

# Frequency domain decomposition-based multisensor data fusion for assessment of progressive damage in structures

Mehrisadat Makki Alamdari<sup>1</sup>  | Ali Anaissi<sup>2</sup> | Nguyen L. D. Khoa<sup>3</sup> | Samir Mustapha<sup>4</sup>

<sup>1</sup>School of Civil and Environmental Engineering, University of New South Wales, Kensington, Australia

<sup>2</sup>Faculty of Engineering and IT, The University of Sydney, Sydney, Australia

<sup>3</sup>CSIRO, Data61, Eveleigh, Australia

<sup>4</sup>Department of Mechanical Engineering, American University of Beirut, Beirut, Lebanon

## Correspondence

Mehrisadat Makki Alamdari, School of Civil and Environmental Engineering, University of New South Wales, NSW 2052, Australia.

Email: mehri.makki@gmail.com

## Summary

In this paper, we focused on the development and verification of a solid and robust framework for structural condition assessment of real-life structures using measured vibration responses, with the presence of multiple progressive damages occurring within the inspected structures. A self-tuning learning method for structural condition assessment was proposed. Damage sensitive features were extracted using a frequency domain decomposition (FDD) approach to fuse all the measured responses, followed by random projection algorithm for dimensionality reduction. An automatic parameter selection method called Appropriate Distance to the Enclosing Surface (ADES) was used for tuning the classifier parameter. The effect of operational conditions on the robustness of the proposed method was also investigated, and it was realized that application of FDD to extract damage sensitive feature reduces the variation in the results. Promising results in the assessment of damage were obtained based on two comprehensive case studies, which included single and multiple damage scenarios. The contributions of the work are threefold. First, through two comprehensive case studies, we demonstrate that the frequency-based feature from a single sensor might not be adequate enough to detect the progress of damage, even if the sensor is in the vicinity of damage. Second, we show that data fusion using FDD can reliably assess the severity of damage, and finally, we propose a new automated approach for tuning the classifier parameter.

## KEYWORDS

damage identification, severity assessment, structural health monitoring, steel reinforced concrete, steel structures

## 1 | INTRODUCTION

Increasing the life span of our infrastructures while maintaining a high level of safety to withstand operational and environmental loading has generated the interests to many researchers and engineering scientists to consider continuous monitoring methods known as structural health monitoring (SHM). SHM can be a complementary approach to the current nondestructive testing methods, such as eddy currents and ultrasonics,<sup>1</sup> which can only provide local monitoring and are associated with a very high cost. Driven by this incentive, a large number of algorithms were developed during the last two decades that aimed to provide a continuous monitoring using vibration measurements.<sup>2,3</sup>

Many researchers working in the area of SHM have focused extensively on data analysis, including feature extraction and data fusion, in order to identify and assess damage. Methods that rely on pattern recognition or more broadly machine learning have proven to be more favorable due to the fact that they afford more robustness.<sup>4</sup> Machine learning is a powerful technique to learn and infer patterns on the basis of observational evidence. A further division of machine learning algorithm can be made between unsupervised learning and supervised learning. Unsupervised learning finds hidden rules and determines underlying structure using unlabeled set of measured data, whereas in supervised learning, a set of labeled data is utilized for training. In the context of SHM, the label can simply be an indication of healthy or damaged state.<sup>5</sup>

Successful applications of one-class machine learning for anomaly detection and fault diagnosis have been reported in the literature.<sup>6</sup> Yin et al<sup>7</sup> designed a robust one-class support vector machine (OCSVM) to eliminate the influence of outliers to the learned boundary and used it to detect damage in a simulated structure. Yin et al<sup>7</sup> and Mahadevan and Shah<sup>8</sup> proposed an approach for fault detection and diagnosis using OCSVM and support vector machine (SVM)-recursive feature elimination. Further, Yin et al<sup>7</sup> and Mahadevan and Shah<sup>8</sup> used OCSVM to detect damage in a rotating machinery, and the results showed that the performance of the proposed method is superior to the state-of-the-art methods.

OCSVM has a sparse model that allows a quick response to new data, and it also generalizes well to new unseen data. Other techniques such as density and cluster-based techniques have limitations in the SHM applications. For clustering-based approach, it is slow because anomalies are just a by-product result of clustering. Also, it is not straightforward to use a clustering technique for an online application (train and test procedure). Density-based technique like kernel density estimation does not work well for high-dimensional sensing data. Moreover, if not impossible, it is challenging to do severity assessment using density or clustering-based techniques.

In this work, even though existing machine learning techniques are applied, they are combined to form a novel procedure of damage detection and severity assessment. Frequency domain decomposition (FDD) is initially applied to fuse data from multiple sensors<sup>9</sup> and extract damage sensitive features. Random projection algorithm is then used in order to reduce the dimensionality of the data followed by a self-tuning OCSVM for anomaly detection. An automatic parameter selection method called Appropriate Distance to the Enclosing Surface (ADES) is used for tuning the Gaussian kernel parameter. The experiments show that the method can reliably detect the presence of damage, even with a very small damage in the presence of varying operational conditions. Additionally, the presented method can successfully monitor the progress of small damage in the structure. All these results are achieved using only a one-class approach, which is more practical in SHM. To the best knowledge of the authors, there is no existing work on machine-learning related SHM to address the problem of detection and assessment of very small progressive structural damage in a one-class and fully automated way without any user-defined parameters.

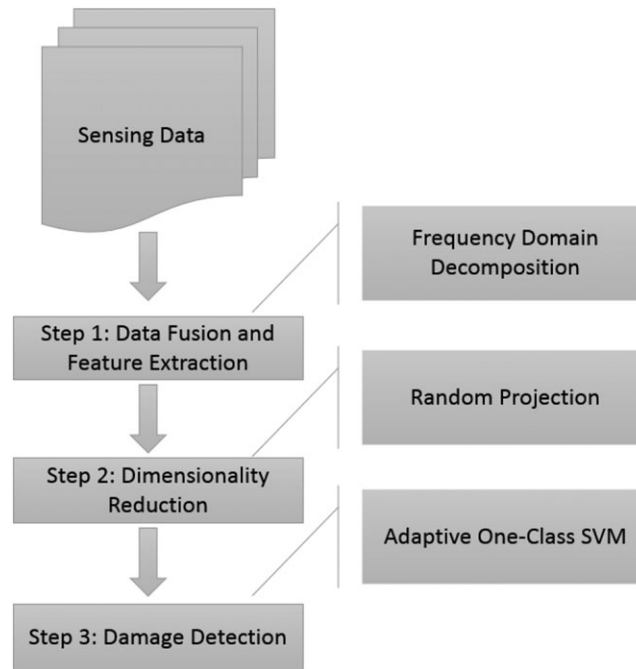
We validate this methodology using two case studies: (a) The first case study is a reinforced concrete structure under progressive cracking inside the concrete. (b) The second case study is a fully fixed steel beam subject to progressive multiple damage with different severity. These two case studies will show the ability of our methodology in detecting and assessing the severity of damage without the need of a supervised learning. In both case studies, vibration measurements will be implemented for learning algorithm.

## 2 | PROPOSED METHOD

This article presents a novel framework to detect and assess the severity of a damage in the structure. Three main steps were involved in this framework as shown in Figure 1. The first step aims to fuse the data and extract features from the sensor networks attached to the structure. The second step applies dimensionality reduction in order to alleviate the curse of dimensionality. The final step builds a one-class classification model to generate health scores for the structure.

### 2.1 | Step 1—Data fusion and feature extraction: FDD

Data fusion and feature extraction plays a critical role in assessing the severity of the detected damage in the structure. FDD was used in this study to fuse data from the sensor networks and extract damage sensitive features. FDD assumed that the vibration responses from  $l$  distinct locations within the structure are available. From a probabilistic point of view, the response process at locations  $p$  and  $q$ , where  $p$  and  $q \in [1:l]$  can be characterized through a correlation function,  $R_{pq}$ , in the time domain as<sup>10</sup>



**FIGURE 1** The framework to detect and assess the severity of damage in the structure

$$R_{pq}(\tau) = E[x_p(t)x_q(t + \tau)], \quad (1)$$

where  $E[\star]$  operator denotes the probabilistic expected value,  $t$  refers to the time, and  $\tau$  is the lag time or the amount of time by which the signal has been shifted. The process can be described in terms of their energy spectra in the frequency domain using a power spectral density (PSD) function, which is the Fourier transform of the correlation function given by

$$S_{pq}(\omega) = \int_{-\infty}^{+\infty} R_{pq}(\tau) \exp^{-i\omega\tau} d\tau, \quad (2)$$

where  $S_{pq}(\omega)$  is the cross PSD of the response at locations  $p$  and  $q$  and frequency  $\omega$ . Once  $p = q$ ,  $S_{pq}(\omega)$  is referred to as the auto-power; otherwise, it is called cross-power.

At each frequency spectra, a symmetric matrix of  $S_{l \times l}(\omega)$  can be populated using an auto- and cross-power information obtained earlier for different pairwise locations. Matrix  $S$  can be decomposed using the singular value decomposition as

$$S(\omega) = U \Sigma U^H, \quad (3)$$

where  $U$  and  $\Sigma$  are  $l \times l$  matrix of singular vectors and diagonal matrix of singular values, respectively, and superscript  $H$  is the conjugate transpose. Singular values are typically in a descending order, and the first singular value is the highest one.

Combining the first singular value obtained at each frequency spectra will result in an  $m$  dimensional vector, which is considered as feature vector  $y$ , where  $m$  refers to the number of spectral lines or attributes. In this work, all frequency spectra has been adopted to construct the feature vector.

## 2.2 | Step 2—Dimensionality reduction: Random projection

Dimensionality reduction is the transformation of a high-dimensional dataset into a lower dimensional one. This transformation is required in this work because we have a low number of observations compared with a large number of features. Principal component analysis (PCA)<sup>11</sup> is a popular dimensionality reduction technique and has been applied

in various domains.<sup>12,13</sup> The main objective of PCA is to determine the directions or components where the data have a maximum variance. This is achieved by finding the eigenvalues and eigenvectors of the covariance matrix of the data. However, PCA has the complexity of  $O(m^3)$  due to the eigendecomposition of the covariance matrix where  $m$  is the dimension of data. This makes it impractical to use for very high-dimensional dataset. Moreover, its performance is sensitive to the number of the selected components. In two-class problems, techniques such as  $k$ -fold cross validation are often used for selecting the optimum number of principal components. However, this method is not applicable for the one-class learning problem because it may overfit the training data.

Random projection<sup>14</sup> technique was used in this study, which is computationally less costly. The rational idea of random projection is to preserve the pairwise Euclidean distances between data points, which is achieved by projecting the high-dimensional data into a random subspace spanned by  $O(\log n)$  columns.<sup>14</sup> Further study carried out by Achlioptas<sup>15</sup> shows that the number of dimensions required for random projection can be calculated using

$$d_{RP} = \log n / \xi^2 \quad (4)$$

where  $d_{RP}$  is the number of dimensions in the low-dimensional space and  $\xi$  is a small positive number with suggested value<sup>15</sup> equal to 0.25.

Given  $X \in \mathbf{R}^{n \times m}$ ,  $\xi > 0$ , and  $d_{RP} = \log n / \xi^2$ . Let  $R_{m \times k}$  be a random matrix where each entry  $r_{ij}$  can be drawn from the following probability distribution:

$$r_{ij} = \begin{cases} +1 & \text{with probability } \frac{1}{2s} \\ 0 & \text{with probability } 1 - \frac{1}{2s}, \\ -1 & \text{with probability } \frac{1}{2s} \end{cases} \quad (5)$$

where  $s$  represents the projection sparsity. With probability at least  $1 - \frac{1}{n}$ , the projection,  $Y = XR$  approximately preserves the pairwise Euclidean distances for all data points in  $X$ .

### 2.3 | Step 3—Damage detection: OCSVM

SVM is a learning process that has a powerful theoretical basis and fast computation. It has been extensively used in both regression and classification problems.<sup>16</sup> For a classification problem, SVM builds a model as a representation of provided training examples and their corresponding binary labels. However, in the context of damage identification, only data instances from one state, that is, healthy state are available, and the samples from other classes, if not impossible, are too difficult or costly to acquire. Consequently, OCSVM<sup>17</sup> was utilized in this work for damage detection. The rational idea of OCSVM is to project one-class data from input domain into the feature space using a kernel function  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ . Then the model constructs a hyperplane that separates the one-class data from the origin with maximum margin.

Let  $X$  be a data point in the feature space extracted from the original sensor data (feature vector), which corresponds to a healthy structure. A hyperplane with maximum margin that separates these data points from the origin must be determined. A feature vector is defined as a vector of  $m$  elements, and each element is called an attribute.

The classification model is a function described by

$$f: \mathbb{R}^m \rightarrow \{-1, +1\} \quad (6)$$

and is written in the form of

$$f(x) = \text{sgn}(w \cdot \phi(x) - \rho), \quad (7)$$

where  $\cdot$  is the dot product.  $w$  and  $\rho$  are the parameters of the model and can be learned from the training data.  $f(x) = +1$ , if  $(w \cdot \phi(x) - \rho) > 0$ , which indicates that the structure is healthy; otherwise,  $f(x) = -1$  which means that the state of the structure has changed.

Given a set of  $n$  training samples,  $X = \{x_i\}_{i=1}^n$ , the training process determines the model parameters  $w$  and  $\rho$  by minimizing the classification error on the training set while still maximizing the margin. Mathematically, it is equivalent to the following minimization problem:

$$\min_{w, \xi, \rho} \frac{1}{2} \|w\|^2 + \frac{1}{\nu n} \sum_{i=1}^n \xi_i - \rho \quad (8)$$

$$s.t \quad w \cdot \phi(x_i) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, n.,$$

where  $\xi_i$  is a slack variable for controlling the amount of training error allowed, and  $\nu \in [0,1]$  is a user-specified variable for controlling the balance between  $\xi_i$  (the training error) and  $w$  (the margin). Furthermore, the problem can be transformed to the dual form using Lagrangian multiplier as

$$\min_{\alpha_1, \alpha_2, \dots, \alpha_n} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \quad (9)$$

$$s.t \quad 0 \leq \alpha_i \leq \frac{1}{\nu n}, \quad \sum_{i=1}^n \alpha_i = 1.$$

This problem can be then solved using quadratic programming.<sup>18</sup> The training dataset,  $X = \{x_i\}_{i=1}^n$ , comprises only of feature vectors, and labeling information is not required. Having obtained a learned model, the decision values for a new data instance  $x$  can be computed as

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i K(x_i, x_{new}) - \rho\right). \quad (10)$$

A negative decision value will indicate an anomaly, which is likely to be damage in this case.

## 2.4 | Self-tuning: Gaussian kernel parameters

Several kernel functions have been used in SVM such as Gaussian and polynomial kernels. However, Gaussian kernel defined in Equation (11) has gained much popularity in the area of machine learning, and it showed a high potential to be an appropriate setting for OCSVM.<sup>19-21</sup> This kernel function has a user defined parameter denoted by  $\sigma$ , which determines the width of the Gaussian kernel and can hysterically affect the performance of the OCSVM. This parameter may yield to overfit the model if it receives small values and underfit in case of large values. Accordingly, an appropriate value for  $\sigma$  should be determined in order to build a reliable generalized OCSVM model.

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad (11)$$

where  $\sigma \in \mathbb{R}$  is the kernel parameter.  $k$ -fold cross validation is often used at a training stage in order to tune  $\sigma$ . However, in case of one-class training, this technique is not possible because it selects the parameter that works only on the training class data and thus lack generalization (overfitting problem). Therefore, alternative approaches have been proposed for tuning this parameter for one class data. The ADES algorithm proposed by Anaissi et al<sup>22</sup> is adopted for tuning the Gaussian kernel parameter based on inspecting the spatial locations of the edge and interior samples and their distances to the enclosing surface of OCSVM. ADES showed successful performances on several datasets and thus was utilized in this work.

Following the objective function  $f(\sigma_i)$  described in Equation 12, the ADES algorithm selects the optimal value of  $\hat{\sigma} = \underset{\sigma_i}{\text{argmax}}(f(\sigma_i))$ , which generates a hyperplane that is maximally distant from the interior samples but close to the edge samples.

$$f(\sigma_i) = \text{mean}(d_N(x_n)_{x_n \in \Omega_{IN}}) - \text{mean}(d_N(x_n)_{x_n \in \Omega_{ED}}), \quad (12)$$

where  $\Omega_{IN}$  and  $\Omega_{ED}$ , respectively, represent the sets of interior and edge samples in the training positive data points identified using a hard margin linear SVM, and  $d_N$  is the normalized distance from these samples to the hyperplane. It is defined as

$$d_N(x_n) = \frac{d(x_n)}{1 - d_\pi}, \quad (13)$$

where  $d_\pi$  is the distance of a hyperplane to the origin described as  $d_\pi = \frac{\rho}{\|w\|}$ , and  $d(x_n)$  is the distance of the sample  $x_n$  to the hyperplane. It is calculated using

$$d(x_n) = \frac{f(x_n)}{\|w\|} = \frac{\sum_{i=1}^n \alpha_i k(x_i, x_n) - \rho}{\sqrt{\sum_{ij} \alpha_i \alpha_j K(x_i, x_j)}}, \quad (14)$$

where  $w$  is a perpendicular vector to the decision boundary,  $\alpha_i$  are the Lagrange multipliers, and  $\rho$  is the bias term. ADES is described in more details in Anaissi et al.<sup>22</sup>

## 2.5 | Training a representative model

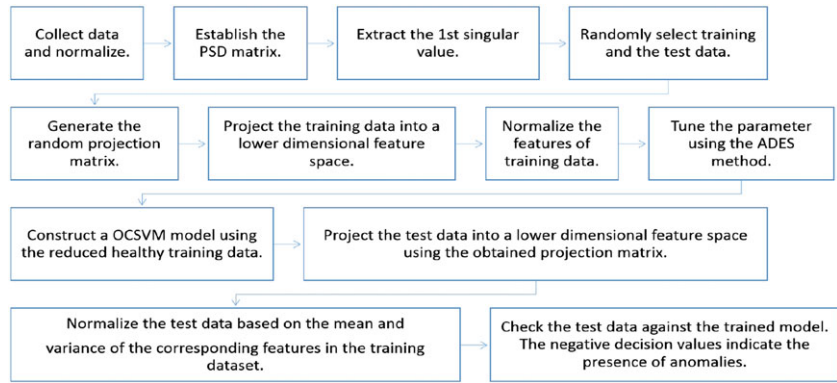
The variation in the healthy data due to changes in environmental and operational conditions may cause false alarms if the learned model is not representative of that change. There are some solutions to resolve this issue, which was also discussed by several researchers in the literature. In the first alternative, a representative training dataset that covers possible variations that a structure might experience is collected. A model is then trained using initial collected data. Due to changes in seasonal or operational conditions, new dataset is collected in addition to the original data. This newly collected data is used to retrain and update the original trained model. To avoid the excessive accumulation of training data, only part of the old training data is stored and used for the model training. In the second alternative, the trained model is adaptively and efficiently updated in time given new information provided by new arriving data. This approach was proposed by Anaissi et al.<sup>23</sup> Further, extracting robust damage sensitive features enhances the performance of the classifier and eventually limits the false alarms. An extensive amount of research in SHM focuses on understanding the effect of temperature and operational changes on the structural responses, in particular on the dynamic behavior. For example, Rainieri et al.<sup>24</sup> proposed the use of Second-Order Blind Identification to model the variability of modal features in presence of changes in environmental and operational conditions. This allows to extract the features that are sensitive to damage but rather insensitive or less sensitive to environmental and operational conditions.

## 3 | CASE STUDIES

In this section, two comprehensive case studies are investigated to illustrate the capability of the proposed method in detecting damage and further assessing its severity. The first case study is a reinforced concrete cantilever beam subjected to increasingly progressive crack that replicates one of the major structural components on the Sydney Harbour Bridge, and the second case study is a two-fixed-end beam subjected to progressive damage at multiple locations.<sup>25</sup> The flowchart of the proposed damage identification and assessment technique, which were carried out for structural condition assessment of the two case studies, has been illustrated in Figure 2.

### 3.1 | Case study I: A reinforced concrete jack arch from the Sydney Harbour Bridge

The first case study is a major structural component from the iconic Sydney Harbour Bridge. There are approximately 800 jack arches distributed over a total distance of 1.2 km in Lane 7; see Figure 3a and 3b. The jack arches are difficult to access and are inspected typically at two yearly intervals according to standard visual inspection practices.



**FIGURE 2** The flowchart of the proposed damage detection and damage severity assessment. ADES: Appropriate Distance to the Enclosing Surface; OCSVM: one-class support vector machine



**FIGURE 3** Illustration of (a) the bus lane on the Sydney Harbour Bridge and (b) one of the concrete jack arches underneath the bus lane

### 3.1.1 | Test setup

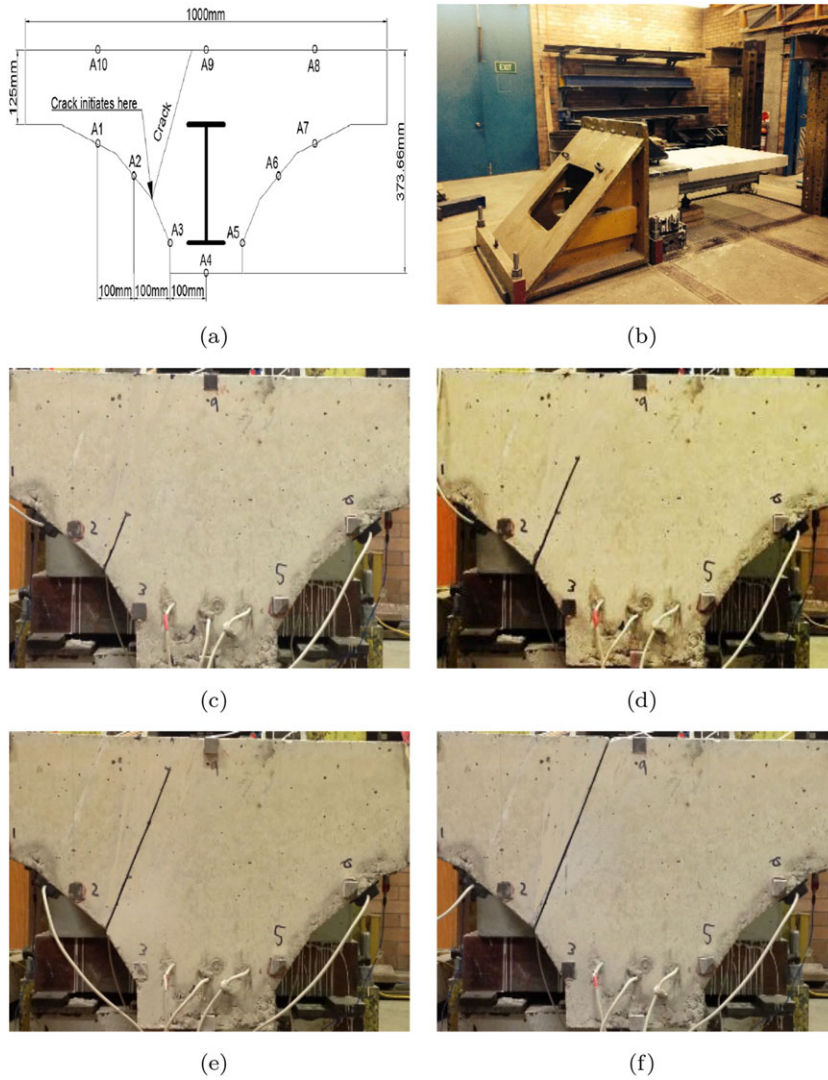
A concrete cantilever beam with an arch section that has a similar geometry to those on the Sydney Harbour Bridge was manufactured and tested, as shown in Figure 4a. The beam consists of a 200UB18 steel I-Beam with a 50-mm concrete cover on both ends. The modulus of elasticity of the steel used for the specimen was 200 GPa. The length of the specimen is 2 m, the width is 1 m, and the depth is 0.375. After pouring the concrete (with Young's modulus of 40 Gpa), the specimen was left to cure for 3 weeks at room temperature before testing. The specimen was fixed at one end using a steel bollard to form a cantilever, where 400 mm along the length of the beam were fully clamped as shown in Figure 4b. In addition, a support was placed at 1,200 mm away from the tip to avoid any cracking occurring in the specimen under its self-weight.

### 3.1.2 | Excitation and response

Ten PCB 352C34 accelerometers were mounted on the specimen to measure the vibration response resulting from impact hammer excitation. Accelerometers were mounted on the front face of the beam. The cross section of the beam and the location of the accelerometers are shown in Figure 4a. A 16-channel NI PCI-6133 data acquisition was used to capture the impact force and the resultant acceleration time histories. The structure was excited using an impact hammer with steel tip, which was applied on the top surface of the specimen just above the location of sensor A9, as shown in Figure 4a. The acceleration response of the structure was collected over a time period of 2 s at a sampling rate of 8 kHz, resulting in 16,000 samples for each event (i.e., a single excitation). A total of 190 impact test responses were collected from the healthy condition.

### 3.1.3 | Introducing damage in the structure

A crack was introduced into the specimen in the location marked in Figure 4a using a cutting saw. The crack is located between sensor locations A2 and A3, and it is progressively increasing towards sensor location A9. The length of the cut was increased gradually from 75 to 270 mm as shown in Figure 4c–f, and the depth of the cut was fixed to 50 mm; a description is provided in Table 1.



**FIGURE 4** The crack introduced into the test specimen: (a) Structure with arrow indicating the cut, (b) support of the structure, (c) Damage Case 1: 75-mm damage cut, (d) Damage Case 2: 150-mm damage cut, (e) Damage Case 3: 225-mm damage cut, and (f) Damage Case 4: 270-mm damage cut

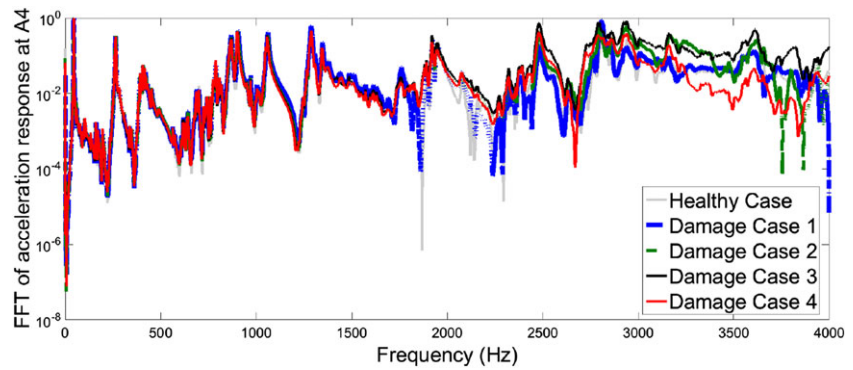
**TABLE 1** Description of the tested cases

Case	Damage description
Benchmark	—
Damage Case 1	75 mm × 50 mm
Damage Case 2	150 mm × 50 mm
Damage Case 3	225 mm × 50 mm
Damage Case 4	225 mm × 50 mm

After introducing each damage case, a total of 190 impact tests were performed on the structure in the location prescribed earlier. The impact of damage was investigated by comparing the frequency response of the structure in healthy and damage cases, as shown in Figure 5. It was observed that the change in the first three natural frequencies between the healthy case and all damage cases is less than 1%, which corresponds to a very small damage. Table 2 compares the natural frequencies for the first three modes in the healthy state and four damage cases, as well as the change in frequency after each damage case relative to the healthy state.

### 3.1.4 | Damage identification results and discussion

Damage identification procedure, shown in Figure 2, was carried out to identify the presence of damage in the structure as well as to assess its progress. This dataset consists of 950 samples (events) separated into two main groups, healthy



**FIGURE 5** Comparison of the frequency response function of response (inertance) between the healthy state and the four damage cases for sensor location A4

**TABLE 2** Comparison of the first three modes of the structure in the healthy state and the four damage cases

$\omega$ (Hz)	Healthy		Damage Case 1		Damage Case 2		Damage Case 3		Damage Case 4	
	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$
$\omega_1$	45.90	—	45.90	0.00%	45.90	0.00%	45.90	0.00%	45.50	0.87%
$\omega_2$	181.6	—	181.4	0.11%	181.2	0.22%	180.8	0.44%	180.0	0.88%
$\omega_3$	265.0	—	264.6	0.15%	264.4	0.23%	264.2	0.30%	262.4	0.98%

state (190 samples) and damaged states (760 samples). Each sample is the measured vibration response of the structure with 8,000 attributes in the frequency domain ( $8 \text{ kHz} \times 2 \text{ s} \times 0.5$  [considering Nyquist frequency]).

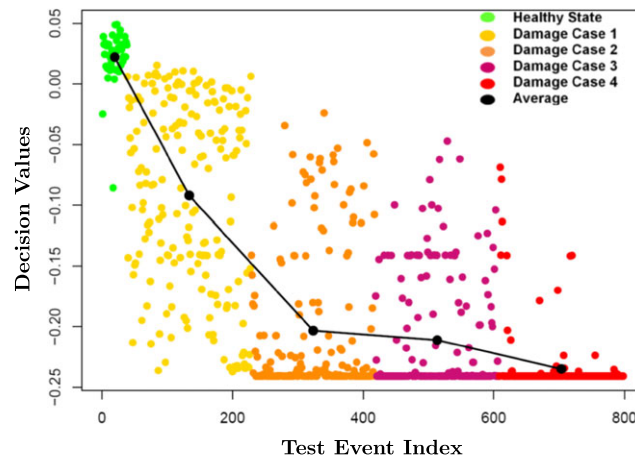
The measured acceleration responses collected from the 10 sensors were utilized to construct the damage sensitive features. Eighty percent of the healthy data were randomly selected for the training stage, whereas the remaining 20% of the healthy samples and all the damage cases were used for testing. The dimension of the data was reduced into 80 attributes using random projection method.

Damage identification results are presented in terms of the number of true positives (TP) and the number of false negatives (FN). TP indicates the number of damaged samples were correctly identified as damage, whereas FN shows the number of damaged samples were incorrectly identified as healthy. From these two numbers, a sensitivity index is calculated based on Equation 15 as

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (15)$$

In the context of SHM, it is important to obtain a sensitivity value of 1 or very close to 1 that indicates the method does not misidentify damage.

Figure 6 shows the obtained decision values for the test samples in healthy and damaged states. The horizontal axis indicates the data instance index or event index, and the vertical axis indicates the magnitude of the decision value. Test samples were grouped into five different categories including healthy state and four different damaged states. Decision values for each category are presented with a different color. The mean of all the decision values for each category was calculated and illustrated in Figure 6. A solid black line was constructed to connect the mean values. Out of 760 damaged samples, 740 were correctly identified as damage with negative decision values, ( $TP = 740$ ), and only 20 of the damaged samples in Damage Case 1 were misidentified with positive decision values, ( $FN = 20$ ). This results in a sensitivity of 0.97, which is a promising result, given the very small size of damage. As can be seen in Figure 6, although no information of damaged events has been employed to construct the OCSVM model and only data from the healthy events have been utilized for the purpose of training, the trained model can successfully identify 97% of damaged events. This illustrates that the method has potential to detect the presence of damage in the structure. Moreover, it can be seen that the method has a capability to identify the progress of the damage in the structure based on the decision values. By increasing the damage severity, the decision values were further decreased (i.e., the data were more deviated from the training data). It should be emphasized that the level of damage in this case study is considerably small with less than 1% reduction in the frequency of the first three modes, as shown in Figure 5 and Table 2. Additionally, the structure is subject to variations in the operational conditions due to high variation in the excitation levels.



**FIGURE 6** Damage identification results using the proposed approach

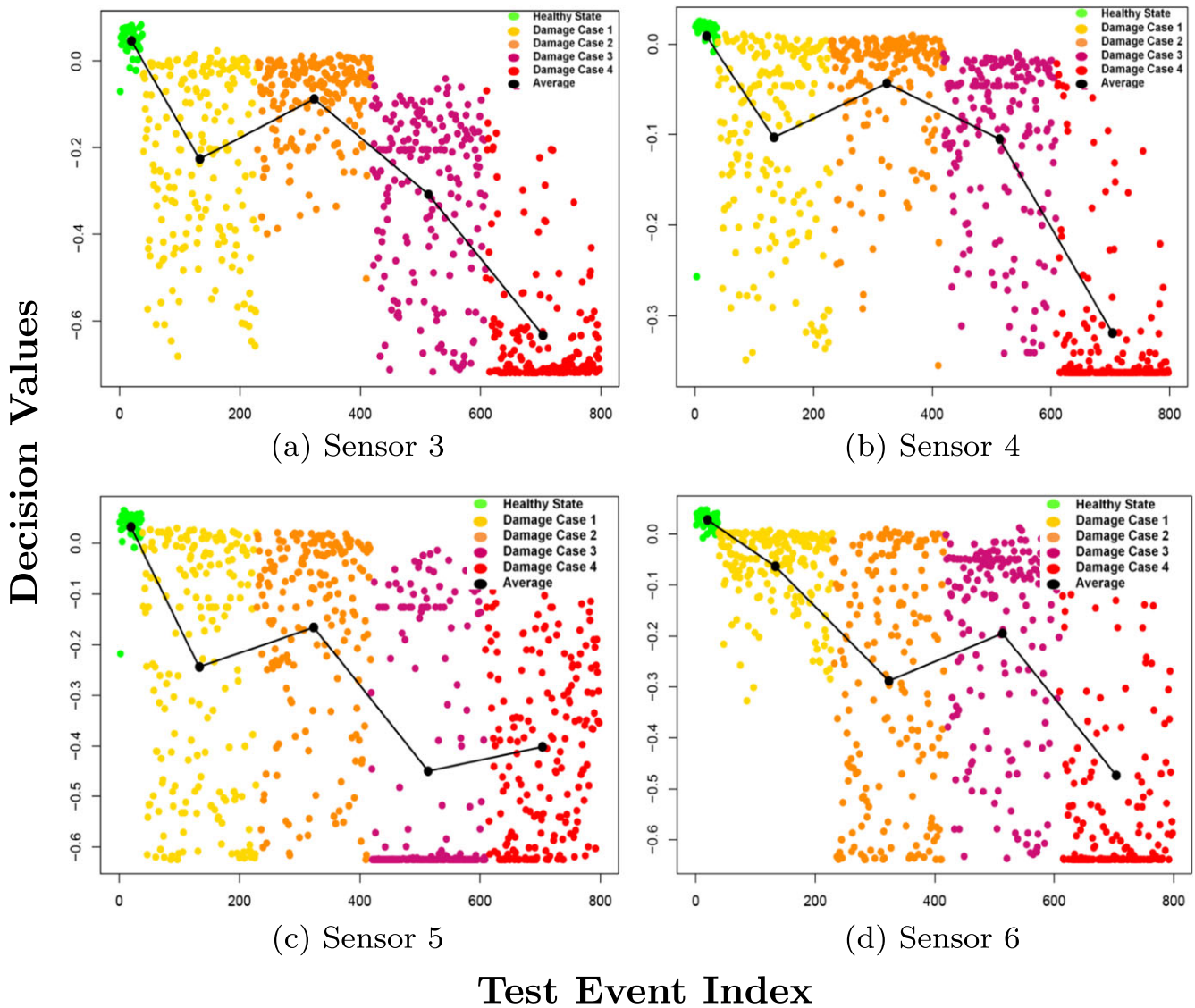
To further investigate the capability of the method, a separate approach was adopted as follows. Instead of using FDD and integrating all sensor information using PSD matrix, only the FFT features of the acceleration response obtained from each sensor were used to construct a separate statistical model for each sensor location using data from the healthy case. Similar damage identification steps were carried out with the only difference that this time, for each sensor location, a separate model was trained and tested. Damage identification results using this approach are presented in Figure 7 for sensors A3, A4, A5, and A6. The obtained sensitivity values for these four sensors are, respectively, 0.94, 0.91, 0.92, and 0.92. Using this approach, not only the sensitivity of the method reduced but also the method failed to monitor the progress of damage. As seen in Figure 7, by increasing damage severity, a decreasing trend in the mean value of decision values is not obtained. Based on this investigation, it can be concluded that compared with the FDD approach, where the information from all sensors are fused, this approach may suffer from the variability in the excitation and lacks the ability to provide reliable information about the presence and severity of damage in the structure.

### 3.1.5 | Impact of sensor array on damage identification results

In order to evaluate the performance and sensitivity of the proposed method to the number of sensors, and to the location of the sensors, an investigation was carried out. Several scenarios were investigated considering only a subset of all sensors to fuse the acceleration responses, using FDD approach, and to generate the feature vector. For each scenario, the proposed damage identification and damage assessment technique was carried out and the obtained decision values together with the number of TP, FN, and the obtained sensitivity index were presented. Five different scenarios were investigated as follows:

- Scenario (1): The responses from all 10 sensors are considered (A1:A10). (This scenario is the same as the one presented in Section 3.1.4).
- Scenario (2): The responses from four sensors in the immediate neighborhood of damage are considered (A2, A3, A9, and A10).
- Scenario (3): The responses from four sensors far from the damage are considered (A5, A6, A7, and A8).
- Scenario (4): The responses from three sensors on the left side of the crack are considered (A1, A2, and A10).
- Scenario (5): The responses from seven sensors on the right side of the crack are considered (A3:A9).

Table 3 lists the number of TP and FN in each scenario together with the obtained sensitivity. Figure 8a–e shows the obtained decision values. From Figure 8a–e, it can be clearly seen that the proposed approach, by fusing responses from multiple sensors, is able to detect progress of damage in the structure in all of the investigated scenarios (more negative decision values are obtained as damage evolves). However, from Table 3, it is realized that the performance and sensitivity of the method is largely affected by the sensor locations. The best sensitivity is obtained for Scenario (4) where only sensors on the left side of the crack are considered. This sensor set up results in the highest sensitivity of 1 with



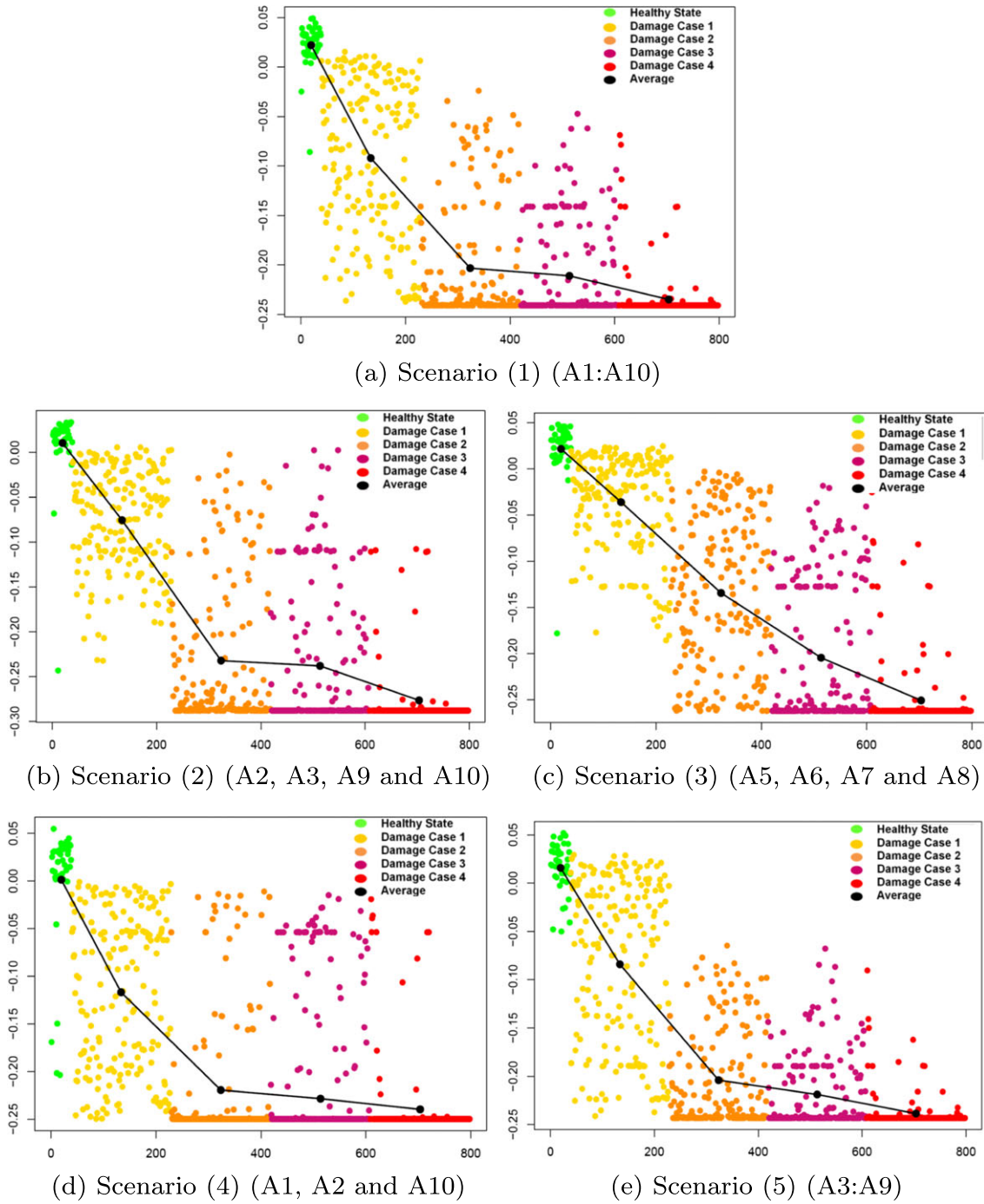
**FIGURE 7** Damage identification results using a separate one-class support vector machine model for each sensor location

**TABLE 3** The number of TP, FN, and sensitivity index for different sensor arrangement

Sensors arrangement	TP	FN	Sensitivity
Scenario (1) (A1:A10)	740	20	0.97
Scenario (2) (A2, A3, A9, and A10)	752	8	0.99
Scenario (3) (A5, A6, A7, and A8)	696	64	0.92
Scenario (4) (A1, A2, and A10)	760	0	1
Scenario (5) (A3:A9)	722	38	0.95

Note. FN: false negatives; TP: true positives.

0 FN (Figure 8d). Although in this scenario only three sensors are involved, it results in the best damage identification. The reason is due to the nature of the excitation (which is happening on top of sensor A9 and in vertical direction) and more importantly, due to the orientation of the crack that allows more movement on the sensors in the left side compared with the sensors on the right side. The second-best identification with a sensitivity of 0.99 is obtained from Scenario (2) where the sensors in the immediate neighborhood of damage are considered. In this scenario, only four sensors are involved, and 8 FN are obtained (Figure 8b). The third-best identification is for the case where all 10 sensors are involved; in this case, a sensitivity of 0.97 with a total of 20 FN are obtained (Figure 8a). It can be seen by involving



### Test Event Index

**FIGURE 8** Damage identification results using different sensor arrangements

more sensors from the right side of damage, the performance of identification deteriorates as the sensors on the right side are less sensitive compared with the sensors in the left side. A sensitivity of 0.95 with a total of 38 FN are obtained for Scenario (5) where all the sensors in the right side of damage are considered (Figure 8e). The lowest sensitivity is for Scenario (3) where only four sensors in the right side of crack and far from damage are considered. In this case, a sensitivity of 0.92 with a total number of 64 FN are obtained (Figure 8c). From this investigation, it can be concluded that by having more number of sensors, the sensitivity of the method does not necessarily increase. However, by having sensors in the immediate neighborhood of damage or at sensitive locations, that is, left side of the crack in this case study, the best sensitivity is obtained. Therefore, location of the sensors is a more critical factor than the number of sensors. In

real-world applications, the most potential damage locations can be identified, that is, through finite element analysis and then sensors are placed in those critical areas.

### 3.2 | Case study II: A steel beam with multiple progressive damage

In the second case study, an experimental test bed was established to further verify the capability of the method in detection and assessment of progressive damage in a steel beam structure.

#### 3.2.1 | Test setup

The test setup is comprised of a two-fixed-end steel beam that has a rectangular cross section with dimension of 1,470 mm in length, 65 mm height, and 8 mm in width. The modulus of elasticity of the steel beam used was 208 GPa. The test setup is illustrated in Figure 9.

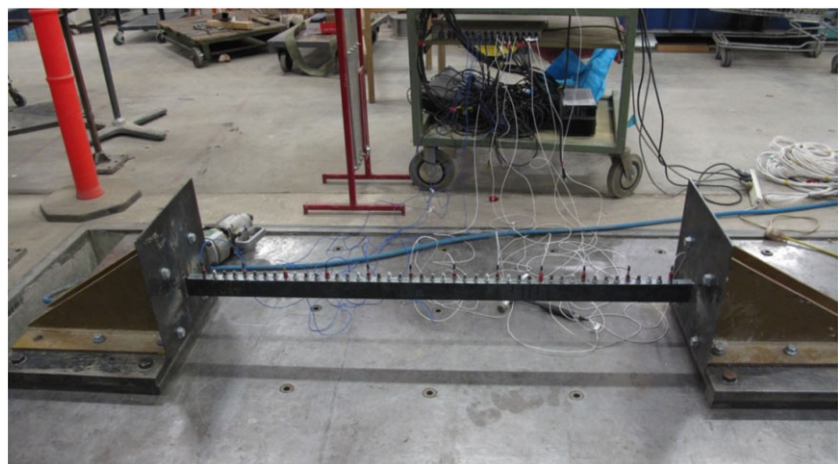
#### 3.2.2 | Excitation and data collection

In order to capture the dynamic behavior of the structure, 12 PCB 352C34 accelerometers were evenly placed along the centerline of the beam to measure the vertical acceleration of the beam with a sampling frequency of 5 KHz. Four different sensor configurations were considered in order to investigate the effect of the sensor location on the damage identification results; these four configurations are labeled as sensor arrangements A, B, C, and D, respectively. Configuration A includes 12 accelerometers whereas configurations B, C, and D include only 11 accelerometers. The schematic of the sensor arrangements is illustrated in Figure 10.

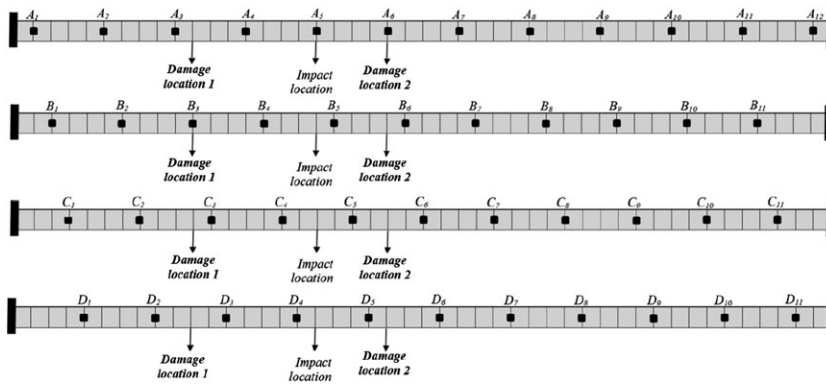
An impact hammer with a rubber tip was used to excite the structure. A rubber tip was chosen because it resulted in a better-quality signals compared with the steel tip. Moreover, at least the first six modes of the structure were captured through the use of the rubber tip. The impact location was identical in all sensor configurations and was set to 40% of the beam span from the left side, as shown in Figure 10. A 16-channel NI PCI-6133 data acquisition system was used to collect time histories from the excitation source and the accelerometers for duration of 30 s following every impact. In this case, the data collection was triggered after a preset threshold on the excitation source was achieved.

#### 3.2.3 | Introducing damage in the structure

The intact steel structure was impacted 10 times at the location prescribed earlier. The impact test was conducted for each sensor configuration A, B, C, and D. As a result of this, testing  $10 \times 4$  data points were created. After testing the structure in the healthy state, damage was progressively introduced in the structure as follows:



**FIGURE 9** The test bed of a two-fixed-end steel beam



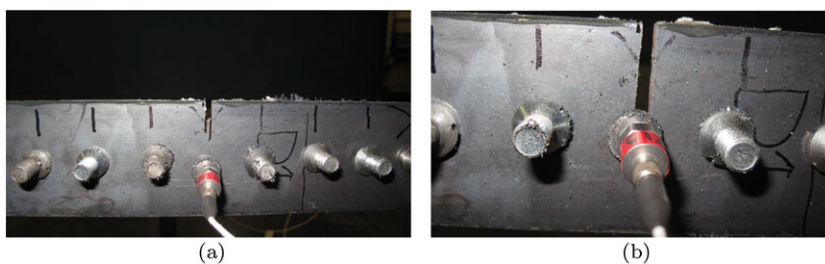
**FIGURE 10** Schematic of the sensor configurations A, B, C, and D

- Damage Case 1: Using a jigsaw machine, a small notch with a depth of 16 mm (25% of the beam height) was induced at about 20% of the beam span. This damage is illustrated in Figure 11a.
- Damage Case 2: Damage Case 1 was increased further by cutting the notch to 32.5 mm as illustrated in Figure 11b.
- Damage Case 3: Damage case 2 is remained the same and a new damage was induced at about 45% of the beam span with a notch depth of 16 mm along the height.
- Damage Case 4: The severity of damage at 45% of the beam span (in Damage Case 3) was increased further to 32.5 mm. Table 4 summarizes these four damage cases.

After introducing each damage case, the dynamic response of the structure was recorded for every sensor configuration (A, B, C, and D) using 10 impact tests with the impact applied at 40% of the beam span. In total, this test resulted in five healthy and damaged cases using four different sensor configurations after 10 impact tests (200 data points were obtained).

In order to experimentally simulate the operational conditions that a structure may undergo in real-life applications, the structure was excited with high level of variability in the excitation magnitude in different impact tests.

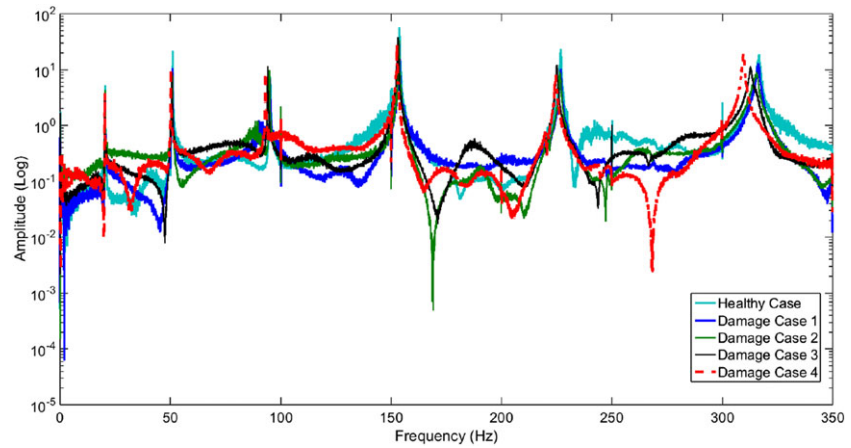
To investigate the impact of the induced damage on the structure, a comparison was made in the frequency domain, using the frequency response function, between the measured responses at sensor location  $A_1$  obtained from the healthy case and four damage cases, as shown in Figure 12. It was observed that the damage effects are more evident at high frequency (the fifth and sixth modes in this case), as the change between the healthy and the damaged structure became more significant. It was realized that the change in the first six natural frequencies between the healthy case and all damage cases was quite small. Table 5 compares the first six modes of the structure in the healthy case and four damage cases. According to the presented results in Table 5, the maximum change for the sixth mode of the structure relative to



**FIGURE 11** The introduced notches in the beam. (a) A notch with a depth of 16 mm (25% of the beam height). (b) A notch with a depth of 32.5 mm (50% of the beam height)

**TABLE 4** Description of the tested cases

Case	Damage description
Benchmark	—
Damage Case 1	Single notch: 16- mm notch placed at 20% of the beam span
Damage Case 2	Single notch: 32.5-mm notch placed at 20% of the beam span
Damage Case 3	Two notches: 32.5-mm notch at 20% of the beam span and 16-mm notch at 45% of the beam span
Damage Case 4	Two notches: 32.5-mm notch at 20% of the beam span and 32.5-mm notch at 45% of the beam span



**FIGURE 12** Frequency response functions (inertance) at sensor location A1 (configuration A) obtained from the healthy state and the four damage cases

**TABLE 5** Comparison of the first six modes of the structure in the healthy case and four damage cases

$\omega$ (Hz)	Healthy		Damage Case 1		Damage Case 2		Damage Case 3		Damage Case 4	
	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$	$\omega$	$\Delta\%$
$\omega_1$	20.53	—	20.50	0.15%	20.47	0.29%	20.30	1.12%	20.28	1.22%
$\omega_2$	51.03	—	50.97	0.12%	50.87	0.31%	50.67	0.71%	50.27	1.49%
$\omega_3$	95.27	—	95.10	0.18%	94.93	0.36%	94.17	1.15%	93.10	2.28%
$\omega_4$	153.8	—	153.7	0.07%	153.6	0.13%	153.0	0.52%	152.6	0.78%
$\omega_5$	226.7	—	226.8	0.04%	226.4	0.13%	225.0	0.75%	224.6	0.93%
$\omega_6$	316.6	—	316.2	0.13%	315.3	0.41%	312.9	1.17%	309.5	2.24%

the healthy state are 0.13%, 0.41%, 1.17%, and 2.24% for Damage Cases 1, 2, 3, and 4, respectively, which demonstrates a very small severity damage in the structure.

### 3.2.4 | Damage identification results and discussion

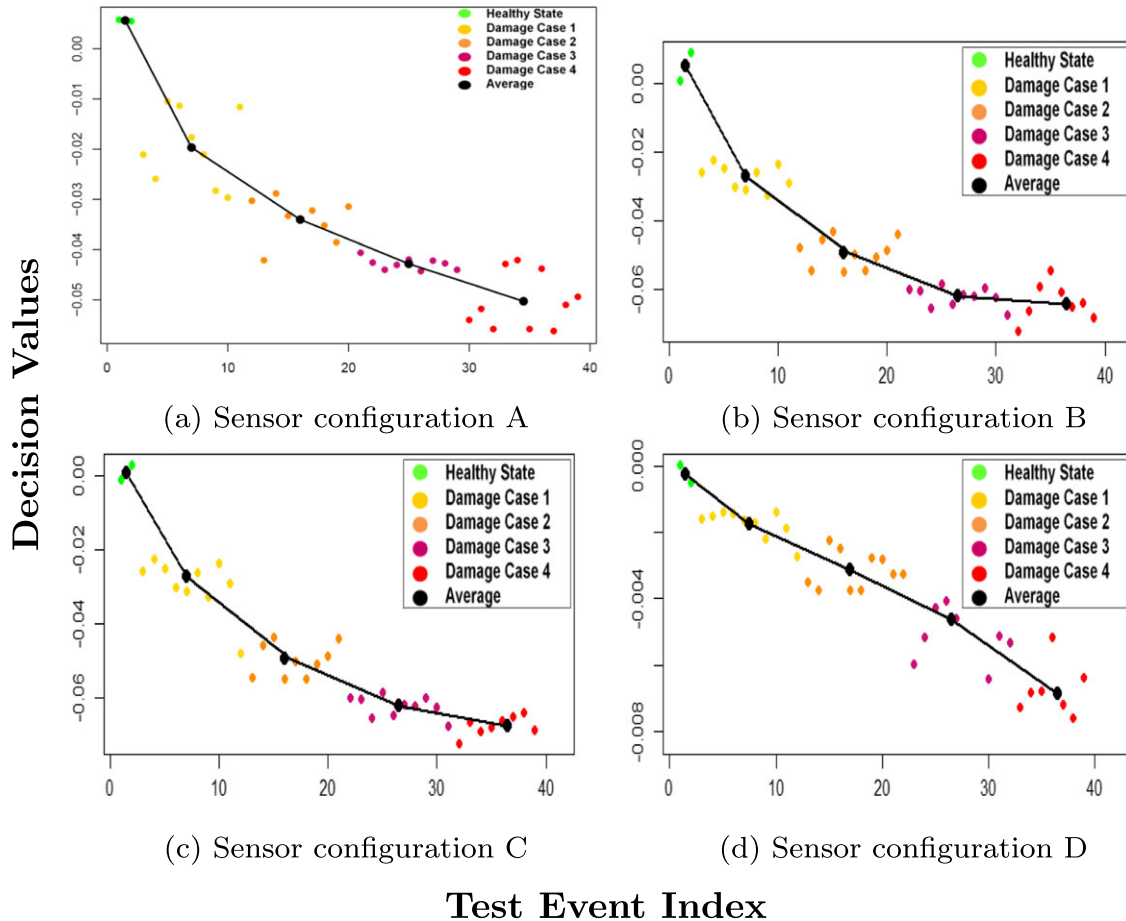
Damage identification process was carried out in the same way that was performed in the previous case study. The damage identification process was repeated for each sensor configuration, A, B, C, and D, separately. This dataset consists of 50 samples, which are classified into two main categories, healthy state (10 samples) and damaged state (40 samples). Each sample is the measured vibration response of the structure with 75,000 attributes in the frequency domain ( $5\text{ kHz} \times 30\text{ s} \times 0.5$  [considering Nyquist frequency]). The damaged state is subcategorized into four different damaged cases containing 10 samples in each.

The measured responses from the 12 sensors (configuration A) and from the 11 sensors (configurations B, C, and D) were utilized to construct the damage sensitive features. The dimension of the data was reduced into 33 attributes using Equation (4).

Eighty percent of the healthy data were randomly selected for training. The remaining 20% of the healthy data was used for testing in addition to the data obtained from the four damage cases.

Figure 13 shows the obtained decision values. As can be seen in Figure 13, although no information of damaged events has been employed to construct the OCSVM model and only data from the healthy events have been utilized for the purpose of training, the trained model can successfully predict all of the damaged events with 0 FN and a maximum sensitivity of 1, regardless the configuration of the sensor locations considered. Further, it can be clearly observed that by increasing the damage severity, the decision values were further decreased (i.e., the data were deviating from the training data). These results again demonstrate the capability of the method in detecting and assessing the progress of damage in the structure based on the decision values.

The same exercise as before was undertaken to illustrate the fact that FDD is less sensitive to conditional changes in the system. To do that, FFT of the acceleration response was utilized as damage sensitive feature, and a separate



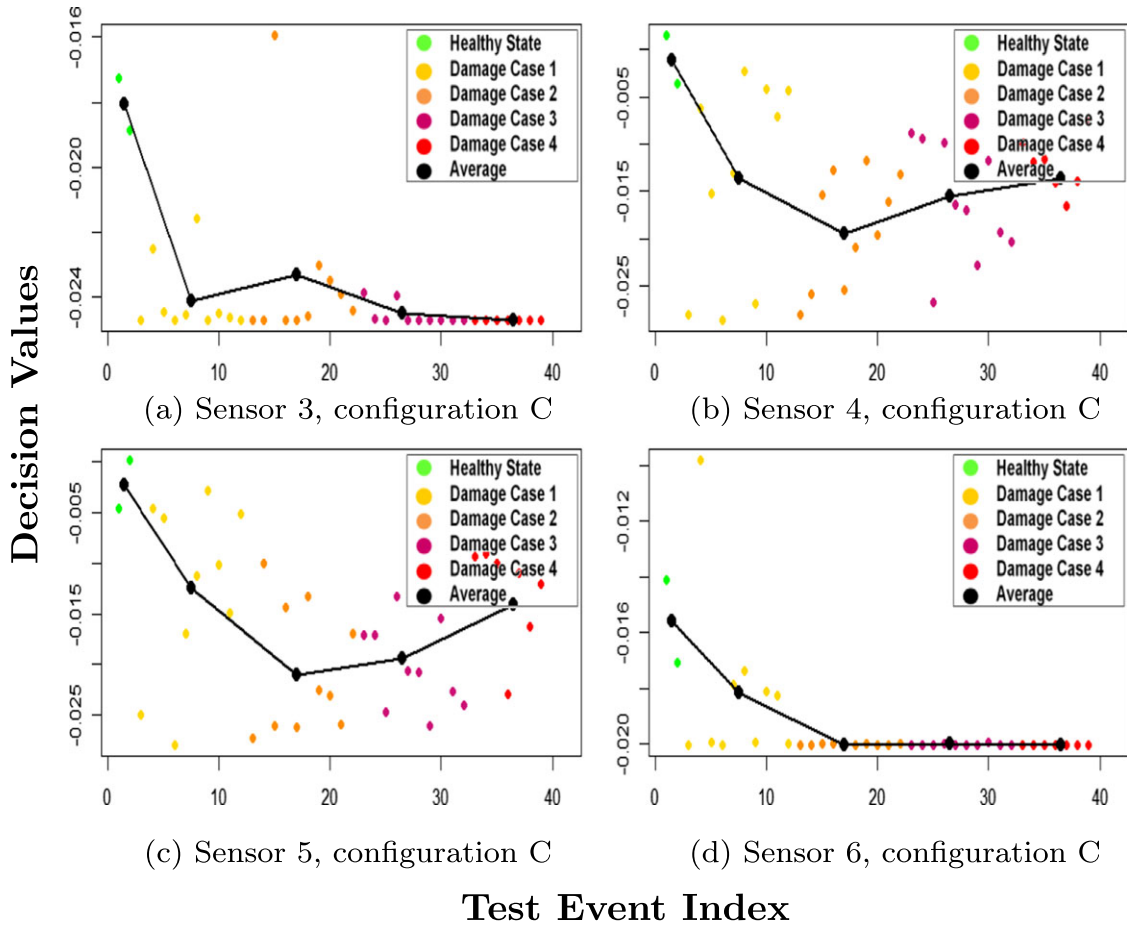
**FIGURE 13** Damage identification results using the proposed framework

OCSVM was constructed for each sensor location. This exercise was repeated for all sensor configurations, and for the sake of space, only the results for Sensors 3, 4, 5, and 6 in configuration C are presented in Figure 14. Based on the results in Figure 14, it can be seen that all the damaged events have been correctly identified, as negative decision values are obtained. This results in a maximum sensitivity of 1 with 0 FN. However, as can be seen in Figure 14, by using the FFT response from one single sensor, the progress of damage in the structure cannot be identified. As illustrated, there is no decreasing trend in the mean value of the decision values as damage progresses. Moreover, it can be clearly seen that using response from individual sensor results in higher number of false positives. It further demonstrates the fact that information provided by individual sensor, even from sensors located very close to damage, is significantly affected by variation in the excitation. Thus, the progress of damage in the structure cannot be identified. Therefore, in order to enhance the robustness of the proposed technique, OCSVM has to use integrated information from different measurement points using FDD.

### 3.2.5 | Impact of sensor array on damage identification results

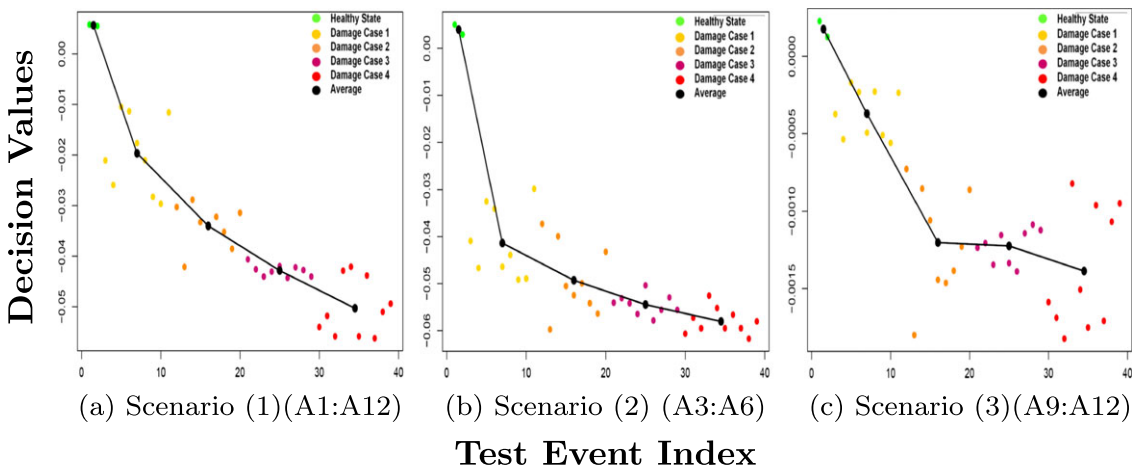
Similar exercise as Section 3.1.5 was performed for the steel beam to further investigate the effect of location of the sensors and the number of sensors in the performance and sensitivity of damage identification. All four configurations (A, B, C, and D) were investigated; however, for the sake of space, only results for configuration A is presented. Three different scenarios are considered as follows:

- Scenario (1): The responses from all 12 sensors are considered (A1:A12). (This scenario is the same as the one presented in Figure 13a).
- Scenario (2): The responses from four sensors in the immediate neighborhood of damage are considered (A3:A6).
- Scenario (3): The responses from four sensors far from damage are considered (A9:A12).



**FIGURE 14** Damage identification results using a separate one-class support vector machine model for each sensor location

Figure 15a–c presents the obtained decision values for these three scenarios. As seen in Figure 15, in all three scenarios, all 37 damaged samples from damaged states were correctly identified ( $FN=0$ ), which results in maximum sensitivity of 1. Additionally, it can be seen that the progress of damage was also successfully identified in all three scenarios. Although, all three scenarios result in maximum sensitivity of 1, the best separation of healthy and damaged states, based on the obtained decision values, happens in Scenario 2. In this scenario, the normalized distance between



**FIGURE 15** Damage identification results using different sensor arrangements

decision values of healthy state and Damage Case 1 (smallest damage) is the furthest, for example,  $\Delta = |d|/|max - min| = 0.66$  (see Figure 2b).  $d$  is the distance between the mean of decision values in healthy state and Damage Case 1, and  $max$  and  $min$  are, respectively, the maximum and minimum of the decision values in each graph to find the range of decision values. The second-best separation is for Scenario 1 with a separation of approximately 0.4 between healthy state and Damage Case 1, and the worst separation belongs to Scenario 3 with a separation of approximately 0.33. It demonstrates that sensors far from damage are likely to misidentify damage, even if large number of sensors are considered. This finding again demonstrates that the location of the sensors is a more effective parameter in the performance and sensitivity of the method than the number of sensors.

## 4 | CONCLUSION

This work presented a robust damage detection and assessment technique based on FDD, random projection, and OCSVM. FDD was initially applied to fuse data from multiple sensors and extract damage sensitive features. Random projection algorithm was then used on the feature data in order to reduce the dimensionality of the data followed by a OCSVM. OCSVM was utilized to build a statistical model using the data from the nominal state of the structure. OCSVM approach is well-suited for real-life applications where data from the damaged state are not available for a supervised learning. Gaussian kernel method was employed in OCSVM, and an automatic parameter selection method called ADES was used for tuning the Gaussian kernel parameter. Two comprehensive case studies were investigated considering different damage scenarios including single and multiple damage states, which were progressively increasing. It was demonstrated that the method has the ability to detect damages with a very small severity in the presence of varying operational conditions. Additionally, it was illustrated that the presented method can successfully monitor the progress of small damage in the structure by providing decreasing negative decision values. The efficiency of the method in detection and assessment of small damage in the structure was compared with another approach where the spectral response at each sensor location was utilized as damage sensitive feature. In this scenario, a separate OCSVM was constructed for each sensor location, and it was demonstrated that compared with the FDD approach, the results were not promising because this approach is very sensitive to varying conditions in the system. These findings indicate that the application of unsupervised learning using OCSVM along with the implementation of FDD can provide a robust method to detect and evaluate the progress of damage, which is of great importance during structural condition assessment. Following the initial work and results presented earlier, we have continued our study of the proposed damage detection and damage assessment method to address some of its limits. More specifically, we have been focusing on the following points:

- Localization of damage using tensor analysis by incorporating temporal, spatial, and feature modes at the same time.
- Minimization of environmental and operational changes using incremental OCSVM.
- Application of the proposed technique on a real-life structure including a cable-stayed bridge.

## ACKNOWLEDGEMENTS

This research was supported by division Data61 within the Commonwealth Scientific and Industrial Research Organization (CSIRO) in Australia. The authors wish to thank the Road and Maritime Services (RMS) in New South Wales and the Center for Built Infrastructure Research (CBIR) at the University of Technology Sydney (UTS) for provision of the support and testing facilities for this research work. Thanks are also extended to Mr. Van Vu Nguyen and Mr. Nick Stephani for their help in collecting and preprocessing the data.

## ORCID

Mehrisadat Makki Alamdari  <https://orcid.org/0000-0001-6587-7493>

## REFERENCES

1. Mutlib NK, Baharom SB, El-Shafie A, Nuawi MZ. Ultrasonic health monitoring in structural engineering: buildings and bridges. *Struct Control Health Monit.* 2016;23(3):409-422.

2. Zhang Q. Statistical damage identification for bridges using ambient vibration data. *Comput Struct*. 2007;85(7):476-485.
3. Makki Alamdari M, Samali B, Li J, Kalhori H, Mustapha S. Spectral-based damage identification in structures under ambient vibration. *J Comput Civ Eng*. 2015;30:4015062.
4. Chen B, Zang C. Artificial immune pattern recognition for structure damage classification. *Comput Struct*. 2009;87(21):1394-1407.
5. Sbarufatti C, Manson G, Worden K. A numerically-enhanced machine learning approach to damage diagnosis using a lamb wave sensing network. *J Sound Vib*. 2014;333(19):4499-4525.
6. Chong JW, Kim Y, Chon KH. Nonlinear multiclass support vector machine-based health monitoring system for buildings employing magnetorheological dampers. *J Intell Mater Syst Struct*. 2014;25(12):1456-1468.
7. Yin S, Zhu X, Jing C. Fault detection based on a robust one class support vector machine. *Neurocomputing*. 2014;145:263-268.
8. Mahadevan S, Shah SL. Fault detection and diagnosis in process data using one-class support vector machines. *J Process Control*. 2009;19(10):1627-1639.
9. Pioldi F, Salvi J, Rizzi E. Refined FDD modal dynamic identification from earthquake responses with soil-structure interaction. *Int J Mech Sci*. 2017;127:47-61.
10. Brincker R, Zhang L, Andersen P. Output-only modal analysis by frequency domain decomposition. In: Proceedings of the international seminar on modal analysis. KU Leuven; 1998, 2001;2:717-724.
11. Jolliffe I. *Principal Component Analysis*, 2nd ed. New York: Springer; 2002.
12. Khoa NL, Zhang B, Wang Y, Chen F, Mustapha S. Robust dimensionality reduction and damage detection approaches in structural health monitoring. *Struct Health Monit*. 2014;13(4):406-417.
13. Anaissi A, Goyal M, Catchpole DR, Braytee A, Kennedy PJ. Case-based retrieval framework for gene expression data. *Cancer Informat*. 2015;14:21-31.
14. Johnson WB, Lindenstrauss J. Extensions of Lipschitz mappings into a Hilbert space. *Contemp Math*. 1984;26(189-206):1.
15. Achlioptas D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *J Comput Syst Sci*. 2003;66(4):671-687.
16. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297.
17. Schölkopf B, Williamson RC, Smola AJ, Shawe-Taylor J, Platt JC. Support vector method for novelty detection. In *Advances in neural information processing systems*. 2000:582-588.
18. Joachims T. Text categorization with support vector machines: learning with many relevant features. In: European Conference On Machine Learning. Springer; 1998; Berlin, Heidelberg:137-142.
19. Chen Y, Zhou XS, Huang TS. One-class SVM for learning in image retrieval. In: Proceedings 2001 International Conference on Image Processing, Vol. 1. Thessaloniki, Greece, Greece: IEEE; 2001:34-37.
20. Kemmler M, Rodner E, Denzler J. One-class classification with Gaussian processes. In: Kimmel R, Klette R, Sugimoto A, eds. *Computer Vision - ACCV 2010. ACCV 2010. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer; 2011:6493.
21. Long J, Buyukozturk O. Automated structural damage detection using one-class machine learning. In: Catbas F., ed. Dynamics of Civil Structures, Volume 4. Conference Proceedings of the Society for Experimental Mechanics Series. Cham: Springer; 2014.
22. Anaissi A, Khoa NLD, Mustapha S, et al. Adaptive one-class support vector machine for damage detection in structural health monitoring. In: Pacific-Asia Conference On Knowledge Discovery And Data Mining. Jeju Island, Korea: Springer; 2017:42-57.
23. Anaissi A, Khoa NLD, Rakotoarivelo T, Alamdari MM, Wang Y. Self-advised incremental one-class support vector machines: an application in structural health monitoring. In: International Conference On Neural Information Processing. Guangzhou, China: Springer; 2017:484-496.
24. Rainieri C, Magalhaes F, Gargaro D, Fabbrocino G, Cunha A. Predicting the variability of natural frequencies and its causes by second-order blind identification. *Struct Health Monit*. 2018. <https://doi.org/10.1177/1475921718758629>.
25. Makki Alamdari M, Li J, Samali B. Damage identification using 2-D discrete wavelet transform on extended operational mode shapes. *Arch Civil Mech Eng*. 2015;15(3):698-710.

**How to cite this article:** Makki Alamdari M, Anaissi A, Khoa NLD, Mustapha S. Frequency domain decomposition-based multisensor data fusion for assessment of progressive damage in structures. *Struct Control Health Monit*. 2019;26:e2299. <https://doi.org/10.1002/stc.2299>