

# Data-Driven Approach to Structural Health Monitoring Using Statistical Learning Algorithms

Debarshi Sen and Satish Nagarajaiah

**Abstract** Condition assessment and prediction of existing infrastructure systems is a major objective in civil engineering. Structural health monitoring (SHM) achieves this goal by continuous data acquisition from an array of sensors deployed on the structure of interest. SHM constructs ubiquitous damage features from the acquired data, ensuring maximum sensitivity to the onset of damage and robustness to noise and variability in environmental and operational conditions. Traditionally, SHM has used a model-based approach, wherein a high-fidelity model of a structure is constructed and studied in detail to aid engineers in detecting the onset of damage based on deviations from an undamaged model of the system. However, given the complexity of structures and the inability to perfectly model all aspects of a system, a data-driven approach becomes an attractive alternative. In data-driven approaches, a surrogate model, constructed using acquired data from a system, is substituted for a real model. Although such models do not necessarily capture all the physics of a system, they are efficient for damage detection purposes. Statistical learning algorithms aid in the construction of such surrogate models, and their use has now been extensively documented in the literature. This chapter provides a brief review of applications of statistical learning algorithms, both supervised and unsupervised, in SHM for real-time condition assessment of civil infrastructure systems.

## Introduction

Structural health monitoring (SHM) is an essential component in both the maintenance of existing civil infrastructure systems and the planning of new systems. It entails condition assessment by data acquisition from an array of sensors deployed on the system of interest [1]. Condition assessment is typically followed by

---

D. Sen (✉) · S. Nagarajaiah  
Rice University, 6100 Main Street, Houston TX 77005, Texas, USA  
e-mail: debarshi.sen@rice.edu

S. Nagarajaiah  
e-mail: satish.nagarajaiah@rice.edu

damage prediction that involves estimation of the remaining useful life [2]. The focus of this chapter is on SHM only. Condition assessment typically employs damage detection that requires quantitative and qualitative appraisal of changes in system properties with respect to a known undamaged state of the system. The comparison can be performed using either an undamaged high-fidelity model of the system or a pseudo-model constructed solely with acquired data from the system. This gives rise to two classes of SHM techniques: model-based [3] and data-driven approaches [4]. Barthorpe [5] compares these approaches and discusses the advantages and limitations of both.

A model-based approach involves the construction of a high-fidelity model of the system, typically a finite element (FE) model. Data acquired from an undamaged state of the system may be used to verify or update such a model. One then performs damage detection by comparing responses from the model and the actual system, and observing discrepancies. The major drawback of such an approach is that evaluation of responses from these models is generally computationally prohibitive. For example, when simulating guided ultrasonic waves for damage detection, a finite element model with a high number of elements is necessary for reliable simulation, which increases the computational effort dramatically [6].

In a data-driven approach, however, a surrogate model constructed from acquired data replaces the high-fidelity model. This derives from the idea that capturing the physics involved in a system becomes more challenging as the complexity of the system increases. The aim of a surrogate model, on the other hand, is not necessarily to replicate system behavior, but only to achieve efficient damage detection by accentuating and quantifying the change in a system due to the onset of damage. Moreover, because of the vast improvements in sensor technology, the volume of acquired data is on the rise as well [7]. In light of this, the surrogate model is typically a statistical model equipped for any of the first three levels of SHM (Rytter [8] defines the four levels of SHM).

Statistical learning algorithms are equipped for construction of such models necessary for data-driven approaches, with the help of data acquired from sensors deployed on systems. Typically, these algorithms are classified as either supervised or unsupervised [9]. Supervised statistical learning algorithms require labels for the data used for surrogate model construction. Unsupervised algorithms, however, do not require data labels, and are used primarily for pattern recognition in the data. In addition, unsupervised algorithms can be used in conjunction with supervised learning algorithms.

This chapter is a brief review of the applications of statistical learning algorithms in data-driven SHM. We discuss the utility of using these algorithms for the purpose of damage detection. We also present a few examples of applications of these algorithms. However, algorithmic details of specific learning algorithms have been avoided.

## Statistical Learning

In this section, we briefly discuss statistical learning and illustrate the applicability of such algorithms for damage detection. As Bousquet et al. [10] state, *The main goal of statistical learning theory is to provide a framework for studying the problem of inference, that is of gaining knowledge, making predictions, making decisions or constructing models from a set of data.* This resonates with the concepts of SHM and damage detection. As discussed earlier, for data-driven approaches to SHM, we use acquired data for model construction, and decisions are made based on predictions from these models.

Data are generally one of two types: one with labels and the other unlabeled. Data labels are typically used to signify the origin of the data from a specific source or source condition. When a statistical learning algorithm uses labeled data for model construction and label predictions for future data, it is called a supervised learning approach. The alternative is an unsupervised learning scheme, where data labels are either unavailable or simply not used. These algorithms, although they do not make direct predictions about data labels, are capable of revealing inherent structures and patterns in the data that may be crucial. We will later show how this class of learning algorithms can be efficiently implemented for damage detection.

Supervised learning can also be divided into two types, namely regression and classification. The key distinction between the two approaches is in the output variables obtained, which are continuous for the former and discrete for the latter. All supervised learning algorithms require two stages. The first is training. This involves the use of part of the data, referred to as the *training data*, for model construction. Predictions are made using the model for the *test data*. These data are used as an input to the constructed model for obtaining the desired predictions or outputs. For the selection of algorithmic parameters, a process known as *cross validation* is performed to optimize the parameters to the given data, using the training data set. For further details, the reader is referred to the books by Hastie et al. [9] and Witten et al. [11].

Based on the above discussion, it is clear that statistical learning algorithms are best suited for data-driven SHM. A primer on the applicability of the various approaches discussed above can be found in Worden and Manson [12]. The following sections will provide a few examples where statistical learning algorithms have been applied for SHM. While these examples are by no means exhaustive, we try to cover a range of applications of a variety of learning algorithms.

## Applications of Statistical Learning Algorithms

### Supervised Learning

#### Classification Techniques

In this section, we discuss a few applications of supervised learning algorithms. As explained earlier, supervised learning algorithms can be either regression or classification problems. We will discuss some examples of classifiers first, followed by illustrations of applications of various regression algorithms in damage detection.

The Naïve-Bayes (NB) classifier is one of the most elemental base learners available in a data scientist's arsenal. Once trained, it classifies data based on conditional probability density functions constructed from the training data. The idea is to define class probabilities. If  $X$  is the data set,  $Y$  is the class variable, and there are  $K$  possible classes, then the classification rule for an NB classifier (probability that variable  $Y = k$ ), based on the Bayes theorem, is defined as:

$$P(Y = k|X) = \frac{P(X|Y = k)P(Y = k)}{\sum_{i=1}^K P(X_i|Y = k)P(Y = k)} \quad (1)$$

Addin et al. [13] used an NB classifier for damage detection in composite materials using Lamb waves. The authors used underlying distinction in Lamb wave signals generated by different kinds of damage. Muralidharan and Sugumaran [14] also used an NB classifier for fault detection in centrifugal pumps; however, they demonstrated that a Bayes net classifier produced higher accuracy for their experiments. An extension of the NB classifier is linear discriminant analysis (LDA) with Fisher's discriminant. The key idea is to evaluate optimum discriminant functions that maximize the difference between various classes of data. For this, LDA assumes a Gaussian distribution for the random variable  $X|Y = k$ , with the same covariance for all classes. Farrar et al. [15] used LDA on acceleration data acquired from a concrete column for damage detection. Guadenzi et al. [16] also used LDA for low-velocity impact damage detection in laminated composite plates.

Support vector machines (SVM) are another classifier that constructs linear or nonlinear functions that separate various classes of data in the sample space. It is based on maximizing the distance of data points, belonging to distinct classes, to these separating hyperplanes. SVM solves the following optimization problem in a sample space:

$$\begin{aligned} \min_{\beta_0, \beta} \quad & \|\beta\|_2^2 \\ \text{subject to} \quad & y_i(x_i^T \beta + \beta_0) \geq 1 \quad \forall i = 1, \dots, n. \end{aligned} \quad (2)$$

where  $x_i$  is a data point in the sample space,  $\beta$  is the slope of a normal to optimal separating hyperplane,  $\beta_0$  is an intercept, and  $y_i$  is a class variable that can take values

of either +1 or -1. This variable essentially aids in the binary classification of the data.

Typically, SVMs are designed for solving two-class classification problems. However, they can be extended to solve multi-class problems as well. He and Yan [17] used an SVM-based damage detection scheme for a spherical lattice dome, where feature extraction was performed using wavelet transform of the ambient vibration data of the structure. Shimada et al. [18] employed SVM for damage detection in power distribution poles. Bulut et al. [19] proposed an SHM system for the Humboldt Bay Bridge using SVM, in which data acquired from a numerical impact test on the bridge were used for damage detection. Chattopadhyay et al. [20] used a nonlinear version of SVM for damage detection in composite laminates. A nonlinear version implies that the optimal separating hyperplane is a nonlinear function. Bornn et al. [21] employed SVM in conjunction with an autoregressive (AR) model for damage detection as well as localization by signal reconstruction and residual estimation. They reconstructed the signal using an AR model and used SVM for damage classification. For nonlinear SVM, they used a Gaussian kernel. Worden and Lane [22] demonstrated the general efficacy of SVM as a statistical learning algorithm for classification problems in engineering.

Trees are another type of popular base learner for classification. They are top-down recursive partitions of the feature or data space via binary splits. Although they are poor predictors, they are adept at making interpretations from data sets that eventually aid in decision making. Kilundu et al. [23] employed decision trees to develop an early detection system for bearing damage, in which they used features extracted from vibration data of bearings to train the trees. Vitola et al. [24] also demonstrated the use of trees for damage detection in aluminum plates using high-frequency waves, and showed the efficacy of *bagging* and *boosting* trees. Bagging, which stands for bootstrap aggregation, as the name suggests, involves bootstrapping (re-sampling) of training data and application of multiple trees to the enlarged data set. The key idea is to reduce the variance in estimates based on the law of large numbers. Boosting, on the other hand, does not involve data duplication. It is a weighted congregation of results from weak base learners trained using a randomly allocated portion of the training data.

## Regression Techniques

Regression algorithms is also used for the purpose of damage detection. As discussed earlier, regression algorithms are generally used to build models consisting of continuous variables using only available data. The output of such regression models is hence not discrete. The most classical form of regression is linear regression or least squares, which attempts to solve the following:

$$\mathbf{Y}_{tr} = \mathbf{X}_{tr}\beta \quad (3)$$

where  $\mathbf{Y}_{tr} \in \mathbb{R}^{n \times 1}$  is the vector of training labels,  $\mathbf{X}_{tr} \in \mathbb{R}^{n \times p}$  is the training data matrix, and  $\beta \in \mathbb{R}^{p \times 1}$  are the coefficients to be estimated by the following optimization problem:

$$\min_{\beta} \quad \|\mathbf{Y}_{tr} - \mathbf{X}_{tr}\beta\|_2^2 \quad (4)$$

where  $\|\cdot\|_2$  is the  $\ell_2$ -norm. The optimal  $\beta$  obtained from the above equation is:

$$\hat{\beta} = (\mathbf{X}_{tr}^T \mathbf{X}_{tr})^{-1} \mathbf{X}_{tr}^T \mathbf{Y}_{tr} \quad (5)$$

The  $\hat{\beta}$  obtained above is then used for making label predictions for a test data set:

$$\mathbf{Y}_{tst} = \mathbf{X}_{tst} \hat{\beta} \quad (6)$$

where the subscript *tst* stands for test data sets. Pan [25] and Shahidi et al. [26] used linear regression for damage detection in structures. They formulated the modal expansion of a dynamic response as a linear regression problem, and used the deviation in slope between the damaged and undamaged models as a damage feature. They demonstrated the use of both single- and multi-variable linear regression.

However, least squares regression suffers drawbacks related to collinearity (results in a singular  $\mathbf{X}_{tr}^T \mathbf{X}_{tr}$ ), high dimensionality of data, and computational issues associated with matrix inversion. To overcome these problems, regularization terms are added to the optimization problem. Two popular regression techniques that can overcome these issues can be found in the literature: Tikhonov regularization or ridge regression, and Sparse Regression.

Ridge regression is based on constraining the elements of vector  $\beta$  evaluated using least squares. The updated optimization problem is as follows:

$$\begin{aligned} \min_{\beta} \quad & \|\mathbf{Y}_{tr} - \mathbf{X}_{tr}\beta\|_2^2 \\ \text{subject to} \quad & \|\beta\|_2^2 \leq t \end{aligned} \quad (7)$$

The more popular Lagrange equivalent form of ridge regression is:

$$\min_{\beta} \quad \|\mathbf{Y}_{tr} - \mathbf{X}_{tr}\beta\|_2^2 + \lambda \|\beta\|_2^2, \quad \lambda \geq 0 \quad (8)$$

where  $\lambda$  is the regularization parameter. The solution to the ridge regression optimization problem is:

$$\hat{\beta}_{ridge} = (\mathbf{X}_{tr}^T \mathbf{X}_{tr} + 2\lambda \mathbf{I})^{-1} \mathbf{X}_{tr}^T \mathbf{Y}_{tr} \quad (9)$$

Sparse regression is another approach for overcoming the drawbacks of least squares regression. A vector of length  $n$  is  $k$ -sparse, if and only if  $k \ll n$ . If the need of a problem statement is a sparse  $\beta$ , the above optimization problem can be further modified by replacing the  $\ell_2$ -norm by an  $\ell_1$ -norm. Although the  $\ell_1$ -norm is a weak

definition of sparsity relative to the  $\ell_0$ -norm, it allows for the optimization problem to remain convex, guaranteeing a global optimum. The use of  $\ell_1$ -norm instead of  $\ell_0$ -norm for sparse regression is referred to as a convex relaxation. The optimization problem associated with sparse regression is:

$$\min_{\beta} \|\mathbf{Y}_{tr} - \mathbf{X}_{tr}\beta\|_2^2 + \lambda\|\beta\|_1, \quad \lambda \geq 0 \quad (10)$$

where  $\lambda$  again is the regularization parameter. The choice of the parameter  $\lambda$  for both ridge and sparse regression is made by performing cross-validation studies on the training data set.

Zhang and Xu [27] compared both regularized algorithms for application to vibration-based damage detection and showed that sparse regularization performed better. Yang and Nagarajaiah [28] used sparse regression to solve the damage detection problem in a sparse representation framework. The key idea was to construct a dictionary of features for a variety of possible damage scenarios. A sparse vector was then used to point to a specific element (a specific damage scenario) of the dictionary which best defined a given test signal. Yang and Nagarajaiah [29] also use sparse regression in a novel sparse component analysis technique that performs blind identification using limited sensor data.

## *Unsupervised Learning*

In this section, we discuss the various applications of unsupervised learning in damage detection. As mentioned earlier, unsupervised learning does not use data labels; rather, they exhibit inherent patterns and structures useful for damage detection in the data. The most common unsupervised algorithms are principal component analysis (PCA), independent component analysis (ICA), and clustering algorithms. Clustering algorithms can range from  $k$ -means clustering to hierarchical clustering.

PCA is well equipped for performing two tasks, namely, data compression and data visualization. The notion of principal components is mathematically defined as the eigenvectors of the covariance matrix of a data set. Physically, they represent the directions of maximum variances. The data, when projected onto these eigenvectors, help to observe patterns in the data that aid in damage detection. For a data matrix  $\mathbf{X} \in \mathbb{R}^{n \times p}$ , the covariance matrix is defined as  $\mathbf{X}^T\mathbf{X}$ . It can be shown that the eigenvectors of the covariance matrix are equivalent to the right singular vectors  $\mathbf{V}$  of  $\mathbf{X}$ , defined as  $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ . The projection of the data in the eigenvector space is then defined by  $\hat{\mathbf{X}} = \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{D}$ .

As discussed earlier, PCA is used in most damage detection studies for reducing the dimensions of the data. This helps increase the numerical efficiency of most algorithms. The idea is to reconstruct the data matrix using eigenvectors with significant singular values only. Zang and Imregun [30] used PCA for frequency response function (FRF) data reduction for application to an artificial neural network (ANN)-based

damage detection scheme. Yan et al. [31, 32] employed PCA for damage detection in bridges under varying environmental conditions. They used a portion of the undamaged data for obtaining the principal component directions, which were subsequently used for estimating changes between damaged and undamaged data from projections along these directions. Tibaduiza et al. [33] included a brief review of applications of PCA for SHM.

Similar to PCA, ICA too is a matrix factorization algorithm. The algorithm was first developed for solving the blind source separation (BSS) problem. It is defined as the retrieval of a set of  $k$  independent signals of length  $n$ ,  $\mathbf{S} \in \mathbb{R}^{k \times n}$  and a corresponding mixing matrix  $\mathbf{A} \in \mathbb{R}^{k \times k}$  from a set of recorded signals  $\mathbf{X} \in \mathbb{R}^{k \times n}$  that are assumed to be a mixture of the independent signals. Mathematically, this can be expressed as:

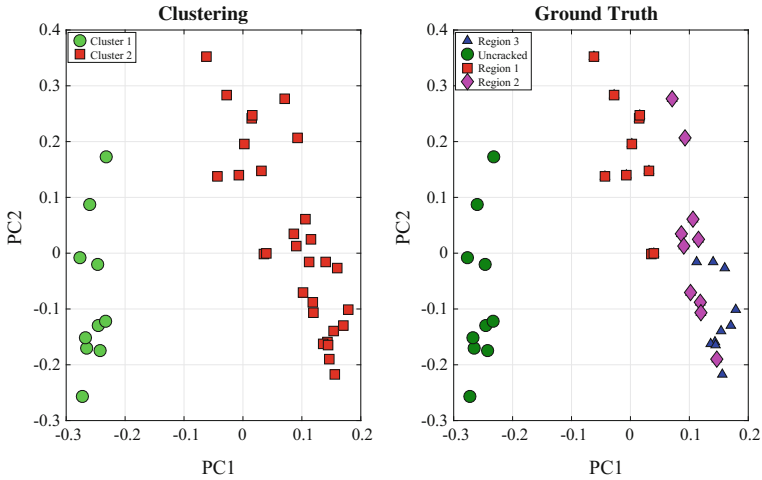
$$\mathbf{X} = \mathbf{AS} \quad (11)$$

The key distinction between PCA and ICA is that PCA evaluates projection directions for maximum variance, whereas ICA evaluates most statistically independent vectors that can construct the given data set. Tibaduiza et al. [34] compared the performance of PCA with ICA for damage detection by active sensing in plates using high-frequency waves. They reported that ICA did not necessarily hold an advantage over PCA for their application. Similar to that with PCA, Zang et al. [35] employed ICA for data reduction for use in FRF damage classification using an ANN. Yang and Nagarajaiah [36] used ICA in conjunction with wavelet transform for blind damage identification.

Another class of unsupervised learning algorithms is the clustering algorithms. Although they cannot perform data reduction like PCA and ICA, they are effective in recognizing different patterns in data that might be useful for damage detection. Da Silva et al. [37] compared the performance of two fuzzy clustering algorithms for damage classification from vibration-based data. Park et al. [38] used a  $k$ -means clustering algorithm for damage detection using electromechanical impedance. The clustering algorithm was used on principal component projections of the impedance data acquired from an experimental beam. Santos et al. [39] employed a mean shift clustering algorithm for damage detection on the data from the Z-24 bridge. They also compared the performance of the proposed algorithm to that of  $k$ -means clustering, fuzzy  $c$ -means clustering, and Gaussian mixture model (GMM)-based techniques. Nair and Kiremidjian [40] demonstrated the use of GMMs for damage detection using vibration data from a benchmark building. Sen and Nagarajaiah [41] used hierarchical clustering for developing a semi-supervised learning approach to damage detection in steel pipes using guided ultrasonic waves. The objective was to detect the presence of damage using a minimum number of piezoelectric actuators and sensors.

Figure 1 shows a typical result of the application of the semi-supervised learning algorithm. The data are acquired from a single actuator and sensor pair. A narrow-band pulse ensures minimal dispersion of Lamb waves propagating in the steel pipes. The central frequency of the pulse used for actuation is appropriately selected to min-





**Fig. 1** Results from Sen and Nagarajaiah [41]. This figure shows the efficacy of hierarchical clustering in detecting the presence of damage in a steel pipe

imize the number of propagating wave modes. In both figures, data are projected onto a two-dimensional plane whose basis vectors are the first two principal components. Figure 1b shows the actual labels of each data point, the damaged and the undamaged data acquired from a steel pipe. The damaged data points are of three types, namely, regions 1, 2, and 3. These regions specify the location of the damage in relation to the location of the actuator and sensor. Figure 1a shows the clusters obtained from hierarchical clustering. Clearly, clusters 1 and 2 consist of all the undamaged data and damaged data, respectively, thus demonstrating the efficacy of the proposed approach.

## Conclusions

In this chapter, we have reviewed some of the most popular techniques from statistical learning that have been used for damage detection, as well as SHM in general. We discussed the application of both supervised and unsupervised learning algorithms in this field. With the advent of a data deluge in the field of SHM, a move from model-based algorithms to data-driven approaches is necessary. The use of statistical learning algorithms not only makes for robust and efficient SHM systems, but also leads to a reduction in computational effort relative to model-based methods, which typically involve high-fidelity modeling of systems. This lays the foundation for developing online SHM systems aimed at performing damage detection in real time, with minimal human interference.

## References

1. Farrar C, Worden K (2007) An introduction to structural health monitoring. *Philos Trans Roy. Soc. A* 365:303–315
2. Farrar CR, Lieven NAJ (2007) Damage prognosis: the future of structural health monitoring. *Philos Trans Roy. Soc. A* 365:632–632
3. Park KC, Reich GW (1999) Model-based health monitoring of structural systems: progress, potential and challenges. In: *Proceedings of the 2nd international workshop on structural health monitoring*, pp. 82–95
4. Farrar CR, Worden K (2012) *Structural health monitoring: a machine learning perspective*. Wiley
5. Barthorpe RJ (2010) *On model-and data-based approaches to structural health monitoring*. PhD thesis, University of Sheffield
6. Gopalakrishnan S (2009) *Modeling aspects in finite elements*. In: *Encyclopedia of structural health monitoring*. Wiley
7. Nagarajaiah S, Yang Y (2017) Modeling and harnessing sparse and low-rank data structure: a new paradigm for structural dynamics, identification, damage detection and health monitoring. *Struct Control Health Monit* 24(1). <http://doi.org/10.1002/stc.1851>
8. Rytter A (1993) *Vibration based inspection of civil engineering structures*. PhD thesis, Aalborg University
9. Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer
10. Bousquet O, Boucheron S, Lugosi G (2004) Introduction to statistical learning theory. *Lect Notes Artif Intel* 3176:169–207
11. James G, Witten D, Hastie T, Tibshirani R (2013) *An introduction to statistical learning*. Springer
12. Worden K, Manson G (2006) The application of machine learning to structural health monitoring. *Proc R Soc A* 365:515–537
13. Addin O, Sapuan SM, Mahdi E, Othman M (2007) A naïve-bayes classifier for damage detection in engineering materials. *Mater Des* 28:2379–2386
14. Muralidharan V, Sugumaran V (2012) A comparative study of naïve bayes classifier and bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. *Appl Soft Comp* 12:2023–2029
15. Farrar CR, Nix DA, Duffey TA, Cornwell PJ, Pardoen GC (1999) Damage identification with linear discriminant operators. In: *Proceedings of the international modal analysis conference*. 599–607
16. Gaudenzi P, Nardi D, Chiapetta I, Atek S, Lampani L, Sarasini F, Tirillo J, Valente T (2015) Impact damage detection in composite laminate plates using an integrated piezoelectric sensor and actuator couple combined with wavelet based features extraction approach. In: *Proceedings of the 7th ECCOMAS thematic conference on smart structures and materials*
17. He H-X, Yan W-M (2007) Structural damage detection with wavelet support vector machine: introduction and applications. *Struct Control Health Monit* 14:162–176
18. Shimada M, Mita A, Feng MQ (2006) Damage detection of structures using support vector machines under various boundary conditions. In: *Proceedings of SPIE*, 6174(61742K-1)
19. Bulut A, Singh AK, Shin P, Fountain T, Jasso H, Yan L, Elgamal A (2004) Real-time nondestructive structural health monitoring using support vector machines and wavelets. In: *Technical report*, University of California, Santa Barbara
20. Chattopadhyay A, Das S, Coelho CK (2007) Damage diagnosis using a kernel-based method. *Insight-Non-Destr Test Cond Monit* 49:451–458
21. Bornn L, Farrar CR, Park G, Farinholt K (2009) Structural health monitoring with autoregressive support vector machines. *J Vibri Acoust* 139:021004
22. Worden K, Lane AJ (2001) Damage identification using support vector machines. *Smart Mater Struct* 10:540–547

23. Kilundu B, Dehombreux P, Letot C, Chimentin X (2009) Early detection of bearing damage by means of decision trees. *J Autom Mob Robot Intel Syst* 3(3):70–74
24. Vitola J, Tibaduiza M, adn Anaya D, Pozo F (2016) Structural damage detection and classification based on machine learning algorithms. In: 2016 Proceedings of the 8th european workshop on structural health monitoring EWSHM
25. Pan Y (2012) Linear regression based damage detection algorithm using data from a densely clustered sensing system. Master's thesis, Lehigh University
26. Shahidi SG, Nigro MB, Pakzad SN, Pan Y (2015) Structural damage detection and localisation using multivariate regression models and two-sample control statistics. *Struct Infrastruct Eng* 11(10):1277–1293
27. Zhang CD, Xu YL (2016) Comparative studies on damage identification with tikhonov regularization and sparse regularization. *Struct Control Health Monit* 23:560–579
28. Yang Y, Nagarajaiah S (2014) Structural damage identification via a combination of blind feature extraction and sparse representation classification. *Mech Syst Sig Process* 45(1):1–23
29. Yang Y, Nagarajaiah S (2013) Output-only modal identification with limited sensors using sparse component analysis. *J Sound Vibration* 332:4741–4765. <http://doi.org/10.1016/j.jsv.2013.04.004>
30. Zang C, Imregun M (2001) Structural damage detection using artificial neural networks and measured frf data reduced via principal component projection. *J Sound Vib* 242(5):813–827
31. Yan A-M, Kerschen G, De Boe P, Golinval J-C (2005) Structural damage diagnosis under varying environmental conditions-part I: a linear analysis. *Mech Syst Sig Process* 19:847–864
32. Yan A-M, Kerschen G, De Boe P, Golinval J-C (2005) Structural damage diagnosis under varying environmental conditions-part II. *Mech Syst Sig Process* 19:865–880
33. Tibaduiza DA, Mujica LE, Rodellar J (2013) Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Struct Control Health Monit* 20:1303–1316
34. Tibaduiza DA, Mujica LE, Anaya M, Rodellar J, Güemes A (2012) Principal component analysis vs independent component analysis for damage detection. In: Proceedings of the 6th European workshop on structural health monitoring
35. Zang C, Friswell MI, Imregun M (2004) Structural damage detection using independent component analysis. *Struct Health Monit* 3(1):69–83
36. Yang Y, Nagarajaiah S (2014), Blind identification of damage in time-varying systems using independent component analysis with wavelet transform. *Mech Syst Signal Process* 47:3–20. <http://doi.org/10.1016/j.ymsp.2012.08.029>
37. da Silva S, Júnior MD, Junior VL, Brennan MJ (2008) Structural damage detection by fuzzy clustering. *Mech Syst Sigl Process* 22:1636–1649
38. Park S, Lee J-J, Yun C-B, Inman DJ (2008) Electro-mechanical impedance-based wireless structural health monitoring using pca-data compression and k-means clustering algorithms. *J Intel Mater Syst Struct* 19:509–520
39. Santos A, Silva M, Santos R, Figueiredo E, Sales C, Costa JCWA (2016) Output-only structural health monitoring based in mean shift clustering for vibration-based damage detection. In: Proceedings of the 8th European workshop on structural health monitoring
40. Nair KK, Kiremidjian AS (2007) Time series based structural damage detection algorithm using gaussian mixtures modeling. *J Dyn Syst Measur Control* 129:285–293
41. Sen D, Nagarajaiah S (2016) Damage detection in steel pipes using a semi-supervised statistical learning algorithm. In: ASNT annual conference. Long Beach, CA