DOI: 10.1111/mice.12511

ORIGINAL ARTICLE

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

Machine learning based novelty detection using modal analysis

Ali I. Ozdagli | **Xenofon Koutsoukos**

Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA

Correspondence
Xenofon Koutsoukos, Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37212, USA.
Email: Xenofon.Koutsoukos@vanderbilt.edu

Funding information
National Science Foundation, Grant/Award Number: CNS-1238959

Abstract

Among many structural assessment methods, the change of modal characteristics is considered a well-accepted damage detection method. However, the presence of environmental or operational variations may pollute the baseline and prevent a dependable assessment of the change. In recent years, the use of machine learning algorithms gained interest within structural health community, especially due to their ability and success in the elimination of ambient uncertainty. This paper proposes an end-to-end architecture to detect damage reliably by employing machine learning algorithms. The proposed approach streamlines (a) collection of structural response data, (b) modal analysis using system identification, (c) learning model, and (d) novelty detection. The proposed system aims to extract latent features of accessible modal parameters such as natural frequencies and mode shapes measured at undamaged target structure under temperature uncertainty and to reconstruct a new representation of these features that is similar to the original using well-established machine learning methods for damage detection. The deviation between measured and reconstructed parameters, also known as novelty index, is the essential information for detecting critical changes in the system. The approach is evaluated by analyzing the structural response data obtained from finite element models and experimental structures. For the machine learning component of the approach, both principal component analysis (PCA) and autoencoder (AE) are examined. While mode shapes are known to be a well-researched damage indicator in the literature, to our best knowledge, this research is the first time that unsupervised machine learning is applied using PCA and AE to utilize mode shapes in addition to natural frequencies for effective damage detection. The detection performance of this pipeline is compared to a similar approach where its learning model does not utilize mode shapes. The results demonstrate that the effectiveness of the damage detection under temperature variability improves significantly when mode shapes are used in the training of learning algorithm. Especially for small damages, the proposed algorithm performs better in discriminating system changes.

1 | INTRODUCTION

In recent years, with the emphasis on reliability and sustainability, the interest in structural health monitoring (SHM) has progressively grown. Operations of maintenance, repair, and replacement (MRR) is an integral part of the structure's life cycle (Rytter, 1993). With the aid of SHM, MRR can be prioritized such that the infrastructure requiring immediate

attention can be serviced first. However, due to the presence of environmental and operational variability, it is challenging to develop a reliable damage detection method that informs the performance of the structure accurately (Farrar & Worden, 2007, 2012). Such variations, if overlooked, may lead to incorrect assessment of the structure and cause unnecessary economic loss and social impact. There is still much need for dependable health monitoring approaches that will ensure sustainable civil infrastructure.

Damage detection, also known as *novelty detection*, is, in essence, a method for discriminating significant deviations of a structure from its initial baseline conditions (Sohn, 2007). While ideally the change in the structure can be detected by inspecting features such as natural frequencies, the environmental or operational variations often pollute the baseline and prevent an accurate assessment of the change. Over the last few years, with the advancements in affordable sensor technologies, SHM entered the era of big data (Matarazzo, Shahidi, Chang, & Pakzad, 2015; Liang et al., 2016; Wang et al., 2018). As a result of this, machine learning algorithms started to gain traction as a promising damage detection tool for explaining and modeling the relationship between structural responses and integrity under temporally changing conditions while harnessing the power of big data (Farrar & Worden, 2012; Lin, Pan, Wang, & Li, 2018; Worden & Manson, 2006).

Damage detection methods employing machine learning can be grouped into two classes: (a) parametric and (b) non-parametric. The parametric approaches often rely on characteristic parameters obtained from structural responses. Such methods often fuse one type of learning algorithms with a preprocessing feature extraction algorithm. For example, system identification can be regarded as a preprocessing algorithm capable of computing features such as natural frequencies, mode shapes, and damping ratios of a structure from raw data. A drastic change of the natural frequency is usually relatable to structural damage. The underlying learning algorithm is expected to capture this damage. Likewise, modal analysis methods, such as cross-correlation functions and frequency response functions can extract other strong features of the structure that provide broader information over time and space (Wirsching, Paez, & Ortiz, 2006). Parametric methods are advantageous over their non-parametric counterparts since they do not need to rely on a numerical model of the structure.

For example, Sohn, Worden, and Farrar (2001) developed a parametric novelty detection method that is capable of taking the variations caused by ambient conditions such as a change in loading, temperature, etc. into account to minimize false positive indicators. The method employs autoassociative neural networks (AANNs) to discriminate critical system changes from ambient induced temperature variations. The network is trained via supervised learning to learn the correlation between the variability in the ambient conditions and inherent changes driven by these conditions. The proposed system is tested for a hard-disk model described as a transfer function and it is hypothesized that it could be applied to civil structures. Worden, Manson, and Allman (2003) used a very similar approach involving an AANN and novelty index, and evaluated the approach using a more realistic structural system such as a plate supported by stringers similar to a bridge deck. In this study, frequency response functions are used as the input to the network. Novelty detection through machine learning is also investigated for detecting damages of wind turbine blades under fatigue loading. For example, Dervilis, Barthorpe, Antoniadou, Staszewski, and Worden (2012) and Dervilis, Choi et al. (2012) developed a noise tolerant AANN to evaluate the condition of CX-100 wind turbine blade. The frequency response functions were used in this study which is a similar approach to Worden et al. (2003). Zhou, Ni, and Ko (2011) developed two neural networks, one back propagation neural network (BPNN) and one AANN to detect the damage for Ting Kau Bridge in Hong Kong. The BPNN is used to create a correlation model between damage-sensitive modal frequencies and temperature and AANN is employed to characterize the healthy state of the bridge. After the field data is analyzed, a finite element (FE) model of the bridge is created and simulated to generate new monitoring data where damage was induced in various regions of the model. In addition, the environmental effects were superimposed to the data. Gu, Gul, and Wu (2017) used the modal frequencies of the target structure and the measured temperatures as the input for AANN to improve the generalization capability. In addition, variations in the temperatures causing a change in the frequencies are considered as the input during the training of the network such that false positives can be eliminated. The study looked at the Euclidian distance between measured and estimated frequencies to calculate a novelty index. Their proposed network was tested on a numerical model and in the laboratory on a small-scale test structure. Deraemaeker and Worden (2018) compared the damage detection performance of Mahalanobis squared distance, and principal component analysis (PCA) using real experimental data from a wooden bridge. The features consist of eigenfrequencies and mode shapes measured under changing environmental conditions. Lee, Lee, Yi, Yun, and Jung (2005) and Mehrjoo, Khaji, Moharrami, and Bahreininejad (2008) considered a hybrid approach where an FE model is established as a baseline and neural network is trained to detect the damage based on the expected output from the FE model. These approaches also utilized natural frequencies and mode shapes.

The non-parametric approaches do not require a baseline to establish from structural parameters prior to deployment and do not depend on the uncertainty of system identification or other modal analysis tools. Non-parametric techniques are advantageous, especially when obtaining a dense array

of structural parameters for complex and large systems are challenging. As an example of non-parametric approaches, Abdeljaber, Avci, Kiranyaz, Gabbouj, and Inman (2017) used decentralized 1D convolutional neural networks (CNNs) to eliminate the feature extraction process of typical system identification methods and perform the damage detection directly on the sensor data in real-time. However, sensor data from the healthy and damaged structure is used to train the network for classification purposes which makes the approach supervised learning. Additionally, this study does not consider operational and environmental variability. The algorithm is tested on a grandstand simulator in the lab. In the study, since the trained neural network was not completely successful for classifying the structural condition, specifically producing false negatives, an index reflecting the likelihood of the damage is proposed by computing the ratio of true positives to the total number of test cases. Gulgec, Takáč, and Pakzad (2017) and Y. Yu, Wang, Gu, and Li (2018) used similar approach utilizing deep CNNS to detect damage from sensor data. They also ignore ambient uncertainties. Multiple signal classification (MUSIC) algorithm is another non-parametric approach based on fuzzy wavelet neural networks known to produce successful damage detection from limited sensor data (Amezquita-Sanchez & Adeli, 2015; Amezquita-Sanchez, Park, & Adeli, 2017; Jiang & Adeli, 2007).

This paper introduces an effective damage detection architecture for structures under environmental uncertainty using machine learning. This study utilizes well-established learning algorithms to extract latent features from modal parameters such as natural frequencies and mode shapes under temperature variations and to reconstruct a new representation of these features that is similar, if not identical, to the original. The difference between original and reconstructed parameters constitutes the essential information for detecting critical changes in the system. While modal parameters are known to be a well-researched damage indicator, to the authors' best knowledge, this research is the first time that unsupervised machine learning components such as PCA and autoencoder (AE) are applied to utilize mode shapes in addition to natural frequencies for effective damage detection under environmental variability.

As stated above, the approach proposed herein uses the natural frequencies and mode shapes resulting from a well-recognized system identification tool, Natural Excitation Technique and Eigensystem Realization Algorithm (NExT/ERA) as the input and produces a target output which is the expected natural frequencies and/or mode shapes of the system (Brownjohn, 2003; Caicedo, Dyke, & Johnson, 2004; Caicedo, Marulanda, Thomson, & Dyke, 2001; James, Carne, & Lauffer, 1993, 1995). The damage detection relies on the concept of novelty index which calculates the mean squared error between input and outputs, for example, actual and expected natural frequencies, respectively (Deraemaeker & Worden, 2018). To achieve this goal, two unsupervised learning approaches are investigated: (a) PCA and (b) AE.

To evaluate and validate the approach along with the learning approaches, a simply supported beam structure is modeled and simulated in OpenSees under ambient vibration conditions. To add uncertainty to the simulation, temperature, which is known to affect material properties nonlinearly, is varied over a range. The resulting response data constitutes the reference basis for the training data of the machine algorithms. The modal properties of the structure are extracted from this data set, and the machine learning model (model set A) is trained using the aforementioned approaches. In parallel, another set of models (model set B) is developed using only natural frequencies as the input as it is prescribed in previous studies. To demonstrate the advantages of fusing frequencies with mode shapes further, we introduce a third model (model set C) utilizing only the mode shapes as the input. Finally, the proposed method is evaluated one more time using the same beam exposed to gradient temperature distribution instead of uniform temperature.

Next, three damage cases are considered where stiffness loss is induced at the midspan at various levels. The structure is again simulated under ambient vibrations, and the resulting modal parameters are fed as the testing data set to the learning model. For the three model sets, novelty index is calculated and the reliability of the results are examined to demonstrate the effectiveness of the proposed approach in detecting the damage.

In addition to the simulations, this study considers a data set containing laboratory experiments of a scaled three-story structure created by Los Alamos National Labs for further validation. The structure is tested under various damage scenarios simulating section loss at single and multiple columns. An approach identical to the analytical study is used for training the machine learning model and obtaining the novelty index for each case. Last, an experimental large-scale three-dimensional, three-story structure is identified in laboratory and modeled under temperature gradient. The detection performance of the proposed method is evaluated under multiple damage conditions.

The overall results of the simulations and lab experiments show that the proposed method has, in general, better performance in detecting damage since it utilizes mode shapes as an input in addition to the natural frequencies. In essence, this modal analysis based novel detection approach has the potential to serve as a reliable and near real-time damage detection tool providing accurate data toward objective-driven decisions for maintenance operations. In theory, the end-to-end pipeline considered in this study is capable of streaming real-time data in the time domain from sensors, extract the modal features from the time domain data in near real-time depending on the availability of the computational resources, and compute the novelty index. This approach would indeed

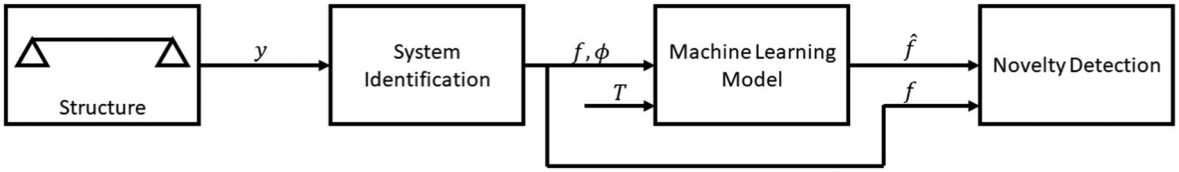

FIGURE 1 Proposed damage detection architecture

accelerate the decision-making process since the state of the target structure is available immediately (Abdeljaber et al., 2017).

In summary, the major contributions of this paper can be summarized as below:

- A new machine learning approach is proposed that relies on natural frequencies and mode shapes.
- This paper streamlines the proposed approach into a pipeline aggregating data collection, system identification, and damage detection.
- For proof of concept, data from simulation and experimental tests are employed. The performance of the proposed method is evaluated by comparing the detection results to those from prior machine learning methods.
- Results demonstrate that the new approach improves the damage detection rate significantly at the presence of environmental variability.

Note: The codes to generate the results presented in this paper can be accessed from https://github.com/aliirmak/ML-SHM.

2 | METHODOLOGY

In this section, the essential components of the proposed approach, illustrated in Figure 1, are explained in detail. First, a general description of the target structure and the problems to be solved are described. Second, the fundamentals of system identification and feature selection are introduced. The third part of this section focuses on the general architecture of the learning components, and the two machine learning models used in this architecture, PCA and AE. The next part which constitutes the final component discusses the implementation of the novelty detection responsible for determining if the structure is damaged. Finally, evaluation criteria to study the performance of the different learning components in detecting the damage are discussed.

2.1 | Structure

In this study, we assume the target structure is a system that can be excited with ambient vibration under a variety of environmental conditions. To minimize errors in damage detection, this study omits the structural responses and the change in the structural mass under service loads. As for environmental variations, it has been shown that temperature plays a major role in affecting dynamic features of the system (Abdeljaber et al., 2017; Sohn et al., 1999; Woon & Mitchell, 1996; Zhou et al., 2011). Hence, this study focuses on the effect of temperature on material properties.

This paper investigates three structures. The first one is an FE model of a simply supported beam. For this beam, a nonlinear relationship between the temperature and the material properties is considered as prescribed by Gu et al. (2017). Additionally, the effect of temperature gradient is investigated to validate the capability of the proposed method further. The second one is a small-scale three-story structure tested by Figueiredo, Park, Figueiras, Farrar, and Worden (2009). Finally, the third one is a three-dimensional, three-story structure tested by Ozdagli (2015). Section 3 presents analytical and experimental structures in detail.

2.2 | System identification and feature selection

System identification is the process of obtaining dynamic and static characteristics of the structure under service, extreme loads, or synthetic excitation. The parameters obtained from identification can be used as an indicator to detect potential damage in the target structure (Doebling, Farrar, Prime, & Shevitz, 1996). In this study, it is assumed that reliable modal parameters can be extracted from the structure under ambient vibration (Farrar, Doebling, Cornwell, & Straser, 1996). To minimize the effect of mass change due to the service load and to minimize the errors in damage detection due to the mass change, this study omits the structural responses and the change in the structural mass under service loads. As a result of this, a well-known modal identification method combination, NExT/ERA is used (Caicedo, 2011; James et al., 1993). This method does not require the external excitation acting on the structure and relies on the ambient vibration measurement which is often available on the field. NExT/ERA takes the structural responses to ambient vibration as the input, which are often accelerations measured at specific locations of the structure with sensors. Then the method produces natural frequencies and the mode shapes defining the dynamic characteristics of the structure for that specific measurement instance. While this study focuses on

one particular system identification method, any approach that is practically applicable in the field can be adopted.

As mentioned above, this component combines NExT with ERA. Essentially, NExT calculates the free response data from ambient data, whereas ERA extracts natural frequencies, mode shapes, and damping ratios from the free response data. Assuming the ambient excitation input is white noise, second-order equation of motion can be written as

$$\boldsymbol{M}\ddot{R}_{\ddot{z},\ddot{z}_i}(\tau) + \boldsymbol{C}\dot{R}_{\ddot{z},\ddot{z}_i}(\tau) + \boldsymbol{K}R_{\ddot{z},\ddot{z}_i}(\tau) = 0 \tag{1}$$

where $\boldsymbol{M}$, $\boldsymbol{C}$, $\boldsymbol{K}$ are mass, damping, and stiffness matrices of the system, $R_{\ddot{z},\ddot{z}_i}(\tau)$ is the cross-correlation function between the acceleration, $\ddot{z}_i$ measured at ith location and a reference acceleration $\ddot{\boldsymbol{z}}$. $R_{\ddot{z},\ddot{z}_i}(\tau)$ has the same form as the free vibration response of the structure to be identified. Here, reference signal, $\ddot{\boldsymbol{z}}$ can be chosen as the acceleration of a node on the structure. By computing the cross-spectral density function with respect to the reference acceleration and applying inverse Fourier transformation, the free vibration response can be obtained in the form of cross-correlation:

$$\ddot{R}_{\ddot{z},\ddot{z}_i} = \frac{1}{N}\sum_{k=0}^{N-1} S_{\ddot{z},\ddot{z}_i}(k) exp\left[j\frac{2\pi kn}{N}\right] \tag{2}$$

where $S_{\ddot{z},\ddot{z}_i}(k)$ is the cross-spectral density function of $\ddot{\boldsymbol{z}}$ and $\ddot{z}_i$, k is the frequency index, and n is the time index. More details on NExT are provided by Caicedo (2011) and Caicedo et al. (2004).

ERA utilizes this free vibration data to determine the modal parameters by first constructing the Hankel matrix:

$$\boldsymbol{H}(k-1) = \begin{bmatrix} Y(k) & Y(k+1) & \dots & Y(k+p) \\ Y(k+1) & Y(k+2) & \dots & Y(k+p+1) \\ & & \ddots & \\ Y(k+r) & Y(k+r+1) & \dots & Y(k+p+r) \end{bmatrix} \tag{3}$$

where $Y(k)$ is the $m \times n$ response matrix at kth time step. A singular value decomposition on Hankel matrix at $k = 1$, $H(0)$ yields:

$$\boldsymbol{H}(0) = \boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{S}^T \tag{4}$$

where $\boldsymbol{R}$ and $\boldsymbol{S}$ are orthonormal matrices, and where $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values. It can be shown that the state-space matrices can be computed as given below:

$$\hat{\boldsymbol{A}} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{R}^T\boldsymbol{H}(1)\boldsymbol{S}\boldsymbol{\Sigma}^{-1/2} \tag{5}$$

$$\hat{\boldsymbol{B}} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{S}^T\boldsymbol{E}_m^T \tag{6}$$

$$\hat{\boldsymbol{C}} = \boldsymbol{E}_n^T\boldsymbol{R}\boldsymbol{\Sigma}^{-1/2} \tag{7}$$

where $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{B}}$, $\hat{\boldsymbol{C}}$ is the estimated state matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, respectively; $\hat{\boldsymbol{D}} = 0$; $\boldsymbol{E}_m^T = [\boldsymbol{I}\ 0]$ and $\boldsymbol{E}_n^T = [\boldsymbol{I}\ 0]$. Juang and Pappa (1985) discuss ERA method in detail.

By applying eigenvalue problem on $\hat{\boldsymbol{A}}$, the modal parameters such as natural frequencies, f, and mode shapes, Φ, can be calculated by

$$(\hat{\boldsymbol{A}} - \lambda I)\Phi = 0 \tag{8}$$

$$w = \left|\frac{\log\lambda}{T_s}\right| \tag{9}$$

where $w = 2\pi f_s$ and T_s is the sampling time.

2.3 | Machine learning model

Often the system identification methods are sensitive to the changes in the system induced by damage or environmental and operational effects. However, it is also a challenging task to differentiate the damage from such variations since the baseline is polluted (Sohn et al., 2001). For instance, studies conducted by Ni, Hua, Fan, and Ko (2005), Liu and DeWolf (2007), Xu, Chen, Ng, Wong, and Chan (2009), Xia, Chen, Weng, Ni, and Xu (2012), Gonzalez (2014), and J. Li (2014) have shown that the temperature can cause significant changes in dynamic properties of structures. With the aid of the unsupervised learning approaches, a higher-fidelity baseline condition of the structure can be extracted from the polluted data set. Here, the objective of the learning model component is to learn a representation of the data set typically through dimension reduction and to reconstruct a new representation that is similar, if not identical, to the original data. In essence, both PCA and AE, also known as AANNs can be used to form this behavior. This study uses those two models interchangeably to extract latent features and evaluates the performance of the architecture by how well the damage is detected. Here, both approaches (PCA and AE) assume that the training data for the learning enabled component contains majorly normal data (from undamaged structure) and very few anomalies (outliers due to instantaneous abnormal events, poor data processing, etc.) (Chalapathy & Chawla, 2019; Chandola, Banerjee, & Kumar, 2009). If there is statistically significant event (caused by damage but not environmental effects) deviating from baseline, then unsupervised approaches are expected to capture this event; thus, the error between actual data and the reconstructed/expected data increases. This process is advantageous especially where human experts have difficulty detecting and observing the damage by looking at the data if there is too much variability. The novelty index presented here is not a damage classification, but rather a signal that something has changed in the system and owner of the

structure may act on this signal considering the risk, operation, and maintenance cost.

The learning component is essentially a mapping process and it can be formalized as

$$\hat{X} = G(X) \tag{10}$$

subjected to

$$min\|\hat{X} - X'\| \tag{11}$$

where X is the input, X' is the subset of X to be reconstructed, and $\hat{X}$ is the output representing the reconstructed X'. $G(\cdot)$ is the mapping function and $\|\cdot\|$ is the normalization operator. To achieve this objective, the mapping function $G(\cdot)$ should be trained with known input X. The input is defined as a set of n natural frequencies where $f = [f_1\ f_2 \dots f_n]$ and n mode shape vectors, $\Phi = [\Phi_1\ \Phi_2 \dots \Phi_n]$. Ambient temperature, T taken during the time of measurement can be also added, if available, to the input since it is considered as a feature in damage detection (Zhou et al., 2011; Gu et al., 2017). A complete input from one measurement instance can be defined as

$$X^{(i)} = vec([T^{(i)}\ f^{(i)}\ \Phi^{(i)}]) \tag{12}$$

where i is the index for ith measurement and $vec(\cdot)$ is the vectorization operator. $f^{(i)}$ and $\Phi^{(i)}$ are obtained through system identification and $T^{(i)}$ is the temperature taken during the system identification measurement. Similarly, the input to be reconstructed and the output are defined as

$$X'^{(i)} = vec(f^{(i)}) \tag{13}$$

$$\hat{X}^{(i)} = vec(\hat{f}^{(i)}) \tag{14}$$

where $\hat{f}^{(i)}$ is the reconstructed representation of the input $f^{(i)}$.

Compared to previous research relying only on natural frequencies (Gu et al., 2017; Zhou et al., 2011), this study considers mode shapes also as a valid input. The modal parameters can be obtained from eigenvalue analysis of K and M as follows:

$$[\boldsymbol{K} - (2\pi f_i)^2 \boldsymbol{M}]\{\Phi_i\} = 0 \tag{15}$$

When the Young's modulus property of the material, E, changes due to the temperature variations, the stiffness matrix, $\boldsymbol{K}$, is affected linearly while $\boldsymbol{M}$ remains same. The relationship between the reference stiffness, $\boldsymbol{K}$, and the temperature-affected stiffness, $\boldsymbol{K'}$, can be simply described by

$$\boldsymbol{K'} = c\boldsymbol{K} \tag{16}$$

where c is a factor defining the linear relationship. When another eigenvalue analysis is applied to $\boldsymbol{K'}$ and $\boldsymbol{M}$, it is observed that the mode shapes stay the same while natural frequencies change as follows:

$$[\boldsymbol{K'} - (2\pi f_i')^2 \boldsymbol{M}]\{\Phi_i\} = 0 \tag{17}$$

As an illustration, a two-story shear frame structure with lumped masses and rigid beams studied by Kim, Christenson, Phillips, and Spencer (2012) and X. Li, Ozdagli, Dyke, Lu, and Christenson (2017) is considered as given in Equation (18):

$$\boldsymbol{M} = \begin{bmatrix} 2.701 & 0 \\ 0 & 2.701 \end{bmatrix} \mathrm{N \cdot mm^{-1} \cdot s^{-2}}$$

$$\boldsymbol{K} = \begin{bmatrix} 558.343 & -279.171 \\ -279.171 & 279.171 \end{bmatrix} \mathrm{N \cdot mm^{-1}} \tag{18}$$

An eigenvalue decomposition on this system using Equation (15) will yield the natural frequencies of $f = [1.00, 2.62]$ Hz and mode shapes such that

$$\Phi = \begin{bmatrix} -0.32 & -0.52 \\ -0.52 & 0.32 \end{bmatrix} \tag{19}$$

Assuming c is 1.05, that is, there is a 5% deviation in the stiffness due to temperature, the new stiffness matrix will be

$$\boldsymbol{K'} = c\boldsymbol{K} = \begin{bmatrix} 586.260 & -293.129 \\ -293.129 & 293.129 \end{bmatrix} \mathrm{N \cdot mm} \tag{20}$$

Using eigenvalue decomposition presented in Equation (17), the natural frequencies of the *shifted* system will be $f' = [1.02, 2.68]$ Hz. However, the mode shapes will remain unchanged and will be equal to Equation (19). This observation indicates that the mode shapes are independent of temperature variations and should always remain the same as long as the structure is not damaged or the mass of the structure does not change. To sum up, the training algorithm considers the persistence of mode shapes as a statistically important feature for developing a proper mapping function $G(\cdot)$. The significance of this observation will be discussed further in Section 2.4.

2.3.1 | Reconstruction using principal component analysis

PCA is a machine learning algorithm that reduces the dimensionality of a data set leading to a simpler representation of it while preserving essential information that defines the data set (Fukunaga & Koontz, 1970; Goodfellow, Bengio, & Courville, 2016). This property of PCA is achieved by computing a linear transformation matrix which can project the original data containing correlated variables to another representation with uncorrelated variables. One main

advantage of this decorrelation is exposing the so-called principal components that explain the dominant patterns in the data (Tibaduiza, Mujica, & Rodellar, 2012; L. Yu, Zhu, & Cheri, 2010; Zang & Imregun, 2001). By selecting the prevailing components, one can compress the data, in other words, reduce the dimension of the data, and expose the most important features that are still faithful to the original.

The linear transformation of the PCA can be represented by

$$Y = X' W_R \tag{21}$$

where X' is the $n \times p$ input data matrix, and n rows and p columns correspond to data points (number of measurement instances containing modal parameters) and features (number of modal parameters), respectively. W_R is the $p \times k$ transformation matrix where k is the number of PCA components to be used that explains the majority of variance for the input data. Y is the PCA projection, that is, a reduced representation of X' with the dimension of $n \times k$. A reconstructed representation of the original input $\hat{X}$ can be obtained by mapping Y back to p dimensions using W_R^T as follows:

$$\hat{X} = Y W_R^T = X' W_R W_R^T \tag{22}$$

The transformation matrix, W_R is the reduced form of W which is derived through a singular value decomposition on X' such that $X' = U \Sigma W^T$. Here, W is $p \times p$ square matrix and W_R contains the first k singular values of W. Goodfellow et al. (2016) discuss the derivation of PCA in detail.

As an alternative to PCA, nonlinear PCA (NLPCA) proposed by Kramer (1991) can be also adopted within this architecture since the environmental variations are defined as nonlinear. While this study considers a nonlinear relationship between temperature and structural dynamic parameters, the results from PCA were satisfactory enough not to pursue this adoption.

2.3.2 | Autoencoder

Autoencoder is a special type of neural network that is trained to reproduce its input as its output (Goodfellow et al., 2016). Usually, autoencoder consists of an encoder and decoder, see Figure 2. Both encoder and decoder are a set of neural network layers.

Here, the encoder takes the input X and translates it to H using the mapping function F described with a set of hidden neural network layers. This mapping can be described as following:

$$H = F(X) \tag{23}$$

The encoder extracts the latent representation of the input that captures the most important features. The decoder function takes H and translates it to $\hat{X}$ using the demapping function G:

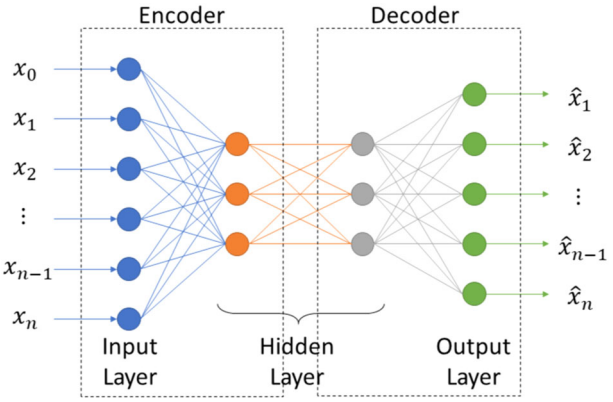

FIGURE 2 An example representation of autoencoder

$$\hat{X} = G(H) \tag{24}$$

The decoder reconstructs a copy of the input by using the latent representation generated by encoder. In summary, the entire autoencoder can be rewritten as

$$\hat{X} = G(F(X)) \tag{25}$$

The training process is performed with an objective function to minimize the error between X and $\hat{X}$ given in Equation (11).

It should be noted that Figure 2 is an example representation. It is difficult to relate the depth of the network and number of neurons at each layer to physical features such as the natural frequencies and mode shapes (Shwartz-Ziv & Tishby, 2017). By rule of thumb, the network parameters are configured such that the resulting model is generalizable and provides an accurate prediction for untrained data as well (Goodfellow et al., 2016). As a result of this, the number of layers and the number of neurons should be tuned depending on the complexity of the system. In addition, the autoencoder can be used to reconstruct the entire input data set or parts of it. For problems where the relationship between the effects of environmental variations is highly nonlinear, AE with nonlinear activation functions is expected to yield more accurate predictions compared to PCA. However, for the examples presented in this study, both PCA and AE provide comparable performance.

2.4 | Novelty detection

The objective of the learning component is recovering an *expected* reconstruction of the original input while eliminating environmental effects. To train the learning component and develop a proper mapping function, the training set is expected to be sampled from the measurements while the

structure is undamaged. After training, when the approach is given data samples from the undamaged structure, the learning component is expected to create a copy of the input as the output. When damage is present, the mapping will generate faulty copies since the new data set is outside of the training data cloud. The novelty detection component quantifies such differences by identifying the existence of new patterns. This component is well researched and often used in past literature (Sohn et al., 2001; Worden, 1997a, b).

The novelty index (*NI*) that describes the similarity between input and the reconstructed copy can be formulated as follows:

$$NI = \|\hat{X} - X'\| \tag{26}$$

This equation is similar to Equation (11) in nature. Novelty index normalizes the difference between the input (original data) X' and the output (reconstructed data) $\hat{X}$. Accordingly, assuming the training of the learning component is performed successfully, *NI* is expected to be zero or close to zero since $\hat{X} \approx X'$. At the presence of damage, *NI* increases since the learning algorithm produces inaccurate results for $\hat{X}$.

2.5 | Evaluation criteria

To quantify the performance of the approach, a modified version of Euclidean distance of novelty index between damaged and undamaged structure is calculated. This criteria can be described as

$$D_{ud,d} = \frac{\|NI_{ud} - NI_d\|}{\mu_{NI_{ud}}} \tag{27}$$

where NI_{ud} and NI_d are the novelty indices for the undamaged and damaged structure, respectively, and $\mu_{NI_{ud}}$ is the mean value of the novelty index for No Damage Case. Here, Euclidean distance is normalized such that a more reliable comparison can be made between architectures and damage case. The larger this distance, the easier it is to detect the damage based on the novelty index.

3 | EVALUATION OF THE PROPOSED METHOD

This study uses three sets of data to verify the proposed approach: (a) an FE model of a simply supported beam; (b) experimental testing of a small-scale three-story structure; and (c) testing and simulation of a large-scale three-dimensional, three-story structure. This section presents and evaluates the results of the structural damage detection performance.

3.1 | Software implementation

The structural responses are obtained from the FE model or the experimental test setup in undamaged and damaged conditions as accelerations, $\ddot{y}$. The accelerations are recorded for some amount of time and saved in a file for each instance of simulation or experiment. Each of these instances containing accelerations is analyzed using NExT/ERA implemented in MATLAB 2018b (MATLAB, 2018). After natural frequencies and mode shapes are obtained from NExT/ERA, this information is vectorized. If the temperature is recorded for an instance, it is also augmented to the vector. Before the training, the data is standardized using *StandardScaler* from scikit-learn toolbox 0.20.2 (Pedregosa et al., 2011) such that each feature has zero mean and unit variance. All the scaled data are saved in their relevant files, based on the condition of the structure.

Next, the machine learning model is trained using the data from the undamaged condition. About half to two-thirds of the data is used for training whereas the remaining data is utilized for testing and validation to make sure overfitting is prevented. Both PCA and AE algorithms are implemented in Python 3.6.7 (Rossum, 1995). The PCA model is trained using scikit-learn toolbox 0.20.2 (Pedregosa et al., 2011). By trial and error, an appropriate number of components are selected to explain the variance of the data. The reconstruction is performed by first transforming the input data to reduced data and then applying an inverse transformation which is explained in Equation (22). AE is trained using Keras 2.2.4 running on TensorFlow 1.12 (Abadi et al., 2015; Chollet, 2015). By trial and error, a neural network with four layers (shown in Figure 2) is developed to capture salient features of the data. The output of the AE model is the natural frequencies to be reconstructed. The models that contain the natural frequencies and mode shapes are called model set A. In parallel, another set (model set B) is developed using only natural frequencies as the input. Additionally, a third model (model set C) is trained which uses only mode shapes.

Finally, the novelty index is obtained by comparing the input natural frequencies with the output for model set A and B or by comparing the input natural frequencies with the output for model set C. Effectively, there is one novelty index for each vector. Data from different damage conditions are tested as well in this last step. This step is also implemented in Python.

It is important to note that, specifically for model set A, while mode shapes could also be a part of the output vector to be reconstructed, the scale of mode shapes is not the same as frequencies; thus, their contributions to the novelty indices may not be as dramatic as natural frequencies. Moreover, the results presented in the following sections demonstrate that the proposed architecture is capable of detecting damage without reconstructing mode shapes. Reducing the dimension of

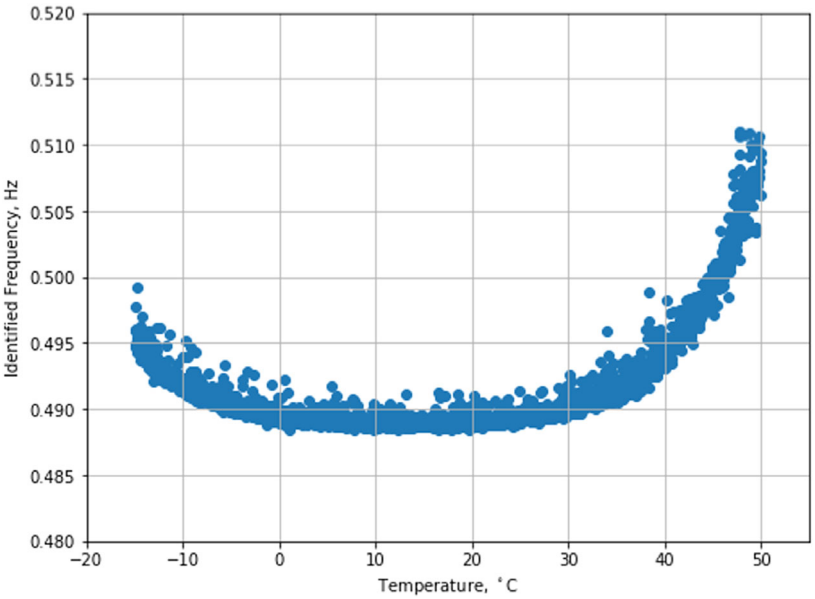

FIGURE 3 Distribution of identified first natural frequency with respect to temperature

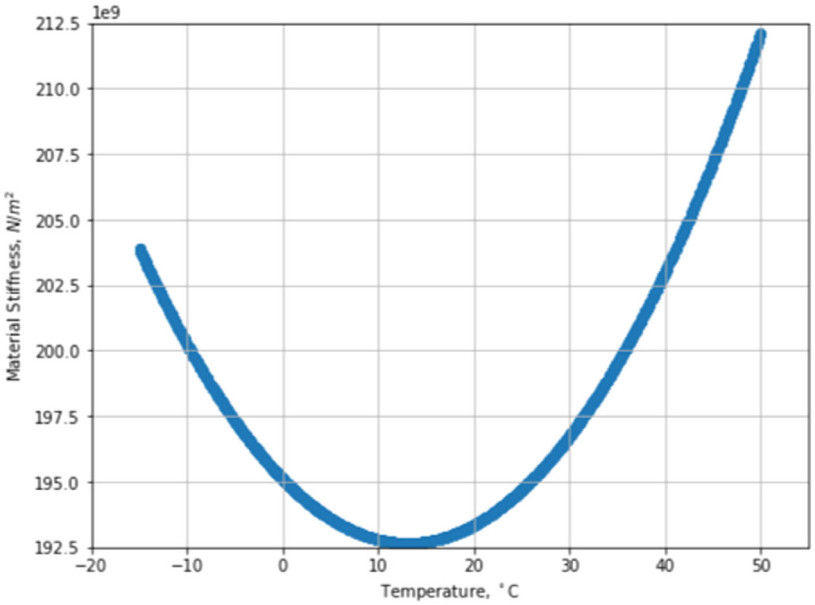

FIGURE 4 Distribution of stiffness with respect to temperature

the output not only accelerates the learning but also reduces the risk of curse of dimensionality (Hughes, 1968).

3.2 | Analytical verification with simply supported beams

A simply supported steel beam used by Gu et al. (2017) with a span length of $L = 5.0$ m is discretized into 40 equally long members having a cross-sectional area of $A = 1.624 \times 10^{-3}$ m^2 and moment of inertia of $I = 1.971 \times 10^{-6}$ m^4. The beam is modeled using finite element modeling (FEM) tool, Open System for Earthquake Engineering Simulation—OpenSees (McKenna, Scott & Fenves, 2010). The members are assumed to be elastic-beam column elements. A nonlinear relationship between material stiffness of the elements, E and temperature, T is described as given below:

$$E = [206.216 - 0.4884\,T + 0.0044\,T^2] \times 10^9\ \text{N} \cdot \text{m}^2 \quad (28)$$

The mass is adjusted such that the structure has the first natural frequency at nearly 0.49 Hz when the temperature is 15°C (∼ 60°F).

TABLE 1 Analytical data matrix

State condition	Description	No. of data
No Damage Case	Baseline condition	2,000
Damage Case 1	5% stiffness reduction at midspan	1,000
Damage Case 2	15% stiffness reduction at midspan	1,000
Damage Case 3	50% stiffness reduction at midspan	1,000

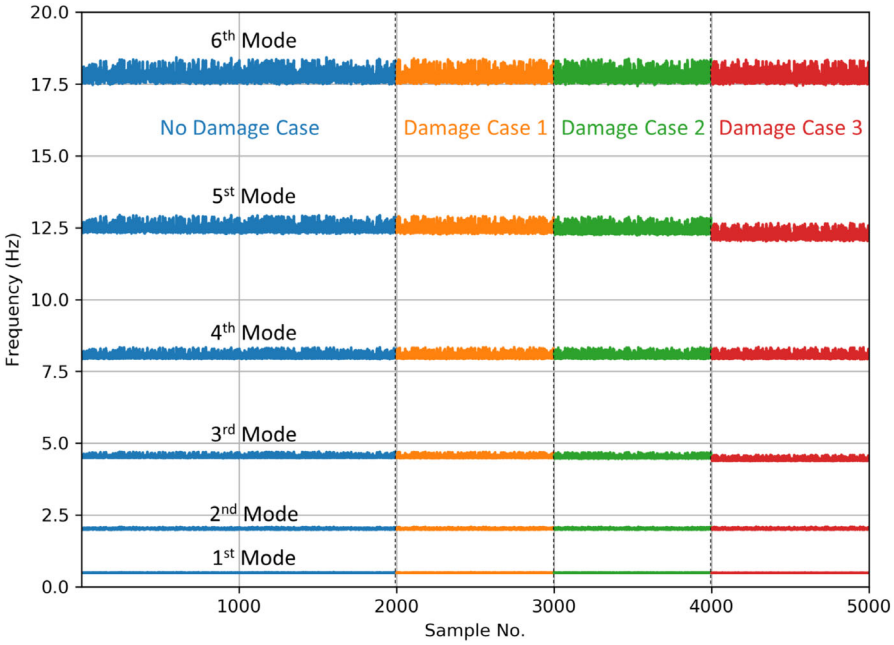

FIGURE 5 Distribution of natural frequencies with varying ambient temperatures for each damage case in analytical data

TABLE 2 Model set properties for analytical data

	Model set A	Model set B	Model set C
Input	1 temperature data 6 natural frequencies 234 mode shape data points	1 temperature data 6 natural frequencies	1 temperature data 234 mode shape data points
Output	6 natural frequencies	6 natural frequencies	234 mode shape data points
PCA component size	100	3	100
AE network structure	241-12-12-6	7-3-3-6	235-50-50-234

Following the architecture discussed in the previous section, the training and validation data set for the undamaged structure is developed by applying ambient vibration made of white noise to the supports of the beam vertically. The input white noise has a bandwidth of 1,024 Hz and the peak displacement is about 0.1 g. The vertical acceleration responses to the ambient excitation at 39 nodes (excluding 2 support responses) are sampled at 200 Hz for 300 s. For each simulation, the ambient temperature governing the material stiffness (see Equation 28) is randomly varied between −15°C and 50°C bounded by a uniform distribution. The temperature range is selected to lay out the nonlinear relationship between temperature, material stiffness, and natural frequencies fully (see Figures 4 and 3). Additionally, the distribution allows the environmental effects to contaminate data over the entire temperature range. Figure 4 illustrates the temperature versus stiffness computed according to Equation (28) for the undamaged case. The difference between the minimum and maximum values of stiffness corresponds to nearly 10% of the minimum stiffness. Figure 3 demonstrates the temperature versus identified frequency distribution for the undamaged case. The difference between the minimum and maximum values of natural frequencies corresponds to nearly 4% of the first natural frequency of the undamaged structure.

A set of damage conditions are defined for this structure (see Table 1). In total, 4,000 simulations are executed. NExT/ERA is performed on the resulting data to extract the first six natural frequencies f and mode shapes for each natural frequency Φ. These first six modes also constitute the features to be used for damage detection in accordance with Gu et al. (2017). From each simulation, including the ambient temperature, six natural frequencies, and 234 mode shape points (6 modes × 39 mode shape points per mode), a vector of 241 data points is created which establishes the input data for the learning enabled component. Out of 2,000 vectors from undamaged case, randomly selected 1,000 vectors are used for training the machine learning component. The remaining data is used for the validation. For the three damage cases considered here, the damage is emulated by reducing the stiffness of the 20th element from the left support (corresponding to the midspan) by 5% (Damage Case 1), 15% (Damage Case 2), and 50% (Damage Case 3). For each damage case, 1,000

simulations are executed under uniformly distributed random ambient vibrations varying between −15°C and 50°C. It should be noted that the temperature range used in the simulations is rather wide and is not observed for most climate conditions. However, this range also introduces relatively large variability to the natural frequencies. The proposed algorithm is expected to robustly detect damage under large temperature variations.

Figure 5 presents the distribution of natural frequencies for the no damage and damage cases using system identification. One can observe that the differences in the frequencies are visually not evident, especially between No Damage Case and Damage Case 1. This can justify machine learning algorithms capable of capturing latent features of the presented data.

Regarding the machine learning component, as mentioned before, two architectures are considered: PCA and AE. For either architecture, the training input is the 1,000 vectors each containing the following data points: (a) For model set A, including temperature, 241 data points (1 temperature data + 6 natural frequencies + 234 mode shape data) are packed as a vector from each simulation. (b) For model set B, only six natural frequencies and temperature data are used. (c) As for model set C, 234 mode shape points and temperature data are utilized. The number of components used for PCA and network architecture for AE are tabulated in Table 2 for each model set.

The novelty index for both architectures is presented in Figure 6. The effect of mode shapes to the performance of the approach is shown by comparing the novelty index of each architecture when mode shapes are used and omitted (model set A, B, and C). It is evident from the visual comparisons that including mode shapes into the learning improves the performance of the detection. When mode shapes are not present, there is an overlap between No Damage Case and Damage Case 1 for both architectures. This overlap may lead to false positives or negatives degrading the performance of detection when the damage is small. However, when the damage is larger, the overlap is not observable anymore. To summarize, the proposed approach is successful in capturing the small damage compared to the primitive model which employs only natural frequencies. For large enough damages, the utilization of mode shapes does not improve the outcome of the detection further since the novelty indices are distinguishable enough for primitive models. The modified Euclidean distances computed using Equation (27) for each damage case and architecture are tabulated in Table 3. Here, for each case, the novelty indices for the No Damage Case from the validation data set relevant to that case are used as the reference, NI_{ud}. The mean of NI_{ud} establishes $\mu_{NI_{ud}}$. For No Damage Case specifically, the comparison is made between the validation and training data. To calculate $D_{ud,d}$, the complete novelty index vector is used. When the mode shape is introduced to the training, distances become smaller for all cases. For PCA, at the absence of mode shapes, the No Damage and Damage Case 1 values are similar for both architectures. However, AE has a higher distance suggesting that it may be still possible to detect damage with AE. At the presence of mode shapes, the distances are much larger which signifies improved damage detection for the given structure. In summary, evaluation of the architecture performance demonstrates that the introduction of mode shapes enhances damage detection. When only mode shapes are considered for reconstruction, it is observed that the relative distances increase. This is due to the fact that more features are reconstructed compared to model sets A and B at the expense of computational complexity. For AE network, reconstructing only mode shapes (model set C) does not improve the detection, whereas for PCA, the sensitivity of model set C is much higher. At this point, it is up to the designer how much sensitivity is desired and what are the computational resources available to reach the desired damage detection sensitivity.

3.3 | Effect of gradient temperature distribution

While in this paper we hypothesized that the mode shapes do not change under uniform temperature distribution, mode shapes will show slight variation if there is temperature gradient. Such variations may cause some degradation in the performance of the proposed method. This section focuses on the effectiveness of the method under temperature gradient. Here, it is assumed that the temperature difference between each end of beams is 10°C and changes linearly across the beam. The same number of inputs are used for all the learning components. In addition, to increase the sensitivity of the AE network for model set A, the system structure is modified to 241-50-50-6. The novelty index for both architectures is presented in Figure 7. In general, the proposed method can detect the damage under temperature gradient. The overall findings are consistent with the results from the uniform distribution.

3.4 | Experimental verification

3.4.1 | Structure 1

For further verification of the proposed approach, a small-scale three-story structure tested by Figueiredo et al. (2009) at Los Alamos National Laboratory is studied. This structure is excited with an electromagnetic shaker attached to its base. The shaker provided a band-limited white noise and the resulting acceleration responses of the structure are recorded for about 25 s at a sampling rate of about 320 Hz (see Figure 8). Figueiredo et al. (2009) indicated that the lab environment is not temperature controlled and some temperature variations observed. However, they also did not record the ambient temperature during the experiments. A set of damage conditions

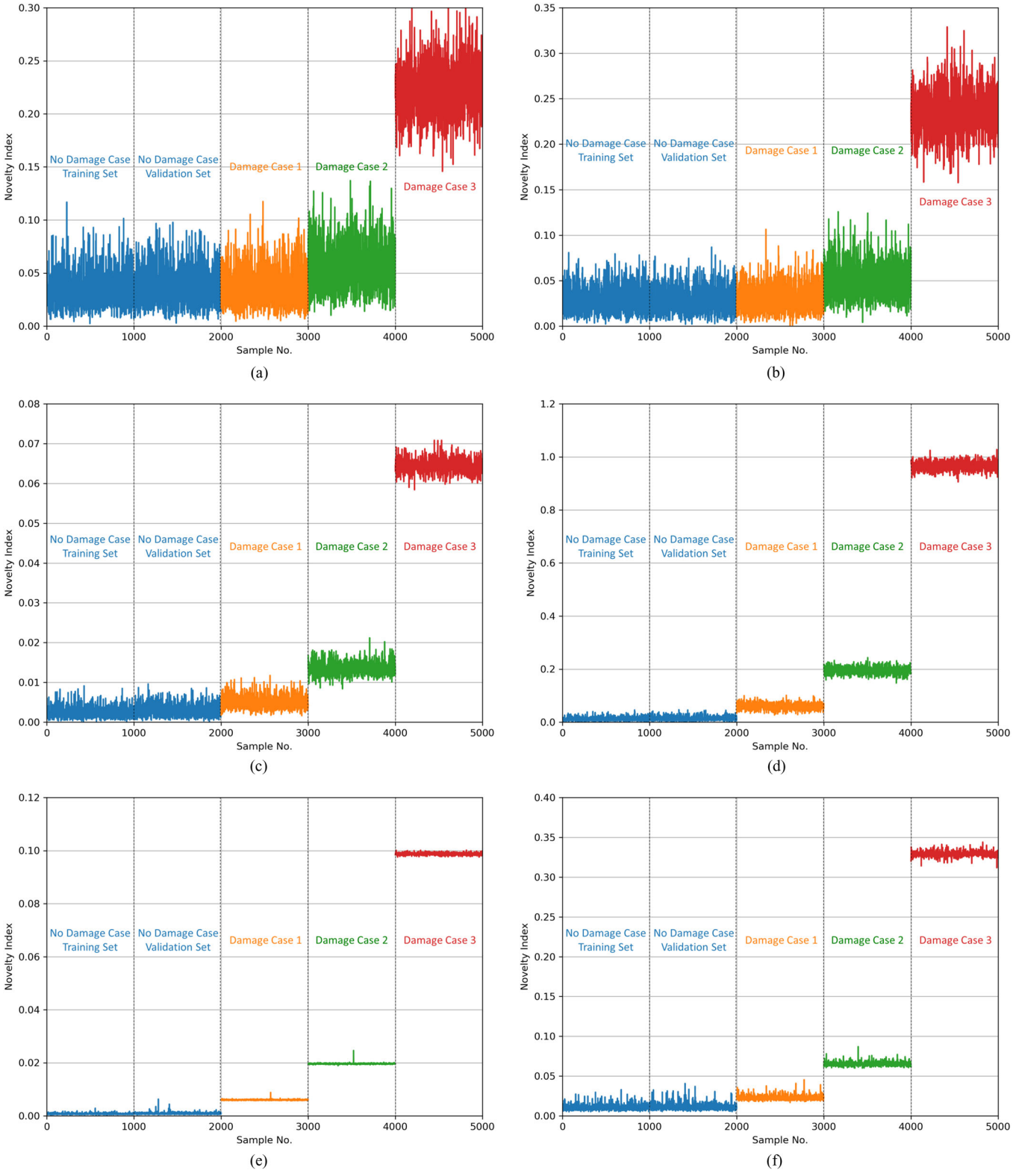

FIGURE 6 Comparison of novelty indices for analytical data: (a) PCA, model set B (mode shapes not included); (b) AE, model set B (mode shapes not included); (c) PCA, model set A (mode shapes included); (d) AE, model set A (mode shapes included); (e) PCA, model set C (only mode shapes included); (f) AE, model set C (only mode shapes included)

TABLE 3 Modified Euclidean distances for damage cases with and without the inclusion of mode shapes for analytical data

	PCA model set A	PCA model set B	PCA model set C	AE model set A	AE model set B	AE model set C
No Damage Case	21.90	23.30	16.28	22.74	22.89	14.59
Damage Case 0	22.78	38.74	174.30	40.57	24.45	29.32
Damage Case 1	29.50	146.43	637.90	158.40	38.31	129.40
Damage Case 2	168.59	820.53	3,333.75	949.47	243.21	770.84

are defined for this structure, see Table 4. Including no damage condition, there are five damage cases for this structure. The damage is introduced by reducing the stiffness of one or two columns at each floor by 87.5%. NExT/ERA is applied to all of the listed experimental data. Since the sampling time is short, the system identification is not able to determine all the dominant modes for all simulations. For the test data, where system identification yields complete dominant modes, natural frequencies and mode shapes are packaged into a vector. Each vector contains three natural frequencies, and 9 mode shape points (3 modes $\times$ 3 mode shape points per mode), summing up to 12 data points. Since temperature was not recorded, this information is excluded in the input. No. of data in Table 4 corresponds to the number of complete data vectors. The distribution of natural frequencies for the no damage and damage cases are shown in Figure 9.

Similar to the analytical investigation, both PCA and AE architectures are considered. For the training of the machine learning model, the baseline condition is used. Out of 50 data, 40 are used for training and 10 for validation. After the data is standardized, the PCA model is trained with six components. The output of the AE model is the three natural frequencies to be reconstructed. A neural network with the dimensions 12-8-8-3 is developed to capture dominant features of the data. Rectified linear units are used as the activation function on all the layers.

Figure 10 presents the novelty index for the experimental data for both architectures with and without the introduction of mode shapes. In general, for all architectures, the damages are distinguished from each other, given the fact that the stiffness degradation was as high as 90%. For all figures, when two columns are damaged (Damage Cases 2 and 4) the index is higher compared to single column damages (Damage Cases 1 and 3). The architectures not relying on mode shapes yield similar indices for Damage Cases 2 and 4, whereas the utilization of mode shapes as input return distinguishable indices. Considering PCA, some instances of the novelty index for the Damage Case 1 at the absence of mode shapes *leak* to No Damage Case region. This behavior is not observed when mode shapes are introduced. One can notice that with the use of mode shapes, the novelty index of Damage Cases 3 and 4 where third columns were damaged is higher for both architectures. The damage at the third floor changes the mode shapes to the point that the novelty index is amplified even though the induced damage is not larger than the first floor. From this observation, it can be concluded that while the proposed approach is a successful damage detection tool, the results may not be definite regarding the magnitude of the damage. To understand the results quantitatively, the modified Euclidean distances for each damage case computed according to Equation (27) are provided in Table 5. The validation data set from the No Damage Case is used as the reference. Since the validation data is limited, the data is repeated to match the size of the target. Figure 10 implies that when mode shapes are introduced novelty indices may decrease slightly. In parallel, the baseline for No Damage Case approaches to zero resembling a flat line. Although the novelty indices reduce, the modified Euclidian distance between the baseline and the damage cases increases which implies that damage is quantitatively more distinguishable. Overall, the distances indicate that the presence of mode shapes in the data improves the reliability of the damage detection for the given test structure. The experimental investigation of the data shows that when mode shapes are used in the machine learning models, the damage detection can be more reliable.

3.4.2 | Structure 2

This section utilizes a linear three-story, three-dimensional frame located at Harbin Institute of Technology (HIT), China (see Figure 11a). The prototype structure has a base plan with dimensions 1.84 m by 2.04 m and each story is 1.2 m tall. The structure is braced in one direction with inverted v-brace (see Figure 11b). A concrete slab weighing approximately 250 kg is attached to each floor. Including the mass of bare structure, total weight sums to 1,066 kg. The columns, beams, and girders are made of structural steel with an elastic modulus estimated to be 220 GPa. More details about the system identification and material properties of the structure are discussed in Ozdagli (2015) and Xi (2014).

To evaluate the performance of the proposed method, temperature gradient in three dimensions over the structure is modeled. To simulate the temperature gradient, first an FE model is established using OpenSees faithful to the experimental structure in terms of boundary conditions and material properties (Xi, 2014). Each member of the model

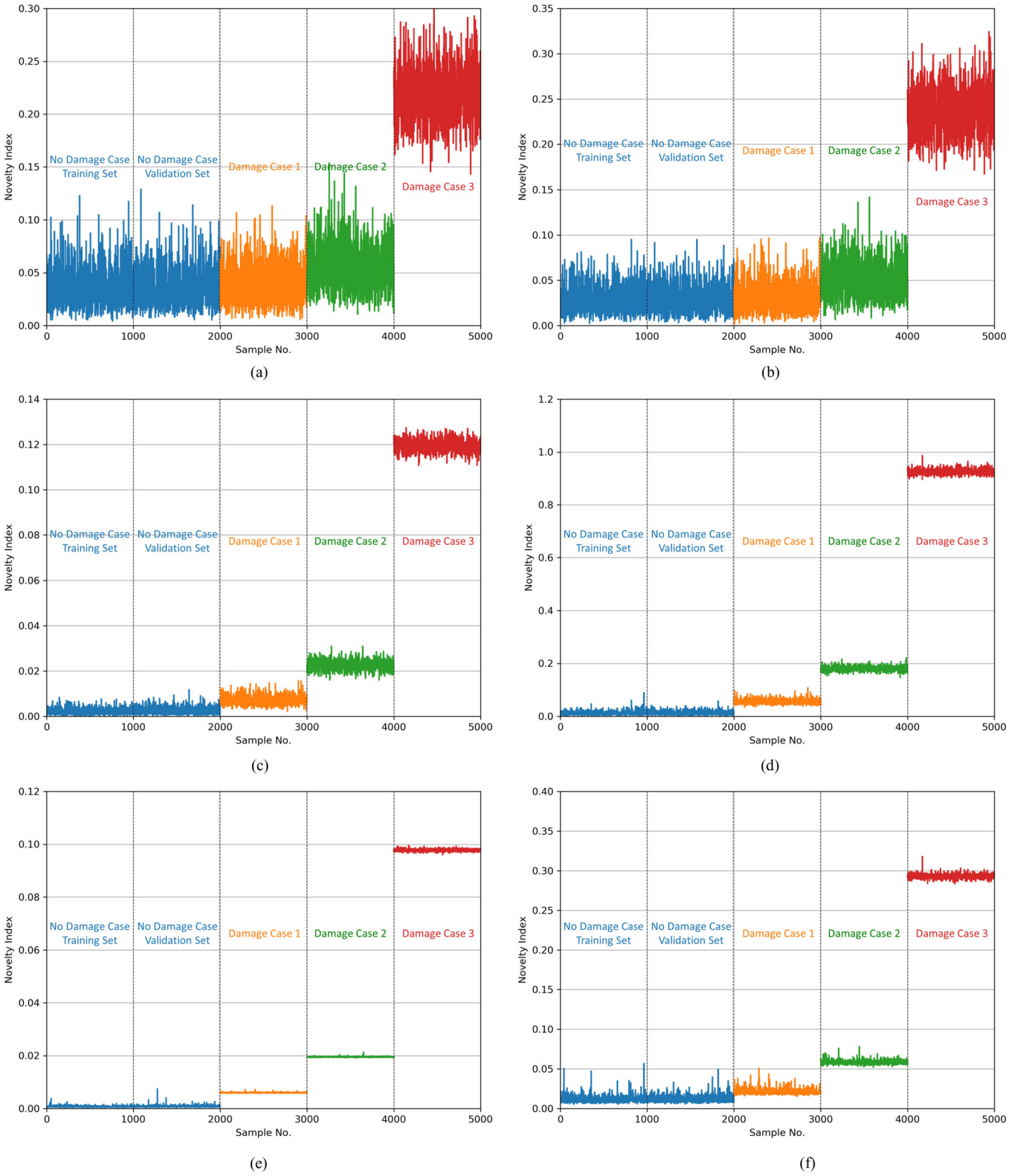

FIGURE 7 Comparison of novelty indices for analytical data under temperature gradient: (a) PCA, model set B (mode shapes not included); (b) AE, model set B (mode shapes not included); (c) PCA, model set A (mode shapes included); (d) AE, model set A (mode shapes included); (e) PCA, model set C (only mode shapes included); (f) AE, model set C (only mode shapes included)

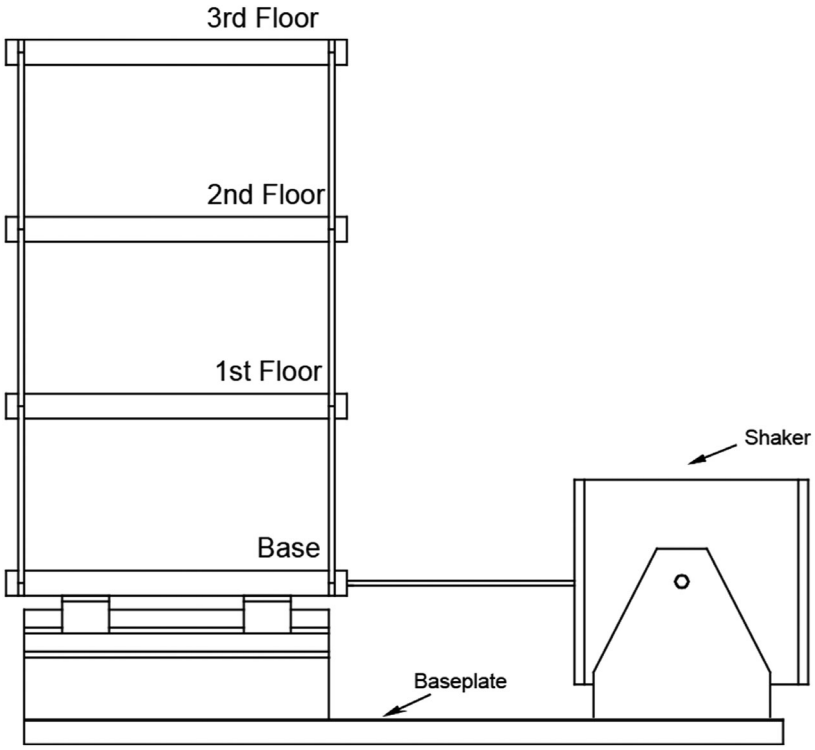

FIGURE 8 Three-story laboratory structure (Structure 1) (Figure modified according to our research from Figueiredo et al., 2009)

TABLE 4 Experimental data matrix (Structure 1)

State condition	Description	No. of data
No Damage Case	Baseline condition	50
Damage Case 1	87.5% stiffness reduction in one first-floor column	49
Damage Case 2	87.5% stiffness reduction in two first-floor columns	50
Damage Case 3	87.5% stiffness reduction in one third-floor column	50
Damage Case 4	87.5% stiffness reduction in two third-floor columns	45

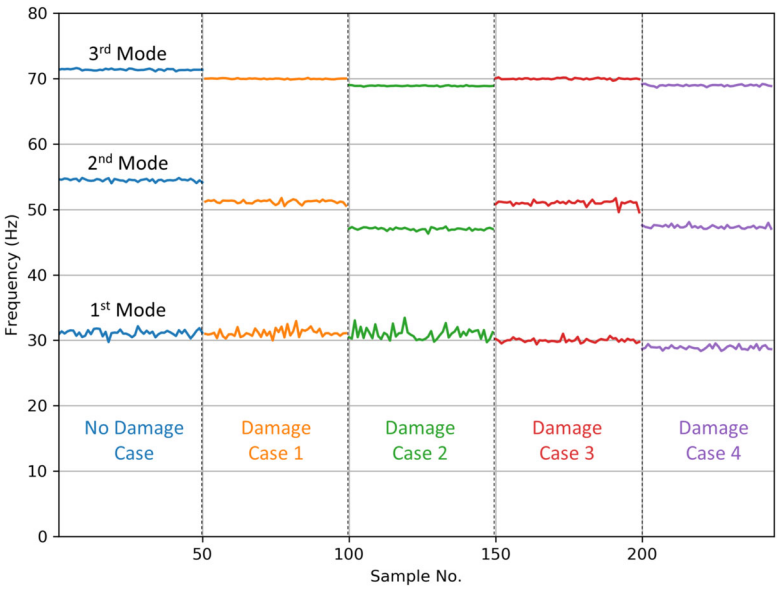

FIGURE 9 Distribution of natural frequencies with varying ambient temperatures for each damage case in experimental data

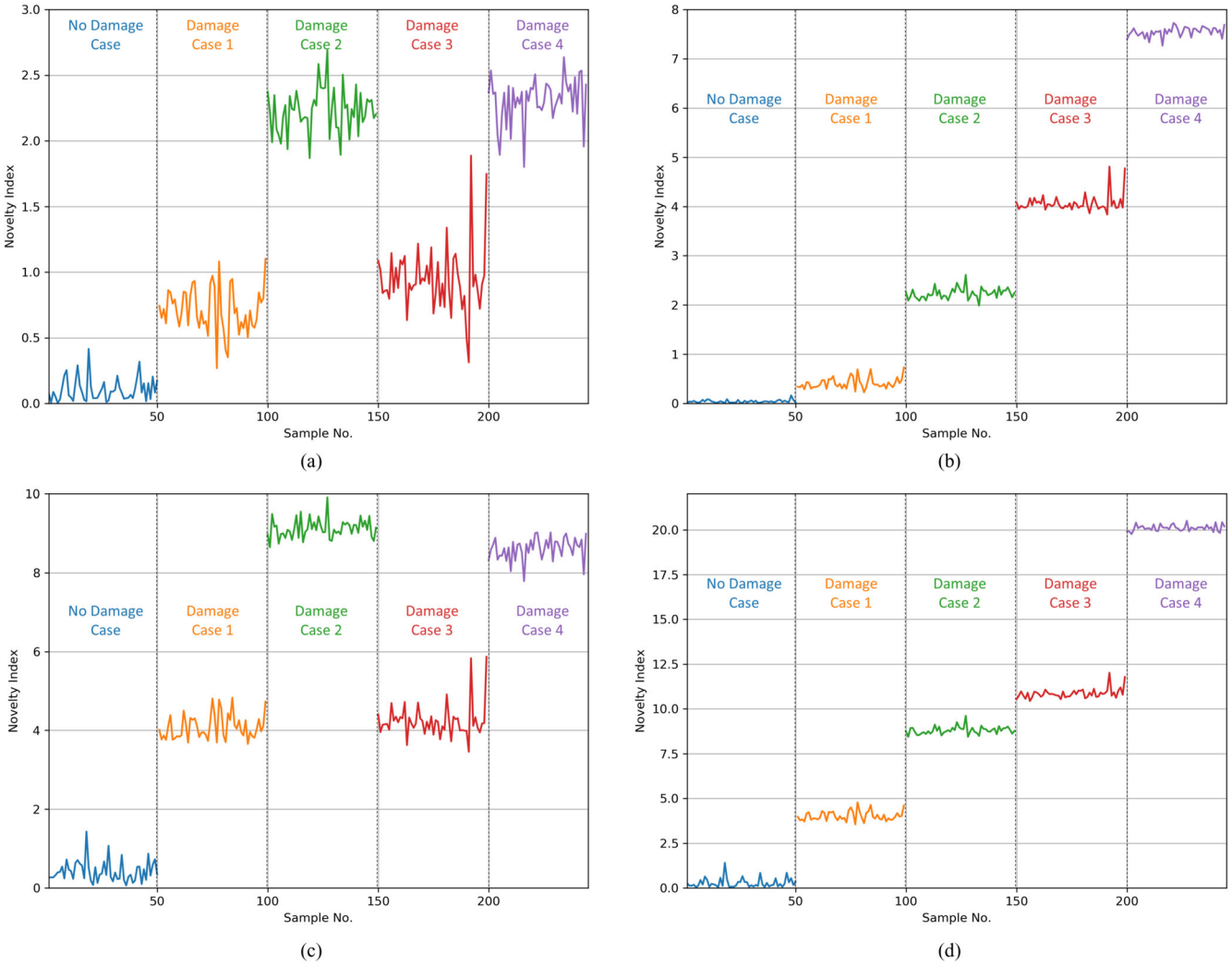

FIGURE 10 Comparison of novelty indices for experimental data: (a) PCA, model set B (mode shapes not included); (b) PCA, model set A (mode shapes included); (c) AE, model set B (mode shapes not included); (d) AE, model set A (mode shapes included)

TABLE 5 Modified Euclidean distances for damage cases with and without the inclusion of mode shapes for experimental data (Structure 1)

	PCA without mode shape	PCA with mode shape	AE without mode shape	AE with mode shape
Damage Case 1	29.79	44.92	53.65	76.87
Damage Case 2	104.48	266.75	129.81	177.53
Damage Case 3	42.28	491.04	56.95	220.72
Damage Case 4	105.50	816.04	112.46	395.06

is discretized into 10 elements, resulting in 360 elements. Equation (28) is used to constitute the relationship between the ambient temperature and the material. The model is calibrated to match the experimental modal properties at 15°C. The natural frequencies and mode shapes of the FE model are presented in Figure 12.

Four different damage conditions are established for this model with damage state varying between 5 and 10%. The damage scenarios are summarized in Table 6 and illustrated in Figure 13. All damage cases consider a reduction of stiffness only at the midspan element to localize the damage.

As for data generation for the validation of the proposed method; it is assumed that temperature gradually increases from bottom node to the top node in the direction of the arrow shown in Figure 13. Here, temperature differences at each node relative to the bottom corner of the structure are shown. The maximum temperature difference between upper and bottom part of the structure is 12.5°C. Since it may be challenging to obtain higher-order modes in reality, only the first four modes are pursued. It is assumed

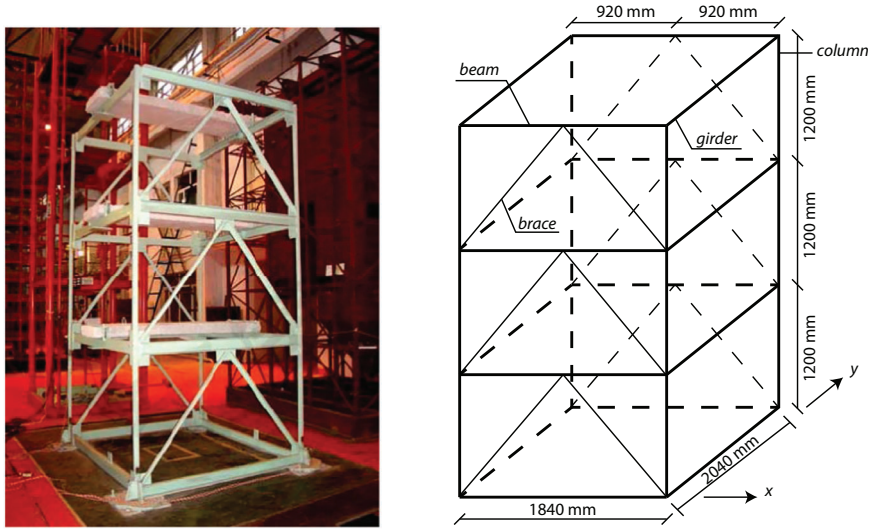

FIGURE 11 Three-story, three-dimensional structure (Structure 2): (a) Experimental prototype; (b) idealization

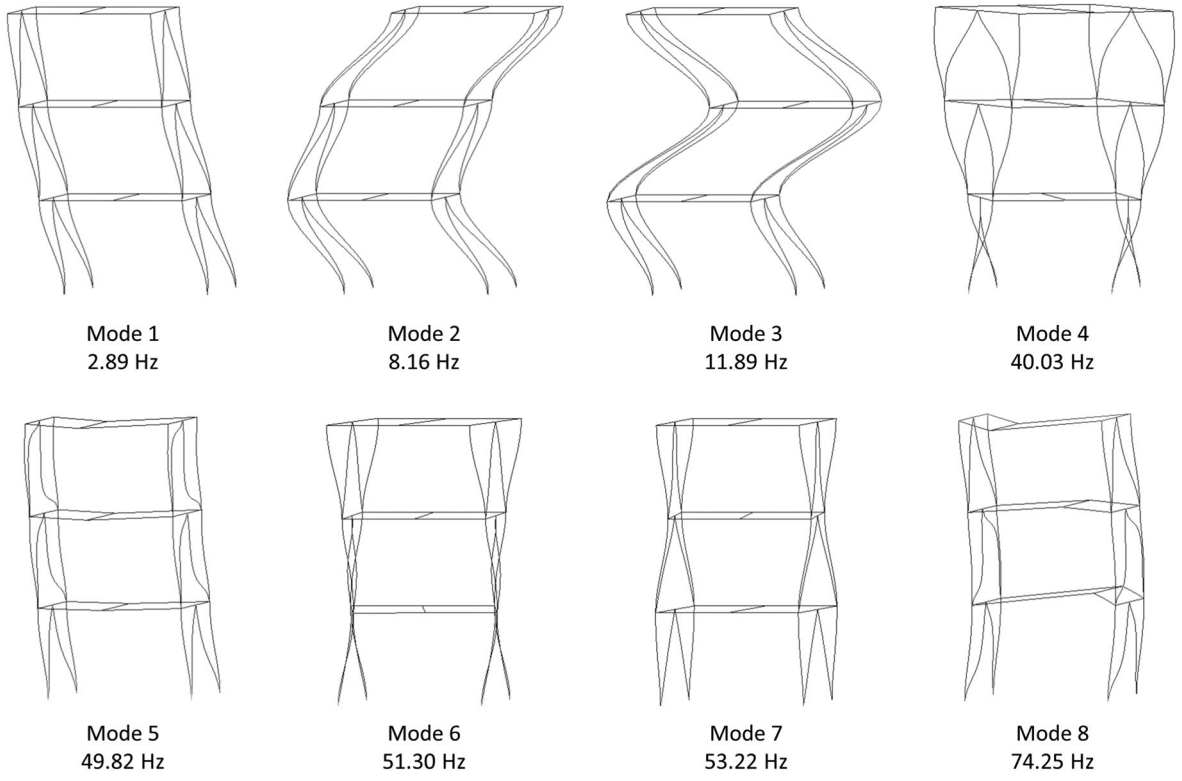

FIGURE 12 Modal properties of the FE model

TABLE 6 Experimental data matrix (Structure 2)

State condition	Description	No. of data
No Damage Case	Baseline condition	2,000
Damage Case 0	5% stiffness reduction at midspan of first floor column	1,000
Damage Case 1	10% stiffness reduction at midspan of second floor beam	1,000
Damage Case 2	10% stiffness reduction at midspan of third floor brace	1,000

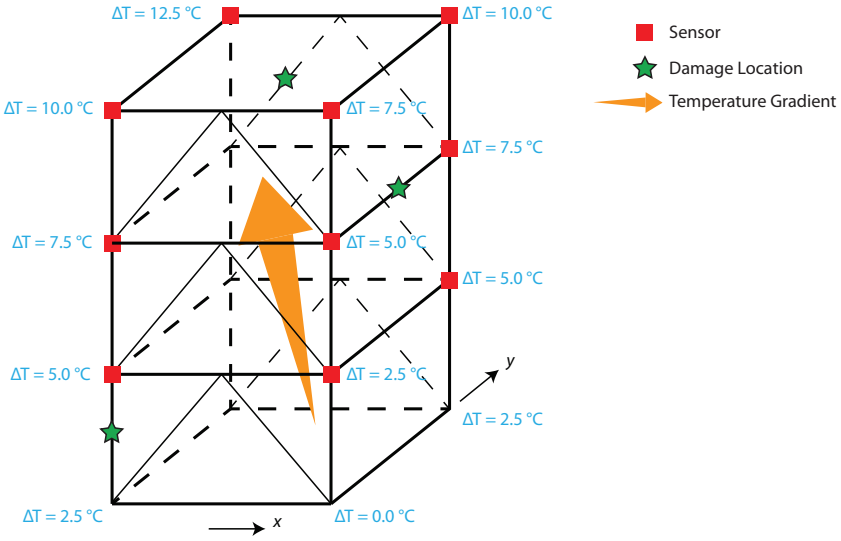

FIGURE 13 Damage conditions and temperature distribution

TABLE 7 Model set properties for experimental data (Structure 2)

	Model set A	Model set B	Model set C
Input	1 temperature data 4 natural frequencies 96 mode shape data points	1 temperature data 4 natural frequencies	1 temperature data 96 mode shape data points
Output	4 natural frequencies	4 natural frequencies	96 mode shape data points
PCA component size	100	3	100
AE network structure	101-50-50-4	5-3-3-4	97-50-50-96

TABLE 8 Modified Euclidean distances for damage cases with and without the inclusion of mode shapes for experimental data (Structure 2)

	PCA model set A	PCA model set B	PCA model set C	AE model set A	AE model set B	AE model set C
No Damage Case	27.09	27.48	11.69	0.01	0.01	20.30
Damage Case 1	298,919.03	7,337.87	124,477.62	73.25	30.62	1,152.85
Damage Case 2	219,040.14	12,039.37	20,180.22	44.31	32.80	237.35
Damage Case 3	33,042.80	2,251.73	2,352.94	13.88	5.65	122.51

that accelerometers at each joint (labeled in red color in Figure 13; the sensors in the background are not labeled) can capture motion along x- and y-axis which will result to about 24 mode shape points per mode. The method is assumed to have access only to the median temperature. Including the median temperature, four natural frequencies, and 96 mode shape points (4 modes × 24 mode shape points per mode), each input contains 101 data points. For each damage scenario, 1,000 inputs are generated whereas for the undamaged case, 2,000 inputs (1,000 input for training + 1,000 input for validation). For all input, the median temperature is uniformly distributed between −15°C and 50°C. Properties of the learning components are summarized in Table 7.

Figure 14 illustrates the novelty indices for all damage cases and model sets. Table 8 summarizes the modified Euclidean distance. Accordingly, using both frequencies and mode shapes improves the detection. The performance is more significant for PCA. For this case, reconstructing more features does not improve the detection of PCA when model sets A and C are considered. As for AE, combination of frequencies and mode shapes improves the detection for all cases. Here reconstructing more features (model set C) indeed improve the detection compared to model set A. However, it should be again noted that model set C requires more parameters to train the AE network and predict the detection; thus, it is computationally more expensive.

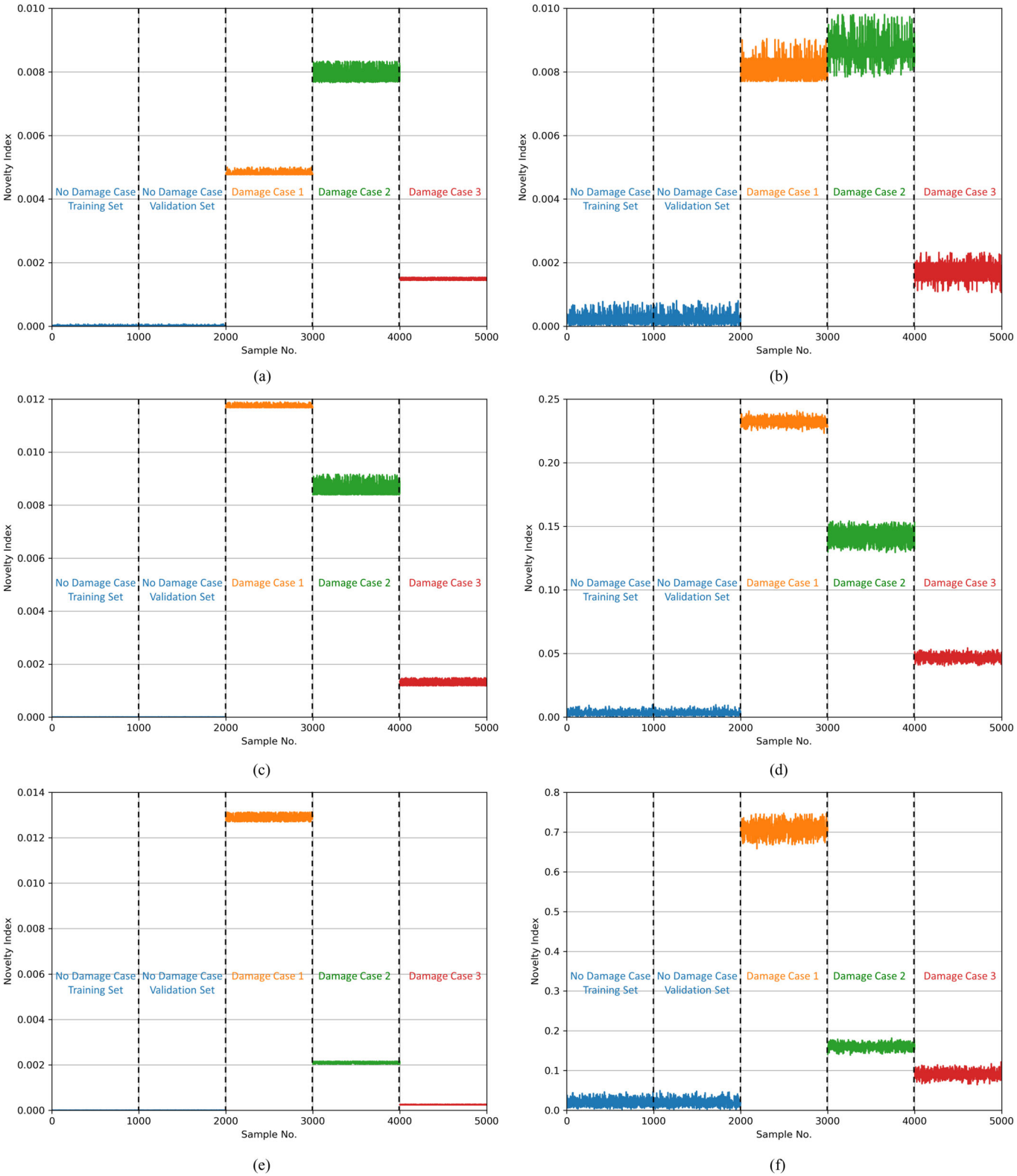

FIGURE 14 Comparison of novelty indices for experimental data (Structure 2): (a) PCA, model set B (mode shapes not included); (b) AE, model set B (mode shapes not included); (c) PCA, model set A (mode shapes included); (d) AE, model set A (mode shapes included); (e) PCA, model set C (only mode shapes included); (f) AE, model set C (only mode shapes included)

4 | CONCLUSIONS

This paper proposes a new machine learning architecture to detect damage in structures reliably by incorporating modal properties such as natural frequencies and mode shapes into the training data for the learning components. While the use of natural frequencies in machine learning algorithms is studied thoroughly in the past literature, it has been shown in this study that mode shapes are independent of temperature variations and remain the same when the structure is not damaged but material proprieties change due to temperature. As a result of this, the learning algorithm considers the persistence of mode shapes as a statistically important feature. To evaluate and validate the proposed approach, this study uses data sets from an FE model of a simply supported beam and experimental testing of one small scale and one large scale three-story structure. Both the analytical and experimental investigation presented herein demonstrate that the introduction of mode shapes improves the detection quality significantly in the presence of environmental variability. Especially for detecting a small amount of damages, the performance of the proposed approach is better compared to the architecture which does not utilize mode shapes. Overall, the findings indicate that the proposed approach has the potential to be used as a viable tool in the field. Regarding the practicality of the method discussed herein, there are some considerations to be given. First, NExT/ERA-based system identification requires long-term data measurements to capture higher modes successfully. Such data may not be always available. However, the pipeline is flexible enough to allow the use of alternative subspace system identification (SSI) methods such as one proposed by Peeters and Roeck (1999) and known to perform well under noisy environments. Second, the success of the method relies on the accuracy of the historical data. A significant change in the system that cannot be identified as damage, such as adding a damper or adding mass, will indeed alter the features that were latent in the historical data. In this case, a new model should be trained. Third, temperature gradients along the structure are common during the field measurements. The training data should consider a wide range of measurements that capture the gradient pattern such that the changes in the mode shapes will not cause novelty indices to increase. While our paper has shown that the proposed method is still effective under temperature gradients, a broader investigation should be performed to ensure the readability of our method. Last, deep networks inherently require big data for the optimization of hyper-parameters. Extracting features from big data with a huge number of features using PCA can be computationally expensive since the data set is processed as a whole. Either, incremental PCA (Ross, Lim, Lin, & Yang, 2008), or AE should be used for a system with large sensor arrays. Ideally, an ensemble of PCA and AE architecture should be considered to maximize the detection sensitivity if computational resources are allowing. Future work should include the investigation of the proposed approach in the presence of multiple environmental and operational variability in the laboratory and in the field. Additionally, the data set should be diversified to include not only responses to ambient noise but also forced vibrations due to service loads. Finally, the proposed architecture should be extended to locate damage by relating mode shapes to the spatial data of the bridge under environmental uncertainty.

ACKNOWLEDGMENTS

The financial support of this research was provided by the National Science Foundation (CNS-1238959). The conclusions of this research represent solely that of the authors.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. Retrieved from https://www.tensorflow.org/

Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., & Inman, D. J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *Journal of Sound and Vibration*, *388*, 154–170.

Amezquita-Sanchez, J. P. & Adeli, H. (2015). A new music-empirical wavelet transform methodology for time–frequency analysis of noisy nonlinear and non-stationary signals. *Digital Signal Processing*, *45*, 55–68.

Amezquita-Sanchez, J. P., Park, H. S., & Adeli, H. (2017). A novel methodology for modal parameters identification of large smart structures using music, empirical wavelet transform, and Hilbert transform. *Engineering Structures*, *147*, 148–159.

Brownjohn, J. M. W. (2003). Ambient vibration studies for system identification of tall buildings. *Earthquake Engineering & Structural Dynamics*, *32*(1), 71–95.

Caicedo, J. M. (2011). Practical guidelines for the natural excitation technique (NExT) and the eigensystem realization algorithm (ERA) for modal identification using ambient vibration. *Experimental Techniques*, *35*(4), 52–58.

Caicedo, J. M., Dyke, S. J., & Johnson, E. A. (2004). Natural excitation technique and eigensystem realization algorithm for phase I of the IASC–ASCE benchmark problem: Simulated data. *Journal of Engineering Mechanics*, *130*(1), 49–60.

Caicedo, J. M., Marulanda, J., Thomson, P., & Dyke, S. J. (2001). Monitoring of bridges to detect changes in structural health. In *Proceedings of the 2001 American Control Conference* (pp. 453–458). Arlington, VA: IEEE.

Chalapathy, R., & Chawla, S. (2019). *Deep learning for anomaly detection: A survey*. arXiv:1901.03407.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 15.

Chollet, F. (2015). *Keras*. Retrieved from https://keras.io

Deraemaeker, A., & Worden, K. (2018). A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mechanical Systems and Signal Processing*, *105*, 1–15.

Dervilis, N., Barthorpe, R. J., Antoniadou, I., Staszewski, W. J., & Worden, K. (2012). Damage detection in carbon composite material

typical of wind turbine blades using auto-associative neural networks. In *Proceedings SPIE 8348, Health Monitoring of Structural and Biological Systems* (834806). San Diego, CA: International Society for Optics and Photonics. https://doi.org/10.1117/12.914710

Dervilis, N., Choi, M., Antoniadou, I., Farinholt, K. M., Taylor, S. G., Barthorpe, R. J., … Farrar, C. R. (2012). Novelty detection applied to vibration data from a CX-100 wind turbine blade under fatigue loading. *Journal of Physics: Conference Series*, *382*(1), 012047.

Doebling, S. W., Farrar, C. R., Prime, M. B., & Shevitz, D. W. (1996). *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review* (Report No. LA-13070). Los Alamos, NM: Los Alamos National Laboratory.

Farrar, C. R., Doebling, S. W., Cornwell, P. J., & Straser, E. G. (1996). *Variability of modal parameters measured on the Alamosa Canyon Bridge* (Technical report). Los Alamos, NM: Los Alamos National Laboratory.

Farrar, C. R., & Worden, K. (2007). An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1851), 303–315.

Farrar, C. R., & Worden, K. (2012). *Structural health monitoring: a machine learning perspective*. New York: John Wiley & Sons.

Figueiredo, E., Park, G., Figueiras, J., Farrar, C., & Worden, K. (2009). *Structural health monitoring algorithm comparisons using standard data sets* (Report No. LA-14393). Los Alamos, NM: Los Alamos National Laboratory.

Fukunaga, K., & Koontz, W. L. G. (1970). Application of the Karhunen–Loeve expansion to feature selection and ordering. *IEEE Transactions on Computers*, *C-19*(4), 311–318.

Gonzalez, I. (2014). *Application of monitoring to dynamic characterization and damage detection in bridges* (PhD thesis), KTH Royal Institute of Technology, Stockholm, Sweden, Structural Engineering and Bridges (QC 20140910).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Retrieved from http://www.deeplearningbook.org

Gu, J., Gul, M., & Wu, X. (2017). Damage detection under varying temperature using artificial neural networks. *Structural Control and Health Monitoring*, *24*(11), e1998.

Gulgec, N. S., Takáč, M., & Pakzad, S. N. (2017). Structural damage detection using convolutional neural networks. In R. Barthorpe, R. Platz, I. Lopez, B. Moaveni, C. Papadimitriou (Eds.), *Model validation and uncertainty quantification, Volume 3* (pp. 331–337). New York: Springer.

Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, *14*(1), 55–63.

James, G. H., Carne, T. G., & Lauffer, J. P. (1993). *The natural excitation technique (NExT) for modal parameter extraction from operating wind turbines* (Report No. SAND92-1666). Albuquerque, NM: Sandia National Laboratories.

James, G. H., Carne, T. G., & Lauffer, J. P. (1995). The natural excitation technique (NExT) for modal parameter extraction from operating structures. *Modal Analysis*, *10*(4), 260.

Jiang, X., & Adeli, H. (2007). Pseudospectra, music, and dynamic wavelet neural network for damage detection of highrise buildings. *International Journal for Numerical Methods in Engineering*, *71*(5), 606–629.

Juang, J., & Pappa, R. S. (1985). An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of Guidance, Control, and Dynamics*, *8*(5), 620–627.

Kim, S. J., Christenson, R., Phillips, B., & Spencer, B. F., Jr. (2012). Geographically distributed real-time hybrid simulation of MR dampers for seismic hazard mitigation. In *Proceedings of the 20th Analysis and Computation Specialty Conference* (pp. 29–31), Chicago, IL.

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, *37*(2), 233–243.

Lee, J. J., Lee, J. W., Yi, J. H., Yun, C. B., & Jung, H. Y. (2005). Neural networks-based damage detection for bridges considering errors in baseline finite element models. *Journal of Sound and Vibration*, *280*(3–5), 555–578.

Li, J. (2014). *Structural health monitoring of an in-service highway bridge with uncertainties* (PhD thesis), Mansfield, CT: University of Connecticut.

Li, X., Ozdagli, A. I., Dyke, S. J., Lu, X., & Christenson, R. (2017). Development and verification of distributed real-time hybrid simulation methods. *Journal of Computing in Civil Engineering*, *31*(4), 04017014.

Liang, Y., Wu, D., Liu, G., Li, Y., Gao, C., Ma, Z. J., & Wu, W. (2016). Big data-enabled multiscale serviceability analysis for aging bridges. *Digital Communications and Networks*, *2*(3), 97–107.

Lin, Z., Pan, H., Wang, X., & Li, M. (2018). Data-driven structural diagnosis and conditional assessment: From shallow to deep learning. In *Proceedings SPIE 10598, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018* (1059814). Denver, CO: International Society for Optics and Photonics. https://doi.org/10.1117/12.2296964

Liu, C., & DeWolf, J. T. (2007). Effect of temperature on modal variability of a curved concrete bridge under ambient loads. *Journal of Structural Engineering*, *133*(12), 1742–1751.

Matarazzo, T. J., Shahidi, S. G., Chang, M., & Pakzad, S. N. (2015). Are today's SHM procedures suitable for tomorrow's bigdata? In C. Niezrecki (Ed.), *Structural health monitoring and damage detection, Volume 7* (pp. 59–65). New York: Springer.

MATLAB (2018). *Version 9.5.0 (R2018b)*. Natick, MA: The MathWorks.

McKenna, F., Scott, M. H., & Fenves, G. L. (2010). Nonlinear finite-element analysis software architecture using object composition. *Journal of Computing in Civil Engineering*, *24*(1), 95–107.

Mehrjoo, M., Khaji, N., Moharrami, H., & Bahreininejad, A. (2008). Damage detection of truss bridge joints using artificial neural networks. *Expert Systems with Applications*, *35*(3), 1122–1131.

Ni, Y. Q., Hua, X. G., Fan, K. Q., & Ko, J. M. (2005). Correlating modal properties with temperature using long-term monitoring data and support vector machine technique. *Engineering Structures*, *27*(12), 1762–1773.

Ozdagli, A. I. (2015). *Distributed real-time hybrid simulation: Modeling, development and experimental validation* (PhD thesis). West Lafayette, IN: Purdue University. https://docs.lib.purdue.edu/open_access_dissertations/531

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peeters, B., & Roeck, G. D. (1999). Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing*, *13*(6), 855–878.

Ross, D. A., Lim, J., Lin, R.-S., & Yang, M.-H. (2008). Incremental learning for robust visual tracking. *International Journal of Computer Vision*, *77*(1–3), 125–141.

Rossum, G. (1995). *Python tutorial* (CS-R9526). Amsterdam, The Netherlands: Centrum voor Wiskunde en Informatica.

Rytter, A. (1993). *Vibrational based inspection of civil engineering structures* (PhD thesis). Aalborg, Denmark: Aalborg University.

Shwartz-Ziv, R. & Tishby, N. (2017). *Opening the black box of deep neural networks via information*. arXiv:1703.00810.

Sohn, H. (2007). Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1851), 539–560.

Sohn, H., Dzwonczyk, M., Straser, E. G., Kiremidjian, A. S., Law, K. H., & Meng, T. (1999). An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake Engineering & Structural Dynamics*, *28*(8), 879–897.

Sohn, H., Worden, K., & Farrar, C. R. (2001). Novelty detection using auto-associative neural network. In *2001 ASME International Mechanical Engineering Congress and Exposition* (pp. 573–580). New York: American Society of Mechanical Engineers.

Tibaduiza, D. A., Mujica, L. E., & Rodellar, J. (2012). Damage classification in structural health monitoring using principal component analysis and self-organizing maps. *Structural Control and Health Monitoring*, *20*(10), 1303–1316.

Wang, T., Bhuiyan, M. Z. A., Wang, G., Rahman, M. A., Wu, J., & Cao, J. (2018). Big data reduction for a smart city's critical infrastructural health monitoring. *IEEE Communications Magazine*, *56*(3), 128–133.

Wirsching, P. H., Paez, T. L., & Ortiz, K. (2006). *Random vibrations: Theory and practice*. North Chelmsford, MA: Courier Corporation.

Woon, C. E., & Mitchell, L. D. (1996). Variations in structural dynamic characteristics caused by changes in ambient temperature: I. Experimental. In *Proceedings of the International Society for Optical Engineering* (pp. 963–971). San Diego, CA: SPIE International Society for Optical Engineering.

Worden, K. (1997a). Damage detection using a novelty measure. In *Proceedings of the International Society for Optical Engineering* (pp. 631–637). San Diego, CA: SPIE International Society for Optical Engineering.

Worden, K. (1997b). Structural fault detection using a novelty measure. *Journal of Sound and Vibration*, *201*(1), 85–101.

Worden, K., & Manson, G. (2006). The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1851), 515–537.

Worden, K., Manson, G., & Allman, D. (2003). Experimental validation of a structural health monitoring methodology: Part I. Novelty detection on a laboratory structure. *Journal of Sound and Vibration*, *259*(2), 323–343.

Xi, W. (2014). *Performance based implementation of seismic protective devices for structures* (PhD thesis). Los Angeles, CA: University of California. Retrieved from https://escholarship.org/uc/item/7m37c94p

Xia, Y., Chen, B., Weng, S., Ni, Y., & Xu, Y. (2012). Temperature effect on vibration properties of civil structures: A literature review and case studies. *Journal of Civil Structural Health Monitoring*, *2*(1), 29–46.

Xu, Y. L., Chen, B., Ng, C. L., Wong, K. Y., & Chan, W. Y. (2009). Monitoring temperature effect on a long suspension bridge. *Structural Control and Health Monitoring*, *17*(6), 632–653.

Yu, L., Zhu, J., & Cheri, L. (2010). Parametric study on PCA-based algorithm for structural health monitoring. In *Prognostics and System Health Management Conference* (pp. 1–6). Macau, China: IEEE.

Yu, Y., Wang, C., Gu, X., & Li, J. (2018). A novel deep learning-based method for damage identification of smart building structures. *Structural Health Monitoring*, *18*(1), 143–163.

Zang, C., & Imregun, M. (2001). Structural damage detection using artificial neural networks and measured FRF data reduced via principal component projection. *Journal of Sound and Vibration*, *242*(5), 813–827.

Zhou, H. F., Ni, Y. Q., & Ko, J. M. (2011). Eliminating temperature effect in vibration-based structural damage detection. *Journal of Engineering Mechanics*, *137*(12), 785–796.

How to cite this article: Ozdagli AI, Koutsoukos X. Machine learning based novelty detection using modal analysis. *Comput Aided Civ Inf*. 2019;34:1119–1140. https://doi.org/10.1111/mice.12511