



Optimal Sensor Placement Considering Both Sensor Faults Under Uncertainty and Sensor Clustering for Vibration-Based Damage Detection

Haichao An¹ · Byeng D. Youn^{1,2,3} · Heung Soo Kim⁴

Received: 12 August 2021 / Revised: 23 December 2021 / Accepted: 25 December 2021 / Published online: 2 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Use of a sensor network to provide adequate and reliable information is paramount for accurate damage detection of structures. However, unavoidably, deployed sensors are occasionally subject to failure faults, which, in turn, cause missing information. Placement of multiple backup sensors in a local region could overcome this difficulty and increase the sensor redundancy; however, this approach leads to a sensor clustering problem and higher costs in sensor deployment. Further, model uncertainty is another important issue that should be considered in a sensor network design. Accordingly, this work is dedicated to presenting a framework for optimization of sensor distribution that considers both sensor faults under uncertainty and sensor clustering for vibration-based damage detection. Based on the effective independence method, the first design objective is newly formulated to consider sensor faults under uncertainty. Moreover, a novel index that is universally applicable for any type of structure is proposed to evaluate sensor clustering, which is treated as the second objective. The non-dominated sorting genetic algorithm II is adopted to solve this multi-objective optimization problem, and Monte Carlo simulation (MCS) is employed for uncertainty analysis in the first objective. To reduce computation costs, real performance evaluations in MCS are replaced with Gaussian process regression models. Based on the vibration information achieved from optimized sensors, an optimization-based damage detection process is applied to validate the optimal sensor layout. Three case studies (i.e., a cantilever beam, a laminated composite structure, and a spatial frame) are presented to demonstrate the effectiveness and applicability of the developed framework.

Keywords Optimal sensor placement · Sensor fault · Sensor clustering · Uncertainty · Damage detection

1 Introduction

Engineered structural systems, like bridges, wind turbines, and aerospace vehicles, among others, are subject to deterioration after a period of service (Ostachowicz et al. 2019).

This deterioration is caused by continuous use and/or harsh working environments, such as extreme temperatures, impact, and related issues. Extensive damage or failure of such systems and components can result in massive property loss and even in human casualties in extreme situations. Instantaneous evaluation of structural conditions is thus essential to conduct proper maintenance and avoid occurrence of hazards. With the goal of enhancing the integrity and safety of structures, structural health monitoring (SHM) and damage detection techniques have gained much attention from both academic and engineering communities (Tan and Zhang 2020). Accurate health condition evaluations are largely dependent on sensor systems that acquire data for analysis (Jayalakshmi et al. 2017; Bigoni et al. 2020). To collect as much information as possible, and also to have a minimal number of sensors for cost savings related to data processing, sensor network design is posed as an optimization problem by seeking the optimal number of sensors and

Responsible Editor: Chao Hu

✉ Byeng D. Youn
bdyoun@snu.ac.kr

¹ Department of Mechanical Engineering, Seoul National University, Seoul 08826, Republic of Korea

² Institute of Advanced Machines and Design, Seoul National University, Seoul 08826, Republic of Korea

³ OnePredict Inc., Seoul 06160, Republic of Korea

⁴ Department of Mechanical, Robotics and Energy Engineering, Dongguk University-Seoul, Seoul 04620, Republic of Korea

their locations from a discrete set of candidate positions. This is a typical combinatorial optimization problem (Yi and Li 2012; Chisari et al. 2017). A great deal of research has been performed to date on the topic of optimal sensor placement (OSP) for structural health monitoring and damage detection; corresponding optimization methodologies are mainly determined by developed SHM techniques, according to the classifications outlined in a review article (Ostachowicz et al. 2019). Vibration-based techniques for damage detection observe the changes in dynamic characteristics caused by any changes in mechanical properties of structures, for further deducing the health conditions (Barman et al. 2021). As the earliest proposed methods, vibration-based techniques have been commonly used in different structures. Optimal sensor deployment for vibration-based damage detection in structures is the main focus of the present work.

In vibration-based techniques for damage detection, accelerometer sensors should be placed in critical positions to capture the key dynamic information of the structure. The most popular and classical approach for optimally placing this type of sensor is the effective independence (EFI) method (Kammer 1991); this method is based on maximizing the determinant of the Fisher information matrix (FIM) to maintain the linear independence of the modal shapes, so that the best set of degrees-of-freedom (DOFs) locations are selected for sensors to measure. This is realized in an iterative process by gradually removing DOFs that contribute less to the linear independence from the modal shape matrix, resulting in a sub-optimal solution (Castro-Triguero et al. 2013). Meanwhile, some other approaches that are based on formal optimization strategies have been proposed; in these approaches, an objective function in terms of an assessment criterion (e.g., the determinant of the FIM or the root mean square for off-diagonal terms of the modal assurance criterion (MAC)) (Sun and Büyüköztürk 2015) is defined and then minimized/maximized. Since the OSP problem is a typical combinatorial optimization, the defined optimization problem is commonly solved using evolutionary algorithms, such as genetic algorithms and particle swarm algorithms (Liu et al. 2008; Rao and Anandakumar 2007), which do not require gradient information and are easily implemented. Those approaches have largely increased the possibility of achieving the global optimum solution for OSP problems. In real-world structures, the manufacturing process always carries different levels of uncertainties; thus, the performances of final products are influenced to a certain extent. These uncertainties should be accounted for in OSP problems. Some researchers have examined OSP techniques under uncertainties. Castro-Triguero et al. (Castro-Triguero et al. 2013) used Monte Carlo simulation to find optimal sensor positions for a truss structure under model parametric uncertainties. For each sample in the MCS, sensor locations

were selected using the EFI method and the most frequently selected locations among all samples were determined as the final optimal solution. Kim et al. (Kim et al. 2018) proposed a stochastic EFI method for optimal sensor placement in a similar truss structure under material uncertainty. Through the use of an interval possibility model, Yang et al. (Yang et al. 2020) solved uncertain OSP problems for sensor number determination and configuration optimization. Those efforts have demonstrated the significance of incorporating uncertainties associated with structures into OSP problems.

Because of their long-term service requirements, and various accidental factors, such as harsh working environments and intrinsic sensing malfunctions, sensor faults may occur in a sensor network; some common sensor fault types include bias (shifted signal values by a constant from the true value), drifting (continuous changes in deviation between the observed signal and the true value), precision degradation (stochastic deviations between the observed signal and the true value), and complete failure (loss of observed signal), among others (Balaban et al. 2009; Jesus et al. 2017; Jäger et al. 2018). It can be seen that sensor functions cannot be properly performed by faulty sensors, and false or null information may be provided, thus making the health evaluation system unreliable (Kullaa 2013). Yet the repair cost of those sensors may be too high considering limited budgets, or maintenance operations may be even impractical in some cases. However, a proper sensor distribution can effectively mitigate the adverse effects of sensor faults on information loss (Salari et al. 2019). Hence, this issue has motivated quite a few prior research efforts on optimal sensor configuration, especially in the field of traffic networks (Li and Ouyang 2011; Danczyk et al. 2016; Zhu et al. 2017); however, quite limited work on OSP has been conducted for structural systems. Staszewski et al. (Staszewski and Worden K et al. 2000) studied the OSP problem for impact detection and location in composite materials, considering the sensor fault of complete failure; in this work, because of the information loss resulting from sensor failure, the worst case among all sensor failure cases were assessed, and then used as the fitness in the genetic algorithm to find near-optimal sensor distributions for damage detection. However, this research team solved the problem only in a deterministic way, neglecting the real uncertainties in structural models. Accordingly, this research gap has formed one motivation of the present work: to optimally design a sensor network that accounts for sensor faults under uncertainty; the sensor fault type of complete failure (meaning complete loss of information) is examined herein. As for other types of sensor faults, such as bias and drifting, sensing performances could be mitigated by calibration and measurand reconstruction as well as some other signal processing techniques based on the observed information (Jesus et al. 2017), and they are not considered in the present work.

By placing multiple sensors in a local region, also known as sensor clustering, duplicated information can be provided for that local region; a previous study has shown that selecting multiple positions to configure sensors in a local region is basically equivalent to selecting one place to deploy a sensor (Friswell and Castro-Triguero 2015). This is indeed a good way to avoid information loss when encountering sensor faults. However, such redundant information can lead to wasted resources when all sensors are working normally; further, the costs in data acquisition and processing increase with sensor clustering. For these reasons, some strategies have been developed to disperse close sensors to collect as much information as possible; relatedly, redundant information elimination methods or indices have been proposed to alleviate sensor clustering configuration issues (Lu et al. 2016; Yang et al. 2019a). The popular EFI method always produces clustering of sensors when the number of placed sensors is larger than the number of target modal shapes. In order to overcome this disadvantage, Lian et al. (Lian et al. 2013) proposed a new fitness function that employed the nearest neighbor index to weight the FIM to avoid information redundancy. Taking the overall sensor configuration into account, Yang et al. (Yang et al. 2020) presented a novel CAD (clustering avoidance distribution) index with the integration of the center of sensor configuration, mean distance between sensors and their center, as well as the standard deviation of all sensor distances to the center. Moreover, by considering the local and global effects of sensor clustering, Yang et al. (Yang et al. 2019b) also developed a sub-clustering strategy including three main procedures: a sub-clustering algorithm, a check step, and the smallest enclosing circle method. However, these prior studies only considered the mean value of the nearest neighbor distance of the sensors; they neglected the effects of its standard deviation. This omission may result in misleading evaluation of sensor layouts in certain cases; this will be illustrated with examples in this paper. Furthermore, these prior approaches are mainly suitable for 2D surface sensor layouts; they lack general applicability for any type of structure. Thus, these two drawbacks provide another motivation for the work described in this paper, which seeks to develop an effective and universally applicable evaluation criterion for assessing sensor clustering. When taking sensor faults into account as well, sensor clustering could be an effective way to mitigate the problem of information loss because of sensor failures. So, it can be obviously seen that sensor clustering and sensor faults are two conflicting objectives when attempting to optimally configuring sensors. That forms the research target of the present work to involve both design objectives in the sensor placement problem, and to the best of the authors' knowledge, this work

simultaneously considers sensor fault and sensor clustering avoidance for the first time.

In the present work, a robust framework is presented for optimal sensor distribution considering both sensor faults under uncertainty and sensor clustering for vibration-based damage detection. Material properties of the observed structure are considered as random variables. Based on the EFI method, a sensor layout that has the maximum determinant of the FIM is sought. With sensors suffering complete failure, the minimal determinant among all possible failures is to be maximized. Under consideration of model uncertainty, the sum of the mean value of that minimal determinant and its standard deviation replaces the original deterministic determinant. Thus, the first design objective is newly formulated, accounting for both sensor faults and model uncertainty. To avoid sensor clustering, a novel index applicable for assessing any type of sensor configuration is proposed, involving the mean value of the nearest neighbor distance for each sensor, its standard deviation value, as well as the minimum volume to enclose all sensor positions; this is then treated as the second design objective. The non-dominated sorting genetic algorithm II (NSGA-II) is adopted as the optimizer to solve this problem and Monte Carlo simulation (MCS) is applied for uncertainty analysis. To save computation costs, a Gaussian process regression (GPR) model is employed to approximate real performance evaluations in the MCS. After achieving optimized sensors, an optimization-based damage detection process is implemented with the measured incomplete modal data from the sensors, where the associated optimization problem is solved with a genetic algorithm (GA). The optimized sensor layout is then validated.

The present work first proposes a novel framework for robust sensor distribution that considers both sensor fault uncertainty and sensor clustering for vibration-based damage detection; then, that framework is applied to three case studies to demonstrate its efficacy. More specifically, the first design objective that considers sensor faults under uncertainty is formulated based on the EFI method outlined in Sect. 2; the second design objective, evaluating sensor clustering, is developed based on a novel index proposed in Sect. 3. Section 4 first constructs the problem formulation and then presents the optimization methodology to address this problem. Section 5 is focused on presenting an optimization-based damage detection process for validation of optimized sensors. Three case studies—a cantilever beam, a stiffened composite panel, and a spatial frame—are presented in Sect. 6, to demonstrate the efficacy of the developed optimization framework. Concluding remarks are provided in Sect. 7.

2 Optimization objective based on the EFI method, considering sensor faults under uncertainty

The original EFI method for optimal sensor placement is only suitable for well-functioning sensors and also only for deterministic structural models. When considering both sensor faults and model uncertainty, the original design objective, in terms of the determinant of the FIM, requires a reformulation. In this section, a new objective function based on the EFI method is defined to incorporate sensor faults and model uncertainty into the OSP problem.

The sensor network attempts to gather as much dynamic information as possible about the measured structure with a limited number of sensors; the amount of information is positively associated with the linear independency of the modal vectors observed by the sensors. The original EFI method is such an approach; it maximizes the linear independency of modal vectors so as to collect maximal information (Kammer 1991). With m target modes to be identified, the output vibration signal from n placed sensors can be expressed as

$$y_s = \Phi q + w \tag{1}$$

where y_s is the vector of the outputs at selected sensor locations; Φ is the modal shape matrix with $n \times m$ dimensions; q is the target modal coordinates; and w is the zero-mean Gaussian white noise with the variance of σ^2 . Evaluating the target modal coordinates using an unbiased estimator and estimating the covariance of error gives the following expression:

$$P = E[(q - \hat{q})(q - \hat{q})^T] = \left[\frac{1}{\sigma^2} \Phi^T \Phi \right]^{-1} = Q^{-1} \tag{2}$$

where Q is the Fisher information matrix (FIM). The best estimation of q is achieved when minimizing the error covariance matrix P , which is equivalent to maximizing the FIM. For simplification, the measurement noise is assumed to be uncorrelated and identical statistical properties are assumed for each sensor; thus, the Fisher information matrix becomes

$$Q = \Phi^T \Phi \tag{3}$$

In this way, the EFI method seeks the best sensor locations by maximizing the following objective:

$$f_1(s) = \det(Q(s)) = \det(\Phi^T(s)\Phi(s)) \tag{4}$$

where s stands for the design variable vector of the sensor location, with its element of s_i ($i = 1, \dots, n$).

Because of their long-term usage and adverse working environments, sensor faults may occur, such as bias, drifting, precision degradation, or even complete failure. The sensor fault of complete failure is so severe that the information

from the measured structure at the location of the faulty sensor is completely lost; this can lead to task failure of accurate monitoring and prevent proper decision-making. However, optimal configuration of sensors is a good way to mitigate the adverse effects of sensor faults on information loss.

Let L be the maximum number of allowable faulty sensors in an OSP problem. For a case in which the number of faulty sensors from an initial sensor distribution of s is k_l ($k_l = 0, 1, \dots, L$), a sensor configuration only consisting of the remaining effective sensors is then obtained as

$$s_f^{(k_l)} = \begin{cases} s, & k_l = 0 \\ s \setminus \{s_{k_1}, \dots, s_{k_j}, \dots, s_{k_l}\}^T \ (k_j \in \{1, \dots, n\}), & k_l = 1, \dots, L \end{cases} \tag{5}$$

where $s \setminus \{s_{k_1}, \dots, s_{k_j}, \dots, s_{k_l}\}^T$ stands for removing all faulty sensors, from s_{k_1} to s_{k_l} , from the initial sensor distribution of s . Obviously, $s_f^{(k_l)}$ only contains sensors without any faults and is a subset of s . Considering all possible faulty cases under the condition in which the maximum number of allowable faulty sensors is L , a set of sensor configurations composed of the remaining well-functioning sensors is achieved as follows:

$$s_f = \{s_f^{(0)}, s_f^{(1)}, \dots, s_f^{(k_l)}, \dots, s_f^{(L)}\}, k_l = 0, 1, \dots, L \tag{6}$$

When accounting for a sensor fault of complete failure in the OSP problem, it is hard to tell which sensor(s) will fail; this implies that the sensor fault presents uncertainty. For this reason, the most severe fault case, which has the minimal value of $f_1(s)$ among the set of well-functioning sensor configurations, should be sought first; subsequently, this minimal value requires maximization to optimally configure the sensor location of s , so as to improve the robustness of the sensor configuration with respect to the sensor fault. Correspondingly, to cope with the sensor fault in the OSP problem, the modified objective to be maximized is formulated as follows:

$$f_2(s) = \min_{s_f} f_1(s) = \min_{s_f} \det(\Phi^T(s)\Phi(s)) \tag{7}$$

Uncertainties always exist in models of real-world structures; these uncertainties arise from several causes, one of which is the complex manufacturing process. Model uncertainty influences structural performance and further sensor configurations. Thus, model uncertainty should be taken into account to reduce the variation of information collected by sensors with respect to the random variables in the OSP problem. Thus, this becomes a robust optimization problem. However, both of the objectives formulated in Eqs. (1) and (7) are only applicable for deterministic models. Thus, reformulation is essential to incorporate the model uncertainty into the objective function. In robust design optimization of structures, the

original deterministic design objective is always replaced with a linear combination of the mean and standard deviation of a performance function. Likewise, in the present OSP problem, based on the formulation in Eq. (7), the new design objective to be maximized is constructed by combining the mean and standard deviation of the determinant of the FIM, as presented below:

$$f_3(s, X) = \min_{s_f} \mu_{\det}(s, X) - \lambda \cdot \sigma_{\det}(s, X) \tag{8}$$

where $\mu_{\det}(s, X)$ and $\sigma_{\det}(s, X)$ represent the mean and standard deviation of the determinant of the FIM, respectively, with respect to random variables of X , under a specific sensor layout of s ; λ is a weight factor to balance the mean and standard deviation (in the present work, $\lambda = 1$).

Using Eq. (8), the most severe faulty case, which has a smaller value of the mean of the determinant of the FIM and also has a larger value of standard deviation of the determinant of the FIM, is sought among the set of well-functioning sensor configurations. By maximizing this severe fault case, i.e., maximizing Eq. (8) with design variables of s , the best sensor configuration is obtained, which is the one most robust to sensor fault and model uncertainty. Consequently, based on the original EFI method, a new objective function of Eq. (8) is formulated to alleviate the adverse effects of sensor faults and model uncertainty, which should be maximized to find a robust sensor layout.

3 Optimization objective based on a novel index for evaluation of sensor clustering

Sensor clustering, or configuring multiple sensors in a local region, provides repeated and redundant information; this can be another effective way to mitigate the negative impacts of sensor faults on information loss. However, placing multiple backup sensors produces high costs in data acquisition and data processing; this approach also causes a waste of resources when no sensor faults occur. Many studies have been performed to disperse sensors and avoid sensor clustering; correspondingly, some indices have been developed to evaluate sensor clustering. In this section, a brief overview of some recent evaluation indices for sensor clustering is provided, pointing out their limitations and disadvantages that could result in misleading assessments in certain cases. Then, a novel evaluation index for sensor clustering is proposed and its effectiveness is demonstrated with examples.

3.1 An overview of representative evaluation indices for sensor clustering and their deficiencies

The EFI method has proved to be one of the most widely used OSP techniques for maximization of the determinant of

the FIM. However, sensor clustering appears when the number of placed sensors surpasses the number of target modes to be identified. Since the objective of the EFI method is to maximize the determinant of the FIM, the spatial correlation is neglected, which results in clustering sensors in two adjacent nodal positions or DOFs. To avoid the redundancy of information caused by sensor clustering in the EFI method, Lian et al. (Lian et al. 2013) proposed a novel fitness function that combines the determinant of the FIM with the nearest neighbor index (NNI). The NNI was employed to evaluate the sensor clustering condition; it compares the distances between the nearest points and expected distances based on chance, which is expressed as follows:

$$f_{NNI} = \frac{D(NN)}{D(\text{ran})} = \frac{\sum_{i=1}^n \frac{\min(D_{ii'})}{n}}{\frac{1}{2} \sqrt{\frac{A}{n}}} \tag{9}$$

where $D(NN)$ is the nearest neighbor distance; $D_{ii'}$ is the distance between sensor position i and i' ($i' = 1, \dots, n$ and $i' \neq i$), and $\min(D_{ii'})$ is the nearest neighbor distance for i ; A is the area of the structural region. A larger value of this index is expected, which indicates that the sensors are dispersed.

The above index in Eq. (9) considers the ratio of the nearest neighbor distance over the structural region, while it neglects the overall sensor configurations. Yang (Yang 2018) demonstrated that this index could cause mistakes in some special cases. Taking the overall sensor configuration into consideration, Yang et al. (2020) developed a CAD (clustering avoidance distribution) index, as shown below:

$$f_{CAD} = \frac{(\mu_s - \sigma_s) \sum_{i=1}^n \min(D_{ii'})}{2A} \tag{10}$$

where μ_s and σ_s are the mean and standard deviation of all located sensor distances to their center. The mean of μ_s and the center of the sensor configuration are expressed in the following two equations:

$$\mu_s = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{11}$$

$$(x_c, y_c) = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i \right) \tag{12}$$

where x_i and y_i are the location coordinates of the i -th sensor, respectively.

Considering both the local and global effects in sensor configuration, Yang et al. (Yang et al. 2019b) developed another index called a redundancy elimination model (REM), based on a sub-clustering strategy, which is suitable for evaluating both local and global sensor distributions. This index is defined as follows:

$$f_{REM} = \frac{r_{\min} \sum_{i=1}^n \min(D_{ii'})}{nA_{\min}} \tag{13}$$

where A_{\min} is the minimum area of a circle that encloses all sensors, i.e., the area of the smallest enclosing circle, and r_{\min} is its corresponding radius.

From those three recent indices listed above, we can see that they share one limitation: they are only applicable for sensor distributions in surface-type structures. Further, those indices mainly focus on the sum or mean of the nearest neighbor distance, neglecting the variance effect, which can lead to mistaken evaluations in certain cases. In order to illustrate this drawback, three cases of sensor layouts, as shown in Fig. 1, where purple dots stand for sensors and dashed blue curves represent the smallest circle that encloses the sensors, are evaluated with those three indices. In each case, 5 sensors are deployed. The three cases of sensor

layouts are assumed to be deployed in a square structural region, with dimensions of 2×2 . From the sensor coordinates presented in Table 1, there is only one difference in sensor locations between Case I and Case II, i.e., the central sensor is moved to the enclosing boundary. In Case III, the sensors are uniformly distributed on the enclosing circle; this enclosing circle is much smaller than that in Case I.

The three cases of sensor layouts are evaluated with the aforementioned three indices, and the evaluation results are tabulated in Table 1. Here, it should be noted that a larger value of any index above indicates a more dispersed sensor distribution; a smaller index value denotes a more clustered sensor deployment. Further, since our goal is to identify the difference between any two cases of sensor layouts, we focus on the relative index value between the cases of sensor layouts, rather than the absolute index value itself. Under one evaluation criterion, the index value is maximized among

Fig. 1 Three cases of sensor layouts using five sensors

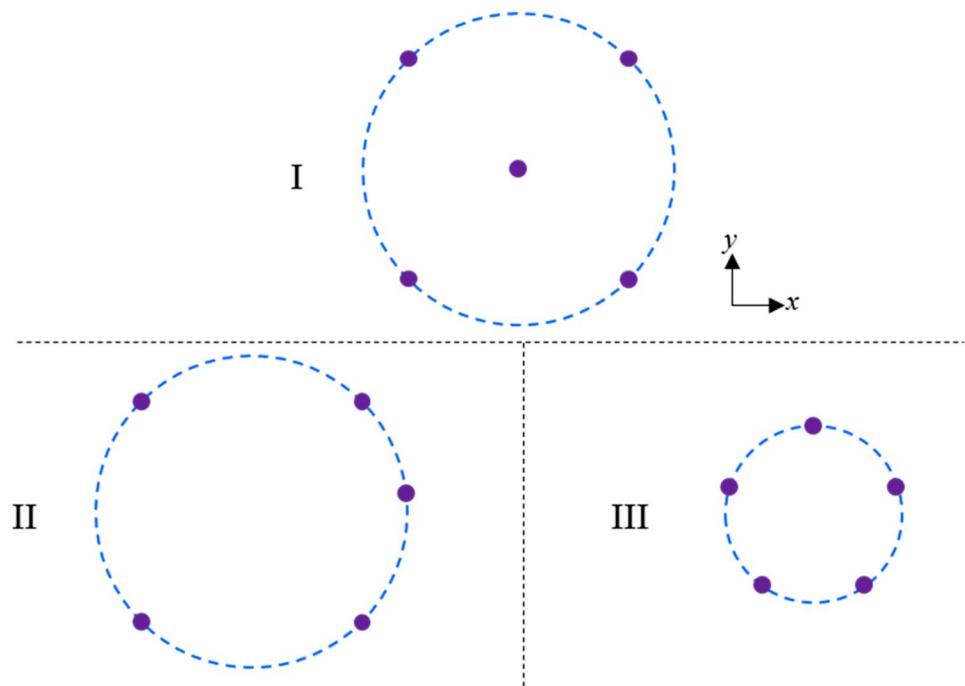


Table 1 Evaluation results of sensor clustering for three example cases

Case	Sensor coordinates	Evaluation indices for sensor clustering			
		f_{NNI} (Lian et al. 2013)	f_{CAD} (Yang et al. 2020)	f_{REM} (Yang et al. 2019b)	f_{SCI} (Present work)
I	(0,0), (1,1), (-1,1), (-1,-1), (1,-1)	3.1623	0.5000	0.3183	1.2676
II	(1.4023,0.18285), (1,1), (-1,1), (-1,-1), (1,-1)	3.1623	1.0215	0.3183	1.0206
III	(0.78456,0.25492), (0,0.82494), (-0.78456,0.25492), (-0.48489,-0.66739), (0.48489,-0.66739)	2.1685	0.5000	0.3742	0.8118

all cases of sensor layouts to find the least clustered case; instead, for each case, the comparison of the index value among all evaluation criteria has no significance. In this way, it can be seen that the NNI index produces the same evaluation result for Case I and Case II. This indicates that this index cannot identify the difference between these two cases, showing the drawback of this index, which has been demonstrated by Yang (2018) using other cases of sensor layouts. When this index is used to compare the clustering condition between Case I (or Case II) and Case III, it can be found that Case I (or Case II) has a larger index value than Case III, which correctly recognizes their difference, as the sensors in Case III are obviously clustered, as compared with Case I (or Case II). When using the CAD index, a larger index value is generated for Case II when comparing Case I and Case II, concluding that Case II should be more dispersed than Case I; however, this conclusion goes against the reality that sensor clustering appears in Case II, while sensors in Case I are evenly distributed, as can be easily seen from Fig. 1. Moreover, the same value is obtained for Case I and Case III, showing that the CAD index cannot tell the difference between these two cases. As for the REM index, it cannot distinguish between Case I and Case II. In addition, it makes a wrong judgment between Case I (or Case II) and Case III, giving a relatively larger index value in Case III, which means Case III has a better scatter of sensor layout than Case I (or Case II). However, this conclusion does not match the facts; it is totally opposite of that determined by the NNI index.

From the observations above, we can determine that the NNI and REM indices cannot identify the difference between Case I and Case II, mainly due to the fact that they only consider the mean of the nearest neighbor distance, while they neglect its standard deviation. Even though the CAD index considers the overall sensor configuration, it produces a misleading judgment between these two cases; when evaluating Case I and Case III, only the NNI index makes a correct assessment, while the CAD index cannot distinguish them and the REM index may even produce an incorrect evaluation. The drawbacks of these indices motivate us to propose a new evaluation index to correctly identify sensor clustering conditions.

3.2 A novel evaluation index for sensor clustering

As stated above, previous indices have a limitation in that they are only applicable for sensor distributions in surface-type structures; they also have drawbacks in that they cannot identify the difference between sensor layouts in some cases and may even produce incorrect evaluations. In order to overcome these shortcomings, a novel sensor clustering index (SCI) is proposed in this work. The proposed SCI can be applicable for curved, surface, and spatial types of

structures and can also correctly evaluate sensor clustering conditions.

By simultaneously accounting for the mean value of the nearest neighbor distance for each sensor and its standard deviation value, as well as the minimum volume to enclose all sensor positions, a novel index to evaluate the clustering condition of sensor layout s is proposed as follows:

$$f_{SCI}(s) = \frac{\mu_{nnd}(s) - \sigma_{nnd}(s) + 1/\rho^\chi(s)}{d_s} \tag{14}$$

where $\mu_{nnd}(s)$ and $\sigma_{nnd}(s)$ are the mean and standard deviation of the nearest neighbor distance of $\min_i(D_{i_i'})$, respectively, related to all sensor locations; $\rho(s)$ is the density of the sensor distribution, and the exponent of χ equals 1, 1/2, and 1/3, respectively, for a curved, surface, and spatial sensor layout; d_s is the feature size of the monitored structure. The density of the sensor distribution is defined by the number of sensors over the minimum volume that encloses all sensors, as expressed below:

$$\rho(s) = \begin{cases} n/L_{min}(s), & \text{for a curved layout} \\ n/A_{min}(s), & \text{for a surface layout} \\ n/V_{min}(s), & \text{for a spatial layout} \end{cases} \tag{15}$$

where $L_{min}(s)$ is the minimum length of the curve that encloses all sensors of s when sensors are configured on a curve-type structure; $A_{min}(s)$ is the minimum area of the enclosing circle for sensors that are placed on a surface-type structure; and $V_{min}(s)$ is the minimum volume of the enclosing sphere if sensors are configured on a spatial structure. The length of the shortest enclosing curve, $L_{min}(s)$, can be easily obtained by computing the distance between any two sensor locations and then finding the maximum. As for the approach to get the minimum enclosing circle and sphere, readers are directed to Refs. (Welzl 1991; Ritter 1990).

When the index proposed in Eq. (14) is used as an optimization objective, it should be maximized, indicating that a larger value of this index means a more dispersed sensor layout. Moreover, as the density of the sensor distribution is concerned in this formulation, maximization of Eq. (14) also means minimizing the number of employed sensors. In order to show the effectiveness of this proposed evaluation index, it is used to assess the three sensor configurations presented in Fig. 1; the evaluation results are shown in the last column of Table 1. Since the dimensions of the structural region are 2×2 , the feature size for this structure, d_s , is selected as 2. Three different evaluation values are generated corresponding to three different sensor layout cases.

For Case I and Case II, even though the mean values of the nearest neighbor distances and the areas of the smallest enclosing circles for both cases are the same, as reflected by the same evaluation value of the NNI and REM indices, the

standard deviation values of the nearest neighbor distances are different. The difference in the standard deviation values is reflected in the formulation of the proposed index of SCI, leading to the result that Case I has a larger value of SCI than that of Case II, which indicates that Case I has a more dispersed sensor layout than Case II. As mentioned above, the only difference between these two cases is that the central sensor in Case I is moved to the enclosing circle boundary of Case II, causing a local sensor clustering in Case II; the proposed index successfully identifies this difference. For Case I and Case III, it is evident that the mean value of the nearest neighbor distance in Case I is larger than that in Case III, while the corresponding standard deviation values for both cases are zero. The occupied area in Case I is also apparently larger than that in Case III; as a consequence, a larger index value of SCI is produced in Case I, also successfully identifying the difference between Case I and Case III. Additionally, when looking at Case II and Case III, a larger index value has been achieved in Case II; this is mainly caused by the larger mean value of the nearest neighbor distance and also the larger occupied circle area. Among all these three cases, Case I has the largest index value, meaning that Case I has the best scattered sensor distribution; this can also be easily observed graphically from Fig. 1.

The successful application of the proposed approach in the three example sensor layouts in Fig. 1 demonstrates that the proposed SCI index can correctly identify sensor clustering conditions. Further, from the formulation of Eqs. (14) and (15), this proposed index can break through the limitation of being applicable only for a specific structure; rather, it is universally applicable for sensor layouts in any type of structure. This novel index is then used as an objective to be maximized in the present OSP problem.

4 Problem formulation and optimization methodology for OSP

Considering both sensor faults under uncertainty and sensor clustering, a robust design problem for sensor placement optimization is first formulated in this section. To solve this design problem, a novel and efficient global optimization strategy is developed; the techniques employed in this optimization strategy are elaborated in this section.

4.1 Problem formulation of OSP, considering sensor faults under uncertainty and sensor clustering

In this work, the OSP problem treats both the number of sensors, n , and corresponding sensor locations, $s = \{s_1, \dots, s_n\}^T$, as design variables. Suppose that DOFs of all nodes in the finite element model of the measured structure compose a set, \mathcal{S} , which stands for the available

sensor placement space. Then, \mathcal{S} is a subset of \mathcal{S} , i.e., $s \subset \mathcal{S}$. The design variable vector, \mathbf{S} , is formed as $\mathbf{S} = \{n; s\}$. To determine a robust sensor placement design with respect to sensor faults and model uncertainty, the new objective function in formulated in Eq. (8) should be maximized; simultaneously, to avoid sensor clustering, the novel evaluation index for sensor clustering proposed in Eq. (14) should be also be maximized. Accordingly, the sensor placement design problem is formulated in a typical form of an optimization problem, which is presented as follows:

$$\begin{cases} \text{find } \mathbf{S} = \{n; s\} = \{n, s_1, \dots, s_n\}^T \\ \text{min } \{f_{\text{SF}} = 1/f_3(\mathbf{S}, \mathbf{X}), f_{\text{SC}} = 1/f_{\text{SCI}}(\mathbf{S})\}^T \\ \text{s.t. } n_{(l)} \leq n \leq n_{(u)}, \\ s \subset \mathcal{S}. \end{cases} \quad (16)$$

where \mathbf{X} denotes the random variable vector of model uncertainty; f_{SF} represents the design objective, considering sensor faults under uncertainty; f_{SC} standards for the design objective, considering sensor clustering; and, $n_{(l)}$ and $n_{(u)}$ are the lower and upper bounds on the number of sensors, respectively. In addition, it should be noted that the objective function of f_{SF} is evaluated under the sensor fault condition, while the other objective function of f_{SC} is evaluated under the assumption that all sensors are normal.

The set of \mathcal{S} representing DOFs of all available nodes is fixed after the finite element model of the structure is constructed. As sensor locations are a subset of \mathcal{S} , all design variables in \mathcal{S} take discrete values. To the best of our knowledge, this is the first attempt to simultaneously consider sensor fault and model uncertainty, as well as sensor clustering, in OSP problems. The solving process for this problem is illustrated in the following subsection.

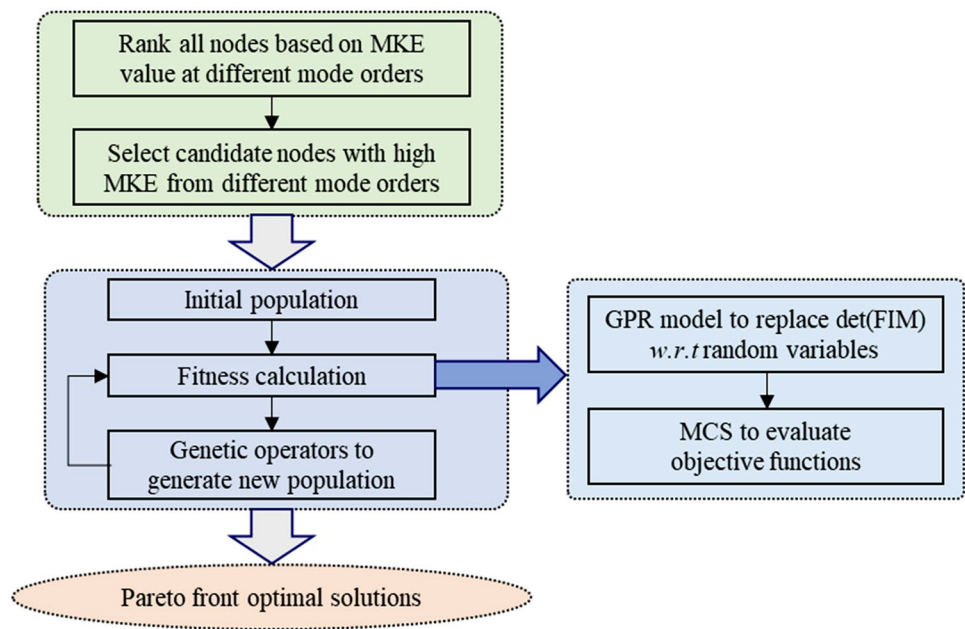
4.2 Optimization methodology

4.2.1 Global optimization strategy

The problem of Eq. (16) involves discrete variables, multiple objectives, and uncertainty analysis. To deal with this problem, a novel and efficient global optimization strategy is developed and presented in Fig. 2.

Equation (16) implies that the design space is composed of all available DOFs of the finite element model associated with the measured structure and optimal sensor locations are a subset of this design space. For a large and complex structure, plentiful nodes and DOFs are generated, especially when a refined finite element mesh is employed, leading to a large design space and further influencing the efficiency in optimization. To shrink the design space to improve the efficiency, while maintaining the optimal solution in the shrunken design space, a modal kinetic energy (MKE)-based

Fig. 2 Proposed global optimization strategy for the OSP problem



index is proposed to remove some unnecessary candidate node locations from the original design space. This modal kinetic index was used directly in previous studies to select optimal sensor locations with high MKE values (Li et al. 2007); this provided a rough measure of the dynamic contribution of candidate sensor locations to target modes but neglected the linear independence of the modal vectors (Gomes et al. 2018). In the present study, instead of directly using MKE to determine optimal sensor locations, MKE is used to choose nodes that have relatively high values of MKE at each target mode; gathering those chosen nodes and combining them together forms a shrunken design space. Nodes with low MKE values, which exhibit low vibrational amplitudes and contribute little in selecting optimal sensor locations, are removed from the original design space; therefore, the narrowed design space still presents enough dynamic information of the structure. Moreover, since the linear independence of the measured modal vectors is considered in the objective, the optimal solution obtained in the shrunken design space maintains both a good level of linear independence for target modes and a high level of MKE values.

To solve the present optimization problem, the non-dominated sorting genetic algorithm II (NSGA-II) is adopted as the optimizer; NSGA-II has the ability to deal with multiple objectives and also discrete variables. In order to evaluate the mean and standard deviation of the determinant of the FIM in the objective function, Monte Carlo simulation (MCS) is employed. MCS can provide an excellent estimation when the number of sampling points is large enough; the analysis result of MCS is often regarded as the exact solution. However, this is a quite time-consuming process,

as it involves numerous deterministic function evaluations at random samples, especially when used together with a population-based evolutionary algorithm. To reduce the computational cost in MCS, a surrogate model, which replaces the real function evaluations in terms of the determinant of the FIM, is constructed based on the Gaussian process regression (GPR) modeling technique. Accordingly, an accurate and efficient uncertainty analysis procedure is proposed.

The global optimization strategy proposed to solve the problem of Eq. (16) is briefly introduced, and the techniques used are illustrated in the following subsections.

4.2.2 MKE-based index to narrow the design space

According to its definition, the computation of modal kinetic energy is shown as follows:

$$MKE_{k,q} = \sum_{p_k} \Phi_{p_k,q} \sum_{l=1}^{N_0} M_{p_k,l} \Phi_{l,q} \tag{17}$$

where $MKE_{k,q}$ denotes the modal kinetic energy value for the k -th node in the q -th target mode; p_k represents the DOF that is related to the k -th node; $\Phi_{p_k,q(l)}$ is the $p_k(l)$ -th component in the q -th modal vector, and $M_{p_k,l}$ is an element in the mass matrix, \mathbf{M} , in terms of the p_k -th row and the l -th column. The MKE can roughly measure the dynamic contribution of candidate sensor locations to target modes; it was previously used to select nodes and DOFs that possess the highest values of MKE as sensor locations. In this way, most of the relevant dynamic information can be captured. However, the linear independence of target modes was not taken into consideration in this approach (Gomes et al. 2018).

In this work, we use MKE as a reference index to narrow the original design space. At each target mode, MKE values for all nodes are first calculated, and nodes with relatively high values of MKE are picked. To this end, the maximum MKE value at each mode is found, and nodes whose associated MKE values are larger than a certain percentage of that maximum value (i.e., $\text{MKE}_{k,q} \geq \xi \cdot \{\text{MKE}_{k,q}\}_{\max}$, where $\xi = 0.2 \sim 0.3$ is recommended) are chosen. Following this process, the nodes selected at each target mode are collected and combined together as a set, and DOFs associated with the nodes in this set produce a shrunken design space. This reduced design space is used as a substitute for the full design space of in Eq. (16). This approach removes DOFs for related nodes that have relatively low values of MKE from the original design space. Since those DOFs present small amplitudes in dynamic signals, they always have little effect on optimal selection of sensors. This procedure of shrinking the design space can save computation costs, especially for large and complex structures with a great number of nodes and DOFs; this approach has been successfully applied in our previous work on optimal sensor placement in composite structures (An et al. 2022).

4.2.3 GPR-model-based NSGA-II to optimize multiple objectives and discrete design variables

The non-dominated sorting genetic algorithm II (NSGA-II) has proved to be a computationally fast and elitist multi-objective evolutionary algorithm that can find a good spread of Pareto front solutions. Further, it has the ability to handle both discrete and continuous design variables. More details about this algorithm can be found in Ref. (Deb et al. 2002). In this study, NSGA-II is employed to solve the problem outlined in Eq. (16).

In order to compute the probabilistic objective of f_{SF} , Monte Carlo simulation (MCS) is used to find the mean and standard deviation of the determinant of the FIM with respect to random variables; this is actually a highly computationally expensive process. For efficiency improvement, the Gaussian process regression (GPR) model is used as a surrogate model to replace the real computation of the determinant of the FIM in the MCS. Gaussian process regression is a supervised learning algorithm in a machine learning task. GPR makes predictions for new inputs based on a series of training datasets of known inputs and outputs; it is a flexible, non-parametric Bayesian model that can make predictions in terms of the posterior mean and variance, which have been widely used in optimization problems (An et al. 2021). For more details about GPR, readers are referred to Refs. (MacKay 1998; Williams and Rasmussen 2006).

The Latin Hypercube sampling (LHS) method is used to generate the training data points for constructing the GPR model. After the GPR model is constructed, its accuracy

is measured with two metrics using test points, which are R -square and relative average absolute error (RAAE) (Jin et al. 2001), as formulated below:

$$R^2 = 1 - \frac{\sum_{j=1}^{J_0} (y_j - \hat{y}_j)^2}{\sum_{j=1}^{J_0} (y_j - \bar{y}_j)^2} \quad (18)$$

$$\text{RAAE} = \frac{\sum_{j=1}^{J_0} |y_j - \hat{y}_j|}{(J_0 * \text{STD})} \quad (19)$$

where y_j is the real response value at the j -th test point; \hat{y}_j denotes the predicted value from the GPR model; \bar{y}_j represents the mean of the real values at all test points; J_0 is the number of all test points, and STD stands for the standard deviation. A larger value of R -square and a smaller value of RAAE denote higher accuracy of the surrogate model.

5 Problem formulation for optimization-based damage detection

Successful damage detection based on the information measured from sensors can validate the optimized sensor layout. Thus, in order to demonstrate the optimized sensors, an optimization-based damage detection process is implemented using data provided by the optimized sensors.

In vibration-based damage detection, modal flexibility has proved to be a damage-sensitive parameter, as compared with the natural frequency and the modal shape (Zhao and DeWolf 1999; Dinh-Cong et al. 2018). The modal flexibility is a derivative of natural frequencies and modal shapes, which is defined as follows:

$$F = [\Phi_1, \dots, \Phi_m] \cdot \text{diag}\{1/\omega_1^2, \dots, 1/\omega_m^2\} \cdot [\Phi_1^T, \dots, \Phi_m^T] \quad (20)$$

where Φ_I and ω_I ($I = 1, \dots, m$) are the modal shape vector and the natural frequency for the I -th mode, respectively. Using this modal flexibility matrix, a design objective is formulated for the optimization-based damage detection process, as shown below:

$$\text{Obj}(\mathbf{x}) = \frac{\|F_{\text{ms}} - F(\mathbf{x})\|_{\text{Frob}}}{\|F_{\text{ms}}\|_{\text{Frob}}} \quad (21)$$

where \mathbf{x} is a variable vector regarding damage, whose components include damage location and damage degree; F_{ms} is the modal flexibility matrix measured with sensors from the damaged model; $F(\mathbf{x})$ is the modal flexibility matrix to deduce damage conditions from a numerical model, and $\|\cdot\|_{\text{Frob}}$ represents the Frobenius norm for a matrix.

Based on the objective function in Eq. (21), the damage detection process is formulated as an optimization problem, as presented below:

$$\begin{cases} \text{find } \mathbf{x} \\ \text{min } Obj(\mathbf{x}) \\ \text{s.t. } \mathbf{x}_{(l)} \leq \mathbf{x} \leq \mathbf{x}_{(u)} \end{cases} \quad (22)$$

where $\mathbf{x}_{(l)}$ and $\mathbf{x}_{(u)}$ are the lower and upper bound vectors for variable \mathbf{x} , respectively. To solve this typical minimization problem, a real-coded genetic algorithm is adopted as the optimizer, and the simulated binary cross-over operator and polynomial mutation are used in this real-coded GA (Deb and Agrawal 1995).

6 Case studies

In this section, the proposed strategy for sensor placement optimization, which considers both sensor faults under uncertainty and sensor clustering, is examined by applying it to three case studies with different structures, i.e., a 1D cantilever beam, a 2D composite plate, and a 3D space frame. Different cases with varying numbers of allowable faulty sensors are tested, and a set of non-dominated solutions is achieved. Based on the dynamic information measured with optimized sensors, optimization-based damage detection is conducted. Results verify the usefulness of the developed optimization strategy.

To shrink the design space, the value of parameter ξ used in the MKE-based index is set as 0.2 for all case studies. Using the NSGA-II to solve the sensor placement problem, the population size and maximum generation are given as 100 and 500, respectively. Considering the randomness of the algorithm, 20 independent runs are repeated for each problem. When using GA to conduct optimization-based damage detection, the probability of cross-over and mutation are specified as 0.8 and 0.2, respectively.

6.1 A cantilever beam – 1D curved sensor layout

The first case study considers a simple structure of a cantilever beam with its left end clamped. A finite element model of this beam is built with 50 elements, and the element numbering is assumed to be from the leftmost gradually to the rightmost, starting from 1 to 50. The beam has a length of 1 m, a width of 20 mm, and a thickness of 2 mm. The mean values for the elastic modulus of material property, Poisson's ratio, and density are 70 GPa, 0.33, and 2700 kg/m³, respectively. All of these measures follow normal distributions, with coefficients of variation of 8%, 5%, and 4%, respectively. Only the in-plane transverse vibrations are considered for this simple structure, and accordingly, the translational DOFs at every node of the finite element model are candidate sensor locations. The first three modes are considered as the target modes, as depicted in Fig. 3, where circles represent nodes of the finite element mesh, and also represent the candidate sensor locations.

6.1.1 Optimal sensor placement

The proposed OSP strategy that considers sensor faults under uncertainty and sensor clustering for better data acquisition is first investigated. For this purpose, the strategy is performed by assuming different numbers of allowable faulty sensors, i.e., 0, 1, and 2, respectively. Using the MKE-based index to narrow the design space, 44 nodes are retained out of the original 50 nodes; the first 6 nodes on the left end are removed from the initial design space. From Fig. 3, we can see that the vibration amplitudes for the three modes of concern are relatively low on the left part, implying that they have small modal kinetic energy values; as a consequence, some nodes on the left side are deleted from the initial design space. To construct the GPR model, 50 data points are generated using the Latin hypercube sampling method; approximation accuracy of the constructed surrogate model is tested with another 200 random points. As a result, the *R*-square and RAAE values for this surrogate

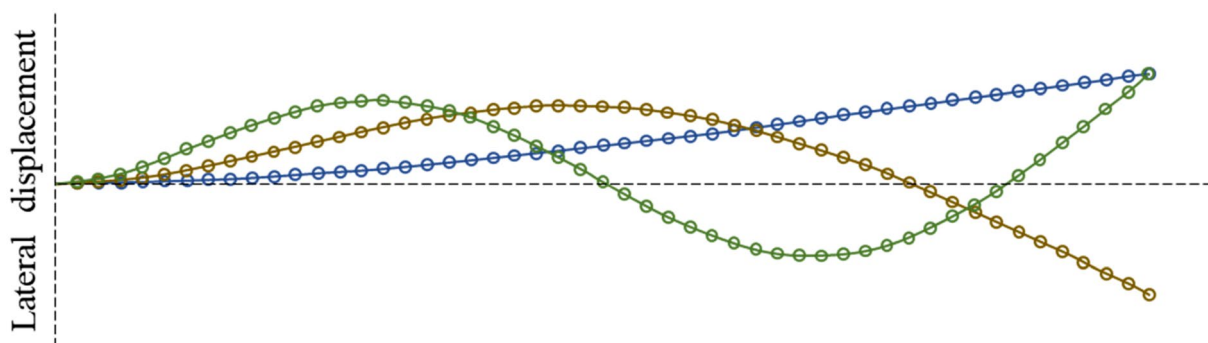


Fig. 3 The first three mode shapes of the discretized cantilever beam (clamped at the left end)

model are 1.0 and 0.0024, respectively, showing a quite high level of accuracy. To evaluate the sensor clustering index of f_{SCI} in Eq. (14), the feature size of this beam structure is assigned as 1 m, which is just the beam length. The lower and upper bounds on the number of sensors are given as 3 and 6, respectively. After optimization, all non-dominated solutions obtained for this problem are summarized in Fig. 4.

From Fig. 4, it can be easily seen that when the maximum number of allowable faulty sensors increases, the objective function values in terms of the determinant of FIM (i.e., $\min_{s_f} \mu_{det}(\mathbf{S}, \mathbf{X}) - \sigma_{det}(\mathbf{S}, \mathbf{X})$) decrease gradually. With the complete failure of sensors, the information obtained from sensors is reduced, further resulting in the decrease of the objective function regarding the determinant of the FIM. As for the range of the sensor clustering index in the non-dominated solutions, there is no significant difference between the case of no faulty sensors and the case of one faulty sensor; both cases are in the range of 0.12~0.71. In contrast, when the number of faulty sensors gets larger, the range gets smaller (it becomes 0.16~0.71); the main change lies in the lower bound. The results show that an increased number of faulty sensors also affects values of the sensor clustering index in the non-dominated solutions, implying that the optimized sensor locations are influenced by sensor faults.

With so many non-dominated solutions, it is hard to determine which solution among them should be selected for next-step damage detection; thus, post-processing of the results is required to assist in selecting an appropriate solution. A common way to treat multiple optimization objectives is to combine them with some simple arithmetic expressions, such as exponents, logarithms, and products (Yang 2021; An et al. 2018). Herein, to properly balance

the sensor placement performance and the sensor clustering index, a combined function is defined by introducing a weighting factor and normalization, as expressed below:

$$\bar{f} = \alpha \frac{f_{SF}}{f_{SF}^*} + (1 - \alpha) \frac{f_{SC}}{f_{SC}^*} \tag{23}$$

where α is the weighting factor that balances the two objectives, it takes a value between 0 and 1; f_{SF}^* and f_{SC}^* are the best values corresponding to the objectives regarding the determinant of the FIM and the sensor clustering index, respectively, which are found from the non-dominated solutions. By minimizing this new combined function of \bar{f} and varying the weighting factor of α , trade-offs between these two objectives are made, and corresponding solutions of sensor placement can be achieved.

When the weighting factor of α takes the value of 0.0, the problem is then transformed into a single-objective optimization, which only considers the sensor clustering index. Under this condition, the optimized sensor distributions for different numbers of allowable faulty sensors are presented in Fig. 5a~c. Since only the sensor clustering is considered, all three cases produce the same sensor layout, leading to the same value of f_{SCI} , as shown in Table 2; the number of employed sensors reaches the lower bound of 3. According to Eqs. (14) and (15), a smaller number of sensors should generate a larger value in the nearest neighbor distance among the sensor locations, and result in a larger value of f_{SCI} . Consequently, when only optimizing the sensor clustering, the number of optimized sensors arrives at the lower bound. From the sensor distributions in Fig. 5a~c, we can see that the three sensors are almost uniformly distributed on the beam length direction. As mentioned before, after using the MKE-based index to narrow the design space, the first 6 nodes on the left end are removed from the initial design space, and the available sensor locations become #7~#50 of nodes. To have the maximum distance between sensors along the beam length direction, the two extreme locations in the shrunken design space, i.e., #7 and #50, are captured to locate two sensors, respectively; the third sensor is located in the middle of these two extreme locations, i.e., #29. As for the performances regarding the determinant of the FIM under sensor faults and model uncertainty, when the most severe fault case occurs, faulty sensors are denoted in Fig. 5a~c with a red color, and corresponding mean and standard deviation values are summarized in Table 2. It can be observed that under the most severe fault case, locations of faulty sensors are gradually from the right end to the left side. Since the right side of the beam always possesses quite high vibration amplitudes for the first three modes, conveying much dynamic information, failure of sensors on the right side results in serious loss of information, which explains why the most severe fault cases occur on the sensors on the right side. From the data for the mean value

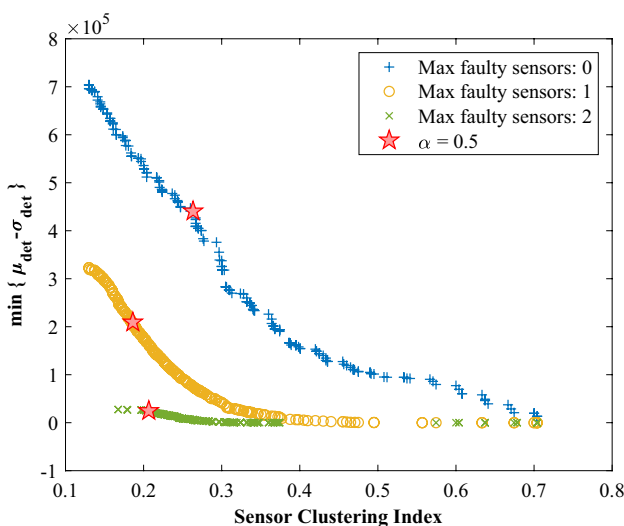


Fig. 4 Non-dominated solutions for the cantilever beam

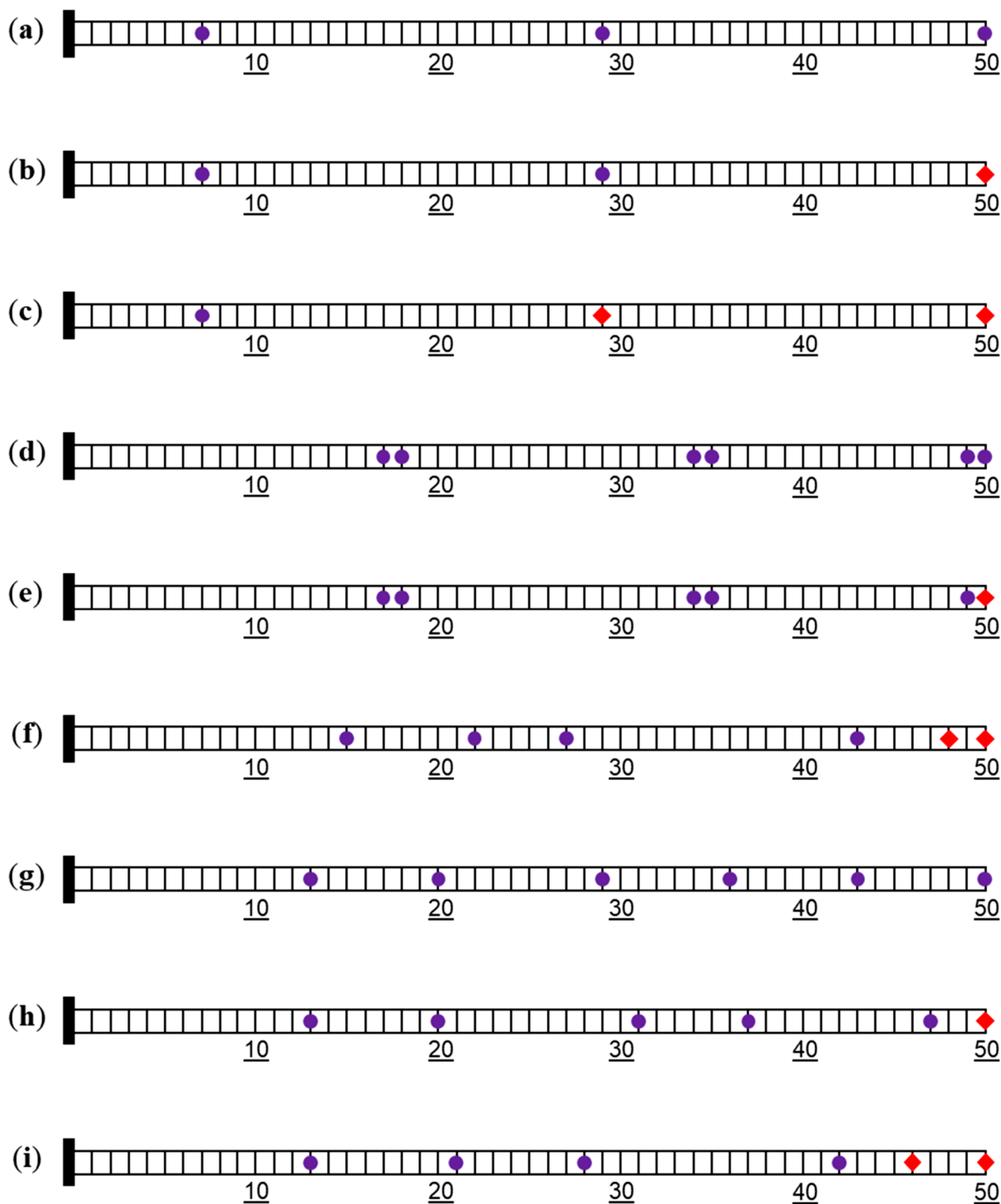


Fig. 5 Selected sensor locations by varying the weighting factor of α (red diamonds represent faulty sensors when the most severe faulty case occurs). **a** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 0. **b** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 1. **c** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 2. **d** $\alpha = 1.0$ for the case in which the number of allowable faulty sensors is 0. **e** $\alpha = 1.0$ for the case in

which the number of allowable faulty sensors is 1. **f** $\alpha = 1.0$ for the case in which the number of allowable faulty sensors is 2. **g** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 0. **h** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 1. **i** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 2

Table 2 Performances of selected sensor placements by varying the weighting factor for the cantilever beam

Weighting factor, α	Maximum number of faulty sensors	After occurrence of a sensor fault (the most severe faulty case)			Sensor clustering index, f_{SCI}
		Locations of faulty sensors	μ_{det}	σ_{det}	
0.0	0	–	1.5110×10^4	1.8263×10^3	0.7039
	1	#50	3.3574×10^{-14}	7.7372×10^{-20}	0.7039
	2	#29, #50	1.0855×10^{-33}	2.3242×10^{-34}	0.7039
1.0	0	–	8.0054×10^5	9.6698×10^4	0.1300
	1	#50	3.6710×10^5	4.4369×10^4	0.1300
	2	#48, #50	3.1235×10^4	3.7701×10^3	0.1674
0.5	0	–	5.0086×10^5	6.0460×10^4	0.2633
	1	#50	2.0896×10^5	2.5252×10^4	0.1862
	2	#46, #50	2.8685×10^4	3.4657×10^3	0.2067

of the determinant of the FIM under the most severe fault case, as listed in Table 2, it can be seen that this mean value decreases gradually as the number of allowable faulty sensors grows; this indicates that the failure of sensors contributes to the reduction of measured information.

The case for $\alpha = 1.0$ is also examined, which means that only the determinant of the FIM that considers sensor faults and model uncertainty is optimized. Corresponding optimized sensor distributions with different numbers of allowable faulty sensors are given in Fig. 5d–f. In Fig. 5d, as the number of allowable faulty sensors is 0, it only involves model uncertainty; to have the maximum determinant of the FIM and also the minimum variation, the number of used sensors gets to the upper bound, i.e., 6, and sensor clustering takes place, as the number of sensors exceeds the number of target modes. With more than one sensor in a local region, repeated information is produced in the local region, and the redundant sensors can be seen as backup sensors. If one of the sensors fails, the backup sensors will supplement the information that is lost because of a sensor fault. Figure 5e demonstrates this point, where a sensor fault is also taken into account. Because Fig. 5d and e have the same sensor layout, they have the same value in the sensor clustering index, as shown in Table 2. The most severe fault case happens when the rightmost sensor fails, as the right side of the beam conveys a large amount of dynamic information. When the maximum number of allowable faulty sensors is increased to 2, another backup sensor should be added in each local region on the basis of the sensor distribution found in Fig. 5d. However, that will make the total number of sensors reach 9, beyond the upper bound of 6. With 6 sensors deployed, the optimal sensor distribution is obtained with a goal of minimizing the effects of sensor faults and model uncertainty, as depicted in Fig. 5f. Likewise, failures of the rightmost two sensors cause the worst fault case. Additionally, the mean value of the

determinant of the FIM decreases as the number of faulty sensors increases, just like the case when $\alpha = 1.0$.

When $\alpha = 0.5$, a trade-off between the two objectives is made, and the locations of corresponding solutions are marked with a red star in Fig. 4. The sensor distributions of those solutions are presented in Fig. 5g–i. All of these three sensor distributions have 6 sensors, reaching the upper bound value to get as much information as possible so as to minimize the effects of sensor faults and model uncertainty. Similar to the results when $\alpha = 0.0$ and $\alpha = 1.0$, the most severe faulty cases occur on the rightmost sensors, as the right side of the beam possesses a large amount of dynamic information; the mean value of the determinant of the FIM declines with more faulty sensors, as more information is lost with failed sensors. When making a comparison between Fig. 5d, g, it can be seen that a more dispersed sensor distribution is achieved in Fig. 5g, which is a result of balancing the two optimization objectives. Similar phenomena can also be observed when making a comparison between Fig. 5e and h, and a comparison between Fig. 5f and i. Since these three results consider both sensor faults under uncertainty and sensor clustering, they are selected for further damage detection processes.

Besides the determinant of the FIM, the modal assurance criterion (MAC) is another commonly used criterion to assess the performance of sensor placement; it also has been widely used as an objective in many OSP problems (Ostachowicz et al. 2019). The MAC value checks the correlation between two mode vectors. A small MAC value indicates two distinguishable vectors. Hence, off-diagonal terms in the MAC matrix are always used to check the linear independence of modal shapes. The (I, J) term of the MAC matrix is defined as follows:

$$MAC_{IJ} = \frac{(\Phi_I^T \Phi_J)^2}{(\Phi_I^T \Phi_I)(\Phi_J^T \Phi_J)} \quad (24)$$

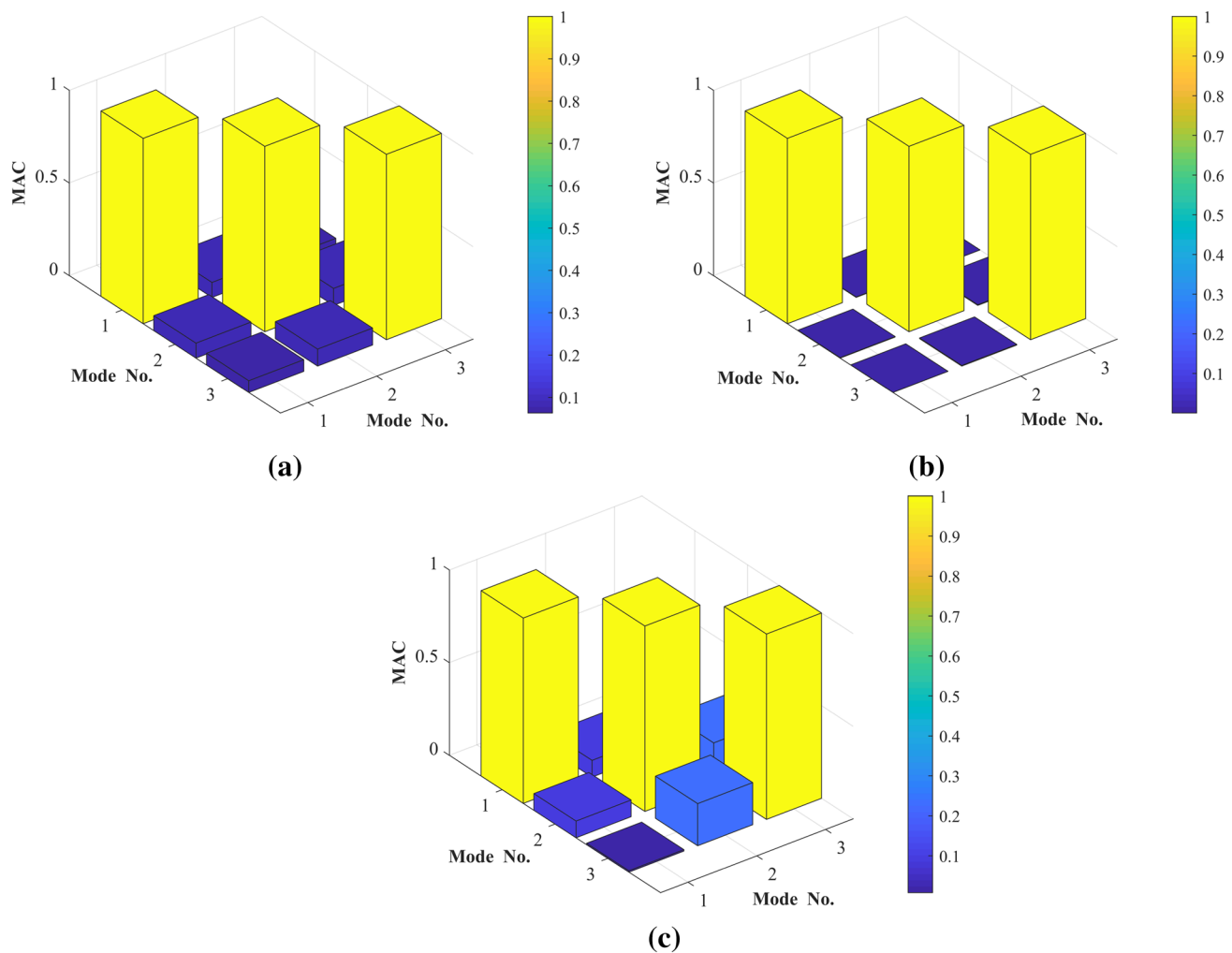


Fig. 6 MAC values for selected sensor locations with $\alpha = 0.5$ under the most severe fault case. **a** MAC for the case in which the number of allowable faulty sensors is 0. **b** MAC for the case in which the

number of allowable faulty sensors is 1. **c** MAC for the case in which the number of allowable faulty sensors is 2

Diagonal terms in the MAC matrix always keep the value of 1, while off-diagonal terms take a value between 0 and 1. Low values in off-diagonal terms represent a good sensor configuration, which indicates less correlation between the associated modal shape vectors. The obtained three sensor distributions when $\alpha = 0.5$ are further evaluated with this criterion of the MAC. When random variables take their mean values in the analysis model, MAC values corresponding to those three sensor distributions under the most severe fault case are graphically given as shown in Fig. 6. The maximum off-diagonal value of the MAC matrix is 0.00915 in Fig. 6a, which is the case in which no sensor fault is allowed; the maximum off-diagonal value is 0.0033 in Fig. 6b, which denotes the case in which one sensor fails. The difference in the MAC values results from the number of deployed sensors and also their locations. From Fig. 5g and h, we can see that most of the sensor positions for these two distributions are the same or

very close to each other; thus, we can say that the differences in the MAC values are mainly caused by the number of sensors; because Fig. 6b is obtained under the evaluation when the rightmost sensor fails, the number of sensors used in Fig. 6b is thus smaller. As stated in Ref. (Yang et al. 2020), redundant vibration information is collected when the number of sensors is larger than the number of target modes, and a linear dependence relationship could be caused by the information from the modal shape vectors measured with additional sensors; as a result, the extra sensor in Fig. 5g results in less independence between the obtained modal shape vectors, which is indicated by the larger off-diagonal MAC values in Fig. 6a than those found in Fig. 6b. When two sensors are allowed to have a fault, MAC values corresponding to the sensor distribution in Fig. 5i are as shown in Fig. 6c. For this case, the two rightmost sensors have failed, causing a loss of much relevant dynamic information; consequently, the off-diagonal MAC values are

much larger than those in Fig. 6a and b, with a maximum value of 0.2273. These analysis results further demonstrate the optimized sensors.

6.1.2 Damage detection with incomplete modal data from optimized sensors

The optimized sensors are further examined via the damage identification in the cantilever beam using the incomplete modal data from the sensors. The three sensor placements obtained when $\alpha = 0.5$ are used for the damage detection process. For this case study of a cantilever beam, only sensor placements after occurrence of sensor faults (which is also the most severe fault case) are considered. In the analysis model, stiffness of the damaged element is reduced to a certain level, i.e., $\mathbf{K} = \sum_{e=1}^{ne} (1 - \gamma_e) \mathbf{k}_e$, where \mathbf{K} and \mathbf{k}_e represent the global and element stiffness matrix, respectively; γ_e denotes the damage ratio of the e -th element, taking a continuous value between 0 and 1; and ne is the total number of finite elements. The cantilever beam was originally discretized into 50 elements; here, a simple case of aggregated damage occurring on two adjacent elements, which are Elements 25 and 26, is assumed, with damage ratios of 0.35 and 0.25, respectively. In order to simplify the damage identification process, it is also assumed that the aggregated damage type in which damages occur on a set of adjacent elements is known beforehand; thus, the optimization-based damage detection process is then to decide how many adjacent elements are damaged among the 50 elements, the locations of these damaged elements, as well as the damage ratios of these elements. Two scenarios of different measured models are tested. In the first scenario of the measured model, random variables take their mean values, and in the second measured model, the values of random variables are randomly generated following corresponding distributions, where the elastic modulus of material property, Poisson's ratio, and density are 70.025 GPa, 0.3238, and 2734.6 kg/m³, respectively. In the damage detection processes for both measured models, the random variables in the numerical model, which is used to deduce the damage condition, take their mean values. Thus, for the first scenario of the measured model, it means that the measured model is completely known, while for the second scenario, it indicates that some parameters of the model are not precisely known.

The genetic algorithm is used to solve the optimization-based damage detection problems, as formulated in Eq. (22), for both scenarios of measured models. Each problem is repeated with 5 independent runs, and the best optimal solution with the minimum objective value is treated as the detection result. Optimization results (i.e., damage detection results) are graphically presented in Fig. 7 for both measured models. For the first scenario of the measured model, we can see that all three optimized sensor placements can

accurately locate the damage, as shown in Fig. 7(a). For the second scenario of the measured model, Fig. 7(b) shows that damage locations on Elements 25 and 26 have been detected, while it also shows that additional damage occurs on Element 27 with a smaller damage ratio, which is inconsistent with the actual case. Table 3 also gives detailed quantitative damage detection results. From Fig. 7 and Table 3, we can see that when random variables in the measured model take their mean values, which means the measured model is completely known, the damage locations and damage ratios can be detected in a quite accurate way, especially when no sensors are allowed to be faulty. When a sensor fault happens, the number available sensors becomes fewer, and less information will be measured; as a consequence, the detected damage ratios show some inaccuracy; however, the results are still quite close to the actual values. When some parameters of the measured model are unknown (to represent this case, values of random variables are generated randomly), and we use the mean values of uncertain parameters in the numerical model for damage prediction, it can be seen that the detected damage ratios get less accurate as the number of faulty sensors increases. An extra element is also identified as being damaged; fortunately, the identified damage ratio of this extra element is not large. Additionally, Fig. 8 plots the convergence processes of GA in solving these inverse problems, with the best objective value in each generation. It can be seen that the best objective value tends to approach the value of zero as the generation number increases for both scenarios of measured models, showing the convergence of the algorithm.

Damage detection results given in Fig. 7 and Table 3 preliminarily demonstrate the effectiveness of optimized sensor placements. The detection result is actually determined by several aspects, like the sensor layout, data used for damage detection, detection algorithms, and related factors. Here, in the present work, we just show the detection results of one case among the many non-dominated solutions that are shown in Fig. 4. More detections are to be conducted in the future work by examining different optimized sensor configurations and using some other recent damage detection approaches, like artificial neural network (Avci et al. 2021).

6.2 A stiffened laminated composite plate—2D surface sensor layout

A composite plate with uniformly distributed T-shaped stiffeners is considered in this section, as shown in Fig. 9. The whole structure is fixed on its one short edge, with three other edges being free. Material properties of the composite layers are $E_{11} = 128\text{GPa}$, $E_{22} = 13\text{GPa}$, $G_{12} = 6.4\text{GPa}$, $\nu_{12} = 0.3$, and density $\rho = 1600\text{kg/m}^3$. Composite stacking sequences in both the panel and stiffeners are assigned as $[90/\pm 45/0_2]_s$; the basic ply thickness is $t_{\text{ply}} = 0.127\text{mm}$.

Fig. 7 Damage detection results for the cantilever beam with different measured models, using observed dynamic information after occurrence of a sensor fault(s) (under the most severe fault case). **a** Values of random variables take their mean in the measured model; **b** Values of random variables are randomly generated in the measured model

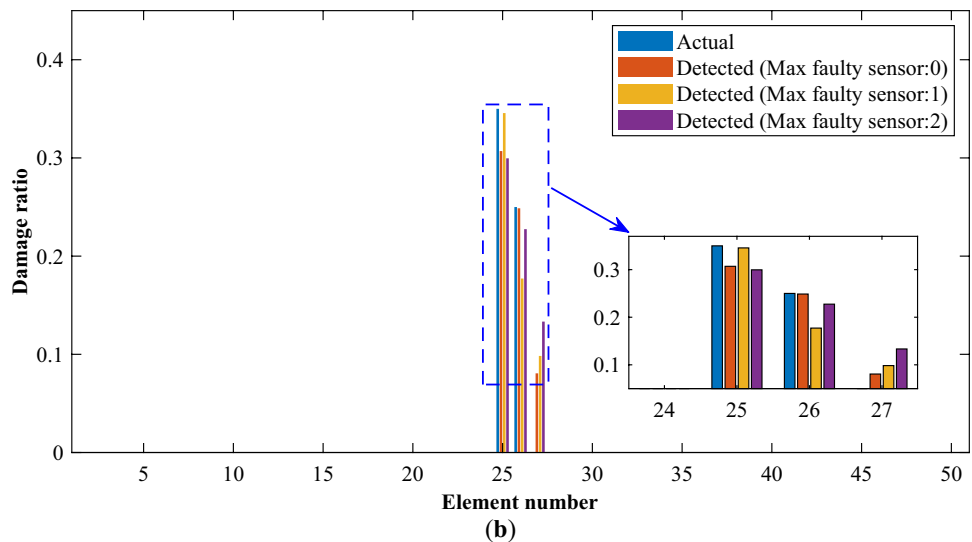
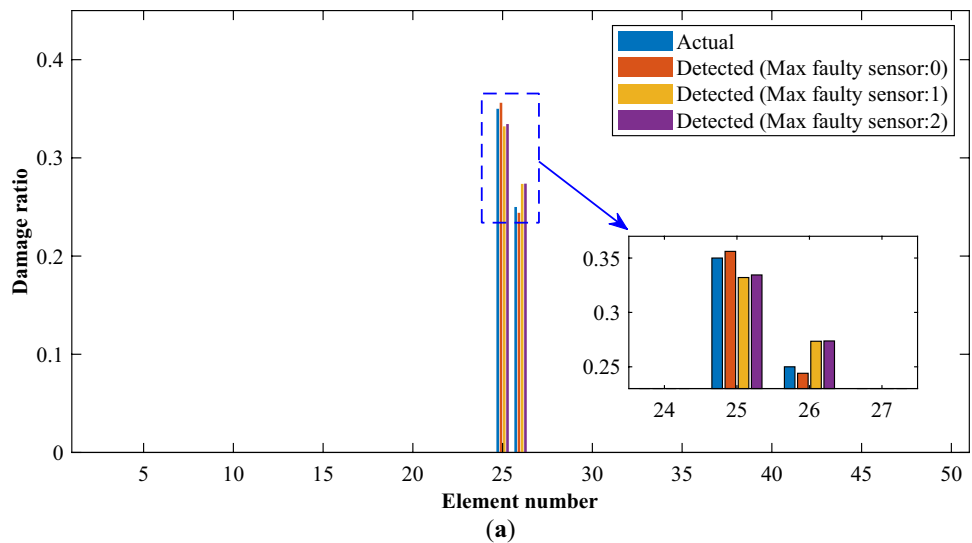


Table 3 Damage detection results for the cantilever beam under different measured models, using observed dynamic information after occurrence of a sensor fault(s) (under the most severe fault case)

Measured model type	Element no	Damage ratios			
		Actual	Detected (Max faulty sensor: 0)	Detected (Max faulty sensor: 1)	Detected (Max faulty sensor: 2)
Mean	1 ~ 24, 27 ~ 50	0	0	0	0
	25	0.35	0.3561	0.3321	0.3344
	26	0.25	0.2440	0.2735	0.2738
Random	1 ~ 24, 28 ~ 50	0	0	0	0
	25	0.35	0.3070	0.3457	0.2996
	26	0.25	0.2487	0.1772	0.2275
	27	0	0.0806	0.0984	0.1333

We focus on the out-of-plane vibrations for this stiffened laminated composite structure; sensors are assumed to measure out-of-plane translational motions. The first six modes are considered as the target modes; Fig. 10 gives

the first six modal shapes and the corresponding natural frequencies. Material properties of E_{11} , E_{22} , G_{12} , and ν_{12} are considered as random variables, with mean values of 128GPa, 13GPa, 6.4GPa, and 0.3, respectively. All of them

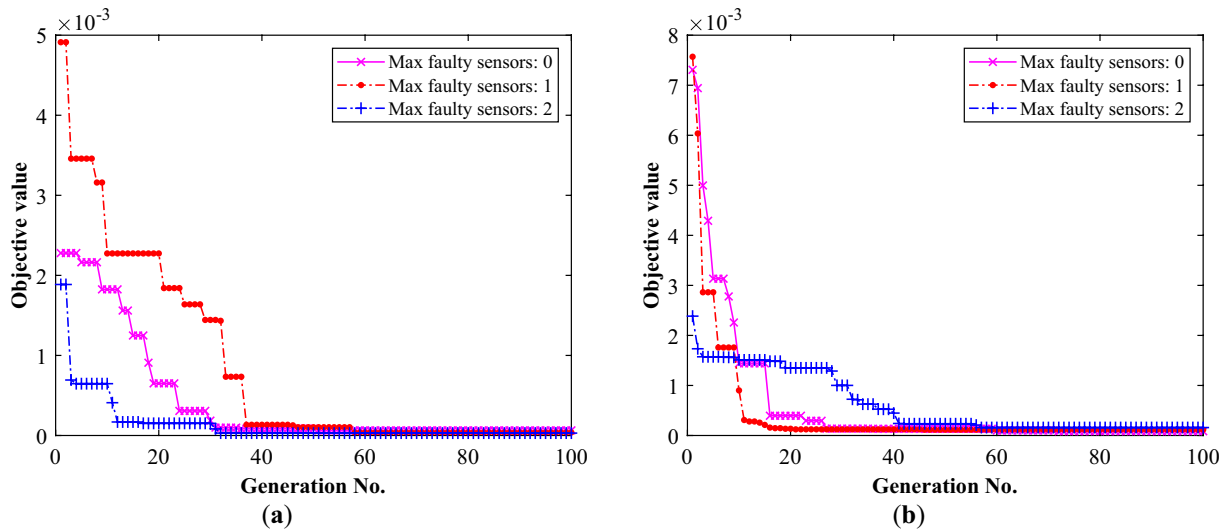


Fig. 8 Iteration history of GA for damage detection in the cantilever beam. **a** Values of random variables take their mean in the measured model; **b** Values of random variables are randomly generated in the measured model

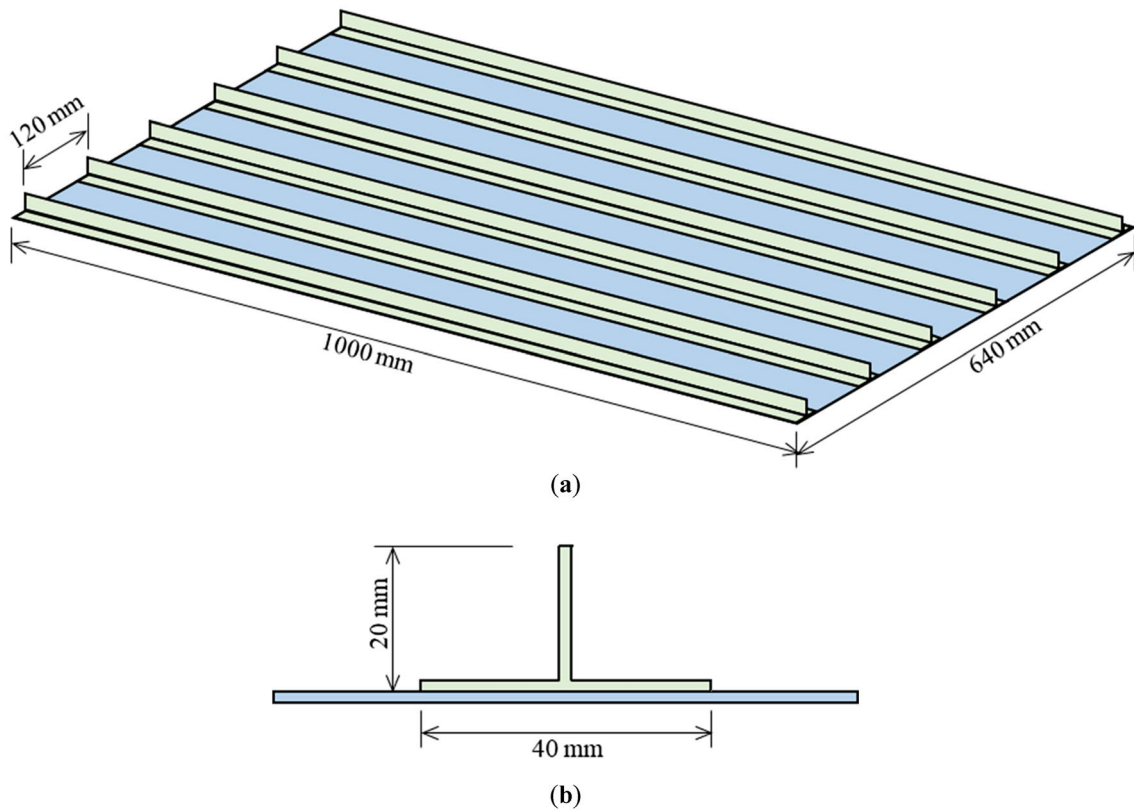


Fig. 9 Geometry of the stiffened laminated composite plate. **a** Composite stiffened plate. **b** T-shaped stiffener

follow normal distributions, with the coefficient of variation of 5%. The sensor placement problem is solved under these uncertainties.

6.2.1 Optimal sensor placement

For this sensor placement problem, as we only focus on out-of-plane vibrations, sensors are assumed to measure

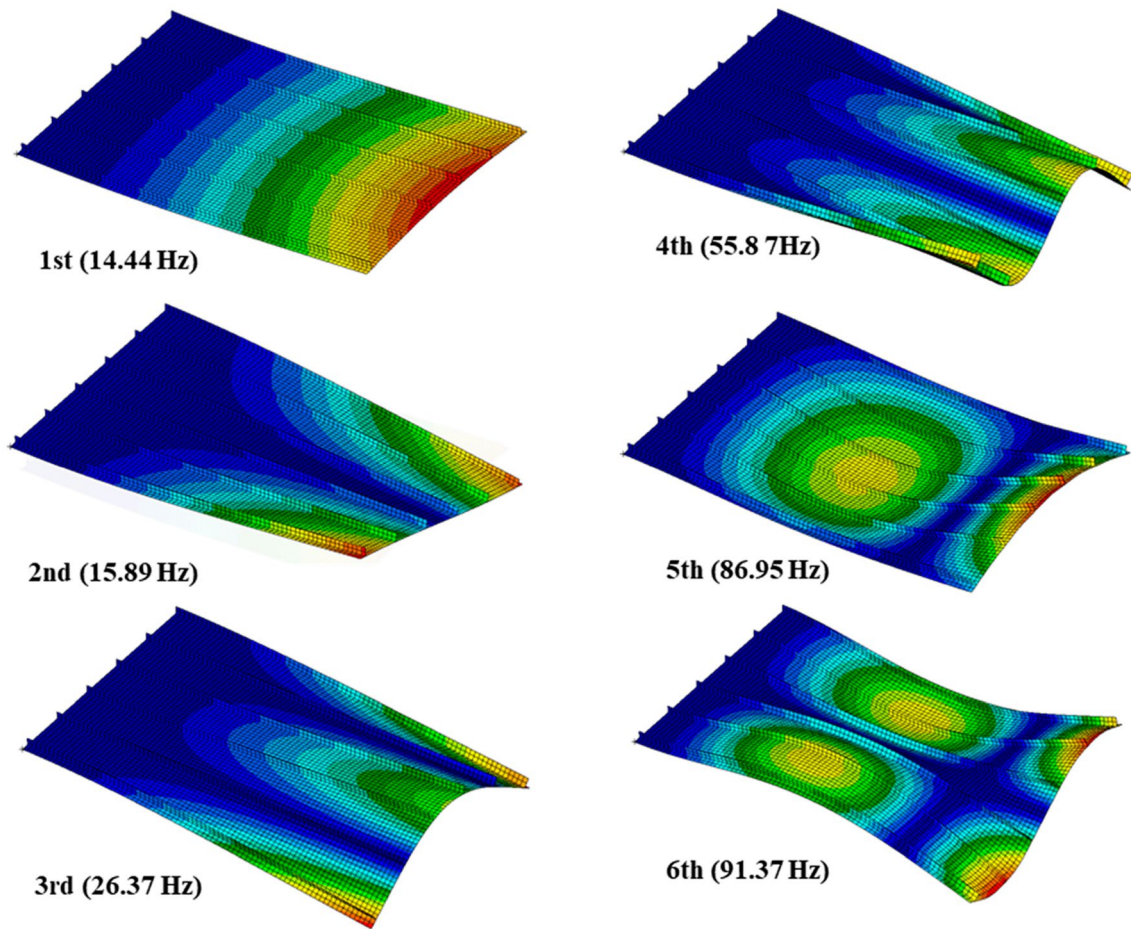


Fig. 10 The first 6 modal shapes and natural frequencies for the composite stiffened plate

out-of-plane translational motions. Available sensor positions are restricted to the surface of the composite panel. After applying the MKE-based index to narrow the design space, 3116 nodes are retained in the design space out of the original 5265 nodes, causing a 40.82% reduction in the design space. When considering sensor faults, the maximum number of allowable faulty sensors assumed to be 0 and 1, respectively. The lower and upper bounds on the number of sensors are given as 6 and 9, respectively. The LHS method is employed to generate sampling points, and 50 data points are produced to build the GPR model. By using another 200 random points, the accuracy of the constructed surrogate model is tested, and the corresponding *R*-square and RAAE values are found to be 1.0 and 0.0022, respectively, presenting a high level of accuracy. The feature size used in Eq. (14) for evaluation of the sensor clustering condition is given as 1000 mm, i.e., the length of the composite panel. All non-dominated solutions obtained for this problem are summarized in Fig. 11.

With so many non-dominated solutions, the combined function formulated in Eq. (23) is used to help select the

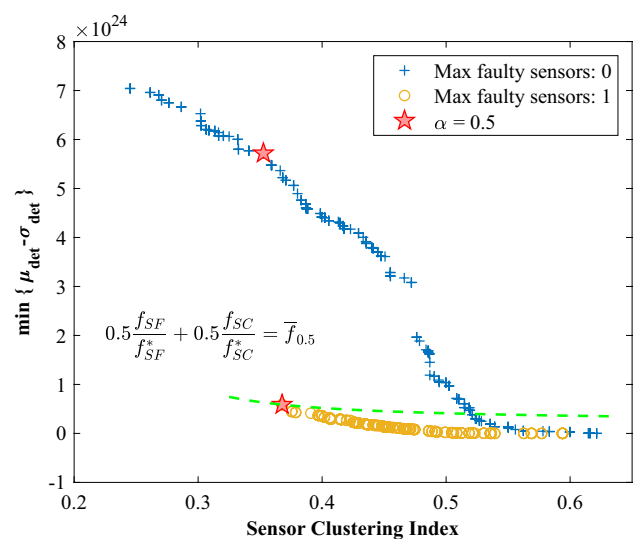


Fig. 11 Non-dominated solutions for the stiffened composite plate

sensor layout for further damage detection. By varying the weighting factor of α , trade-offs between the two objectives are made, and corresponding solutions of sensor placement can be obtained. From Fig. 11, we can see that the relation between the two considered objectives are not linear; the value ranges for these two objectives are quite different, with different orders of magnitude. Further, compared with the case of no faulty sensors, the value ranges of the two objectives are much smaller for the case of maximum one faulty sensor. As a result, even though in Eq. (23) two objectives have been normalized, it is still found that the same solution is obtained for the latter case when α is in the range of 0.3~1.0. When $\alpha = 0.5$, a contour line related to Eq. (23) is shown with a dashed green line in Fig. 11. The positions of solutions when $\alpha = 0.5$ are also indicated with red stars in Fig. 11. Corresponding to three values of weight factor, i.e., 0.0, 0.5 and 1.0, optimized sensor distributions and their performances are presented in Fig. 12 and Table 4.

When $\alpha = 0.0$, the selecting problem is equivalent to optimizing the sensor clustering index only. As a consequence, for both cases, the number of employed sensors reaches the lower bound of 6, and sensors are located as far away from each other as possible in the shrunken design space, as shown in Fig. 12a and d, leading to large values of f_{SCI} . Considering the sensor fault, it can be easily seen from Table 4 that the mean value of the determinant of the FIM significantly falls off after the most severe faulty case occurs; this is because the faulty sensor results in a smaller amount of measured information. From Table 4, we can see that before the occurrence of a sensor fault, the mean value of the determinant of the FIM for the case of one allowable faulty sensor is even larger than that of the case with no sensor fault allowed; however, this mean value declines drastically when the most severe faulty case occurs.

When $\alpha = 1.0$, this means only the determinant of the FIM that considers sensor faults and model uncertainty is involved. To get a robust design with respect to both sensor faults and model uncertainty, the number of sensors used reach the upper bound of 9 for both cases. From Fig. 12c, sensor clustering happens for the case in which the number of allowable faulty sensors is 0. When the number of allowable faulty sensors is given as 1, sensor clustering has been mitigated, as seen in Fig. 12f. Since the first six modes are to be identified, the minimum number of sensors should be 6, which is determined from the perspective of control theory and also the perspective of mathematics (Yang et al. 2020); considering sensor faults, these original 6 sensors should be backed up with another 6 sensors around their local locations, resulting in the use of 12 sensors; however, as the upper limit of the number of sensors is 9 in this study, the optimized sensor locations are obtained without

sensor clustering to collect as much information as possible, as seen in Fig. 12f, which is quite similar to the result of Fig. 5f. When $\alpha = 0.5$, even the same sensor configuration is obtained, as shown in Fig. 12e and f. For the case in which no sensor fault is allowed, a different sensor layout is achieved, as presented in Fig. 12b, which is a result of a trade-off between the performance regarding the determinant of the FIM and the sensor clustering.

The obtained sensor layouts for both cases when $\alpha = 0.5$ consider sensor faults under uncertainty and sensor clustering; these two sensor distributions are selected for further damage detection processes. The criterion of MAC is also used to evaluate the sensor deployments. When random variables take their mean values in the analysis model, MAC values corresponding to these sensor layouts before and after sensor faults occur (under the most severe fault case) are graphically shown in Fig. 13. For the case in which the maximum number of allowable faulty sensors is 0, the maximum value in the off-diagonal MAC is around 0.32. When the maximum number of allowable sensors becomes 1, before any sensor is faulty, the maximum off-diagonal MAC value is around 0.34. When the worst sensor fault case happens, the maximum value of the off-diagonal MAC becomes 0.51. This sensor fault makes the number of effective sensors fewer, and further results in a smaller amount of observed information, consequently making the linear independence of the measured modal data become weaker.

6.2.2 Delamination detection with incomplete modal data from optimized sensors

Damage from delamination between adjacent composite layers can be encountered anytime in the complicated manufacturing process or during the service period. The obtained sensor distributions when $\alpha = 0.5$, before or after a sensor fault happens, are further examined in the delamination detection processes. Here, an embedded delamination, which is located in the center of the composite panel, is assumed with a rectangular shape, as shown in Fig. 14. For this kind of delamination, the delamination length, d_L , delamination width, d_w , as well as its through-the-thickness position, i.e., the interface at which the delamination is located, are to be identified. These delamination parameters are determined by the optimization-based damage detection process through the use of incomplete dynamic data from optimized sensors. The delamination is assumed to have a length of $d_L = 100\text{mm}$ and a width of $d_w = 80\text{mm}$. The through-the-thickness position is at Interface 1. Two different models where random variables take different values are involved in each case. In one of the measured models, random variables take their mean

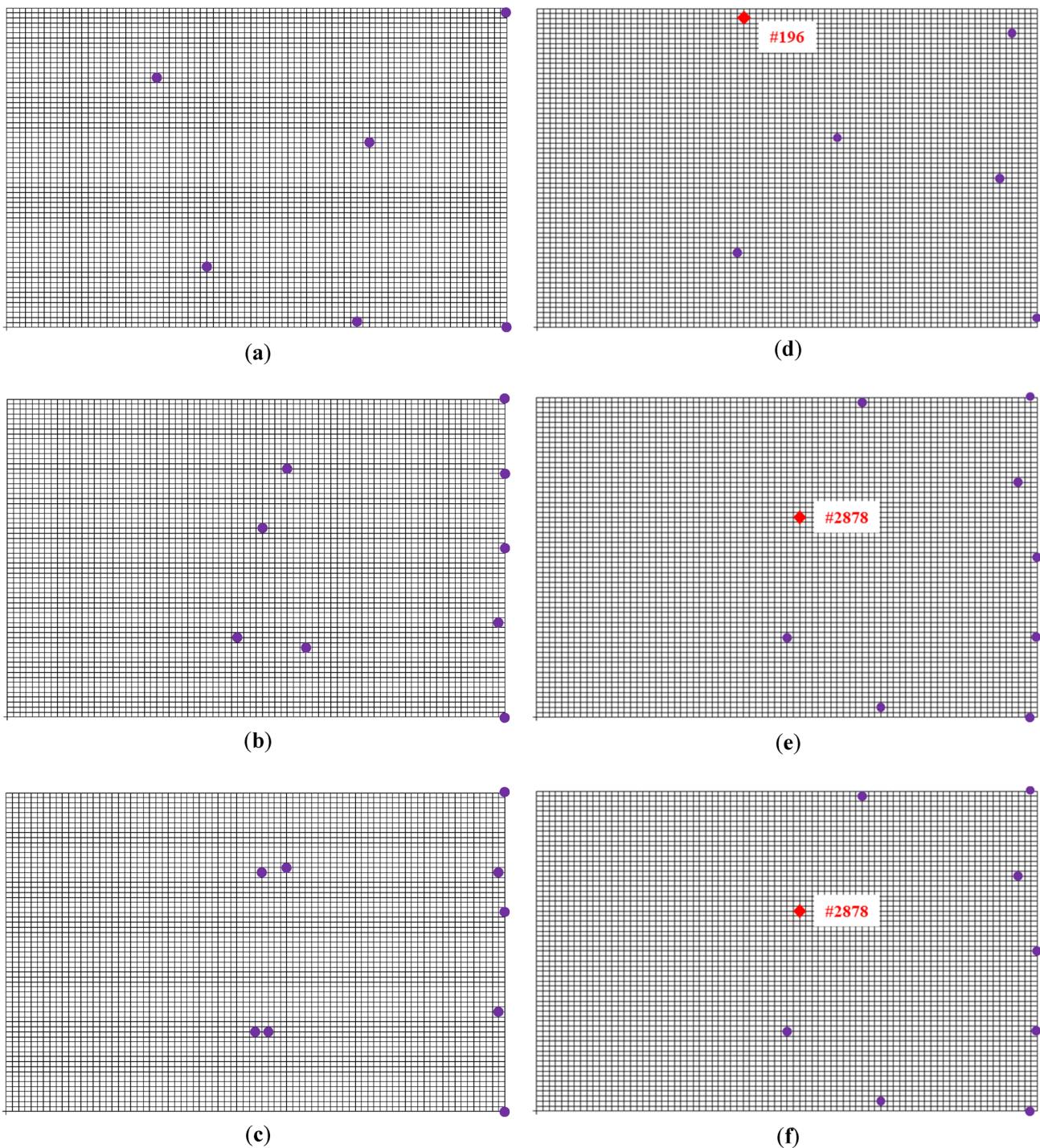


Fig. 12 Sensor locations selected by varying the weighting factor of α (red diamonds represent faulty sensors when the most severe fault case occurs). **a** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 0. **b** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 0. **c** $\alpha = 1.0$ for the case in which the num-

ber of allowable faulty sensors is 0. **d** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 1. **e** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 1. **f** $\alpha = 1.0$ for the case in which the number of allowable faulty sensors is 1

values; in the other, values for random variables are generated randomly following their distributions, which are: $E_{11} = 127.88\text{GPa}$, $E_{22} = 13.073\text{GPa}$, $G_{12} = 6.4275\text{GPa}$, $\nu_{12} = 0.2968$. As random variables in the numerical

prediction model take their mean values, the two models considered above represent the completely known model and the partly unknown model with some uncertain parameters, respectively.

Table 4 Performances of selected sensor placements calculated by varying the weighting factor for the stiffened composite plate

Weighting factor, α	Maximum number of faulty sensors	Before occurrence of a sensor fault		After occurrence of a sensor fault (the most severe faulty case)			Sensor clustering index, f_{SCI}
		μ_{det}	σ_{det}	Location of faulty sensor	μ_{det}	σ_{det}	
0.0	0	2.9988×10^{20}	1.8436×10^{18}	—	—	—	0.6214
	1	1.0590×10^{22}	9.9731×10^{18}	#196	1.3690×10^6	9.8651×10^5	0.5941
0.5	0	5.7261×10^{24}	6.3614×10^{21}	—	—	—	0.3527
	1	3.3076×10^{24}	3.6543×10^{21}	#2878	5.8793×10^{23}	6.5237×10^{20}	0.3677
1.0	0	7.0474×10^{24}	5.5350×10^{21}	—	—	—	0.2451
	1	3.3076×10^{24}	3.6543×10^{21}	#2878	5.8793×10^{23}	6.5237×10^{20}	0.3677

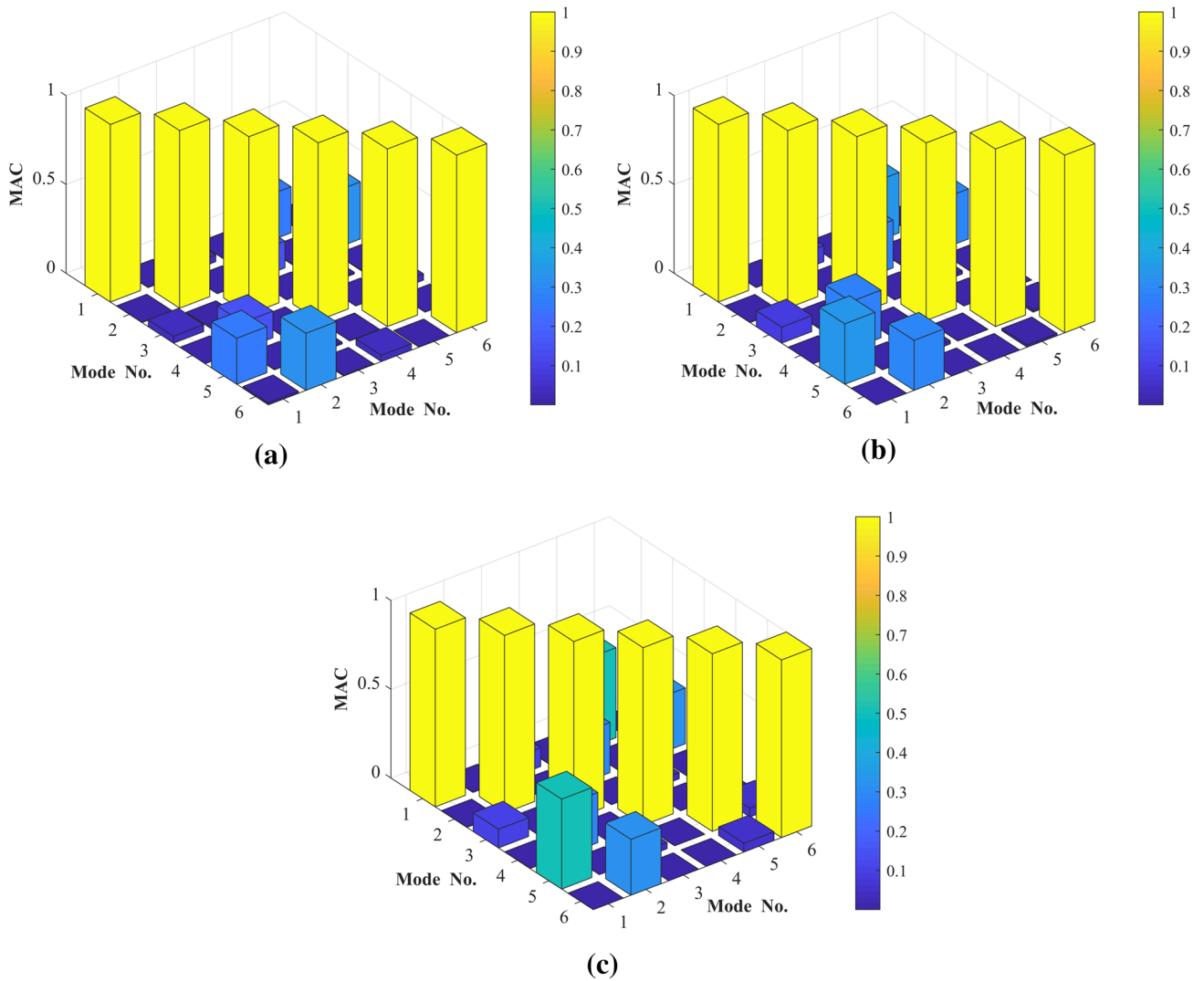


Fig. 13 MAC values for selected sensor locations with $\alpha = 0.5$. **a** MAC for the case in which the number of allowable faulty sensors is 0. **b** MAC for the case in which the number of allowable faulty

sensors is 1 (before a sensor fault occurs). **c** MAC for the case in which the number of allowable faulty sensors is 1 (after a sensor fault occurs—the most severe faulty case)

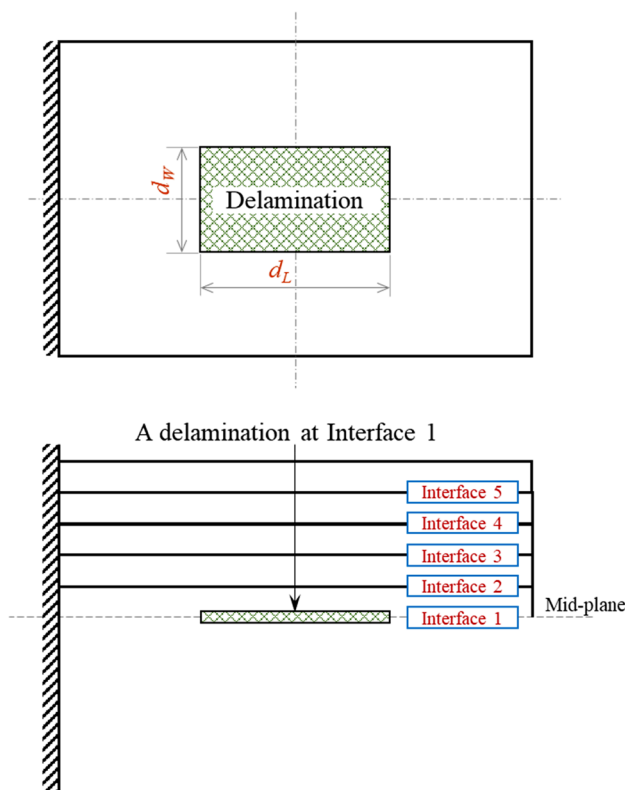


Fig. 14 An embedded delamination with rectangular shape in the composite panel

Next, the inverse problem of optimization-based delamination detection is solved with the genetic algorithm as the optimizer. Table 5 summarizes the optimization results, which are also the delamination detection results. For the scenario of the mean model, where the observed model can be completely known, it can be seen that the delamination can be identified with a considerably high level of accuracy, with all relative errors smaller than 3%, regardless of the occurrence of sensor faults. As for the scenario of the random model, where values for random variables in the measured model are randomly generated, representing the situation

of modeling error, the delamination result under the optimized sensor layout without allowance of sensor faults is quite similar to actual damage condition; however, when the maximum number of allowable faulty sensors is 1, there is a certain level of bias in predicting the delamination of length (before the occurrence of sensor fault) or width (after the occurrence of sensor fault), which is mainly caused by the modeling error. Nevertheless, it should be noted that delamination thickness-wise locations can be accurately detected, without any errors. The delamination detection results indicate that the optimized sensors are reasonable and that the solving procedure for the inverse problem is effective. The performance of damage detection depends not only on sensor distributions but also on detection approaches. More detection processes will be implemented in future work by examining different numbers of sensors and sensor locations that are from the obtained non-dominated solutions, and also by exploring some other damage detection methods. In addition, for vibration-based delamination assessment, it has been found that even though numerical predictions based on frequency changes could achieve high levels of accuracy, experimental validation showed low prediction accuracies, especially when predicting the through-the-thickness position (Zhang et al. 2018), as more uncertainties could be involved in the experimental implementation. Thus, future work will also focus on the investigation of experimental validation.

6.3 A three-story, two-bay spatial frame structure—3D spatial sensor layout

For further validation of the proposed OSP strategy in 3D structures, a three-story, two-bay steel frame structure is studied here. As shown in Fig. 15, this spatial frame structure is composed of 39 frame members and 24 nodes, with the bottom 6 nodes fixed. Each frame member has a length of 0.6 m, and a circular cross-section with a radius of 0.015 m. The mean values for the elastic modulus of material property, Poisson’s ratio, and density are 196

Table 5 Delamination detection results for the stiffened composite plate under different measured models

Measured model type	Delamination parameters	Actual	Detected (Max faulty sensor: 0) [error]	Detected (Max faulty sensor: 1) [error]	
				Before fault	After fault (The most severe case)
Mean	Length d_L , mm	100	100.95 [0.95%]	101.42 [1.42%]	99.61 [0.39%]
	Width d_w , mm	80	81.76 [2.21%]	79.98 [0.03%]	81.62 [2.03%]
	Layer interface	1	1 [0%]	1 [0%]	1 [0%]
Random	Length d_L , mm	100	100.59 [0.59%]	126.98 [26.98%]	100.82 [0.82%]
	Width d_w , mm	80	83.09 [3.86%]	82.36 [2.95%]	101.48 [26.85%]
	Layer interface	1	1 [0%]	1 [0%]	1 [0%]

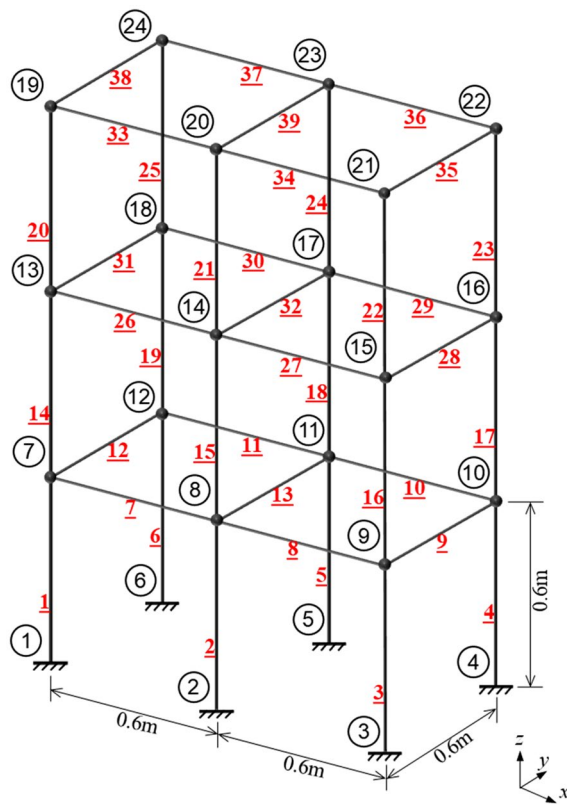


Fig. 15 Geometry and boundary condition of a three-story, two-bay frame structure

GPa, 0.26, and 7880 kg/m^3 , respectively. All of these parameters follow normal distributions, with the coefficient of variation of 5%. For modal analysis, the finite element model is established with beam elements and each frame member is discretized into 5 beam elements. The first 6 modes are considered as the target modes. From Fig. 16, we can see that all the target modes are bending modes along the x - or y - direction, or torsion modes along the z - direction. Accordingly, the 36 translational DOFs (along the x - and y -directions) at Nodes #7 - #24 are considered as candidate sensor locations. Since we use uniaxial sensors, one sensor can only measure dynamic information in a specific direction; when two sensors are placed in the same node but instrumented in different directions, this condition is not considered to be sensor clustering, as the measured information is coming from different directions. Consequently, the sensor clustering condition is evaluated in different directions separately, and then the evaluation values from different directions are summed to obtain the final value regarding the sensor clustering index.

6.3.1 Optimal sensor placement

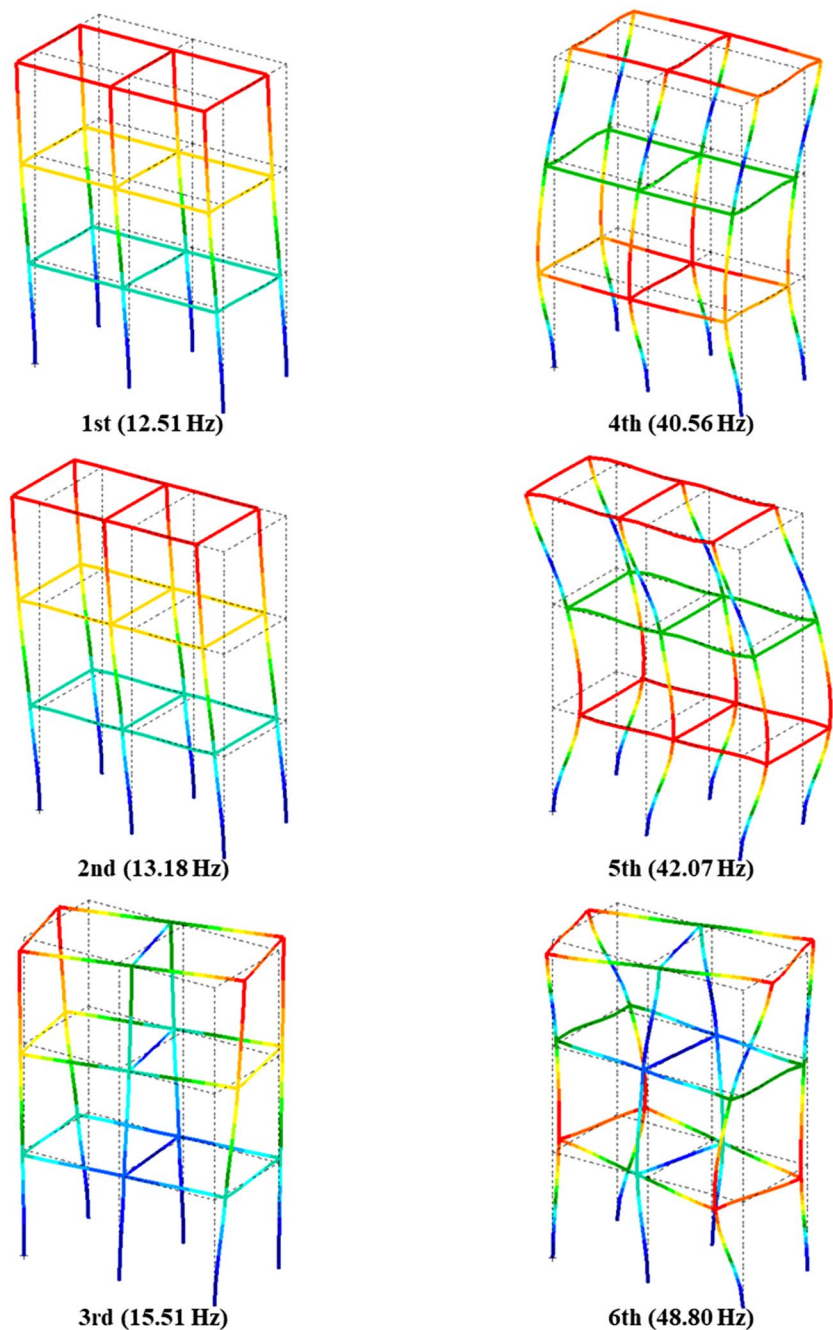
For this case study, two cases are considered; in each case, the maximum number of allowable faulty sensors is assumed to be 0 and 1, respectively. The lower and upper bound on the number of sensors is given as 6 and 10, respectively. Based on the LHS method, 50 data points are generated to build the GPR model. The accuracy of the constructed surrogate model is then tested with another 200 random points and the corresponding R-square and RAAE values are 1.0 and 0.0035, respectively, showing a high level of accuracy. The feature size used in Eq. (14) for evaluation of the sensor clustering condition is specified as 1.8 m, which is the height of the spatial frame structure. All non-dominated solutions obtained for this problem are summarized in Fig. 17.

The defined combined function in Eq. (23) assists in the selection of sensor layout for the next-step damage detection process. Varying the weighting factor of α , balances between the two objectives are made, and corresponding solutions are obtained. Three different values of α , which are 0.0, 0.5, and 1.0, are considered. Figure 17 shows the position of the solution when $\alpha=0.5$ with a red star. Corresponding to these three weight factor values, optimized sensor distributions and their performances are given in Fig. 18 and Table 6.

When the weighting factor is assigned to be 0, only the sensor clustering index is considered. From Table 6, we can see that the same value of the sensor clustering index is obtained for the cases in which the maximum number of allowable faulty sensors is 0 and 1, even though the sensor distributions for these two cases are not totally the same, as shown in Fig. 18a and d. As aforementioned, the sensor clustering index of f_{SCI} is computed by summing the corresponding index values in the different directions. From Fig. 18a and d, it can be easily seen that $(f_{\text{SCI}})_y$ (f_{SCI} in the y -direction) in Fig. 18a is equal to the $(f_{\text{SCI}})_x$ (f_{SCI} in the x -direction) in Fig. 18d, as they have the same sensor positions regarding these two directions; with structural symmetry, $(f_{\text{SCI}})_x$ in Fig. 18a also equals $(f_{\text{SCI}})_y$ in Fig. 18d, as the two concerned sensors in the related directions are located in the diagonal vertices of the structure. Consequently, the summed value of f_{SCI} is the same for both cases. For the case in which the maximum number of allowable faulty sensors is 1, the most severe faulty case occurs when the sensor instrumented in the x -direction of Node #19 fails. Because of this sensor fault, the mean value of the determinant of the FIM is dramatically reduced from 3.7376×10^{-10} to 1.4188×10^{-27} , as shown in Table 6, showing the severe impact of the sensor fault.

Gradually increasing the value of weighting factor, α , to 0.5 and 1.0, which corresponds to a trade-off between the two objectives and only optimizing the performance regarding the determinant of the FIM, respectively, relevant

Fig. 16 The first 6 modal shapes and natural frequencies for the spatial frame structure



solutions are achieved. When the maximum number of allowable faulty sensors is 0, i.e., no sensor fault is allowed, sensors in the y -direction are distributed in the upper, middle and lower layers of nodes in the structure, while sensors in the x -direction are deployed only in the upper and lower layers of nodes. From the first 6 modal shapes presented in Fig. 16, relatively large vibration amplitudes representing main dynamic information are mainly exhibited in the upper and lower layers of nodes. This also results in sensor distributions only in the lower and upper layers of nodes for the case in which the maximum number of allowable faulty

sensors is 1, for both $\alpha = 0.5$ and $\alpha = 1.0$. Considering the torsion modes, which are the 3rd and 6th modes, as shown in Fig. 16, vibration amplitudes in the y -direction are larger than the x -direction, while for the bending modes, there is no such difference in the two directions. As a result, the number of employed sensors in the y -direction is larger than that in the x -direction, as can be seen in Fig. 18b, c, e and f; this kind of dynamic characteristic also leads to the occurrence of the most severe fault case in the y -direction, as shown in Fig. 18e, f, as well as Table 6. When considering a sensor fault, compared with the performance before the occurrence

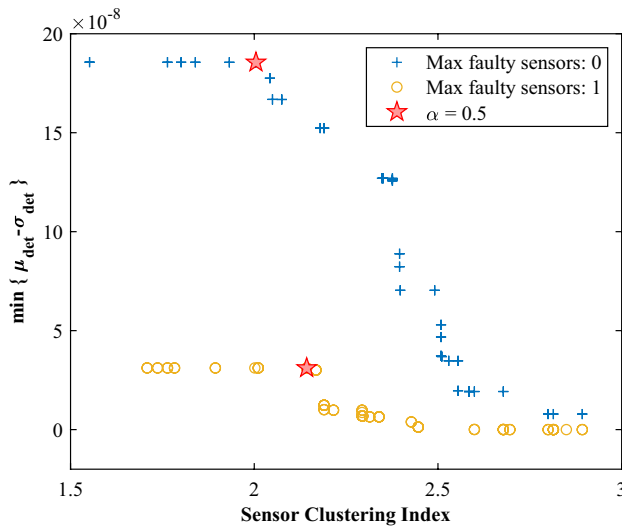


Fig. 17 Non-dominated solutions for the spatial frame structure

of sensor fault, the mean value of the determinant of the FIM goes down significantly after a sensor fault occurs. For example, when $\alpha = 0.5$, the mean of the determinant of the FIM goes from 1.2179×10^{-7} before a sensor fault happens to 3.2751×10^{-8} after the occurrence of a sensor fault (under the most severe fault case), meaning that the sensor fault causes some information loss.

Since the solutions corresponding to $\alpha = 0.5$ allow for both design objectives, these optimized sensor distributions are then used for the damage detection processes, and they are also evaluated with the criterion of the MAC. When random variables take their mean values in the analysis model, MAC values corresponding to the sensor layouts are graphically shown in Fig. 19. For the case in which the maximum number of allowable faulty sensors is 0, the maximum value in the off-diagonal MAC is around 0.19. When the maximum number of allowable sensors becomes 1, before the sensor fault occurs, the maximum off-diagonal MAC value is about 0.20 and this maximum value turns to 0.23 after a sensor fault happens (for the most severe fault case), showing that information loss because of the sensor fault leads to a decrease in the linear independency of the measured modal shapes.

6.3.2 Damage detection with incomplete modal data from optimized sensors

Based on the incomplete modal data from the sensors, damage detection processes are implemented in the spatial frame structure to verify the optimized solutions of sensor configurations when $\alpha = 0.5$. Similarly, the simulation model that is used in the beam structure is also applied here to simulate the damage, i.e., by reducing the

stiffnesses of the damaged frame members to a certain level. Here, it is assumed that the frame member of 14 is encountered with a damage ratio of 0.3. The optimization-based damage detection process is to decide the number of damaged frame members, the locations of these damaged members, and the damage ratios on these members. Likewise, two scenarios of different measured models are tested. In the first scenario of the measured model, random variables take their mean values; in the second model, the values of random variables are randomly generated following their corresponding distributions, where the elastic modulus of material property, Poisson's ratio, and density are 196.56 GPa, 0.2682, and 8061.0 kg/m^3 , respectively. In the damage detection processes for both scenarios, the random variables in the numerical model to deduce the damage condition take their mean values. In this way, for the first scenario of the measured model, the measured model is completely known; as for the second scenario, this means that some parameters of the model are not exactly known with certain levels of error.

The optimization-based damage detection problems are then solved with GA for both measured model scenarios. Damage detection results are graphically presented in Fig. 20 for both measured models and the detailed damage detection results are quantitatively tabulated in Table 7. For the first measured model scenario, we can see that not only the frame member of 14 is identified as damaged, but also several other members are deemed to be damaged. Even so, the identified damage ratios on frame member 14 are quite close to the actual value, while some other identified damaged members have quite small values of damage ratios; the largest one is 0.0126, which occurs just when the most severe fault case happens. As for the positions of the identified damaged members with small damage ratios, it can be found that they are adjacent to or quite near the frame member of 14 from Fig. 15 and Table 7. Further, based on the data of the damage ratios, we can see that the highest detection accuracy is reached when the maximum number of allowable faulty sensors is 0, i.e., no sensor fault is allowed. The detection accuracy goes down when the maximum number of allowable faulty sensors is varied from 0 to 1. Under the condition in which the maximum number of allowable faulty sensors is 1, the detection accuracy becomes worse after the most severe fault case occurs, as compared with the result before the occurrence of a sensor fault. A similar phenomenon can also be observed for the second scenario of the measured model regarding the detection accuracy of the damage ratios. However, the number of damaged frame members and the damage location have been accurately identified without any errors. These damage detection results have preliminarily demonstrated the effectiveness of optimized sensors. More detections are to be conducted in the future work by using some other recent advanced damage detection approaches.

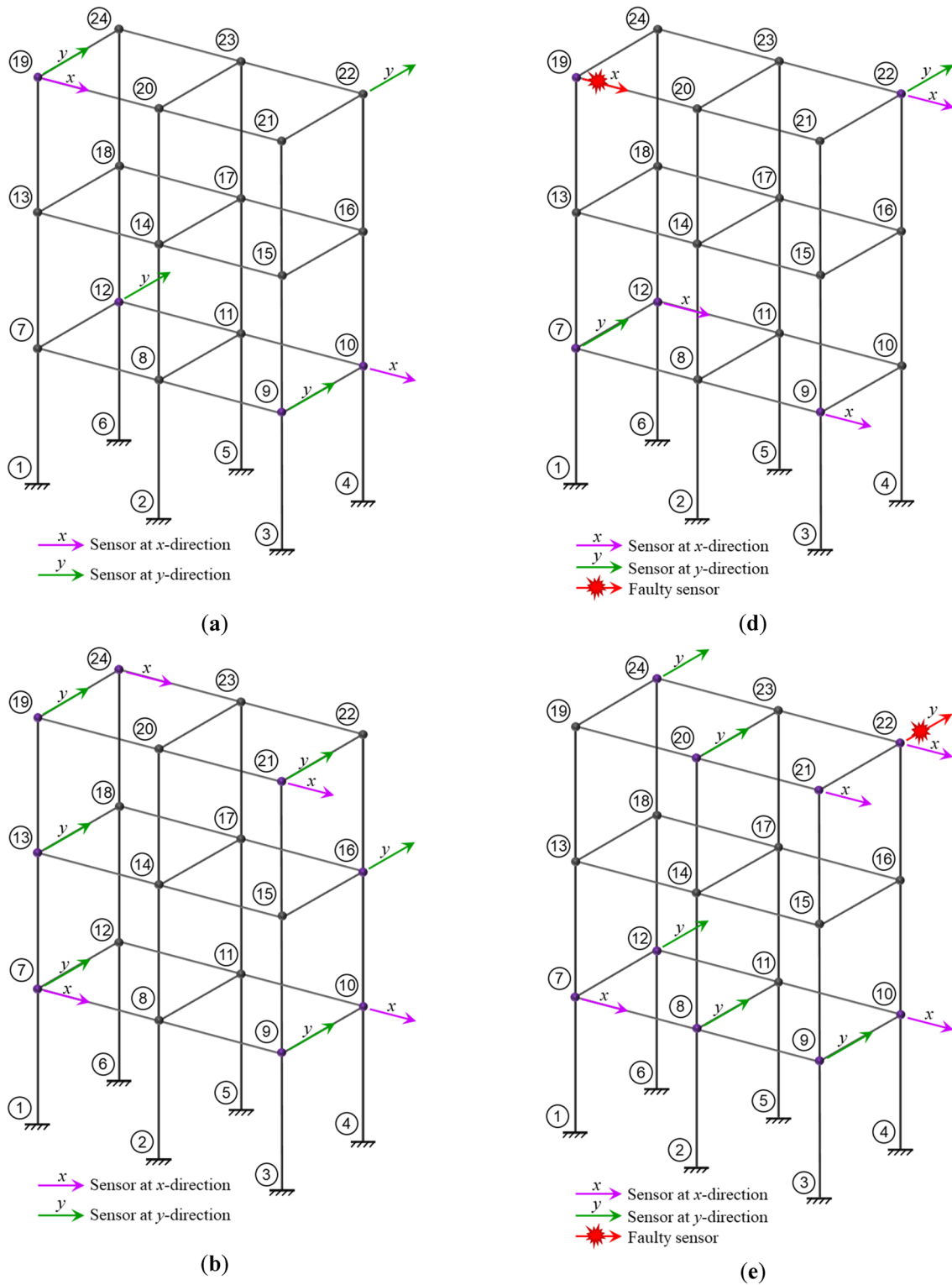


Fig. 18 Sensor locations selected by varying the weighting factor of α (red arrows with explosion shapes represent faulty sensors when the most severe fault case occurs). **a** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 0. **b** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 0. **c** $\alpha = 1.0$ for the

case in which the number of allowable faulty sensors is 0. **d** $\alpha = 0.0$ for the case in which the number of allowable faulty sensors is 1. **e** $\alpha = 0.5$ for the case in which the number of allowable faulty sensors is 0. **f** $\alpha = 1.0$ for the case in which the number of allowable faulty sensors is 1

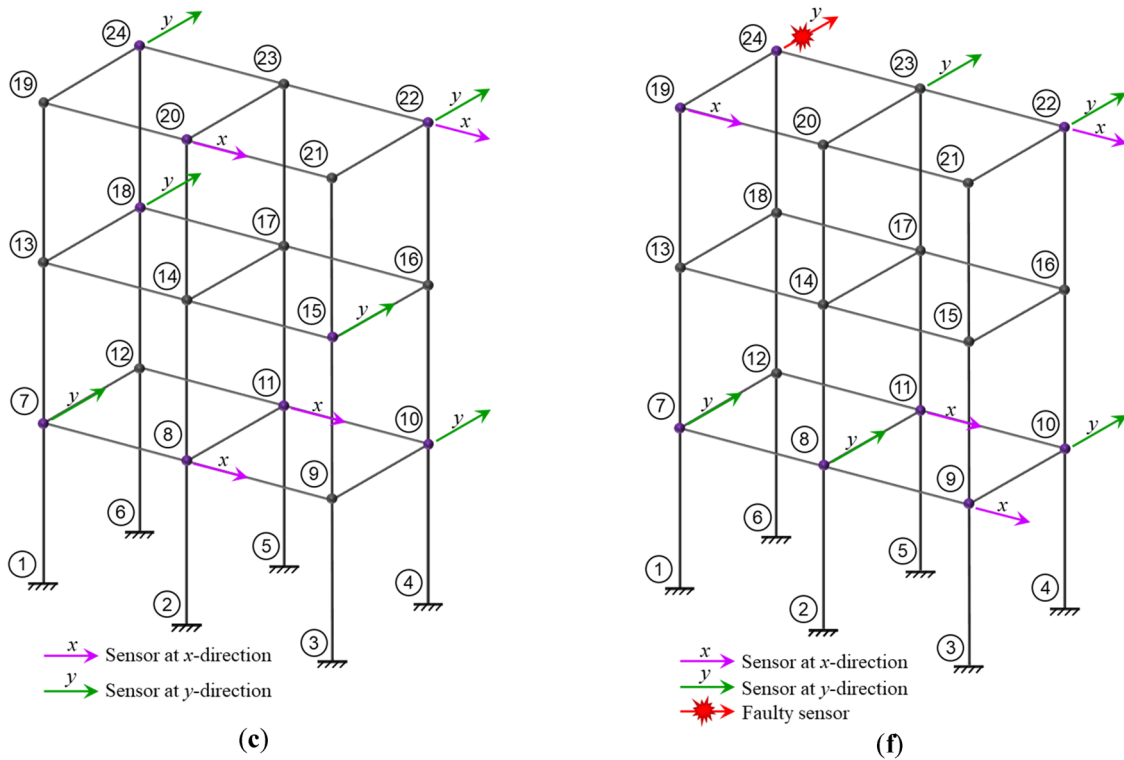


Fig. 18 (continued)

Table 6 Performances of selected sensor placements calculated by varying the weighting factor for the spatial frame structure

Weighting factor, α	Maximum number of faulty sensors	Before occurrence of a sensor fault		After occurrence of a sensor fault (the most severe faulty case)	Sensor clustering index, f_{SCI}
		μ_{det}	σ_{det}		
0.0	0	3.7388×10^{-10}	1.8304×10^{-11}	–	2.8928
	1	3.7376×10^{-10}	1.8333×10^{-11}	#19(x)	2.8928
0.5	0	1.9511×10^{-7}	9.5319×10^{-9}	–	2.0046
	1	1.2179×10^{-7}	5.9495×10^{-9}	#22(y)	2.1427
1.0	0	1.9511×10^{-7}	9.5290×10^{-9}	–	1.5518
	1	1.2182×10^{-7}	5.9624×10^{-9}	#24(y)	1.7084

In this work, lower and upper bounds are assigned on the number of sensors, which means available and limited budgets for sensor placement and data processing. Non-dominated optimal solutions are then achieved within this range. In order to select an appropriate solution from those non-dominated solutions for further damage detection processes, a combined function is defined and suggested by introducing a weighting factor and combining both objectives. When the weighting factor takes the value of 0.5, a solution is chosen and then validated for damage detection processed. For this selected solution, the number of sensors often reaches the corresponding upper bound, which can be

seen from Figs. 5, 12 and 18, meaning sufficient sensors. As a result, even if sensor fault occurs, the damage detection results seem still good because of redundant sensors. Meanwhile, it can also be seen that the damage detection performance goes worse in the case of occurrence of sensor faults, especially when the number of allowable faulty sensors goes up. In addition, since only one solution when weighting factor takes the value of 0.5 is selected in damage detections, more solutions decided by assigning the weighting factor with different values are to be examined. Actually, when larger weights are assigned on the aspect of avoiding sensor clustering, for example, when the weighting factor

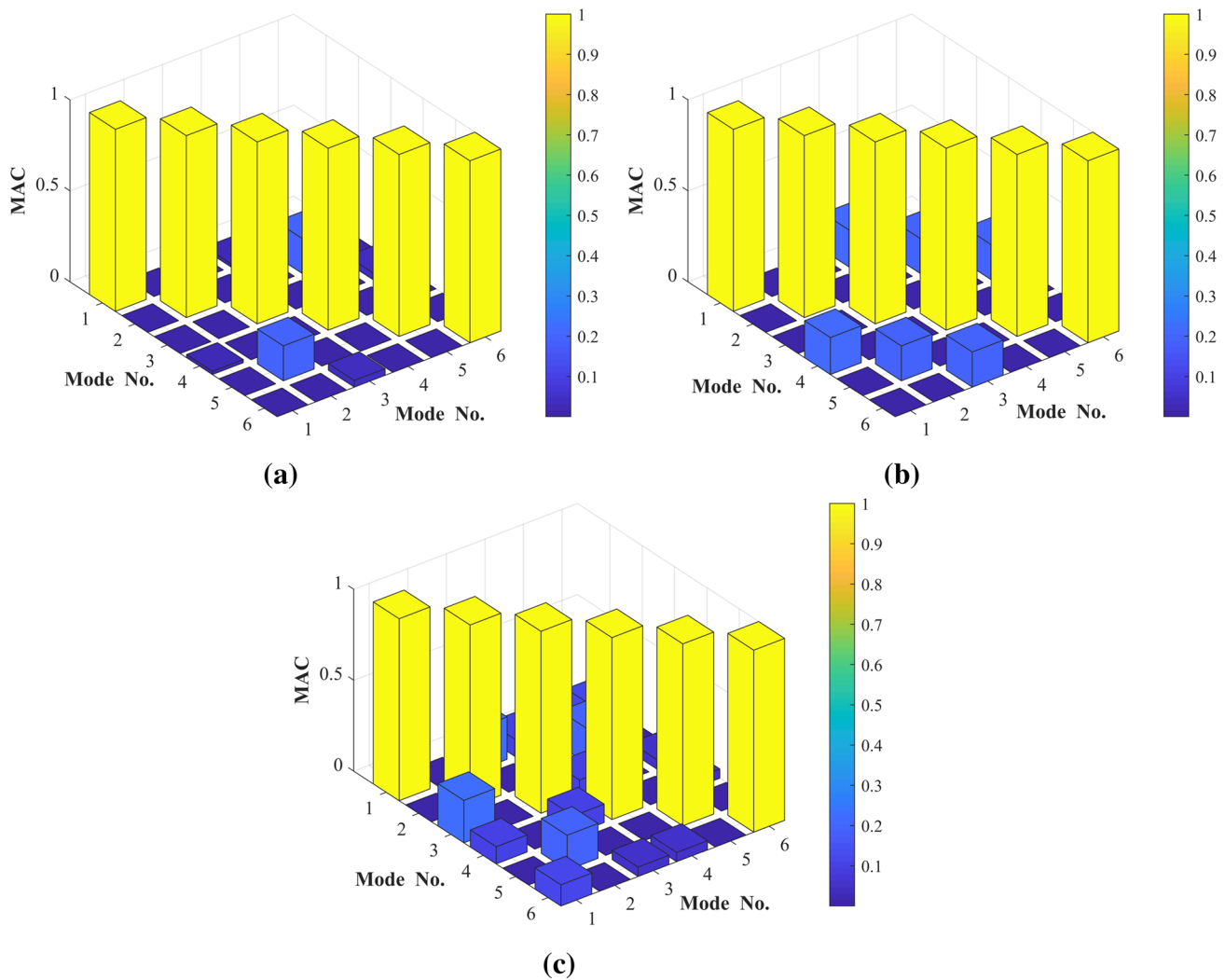


Fig. 19 MAC values for selected sensor locations with $\alpha=0.5$. **a** MAC for the case in which the number of allowable faulty sensors is 0. **b** MAC for the case in which the number of allowable faulty

sensors is 1 (before a sensor fault occurs). **c** MAC for the case in which the number of allowable faulty sensors is 1 (after a sensor fault occurs—the most severe faulty case)

of α equals 0.0, fewer sensors are obtained, which can also be observed from Figs. 5, 12 and 18, and the redundancy of sensors becomes less. Furthermore, mean values of the determinant of Fisher information matrix for those solutions, as presented in Tables 2, 4 and 6, become smaller along with smaller values of α , which means less observed information and less redundancy. So, it could be expected that when these solutions are used for damage detection, the detection results could be worse than those solutions when α equals 0.5, which will be examined in our future work.

If the number of available sensors becomes even smaller, for example, the number of available sensors is smaller than the number of observed target modes, the phenomenon of sensor clustering will not occur. Accordingly, when applying the proposed framework for optimal

sensor placement considering uncertainty and sensor faults, the design objective regarding the sensor clustering can be removed, and only the objective related to the determinant of Fisher information matrix remains. In this way, the proposed framework is still applicable after a minor modification, regardless of the complexity of the observed structure and system. With smaller number of sensors, it is true that the damage detection results could possibly be not that accurate and reliable. However, that is mainly resulted from the limited sensors, and optimization of sensor placements is just trying to produce the maximum benefits of the limited resources.

Fig. 20 Damage detection results for the spatial frame structure with different measured models. **a** Values of random variables take their mean in the measured model; **b** Values of random variables are randomly generated in the measured model

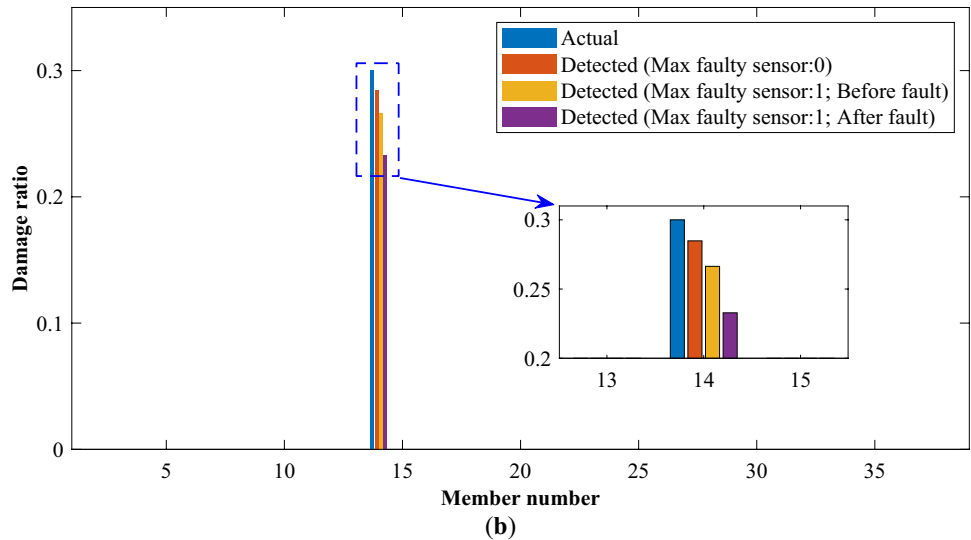
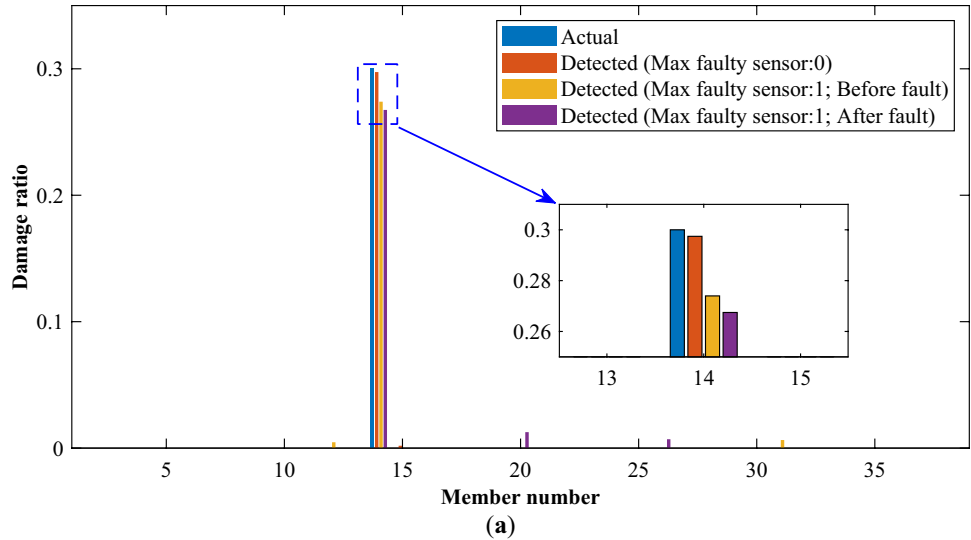


Table 7 Damage detection results for the spatial frame structure under different measured models

Measured model type	Member No	Damage ratios			
		Actual	Detected (Max faulty sensor: 0)	Detected (Max faulty sensor: 1)	
				Before fault	After fault (The most severe case)
Mean	1 ~ 11, 13, 16 ~ 19, 21 ~ 25, 27 ~ 30, 32 ~ 39	0	0	0	0
	12	0	0	0.0046	0
	14	0.3	0.2974	0.2740	0.2675
	15	0	0.0019	0	0
	20	0	0	0	0.0126
	26	0	0	0	0.0070
	31	0	0	0.0064	0
Random	1 ~ 13, 15 ~ 39	0	0	0	0
	14	0.3	0.2848	0.2664	0.2328

7 Conclusions

Under the consideration of sensor faults and model uncertainty, as well as sensor clustering, this study investigates optimal sensor placement for vibration-based damage detection. A new objective is formulated, considering sensor faults and model uncertainty, and a novel sensor clustering index is proposed to evaluate the sensor clustering condition. Based on these two objectives, the design problem is formulated to seek the optimal number of sensors and their locations. As any available place in the structure can represent a potential candidate sensor position, which causes a large design space, a modal kinetic energy-based index is developed to narrow the design space to enhance the optimization search capability. To solve this multi-objective optimization problem under a shrunken design space, the non-dominated genetic algorithm II is adopted as the optimizer and MCS is used to evaluate the design objective involving uncertainty. To reduce the computation cost of the MCS, a GPR model is employed to approximate the real performance evaluations. For the validation of optimized sensors, an optimization-based damage detection process is conducted based on the incomplete dynamic data from the sensors. Three case studies, including a cantilever beam, a stiffened laminated composite plate, and a spatial frame structure are presented to verify the developed optimization framework and the optimized sensors. Some conclusions are summarized as follows:

- (1) Based on the EFI method, a new objective is formulated to account for sensor faults and model uncertainty. Under uncertainty, a compound function is constructed, combining the mean and standard deviation of the determinant of the FIM. The minimum value of this compound function is sought under all possible sensor fault cases. After maximizing this minimum value, a sensor configuration that can be robust to both sensor faults and model uncertainty is obtained.
- (2) To overcome the limitations and disadvantages of previous sensor clustering indices, a novel evaluation index for sensor clustering is proposed based on the mean and standard deviation of the nearest neighbor distance, as well as the density of the sensor distribution. Simple examples and case studies demonstrate that the proposed index can correctly identify sensor clustering conditions and can also be universally applicable for any type of structures, i.e., 1D, 2D and 3D type of structures.
- (3) Through the use of non-dominated genetic algorithm II as the optimizer and a GPR-based surrogate model, non-dominated solutions for sensor distributions are achieved in an efficient way.

- (4) Optimized sensor configurations are obtained for the three case studies that correspond to 1D, 2D and 3D sensor layout types. As the maximum number of allowable faulty sensors increases, the linear independency of the measured modal data can be decreased; when a sensor fault occurs, the number of effective sensors decreases and information loss happens, leading to an increase in the linear dependency of the observed modal shapes, as compared to the case before the occurrence of the sensor fault. Successful optimization-based damage processes verify the optimized sensor distributions, and also show that because of the information loss resulting from faulty sensors, the damage detection accuracy may go down after the occurrence of a sensor fault.
- (5) The proposed methodology, which combines a robust optimal sensor placement strategy and a damage detection approach, can be considered a promising tool in practical SHM applications. More recent advanced damage identification methods, such as machine-learning-based methods, will be explored in the next-step work. Moreover, only sensor fault of complete failure is examined in the present work, and more types of sensor faults, like signal bias and drifting, will be considered in our future work.

Acknowledgements This work was supported by the Brain Korea 21 FOUR Project in 2020.

Authors contribution H. An wrote the manuscript using the computational results. BD. Youn and HS. Kim (heungsoo@dgu.edu) supervised the project, and they are co-corresponding authors of this paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Replication of results Data and materials for replication of the case studies will be available from the corresponding author upon request.

References

- An H, Chen S, Huang H (2018) Multi-objective optimization of a composite stiffened panel for hybrid design of stiffener layout and laminate stacking sequence. *Struct Multidisc Optim* 57(4):1411–1426
- An H, Youn BD, Kim HS (2021) Reliability-based design optimization of laminated composite structures under delamination and material property uncertainties. *Int J Mech Sci* 205:106561
- An H, Youn BD, Kim HS (2022) A robust design framework to determine the optimal number and placement of sensors under uncertainty for vibration-based damage detection of composite structures. *Compos Struct* 279:114863
- Avci O, Abdeljaber O, Kiranyaz S, Hussein M, Gabbouj M, Inman DJ (2021) A review of vibration-based damage detection in civil

- structures: From traditional methods to Machine Learning and Deep Learning applications. *Mech Syst Signal Proc* 147:107077
- Balaban E, Bansal P, Stoelting P, Saxena A, Goebel KF, Curran S. A diagnostic approach for electro-mechanical actuators in aerospace systems. In: 2009 IEEE Aerospace conference, BigSky, MT; 2009, pp. 1–13. doi: <https://doi.org/10.1109/AERO.2009.4839661>
- Barman SK, Mishra M, Maiti DK, Maity D (2021) Vibration-based damage detection of structures employing Bayesian data fusion coupled with TLBO optimization algorithm. *Struct Multidisc Optim*. 64: 2243–2266. <https://doi.org/10.1007/s00158-021-02980-6>
- Bigoni C, Zhang Z, Hesthaven JS (2020) Systematic sensor placement for structural anomaly detection in the absence of damaged states. *Comput Methods Appl Mech Engrg* 371:113315
- Castro-Triguero R, Murugan S, Gallego R, Friswell MI (2013) Robustness of optimal sensor placement under parametric uncertainty. *Mech Syst Signal Proc* 41:268–287
- Chisari C, Macorini L, Amadio C, Izzuddin BA (2017) Optimal sensor placement for structural parameter identification. *Struct Multidisc Optim* 55(2):647–662
- Danczyk A, Di X, Liu HX (2016) A probabilistic optimization model for allocating freeway sensors. *Transp Res Pt C-Emerg Technol* 67:378–398
- Deb K, Agrawal RB (1995) Simulated binary crossover for continuous search space. *Complex Sys* 9(2):115–148
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* 6(2):182–197
- Dinh-Cong D, Dang-Trung H, Nguyen-Thoi T (2018) An efficient approach for optimal sensor placement and damage identification in laminated composite structures. *Adv Eng Softw* 119:48–59
- Friswell MI, Castro-Triguero R (2015) Clustering of sensor locations using the effective independence method. *AIAA J* 53(5):1388–1390
- Gomes GF, Cunha SS da da Silva Lopes Alexandrino P, de Sousa BS, Anceletti AC (2018) Sensor placement optimization applied to laminated composite plates under vibration. *Struct Multidisc Optim* 58:2099–2118
- Jäger G, Zug S, Casimiro A (2018) Generic sensor failure modeling for cooperative systems. *Sensors* 18(3):925
- Jayalakshmi V, Rao RM A (2017) Simultaneous identification of damage and input dynamic force on the structure for structural health monitoring. *Struct Multidisc Optim* 55(6):2211–2238
- Jesus G, Casimiro A, Oliveira A (2017) A survey on data quality for dependable monitoring in wireless sensor networks. *Sensors* 17(9):2010
- Jin R, Chen W, Simpson TW (2001) Comparative studies of metamodeling techniques under multiple modelling criteria. *Struct Multidisc Optim* 23(1):1–13
- Kammer DC (1991) Sensor placement for on-orbit modal identification and correlation of large space structures. *J Guid Control Dyn* 14(2):251–259
- Kim T, Youn BD, Oh H (2018) Development of a stochastic effective independence (SEFI) method for optimal sensor placement under uncertainty. *Mech Syst Signal Proc* 111:615–627
- Kullaa J (2013) Detection, identification, and quantification of sensor fault in a sensor network. *Mech Syst Signal Proc* 40(1):208–221
- Li X, Ouyang Y (2011) Reliable sensor deployment for network traffic surveillance. *Transp Res Pt B-Methodol* 45(1):218–231
- Li DS, Li HN, Fritzen CP (2007) The connection between effective independence and modal kinetic energy methods for sensor placement. *J Sound Vibr* 305(4–5):945–955
- Lian J, He L, Ma B, Li H, Peng W (2013) Optimal sensor placement for large structures using the nearest neighbour index and a hybrid swarm intelligence algorithm. *Smart Mater Struct* 22(9):095015
- Liu W, Gao W-C, Sun Y, Xu M-J (2008) Optimal sensor placement for spatial lattice structure based on genetic algorithms. *J Sound Vibr* 317(1–2):175–189
- Lu W, Wen R, Teng J, Li X, Li C (2016) Data correlation analysis for optimal sensor placement using a bond energy algorithm. *Measurement* 91:509–518
- MacKay DJC (1998) Introduction to Gaussian processes. *NATO ASI Ser F Comput Syst Sci* 168:133–166
- Ostachowicz W, Soman R, Malinowski P (2019) Optimization of sensor placement for structural health monitoring: a review. *Struct Health Monit* 18(3):963–988
- Rao ARM, Anandakumar G (2007) Optimal placement of sensors for structural system identification and health monitoring using a hybrid swarm intelligence technique. *Smart Mater Struct* 16:2658–2672
- Ritter J (1990) An efficient bounding sphere. *Graph Gems* 1:301–303
- Salari M, Kattan L, Lam WHK, Lo HP, Esfeh MA (2019) Optimization of traffic sensor location for complete link flow observability in traffic network considering sensor failure. *Transp Res Pt B-Methodol* 121:216–251
- Staszewski WJ, Worden K, Wardle R, Tomlinson GR (2000) Fail-safe sensor distributions for impact detection in composite materials. *Smart Mater Struct* 9(3):298–303
- Sun H, Büyükköztürk O (2015) Optimal sensor placement in structural health monitoring using discrete optimization. *Smart Mater Struct* 24:125034
- Tan Y, Zhang L (2020) Computational methodologies for optimal sensor placement in structural health monitoring: A review. *Struct Health Monit* 19(4):1287–1308
- Welzl E (1991) Smallest enclosing disks (balls and ellipsoids). In: Maurer H (eds) *New results and new trends in computer science*. Springer Berlin, Heidelberg, 359–370
- Williams CKI, Rasmussen CE (2006) *Gaussian processes for machine learning*. MIT press Cambridge, MA.
- Yang C (2018) Sensor placement for structural health monitoring using hybrid optimization algorithm based on sensor distribution index and FE grids. *Struct Control Health Monit* 25(6):e2160
- Yang C (2021) An adaptive sensor placement algorithm for structural health monitoring based on multi-objective iterative optimization using weight factor updating. *Mech Syst Signal Proc* 151:107363
- Yang C, Zheng W, Zhang X (2019a) Optimal sensor placement for spatial lattice structure based on three-dimensional redundancy elimination model. *Appl Math Model* 66:576–591
- Yang C, Liang K, Zhang X, Geng X (2019b) Sensor placement algorithm for structural health monitoring with redundancy elimination model based on sub-clustering strategy. *Mech Syst Signal Proc* 124:369–387
- Yang C, Liang K, Zhang X (2020) Strategy for sensor number determination and placement optimization with incomplete information based on interval possibility model and clustering avoidance distribution index. *Comput Methods Appl Mech Engrg* 366:113042
- Yi T-H, Li HN (2012) Methodology developments in sensor placement for health monitoring of civil infrastructures. *Int J Distrib Sens Netw* 8:612726
- Zhang Z, He M, Liu A (2018) Vibration-based assessment of delaminations in FRP composite plates. *Compos Part B-Eng* 144:254–266
- Zhao J, DeWolf JT (1999) Sensitivity study for vibrational parameters used in damage detection. *J Struct Eng* 125(4):410–416
- Zhu N, Ma S, Zheng L (2017) Travel time estimation oriented freeway sensor placement problem considering sensor failure. *J Intell Transport Syst* 21(1):26–40