

Localizing structural damage based on auto-regressive with exogenous input model parameters and residuals using a support vector machine based learning approach

Burcu GUNES*

Department of Civil Engineering, Istanbul Technical University, Istanbul 34467, Turkey

**Corresponding author. E-mail: bgunes@itu.edu.tr*

© The Author(s) 2024. This article is published with open access at link.springer.com and journal.hep.com.cn

ABSTRACT Machine learning algorithms operating in an unsupervised fashion has emerged as promising tools for detecting structural damage in an automated fashion. Its essence relies on selecting appropriate features to train the model using the reference data set collected from the healthy structure and employing the trained model to identify outlier conditions representing the damaged state. In this paper, the coefficients and the residuals of the autoregressive model with exogenous input created using only the measured output signals are extracted as damage features. These features obtained at the baseline state for each sensor cluster are then utilized to train the one class support vector machine, an unsupervised classifier generating a decision function using only patterns belonging to this baseline state. Structural damage, once detected by the trained machine, a damage index based on comparison of the residuals between the trained class and the outlier state is implemented for localizing damage. The two-step damage assessment framework is first implemented on an eight degree-of-freedom numerical model with the effects of measurement noise integrated. Subsequently, vibration data collected from a one-story one-bay reinforced concrete frame inflicted with progressive levels of damage have been utilized to verify the accuracy and robustness of the proposed methodology.

KEYWORDS structural health monitoring, damage localization, auto-regressive with exogenous input models, one-class support vector machine, reinforced concrete frame

1 Introduction

Detection and potential localization of the structural damage at the earliest stage possible in an automated manner eliminating the need for periodic-inspections is the main objective of any structural health monitoring (SHM) scheme. While several vibration-based damage methods relying on identification of modal parameters and finite element model updating approaches have shown promise, the low sensitivity of the global parameters and the model-based nature of their application are regarded among the major hindrances in their potential applications for complex structural systems.

Time series models as representation of dynamic processes, provide solely data-driven means to detect

damage by monitoring the system response with sensors deployed throughout the structure. The model parameters extracted using only the measured data provide damage sensitive features that can be employed in a data-driven methodology to detect damage [1–4]. Machine learning (ML) technologies, further enable these approaches to be applied in an unsupervised automated manner to assess the structural health in the form of an anomaly detection. The idea hinges on the fact that the features representing the undamaged conditions can be used in the context of statistical pattern recognition paradigm to test for abnormality as the indication of damage presence.

Based on the type and the nature of the available vibration data, different time series models exploiting input-output or output only measurements provided that input loading meet specific characteristics, can be used to extract damage sensitive features. For civil engineering structures subjected to ambient vibration, measuring the

excitation signal is often difficult, if not impossible. Hence, within the framework of health monitoring, an output-only SHM method, not requiring the knowledge of the input and can operate with only the response measurements would be ideal for regular health monitoring applications of in-service structures.

As parametric time series representations of the system response, different type autoregressive models with its parameters providing a link between the mechanical properties of the system have been used for SHM purposes [5–8]. Among the available options, pure autoregressive (AR) models with the assumption that the structural response is stationary, stands out to be a popular option since it predicts future values of the system output based solely on the previous observations without demanding the measured input excitation values. The computational ease in the application of least squares minimization to identify the parameters of the model and the robustness and low uncertainty in the model help enhance the popularity of the model. The major limitation of these models for SHM applications, however arise due to the fact that AR models are only pole functions and to mimic the system response that includes measurement noise it is certain that the model, in addition to the physical poles, will include spurious poles. These poles, in the next stage are to be exploited as the damage sensitive features. However, under noisy conditions the correlation between the physical properties of the structural system and the parameters of the time-series model loses its strength as the spurious ones are neither not linked to the physical properties of the system nor to the changes in them.

Polynomial parametric models of the auto-regressive with exogenous input (ARX) structure are the simplest models for representing a dynamic process driven by an input in presence of uncertainties. The stability of their optimal predictors and their robustness when applied to inverse problems for experimental model identification offers advantages that makes them appealing for SHM applications. Bernal et al. [9] validated with experimental data that the residual error of the ARX model can be utilized as a reliable damage detection parameter. Roy et al. [10] explored the performance of ARX model in the presence of parametric uncertainty in damage localization. Unlike the AR-model however, ARX model is an input-output model requiring both the input and past output observations to predict future output measurements. In most typical SHM applications, it is only possible to rely on output measurements acquired from several positions within the structure. In the absence of the measured input signal, estimation of the model through a ‘pseudo-input’ signal may be exploited. A two-step manner estimation of AR-ARX models [1,11,12] or utilizing a single or a set of output sensor measurements for input in creating the ARX models [7,10,13] have been

employed by different researchers to overcome the unmeasured input problem. The numerical simulation examples and experimental data in Refs. [7,10,13] demonstrated that with response at one location chosen as the ‘input’ of the prediction model, based on some simplifications and satisfying certain assumptions regarding the nature of the input excitation, damage locations can be predicted. The two-step methodology has also been reported to be successful with experimental data [1,12] as well as numerical simulations [1,11,12] for locating damaged region.

Recent advances in the sensor and chip technology coupled with improvements in computing power enabled implementation of ML approaches for SHM applications exploiting vibration data. ML approaches are quite effective in extracting relevant features from the identified parameters of the time series models for detecting and at a later stage locating damage [14–16]. These approaches can be grouped in two main categories of supervised and unsupervised techniques. Artificial Neural Networks (ANNs), Expectation-Maximization (EM), Principal Component Analysis (PCA), Support Vector Machines (SVMs), Self-Organizing Maps (SOMs), Convolutional Neural Networks (CNNs) are among the tools that have been utilized for the wide variety of the methods developed for this purpose. Within the ML framework, supervised approaches received considerable attention for damage diagnosis since they have the capability to unravel the detection and localization aspects of the damage identification problem simultaneously. In supervised learning, the ML model is trained using a set of training data with label target values. However, this labeled data requirement of the supervised techniques from all possible damage states for training the machine constitutes a major limitation for SHM applications. Although for certain mechanical systems with limited damage scenarios this may be feasible, for complex structural systems and civil infrastructures collecting such information is very difficult, if not impossible. To overcome such a limitation, unsupervised approaches that do not require labeled data sets from the damage states have been proposed [17–19].

For unsupervised learning, clustering analysis has been interrogated in the recent years for its implementation potential in damage detection problems [20–23]. Techniques including k-means, hierarchical, fuzzy c-means, Gaussian mixture model (GMM), and spectral clustering approaches for categorizing the data in groups or partitions have been explored by several researchers [24–29]. One-class SVM that can produce a binary decision in terms of identifying an unlabeled data as ‘belonging’ or ‘not belonging’ the trained ‘healthy/baseline’ state of the structure, are powerful classifiers and have shown to be a promising tool for as fault

detection especially when trained properly on appropriate damage sensitive features [30–32].

A sub-branch of ML paradigm is the concept of deep learning that fits quite well for SHM applications dealing with large amounts of data [33,34]. The deep-learning models have the capability to capture and learn information that is hidden in the data through deep mining of the monitoring data. The state of the art reviews on data-driven algorithms and applications through ML [35,15] and deep learning [36–38] reveal the notable progress of ML applications in health monitoring of civil engineering systems.

It should be noted that in long-term monitoring applications of SHM, one of the challenging issues is the variability in data resulting from environmental and/or operational conditions. Discriminating between the changes resulting due to damage versus operational variabilities is crucial in limiting the decision making errors to a minimum. Several coping strategies including probabilistic approaches [39,40], anomaly detectors [41–43], neural networks [44,45] and sensitivity based [46] approaches are proposed in the literature for addressing this challenge. The present study, however, assumes that the training data captures all possible environmental and/or operational variations in the measurements collected during the training period.

To enhance the information acquired as the outcome of unsupervised learning beyond the ‘damage detection’ level and facilitate ‘damage localization’ task, statistical measures computed from the extracted damage sensitive features through gathering and fusing data from multiple sensors can have to be employed. For example, Lu and Gao [13] and Nair et al. [5] proposed damage localization indices on the basis of the AR-parameters extracted from Auto Regressive Moving Average models. Itakura distance and cepstral distance measures were explored by Zheng and Mita [6]. Gul and Catbas [7] suggested a damage index based on the fit ratios of the ARX models. Mousavi and Gandomi [44] exploited Mahalanobis distance with AR coefficients and Yao and Pakzad [8] exploited spectral distance with the AR model spectra to analyze damaged states. Roy et al. [10] explored the sensitivity of the statistical distance of ARX model residual error with damage location. Entezami and Shariatmadar [18] developed two damage indices that rely on parameters and residuals of the AR models. Entezami et al. [47] proposed a statistical distance method called partition-based Kullback–Leibler divergence to measure the discrepancy between the extracted features and the undamaged scenario.

The major shortcoming of the unsupervised techniques is the low success ratio for detecting ‘minor’ damage. Concerned with this issue, this paper proposes a novel damage detection and localization approach utilizing both the parameters and the residuals of the ARX model fit to

the response data. More specifically, the coefficients of the ARX model are exploited as damage sensitive features to train the one-class SVM at the baseline state of the structure. The novelty at this stage lies in the fact that the original ARX model representation requiring both input excitation and output response measurements is constructed using only the measured acceleration responses.

Furthermore, the predictions of the ARX model are obtained based on output error residuals rather than equation error ones to detect and locate damage. The theoretical framework allowing for such a manipulation is given in the next section. Following the discussion on theoretical background, the paper outlines the proposed strategy and proposes a statistical residual-based damage index is proposed for damage localization. Damage assessment results of numerical simulations with different damage states is followed by investigation of a 1/2-scale one-story one-bay plane frame subjected to progressive damage through cyclic lateral loading as a case study validating the applicability of the proposed approach for real life applications.

2 Description of auto-regressive with exogenous input model

In the most general sense, time series modeling is a statistical method that postulates the relationship between a number of time series data to predict future data. Invoking the linear system theory stating that the output response to a linear combination of inputs is the linear combination of output responses of the individual inputs, the input–output map for a finite dimensional system displaying linear behavior, can be expressed [48] as follows:

$$y_k = \sum_{i=1}^{n_a} a_i y_{k-i} + \sum_{j=1}^{n_b} b_j u_{k-j+1}, \quad (1)$$

where a_i and b_j are the model coefficients, n_a and n_b are the orders of the autoregressive and exogenous parts, respectively, and the map is known as the ARX model and is represented by ARX (n_a, n_b). Although one can utilize Akaike’s Information Criterion (AIC) for optimal order selection in high order ARX models, in practice it does not always select an appropriate model. With the aim of not using a model order higher than necessary, the order of the system can be selected by analyzing the improvement in percent fit of a function with increase in model order.

Now, assume that an input–output map for the system at the reference ‘undamaged’ state as is obtained by fitting the ARX model given in Eq. (1) to the measured data. For a set of data obtained subsequently from an unknown state of the system, the model predictions based on ‘output-error residual’ approach can be expressed as:

$$\tilde{y}_k = \sum_{i=1}^{n_a} a_i \tilde{y}_{k-i} + \sum_{j=1}^{n_b} b_j u_{k-j+1}, \quad (2)$$

where the superscript ‘~’ added to underline the fact that the ‘output-error’ approach uses the measured values for only the input whereas for the output term the predicted values, not the measured ones, are implemented.

The residuals between the measurements, y_m and the predictions, y can then be computed as:

$$\varepsilon_{yk} = y_{mk} - \sum_{i=1}^{n_a} a_i \tilde{y}_{k-i} - \sum_{j=1}^{n_b} b_j u_{k-j+1}. \quad (3)$$

Using measured values for the output term in the predicted values leads to ‘equation error’ residual approach. The mathematical framework supporting the fact that the output error residuals are superior than the equation error ones for damage detection purposes, is given by Ref. [9].

For cases where input excitation is not available but originating from a single source as force excitation with a single forcing function or base excitation, Lu and Gao [13] have shown that it is possible to eliminate the external excitation input from the ARX model with one of the measured response signals employed as the input to the ARX model. Gul and Catbas [7] exploited the concept of using the outputs from a sensors cluster as the input signals to the ARX model for predicting the response of the reference channel included in that sensor cluster as part of the input. The methodology implemented in this paper also constructs the time-series signature using only the measured acceleration response signals in the form of an ARX model, however the reference channel for which the predictions are made for is excluded from the data set substituted in lieu of input signal.

For the multiple output-single input case, partitioning the output vector for the collocated sensor and the uncollocated ones such that

$$y = \begin{Bmatrix} y^c \\ y^u \end{Bmatrix}, \quad (4)$$

where y^c and y^u depict the measurements at the collocated and uncollocated locations, respectively; one can set-up the ARX model at time instant k , for the output collocated with the input as:

$$y_k^c = \sum_{i=1}^{n_a} a_i y_{k-i}^c + \sum_{j=1}^{n_b} b_j u_{k-j+1}. \quad (5)$$

Re-arranging such that

$$y_k^c = \sum_{i=1}^{n_a} a_i y_{k-i}^c + b_1 u_k + \sum_{j=2}^{n_b} b_j u_{k-j+1}, \quad (6)$$

which leads to

$$u_k = \frac{1}{b_1} \left(y_k^c - \sum_{i=1}^{n_a} a_i y_{k-i}^c - \sum_{j=2}^{n_b} b_j u_{k-j+1} \right). \quad (7)$$

It should be noted that using Eq. (7) in a recursive

manner starting with zero initial conditions, input for which explicit expressions for the first three steps are given below, can be expressed in terms of the collocated sensor output y^c :

$$u_1 = \frac{1}{b_1} y_1^c,$$

$$u_2 = \frac{1}{b_1} \left(y_2^c - \left(a_1 + \frac{b_2}{b_1} \right) y_1^c \right),$$

$$u_3 = \frac{1}{b_1} \left(y_3^c - \left(a_1 + \frac{b_2}{b_1} \right) y_2^c + \left(\frac{b_2^2}{b_1^2} + \frac{b_2 a_1 - b_3}{b_1} \right) y_1^c \right). \quad (8)$$

Since the input excitation is assumed to be coming from a single source, same time variation applies for the remaining uncollocated sensor positions. Hence one can write:

$$y_k^u = \sum_{i=1}^{n_a} \bar{a}_i y_{k-i}^u + \bar{b}_1 u_k + \sum_{j=2}^{n_b} \bar{b}_j u_{k-j+1}. \quad (9)$$

3 One-class support vector machines for detection of damage

One-class SVM is a promising unsupervised learning approach that requires only one-class of data which is from the baseline state of the system for detecting an outlier state. Since acquiring training data from possible states of damage may not be possible or practical for existing real-life structural systems, the one-class SVM is a well-suited candidate as a technique to be employed in the area of SHM for damage detection purposes.

The rationale behind the one-class SVM [49] is to estimate a decision boundary within which most of the training samples, x_i , belonging to one-class are included. Unlike the binary version of SVM, one-class SVM does not have any data for the second class. Hence, treating the origin as the only member of the second class, one-class SVM seeks to find a classification boundary that separates the training data from the origin with the maximal margin. In the linear case, the equation of the hyperplane with a normal vector w and intercept b for a given point vector $w(x)$ can be written as

$$H : w^T \phi(x) + b = 0. \quad (10)$$

Schölkopf and Smola [49] developed an algorithm returning a function $f(x)$ that takes the value of +1 in the region capturing most of the data and -1 elsewhere.

$$f(x) = \text{sgn}((w \cdot \phi(x) + b)). \quad (11)$$

The objective is then to find the maximum-margin hyperplane or the decision surface separating the data

points from the origin with the largest distance, with origin being the only member of the second class. The distance of the hyperplane Eq. (10) can be expressed as:

$$d_H = \frac{|w^T \phi(x) + b|}{\|w\|_2}, \quad (12)$$

where $\|w\|_2$ is the Euclidean norm for the length of w .

To maximize the margin with the closest points called the support vectors, one needs to minimize $\|w\|$. This problem can be cast as a linear quadratic programming problem using a quadratic function, such that the decision boundary is obeyed:

$$\min_w \frac{1}{2} \|w\|_2^2; \text{ s.t. } y_n [w^T \phi(x_n) + b] \geq 1, \forall n. \quad (13)$$

The above expression is called the primal formulation for cases where data are perfectly separable and there is no misclassification. To allow the model to be more flexible in allowing for the possibility of outliers in the training set and to relax the constraints, it is possible to add a slack variable as a penalty for every misclassification for each data point represented by β . So, no penalty means the data point is correctly classified, $\beta = 0$, and at any miss classification $\beta > 1$, as a penalty. Hence soft-margin linear SVM formulation can be expressed as:

$$\min_{w,b,\beta_n} \frac{1}{2} \|w\|_2^2 + C \sum_n \beta_n, \quad (14)$$

such that

$$y_n [w^T \phi(x_n) + b] \geq 1 - \beta_n; \forall n, \quad (15)$$

$$\beta_n \geq 0; \forall n. \quad (16)$$

It should be noted that slack for every variable should be as low as possible and further regularized by hyperparameter C . The above convex optimization problem can be solved by the help of Lagrangian multiplier method that allows for find the local minima or maxima of a function subject to the condition that one or more equations have to be satisfied exactly by the chosen values of variables. The Lagrangian form can be expressed as:

$$L(w, b, \beta_n, \alpha_n, \mu_n) = \frac{1}{2} \|w\|_2^2 + C \sum_n \beta_n - \sum_n \alpha_n \{y_n [w^T \phi(x_n) + b] - 1 + \beta_n\} - \sum_n \mu_n \beta_n, \quad (17)$$

where $\alpha_n \geq 0$ and $\mu_n \geq 0$ are Lagrange multipliers.

After substituting the Karush–Kuhn–Tucker conditions [50] into the above Lagrangian formulation, the dual Lagrangian form, L_D can be obtained as:

$$L_D = \sum_n \alpha_n - \frac{1}{2} \sum_{l=1}^n \sum_{j=1}^n \alpha_l \alpha_j y_l y_j \phi(x_l)^T \phi(x_j). \quad (18)$$

Then the dual problem is recast as:

$$\max_{\alpha_n} L_D \text{ s.t. } 0 \leq \alpha_n \leq C, \sum_n \alpha_n y_n = 0. \quad (19)$$

Solving the dual problem for the Lagrangian multipliers, one can obtain.

$$w = \sum_n \alpha_n y_n \phi(x_i), \quad b = \frac{1}{2} (x_{+1} + x_{-1}) w, \quad (20)$$

where x_{+1} and x_{-1} are the two support vectors belonging to different classes and classification can then follow using the rule given in Eq. (11). The SVMs can also be easily generalized to the nonlinear case by using the so called ‘kernel trick’. This means, one can apply kernels, $k(x_i, x_j)$ to transform all the data into a high dimensional Hilbert feature space to use linear algorithms in the transformed space. By the definition of the kernel, one can substitute an appropriate kernel such as Gaussian to Eq. (18):

$$\phi(x_i)^T \phi(x_j) = k(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}, \quad (21)$$

and obtain the solution independent of complex feature basis Φ .

4 Proposed damage detection and localization methodology

This section describes the proposed unsupervised learning methodology for detecting and locating structural damage. The method uses a one-class SVM in conjunction with the ARX model requiring only the acceleration response data collected from an intact structure to train the machine. The general flow and the key steps of the proposed methodology is presented in Fig. 1. The training stage of the methodology starts with collection of sets of acceleration data from the intact/baseline state of the structure set. The proposed approach to damage detection can be summarized as follows.

1) Acceleration response time histories collected at the baseline state of the system from each sensor are fitted with an ARX model.

2) In fitting an ARX model, each sensor is treated as the reference channel one at a time, the response measurements and the response measured at the sensor designated as the reference is treated as the output while the response measurements collected at the neighboring sensors are treated as the input to the ARX model. This arrangement of input–output pairing is based on the system and the structural connectivity between the sensor locations.

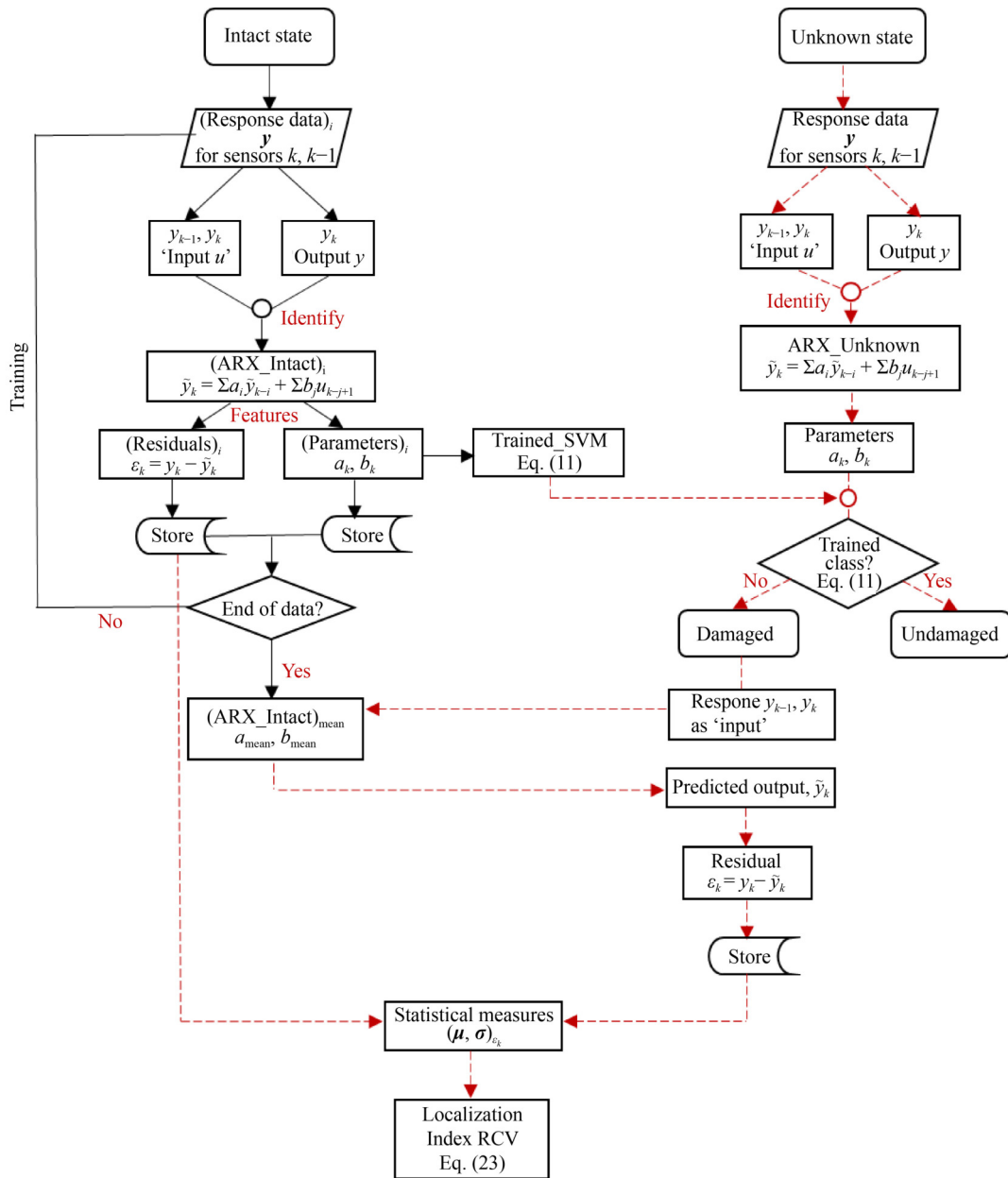


Fig. 1 Unsupervised damage detection methodology employed in this study.

3) The coefficients of the ARX model, referred to as the model parameters, with the predetermined order for each selected reference and the residuals of the predictions are stored in the database as damage sensitive features.

4) The identified model parameters are used to train the one-class SVM for the baseline ‘undamaged’ state of the system. This completes the training stage of the machine for detecting damage.

5) Collection of acceleration response time histories from an ‘unknown’ state of the system initiates the testing stage of the methodology. ARX models are fit to the newly acquired unlabeled data and ARX model parameters are extracted in the same manner. In

extracting the residuals however, the predictions are based on the mean values of the estimated parameters stored in the training stage for the baseline state of the system.

6) The unlabeled ARX model parameters are then fed into the trained SVM for testing whether they fit with the trained class or should be labeled as an ‘outlier’. Outlier classification is interpreted as being indicative of damage presence and localization stage ensues.

Residual analysis is exploited for the localization task since the results of the previous research [8] suggested the ratio of standard deviation of the residual errors as a suitable feature for localizing damage. To compare the variation within the data without depending on the

measurement unit, normalizing the standard deviation with the population mean; more specifically utilizing coefficient of variation (CV), is proposed. The damage localization index in terms of relative coefficient of variation (RCV), is proposed as follows:

$$RCV = \frac{(CV_U - CV_D)^2}{CV_U^2}. \quad (22)$$

The subscripts ‘U’ and ‘D’ denote the baseline (undamaged) and damaged states, respectively. The terms in the ratio are squared to eliminate the possibility of encountering negative values of the index. Substituting the appropriate means and standard deviations, the expression in Eq. (16) can be written as follows:

$$RCV = \frac{\mu_D^2 \sigma_U^2 + \mu_U^2 \sigma_D^2 - 2\mu_U \mu_D \sigma_U \sigma_D}{\mu_D^2 \sigma_U^2}. \quad (23)$$

Hence through comparison of the RCV values obtain at all sensor locations, it may be possible to localize damage with an increase in this index close to the damaged region.

It should be noted that the proposed methodology is entirely data-driven and has no requirement of a numerical model of the system. Damage sensitive features are extracted from a time series model which is originally developed for analyzing long sequences of regularly sampled data. It operates solely using the measured acceleration response data and applicable for cases where input excitation is not available but known to be originating from a single source. Furthermore, the technique requires training data only from the baseline state of the structure representing the baseline condition. Hence, the proposed technique by design is well suited to large scale structural monitoring and can be applied to real life structures with no practical difficulty as long as the machine is trained properly with data covering all possible environmental and/or operational variations. The damage localization information will be within the resolution of the spatial distribution of the sensors, as can be expected.

5 Numerical simulations

Simulated data from an eight-story shear building have been utilized to evaluate the performance of the proposed methodology. With the mass of each story taken as 2×10^4 kg, the undamped natural frequencies at the baseline (healthy) state are summarized in Table 1 for the specified distribution of the story stiffnesses presented in the table. The damping ratio is assumed to be proportional and taken as 2% for all the modes.

The data generation procedure starts with the baseline state. A series of acceleration responses are simulated as

Table 1 Structural parameters of the numerical model

Floor	Story stiffness ($\times 10^7$ N/m)	Mode	f (Hz)
1	40	1	0.12
2	40	2	0.35
3	30	3	0.57
4	30	4	0.76
5	30	5	0.94
6	30	6	1.08
7	30	7	1.19
8	30	8	1.26

free vibration data with a prescribed initial velocity at a selected degree-of-freedom. These responses are generated for a duration of 100 s. with a sampling frequency of 50 Hz ($dt = 0.02$ s). The simulated data are then polluted with different noise levels to represent various operational conditions. In each simulation, the RMS value for the additive measurement noise is selected randomly from a uniform distribution ranging between 2%–5%. They are then standardized to remove the data trends using Eq. (24) in which z is the standardized signal, μ_y and σ_y are the mean and standard deviation, respectively.

$$z = \frac{y - \mu_y}{\sigma_y}. \quad (24)$$

Sample acceleration data simulated in this manner are displayed in Fig. 3.

Once the signal pool is generated in this format, ARX models will be created using only the response signals, with output at certain response locations treated as the ‘input’ while the response of reference channel is treated as the ‘output’. More specifically, the sensor location at which an initial velocity is prescribed, is treated as the reference position and acceleration response at this sensor is treated as the output vector. Based on the nodal connectivity, the acceleration responses of the degrees-of-freedom that are banded around the reference channel are defined as the input vector. Figure 2 displays the arrangement of sensor clusters as input and output vectors for different reference sensor positions. Note that with each floor sensor being treated as the reference channel, a total of eight sets of sensor clusters are created for generation of the ARX models.

Selecting an appropriate order for the regression models is a decision that is directly correlated with the accuracy of these models. While higher model orders may enable better fit of the data that is being utilized, they may not generalize well to other data sets and it is desirable to keep the order at the minimum level. Furthermore, since the aim here is to utilize these models in the context of damage detection, consistency in the model order is more critical than selection of the correct

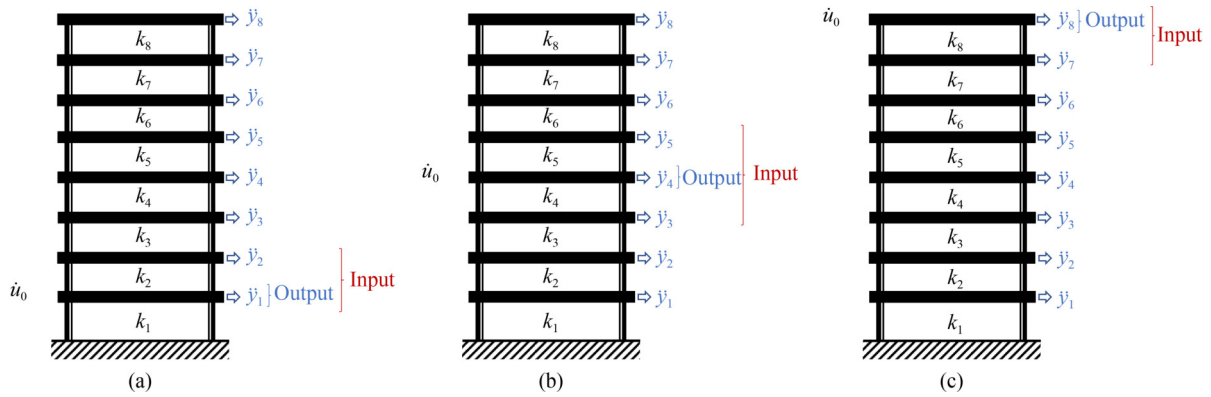


Fig. 2 Arrangement of input and output vectors for the ARX models with the reference sensor at: (a) the first floor; (b) an intermediate floor; (c) the top floor.

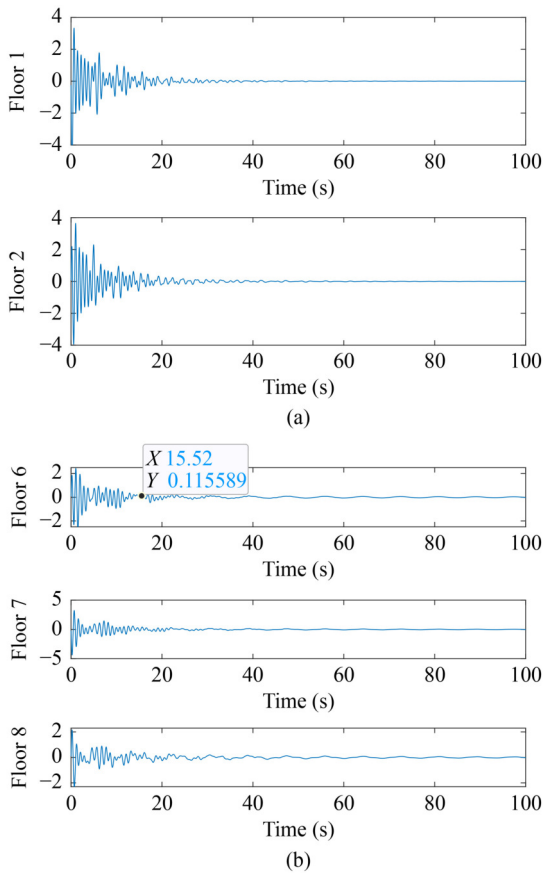


Fig. 3 Sample acceleration response data (cm/s^2) for: (a) floors 1 and 2 with impact at floor 1; (b) floors 6, 7, and 8 with impact at floor 7.

order since the changes in the system is monitored over time. In any case, for real-life systems ‘correct’ model order is an unachievable concept [51].

With these considerations, the *fit* index given in Eq. (25) [48] is adopted here for ARX model order decision in an effort to ensure the accuracy of the model yet avoiding over-parameterization. For the numerically simulated data set, ARX model orders of $n_a = 2$ and $n_b =$

2 and a time delay of 1 leading to a fit index of 90% are found to be appropriate.

$$fit = 100\% \times \left(1 - \frac{\|\hat{y} - y\|}{\|y - \bar{y}\|} \right). \quad (25)$$

A total of 100 simulations are performed at the baseline state of the structure and $ARX(2,2)$ models (with orders $n_a = n_b = 2$) are fit to the acquired acceleration responses following input–output clustering approach discussed previously. The extracted parameters are divided into two data sets. The first data set that includes 80% of the features is used to train the SVM for the healthy state (one-class) of the structure. The second set consisting of the remaining 20%, is used for false positive testing.

The next stage includes simulation of the data corresponding to the ‘damage’ states. Damage is imposed as a reduction in the story stiffness; one floor at a time for each floor resulting in a total of eight different damage scenarios. With 30% reduction of stiffness imposed as damage, 30 simulations are performed for each damage scenario. In these 240 simulations, sensor noise is contemplated randomly in a similar manner to that of the nominal state. $ARX(2,2)$ model fits are performed in the same manner and the parameters of the fitted model are fed into the trained SVM for labeling the state of the structure as damaged or intact. The variation in the natural frequencies of the system for the investigated damage scenarios are summarized in Table 2. It should be noted that the changes, especially in the first few modes are minimal which makes the proposed approach more attractive than alternative approaches tracking the changes in the identified modal parameters.

Damage detection accuracy of the one-class SVM created with 80% of the data simulated for the baseline state of the structure for training, are listed in Table 3. The 20% of the data reserved for validation from the healthy state verified false detection quality of the trained machine with 10% false-positive rate. Table 3 shows that the performance of the trained machine to classify

Table 2 The natural frequencies of the system for the baseline state and the damage scenarios

		Natural frequencies (Hz)							
Healthy state		Damage location							
Mode	f_n	Floor 1	Floor 2	Floor 3	Floor 4	Floor 5	Floor 6	Floor 7	Floor 8
1	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
2	0.35	0.34	0.35	0.35	0.35	0.34	0.34	0.34	0.35
3	0.57	0.55	0.56	0.56	0.54	0.56	0.56	0.54	0.55
4	0.76	0.74	0.76	0.72	0.76	0.74	0.74	0.75	0.72
5	0.94	0.92	0.92	0.91	0.91	0.92	0.90	0.93	0.90
6	1.08	1.07	1.05	1.08	1.04	1.05	1.08	1.03	1.06
7	1.19	1.18	1.15	1.18	1.18	1.16	1.14	1.16	1.18
8	1.26	1.24	1.21	1.22	1.24	1.25	1.26	1.26	1.26

Table 3 Accuracy of the one-class SVM for classifying the health state of the structure

Damage scenarios: damage at floor	Damage detection performance of the SVM (%)	Damage localization performance of RCV (%)
1	93.3	100
2	100	100
3	100	100
4	100	100
5	100	100
6	96.7	100
7	100	100
8	93.3	100

damaged cases is excellent; the largest value for the false negative indication is 6.7% (2 cases out of 30). In all the cases classified as damaged, damage localization was performed successfully. Examination of *RCV* values plotted in Fig. 4 reveals that although the damage index *RCV* attains a value larger than 0.9 at the damaged floor. However, when there is damage at an intermediate floor, with a reasonable threshold limit applied for the damage index, the neighboring floor is also included in the potential damage locations. Nevertheless, for all the damage scenarios examined even if the methodology does not single out the damaged floor, it always identifies the truly damaged floor in the potential damage location.

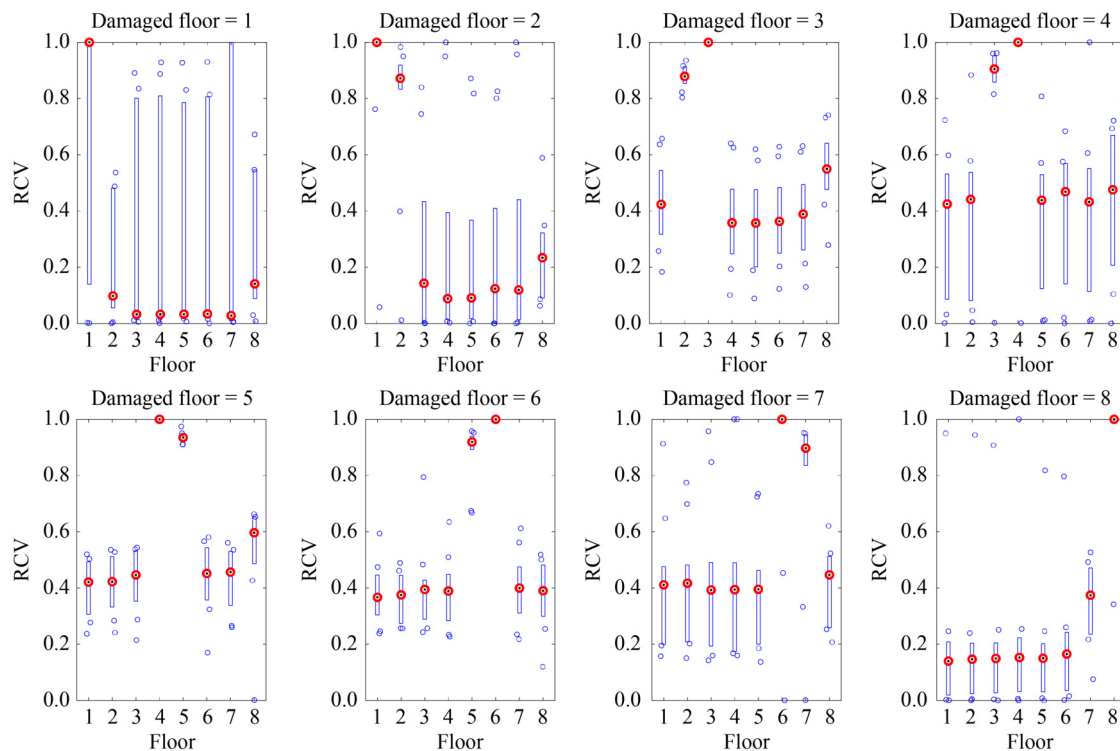


Fig. 4 Damage localization using damage index, RCV for ‘single floor’ damage scenarios.

6 Experimental validation

A 1/2 scale one-story one bay reinforced concrete frame with a height of 1.5 m and a span length of 2 m, shown in Fig. 5, was subjected to cyclic lateral loading in a displacement-controlled mode with a ‘pushover’ setting and damage was inflicted in a progressive manner by

increasing the imposed drift levels [52]. The loading protocol was applied in three repetitions of load-unload cycles with each incremental increase in the drift level. Following the pushover test at each drift level, the lateral loading setup was detached from the frame and vibration tests with impact excitation were performed. The aim of successive loading and vibration tests was to monitor the

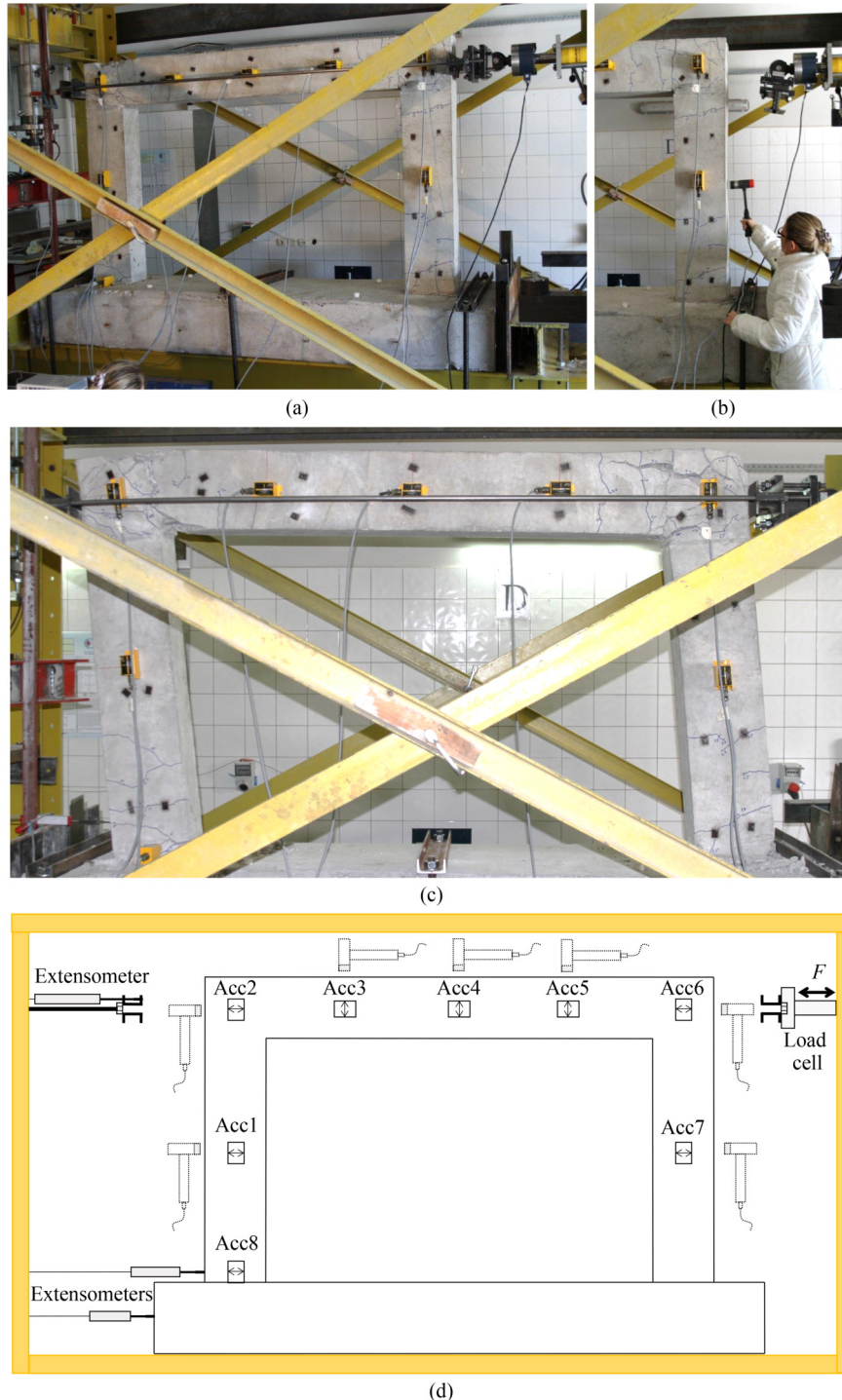


Fig. 5 Experimental set-up: (a) loading frame; (b) impact testing; (c) test specimen during push-over; (d) accelerometer arrangement and impact locations.

progression of damage and the inflicted damage state. The frame was instrumented with eight accelerometers at the designated locations shown in Fig. 5. Hammer impacts were performed at each of these locations except the one located at the foundation level. Quarter point of the span and acceleration data were collected with a sampling frequency of 500 Hz ($dt = 0.002$ s).

The impact tests at the initial ‘baseline’ state of the structure was repeated 10 times over the course of 10 days and the measured pool of response signals were processed to train the one-class SVM utilizing the parameters of the ARX models. In creating these models, the measured acceleration responses at seven accelerometer locations are paired as ‘input–output’ following the clustering strategy depicted in Fig. 6. One can note that the impact location is selected as the ‘reference output’ sensor with the neighboring sensors measuring in the same direction are treated as the ‘input’. The last accelerometer, Acc8, is mounted on the footing to measure the accelerations at the base of the structure.

Table 4 lists the selected drift levels corresponding to

different damage states imposed through the cyclic pushover tests and the variation in the identified modal frequency corresponding mainly to the sway mode. These frequencies are the average values of the repeated tests at each drift level that were identified in Ref. [52]. The observed damage summarized in the table are displayed in Fig. 7.

The procedure depicted in Fig. 1 was employed for all four different states. The one-class SVM trained with ARX model orders of $n_a = 2$ and $n_b = 7$ (ARX (2,7)) and a time delay of 1, which resulted in at least 80% fit in the predicted response was found to represent the ‘intact’ state. The data collected from four different states were then questioned first for existence of damage. The trained one-class SVM classified all 12 data sets corresponding to four different damage states with three repetitions at each stage, including the 0.25% drift state in which no observable damage was reported, as ‘damaged’. Hence, damage localization stage was invoked for all cases examined and normalized RCV values displayed in Fig. 8 were obtained for the damage cases examined. Instead of

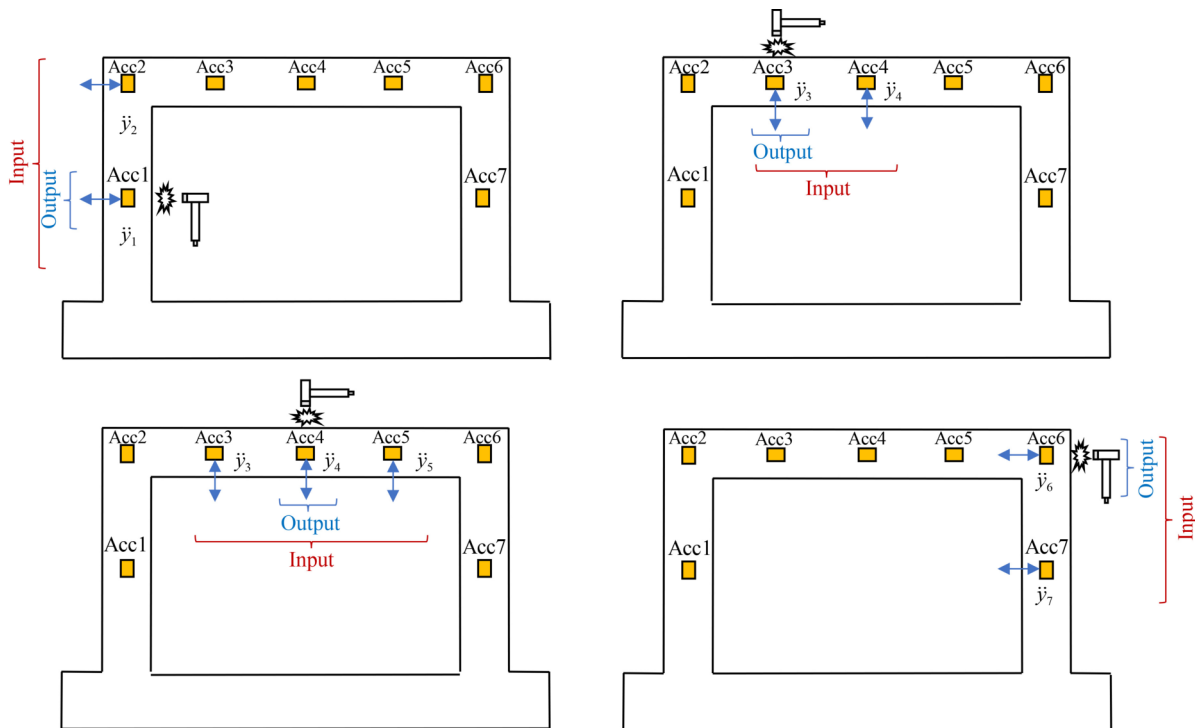


Fig. 6 Arrangement of input and output vectors for the ARX models with a rowing reference sensor.

Table 4 The push-over load test cases and the associated damage states

State	Push level	Drift (%)	$f_{1,ID}$ (Hz)	Observed damage
Healthy	0	0	40.91	initial baseline state
D1	1	0.25	40.22	no damage observed
D2	6	1.5	26.04	cracks around joint 1
D3	7	2.0	21.54	cracks around both joints
D4	9	3.2	17.60	damaged beam–column joints and column base



Fig. 7 Progression of damage in the tested frame for the selected drift ratios.

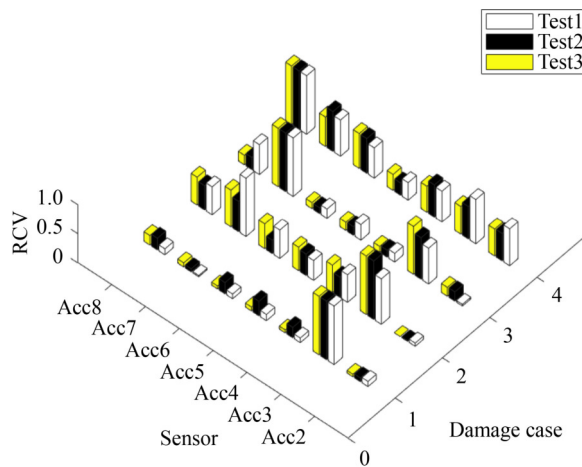


Fig. 8 Damage localization results.

defining threshold limits for the damage index which may vary with the type of structural system and the type of inflicted damage, it is recommended that the general trend in the values should be inspected for decision making.

Investigation of the computed normalized RCV index for the first state D1 at 0.25% drift level reveals potential damage location around accelerometer ‘Acc3’. This case, in fact, is believed to be a misclassification at the detection stage; a ‘false-positive’ detection of damage by the trained one-class SVM based on visual observations. The potential damage location identified by the damage index, however, appears to be consistent with the observed damage at later push-over stages. With

increasing drift levels of 1.5% (D2) and 2% (D3), the sensors located at the beam column connection points are identified as potential damage locations and the progression of damage around the supports becomes pronounced at 3.2% (D4) drift level. Although damage index varies slightly for the repeated tests within a given damage state, the trend in the values within the spatial distribution of the sensors remain consistent and compatible with visual inspections.

7 Concluding remarks

This paper presents a machine-learning based damage identification methodology that can be implemented using measured vibration response of structural systems. The ARX model parameters and residuals extracted from the measured data are utilized as damage sensitive features. Within this framework, the one-class SVM learning strategy, requiring training data only from the baseline healthy state, is selected for binary classification purpose; healthy or damaged. This decision hinges on the fact that for structural systems it is impractical if not impossible to expect that training data from all possible damage scenarios will be available. Unsupervised classification not needing class label information therefore, emerges as the only viable approach that can be implemented for health monitoring purposes for real structural systems. Following the detection stage of the damage identification problem, localization task using the proposed damage index, RCV commences. The

dispersion of the ARX model residual error between the baseline undamaged state and damaged states that is geared toward highlighting possible damage location within the given sensor resolution, is devised for this purpose.

The methodology is implemented on an eight-story shear building model with numerically simulated data generated for damage at different floors and the one-class SVM, eliminating the need for labeled damaged information, with its simplistic classification model allowing faster training and testing verified to be successful in detecting the existence of damage. The localization task carried out with the proposed damage index relying on residual analysis has shown promise in locating the damaged floor. Although for certain damage scenarios, the damage index is unable to single out the damaged floor, damaged region is successfully localized to its true neighborhood. The experimental data obtained from the reinforced concrete frame subjected to progressive damage states have further helped to assess the performance of the approach when the inflicted damage extends to multiple locations. Although with the experimental data false detection quality of the proposed method has not been verified, the identified damaged locations are compatible with the visual observations. Notwithstanding the display of the potential of the proposed algorithm with progressive damage states included in the experimental data, the extension of the proposed damage index for handling multiple damage sites by including additional damage sensitive features and establishing thresholds discriminating between the potential damage locations from the remaining ones, may be a viable improvement for assessment performance of the methodology.

The use of ARX model parameters and residuals in conjunction with one-class SVM and the proposed damage index has proven useful for detecting and localizing damaged region within the given sensor resolution. The experimental data validated the applicability of the presented approach even for progressive damage scenarios where multiple damaged locations can be detected. With the performance of the proposed methodology validating the technique to be potentially viable for unmeasured input excitation cases, a comparative analysis with regards to ML techniques employing features extracted from frequency domain approaches such as the exploitation of the concept of transmissibility for unmeasured input cases as well as deep learning approaches that would eliminate extraction of damage sensitive features entirely in regards to the predictive power of the methodology should be addressed in further research.

Various uncertainties that may affect the accuracy of the proposed methodology includes inconsistent boundary conditions, ambient conditions including

changes in temperature and humidity, during normal operating conditions. The present study assumes that the training data captures all possible environmental and/or operational variations in the measurements collected during the training period. Although this can be justified in a long-term monitoring program, new variabilities that may lead to false identifications may arise in this period. Therefore, enhancement of the proposed methodology to distinguish anomalies due to damage versus variability in the operating conditions is recommended for future work.

It should be noted that, with the conviction that detection and localization of structural damage is the most exacting task of the damage assessment problem, the implemented methodology focuses on these aspects and does not address the issue of damage quantification. Leveraging quantification information from the data without a numerical model requires a larger set of training data and a different architecture to be employed in a supervised manner.

Funding note Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Competing interests The authors declare that they have no competing interests.

References

1. Sohn H, Farrar C R. Damage diagnosis using time series analysis of vibration signals. *Smart Materials and Structures*, 2001, 10(3): 446–451
2. Fassois S D, Sakellariou J S. Time-series methods for fault detection and identification in vibrating structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2007, 365(1851): 411–448
3. Avendaño-Valencia L D, Chatzi E N, Koo K Y, Brownjohn J M. Gaussian process time-series models for structures under operational variability. *Frontiers in Built Environment*, 2017, 3: 69
4. Fassois S D, Kopsaftopoulos F P. *Statistical Time Series Methods for Vibration Based Structural Health Monitoring*. Vienna: Springer Vienna, 2013, 209–264

5. Nair K K, Kiremidjian A S, Law K H. Time series-based damage detection and localization algorithm with application to the ASCE benchmark structure. *Journal of Sound and Vibration*, 2006, 291 (1–2): 349–368
6. Zheng H, Mita A. Localized damage detection of structures subject to multiple ambient excitations using two distance measures for autoregressive models. *Structural Health Monitoring*, 2009, 8(3): 207–222
7. Gul M, Catbas F N. Structural health monitoring and damage assessment using a novel time series analysis methodology with sensor clustering. *Journal of Sound and Vibration*, 2011, 330(6): 1196–1210
8. Yao R, Pakzad S N. Damage and noise sensitivity evaluation of autoregressive features extracted from structure vibration. *Smart Materials and Structures*, 2014, 23(2): 025007
9. Bernal D, Zonta D, Pozzi M. ARX residuals in damage detection. *Structural Control and Health Monitoring*, 2012, 19(4): 535–547
10. Roy K, Bhattacharya B, Ray-Chaudhuri S. ARX model-based damage sensitive features for structural damage localization using output-only measurements. *Journal of Sound and Vibration*, 2015, 349: 99–122
11. Lei Y, Kiremidjian A S, Nair K K, Lynch J P, Law K H, Kenny T W, Carryer E, Kottapalli A. Statistical damage detection using time series analysis on a structural health monitoring benchmark problem. In: *Proceedings of the 9th International Conference on Applications of Statistics and Probability in Civil Engineering*. San Francisco, CA: A.A. Balkema, 2003: 6–9
12. Silva S D, Dias Júnior M, Lopes V Junior. Damage detection in a benchmark structure using AR-ARX models and statistical pattern recognition. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 2007, 29(2): 174–184
13. Lu Y, Gao F. A novel time-domain auto-regressive model for structural damage diagnosis. *Journal of Sound and Vibration*, 2005, 283(3–5): 1031–1049
14. Flah M, Nunez I, Ben Chaabene W, Nehdi M L. Machine learning algorithms in civil structural health monitoring: A systematic review. *Archives of Computational Methods in Engineering*, 2021, 28(4): 2621–2643
15. Avci O, Abdeljaber O, Kiranyaz S, Hussein M, Gabbouj M, Inman D J. A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing*, 2021, 147: 107077
16. Cunha B Z, Droz C, Zine A M, Foulard S, Ichchou M. A review of machine learning methods applied to structural dynamics and vibroacoustic. *Mechanical Systems and Signal Processing*, 2023, 200: 110535
17. Figueiredo E, Park G, Farrar C R, Worden K, Figueiras J. Machine learning algorithms for damage detection under operational and environmental variability. *Structural Health Monitoring*, 2011, 10(6): 559–572
18. Entezami A, Shariatmadar H. An unsupervised learning approach by novel damage indices in structural health monitoring for damage localization and quantification. *Structural Health Monitoring*, 2018, 17(2): 325–345
19. 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*, 2020, 140: 106495
20. Diez A, Khoa N L D, Makki Alamdari M, Wang Y, Chen F, Runcie P. A clustering approach for structural health monitoring on bridges. *Journal of Civil Structural Health Monitoring*, 2016, 6(3): 429–445
21. Yu L, Zhu J H, Yu L L. Structural damage detection in a truss bridge model using fuzzy clustering and measured FRF data reduced by principal component projection. *Advances in Structural Engineering*, 2013, 16(1): 207–217
22. Zhou Y L, Maia N M, Sampaio R P, Wahab M A. Structural damage detection using transmissibility together with hierarchical clustering analysis and similarity measure. *Structural Health Monitoring*, 2017, 16(6): 711–731
23. Cha Y J, Wang Z. Unsupervised novelty detection-based structural damage localization using a density peaks-based fast clustering algorithm. *Structural Health Monitoring*, 2018, 17(2): 313–324
24. Langone R, Reynders E, Mehrkanoon S, Suykens J A. Automated structural health monitoring based on adaptive kernel spectral clustering. *Mechanical Systems and Signal Processing*, 2017, 90: 64–78
25. Silva M, Santos A, Santos R, Figueiredo E, Sales C, Costa J C. Agglomerative concentric hypersphere clustering applied to structural damage detection. *Mechanical Systems and Signal Processing*, 2017, 92: 196–212
26. Sarmadi H, Entezami A, Salar M, De Michele C. Bridge health monitoring in environmental variability by new clustering and threshold estimation methods. *Journal of Civil Structural Health Monitoring*, 2021, 11(3): 629–644
27. 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(6): 04021090
28. Santos J P, Crémona C, Calado L, Silveira P, Orcesi A D. On-line unsupervised detection of early damage. *Structural Control and Health Monitoring*, 2016, 23(7): 1047–1069
29. Qiu L, Fang F, Yuan S. Improved density peak clustering-based adaptive Gaussian mixture model for damage monitoring in aircraft structures under time-varying conditions. *Mechanical Systems and Signal Processing*, 2019, 126: 281–304
30. Anaissi A, Khoa N L D, Mustapha S, Alamdari M M, Braytee A, Wang Y, Chen F. Adaptive one-class support vector machine for damage detection in structural health monitoring. In: *Proceedings of Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Cham: Springer International Publishing, 2017, 42–57
31. Wang Z, Cha Y J. Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect damage. *Structural Health Monitoring*, 2021, 20(1): 406–425
32. Gunes B. One-class machine learning approach for localized damage detection. *Journal of Civil Structural Health Monitoring*, 2022, 12(5): 1115–1131
33. Rafiei M H, Adeli H. A novel unsupervised deep learning model for global and local health condition assessment of structures.

- Engineering Structures, 2018, 156: 598–607
34. Sun L, Shang Z, Xia Y, Bhowmick S, Nagarajaiah S. Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection. *Journal of Structural Engineering*, 2020, 146(5): 04020073
 35. Tibaduiza Burgos D A, Gomez Vargas R C, Pedraza C, Agis D, Pozo F. Damage identification in structural health monitoring: A brief review from its implementation to the use of data-driven applications. *Sensors*, 2020, 20(3): 733
 36. Azimi M, Eslamlou A D, Pekcan G. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review. *Sensors*, 2020, 20(10): 2778
 37. Tian Y, Xu Y, Zhang D, Li H. Relationship modeling between vehicle-induced girder vertical deflection and cable tension by BiLSTM using field monitoring data of a cable-stayed bridge. *Structural Control and Health Monitoring*, 2021, 28(2): e2667
 38. Xu Y, Qian W, Li N, Li H. Typical advances of artificial intelligence in civil engineering. *Advances in Structural Engineering*, 2022, 25(16): 3405–3424
 39. Sarmadi H, Yuen K V. Structural health monitoring by a novel probabilistic machine learning method based on extreme value theory and mixture quantile modeling. *Mechanical Systems and Signal Processing*, 2022, 173: 109049
 40. Sarmadi H, Entezami A, De Michele C. Probabilistic data self-clustering based on semi-parametric extreme value theory for structural health monitoring. *Mechanical Systems and Signal Processing*, 2023, 187: 109976
 41. Dervilis N, Cross E J, Barthorpe R J, Worden K. Robust methods of inclusive outlier analysis for structural health monitoring. *Journal of Sound and Vibration*, 2014, 333(20): 5181–5195
 42. Entezami A, Shariatmadar H, Mariani S. Early damage assessment in large-scale structures by innovative statistical pattern recognition methods based on time series modeling and novelty detection. *Advances in Engineering Software*, 2020, 150: 102923
 43. Entezami A, Sarmadi H, Behkamal B. A novel double-hybrid learning method for modal frequency-based damage assessment of bridge structures under different environmental variation patterns. *Mechanical Systems and Signal Processing*, 2023, 201: 110676
 44. Mousavi M, Gandomi A H. Prediction error of Johansen cointegration residuals for structural health monitoring. *Mechanical Systems and Signal Processing*, 2021, 160: 107847
 45. Cadini F, Lomazzi L, Roca M F, Sbarufatti C, Giglio M. Neutralization of temperature effects in damage diagnosis of MDOF systems by combinations of autoencoders and particle filters. *Mechanical Systems and Signal Processing*, 2022, 162: 108048
 46. Avendaño-Valencia L D, Chatzi E N. Sensitivity driven robust vibration-based damage diagnosis under uncertainty through hierarchical Bayes time-series representations. *Procedia Engineering*, 2017, 199: 1852–1857
 47. Entezami A, Shariatmadar H, Karamodin A. Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods. *Structural Health Monitoring*, 2019, 18(5–6): 1416–1443
 48. Ljung L. *System Identification: Theory for the User*. 2nd Edition. Upper Saddle River: Prentice Hall PTR, 1999
 49. Schölkopf B, Smola A J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT press, 2002
 50. Gale D, Kuhn H W, Tucker A W. *Linear programming and the theory of games*. *Activity Analysis of Production and Allocation*, 1951, 13: 317–335
 51. Horner M, Pakzad S N, Gulgec N S. Parameter estimation of autoregressive-exogenous and autoregressive models subject to missing data using expectation maximization. *Frontiers in Built Environment*, 2019, 5: 109
 52. Gunes B, Gunes O. Vibration-based damage evaluation of a reinforced concrete frame subjected to cyclic pushover testing. *Shock and Vibration*, 2021, 2021: 1–16