

Non-parametric empirical machine learning for short-term and long-term structural health monitoring

Structural Health Monitoring
2022, Vol. 0(0) 1–19
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14759217211069842
journals.sagepub.com/home/shm


Alireza Entezami^{1,2} , Hashem Shariatmadar²  and Carlo De Michele¹

Abstract

Early damage detection is an initial step of structural health monitoring. Thanks to recent advances in sensing technology, the application of data-driven methods based on the concept of machine learning has significantly increased among civil engineers and researchers. On this basis, this article proposes a novel non-parametric anomaly detection method in an unsupervised learning manner via the theory of empirical machine learning. The main objective of this method is to define a new damage index by using some empirical measure and the concept of minimum distance value. For this reason, an empirical local density is initially computed for each feature and then multiplied by the minimum distance of that feature to derive a new damage index for decision-making. The minimum distance is obtained by calculating the distances between each feature and training samples and finding the minimum quantity. The major contributions of this research contain developing a novel non-parametric algorithm for decision-making under high-dimensional and low-dimensional features and proposing a new damage index. To detect early damage, a threshold boundary is computed by using the extreme value theory, generalized Pareto distribution, and peak-over-threshold approach. Dynamic and statistical features of two full-scale bridges are used to verify the effectiveness and reliability of the proposed non-parametric anomaly detection. In order to further demonstrate its accuracy and proper performance, it is compared with some classical and recently published anomaly detection techniques. Results show that the proposed non-parametric method can effectively discriminate a damaged state from its undamaged condition with high damage detectability and inconsiderable false positive and false negative errors. This method also outperforms the anomaly detection techniques considered in the comparative studies.

Keywords

Structural health monitoring, non-parametric anomaly detection, empirical machine learning, environmental variability, bridges

Introduction

Health and safety of vital civil structures such as bridges, high-rise buildings, and dams are of paramount importance to every society and government due to the high influences of such structural systems on social life and economics. Most of these structures were designed and constructed several years or decades ago with classical design codes and technologies. Furthermore, aging, material deterioration, corrosion, settlement, and fatigue are inevitable in their lifetime. These structures may also experience natural hazards such as strong or weak seismic ground motions, strong wind, flood, and hurricane that threaten their safety and serviceability. To avoid any catastrophic event caused by the occurrence of damage, structural health monitoring (SHM) presents a new technology for assessing the safety and integrity of civil structures in a short-term or long-term

(continuous) fashion.^{1–4} For this practical process, one attempts to apply ambient or artificial sources for exciting a civil structure and acquire its static or dynamic responses via various sensors from classical wireless sensor networks of accelerometers to images or videos from professional cameras.^{5,6} Hence, the main objective of SHM can be summarized in four levels including early damage

¹Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy

²Department of Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

Corresponding author:

Alireza Entezami, Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milano, MI 20133, Italy.
Email: alireza.entezami@polimi.it

detection, damage localization, damage quantification, and damage prognosis.⁷

Using field monitoring data on real structures and thanks to recent development of sensing and data acquisition systems, civil engineers and researchers have paid more attention to data-driven SHM methods. In general, these methods rely on applying raw measured data under the concepts of statistical pattern recognition and machine learning algorithms.^{7,8} In contrast to model-driven techniques,⁹ the great merit of data-driven methods is to facilitate SHM of full-scale structures. Furthermore, such approaches are more suitable for the first level of SHM; that is, early damage detection.¹⁰ Feature extraction and feature classification are two main parts of a data-driven method.⁴¹ The former aims at discovering meaningful information (damage-sensitive features) from raw measured data, while the latter refers to the process of utilizing extracted features and making an accurate decision on the problem under study via the concept of machine learning.

Machine learning is a branch of artificial intelligence that aims to automatically recognize or classify information (i.e., the extracted damage-sensitive features) based on a learned pattern or statistical/computational model through training data. Depending on the type of training data and the availability of damaged information, a machine learning algorithm can be categorized as supervised, semi-supervised, and unsupervised learning classes.^{7,8,11–13} For complex and full-scale civil engineering structures, it may not be practical to impose intentional damage patterns in an effort to prepare full-labeled training data including undamaged and damaged information. Therefore, it seems that unsupervised learning is more suitable than the other machine learning algorithms for SHM of civil structures, particularly for early damage detection. In this algorithm, the main problem is to discover the relation in the structure of data and implement the process of decision-making under the concept of anomaly detection or novelty detection.¹⁴

Depending upon the type of statistical/computational model, unsupervised learning can be divided into *non-parametric*, *semi-parametric*, and *parametric* methods.¹⁵ The separation of these methods relies on the main characteristics of statistical/computational unsupervised model or novelty detector. If this model or detector does not need any unknown parameters, called *hyperparameters*,¹⁵ it falls into a non-parametric framework that provides the simplest and most effective unsupervised learning tool without estimating prior knowledge. Distance-based anomaly detection methods based on Mahalanobis distance,^{16,17,43} Kullback-Leibler divergence,^{18–20} dynamic time warping,²¹ singular vector decomposition (SVD),^{15,22} and correlation coefficient²³ are some successful examples of non-parametric anomaly detection or unsupervised learning methods. In contrast, if the model or novelty detector depends on some hyperparameters, the anomaly detection

method is parametric. Data clustering methods based on the *k*-means,²⁴ *k*-medoids,²⁵ fuzzy *c*-means,¹⁵ and Gaussian mixture model,²⁶ and artificial neural networks based on an auto-associative neural network²⁷ and auto-encoder²⁸ are some well-known examples of this kind of unsupervised learning algorithm. Finally, a semi-parametric novelty detector is based on a combination of the non-parametric and parametric algorithms in order to enhance the performance of non-parametric detectors.^{15,27,29}

Despite effective unsupervised learning methods for health monitoring of full-scale structures, some challenges are still open problems that need to be dealt with appropriately. In the context of data-driven SHM under the concept of machine learning, a high rate of decision-making errors (i.e., false positive and false negative) is a major challenging issue. This challenge may result from variations in measured data or extracted features due to environmental and operational conditions.^{17,42} This problem makes sense that the structure is undamaged, but the SHM method mistakenly alarms the occurrence of damage or the structure really suffered from damage, but the method of interest fails in accurately alarming damage. The first subject is related to the false positive error that pertains to an economic issue, while the second subject refers to the false negative error that is concerned with a safety issue.¹¹ Hence, it is critical to deal with the problem of environmental and/or operational variability in any real-world SHM application. The other important challenge relates to the general aspect of machine learning. Although it has been described that unsupervised learning is an effective and efficient tool for early damage detection, a robust novelty detector or unsupervised statistical/computational model should have a proper generality and reliable performance under any type and size of data in terms of statistical versus dynamic features and high-dimensional (large-size) vs. low-dimensional (small-size) data. Moreover, high classification accuracy or minimum total error is an important factor for accurate decision-making. Finally, the other challenge pertains to the complexity of a machine learning model, particularly a parametric approach, that may need hyperparameter(s) for an accurate learning process and decision-making.

Accordingly, this article proposes an innovative non-parametric anomaly detection method based on the concept of empirical machine learning.³⁰ In this method, the main aim is to define a new damage index (novelty score) by computing an empirical local density for each feature and finding the minimum distance value of this feature and all training samples. In this case, the damage index of interest is obtained by multiplying the local density by the minimum distance value by getting idea from density peak clustering³¹ and its center cluster selection strategy.³² This process is carried out by using all training and test samples so as to determine their damage indices for decision-making. An effective approach through the extreme value theory (EVT),

generalized Pareto (GP) distribution, and peak-over-threshold (POT) is considered to determine a threshold boundary. The main contributions of this article include developing a novel non-parametric anomaly detection approach under the theory of empirical machine learning and proposing a new damage index. Simplicity, computational efficiency and a non-parametric characteristic are the main methodological advantages of the proposed method. The engineering benefits of this method consist of dealing with the major challenge of the environmental and/or operational variability conditions and making an accurate decision under high-dimensional and low-dimensional features in long-term and short-term monitoring programs. To verify the effectiveness and reliability of the proposed method, dynamic and statistical features of two full-scale bridge structures are incorporated. This method is also compared with the classical non-parametric anomaly detection techniques based on the Mahalanobis distance and a SVD algorithm as well as recently published anomaly detection techniques. The effect of the number of training samples on anomaly detection is also evaluated. Results demonstrate that the method presented here is highly able to detect early damage and properly distinguish a damaged state from an undamaged condition with high classification rates. Moreover, it can be observed that the proposed non-parametric method outperforms the classical and recently published anomaly detection techniques in terms of having higher damage detectability and smaller false positive and false negative errors.

Empirical machine learning

The empirical machine learning is a new branch of data mining that is based entirely on the empirical observations of data samples and the relative proximity of these samples in the data space.³⁰ The term “*empirical*” is mentioned to distinguish the current machine learning method from conventional machine learning techniques that rely strongly on several restrictive prior assumptions, generative probabilistic or non-probabilistic models, a learning process, infinite amounts of observations/data samples, and some important hyperparameters and parameter estimation approaches. In this regard, the major benefit of the empirical machine learning is that statistical models or tools under this theory do not need any prior assumptions and learning processes, which may be unrealistic and restrictive, or hyperparameters and parameter estimation. To put it another way, this theory presents a novel non-parametric approach to engineering problems.

Having considered the only measured data samples or features extracted from them, the central core of the

empirical machine learning concentrates on the concept of *cumulative proximity*.³⁰ Given the data samples in the matrix $\mathbf{X}=[\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$, the cumulative proximity of \mathbf{x}_i can be derived from the other data samples as follows

$$q(\mathbf{x}_i) = \sum_{j=1}^n d(\mathbf{x}_i, \mathbf{x}_j) \quad (1)$$

where $i = 1, 2, \dots, n$ and d refers to a distance measure. Equation (1) is an initial empirical measure that is simply derived from data samples and plays an important role in the other empirical measures. The cumulative proximity is important because it provides centrality information on any particular data without any prior assumptions about the data generation model. An advantage of the cumulative proximity is its generality to any distance value, either univariate (e.g., Euclidean) or multivariate (i.e., Mahalanobis). By considering the well-known Euclidean distance and also correlation among variables in \mathbf{X} , the cumulative proximity based on this statistical metric is rewritten as follows

$$q(\mathbf{x}_i) = n \left(\|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 - \|\bar{\mathbf{x}}\|^2 + \gamma \right) \quad (2)$$

where $\|\cdot\|$ denotes the l_2 - or Euclidean norm and $\bar{\mathbf{x}} \in \mathbb{R}^p$ is the mean vector of the matrix \mathbf{X} . In this equation, γ is a scalar value defined as

$$\gamma = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i^T \Sigma^{-1} \mathbf{x}_i \quad (3)$$

where $\Sigma \in \mathbb{R}^{p \times p}$ is the covariance matrix of \mathbf{X} , which is expressed in the following form

$$\Sigma = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (4)$$

From the cumulative proximity, one can derive the other empirical measure called *local eccentricity*.³⁰ This measure can be defined as a normalized cumulative proximity that represents the ensemble properties of the data samples that are far away from the peak. The local eccentricity of \mathbf{x}_i is expressed as

$$e(\mathbf{x}_i) = \frac{2q(\mathbf{x}_i)}{\sum_{j=1}^n q(\mathbf{x}_j)} = \frac{\|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 - \|\bar{\mathbf{x}}\|^2 + \gamma}{n \left(\gamma - \|\bar{\mathbf{x}}\|^2 \right)} \quad (5)$$

or

$$e(\mathbf{x}_i) = \frac{1}{n} \left(1 + \frac{\|\mathbf{x}_i - \bar{\mathbf{x}}\|^2}{v} \right) \quad (6)$$

where $v = \gamma - \|\bar{\mathbf{x}}\|^2$. Furthermore, in equation (5)

$$\sum_{j=1}^n q(\mathbf{x}_j) = 2n^2 \left(\gamma - \bar{\mathbf{x}}^2 \right) \quad (7)$$

Based on the concept of the local eccentricity, one can derive the other empirical measure called local data density. Generally, data density is simply defined as the number of items (i.e., here data samples) per unit of area. In the context of the empirical machine learning, the local data density is described by the mutual proximity of data samples as the inverse of the local eccentricity. In this regard, the local data density of \mathbf{x}_i is expressed as follows

$$\lambda(\mathbf{x}_i) = \frac{1}{e(\mathbf{x}_i)} = \frac{n}{1 + \|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 / v} \quad (8)$$

Because the local density is derived from the cumulative proximity, it can be an appropriate measure for anomaly detection. This is due to the fact that this measure is equivalent to the inverse of the distance between two points. This means that as these points have a further distance, they are far away from together, in which case one can reach a lower density.

Proposed non-parametric method

The proposed non-parametric method is developed by using the main concepts of unsupervised learning anomaly detection and empirical machine learning. The main goal is to define a damage index by using the local data density and minimum distance value. To detect early damage in short-term or long-term SHM frameworks, the proposed method is implemented in the training and inspection phases. The first phase contains all available data or features concerning the undamaged or normal condition of the structure, while the second phase refers to the unknown structural state that attempts to make an accurate decision on it. To fulfill this objective, once all damage indices of the training samples in the training phase have been determined, a threshold boundary is defined to detect whether the damage index of a test sample belongs to the normal condition or it is related to the damaged state.

Damage index

The choice of an effective damage index significantly affects the performance and classification accuracy of any unsupervised anomaly detection. In other words, in the context of SHM, if the novelty detector has high damage detectability, it means that corresponding damage indices allow one to appropriately discriminate the damaged state from the normal condition. Accordingly, one can reduce the decision-making errors or increase the accuracy of feature classification. Inspired by the concept of density peak clustering,²⁸ this article proposes a new damage index based

on the local data density (λ) from the theory of empirical machine learning and the minimum distance value (δ). The fundamental principle of this damage index originates from the process of center cluster selection in density peak clustering.³² The damage index is then expressed as follows

$$DI = \lambda \delta \quad (9)$$

The minimum distance value indicates the distance of the nearest neighbor of each sample, in which case one can find the most representative feature of that feature sample. With this strategy, it is possible to deal with the negative effect of the environmental and operational variability conditions on decision-making outputs. By performing the calculation of the damage index through the training and test data, one can determine the decision-making outputs for SHM and early damage detection.

Training phase

Assume that the matrix $\mathbf{X}=[\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{p \times n}$ includes n feature vectors of the normal condition, each of which consists of p variables. Based on the descriptions in empirical machine learning Section 2, one can determine a local density value for each of the feature vectors of the training matrix $\lambda_{x_1}, \dots, \lambda_{x_n}$, which are equivalent to $\lambda(\mathbf{x}_1), \dots, \lambda(\mathbf{x}_n)$. To obtain the minimum distance value, it is necessary to compute the distance of each feature vector with the remaining feature vectors without itself. Having considered the i^{th} feature \mathbf{x}_i , its distance with the remaining features $\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n$ is calculated via the Euclidean distance as follows

$$d(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) \quad (10)$$

where $j=1, 2, \dots, n-1$. In this case, one can determine $n-1$ distance values $d_1^{(i)}, \dots, d_{n-1}^{(i)}$, where $d_1^{(i)} = d(\mathbf{x}_i, \mathbf{x}_1)$ and $d_{n-1}^{(i)} = d(\mathbf{x}_i, \mathbf{x}_n)$. Therefore, the minimum distance value for the i^{th} feature vector can simply be selected as

$$\delta_{x_i} = \min(d_1^{(i)}, \dots, d_{n-1}^{(i)}) \quad (11)$$

As such, it can be determined n minimum distance values $\delta_{x_1}, \dots, \delta_{x_n}$ for all n feature vectors of the training matrix. Eventually, the damage indices of the training samples are computed by multiplying their local density quantities by their minimum distance value based on equation (9); that is, $DI_{x_1}, \dots, DI_{x_i}, \dots, DI_{x_n}$, where $DI_{x_i} = \lambda_{x_i} \delta_{x_i}$ and $i = 1, 2, \dots, n$. Using these damage indices, it is possible to estimate a threshold boundary (see *Threshold boundary*). Because the aforementioned damage indices are concerned with the normal condition of the structure, it is expected that all of them fall below the threshold boundary implying the safe condition. Any

deviation of the normal damage index from the threshold is also indicative of a false positive error.¹¹

Inspection phase

During the inspection phase, one supposes that the matrix $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_m] \in \mathbb{R}^{p \times m}$ consists of m test samples, each of which is a p -dimensional vector. The main objective in this stage is to determine a new damage index for the test sample \mathbf{z}_k via the same procedures as the training samples, where $k = 1, 2, \dots, m$. Hence, one needs to calculate its local density in the empirical machine learning framework and then select its minimum distance value among all training samples. For this aim, it is only necessary to replace the feature vector \mathbf{z}_k with \mathbf{x}_i in the equations related to the cumulative proximity or equation (2), local eccentricity or equation (6) and local data density or equation (8). In this regard, one can determine m local density values $\lambda_{z_1}, \dots, \lambda_{z_m}$ for all test samples.

To obtain the minimum distance value of the test sample \mathbf{z}_k , the distances between this feature and the training samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ are computed, which lead to n distance values $\tilde{d}_1^{(k)}, \dots, \tilde{d}_n^{(k)}$, where $\tilde{d}_1^{(k)} = d(\mathbf{z}_k, \mathbf{x}_1)$ and $\tilde{d}_n^{(k)} = d(\mathbf{z}_k, \mathbf{x}_n)$. Similarly, the distance between \mathbf{z}_k and each of the training samples is calculated by the Euclidean distance as same as equation (10). Hence, the minimum distance value regarding the feature vector \mathbf{z}_k can simply be selected as

$$\delta_{z_k} = \min \left(\tilde{d}_1^{(k)}, \dots, \tilde{d}_n^{(k)} \right) \quad (12)$$

Having considered all test samples in \mathbf{Z} , one can then determine m minimum distance values $\delta_{z_1}, \dots, \delta_{z_m}$. Finally, the damage indices of the test samples are calculated by multiplying their local density quantities by their minimum distance values; that is, $DI_{z_1}, \dots, DI_{z_k}, \dots, DI_{z_m}$, where $DI_{z_k} = \lambda_{z_k} \delta_{z_k}$. To recognize the current state of the structure between the undamaged or damaged conditions, one should compare each of the damage index with the threshold boundary. If the damage index exceeds the threshold, this implies that the structure suffered from damage; otherwise, it is still in its normal phase.

Threshold boundary

Determination of a threshold boundary is an important process in an unsupervised anomaly detection framework. In contrast to supervised learning, in which the decision-making can be performed by entirely checking the classification accuracy due to accessibility to full-labeled data⁴⁴, an accurate threshold boundary plays a crucial role in decision-making under unsupervised learning. This procedure is often carried out by considering the probabilistic characteristics of damage indices concerning the training samples.¹⁶ Central limit theory (CLT) and EVT provide two general approaches to determining the threshold boundary.

The former supposes the Gaussianity of damage indices and utilizes standard confidence intervals under various probabilities (i.e., significance levels).^{19,21} The latter relies on the probabilistic properties of extreme values (i.e., maximum or minimum samples) under non-Gaussianity assumption and tail modeling.^{16,25} Unlike the CLT, which requires a large number of data samples, the EVT only focuses on a few maximum or minimum samples that provide the best fit to some extreme value distribution.³³

Despite the applicability of both methods for threshold determination, it was demonstrated by some authors that the EVT-based methods are more reliable for threshold determination. In this regard, Sarmadi and Karamodin¹⁶ proposed an EVT-based technique based on a generalized extreme value (GEV) distribution and block-maxima approach. They compared their method with the classical CLT-based algorithm under a standard confidence interval and Bootstrap confidence interval. They demonstrated that the EVT is significantly reliable than the other techniques. In a comparative study, Sarmadi¹⁵ showed that the EVT highly outperformed the CLT by considering various machine learning methods and two different features from two full-scale structures. Entezami et al.³³ also reached the same conclusion by using time series features from laboratory and full-scale structures and a non-parametric machine learning method. With such evidence, this article considers the EVT for the threshold determination.

In statistics, the EVT is a branch of probability theory that generally intends to analyze extreme values and model them by some distribution functions. In a univariate case, the well-known extreme value distributions come from the Gumbel, Frechet, and Weibull families.³⁴ It is possible to integrate these distributions into single models and produce the GEV and GP distributions. The modeling process in the EVT is often carried out by the block-maxima and POT methods. The former is based on determining an optimal number of blocks, dividing all data samples into the determined blocks, selecting the maximum value of each block, and modeling the maximum samples by the triple extreme value distributions or the GEV model.¹⁶ The latter relies on determining a threshold level, selecting all samples over this level, and modeling such selected samples via the GP distribution.^{10,35} In both techniques, the inverse of the cumulative density function (CDF) of the extreme value distribution under a significance level is representative of a threshold boundary. Although both the BM and POT methods are suitable and reliable, it seems that the POT is more beneficial than the BM due to considering all possible maximum samples.³⁵ Therefore, this article utilizes the POT-based threshold determination.

Assume that Y_1, \dots, Y_s are the maximum samples (i.e., often called *exceedances*) over the threshold level u . On this basis, the GP distribution can model these maximum samples in the following form³⁴

$$H(y) = 1 - \left(1 + \frac{\zeta y}{\sigma}\right)^{-1/\zeta} \quad (13)$$

where ζ and σ are the main parameters of the GP distribution; $y = Y - u > 0$, and $1 + \zeta y / \sigma > 0$; $H(y)$ refers the CDF of the GP model. As such, a threshold boundary can be derived by inverting the CDF under the significance level α as follows

$$\tau = \begin{cases} u + \frac{\sigma}{\zeta} (\alpha^{-\zeta} - 1), & \zeta \neq 0 \\ u - \sigma \log \alpha, & \zeta = 0 \end{cases} \quad (14)$$

Since the determination of a threshold boundary depends on the unknown parameters of the GP distribution, one should estimate them by one of the computational techniques. In statistics, the maximum likelihood estimation is a popular and effective approach to this objective. Accordingly, one needs to define the log-likelihood function of the GP distribution and maximize this function with respect to the unknown parameters (i.e., ζ and σ).³⁴ The other prominent parameter for threshold determination is to obtain the threshold level u . At first, it needs to clarify that this parameter entirely differs from the threshold boundary τ applied to makes a decision. To determine this level, it is

possible to use graphical and numerical approaches. As their names indicate, the former is based on a graphical function (e.g., excess sample mean function) and user expertise in interpreting the mean excess plot.³⁵ The latter relies on numerical evidence such as a statistical hypothesis tests for making sure of accurately modeling the GP distribution.¹⁰ As Sarmadi and Yuen¹⁰ demonstrated, a numerical approach is superior to a graphical tool due to avoiding choosing exceedances with user expertise and some laborious attempts finding the linear part of the plot by a visual observation. Therefore, this article exploits their iterative numerical approach¹⁰ to choose the best threshold level u .

Case studies

A box-girder concrete bridge

To demonstrate the effectiveness and performance of the proposed non-parametric method, the modal-based features of the well-known Z24 bridge are applied. This structure was a classical post-tensioned concrete box-girder bridge located in Canton Bern, Switzerland.³⁶ It consisted of a main span of 30 m and two side-spans of 14 m as shown in Figure 1. In 1998, the Z24 bridge was demolished to build a

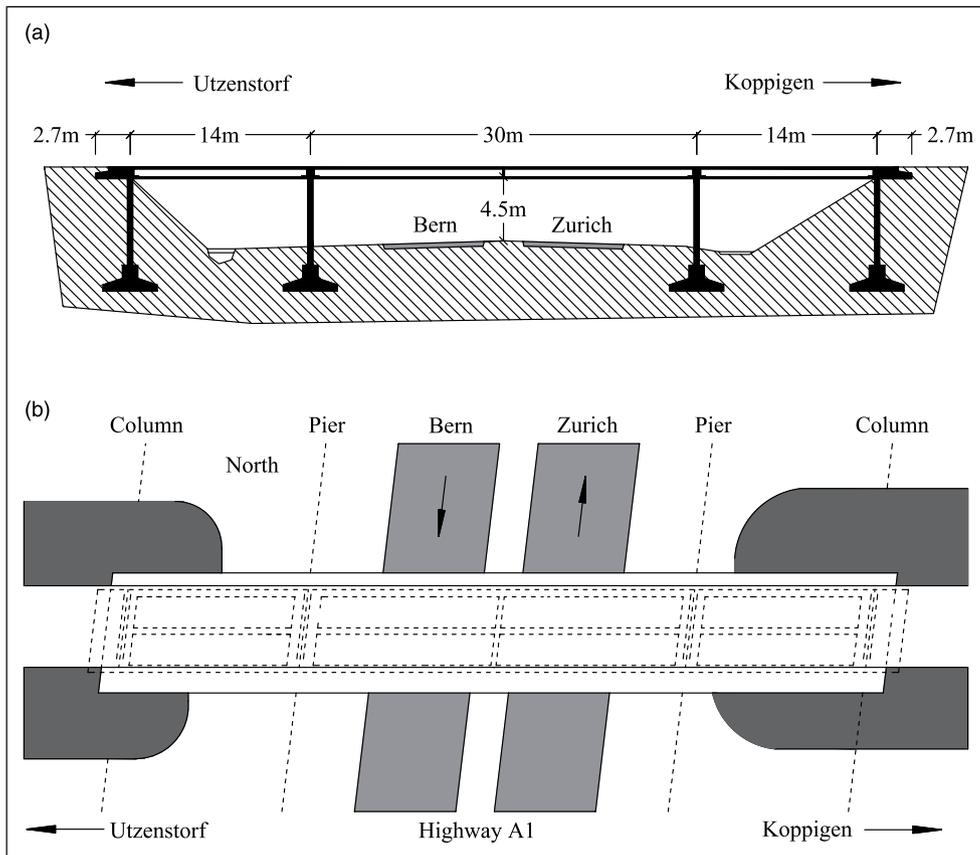


Figure 1. The Z24 bridge: (a) longitudinal section and (b) top view.

new bridge structure with a larger side span. Before this process, a long-term monitoring program was performed to measure environmental and vibration data. In the month before demolition, some realistic damage patterns in a controlled way were considered to define the damaged state of this bridge. Using an automated operational modal identification, natural frequencies in four modes were extracted from acceleration time histories under actual environmental variability conditions. Totally, the modal frequencies included 5652 samples. A data pre-processing procedure was then applied to eliminate some missing data, in which the total number of real-valued dynamic features corresponds to 3932, where the first 3475 samples are associated with the normal condition of the Z24 bridge and the remaining 457 samples belong to the damaged state.

In the context of SHM, the Z24 bridge is one of the famous benchmark structures due to having one-year monitoring data and existing strong environmental variability. For example, Figure 2 depicts the identified modal frequencies of the first and fourth modes. As can be seen, there are considerable sudden jumps in the modal frequencies of the normal condition, which imply the strong influences of the environmental fluctuations. Therefore, it is prominent to address this critical issue by a robust machine learning method. To detect damage by the

proposed non-parametric method, it is initially necessary to define the training and test matrices. Accordingly, 80% of the modal frequencies of the normal condition, Samples 1–2780, are taken to make the training matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{2780}] \in \mathbb{R}^{4 \times 2780}$, where $p = 4$ and $n = 2780$. Subsequently, the remaining 20% of the modal frequencies of the normal condition and all samples regarding the damaged states, Samples 3476–3932, are gathered to produce the test matrix $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_{1152}] \in \mathbb{R}^{4 \times 1152}$, where $m = 1152$. Notice that one supposes that only one test sample is available at each time to regard online damage detection.

Based on the proposed method, the first step in the training phase is to determine the local density and minimum distance value of each training sample. In this regard, Figure 3 shows the amounts of $\lambda_{x_1}, \dots, \lambda_{x_{2780}}$ and $\delta_{x_1}, \dots, \delta_{x_{2780}}$. An important note in Figure 3(a) is related to sudden reductions in the local density values between the samples 400–800 and 1200–1600, where their modal frequencies were highly influenced by the environmental fluctuations. This result means that the local data density originated from the theory of empirical machine learning can be an indicator for finding low dense areas in sampling data. Since such areas are far away from the other data samples, the local density is an appropriate tool for outlier detection. From Figure 3(b), the other important note is that the rationale

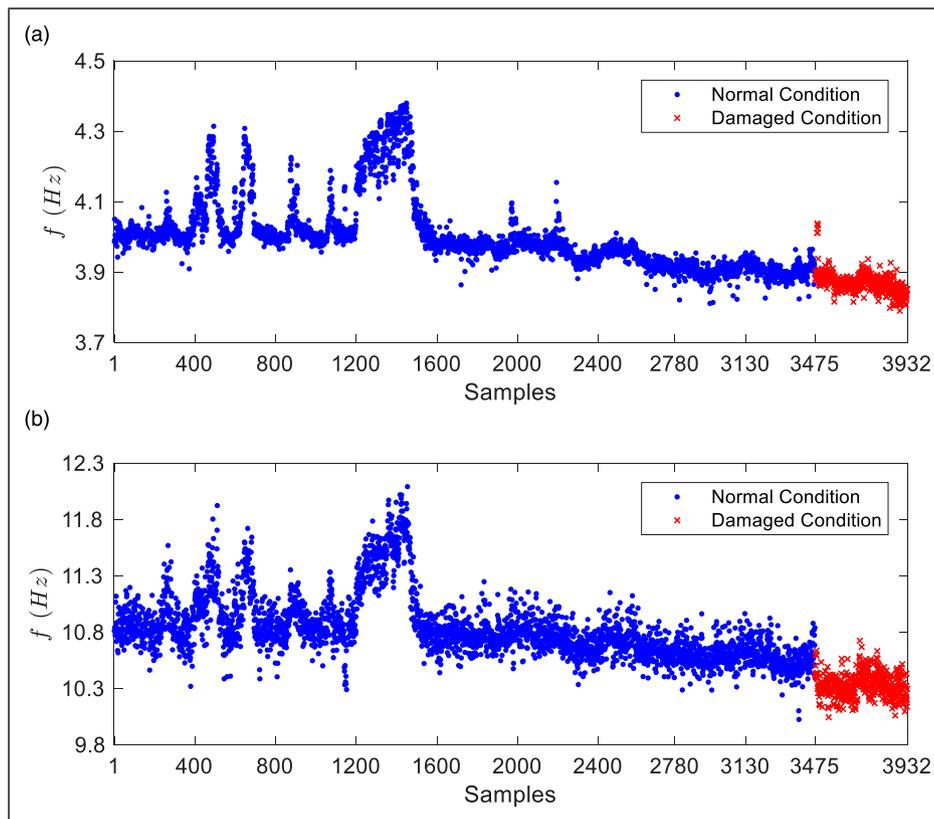


Figure 2. The modal frequencies of the Z24 bridge: (a) the first mode and (b) the fourth mode.

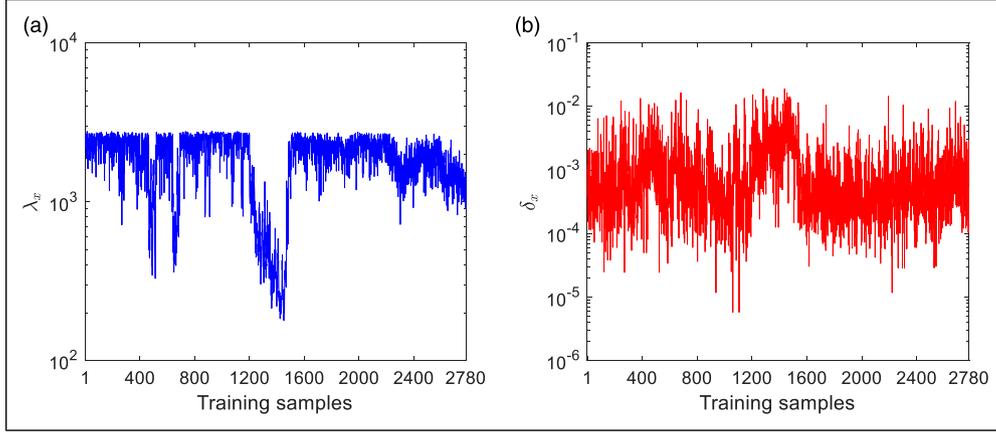


Figure 3. The main components for obtaining the damage indices of the training samples regarding the Z24 bridge: (a) the local densities $\lambda_{x_1}, \dots, \lambda_{x_{2780}}$ and (b) the minimum distance values $\delta_{x_1}, \dots, \delta_{x_{2780}}$.

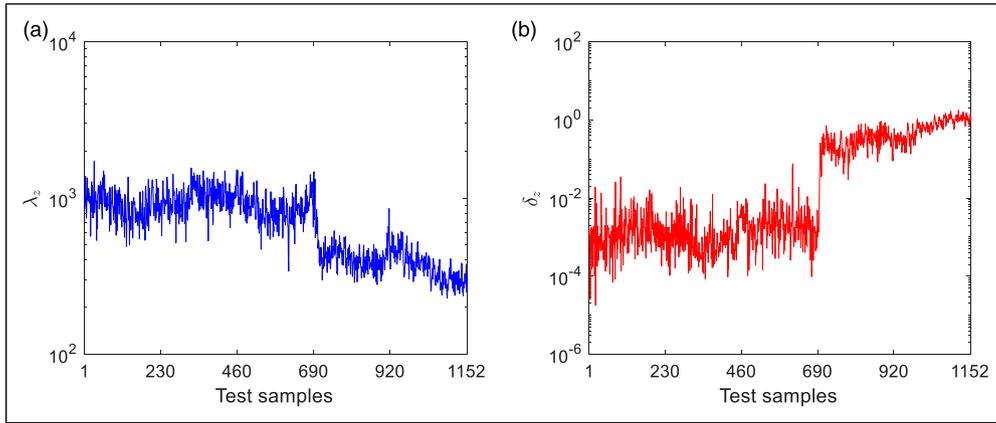


Figure 4. The main components for obtaining the damage indices of the test samples regarding the Z24 bridge: (a) the local densities $\lambda_{z_1}, \dots, \lambda_{z_{1152}}$ and (b) the minimum distance values $\delta_{z_1}, \dots, \delta_{z_{1152}}$.

behind choosing the minimum distance value is able to reduce the effect of the environmental variability. In other words, a simple comparison between Figure 2 and 3(b) visibly demonstrates that this choice can relatively remove the environmental fluctuations. Nonetheless, there is still a peak between the samples 1200–1600 indicating the existence of such fluctuations.

Having considered all test samples, the main steps of the proposed non-parametric method are implemented to determine the local densities $\lambda_{z_1}, \dots, \lambda_{z_{1152}}$ and minimum distance values $\delta_{z_1}, \dots, \delta_{z_{1152}}$ of these samples. Similarly, Figure 4 displays the amounts of local density and minimum distance. From Figure 4(a), one can also discern that the local densities of the damaged states regarding the test samples z_{691}, \dots, z_{1152} , which are equivalent to the modal frequencies of the samples 3476–3932, decrease indicating low dense areas in sampling data. However, it is not feasible to directly use the local density as a damage indicator. This is because the changes caused by the environmental

variability and damage behave similarly, for which the local density decreases in both conditions. This conclusion also confirms the importance of proposing the damage index obtained from multiplying the local density by the minimum distance value. In fact, these two components assist us in reaching a more reliable damage index by reducing the impact of the environmental variability and providing a meaningful indicator for damage detection. Therefore, the local densities are multiplied by the minimum distance values to obtain $DI_{x_1}, \dots, DI_{x_{2780}}$ and $DI_{z_1}, \dots, DI_{z_{1152}}$ regarding the training and test samples, respectively.

The second step in the training phase is to determine a threshold boundary for decision-making. Using the obtained damage indices $DI_{x_1}, \dots, DI_{x_{2780}}$ one should compute a threshold value based on the GP distribution modeling and the POT technique. Based on the iterative algorithm of adequate exceedance selection proposed by Sarmadi and Yuen,¹⁰ seven maximum damage indices from all 2780 samples provide the best fit to the GP distribution model, in

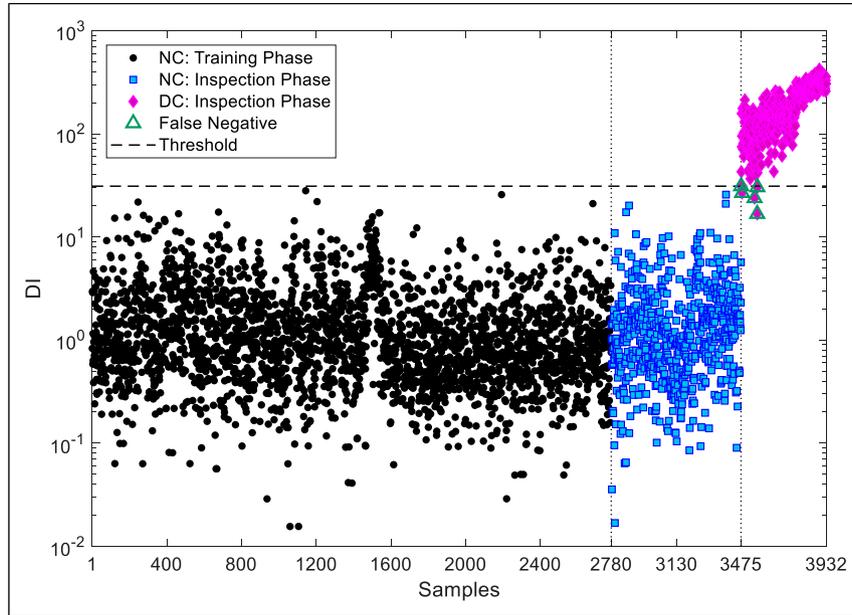


Figure 5. Early damage detection in the Z24 bridge by the proposed anomaly detection method (NC: Normal Condition and DC: Damaged Condition).

which case the best threshold level (u) is equal to ... Notice that the threshold level u is equivalent to the eighth maximum damage index; that is, $u = 17.0569$. Subsequently, the maximum likelihood estimation is applied to estimate the shape and scale parameter of the GP model, which are $\xi = -1.9049$ and $\sigma = 26.6746$. Hence, the threshold boundary at a 5% significance level ($\alpha = 0.05$) corresponds to $\tau = 31.013$. To interpret the result of damage detection, the damage indices of the training and test samples are collected into a vector, whose scores are compared to the obtained threshold boundary for decision-making. On this basis, Figure 5 shows the result of damage detection in the Z24 bridge, where the horizontal line refers to the threshold boundary.

As can be observed, all damage indices of the training samples are below the threshold line. The same result can be drawn for the damage indices of the validation data in the samples 2781–3475. These results convey two important conclusions. First, the proposed non-parametric method and proposed damage index could appropriately deal with the challenge of environmental variability. In this regard, one can perceive that no sudden increases or jumps are available between the samples 400–800 and 1200–1600. Second, the EVT-based threshold determination enables the proposed non-parametric method to make an accurate decision about the training and validation samples and entirely label them as the data related to the normal condition. On the other hand, it is seen in Figure 5 that the majority of the damage indices of the damaged state exceed the threshold except for five points among 457 samples. This demonstrates that the

proposed could correctly discriminate the damaged state from the normal condition and provide a reliable and reasonable result for early damage detection with an inconsiderable error.

Despite the effectiveness and reliability of the proposed non-parametric method, it is necessary to conduct a comparison with some existing non-parametric techniques. In the context of SHM, the anomaly detection based on the Mahalanobis distance is one of the well-known data-driven methods for early damage detection.^{16,17,22,37} This is due to its some advantages such as simplicity, computational efficiency, and non-parametric characteristic. The other non-parametric anomaly detection for SHM applications is based on the SVD-based algorithm.^{15,22} Hence, the comparison process is based on assessing the performance of the proposed non-parametric method against the anomaly detection algorithms developed from the Mahalanobis distance and SVD in terms of damage detectability without considering the threshold boundary. Accordingly, Figure 6 illustrates the variations in the damage indices obtained from the proposed method, Mahalanobis distance metric (DI_m), and SVD (DI_v). Notice that the full discussions and main formulations of these methods can be found in Sarjadi¹⁵ and Figueiredo et al.²²

From Figure 6(a), it is clear that the scales of the damage indices of the damaged state are considerably larger than the corresponding scales regarding the damage indices of the normal condition. As the damaged state is appropriately distinguishable from the normal condition, one can conclude that the proposed non-parametric method can properly

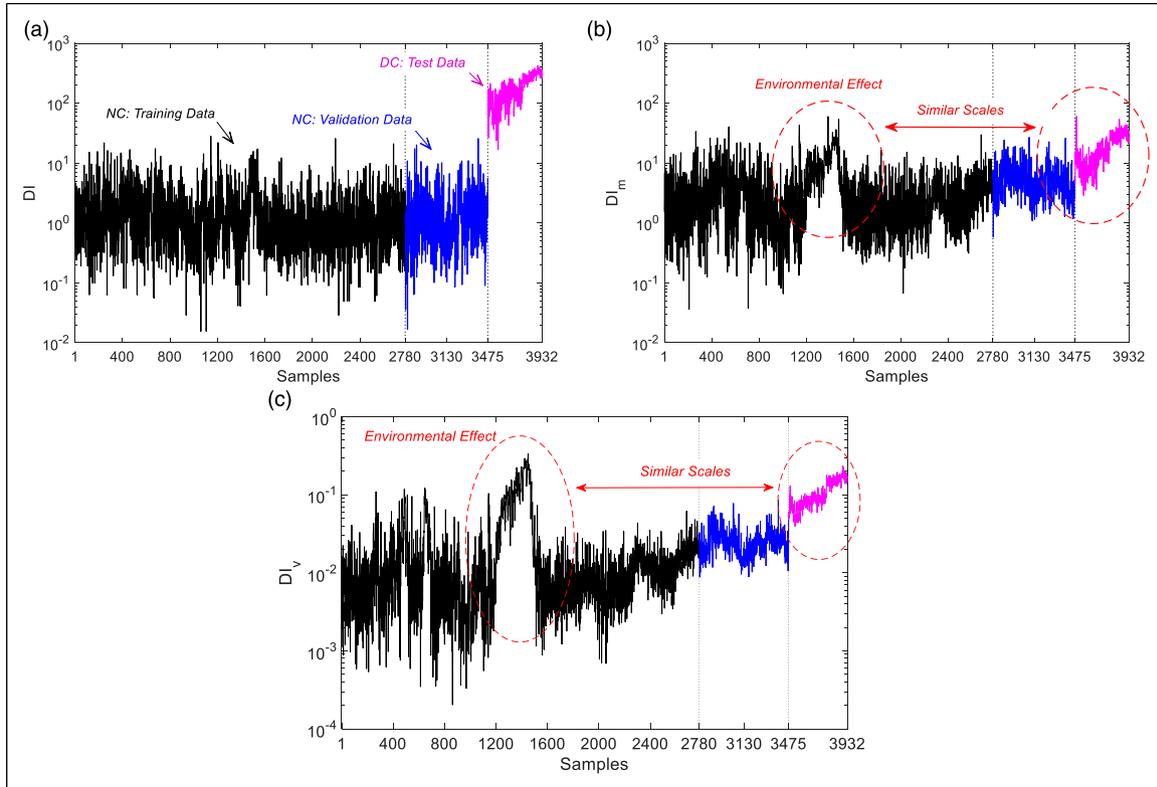


Figure 6. Evaluation of non-parametric anomaly detectors: (a) the proposed method, (b) Mahalanobis distance, and (c) SVD.

Table 1. Comparison of the proposed method with the classical and recently published anomaly detection techniques based on the decision-making errors.

Method (%)	False positive	False negative	Total
Proposed	0 (0.00)	5 (0.87)	5 (0.12)
MD	0 (0.00)	457 (100)	457 (11.62)
SVD	0 (0.00)	457 (100)	457 (11.62)
AANN-MD ¹⁵	0 (0.00)	8 (1.75)	8 (0.20)
AANN-SVD ¹⁵	3 (0.09)	11 (2.40)	14 (0.35)

SVD: singular vector decomposition; AANN-MD: auto-associative neural networks-Mahalanobis distance; AANNA-SVD: auto-associative neural networks-singular vector decomposition; MD: Mahalanobis distance.

provide high damage detectability. By contrast, as Figure 6(b) and (c) appears, the sudden jump in the modal frequencies of the normal condition between the samples 1200–1600 (see Figure 2) is still available in the damage indices obtained from the Mahalanobis distance and SVD. On the other hand, there is not a proper difference between the distance quantities (i.e., DI_m and DI_v) of the damaged and normal conditions. Furthermore, in Figure 6(b) and (c), some distance values of the damaged state have similar or smaller scales than the distance quantities of the normal condition. This implies low damage detectability of the classical Mahalanobis distance and the SVD-based anomaly

detection. This comparison proves that these classical techniques are not reliable for decision-making under strong variability in data or features. Therefore, one can conclude that the proposed method is highly superior to these classical techniques.

To further evaluate, Table 1 lists the numbers and percentages of false positive, false negative, and total errors in early damage detection of the Z24 bridge regarding the proposed non-parametric method, the classical Mahalanobis distance (MD), and SVD. Apart from these techniques, this comparison introduces two other machine learning methods proposed by Sarmadi.¹⁵ These methods are based on combinations of auto-associative neural networks (AANNs) with the Mahalanobis distance and SVD, called AANN-MD and AANN-SVD, which were developed to reduce the decision-making errors by dealing with the problem of environmental/operational variability. From the data in Table 1, it is seen that the proposed method with the aid of the EVT-based threshold estimator presents the best performance with the smallest rates (i.e., the number and percentage) of the decision-making errors. Having considered the same threshold estimator, AANN-MD and AANN-SVD are the next effective techniques for early damage detection. Nonetheless, the classical Mahalanobis distance and SVD fail in providing reasonable results. Although they have no false positive errors, all damage

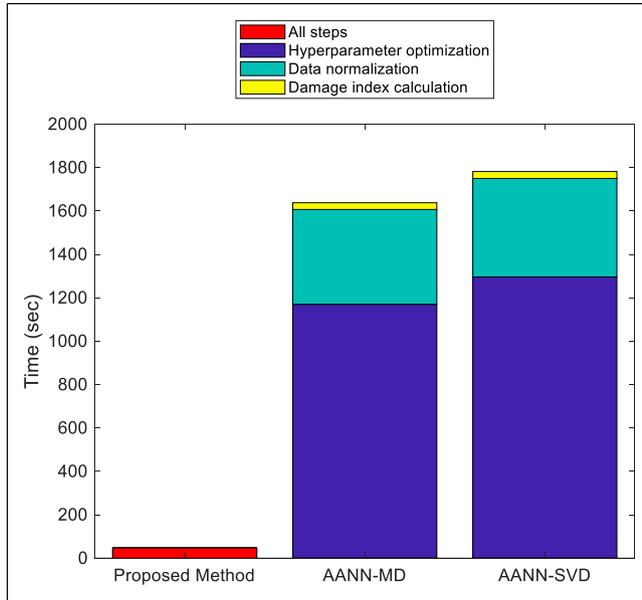


Figure 7. Computational time for decision-making by the proposed, AANN-MD, and AANN-SVD methods in early damage detection of the Z24 bridge.

indices related to the damaged state fall below the threshold, in which case all of them are incorrectly detected as the normal indices (i.e., an entire false negative error). This means that the sudden jump in the modal frequencies of the Z24 bridge between the samples 1200–1600 are still available in the damage indices obtained from the Mahalanobis distance and SVD. This conclusion is also observable in Figure 6(b) and (c). Regarding the total error, one can conclude that the proposed non-parametric method yields the best performance with the smallest error rate. After that, AANN-MD and AANN-SVD give the reasonable results. However, both the classical Mahalanobis distance and SVD cannot provide reliable outputs of decision-making due to their considerable errors.

In the other comparison, the computational time for decision-making by the proposed method, AANN-MD, and AANN-SVD is measured to compare their efficiency. The result of this comparison is presented in Figure 7. This procedure is conducted by the MATLAB function “tic-toc” and a computer with the specification of Intel® Core i5, 2.20 GHz CPU and 8 G RAM. Regarding AANN-MD and AANN-SVD, both of them contain three main steps for decision-making called hyperparameter optimization (i.e., selection of the neurons of the hidden layers of the AANN), data normalization (i.e., the removal of the effects of environmental/operational variability from initial features) by an AANN, and damage index calculation via the classical Mahalanobis distance and SVD. More details about these techniques are available in Sarmadi.¹⁵ As can be seen in Figure 7, the proposed method requires very short time

for decision-making (i.e., ~47 s). By contrast, both AANN-MD and AANN-SVD take longer time due to the procedures of hyperparameter optimization and data normalization. Therefore, one can conclude that despite relatively better performance in decision-making, the proposed method is more efficient than AANN-MD and AANN-SVD. Notice that the computational time presented in Figure 7 will be decreased by applying better computer systems.

The other comparison is to consider smaller training samples in the process of feature classification. In most of the machine learning methods, a good performance is obtained by using adequate or relatively large training samples.^{16,17,36} For this reason, it is attempted to investigate the performance of the proposed method with smaller fractions of normal samples equal to 30% and 50% for making the training data. Accordingly, one can generate new training matrices including 1042 and 1738 training samples. The remaining feature samples of the normal condition (i.e., the samples 1043–3475 and 1739–3475 for the 30% and 50% training sample ratios) are collected to the features of the damaged state (i.e., the samples 3476–3932) to determine two different matrices of test samples. Using the new training and test matrices, all steps of the proposed method along with the EVT-based threshold estimator are implemented to obtain two types of damage indices for early damage detection. In this regard, Figure 8 shows the results of early damage detection of the Z24 bridge via 30% and 50% training sample ratios.

From Figure 8(a), one can discern that the proposed method provides reliable damage detectability with large damage indices of the damaged state over the threshold boundary. Moreover, no false positive is observable in the training samples. However, there are numerous false positive errors in the validation samples. Apart from this issue, the sudden jump in the modal frequencies between the samples 1200–1600 are also available in the validation data. The result in Figure 8(a) clearly demonstrates the serious influence of taking a small value of the training sample ratio on the performance of the proposed method. Although Figure 8(b) shows that the use of 50% of normal features for making the training set improves the decision-making process, there are some false negative errors in the damage indices of the damaged condition. For better clarification, Table 2 presents the decision-making errors in early damage detection under 80%, 50%, and 30% of training sample ratios. As expected, the best and worst performances are related to the ratios equal to 80% and 30%. On the other hand, although the ratio 50% outperforms 30%, it cannot yield the good performance of 80%. Thus, similar to most of the machine learning techniques, it can be concluded that the proposed method requires relatively large training samples by considering all possible

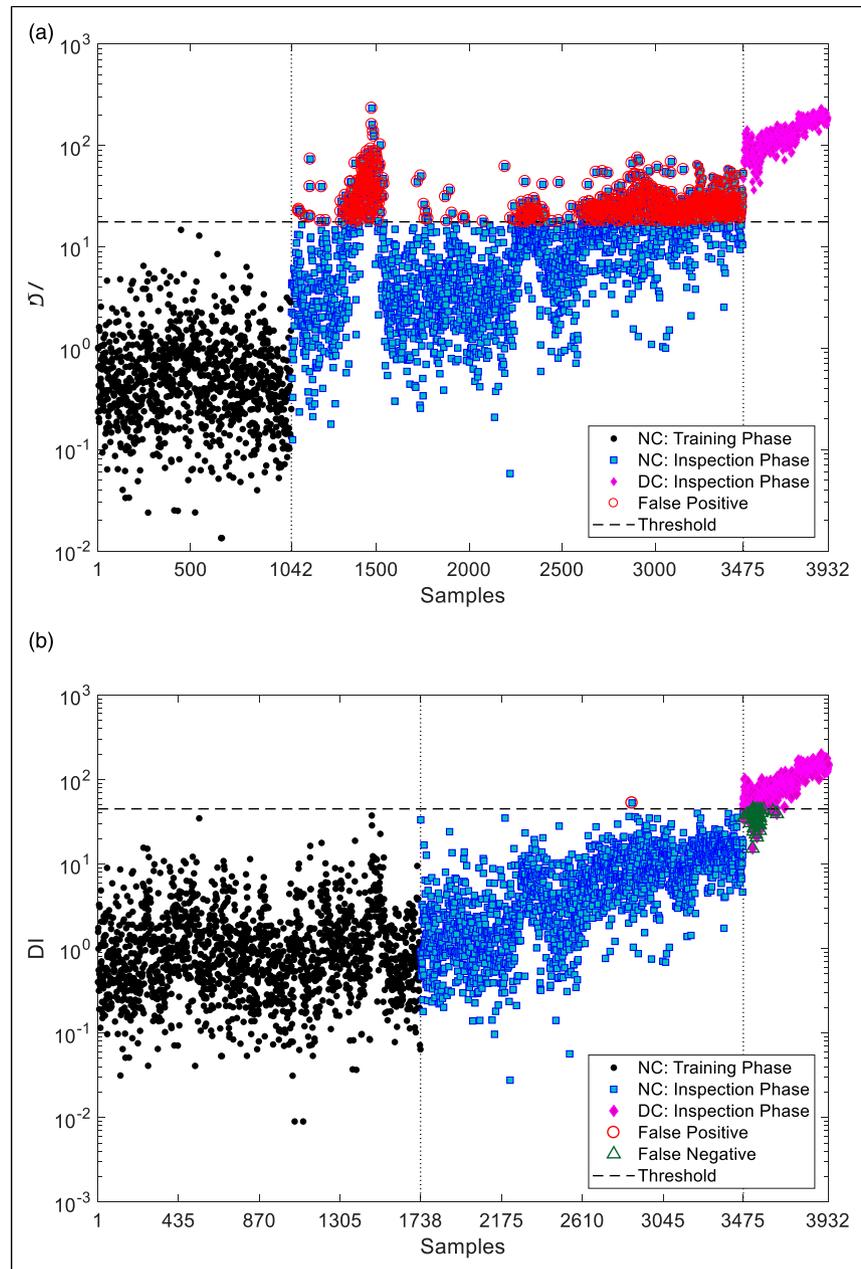


Figure 8. Early damage detection in the Z24 bridge by the proposed anomaly detection method by using smaller training sample ratios: (a) 30% and (b) 50%.

Table 2. Performance evaluation of the proposed method under different training sample ratios.

Training sample ratio, (%)	False positive	False negative	Total
80	0 (0.00)	5 (0.87)	5 (0.12)
50	1 (0.03)	35 (7.65)	36 (0.91)
30	752 (21.64)	0 (0.00)	752 (19.12)

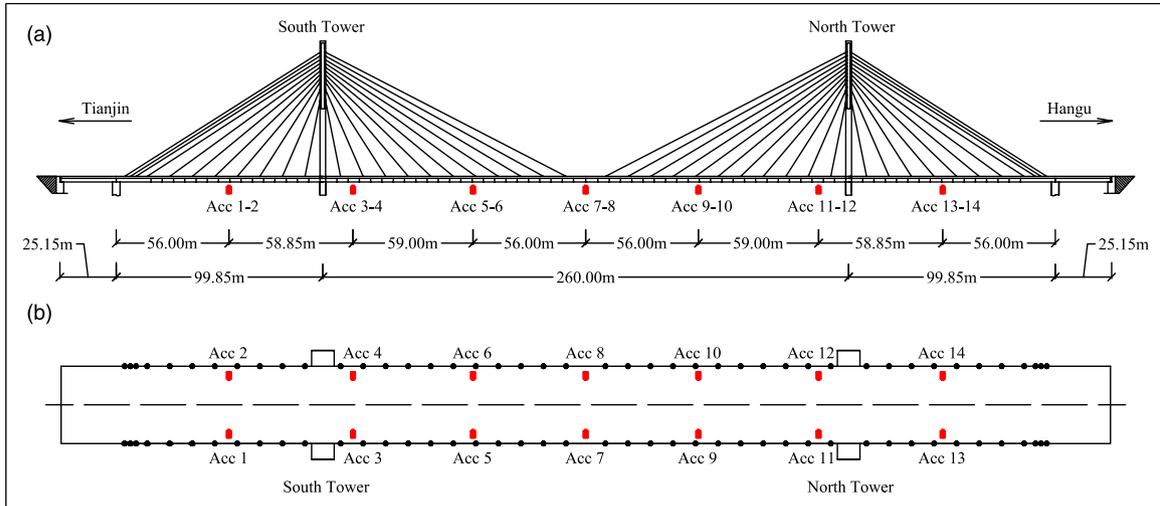


Figure 9. The cable-stayed bridge: (a) the elevation view and main dimensions and (b) the plan view and sensor locations and labels.

environmental/operational variability conditions to provide the best performance for decision-making.

A cable-stayed bridge

In the previous case study, the effectiveness and good performance of the proposed non-parametric method were proposed by the relatively large-size modal features in a long-term scheme. In this section, one attempts to verify the proposed method by small-size or low-dimensional features from a short-term monitoring process. For this reason, the measured vibration responses and statistical features extracted from them related to a cable-stayed bridge are considered. This structure is one of the earliest cable-stayed bridges with continuous pre-stressed box-girder constructed in China. In some literature, it is called the Tianjin-Yonghe bridge, which is still under operation. Figure 9 shows this bridge along with its main dimensions (i.e., the main span of 260 m and two side-spans of 25.15 and 99.85 m, and the height of 60.5 m for both concrete towers). In 2005, after 19 years of operation, some cracks were found at the bottom of a girder segment over the mid-span. In addition, some cables near the anchors were corroded severely. A major rehabilitation program was performed to replace the damaged girder and all cables between 2005–2007. A sophisticated SHM system organized by the Center of Structural Monitoring and Control (SMC) at the Harbin Institute of Technology in China was designed and installed on the bridge so as to measure environmental and dynamic data. During a routine inspection in August 2008, new damage patterns were identified in the girders of the bridge. Due to the equipment of the bridge with different sensors, it is possible to use the measured vibration responses, acquired from ambient vibration, for validating the proposed method.

As Li et al.³⁸ reported, the vibration responses (acceleration time histories) on 12 days (i.e., 1 January, 17 January, 3 February, 19 March, 30 March, 19 April, 5 May, 18 May, 31 May, 7 June, 16 June, and 31 July) are available to utilize them for damage detection. The vibration measurements consist of 24 sets of 1-h acceleration time histories with 360,000 data points per measurement (h) acquired from 14 single-axis accelerometers, as shown in Figure 9. The sampling frequency and time interval of the acceleration responses are identical to 100 Hz and 0.01 s, respectively. Because the data samples of the 10th sensor are meaningless, it is neglected to apply to the SHM process. On the other hand, the vibration responses of 3 days including 31 May, 7 June, and June 16 are excluded due to poor excitation conditions or lacking stability in the consecutive sets.^{39,40} Accordingly, the vibration data of 9 days (i.e., 1 January, 17 January, 3 February, 19 March, 30 March, 19 April, 5 May, 18 May, and 31 July) from 13 sensors are incorporated into the SHM process. In this case, the first 8 days refer to the normal condition of the bridge and the last day (i.e., 31 July) is a damaged state.³⁸

Feature extraction on the vibration time-domain responses of the bridge was carried out by time series modeling through autoregressive moving average (ARMA) models as fully discussed in Entezami et al.⁴⁰ Hence, the final feature sets of this feature extraction are considered here. These sets include the variances of the ARMA residuals from 13 sensors leading to a feature matrix of 13×192 , where the number 192 is obtained from multiplying 24 (the total measurement time for each day) by 8 (the number of days with the normal condition). The information of the ARMA models fitted to the 8 days is also used in the vibration data of the last day to extract a new feature matrix of the same size. To detect damage via the proposed non-parametric method, such small-size feature samples are

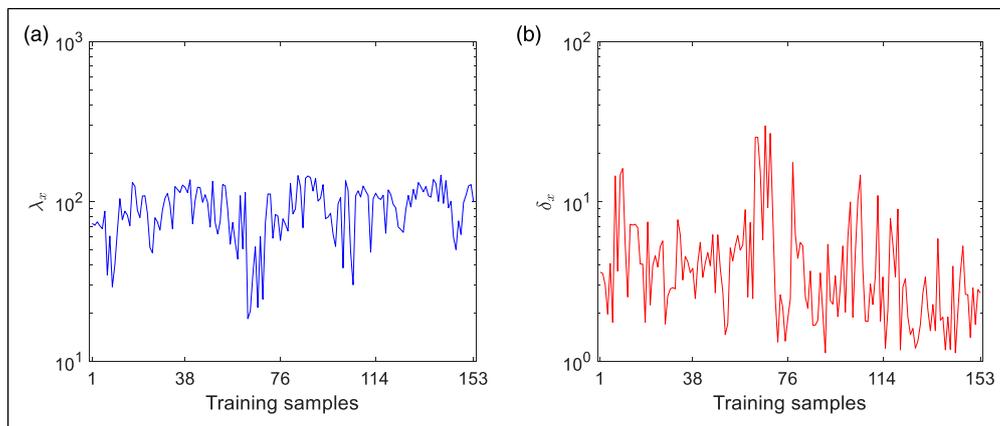


Figure 10. The main components for obtaining the damage indices of the training samples regarding the cable-stayed bridge: (a) the local densities $\lambda_{x_1}, \dots, \lambda_{x_{153}}$ and (b) the minimum distance values $\delta_{x_1}, \dots, \delta_{x_{153}}$.

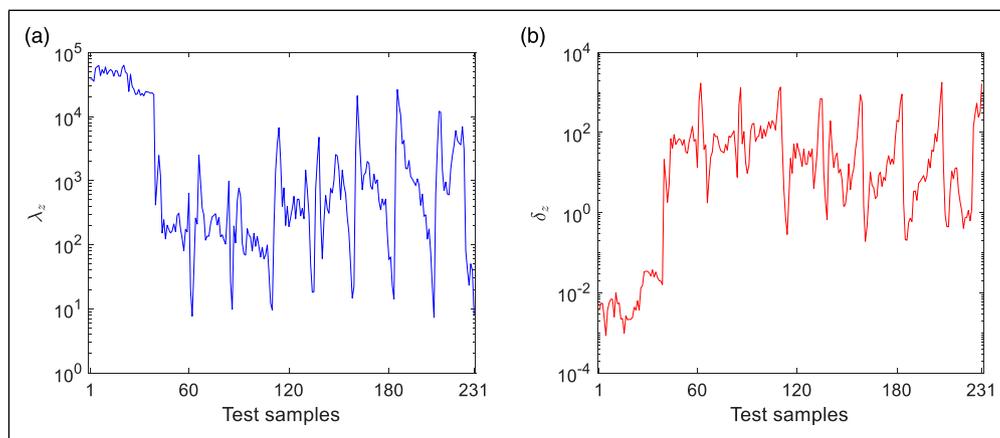


Figure 11. The main components for obtaining the damage indices of the test samples regarding the cable-stayed bridge: (a) the local densities $\lambda_{z_1}, \dots, \lambda_{z_{231}}$ and (b) the minimum distance values $\delta_{z_1}, \dots, \delta_{z_{231}}$.

divided into the training and test matrices. In this regard, 80% of the values of the first feature matrix related to the normal condition is taken to make the training matrix $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{153}] \in \mathbb{R}^{13 \times 153}$, where $p = 13$ and $n = 153$. Moreover, the remaining 20% of the variance amounts of the first feature matrix and all variances of the second feature matrix related to the damaged state are gathered to produce the test matrix $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_{231}] \in \mathbb{R}^{13 \times 231}$, where $m = 231$.

Based on the main steps of the proposed non-parametric method in the training phase, one initially needs to compute the local density of each feature vector in \mathbf{X} . In this case, it can be determined the local densities $\lambda_{x_1}, \dots, \lambda_{x_{153}}$ as shown in Figure 10(a). Using the Euclidean distance, the minimum distance values $\delta_{x_1}, \dots, \delta_{x_{153}}$ of all training samples are then determined, see Figure 10(b). Unlike the previous case study, it is seen that no considerable reduction is available in the local densities, which means that the time series features were not influenced by the environmental and/or operational variability conditions. According to the research

conducted by Sarmadi,¹⁵ who fully discussed this problem by proposing an automated prediction level of the environmental and operational variability conditions, one can state that the aforementioned conclusion is reasonable. Finally, the damage indices of the training samples $DI_{x_1}, \dots, DI_{x_{153}}$ are calculated by multiplying these amounts.

In the inspection phase, the procedures of the local density and minimum distance value calculations are repeated by using the all training samples to determine $\lambda_{z_1}, \dots, \lambda_{z_{231}}$ and $\delta_{z_1}, \dots, \delta_{z_{231}}$ as shown in Figure 11(a) and (b), respectively. Similarly, these amounts are multiplied to obtain the damage indices of the test samples $DI_{z_1}, \dots, DI_{z_{231}}$. To evaluate the cable-stayed bridge for the problem of damage occurrence, the EVT-based threshold determination approach is used to estimate a threshold value via the GP distribution modeling and the POT framework. In this regard, based on the method proposed by Sarmadi and Yuen,¹⁰ the five maximum samples from $DI_{x_1}, \dots, DI_{x_{153}}$ are selected as the adequate exceedances. Accordingly, the

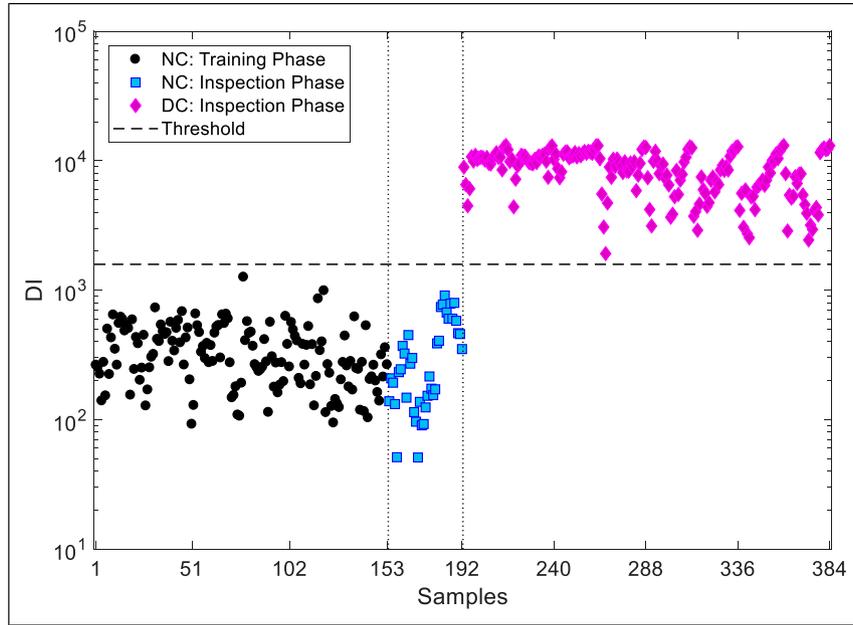


Figure 12. Early damage detection in the cable-stayed bridge by the proposed anomaly detection method.

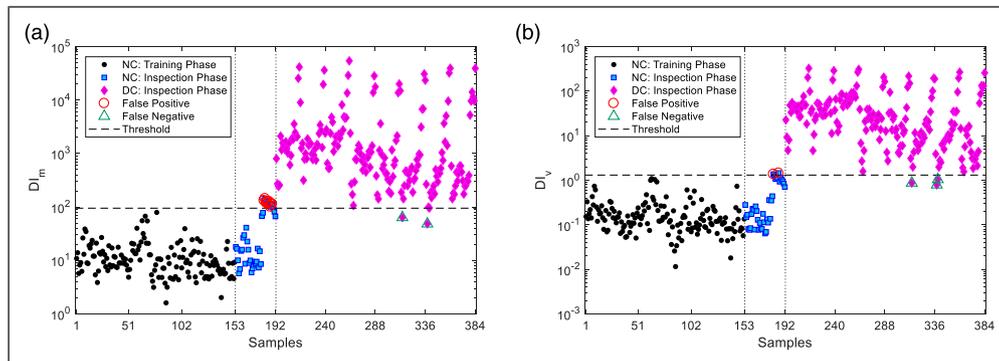


Figure 13. Early damage detection of the cable-stayed bridge by the classical non-parametric anomaly detection methods: (a) Mahalanobis distance and (b) singular vector decomposition.

shape and scale of the GP distribution are identical to $\xi = -1.3564$ and $\sigma = 1281$, respectively. By choosing the sixth maximum sample as the threshold level $u = 654.2939$ and 5% significance level $\alpha = 0.05$, the threshold boundary for the problem of damage detection of the cable-stayed bridge corresponds to $\tau = 1582$.

All damage indices of the training and test samples are then collected into a vector to compare the damage indices with the threshold boundary. Figure 12 shows the result of early damage detection of the cable-stayed bridge, where the horizontal line is indicative of the threshold. As can be observed, all damage indices concerning the normal condition, both the training and validation samples, are below the threshold line. This result confirms that the proposed non-parametric method is still reliable when a small set of

features is incorporated in the process of damage detection with no false positive. Moreover, one can ensure that the EVT-based method for threshold determination is also successful in giving an accurate threshold boundary with a such feature set. Furthermore, the damage indices of the damaged state in the sample 193–384 exceed the threshold implying accurate damage detection without any false negative.

To further demonstrate the effectiveness and reliability of the proposed method, it is compared with the classical anomaly detection techniques based on the Mahalanobis distance and SVD. Unlike the previous case study, where the severe influence of the environmental variability could not allow us to compare the non-parametric approaches via the threshold limits, the current comparative study is based

on comparing the damage indices (distance values) obtained from the Mahalanobis distance and SVD with two threshold boundaries gained by their damage indices in the normal condition. It needs to mention that the same EVT-based approach is applied to estimate these two threshold values.

Hence, Figure 13 presents the results of early damage detection by using the classical non-parametric methods, where the horizontal lines refer to the aforementioned threshold boundaries. Due to the weak effect of the

Table 3. Comparison of the proposed method with the classical and recently published anomaly detection techniques based on the decision-making errors.

Method (%)	False positive	False negative	Total
Proposed	0 (0.00)	0 (0.00)	0 (0.00)
MD	9 (4.68)	2 (1.04)	11 (2.86)
SVD	2 (1.04)	3 (1.56)	5 (1.30)
AANN-MD ¹⁵	0 (0.00)	0 (0.00)	0 (0.00)
AANN-SVD ¹⁵	0 (0.00)	0 (0.00)	0 (0.00)

MD: Mahalanobis distance; SVD: singular vector decomposition; AANN-MD: auto-associative neural networks-Mahalanobis distance; AANN-SVD: auto-associative neural networks-singular vector decomposition.

environmental and/or operational variability conditions on the time series features (i.e., the variances of the ARMA residuals),¹⁵ one can see that the distance values related to the damaged state in the samples 193–384 are far away from the corresponding values regarding the normal condition. On the other hand, although most of the distance quantities of the training and validation samples are under the threshold lines, some false positive errors, that is, 9 out of 192 in Figure 13(a) and 2 out of 192 in Figure 13(b), are observable in the validation samples. Furthermore, both classical techniques yielded few false negative errors, that is, 2 out of 192 in Figures 13(a) and 3 out of 192 in Figure 13(b). An important observation in Figure 13 is that the variability in distance values obtained from the Mahalanobis distance and SVD is much larger than the damage indices of the proposed method. All the conclusions confirm that this method is superior to the classical non-parametric techniques in spite of low effects of the environmental and/or operational variability conditions on the time series features (i.e., the variances of the ARMA residuals).

For further investigation, Table 3 lists the numbers and percentages of the false positive, false negative, and total errors in early damage detection of the cable-stayed bridge

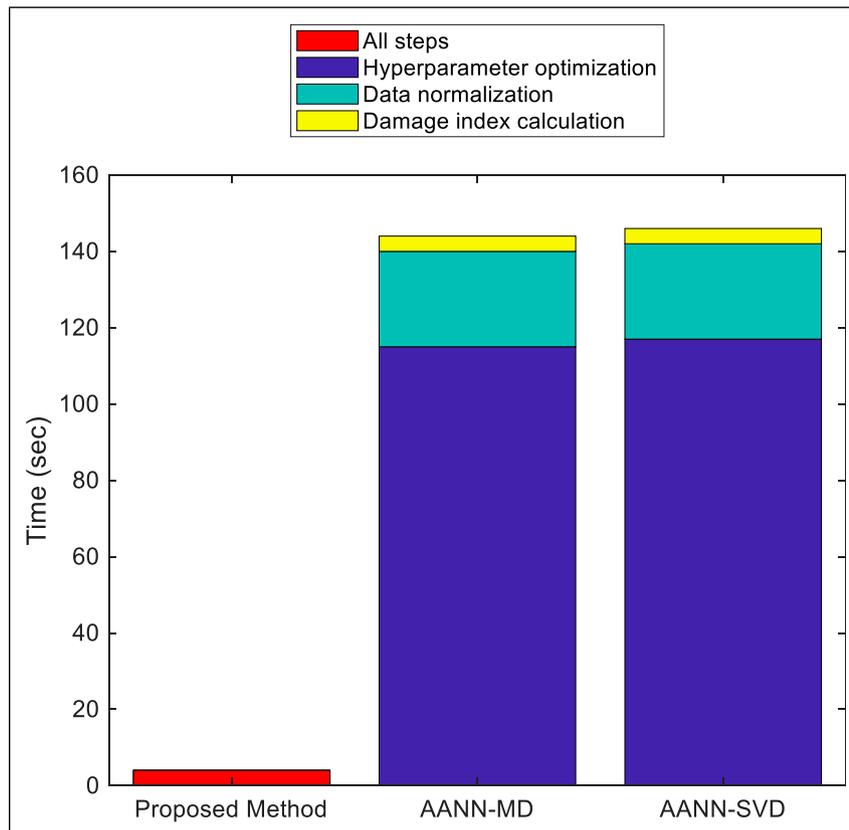


Figure 14. Computational time for decision-making by the proposed, AANN-MD, and AANN-SVD methods in the SHM problem of the cable-stayed bridge.

by using the EVT-based threshold estimator. Similar to the previous case study, the amounts in this table intend to compare the proposed non-parametric method with the classical Mahalanobis distance and SVD as well as AANN-MD and AANN-SVD. Compared with the classical approaches, it is clear that the proposed non-parametric method along with AANN-MD and AANN-SVD provide the best performances of decision-making without any decision-making errors. Although the error rates regarding the classical Mahalanobis distance and SVD are not as inappropriate as the corresponding rates in the SHM problem of the Z24 bridge, they cannot still yield the good performance of the proposed method.

Furthermore, Figure 14 shows the comparison between the proposed method and AANN-MD and AANN-SVD in terms of computational time. Using the same computer system as the previous case study, it is observed that the proposed method is more efficient than AANN-MD and AANN-SVD (i.e., its computational time is about 4 s). In contrast, the procedures of hyperparameter optimization and data normalization take the longest computational time for AANN-MD and AANN-SVD. Thus, this comparison reveals that the proposed method is much more efficient than AANN-MD and AANN-SVD despite their similar outputs (i.e., error rates) of decision-making.

Conclusions

To detect early damage in short-term and long-term monitoring schemes, this article proposed a novel non-parametric anomaly detection method based on the theory of empirical machine learning. Accordingly, a new damage index was proposed by multiplying the local data density of each feature obtained from the empirical machine learning framework by the minimum distance value of that feature. By implementing this process for each of the training and test samples in the training and inspection phase, their damage indices were computed to detect damage. An EVT-based method based on the GP distribution modeling and the POT framework was considered to determine a threshold boundary for decision-making. The dynamic (modal) and statistical (time series) features of two full-scale bridge structures were applied to validate the effectiveness and reliability of the proposed non-parametric method. Finally, this method was compared with two classical non-parametric anomaly detection based on the Mahalanobis distance and the SVD algorithm.

The results on both structures demonstrated that the proposed method is highly successful in discriminating the damaged state from the normal condition in long-term and short-term procedures with the strong and weak environmental variability conditions, respectively. This method

could also deal with the major challenge regarding the environmental variability and significantly reduced its negative influences. In the problem of the Z24 bridge, it was indicated that the proposed local data density cannot alone consider as a damage index. In this case, the idea of finding the minimum distance value of each feature in the training samples helped to define a more effective and reliable damage index. In the SHM problem of both bridges, the comparative studies revealed that the proposed non-parametric anomaly detector outperforms the classical Mahalanobis distance and SVD techniques in terms of providing higher damage detectability and smaller false positive and false negative errors. The comparisons between the proposed method with AANN-MD and AANN-SVD in both bridges showed that all three methods yield reasonable decision-making outputs. However, the proposed method gives better performance under strong environmental/operational variability related to the long-term SHM problem of the Z24 bridge. Moreover, this method is much more efficient than AANN-MD and AANN-SVD in terms of computational time. Nonetheless, the use of a very small ratio for making training samples seriously reduces the effectiveness of the proposed method. In this case, one should consider a relatively large training data by capturing all possible environmental/environmental variability conditions.

Despite reasonable and good performance of the proposed non-parametric method for early damage detection, its main limitations include the necessity of applying adequate training samples and its offline learning scheme in the training phase. Accordingly, it is recommended to develop the proposed method for smaller training samples in an online learning manner. Another recommendation for further research is to evaluate the performance of the proposed method or its developed version on other types of structures as well as optimum sensor configurations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Alireza Entezami  <https://orcid.org/0000-0002-4864-2120>

Hashem Shariatmadar  <https://orcid.org/0000-0001-5966-3317>

References

1. Cawley P. Structural health monitoring: closing the gap between research and industrial deployment. *Struct Health Monit* 2018; 17: 1225–1244.

2. Vagnoli M, Remenyte-Prescott R and Andrews J. Railway bridge structural health monitoring and fault detection: state-of-the-art methods and future challenges. *Struct Health Monit* 2018; 17: 971–1007.
3. Sun L, Shang Z, Xia Y, et al. Review of bridge structural health monitoring aided by big data and artificial intelligence: from condition assessment to damage detection. *J Struct Eng* 2020; 146: 04020073.
4. Gu H, Yang M, Gu C, et al. A comprehensive evaluation method for concrete dam health state combined with gray-analytic hierarchy-optimization theory. *Struct Health Monit*. Epub ahead of print 05 March 2021. DOI: [10.1177/1475921721993388](https://doi.org/10.1177/1475921721993388).
5. Abdulkarem M, Samsudin K, Rokhani FZ, et al. Wireless sensor network for structural health monitoring: a contemporary review of technologies, challenges, and future direction. *Struct Health Monit* 2020; 19: 693–735.
6. Dong C-Z and Catbas FN. A review of computer vision-based structural health monitoring at local and global levels. *Struct Health Monit* 2021; 20: 692–743. DOI: [10.1177/1475921720935585](https://doi.org/10.1177/1475921720935585).
7. Malekloo A, Ozer E, AlHamaydeh M, et al. Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Struct Health Monit* Epub ahead of print 16 August 2021. DOI: [10.1177/14759217211036880](https://doi.org/10.1177/14759217211036880).
8. Mohri M, Rostamizadeh A and Talwalkar A. *Foundations of Machine Learning*. Cambridge, MA, USA: MIT Press, 2012, p. 23–52.
9. Sarmadi H, Entezami A and Ghalehnovi M. On model-based damage detection by an enhanced sensitivity function of modal flexibility and LSMR-Tikhonov method under incomplete noisy modal data. *Eng Comput* Epub ahead of print 25 May 2020. DOI: [10.1007/s00366-020-01041-8](https://doi.org/10.1007/s00366-020-01041-8).
10. Sarmadi H and Yuen KV. Early damage detection by an innovative unsupervised learning method based on kernel null space and peak-over-threshold. *Computer-Aided Civil Infrastructure Eng* 2021; 36(9): 1150–1167, DOI: [10.1111/mice.12635](https://doi.org/10.1111/mice.12635).
11. Sarmadi H and Entezami A. Application of supervised learning to validation of damage detection. *Archive Appl Mech* 2021; 91: 393–410. DOI: [10.1007/s00419-020-01779-z](https://doi.org/10.1007/s00419-020-01779-z).
12. Tan X, Sun X, Chen W, et al. Investigation on the data augmentation using machine learning algorithms in structural health monitoring information. *Struct Health Monit* 2021; 20: 2054–2068. DOI: [10.1177/1475921721996238](https://doi.org/10.1177/1475921721996238).
13. Bull LA, Worden K and Dervilis N. Towards semi-supervised and probabilistic classification in structural health monitoring. *Mech Syst Signal Process* 2020; 140: 106653. DOI: [10.1016/j.ymssp.2020.106653](https://doi.org/10.1016/j.ymssp.2020.106653).
14. Pimentel MAF, Clifton DA, Clifton L, et al. A review of novelty detection. *Signal Process* 2014; 99: 215–249. DOI: [10.1016/j.sigpro.2013.12.026](https://doi.org/10.1016/j.sigpro.2013.12.026).
15. Sarmadi H. Investigation of machine learning methods for structural safety assessment under variability in data: comparative studies and new approaches. *J Perform Constructed Facil* 2021; 35: 04021090. DOI: [10.1061/\(ASCE\)CF.1943-5509.0001664](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001664).
16. Sarmadi H and 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. *Mech Syst Signal Process* 2020; 140: 106495, DOI: [10.1016/j.ymssp.2019.106495](https://doi.org/10.1016/j.ymssp.2019.106495).
17. Sarmadi H, Entezami A, Saeedi Razavi B, et al. Ensemble learning-based structural health monitoring by Mahalanobis distance metrics. *Struct Contr Health Monit* 2021; 28: e2663. DOI: [10.1002/stc.2663](https://doi.org/10.1002/stc.2663).
18. Entezami A and Shariatmadar H. An unsupervised learning approach by novel damage indices in structural health monitoring for damage localization and quantification. *Struct Health Monit* 2018; 17: 325–345.
19. Entezami A, Shariatmadar H and Karamodin A. Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods. *Struct Health Monit* 2019; 18: 1416–1443.
20. Entezami A, Shariatmadar H and Mariani S. Fast unsupervised learning methods for structural health monitoring with large vibration data from dense sensor networks. *Struct Health Monit* 2020; 19: 1685–1710.
21. Entezami A and Shariatmadar H. Structural health monitoring by a new hybrid feature extraction and dynamic time warping methods under ambient vibration and non-stationary signals. *Measurement* 2019; 134: 548–568. DOI: [10.1016/j.measurement.2018.10.095](https://doi.org/10.1016/j.measurement.2018.10.095).
22. Figueiredo E, Park G, Farrar CR, et al. Machine learning algorithms for damage detection under operational and environmental variability. *Struct Health Monit* 2011; 10: 559–572.
23. Catbas FN, Gokce HB and Gul M. Nonparametric analysis of structural health monitoring data for identification and localization of changes: concept, lab, and real-life studies. *Struct Health Monit* 2012; 11: 613–626.
24. Santos JP, Crémona C, Calado L, et al. On-line unsupervised detection of early damage. *Struct Control Health Monit* 2016; 23: 1047–1069.
25. Sarmadi H, Entezami A, Salar M, et al. Bridge health monitoring in environmental variability by new clustering and threshold estimation methods. *J Civil Struct Health Monit* 2021; 11: 1–16. DOI: [10.1007/s13349-021-00472-1](https://doi.org/10.1007/s13349-021-00472-1).
26. Qiu L, Fang F, Yuan S, et al. An enhanced dynamic gaussian mixture model-based damage monitoring method of aircraft structures under environmental and operational conditions. *Struct Health Monit* 2019; 18: 524–545. DOI: [10.1177/1475921718759344](https://doi.org/10.1177/1475921718759344).
27. Daneshvar MH, Gharighoran A, Zareei SA, et al. Early damage detection under massive data via innovative hybrid methods: application to a large-scale cable-stayed bridge.

- Struct Infrastructure Eng* 2021; 17: 902–920. DOI: [10.1080/15732479.2020.1777572](https://doi.org/10.1080/15732479.2020.1777572).
28. Wang Z and Cha Y-J. Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect damage. *Struct Health Monit* 2021; 20: 406–425. DOI: [10.1177/1475921720934051](https://doi.org/10.1177/1475921720934051).
 29. Gharibnezhad F, Mujica LE and Rodellar J. Applying robust variant of principal component analysis as a damage detector in the presence of outliers. *Mech Syst Signal Process* 2015; 50–51: 467–479. DOI: [10.1016/j.ymsp.2014.05.032](https://doi.org/10.1016/j.ymsp.2014.05.032).
 30. Angelov PP and Gu X. *Empirical Approach to Machine Learning*. Cham, Switzerland: Springer Nature, 2019, p. 437.
 31. Cha Y-J and Wang Z. Unsupervised novelty detection-based structural damage localization using a density peaks-based fast clustering algorithm. *Struct Health Monit* 2018; 17: 313–324. DOI: [10.1177/1475921717691260](https://doi.org/10.1177/1475921717691260).
 32. Seyedi SA, Lotfi A, Moradi P, et al. Dynamic graph-based label propagation for density peaks clustering. *Expert Syst Appl* 2019; 115: 314–328. DOI: [10.1016/j.eswa.2018.07.075](https://doi.org/10.1016/j.eswa.2018.07.075).
 33. Entezami A, Sarmadi H, Salar M, et al. A novel data-driven method for structural health monitoring under ambient vibration and high-dimensional features by robust multidimensional scaling. *Struct Health Monit* Epub ahead of print 22 April 2021. DOI: [10.1177/1475921720973953](https://doi.org/10.1177/1475921720973953).
 34. Coles S, Bawa J, Trenner L, et al. *An Introduction to Statistical Modeling of Extreme Values*. London, United Kingdom: Springer, 2001, p. 208.
 35. Rébillat M, Hmad O, Kadri F, et al. Peaks over threshold-based detector design for structural health monitoring: application to aerospace structures. *Struct Health Monit* 2018; 17: 91–107. DOI: [10.1177/1475921716685039](https://doi.org/10.1177/1475921716685039).
 36. Reynders E, Wursten G and De Roeck G. Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. *Struct Health Monit* 2014; 13: 82–93.
 37. Entezami A, Shariatmadar H and Mariani S. Early damage assessment in large-scale structures by innovative statistical pattern recognition methods based on time series modeling and novelty detection. *Adv Eng Softw* 2020; 150: 102923–102939. DOI: [10.1016/j.advengsoft.2020.102923](https://doi.org/10.1016/j.advengsoft.2020.102923).
 38. Li S, Li H, Liu Y, et al. SMC structural health monitoring benchmark problem using monitored data from an actual cable-stayed bridge. *Struct Control Health Monit* 2014; 21: 156–172.
 39. Nguyen T, Chan TH and Thambiratnam DP. Field validation of controlled monte carlo data generation for statistical damage identification employing mahalanobis squared distance. *Struct Health Monit* 2014; 13: 473–488.
 40. Entezami A, Sarmadi H, Behkamal B, et al. Big data analytics and structural health monitoring: a statistical pattern recognition-based approach. *Sensors* 2020; 20: 2328. DOI: [10.3390/s20082328](https://doi.org/10.3390/s20082328).
 41. Rezaiee-Pajand M, Sarmadi H and Entezami A. A hybrid sensitivity function and Lanczos bidiagonalization-Tikhonov method for structural model updating: Application to a full-scale bridge structure. *Appl Math Model* 2021; 89: 860–884. DOI: [10.1016/j.apm.2020.07.044](https://doi.org/10.1016/j.apm.2020.07.044).
 42. Entezami A, Shariatmadar H and Karamodin A. Improving feature extraction via time series modeling for structural health monitoring based on unsupervised learning methods. *Sci Iran* 2020; 27: 1001–1018. doi: [10.24200/SCI.2018.20641](https://doi.org/10.24200/SCI.2018.20641).
 43. Behkamal B, Naghibzadeh M, Pagnani A, Saberi MR and Al Nasr K. Solving the α -helix correspondence problem at medium-resolution Cryo-EM maps through modeling and 3D matching. *J Mol Graph Model* 2021; 103: 107815. doi: [10.1016/j.jmgm.2020.107815](https://doi.org/10.1016/j.jmgm.2020.107815).
 44. Entezami A, Shariatmadar H and Ghalehnovi M. Damage detection by updating structural models based on linear objective functions. *J Civ Struct Health Monit* 2014; 4: 165–176. doi: [10.1007/s13349-014-0072-9](https://doi.org/10.1007/s13349-014-0072-9).