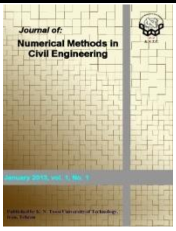
Numerical Methods in Civil Engineering, 6-4 (2022) 16-28



Numerical Methods in Civil Engineering

Journal Homepage: https://nmce.kntu.ac.ir/

A novel method for detecting structural damage based on data-driven and similarity-based techniques under environmental and operational changes

Masoud Haghani Chegeni*, Mohammad Kazem Sharbatdar, Reza Mahjoub***, Mehdi Raftari*****

ARTICLE INFO

RESEARCH PAPER

Article history:
Received:
August 2021.
Revised:
November 2021.
Accepted:
December 2021.

Keywords:
Structural damage detection
damage localization
damage level estimation
autoregressive model
independent residuals

Abstract:

The applications of time series modeling and statistical similarity methods to structural health monitoring (SHM) provide promising and capable approaches to structural damage detection. The main aim of this article is to propose an efficient univariate similarity method named as Kullback similarity (KS) for identifying the location of damage and estimating the level of damage severity. An improved feature extraction technique based on autoregressive (AR) model is presented to extract independent residuals of the AR model as damage-sensitive features. This technique emphasizes to choose a sufficient order such that the model residuals be independent. The proposed univariate similarity approach is a new application of the well-known KS method that attempts to measure a difference between two randomly distributed variables. The major contribution of the proposed KS method is that it only requires one measurement of undamaged and damaged conditions to compute the similarity between them. For the process of damage localization, the sensor location associated with the largest KS quantity is identified as the damaged area. In the damage level estimation, it is necessary to compare at least two different damaged conditions and find the maximum KS value in these conditions as the highest level of damage severity. The performance and capability of the improved and proposed methods is successfully verified by an experimental laboratory frame belonging to the Los Alamos National Laboratory. Results show that the methods are powerful and reliable tools for identifying the location of damage and estimating the level of damage severity.

1. Introduction

Structural health monitoring (SHM) is an implementing process that aims to evaluate the health and safety of engineering systems and detect any probable structural damage by vibration data [1].

All of these factors may cause permanent changes in structural stiffness, undesirable stresses and displacements, inappropriate vibrations, adverse dynamic behavior, failure, and even collapse in structures. To prevent such dramatic events and decrease high costs of maintenance and rehabilitation, it is imperative to use SHM systems for assessing the safety of structures and detect structural damage.

Damage detection process is normally carried out by model-based or data-based methods. The main premise associated with the model-based methods is to use a finite element model and its inherent physical properties such as mass, stiffness and damping [2, 3]. A main limitation of such methods is that the finite element model of the structure may not be accurate; therefore, a model updating strategy is typically employed to calibrate or update the physical properties of the structure [4-7]. On the contrary,

** Ph.D. Candidate, Department of Civil Engineering, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran.*
*** Corresponding author: Professor, Faculty of Civil Engineering, Semnan University, Semnan, Iran. Email: msharbatdar@semnan.ac.ir.*
**** Assistant Professor, Department of Civil Engineering, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran.*

the data-based methods rely on applying raw vibration data on the basis of statistical pattern recognition paradigm that attempts to extract patterns or features from measured vibration responses and then analyze the extracted features to make a meaningful decision about the condition of structure and damage occurrence [8-12].

The process of feature extraction is concerned with finding out features that should be sensitive to damage, called damage-sensitive features [9, 13-15], and not to operational and environmental variability [16-18]. In the SHM community, coefficients and residuals of time series models are normally chosen as the damage-sensitive features. On this basis, Sohn et al. [19] extracted the coefficients of autoregressive (AR) model as the damage-sensitive features and applied X-bar control charts based on statistical process control for damage detection. In a similar way, Fugate et al. [20] utilized the residual errors of AR model as the damage-sensitive features and then used some statistical process control methods such as X-bar and S control charts in order to identify damage in a concrete column. Nair et al. [21] used an autoregressive moving average (ARMA) model and applied the first three AR coefficients, which were used into two damage localization indices for the identification of structural damage in the ASCE benchmark structure. Mei et al. [22] proposed an improved substructure-based damage detection method to locate and quantify damage in a shear building structure. In their article, both coefficients and residuals of ARMAX models in undamaged and damaged conditions were incorporated as damage features, rather than only considering the model coefficients or residuals, to capture more complete structural damage information. Hoell and Omenzetter [23] conducted a research to enhance an improved selection of coefficients regarding the AR or ARMA models as widely-used damage-sensitive features in structural damage detection.

Even though time series modeling has extensively been used to extract the damage-sensitive features, there are some limitations and issues in using time series models that should be dealt with. One of the significant issues of such a modeling technique is to select sufficient model orders which still remains as a challenging issue in the SHM community. Figueiredo et al. [24] assessed the influence of selecting different orders of AR model on damage detection. They introduced four information criterion techniques to select the most appropriate model order. Based on their results, one can realize that the selection of inappropriate orders may cause insensitive features for damage detection problems. Stull et al. [25] employed information-gap decision theory and robustness curves based on area under receiver operating characteristic (ROC) curve, which is abbreviated as AUC, versus uncertainty to choose a robust order of AR model. From a statistical viewpoint, a time series model should extract independent residuals to ensure the model accuracy and adequacy [13, 26-29]. In addition to the statistical sense, it is essential to fit an accurate time series model to vibration data in order to capture the entire physical and dynamic characteristics of the structure. To address these issues, in the first step an improved feature extraction technique using AR model is presented to choose a sufficient order and then extract the independent residuals of the model as the damage-sensitive features.

Another important challenge of feature classification method for making a right decision about the current status of structure is to utilize a reliable statistical technique for measuring the discrepancy of damage features. Among a large number of methods, statistical similarity techniques are widely used in the context of SHM to implement the damage detection levels. A statistical similarity method generally calculates a difference or similarity between two objects or variables [12, 30-36]. In the decision-making phase, the similarity methods can be applied to detect early damage, identify damage location, and then estimate damage severity. Gul and Catbas [37] applied Mahalanobis similarity method by using the coefficients of AR model in order to detect and locate structural damage in the different laboratory structures. Mosavi et al. [38] proposed a novel feature extraction methodology by measuring the similarity of coefficients of vector autoregressive (VAR) model through Mahalanobis similarity method and employed Fisher criterion to identify the location of damage. Balsamo et al. [39] utilized squared Mahalanobis similarity method using Mel-Frequency Cepstral Coefficients, as new damage features, to detect damage.

The main objective of this study is to propose a novel statistical similarity method for the identification of damage location and estimation of damage severity. The main novelty of the proposed KS method is to avoid exploiting any time-consuming and complex partitioning processes for measuring the similarity between two random variables. This characteristic makes it particularly suited for SHM applications. Another contribution of the proposed KS method is that it only requires one measurement of undamaged and damaged conditions to compute the similarity between them. In addition to the the proposed KS method, an improved residual-oriented feature extraction technique on the basis of time series modeling is presented to deal with some limitations in the feature extraction process such as the complexity of time series modeling, the overfitting problem, and the lack of sufficient efficiency of extracted features for SHM applications. In this technique, a sufficient order of AR model is chosen so that the model is able to extract independent residuals as the main criterion of sufficiency and accuracy of time series models. The performance and

[DOR: 20.1001.1.23454296.2022.6.4.1.8]
[DOI: 10.52547/nmce.6.4.16]

[DOR: 20.1001.1.23454296.2022.6.4.1.8]

[DOI: 10.52547/nmce.6.4.16]

robustness of the improved and proposed methods are verified by an experimental laboratory frame belonging to the Los Alamos National Laboratory (LANL). In this frame, operational and environmental variability are simulated in undamaged and damaged conditions by adding mass or reducing structural stiffness. Results demonstrate that the proposed KS method using the residuals of the sufficient AR model is influentially able to identify the location of damage and estimate the level of damage severity.

The layout of this paper is as follows; Section 2 presents a brief discussion on the time series analysis by AR model. Section 3 describes the improved residual-oriented feature extraction technique. Section 4 entirely discusses the proposed KS method. Section 5 provides a full discussion of damage localization and damage level estimation on the laboratory frame. In this section, a comparative study on the process of damage localization between the proposed KS and the well-known Euclidean similarity method is carried out to reveal the robustness and performance of the proposed KS method in SHM applications. Eventually, the conclusions of this study are remarked in Section 6.

2. Feature extraction by time series modeling

2.1 Autoregressive model

Time series analysis is a statistical method that attempts to fit a mathematical model to the time series data for extracting some statistical features [40-42]. Autoregressive (AR) model is a widely used linear stationary model, which is generally applied to analyze stationary time series data. A stationary procedure is a stochastic process in which the statistical moments of time series data such as the mean, variance and higher order moments are time invariant [26]. The basic formulation of an AR model is expressed in the following form:

$$x(t)=\sum_{\mathrm{i}=1}^{p}a_{\mathrm{i}}x(t-\mathrm{i})+e(t) \tag{1}$$

where $x(t)$ is the measured vibration time-domain signal at the time t; $\mathbf{a} = [a_1,a_2,\ldots,a_p]$ denotes the vector of AR coefficients (parameters), and p represents the order of AR model. In addition, $e(t)$ is an independent residual (an unobservable random error) at the time t implying the difference between the measured vibration response and the predicted one gained by the model expressed as follows:

$$e(t)=x(t)-\hat{x}(t) \tag{2}$$

It is important to mention that there are several reasons which confirm the capability of the AR model in the SHM applications. The most significant reason that makes the AR model particularly applicable for SHM is that statistical features extracted from this model are sensitive to damage. Another reason is that the AR model coefficients and residuals reflect the inherent structural properties, which is highly beneficial to SHM. In addition, the implementation of this model is simple [43].

2.2 Identification of stationary time series models

There are some factors to identify a specific type of time series model such as availability of time series data (input-output versus output-only), nature of time series data (stationary versus non-stationary, Gaussian versus non-Gaussian, and linear versus nonlinear), and type of application. For example, an ARX model may lead to a better performance than an AR model when the input-output time series data is available. As another example, in the non-stationary time series data, an integrated time series model such as an autoregressive integrated moving average (ARIMA) model may provide more proper statistical features compared with an ARMA model.

Regardless of the type of the time series data and models, there are engineering reasons that the AR model may be an appropriate choice for using in the SHM applications. The first reason is that the statistical features extracted from the AR model are sensitive to damage. Second, this model only depends on the response or output of the structure without regard to excitation sources. The third reason is that the parameters (coefficients) of AR model reflect the inherent properties of the structure so that excitation fluctuations do not have any influences on the model parameters. Eventually, the implementation of this model is simple and easy.

2.3 Model order selection

After the identification of a time series model, it is important to determine the model order(s) because an inadequate order selection can result in an inappropriate time series model. The number of orders required to a model specifies how many unknown parameters should be chosen for the mathematical equation of the time series model to predict the response of the structure. The selection of model orders can be implemented by Akaike and Bayesian information criteria and checking the autocorrelation function of the model residuals [26]. However, the model order determination depends strongly on the property of residuals. This means that a sufficient and robust order is one that enables the time series model to extract or generate independent residuals with zero mean [44]. Any time series model that does not satisfy these requirements should be modified. As a result, the extraction of independent residuals is of paramount importance in time series modeling, and one should consider it as the first and main factor to ensure that the time series models are adequate and accurate [44].

3. An improved residual-oriented feature extraction technique

In the SHM community, the residuals of the AR model are chosen as the damage-sensitive features [20, 24, 45]. Unlike the process of feature extraction by the model coefficients, the residual-oriented feature extraction algorithm relies on using an AR model, along with its coefficients estimated from the undamaged condition, to predict the response of the structure in the damaged state. The fact beyond this algorithm is that the linear AR model used in the undamaged structure will no longer correctly predict the response of the damaged structure; therefore, the residual errors regarding this structure will increase [24]. In this case, the increase in the model residuals is an indicator of damage occurrence.

Due to the importance of extracting the independent residuals from engineering and statistical viewpoints, the improved residual-oriented feature extraction technique using the AR model is described in this section. This technique emphasizes to choose a sufficient AR order at each sensor location in such a way that the model is able to extract the independent residuals. Concisely, the algorithm of the improved residual-oriented feature extraction technique is described in the following steps:

Step 1: Different AR models are fitted to time series signals acquired from all sensors in the undamaged condition in such a way that each sensor has a separate model. An information criterion technique is applied to estimate an initial order (p_0) of the AR model at each sensor location.

Step 2: Although the information criteria are normally applied to determine the orders of time series models, the residuals extracted from the initial model, AR(p_0), may not fully be independent. In order to extract the independent residuals, the initial estimation of the model order is developed to achieve an improved order (p_m). It is significant to point out that the primary criterion of choosing the improved order is to check the correlation of the model residuals by the autocorrelation function. In time series modeling, the residual analysis through the autocorrelation function is known as an efficient graphical tool.

Step 3: For the damage detection problems, it is better to use equal numbers of the damage-sensitive features, either coefficients or residuals, in statistical approaches. To deal with the feature inequality in the improved feature extraction technique, the maximum number of the improved orders is chosen as a sufficient order (p_r).

Step 4: The sufficient order leads to a sufficient AR model and sufficient model coefficients. On this basis, in the undamaged condition, the sufficient AR model is fitted to all vibration time-domain signals and the coefficients of the sufficient model are estimated by one of the computational techniques such as least squares, Burg, forward-backward, and Yule-Walker methods [26].

Step 5: Eventually, the residuals of the AR models are extracted as the damage-sensitive features.

4. An efficient statistical method: Kullback similarity

In statistics, the Kullback similarity is a non-symmetric measure of the difference between two probability distributions [46]. The results of KS method are always positive; therefore, a zero-similarity value obtained by this method is indicative of the presence of a full similarity between the distributions. Due to measuring the distance of two data distributions or variables, the Kullback similarity method falls into the category of univariate statistical distance methods. It is worth mentioning that the statistical distances are mostly not metrics and do not need to be symmetric; therefore, some kinds of distance measures are referred to as the statistical divergences.

Assume that n-dimensional distributed sets $\mathbf{X}$ and $\mathbf{Y}$ are discrete probability distributions. The classical formulation of KS for such data sets is given by:

$$D_{\mathrm{KS}}(\mathbf{X}\,||\,\mathbf{Y})=\sum_{\mathrm{i}=1}^{n}\mathbf{X}(\mathrm{i})\ln\frac{\mathbf{X}(\mathrm{i})}{\mathbf{Y}(\mathrm{i})} \tag{3}$$

In the Equation 3, the D_{KS} is an expectation of the logarithmic difference between the $\mathbf{X}$ (as the reference data) and $\mathbf{Y}$ (as the target data). In order to identify the location of damage, one needs to compute the D_{KS} quantity at each sensor location between n-dimensional distributed sets $\mathbf{X}$ and $\mathbf{Y}$.

For the continuous probability distributions $\mathbf{X}$ and $\mathbf{Y}$, on the other hand, the classical KS equation is defined in the integral form as:

$$D_{\mathrm{KS}}(\mathbf{X}\,||\,\mathbf{Y})=\int_{-\infty}^{\infty}\rho(\mathrm{x})\ln\frac{\rho(\mathrm{x})}{\rho(\mathrm{y})}\mathrm{dx} \tag{4}$$

where $\rho(\mathrm{x})$ and $\rho(\mathrm{y})$ denote the density functions of the distributions $\mathbf{X}$ and $\mathbf{Y}$, respectively. To identify the location of damage and estimate the level of damage severity, one needs to compute the KS value at each sensor location using the residuals of the sufficient AR model in the undamaged and damaged conditions. On this basis, a similarity vector is formulated as follows:

$$\mathbf{D}=[D_{\mathrm{KS}}(\mathbf{X}_1\,||\,\mathbf{Y}_1)\quad D_{\mathrm{KS}}(\mathbf{X}_2\,||\,\mathbf{Y}_2)\quad \mathrm{L} \\ D_{\mathrm{KS}}(\mathbf{X}_s\,||\,\mathbf{Y}_s)] \tag{5}$$

where s denotes the number of sensors mounted on the structure. Each element in the vector **D** represents a similarity value at each sensor location obtained by the reference and target variables. The major advantage of this vector is to detect single and multiple damage cases in such a way that the sensor location in association with the largest KS value is identified as the damage location for the single damage scenario. In the multiple damage cases, the sensor locations that have more substantial KS quantities than the other sensors are indicative of damage locations. At these areas, the maximum similarity value is representative of the highest level of damage severity.

5. Experimental validation

In this section, an experimental model is applied to validate the robustness and capability of the improved and proposed methods. This model is a three-story laboratory frame constructed at the Los Alamos National Laboratory [43]. The frame schematic and sensor locations are depicted in Fig. 1. This three-story laboratory frame was constructed from aluminum columns (height 177 mm, width 25 mm, and thickness 6 mm) and aluminum plates (length 305 mm × 305 mm, and thickness 25 mm) [43]. At each floor, four aluminum columns were connected to the top and bottom of the aluminum plate assembled using bolted joints. A random vibration load was applied by means of an electrodynamics shaker to the base floor along the centerline of the frame. The structure was instrumented with four accelerometers (channels 2-5) mounted at the centerline of each floor on the opposite side from the excitation source to measure acceleration time histories. The sensor signals were sampled at 320 Hz for 25.6 sec in duration, which were discretized into 8192 data samples at 3.125 microsecond intervals. A comprehensive documentation concerning this model is available in [43]. Moreover, the experimental data can be downloaded free from [47].

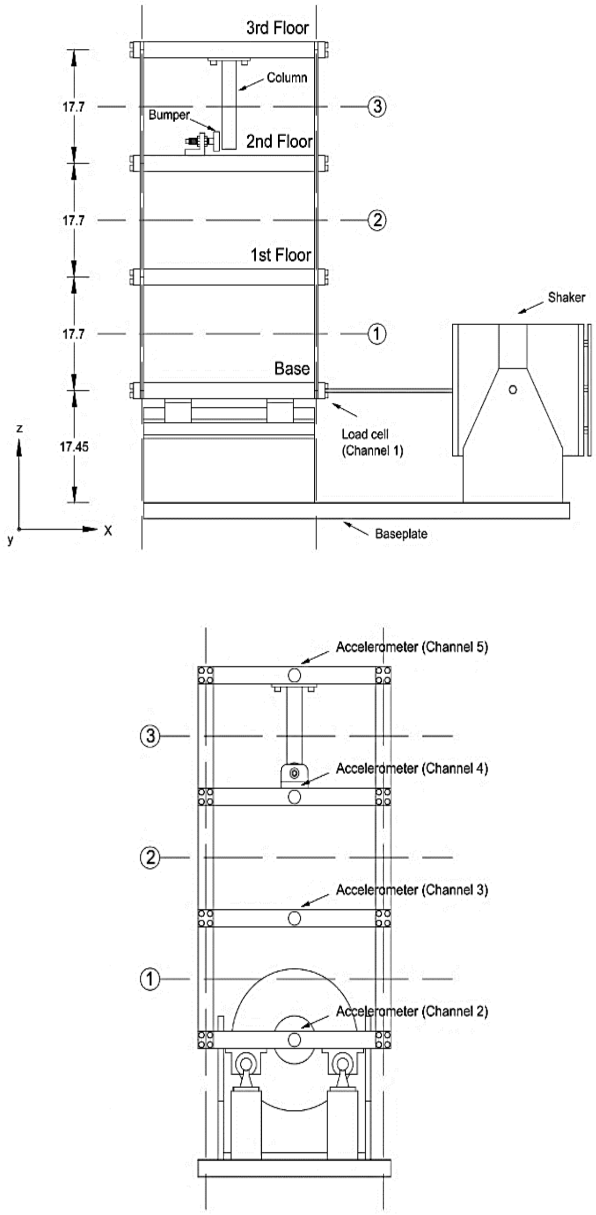


Fig. 1: The three-story laboratory benchmark frame [43]

To induce nonlinear damage, a center column (height 150 mm and cross section 25 mm × 25 mm) was suspended from the third floor. This column was connected to a bumper mounted on the second floor, the position of which could be adjusted to define diverse structural damage. The source of damage is a simulation of breathing cracks to produce nonlinear behavior through opening and closing under excitation forces. The acceleration time-domain responses at all floors and base were measured under 17 structural state conditions as shown in Table 1.

[DOR: 20.1001.1.23454296.2022.6.4.1.8]
[DOI: 10.52547/nmce.6.4.16]

[DOR: 20.1001.1.23454296.2022.6.4.1.8]

[DOI: 10.52547/nmce.6.4.16]

Table. 1: The structural state conditions in the laboratory frame[43]

State	Condition	Description
1	Undamaged	Baseline without damage and environmental and operational variability
2	Undamaged	Added mass of 1.2 kg at the base
3	Undamaged	Added mass of 1.2 kg at the 1st floor
4	Undamaged	87.5% stiffness reduction in one column of the 1st inter-story
5	Undamaged	87.5% stiffness reduction in two columns of the 1st inter-story
6	Undamaged	87.5% stiffness reduction in one column of the 2nd inter-story
7	Undamaged	87.5% stiffness reduction in two columns of the 2nd inter-story
8	Undamaged	87.5% stiffness reduction in one column of the 3rd inter-story
9	Undamaged	87.5% stiffness reduction in two columns of the 3rd inter-story
10	Damaged	Nonlinear damage; distance between bumper and column tip 0.20 mm
11	Damaged	Nonlinear damage; distance between bumper and column tip 0.15 mm
12	Damaged	Nonlinear damage; distance between bumper and column tip 0.13 mm
13	Damaged	Nonlinear damage; distance between bumper and column tip 0.10 mm
14	Damaged	Nonlinear damage; distance between bumper and column tip 0.05 mm
15	Damaged	Bumper 0.20 mm from column tip, 1.2 kg added at the base
16	Damaged	Bumper 0.20 mm from column tip, 1.2 kg added at the 1st floor
17	Damaged	Bumper 0.10 mm from column tip, 1.2 kg added at the 1st floor

The structural state conditions in the laboratory frame were categorized into the four main groups including an undamaged condition (state 1), undamaged conditions with operational and environmental variability (states 2-9), damaged conditions (states 10-14), and damaged conditions with the environmental and operational variability (states 15-17). In state 1, which refers to a baseline condition, there is no change in the laboratory frame. This state implies an ideal condition in the SHM community since there are neither nonlinear changes caused by damage nor linear changes due to the operational and environmental variability in the frame [43]. The states 2-9 provide linear changes to the laboratory frame by adding a concentrated mass or decreasing the stiffness of the frame so as to simulate the environmental and operational variabilities in the undamaged condition.

In order to assess the effects of the environmental and operational variability, the processes of damage localization and damage level estimation are implemented by three different undamaged conditions to extract independent residuals from such conditions as reference data sets. On this basis, the independent residuals of sufficient AR models extracted from the states 1, 3, and 7 are used as the reference data Type I, Type II, and Type III, respectively. Table 2 entirely gives the reference and target data sets for using in the proposed KS method.

Table. 2: The reference and target data sets applied to the proposed KS method

Input data	Data type	Structural conditions
Reference	Type I	1
	Type II	3
	Type III	7
Target	-	10-17

5.1 *Residual-oriented feature extraction by AR model*

In order to extract the independent residuals of AR model as the damage-sensitive features, four AR models are separately fitted to the acceleration time histories acquired from the channels 2-5 in the baseline condition. According to the improved feature extraction technique, the initial orders of AR models are estimated by Bayesian information criterion. Next, the improved orders are determined such that the AR models are able to produce independent residuals at all channels. Table 3 shows the initial and improved estimation of the orders of the AR models.

Table. 3: The initial and improved orders of the AR models at all channels in the baseline condition

Channel No.	2nd	3rd	4th	5th
Initial estimation	36	28	12	16
Improved estimation	45	40	31	35

The sufficient order is 45 because the sufficient model, AR(45), can extract the independent residuals from all channels. It is desirable to verify the accuracy and adequacy of AR(45) by the goodness-of-fit statistics. Table 4 presents the results of numerical statistics for the sufficient AR(45) at all channels.

Table. 4: The goodness-of-fit statistics for the AR(45) in the baseline condition

Goodness-of-fit	Channel No.			
	2^{nd}	3^{rd}	4^{th}	5^{th}
SSE	0.0017	0.0005	0.0001	0.0003
R-square (%)	94.67	97.61	99.68	98.72
Adjusted R-square	0.9467	0.9761	0.9968	0.9872
RMSE	0.0081	0.0063	0.0009	0.0021

In this table, all SSE quantities are roughly equal to zero in the sense that AR(45) is an adequate time series model for the prediction procedure. The values of R-square statistic are close to 100%; therefore, the sufficient AR model is not underfitting. The amounts of adjusted R-square statistic are positive and approximately identical to 1, which mean that the sufficient AR model has appropriately been matched with the acceleration time histories. Eventually, the values of RMSE statistic confirm that the sufficient AR(45) is not overfitting.

Even though the use of high-order models may be a limitation in the time series modeling, it is important to note that the accuracy and sufficiency of a time series model is extracting the independent residuals and providing significant necessity for applying accurate time series models. As stated earlier, a correct time series model should generate the independent residuals; otherwise, the model should be modified. For a comparative evaluation, Fig. 2 illustrates the autocorrelation functions of the residuals for the initial and sufficient AR models at channel 4.

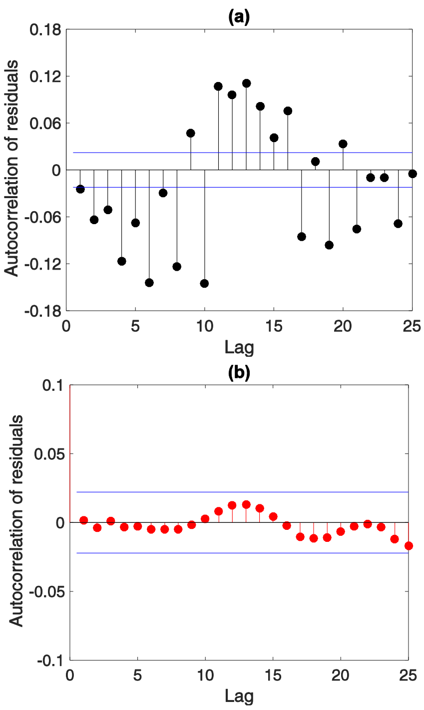


Fig. 2: The autocorrelation functions of residuals at channel 4 in the baseline condition with 95% confidence intervals: (a) the initial model-AR(12), (b) the sufficient model-AR(45)

As Fig. 2(a) reveals, there are considerable correlation patterns that exceed the correlation bounds in the sense that the initial model AR(12) fails to extract the independent residuals. In other words, the initial order of model at channel 4 is inadequate and needs to be improved. By contrast, the samples of autocorrelation function for the residuals of AR(45) are roughly within the correlation bounds, which means that this model is sufficient and its residuals are independent. Note that the process of time series modeling is repeated for the states 3 and 7 to extract the independent residuals as the reference data Type II and Type III. Another point is that the sufficient model of AR(45) is also valid for these states.

Considering the residuals of the sufficient AR models as the damage-sensitive features, a question arises as how to understand their sensitivity to damage. One way to address this problem is to utilize the norm of the AR residuals at the damaged area in all structural conditions. For the laboratory frame, channel #4 is the location of damage; therefore, the l_2-norms of the vector of AR residuals, $\|\mathbf{e}(\mathbf{t})\|_2$, at this channel are computed in the states 1-17 as shown in Fig. 3. Notice that in this figure, UC denotes the undamaged condition (state 1), UCEOV implies the undamaged conditions with environmental and operational variability (states 2-9), DC is abbreviated to the damaged conditions (states 10-14), and DCEOV denotes the damaged condition with environmental and operational variability (states 15-17).

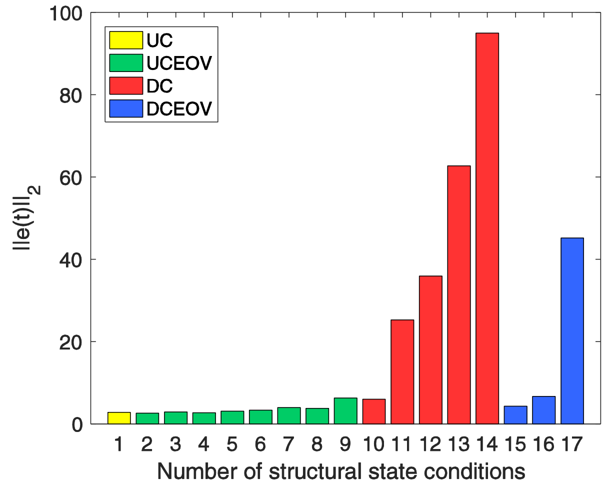


Fig. 3: Evaluating the sensitivity of the AR residuals to damage at the channel 4

As can be observed in Fig. 3, the l_2-norms of the AR residuals in the states 1-9 are invariant, whereas there are considerable variations in the states 10-14 and 15-17. Among all structural conditions, state #14 has the maximum norm value and the rate of changes in the norms of the AR residuals increases with increasing the level of damage from state 10 to state 14. All results obtained from Fig. 3 lead to the conclusion that the residuals of the sufficient AR model extracted from the improved feature extraction technique are sensitive to damage.

5.2 Damage localization

The residuals of AR(45) at the different sensor locations in both undamaged and damaged conditions are chosen as the reference and target variables to use in the proposed KS method. To perform the process of damage localization, it is only necessary to form the similarity vector **D**, in which each quantity represents the KS value at one sensor location.

5.2.1 The KS-based damage localization in the laboratory frame

The amounts obtained from the previous step are applied to the proposed equation of KS, Eq. (5), for identifying the location of damage in the laboratory frame. Before the analysis of results, it is important to mention that the location of channel 4 is equivalent to the damaged area in the laboratory frame owing to the existence of gap as the source of nonlinear damage at this channel. Therefore, the results of damage localization in this section are intended to demonstrate whether the improved and proposed methods are able to identify this channel as the damage location. On this basis, the largest KS value should be obtained at channel 4 in all reference data sets (Type I, II, and III) as shown in Figs. 4-6.

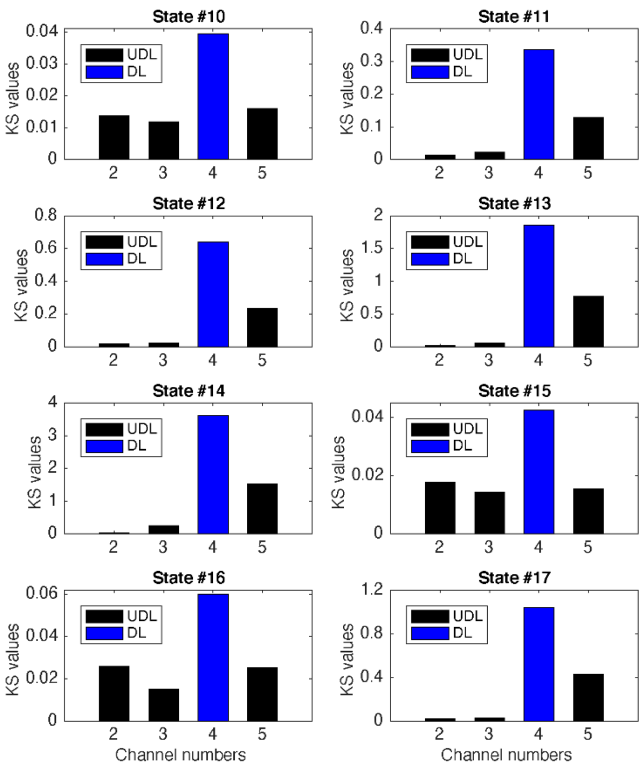


Fig. 4: Damage localization by the proposed KS method in the laboratory frame based on the reference data Type I

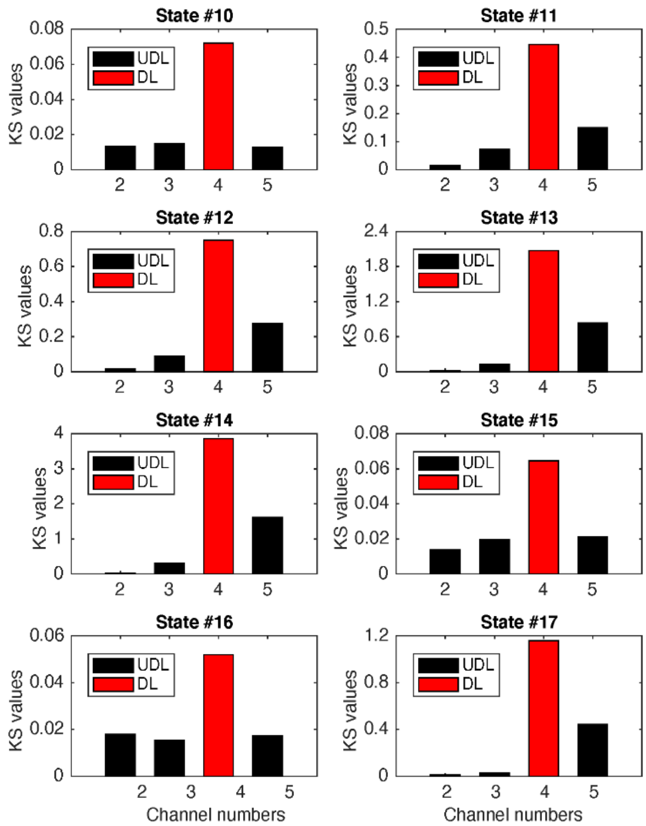


Fig. 5: Damage localization by the proposed KS method in the laboratory frame based on the reference data Type II

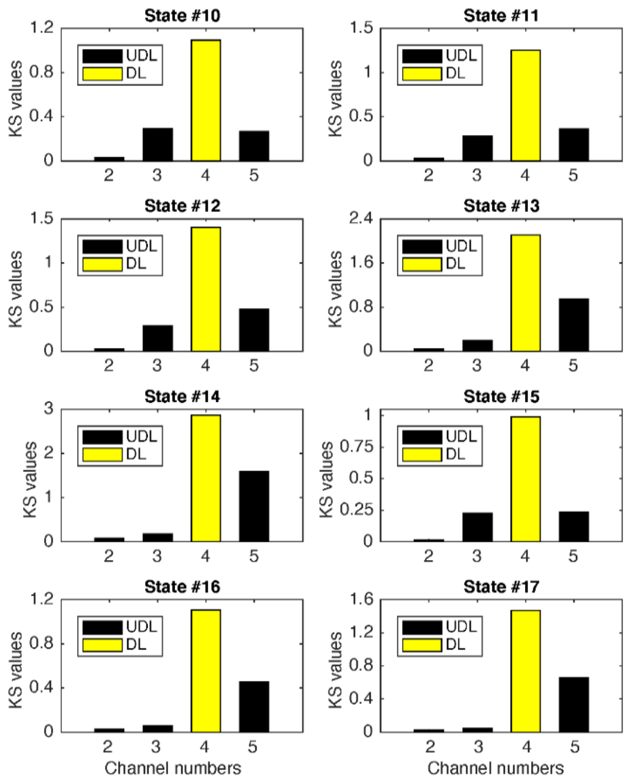


Fig. 6: Damage localization by the proposed KS method in the laboratory frame based on the reference data Type III

[DOR: 20.1001.1.23454296.2022.6.4.1.8]

[DOI: 10.52547/nmce.6.4.16]

Fig. 4 illustrates the results of damage localization in all damaged conditions (states 10-17) when the residuals of the sufficient AR model in the baseline condition are used as the reference data Type I. In this figure, it is obvious that the KS value at channel 4 is maximum; hence, the location of this sensor is indicative of the damaged area in the laboratory frame. By comparing the KS quantities at the DL and UDLs, it can be observed that the similarity value associated with channel 4 entirely differs from the other channel. Therefore, there is enough evidence to identify this channel as the damage location (DL).

The observations of identifying the damage location using the reference data Type II and III are shown in Figs. 5 and 6, respectively. From these figures, it is possible to assess the effects of the environmental and operational variability in both reference and target data sets for the identification of the damage location. As Figs. 5 and 6 indicate, channel #4 has the largest KS quantity implying the only damaged area in the laboratory frame. The observations in these figures also confirm that the operational and environmental variability do not have any adverse influences on the process of damage localization. In particular, the results concerning the states 15-17 entirely verify this claim because in these structural conditions, both the reference and target data sets incorporate the operational and environmental variability.

In addition, the results of damage localization in the states 10, 15, and 16 demonstrate that the improved and proposed methods can precisely identify the location of small damage. These structural conditions have the same damage pattern (the gap similarity is 0.20 mm) with the exception that different linear changes caused by the operational and environmental variability are applied to the structural states 15 and 16. To sum up, the results shown in Fig. 4-6 lead to the conclusion that the improved feature extraction technique and the proposed KS method are successfully able to identify the location of damage even under varying operational and environmental conditions.

5.2.2 A comparative study on the damage localization process

One way to emphasize the capability and reliability of the proposed KS method is to compare it with a well-known similarity approach. On this basis, this research has attempted to provide a comparative analysis on the process of damage localization between the proposed KS method and the well-known Euclidean similarity (EUD) method. Given the vectors of residuals extracted from the sufficient AR model in the undamaged (**X**) and damaged (**Y**) conditions, the EUD method computes the similarity between these vectors in the following form:

$$D_{\mathrm{E}}\left(\mathbf{X}\,||\,\mathbf{Y}\right)=\left(\mathbf{X}-\mathbf{Y}\right)\left(\mathbf{X}-\mathbf{Y}\right)^{\mathrm{T}} \tag{6}$$

At each sensor location, the EUD method yields a scalar quantity implying the similarity between the residuals of the undamaged and damaged conditions. In a similar manner to the proposed KS method, the computation of the EUD method leads a similarity vector, **D**, which consists of *s* elements representing the number of sensors mounted on the structure. Thus:

$$\mathbf{D}=[D_{\mathrm{E}}\left(\mathbf{X}_1\,||\,\mathbf{Y}_1\right)\quad D_{\mathrm{E}}\left(\mathbf{X}_2\,||\,\mathbf{Y}_2\right)\quad \ldots \quad D_{\mathrm{E}}\left(\mathbf{X}_s\,||\,\mathbf{Y}_s\right)] \tag{7}$$

Taking the reference data Type II and III into account, the results of damage localization for the damaged conditions 10-17 in the laboratory frame are shown in Figs. 7 and 8.

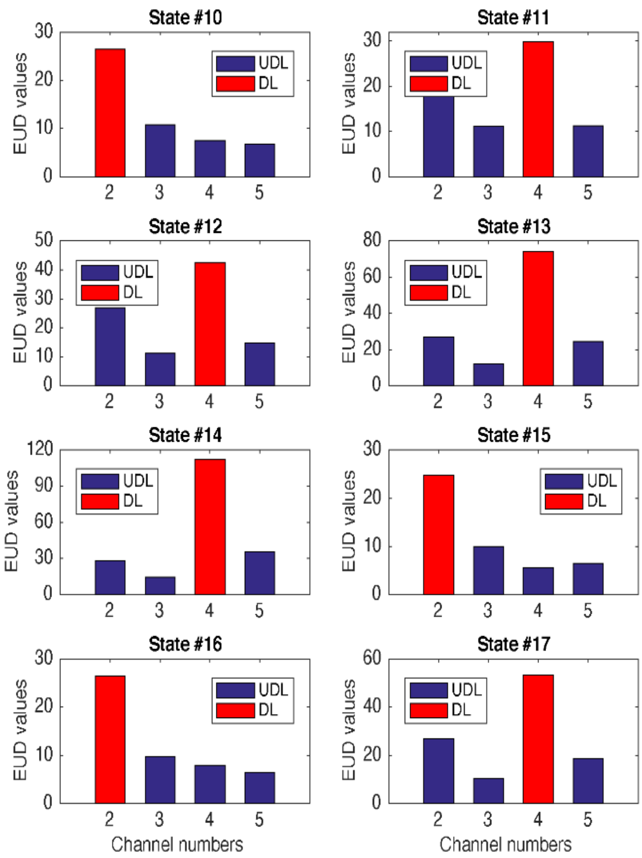


Fig. 7: Damage localization by the EUD method in the laboratory frame based on the reference data Type II

Fig. 7 indicates the results of damage identification using the reference data Type II. From this figure, one can realize that the EUD method is approximately able to identify the damage location with the exception of the small damage scenarios such as the structural states 10, 15, and 16. In other words, the EUD method fails to find channel 4 as the damage location in these states.

Fig. 8 obviously reveals that the operational and environmental variability adversely affect the damage localization results. In all structural conditions with the exception of states 13 and 14, the EUD method cannot satisfactorily identify the location of damage occurred at channel 4. By comparing the results of damage localization between the proposed KS method and the well-known EUD approach, one can conclude that the KS method is an influential and promising similarity technique locating the damaged area even in the presence of the operational and environmental variability in both reference and target data sets.

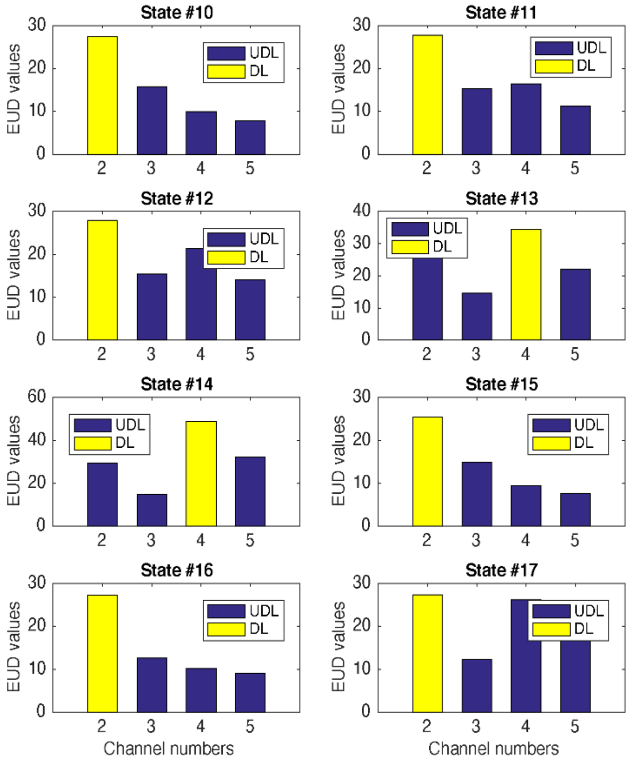


Fig. 8: Damage localization by the EUD method in the laboratory frame based on the reference data Type III

5.3 Damage level estimation

After identifying the precise location of damage, it is investigated whether the improved and proposed methods are capable of estimating the level of damage severity in the damaged conditions. In the laboratory frame, the severity of nonlinear damage increases with reducing the gap similarity between the suspended column and bumper. In this case, the damaged states 10, 15, and 16 give the lowest level of damage severities, while state 14 introduces the highest damage level as remarked in [43]. In order to demonstrate the results of damage level estimation, the KS values at the damage location in the laboratory frame (channel4) for the damaged conditions 10-17 are evaluated based on the three types reference data sets. Fig. 8 shows the results of estimating the level of damage severity.

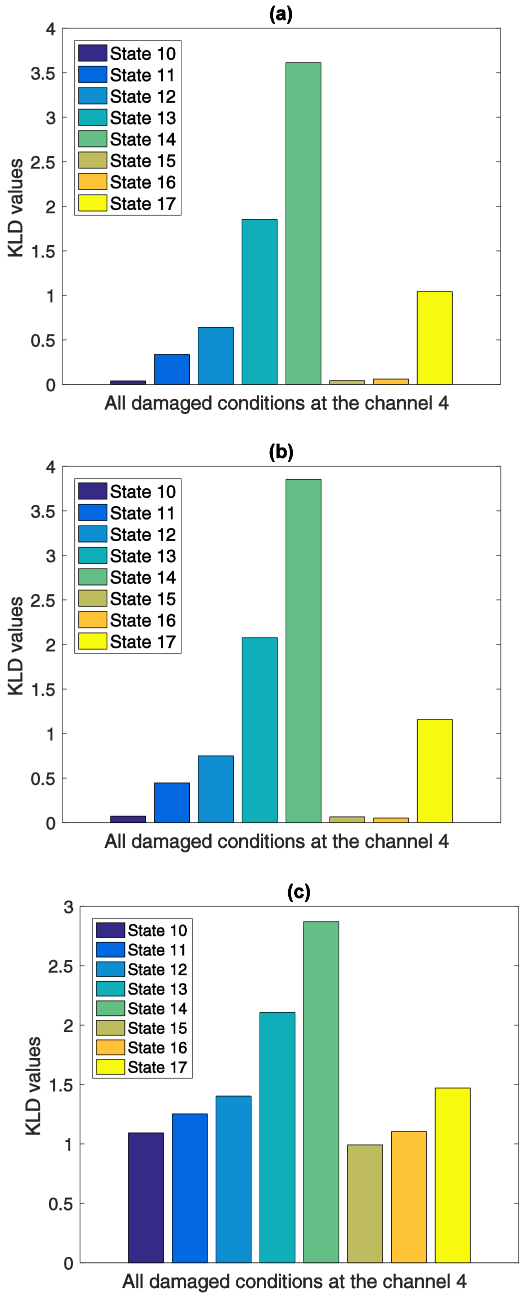


Fig. 9: Estimating the level of damage severity by the proposed KS method at the channel 4: (a) the reference data Type I, (b) the reference data Type II, (c) the reference data Type III

In this figure, it is seen that the values of KS increase with increasing the damage level from state 10 to state 14. In all reference data sets, the structural state 14 provides the largest similarity quantity in the sense that this condition is the highest level of damage severity occurred in the laboratory frame. In addition, the similarity values regarding the states 10, 15, and 16 indicate that these structural conditions have the lowest damage levels. As another conclusion, these conditions give the same

[DOR: 20.1001.1.23454296.2022.6.4.1.8]

[DOI: 10.52547/nmce.6.4.16]

similarity value in each reference data set, which means that the presence of the operational and environmental variability in the target data set do not have any considerable effects on the KS quantities.

Despite the reliable and accurate results in estimating the level of damage severity, there are a few drawbacks in which, most of them pertain to the environmental and operational conditions. In the structural state conditions presented in Table 1, the states 13 and 17 have the same damage pattern (the gap similarity corresponds to 0.1 mm). However, the results of damage level estimation in Fig. 9 reveal a computational error resulting from a discrepancy between the similarity values of these states. In the reference data Type III, on the other hand, the KS amounts have risen compared with the other reference data sets. This situation depends on the presence of the high-level operational and environmental variability (87.5% stiffness reduction in two columns of the second floor) in the reference and target data sets. Therefore, it would be very appropriate to remove such variability, either mass increasing or stiffness decreasing from the damage features, for better damage level estimation.

6. Conclusions

In this study, a new formulation of KS method has been proposed to identify the location of damage and estimate the level of damage severity. An improved residual-oriented feature extraction technique using AR model has been presented to determine a sufficient order of AR model such that the model is able to extract independent residuals. An experimental laboratory frame has been employed to validate the accuracy and reliability of the improved and proposed methods in the context of SHM. In this frame, the effects of operational and environmental variability have been considered in the undamaged and damaged conditions. Furthermore, the goodness-of-fit statistics and the graphical residual analysis by the autocorrelation function confirmed that increasing the initial order of AR model leads to an adequate and accurate time series model without the overfitting or underfitting problems. The experimental results demonstrated that the proposed KS method is a robust and efficient univariate statistical method that can precisely identify the location of damage even in the presence of the operational and environmental variability. Furthermore, it was observed that this method could accurately estimate the level of damage severity in the different damaged conditions. The observations in the experimental model showed that the different types of the operational and environmental conditions do not have any influences on the results of damage localization without false alarms. In all reference data types, channel 4 has been identified as the damage location. The comparative study of the initial and sufficient AR models showed that the sufficient order enables the AR model to extract independent residuals, whereas the residuals of the initial AR model are correlated. In another comparative study, it was observed the KS method provides more precise and robust results in comparison with the well-known EUD method, particularly when the high-level operational and environmental variability are available.

Acknowledgement

The authors would like to acknowledge the Engineering Institute at Los Alamos National Laboratory for accessing the entire experimental data of the three-story laboratory frame.

References

[1] Farrar CR, Worden K. An introduction to structural health monitoring. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2007;365:303-15.

[2] Entezami A, Shariatmadar H, Sarmadi H. Structural damage detection by a new iterative regularization method and an improved sensitivity function. Journal of Sound and Vibration. 2017;399:285-307.

[3] Entezami A, Shariatmadar H, Ghalehnovi M. Damage detection by updating structural models based on linear objective functions. Journal of Civil Structural Health Monitoring. 2014;4:165-76.

[4] Mottershead J, Friswell M. Model updating in structural dynamics: a survey. Journal of sound and vibration. 1993;167:347-75.

[5] Rezaiee-Pajand M, Sarmadi H, Entezami A. A hybrid sensitivity function and Lanczos bidiagonalization-Tikhonov method for structural model updating: Application to a full-scale bridge structure. Applied Mathematical Modelling. 2020;89:860-84.

[6] Rezaiee-Pajand M, Entezami A, Sarmadi H. A sensitivity-based finite element model updating based on unconstrained optimization problem and regularized solution methods. Structural Control and Health Monitoring. 2019;27:1-29.

[7] Sarmadi H, Karamodin A, Entezami A. A new iterative model updating technique based on least squares minimal residual method using measured modal data. Applied Mathematical Modelling. 2016;40:10323-41.

[8] Farrar CR, Doebling SW, Nix DA. Vibration-based structural damage identification. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2001;259:131–49.

[9] Farrar CR, Worden K. Structural Health Monitoring : A Machine Learning Perspective: John Wiley & Sons Ltd; 2013.

[10] Entezami A, Sarmadi H, Behkamal B, Mariani S. Big Data Analytics and Structural Health Monitoring: A Statistical Pattern Recognition-Based Approach. Sensors. 2020;20:2328.

[11] Entezami A, Sarmadi H, Salar M, De Michele C, Arslan AN. A novel data-driven method for structural health monitoring under ambient vibration and high-dimensional features by robust multidimensional scaling. Structural Health Monitoring. 2020.

[12] Sarmadi H, Entezami A, Saeedi Razavi B, Yuen K-V. Ensemble learning-based structural health monitoring by Mahalanobis distance metrics. Structural Control and Health Monitoring. 2020:e2663.

[13] Entezami A, Shariatmadar H. An unsupervised learning approach by novel damage indices in structural health monitoring for damage localization and quantification. Structural Health Monitoring. 2017;17:325-45.

[14] Entezami A, Shariatmadar H, Karamodin A. Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods. Structural Health Monitoring. 2018;18:1416-43.

[15] Chegeni MH, Sharbatdar MK, Mahjoub R, Raftari M. New supervised learning classifiers for structural damage diagnosis using time series features from a new feature extraction technique. Earthquake Engineering and Engineering Vibration. 2022;21:169-91.

[16] Sohn H. Effects of environmental and operational variability on structural health monitoring. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 2007;365:539-60.

[17] Sarmadi H. Investigation of Machine Learning Methods for Structural Safety Assessment under Variability in Data: Comparative Studies and New Approaches. Journal of Performance of Constructed Facilities. 2021;35:04021090.

[18] Sarmadi H, Yuen K-V. Early damage detection by an innovative unsupervised learning method based on kernel null space and peak-over-threshold. Computer-Aided Civil and Infrastructure Engineering. 2020.

[19] Sohn H, Czarnecki JA, Farrar CR. Structural health monitoring using statistical process control. Journal of Structural Engineering. 2000;126:1356-63.

[20] Fugate ML, Sohn H, Farrar CR. Vibration-based damage detection using statistical process control. Mechanical Systems and Signal Processing. 2001;15:707-21.

[21] Nair KK, Kiremidjian AS, Law KH. Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure. Journal of Sound and Vibration. 2006;291:349-68.

[22] Mei L, Mita A, Zhou J. An improved substructural damage detection approach of shear structure based on ARMAX model residual. Structural Control and Health Monitoring. 2016;23:218-36.

[23] Hoell S, Omenzetter P. Optimal selection of autoregressive model coefficients for early damage detectability with an application to wind turbine blades. Mechanical Systems and Signal Processing. 2016;70–71:557-77.

[24] Figueiredo E, Figueiras J, Park G, Farrar CR, Worden K. Influence of the Autoregressive Model Order on Damage Detection. Computer-Aided Civil and Infrastructure Engineering. 2011;26:225-38.

[25] Stull CJ, Hemez FM, Farrar CR. On assessing the robustness of structural health monitoring technologies. Structural Health Monitoring. 2012;11:712-23.

[26] Box GE, Jenkins GM, Reinsel GC. Time Series Analysis: Forecasting and Control (Fourth Edition) New Jersey: John Wiley & Sons, Inc; 2008.

[27] Entezami A, Shariatmadar H. Structural health monitoring by a new hybrid feature extraction and dynamic time warping methods under ambient vibration and non-stationary signals. Measurement. 2019;134:548-68.

[28] Entezami A, Shariatmadar H. Damage localization under ambient excitations and non-stationary vibration signals by a new hybrid algorithm for feature extraction and multivariate distance correlation methods. Structural Health Monitoring. 2019;18:347-75.

[29] Entezami A, Shariatmadar H, Mariani S. Structural Health Monitoring for Condition Assessment Using Efficient Supervised Learning Techniques. Proceedings. 2020;42:17.

[30] Entezami A, Sarmadi H, Razavi BS. An innovative hybrid strategy for structural health monitoring by modal flexibility and clustering methods. Journal of Civil Structural Health Monitoring. 2020;10:845-59.

[31] Sarmadi H, Entezami A, Daneshvar Khorram M. Energy-based damage localization under ambient vibration and non-stationary signals by ensemble empirical mode decomposition and Mahalanobis-squared distance. Journal of Vibration and Control. 2019;26:1012-27.

[32] Sarmadi H, Karamodin A. A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects. Mechanical Systems and Signal Processing. 2019;140:1-24.

[33] Entezami A, Shariatmadar H, Karamodin A. Improving feature extraction via time series modeling for structural health monitoring based on unsupervised learning methods. Scientia Iranica. 2020;27:1001-18.

[34] Behkamal B, Naghibzadeh M, Pagnani A, Saberi MR, Al Nasr K. Solving the α-helix correspondence problem at medium-resolution Cryo-EM maps through modeling and 3D matching. Journal of Molecular Graphics and Modelling. 2021;103:107815.

[35] Behkamal B, Naghibzadeh M, Saberi MR, Tehranizadeh ZA, Pagnani A, Al Nasr K. Three-Dimensional Graph Matching to Identify Secondary Structure Correspondence of Medium-Resolution Cryo-EM Density Maps. Biomolecules. 2021;11:1773.

[36] Behkamal B, Naghibzadeh M, Moghadam RA. Using pattern detection techniques and refactoring to improve the performance of ASMOV. 2010 5th International Symposium on Telecommunications: IEEE; 2010. p. 979-84.

[37] Gul M, Necati Catbas F. Statistical pattern recognition for Structural Health Monitoring using time series modeling: Theory and experimental verifications. Mechanical Systems and Signal Processing. 2009;23:2192-204.

[38] Mosavi AA, Dickey D, Seracino R, Rizkalla S. Identifying damage locations under ambient vibrations utilizing vector autoregressive models and Mahalanobis distances. Mechanical Systems and Signal Processing. 2012;26:254-67.

[39] Balsamo L, Betti R, Beigi H. A structural health monitoring strategy using cepstral features. Journal of Sound and Vibration. 2014;333:4526-42.

[40] Fassois SD, Sakellariou JS. Statistical time series methods for SHM. Encyclopedia of Structural Health Monitoring. 2009.

[41] Entezami A. Feature Extraction in Time Domain for Stationary Data. Structural Health Monitoring by Time Series Analysis and Statistical Distance Measures: Springer; 2021. p. 17-45.

[42] Entezami A. An Introduction to Structural Health Monitoring. Structural Health Monitoring by Time Series Analysis and Statistical Distance Measures: Springer; 2021. p. 1-15.

[43] Figueiredo E, Park G, Figueiras J, Farrar C, Worden K. Structural health monitoring algorithm comparisons using standard data sets. Los Alamos National Laboratory: LA-14393; 2009.

[44] Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice: OTexts; 2014.

[45] Entezami A, Shariatmadar H, Sarmadi H. Condition Assessment of Civil Structures for Structural Health Monitoring Using Supervised Learning Classification Methods. Iranian Journal of Science and Technology, Transactions of Civil Engineering. 2020;44:51-66.

[46] Kullback S, Leibler RA. On information and sufficiency. The annals of mathematical statistics. 1951;22:79-86.

[47] Los Alamos National Laboratory: SHM data sets <http://www.lanl.gov/projects/national-security-education-center/engineering/software/shm-data-sets-and-software.php>.