



Sharif University of Technology  
Scientia Iranica  
Transactions A: Civil Engineering  
<http://scientiairanica.sharif.edu>



# Improving feature extraction via time series modeling for structural health monitoring based on unsupervised learning methods

A. Entezami, H. Shariatmadar\*, and A. Karamodin

Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad, P.O. Box 91775-1111, Iran.

Received 12 February 2017; received in revised form 12 March 2018; accepted 2 July 2018

## KEYWORDS

Structural health monitoring;  
Statistical pattern recognition;  
Feature extraction;  
Time series modeling;  
Residual extraction;  
Unsupervised learning;  
Andrews plot;  
Clustering analysis.

**Abstract.** Feature extraction by time series modeling based on statistical pattern recognition is a powerful approach to Structural Health Monitoring (SHM). Determination of an adequate order and identification of an appropriate model play prominent roles in extracting sensitive features to damage from time series representations. Early damage detection under statistical decision-making via high-dimensional features is another significant issue. The main objectives of this study were to improve a residual-based feature extraction method by time series modeling and to propose a multivariate data visualization approach to early damage detection. A simple graphical tool based on Box-Jenkins methodology was adopted to identify the most compatible time series model with vibration time-domain measurements. Furthermore, *k*-means and Gaussian Mixture Model (GMM) clustering techniques were utilized to examine the performance of the residuals of the identified model in damage detection. A numerical concrete beam and an experimental benchmark model were applied to verifying the improved and proposed methods along with comparative analyses. Results showed that the approaches were successful and superior to a state-of-the-art order determination technique in obtaining a sufficient order, generating uncorrelated residuals, extracting sensitive features to damage, and accurately detecting early damage by high-dimensional data.

© 2020 Sharif University of Technology. All rights reserved.

## 1. Introduction

Early detection of damage is the initial step and key component of Structural Health Monitoring (SHM) in civil engineering systems, because there is no engineering and economic justification to reconstruct most of the large and complex infrastructures such as bridges, towers, and dams. The process of SHM initially aims

to evaluate the global structural condition and detect early damage. Subsequently, one needs to locate and quantify damage, which are local SHM procedures. In the engineering literature, damage is defined as an adverse change in a structure that leads to undesirable alterations in the structural behavior and performance. It may appear as cracks in concrete, loose bolts and broken welds in steel connections, corrosions, fatigue, etc. All of them may cause unfavorable stresses and displacements, inappropriate vibration, failure, and even collapse [1].

Currently, most of the methods in SHM focus on the statistical pattern recognition paradigm. The reason is that all SHM problems are subject to various degrees of uncertainty. Therefore, the statistical

\*. Corresponding author. Tel.: +98 51-38805129  
E-mail addresses: entezami@mail.um.ac.ir (A. Entezami);  
shariatmadar@um.ac.ir (H. Shariatmadar);  
a-karam@um.ac.ir (A. Karamodin)

approach to pattern recognition appears to stand out as a natural way for SHM applications. On the other hand, this paradigm can assist engineers in applying raw vibration data and implementing early damage detection, localization, and quantification based on data-driven methods for damage classification. These methods are run in the four steps of operational evaluation, data acquisition, feature extraction, and statistical decision making [1]. Operational evaluation is concerned with the lifestyle, economic justification, limitations, and possibility of performing the SHM. Data acquisition process includes choosing the excitation tools (force or ambient vibration), the sensor types (wired or wireless), the number and placement of sensors, and the design of sensing systems. However, the majority of technical studies focus on the steps of feature extraction and statistical decision-making. A comprehensive review of these steps can be found in [2,3].

From the statistical pattern recognition perspective, feature extraction is intended to find out meaningful information from raw vibration data through advanced signal processing techniques [4]. Sensitivity to damage is the main characteristic of such information, which can be translated to damage-sensitive features. On the other hand, statistical decision-making utilizes statistical methods based on machine learning algorithms to discriminate the damaged state of the structure from the normal condition. In general, the machine learning algorithms are categorized in supervised and unsupervised learning classes, both of which mainly aim to train statistical models or classifiers by training data and to make a decision about the problem of damage (e.g. early damage detection) via testing data [1]. In the supervised learning class, one needs to use features from both the undamaged and damaged conditions, whereas the unsupervised learning class only requires the features of undamaged state to learn the classifier of interest. This noteworthy specification of unsupervised learning class makes it more beneficial than the supervised learning strategy for SHM. This is because it may not be conceivable to detect damage in large and complex structures in an effort to establish a supervised learning framework. Novelty detection [1,5,6] and clustering analysis [7,8] are well-known unsupervised learning techniques for early damage detection.

Time series modeling is one of the powerful and promising approaches to feature extraction [9]. The major advantage of this approach is utilizing raw time series data (i.e., excitation and/or responses) for finding out the damage-sensitive features. This benefit provides the great opportunity of neglecting the transformation of raw time series data into frequency and/or modal domains. As another advantage, time series representations use a few samples of time series

dataset to describe it [10]. Depending upon the nature, type, and dimension of time series data, there is a wide range of models that give an appropriate diversity for feature extraction. Eventually, the main merit of time series modeling in the context of SHM is that some statistical characteristics of time series representations (e.g. the model parameters and residuals) are sensitive to damage. In case of measuring linear and stationary (time-invariant) vibration data, some widely used time series models for feature extraction are autoregressive (AR) [11], autoregressive with exogenous input (ARX) [9], autoregressive-autoregressive with exogenous input (ARARX) [12], autoregressive moving average (ARMA) [13], and autoregressive moving average with exogenous input (ARMAX) [14].

In spite of numerous applications of time series modeling to feature extraction, some important issues and limitations should be dealt with. The significant issue in time series modeling is concerned with the determination of an adequate and correct order. Because it directly affects the model sufficiency and accuracy, one essentially needs to select a robust order that enables the model of interest to generate uncorrelated residuals. From a statistical viewpoint, the residual sequences of a time series representation should not be dependent or correlated, in which case one can realize the insufficiency and inaccuracy of model order [15]. In this regard, it is significant to improve the time series modeling and guarantee the model accuracy and adequacy by generating uncorrelated residuals. From the engineering aspect, an improper order does not allow the time series model to capture the underlying dynamics of structure, which may lead to extract-insensitive features to damage and weak detectability of damage [16]. Therefore, a reliable feature extraction via time series modeling depends strongly on obtaining an accurate and sufficient order. The main limitation in this aspect is that the classical order selection techniques may not yield a proper order and cause time series models to produce uncorrelated residuals.

The other important issue in time series modeling is to identify an appropriate model. Although diverse types of time series models have been utilized to extract the damage-sensitive features for SHM applications, the limitation is that little attention has been paid to model identification for feature extraction. The main reason for the importance of this issue is the availability of a broad range of time series representations that seem suitable for modeling; however, they may yield poor fits resulting from incompatibility with the nature of time series data or may lead to a time-consuming and complex process by incorporating redundant orders.

Considering the above-mentioned issues and limitations, this study proposes an improved residual-based feature extraction method via time series modeling based on improvements in order determination and

selection. Box-Jenkins methodology is presented to design a simple graphical tool for model identification. The main advantage of the improved method is extracting sensitive features (the model residuals) to damage with an accurate and sufficient order that guarantees extracting uncorrelated residual samples. A multivariate data visualization approach called Andrews plot is proposed to detect early damage using the model residuals obtained from the undamaged and damaged conditions. The main contribution of this approach is establishing a graphical and simple decision-making framework for damage detection. Appropriateness to high-dimensional data is the great benefit of the Andrews plot. Additionally, two efficient and successful clustering techniques, namely  $k$ -means and Gaussian Mixture Model (GMM), are applied to examining the reliability and performance of the damage-sensitive features extracted from the improved feature extraction method for early damage detection. A numerical concrete beam and an experimental benchmark laboratory frame are utilized to demonstrate accuracy and effectiveness of the improved and proposed methods. As will be shown, these methods are successful in obtaining the accurate and adequate order, extracting sensitive features to damage, and efficiently detecting early damage.

## 2. Review of literature

Although it is common to utilize non-time series-based techniques such as approaches based on the extraction of modal data in either data-driven or model-driven framework, some significant limitations of these techniques, including low sensitivity to local damage, computational difficulties, and uncertainties for complex structures, raise time series modeling as a powerful data-driven strategy for feature extraction [9,17]. In this regard, Farrar and Jauregui [18,19] conducted comparative studies to demonstrate major limitations related to the use of modal data for damage detection.

A key component of employing time series representations in feature extraction is to determine an adequate and correct order. Figueiredo et al. [16] investigated the effect of different orders of AR model on damage detection by four classical order determination techniques. They selected three types of AR order and concluded that an inappropriate choice would lead to weak damage detectability. Gul and Necati Catbas [5] utilized the simple partial autocorrelation function for obtaining the order of AR model without any investigation into the model accuracy and adequacy for generating uncorrelated residuals. The applications of the well-known information criteria for order selection such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) can be found in [9,11]. The main limitation on using such

techniques is that they may suggest different model orders and should be used with careful judgment [13]. On the other hand, as mentioned in the previous section, such approaches may not determine sufficient orders for extracting the uncorrelated residuals.

In relation to the residual-based feature extraction algorithm, Fugate et al. [20] utilized the residuals of AR model and statistical control charts for early damage detection. Gul and Necati Catbas [21] presented a new sensor clustering method based on ARX model and applied its residual sequences with the aid of a damage indicator to damage detection. For locating structural damage, Roy et al. [22] proposed an ARX-model-based damage localization framework by using the model residuals and some damage indicators based on statistical hypothesis tests. The main difference between the improved residual-based feature extraction method presented here and the conventional technique used in the above-mentioned studies is that the former guarantees that the residual sequences extracted from the identified model and the improved order are uncorrelated. This does not only ensure the model accuracy and adequacy, but also lead to the extraction of sensitive features to damage with an accurate damage detection. As the other difference, in the improved method, the maximum amount of improved orders at all sensors is fitted to the vibration time-domain responses. This process makes sure of extracting the uncorrelated residuals from all sensors.

## 3. Time series modeling

In statistics, time series modeling is a method that attempts to fit a mathematical equation to time series data for some special issues such as data analysis, model identification, parameter estimation, and forecasting [15]. On the other hand, time series modeling is a powerful tool for feature extraction. Since there are various types of time series data (i.e., stationary versus non-stationary, linear versus nonlinear, seasonal versus non-seasonal, etc.) [23], it is necessary to select an appropriate model, estimate its parameters, and validate the model sufficiency and accuracy by generating uncorrelated residuals [15].

Considering the linear and stationary vibration time-domain measurements, time series modeling is usually carried out by time-invariant linear representations such as AR, ARX, ARARX, ARMA, and ARMAX. In general, they are comprised of AR or output, exogeneous (X) or input, and Moving Average (MA) or error terms. The AR model is known as the simplest time series representation that linearly depends on the output data (the vibration response). In the availability of both the input (the measurable and known excitation force) and output data, the ARX model is usable. It is possible to combine these

models with the MA term and produce the ARMA representation for the output-only cases and ARMAX model for the input-output conditions [15]. The general formulation of a time series model by incorporating the input, output, and error terms is expressed as follows:

$$\begin{aligned} y(t) + \theta_1 y(t-1) + \dots + \theta_p y(t-p) &= \varphi_1 u(t-1) \\ &+ \dots + \varphi_q u(t-q) + e(t) + \psi_1 e(t-1) + \dots \\ &+ \psi_r e(t-r), \end{aligned} \quad (1)$$

where  $u(t)$  and  $y(t)$  denote the input and output data at time  $t$ ;  $e(t)$  is the residual sequence, which corresponds to the difference between the measured time series data and the predicted data obtained by the model. In Eq. (1),  $\theta = [\theta_1 \dots \theta_p]$ ,  $\Phi = [\varphi_1 \dots \varphi_q]$ , and  $\Psi = [\psi_1 \dots \psi_r]$  represent the unknown parameters of the model. Moreover, the orders of output, input, and error terms are defined as  $p$ ,  $q$ , and  $r$ , respectively. It is possible to rewrite Eq. (1) in a more compact form as follows:

$$A(z)\mathbf{y}(t) = B(z)\mathbf{u}(t) + D(z)\mathbf{e}(t), \quad (2)$$

where,  $A(z)$ ,  $B(z)$ , and  $D(z)$  are the polynomials in the delay operator  $z^{-1}$ , which can be formulated as:

$$\begin{aligned} A(z) &= 1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_p z^{-p}, \\ B(z) &= 1 + \varphi_1 z^{-1} + \varphi_2 z^{-2} + \dots + \varphi_q z^{-q}, \\ D(z) &= 1 + \psi_1 z^{-1} + \psi_2 z^{-2} + \dots + \psi_r z^{-r}. \end{aligned} \quad (3)$$

It would be interesting to know that Eq. (1) refers to the formulation of ARMAX. Any change in the terms of this formulation leads to obtaining the other types of time-invariant linear representations. For example, if  $r = 0$ , in which case the error term or  $D(z)$  is removed from Eq. (1), the model becomes ARX as follows:

$$A(z)\mathbf{y}(t) = B(z)\mathbf{u}(t) + \mathbf{e}(t). \quad (4)$$

The ARMA model is obtained by setting  $q$  to zero or removing  $B(z)$  from Eq. (1) as:

$$A(z)\mathbf{y}(t) = D(z)\mathbf{e}(t). \quad (5)$$

Finally, the AR model is generated by  $q = r = 0$  and eliminating  $B(z)$  and  $D(z)$ , that is:

$$A(z)\mathbf{y}(t) = \mathbf{e}(t). \quad (6)$$

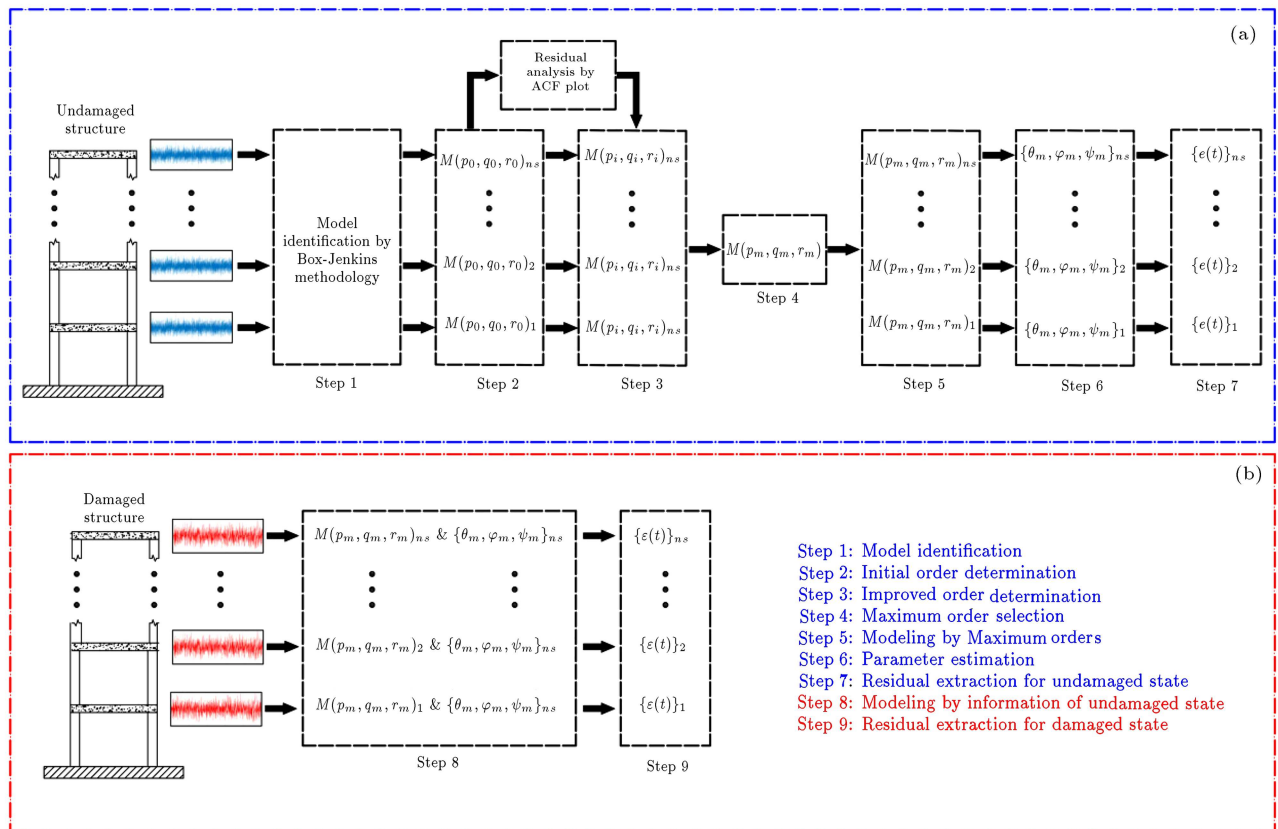
#### 4. An improved residual-based feature extraction method

In the SHM community, the residuals of time series

models are chosen as the damage-sensitive features [9]. Unlike the process of feature extraction by the model parameters, the residual-based feature extraction algorithm exploits the model orders and parameters obtained from the normal condition of the structure in an effort to extract the residual sequences of the damaged state. By obtaining the model information from the only normal condition, one can realize that this algorithm acts in an unsupervised learning manner. The fact beyond the residual-based feature extraction approach is that the model (i.e., its orders and parameters) used in the normal condition will no longer provide a good fit and do not correctly predict the response of the damaged state. Therefore, the residual samples associated with this state will increase [16]. In this case, the increase in the model residuals is an indicator of damage occurrence. The main merit of using the residual-based feature extraction algorithm is that one does not require any order determination and parameter estimation for the damaged structure.

The improved residual-based feature extraction method presented here consists of two stages. The first one belongs to the normal or undamaged condition of the structure. At this stage, one attempts to identify an appropriate model based on the Box-Jenkins methodology, determine adequate and accurate orders, estimate the model parameters, and then extract the uncorrelated residuals of the identified model at each sensor as the damage-sensitive features of the normal condition. On the contrary, the second stage is concerned with the damaged state of the structure. At this stage, the obtained model characteristics (i.e., the orders and parameters) are applied to extracting the residual sequences associated with the damaged state. For the sake of convenience, Figure 1 depicts the flowchart of the improved residual-based feature extraction method. In the following, all steps in this figure are described in details.

**Step 1. Model identification:** This is the initial step of time series modeling. One promising and straightforward way of identifying the most proper time series representation for the stationary data is Box-Jenkins methodology [15]. It relies on using Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF), which are known as important statistical tools for measuring the correlation between time series samples. Under the Box-Jenkins methodology, if the plot of ACF tails off in an exponential decay or a damped sine wave and the plot of the PACF becomes zero after a lag, time series conforms to the AR and ARX models for the output-only and input-output cases, respectively. On the contrary, if the plot of PACF tails off in an exponential decay or a damped sine wave and the plot of the ACF cuts off after a lag, one can select the



**Figure 1.** Flowchart of the improved residual-based feature extraction method: (a) The first stage for an undamaged state and (b) the second stage for a damaged state.

MA representation. Eventually, if the plots of both ACF and PACF tail off in an exponential decay or damped sine waves, ARMA and ARMAX are chosen as the most proper time series models for the output-only and input-output conditions, respectively. In Figure 1, the identified model is designated by “ $M$ ”, which can be one of the time-invariant linear models;

**Step 2. Initial order determination:** In this step, it is attempted to determine the initial orders ( $p_0$ ,  $q_0$ , and  $r_0$ ) at each sensor via one of the state-of-the-art order selection techniques based on the concept of information criterion, such as AIC and BIC [15]. It is significant to point out that  $ns$  denotes the number of sensors mounted on the structure. Moreover, it should be mentioned that the AIC often tends to the overfitting problem (i.e., determining redundant orders that make an elaborate model with a poor forecasting [15]), while the BIC enhances it by adding a rigorous penalty term. For this reason, in this article, the BIC technique is applied to choosing the initial model order. Given an  $n$ -dimensional time series dataset and a model with  $\nu$  parameters, in which  $\nu$  denotes the sum of the model orders (i.e.,  $\nu = p + q + r$  for ARMAX,  $\nu = p + r$  for ARMA,  $\nu = p + q$  for ARX, and  $\nu = p$  for AR), the BIC is

given by:

$$BIC = n \ln(\hat{\sigma}_e^2) + \nu \ln(n), \quad (7)$$

where  $\hat{\sigma}_e^2$  denotes the estimate of the residual variance. To gain the initial orders, one should examine a wide range of orders (e.g. 1 – 100) and choose a number for each of  $p_0$ ,  $q_0$ , and  $r_0$  with the minimum BIC value;

**Step 3. Improved order determination:** Although the information criteria are usually applied to choosing the orders of time series representations, the uncorrelatedness of the residual samples gained by them may not be fully satisfied. Hence, the initial orders are developed to achieve the improved orders ( $p_i$ ,  $q_i$ , and  $r_i$ ). The development is based on observing the correlation of residual sequences by the ACF. If the values of ACF are roughly located between the upper and lower bounds of a confidence interval, one can understand that the model residuals are uncorrelated and  $p_i$ ,  $q_i$ , and  $r_i$  are chosen as the improved orders; otherwise, they should be improved;

**Step 4. Maximum order selection:** For the damage detection problems, it is better to use features (either the model parameters or the model residuals) with the same dimensions for both the

undamaged and damaged states. To do so and deal with the inequality of feature dimensions, the maximum number of each improved order is selected to utilize in the processes of parameter estimation and residual extraction. In this step,  $p_m$ ,  $q_m$ , and  $r_m$  denote the maximum orders for the output, input, and error terms, respectively;

**Step 5. Modeling by the maximum orders:** The identified model in the first step with the maximum orders (e.g.  $AR(p_m)$ ,  $ARMA(p_m, r_m)$ , etc.) is fitted to the vibration responses of all sensors. The main property (advantage) of this type of modeling is guaranteeing the extraction of uncorrelated residuals from all sensors;

**Step 6. Parameter estimation:** The unknown model parameters ( $\theta_m$ ,  $\Phi_m$ , and  $\Psi_m$ ) are estimated by one of the well-known computational techniques such as Least Squares, Burg, Forward-Backward, and Yule-Walker [15]. In this article, the Burg approach is applied to estimating the identified model;

**Step 7. Residual extraction for the undamaged state:** Eventually, the uncorrelated model residuals at each sensor are extracted as the damage-sensitive features for the normal condition;

**Step 8. Modeling by the information of the undamaged state:** In this step, the maximum orders and the model parameters obtained from the normal condition are employed to model the vibration responses of the damaged structure;

**Step 9. Residual extraction for the damaged state:** Similarly to Step 7, the model residuals at each sensor are extracted as the damage-sensitive features for the damaged condition.

## 5. A multivariate data visualization method

In the multivariate data visualization, Andrews plot or Andrews curve is a graphical tool to visualize high-dimensional multivariate data [24]. The function of Andrews plot is based on Fourier series in which variables are the high-dimensional samples. More precisely, this graphical tool calculates a periodic function  $f(t)$ , which is composed of sine and cosine components, to depict each observation of the multivariate dataset. Apart from appropriateness to the high-dimensional data, the other merit of the Andrews plot is the ability to detect outliers or adverse changes in time series data, which makes it a simple and efficient approach to damage detection.

Assuming that  $\mathbf{Y}_{i,j} \in \mathfrak{R}^{nm \times ns}$  is a multivariate time series dataset, where  $i = 1, 2, \dots, nm$  denotes the number of observations and  $j = 1, 2, \dots, ns$  refers to the number of variables, the function  $f(t)$  for the  $i$ th

observation can be formulated in the following forms:

$$f_i(t) = \frac{Y_{i,1}}{\sqrt{2}} + A_{i,2} \sin(t) + Y_{i,3} \cos(t) + \dots + Y_{i,ns-1} \sin\left(\frac{nm-1}{2}t\right) + Y_{i,ns} \cos\left(\frac{nm-1}{2}t\right), \quad (8)$$

$$f_i(t) = \frac{Y_{i,1}}{\sqrt{2}} + Y_{i,2} \sin(t) + Y_{i,2} \cos(t) + \dots + Y_{i,ns} \sin\left(\frac{nm}{2}t\right), \quad (9)$$

where  $t \in [-\pi, \pi]$ . The application of Eq. (8) is related to the cases in which  $ns$  is an odd number. By contrast, if  $ns$  is an even number, one should apply Eq. (9). For the problem of damage detection, it is only necessary to collect the residual sets of both the undamaged and damaged conditions to generate a multivariate time series dataset. Accordingly,  $nm$  is obtained by multiplying the number of undamaged and damaged conditions (the default value is 2) by the number of samples in each residual vector. Under the theory of Andrews plot, it is possible to detect early damage by discerning the curve deviations regarding the damaged state from the normal condition.

## 6. Clustering analysis

Clustering is an unsupervised learning approach intended to arrange large quantities of multivariate data into natural groups or clusters. This approach consists of various algorithms such as hierarchical, partitioning, self-organizing maps, etc. each of which seeks to organize a given dataset into homogeneous clusters [25].

### 6.1. *k*-means clustering

The *k*-means clustering is one of the well-known unsupervised learning methods that falls into a non-hierarchical or partitioning clustering strategy [25,26]. This method simply splits a multivariate dataset into  $k$  predetermined groups or clusters so that the samples within a cluster are similar, whereas the samples from different clusters are quite dissimilar. The *k*-means algorithm is an iterative procedure that assigns the observations of the multivariate dataset to exactly one of the  $k$  clusters defined by centroids. Once the centroids of  $k$  clusters have been determined, the distance of each observation in each cluster from its centroid is computed by means of a distance method, such as Euclidean-Squared Distance (ESD) in the following form:

$$ESD_k = (x_i - c_k)^T(x_i - c_k), \quad (10)$$

where  $x_i$  is the  $i$ th data sample of the  $k$ th cluster. Moreover,  $c_k$  represents the centroid of the  $k$ th cluster, which corresponds to the mean of all data points in this cluster. One merit of the  $k$ -means clustering method is establishing an unsupervised learning manner by partitioning the multivariate datasets without any effort to train a classifier. In other words, one can state that it is a non-model clustering approach [25].

For the process of early damage detection, it is initially necessary to establish an  $nm$ -by- $ns$  multivariate dataset ( $\mathbf{Y} \in \mathbb{R}^{nm \times ns}$ ) and then, perform the  $k$ -means clustering in an effort to determine an  $nm$ -by- $k$  ESD matrix. In each cluster, the distance values of the damaged state are separable from the corresponding quantities of the normal condition.

**6.2. Gaussian Mixture Model (GMM)**

A GMM is a probabilistic model under the assumption that all samples in a dataset are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. This model is composed of  $k$  multivariate normal density components or clusters. The GMM creates a model-based clustering approach, which utilizes multivariate finite mixture models aiming at determining the main clusters/components of datasets. In order to perform a clustering algorithm by the GMM, it is necessary to learn a mixture model (a classifier) by using the features of the undamaged condition of the structure as the training data. Suppose that  $\mathbf{X} \in \mathbb{R}^{n \times ns}$  is an  $n$ -dimensional multivariate dataset. This matrix is equivalent to the model residuals of the normal condition obtained from all sensors. Using  $k$  components, the trained GMM is parameterized by the mean vector ( $\boldsymbol{\mu}_k$ ) and covariance matrix ( $\boldsymbol{\Sigma}_k$ ) for each component in the following form [27]:

$$GMM(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}_k|^{1/2}} \left( -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{x} - \boldsymbol{\mu}_k) \right). \tag{11}$$

The unknown parameters (the mean vector and covariance matrix) are estimated by the classical maximum likelihood estimation based on the expectation maximization algorithm. For the process of early damage detection and clustering the distance of each observation in the testing dataset, an  $nm$ -by- $ns$  multivariate dataset ( $\mathbf{Y} \in \mathbb{R}^{nm \times ns}$ ) from each component of the

GMM is calculated by a distance method such as Mahalanobis-Squared Distance (MSD) as follows:

$$MSD_k = (\mathbf{y} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{y} - \boldsymbol{\mu}_k). \tag{12}$$

For the process of early damage detection, an  $nm$ -by- $k$  MSD matrix can be obtained. In each component, the distance quantities of the damaged state are detectable from the corresponding values of the undamaged condition.

**6.3. Choosing the optimal number of clusters**

Selection of the optimal number of clusters is a key element of a clustering method. In most cases, this process is typically conducted by the training dataset; that is, the damage-sensitive features of the undamaged condition are applied to selecting  $k$ . One reliable way is to employ the methodology based on Silhouette value [28]. It is a measure of how similar an observation of a multivariate dataset is to its own cluster compared to other clusters. The silhouette value varies from  $-1$  to  $1$ . Let  $a_i$  be the average distance between the  $i$ th sample and all other data points within the same cluster. Furthermore,  $b_i$  is defined as the lowest average distance of the  $i$ th sample to all data points in any other cluster. With these definitions, the formulation for obtaining the silhouette value is given by:

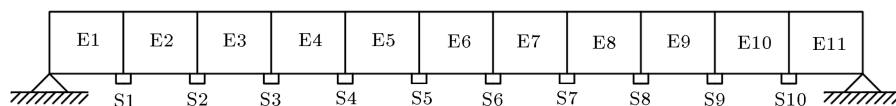
$$S = \frac{b_i - a_i}{\max(a_i, b_i)}. \tag{13}$$

A high silhouette value indicates that  $S$  well matches its own cluster and poorly matches the neighboring clusters. If the majority of data points have high silhouette values, the clustering solution is appropriate; otherwise, one can deduce that the clustering solution may have improper performance with either too many or too few clusters.

**7. Applications**

**7.1. A numerical concrete beam**

In order to verify the accuracy and capability of the improved and proposed methods, a numerical model of the concrete beam is simulated as shown in Figure 2. This model is constructed by the finite element method under Bernoulli-Euler beam theory with the aid of an in-house code implemented in MATLAB environment. Based on this theory, each element of the beam includes four Degrees of Freedom (DoFs), in which case it is discretized by 11 elements (E1-E11), 12 nodes, and 22



**Figure 2.** Numerical model of the concrete beam.

DoFs. Assume that similar damping mechanisms are distributed throughout the beam; hence, the classical damping is an appropriate idealization. Furthermore, Rayleigh damping approach is utilized to construct the damping matrix by using 5% damping ratio for all modes.

The geometry of the beam element is length 300 mm, height 250 mm, and width 250 mm in the cross-section. The material properties of the beam are the modulus of elasticity 22.3 GPa, material density 2400 kg/m<sup>3</sup>, and Poisson coefficient of 0.2. It is assumed that the beam is equipped with ten sensors (i.e., S1-S10 as shown in Figure 2) at the bottom edge to acquire acceleration time histories in the vertical direction. The vibration response of each sensor is measured at 25 sec in 0.003125 sec time intervals (320 Hz sampling frequency), which leads to 8000 data samples. The beam is subjected to different Gaussian white noise signals in the vertical direction to simulate random excitation forces. Furthermore, Newmark method [29] is utilized to implement the simulations for obtaining acceleration time histories.

A single damage as a flexural crack is simulated by reducing the concrete flexural rigidity at the middle-

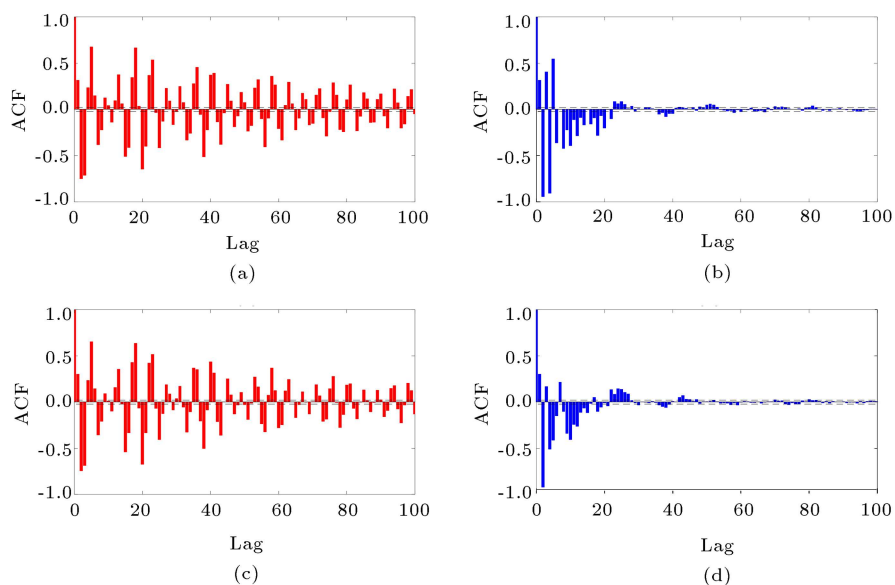
span of the beam (Element 6 or E6). Based on this damage pattern, three incremental damage scenarios are defined at the location of damage. This pattern is a realistic simulation of cracks in the reinforced concrete beams, which is introduced as a common way to use in the numerical applications [30]. Table 1 represents the undamaged and damaged cases of the numerical beam.

At each sensor, the most appropriate time series model is identified using the Box-Jenkins methodology. By using the vibration responses of the beam (the output-only condition), one does not need to apply time series models that require the input terms such as ARX or ARMAX. Hence, the only remaining time-invariant linear models are AR and ARMA. Figure 3 shows the ACF and PACF of the acceleration time histories for sensor 5 in the first and fourth cases, respectively.

From these figures, it is clear that the PACFs become approximately zero after the 30th lag, whereas the ACFs have exponentially decreasing forms without any inclination toward zero. According to the Box-Jenkins methodology, such observations confirm that the acceleration time histories conform to the AR process. Therefore, one should choose the AR

**Table 1.** Structural state conditions in the numerical model of the beam.

Case	Condition	Location	Structural change	Index (%)
1	Undamaged	—	—	0
2	Damaged	E6	Reduction in concrete flexural rigidity ( $EI_c$ )	-10
3	Damaged	E6		-20
4	Damaged	E6		-40



**Figure 3.** Model identification by the Box-Jenkins methodology at sensor 5 of the numerical beam: (a) Auto-Correlation Function (ACF) in Case 1, (b) Partial Auto-Correlation Function (PACF) in Case 1, (c) ACF in Case 4, and (d) PACF in Case 4.



representation to use in the feature extraction. The same conclusion is obtainable for the other sensors and cases.

In order to extract the residuals of AR model as the damage-sensitive feature, the initial and improved AR orders of Case 1 (the normal condition of the beam) should be determined as presented in Table 2. In this regard, the initial orders are gained by the state-of-the-

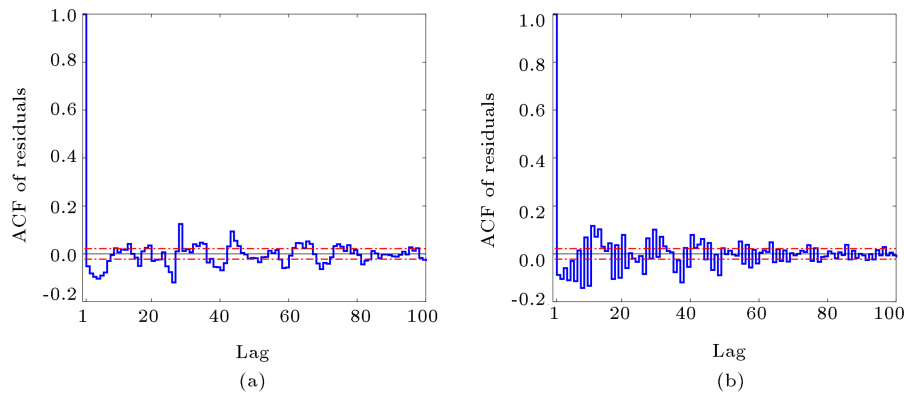
**Table 2.** Determining the initial and improved autoregressive (AR) orders for Case 1.

Sensor no.	Initial orders	Improved orders
1	17	27
2	19	29
3	15	22
4	15	23
5	12	24
6	14	25
7	11	19
8	14	23
9	17	26
10	17	29

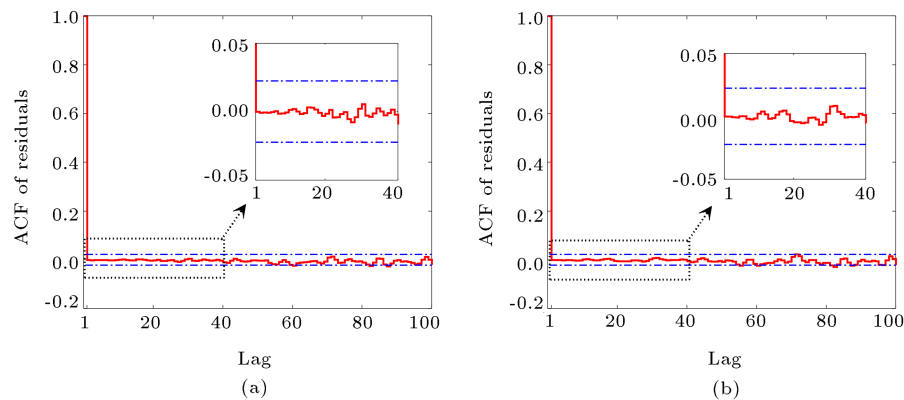
art BIC technique. In addition, Figures 4 and 5 compare the ACFs of the model residuals obtained from these orders in terms of extracting the uncorrelated residuals at sensors 3 and 8. In these figures, the dashed lines are the upper and lower bounds of 95% confidence interval.

From Figure 4, it is apparent that there are several violations of the samples of ACF on the upper and lower limits. This confirms that the initial order gained by the BIC technique fails in obtaining the uncorrelated residuals as the main factor for the model adequacy and accuracy. On the contrary, the observations in Figure 5 clearly show that the samples of ACF for the model residuals obtained from the improved order are within the limitations of upper and lower bounds, implying the uncorrelatedness of the residual sequences. It is important to note that the amount of ACF at the first lag always corresponds to one.

Based on the third step of the improved feature extraction method (Figure 1), the maximum order ( $p_m$ ) is 29, which should be applied to all sensors for parameter estimation and residual extraction. Removing the first 29 residual samples because they are zero, the final residual vectors for



**Figure 4.** Auto-Correlation Function (ACF) of the autoregressive (AR) model residuals using the initial order obtained by the state-of-the-art Bayesian Information Criterion (BIC) technique: (a) Sensor 3 and (b) Sensor 8.



**Figure 5.** Auto-Correlation Function (ACF) of the autoregressive (AR) model residuals using the improved order: (a) Sensor 3 and (b) sensor 8.

each of the undamaged and damaged cases contain 7971 observations (rows) and 10 variables (columns). To use the Andrews plot for early damage detection by considering Cases 1–4, a multivariate dataset is made including 31884 samples in rows ( $nm = 31884$ , which is calculated by multiplying the number of cases by the number of observations of the residual set) and 10 variables in columns ( $ns = 10$ ). Figure 6 indicates the results of the early damage detection by the Andrews plot in Cases 1–4 based on the residual datasets obtained from the initial and improved orders. This figure is intended to compare the performance of the AR model residuals extracted from these orders in the early damage detection. Note that since  $ns$  is an even number, Eq. (9) is used to plot the Andrews function.

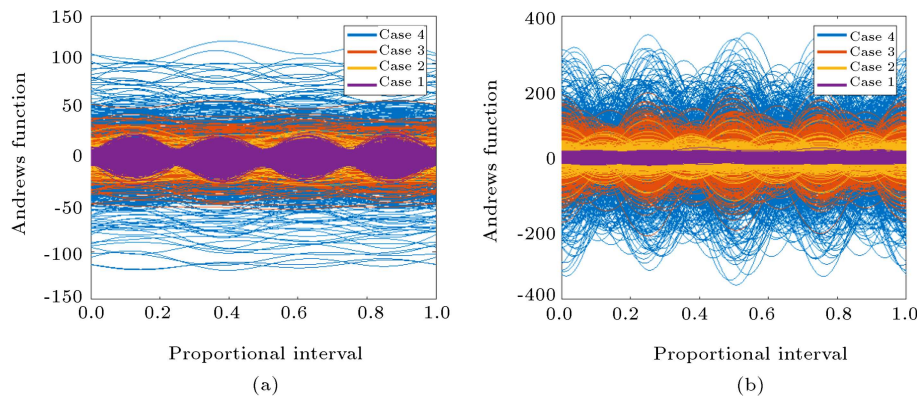
As can be seen in Figure 6(a), it is difficult to recognize the difference between Cases 1 and 2 for damage detection. Almost the same conclusion can be drawn in Cases 1 and 3; however, a discrepancy is clearly observed between Cases 1 and 4. Therefore, it can be deduced that the residual sequences of AR model with the initial order cannot provide reliable results for damage detection. In contrast, Figure 6(b) shows clear discrepancies between the undamaged and

damaged cases. Such observations confirm the positive effect of using the improved order on damage detection. In addition, in Figure 6(b), it is seen that the Andrews plot not only detects early damage in the beam but also estimates the level of damage severity with the rises in the curves for Cases 1–4.

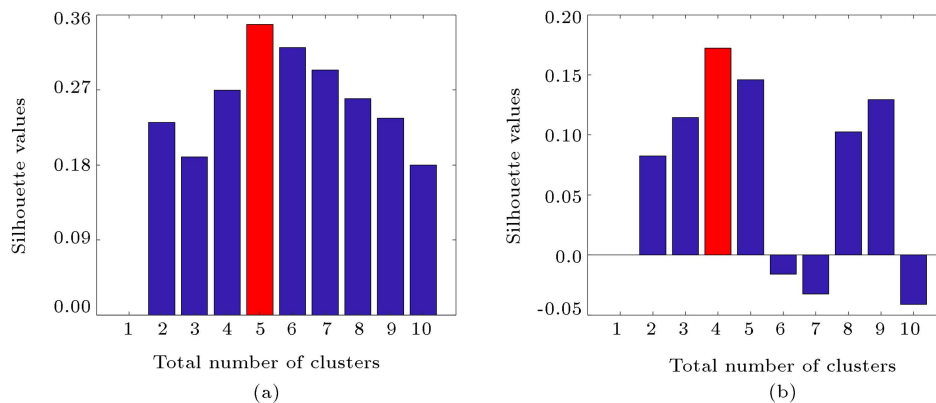
For the  $k$ -means and GMM clustering approaches, the methodology of Silhouette value is applied to determining the optimal number of clusters as shown in Figure 7. To achieve this purpose, the uncorrelated residuals of AR(29) fitted to the acceleration responses of all sensors in the undamaged condition (Case 1) are chosen as the training data. In Figure 7, the cluster numbers for the  $k$ -means and GMM methods are 5 and 4 based on the maximum Silhouette values.

Finding the optimal number of clusters, Figure 8 illustrates the results of early damage detection in each cluster via the ESD technique used in the  $k$ -mean clustering. The multivariate dataset ( $\mathbf{Y}$ ) for this clustering approach consists of 31884 observations and 10 variables.

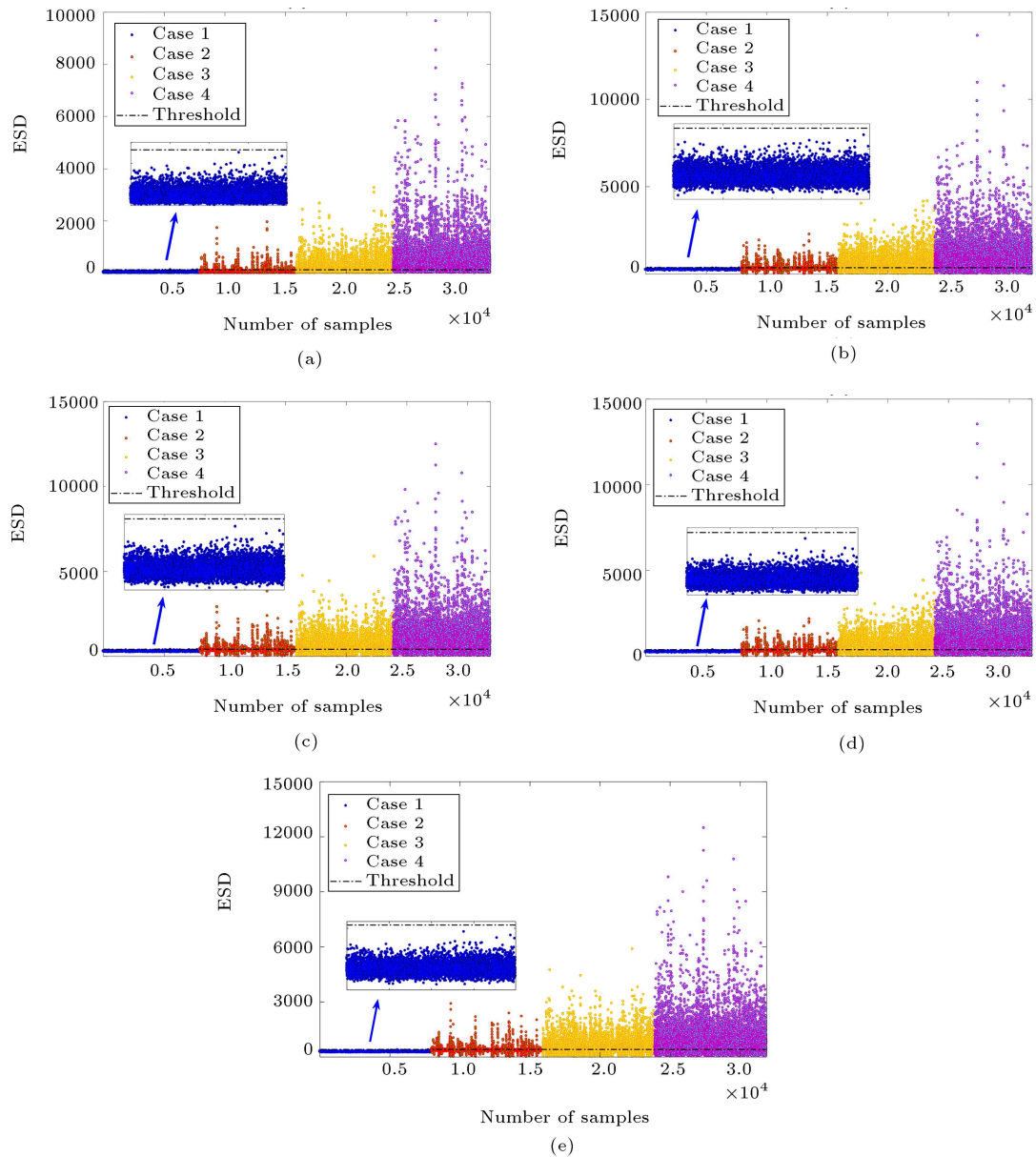
As Figure 8 reveals, one can conclude that the  $k$ -mean clustering method by means of the residual datasets of AR model extracted from the improved iterative feature extraction technique is able to accu-



**Figure 6.** Early damage detection by the Andrews plot in the numerical beam based on (a) the residual datasets obtained from the initial order, and (b) the residual datasets obtained from the improved order.



**Figure 7.** The optimal number of clusters in the beam: (a)  $k$ -means and (b) Gaussian Mixture Model (GMM).

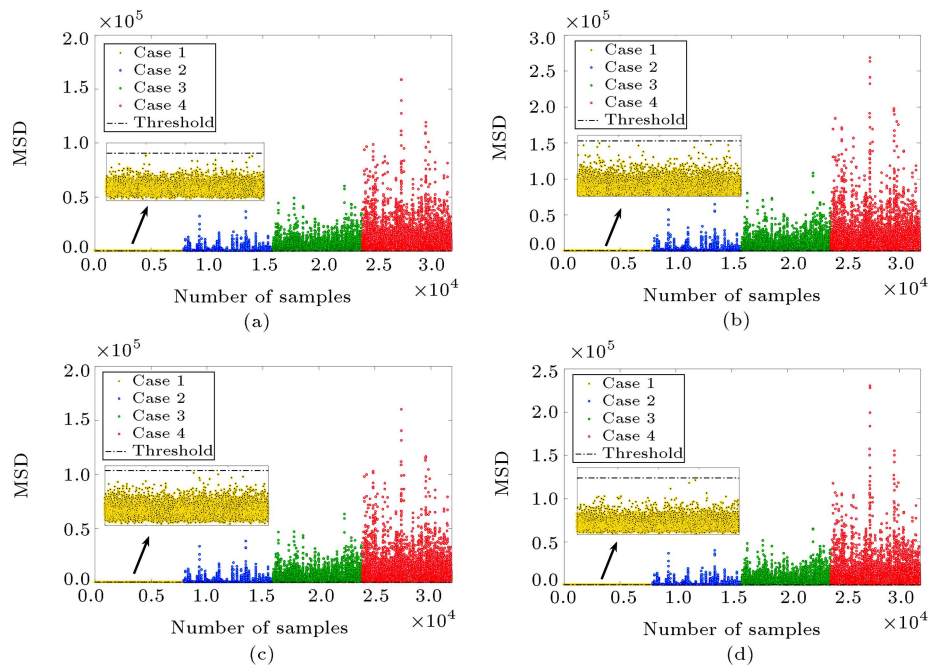


**Figure 8.** Early damage detection in the numerical beam using the *k*-means clustering technique: (a) 1st cluster, (b) 2nd cluster, (c) 3rd cluster, (d) 4th cluster, and (e) 5th cluster.

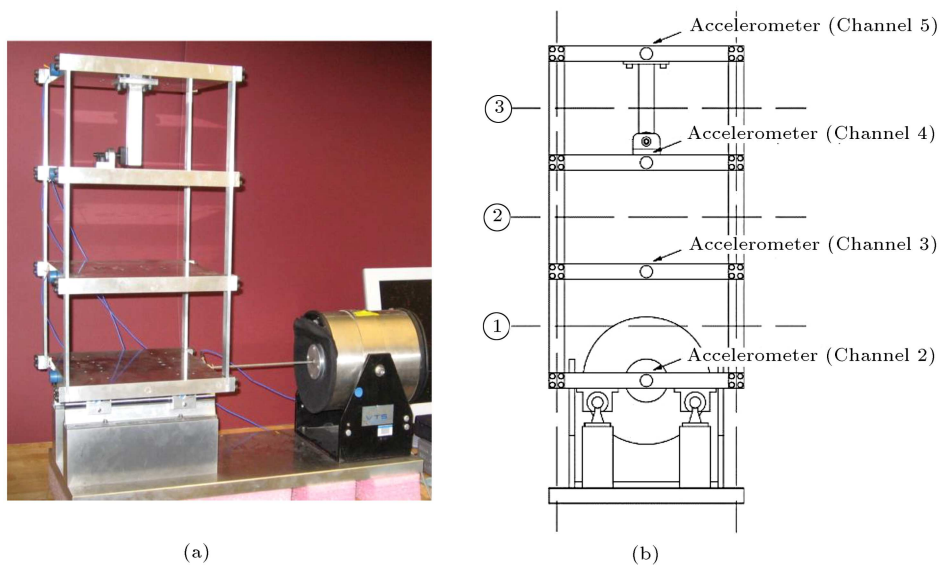
rately distinguish the undamaged condition from the damaged one. In this figure, it can be observed that Cases 2–4 indicate the damaged conditions of the beam since the values of ESD in each cluster exceed the threshold values (the horizontal lines). In contrast, all of the distance quantities in Case 1 are below this value. Note that the threshold limit is based on the upper-bound 95% confidence interval of the ESD values associated with the undamaged condition. Furthermore, it is seen that the distance quantities in each cluster increase with increasing the level of damage severity. In this regard, the distances of Case 4 are higher than the corresponding values in the other damaged conditions (Cases 2 and 3). This means that

the severest damage scenario gives the largest distance values.

Figure 9 presents the results of the early damage detection by the GMM clustering technique. In this figure, each sample indicates an MSD value. Unlike the *k*-means clustering, the GMM clustering approach needs to train a model using the training data (**X**) containing the uncorrelated residuals of the undamaged condition with 7971 residual samples in 10 variables. On the other hand, the testing data (**Y**) consists of 31884 observations (*nm*) including the residual sequences of all cases and the same variables as **X**. Obtaining the MSD values in each cluster, it is seen that the GMM clustering approach gives the same



**Figure 9.** Early damage detection in the numerical beam using the Gaussian Mixture Model (GMM) clustering technique: (a) 1st cluster, (b) 2nd cluster, (c) 3rd cluster, and (d) 4th cluster.



**Figure 10.** (a) Three-story laboratory frame [16]. (b) Sensor (channel) locations.

results of the early damage detection as the *k*-means. Eventually, all observations in Figures 8 and 9 lead to the conclusion that the improved residual-based feature extraction and the mentioned clustering methods are reliable tools for SHM applications.

### 7.2. An experimental laboratory benchmark frame

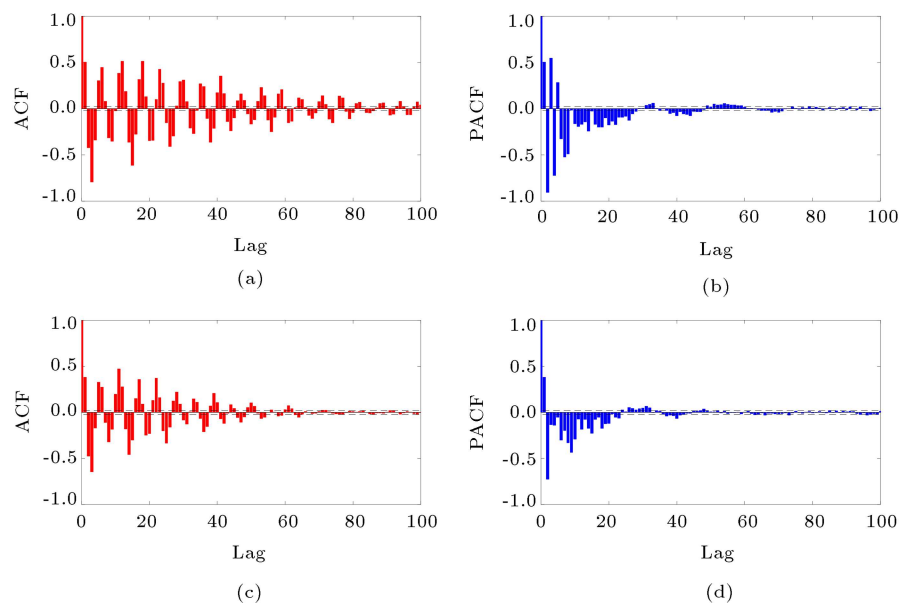
Another validation of the improved and proposed methods is carried out by a set of experimental datasets from a laboratory benchmark model [31]. It is a

three-story aluminum frame as shown in Figure 10(a). Four accelerometers were mounted on the floors to measure acceleration time histories at each floor. A random vibration load was applied by means of an electrodynamic shaker to the base along the centerline of the frame. The sensor signals were sampled at 320 Hz for 25.6 sec in duration, which was discretized into 8192 data points with 0.003125 sec time intervals.

Damage was simulated as the breathing crack by a bumper and a suspended column between the second and third floors. Different levels of damage

**Table 3.** The structural state conditions of the laboratory frame.

State(s)	Condition	Description
1	Undamaged	Baseline condition
2–3	Undamaged	Simulated operational variability by adding a concentrated mass (1.2 kg) on the base and first floors
4–9	Undamaged	Simulated environmental variability by decreasing structural stiffness at the first, second, and third floors
10–14	Damaged	Nonlinear damage (gap = 0.20, 0.15, 0.13, 0.10, and 0.05 mm)
15–17	Damaged	Nonlinear damage (gap = 0.20, 0.20, and 0.10 mm) with simulated operational variability at the base and first floors



**Figure 11.** Model identification by the Box-Jenkins methodology at sensor 5 of the laboratory frame: (a) Auto-Correlation Function (ACF) in state 1, (b) Partial Auto-Correlation Function (PACF) in state 1, (c) ACF in state 14, and (d) PACF in state 14.

severity were considered by the diverse gap between the suspended column and the bumper. This type of damage is a simulation of fatigue crack with nonlinear behavior that is able to open and close under loading conditions or loose connections in the structure. Table 3 summarizes the structural state conditions of the test structure with five damage levels from the gap of 0.20 mm (the lowest level of nonlinear damage) to 0.05 mm (the highest level of nonlinear damage). More details about the laboratory frame and full descriptions of the structural conditions can be found in [31]. In this study, the states 1, 5, 10, 14, and 17 are utilized to examine the improved and proposed methods.

In a similar manner to the numerical example, the Box-Jenkins methodology is used to identify the most compatible time series model with the acceleration time histories. As a sample, Figure 11 depicts the ACF and PACF of sensor 5 for the states 1 and 14. It is observed that the plots of ACFs have exponentially decreasing forms, whereas the plots of PACFs roughly become zero

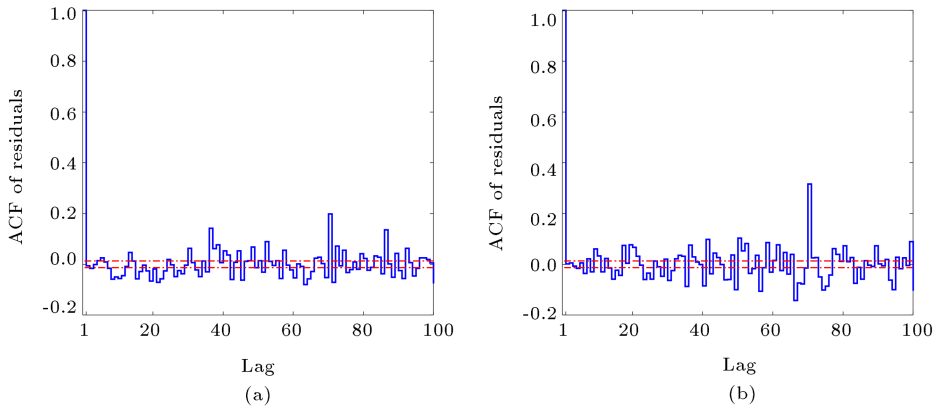
**Table 4.** The initial and improved orders of the AR models at all sensors of State 1.

Sensor no.	2	3	4	5
Initial orders	36	28	12	16
Improved orders	46	40	31	35

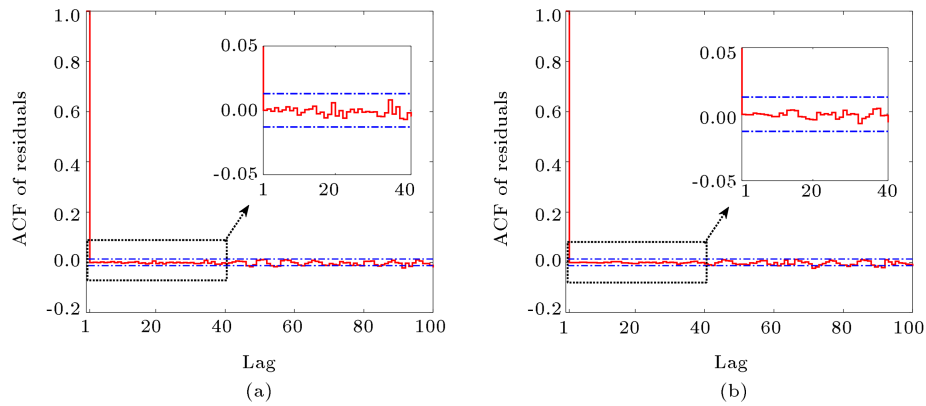
after the 30th lag. Therefore, one can argue that the selection of AR model is accurate and reasonable.

Table 4 presents the initial and improved orders of AR models for sensors 2–5 in the first structural state. The initial orders are determined by the state-of-the-art BIC technique. The maximum order is 46, in which case AR(46) is fitted to the acceleration responses acquired from all sensors in the selected undamaged and damaged conditions for the extraction of the model residuals as the damage-sensitive features.

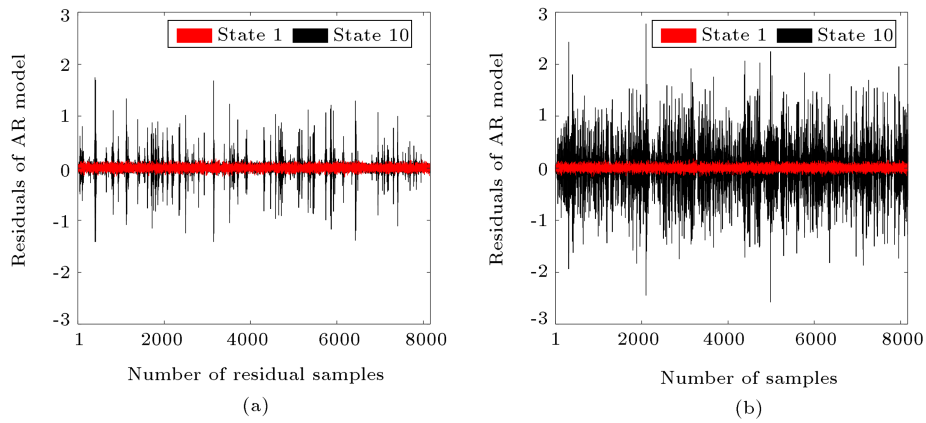
A comparative analysis is carried out to evaluate the correlation of the residual sequences obtained from the initial and improved AR orders. For this



**Figure 12.** Auto-Correlation Function (ACF) of the AR model residuals using the initial order obtained by the state-of-the-art BIC technique: (a) Sensor 3 and (b) sensor 5.



**Figure 13.** Auto-Correlation Function (ACF) of the AR model residuals using the improved order: (a) Sensor 3 and (b) sensor 5.



**Figure 14.** Comparison of the variations in the residuals of AR(46) between (a) states 1 and 10, and (b) states 1 and 14.

comparison, Figures 12 and 13 indicate the plots of ACF regarding the sequences of the residuals for sensors 3 and 5 gained by the initial and improved orders, respectively. Considering the numerous ACF samples that exceed the upper and lower bounds in Figure 12, one can conclude that the state-of-the-art BIC technique fails in fully extracting the uncorrelated residuals. On the contrary, it is seen that the improved

order gained by the improved method enables the AR model to generate uncorrelated residual samples.

To demonstrate sensitivity of the AR model residuals extracted from the improved feature extraction method, Figure 14 compares the variations in the residual sequences of AR(46) at sensor 4 (the location of damage) between the states 1 and 10 (the lowest damage severity) and the states 1 and 14 (the highest

damage severity). As can be seen, the occurrence of damage leads to increase in the residual samples so that state 14 has more increase than state 10. This indicates the sensitivity of the AR residuals to damage.

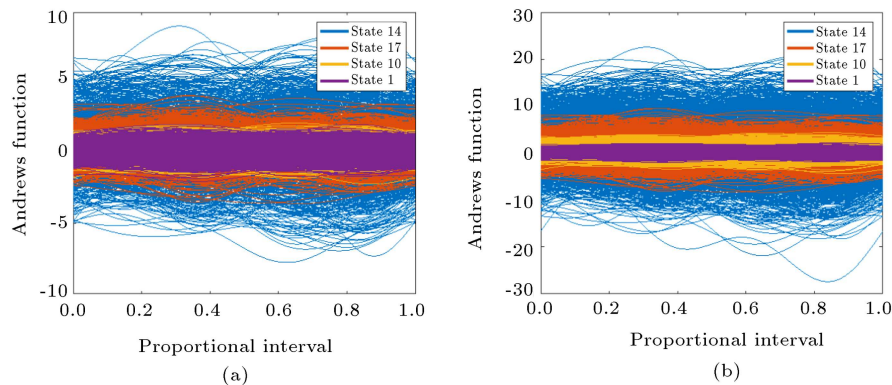
In another comparison, Figure 15 illustrates the results of early damage detection by using the AR residuals obtained from the initial and improved orders. To attain this goal, the residual sets of states 1, 10, 17, and 14 are applied to making a multivariate dataset ( $\mathbf{Y}$ ) with 32584 observations (i.e., by removing the first 46 samples from the 8192 residual sequences because they are zero) and 4 variables; that is,  $\mathbf{Y} \in \mathfrak{R}^{32584 \times 4}$ .

It is noteworthy that Eq. (9) is used to plot the Andrews function since  $ns$  is an even number. Based on Figure 15(a), it is perceived that the residual sets gained from the initial order by the BIC technique cannot distinguish the differences between states 1 and 10. In fact, it is difficult to recognize the curves of state 10. By contrast, Figure 15(b) obviously demonstrates that the residual sets obtained from the improved order through the enhanced feature extraction approach influentially succeed in detecting damage.

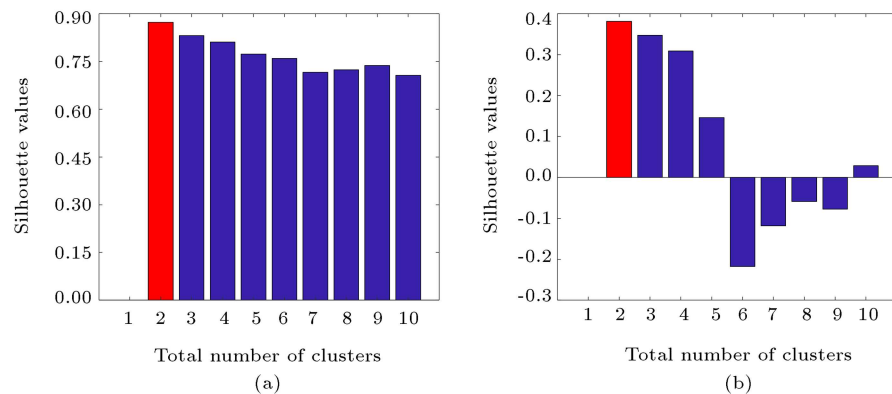
Applying the methodology of Silhouette value, the optimal number of clusters for  $k$ -means and GMM

clustering approaches is determined as shown in Figure 16. For this purpose, the uncorrelated residuals of states 1 and 5 are employed as the training dataset, which includes 16292 observations and 4 variables. Notice that structural state 5 is representative of an undamaged condition along with the operational and environmental variability, which efficiently assist in investigating the influence of such variability on the clustering process.

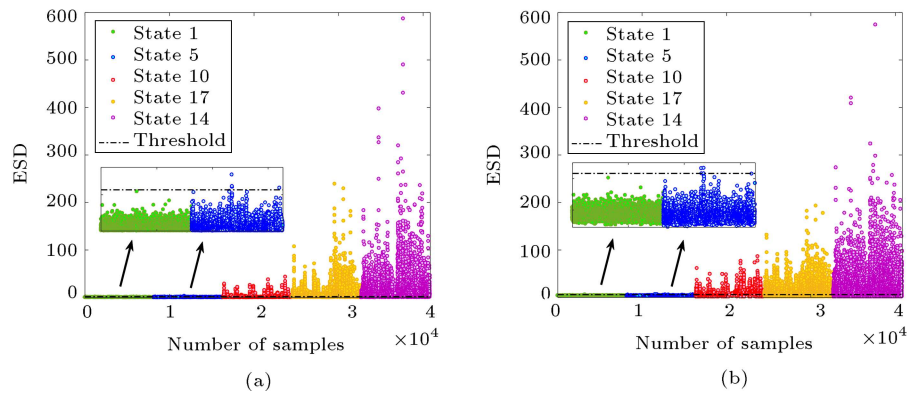
For the early damage detection process by the  $k$ -means clustering, the residual sets regarding states 1, 5, 10, 17, and 14 are applied to making a multivariate dataset ( $\mathbf{Y}$ ) containing 40730 observations (multiplying 5 by 8146) and 4 variables. Since the GMM clustering approach requires training data from the undamaged conditions, the multivariate dataset  $\mathbf{X} \in \mathfrak{R}^{16292 \times 4}$  is initially determined by using the uncorrelated residual sets of states 1 and 5. Furthermore, the testing data are determined by employing the residual sets of states 1, 5, 10, 17, and 14, which have the same dimension as the multivariate dataset for the  $k$ -means. Figures 17 and 18 depict the results of damage detection by the  $k$ -means and GMM clustering techniques, respectively. It is important to point out that the threshold values



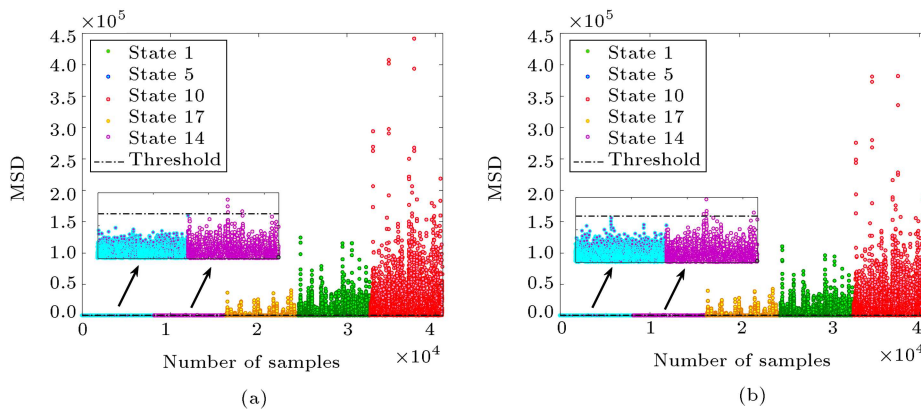
**Figure 15.** Early damage detection by Andrews plot in the numerical beam based on (a) the residual datasets obtained from the initial order and (b) the residual datasets obtained from the improved order.



**Figure 16.** The optimal number of clusters in the laboratory frame: (a)  $k$ -means and (b) Gaussian Mixture Model (GMM).



**Figure 17.** Early damage detection in the laboratory frame by the *k*-means clustering method: (a) 1st cluster and (b) 2nd cluster.



**Figure 18.** Early damage detection in the laboratory frame by the Gaussian Mixture Model (GMM) clustering method; (a) 1st cluster and (b) 2nd cluster.

(the dashed lines) in these figures are based on the 95% confidence intervals of the ESD and MSD values in the undamaged conditions (states 1 and 5).

In both figures, the distance values of the damaged conditions cross the threshold limits, indicating the occurrence of damage in the laboratory frame, while most of the ESD and MSD quantities belonging to states 1 and 5 are under these limits, implying the normal conditions of the frame. Among all states, it is observed that state 14 is associated with the highest level of damage having the largest distance amounts, whereas state 10 indicates the lowest level of damage. The other significant result pertains to the effect of the operational and environmental variability on the early damage detection. From Figures 17 and 18, one can see that the ESD and MSD amounts of state 5 are roughly similar to the baseline condition in spite of a few distance values that exceed the threshold limits. Furthermore, the distances of state 17 are smaller than those of state 14 even in the presence of operational and environmental conditions.

### 8. Conclusions

This study focused on the steps of feature extraction

and statistical decision-making regarding the statistical pattern recognition paradigm. An improved residual-based feature extraction method by time series modeling was proposed to determine a sufficient order and extract the model residuals as the damage-sensitive features. The Box-Jenkins methodology was also applied to identifying the most compatible time series representation with the vibration time-domain responses. A multivariate data visualization approach called Andrews plot was proposed to detect damage by using the high-dimensional features (the model residuals). The well-known *k*-means and GMM clustering techniques were also utilized to assess the reliability and performance of these features in the early damage detection based on the estimation of threshold limits. Eventually, a numerical concrete beam and an experimental benchmark frame were employed to validate the improved and proposed methods.

The main conclusions that can be drawn are:

1. The Box-Jenkins methodology establishes a simple and efficient graphical tool for identifying the most appropriate time series representations. Accordingly, the AR model is the most appropriate time series representation for feature extraction;



2. The improved order determination algorithm provides a sufficient and accurate order so that it enables the AR model to generate the uncorrelated residuals;
3. This algorithm is superior to the state-of-the-art BIC technique in terms of obtaining the uncorrelated residual sequences;
4. The residual sets extracted from the improved feature extraction method are sensitive to damage;
5. Utilizing the model characteristics obtained from the normal condition in the damaged state, the increases in the values of the AR residuals are representative of damage occurrence;
6. The Andrews plot succeeds in detecting adverse changes caused by damage even in the presence of the operational and environmental variability and high-dimensional features;
7. Determination of an adequate order not only plays an important role in the feature extraction but also highly affects the results of damage detection as shown in the comparative analysis of the Andrews plots;
8. Both the  $k$ -means and GMM clustering techniques along with the threshold limits are able to distinguish the undamaged state from the damaged one even under varying operational and environmental conditions;
9. These approaches not only confirm the sensitivity of the AR residuals extracted from the improved feature extraction method to damage but also estimate the level of damage.
5. Gul, M. and Necati Catbas, F. “Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications”, *Mech. Syst. Sig. Process.*, **23**(7), pp. 2192–2204 (2009).
6. Yao, R. and Pakzad, S.N. “Autoregressive statistical pattern recognition algorithms for damage detection in civil structures”, *Mech. Syst. Sig. Process.*, **31**, pp. 355–368 (2012).
7. Park, S., Lee, J.-J., Yun, C.-B., and Inman, D.J. “Electro-mechanical impedance-based wireless structural health monitoring using PCA-data compression and  $k$ -means clustering algorithms”, *J. Intell. Mater. Syst. Struct.*, **19**(4), pp. 509–520 (2008).
8. Silva, M., Santos, A., Santos, R., Figueiredo, E., Sales, C., and Costa, J.C.W.A. “Agglomerative concentric hypersphere clustering applied to structural damage detection”, *Mech. Syst. Sig. Process.*, **92**, pp. 196–212 (2017).
9. Kopsaftopoulos, F. and Fassois, S. “Vibration based health monitoring for a lightweight truss structure: experimental assessment of several statistical time series methods”, *Mech. Syst. Sig. Process.*, **24**(7), pp. 1977–1997 (2010).
10. Datteo, A. and Lucà, F. “Statistical pattern recognition approach for long-time monitoring of the G. Meazza stadium by means of AR models and PCA”, *Eng. Struct.*, **153**, pp. 317–333 (2017).
11. de Lautour, O.R. and Omenzetter, P. “Nearest neighbor and learning vector quantization classification for damage detection using time series analysis”, *Struct. Control Health Monit.*, **17**(6), pp. 614–631 (2010).
12. Farahani, R.V. and Penumadu, D. “Full-scale bridge damage identification using time series analysis of a dense array of geophones excited by drop weight”, *Struct. Control Health Monit.*, **23**(7), pp. 982–997 (2016).
13. Carden, E.P. and Brownjohn, J.M. “ARMA modelled time-series classification for structural health monitoring of civil infrastructure”, *Mech. Syst. Sig. Process.*, **22**(2), pp. 295–314 (2008).
14. Mei, L., Mita, A., and Zhou, J. “An improved substructural damage detection approach of shear structure based on ARMAX model residual”, *Struct. Control Health Monit.*, **23**, pp. 218–236 (2016).
15. Box, G.E., Jenkins, G.M., and Reinsel, G.C., *Time Series Analysis: Forecasting and Control*, 4th Edn., John Wiley & Sons, Inc, New Jersey (2008).
16. Figueiredo, E., Figueiras, J., Park, G., Farrar, C.R., and Worden, K. “Influence of the autoregressive model order on damage detection”, *Comput.-Aided Civ. Infrastruct. Eng.*, **26**(3), pp. 225–238 (2011).

## Acknowledgment

The authors would like to express their sincere appreciation to the Los Alamos National Laboratory in the USA for providing the experimental datasets of the three-story frame.

## References

1. Farrar, C.R. and Worden, K., *Structural Health Monitoring: A Machine Learning Perspective*, John Wiley & Sons (2013).
2. Amezcua-Sanchez, J. and Adeli, H. “Feature extraction and classification techniques for health monitoring of structures”, *Sci. Iran., A.*, **22**(6), pp. 1931–1940 (2015).
3. Qarib, H. and Adeli, H. “Recent advances in health monitoring of civil structures”, *Sci. Iran., A.*, **21**(6), pp. 1733–1742 (2014).
4. Amezcua-Sanchez, J.P. and Adeli, H. “Signal processing techniques for vibration-based health monitoring of smart structures”, *Arch. Comput. Methods Eng.*, **23**(1), pp. 1–15 (2016).

17. Mosavi, A.A., Dickey, D., Seracino, R., and Rizkalla, S. "Identifying damage locations under ambient vibrations utilizing vector autoregressive models and Mahalanobis distances", *Mech. Syst. Sig. Process.*, **26**, pp. 254–267 (2012).
18. Farrar, C.R. and Jauregui, D.A. "Comparative study of damage identification algorithms applied to a bridge: I. Experiment", *Smart Mater. Struct.*, **7**(5), p. 704 (1998).
19. Farrar, C.R. and Jauregui, D.A. "Comparative study of damage identification algorithms applied to a bridge: II. Numerical study", *Smart Mater. Struct.*, **7**(5), p. 720 (1998).
20. Fugate, M.L., Sohn, H., and Farrar, C.R. "Vibration-based damage detection using statistical process control", *Mech. Syst. Sig. Process.*, **15**(4), pp. 707–721 (2001).
21. Gul, M. and Necati Catbas, F. "Structural health monitoring and damage assessment using a novel time series analysis methodology with sensor clustering", *J. Sound Vib.*, **330**(6), pp. 1196–1210 (2011).
22. Roy, K., Bhattacharya, B., and Ray-Chaudhuri, S. "ARX model-based damage sensitive features for structural damage localization using output-only measurements", *J. Sound Vib.*, **349**, pp. 99–122 (2015).
23. Kitagawa, G., *Introduction to Time Series Modeling*, CRC Press, Taylor & Francis Group, Boca Raton (2010).
24. Andrews, D.F. "Plots of high-dimensional data", *Biometrics*, **28**(1), pp. 125–136 (1972).
25. Izenman, A.J., *Modern Multivariate Statistical Techniques: Regression, Classification and Manifold Learning*, Springer, New York, NY (2008).
26. Wu, J., *Advances in k-Means Clustering: A Data Mining Thinking*, Springer, Science & Business Media (2012).
27. McLachlan, G. and Peel, D., *Finite Mixture Models*, John Wiley & Sons (2004).
28. Kaufman, L. and Rousseeuw, P.J., *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons (2009).
29. Newmark, N.M. "A method of computation for structural dynamics", *J. Eng. Mech. Div. ASCE*, **85**(3), pp. 67–94 (1959).

30. Friswell, M.I. and Penny, J.E. "Crack modeling for structural health monitoring", *Struct. Health Monit.*, **1**(2), pp. 139–148 (2002).
31. Figueiredo, E., Park, G., Figueiras, J., Farrar, C., and Worden, K., *Structural Health Monitoring Algorithm Comparisons Using Standard Data Sets*, LA-14393, Los Alamos National Laboratory, Los Alamos, NM (2009).

## Biographies

**Alireza Entezami** is a PhD candidate of Structural Engineering at Ferdowsi University of Mashhad (FUM). He received his MSc degree with honors based on GPA among graduated MSc students of Civil Engineering (structure) from FUM in 2014. Furthermore, he was selected as a superior researcher among all MSc students in the Civil Engineering Department at FUM and as a distinguished student, who could receive the Iranian Top Student Award by the first vice president of Iran, among all Iranian MSc students in 2015. In 2014, he began his PhD studies as an honorary admission of Brilliant Students at FUM. His research interests are structural health monitoring, machine learning, data mining, statistical signal processing, and statistical pattern recognition. He has published several research papers in scientific journals.

**Hashem Shariatmadar** received his BSc degree in 1989 from Tabriz University, MSc degree in 1993, and PhD degree in 1997 from McGill University in Canada in Structural Engineering. Currently, he is an Associate Professor at Ferdowsi University of Mashhad (FUM). His research studies are structural control, earthquake engineering, and structural health monitoring and he has authored several papers and publications.

**Abbas Karamodin** received his BSc and MSc degrees in Structural Engineering from University of Tehran (UT) in 1986, and his PhD degree in Structural Engineering with the specialty of Structural Control from Ferdowsi University of Mashhad (FUM) in 2009. He has served at FUM as faculty member since 1987. He has had 3 books published in his related field and more than 10 papers in respected journals and conference proceedings. His areas of research are earthquake engineering and structural control.