

# Chapter 20

## Structural Health Monitoring Using Machine Learning Techniques and Domain Knowledge Based Features



Nguyen Lu Dang Khoa, Mehrisadat Makki Alamdari,  
Thierry Rakotoarivelo, Ali Anaissi and Yang Wang

**Abstract** Structural Health Monitoring (SHM) is a condition-based maintenance technology using sensing systems. In SHM, the use of domain knowledge is essential: it motivates the use of machine learning approaches; it can be used to extract damage sensitive features and interpret the results by machine learning. This work focuses on two SHM problems: damage identification and substructure clustering. Our solutions to address them are based on machine learning techniques and robust feature extraction using domain knowledge. In the first problem, damage sensitive features were extracted using a frequency domain decomposition, followed by a robust one-class support vector machine for damage detection. In the second problem, a novel clustering technique and spectral moment feature were utilised for substructure grouping and anomaly detection. These methods were evaluated using data from lab-based structures and data collected from the Sydney Harbour Bridge. We obtained high damage detection accuracies and were able to assess damage severity. Furthermore, the clustering technique was able to group substructures of similar behaviour and detect spatial anomalies.

---

N. L. D. Khoa (✉) · T. Rakotoarivelo · Y. Wang  
DATA61, CSIRO, 13 Garden Street, Eveleigh, NSW 2015, Australia  
e-mail: khoa.nguyen@data61.csiro.au

T. Rakotoarivelo  
e-mail: thierry.rakotoarivelo@data61.csiro.au

Y. Wang  
e-mail: yang.wang@data61.csiro.au

M. Makki Alamdari  
School of Civil and Environmental Engineering, University of New South Wales,  
Kensington, NSW 2052, Australia  
e-mail: m.makkialamdari@unsw.edu.au

A. Anaissi  
School of IT, The University of Sydney, Sydney, NSW 2006, Australia  
e-mail: ali.anaissi@sydney.edu.au

## 20.1 Introduction

Most structural and mechanical system maintenance is time-based, i.e. an inspection is carried out after a predefined amount of time. Structural health monitoring (SHM) is a condition-based approach to monitor infrastructure using sensing systems. SHM systems promise significant safety and economic benefits [21], and thus they have been the focus of several studies and activities with sometime real-world deployments [21, 24, 62].

One of the key problems in SHM is damage identification, which can be classified into different levels of complexity [49]:

- Level 1 (Detection): to detect if damage is present in the structure.
- Level 2 (Localisation): to locate the position of the damage.
- Level 3 (Assessment): to estimate the extent of the damage.
- Level 4 (Prediction): to give information about the safety of the structure, e.g. remaining life estimation. This level requires an understanding of the physical damage progression in the structure.

A typical engineering approach in SHM adopts a physic-based model of the structure, usually based on finite element analysis. The differences between measured data and the data generated by the model are used to identify any damage [18]. However, a numerical model may not always be available in practice and does not cater well to uncertainties due to changes in environmental and operational conditions. This challenge motivates the use of a data-driven approach which establishes a model by learning from measured data and then makes a comparison between the data model and new measured responses to detect damage. This approach normally uses techniques in machine learning [62].

Farrar and Worden defined the SHM process in terms of a four-step statistical pattern recognition paradigm [21]: (1) operational evaluation; (2) data acquisition, normalisation and cleansing; (3) feature extraction and information condensation; (4) statistical model development. Among the four, feature extraction and information condensation in Step 3 is an important step to help the statistical modelling using machine learning in Step 4 to identify damage.

Feature extraction is a process of extracting meaningful indicative information from the measured response to determine the structural health state of the system and identify the presence, location and severity of any possible damage. Features may or may not have explicit physical meaning. However, the features that represent the underlying structural physic are preferred for SHM from the point of view that they can provide more effective insight into the condition of the structure. An ideal feature should be sensitive to damage and correlated with the severity of damage but insensitive to environmental and operational effects. The reason is that in real-world SHM applications the effect of environmental and operational changes on features might camouflage damage-related changes and also alter the correlation between the magnitude of changes in the features and associated damage levels [51], and this is one of the main challenges in SHM [21].

All the aforementioned challenges highlight the role of domain experts in solving SHM problems and in this chapter domain knowledge is used in all stages of the data analysis. First, domain knowledge shows data-driven machine learning approaches are suitable for forming an SHM problem. Second, it shows robust feature extraction techniques using domain knowledge are essential in order to extract damage sensitive features. Last, domain knowledge is also used to explain the results found by machine learning techniques.

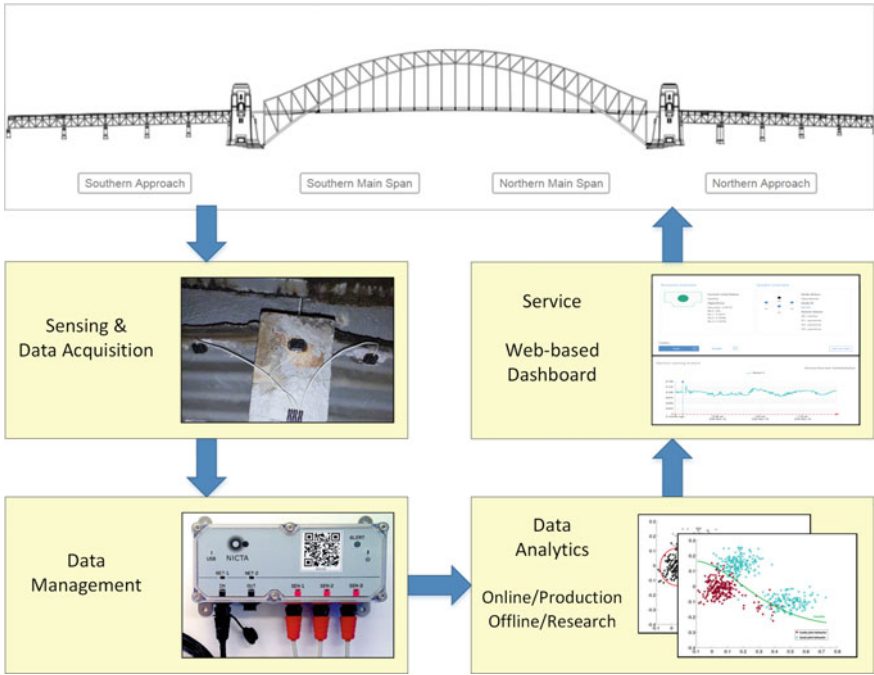
This work is part of our ongoing efforts to apply data driven SHM to the Sydney Harbour Bridge (SHB), one of the iconic structures in Australia. We tackle two different problems faced by a civil infrastructure: damage detection and substructure clustering. Our approaches to these problems are based on machine learning techniques and robust feature extraction using domain knowledge. The first problem is identifying damage in components of a structure over time. In this case, we fused and extracted damage sensitive features from multiple sensors using a frequency domain decomposition (FDD), and then applied a novel self-tuning one-class support vector machine (SVM) for damage detection. The second problem is detecting similar characteristics of a structure's components by comparing and grouping them across locations. In this case, we extended a robust clustering technique and utilised a novel spectral moment feature for substructure grouping and anomaly detection. These methods were evaluated using data from controlled lab-based structures and data collected from a real world deployment on the SHB.

The remainder of this chapter is organised as follows. Section 20.2 provides information about the SHM system of the SHB. Section 20.3 presents a review on feature extraction and fusion in SHM, which is based on domain knowledge. Then the proposed approaches to extract features, to identify damage and to group substructures are introduced in Sect. 20.4. Section 20.5 presents the results of our proposed techniques in two case studies. Finally, there are concluding remarks in Sect. 20.6.

## 20.2 A Large Scale SHM on the Sydney Harbour Bridge

The SHB supports eight lanes of road traffic and two railway lines. Lane 7 on its eastern side is dedicated to buses and taxis. This lane is supported by 800 concrete and steel jack arches, which may develop cracks due to the ageing of the structure and traffic loadings on the lane. It is critical to detect such a deterioration as early as possible. However, they are currently visually inspected once every two years and some locations are difficult to access.

We have developed and deployed a SHM system on the SHB which acquires, integrates, and analyses a large amount of data from about 2400 sensors distributed underneath Lane 7 of the infrastructure [48]. Our SHB system is composed of four layers, as described in Fig. 20.1. First at the Sensing and Data Acquisition layer, we have deployed three tri-axial accelerometers on each of the 800 jack arches. These



**Fig. 20.1** Overview of the SHM system deployed on the SHB

sensors are low-cost MEMS (Microelectromechanical systems) and they record the vibrations of the structure.

At the Data Management layer, we have smart nodes and gateways, which concentrate the data from the sensors. Vibration data are captured at 250Hz from the three sensors on a given jack arch, when a vehicle drives over it. Each node also collects continuous ambient vibration at midnight for 10 min at 1500Hz. The data are transmitted and used by the next Data Analytics layer.

At the third Data Analytics layer, we can deploy several algorithms to derive actionable information from the data. Some algorithms are online and in production, i.e. they operate on real-time data to produce information for the bridge manager and engineers. Other algorithms are offline and in research phase, i.e. they operate on past collected data for a research purpose.

Finally at the Service layer, we developed a secure web-based visualisation dashboard, which allows the bridge manager and engineers to monitor all the jack arches in real time so that they can optimise the maintenance schedule.

## 20.3 Feature Extraction Using Domain Knowledge: A Review

As a result of damage occurrence in the structure, the physical characteristics of the structure (e.g. stiffness, mass or damping) change, which consequently induces a change to the dynamic response [39]. Therefore, one of the key factors in a successful implementation of any vibration-based SHM technique is an appropriate selection of damage sensitive feature from the measured vibration response of the structure [55]. The efforts of previous researchers have been directed to damage sensitive features in modal domain [20], frequency domain [38], time domain [11] and time-frequency domain [43].

Examples of the early features introduced and adopted for SHM applications are modal parameters (e.g. natural frequencies [50], damping [14], and mode shapes), and their derivatives such as modal strain energy [53] and flexibility matrix [44]. Although successful applications of these features have been widely reported in the literature (as discussed in [7]), the use of modal-based features to identify damage in real-world applications has been highly debated in the last few years. Modal-based features are suffering from several problems. Firstly, they are not broadband data and they only provide information at limited frequency resonances. Secondly, they are error prone by nature as they are not directly-measured data and thus complicated modal analysis should be carried out to extract these features from the measured time responses, which may lead to computational errors [40]. Moreover, in real-world applications, it is not possible to capture a complete set of modal parameters from the measurements because only a limited number of lower modes are measured and the information related to higher modes, which is more sensitive to minor changes in the structural integrity, is missed. Finally, it has been demonstrated that modal parameters and in particular natural frequencies are quite sensitive to environmental changes, which is not desirable [45]. These major shortcomings make modal-based approaches less suitable for practical applications.

SHM schemes based on time-domain features have also attracted attention in recent years since no domain transformation is required, which leads to faster monitoring applications [11]. In such a case, damage identification is directly sought based on discrepancies of the measured responses in time domain. Basically, time domain-based features can be treated as data-based features rather than physics-based features and the adopted features might not have an explicit physical meaning. Damage is identified by comparison of a current characteristic quantity with its baseline in a statistical sense. Statistical properties of a time series (e.g. mean and variance) were amongst the earliest statistical frameworks employed for monitoring the acceleration measurements in order to identify data that are inconsistent with the past data (e.g. undamaged state) [22]. Features based on autoregressive models have also been adopted in various SHM applications [54]. In this regard, features are either based on the residues between the prediction from an autoregressive model and the actual measured time history at each time interval, or they are simply based on autoregressive model coefficients [63].

Frequency-based features such as power spectral density (PSD) [34], frequency response functions and their derivatives [33] can be derived from the response in the frequency domain. Unlike modal parameters, frequency data are broadband data which contain a wide range of frequencies [2]. Spectral-based methods in the frequency domain have become another alternative to extract features in mechanical components under stochastic loadings [8]. Applications of spectral methods in the context of damage detection have been found in the literature [5]. Spectral-based methods use spectral moments which can be evaluated directly from the PSD of time responses. Spectral moments represent some major statistical properties of a stochastic process; for example, the variance of a random process is the zero-order spectral moment of that observation [46]. Spectral moments are useful for characterisation of non-Gaussian signals buried in a Gaussian background such as noisy environment [59]. The early efforts in this field were conducted by Vanmarcke to estimate modal parameters (natural frequency and damping) from ambient response measurements of dynamically excited structures [60]. Zero, first and second moments were applied to identify modal parameters. Later on, some researchers used spectral moments to predict the fatigue damage evaluation and estimate the rate of damage accumulation in structures subjected to random processes [8]. Several researchers have applied higher order spectral moments such as spectral kurtosis of the time series data for health assessment of rotary structures [5].

Further, features can be extracted by time-frequency analysis of the measured response using wavelet analysis [43]. Wavelet transform has emerged as a powerful tool for capturing changes in structural properties induced by damage. Wavelet analysis allows the study of local data with a “zoom lens having an adjustable focus” to provide multiple levels of details and approximations of the original signals. Therefore, transient behaviour of the data can be retained [23]. Wavelet analysis not only can detect any subtle differences in the signals but also can localise them in time, and therefore it is quite useful for studying non-stationary systems. Promising applications of wavelet transform approaches to SHM have been reported in the literature [32, 58].

In addition to feature extraction from one single sensor, data fusion which is the process of integrating information from multiple sensors, needs to be considered. An appropriate fusion process can reduce imprecision, uncertainties and incompleteness and achieve more robust and reliable results than a single source approach [26, 57]. Various data fusion methods have been used in SHM [37, 56]. Fusion can be executed in three levels: data-level fusion, feature-level fusion, and decision-level fusion [35]. In data-level, raw data from multiple sensors are combined to produce new raw data that are expected to be more informative than data from a single sensor. In feature-level, features obtained from individual sensors are fused to obtain more relevant information [26]. Data fusion in feature-level can be performed in an unsophisticated manner by simply concatenating features obtained from different sensors. However, more advanced methods including Principle Component Analysis (PCA), neural networks and Bayesian methods have been adopted at this level. Fusion at decision-level can be achieved through various techniques such as voting or fuzzy logic to obtain an ultimate decision based on each decision obtained from individual sensors.

In this study, we adopt a spectral-based approach using the concept of spectral moment to extract the damage sensitive feature from the measured acceleration response. Spectral moment correlates to the energy of the signal in the frequency domain and is computed from the PSD of a signal. Moreover, we also adopt a feature extraction and data fusion approach using FDD to integrate frequency data from multiple sensors. The next section describes in detail our feature extraction and fusion methods.

## 20.4 Damage Identification and Substructure Grouping

In this section, we discuss how domain knowledge is used to phrase a general SHM problem as a machine learning problem and the importance of domain knowledge for feature extraction. Then two typical problems faced by a civil infrastructure are presented: damage detection and substructure clustering. We propose solutions for these two problems which utilise machine learning techniques and robust features extracted using domain knowledge. Specifically, FDD is used with a self-tuning one-class SVM for damage identification; and a spectral moment feature is used with  $k$ -means— for substructure grouping.

### 20.4.1 *Machine Learning Approach for SHM Using Domain Knowledge*

Any change in the structural integrity reflects the vibration characteristic, e.g. natural frequency of the structure. In the context of vibration-based SHM, the main objective is thus to identify any change in these characteristics with respect to a benchmark state. To achieve this, either a physics-based model of the structure or a statistical-based model of the system under study is developed to build a representative model of the structure in the benchmark state. In the first approach, finite element method and optimisation techniques are adopted to establish and calibrate a numerical model of the structure. Future measured response of the structure is then compared with the numerical model prediction to identify any potential change in the system. Although this approach is capable of providing additional useful information about any potential change in the structure, e.g. location and severity, its capability is quite limited to small scale structures in a controlled environment. The main reason is that obtaining a detailed, reliable and calibrated model of the structure is not straightforward, especially in the case of large infrastructures and in the presence of practical uncertainties.

In contrast, a data-based or machine learning model relies solely on measured data. The massive data obtained from monitoring are transformed into meaningful information using domain knowledge as reviewed in Sect. 20.3. It is a more promising

alternative for real-world SHM applications. Not only is establishing the model more straightforward, but also it is capable of overcoming problems associated with environmental and operational variability in SHM since the measured data from many different conditions can be employed for learning the model, which is not the case for a physics-based approach.

Most of the vibration-based SHM techniques require both input and output signals in order to identify possible structural damage. This technique is applied only to small and moderate sized structures and often requires disruption of traffic and human activities for structures under in-service condition. These drawbacks make this approach less practical, specifically in the case of large infrastructures. In contrast, methods based on output-only dynamic test where the structure is excited by natural or randomly varying environmental excitations such as traffic, winds, waves or human movements are more practical for SHM applications. In this approach, structural integrity assessment is performed based on only response measurement data without any knowledge of the input driving forces. Hence, a smaller number of operators and equipment is required, which makes this approach more attractive over measured input vibration. In order to extract the vibration characteristics of the structure, a special procedure named output-only modal identification needs to be considered [41]. It highlights the role of domain knowledge experts in extracting the most characteristic features from the measured response. In the following sections, two different features have been employed based on the domain knowledge about output-only modal identification.

## 20.4.2 Damage Identification

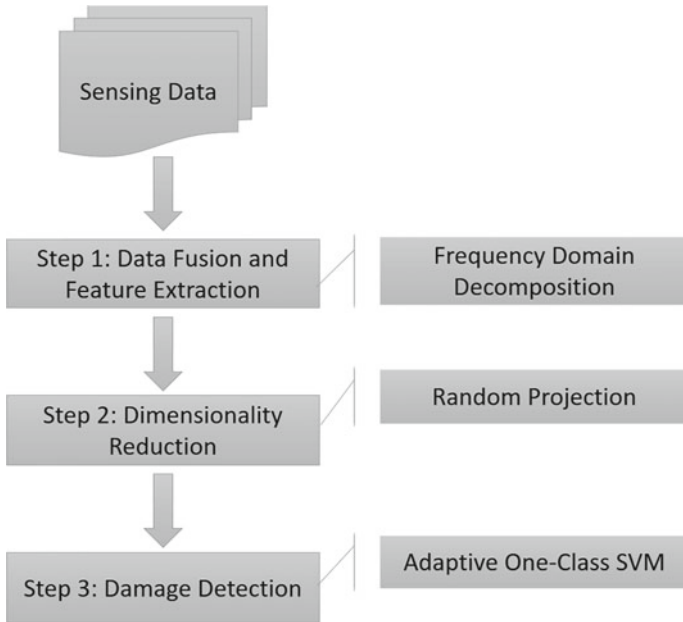
This section presents an approach to identifying damage in components of a structure over time. A flowchart of the approach is shown in Fig. 20.2. First, damage sensitive features are extracted using FDD followed by a dimensionality reduction using random projection. Then an adaptive (self-tuning) one-class SVM is used on the reduced dimensional space for damage detection.

### 20.4.2.1 Data Fusion and Feature Extraction: Frequency Domain Decomposition

FDD was used in this study to fuse data from a sensor network in a data-level. FDD assumes 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  $q$  ( $p$  and  $q \in [1 : l]$ ) can be characterised through a correlation function,  $R_{pq}$ , in the time domain as [10],

$$R_{pq}(\tau) = E[x_p(t)x_q(t + \tau)] \quad (20.1)$$





**Fig. 20.2** The flowchart of the proposed damage detection and severity assessment

where  $E[\ ]$  and  $\tau$  are, respectively, the probabilistic expected value operator and the lag operator.  $R_{pq}(\tau)$  function defines how a signal is correlated with the other, with a time separation  $\tau$ .

The frequency characterisation of such a random stationary process can be computed using the PSD function which is calculated by taking the Fourier transform as,

$$S_{pq}(\omega) = \int_{-\infty}^{+\infty} R_{pq}(\tau) \exp^{-i\omega\tau} d\tau \tag{20.2}$$

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 pair-wise locations. Matrix  $S$  can be decomposed using the singular value decomposition (SVD) as,

$$S(\omega) = U \sum U^H \tag{20.3}$$

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 a feature vector for further analysis, where  $m$  refers to the number of spectral lines or attributes. In this way, information from  $l$  signals obtained from  $l$  sensors is fused into a single feature vector.

### 20.4.2.2 Dimensionality Reduction: Random Projection

Dimensionality reduction aims to extract an intrinsic low dimensional information from a high dimensional dataset. It transforms a high-dimensional data set into a lower dimensional one which represents the most important variables that can explain the original data. This feature extraction step is required in this work since we have a low number of observations compared to a large number of features. In [31], the authors discussed an effectiveness of dimensionality reduction approaches in SHM applications.

PCA [29] is one of the most popular and widely used techniques proposed for dimensionality reduction. The main objective of PCA is to calculate eigenvalues and eigenvectors of a covariance matrix computed from a given dataset to determine the components where the data have a maximum variance. However, PCA has a complexity of  $O(m^3)$  due to the eigen decomposition of the covariance matrix where  $m$  is the dimension of data. This makes it impractical to use for very high dimensional datasets, a common issue in SHM sensing data. Moreover, its performance is sensitive to the number of the selected components.

Random projection is an alternative and less expensive method to reduce the dimensionality of extremely high dimensional data [1]. Using random projection, the dimension of the projected space only depends on the number of data points  $n$ , no matter how high the original dimension  $m$  of the data is. It is an effective and efficient dimensionality reduction method for high-dimensional data [9]. 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 [28]. Further study, carried out by Achlioptas [1], shows that the number of dimensions required for random projection can be calculated using:

$$k = \log n / \xi^2 \quad (20.4)$$

where  $k$  is the number of dimensions in the low-dimensional space and  $\xi$  is a small positive number.

Given  $X \in \mathbf{R}^{n \times m}$ ,  $\xi > 0$ , and  $k = \log n / \xi^2$ . Let  $R_{m \times k}$  be a random matrix where each entry  $r_{ij}$  can be drawn from the following probability distribution [1]:

$$r_{ij} = \begin{cases} +1 & \text{with probability } \frac{1}{2s} \\ 0 & \text{with probability } 1 - \frac{1}{2s} \\ -1 & \text{with probability } \frac{1}{2s} \end{cases} \quad (20.5)$$

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  $X$ .

In practice,  $k$  is usually a small number. Venkatasubramanian and Wang [61] suggested that  $k_{RP} = 2 \ln n / 0.25^2$ .

### 20.4.2.3 Damage Detection: Self-tuning One-Class Support Vector Machine

In practice, events corresponding to damaged states of structures are often unavailable for a supervised learning approach. Therefore, a one-class approach using only data from a healthy structure is more practical. In this work, we use one-class SVM [52] as an anomaly detection method.

Given a set of data  $X = \{x_i\}_{i=1}^n$  extracted from the original sensor data (feature vector) collected from a healthy structure and where  $n$  is the number of training samples, one-class SVM maps these samples into a high dimensional feature space using a function  $\phi$  through the kernel  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ . Then one-class SVM learns a hyperplane that separates these data points from the origin with a maximum margin. 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  $f : \mathbf{R}^m \rightarrow \{-1, +1\}$  and is written in the form of

$$f(x) = \text{sgn}(w \cdot \phi(x) - \rho) \quad (20.6)$$

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.

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

$$\min_{w, \xi, \rho} \frac{1}{2} \|w\|^2 + \frac{1}{\nu n} \sum_{i=1}^n \xi_i - \rho \quad (20.7)$$

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

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). The problem can be transformed to a dual form using Lagrangian multiplier as,

$$\min_{\alpha_1, \alpha_2, \dots, \alpha_n} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \tag{20.8}$$

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

This problem can then be solved using quadratic programming [27]. Having obtained a learned model, the decision values for a new data instance  $x_{new}$  can be computed as,

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i K(x_i, x_{new}) - \rho\right) \tag{20.9}$$

A negative decision value indicates an anomaly, which likely corresponds to a structural damage.

*Self-tuning Gaussian Kernel:*

Gaussian kernel defined in Eq.20.10 has gained much popularity in the area of machine learning and it turned out to be an appropriate setting for one-class SVM [13, 30, 36]. It has a parameter denoted  $\sigma$  which may severely affect the performance of a one-class SVM. An inappropriate choice of  $\sigma$  may lead to overfitting or underfitting.

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{20.10}$$

where  $\sigma \in \mathbf{R}$  is the kernel parameter.

$K$ -fold cross validation is often used at a training stage in order to tune  $\sigma$ . However, in case of a one-class learning, this technique is not possible because it selects  $\sigma$  that works only on the training class data and thus it is lack of generalisation capability (overfitting problem). Therefore, alternative approaches have been proposed for tuning  $\sigma$  in one-class SVM. The Appropriate Distance to the Enclosing Surface (ADES) algorithm [4] is our recent proposed method for tuning  $\sigma$  based on inspecting the spatial locations of the edge and interior samples, and their distances to the enclosing surface of one-class SVM. ADES showed successful performances on several datasets and thus was adopted for tuning  $\sigma$  in this work.

Following the objective function  $f(\sigma_i)$  described in Eq.20.11, the ADES algorithm selects the optimal value of  $\hat{\sigma} = \underset{\sigma_i}{\text{argmax}}(f(\sigma_i))$ , which generates a hyperplane that is the furthest from the interior samples and the closest to the edge samples,

using a normalised distance function.

$$f(\sigma_i) = \text{mean}(d_N(x_n)_{x_n \in \Omega_{IN}}) - \text{mean}(d_N(x_n)_{x_n \in \Omega_{ED}}) \quad (20.11)$$

where  $\Omega_{IN}$  and  $\Omega_{ED}$ , respectively, represent sets of interior and edge samples in the healthy training 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:

$$d_N(x_n) = \frac{d(x_n)}{1 - d_\pi} \quad (20.12)$$

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:

$$d(x_n) = \frac{f(x_n)}{\|w\|} = \frac{\sum_{i=1}^n \alpha_i K(x_i, x_n) - \rho}{\sqrt{\sum_{ij} \alpha_i \alpha_j K(x_i, x_j)}} \quad (20.13)$$

where  $w$  is a perpendicular vector to the decision boundary,  $\alpha_i$  are the Lagrange multipliers, and  $\rho$  is the bias term. More details on the ADES method can be found in [4].

### 20.4.3 Substructure Grouping

This section proposes a robust clustering technique, which uses spectral moment features for substructure grouping and anomaly detection. The proposed approach follows the following steps, which are further detailed in the remainder of this section:

- a structurally meaningful feature is extracted using spectral moment from the measured acceleration for each jack arch for many time windows,
- a modified  $k$ -means— clustering algorithm is applied to this feature data to identify groups of similar substructures and potential anomalies,
- a multi-indices criterion is used to select the best grouping outcome,
- under the assumption that near-by substructures should have similar behaviours and thus should belong to the same cluster groups, any substructure which is identified as an outlier or which belongs to a one-member group, is then marked as an anomaly.

#### 20.4.3.1 Feature Extraction Using Spectral Moment

In this study, a frequency-based feature using spectral moments of the measured acceleration responses is adopted as a damage sensitive feature. PSD of the response signal is required to calculate spectral moment. For a stationary random process,

PSD contains some major characteristics of the system that can be extracted. In a classical Fourier analysis, the power of a signal can be obtained by integrating the PSD, i.e., the square of the absolute value of the Fourier-transform coefficients [15].

The energy contents of a signal within a frequency band of interest can also be quantified using PSD. The calculation of PSD is computationally efficient, as it has a low processing cost compared to modal analysis. Moreover unlike modal data, PSD does not suffer the lack of information and provides an abundance of information in a wider frequency range.

The spectral moment of a random stationary signal provides some important information about its statistical properties. They explicitly depend on the frequency content of the original signal, which makes them suitable to SHM applications. Spectral moment captures information from entire spectra and hence they can distinguish any subtle difference between normal and distorted signals.

As described in Sect. 20.4.2.1, the frequency characterisation of a random stationary process can be computed using the PSD function as,

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-i\omega\tau} d\tau \quad (20.14)$$

For a given PSD, the  $n$ th-order spectral moment can be then computed as,

$$\lambda_x^n = \int_{-\infty}^{\infty} |\omega|^n S_{xx}(\omega) d\omega \quad (20.15)$$

where  $n$  is the order of spectral moment. Finally, for a discretised signal  $x$ , the  $n$ th-order spectral moment  $\lambda_x^n$  can be obtained using,

$$\lambda_x^n = \frac{2}{N^{n+1}} \sum_0^{\lfloor N/2 \rfloor} S_{xx}(j) \left( \frac{j}{\Delta t} \right)^n \quad j \in [1 : N/2] \quad (20.16)$$

where  $S_{xx}$  and  $\Delta t$  are, respectively, the discrete spectral density and the sampling period.

The zero-th order moment refers to the area under the spectral curve which represents the significance of the response. Higher order moments assign more weight to frequency components. Past research studies have concluded that spectral moments with orders 1–4 provide useful information about the system, whereas higher order moments usually do not provide further information as they are highly masked by noise [17].

### 20.4.3.2 *k*-means– Clustering

Clustering is a popular method in data mining applications [25]. The goal of clustering is to partition a set of data objects into groups of similar objects based on a given set of features. *k*-means is a widely used clustering algorithm, which groups data into *k* clusters  $C = \{C_1, \dots, C_k\}$  with the goal of minimising the within-cluster sum of squares, i.e.

$$\arg \min_C \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (20.17)$$

where  $\mu_i$  is the centre of cluster *i* (mean of data points in  $C_i$ ). This optimisation function can be solved in an iterative manner, which converges after no further assignment changes between iterations.

However, the *k*-means method may converge to a sub-optimal partitioning, as it is sensitive to the initial selection of cluster centres. The *k*-means++ algorithm [6] is an alternative method, which uses a specific mechanism to select the initial set of centres, before applying the original *k*-means steps. *k*-means++ only selects one initial centre uniformly at random from all data points (as opposed to all the initial centres for *k*-means). Each subsequent cluster centre is then selected from the remaining data points with a probability proportional to its squared distance to the closest existing centre.

Outliers in the data can skew the selection of cluster centres and thus can lead both *k*-means and *k*-means++ to sub-optimal solutions. The recent *k*-means—– alternative [12] proposes a mechanism to detect such outliers (e.g. potential anomalies). In the previous methods, such anomalies were likely located in significantly small clusters as a by-product of the iterative process. In contrast, in *k*-means—–, these anomalies are explicitly detected and isolated before the iterative cluster update process.

We propose the following extension to the original *k*-means—– algorithm. When convergence is achieved, any group with a single member is removed from the cluster set and its data point is added to the set of anomalies. This additional step prevents biases when selecting the best cluster result, as described in the next subsection. Our extended *k*-means—– is described in Algorithm 3. It follows the iterative steps of *k*-means, but first selects *o* anomalies in the data before assigning the remaining points into *k* clusters. Thus, these *o* data points that are furthest from their closet centres are isolated and are not used to recompute the centres in the update step and subsequent iterations.

### 20.4.3.3 Selection of the Best Clustering Result

Due to the random choice of the initial first centres in Algorithm 3, multiple runs over the same data set will produce different clustering results. This can be addressed by using a high number of replications, such as 50. However, different settings of

---

**Algorithm 3** A modified  $k$ -means— clustering.
 

---

**Input:** Matrix of  $X$  data points, number of clusters  $k$ , number of anomalies  $o$

**Output:**  $o$  anomalies in  $L$ , cluster ID for each data point in  $X - L$

---

- 1: **Initialisation:** using  $k$ -means++ to find  $k$  initial centres
  - 2: **Assignment:** assign each data point to the nearest centre, and set their cluster ID accordingly
  - 3: **Anomaly detection:** find the  $o$  points, which are furthest from their cluster centres (i.e. anomalies), and assign them to  $L$
  - 4: **Update:** recompute the centre for each cluster (excluding the found anomalies)
  - 5: **Iterate:** repeat steps 2 to 4 until the algorithm converges, i.e. no further changes in the data point assignments
  - 6: **Finalise:** convert clusters with only one data point to anomalies
- 

$k$  and  $o$  will also produce different clustering results. To address this issue, we limit the choice for  $k$  to a fixed maximum arbitrary value. In practice, this selection of the maximum  $k$  should be guided by domain knowledge of the application at hand. In the case of SHM such as the application of our scheme to a bridge, the maximum  $k$  value could be set equal to the number of structural spans of a bridge. For example,  $k$  could be set to 6 for a bridge which has 6 different structural spans. The  $o$  parameter may remain arbitrarily low, such as less than 5.

We then propose the following mechanism to select the most informative clustering and anomaly detection results. For each pair of input parameters  $(k, o)$ , we compute the values of the Silhouette [47], the Davies-Bouldin [16], and the Dunn [19] indices over the resulting cluster set. Each index measures a specific characteristic of such a resulting cluster set. Indeed, the Silhouette index measures the averaged dissimilarity of each point against its assigned cluster, and then compares these measurements against the dissimilarity of the points within their nearest neighbouring clusters. On the other hand, the Davies-Bouldin index reports on the compactness and separation of the clusters, through the ratio between the similarities within a group and the differences between groups. The Dunn index computes the ratio between the closest points across different groups and the furthest points within groups.

We then select the  $(k, o)$  results which have extremum values for each of the computed indices, i.e. maximum value for Silhouette and Dunn; and minimum value for Davies-Bouldin. Within this set of results, we select the logical intersection of all identified anomalies as the final set of anomalies, i.e. points which have instrumentation issues or indicate structural damage. Any empty set of identified anomalies is treated as the identity element for this operation (i.e. does not influence the outcome). As the three indices report on different aspects of the cluster groups, using their intersection may lead to a more accurate set of anomalies. This is confirmed through experimental results in the next Sect. 20.5.



## 20.5 Case Studies and Results

### 20.5.1 Damage Identification

#### 20.5.1.1 Case Study: The Sydney Harbour Bridge Specimen

A concrete cantilever beam, which has an arch section with a similar geometry to those on the SHB, was manufactured and tested, as shown in Fig. 20.3. The beam consists of a 200UB18 steel I-Beam with a 50 mm concrete cover on both ends. The length of the specimen is 2 m, the width is 1 m and the depth is 0.375 m. 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. In addition, a support was placed at 1200 mm away from the tip to avoid any cracking occurring in the specimen under its self-weight [42].

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 locations of the accelerometers are shown in Fig. 20.3. The structure was excited using an impact hammer with steel tip, which was applied on the top surface of the specimen and just above the location of sensor A9. The acceleration response of the structure was collected over a time period of 2 s at a sampling rate of 8 kHz, resulting in 16000 samples for each event (i.e. a single excitation). A total of 190 impact tests were collected from a healthy condition of the specimen.

A crack was introduced into the specimen in the location marked in Fig. 20.3 using a cutting saw. The crack is located between sensor locations A2 and A3 and progressively increases towards sensor location A9. The length of the cut was increased gradually from 75 to 150 mm, 225 and 270 mm, and the depth of the cut was fixed to 50 mm. After introducing each damage case, a total of 190 impact tests were performed on the structure in the location described earlier.

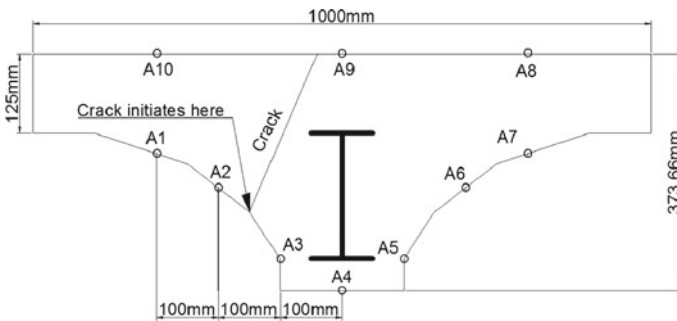
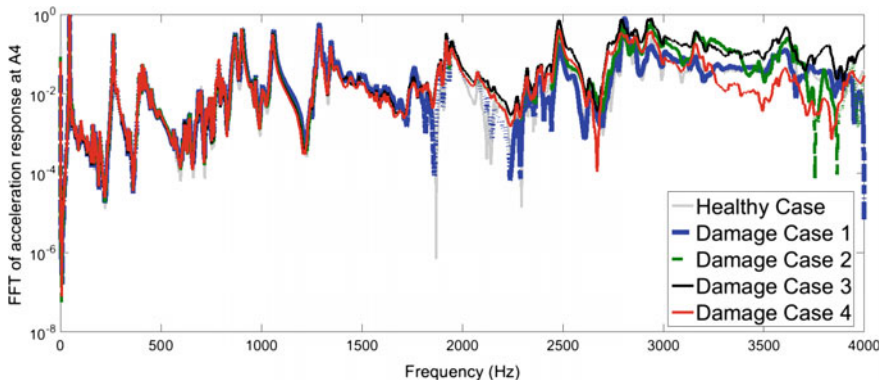


Fig. 20.3 A laboratory specimen with cracking



**Fig. 20.4** Comparison of the frequency response function (inertance) between the healthy state and the four damage cases for sensor location A4

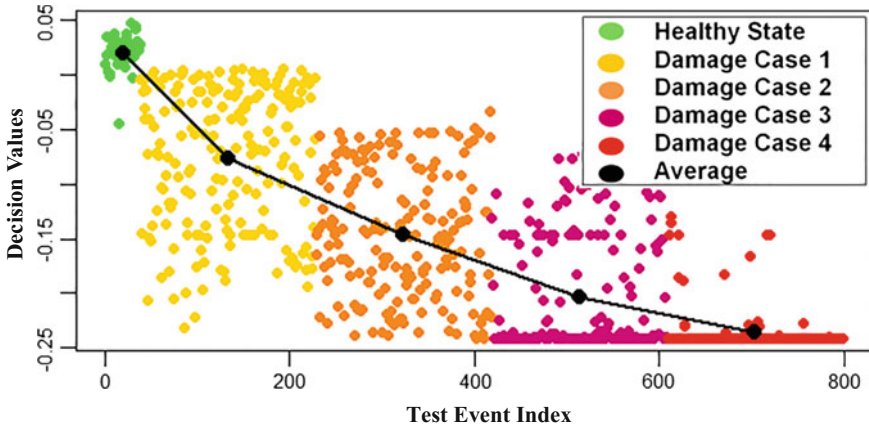
**Table 20.1** Comparison of the first three modes of the structure in the healthy state and the four damage cases

Natural Fre- quency (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

We further investigated the impact of damage by comparing the frequency response function (FRF) of the structure between the measured responses obtained from the healthy case and four damage cases as shown in Fig. 20.4. It was observed that the damage effects are more evident at high frequency, as the change between the healthy and the damaged structure became more significant. Table 20.1 compares the natural frequencies for the first three modes in the healthy state and three damage cases, as well as the change in frequency of each damage case relative to the healthy state. From Table 20.1, it can be clearly seen that once the severity of damage increases, a higher discrepancy in the first three modal frequencies with respect to the healthy state is obtained.

**20.5.1.2 Results**

We have applied our proposed damage detection and severity assessment framework (described in Fig. 20.2) onto our specimen dataset. A total of 950 samples were collected in this experiment, where each sample is a measured vibration response of

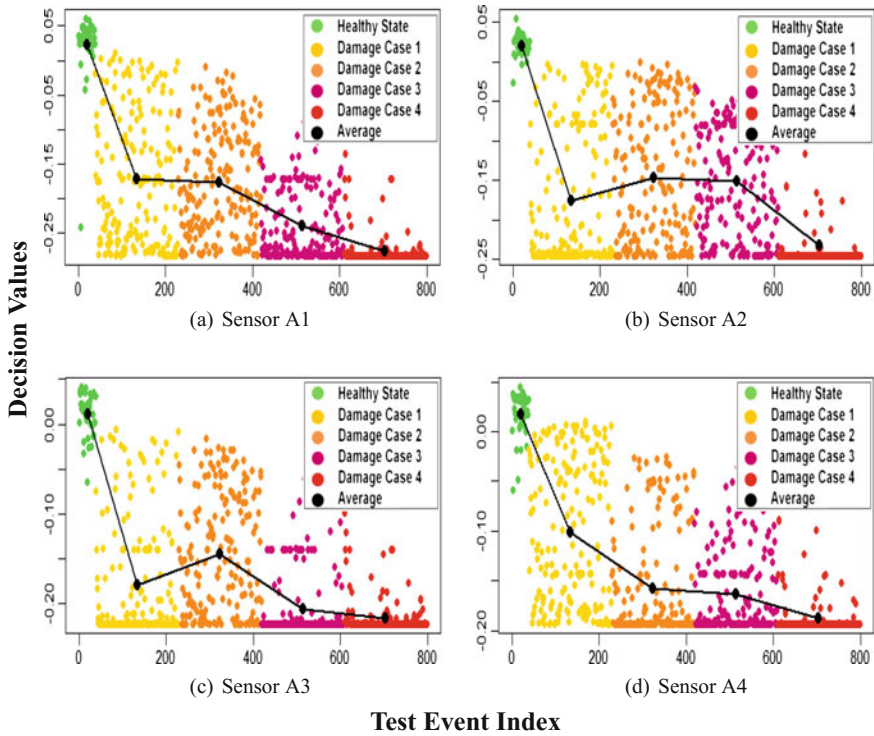


**Fig. 20.5** Damage identification results using FDD for feature fusion and extraction

the structure with eight thousand attributes in the frequency domain ( $8\text{ kHz} \times 2\text{ s} \times 0.5$  (considering Nyquist frequency)). We separated the data samples into two main groups, healthy samples (190 samples) and damaged samples (760). 80% of the healthy cases data were randomly selected for a training stage, while the remaining 20% of healthy samples and all the damaged cases were used as a test data for validating the proposed approach. Feature extraction and fusion from ten sensors using FDD were initially applied on the training data, and random projection was used for dimensionality reduction. This was followed by calculating the optimal value of  $\sigma$  using the ADES method defined in Eq. 20.11 and constructing a one-class SVM as a damage detection model.

The constructed model was then validated using the test data. Similar to the training steps, the FDD method was initially applied to the test data followed by dimensionality reduction algorithm. The final step was to present the test data onto the constructed one-class SVM model to evaluate its performance in terms of damage detection and severity assessment. As expected, the constructed model was able to successfully detect the damaged cases and produced an F1-score of 0.95. A detailed summary of the results is presented in Fig. 20.5. The figure shows the decision values of all test data, where the black dots represent average decision values for healthy and each damaged cases.

Only three events from the healthy samples were misclassified as damaged. On the other hand, all the damaged samples were correctly classified except for four events in Damage Case 1 that had positive decision values (false negative). This suggests that the model is well generalised on unseen samples and has the ability to detect damaged and healthy samples. It should be emphasised that the level of damage in this case study is considerably small. Moreover, the method also shows a capability to assess a progression of damage (as shown by decreasing decision values for Damage Cases 1 to 4) despite variations in operational conditions. Moreover, the obtained machine learning results match very well with the findings from domain



**Fig. 20.6** Damage identification results using a separate one-class SVM model for each sensor location

knowledge presented in Table 20.1. A decreasing trend in the ML scores indicates progressive damage in the structure.

To further investigate the effectiveness of feature fusion using FDD, an alternative approach was adopted without using FDD for sensor fusion. Only the frequency features (using FFT) of the acceleration response obtained from each sensor were used to construct a separate damage detection model for each sensor using data from the healthy case.

Damage identification results using this approach are presented in Fig. 20.6 for sensors A1, A2, A3 and A4 (results for other sensors were similar). It can be realised that this approach does not have the capability to monitor the progress of damage. The decision values did not consistently follow the trend of the damage as shown in Fig. 20.6b, c. Based on this, it can be concluded that FDD is robust against excitation variations and can provide reliable information about the severity of damage in the structure.

## 20.5.2 Substructure Clustering and Anomaly Detection

### 20.5.2.1 Case Study: The Sydney Harbour Bridge

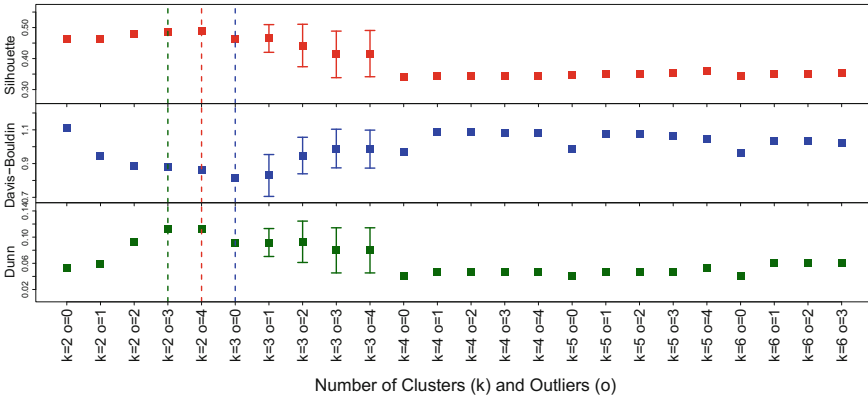
The goal of this study was to group substructures (i.e. jack arches) with a similar behaviour and then identify substructures with potential anomalies. We used a set of 85 nodes over five structural sections of the SHB, i.e. five different spans of the bridge. These spans were located on the Northern Main Span and the Northern Approach, as illustrated on Fig. 20.1. For each node, we collected 10 min of continuous acceleration data at 1500Hz over 22 days in July 2015 (as described in Sect. 20.2). We pre-processed this data to identify a continuous 1 min of *ambient* response, i.e. a period where no vehicle was driving over the node. For each of these periods, we computed the spectral moment feature as described in Sect. 20.4.3.1 for accelerations in  $x$ ,  $y$  and  $z$  direction (denoted  $SM_x$ ,  $SM_y$  and  $SM_z$ ), and we averaged them for each node over the 22 days.

We applied our extended  $k$ -means – method and its outcome selection criteria (Sect. 20.4.3.2) to this set of spectral moment features. We varied the parameter  $k$  (i.e. number of clusters) from 2 to 6, as the studied nodes were spread across five structural sections, and the parameter  $o$  (i.e. number of anomalies) from 0 to 4. Finally, we replicated this experiment 10 times. The following subsection reports on the results related to the second order spectral moment. The first and third order moments produced similar results and were not included here.

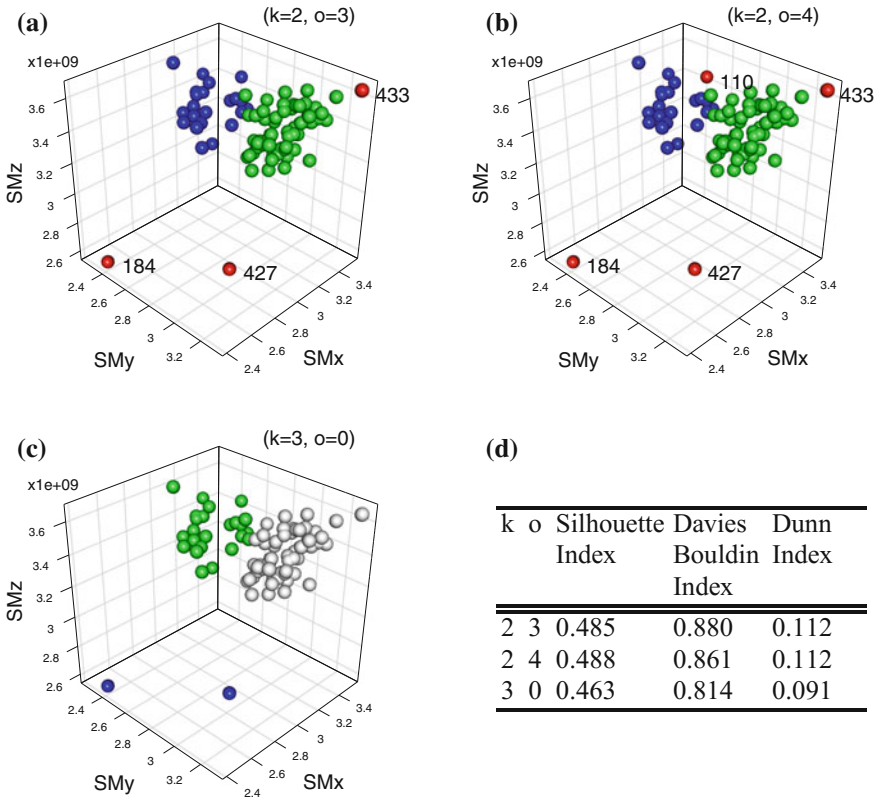
### 20.5.2.2 Results

Figure 20.7 shows the Silhouette, Davies-Bouldin, and Dunn indices for each  $(k, o)$  pair. Using our selection criteria, we retained the pairs  $(k = 2, o = 3)$ ,  $(k = 2, o = 4)$ , and  $(k = 3, o = 0)$  as they corresponded to the required extremum values. For these pairs, Fig. 20.8 shows the 3D scatterplots for the second order spectral moment in  $x$ ,  $y$  and  $z$ , and Fig. 20.8d shows the related index values. The nodes 184, 427, and 433 formed the set of anomalies resulting from the intersection of these pairs as described in Sect. 20.4.3.2.

For  $(k = 3, o = 0)$ , the nodes 184 and 427 were in a well-separated group in the 3D feature space, and node 433 was included into one of the other two clusters. This outcome is due to the setting  $o = 0$ , i.e. the clustering algorithm had to reject any outright outliers (i.e. by-pass step 3 of Algorithm 3). Limiting the range of  $o$  to strictly positive integers (e.g.  $o \in [1, 4]$ ) would result in node 433 being identified as an anomaly. However, having  $o > 1$  may provide more false positives, as it will force the clustering process to mark the most distant point in a dataset as an anomaly, even if that point is *well* matched to a group. This may be a better decision for a bridge manager, as it could be safer to discard a false positive after a visual engineering inspection than letting a false negative remain undetected.

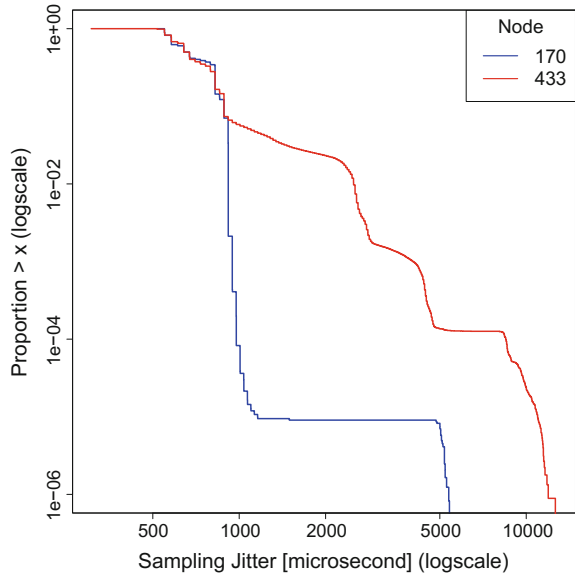


**Fig. 20.7** Silhouette, Davies-Bouldin, and Dunn indices for different  $(k, o)$  parameters



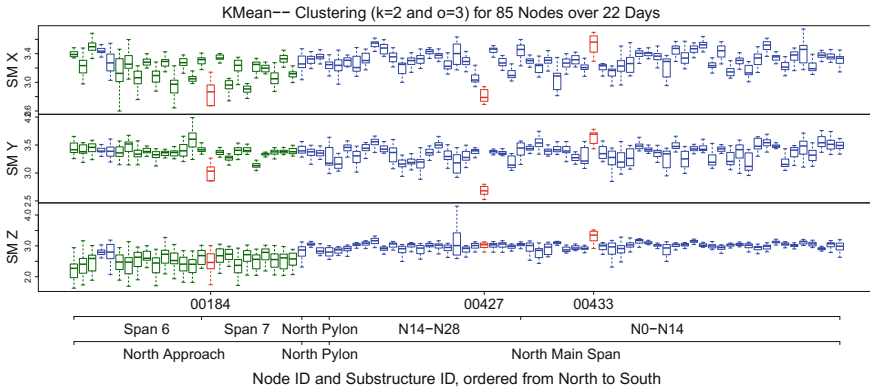
**Fig. 20.8** **a, b, c** Selected 3D scatter plots of spectral moments (SM) for each node, which are coloured based on their cluster membership for specific parameters, and **d** their corresponding performance index values. Cluster groups are coloured in blue, green, and grey, anomalies are coloured in red

**Fig. 20.9** Difference between the time interval jitter for the data of a healthy working node 170 and node 433



Further engineering investigations of the nodes in the resulting set of anomalies (i.e. 184, 427, and 433) showed that they were all having instrumentation issues during the 22-day period of this study (i.e. sensor defect for 184 and 433, power unit defect for 427) [3]. As an example for node 433, Fig. 20.9 presents the log-scale ECDF of the time interval jitter between two collected data points, as compared to the healthy working node 170. This jitter should be as close to 0 as possible, i.e. for node 170 only 0.01% of the data points had a jitter greater than 1 ms. Node 433 produced in contrast a higher jitter distribution, i.e. more than 1% of the data points had a jitter greater than 2 ms. From a hardware perspective, the cause of such a high jitter could be a failure of the oscillator-based clock of the sensor producing the data. This sensor was marked for replacement.

Figure 20.10 shows the boxplots of second order spectral moment values for each direction and each node in the case of ( $k = 2, o = 3$ ). The nodes are ordered on the x-axis according to their physical location on the SHB from north (left) to south (right). The boxplot for a node is coloured based on its group membership, with the anomalies marked in red. This figure confirms that the nodes that are located on a given structural section are mostly grouped into the same cluster. Indeed most of the North Approach nodes are in the green group, whereas all the Northern Main Span nodes are in the blue group.



**Fig. 20.10** Boxplots of second order X, Y, and Z spectral moment values for  $(k = 2, o = 3)$  for each node. On the x-axis, the nodes are ordered based on their location from north to south. The colours indicate the assigned cluster groups (blue or green) and the anomalies (red)

## 20.6 Conclusion

This work presents damage identification and substructure grouping approaches for SHM applications using machine learning techniques and features extracted using domain knowledge. The two approaches performed successfully in two case studies using data from a laboratory structure and real data collected from the SHB. Domain knowledge is used in this chapter to show how an SHM problem is formed as a machine learning problem using domain knowledge. It also shows the importance of domain knowledge in extracting damage sensitive features as well as interpreting the results found by machine learning approaches.

In the first approach, a structural benchmark model was built using a self-tuning one-class SVM on a feature space fused and extracted from multiple sensors by FDD, followed by random projection for dimensionality reduction. Then new events were tested against the benchmark model to detect damage. The approach detected damage well with high accuracy and low false positives, even for a small damage case. Moreover, this proposed approach also achieved damage severity assessment using data fusion and decision values from the SVM. In the second approach, a robust clustering technique was utilised on spectral moment features for substructure grouping and anomaly detection. The technique was able to group substructures of similar behaviour on the SHB and to detect anomalies spatially, which were associated with sensor issues from the instrumented substructures.

This work is part of our ongoing effort to build Smart Infrastructures, which bring together data acquisition, data management, and data analytics techniques to optimise their maintenance and services. Our future works include an implementation of the proposed approaches on our production system on the SHB, and applying them using data collected from other structures.



**Acknowledgements** The authors wish to thank the Roads and Maritime Services (RMS) in New South Wales, Australia for provision of the support and testing facilities for this research work. NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program. CSIRO's Digital Productivity business unit and NICTA have joined forces to create digital powerhouse Data61.

## References

1. Achlioptas, D.: Database-friendly random projections: Johnson-lindenstrauss with binary coins. *J. Comput. Syst. Sci.* **66**(4), 671–687 (2003)
2. Alamdari, M.M., Li, J., Samali, B.: Frf-based damage localization method with noise suppression approach. *J. Sound Vib.* **333**(14), 3305–3320 (2014)
3. Alamdari, M.M., Rakotoarivelo, T., Khoa, N.L.D.: A spectral-based clustering for structural health monitoring of the Sydney Harbour Bridge. *Mech. Syst. Signal Process.* **87**, 384–400 (2017)
4. Anaissi, A., Khoa, N.L.D., Mustapha, S., Alamdari, M.M., Braytee, A., Wang, Y., Chen, F.: Adaptive one-class support vector machine for damage detection in structural health monitoring. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 42–57. Springer (2017)
5. Antoni, J., Randall, R.: The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process.* **20**(2), 308–331 (2006)
6. Arthur, D., Vassilvitskii, S.: K-means++: The advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '07. Society for Industrial and Applied Mathematics, pp. 1027–1035. Philadelphia, PA, USA (2007)
7. Banks, H., Inman, D., Leo, D., Wang, Y.: An experimentally validated damage detection theory in smart structures. *J. Sound Vib.* **191**(5), 859–880 (1996)
8. Benasciutti, D., Cristofori, A., Tovo, R.: Analogies between spectral methods and multiaxial criteria in fatigue damage evaluation. *Probab. Eng. Mech.* **31**, 39–45 (2013)
9. Bingham, E., Mannila, H.: Random projection in dimensionality reduction: applications to image and text data. In: Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '01, pp. 245–250. ACM, New York, NY, USA (2001). DOI <https://doi.org/10.1145/502512.502546>
10. Brincker, R., Zhang, L., Andersen, P.: Output-only modal analysis by frequency domain decomposition. In: The International Conference on Noise and Vibration Engineering, pp. 717–723 (2001)
11. Cattarius, J., Inman, D.: Time domain analysis for damage detection in smart structures. *Mech. Syst. Signal Process.* **11**(3), 409–423 (1997)
12. Chawla, S., Gionis, A.: k-means-: A unified approach to clustering and outlier detection. In: Proceedings of the 13th SIAM International Conference on Data Mining, 2–4 May 2013. Austin, Texas, USA, pp. 189–197 (2013)
13. Chen, Y., Zhou, X.S., Huang, T.S.: One-class svm for learning in image retrieval. In: 2001 International Conference on Image Processing, vol. 1, pp. 34–37. IEEE (2001)
14. Curadelli, R., Riera, J., Ambrosini, D., Amani, M.: Damage detection by means of structural damping identification. *Eng. Struct.* **30**(12), 3497–3504 (2008)
15. CusidÓcusido, J., Romeral, L., Ortega, J.A., Rosero, J.A., Espinosa, A.G.: Fault detection in induction machines using power spectral density in wavelet decomposition. *IEEE Trans. Ind. Electron.* **55**(2), 633–643 (2008)
16. Davies, D., Bouldin, D.: A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1**, 224–227 (1979)
17. Di-Paola, M., Muscolino, G.: On the convergent parts of high order spectral moments of stationary structural responses. *J. Sound Vib.* **110**, 233–245 (1986)

18. Doebling, S.W., Farrar, C.R., Prime, M.B., Shevitz, D.W.: Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. Technical report. Los Alamos National Laboratory, NM, USA (1996)
19. Dunn, J.: A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *J. Cybern.* **3**, 32–57 (1973)
20. Farrar, C.R., Doebling, S.W., Nix, D.A.: Vibration-based structural damage identification. *Philos. Trans. R. Soc. Lond. A: Math. Phys. Eng. Sci.* **359**(1778), 131–149 (2001)
21. Farrar, C.R., Worden, K.: An introduction to structural health monitoring. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 365, pp. 303–315. The Royal Society (2007)
22. Fugate, M.L., Sohn, H., Farrar, C.R.: Vibration-based damage detection using statistical process control. *Mech. Syst. Signal Process.* **15**(4), 707–721 (2001)
23. Hou, Z., Noori, M., Amand, R.S.: Wavelet-based approach for structural damage detection. *J. Eng. Mech.* **126**(7), 677–683 (2000)
24. H.Sohn, Farrar, C., Hemez, F., Shunk, D., Stinemates, D., Nadler, B., Czarnecki, J.: A review of structural health monitoring literature: 1996–2001. Los Alamos National Laboratory, USA (2003)
25. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. *ACM Comput. Surv.* **31**, 264–323 (1999)
26. Jiang, S.F., Zhang, C.M., Koh, C.: Structural damage detection by integrating data fusion and probabilistic neural network. *Adv. Struct. Eng.* **9**(4), 445–458 (2006)
27. Joachims, T.: Text categorization with support vector machines: Learning with many relevant features. In: *European Conference on Machine Learning*, pp. 137–142. Springer (1998)
28. Johnson, W.B., Lindenstrauss, J.: Extensions of lipschitz mappings into a hilbert space. *Contem. Math.* **26**(189–206), 1 (1984)
29. Jolliffe, I.: *Principal component analysis*. Wiley Online Library (2002)
30. Kemmler, M., Rodner, E., Denzler, J.: One-class classification with gaussian processes. In: *Asian Conference on Computer Vision*, pp. 489–500. Springer (2010)
31. Khoa, N.L.D., Zhang, B., Wang, Y., Chen, F., Mustapha, S.: Robust dimensionality reduction and damage detection approaches in structural health monitoring. *Struct. Health Monit.* **13**(4), 406–417 (2014)
32. Law, S., Li, X., Zhu, X., Chan, S.: Structural damage detection from wavelet packet sensitivity. *Eng. Struct.* **27**(9), 1339–1348 (2005)
33. Lee, U., Shin, J.: A frequency response function-based structural damage identification method. *Comput. & Struct.* **80**(2), 117–132 (2002)
34. Liberatore, S., Carman, G.: Power spectral density analysis for damage identification and location. *J. Sound Vib.* **274**(3), 761–776 (2004)
35. Liu, Q.C., Wang, H.P.B.: A case study on multisensor data fusion for imbalance diagnosis of rotating machinery. *AI EDAM* **15**(3), 203–210 (2001)
36. Long, J., Buyukozturk, O.: Automated structural damage detection using one-class machine learning. In: *Dynamics of Civil Structures*, Vol. 4, pp. 117–128. Springer (2014)
37. Lu, Y., Michaels, J.E.: Feature extraction and sensor fusion for ultrasonic structural health monitoring under changing environmental conditions. *IEEE Sens. J.* **9**(11), 1462–1471 (2009)
38. Maia, N., Silva, J., Almas, E., Sampaio, R.: Damage detection in structures: from mode shape to frequency response function methods. *Mech. Syst. Signal Process.* **17**(3), 489–498 (2003)
39. Makki Alamdari, M.: *Vibration-based structural health monitoring*. Ph.D. thesis, University of Technology Sydney (2015)
40. Makki-Alamdari, M., Li, J., Samali, B.: A comparative study on the performance of the damage detection methods in the frequency domain. In: *Australasian Conference on the Mechanics of Structures and Materials*. CRC press/Balkema (2013)
41. Makki Alamdari, M., Samali, B., Li, J., Kalhori, H., Mustapha, S.: Spectral-based damage identification in structures under ambient vibration. *J. Comput. Civil Eng.* p. 04015062 (2015)
42. Mustapha, S., Hu, Y., Nguyen, K., Alamdari, M.M., Runcie, P., Dackermann, U., Nguyen, V., Li, J., Ye, L.: Pattern recognition based on time series analysis using vibration data for structural health monitoring in civil structures. Special Issue: *Electron. J. Struct. Eng.* **14**(1) (2015)

43. Nagarajaiah, S., Basu, B.: Output only modal identification and structural damage detection using time frequency & wavelet techniques. *Earthquake Eng. Eng. Vib.* **8**(4), 583–605 (2009)
44. Pandey, A., Biswas, M.: Damage detection in structures using changes in flexibility. *J. Sound Vib.* **169**(1), 3–17 (1994)
45. Peeters, B., Maeck, J., De Roeck, G.: Vibration-based damage detection in civil engineering: excitation sources and temperature effects. *Smart Mater. Struct.* **10**(3), 518 (2001)
46. Petrucci, G., Zuccarello, B.: Fatigue life prediction under wide band random loading. *Fatigue & Fract. Eng. Mater. Struct.* **27**(12), 1183–1195 (2004)
47. Rousseeuw, P.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987)
48. Runcie, P., Mustapha, S., Rakotoarivelo, T.: Advances in structural health monitoring system architecture. In: *International Symposium on Life-Cycle Civil Engineering*, pp. 1064–1071 (2014)
49. Rytter, A.: Vibration-based inspection of civil engineering structures. Ph.D. thesis, University of Aalborg (1993)
50. Salawu, O.: Detection of structural damage through changes in frequency: a review. *Eng. Struct.* **19**(9), 718–723 (1997)
51. Santos, A., Figueiredo, E., Silva, M., Sales, C., Costa, J.: Machine learning algorithms for damage detection: Kernel-based approaches. *J. Sound Vib.* **363**, 584–599 (2016)
52. Schölkopf, B., Williamson, R.C., Smola, A.J., Shawe-Taylor, J., Platt, J.C., et al.: Support vector method for novelty detection. In: *NIPS*, vol. 12, pp. 582–588. Citeseer (1999)
53. Shi, Z., Law, S., Zhang, L.: Structural damage detection from modal strain energy change. *J. Eng. Mech.* **126**(12), 1216–1223 (2000)
54. Sohn, H., Farrar, C.R.: Damage diagnosis using time series analysis of vibration signals. *Smart Mater. Struct.* **10**(3), 446 (2001)
55. Sohn, H., Farrar, C.R., Hemez, F.M., Czarnecki, J.J.: A review of structural health review of structural health monitoring literature 1996–2001. Technical Report, Los Alamos National Laboratory (2002)
56. Sophian, A., Tian, G.Y., Taylor, D., Rudlin, J.: A feature extraction technique based on principal component analysis for pulsed eddy current ndt. *NDT & e Int.* **36**(1), 37–41 (2003)
57. Su, Z., Wang, X., Cheng, L., Yu, L., Chen, Z.: On selection of data fusion schemes for structural damage evaluation. *Struct. Health Monit.* **8**(3), 223–241 (2009)
58. Sun, Z., Chang, C.: Structural damage assessment based on wavelet packet transform. *J. Struct. Eng.* **128**(10), 1354–1361 (2002)
59. Sunder, S.S., Grewatz, S.E., Ting, S.K.: Modal identification using spectral moments. *Struct. Saf.* **3**(1), 1–11 (1985)
60. Vanmarcke, E.H.: Properties of spectral moments with applications to random vibration. *J. Eng. Mech. Div.* **98**(2), 425–446 (1972)
61. Venkatasubramanian, S., Wang, Q.: The johnson-lindenstrauss transform: An empirical study. In: M. Mller-Hannemann, R.F.F. Werneck (eds.) *ALENEX*, pp. 164–173. SIAM (2011)
62. Worden, K., Manson, G.: The application of machine learning to structural health monitoring. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **365**(1851), 515–537 (2007)
63. Yao, R., Pakzad, S.N.: Autoregressive statistical pattern recognition algorithms for damage detection in civil structures. *Mech. Syst. Signal Process.* **31**, 355–368 (2012)