Engineering Structures 334 (2025) 120253

Contents lists available at ScienceDirect

Engineering Structures

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

A two-step failure identification approach using a stochastic optimization-based ensemble learning model for beams without pristine data

Long Viet Ho [a], Thanh Bui-Tien [a], Magd Abdel Wahab [b,c,d,*]

[a] University of Transport and Communications, Vietnam
[b] Soete Laboratory, Department of Electrical Energy, Metals, Mechanical Constructions and Systems, Faculty of Engineering and Architecture, Ghent University, Ghent 9000, Belgium
[c] College of Engineering, Yuan Ze University, Taiwan
[d] Parallel and Distributed Systems Laboratory, Jožef Stefan Institute, Jamova cesta 39, Ljubljana SI-1000, Slovenia

ARTICLE INFO

Keywords:
Dynamic properties
Gapped smoothing method
Optimization
Ensemble learning model

ABSTRACT

Most vibration-based structural health monitoring techniques evaluate structural conditions by relying on data from both intact and damaged states. Data from a healthy structure is highly valuable for detecting damage. While it is often abundant in newly monitored structures, it may be limited or unavailable in existing bridges, where monitoring usually starts after years of service-related deterioration. Additionally, noise is inherently present in measured data. Therefore, combining data from both healthy and deteriorated states introduces a higher risk of inaccuracies. To address this, the present study proposes a two-step approach for damage identification that relies solely on the damaged state. First, a modified mode shape-based damage index is employed to localize failures using only the available damaged-state data. Then an ensemble learning model improved by a stochastic optimization process is implemented to estimate damage extent. In this study, damage scenarios involving single and multiple cuts are examined instead of assuming stiffness loss through Young's modulus reduction. A flat beam, developed based on experimental data, and a concrete T-girder bridge are utilized to validate the applicability of the proposed approach. Numerical results indicate that damage locations can be identified using the first four modes, even under noisy conditions. Furthermore, optimization-based ensemble learning models demonstrate greater efficiency and stability compared to traditional models in extent estimation through cut length.

1. Introduction

Vibration-based damage detection methods are diverse, utilizing changes in time series data and dynamic properties such as frequencies, mode shapes, and damping ratios for identifying damage [1,2]. Although damping ratio-based methods have attracted research attention, their practical application remains challenging due to high sensitivity to environmental and operational variations, difficulty in accurate measurement, and the need for complex instrumentation. In contrast, time series data and dynamic characteristics such as frequencies and mode shapes are both economically and technically feasible, making them more suitable for this study. Nonetheless, to achieve high accuracy in identifying faults consisting of damage location and quantification, these data typically need to be integrated with optimization algorithms (OAs) or machine learning (ML) techniques [3,4], which often require iterative analysis or complex training processes with a proper number of datasets.

In the time series-based damage detection approach, raw data can be extracted from experiments or finite element (FE) models [5,6] or both [7], and then an ML model for classification problems is implemented to detect and classify damage. However, this approach often identifies the presence and location of damage but does not determine the damage level. Shirazi et al. evaluated the effectiveness of the 1D-CNN network for crack classification [8]. Huang et al. applied DCNN-LSTM to predict abnormal bearing displacement in the observed structure due to temperature variations [9]. Nguyen-Ngoc et al. applied a long short-term

* Corresponding author at: Soete Laboratory, Department of Electrical Energy, Metals, Mechanical Constructions and Systems, Faculty of Engineering and Architecture, Ghent University, Ghent 9000, Belgium.
E-mail address: magd.abdelwahab@ugent.be (M. Abdel Wahab).

https://doi.org/10.1016/j.engstruct.2025.120253
Received 22 February 2025; Accepted 31 March 2025
Available online 11 April 2025

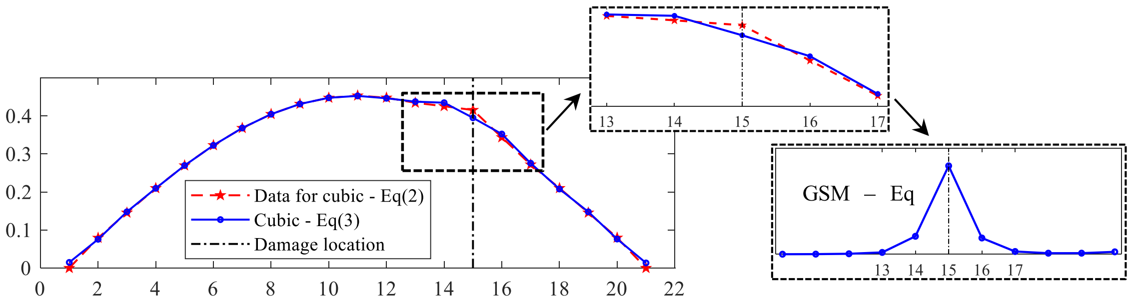

Fig. 1. Calculation of the difference function for node $k = 15$.

memory (LSTM) network combined with time series data for damage detection and localization in a truss bridge [10]. Computational time was one of the crucial criteria that was used to evaluate the effectiveness of the algorithms investigated. Tran-Ngoc et al. implemented a working parallel method combining ANN and optimization algorithms to identify damaged location and severity based on changes in natural frequencies [11]. Generally, while this approach leverages optimization algorithms to determine damage, it also requires considerable time to analyze damage location, particularly when integrating the estimation of damage severity [12].

This drawback drives researchers to develop damage indices derived from modal properties for non-iterative damage localization. According to this approach, two well-known methods, modal curvature method (MCM), and modal flexibility method (MFM), were developed by Pandey et al. [13,14]. The effectiveness of the indicator and its variants in failure identification has been validated through research [15–20]. However, these methods often require data from both damaged and intact states. Unfortunately, information about the intact state is sometimes unavailable, and incorporating data from both damaged and undamaged states may introduce inaccuracies in damage identification.

As a result, developing an indicator that does not require pristine data has a significant meaning. Ratcliffe developed this approach using only displacement mode shapes to determine indicators for damage localization [21]. Success in damage localization confirmed the effectiveness of this method [15], [22–24]. However, the findings also indicate that, in addition to the large peaks at the damaged locations, false peaks may emerge in undamaged areas, particularly near the ends of the test structures or under noisy conditions. This may lead to difficulties in accurately interpreting the location of damage. Therefore, to enhance the precision of damage localization while retaining the advantages of the existing approach, a truncation technique is applied to mitigate unexpected peaks in undamaged regions. This refinement enables clearer identification of damage locations [25]. The effectiveness of the improved index was further investigated in simply supported beam structures, focusing on accurately determining cut locations. This approach provides a more practical representation of structural damage compared to the common assumption of stiffness reduction through the change in Young's modulus, which alters stiffness without reflecting the corresponding mass loss typically associated with real damage.

In the damage index-based approach, the extent of damage is assessed only at known locations, which significantly reduces analysis time. Quantifying damage severity has also attracted the attention from researchers because the relationship between the damage index and damage level is less well-defined than that between natural frequencies and the corresponding levels. Popular methods used to estimate fault levels include model updating and machine learning. The former relies on an iterative process controlled by an optimization algorithm, where variables, such as damage level, are updated until a specified objective function is optimized. Therefore, this method often requires significant time to analyze and assess the severity of damage. This time is proportional to the complexity involved in analyzing the dynamic characteristics of the investigated structures. The latter concentrates on building a surrogate model to predict damage severities based on unseen data. Neural networks (NNs), including feedforward neural networks (FNNs) for direct information flow and artificial neural networks (ANNs) trained with iterative feedforward and backpropagation, are commonly used in this field. Authors in [26] applied ANNs to predict the cracks in the beam using both simulated and experimental data. Khatir and his colleagues successfully conducted a study using an RBF neural network to identify the extent level, measured by the depth and width of notches in an experimental RC beam [27]. Recently, an ensemble-based model, namely Least squares boosting (LSBoost), has been used thanks to its simple structure, ease of implementation [28], and ability to handle missing data while maintaining high accuracy [29], particularly against noisy data [30]. This algorithm was used in the civil engineering field such as predicting the compressive strength of concrete [31], estimation of displacement of brick infill under the earthquake effect [32], or prediction of erosion rate [33]. In other fields like the biogeochemistry field [34,35] or mechanics field [36,37], LSBoost also received the attention of researchers. The results obtained confirmed the effectiveness of LSBoost. However, to the authors' knowledge, there has been limited investigation into the efficiency of LSBoost for damage identification. Consequently, LSBoost serves as the ML model to predict the dimension of the cuts in the observed structures. Moreover, to enhance the model's performance, optimization algorithms are applied instead of empirical experience to fine-tune hyperparameters that significantly impact prediction accuracy. With the rapid advancement of optimization algorithms, the hybrid OAs-LSB model presents significant application potential in engineering, particularly in the field of structural health monitoring (SHM).

This paper introduces a two-step method for damage identification without reference to pristine data. In the first step, the cut location is directly identified using a developed damage index which is a modified version of the gapped smoothing method. The damage extent i.e. cut length in the investigated structures is estimated based on frequency changes using a stochastic optimization-based ensemble learning model. This approach helps reduce computational costs while ensuring accuracy.

2. Methodology of the two-step approach

2.1. Localization using damage index

a. Gapped smoothing method (GSM)

Ratcliffe developed the gapped smoothing method (GSM) as a vibration-based damage detection that requires only the damaged state [21]. The displacement mode shapes of the damaged structure are extracted to determine the GSM indicator:

$$GSM_k = \left(\phi^{//}_{k,un} - CP_k\right)^2_{real} + \left(\phi^{//}_{k,un} - CP_k\right)^2_{imaginary} \quad (1)$$

Where, $\phi^{//}_{k,un}$ *and* CP_k indicate curvature and cubic polynomial function at point k of damaged structure. The difference function in Eq. (1)

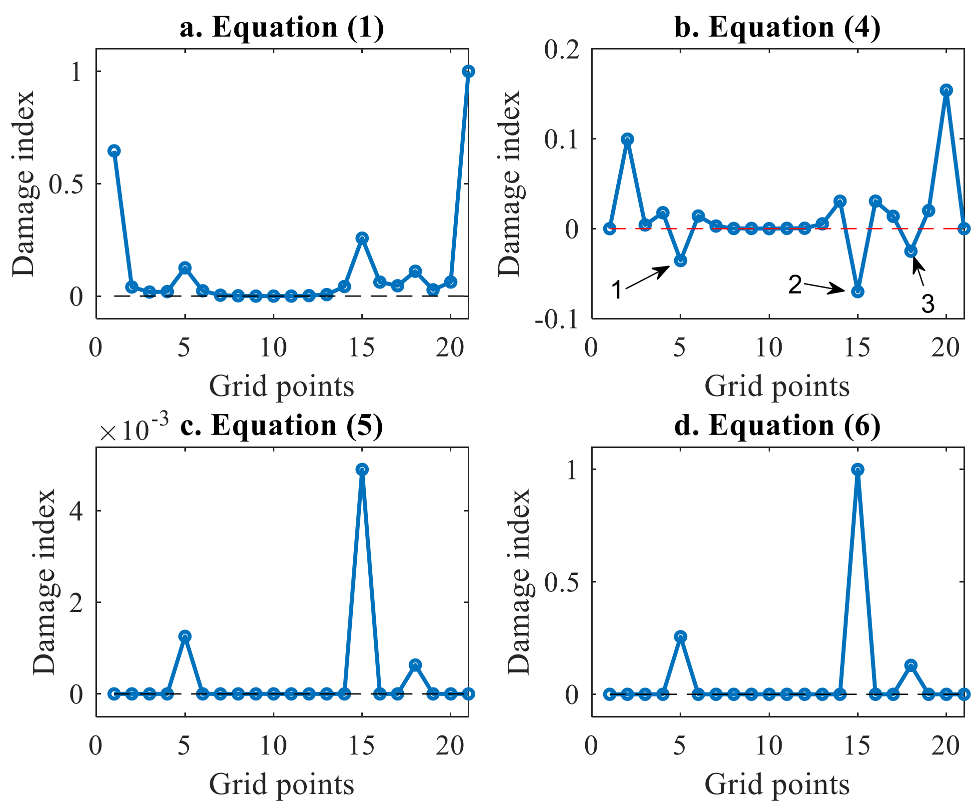

Fig. 2. Illustration of damage localization using the first mode.

identifies the damage location by analyzing the deviation between the modal curvature and the interpolated cubic polynomial function, as shown in Fig. 1.

The central differences approximation is used to identify the second derivative or the curvature $\phi^{//}_{k,un}$ of the collected mode shapes of the damaged structure at point k, $\phi_{k,un}$:

$$\phi^{//}_{k,un} = \frac{(\phi_{k-1,un} - 2\phi_{k,un} + \phi_{k+1,un})}{{s_p}^2} \tag{2}$$

Where s_p represents the uniform distance between two consecutive points along the girder [13]. The cubic polynomial function CP_k at the point k (or at a distance x_k along the investigated structure) can be achieved as in Eq. (3). The polynomial coefficient c_0, c_1, c_2, and c_3 at point k is determined by using four neighboring curvatures obtained from the damaged structure.

$$CP_k = c_0 + c_1 x_k + c_2 x_k^2 + c_3 x_k^3 \tag{3}$$

b. Modified damage index, IGSM

Interpolation of a point CP_k from Eq. (3) in the polynomial curve is most effective when there are at least two spatial coordinates on each side of the considered point k. However, when the point is near one end of the beam, interpolation accuracy decreases due to the "end effect" arising from the numerical operations. This effect reduces the effectiveness of using the difference function in determining the damage location, as spurious peaks tend to appear near the beam's ends. Additionally, minor false peaks also emerge around the actual damage location (see Fig. 2a).

To enhance the identification of potential damage locations, central difference approximation estimates the curvature of the GSM index (see Fig. 2b). Since damage typically appears as sharp variations in the damage index, curvature analysis enhances the detection of these critical regions. This method helps isolate significant structural irregularities while minimizing the influence of minor fluctuations. The curvature of the GSM index at point k is computed as follows:

$$CGSM_k = \frac{(GSM_{k-1,un} - 2GSM_{k,un} + GSM_{k+1,un})}{{s_p}^2} \tag{4}$$

Then, the damage location is identified using a threshold at each point k, as defined in Eq. (5), to suppress peaks in non-damaged regions and focus on the actual damaged areas as in Fig. 2c.

$$IGSM_k = [\min(CGSM_k, 0)]^2 \tag{5}$$

Finally, to enhance the accuracy of damage localization, especially in cases of multiple damages, average values of the damage indices, *norIGSM*, across N modes should be calculated (see Fig. 2d).

$$norIGSM = \frac{\sum_{l=1}^{N} normalize(IGSM)_l}{N} \tag{6}$$

where $l = 1$ *to* N; *normalize(IGSM)* represents the *CGSM* indices standardized to a range of [0 −1].

2.2. *Least-squares boosting (LSBoost)*

Least-squares boosting (LSBoost) is a tree-based algorithm introduced by Friedman [38]. Boosting is a popular technique in ensemble learning that builds models sequentially. In LSBoost, each new tree is constructed to fit the residuals of the previous trees, allowing to gradually improve the model's overall accuracy. Generally, by adding new trees (or weak learners) that are fitted to the observed residuals of the previous iterations, LSBoost iteratively enhances the predictive model.

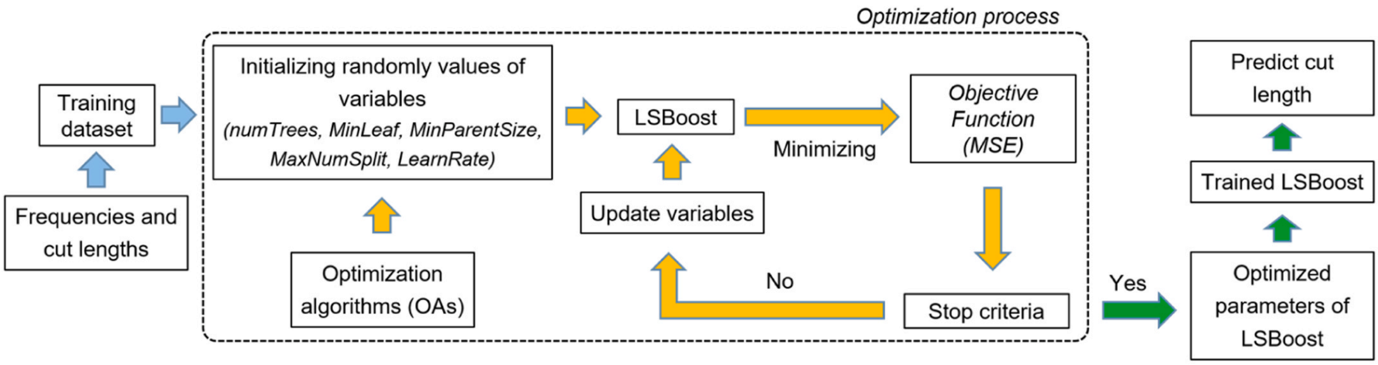

Fig. 3. Optimizing hyperparameters of LSBoost using an optimization process.

In each iteration, a weak learner is trained to predict the residuals rather than the target variables, resulting in a stronger overall model that minimizes the least squares error over time. This approach focuses on reducing bias by combining multiple weak learners.

Pseudocode for LSBoost.

LSBoost algorithm: Define a set of random input or explanatory variables $x = \{x_1, \dots x_n\}$.

Define a set of corresponding output or target variables $y = \{y_1, \dots y_n\}$. Define a loss function,

$L(y, F) = (y - F)^2/2$. With the number of trees (or weak learner) T, data points Ns,

1. Initial model $F_0(x) = y_{ave}$, where y_{ave} is the average of the target values in the training dataset.

2. For $t = 1$ to T do:

3. Compute residual: $y_{res} = \mathrm{y}_i - F_{t\text{-}1}(x_i)$, with $i = 1$ to Ns

4. Fit a new weak learner to the residual: $(\eta_{opt}, a_{opt}) = argmin_{a,\eta}.\Sigma^{Ns}_{i=1}\,[y_{res,i} - \eta.h(x_i\,;\,a)]^2$

5. Update the model: $F_t(x) = F_{t\text{-}1}(x) + \eta_{opt}.h(x_i\,;\,a_{opt})$

In addition, key parameters, such as number of ensemble learning cycles (NumTrees), minimum number of leaves (MinLeaf), maximum number of splits (MaxNumSplit), minimum parent node size (MinParsentSize), and learning rate (LearnRate) should be considered when building an LSBoost model, as they significantly impact its performance. For instance, NumTrees determines number of weak learners, which directly affects model complexity and predictive accuracy. MinLeaf and MaxNumSplit control tree depth and branching, helping to prevent overfitting while maintaining sufficient model flexibility. MinParentSize plays a crucial role in managing the bias-variance trade-off. Meanwhile, LearnRate regulates contribution of each tree to final prediction as well as balances convergence speed and model stability. Therefore, hyperparameter optimization ensures that the LSBoost achieves high accuracy while preventing overfitting and minimizing computational costs.

2.3. *Optimization algorithms–based LSboost (OAs-LSBoost) for damage quantification*

For free, undamped vibration of a structure with stiffness matrix K and mass matrix M, the governing equation of motion is expressed as:

$$M \times \ddot{x} + K \times x = 0 \quad (7)$$

where $\ddot{x}$ *and* x represent the acceleration and displacement vectors as functions of time *t*.

In free vibration, the displacement $x(t)$ exhibits periodic motion, where: $\ddot{x} = -\omega^2 \times x$. Substituting this into the equation of motion:

$$(K - \omega^2 \times M) \times x = 0 \quad (8)$$

Since the angular frequency $\omega = 2\pi f$, the natural frequency f can be expressed as:

$$f = \frac{1}{2\pi}\sqrt{\frac{K}{M}} \quad (9)$$

Eq. (9) shows that a reduction in stiffness due to damage results in a decrease in the natural frequency. In this study, a stochastic optimization-based ensemble learning model is employed to capture the

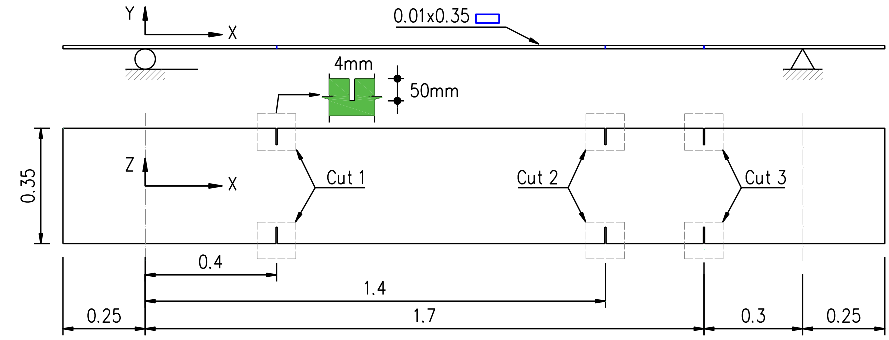

Fig. 4. Schematic of a 2.5-meter flat beam and damage scenarios of Case 1 (unit: m).

relationship between frequency changes and damage extents using training data. Then, the trained model can effectively generalize to unseen data and accurately assess the extent of damage.

The stochastic process optimizes hyperparameters for LSBoost based on the loss function (*OF*) for both training and test processes (see Eq. (10)). The LSBoost implementation is performed using the "fitensemble" function in MATLAB. The step-by-step procedure of the proposed method is presented in Fig. 3. The training dataset encompasses frequencies and corresponding cut scenarios. The optimization process starts by randomly initializing the investigated hyperparameters such as numTrees, Minleaf, MinParentSize, MaxNumSplit, and LearnRate.

Each set of hyperparameters is input into LSBoost to calculate the objective function. The optimization process continues until meeting the stop condition, e.g. reaching the maximum number of iterations. The result of the optimization process is a set of optimized hyperparameters, which is then used to predict the cut length.

$$OF = MSE = \frac{\sum_{i=1}^{Ns}\left(y^i_{predict} - y^i_{target}\right)^2}{Ns} \tag{10}$$

Where $y_{predict}$, y_{target} represent the predicted outputs and desired targets, respectively. *Ns* denotes the number of observed data points.

Several stochastic optimization algorithms, including Antlion Optimizer (ALO) [39], Artificial Gorilla Troop Optimizer (AGTO) [40], and Marine Predator Algorithm (MPA) [41,42], were examined due to their distinct advantages and promising performance. ALO is known for its strong exploration ability, effective handling of exploitation, and reliable convergence speed in solving complex optimization problems. AGTO is recognized for balancing exploration and exploitation, with the ability to escape local optima effectively. MPA demonstrates high search efficiency, inspired by the hunting behavior of marine predators and prey.

2.4. Quality evaluation metrics

Several statistical metrics in Eq. (11–18) are used to evaluate the performances of the proposed approach because they can provide a comprehensive understanding of how well the model fits the data and predicts outcomes. Mean square error (*MSE*) and mean absolute percentage error (*MAPE*) assess the accuracy of predictions, while determination coefficient (R^2) and adjusted R^2 measure how much of the

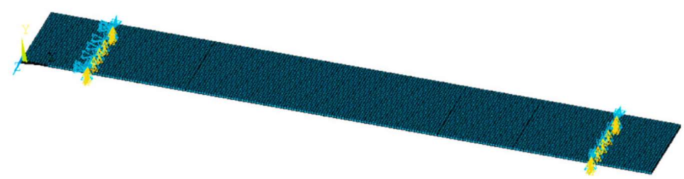

Fig. 7. The FE model of the flat beam.

variability in the data is explained by the model. Normalized mean square error (*NMSE*) and variance accounted for (*VAF*) offer insights into the model's performance relative to the variance in the data. Additionally, mean error (*Mean*) and standard deviation (*Std*) help understand the bias and spread of errors between the outputs and targets. Using these metrics ensures a well-rounded evaluation, considering different aspects of model performance such as accuracy, and reliability.

$$R^2 = 1 - \frac{\sum_{i=1}^{Ns}\left(y^i_{target} - y^i_{predict}\right)^2}{\sum_{i=1}^{Ns}\left(y^i_{target} - \bar{y}_{target}\right)^2} \tag{11}$$

$$AdjustedR^2 = 1 - \frac{(1 - R^2)(Ns - 1)}{Ns - p - 1}, \tag{12}$$

with *p* is the number of input variables

$$Mean = \frac{\sum_{i=1}^{Ns}\left(y^i_{target} - y^i_{predict}\right)}{Ns} \tag{13}$$

$$std = \sqrt{\frac{\sum_{i=1}^{Ns}\left(\left[y^i_{target} - y^i_{predict}\right] - Mean\right)^2}{Ns}} \tag{14}$$

$$MAPE = \frac{1}{Ns}\sum_{i=1}^{Ns}\left|\frac{y^i_{target} - y^i_{predict}}{y^i_{target}}\right| \tag{15}$$

$$NMSE = \frac{MSE}{\sigma_t^2} \tag{16}$$

where σ_t^2 indicates variance of targets:

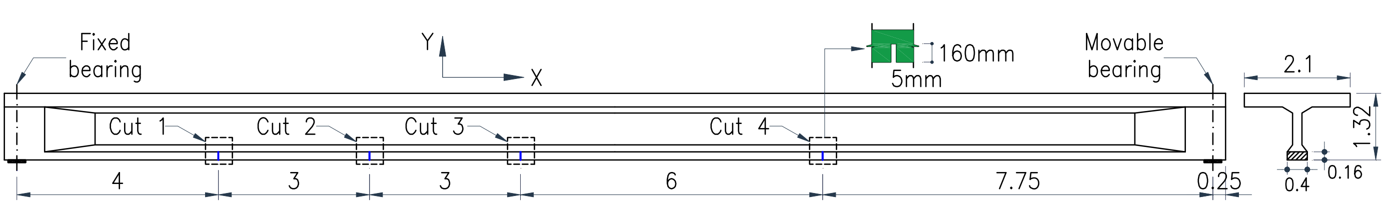

Fig. 5. Schematic of a 24.25-meter concrete T-girder (A) and damage scenarios of Case 2 (unit: m).

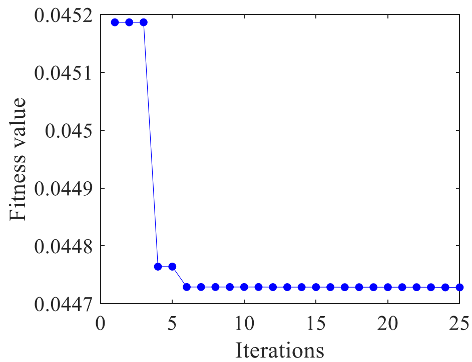

Fig. 6. Evolution of the fitness function using MPA with 10 search agents and 25 iterations.

Table 1
Comparison of natural frequencies (**Error = (Measurement − FEM)× 100 %/ Measurement)*.

Mode	Measurement	$FEM_{current}$	Error (%)*	FEM [43]	Error (%)*
1	5.74	5.735	**0.09** ↓	5.69	0.87
2	21.89	22.125	−1.07 ↑	21.94	−0.23
3	45.95	45.95	**0** ↓	45.49	1
4	72.35	69.953	**3.31** ↓	69.04	4.57

$$\sigma_t^{\,2} = \frac{1}{Ns}\sum_{i=1}^{Ns}\left(y^i_{target} - \overline{y}_{target}\right)^2 \tag{17}$$

where $\overline{y}_{target}$ indicates the mean of targets.

$$VAF = \left(1 - \frac{var\left(y_{target} - y_{predict}\right)}{\sigma_t^{\,2}}\right) \times 100 \tag{18}$$

3. Application of the two-step approach to damage detection

3.1. Case study description and damage scenarios

This research investigates two structures. The first is a simply supported flat beam with a rectangular cross-section as introduced in [43]. The published modal properties of the beam are utilized to build the FE model in the current study. The second structure is a simply supported concrete double T-girder bridge, with assumed material properties [44]. Detailed descriptions of the two cases and damage scenarios are presented in Fig. 4 and Fig. 5.

Three damage scenarios are introduced in the flat beam, with cuts of 4 × 50 mm created on both sides and along the beam as in Fig. 4. CS1: single damage with cut 1. CS2: multiple damage with cuts 1 and 2. CS3: multiple damage with cuts 1, 2, and 3.

Fig. 5 presents four open cuts (5 × 160 mm) in the bottom flange of the girder (A). This study examines two single-damage scenarios (CS1, CS2) and two multiple-damage scenarios (CS3, CS4). CS1 and CS2 correspond to cuts 1 and 2, respectively, while CS3 includes both cuts 1 and 2, and CS4 consists of cuts 3 and 4.

3.2. Finite element model for modal analysis

a. Experiment-based flat beam

Due to its thin thickness ($t = 0.01$ m), the flat beam exhibits plate-like behavior. Consequently, SHELL181 elements, which feature three translational and three rotational degrees of freedom per node, are employed in ANSYS [45] is used to ensure accurate analysis while maintaining computational efficiency (see Fig. 7). Additionally, using SHELL181 facilitates generating cuts on the beam for further damage identification studies (refer to Fig. 4). The mesh properties of the FE model are as follows: element size of 5 mm, hexahedral (brick) element shape, and a mapped mesh. Modal analysis of FE models is performed using the Block Lanczos method.

The material properties in the FE model are obtained through a model updating process that utilizes the measured dynamic properties of the intact beam. For further details, please refer to the reference [43]. The objective function of the optimization process is the difference in frequencies using Eq. (19), with Young's modulus (E) and weight density (γ) being the two variables involved in this process. Young's modulus ranges from 1.9×10^5 to 2.1×10^5 (MPa), while the weight density is investigated within the range of 7750 and 8050 Kg/m^3. Poisson's ratio $\nu_s = 0.3$.

$$fitness = \sum_{i=1}^{N}\frac{\left|f^i_m - f^i_c\right|}{f^i_m}, \tag{19}$$

with f_m and f_c are measured and calculated frequencies.

MPA is employed to optimize the values of Young's modulus and weight density. Values of two variables are randomly initialized to determine the fitness values. Throughout the optimization process, the fitness function converges to a minimum value of 0.04472 after 6 iterations, as shown in Fig. 6. The optimized objective function yields values of $E = 2.0481 \times 10^5$ *MPa* and $\gamma = 7826$ *Kg/m*^{3}.

Differences in frequencies between the FE model and measurement are shown in Table 1. The maximum absolute error between simulation

Fig. 8. The first two vertical bending modes of the intact beam.

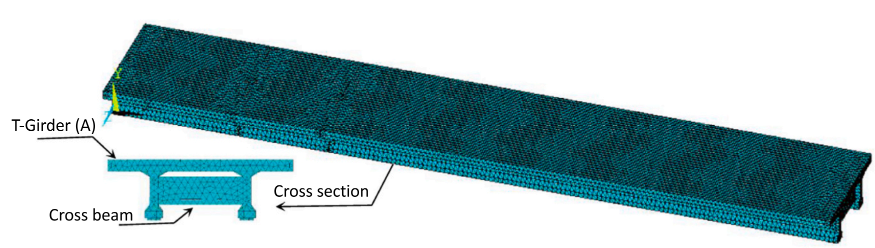

Fig. 9. FE model of the simply supported T-girder bridge.

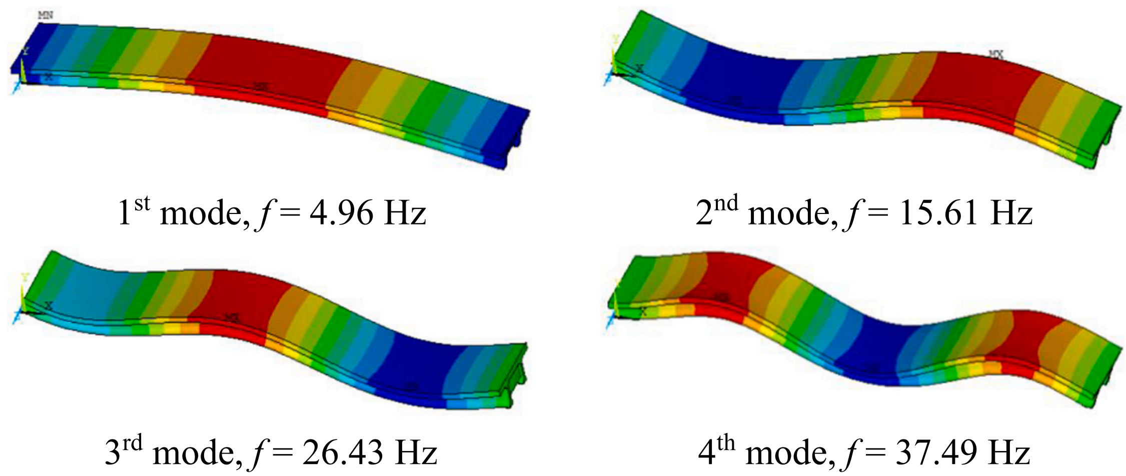

Fig. 10. The first four vertical bending modes of the intact T-girder bridge.

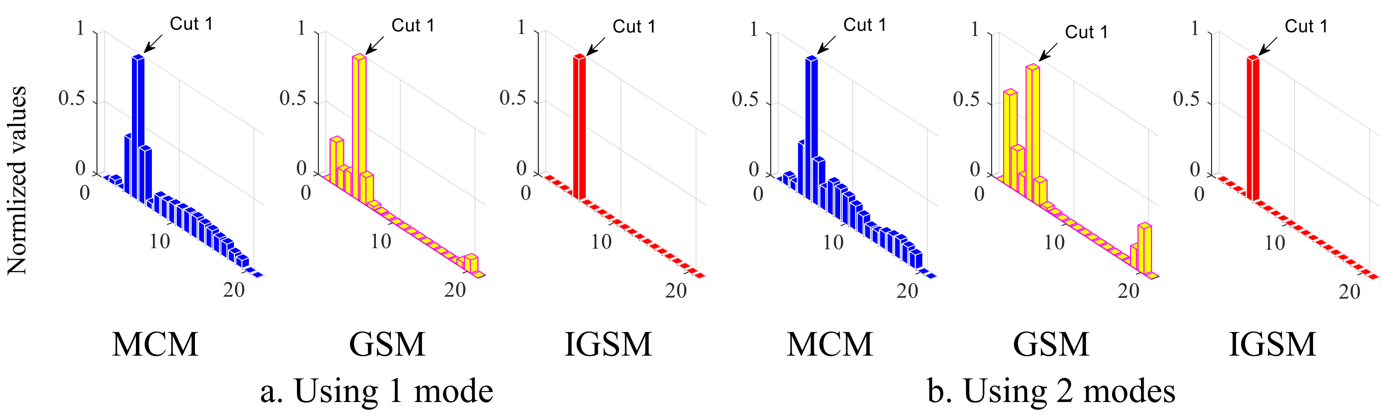

Fig. 11. Damage localization of CS1 - Case 1.

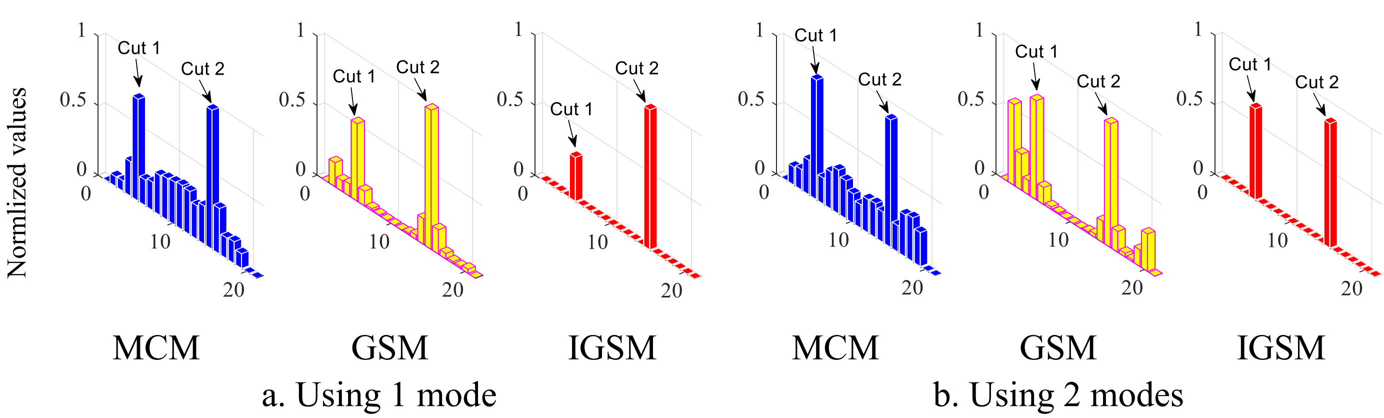

Fig. 12. Damage localization of CS2 - Case 1.

and measurement is below 3.31 %. Consequently, these material properties are used as inputs for the FE model. The first two vertical bending mode shapes are extracted from the FE analysis for the damage identification procedure (see Fig. 8)

b. Concrete T-girder bridge

The concrete T-girder is modeled by SOLID185 elements as shown in Fig. 9. Mesh properties of the FE model of the bridge: element sizes of 0.05 m and 0.15 m, tetrahedral elements, and free meshing. The bridge cross-section consists of two girders based on published paper [44]. The T-Girder (A) is assumed to undergo damage scenarios. The material properties in the FE model include a concrete compressive strength of $f'_c = 40$ Mpa, Young's modulus $E_c = 3.61 \times 10^4$ MPa, weight density $\gamma_c = 2600$ kg/m^3, and Poisson's ratio $\nu_c = 0.2$. Modal analysis of FE models is also conducted using the Block Lanczos method. The first four modes are presented in Fig. 10.

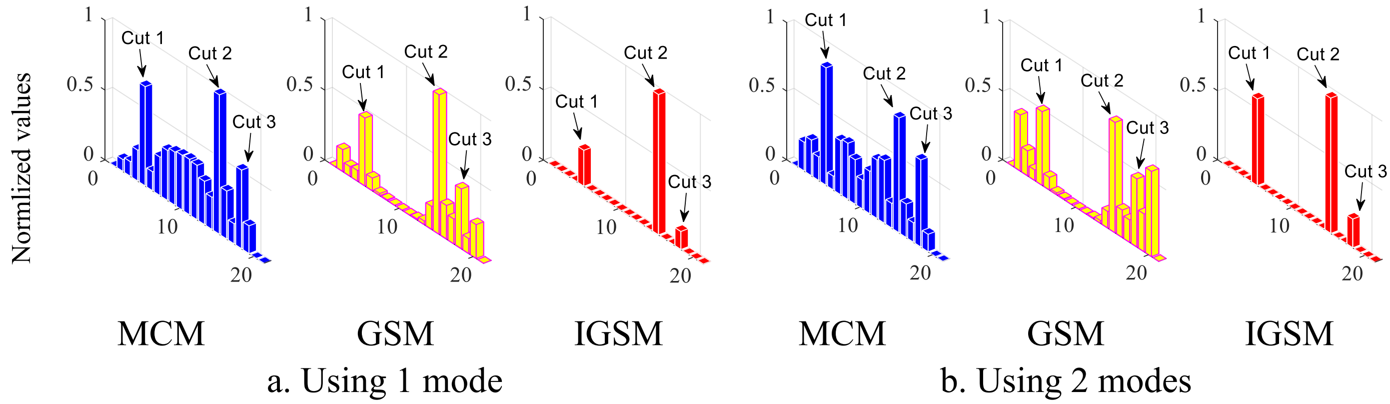

Fig. 13. Damage localization of CS3 - Case 1.

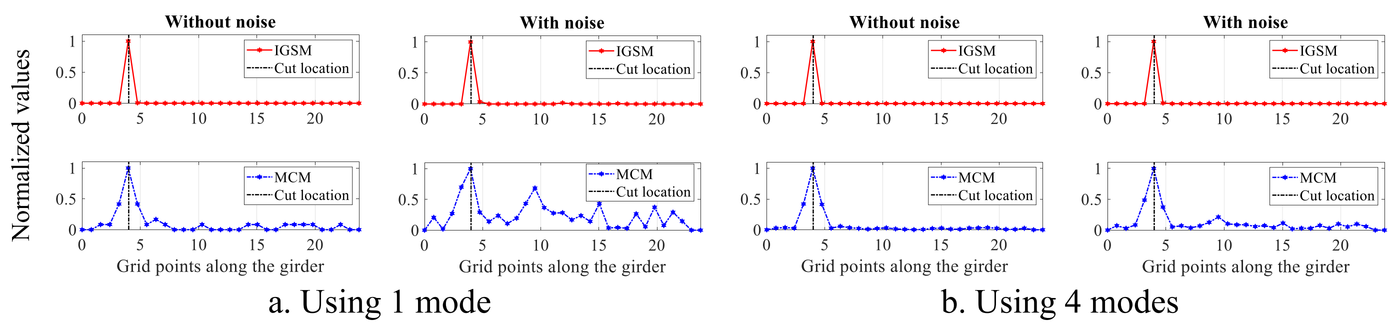

Fig. 14. Damage localization of CS1 - Case 2.

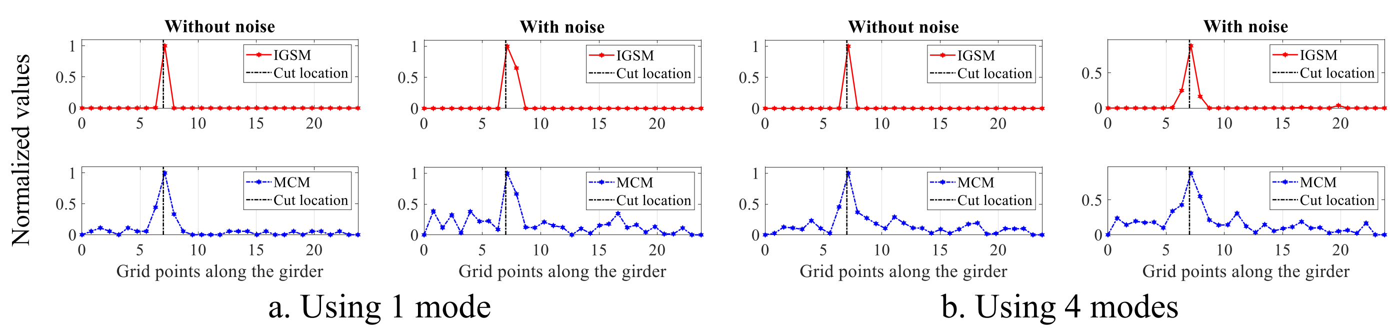

Fig. 15. Damage localization of CS2 - Case 2.

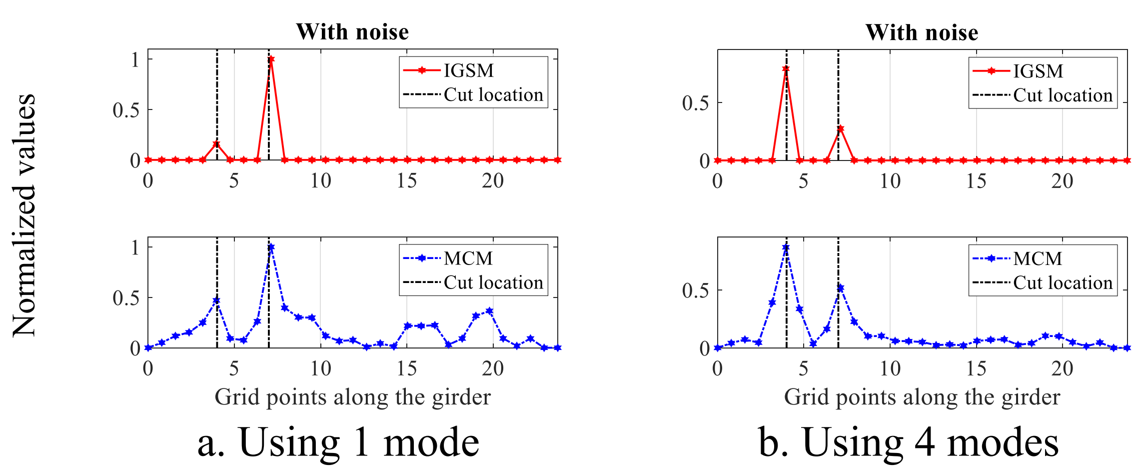

Fig. 16. Damage localization of CS3 - Case 2.

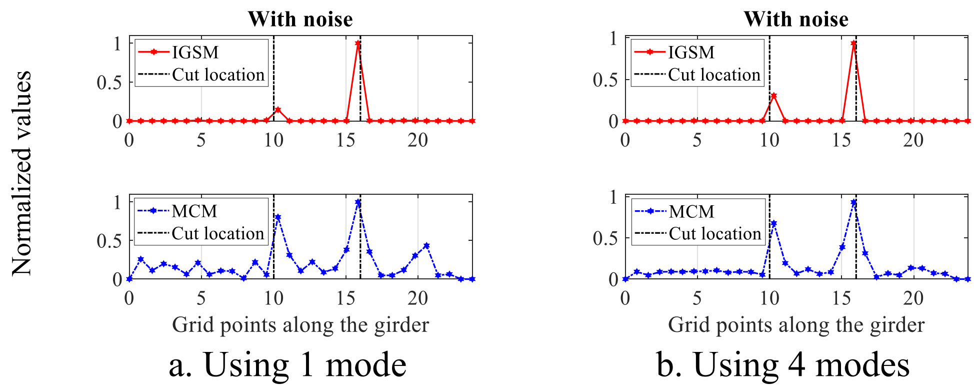

Fig. 17. Damage localization of CS4 - Case 2.

4. Damage localization using damage index

4.1. Damage localization for Case 1

a. Single damage scenario, CS1

Generally, the first five modes significantly affect the global structural response and can be utilized for damage identification. In the first case, the mode shapes of the first and second modes of the damaged beam are collected to evaluate the proposed damage index. This evaluation aims to assess the effectiveness of the proposed approach in scenarios where only lower-order modes are obtained. Localization of the cut is performed through three indicators: MCM [13], GSM [21,22], and IGSM. To facilitate comparison, these damage indices are normalized to the range $[0-1]$.

By determining irregularities in the calculated indices, GSM can indicate the position of cuts in the beam. However, GSM also shows peaks in other areas of the beam where no cut is presented, which could lead to false detections (please see Fig. 11.). After applying Eqs. (4) and (5) the developed damage indicators are clearly highlighted at the intended damage location. This failure location is also verified by MCM using dynamic properties from the damaged and undamaged states. Both GSM and MCM exhibit peaks around the damage location, whereas IGSM displays a peak only at the predetermined cut location. This indicates that IGSM has a superior ability to pinpoint the damaged zone compared to the other methods.

b. Multiple damage scenarios, CS2 and CS3

The same procedure is implemented across multiple damage scenarios. Generally, successful identification of cuts at different locations is achieved by each investigated index (refer to Fig. 12 and Fig. 13). The clarity of damage location varies across the different methods. GSM, while determining the correct location of damage, also exhibits some peaks at undamaged positions, particularly near the bearings. MCM performs similarly to IGSM in locating damage across multiple scenarios. However, IGSM more clearly highlights the cuts, minimizing minor peaks in undamaged areas.

Additionally, incorporating mode 2 using Eq. (6) enhances the accuracy of damage location identification by increasing the magnitude of the damage indicator, as shown in Fig. 12b and Fig. 13b. Based on the results from Case 1, IGSM demonstrates superior performance in damage localization, followed by MCM and GSM.

4.2. Damage localization for Case 2

a. Single damage scenarios CS1 and CS2

In this complex structure, the effectiveness of the proposed index in locating damage is verified using mode 1 and the first four modes. The mode shapes of the girder (A) are collected and randomly corrupted by white Gaussian noise using the "awgn" function in MATLAB, with a signal-to-noise ratio (SNR) of 70. This process aims to evaluate the practical applicability of the proposed index when noise is present in the measured data. Notably, only the two indices, IGSM and MCM, are used to identify the cuts in the T-girder as they demonstrated superior efficiency in the previous section.

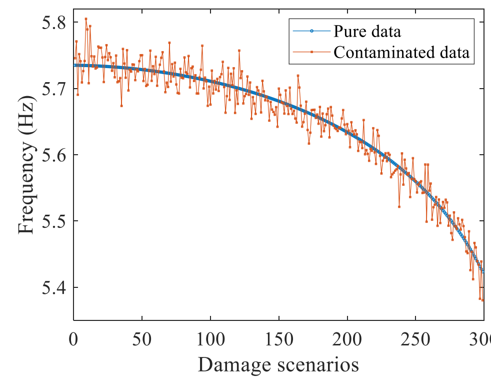

Fig. 18. Relationship between the first frequency and the increase in cut length.

At first glance, both the MCM and IGSM methods accurately identify locations of the cut (at 4 m and 7 m from the bearings). However, the presence of noise in the mode shape data adversely affects the accuracy of damage detection. With MCM, more peaks are observed at undamaged locations compared to IGSM. This discrepancy arises because MCM requires mode shape data from both damaged and undamaged states, which introduces additional errors in pinpointing damage locations. In contrast, IGSM only requires mode shape data from a single state and employs a truncation technique, resulting in fewer false peaks at non-damage zones (Fig. 14a) or near the cut location (Fig. 15a). Overall, IGSM consistently demonstrates superior accuracy in highlighting the cut position compared to MCM. Additionally, utilizing higher mode shapes enhances the accuracy of identifying cut locations in beams by minimizing erroneous peaks in intact regions (Figs. 14b and 15b).

b. Multiple damage scenarios CS3 and CS4

In this multiple-damage scenario, only noisy mode shape data is used to evaluate the performance of the two indicators. Fig. 16 and Fig. 17 present the results of determining cuts located at the bottom flange, specifically at positions (4 m, 7 m) and (10 m, 16 m) from the bearing, respectively.

The influence of noise is noticeable in both cases. When using MCM

Table 2
Hyperparameters used in ensemble models for CS1 - Case 1, CS4-Case 2.

Model	numTrees	MinLeaf	MinParentSize	MaxNumSplit	LearnRate
LSBoost	250	5	Default	Default	0.01
OAs-LSBoost	1 – 500	1 – 50	1 – 50	1 – 100	0.001 – 1

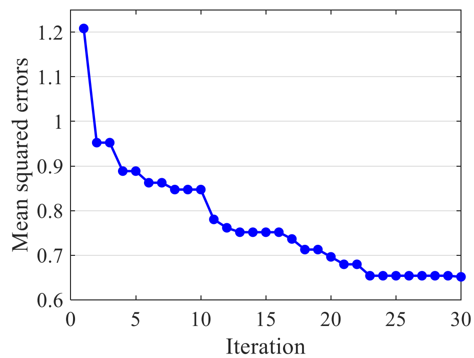

Fig. 19. Convergence curve using MPA-LSBoost.

with only the first mode, identifying damaged locations becomes challenging due to the appearance of peaks at undamaged locations (see Figs. 16a and 17a). However, this effect diminishes as the number of modes used increases to four, as shown in Figs. 16b and 17b. In contrast, by detecting sharp curvature changes in the GSM index and applying a threshold as in Eq. (5), the IGSM indicator, despite noise, effectively identifies damage locations and removes most erroneous peaks around the cut as well as non-damaged zones. Additionally, averaging values from four modes further enhances damage localization. In this case study, IGSM proves more consistent and effective than MCM in localizing damage within a smaller predicted area while minimizing the impact of noise in mode shapes.

5. Damage quantification

In this step, the extent of damage, specifically the cut length in CS1-Case 1 and CS4-Case 2, is determined using the ensemble learning model.

5.1. CS1-Case 1: damage quantification for the single cut located 4 m from the left bearing

a. Data preparation

Several length scenarios of the cut are simulated in the beam to generate a training dataset. The cut lengths range from 0 to 150 (mm), with increments of 0.5 mm. The dataset consists of 299 samples with 80 % allocated for training and the remaining 20 % reserved for testing. The “awgn” function in MATLAB is used to introduce white Gaussian noise to both the input and output training data, with a signal-to-noise ratio (SNR) of 70. Fig. 18 shows the variation in the first frequencies of the input data, accounting for the presence of Gaussian noise.

b. *Training process*

LSBoost models’ structure consists of five inputs and one output. In this application, natural frequencies of the first five vertical bending modes of the beam serve as inputs, while the output is the cut length.

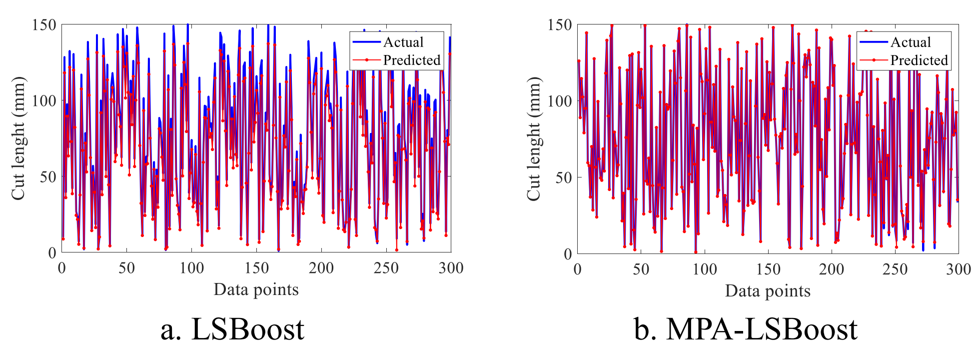

Fig. 20. Error observation.

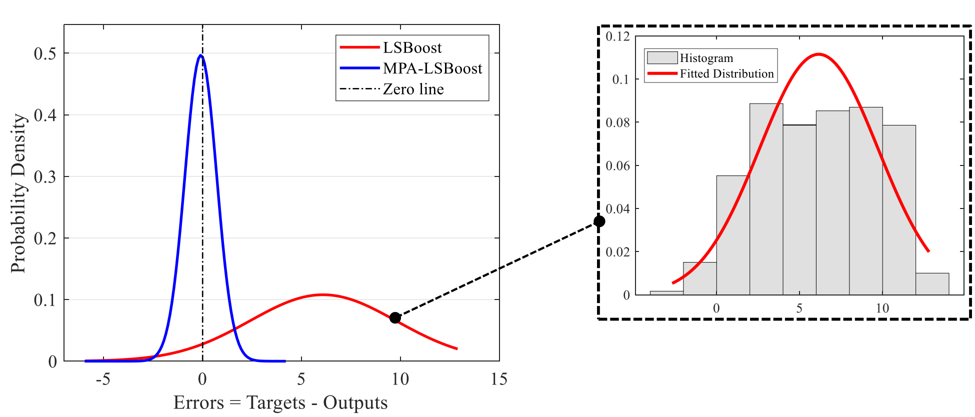

Fig. 21. Error distribution for CS1-Case 1.

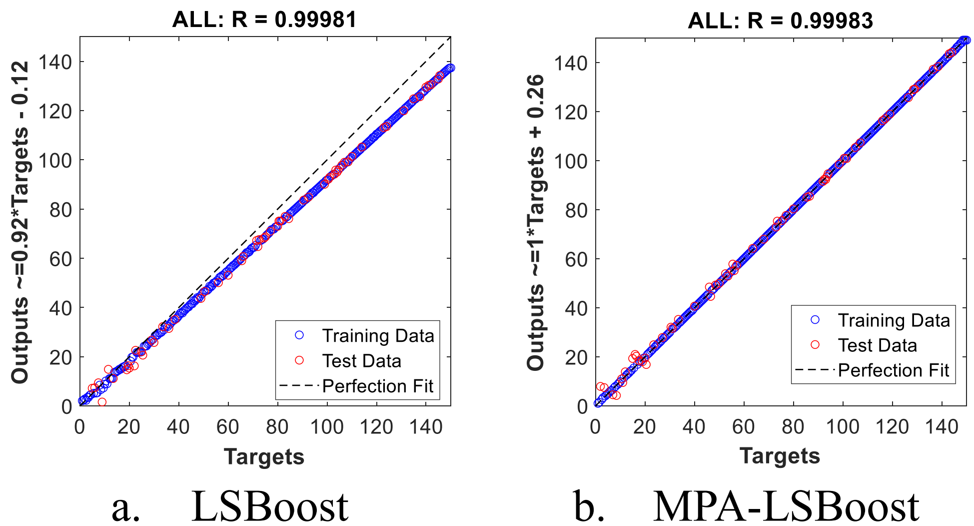

Fig. 22. Regression plots for CS1 – Case 1.

Table 3
Evaluation metrics for CS1 – Case 1.

Model	R^2	Adjusted R^2	Mean	Std	MSE	MAPE	NMRE	VAF
LSBoost	0.9996	0.9996	6.17	3.55	50.62	9.79	0.0271	99.33
MPA-LSBoost	**0.9997**	**0.9997**	**−0.091**	**0.8**	**0.65**	**2.51**	**3.49×10^{-4}**	**99.97**

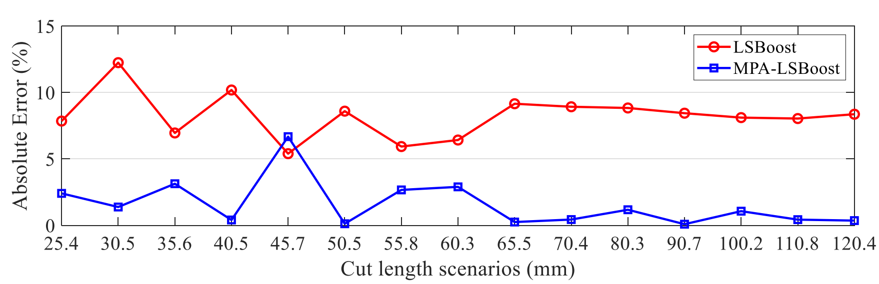

Fig. 23. Differences in cut length estimation using LSBoost and MPA-LSBoost.

Marine predator algorithm (MPA) is utilized to control the stochastic optimization procedure. In the first iteration, each of the 50 particles holds a set of LSBoost hyperparameters selected from the range in Table 2, which is used to calculate MSE between the predicted and actual values in the training dataset.

The optimization process continues until it exceeds 30 iterations, ultimately achieving an optimized MSE of 0.65. The evolution of the fitness values is presented in Fig. 19.

c. *Performance evaluation*

LSBoost serves as a reference to assess the effectiveness of MPA-LSBoost in quantifying damage. In LSBoost, the NumTree, Minleaf, and LearnRate are set as 250, 5, and 0.01, respectively. Other parameters are as defaults. The optimized hyperparameters for MPA-LSBoost: numTrees, Minleaf, MinParentSize, MaxNumSplit, and LearnRate are (498, 7, 38, 51, 0.089).

Fig. 20 shows the deviation between the estimated and desired values across the entire dataset. On initial inspection, MPA-LSBoost provides a better match between actual and predicted values. In contrast, the lower predicted values of cut length are observed with LSBoost. To gain further insights into the model's performance, several statistical metrics are determined using the obtained results.

- First, the error distribution for each model is illustrated with a histogram in Fig. 21, where the distribution shape, peak height and spread, and error magnitude are discussed. The MPA-LSBoost's error distribution is concentrated around the zero line, while LSBoost has a broader spread and skews to the right. This suggests that MPA-LSBoost generates smaller errors or is less biased compared to LSBoost. Besides, MPA-LSBoost shows a higher peak in probability density than LSBoost, indicating that more predictors are closer to the actual values. The wider spread or range of errors in LSBoost's distribution shows that this model has more variability in its error which may make it less reliable compared to MPA-LSBoost.

- Fig. 22 visualizes the relationship between the actual and predicted values derived from the two prediction models. Both models show high correlation coefficients such as 0.99981 and 0.99983, meaning they effectively capture the trend data. However, LSBoost has a slope of 0.92, which is slightly below 1. Therefore, LSBoost tends to produce lower predicted values for cut length. The slope of 1 in MPA-LSBoost is closer to the ideal slope of the fit line, indicating less bias in the predictions. Additionally, nearly all data points are closely aligned with the perfect fit line. This alignment indicates better prediction accuracy and consistency between training and testing datasets.

- Evaluation metrics are shown in Table 3. It can be seen that the residual errors, measured by Mean, MSE, MAPE, and NMRE, obtained by MPA-LSBoost are significantly smaller than LSBoost. These findings prove that MPA-LSBoost provides more precise predictions relative to

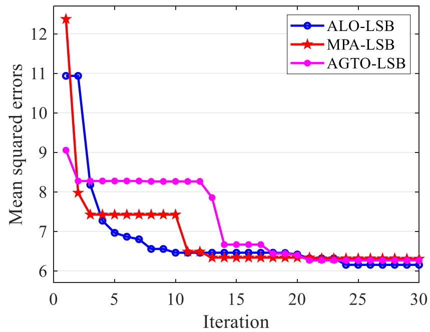

Fig. 24. Convergence curve of OAs-LSBoost with 30 search particles and 30 iterations.

the actual values. Moreover, higher values of R^2, Adjusted R^2, and VAF show that MPA-LSBoost can explain nearly all the variance in the target data. A lower standard deviation std of 0.8, compared to 3.55 for LSBoost, reveals that MPA-LSBoost produces more consistent predictions.

To evaluate the trained model's ability to estimate cut length, fifteen cut length scenarios were generated such as 25.4, 30.5, 35.6, 40.5, 45.7, 50.5, 55.8, 60.3, 65.5, 70.4, 80.3, 90.7, 100.2, 110.8, 120.4 (mm). The modal properties from these damage scenarios were corrupted with an SNR of 70 before being input into the trained model for damage quantification. Due to the consideration of noise, the cut scenario of 45.7 mm predicted by MPA-LSBoost shows differences of approximately 6.7 % (see Fig. 23). However, for the remaining scenarios, MPA-LSBoost consistently produces lower errors, staying below 3 %, while LSBoost exhibits a higher error range of 5.4–12.3 %.

5.2. CS4-Case 2: damage quantification for the two cuts located 10 m and 16 m from the left bearing

a. *Data preparation*

Unlike Case 1, this scenario examines the prediction model's effectiveness when training data is limited. Instead of collecting cut length changes at 0.5 mm intervals as in Case 1, larger increments are used. Therefore, cut-length data, ranging from 0 to 155 mm in increments of 3 mm, were collected to build an ML model. In this case, two cuts are assumed to have the same cut length, rather than different cut lengths. Additionally, the noise incorporated into the training and test data remains consistent with Case 1. The first five frequencies of the bridge and the corresponding cut length were collected as training datasets. The data were split into training and test sets with an 80–20 proportion.

Additionally, an artificial neural network (ANN) with an architecture of $I_{node} \times H_{node} \times O_{node} = 5 \times H_{node} \times 1$ also utilized for verification purposes using the same training and testing proportions of 80 % and 20 %, respectively. The hidden size is assumed as $H_{node} = 3 \times I_{node} = 15$, where H_{node}, I_{node}, and O_{node} denote the number of nodes of the hidden, input, and output layers, respectively. The activation function and training algorithm are sigmoid and Levenberg-Marquardt.

b. *Training process*

The structure of LSBoost is the same as in Case 1. In Case 2, MPA, ALO, and AGTO are employed to reinforce the efficiency of the proposed approach. Each algorithm has its strategy for randomly initializing values of the investigated variables, resulting in different procedures for identifying the objective function (refer to Fig. 24). ALO-LSBoost achieves rapid convergence, while AGTO-LSBoost begins to converge around the halfway point. The convergence curve of MPA-LSBoost consists of three distinct phases, each spanning approximately one-third of the iterations [41,42]. After 25 iterations, the convergence curves of ALO-LSBoost, MPA-LSBoost, and AGTO-LSBoost stabilize, reaching values of 6.15, 6.31, and 6.27, respectively, by the 30th iteration. The optimized parameters e.g. numTrees, Minleaf, MinParentSize, MaxNumSplit, and LearnRate for LSBoost are (238, 2, 5, 44, 0.1598), (182, 2, 6, 62, 0.2055), and (81, 2, 6, 34, 0.1468) respectively.

c. *Performance evaluation*

Fig. 25 sheds light on the distribution of errors in the 5 investigated models through a probability density plot. ANN produces larger positive errors compared to other models, with an error range from (–9.87)–(28.66) (see Table 4). The LSBoost model follows, with its error range skewed to the right. The three OAs-LSBoost models have error distributions that are more centered around zero and narrower. This means these models perform more accurately and consistently in minimizing errors.

The correlation coefficients in Fig. 26 are relatively high, indicating that these models capture the overall trend in the data well. The highest coefficients are achieved by OAs-LSBoost models, with all values exceeding 0.998, followed by LSBoost and ANN at 0.997 and 0.991, respectively. In the scatter plot, the data predicted by ANN lies slightly below the perfect fit line, while the LSBoost predictions deviate further downward. In contrast, the OAs-LSBoost models align data points almost precisely along the 45-degree line for both training and testing datasets.

Table 4
Error ranges of 5 models.

LSBoost	ANN	ALO-LSBoost	MPA-LSBoost	AGTO-LSBoost
−3.55 ÷ 22.38	−9.87 ÷ 28.66	−9.4 ÷ 8.65	−9.85 ÷ 8.59	−9.98 ÷ 8.34

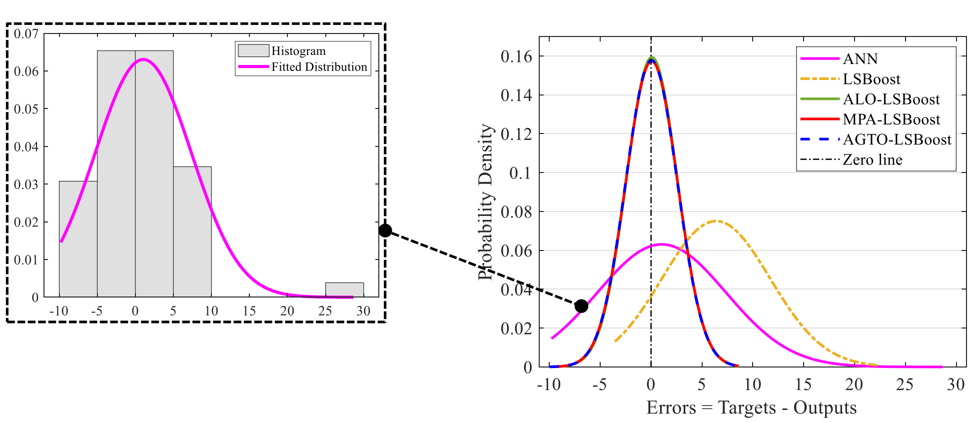

Fig. 25. Error distribution for CS4-Case 2.

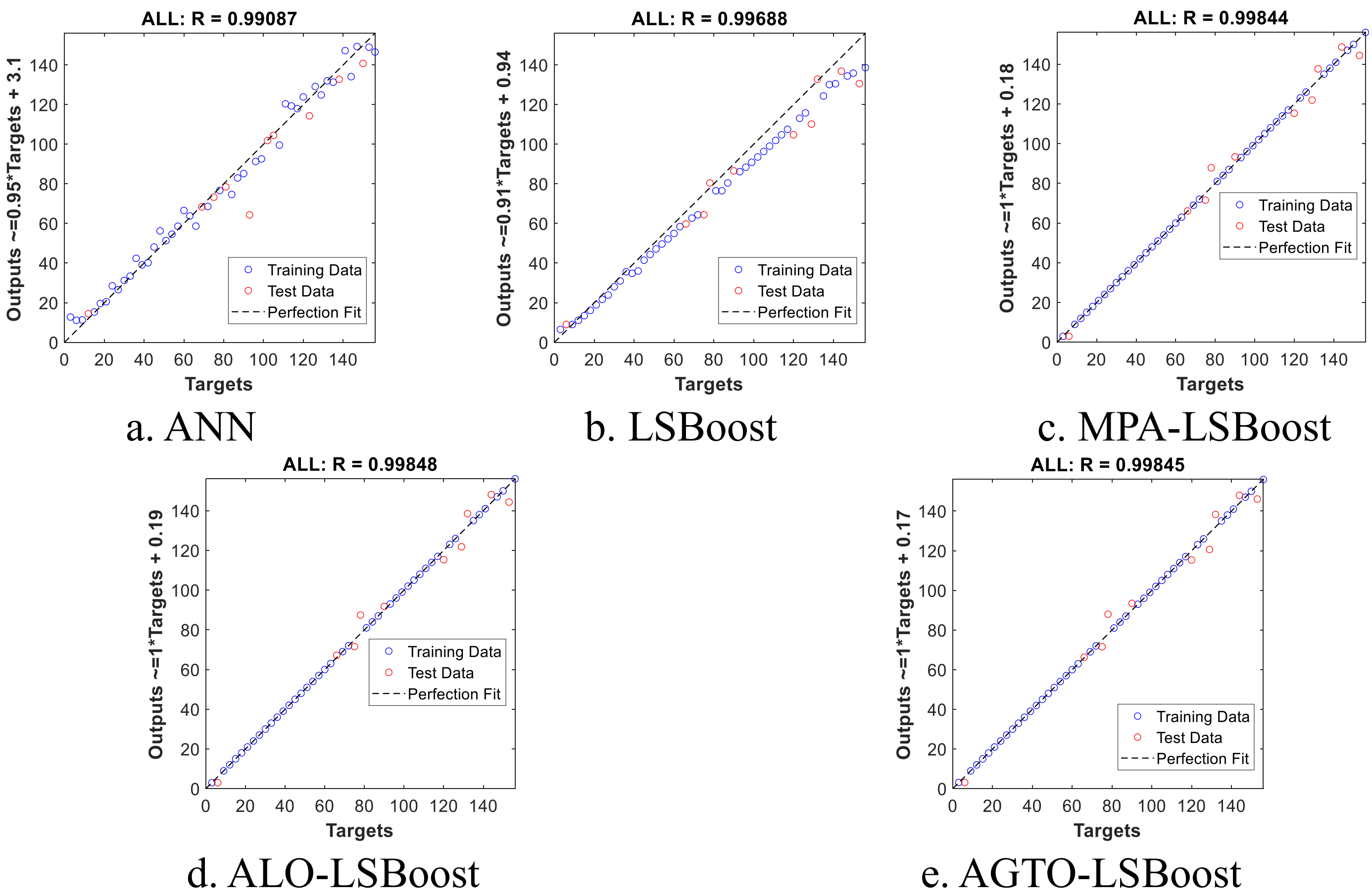

Fig. 26. Regression plots for all datasets CS4 - Case 2.

Table 5
Evaluation metrics for CS4 - Case 2.

Model	R^2	Adjusted R^2	Mean	Std	MSE	MAPE	NMRE	VAF
ANN	0.9818	0.9815	1.05	6.33	40.34	13.86	0.0195	98.06
LSBoost	0.9938	0.9936	6.39	5.31	68.51	11.67	0.0331	98.63
ALO-LSBoost	**0.9970**	**0.9969**	**0.0797**	**2.50**	**6.15**	**1.79**	**0.0030**	**99.70**
MPA-LSBoost	**0.9969**	**0.9968**	**0.0606**	**2.53**	**6.31**	**1.80**	**0.0031**	**99.69**
AGTO-LSBoost	**0.9969**	**0.9968**	**0.0459**	**2.53**	**6.27**	**1.88**	**0.0030**	**99.69**

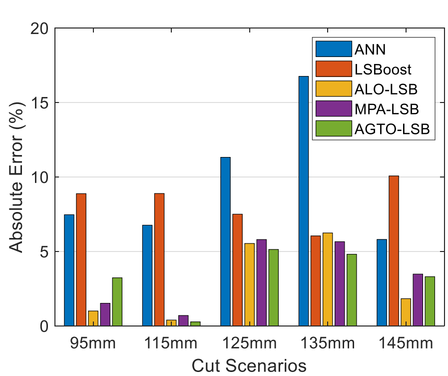

Fig. 27. The absolute errors between predicted and actual values under noisy condition.

This means OAs-LSBoost models not only show consistency between training and testing data but also yield minimal errors.

In Table 5, although all models achieve R^2 values exceeding 98 %, indicating strong predictive capabilities, notable improvements are observed in evaluation metrics. Specifically, OAs-LSBoost models outperform the others across all metrics, with R^2, adjusted R^2, and VAF values exceeding 0.996. These models also achieve the lowest mean errors, approximately 0.0459–0.0797, and the smallest error standard deviations, around 2.50–2.53. Furthermore, MAPE and NMRE values for these models are lower than those of the alternatives, highlighting their superior predictive accuracy.

ANN and LSBoost perform reasonably well but fall short compared to the top three models. Their higher mean, standard deviation, MSE, MAPE, and NMRE suggest a larger error spread, making them less reliable and less stable than the OAS-LSBoost models.

Finally, the trained models' efficiency then is verified by several cut length scenarios e.g. 95, 115, 125, 135, 145 (mm). Errors of cut length estimation are presented in Fig. 27. The errors of the OAs-LSBoost models fluctuate between 0.26 % and 6.23 %. The errors observed by the LSBoost range from 6.0 % to 10 %. ANN exhibits the largest errors, falling within the range of 5.8–16.75 %. These findings reinforce the effectiveness of the proposed approach in enhancing LSBoost's performance.

6. Conclusions

In this study, a modified damage index, IGSM, is developed to localize cut locations within the investigated structures. Subsequently, ensemble models optimized with advanced algorithms are employed to quantify the cut length. The key findings are outlined below:

- The proposed approach accurately identifies cut locations using only the displacement mode shapes of damaged structures. While the fundamental mode sufficiently supports the method, averaging values across higher modes can further enhance the reliability of cut identification in beams. The index accurately identifies the positions of both single and multiple cuts, even in the presence of white Gaussian noise in the mode shape data. Applying a truncate technique reduces false peaks in undamaged zones, enhancing the clarity of cut localization. Comparisons with MCM and GSM further confirm the efficiency of the improved damage index.
- In the current study, the OAs-LSBoost models effectively enhance LSBoost's performance in damage quantification. Hyperparameter optimization through a stochastic process minimizes errors between predicted and actual values, while maintaining LSBoost's predictive consistency under noisy conditions. The effectiveness of the proposed approach is demonstrated through visual representations, including probability density plots, regression plots, and statistical metrics.

CRediT authorship contribution statement

Magd Abdel Wahab: Investigation, Supervision, Validation, Writing – review & editing. **Ho Viet Long:** Conceptualization, Data curation, Methodology, Software, Writing – original draft. **Bui-Tien Thanh:** Supervision, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research is funded by University of Transport and Communications (UTC) under grant number T2023-PHII-007TĐ.

Data availability

Data will be made available on request.

References

[1] Chang K-C, Kim C-W. Modal-parameter identification and vibration-based damage detection of a damaged steel truss bridge. Eng Struct 2016;122:156–73. https://doi.org/10.1016/j.engstruct.2016.04.057.
[2] Soo Lon Wah W. Damage detection of structures under changing environmental and operational conditions using the COVRATIO statistic. Eng Struct 2023;281:115704. https://doi.org/10.1016/j.engstruct.2023.115704.
[3] Huang M-S, Gül M, Zhu H-P. Vibration-based structural damage identification under varying temperature effects. J Aerosp Eng 2018;31(3):04018014. https://doi.org/10.1061/(ASCE)AS.1943-5525.0000829.
[4] M. Huang, J. Zhang, J. Li, Z. Deng, and J. Luo, "Damage identification of steel bridge based on data augmentation and adaptive optimization neural network," Structural Health Monitoring, p. 14759217241255042, 2024, doi: 10.1177/14759217241255042.
[5] Zhang J, Huang M, Wan N, Deng Z, He Z, Luo J. Missing measurement data recovery methods in structural health monitoring: the state, challenges and case study. Measurement 2024;231:114528. https://doi.org/10.1016/j.measurement.2024.114528.
[6] Huang M, Wan N, Zhu H. Reconstruction of structural acceleration response based on CNN-BiGRU with squeeze-and-excitation under environmental temperature effects. J Civ Struct Health Monit 2024. https://doi.org/10.1007/s13349-024-00859-w.
[7] Lin Y, Nie Z, Ma H. Dynamics-based cross-domain structural damage detection through deep transfer learning. Comput Aided Civ Eng 2022;37(1):24–54. https://doi.org/10.1111/mice.12692.
[8] Shirazi MI, Khatir S, Boutchicha D, Abdel Wahab M. Feature extraction and classification of multiple cracks from raw vibrational responses of composite beams using 1D-CNN network. Compos Struct 2024;327:117701. https://doi.org/10.1016/j.compstruct.2023.117701.
[9] Huang M, Zhang J, Hu J, Ye Z, Deng Z, Wan N. Nonlinear modeling of temperature-induced bearing displacement of long-span single-pier rigid frame bridge based on DCNN-LSTM. Case Stud Therm Eng 2024;53:103897. https://doi.org/10.1016/j.csite.2023.103897.
[10] Nguyen-Ngoc L, et al. A two-step approach for damage identification in bridge structure using convolutional Long Short-Term Memory with augmented time-series data. Adv Eng Softw 2024;198:103795. https://doi.org/10.1016/j.advengsoft.2024.103795.
[11] H. Tran-Ngoc et al., "A Novel Approach to Damage Assessment in Structures Based on Artificial Neural Network Working Parallel With a Hybrid Stochastic Optimization," Dec. 29, 2021. doi: 10.21203/rs.3.rs-1191581/v1..
[12] Huu QN, Ngoc LN, Tien TB, Ngoc HT, Tran HN, Xuan TN. An improved Artificial Rabbit Optimization for structural damage identification. Lat Am J Solids Struct 2024;21(1):e523. https://doi.org/10.1590/1679-78257810.
[13] Pandey AK, Biswas M, Samman MM. Damage detection from changes in curvature mode shapes. J Sound Vib 1991;145(2):321–32. https://doi.org/10.1016/0022-460X(91)90595-B.
[14] Pandey AK, Biswas M. Experimental verification of flexibility difference method for locating damage in structures. J Sound Vib 1995;184(2):311–28. https://doi.org/10.1006/jsvi.1995.0319.
[15] Nguyen DH, Nguyen QB, Bui-Tien T, De Roeck G, Abdel Wahab M. Damage detection in girder bridges using modal curvatures gapped smoothing method and Convolutional Neural Network: application to Bo Nghi bridge. Theor Appl Fract Mech 2020;109:102728. https://doi.org/10.1016/j.tafmec.2020.102728.
[16] Dawari VB, Vesmawala GR. Modal curvature and modal flexibility methods for honeycomb damage identification in reinforced concrete beams. Procedia Eng 2013;51:119–24. https://doi.org/10.1016/j.proeng.2013.01.018.
[17] Porcu MC, Patteri DM, Melis S, Aymerich F. Effectiveness of the FRF curvature technique for structural health monitoring. Constr Build Mater 2019;226:173–87. https://doi.org/10.1016/j.conbuildmat.2019.07.123.
[18] Ho LV, Nguyen DH, De Roeck G, Bui-Tien T, Wahab MA. Damage detection in steel plates using feed-forward neural network coupled with hybrid particle swarm optimization and gravitational search algorithm. J Zhejiang Univ Sci A 2021;22(6):467–80. https://doi.org/10.1631/jzus.A2000316.
[19] Ho Viet L, Trinh Thi T, Ho Xuan B. Swarm intelligence-based technique to enhance performance of ANN in structural damage detection. vn 2022;73(1):1–15. https://doi.org/10.47869/tcsj.73.1.1.
[20] Ho LV, Trinh TT, De Roeck G, Bui-Tien T, Nguyen-Ngoc L, Abdel Wahab M. An efficient stochastic-based coupled model for damage identification in plate structures. Eng Fail Anal 2022;131:105866. https://doi.org/10.1016/j.engfailanal.2021.105866.
[21] Ratcliffe CP. Damage detection using a modified laplacian operator on mode shape data. J Sound Vib 1997;204(3):505–17. https://doi.org/10.1006/jsvi.1997.0961.
[22] Yoon MK, Heider D, Gillespie JW, Ratcliffe CP, Crane RM. Local damage detection using the two-dimensional gapped smoothing method. J Sound Vib 2005;279(1–2):119–39. https://doi.org/10.1016/j.jsv.2003.10.058.
[23] M. Thiene, M. Zaccariotto, U. Galvanetto, and M. Gherlone, "Application of smoothing techniques to damage detection using Proper Orthogonal Decomposition," p. 12.
[24] Yoon MK, Heider D, Gillespie JW, Ratcliffe CP, Crane RM. Local damage detection with the global fitting method using mode shape data in notched beams. J Nondestruct Eval 2009;28(2):63–74. https://doi.org/10.1007/s10921-009-0048-6.
[25] Ho Viet L, Ho Vinh H, Vu Van T, Ho Xuan B. An improved version of mode shape based indicator for structural damage identification. vn 2023;74(7):790–804. https://doi.org/10.47869/tcsj.74.7.3.
[26] Seguini M, Djamel N, Djilali B, Khatir S, Abdel Wahab M. Crack prediction in beam-like structure using ANN based on frequency analysis. Frat Ed Integrità Strutt 2021;16(59):18–34. https://doi.org/10.3221/IGF-ESIS.59.02.
[27] A. Khatir, R. Capozucca, E. Magagnini, S. Khatir, and E. Bettucci, "Structural Health Monitoring for RC Beam Based on RBF Neural Network Using Experimental Modal Analysis," in Proceedings of the International Conference of Steel and Composite for Engineering Structures, vol. 317, R. Capozucca, S. Khatir, and G. Milani, Eds., in Lecture Notes in Civil Engineering, vol. 317., Cham: Springer International Publishing, 2023, pp. 82–92. doi: 10.1007/978-3-031-24041-6_7..
[28] Zhang Y, Xu X. Modulus of elasticity predictions through LSBoost for concrete of normal and high strength. Mater Chem Phys 2022;283:126007. https://doi.org/10.1016/j.matchemphys.2022.126007.
[29] Alawi OA, et al. Design optimization of solar collectors with hybrid nanofluids: an integrated ansys and machine learning study. Sol Energy Mater Sol Cells 2024;271:112822. https://doi.org/10.1016/j.solmat.2024.112822.
[30] Maleki M, Dehghani MR, Akbari A, Kazemzadeh Y, Ranjbar A. Investigation of wettability and IFT alteration during hydrogen storage using machine learning. Heliyon 2024;10(19):e38679. https://doi.org/10.1016/j.heliyon.2024.e38679.
[31] Kina C, Tanyildizi H, Turk K. Forecasting the compressive strength of GGBFS-based geopolymer concrete via ensemble predictive models. Constr Build Mater 2023;405:133299. https://doi.org/10.1016/j.conbuildmat.2023.133299.

[32] Onat O, Tanyıldızı H. Machine learning-based estimation of the out-of-plane displacement of brick infill exposed to earthquake shaking. Eng Appl Artif Intell 2024;136:109007. https://doi.org/10.1016/j.engappai.2024.109007.
[33] Zhang Y, Xu X. Solid particle erosion rate predictions through LSBoost. Powder Technol 2021;388:517–25. https://doi.org/10.1016/j.powtec.2021.04.072.
[34] Tahraoui H, Amrane A, Belhadj A-E, Zhang J. Modeling the organic matter of water using the decision tree coupled with bootstrap aggregated and least-squares boosting. Environ Technol Innov 2022;27:102419. https://doi.org/10.1016/j.eti.2022.102419.
[35] Saccotelli L, et al. Enhancing estuary salinity prediction: a Machine Learning and Deep Learning based approach. Appl Comput Geosci 2024;23:100173. https://doi.org/10.1016/j.acags.2024.100173.
[36] Behnamnia M, Mozafari N, Dehghan Monfared A. Rigorous hybrid machine learning approaches for interfacial tension modeling in brine-hydrogen/cushion gas systems: Implication for hydrogen geo-storage in the presence of cushion gas. J Energy Storage 2023;73:108995. https://doi.org/10.1016/j.est.2023.108995.
[37] Godwin DJ, Varuvel EG, Martin MLJ. Prediction of combustion, performance, and emission parameters of ethanol powered spark ignition engine using ensemble Least Squares boosting machine learning algorithms. J Clean Prod 2023;421:138401. https://doi.org/10.1016/j.jclepro.2023.138401.
[38] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29(5). https://doi.org/10.1214/aos/1013203451.
[39] Mirjalili S. The Ant Lion Optimizer. Adv Eng Softw 2015;83:80–98. https://doi.org/10.1016/j.advengsoft.2015.01.010.
[40] Abdollahzadeh B, Soleimanian Gharehchopogh F, Mirjalili S. Artificial gorilla troops optimizer: a new nature-inspired metaheuristic algorithm for global optimization problems. Int J Intell Syst 2021;36(10):5887–958. https://doi.org/10.1002/int.22535.
[41] Faramarzi A, Heidarinejad M, Mirjalili S, Gandomi AH. Marine Predators Algorithm: a nature-inspired metaheuristic. Expert Syst Appl 2020;152:113377. https://doi.org/10.1016/j.eswa.2020.113377.
[42] Ho LV, et al. A hybrid computational intelligence approach for structural damage detection using marine predator algorithm and feedforward neural networks. Comput Struct 2021;252:106568. https://doi.org/10.1016/j.compstruc.2021.106568.
[43] Nguyen DH, Abdel Wahab M. Damage detection in slab structures based on two-dimensional curvature mode shape method and Faster R-CNN. Adv Eng Softw 2023;176:103371. https://doi.org/10.1016/j.advengsoft.2022.103371.
[44] L.V. Ho, T. Bui-Tien, and M. Abdel Wahab, “Application of Gorilla Troops’ Social Intelligence in Damage Detection for a Girder Bridge,” in Proceedings of the 5th International Conference on Numerical Modelling in Engineering, vol. 311, M. Abdel Wahab, Ed., in Lecture Notes in Civil Engineering, vol. 311., Singapore: Springer Nature Singapore, 2023, pp. 11–30. doi: 10.1007/978-981-19-8429-7_2..
[45] ANSYS, Inc. Southpointe, 275 Technology Drive, Canonsburg, PA 15317, Release 17.2..