrument and the target structure is essential.

The instrument must be placed in an area free from vibrations.

The geometry of the instrument's placement in relation to the target must be clear and well-defined.

In the second phase of the study, the effectiveness of IBIS-FS in combination with advanced machine-learning approaches, including Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and XGBoost, in damage detection was evaluated. Using the SAP2000 open application programming interface (API) with Python, a finite element model of the bridge was developed to generate data for training these algorithms. Two main scenarios for damage detection were considered. In the first scenario, complete mode shapes were used, and the results demonstrated the capability of RF, and XGBoost methods in accurately detecting the location and severity of damage. In the second scenario, incomplete mode shapes were employed, and these methods showed remarkable accuracy. To specifically assess the capability of IBIS-FS in damage detection, three non-torsional modes and associated modal values at only seven nodes of the bridge with a single degree of freedom (in Z-direction) were analysed. The findings indicate that IBIS-FS, when used in conjunction with these advanced machine-learning methods, can effectively locate and assess the severity of damage. This article introduces a novel approach to bridge damage detection using IBIS-FS in tandem with machine-learning techniques. While this method is still nascent compared to traditional non-destructive techniques, it shows considerable promise for future explorations. Future

research directions include achieving three-dimensional displacement measurements of bridges by employing multiple IBIS-FS units. This could significantly enhance the accuracy of modal parameters and capture additional mode shapes, thereby improving the precision of damage detection. Efforts to minimise errors could involve refining Operational Modal Analysis (OMA) techniques, such as NeXT, as well as accounting for finite element modeling inaccuracies to improve overall results. Further to these advancements, the proposed machine learning (ML)-based damage detection methods should be compared with traditional methods (e.g., the cross-correlation damage index proposed by Ref. [46], cross-model cross-mode by Ref. [47]), as well as with other ensemble-based methods like LightGBM. However, specific challenges must be addressed to ensure the reliability and effectiveness of these novel approaches. Environmental factors such as temperature, humidity, and operational conditions, including traffic load on the bridge, can impact both the IBIS-FS system's performance and machine-learning analysis. Developing models robust enough to adjust to these variabilities is essential for dependable damage detection. Therefore, identifying the most strategic arrangement of IBIS-FS to maximise coverage and minimise blind spots, especially in bridges with complex geometries, is a spatial challenge that demands planning and simulation. Furthermore, the integration of IBIS-FS data with pre-existing monitoring systems presents a unique challenge. It is crucial to merge data from IBIS-FS with these systems seamlessly, ensuring that the machine-learning models can interpret and analyse data from diverse sources effectively, without causing any disruptions to ongoing monitoring operations.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. The financial support from the University of Melbourne through the Melbourne Research Scholarship (468777) is gratefully acknowledged.

Data availability Data sets generated during the current study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Hou R, Xia Y (2021) Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. *J Sound Vib* 491:115741. <https://doi.org/10.1016/j.jsv.2020.115741>
- Cao MS, Sha GG, Gao YF, Ostachowicz W (2017) Structural damage identification using damping: A compendium of uses and features. *Smart Mater Struct* 26. <https://doi.org/10.1088/1361-665X/aa550a>
- Meng X, Dodson AH, Roberts GW (2007) Detecting bridge dynamics with GPS and triaxial accelerometers. *Eng Struct* 29:3178–3184. <https://doi.org/10.1016/j.engstruct.2007.03.012>
- Kalooop MR, Li H (2009) Monitoring of bridge deformation using GPS technique. *KSCE J Civ Eng* 13:423–431. <https://doi.org/10.1007/s12205-009-0423-y>
- Kalooop MR, Li H (2011) Sensitivity and analysis GPS signals based bridge damage using GPS observations and wavelet transform. *Measurement (Lond)* 44:927–937. <https://doi.org/10.1016/j.measurement.2011.02.008>
- Yi TH, Li HN, Gu M (2013) Experimental assessment of high-rate GPS receivers for deformation monitoring of bridge. *Measurement (Lond)* 46:420–432. <https://doi.org/10.1016/j.measurement.2012.07.018>
- Feng C, Zhang H, Wang S et al (2019) Structural damage detection using deep convolutional neural network and transfer learning. *KSCE J Civ Eng* 23:4493–4502. <https://doi.org/10.1007/s12205-019-0437-z>
- Zhang X, Wang J, Wang Q, et al (2022) Service performance evaluation of long-span cable-stayed bridge based on health monitoring data. *J of Vibroeng* 24:651–665. <https://doi.org/10.21595/jve.2022.22756>
- Chen H-P (2018) *Structural health monitoring of large civil engineering structures*. John Wiley & Sons Ltd
- Im SB, Hurlbaas S, Kang YJ (2013) Summary Review of GPS technology for structural health monitoring. *J Struct Eng* 139:1653–1664. [https://doi.org/10.1061/\(asce\)st.1943-541x.0000475](https://doi.org/10.1061/(asce)st.1943-541x.0000475)
- Castellini P, Martarelli M, Tomasini EP (2006) Laser Doppler Vibrometry: Development of advanced solutions answering to technology's needs. *Mech Syst Signal Process* 20:1265–1285. <https://doi.org/10.1016/j.ymsp.2005.11.015>
- Siringoringo DM, Fujino Y (2006) Experimental study of laser Doppler vibrometer and ambient vibration for vibration-based damage detection. *Eng Struct* 28:1803–1815. <https://doi.org/10.1016/j.engstruct.2006.03.006>
- Gentile C, Bernardini G (2010) Radar-based measurement of deflections on bridges and large structures. *Euro Eur J Environ Civ Eng* 14:495–516. <https://doi.org/10.1080/19648189.2010.9693238>
- Mayer L, Yanev BS, Olson LD, Smyth AW (2010) Monitoring of Manhattan bridge for vertical and torsional performance with GPS and Interferometric Radar systems. In: Transportation Research Board 89th Annual Meeting, Washington DC, United States
- Sofi M, Lumantarna E, Zhong A et al (2018) Determining dynamic characteristics of high rise buildings using interferometric radar system. *Eng Struct* 164:230–242. <https://doi.org/10.1016/j.engstruct.2018.02.084>
- Pieraccini M, Fratini M, Parrini F et al (2008) Interferometric radar vs. accelerometer for dynamic monitoring of large structures: an experimental comparison. *NDT and E Intl* 41:258–264. <https://doi.org/10.1016/j.ndteint.2007.11.002>
- Gentile C, Bernardini G (2008) Output-only modal identification of a reinforced concrete bridge from radar-based measurements. *NDT and E Intl* 41:544–553. <https://doi.org/10.1016/j.ndteint.2008.04.005>
- Sofi M, Lumantarna E, Mendis PA et al (2017) Assessment of a pedestrian bridge dynamics using Interferometric Radar system IBIS-FS. *Procedia Eng* 188:33–40. <https://doi.org/10.1016/j.proeng.2017.04.454>
- Clinici TS, Saracin A, Negrila AF (2017) Interferometric Radar technology in construction monitoring. In: SGEM 453–461
- Saracin A, Negrila AFC, Clinici TS (2019) Possibilities for building monitoring, using terrestrial radar interferometry *J Geod Cartogr Cadastre* 27–33
- Silva L, Barros RC, Paiva F, Henriques J (2016) Structural monitoring of a telecommunication mast by Radar Interferometry. In: Proceedings of the 5th Publ. INEGI/FEUP 1455–1466
- Quirkel P, Barrias A (2020) Validation of finite element light rail bridge model using dynamic bridge deflection measurement. *Civ Eng Res in Ireland* 52–57
- Castagnetti C, Bassoli E, Vincenzi L, Mancini F (2019) Dynamic assessment of masonry towers based on terrestrial radar interferometer and accelerometers. *Sensors (Switzerland)* 19. <https://doi.org/10.3390/s19061319>
- Camassa D, Castellano A, Fraddosio A, et al (2021) Dynamic identification of tensile force in tie-rods by Interferometric Radar measurements. *Applied Sciences (Switzerland)* 11. <https://doi.org/10.3390/app11083687>
- Castellano A, Camasso D, Fraddosio A, Piccioni MD (2022) Radar Interferometric experimental reconstruction of three-dimensional displacement vectors and mode shapes for masonry constructions. *J Phys Conf Ser* 2204:1–6. <https://doi.org/10.1088/1742-6596/2204/1/012055>
- George Hearn B, Testa RB (1991) modal analysis for damage detection in structures. *J struct Eng* 117:3042–3063
- Frigui F, Faye JP, Martin C et al (2018) Global methodology for damage detection and localization in civil engineering structures. *Eng Struct* 171:686–695. <https://doi.org/10.1016/j.engstruct.2018.06.026>
- Avci O, Abdeljaber O, Kiranyaz S, et al (2021) A review of vibration-based damage detection in civil structures: from traditional methods to machine learning and deep learning applications. *Mech Syst Signal Process* 147. <https://doi.org/10.1016/j.ymsp.2020.107077>
- Zhou Q, Ning Y, Zhou Q et al (2013) Structural damage detection method based on random forests and data fusion. *Struct Health Monit* 12:48–58. <https://doi.org/10.1177/1475921712464572>
- Zhou Q, Zhou H, Zhou Q et al (2014) Structure damage detection based on random forest recursive feature elimination. *Mech Syst Signal Process* 46:82–90. <https://doi.org/10.1016/j.ymsp.2013.12.013>
- Lu Q, Zhu J, Zhang W (2019) Quantification of fatigue damage of structural details in slender coastal bridges using machine learning based methods. In: Structures Congress, Orlando, Florida, 122–133
- Chencho, Li J, Hao H, et al (2021) Development and application of random forest technique for element level structural damage quantification. *Struct Control Health Monit* 28. <https://doi.org/10.1002/stc.2678>
- Chaiyasarn K, Sharma M, Ali L, et al (2018) Crack detection in historical structures based on Convolutional Neural Network. *Intl J GEOMATE* 15:240–251. <https://doi.org/10.21660/2018.51.35376>
- Huang D, Hu D, He J, Xiong Y (2018) Structure damage detection based on ensemble learning. In: Proceedings of 2018 9th International Conference on Mechanical and Aerospace Engineering, In: Proceedings of 9th ICMAE 219–224

35. Garg Y, Masih A, Sharma U (2021) Predicting bridge damage during earthquake using machine learning algorithms. In: Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering. Institute of Electrical and Electronics Engineers Inc. 725–728
36. Li Q, Song Z (2022) Ensemble-learning-based prediction of steel bridge deck defect condition. *Applied Sciences (Switzerland)* 12. <https://doi.org/10.3390/app12115442>
37. Sony S, Dunphy K, Sadhu A, Capretz M (2021) A systematic review of convolutional neural network-based structural condition assessment techniques. *Eng Struct* 226:111347. <https://doi.org/10.1016/j.engstruct.2020.111347>
38. Sun L, Shang Z, Xia Y, et al (2020) Review of bridge structural health monitoring aided by big data and artificial intelligence: From Condition Assessment to Damage Detection. *Journal of Structural Engineering* 146. [https://doi.org/10.1061/\(asce\)st.1943-541x.0002535](https://doi.org/10.1061/(asce)st.1943-541x.0002535)
39. Rashidi Nasab A, Elzarka H (2023) Optimizing machine learning algorithms for improving prediction of bridge deck deterioration: A Case Study of Ohio Bridges. *Buildings* 13. <https://doi.org/10.3390/buildings13061517>
40. SAP2000, C. S. I. (2015). Computer & Structures Inc. Linear and nonlinear static and dynamic analysis of three-dimensional structures. Berkeley (CA): Computer & Structures
41. Bernardini G, Gallino N, Gentile G, Ricci P (2007) Dynamic monitoring of civil engineering structures by microwave interferometer. In: 4th Conceptual Approach to Structural Design Venice
42. Rainieri C, Fabbrocino G (2014) Operational modal analysis of civil engineering structures an introduction and guide for applications. Springer, US
43. Yang Z, Gao W, Chen L, et al (2022) A novel electromechanical impedance-based method for non-destructive evaluation of concrete fiber content. *Constr Build Mater* 351. <https://doi.org/10.1016/j.conbuildmat.2022.128972>
44. Ribeiro RR, Veloso LACM, Lameiras RM (2022) Machine learning vibration-based damage detection and early-developed damage indicators. In: Dilworth BJ, Mains M (eds) *Topics in Modal Analysis & Testing*. Springer International Publishing, Cham, pp 23–32
45. Alva RE, González-Drigo JR, Luzi G, et al (2019) Remote ambient vibration measurements with real-aperture radar to estimate buildings dynamic properties. In: *COMPDYN Proceedings*. National Technical University of Athens, 1797–1808
46. Khosraviani MJ, Bahar O, Ghasemi SH (2021) Global and local damage detection in continuous bridge decks using instantaneous amplitude energy and cross-correlation function methods. *KSCE J Civ Eng* 25:603–620. <https://doi.org/10.1007/s12205-020-0622-0>
47. Li H, Wang J, James Hu SL (2008) Using incomplete modal data for damage detection in offshore jacket structures. *Ocean Eng* 35:1793–1799. <https://doi.org/10.1016/j.oceaneng.2008.08.020>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.