

VIBRATION-BASED ANOMALY DETECTION USING SPARSE AUTO-ENCODER AND CONTROL CHARTS

Rafaelle P. Finotti¹, Carmelo Gentile², Flávio S. Barbosa¹ and Alexandre A. Cury¹

¹UFJF - Federal University of Juiz de Fora
Faculty of Engineering
Department of Applied and Computational Mechanics
e-mail: {rafaelle.finotti, flavio.barbosa, alexandre.cury}@engenharia.ufjf.br

² Politecnico di Milano
Department of Architecture, Built Environment and Construction Engineering (ABC)
e-mail: carmelo.gentile@polimi.it

Keywords: Structural Health Monitoring, Damage Detection, Vibration Signals, Deep Learning.

Abstract. *Several approaches can be found in the scientific literature when the subject is damage detection based on vibration signals. In the last few years, increasing attention has been given to the application of Computational Intelligence algorithms in structural novelty identification. In more details, the powerful data mapping capability of computational deep learning methods has been recently exploited to develop strategies of structural health monitoring through appropriate characterization of dynamic responses. Therefore, the present work is aimed at investigating the capability of a deep learning algorithm called Sparse Auto-Encoder (SAE) to identify structural alterations of the Z24 bridge, a classical benchmark for integrity assessment studies. The main idea is to characterize the Z24 dynamic responses via SAE models and, subsequently, to detect the onset of abnormal behavior through the well-known Shewhart T control chart (T^2 -statistic), calculated with SAE extracted features. An advantage of the proposed methodology is that data are processed directly in the time domain, avoiding modal parameters estimation and tracking analysis. Moreover, control charts are considered suitable tools for continuous monitoring due to their relatively simple implementation. The obtained results demonstrate that the proposed strategy based on SAE and Shewhart T control chart has potential to be explored in structural damage detection problems, since it is able to distinguish between the two investigated scenarios (i.e., undamaged and damaged) of Z24 bridge.*

1 INTRODUCTION

The proper functioning of structural systems and the safety of users are among the main concerns of engineers throughout the life cycle of any civil engineering construction. In order to avoid catastrophic failures, it is important to continuously monitor the structural condition and detect any abnormal behavior at an early stage, especially when dealing with large structures, such as bridges, viaducts, tall buildings, and towers.

Among the different methodologies available to identify the occurrence of structural anomalies, vibration-based damage detection has been extensively investigated in the scientific literature (see e.g. [1]). Starting from dynamic time histories obtained through Structural Health Monitoring (SHM) systems, these approaches are applied to investigate novelties in terms of structural behavior. In summary, essential features are extracted from measured data and are successively compared to deduce if a structural change has occurred.

Typically, the structural integrity investigation is performed by employing modal analysis [2, 3]. The degrading process alters the physical properties of the structure, such as mass and stiffness, which influence its natural frequencies, mode shapes, and damping ratios. In the last few years, due to the evolution of computer and information technologies, increasing attention has been given to the application of Computational Intelligence (CI) in structural novelty identification [4, 5, 6]. Among all possible CI algorithms, those based on deep learning appear as promising alternatives to traditional techniques. In this paper, a special highlight is given to the Sparse Auto-Encoder (SAE), a deep neural network algorithm that automatically extracts features from data. The SAE reconstructs its inputs through an internal coding - modeled by linear and nonlinear functions - that transforms them into a “new” group of variables (features) [7]. Auto-encoders are known not only for the ability to deal with large volumes of data but also for their capability to provide optimal solutions, particularly for nonlinear problems, such as structural anomaly detection [8, 9]. Therefore, SAE may be an appropriate method for handling vibration signals. It is important to notice that deep learning algorithms are recent tools in the SHM area, and studies focused on evaluating them to solve novelty detection problems are ongoing.

In this context, the present work is aimed at investigating the performance of the SAE algorithm when applied to the identification of structural alterations. The fundamental idea here is to characterize the structural dynamic responses via SAE models and, subsequently, to detect the onset of abnormal behavior through the well-known Shewhart T control chart (T^2 -statistic) [10], calculated with SAE extracted features. The Shewhart chart is a sort of statistical process control technique frequently used in SHM strategies and applied to the residuals between measured and predicted (via regression analysis) quantities. Due to the relatively simple implementation, control charts are considered suitable tools for continuous monitoring. The anomaly detection approach is exemplified using data collected on the Z24 bridge [11], before and after damaging the structure.

An advantage of the proposed methodology is that data are processed directly in the time domain, avoiding modal parameters estimation and tracking. Another interesting aspect of such an approach is the unsupervised analysis, which means that it does not use previously labeled observations or desired output variables. Although many methods in the literature are based on the pre-establishment of different degradation levels (supervised analysis), it is difficult to have prior knowledge of the structure’s health condition in actual SHM systems.

2 THEORETICAL BACKGROUND

2.1 Deep Learning and Sparse Auto-Encoder

Deep Learning encompasses a variety of machine learning techniques based on Artificial Neural Network (ANN) theory (see e.g. [12]), mainly characterized by their multiple processing layers. As already stated, the present work is focused on the deep learning algorithm called Sparse Auto-Encoder (SAE). In general terms, an Auto-Encoder (AE) is an Artificial Neural Network (ANN) built to return an approximation of its input. According to Goodfellow *et al.* (2016) [7], this network consists of an internal coding layer described by the function $\mathbf{h} = f(\mathbf{x})$, which learns the characteristics $\mathbf{h}$ of input data $\mathbf{x}$, and a decoding layer defined by $\mathbf{y} = g(\mathbf{h})$, which rebuilds the vector $\mathbf{x}$ from the self-learned vector features $\mathbf{h}$, as suggested by the network architecture shown in Figure 1. Since the cost function evaluates the difference between $\mathbf{x}$ and its own output $\mathbf{y} \approx \mathbf{x}$ ($\mathbf{x}$ reconstructed), the learning of an AE is unsupervised. It is worth mentioning that an AE may have several layers between $\mathbf{x}$ and $\mathbf{h}$, as well as between $\mathbf{h}$ and $\mathbf{y}$ (the AE structure has to be symmetric). However, for didactic purposes, Figure 1 represents an AE with only one encoder and decoder layer.

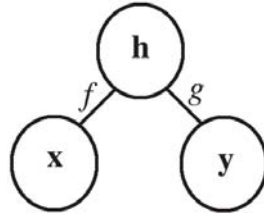


Figure 1: The basic structure of an Auto-Encoder [7].

Nevertheless, the main interest is not in the replicated output of the AE, but in exploring its ability to extract characteristics from “raw data”. If the AE is used for data mapping only, without feature reconstruction, it is designed to produce $\mathbf{h}$ with a smaller dimension than $\mathbf{x}$ and is known as Undercomplete Auto-Encoder. The great advantage of reducing the data from $\mathbf{x}$ vector dimension to $\mathbf{h}$ vector dimension is the identification of relevant parameters at a high-level of abstraction, helpful for recognizing patterns in datasets. In this context, the learning of the AE network is accomplished by minimizing a function $Z(\mathbf{x}, g(f(\mathbf{x})))$ that penalizes the differences between $\mathbf{x}$ and $g(f(\mathbf{x}))$. In cases where the coding function is linear, and Z is the mean quadratic error, the Undercomplete Auto-Encoder behaves like the Principal Component Analysis (PCA) [13]. Conversely, when employing nonlinear functions for f and g , the Undercomplete Auto-Encoder may be more powerful than PCA to reduce the dimensionality of a problem [14].

Despite their efficiency in characterizing data, the encoders and decoders may acquire an excessive ability to approximate $\mathbf{y} \approx \mathbf{x}$, resulting in a vector $\mathbf{h}$ with high dimensionality, which many times is not able to provide interesting parameters for modeling the problem. In order to improve the performance of this deep machine learning technique, the Sparse Auto-Encoder (SAE) was proposed. The SAE is an undercomplete auto-encoder where a sparse penalty $\Gamma(f(\mathbf{x}))$ is incorporated into the function Z in the training process. In summary, this penalty allows the AE model to represent large datasets with a small number of $\mathbf{h}$ components, controlling the amount of active neurons in the layers (most weights equal to zero). Consequently, the addition of sparsity-inducing term to auto-encoders usually leads to an increase in the model performance and to a reduction of processing time.

2.2 Shewhart T Control Chart

The Control Chart is a graphical statistical tool used to monitor the variability of a problem's parameters over time. The charts usually depict several data points, which are formed by a specific statistical characteristic and horizontal lines (control limits) responsible for indicating the extreme values of such characteristic when the problem is in-control state. Any point that is beyond these predetermined limits, on the other hand, reveals unusual sources of variability, suggesting an out-of-control situation. [10].

Due to their relatively simple implementation, intuitive interpretation, and effective results, control charts are considered suitable tools for structural on-line monitoring and anomaly detection. The multivariate control technique used herein is the Shewhart T^2 Control Chart. The characteristic plotted in this chart is the Hotelling's T^2 -statistic. The T^2 -statistic represents the distance between a new data observation and the corresponding sample mean vector - the higher T^2 value, the greater the distance of the new data from the mean. This metric is based on the relationship among the variables and on the scatter of data (covariance matrix). By assuming that matrix $\mathbf{H}_{n \times m}$ represents a dataset during a certain time period (which in this paper are the SAE extracted features), the T^2 -statistic may be calculated as follows:

$$T^2 = r(\bar{\mathbf{h}} - \bar{\bar{\mathbf{h}}})^T \mathbf{S}^{-1} (\bar{\mathbf{h}} - \bar{\bar{\mathbf{h}}}) \quad (1)$$

where $\bar{\mathbf{h}}$ is the sample mean vector of the m available features, obtained from a submatrix of $\mathbf{H}$ with r observations ($\mathbf{H}_{r \times m}$, $r < n$); $\bar{\bar{\mathbf{h}}}$ and $\mathbf{S}$ are the vector of reference averages and the mean of the reference covariance matrices, respectively, both estimated using s preliminary submatrices collected during the in-control state of the problem. In this work, the Upper Control Limit (UCL) is defined as the 95th percentile of the T^2 values of the training data (values greater than UCL may be observed only 5% of the time by chance). The Lower Control Limit (LCL) is zero.

3 THE VIBRATION-BASED ANOMALY DETECTION APPROACH

The present work proposes a structural assessment framework based on SAE models for automatic vibration-data feature learning. Such a framework includes the use of the Hotelling's T^2 control chart applied to SAE extracted characteristics to investigate novelties in terms of structural behavior. Figure 2 shows a general scheme of the suggested anomaly detection methodology.

SHM systems usually comprise a number of accelerometers that record dynamic responses over time. Therefore, it is initially necessary to rearrange the acceleration time histories collected in each measurement point by setting an appropriate window length (duration time) to analyze each measured signal. The vibration measurements are organized in a matrix considering the structural responses of each accelerometer k ($1 \leq k \leq K$) separately, as illustrated in Figure 3. The rows I indicate each rearranged signal, and the columns J represent the respective dynamic response sampled in time. The definition of the signal duration time should consider the measurements' capacity to represent the current structural condition and the amount of available data. Moreover, since the present approach is focused on continuous monitoring, it is also important to keep the chronological order of the collected signals. In order to make the detection model less sensitive to data scale, after the input matrix is assembled, all data values are divided by the maximum absolute amplitude, leading to acceleration signals between $[-1;1]$.

The developed strategy relies on creating a model with data from the system working in normal conditions, which is named as SAE/ T^2 -statistic reference model. At this point, the reference matrix previously organized is divided into training and testing data, respecting their

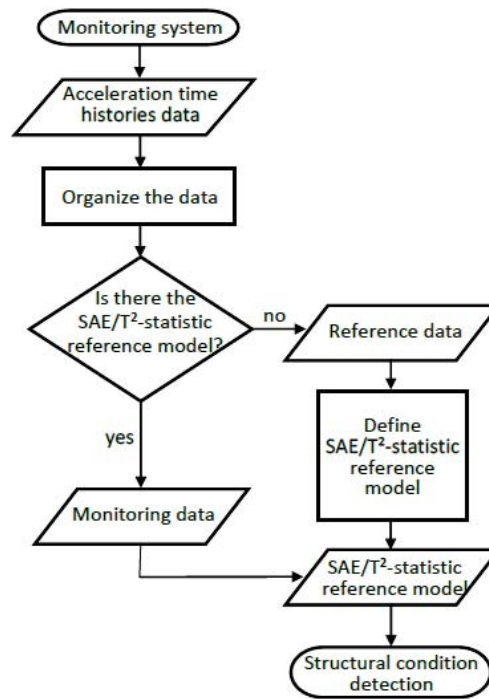


Figure 2: General scheme of the proposed methodology.

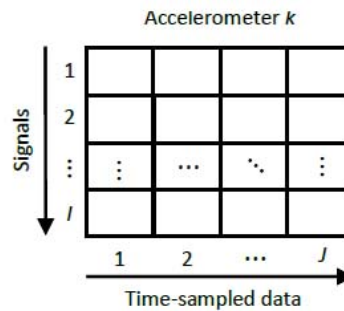


Figure 3: Input data organization for structural anomaly detection approach.

position in time. These two submatrices are used to construct and adjust the SAE/T^2 model through an iterative procedure, described in the following steps:

1. Firstly, the training data are randomized, as usual in algorithms based on ANN theory. This shuffling has the objective of reducing data variance, guaranteeing a greater generalization power for the artificial intelligence model. It is important to emphasize that the randomization is herein performed only on data used to create the artificial intelligence model. After the training task, the data, or rather the SAE extracted features, must be placed in chronological order again;
2. Initial training parameters for the SAE model are defined. These parameters are number of processing layers, number of features to be extracted from time domain responses (number of neurons), optimization method, activation and error functions, maximum number of training iterations (epochs), sparsity proportion ρ and the coefficients β and

λ , related to the sparsity and weights regularization terms of the cost function, respectively. The parameters ρ , β , and λ are related to the sparse penalty function Γ (defined in section 2.1) and assist in the determination of the best solution by the SAE (see e.g. [15]);

3. By using the training data, the SAE model is generated. In the proposed approach, the SAE model is applied to transform the structural signals into a few representative characteristics;
4. The m extracted SAE features are reorganized in chronological sequence (as mentioned in step 1) and used to calculate the T^2 -statistic points and the UCL, as shown in Figure 4;

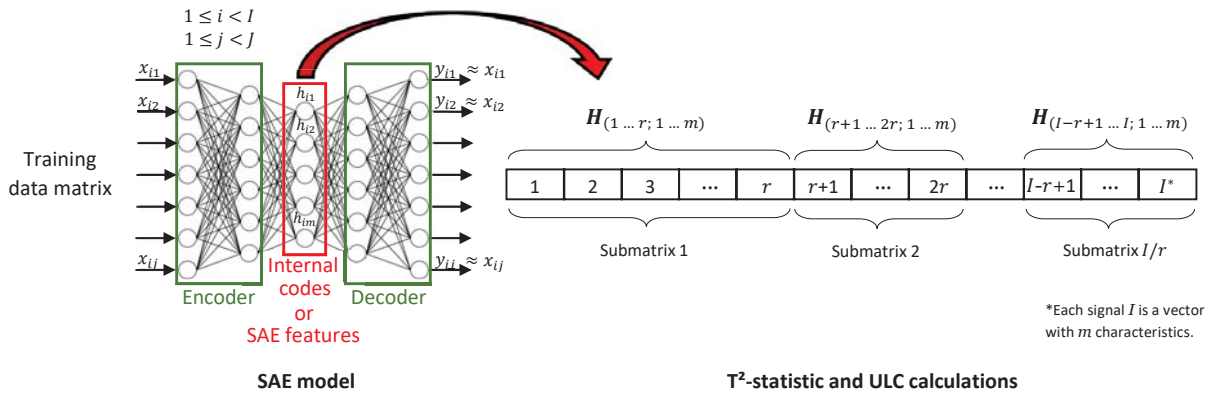


Figure 4: Graphical representation of the T^2 -statistic calculation procedure.

5. In order to evaluate the training performance, the SAE/ T^2 model is fed with testing data. Since the SAE features are also derived from time measurements collected when the structural system is assumed to be under normal conditions, it is expected that the corresponding T^2 -statistic points are below the UCL value. If this statement is not verified, the SAE training parameters are readjusted (empirically), and a new model is created by returning to step 1.

The reference model definition is complete when the SAE/ T^2 model developed with the training data is able to correctly represent the current structural behavior by using the testing data. It means that the steps above shall be repeated until all or practically all reference data produce T^2 -statistic values bellow the UCL.

Finally, once the SAE/ T^2 reference model is properly established, newly acquired data may be classified as being from the normal or abnormal structural condition. SAE features and their T^2 -statistics are calculated from monitoring data and plotted in the control chart for comparing the reference structural responses with the newly collected ones. The entire anomaly-detection approach was developed using toolboxes and built-in functions available in Matlab® R2017a.

4 APPLICATION I: Z24 BRIDGE

The Z24 bridge was a highway overpass located at Canton Bern near Solothurn, in Switzerland, built in the early 1960s to connect Koppigen and Utzenstorf. This post-tensioned structure was 58m long composed of three continuous spans with 14m, 30m and 14m, supported on four

piers, as it can be seen in Figure 5. Due to the construction project of a new railway line underneath the highway, the bridge had to be demolished. Therefore, before the structure was completely knocked down, it was instrumented and gradually damaged [11].

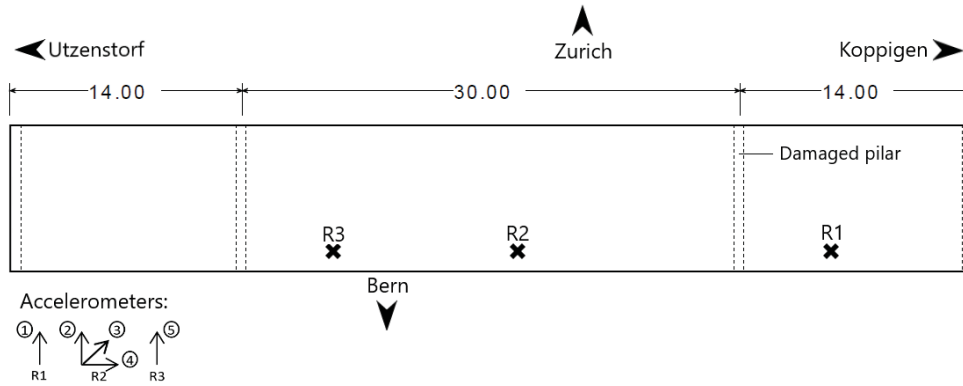


Figure 5: The Z24 bridge - Top view and experimental setup.

The present study evaluates the proposed SHM approach through dynamic signals of forced vibration tests conducted within two structural conditions: undamaged and damaged (settlement of pier - 40mm). Two vertical shakers were used to generate a flat force spectrum excitation with 3-30Hz bandwidth. In each damage scenario, nine dynamic tests were performed considering five accelerometers installed at three measurement points, R1, R2 and R3, as described in Figure 5. Vibration responses have 65535 data points, acquired at a frequency sampling of 100Hz for approximately 11 minutes. The structure's natural frequencies and the temperature variation - before and after the damage - are given in Table 1 for information purposes. The evolution of the eigenfrequencies, as well as the structural modes (see Figure 6), are derived from data of ambient vibration tests by the stochastic subspace identification method [11].

Structural condition	Temperature	1st natural frequency	2nd natural frequency	3rd natural frequency	4th natural frequency	5th natural frequency
Undamaged	17°C	3.92Hz	5.12Hz	9.93Hz	10.52Hz	12.69Hz
Damaged	29°C	3.86Hz	4.93Hz	9.74Hz	10.25Hz	12.48Hz

Table 1: Variation of the Z24 eigenfrequencies (values extracted from the work of De Roeck *et al.* (2000) [11]).

For this application, each accelerometer measurement is rearranged into 10-second signals, providing an input matrix $[1070 \times 1000]$: 1070 dynamic time histories (65 signals per accelerometer $\times$ 9 dynamic tests $\times$ 2 structural conditions = 1070) composed of 1000 data points ($100\text{Hz} \times 10\text{s} = 1000$). Henceforth, maintaining its chronological order, the input matrix is subdivided as follows:

- Reference data (undamaged state):
Training data $\rightarrow$ matrix $[450 \times 1000]$
Testing data $\rightarrow$ matrix $[135 \times 1000]$
- Monitoring data (damaged state) $\rightarrow$ matrix $[585 \times 1000]$

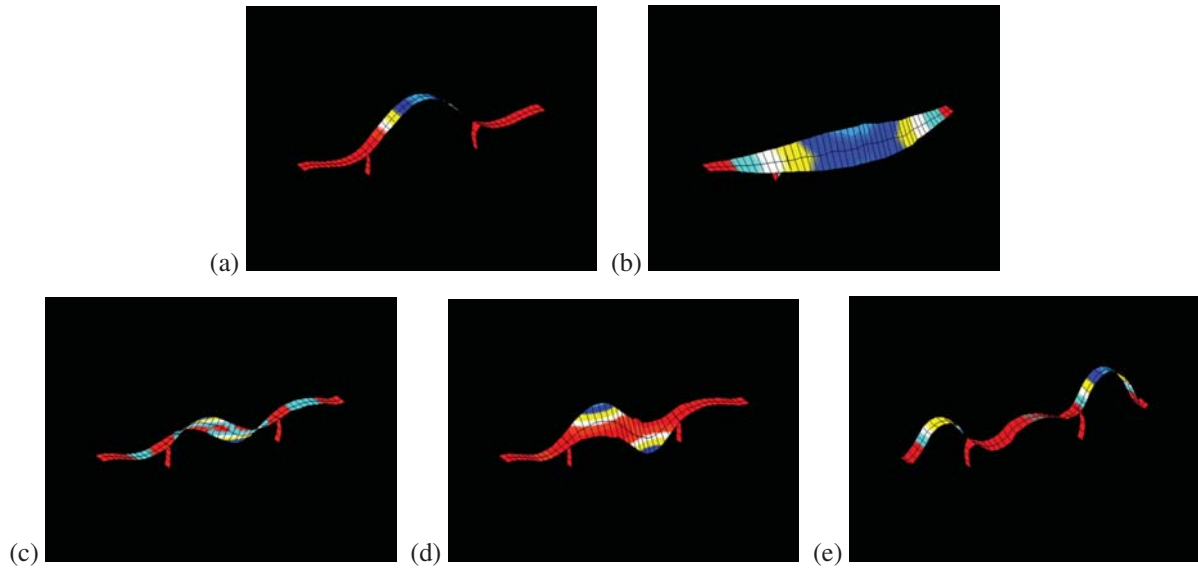


Figure 6: Reference mode shapes of the Z24 bridge - undamaged condition (draws extracted from the work of De Roeck *et al.* (2000) [11]). (a) 1st mode - symmetrical bending mode. (b) 2nd mode - transversal mode. (c) 3rd mode - combined torsional/antisymmetrical bending mode. (d) 4th mode - combined torsional/antisymmetrical bending mode. (e) 5th mode - symmetrical bending mode.

The anomaly detection approach is investigated for each accelerometer channel separately. For all analyses, the achieved SAE model was implemented employing one encoder and decoder layer constituted by 100 neurons (vector $\mathbf{h}$ with 100 components), which reduces the dimensionality of the problem from 1000 data points to 100 features. The other parameters were set to: the sparsity proportion (ρ) = 0.050; the sparsity regularization (β) = 4.000; the weight regularization (λ) = 0.001; the training function = Scaled Conjugate Gradient (SCG) optimization method [16] with gradient maximum value of 1.00×10^{-6} ; the encoder and decoder activation functions = logarithmic sigmoid and linear function, respectively; the error metric = mean square; and the maximum number of training epochs = 1000. The Hotelling's T^2 -statistic was calculated using data subgroups composed of 15 observations ($r = 15$) with the 100 SAE characteristics ($m = 100$). The UCL was estimated considering 30 subgroups of data ($s = 30 = 450$ training examples / 15 observations).

4.1 Results

The results of the proposed approach for the Z24 bridge are shown from Figure 7 to Figure 11. In total, 78 data subgroups were evaluated for each accelerometer channel, considering training (30 subgroups - blue points), testing (9 subgroups - green points), and monitoring data (39 subgroups - red points). In addition to the control chart plotted for one SAE/ T^2 model, another control chart was also constructed representing the T^2 -statistic for 30 different SAE models. In this case, the UCL is calculated considering the UCL's mean from the 30 models. This practice of examining the performance of various models is usual in ANN-based algorithms. It aims to identify that a specific data ordering (possible correlated samples) is not influencing the model performance.

By analyzing the results, a good performance was achieved for dynamic responses from accelerometers 1 (Figure 7) and 5 (Figure 11). In both cases, during the training and testing periods (undamaged condition), most T^2 -statistic values are below the limit line (a small number of false alarms). In the monitoring period (damaged condition), the T^2 -statistic abruptly

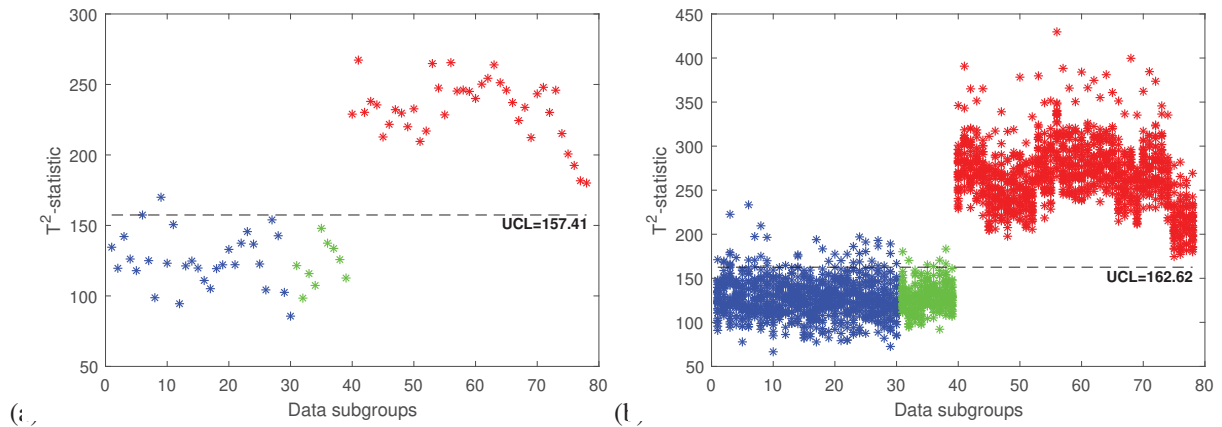


Figure 7: Anomaly detection approach - Output for the accelerometer 1. (a) Control chart for one SAE/ T^2 model. (b) Control for 30 SAE/ T^2 models.

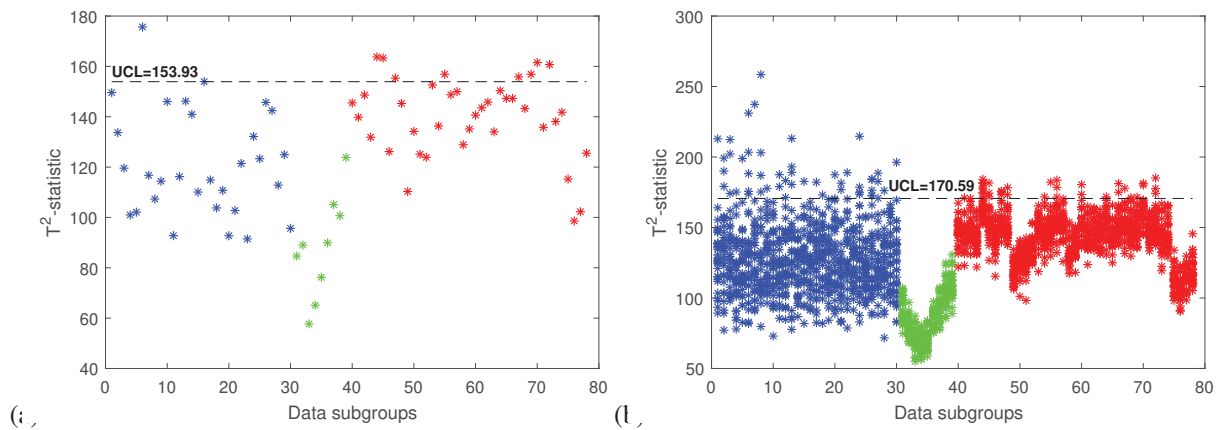


Figure 8: Anomaly detection approach - Output for the accelerometer 2. (a) Control chart for one SAE/ T^2 model. (b) Control for 30 SAE/ T^2 models.

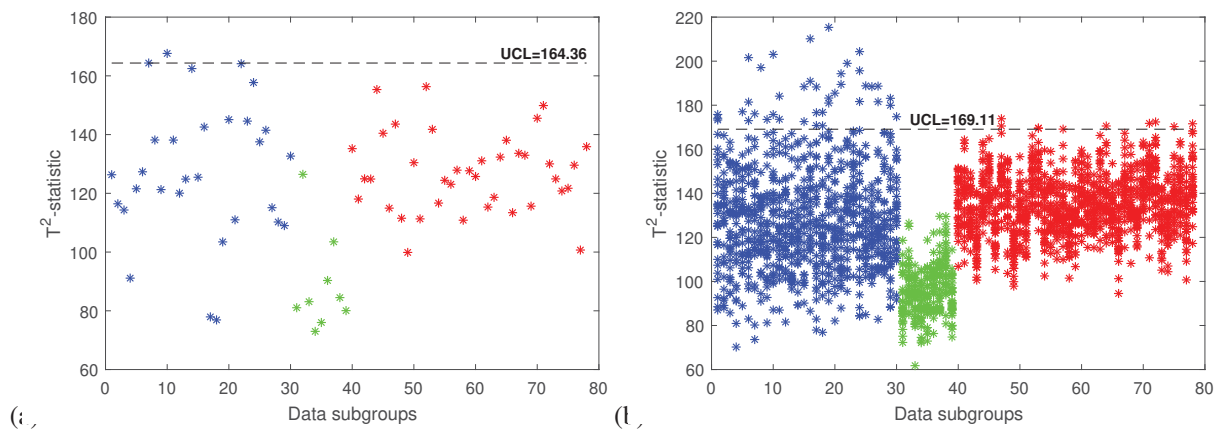


Figure 9: Anomaly detection approach - Output for the accelerometer 3. (a) Control chart for one SAE/ T^2 model. (b) Control for 30 SAE/ T^2 models.

exceeds the UCL value, laying outside of the in-control region and correctly indicating the presence of the structural anomaly. These outcomes may be attributed to the relation between the position of sensors and the structure's mode shapes. In general, higher natural frequencies are more affected in absolute terms by damage, as seen in Table 1. For this reason, since ac-

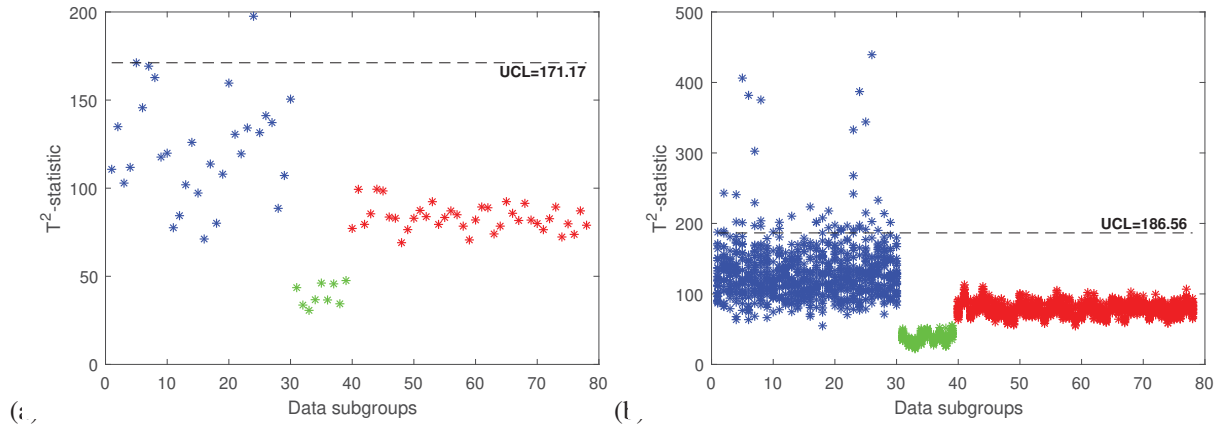


Figure 10: Anomaly detection approach - Output for the accelerometer 4. (a) Control chart for one SAE/ T^2 model. (b) Control for 30 SAE/ T^2 models.

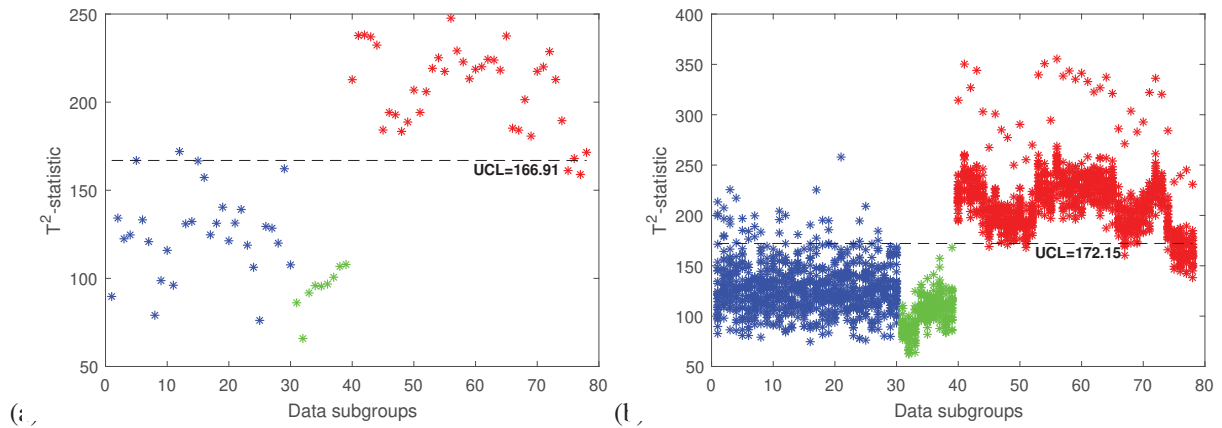


Figure 11: Anomaly detection approach - Output for the accelerometer 5. (a) Control chart for one SAE/ T^2 model. (b) Control for 30 SAE/ T^2 models.

celerometers 1 and 5 are in positions where the amplitudes related to higher magnitude modes are more significant (absolute difference of 0.27Hz for the 4th frequency and of 0.21Hz for the 5th frequency), they may have had better performance.

The same idea can be used to justify the T^2 -statistic of data collected by the accelerometer 2. Even though this channel is more affected by the first vibration mode, due to its location at the midspan, the absolute difference between the natural frequencies is small (0.06Hz), maybe not sufficient for the SAE/ T^2 model detects the damage occurrence. Regarding the sensors 3 and 4, they were positioned to capture transversal and longitudinal movements, respectively. Therefore, since the shakers generated components predominantly in the vertical direction and, the structure was considerably rigid on the longitudinal direction, it was expected that the control charts related to these measurement positions would have inconclusive results.

The Mean Square Error (MSE) between original and reconstructed dynamic signals for the 30 SAE models are exhibited in Table 2. It should be noted that the MSE values of training data are smaller than the MSE values of testing and monitoring data, as expected, since the first group of data was used to train the SAE, and the second and third ones were unknown by the created models. Figure 12 displays an example of the SAE reconstructed response in comparison with its respective original signal. Instead of perfectly reconstructs signals, the idea behind the SAE consists of modeling key features of data that are sensitive to structural novelties. Based on

the assumption that lower MSEs are evidence of better signal reconstruction, the slight upward trend in the T^2 -statistic verified for testing data of accelerometer 1 may be associated with these values. Table 2 reveals that the responses related to accelerometer 1 have a reconstruction error almost twice as large as the error found for the responses of accelerometer 5, the other channel with good results, where this trend does not appear.

Accelerometer	1	2	3	4	5
Training data	0.0129	0.0074	0.0057	0.0021	0.0076
Testing and monitoring data	0.0222	0.0120	0.0092	0.0026	0.0138

Table 2: MSE values between original structural responses and corresponding reconstructed signals for the 30 SAE models (units: m^2/s^4).

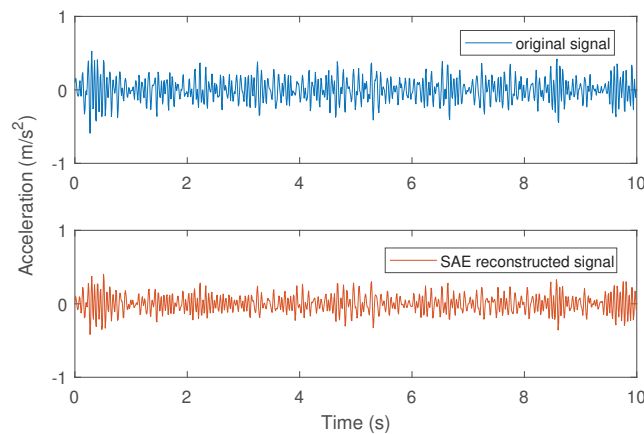


Figure 12: An original response of the Z24 bridge and its respective signal reconstructed by SAE.

5 CONCLUSIONS

The paper presented a structural anomaly detection approach based on Sparse Auto-Encoder and Shewhart T control chart. Features extracted by SAE models, directly from time-domain accelerations, were passed along as input variables of the control chart to detect the onset of abnormal behavior in structures. The developed method was exemplified using experimental data from dynamic tests performed on the Z24 bridge.

According to the obtained control charts, the T^2 -statistic calculated with the SAE extracted characteristics were efficient in detecting the two different structural states of the Z24 bridge. The results are in agreement with the respective structural scenarios since the damage imposed on the bridge was clearly identified. Nevertheless, more investigation is required to validate the proposed anomaly detection strategy. In particular, the next steps in this work include analyzing the temperature influence on SAE characteristics and applying the methodology in other structures.

Acknowledgements

The authors would like to thank UFJF (Universidade Federal de Juiz de Fora - Programa de Pós-Graduação em Modelagem Computacional), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) and Politecnico di Milano.

REFERENCES

- [1] S. W. Doebling, C. R. Farrar, M. B. Prime, A summary review of vibration-based damage identification methods, *Shock and Vibration Digest*, 30(2), 91–105, 1998.
- [2] A. Alvandi, C. Cremona, Assessment of vibration-based damage identification techniques, *Journal of Sound and Vibration*, 292(1-2), 179–202, 2006.
- [3] G. Marrongelli, C. Gentile, A. Saisi, Anomaly detection based on automated OMA and mode shape changes: application on a historic arch bridge, In *Proceedings of ARCH 2019, 9th International Conference on Arch Bridges*, 447–455, Springer, Cham, 2019.
- [4] H. Salehi, R. Burgueno. Emerging artificial intelligence methods in structural engineering. *Engineering Structures*, 171, 170–189, 2018.
- [5] R. P. Finotti, A. A. Cury, F. S. Barbosa, An SHM approach using machine learning and statistical indicators extracted from raw dynamic measurements, *Latin American Journal of Solids and Structures*, 16(2), e165, 2019.
- [6] R. Almeida Cardoso, A. Cury, F. Barbosa, C. Gentile, Unsupervised real-time SHM technique based on novelty indexes, *Structural Control and Health Monitoring*, 26, e2364, 2019.
- [7] I. Goodfellow, Y. Bengio, A. Courville. Deep Learning. MIT press, 2016.
- [8] C. S. N. Pathirage, J. Li, L. Li, H. Hao, W. Liu, R. Wang. Development and application of a deep learning–based sparse autoencoder framework for structural damage identification. *Structural Health Monitoring*, 18(1), 103–122, 2019.
- [9] Y. Bao, Z. Tang, H. Li, Y. Zhang, Computer vision and deep learning–based data anomaly detection method for structural health monitoring, *Structural Health Monitoring*, 18(2), 401–421, 2019.
- [10] D. Montgomery, *Introduction to Statistical Quality Control*, John Wiley & Sons, 2009.
- [11] G. De Roeck, B. Peeters, J. Maeck, Dynamic monitoring of civil engineering structures, *Computational Methods for Shell and Spatial Structures*, 2000.
- [12] J. C. Principe, N. R. Euliano, W. C. Lefebvre, *Neural and adaptive systems: fundamentals through simulations*, John Wiley & Sons, 2000.
- [13] P. Baldi, K. Hornik, Neural networks and principal component analysis: learning from examples without local minima, *Neural networks*, 2(1), 53–58, 1989.

- [14] Q. Meng, D. Catchpoole, D. Skillicom, P. J. Kennedy, Relational autoencoder for feature extraction, In *2017 International Joint Conference on Neural Networks (IJCNN)*, 364–371, IEEE, 2017.
- [15] A. Ng, Sparse autoencoder, *CS294A Lecture Notes*, 72, 1–19, 2011.
- [16] M. F. Møller, A scaled conjugate gradient algorithm for fast supervised learning, *Neural Networks*, 6(4), 525–533, 1993.