A Non-Parametric Mixed Learning Technique for Mitigating Environmental Effects on Structural Modal Frequencies

ALIREZA ENTEZAMI and STEFANO MARIANI

ABSTRACT

Health monitoring of civil structures via machine learning is a powerful approach to the early detection of any damage pattern. Besides structural damage, also environmental and operational variabilities are known to affect the inherent structural properties. Although the induced variations in the monitored properties are not harmful, their confounding influence can lead to economic and human losses. For these reasons, a novel unsupervised learning strategy is here proposed, aiming to properly account for the environmental effects on the structural modal frequencies. The offered solution is a non-parametric mixed learning strategy resting on hierarchical clustering, local non-negative matrix factorization, and Mahalanobis-squared distance (MSD). By means of the hierarchical clustering, training data consisting of modal frequencies relevant to the undamaged condition are subdivided into local clusters, which are then exploited in order to get rid of the environmental effects. The reconstructed data are finally used to train a non-parametric novelty detector based on the MSD, to obtain scores for decision making regarding the current state. To validate the proposed method, a set of modal frequencies of a steel arch bridge in its long-term monitoring has been considered; results show that the proposed methodology is effective in taking aside the environmental variability from the time history of the collected modal frequencies of the structure.

INTRODUCTION

Health monitoring of civil structures is of utmost importance in modern society, owing to their critical significance in our everyday life. It is therefore indispensable to avoid catastrophic events, such as partial or global collapses causing huge economic losses and human causalities. To prevent them, actions are needed to allow the prompt detection of a structural damage in its early stage. After having ascertained the occurrence of such a damage, one should next identify its location and quantify its severity, in order to make a decision on whether to repair or retrofit the damaged areas, or even replace the entire structural elements [1]. A structural health monitoring (SHM) project can be thus implemented to feature three main stages of early damage detection [2], damage localization [3, 4], and damage quantification [5].

Alireza Entezami, Stefano Mariani, Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy. Email: alireza.entezami@polimi.it

The prerequisites for the aforementioned stages of SHM include: (i) select a measurement technique; (ii) deploy sensors and record relevant data; (iii) define meaningful information brought by the measured data; and (iv) implement the SHM procedure via model-based or data-driven techniques. The first two steps pertain to the measurement science and sensing technologies [6]. The third step is termed feature extraction and plays a significance role in SHM. The fourth step exploits the information from the previous stage to properly implement the SHM strategy. Model-based SHM techniques usually rest on a numerical (finite element) model of the monitored structure, supplemented by the measured experimental data for model updating purposes [7]. In contrast, data-based SHM techniques take advantage of measured data only, in principle without any numerical model to be exploited for damage assessment. Hybrid solutions, trying to take advantage of the strengths of both the aforementioned approaches, have been recently proposed; interested readers can find thorough discussions in, e.g. [8, 9].

One of the widely-used and promising feature extraction methods for the SHM of large-scale civil structures is the operational modal analysis (OMA), which allows identifying properties like modal frequencies, mode shapes, and damping ratios when excitation forces are not recorded [10]. Due to the benefits of the data-based SHM, as compared to the model-based one, machine learning (ML) has become the main tool to perform data analytics [11]. ML is a branch of artificial intelligence that aims at developing a computational and intelligent model, by exploiting training data to solve complex problems. Supervised learning and unsupervised learning are two different strategies in ML. In the SHM realm, fully labeled data are needed when a supervised learning strategy is adopted, including the features related to the damaged structural state; on the contrary, an unsupervised learning strategy exploits unlabeled data regarding the undamaged state only. Because the (un)availability of fully labeled data for real-world SHM, most of the developed proposals have focused so far on unsupervised learning schemes, especially when early damage detection is the target [12, 13].

A cooperative integration of OMA and unsupervised learning can lead to a promising method for early damage detection, see [14]. In this way, a set of modal frequencies identified from measured vibration data are exploited, and an OMA algorithm is defined to provide the training data; such a dataset is then fed into an unsupervised learner. When modal frequencies related to the unknown state of the structure are obtained from the current measurements, those are fed on their own into the trained digital model for decision-making: any deviation from the formerly set baseline is therefore indicative of damage occurrence. A big challenge in such a strategy is that structural damage is not the only cause of variations of the modal frequencies: in real-world scenarios, especially when bridges are involved, environmental and operational conditions affect the collected data too [15]. These conditions are in fact able to alter inherent physical properties of the structure, such as mass and stiffness, and thus change the structural response in a way similar to, or undistinguishable from what induced by damage. Accurate and sensitive data analytics solutions have to be adopted in the analysis. Hence, whenever the confounding influences by the environmental and operational variabilities show up, false positive and false negative errors can show up; these errors are finally related to possible economic and human losses, so that it becomes indispensable to get rid of such variability conditions [16].

Locally unsupervised mixed learning methods have been recently proposed to address this problem. These methods avoid to handle the training dataset as a whole, and take instead advantage of local information (namely, related to a portion of the sampling data) on the training space, and combinations of different unsupervised learning methods to address the issue linked to the confounding influences. In this regard, Entezami et al. [17] proposed an unsupervised meta-learning method that consists of a data segmentation via spectral clustering, and local damage detection based on the Mahalanobis-squared distance (MSD). Daneshvar et al. [18] developed a locally unsupervised hybrid learning method via Gaussian mixture to define the needed local information, and a discriminative reconstruction-based dictionary learning model to remove the environmental effects. Entezami et al. [19] also proposed a multi-task unsupervised learning method based on the density-based spatial clustering of applications with noise (DBSCAN) to remove outliers from the training data, spectral clustering to set the local information, and local empirical measures for anomaly detection.

In spite of the achieved results and of the noteworthy performance in removing environmental and operational variabilities, a major limitation of these techniques is related to their parametric nature. In other words, some stages of these methodologies rest on parametric algorithms, whose hyperparameters must be estimated/set beforehand. Therefore, the main objective of this work is the proposal of a new unsupervised mixed learning method, within a non-parametric framework. The proposed method consists of three non-parametric algorithms: hierarchical clustering, to provide local subsets of the training data; local non-negative matrix factorization, to develop a subspace learning algorithm; and MSD, to allow novelty detection. To evaluate the performance of the proposed method, a set of modal frequencies of a bridge in its undamaged state is considered. Results show that the proposed method can effectively remove the environmental effects from the modal frequencies, so that reliable information can be next moved onto the decision-making stage.

PROPOSED METHOD

HIERARCHICAL AGGLOMERATIVE CLUSTERING

A hierarchical clustering method is a non-parametric approach to subdivide unlabeled data into a hierarchy of clusters, called *dendrogram*. By means of this procedure, it is possible to represent the relationships between data in the clustering process, where clusters are arranged similar to a tree. Data clustering can be based on *agglomerative* (bottom-up) or *divisive* (top-down) algorithms: the former ones start by taking singleton clusters, each one containing one datum only, and keep merging two clusters at a time to build a *bottom-up* hierarchy; the latter ones start instead from all the data merged in a single macro-cluster, and then split it continuously into two groups in a *top-down* hierarchy. The main advantage of hierarchical clustering methods, as compared to the partition-based ones like k-means, k-medoids, fuzzy c-means, or Gaussian mixture, is the lack of a constraint in setting the number of clusters [20].

The implementation of the hierarchical agglomerative clustering consists of three steps. Generally, the process starts by clustering individual data points in a singleton cluster, which is then continuously merged based on similarity, until when it forms one big cluster containing all the data points. First, a dissimilarity (distance) measure is defined and used to compute the distance between any pair of data points in the training data, to generate a distance matrix for all the data points to be represented at the bottom of the dendrogram. Second, the closest sets of the clusters are iteratively merged by using a *linkage* function to define the distance between two clusters, and the distance matrix is accordingly updated. The said linkage function exploits the distance measures generated in the first step, to determine the proximity of points to each other. Commonly used algorithms for linkage are based on single, complete, average, centroid, median, weighted, of Ward's functions. As data points are paired into binary clusters, the newly formed clusters are grouped into larger clusters until a hierarchical tree is formed. Third, a strategy is defined to set where the hierarchical tree is cut. One can then prune branches off the bottom of the hierarchical tree and assign all the objects below each cut to a single cluster, leading to a partition of the data [20].

Let us suppose that $\mathbf{X} \in \mathbb{R}^{p \times n}$ is the training matrix, containing n vectorial feature points each of p variables. In concrete terms and in relation to the modal-based SHM of civil structures, $\mathbf{X}$ is a matrix of n modal frequency samples from p identified modes, so that the mentioned samples represent the training data. Using the weighted average distance as the linkage function, the hierarchical agglomerative clustering subdivides $\mathbf{X}$ into the two clusters $\mathbf{C}_1 \in \mathbb{R}^{p \times r_1}$ and $\mathbf{C}_2 \in \mathbb{R}^{p \times r_2}$, where r_1 and r_2 denote the number of clustered features respectively belonging to the first and second clusters. The main reasoning behind handling two clusters only, departs from the real structural conditions. In fact, when the structure keeps its undamaged state, two types of features can be drawn: one is related to the undamaged state without the environmental and operational effects; the other pertains to the same state, with the environmental and operational effects instead included. Simply speaking, the hierarchical agglomerative clustering allows grouping the available undamaged features into the two groups.

LOCAL NON-NEGATIVE MATRIX FACTORIZATION

Non-negative matrix factorization (NMF) is a dimensionality reduction technique based on a low-rank approximation of the feature space [21]. The major purpose of this technique is the factorization of a matrix into two non-negative matrices. By means of these matrices, having defined a factor rank, it is next possible to reconstruct the entire original data. As reconstruction-based, output-only data normalization is the main unsupervised learning method to remove the environmental and operational variability conditions, NMF can be used to that purpose. Given the clustered feature sets $\mathbf{C}_1$ and $\mathbf{C}_2$ as well as a rank factor f, NMF provides the non-negative matrices $\mathbf{W}_l \in \mathbb{R}^{p \times f}$ and $\mathbf{H}_l \in \mathbb{R}^{f \times r_l}$, being $l = 1{,}2$, to minimize the Frobenius norm $\|\mathbf{C}_l - \mathbf{W}_l\mathbf{H}_l\|_F^2$ handled as the root mean square error (RMSE). Moreover, the rank factor f is simultaneously set by minimizing the said RMSE under these two conditions: (i) $1 \leq f \leq (p - 1)$ when $p \leq r_l$ and (ii) $1 \leq f \leq (r_l - 1)$ when $p > r_l$. With the non-negative matrices $\mathbf{W}_l$ and $\mathbf{H}_l$, the clustered features can be reconstructed via:

$$\hat{\mathbf{C}}_l = \mathbf{W}_l\mathbf{H}_l \tag{1}$$

The reconstructed clusters $\hat{\mathbf{C}}_1 \in \mathbb{R}^{p \times r_1}$ and $\hat{\mathbf{C}}_2 \in \mathbb{R}^{p \times r_2}$ then represent the output of the second step of the proposed method.

MAHALANOBIS-SQUARED DISTANCE

Novelty detection is the unsupervised learning stage for feature analysis. In fact, the unsupervised learner or novelty detector is trained by the means of unlabeled data, to finally make a decision on the test data regarding the current, unknown state. If the test data belongs to an abnormal situation characterized by a damaged structural state, the novelty detector should immediately detect it. In this work, a non-parametric novelty detector based on the MSD is adopted to investigate the evolution of the model output, namely of the novelty scores, and also understand whether the environmental and operational effects are mitigated.

Given the reconstructed features in $\{\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2\}$, the parameters handled by the MSD-based novelty detector are the mean vectors $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2\}$ and the relevant covariance matrices $\{\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2\}$ of the reconstructed features. Each mean vector gathers p elements, while each covariance matrix has dimensions $p \times p$. To define the novelty indices, each training feature is exploited in the following form:

$$\mathrm{D}_m(\mathbf{x}_i) = \min\left(\left(\mathbf{x}_i - \boldsymbol{\mu}_j\right)^T \boldsymbol{\Sigma}_j^{-1} \left(\mathbf{x}_i - \boldsymbol{\mu}_j\right)\right) \tag{2}$$

where $i = 1,..,n$ and $j = 1,2$. In the testing (inspection) period, the vectors $\mathbf{x}_i$ are substituted in Eq. (2) by the new received test vectors $\mathbf{z}_k$, being $k = 1,..,m$, as follows:

$$\mathrm{D}_m(\mathbf{z}_k) = \min\left(\left(\mathbf{z}_k - \boldsymbol{\mu}_j\right)^T \boldsymbol{\Sigma}_j^{-1} \left(\mathbf{z}_k - \boldsymbol{\mu}_j\right)\right) \tag{3}$$

For m test features, m novelty indices can be obtained as well. Accordingly, if the test features belong to the damaged state of the structure, there should be a clear distance or difference between the novelty indices relevant to these features and the novelty indices obtained for the training features, which are related to the undamaged condition. On the contrary, in case of test features belonging to the undamaged state, the corresponding novelty indices should be similar to those pertinent to training.

EXPERIMENTAL INVESTIGATION: THE KW51 BRIDGE

To assess the performance of the proposed method, the steel arch bridge called KW51 [22] is here considered, see Figure 1. This structure is a railway bridge that connects Leuven and Brussels in Belgium, along the railway line L36N. The bridge is 115 m long and 12.4 m large. Since 2 October 2018, the bridge is equipped with an SHM system consisting in vibration and environmental sensors, to acquire acceleration time histories and environmental data too. The time histories of the modal frequencies between 2 October 2018 and 15 May 2019 are here considered, with reference to the normal condition; therefore, the effects of environmental factors on those modal frequencies are evaluated next.

An automated OMA was implemented by Maes and Lombaert [23], to identify the modal properties of the bridge including the relevant natural frequencies. The automated OMA allowed to yield information in time concerning 14 vibration modes. To avoid issues related to missing information regarding some of those modal features, only the modes 6, 10, 12, and 13 are here considered.

Figure 1. The KW51 Bridge.

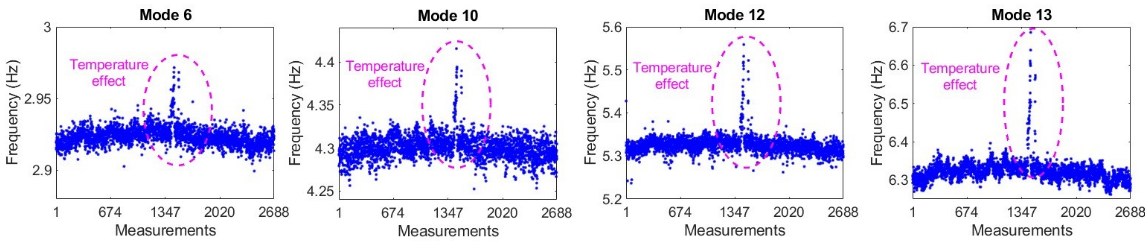

Figure 2. Time evolution of the modal frequencies of the KW51 Bridge.

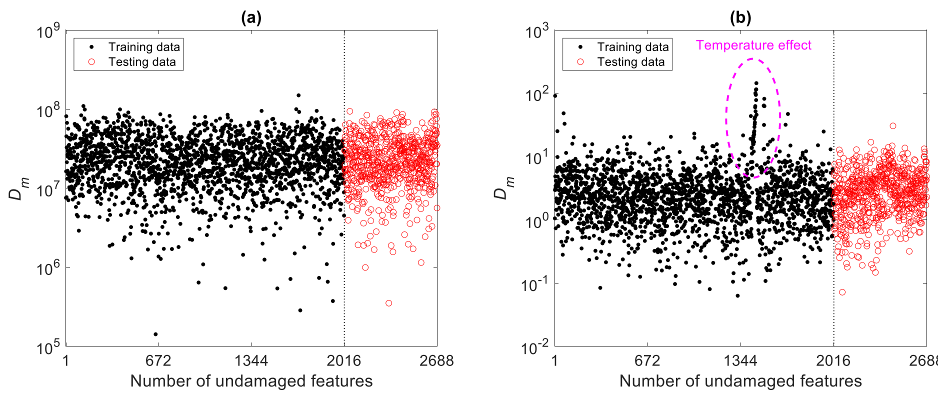

Figure 3. Evolution of novelty indices related to the training (black points) and testing (red points) data, as provided by: (a) the proposed method, (b) the direct use of the MSD-based novelty detection without data clustering and feature reconstruction.

The total number of samples in the time histories of the modal frequencies of the undamaged state are 2688: Figure 2 shows the said histories of the dynamic features. As can be seen, environmental effects show up as sudden jumps in the range of samples 1345-2017. Since such variability in any novelty detector provides a footprint similar to damage, the aim of the proposed methodology is to remove the aforementioned sudden jumps in the time histories. Therefore, the existence of two variability forms including the usual and sudden ones, provides a rationale to ascertain the accuracy of the two clusters adopted for the hierarchical clustering.

All the features in the time series of Figure 3 are divided into training and testing datasets, with a ratio of 75% : 25%. In the first step, the training data $\mathbf{X} \in \mathbb{R}^{4\times2016}$ are partitioned into the two clusters $\mathbf{C}_1\in\mathbb{R}^{4\times1979}$ and $\mathbf{C}_1\in\mathbb{R}^{4\times37}$, so resulting into r_1=1979 and r_2=37. The clustered features are then exploited by the local NMF methodology, to reconstruct them. Regarding this stage of the analysis, on the basis of the first condition needed for the rank factor, the parameter f can take a value between 1 and 3; its optimal value is selected as the one providing the minimum RMSE value. The result is shown in Figure 3(b), where the optimal rank for the first and second clusters are identified as 3 and 2, respectively. The clusters are then reconstructed to be adopted in the MSD-based novelty detection stage. Figure 3(a) shows the corresponding time evolution of the novelty indices related to the training and testing data points. For a comparison, the same process is repeated by using the MSD-based novelty detection without prior data clustering and feature reconstruction, so by directly handling the training and testing matrices via the MSD. The results of this latter strategy are displayed in Figure 3(b). As can be seen, the proposed method succeeds in removing the environmental effects caused by the freezing air temperature: sudden jumps in the modal frequencies are no longer visible in the novelty indices furnished by the proposed method. The other way around, the said jump is still clearly visible in Figure 3(b). It can be thus concluded that the proposed method, thanks to its mixed and non-parametric properties, is perfectly able to deal with the confounding influences induced by the environmental variability. It also turns out to be superior to the classical MSD-based novelty detection one.

CONCLUSIONS

In this paper, a non-parametric mixed learning technique has been proposed, within an unsupervised learning strategy, to address the confounding effects caused by environmental and/or operational variability. The proposed method consists of the three steps of data clustering via the hierarchical agglomerative clustering, clustered feature reconstruction through local NMF, and novelty detection based on the MSD metric.

The time histories of the modal frequencies of the KW51 Bridge have been adopted to evaluate the performance and effectiveness of the proposed method. The obtained results have demonstrated that the proposed method can effectively remove the environmental variability, for instance showing up as sudden jumps in the structural vibration frequencies caused by a freezing air temperature. This method has been also shown to outperform the classical MSD-based novelty detection technique. This conclusion has confirmed the proposed unsupervised mixed learning strategy, if based on an ad-hoc defined data clustering to deal with the said variability effects, can properly lead to robust SHM strategies.

REFERENCES

1. Li, H. 2020. “Structural assessment of concrete cable-stayed bridge after replacement of closure segment: The service stage,” *Pract. Periodical Struct. Des. Constr.*, 25(3): 04020023.
2. Giordano, P., Z. Turksezer, M. Previtali, and M. Limongelli. 2022. “Damage detection on a historic iron bridge using satellite DInSAR data,” *Struct. Health Monit.*, 21(5): 2291-2311.
3. Entezami, A., H. Sarmadi, and C. De Michele. 2022. “Probabilistic damage localization by empirical data analysis and symmetric information measure,” *Meas.*, 198: 111359.
4. Giordano, P. F., S. Quqa, and M. P. Limongelli. 2021. “Statistical Approach for Vibration-Based Damage Localization in Civil Infrastructures Using Smart Sensor Networks,” *Infrastructures*, 6(2): 22.
5. Entezami, A., H. Shariatmadar, and A. Karamodin. 2019. “Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods,” *Struct. Health Monit.*, 18(5-6): 1416-1443.

6. Wang, M. L., J. P. Lynch, and H. Sohn. 2014. *Sensor Technologies for Civil Infrastructures: Applications in Structural Health Monitoring*. Woodhead Publishing (Elsevier).
7. Cabboi, A., C. Gentile, and A. Saisi. 2017. “From continuous vibration monitoring to FEM-based damage assessment: Application on a stone-masonry tower,” *Constr. Build. Mater.*, 156: 252-265.
8. Entezami, A., H. Sarmadi, and B. Saeedi Razavi. 2020. “An innovative hybrid strategy for structural health monitoring by modal flexibility and clustering methods,” *J. Civ. Struct. Health Monit.*, 10(5): 845-859.
9. Kim, Y., J.-c. Park, and S. Shin. 2018. “Development of a hybrid SHM of cable bridges based on the mixed probability density function,” *J. Civ. Struct. Health Monit.*, 8(4): 569-583.
10. García-Macías, E. and F. Ubertini. 2020. “MOVA/MOSS: Two integrated software solutions for comprehensive Structural Health Monitoring of structures,” *Mech. Syst. Sig. Process.*, 143: 106830.
11. Flah, M., I. Nunez, W. Ben Chaabene, and M. L. Nehdi. 2021. “Machine Learning Algorithms in Civil Structural Health Monitoring: A Systematic Review,” *Arch. Comput. Methods Eng.*, 28(4): 2621-2643.
12. Entezami, A., H. Shariatmadar, and S. Mariani. 2020. “Early damage assessment in large-scale structures by innovative statistical pattern recognition methods based on time series modeling and novelty detection,” *Adv. Eng. Softw.*, 150: 102923.
13. Entezami, A., H. Sarmadi, B. Behkamal, and S. Mariani. 2021. “Health Monitoring of Large-Scale Civil Structures: An Approach Based on Data Partitioning and Classical Multidimensional Scaling,” *Sensors*, 21(5): 1646.
14. Chalouhi, E. K., I. Gonzalez, C. Gentile, and R. Karoumi. 2017. “Damage detection in railway bridges using Machine Learning: application to a historic structure,” *Procedia Eng.*, 199: 1931-1936.
15. Gentile, C., M. Guidobaldi, and A. Saisi. 2016. “One-year dynamic monitoring of a historic tower: damage detection under changing environment,” *Meccanica*, 51(11): 2873-2889.
16. Wang, Z., D.-H. Yang, T.-H. Yi, G.-H. Zhang, and J.-G. Han. 2022. “Eliminating environmental and operational effects on structural modal frequency: A comprehensive review,” *Struct. Contr. Health Monit.*, 29(11): e3073.
17. Entezami, A., H. Sarmadi, and B. Behkamal. 2023. “Long-term health monitoring of concrete and steel bridges under large and missing data by unsupervised meta learning,” *Eng. Struct.*, 279: 115616.
18. Daneshvar, M. H., H. Sarmadi, and K.-V. Yuen. 2023. “A locally unsupervised hybrid learning method for removing environmental effects under different measurement periods,” *Meas.*, 208: 112465.
19. Entezami, A., H. Sarmadi, B. Behkamal, and C. De Michele. 2023. “On continuous health monitoring of bridges under serious environmental variability by an innovative multi-task unsupervised learning method,” *Struct. Infrastruct. Eng.*, In Press: 1-19.
20. Aggarwal, C. C. and C. K. Reddy. 2016. *Data Clustering: Algorithms and Applications*. CRC Press, Boca Raton, Florida, United States.
21. Berry, M. W., M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons. 2007. “Algorithms and applications for approximate nonnegative matrix factorization,” *Comput. Stat. Data Anal.*, 52(1): 155-173.
22. Maes, K. and G. Lombaert. 2021. “Monitoring Railway Bridge KW51 Before, During, and After Retrofitting,” *J. Bridge Eng.*, 26(3): 04721001.
23. Maes, K., L. Van Meerbeeck, E. P. B. Reynders, and G. Lombaert. 2022. “Validation of vibration-based structural health monitoring on retrofitted railway bridge KW51,” *Mech. Syst. Sig. Process.*, 165: 108380.