



An innovative hybrid strategy for structural health monitoring by modal flexibility and clustering methods

Alireza Entezami^{1,2} · Hassan Sarmadi^{2,3} · Behzad Saeedi Razavi⁴

Received: 7 December 2019 / Revised: 25 June 2020 / Accepted: 1 July 2020 / Published online: 11 July 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Structural health monitoring is usually implemented by model-driven or data-driven methods. Both of them have their advantages and disadvantages. This article proposes an innovative hybrid strategy as a combination of model-driven and data-driven approaches to detecting and locating damage in civil structures. In this regard, modal flexibility matrices of the undamaged and damaged conditions are initially derived from their modal frequencies and mode shapes. Subsequently, the discrepancy between these matrices is proposed as a damage-sensitive feature. To increase damage detectability and localizability, the modal flexibility discrepancy matrix is expanded by the Kronecker product and then converted into a vector by a simple vectorization algorithm yielding vector-style feature samples. To detect and locate damage, this article introduces the k -medoids and density-based spatial clustering of applications with noise techniques. The vector-style feature samples are incorporated into these clustering methods to obtain two different damage indices including the direct clustering outputs and their Frobenius norms. The great novelty of this article is to develop an innovative hybrid strategy for damage detection and localization under noise-free and noisy conditions so that the damage-sensitive feature is obtained from a model-driven scheme and the decision-making is carried out by a data-driven strategy. A shear-building frame and the numerical model of the ASCE benchmark structure are used to validate the accuracy and performance of the proposed methods. Results demonstrate that the hybrid strategy presented here is influentially able to detect and locate damage in the presence of noisy modal data.

Keywords Structural health monitoring · Modal flexibility · Noisy modal data · Clustering · k -medoids, DBSCAN

1 Introduction

Structural health monitoring (SHM) is an important and practical process for ensuring the integrity and serviceability of civil structures [1–5]. Aging, material deterioration, and unpredictable excitation loads always threaten the health and safety of such structural systems and may cause irreparable damage, failure, and even collapse. Damage is an adverse phenomenon that causes changes in the inherent properties of a structure, in most cases its stiffness and dynamic characteristics such as modal data. Generally, this phenomenon is defined as intentional or unintentional changes in the material and/or geometric properties of structural systems, changes in the boundary conditions and system connectivity, which may appear as cracks in concrete elements and connections, bolt loosening and weld cracking in steel connections, failure in steel elements, scour of a bridge pier, etc. The term regarding damage does not necessarily imply a total loss of the structural system functionality but rather that

✉ Hassan Sarmadi
hassan.sarmadi@mail.um.ac.ir; ipesfp.co@gmail.com

¹ Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy

² Department of Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, P.Box: 9177948974, Mashhad, Iran

³ Head of Research and Development, Ideh Pardazan Etebar Sazeh Fanavar Pooya (IPESFP) Company, 29th Reza St., Reza Blvd, P.Box: 9176768544, Mashhad, Razavi Khorasan, Iran

⁴ Department of Construction and Mineral Engineering, Faculty of Technical and Engineering, Standard Research Institute, Tehran, Iran

the system of interest is no longer operating in its optimal and normal manner [6]. To preserve civil structures from unfavorable events caused by damage, it is essential to initially evaluate the global state of the structure in terms of early damage detection and then locate and quantify damage in local manners. Early damage detection is the first level of SHM, which enables civil engineers to know whether the damage is available throughout the structure. Subsequently, it is attempted to implement the second and third levels of SHM, that is, damage localization and quantification [2].

In general, a SHM strategy can be implemented by model-driven and data-driven approaches. A model-driven method is based on constructing an elaborate finite element (FE) model of the real structure and utilizing the main concept of model updating for damage detection, localization, and quantification [7–10]. Model updating is a computational approach that is intended to adjust or correct the FE or analytical model of the real structure [11–13]. For SHM applications, one supposes that the FE model of the structure behaves as an undamaged condition and the real structure is a potentially damaged state [12]. Any dissimilarity between the FE and real models is indicative of damage occurrence. In most cases, the model-driven strategy for SHM is an inverse problem and one needs to apply mathematical techniques for solving damage equations [8, 13–15]. On the other hand, a data-driven method is based on using raw measured vibration data without any FE modeling and model updating. Most of the data-driven techniques rely on the statistical pattern recognition paradigm [16–23]. In the context of SHM, this paradigm consists of sensing and data acquisition, feature extraction, and statistical decision-making via machine learning algorithms [24]. Both the model-driven and data-driven methods have their advantages and disadvantages. Therefore, it is difficult to distinguish which of them prevails against the other one. To address this limitation, one can combine model-driven and data-driven algorithms to develop a hybrid method for SHM [25, 26].

When a structure suffers from damage, the stiffness decreases leading to adverse changes in structural responses and dynamic characteristics. Modal data (i.e., natural frequencies and mode shapes) are useful features for SHM because those are directly related to the structural properties. Despite the simplicity of the measurement of the modal frequencies, those are global dynamic characteristics and cannot present spatial information, which is very important for damage localization. Although the use of mode shapes can address this drawback, it is difficult to obtain a complete set of the modal displacements due to practical and economical limitations regarding dense sensor networks and sensor placement. Therefore, one needs to present a new feature for SHM. As an alternative, modal flexibility is a function of both the modal frequencies and mode shapes. This function is an inverse of the structural stiffness [27], in

which case one can exploit it as a proper damage-sensitive feature. The great advantage of the modal flexibility function is that it is only derived from the modal data, even in a few modes, without any requirement of using the structural properties. Another merit of this function is to deal with the drawbacks of the modal frequencies and mode shapes for damage detection and localization. Therefore, one can consider the modal flexibility as a better damage-sensitive feature than the modal data. Due to such remarkable merits, many researchers proposed this function for damage detection [9, 27–32]. Most of these studies are based on solving inverse problems or applying linear and non-linear optimization techniques. When the modal data used in the flexibility matrix are contaminated with noise, those problems become ill posed [8, 13, 32], which are susceptible to an unstable solution with erroneous results.

Rather than using mathematical and optimization techniques for solving the ill-posed inverse problems, this article proposes statistical methods for damage detection and localization using the modal flexibility. These methods can be utilized in an unsupervised learning manner. Among them, cluster analysis presents a data-driven unsupervised learning method that does not require rigorous learning schemes [33]. In this regard, Diez et al. [22] presented a clustering-based method to group substructures or joints with the same behavior on bridges and then detect abnormal or damaged areas by *k*-means clustering. In another article, Mahato and Chakraborty [34] utilized this clustering technique with the aid of the wavelet transform for modal identification. da Silva et al. [35] proposed a three-stage damage detection framework and used fuzzy clustering methods in the third stage to detect damage. Figueiredo and Cross [21] proposed the Gaussian mixture model for detecting damage under strong environmental variability conditions and concluded that this kind of clustering analysis outperformed some linear approaches to SHM. Silva et al. [36] proposed agglomerative concentric hypersphere clustering for early damage detection under varying environmental variations. In another study, Langone et al. [37] proposed kernel spectral clustering for SHM in bridge structures.

The main objective of this article is to propose an innovative hybrid method as the combination of the model-driven and data-driven strategies for detecting and locating damage under noise-free and noisy modal data. In this regard, the discrepancy between the modal flexibility matrices of the undamaged and damaged states of the structure is introduced as the main damage-sensitive feature. To increase damage detectability and localizability, the discrepancy matrix of modal flexibility is initially expanded by the Kronecker product and subsequently converted into a vector using a simple vectorization algorithm. In the second stage regarding the data-driven strategy, this article introduces the *k*-medoids and a new non-parametric clustering method

called density-based spatial clustering of applications with noise (DBSCAN) to apply the vector-style damage feature for detecting and locating damage. Two kinds of damage indices including the direct clustering outputs and their Frobenius norms are used to present the results of damage detection and localization. The accuracy and performance of the proposed methods are numerically verified by a simple shear-building model and the ASCE benchmark structure in the first phase. Results show that both clustering methods are capable of detecting and locating damage even when the modal parameters are contaminated with different levels of noise. Furthermore, it is observed that the modal flexibility discrepancy matrix, which is used as the damage-sensitive feature, assists the clustering methods in providing reliable results of damage detection and localization.

2 Modal flexibility

The modal flexibility or structural flexibility matrix \mathbf{F} is a function derived from the modal frequencies (eigenvalues) and mode shapes (eigenvectors). In a FE model of a structure with n degrees-of-freedom (DOFs), which this model refers to the undamaged state of that structure, the modal flexibility function (\mathbf{F}_u) is expressed as follows:

$$\mathbf{F}_u = \sum_{i=1}^n \frac{1}{\lambda_i} \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^T, \quad (1)$$

where λ_i and $\boldsymbol{\varphi}_i = [\varphi_{i,1} \dots \varphi_{i,n}]^T$ denote the eigenvalue and eigenvector of the i th mode, respectively. Because the global mass and stiffness matrices of the FE model or undamaged state of the structure are available, the modal data can simply be obtained by the eigenvalue problem. On the other hand, it is possible to define the same formulation of the modal flexibility for the real structure, which is considered as the potentially damaged condition. Here, one assumes that no FE model is needed for this condition and its modal frequencies and mode shapes are representative of the measured data. The main limitation is that the measured modal parameters are usually incomplete. In reality, one cannot measure all modes of the structure in an experimental program due to some practical and economic limitations. Furthermore, it is not necessarily feasible to measure all modal displacements at all DOFs. Under such circumstances, the measured modal parameters of the damaged state are incomplete, which means that those are only available and measurable in a few modes and DOFs.

The other important note is that the mode shape vectors of the undamaged and damaged conditions of the structure should originate from the same physical condition [38]. This means that both of them should be scaled and mass-normalized. Due to the availability of the mass matrix of

the FE model, the mode shapes of the undamaged state are always mass-normalized. However, the inherent properties of the real structure are often unknown, in which case it is necessary to scale or normalize the measured modal displacements. With these descriptions, the mode shapes of the real model are initially normalized and then expanded to obtain a complete set of the mass-normalized measured modal displacements. In this article, the procedures of normalization and mode expansion are carried out by the modal scale factor and SEREP technique [8].

Assume that the only m measured modes of the modal frequencies and mode shapes of the damaged state are available. This means that limited identifiable frequencies in the first few modes are measurable. Hence, the modal flexibility matrix (\mathbf{F}_d) of the real structure or damaged condition is derived from the truncated lower modes of the measured modal frequencies and mode shapes as follows:

$$\mathbf{F}_d = \sum_{j=1}^m \frac{1}{\hat{\lambda}_j} \hat{\boldsymbol{\varphi}}_j \hat{\boldsymbol{\varphi}}_j^T, \quad (2)$$

where $\hat{\lambda}_j$ and $\hat{\boldsymbol{\varphi}}_j = [\hat{\varphi}_{j,1} \dots \hat{\varphi}_{j,n}]^T$ are the j th modal frequency and normalized mode shape vector of the damaged state, respectively. Despite the difficulty or impossibility of measuring all modes, the truncated modal flexibility in Eq. (2) using the first lower modes is sufficiently accurate and can be correctly estimated [27]. Furthermore, it should be mentioned that both \mathbf{F}_u and \mathbf{F}_d are n -by- n square matrices despite using incomplete modes associated with the damaged state. Each of the modal flexibility matrices cannot alone express the changes in the structure caused by damage. The best way is to calculate their difference or residual and obtain the discrepancy matrix of modal flexibility, that is, $\Delta \mathbf{F} = |\mathbf{F}_d - \mathbf{F}_u|$, where $|\cdot|$ stands for the absolute operator.

3 Cluster analysis

Clustering is a statistical tool for arranging and dividing data samples into groups or clusters [39]. Because this method conforms to the unsupervised learning class, it is widely used in SHM applications. Cluster analysis consists of various clustering algorithms, each of which seeks to organize a given data set into homogeneous clusters. A cluster is generally defined as a group of observations such as objects or data points that have large similarities [39]. Clustering methods depend on how to separate similar or dissimilar observations from each other. In other words, the same observations are treated as a homogeneous cluster, while dissimilar observations make additional clusters. There are a large number of clustering algorithms for data analysis. The choice of a clustering algorithm depends on the type of data and the purpose and application of clustering. If cluster analysis is

used as a descriptive or exploratory tool, it is possible to try several algorithms on the same data to observe what the data may disclose. Generally, most of the clustering algorithms can be classified into (1) hierarchical, (2) partitioning, (3) density-based, (4) grid-based, and (5) model-based methods. A full discussion on each of these methods is beyond the scope of this article. For more information, the reader can refer to [39, 40].

3.1 k-medoids clustering

The k -medoids clustering is a partitioning method that divides the data of s samples into k clusters. In this method, there is no hierarchical relation between the k -cluster solution and the $(k + 1)$ cluster solution; therefore, it is an appropriate clustering method for large data sets [39]. This approach is related to the k -means clustering, that is, both of them are intended to divide a set of observations into k clusters in such a way that the clusters minimize an error sum-of-squares (ESS) criterion between an observation and a center of the cluster.

A distinction between these methods is that the center of observations in the k -means clustering is the mean of observations, which is usually called centroid as a reference point, whereas the k -medoids clustering chooses a representative observation as a reference point. To put it another way, the medoid of a cluster in the k -medoids clustering is defined as an observation (i.e., the most centrally located observation in the cluster) that minimizes the total dissimilarity to all other observations within that cluster, while the centroid in the k -means clustering is the

mean or average of the observations. In other words, the key merit of the k -medoids clustering is to use representative observations as reference points rather than taking the mean of observations in each cluster. Due to such a merit, this method is more robust to noise and outliers compared to the k -means clustering [39]. Figure 1 depicts the difference between these methods when one of the data points (i.e., one that is far away from the others) is an outlier. As can be observed, the k -means clustering is sensitive to the outlier and chooses the cluster mean by considering that data point. On the contrary, the k -medoids clustering finds the cluster center or medoid without affecting the outlier.

The k -medoids clustering starts with defining a proximity metric such as the Euclidean distance and finds a medoid within each cluster that minimizes the total dissimilarity to all other medoids within that cluster. For the s -dimensional vectors \mathbf{x}_i and \mathbf{x}_j , the proximity between these vectors ($D_{i,j}$) based on the Euclidean distance is expressed as:

$$D_{i,j} = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)}. \quad (3)$$

In the k -medoids clustering, the ESS criterion is formulated based on the pre-defined proximity metric as follows [39]:

$$ESS = \sum_{r=1}^k \sum_{c(i)=k} \sqrt{(\mathbf{x}_i - \mathbf{c}_k)^T (\mathbf{x}_i - \mathbf{c}_k)^T}, \quad (4)$$

where \mathbf{c}_k is the k th cluster medoid and $\mathbf{c}(i)$ is the cluster containing \mathbf{x}_i . For more information, Fig. 2 illustrates the flowchart of the k -medoids clustering.

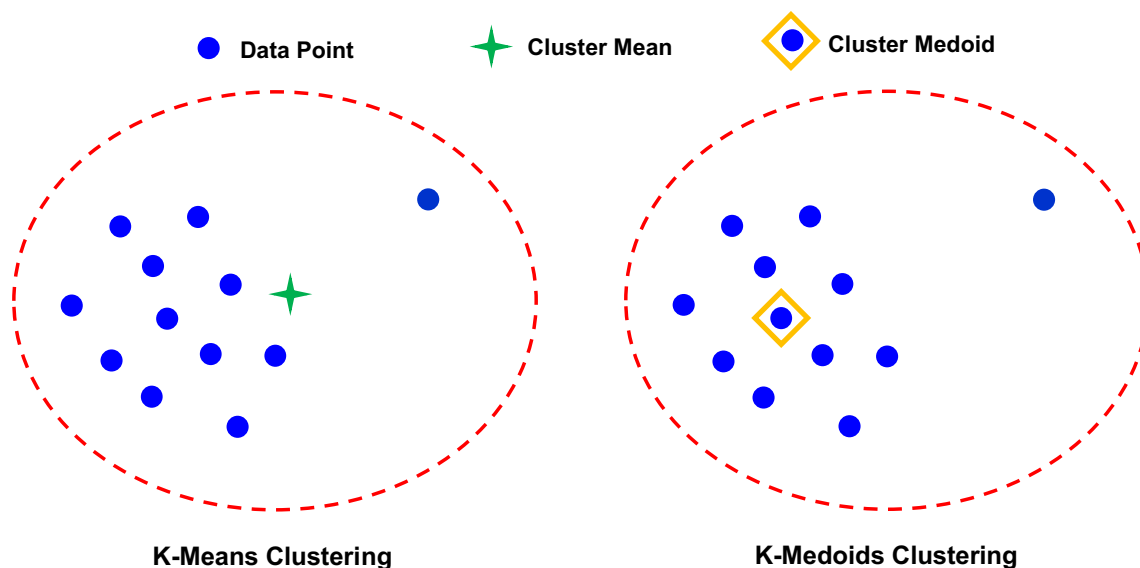


Fig. 1 The graphical representation of the difference between the k -means and k -medoids clustering methods

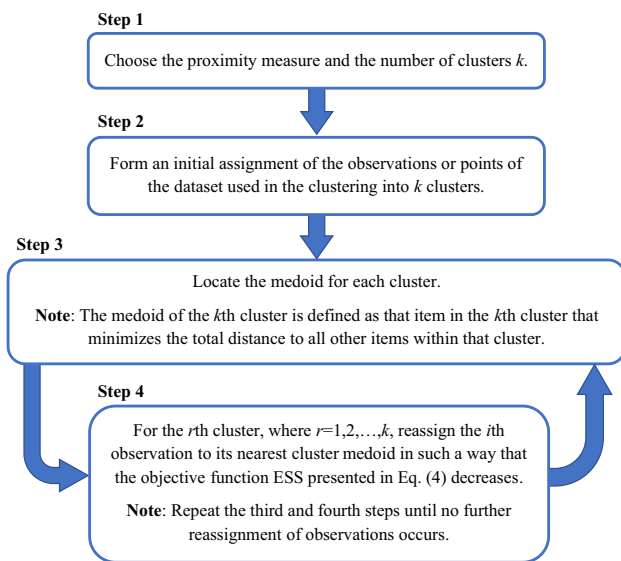


Fig. 2 The flowchart of the k -medoids clustering

3.2 DBSCAN

DBSCAN is a density-based clustering approach, which was initially proposed by Ester et al. [41]. This method clusters a set of observations into spaces with different sizes and shapes under the assumption that a cluster is a region in the data space with a high density. Indeed, the DBSCAN method groups the data points that are dense and nearby into a single cluster [40]. The algorithm of this method has computational similarities to centroid-based clustering techniques such as the k -means and k -medoids clustering. However, DBSCAN utilizes the density of the data points in the feature space to identify clusters rather than the locations of the centroids or medoids. Most of the clustering methods including hierarchical and partitioning methods require the number of clusters that should be known and determined by some computational techniques. However, the DBSCAN algorithm only

needs two input parameters including the minimum number of points, which is required to insert into a region for making it dense (MinPts), and the maximum distance between points in a cluster as a radius (Eps). More precisely, this technique adopts the radius value Eps based on a user-defined distance measure and the value MinPts for collecting the points in the region made by Eps. Using these scenarios, it is not necessary to specify the number of clusters.

Given the vector $\mathbf{x} = [x_1 \dots x_s]^T$, DBSCAN yields a local density denoted as $\rho(x_i)$ in the neighborhood of the i th point x_i , which is the total number of points in its neighborhood. Note that the neighborhood is a distance measure for two points x_i and x_j , denoted by $d_{i,j}$. The algorithm of DBSCAN classifies the data points $x_1 \dots x_s$ as core, border, density-reachable, and outlier points. Figure 3 depicts a graphical representation of these points. Having considered the minimum number of points in the neighborhood (MinPts), x_i is defined as a core point ($x_i = x_{\text{core}}$) if $\rho(x_i) \geq \text{MinPts}$ as can be observed in Fig. 3b. Moreover, x_i can be a border point ($x_i = x_{\text{border}}$) if $\rho(x_i) < \text{MinPts}$, in which case a core point exists so that $x_{\text{border}} \in N_{\text{Eps}}(x_{\text{core}})$. This means that the border point x_{border} belongs to the neighborhood of x_{core} and the local density is less than MinPts. It is worth remarking that N_{Eps} represents the nearest points in the neighborhood of the radius Eps of the i th point, which is defined as $N_{\text{Eps}}(x_i) = \{x_j | \forall j, d_{i,j} < \text{Eps}\}$. The other important definition used in the DBSCAN algorithm is related to the density-reachable points. Given two points x_1 and x_s , these are the density-reachable points if a chain of points $x_1, \dots, x_i, x_{i+1}, \dots, x_s$ exists, where $i \geq 1$ and $s \geq 2$. In such a case, for all $i < s$, x_i is a core point ($\rho(x_i) \geq \text{MinPts}$), and x_{i+1} is a neighbor of x_i ($x_{i+1} \in N_{\text{Eps}}(x_i)$). Eventually, a point is an outlier (x_{outlier}) if its local density is less than MinPts. In Fig. 3d, the three black points out of the clusters are the outlier points.

It should be mentioned that the DBSCAN algorithm begins with an arbitrary starting point from the vector \mathbf{x} and retrieves all neighbors of that point within Eps. If it is a core point, the algorithm creates a new cluster and assigns that

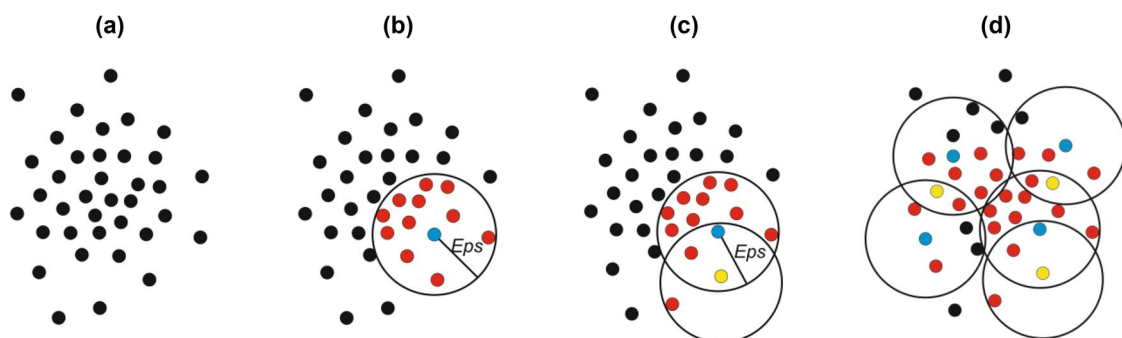


Fig. 3 The graphical representation of the key components of the DBSCAN algorithm: **a** a cluster, **b** a core point (the blue point), **c** a border point (the yellow point), **d** density-reachable points

point and its neighbors into this new cluster. Subsequently, the algorithm collects the neighbors within Eps from all core points in an iterative manner. The process is repeated until all of the data points in \mathbf{x} are considered.

4 Proposed SHM strategy

The k -medoid and DBSCAN methods utilize the vector $\mathbf{x} \in \mathcal{R}^s$ (input) to divide its data points into clusters and yield clustering outputs. As explained in Sect. 2, the modal flexibility discrepancy matrix $\Delta \mathbf{F} \in \mathcal{R}^{n \times n}$ is considered as the damage-sensitive feature for damage detection and localization. In such a case, it is necessary to convert it into a vector and use the vector-style feature as the input data in the above-mentioned clustering methods. Before this procedure, it is better to expand $\Delta \mathbf{F}$ by the Kronecker product to increase damage detectability and localizability. Using this operator, one can obtain the expanded matrix $\Delta \mathbf{F}^* = \Delta \mathbf{F} \otimes \Delta \mathbf{F} \in \mathcal{R}^{n^2 \times n^2}$. In the following, this matrix is converted into a vector using the vectorization process leading to the vector $\Gamma \in \mathcal{R}^{n^4}$. In mathematics and linear algebra, the vectorization of a matrix is a linear transformation that converts it into a column vector [12]. Therefore, the vector

Γ , which is equivalent to the vector \mathbf{x} (i.e., $s = n^4$), is fed into the k -medoids and DBSCAN methods for damage detection and localization.

5 Applications

This section presents the results of damage detection and localization via the proposed methods using two numerical models including a shear building frame and the numerical ASCE benchmark structure related to the first phase.

5.1 The shear building model

As the first example, a simple undamped shear building frame with six stories is modeled to evaluate the robustness and performance of the proposed methods. This model is a simulation of a dynamic discrete system with six DOFs as shown in Fig. 4. The initial structural properties of the model such as the mass and stiffness of each story are listed in Table 1. In this model, one assumes that each floor at each DOF, except for the ground floor, equips with a sensor; therefore, it is possible to measure all modal displacements at all DOFs. Several damage scenarios are defined to simulate structural damages in the model based on the stiffness reduction of some stories. Accordingly, it is presumed that the mass matrix remains invariant in these scenarios. Table 2 lists the three damage cases applied to the shear-building model. Moreover, one can simply observe these cases in Fig. 4.

Given the initial structural parameters of the shear building model in the undamaged and damaged conditions, the

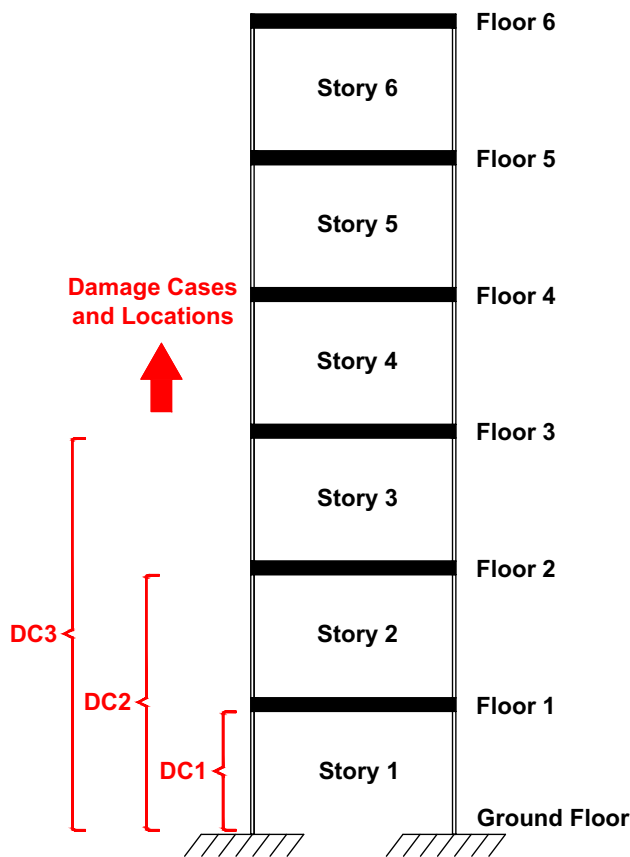


Fig. 4 The shear-building model

Table 1 The structural characteristics of the shear-building model

Story no.	Mass (kg)	Stiffness (kN/m)
1–2	3000	1200
3–4	2000	1000
5–6	1000	800

Table 2 The damage scenarios in the shear-building model

Damage cases	Location (story)	Stiffness reduction (%)
DC1	1	– 5
DC2	1	– 10
	2	– 25
DC3	1	– 15
	2	– 30
	3	– 55

global mass and stiffness matrices are simply obtained using their simple formulations regarding the discrete dynamical systems. The generalized eigenvalue problem is applied to extract the modal parameters of the undamaged and damaged conditions. As described earlier, one assumes that the modal data of the undamaged state are related to the FE model of the shear building. Furthermore, it is supposed that the modal parameters of the damaged conditions (i.e., DC1–DC3) are measurable and available in a few modes. In most cases, noise contaminates measured modal data regarding the damaged states and causes undesirable perturbations in such data. On this basis, four levels of noise such as 1, 3, 5, and 10% are introduced to the measured modal parameters in the following forms:

$$\hat{\Phi}_j^* = \hat{\Phi}_j + \alpha \mathbf{v}, \tag{5}$$

$$\hat{\lambda}_j^* = \hat{\lambda}_j + \alpha \nu, \tag{6}$$

where α is the noise level; \mathbf{v} and ν represent the vector of random samples with the standard normal distribution and a scalar value with zero mean and unit standard deviation, respectively. Once the inherent physical properties of the undamaged and damaged states, as well as their modal parameters, have been obtained, the modal flexibility matrices \mathbf{F}_u , \mathbf{F}_d , and $\Delta \mathbf{F} \in \mathcal{R}^{6 \times 6}$ are determined. As a sample, Fig. 5 indicates the modal flexibility discrepancy matrix related to DC2. It is clear from this figure that most of the changes have occurred in the first and second stories. An important note in Fig. 5 is that the reduction in the first story is much less than the second story despite the presence of a more intensive severity of damage in the second story (the stiffness reduction equal to -25%) than the first one (the

stiffness reduction equal to -10%). Therefore, one can conclude that although the discrepancy matrix of modal flexibility is sensitive to damage, it is not sufficiently suitable for damage detection and localization. In fact, this conclusion emphasizes the importance of applying the clustering methods.

Using the Kronecker product, the modal flexibility discrepancy converts into a 36-by-36 matrix. Finally, this expanded matrix is converted into a vector with 1296 samples through the vectorization algorithm. It is important to note that each story of the shear-building model has 216 samples (observations), in which case the vector-style feature used in the k -medoids and DBSCAN algorithms divides into 6 clusters with 216 samples. The results of damage detection by DBSCAN and k -medoids clustering in DC1–DC3 are shown in Figs. 6, 7, 8, respectively. Note that the expression “ST” stands for the abbreviation of “Story”, which is used throughout this article. For the k -medoids clustering, the number of clusters (k) is set as 6. From Fig. 6, one can observe that most dispersion of the clustering outputs (the amounts of the vector Γ) is related to the first 216 samples, which belong to the first story of the shear-building model. For the other stories, it is seen that the outputs of the clustering methods are zero implying the undamaged areas in the shear-building model. In Fig. 7, the first and second 216 samples of the clustering outputs are more than the other stories. Therefore, it can be argued that the first and second stories of the shear-building model have suffered from damage. Eventually, the observations in Fig. 8 demonstrate that the first, second, and third stories of the model are the damaged areas due to considerable dispersion of the clustering outputs at these stories. All the obtained results in Figs. 6, 7, 8 lead to the conclusion that the proposed clustering methods in conjunction with the modal flexibility discrepancy are highly capable of detecting and locating damage in the noise-free condition.

To assess the robustness of the proposed methods in the presence of noisy modal data, Fig. 9 shows the results of damage detection in DC3 for the high level of noise (10%). As can be discerned, the first, second, and third stories of the shear building are identified as the damaged areas due to substantial dispersion of the clustering outputs. This conclusion is similar to the corresponding result in Fig. 8 without any noise in the modal data. Therefore, it can be realized that the proposed methods are still successful in detecting damage even under the noisy modal data. Furthermore, one can observe that there are some erroneous results at the locations of undamaged areas (the stories 4–6) due to the negative effect of noise. By comparing the clustering outputs at these stories, it can be perceived that DBSCAN outperforms the k -medoids clustering owing to smaller computational errors. In Fig. 9a, the maximum output of the k -medoids clustering at the mentioned undamaged stories is about 0.04, whereas

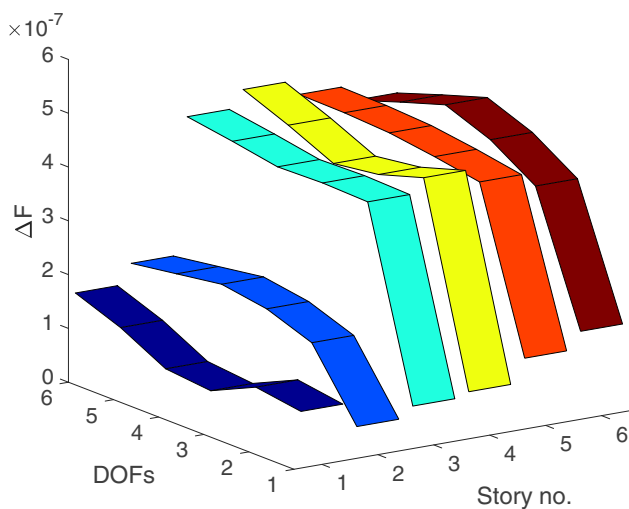


Fig. 5 The discrepancy matrix of modal flexibility in DC2

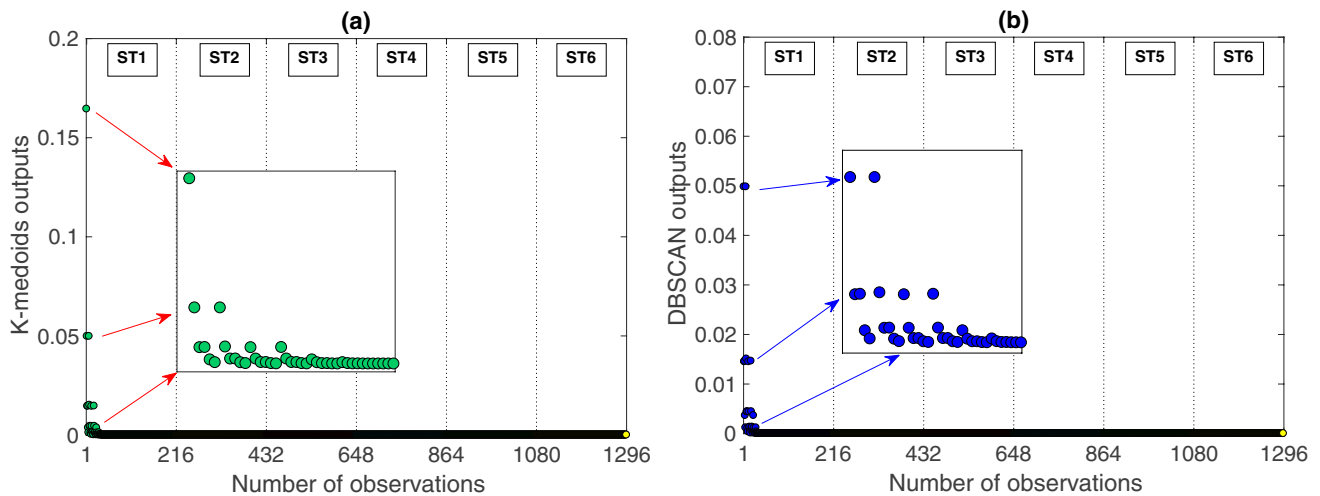


Fig. 6 Damage detection and localization in the shear-building model for the noise-free condition in DC1: **a** *k*-medoids clustering, **b** DBSCAN

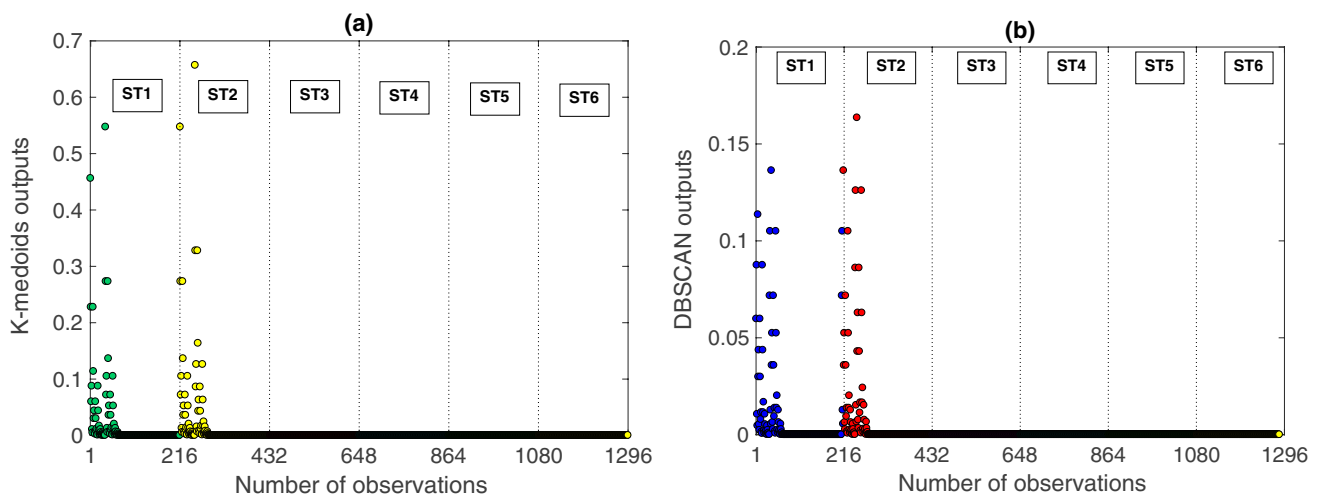


Fig. 7 Damage detection and localization in the shear-building model for the noise-free condition in DC2: **a** *k*-medoids clustering, **b** DBSCAN

Fig. 9b indicates that the maximum DBSCAN output (error) in these areas corresponds to 0.001.

The other way for verifying the accuracy and effectiveness of the proposed methods in detecting and locating damage is to determine the norms of the clustering outputs concerning DC1–DC3 in the different noise levels as shown in Figs. 10, 11, 12, respectively. From these figures, it can be understood that the Frobenius norms of the clustering outputs in the noisy modal data are approximately identical to the corresponding values in the noise-free condition. This conclusion is also valid for both clustering methods. Thus, it can be concluded that the *k*-medoids clustering and DBSCAN methods along with the modal flexibility discrepancy succeed in detecting and locating damage under noise-free and noisy conditions. An important note in Figs. 10, 11,

12 is that the use of the Frobenius norms of the clustering output provides an appropriate tool for locating damage. It is simply seen in these figures that the norms of the clustering outputs precisely identify the damaged areas of the shear building. In this regard, the comparison between the results of damage detection and localization in Figs. 6–8 and 10–12 reveals that the Frobenius norms present more obvious results of damage localization than the direct use of the clustering outputs.

5.2 The ASCE benchmark structure

Another verification example is the numerical model of the ASCE benchmark structure [42, 43]. It is a four-story steel frame in the model scale as shown in Fig. 13. This structure

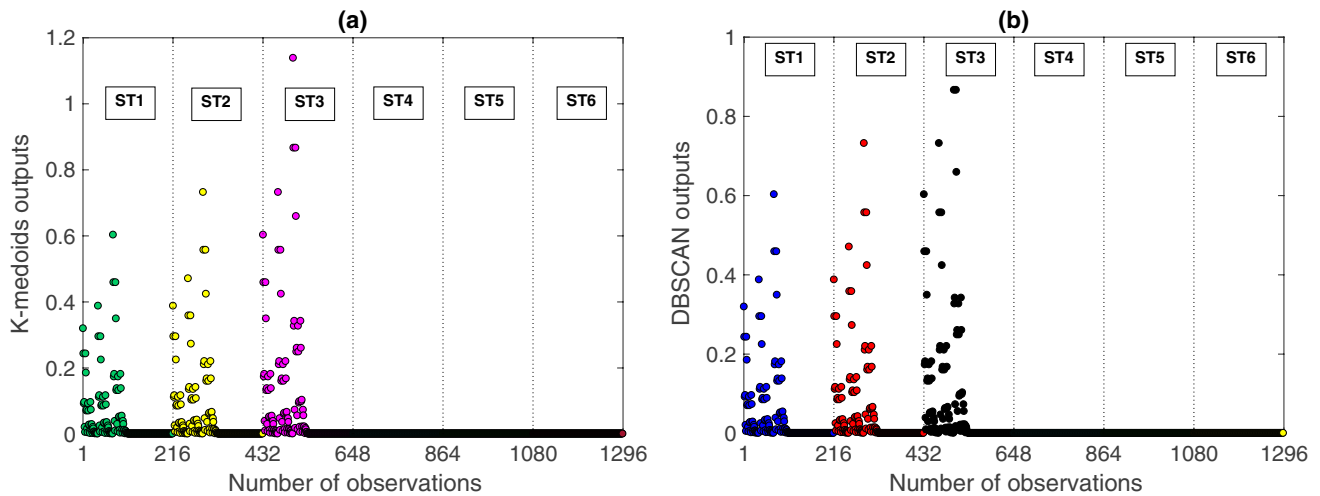


Fig. 8 Damage detection and localization in the shear-building model for the noise-free condition in DC3: **a** *k*-medoids clustering, **b** DBSCAN

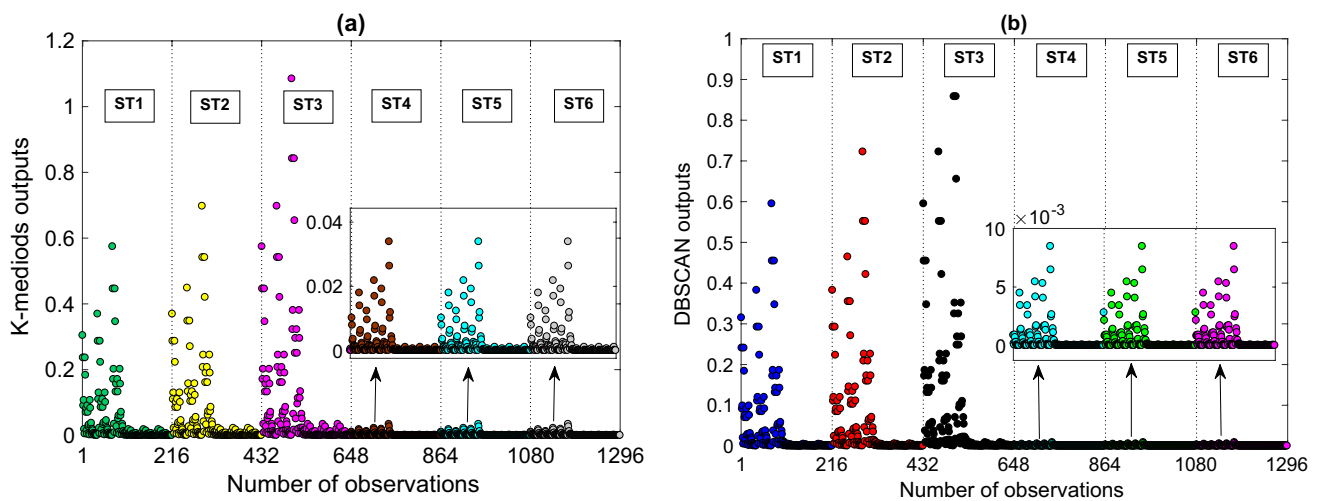


Fig. 9 Effect of noisy modal data on the results of damage detection and localization in DC3: **a** *k*-medoids clustering, **b** DBSCAN

includes a plan of 2.5 m-by-2.5 m and a height of 3.6 m. The members are hot rolled grade 300 W steel with nominal yield stress 300 MPa. The columns and floor beams are B100 × 9 and S75 × 11 sections, respectively. There are two types of FE models of the ASCE structure including 12-DOF and 120-DOF models. This article considers the first model for damage detection and localization. For this structure, six damage patterns were defined as reductions in the structural stiffness by removing brace systems from some stories as illustrated in Fig. 14. Accordingly, the third and fourth damage patterns are used here to evaluate the performance and effectiveness of the proposed methods.

Since the main objective of modeling the numerical ASCE benchmark structure is to simulate acceleration time histories at simulated sensor locations, it is possible

to access the mass and stiffness matrices of the undamaged and damaged conditions. Similar to the previous example, it is assumed that the FE model of the ASCE structure refers to its undamaged condition and the damaged state is indicative of the real model. The generalized eigenvalue problem is employed to extract the modal parameters of the undamaged and damaged conditions. In order to simulate realistic situations, one supposes that the only five modes regarding the damaged state are available and measurable. Since the mode shapes of the undamaged condition are mass-normalized, the modal scale factor is applied to scale the modal displacements of the damaged condition. Furthermore, the mode expansion technique is used to expand the mass-normalized modal displacements of this condition using the SEREP technique. The same

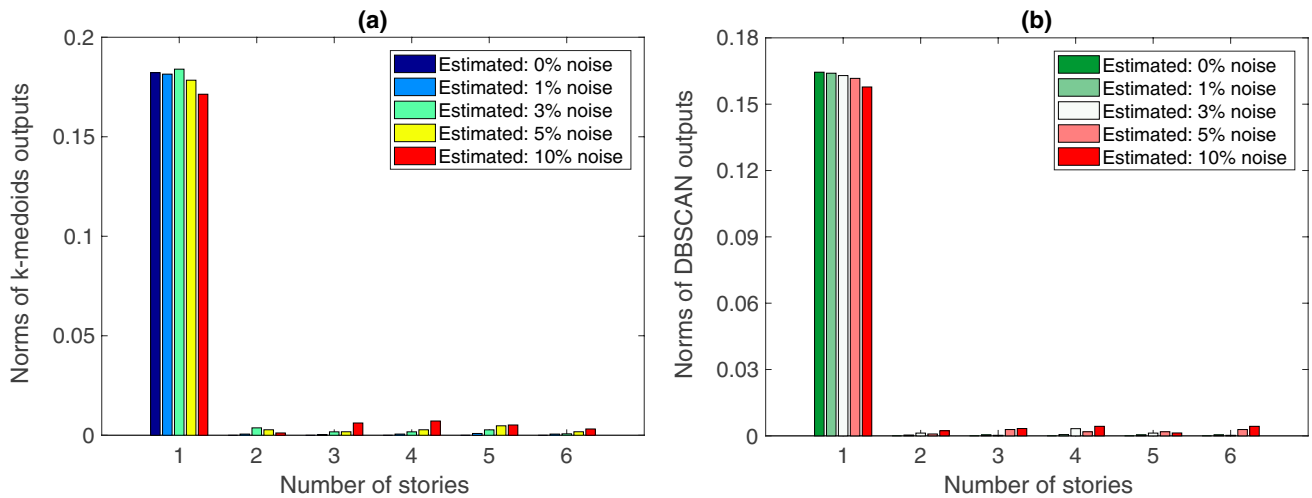


Fig. 10 Damage detection and localization using the Frobenius norms of the clustering outputs in the different noise levels for DC1: **a** *k*-medoids clustering, **b** DBSCAN

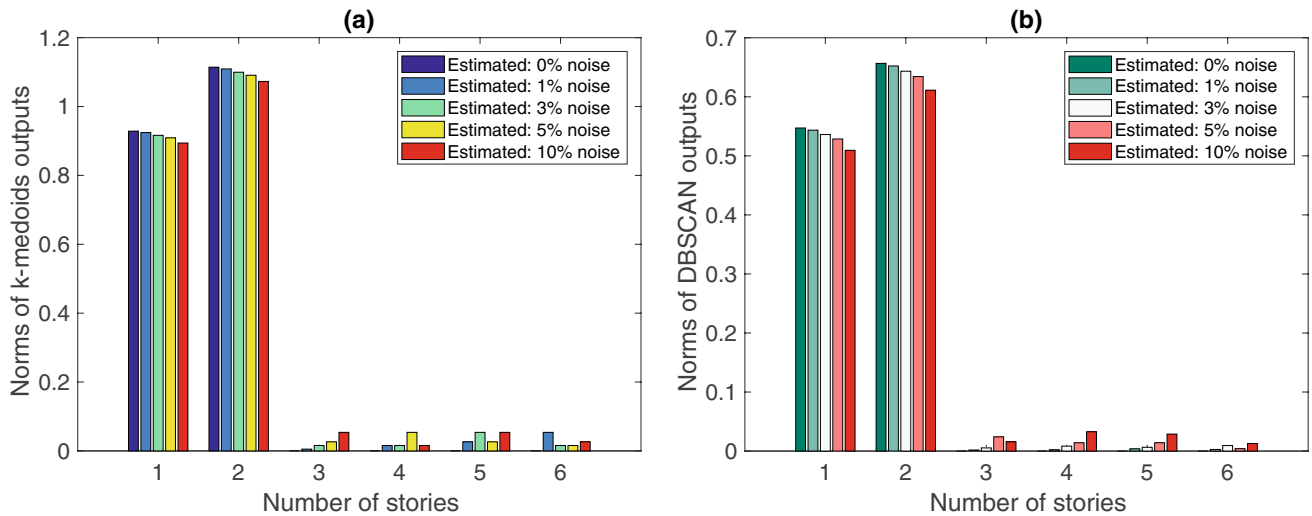


Fig. 11 Damage detection and localization using the Frobenius norms of the clustering outputs in the different noise levels for DC2: **a** *k*-medoids clustering, **b** DBSCAN

noise levels as the previous numerical example are considered to simulate noisy modal data.

Using the structural properties and modal parameters of the undamaged and damaged conditions, the modal flexibility matrices F_u and $F_d \in \mathcal{R}^{12 \times 12}$ are obtained to determine the modal flexibility discrepancy matrix $\Delta F \in \mathcal{R}^{12 \times 12}$. Based on the Kronecker product, the modal flexibility discrepancy matrix is expanded into a 144-by-144 matrix and this expanded matrix is then converted to a vector with 20,736 samples, which serves as the input data for the clustering methods. It is important to point out that each story of the ASCE structure consists of 5184 observations (samples). The results of damage detection and localization for the third

and fourth damage patterns on the basis of the outputs of the *k*-medoids clustering and DBSCAN under the noise-free condition are shown in Figs. 15 and 16, respectively. Note that the number of clusters required for the *k*-medoids clustering is equal to 4. As Fig. 15 shows, it can be observed that the first story of the ASCE structure is representative of the damaged area because the first 5184 samples of the clustering outputs have the most dispersion in comparison with the other areas. This is a reasonable result since the damage pattern of the third case has been simulated at this story. From Fig. 16, one can discern that the first and third 5184 samples of the clustering outputs include most dispersion compared to the other samples. Therefore, it is deduced that the first

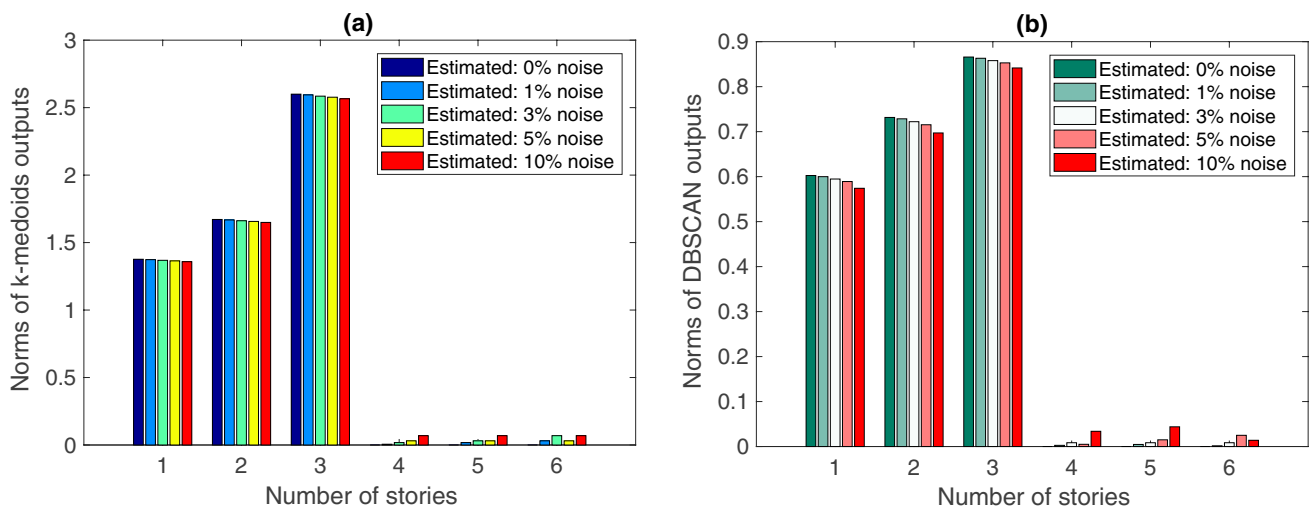


Fig. 12 Damage detection and localization using the Frobenius norms of the clustering outputs in the different noise levels for DC3: **a** *k*-medoids clustering, **b** DBSCAN

and third stories of the ASCE structure have suffered from damage, which is an accurate conclusion according to the fourth damage pattern.

All the previous results have been achieved using the vector-style input data obtained from the vectorization of the expanded matrix of the modal flexibility discrepancy through the Kronecker product. The main premise is that the use of this operator increases the damage detectability and localizability. As a comparison, one attempts to evaluate the result of damage detection without the Kronecker product. In other words, the vectorization procedure is only carried out on the modal flexibility discrepancy matrix $\Delta F \in \mathcal{R}^{12 \times 12}$. Hence, the new vector-style input data needed for the clustering methods consists of 144 samples. Figure 17 shows the outputs of DBSCAN in the third and fourth damage

patterns without applying the Kronecker product. By comparing the results in Figs. 15, 16, 17, one can realize that the Kronecker product gives more clear results than the situation without applying it. This issue may be significant when small damage occurs in the structure. In such a case, the clustering methods may not be able to detect and locate damage accurately.

Similar to the previous numerical example, the other result of damage detection and localization is based on computing the Frobenius norms of the clustering outputs as shown in Fig. 18, where Fig. 18a, b is related to the results obtained from the *k*-medoids clustering, and Fig. 18c, d shows the results of the DBSCAN method. All observations in Fig. 18 confirm the great ability of the proposed methods to detect and locate damage even under the noisy modal data. Some inconsiderable norm quantities are also observable in the undamaged areas (i.e., the stories 2–4 for the third pattern and the stories 2 and 4 concerning the four pattern), which can be neglected them. Furthermore, it can be seen that the Frobenius norms of the clustering outputs in the noisy modal parameters are roughly similar to the corresponding values in the noise-free condition. This means that the proposed methods are satisfactorily able to detect and locate damage under noise-free and noisy modal data.

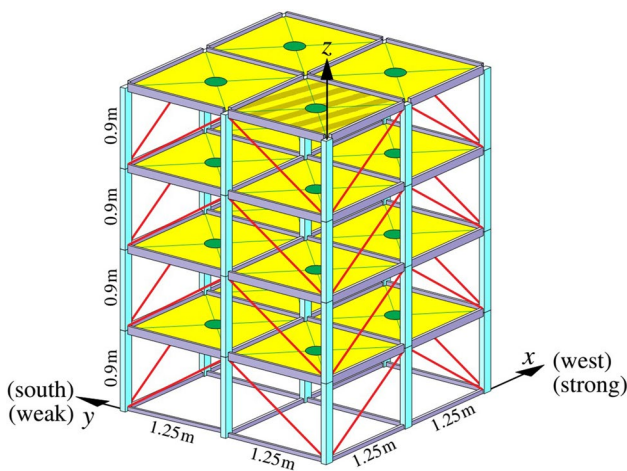


Fig. 13 The numerical model of the ASCE benchmark structure [42]

6 Conclusions

In this article, an innovative hybrid method was proposed to detect and locate damage under noisy modal data. The discrepancy between the modal flexibility matrices regarding the undamaged and damaged conditions was selected as the main damage-sensitive feature. The modal flexibility

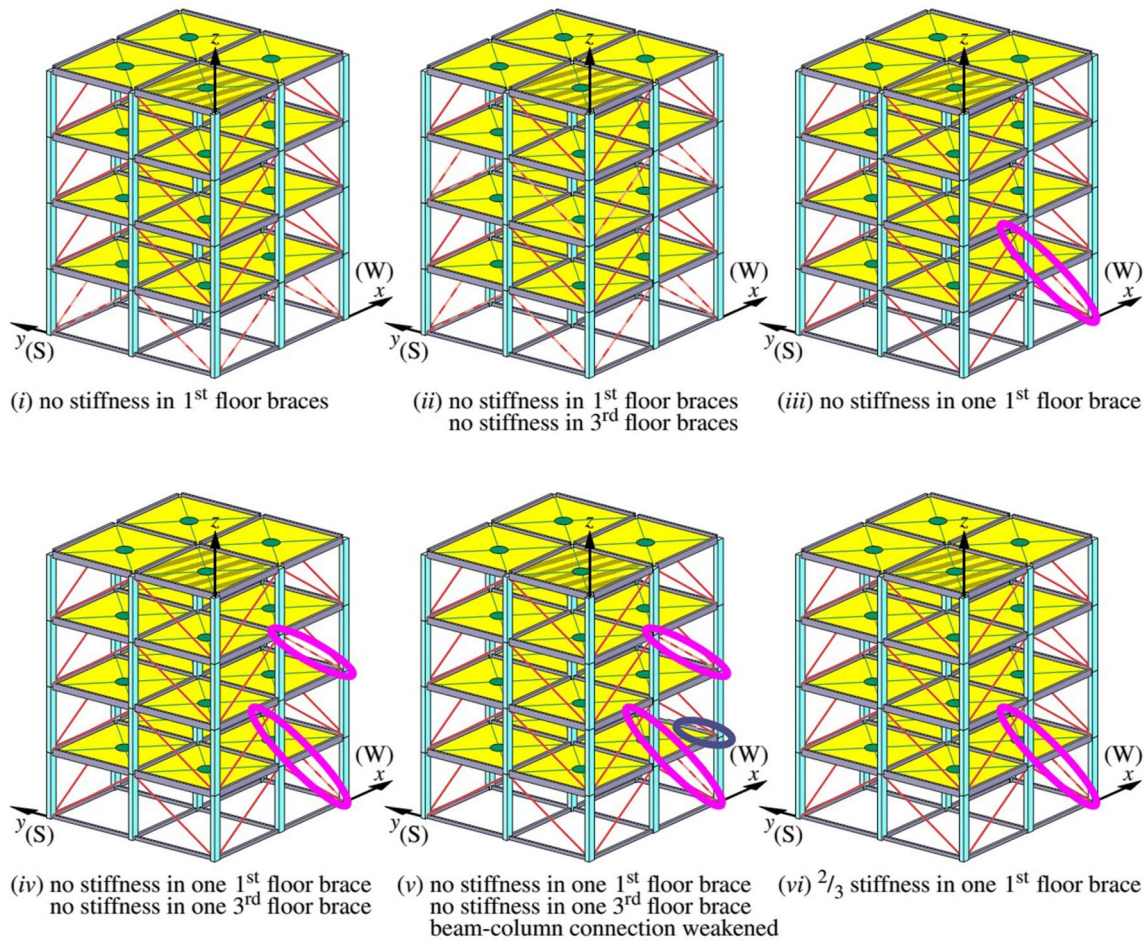


Fig. 14 The six damage patterns of the numerical problem of the ASCE benchmark model [42]

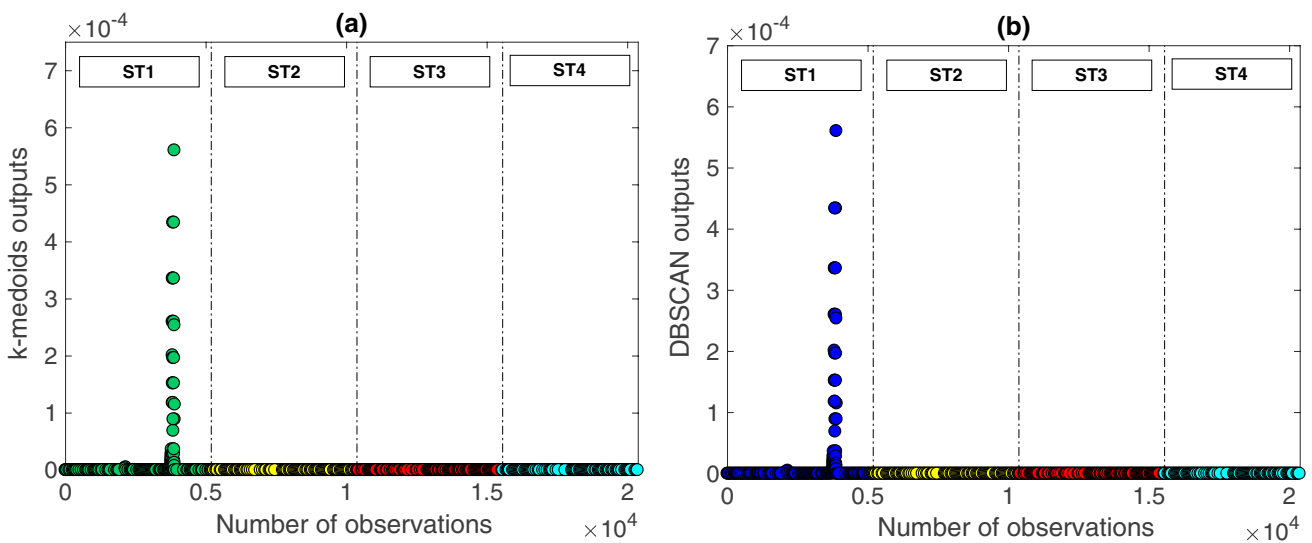


Fig. 15 Damage detection and localization in the ASCE benchmark structure in the third damage pattern: **a** *k*-medoids clustering, **b** DBSCAN

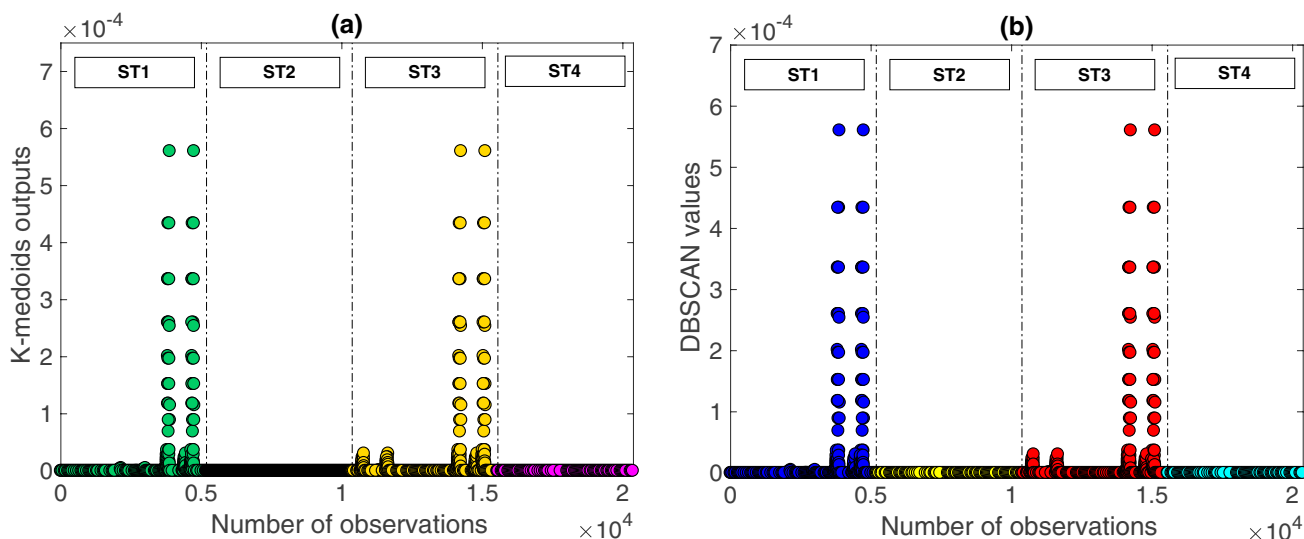


Fig. 16 Damage detection and localization in the ASCE benchmark structure in the fourth damage pattern: **a** *k*-medoids clustering, **b** DBSCAN

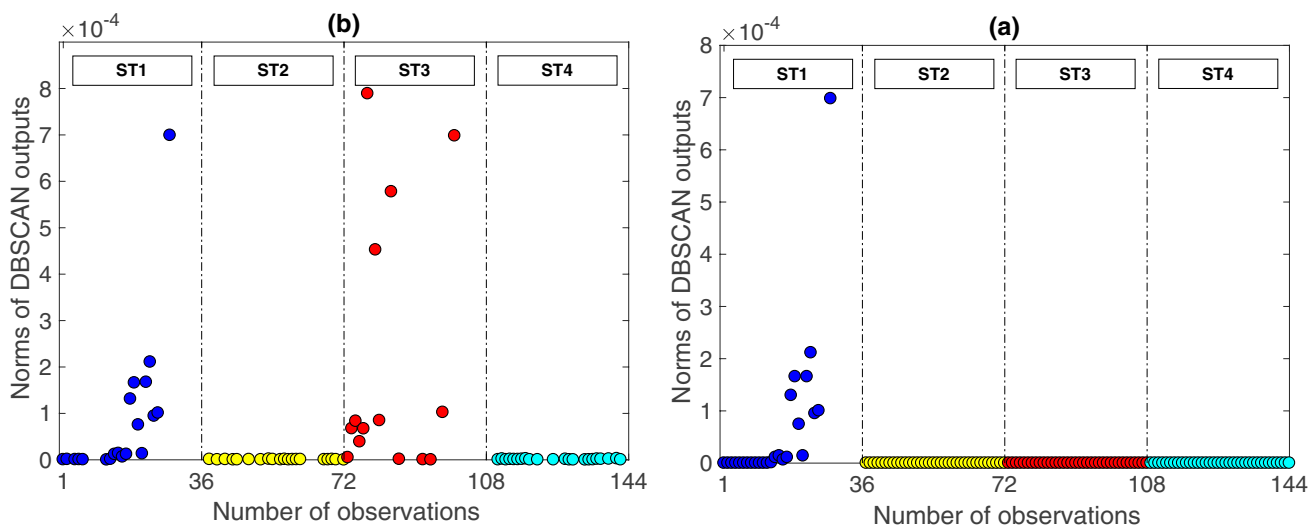


Fig. 17 Damage detection and localization using the DBSCAN method without applying the Kronecker product: **a** Pattern 3, **b** Pattern 4

matrices of these conditions were determined based on the fundamental principle of the model-driven scheme. Using the Kronecker product and the vectorization algorithm, a vector-style feature set was obtained and fed into the *k*-medoids clustering and DBSCAN. Finally, the direct clustering outputs and their Frobenius norms were utilized as indices for damage detection and localization. The effectiveness and performance of the proposed methods were verified numerically by a shear-building model and the ASCE benchmark structure.

Based on the numerical structures, the following conclusions are drawn. (1) The proposed hybrid strategy is an effective tool for extracting a reliable damage-sensitive

feature in a model-driven manner as well as detecting and locating damage using the *k*-medoids clustering and DBSCAN methods on the basis of a data-driven strategy. (2) The modal flexibility discrepancy matrix is sensitive to damage so that the damaged areas have the most reductions. (3) The use of the Kronecker product in providing the main damage-sensitive feature and the input data of the clustering methods increases damage detectability and localizability. This conclusion confirms its positive effect on presenting more obvious results of damage detection and localization compared to the situation without applying it. (4) Both clustering outputs and their Frobenius norms are able to give accurate and obvious results of

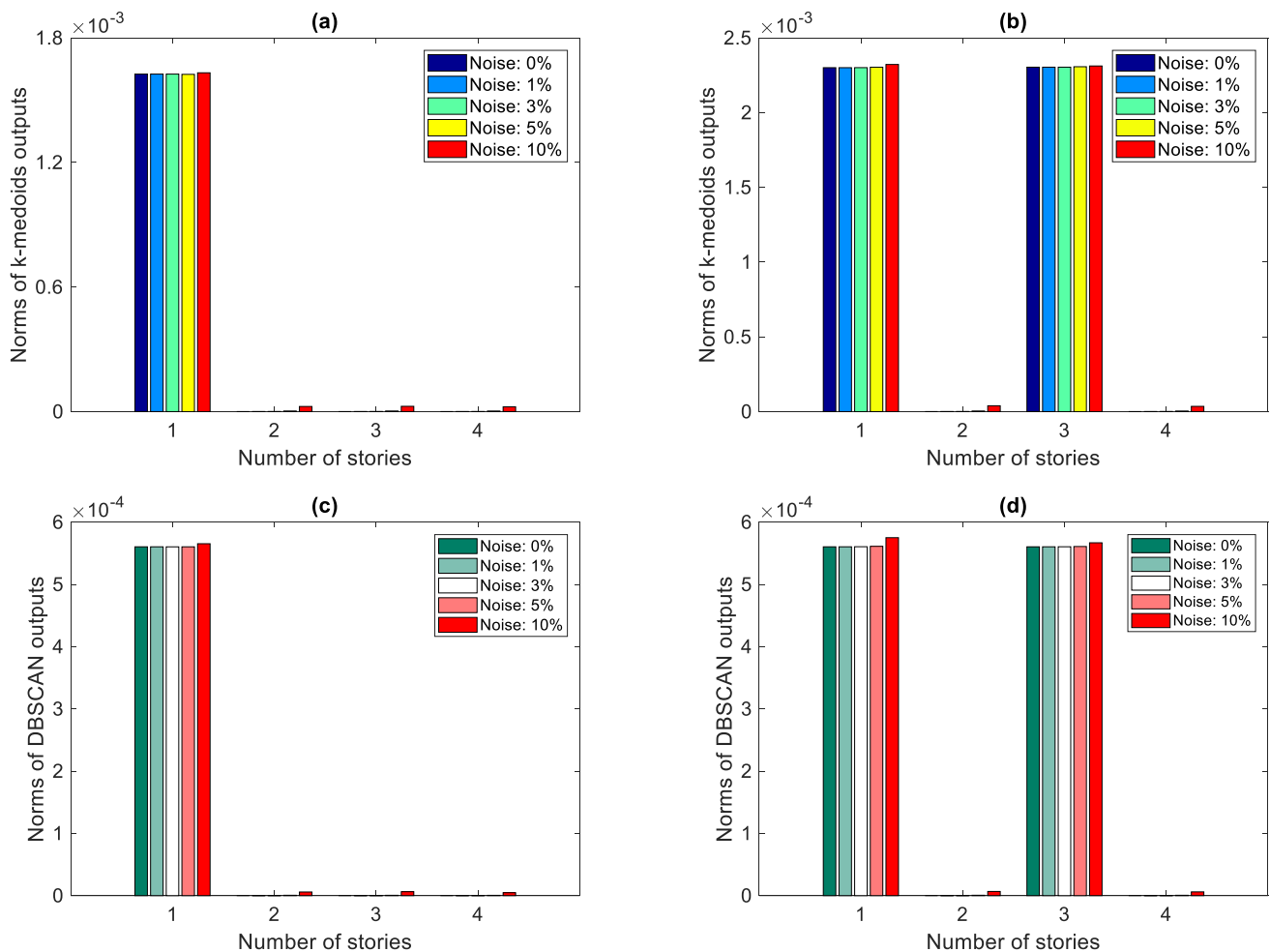


Fig. 18 Damage detection and localization using the Frobenius norms of the clustering outputs in the different noise levels: **a** *k*-medoids clustering in Pattern 3, **b** *k*-medoids clustering in Pattern 4, **c** DBSCAN in Pattern 3, **d** DBSCAN in Pattern 4

damage detection and localization. However, it is recommended to apply the Frobenius norm for locating damage. (5) The DBSCAN method outperforms the *k*-medoids clustering in terms of smaller computational errors in the undamaged areas under noisy conditions.

Compliance with ethical standards

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

References

- Brownjohn JMW, De Stefano A, Xu Y-L, Wenzel H, Aktan AE (2011) Vibration-based monitoring of civil infrastructure: challenges and successes. *J Civ Struct Health Monit* 1(3):79–95. <https://doi.org/10.1007/s13349-011-0009-5>
- Mesquita E, Antunes P, Coelho F, André P, Arêde A, Varum H (2016) Global overview on advances in structural health monitoring platforms. *J Civ Struct Health Monit* 6(3):461–475
- Li H, Ou J (2016) The state of the art in structural health monitoring of cable-stayed bridges. *J Civ Struct Health Monit* 6(1):43–67
- Bukenya P, Moyo P, Beushausen H, Oosthuizen C (2014) Health monitoring of concrete dams: a literature review. *J Civ Struct Health Monit* 4(4):235–244
- Das S, Saha P, Patro S (2016) Vibration-based damage detection techniques used for health monitoring of structures: a review. *J Civ Struct Health Monit* 6(3):477–507
- Farrar CR, Worden K (2007) An introduction to structural health monitoring. *Philos Trans R Soc A Math Phys Eng Sci* 365(1851):303–315
- Entezami A, Shariatmadar H, Ghalehnovi M (2014) Damage detection by updating structural models based on linear objective functions. *J Civ Struct Health Monit* 4(3):165–176. <https://doi.org/10.1007/s13349-014-0072-9>
- Entezami A, Shariatmadar H, Sarmadi H (2017) Structural damage detection by a new iterative regularization method and an improved sensitivity function. *J Sound Vib* 399:285–307. <https://doi.org/10.1016/j.jsv.2017.02.038>

9. Katebi L, Tehranizadeh M, Mohammadgholibeiki N (2018) A generalized flexibility matrix-based model updating method for damage detection of plane truss and frame structures. *J Civ Struct Health Monit* 8(2):301–314. <https://doi.org/10.1007/s13349-018-0276-5>
10. Krishnanunni CG, Raj RS, Nandan D, Midhun CK, Sajith AS, Ameen M (2019) Sensitivity-based damage detection algorithm for structures using vibration data. *J Civ Struct Health Monit* 9(1):137–151. <https://doi.org/10.1007/s13349-018-0317-0>
11. Sehgal S, Kumar H (2016) Structural dynamic model updating techniques: a state of the art review. *Arch Comput Methods Eng* 23(3):515–533
12. Sarmadi H, Karamodin A, Entezami A (2016) A new iterative model updating technique based on least squares minimal residual method using measured modal data. *Appl Math Model* 40(23):10323–10341. <https://doi.org/10.1016/j.apm.2016.07.015>
13. Rezaiee-Pajand M, Entezami A, Sarmadi H (2020) A sensitivity-based finite element model updating based on unconstrained optimization problem and regularized solution methods. *Struct Control Health Monit* 27(5):e2481. <https://doi.org/10.1002/stc.2481>
14. Yin T, Jiang Q-H, Yuen K-V (2017) Vibration-based damage detection for structural connections using incomplete modal data by Bayesian approach and model reduction technique. *Eng Struct* 132:260–277
15. Yuen KV, Beck JL, Katafygiotis LS (2006) Efficient model updating and health monitoring methodology using incomplete modal data without mode matching. *Struct Control Health Monit* 13(1):91–107
16. Sarmadi H, Karamodin A (2020) A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects. *Mech Syst Signal Process* 140:106495. <https://doi.org/10.1016/j.ymssp.2019.106495>
17. Sarmadi H, Entezami A, Daneshvar Khorram M (2020) Energy-based damage localization under ambient vibration and non-stationary signals by ensemble empirical mode decomposition and Mahalanobis-squared distance. *J Vib Control* 26(11–12):1012–1027. <https://doi.org/10.1177/1077546319891306>
18. Entezami A, Sarmadi H, Behkamal B, Mariani S (2020) Big data analytics and structural health monitoring: a statistical pattern recognition-based approach. *Sensors* 20(8):2328. <https://doi.org/10.3390/s20082328>
19. Entezami A, Shariatmadar H, Karamodin A (2019) Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods. *Struct Health Monit* 18(5–6):1416–1443
20. Entezami A, Shariatmadar H (2019) Structural health monitoring by a new hybrid feature extraction and dynamic time warping methods under ambient vibration and non-stationary signals. *Measurement* 134:548–568. <https://doi.org/10.1016/j.measurement.2018.10.095>
21. Figueiredo E, Cross E (2013) Linear approaches to modeling nonlinearities in long-term monitoring of bridges. *J Civ Struct Health Monit* 3(3):187–194
22. Diez A, Khoa NLD, Alamdari MM, Wang Y, Chen F, Runcie P (2016) A clustering approach for structural health monitoring on bridges. *J Civ Struct Health Monit* 6(3):429–445
23. Neves A, Gonzalez I, Leander J, Karoumi R (2017) Structural health monitoring of bridges: a model-free ANN-based approach to damage detection. *J Civ Struct Health Monit* 7(5):689–702
24. Farrar CR, Worden K (2013) *Structural health monitoring: a machine learning perspective*. Wiley, Chichester, United Kingdom
25. Ghorbani E, Buyukozturk O, Cha Y-J (2020) Hybrid output-only structural system identification using random decrement and Kalman filter. *Mech Syst Signal Process* 144:106977. <https://doi.org/10.1016/j.ymssp.2020.106977>
26. Ghannadi P, Kourehli SS (2019) Data-driven method of damage detection using sparse sensors installation by SEREPA. *J Civil Struct Health Monit* 9(4):459–475
27. Duan Z, Yan G, Ou J, Spencer BF (2007) Damage detection in ambient vibration using proportional flexibility matrix with incomplete measured DOFs. *Struct Control Health Monit* 14(2):186–196
28. Li J, Wu B, Zeng Q, Lim CW (2010) A generalized flexibility matrix based approach for structural damage detection. *J Sound Vib* 329(22):4583–4587
29. Sung S, Koo K, Jung H (2014) Modal flexibility-based damage detection of cantilever beam-type structures using baseline modification. *J Sound Vib* 333(18):4123–4138
30. Yan W-J, Ren W-X (2014) Closed-form modal flexibility sensitivity and its application to structural damage detection without modal truncation error. *J Vib Control* 20(12):1816–1830
31. Zare Hosseinzadeh A, Ghodrati Amiri G, Seyed Razzaghi SA, Koo KY, Sung SH (2016) Structural damage detection using sparse sensors installation by optimization procedure based on the modal flexibility matrix. *J Sound Vib* 381(Supplement C):65–82. <https://doi.org/10.1016/j.jsv.2016.06.037>
32. Sarmadi H, Entezami A, Ghalehnovi M (2020) On model-based damage detection by an enhanced sensitivity function of modal flexibility and LSMR-Tikhonov method under incomplete noisy modal data. *Eng Comput*. <https://doi.org/10.1007/s00366-020-01041-8>
33. Aghabozorgi S, Shirkhorshidi AS, Wah TY (2015) Time-series clustering—a decade review. *Inf Syst* 53:16–38
34. Mahato S, Chakraborty A (2019) Sequential clustering of synchrosqueezed wavelet transform coefficients for efficient modal identification. *J Civ Struct Health Monit* 9(2):271–291. <https://doi.org/10.1007/s13349-019-00326-x>
35. da Silva S, Dias Júnior M, Lopes Junior V, Brennan MJ (2008) Structural damage detection by fuzzy clustering. *Mech Syst Signal Process* 22(7):1636–1649. <https://doi.org/10.1016/j.ymssp.2008.01.004>
36. Silva M, Santos A, Santos R, Figueiredo E, Sales C, Costa JC (2017) Agglomerative concentric hypersphere clustering applied to structural damage detection. *Mech Syst Signal Process* 92:196–212
37. Langone R, Reynders E, Mehrkanoon S, Suykens JA (2017) Automated structural health monitoring based on adaptive kernel spectral clustering. *Mech Syst Signal Process* 90:64–78
38. Mottershead JE, Link M, Friswell MI (2011) The sensitivity method in finite element model updating: a tutorial. *Mech Syst Signal Process* 25(7):2275–2296. <https://doi.org/10.1016/j.ymssp.2010.10.012>
39. Izenman AJ (2009) *Modern multivariate statistical techniques: regression, classification, and manifold learning*. Springer, New York
40. Aggarwal CC, Reddy CK (2016) *Data clustering: algorithms and applications*. CRC Press
41. Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, US, vol 34, pp 226–231. <https://asc.library.org/doi/10.1061/%28ASCE%290733-9399%282004%29130%3A1%283%29>
42. Johnson EA, Lam HF, Katafygiotis LS, Beck JL (2004) Phase I International Association of Structural Control-American Society of Civil Engineer structural health monitoring benchmark problem using simulated data. *J Eng Mech* 130(1):3–15
43. Yuen K-V, Au SK, Beck JL (2004) Two-stage structural health monitoring approach for phase I benchmark studies. *J Eng Mech* 130(1):16–33

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