VNUHCM Journal of Engineering and Technology 2026, 9(2):2941-2952

Open Access Full Text Article

Research Article

A XGBoost-aided metamodel for damage detection of beams using optimized sensor locations based on reduced-order model

Qui X. Lieu[1,2,*], Quan M. Lieu[1,2]

[1]*Faculty of Civil Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, Dien Hong Ward, Ho Chi Minh City, Vietnam*

[2]*Vietnam National University Ho Chi Minh City (VNU-HCM), Linh Xuan Ward, Ho Chi Minh City, Vietnam*

Correspondence

Qui X. Lieu, Faculty of Civil Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, Dien Hong Ward, Ho Chi Minh City, Vietnam

Vietnam National University Ho Chi Minh City (VNU-HCM), Linh Xuan Ward, Ho Chi Minh City, Vietnam

Email: lieuxuanqui@hcmut.edu.vn

History

- Received: 22-8-2025
- Revised: 03-10-2025
- Accepted: 09-3-2026
- Published Online: 28-06-2026

DOI : https://doi.org/10.32508/vnuhcmjet.v9i2.1538

ABSTRACT

This study aims to propose an Extreme gradient boosting (XGBoost)-driven metamodel for damage detection of beams using free vibration signals optimally measured at a limited number of sensors based on reduced-order model. Instead of utilizing an exceedingly huge and complicated learning model constructed by big data, this methodology only builds a learning model via a series of smaller and simpler XGBoost models with only small and moderate data. The sensor placement locations are found by adaptive hybrid evolutionary firefly algorithm (AHEFA) via the optimization problem based on a reduced-order model. Then, to obtain data for feeding into the above learning models, the isogemetric analysis (IGA) combined with the third-order shear deformation (TSDT) is employed. In which, the eigenvector values at a number of degrees of freedom (DOFs) corresponding to the optimized sensor positions are treated as the input data, whilst the randomly assumed damage ratios of beam elements are considered as the outputs. A modal strain energy-based index (MSEI) is applied to remove low-risk elements before employing the suggested metamodel built by a series of XGBoost models. A simply supported beam with two damage scenarios is investigated. The results obtained by the present methodology have shown the reliability and efficiency in identifying the locations and ratios of damaged elements in beam structures.

Key words: Metamodel, Extreme gradient boosting (XGBoost), Damage detection, Sensor location optimization, Isogemetric analysis (IGA), Third-order shear deformation (TSDT)

INTRODUCTION

The structural health monitoring (SHM) field is very crucial and has a wide range of applications, especially in the structural engineering with the purpose of detecting, localizing and evaluating the existing and/or implicit damage of a certain structural component, even for the whole structural system at an early stage. This helps engineers discover the harm and provide proper and timely solutions for maintaining, reinforcing, repairing, or even replacing a certain structural member to extend the service life of structures. For those prominent advantages, this area has been intensively and increasingly studied by a large number of scholars all over the world. Accordingly, a variety of approaches have been developed to handle such key matters. Regarding the respect of numerical methods, the model updating approaches based on the inverse optimization problem have become more popular. In these paradigms, the damage locations and ratios of structural members are often determined by resolving an inverse optimization problem via metaheuristic algorithms. Typically, Seyedpoor [1] utilized particle swarm optimization (PSO) for the damage estimation of trusses based on free vibration behavior. Kaveh and Zolghadr [2] employed Charged System Search (CSS) algorithm for the truss damage diagnosis based on eigenvectors and eigenvalues. Ding et al. [3] utilized artificial bee colony (ABC) algorithm, while Kim et al. [4] used differential evolution (DE). Dang et al. [5] utilized history acceleration signals for the damage evaluation of truss members by using adaptive hybrid evolutionary firefly algorithm (AHEFA), and so on. Although these methods have recorded salient achievements, they also encounter many limitations relating to: (i) divergent optimization solutions; (ii) insufficiently robust optimizers; (iii) high noise effects; (iv) enormous computational cost of model updating, and so on.

For the above reasons, the so-called data-driven methods have been recently proposed. In this regard, the machine learning (ML)-based learning models in artificial intelligence (AI) have been widely applied. Instead of using inverse optimization problems, these approaches use the data experimentally measured from a monitored structure and numerically simulated from finite element model. Learning models are trained and tested via such data. Then, they are utilized as surrogate models for real-time damage prediction. Several publications on employing learning models such as deep neural networks (DNNs) can

Cite this article : Lieu Q X, Lieu Q M. **A XGBoost-aided metamodel for damage detection of beams using optimized sensor locations based on reduced-order model**. *VNUHCM J. Eng. Technol.* 2026; 9(2):2941-2952.

be found in Refs.[6-10], etc.

However, the neural network structures of learning models are often very complicated since they need to be constructed based on big data, aiming to achieve sufficiently high accuracy and reliability. As a result, the computational resources are extremely expensive and huge. Moreover, collecting the data for experimental measurements is very expensive and time-consuming, especially for large-scale structures due to the limitation on the number of sensors. Therefore, many techniques on model order reduction (ROM) have been applied to the SHM area. More specifically, in place of measuring the signals at all degrees of freedom (DOFs) of a structural system, these methods enable to numerically infer the related information on unmeasured DOFs from those recorded at a given number of sensors. Many researchers have successfully applied these strategies to damage identification based on model updating methods[11-13]. More recently, the author has also succeeded in combining ROMs with learning models such as DNN[14] and XGBoost[15] for the trusses' damage detection. Nevertheless, most of the current works did not take into account the optimal locations of placing measurement sensors. Therefore, the data collected from the sensors' positions may not be the most sensitive to the damage. Consequently, the accuracy and efficiency of the ML models learned from such data may not be high.

In this study, the best locations of positioning a limited number of sensors to measure the most damage-sensitive data at important DOFs are found via an optimization problem resolved by the AHEFA[16]. Then, the remaining signals of unrecorded DOFs can be numerically inferred using the second-order Neumann series expansion (SNSE)-based ROM. In the first step, a MSEI is utilized to extract doubtfully flawed elements, aiming to reduce the number of outputs in the data to feed into the XGBoost. In the next stages, a strategy originally developed by the author[14] is applied. More concretely, a series of XGBoost models is constructed based upon datasets continuously refined after each step. Consequently, the accuracy of the subsequently upgraded model is enhanced remarkably. To the best knowledge of the authors, there have been no such researches to be published concerning the damage diagnosis of beam structures thus far, especially considering the optimal placements of a given number of sensors.

RESEARCH METHODOLOGY

Third-order shear deformation beam theory

Assume that a beam with the length L, the width b and the height h, which respectively correspond to the x-, y- and z-axes, is considered. Then, the displacement across the beam height can be represented by[17,18].

$$\begin{cases} u(x,z) = u_0(x) - z w_{0,x}(x) + \xi(z)\theta_0(x), \\ w(x,z) = w_0(x), \end{cases} \quad \left(-\frac{h}{2} \le z \le \frac{h}{2}\right), \tag{1}$$

where $u_0(x)$ and $w_0(x)$ denote the straight displacements in the x- and z-directions at $z = 0$; $\theta_0(x)$ is the bending rotation about the y-axis at $z = 0$, and $\xi(z) = z - \frac{4z^3}{3h^2}$ stands for the third-order polynomial used to describe the strain and stress through the beam height.

Isogeometric analysis

Herein, the finite element model of a typical element for free vibration analysis using the IGA[19] can be expressed as follows

$$(\mathbf{K}_e - \omega^2 \mathbf{M}_e)\mathbf{\Psi}_e = 0, \tag{2}$$

where ω and $\mathbf{\Psi}_e$ denote the eigenvalue and eigenvector, respectively; $\mathbf{K}_e$ and $\mathbf{M}_e$ are the stiffness and mass matrices of a typical beam element, and are respectively given by

$$\mathbf{K}_e = b\int_{l^e} \sum_{i=1}^{nCPs} \left[\left(\mathbf{\Theta}_i^{\mathrm{mb}}\right)^{\mathrm{T}} \mathbf{\Xi}_{\mathrm{mb}} \mathbf{\Theta}_i^{\mathrm{mb}} + \left(\mathbf{\Theta}_i^{\mathrm{s}}\right)^{\mathrm{T}} \Xi_{\mathrm{s}} \mathbf{\Theta}_i^{\mathrm{s}} \right] \mathrm{d}x, \tag{3}$$

and

$$\mathbf{M}_e = b\int_{l^e} \sum_{i=1}^{nCPs} \mathbf{\Lambda}^{\mathrm{T}} \mathbf{m} \mathbf{\Lambda} \mathrm{d}x, \tag{4}$$

with $nCPs$ stands for the number of control points of the eth element, and

$$\mathbf{\Theta}_i^{\mathrm{mb}} = \begin{bmatrix} B_{i,x} & 0 & 0 \\ 0 & -B_{i,xx} & 0 \\ 0 & 0 & B_{i,x} \end{bmatrix}, \quad \mathbf{\Theta}_i^{\mathrm{s}} = \{0 \quad 0 \quad B_i\},$$

$$\mathbf{\Xi}_{\mathrm{mb}} = \begin{bmatrix} \Xi_1 & \Xi_z & \Xi_\xi \\ \Xi_z & \Xi_{z^2} & \Xi_{z\xi} \\ \Xi_\xi & \Xi_{z\xi} & \Xi_{\xi^2} \end{bmatrix}, \quad \Xi_{\mathrm{s}} = \int_{-h/2}^{+h/2} \frac{E}{2(1+\nu)} \left(\frac{\partial \xi}{\partial z}\right)^2 \mathrm{d}z,$$

$$\mathbf{\Lambda} = \begin{bmatrix} B_i & 0 & 0 \\ 0 & -B_{i,x} & 0 \\ 0 & 0 & B_i \\ 0 & B_i & 0 \end{bmatrix}, \quad \mathbf{m} = \begin{bmatrix} I_1 & I_z & I_\xi & 0 \\ I_z & I_{z^2} & I_{z\xi} & 0 \\ I_\xi & I_{z\xi} & I_{\xi^2} & 0 \\ 0 & 0 & 0 & I_1 \end{bmatrix}, \tag{5}$$

in which

$$\left\{\Xi_1, \Xi_z, \Xi_{z^2}, \Xi_\xi, \Xi_{z\xi}, \Xi_{\xi^2}\right\} = \int_{-h/2}^{+h/2} E\left(1, z, z^2, \xi, z\xi, \xi^2\right) \mathrm{d}z,$$
$$\left\{I_1, I_z, I_{z^2}, I_\xi, I_{z\xi}, I_{\xi^2}\right\} = \int_{-h/2}^{+h/2} \rho\left(1, z, z^2, \xi, z\xi, \xi^2\right) \mathrm{d}z, \tag{6}$$

and B_i symbolizes the ith B-spline function[20]; E and ρ stand for the Young's modulus and the material density, respectively.

Second-order Neumann series expansion

In this work, SNSE[21] is applied to condense the free vibration properties of the beam structures. Before implementing this, each of all the element stiffness and mass matrices needs to be assembled into their global system. Then, the algebraic equation system for the free vibration analysis of a beam in the full system can be arranged as follows

$$\left(\begin{bmatrix}\mathbf{K}_{\mathrm{rr}} & \mathbf{K}_{\mathrm{rc}}\\ \mathbf{K}_{\mathrm{cr}} & \mathbf{K}_{\mathrm{cc}}\end{bmatrix} \quad \boldsymbol{\omega}^2\begin{bmatrix}\mathbf{M}_{\mathrm{rr}} & \mathbf{M}_{\mathrm{rc}}\\ \mathbf{M}_{\mathrm{cr}} & \mathbf{M}_{\mathrm{cc}}\end{bmatrix}\right)\begin{Bmatrix}\boldsymbol{\Psi}_{\mathrm{r}}\\ \boldsymbol{\Psi}_{\mathrm{c}}\end{Bmatrix} \quad \mathbf{0}, \tag{7}$$

where “r” and “c” indicate the retained and curtained DOFs, respectively.

Eq. (7) can be reduced as follows

$$\left(\mathbf{K}_{\mathrm{SNSE}} - \boldsymbol{\omega}^2_{\mathrm{SNSE}}\mathbf{M}_{\mathrm{SNSE}}\right)\boldsymbol{\Psi}_{\mathrm{r}} = \mathbf{0}, \tag{8}$$

where

$$\mathbf{K}_{\mathrm{SNSE}} = \Im^{\mathrm{T}}\mathbf{K}\Im \quad \text{and} \quad \mathbf{M}_{\mathrm{SNSE}} = \Im^{\mathrm{T}}\mathbf{M}\Im, \tag{9}$$

$$\Im = \begin{bmatrix}\mathbf{I}_{\mathrm{rr}}\\ -\left[\boldsymbol{\kappa}_1 + \mathbf{K}_{\mathrm{cc}}^{-1}\mathbf{M}_{\mathrm{cc}}\left(\chi_1\chi_4 + \chi_1\chi_5\right)\right]^{-1}\left[\boldsymbol{\kappa}_2 + \mathbf{K}_{\mathrm{cc}}^{-1}\mathbf{M}_{\mathrm{cc}}\left(\chi_1\chi_2 + \chi_1\chi_3\right)\right]\end{bmatrix}, \tag{10}$$

$$\begin{aligned}\boldsymbol{\kappa}_1 &= \mathbf{I}_{\mathrm{rr}} + \chi_1\mathbf{K}_{\mathrm{rc}},\\ \boldsymbol{\kappa}_2 &= \mathbf{K}_{\mathrm{cc}}^{-1}\mathbf{K}_{\mathrm{cr}} + \chi_1\mathbf{K}_{\mathrm{rr}},\\ \chi_1 &= \mathbf{K}_{\mathrm{cc}}^{-1}\mathbf{M}_{\mathrm{cc}}\mathbf{K}_{\mathrm{cc}}^{-1}\mathbf{K}_{\mathrm{cr}}\mathbf{M}_{\mathrm{rr}}^{-1},\\ \chi_2 &= \mathbf{K}_{\mathrm{rr}}\mathbf{M}_{\mathrm{rr}}^{-1}\mathbf{K}_{\mathrm{rr}},\\ \chi_3 &= \mathbf{K}_{\mathrm{rc}}\mathbf{M}_{\mathrm{cc}}^{-1}\mathbf{K}_{\mathrm{cr}},\\ \chi_4 &= \mathbf{K}_{\mathrm{rr}}\mathbf{M}_{\mathrm{rr}}^{-1}\mathbf{K}_{\mathrm{rc}},\\ \chi_5 &= \mathbf{K}_{\mathrm{rc}}\mathbf{M}_{\mathrm{cc}}^{-1}\mathbf{K}_{\mathrm{cc}}.\end{aligned} \tag{11}$$

Sensor placement optimization

With a given number of measurement sensors, the problem aims to determine which DOFs are necessary to achieve the best ROM. For that purpose, the objective function, which is computed from the correlation between eigenvectors obtained by the full model (FM) and the ROM via Modal Assurance Criterion (MAC)[22], is maximized.

$$\begin{aligned}&\text{maximize}: f_{\mathrm{objective}}(\mathbf{x}) = \frac{1}{nom}\sum_{i=1}^{nom}\frac{\left[\boldsymbol{\Psi}^i_{\mathrm{FM}}\boldsymbol{\Psi}^i_{\mathrm{SNSE}}(\mathbf{x})\right]^2}{\left[\left(\boldsymbol{\Psi}^i_{\mathrm{FM}}\right)^{\mathrm{T}}\boldsymbol{\Psi}^i_{\mathrm{FM}}\right]\left[\left(\boldsymbol{\Psi}^i_{\mathrm{SNSE}}(\mathbf{x})\right)^{\mathrm{T}}\boldsymbol{\Psi}^i_{\mathrm{SNSE}}(\mathbf{x})\right]},\\ &\text{subjected to}: \begin{cases}\left(\mathbf{K}_{\mathrm{SNSE}} - \boldsymbol{\omega}^2\mathbf{M}_{\mathrm{SNSE}}\right)\boldsymbol{\Psi}_{\mathrm{SNSE}} = \mathbf{0},\\ \mathbf{x} \in \mathbb{R}^{1\times nor},\\ x_j \in \mathbb{Z}^+,\end{cases}\end{aligned} \tag{12}$$

where $\boldsymbol{\Psi}^i_{\mathrm{FM}}$ and $\boldsymbol{\Psi}^i_{\mathrm{SNSE}}(\mathbf{x})$ denote the ith eigenvectors given by the FM and the ROM, respectively; nom=10 stands for the number of mode shapes used to construct the objective function; $\{x_1, \ldots, x_j, \ldots, x_{nor}\}$is the vector consisting of nor retailed DOFs which serve as design variables in the optimization process.

In this work, the AHEFA which was previously developed for truss optimization[16] is used as an optimizer to resolve the above-defined problem. Nonetheless, for the sensor location optimization, parameters defined in the above algorithm follow the study[23].

Damage model of beam elements

In practice, there are many reasons that can lead to the damage of a beam. In this study, in order to simulate such damage for a certain element in a beam, the eth element stiffness matrix is assumed to be numerically declined by a damage ratio β via the following formulation,

$$\mathbf{K}^{\mathrm{d}}_e = (1-\beta)\mathbf{K}^{\mathrm{h}}_e, \tag{13}$$

where $\mathbf{K}^{\mathrm{h}}_e$ and $\boldsymbol{K}^{\mathrm{d}}_e$ denote the eth element healthy and damaged stiffness matrices, respectively.

Modal strain energy-based index

Before using the proposed metamodel constructed by a series of XGBoost models, the MSEI proposed by Seyedpoor[1] is applied to exclude low-risk elements, aiming to reduce the number of outputs for constructing the learning model. This indicator for the eth element is computed by

$$MSEI_e = \begin{cases}\dfrac{\overline{MSE}^{\mathrm{d}}_e - \overline{MSE}^{\mathrm{h}}_e}{\overline{MSE}^{\mathrm{h}}_e} > 0, & \text{high-risk element,}\\[2ex] \dfrac{\overline{MSE}^{\mathrm{d}}_e - \overline{MSE}^{\mathrm{h}}_e}{\overline{MSE}^{\mathrm{h}}_e} \le 0, & \text{low-risk element,}\end{cases} \tag{14}$$

where

$$\overline{MSE}^{\mathrm{d/h}}_e = \frac{\sum_{i=1}^{nom}\overline{MSE}^{\mathrm{d/h}}_{e,i}}{nom}, \tag{15}$$

in which

$$\overline{MSE}^{\mathrm{d/h}}_{e,i} = \frac{MSE^{\mathrm{d/h}}_{e,i}}{\sum_{e=1}^{noe}MSE^{\mathrm{d/h}}_{e,i}}, \tag{16}$$

with

$$MSE^{\mathrm{d/h}}_{e,i} = \frac{1}{2}\left(\boldsymbol{\Psi}^{\mathrm{d/h}}_{e,i}\right)^{\mathrm{T}}\mathbf{K}^{\mathrm{h}}_e\boldsymbol{\Psi}^{\mathrm{d/h}}_{e,i}, \tag{17}$$

where $\mathbf{K}^{\mathrm{h}}_e$ symbolizes the eth global stiffness matrix at the healthy state; $\boldsymbol{\Psi}^{\mathrm{d}}_{e,i}$ and $\boldsymbol{\Psi}^{\mathrm{h}}_{e,i}$ denote the eth eigenvector of the ith mode shape obtained from the measured and simulated data, respectively.

Extreme gradient boosting-aided metamodel

In the previously published author’s works[14,24], a DNN-assisted metamodel was developed for damage detection of truss structures using time-history acceleration signals. Herein, the Extreme gradient boosting (XGBoost)[25] is employed as a learning model to replace the above DNN. As known, XGBoost is an ensemble learning model which is built as an ensemble of decision trees with sequentially handled manners to improve its predictive implementation. Numerous problems concerning classification, regression and ranking tasks have been applied owing to its efficiency. In the proposed metamodel, XGBoost is trained and tested for the damage identification of beam elements using the values at DOFs with regard to the optimal sensor locations of the first five modal shapes simulated from the TSDT-based IGA as the inputs. Meanwhile, the outputs are the randomly created damage ratios of beam elements. However, instead of utilizing a unique learning XGBoost constructed by big data and a complicated network, the present methodology uses a series of XGBoost models whose accuracy is continuously enhanced. More concretely, the first model is built for high-risk beam elements discovered by the MSEI. Then, the damage ratio of all elements is predicted by the current learning model. If the predicted ratio of a certain element is less than or equal to a suggested damage threshold of 5%, it is treated as a healthy element. It is removed and not included in the outputs for upgrading the next model. On the other hand, when the predicted ratio is larger than the allowable threshold, it is taken into account as a high-risk element. Then, only the damage ratio of these elements is randomly generated as the outputs of the subsequent learning model. By repeating such a procedure, the output data dimension can be continuously reduced. The number of samples required for each stage to feed into the XGBoost is 2000. As a consequence, the accuracy of the subsequently upgraded models is improved dramatically. And this leads to the terminology of the so-called metamodel. This process is terminated when the number of elements detected at two consecutive steps is unchanged. Hyperparameters defined in the XGBoost models are automatically established by Bayesian optimization.

RESULTS AND DISCUSSION

Verification

Firstly, in order to verify the reliability of code structures in using the TSDT-based IGA for the free vibration analysis of beams, a simply supported beam which was previously studied in Refs.[18,26] is examined. Following this, the boundaries at its two ends are set up as $u_0(0) = v_0(0) = 0,\ v_0(L) = 0$. This beam is of $\rho = 3960$ kg/m³, $E = 380$ MPa and $\nu = 0.3$. For comparison, the dimensionless frequencies based on the formula $\bar{\omega} = \omega \frac{L^2}{h}\sqrt{\frac{\rho_m}{E_m}}$ with $\rho_m = 2702$kg/m³ and $E_m = 70$ MPa are used for the report. It can be seen from Table 1, the outcomes obtained from the TSDT-based IGA are in agreement with those of the reference solutions in the case of utilizing 20 cubic elements. This mesh size is adequate to result in highly reliable and accurate outcomes, and it is hence adopted for use in the next investigations.

Table 1. Comparison of the first dimensionless frequencies.

L/h	Reference	Mode							
		1	2	3	4	5	6	7	8
5	TSDT-based IGA								
	14 elements	5.1528	15.1157	17.8820	34.2203	45.3470	52.1041	70.8659	75.5785
	16 elements	5.1528	15.1157	17.8816	34.2157	45.3470	52.0697	70.6858	75.5784
	18 elements	5.1528	15.1157	17.8815	34.2134	45.3470	52.0525	70.5980	75.5784
	20 elements	**5.1528**	**15.1157**	**17.8814**	**34.2121**	**45.3470**	**52.0431**	**70.5512**	**75.5784**
	22 elements	5.1527	15.1157	17.8813	34.2113	45.3470	52.0377	70.5243	75.5784
	Ref.[18]	5.1527	17.8812	34.2097	-	-	-	-	-
	Ref.[26]	5.1528	17.8817	34.2143	-	-	-	-	-
20	TSDT-based IGA								
	14 elements	5.4603	21.5739	47.6009	60.4627	82.4930	125.1102	174.4646	181.3882
	16 elements	5.4603	21.5736	47.5975	60.4627	82.4716	125.0148	174.1215	181.3882
	18 elements	5.4603	21.5735	47.5958	60.4627	82.4608	124.9677	173.9566	181.3882
	20 elements	**5.4603**	**21.5734**	**47.5948**	**60.4627**	**82.4548**	**124.9421**	**173.8693**	**181.3882**
	22 elements	5.4603	21.5733	47.5942	60.4627	82.4512	124.9273	173.8195	181.3882
	Ref.[18]	5.4603	21.5732	47.5930	-	-	-	-	-
	Ref.[26]	5.4603	21.5732	47.5999	-	-	-	-	-

Source: Authors' own work

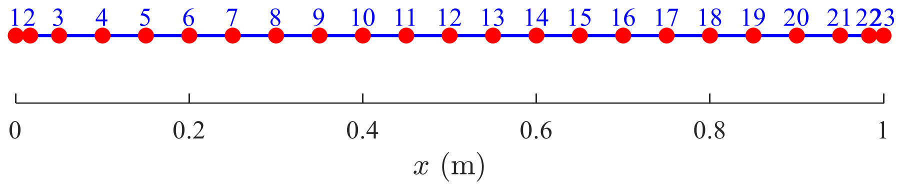


Fig. 1. A control point net using 20 cubic elements for the beam model. (Source: Authors' own work)

Present study

Now, assume that only 15 sensors are used for measurement, and their locations are optimized. For this, the same parameters of the above beam with $L/h = 5$ (L=1 m and b = 0.2 m) is examined as a specific case demonstrated in this study. Accordingly, for a mesh of 20 cubic elements, the total number of DOFs is 69, including ones belong to the boundary conditions. Fig. 1 plots the control point net of the beam given by the IGA. By resolving the optimization problem defined in Section 2.4, DOFs optimized by the AHEFA for the best run are tabulated in Table 2. The convergence history of the optimization process is shown in Fig. 2. As found, the objective function's value reaches 1. The correlation of the first five mode shapes obtained by the FM and the ROM is shown in Fig. 3. It is clear that the terms on the diagonal of the MAC matrix also reach 1. The first five mode shapes given by the SNSE based on optimal DOFs are plotted in Fig. 4.

Table 2. DOFs with regard to measurement sensors optimized by the AHEFA.

DOFs	Control points
u_0	4 7 10 20
v_0	5 6 7 9 14 22
β_0	1 3 4 6 12 15 16 19 21 22

Source: Authors' own work

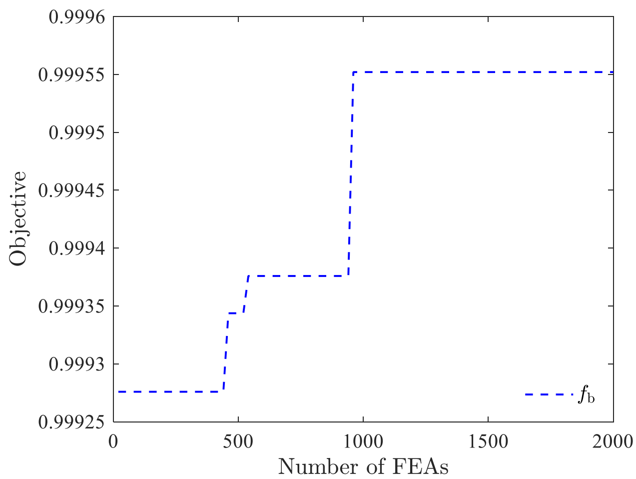


Fig. 2. The convergence history attained in the optimization process of sensor placements. (Source: Authors' own work)

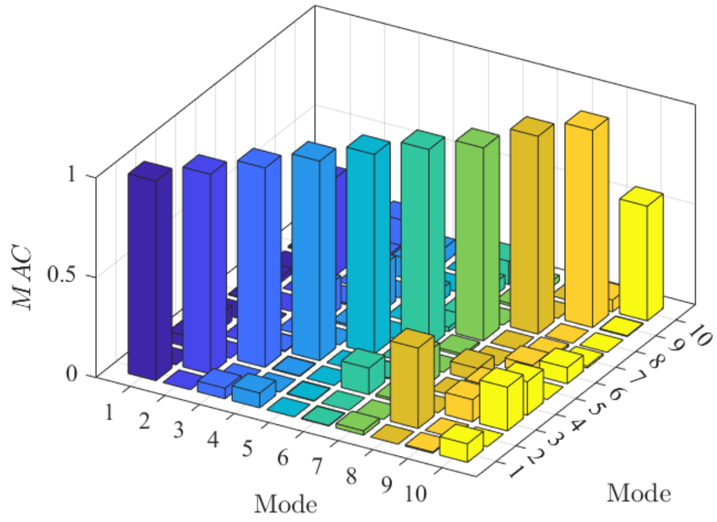


Fig. 3. The MAC optimized by the AHEFA. (Source: Authors' own work)

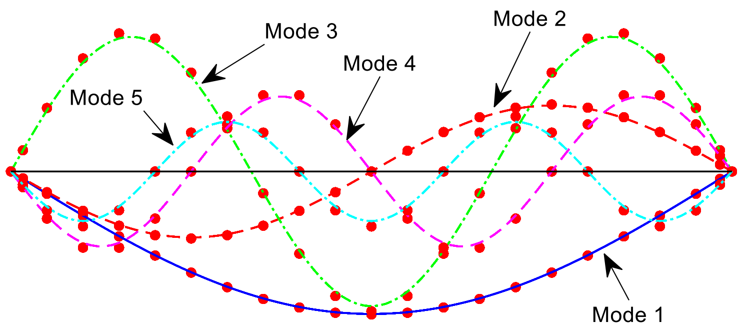


Fig. 4. The first five mode shapes of the beam attained by the SNSE based on optimal DOFs. (Source: Authors' own work)

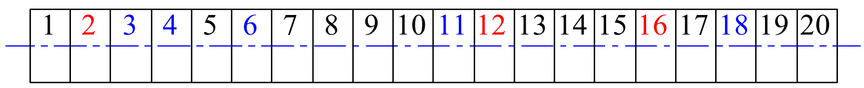


Fig. 5. Labels of beam elements. (Source: Authors' own work)

Table 3. Two damage scenarios tested in this work.

Scenario	1			2				
Beam element	2	12	16	3	4	6	11	18
Damage ratio	0.40	0.30	0.35	0.30	0.20	0.35	0.30	0.30

Source: Authors' own work

Now, for the damage evaluations of the above beam structure, its elements are numbered as Fig. 5. Next, two damage scenarios as assumed in Table 3 are tested. In order to build the proposed metamodel, the values at previously optimized DOFs which are simulated from the TSDT-based IGA are employed as inputs. In which, only the first five mode shapes are used instead of all of them. In the first step, the MSEI discovers high-risk elements as reported in Fig. 6. Next, the number of outputs is continuously refined based on the results identified by the previous learning model. In the first XGBoost, the damage ratio of 20 beam elements is randomly generated to attain those eigenvectors. The damage outcome obtained in case 1 is given in Fig. 7a). It can be recognized that if a damage threshold is applied, there are only 3 suspiciously diagnosed elements to be kept for establishing the next model. Then, a newly refined dataset with inputs and outputs is created to feed into the learning model. As a consequence, the second XGBoost is of higher accuracy than the first one, as indicated in Fig. 7b). It is obvious that the newly upgraded XGBoost can yield better damage detection results than the first one. This convergence history acquired by the process is plotted in Fig. 8.

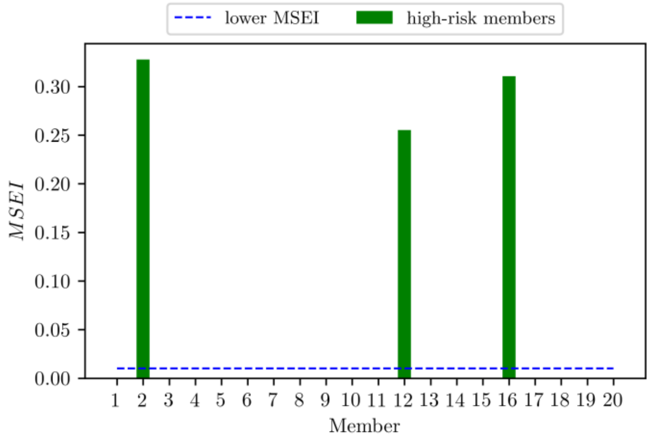
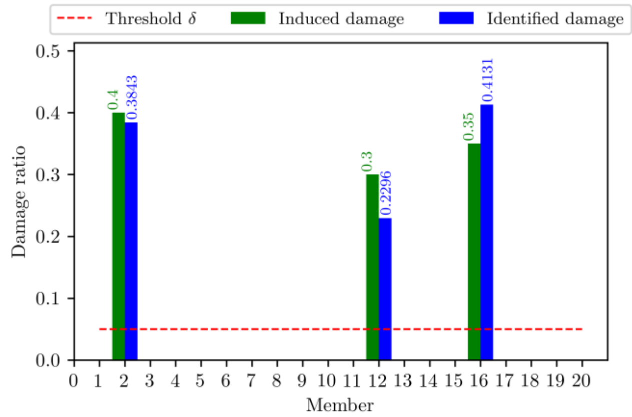
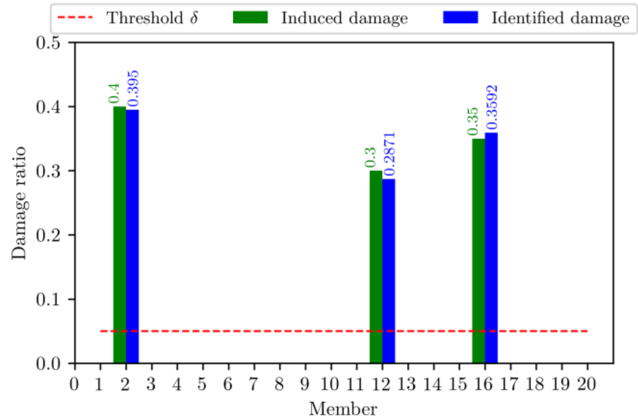


Fig. 6 High-risk elements discovered by the MSEI in the first stage for scenario 1. (Source: Authors' own work)



a) The 1st XGBoost



b) The 2nd XGBoost

Fig. 7 The damage results detected by the XGBoost-based metamodel for scenario 1. (Source: Authors' own work)

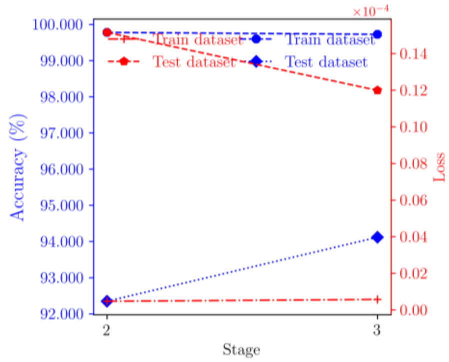


Fig. 8 The convergence history attained by the present paradigm for scenario 1. (Source: Authors' own work)

Analogously, the MSEI is used to indicate the potentially damaged elements as shown in Fig. 9. The damage outcomes diagnosed by the present methodology for scenario 2 are reported in Fig. 10. As expected, the XGBoost-based metamodel can still detect the damage to multiple elements with high reliability only after several steps. Moreover, the accuracy of the learning models is remarkably enhanced, while the value of the loss function is reduced after each step. These remarks can be easily observed via Fig. 11.

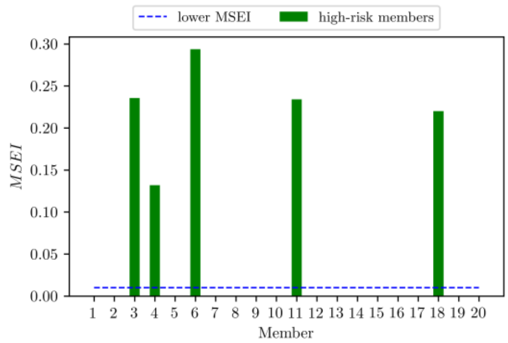
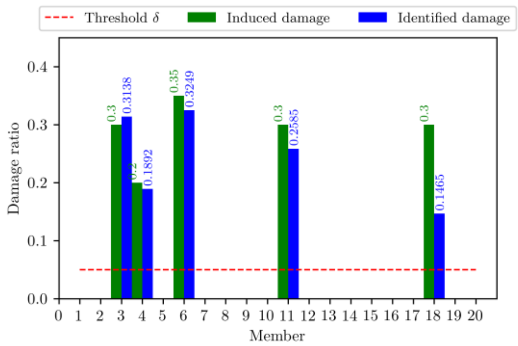
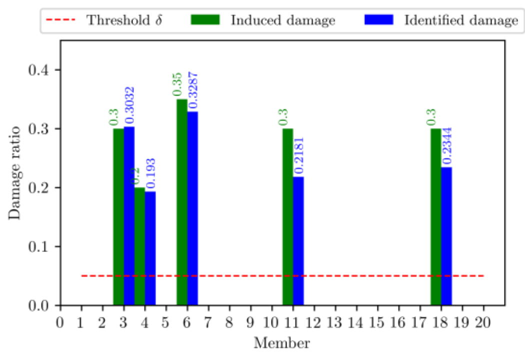


Fig. 9 High-risk elements discovered by the MSEI in the first stage for scenario 2. (Source: Authors' own work)



a) The 1st XGBoost



b) The 2nd XGBoost

Fig. 10 The damage results detected by the XGBoost-based metamodel for scenario 2. (Source: Authors' own work)

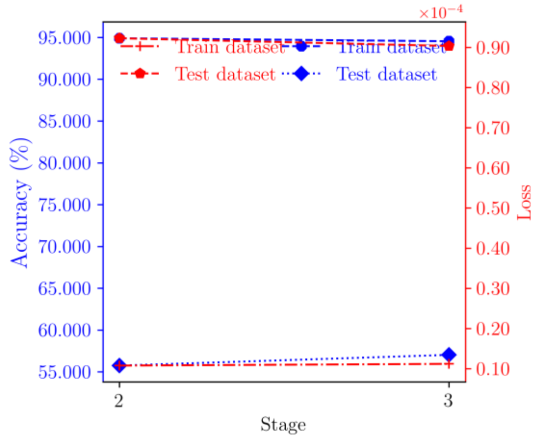
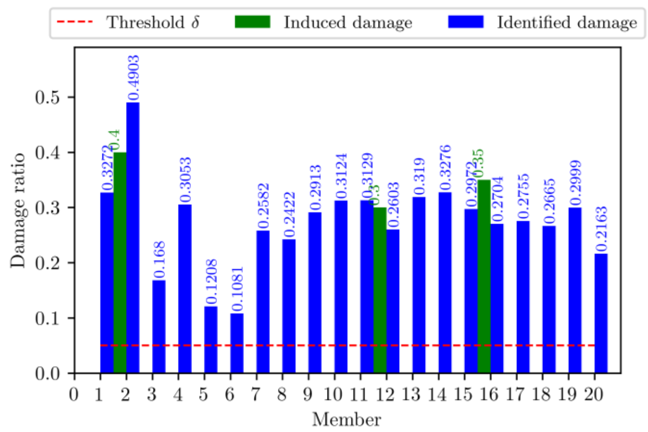
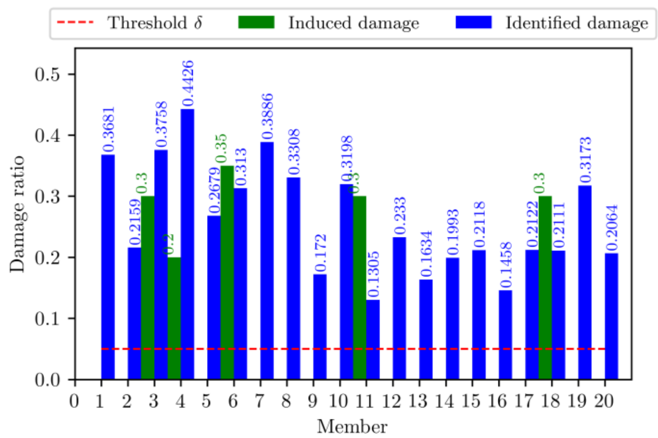


Fig. 11 The convergence history attained by the present paradigm for scenario 2. (Source: Authors' own work)

It is clear that if the MSEI is not used in the first step, the XGBoost-based metamodel can not recognize the damage of beam elements with only 2000 samples as indicated in Fig. 12. This proves that the accuracy of the learning model is higher and more reliable if it is constructed via the properly refined data instead of randomly selected data.



a) Scenario 1



b) Scenario 2

Fig. 12 The damage results detected by the XGBoost-based metamodel without MSEI. (Source: Authors' own work)

CONCLUSIONS

This article has successfully suggested a XGBoost-aided metamodel for damage detection of beams using the free vibration signals which are measured from the sensor placement optimization and simulated from the TSDT-based IGA. In which, the locations of positioning sensors for measuring the signals of a monitored beam are searched via the optimization problem established from the ROM. The inputs of learning modes during the process of constructing a metamodel are the eigenvectors' values at DOFs with regard to sensors' locations, while the outputs are the randomly given damage ratios of beam elements. In the suggested methodology, in place of utilizing big data to build an extremely large and complicated learning model, only proportionally sized datasets are employed to a metamodel via a series of XGBoost-based learning models. Nevertheless, in order to further reduce the number of outputs for the XGBoost, a MSEI is employed to curtail low-risk elements. Then, the accuracy of the subsequently upgraded model is continuously and remarkably enhanced by excluding low-risk elements via a damage threshold. As a result, the proposed metamodel enables to reliably and accurately detect the damage of beam elements with only moderate data and low computational cost. The proposed paradigm is very promising to extend its applications to examine other structures in the future.

COMPETING INTERESTS

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.

AUTHORS' CONTRIBUTIONS

Qui X. Lieu: Conceptualization, Methodology, Investigation, Data Curation, Validation, Resources, Software, Writing-Original Draft, Writing-Review and Editing, Funding acquisition, Project administration. **Quan M. Lieu:** Methodology, Investigation, Validation, Resources, Software, Writing-Original Draft, Writing-Review and Editing.

ACKNOWLEDGEMENTS

This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number DS2025-20-11.

REFERENCES

1. Seyedpoor SM (2012) A two stage method for structural damage detection using a modal strain energy based index and particle swarm optimization. Int J Non-Linear Mech 47:1–8. https://doi.org/10.1016/j.ijnonlinmec.2011.07.011
2. Kaveh A, Zolghadr A (2015) An improved CSS for damage detection of truss structures using changes in natural frequencies and mode shapes. Adv Eng Softw 80:93–100. https://doi.org/10.1016/j.advengsoft.2014.09.010
3. Ding ZH, Huang M, Lu ZR (2016) Structural damage detection using artificial bee colony algorithm with hybrid search strategy. Swarm Evol Comput 28:1–13. https://doi.org/10.1016/j.swevo.2015.10.010
4. Kim N-I, Kim S, Lee J (2019) Vibration-based damage detection of planar and space trusses using differential evolution algorithm. Appl Acoust 148:308–321. https://doi.org/10.1016/j.apacoust.2018.08.032
5. Dang KD, Nguyen NH, Lee S, et al (2023) A novel model order reduction-based two-stage damage detection paradigm for trusses using time–history acceleration. Adv Eng Softw 176:103374. https://doi.org/10.1016/j.advengsoft.2022.103374
6. Lee S, Park S, Kim T, et al (2021) Damage quantification in truss structures by limited sensor-based surrogate model. Appl Acoust 172:107547. https://doi.org/10.1016/j.apacoust.2020.107547
7. Lee S, Kim T, Lieu QX, et al (2023) A novel data-driven analysis for sequentially formulated plastic hinges of steel frames. Comput Struct 281:107031. https://doi.org/10.1016/j.compstruc.2023.107031
8. Truong TT, Dinh-Cong D, Lee J, Nguyen-Thoi T (2020) An effective deep feedforward neural networks (DFNN) method for damage identification of truss structures using noisy incomplete modal data. J Build Eng 30:101244. https://doi.org/10.1016/j.jobe.2020.101244
9. Dang KD, Ly NM, Truong HH, et al (2024) Detection of Plastic Hinges in Inelastic Nonlinear Steel Frames Using Deep Learning. In: Reddy JN, Wang CM, Luong VH, Le AT (eds) Proceedings of the Third International Conference on Sustainable Civil Engineering and Architecture. Springer Nature, Singapore, pp 1340–1348
10. Nguyen DD, Dang KD, Nguyen AH, et al (2024) Damage Detection of Frame Structures Based on Acceleration Using Deep Learning. In: Reddy JN, Wang CM, Luong VH, Le AT (eds) Proceedings of the Third International Conference on Sustainable Civil Engineering and Architecture. Springer Nature, Singapore, pp 1333–1339
11. Kaveh A, Zolghadr A (2017) Cyclical Parthenogenesis Algorithm for guided modal strain energy based structural damage detection. Appl Soft Comput 57:250–264. https://doi.org/10.1016/j.asoc.2017.04.010
12. Tiachacht S, Khatir S, Thanh CL, et al (2022) Inverse problem for dynamic structural health monitoring based on slime mould algorithm. Eng Comput 38:2205–2228. https://doi.org/10.1007/s00366-021-01378-8
13. Ho LV, Trinh TT, De Roeck G, et al (2022) An efficient stochastic-based coupled model for damage identification in plate structures. Eng Fail Anal 131:105866. https://doi.org/10.1016/j.engfailanal.2021.105866
14. Lieu QX (2023) A novel multistage damage detection method for trusses using time-history data based on model order reduction and deep neural network. Mech Syst Signal Process 200:110635. https://doi.org/10.1016/j.ymssp.2023.110635
15. Qui LX (2023) Damage identification of trusses using limited modal features and ensemble learning. J Sci Technol Civ Eng STCE - HUCE 17:9–20. https://doi.org/10.31814/stce.huce2023-17(2)-02
16. Lieu QX, Do DTT, Lee J (2018) An adaptive hybrid evolutionary firefly algorithm for shape and size optimization of truss structures with frequency constraints. Comput Struct 195:99–112. https://doi.org/10.1016/j.compstruc.2017.06.016
17. Reddy JN (1984) A Simple Higher-Order Theory for Laminated Composite Plates. J Appl Mech 51:745–752. https://doi.org/10.1115/1.3167719
18. Thai H-T, Vo TP (2012) Bending and free vibration of functionally graded beams using various higher-order shear deformation beam theories. Int J Mech Sci 62:57–66. https://doi.org/10.1016/j.ijmecsci.2012.05.014
19. Hughes TJR, Cottrell JA, Bazilevs Y (2005) Isogeometric analysis: CAD, finite elements,

NURBS, exact geometry and mesh refinement. Comput Methods Appl Mech Eng 194:4135–4195. https://doi.org/10.1016/j.cma.2004.10. 008
20. Boor C de (1972) On calculating with B-splines. J Approx Theory 6:50–62
21. Yang QW (2009) Model reduction by Neumann series expansion. Appl Math Model 33:4431–4434. https://doi.org/10.1016/j.apm.2009.02.012
22. Pastor M, Binda M, Harčarik T (2012) Modal Assurance Criterion. Procedia Eng 48:543–548. https://doi.org/10.1016/j.proeng.2012.09.551
23. Lieu QX (2024) An optimization-assisted reduced order model for dynamics of plates using isogeometric analysis. Thin-Walled Struct 202:112078. https://doi.org/10.1016/j.tws.2024.112078
24. Lieu QX (2023) A deep neural network-assisted metamodel for damage detection of trusses using incomplete time-series acceleration. Expert Syst Appl 233:120967. https://doi.org/10.1016/j.eswa.2023.120967
25. Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp 785–794
26. Nguyen T-K, Truong-Phong Nguyen T, Vo TP, Thai H-T (2015) Vibration and buckling analysis of functionally graded sandwich beams by a new higher-order shear deformation theory. Compos Part B Eng 76:273–285. https://doi.org/10.1016/j.compositesb.2015.02.032

VNUHCM Journal of Engineering and Technology 2026, 9(2):2941-2952

Open Access Full Text Article

Bài nghiên cứu

Một siêu mô hình được hỗ trợ bởi XGBoost để phát hiện hư hỏng của dầm bằng cách sử dụng vị trí cảm biến tối ưu dựa trên mô hình giảm bậc

Liêu Xuân Quí[1,2,*], Liêu Minh Quân[1,2]

TÓM TẮT

Nghiên cứu này nhằm mục đích đề xuất một siêu mô hình dựa trên thuật toán tăng cường độ dốc cực trị (XGBoost) để phát hiện hư hỏng dầm bằng cách sử dụng tín hiệu dao động tự do được đo tối ưu tại một số lượng cảm biến hạn chế dựa trên mô hình giảm bậc. Thay vì sử dụng một mô hình học siêu to và phức tạp được xây dựng từ dữ liệu lớn, phương pháp này chỉ xây dựng mô hình học thông qua một loạt các mô hình XGBoost nhỏ và đơn giản hơn với lượng dữ liệu nhỏ và vừa phải. Vị trí đặt cảm biến tối ưu được tìm thấy bằng thuật toán đom đóm tiến hóa lai tương thích (AHEFA) thông qua bài toán tối ưu hóa dựa trên mô hình giảm bậc. Sau đó, để thu được dữ liệu đưa vào các mô hình học, phương pháp phân tích đẳng hình (IGA) kết hợp với biến dạng cắt bậc ba (TSDT) được sử dụng. Trong đó, các giá trị vectơ riêng tại một số bậc tự do (DOFs) tương ứng với các vị trí cảm biến tối ưu được xem là dữ liệu đầu vào, trong khi tỷ lệ hư hỏng được giả định ngẫu nhiên của các phần tử dầm là đầu ra. Chỉ số năng lượng biến dạng dựa trên dao động (MSEI) được áp dụng để loại bỏ các phần tử hư hỏng rủi ro thấp trước khi sử dụng siêu mô hình được xây dựng bởi một loạt các mô hình XGBoost. Một dầm đơn giản với hai kịch bản hư hỏng khác nhau được khảo sát. Kết quả thu được từ phương pháp đã chứng minh độ tin cậy và hiệu quả trong việc xác định vị trí và tỷ lệ các phần tử bị hư hỏng trong kết cấu dầm.

Từ khoá: Siêu mô hình, Tăng cường độ dốc cực trị (XGBoost), Phát hiện hư hỏng, Tối ưu vị trí cảm biến, Phân tích đẳng hình (IGA), Biến dạng cắt bậc ba (TSDT)

[1]*Khoa Kỹ thuật Xây dựng, Trường Đại học Bách khoa, 268 Lý Thường Kiệt, Phường Diên Hồng, Thành phố Hồ Chí Minh, Việt Nam*

[2]*Đại học Quốc gia Thành phố Hồ Chí Minh, Phường Linh Xuân, Thành phố Hồ Chí Minh, Việt Nam*

Liên hệ

Liêu Xuân Quí, Khoa Kỹ thuật Xây dựng, Trường Đại học Bách khoa, 268 Lý Thường Kiệt, Phường Diên Hồng, Thành phố Hồ Chí Minh, Việt Nam

Đại học Quốc gia Thành phố Hồ Chí Minh, Phường Linh Xuân, Thành phố Hồ Chí Minh, Việt Nam

Email: lieuxuanqui@hcmut.edu.vn

Lịch sử

- Ngày nhận: 22-8-2025
- Ngày sửa đổi: 03-10-2025
- Ngày chấp nhận: 09-3-2026
- Ngày đăng: 28-06-2026

DOI : https://doi.org/10.32508/vnuhcmj-et.v9i2.1538

Trích dẫn bài báo này: Xuân Quí L, Minh Quân L. **Một siêu mô hình được hỗ trợ bởi XGBoost để phát hiện hư hỏng của dầm bằng cách sử dụng vị trí cảm biến tối ưu dựa trên mô hình giảm bậc**. *VNUHCM J. Eng. Technol.* 2026; 9(2):2941-2952.