

On the advances of automatic modal identification for SHM

Rharã Cardoso¹, Alexandre Cury^{2,a}, and Flávio Barbosa²

¹Federal University of Ouro Preto, Department of Civil Engineering, Ouro Preto, Brazil.

²Federal University of Juiz de Fora, Department of Applied Computational Mechanics, Juiz de Fora, Brazil.

Abstract. Structural health monitoring of civil infrastructures has great practical importance for engineers, owners and stakeholders. Numerous researches have been carried out using long-term monitoring, for instance the Rio-Niterói Bridge in Brazil, the former Z24 Bridge in Switzerland, the Millau Bridge in France, among others. In fact, some structures are monitored 24/7 in order to supply dynamic measurements that can be used for the identification of structural problems such as the presence of cracks, excessive vibration, damage or even to perform a quite extensive structural evaluation concerning its reliability and life cycle. The outputs of such an analysis, commonly entitled modal identification, are the so-called modal parameters, i.e. natural frequencies, damping ratios and mode shapes. Therefore, the development and validation of tools for the automatic identification of modal parameters based on the structural responses during normal operation is fundamental, as the success of subsequent damage detection algorithms depends on the accuracy of the modal parameters estimates. The proposed methodology uses the data driven stochastic subspace identification method (SSI-DATA), which is then complemented by a novel procedure developed for the automatic analysis of the stabilization diagrams provided by the SSI-DATA method. The efficiency of the proposed approach is attested via experimental investigations on a simply supported beam tested in laboratory and on a motorway bridge.

1 Introduction and state-of-the-art

Firstly, for the sake of clarity, one should clearly distinguish modal parameter estimation (MPE) from modal tracking. The former consists in the estimation of modal parameters from a single record of measured data and the latter corresponds to tracking the evolution of the modal parameters of a structure through repeated MPE. This paper focuses on the MPE process, which means that structural responses are evaluated independently. Therefore, what would naturally be the second step, related to modal tracking, will not be considered here.

Thus, concerning the parametric system identification techniques, one can enumerate the following two steps for the MPE:

1. Computation of the modes estimates for several orders of the parametric model, by fitting them to the response series;
2. Interpretation of the modes estimates data by detecting spurious modes and selection of the physical ones.

In fact, without the automation of the MPE process, a lot of user interaction over a large amount of measured vibration data would be required, leading to an impossibility of practical applications such as the health monitoring of crucial infrastructures. Therefore, especially over the last decade, several methods aiming the reduction of the number of user-defined parameters are being developed [1-4], some of which reached the

complete elimination of all manually specified parameters: these methods are usually labelled as fully automated.

For instance, in 2012, Reynders *et al.* [1] published a fully automated approach for the interpretation of a stabilization diagram. It was cleverly based on clustering techniques through three stages: a diagram pre-cleaning by means of a classification of all modes into two categories (possible physical or certainly spurious); a hierarchical clustering of the possible physical ones to group them together (automatic detection of vertical lines on diagram); and a final classification of the formed clusters into a physical or spurious condition. No user-defined parameter is needed whatsoever in the entire process.

On the other hand, the structural dynamic monitoring practice has been showing that, even the most well-crafted automated process of MPE has to be, at a first use, manually checked or tuned to fit a possible specific behaviour of the analysed structure in order to obtain satisfactory results. As a consequence, the development focus of new methods does not need to be on the elimination of all the user-defined parameters (fully automation), but on the construction of a robust automated approach with some few easy configurable user settings, which are defined once before starting the repetitive automated process itself. Thereby, this paper presents an automated method for interpreting the stabilization diagram with just two easily user-defined

^a Corresponding author: alexandre.cury@ufjf.edu.br

parameters. In other words, the main objective here is all about the second step of the MPE previously stated.

Naturally, in order to obtain the data related to the stabilization diagram, one must first carry out the first step. This is done by the parametric system identification algorithm. Thus, in this paper, one utilizes the data-driven Stochastic Subspace Identification method (SSI-DATA), which is briefly discussed in Section 2.

A benchmark method for the automatic interpretation of the stabilization diagram according to Magalhaes [2] is treated in section 3. Moreover, this section also presents the novel approach developed in this paper. Section 4 shows the results and comparisons between the methods subjected to practical applications. Finally, in section 5, some conclusions about the results are drawn.

2 Data-driven stochastic subspace identification

Since the work published by Van Overschee and De Moor in 1996 [5], the application of data-driven SSI algorithms to determine the modal parameters of structures has been quite spread over the ambient vibration tests. They compute state space models from output data (structure response) and can have its problem stated as following.

Given s measurements of the output $\mathbf{y}_k \in \mathbb{R}^l$ generated by the unknown stochastic system of order n :

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \quad (1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \quad (2)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector at a discrete time instant k , \mathbf{w}_k and $\mathbf{v}_k \in \mathbb{R}^n$ are zero mean, white noise vector sequences with covariance matrix (at instants p and q with δ meaning the Kronecker delta):

$$\mathbb{E} \left[\begin{pmatrix} \mathbf{w}_p \\ \mathbf{v}_p \end{pmatrix} \begin{pmatrix} \mathbf{w}_q^T & \mathbf{v}_q^T \end{pmatrix} \right] = \begin{pmatrix} \mathbf{Q} & \mathbf{S}_v \\ \mathbf{S}_v^T & \mathbf{R} \end{pmatrix} \delta_{pq} \geq 0 \quad (3)$$

Determine:

- The order n of the unknown system;
- The system matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{C} \in \mathbb{R}^{l \times n}$ with the covariance matrices $\mathbf{Q} \in \mathbb{R}^{n \times n}$, $\mathbf{S}_v \in \mathbb{R}^{n \times l}$ and $\mathbf{R} \in \mathbb{R}^{l \times l}$.

For this purpose, the SSI-DATA algorithm must consist of the following 8 steps:

1. Organize the response series \mathbf{y} (the only measured quantity) in a Hankel matrix \mathbf{Y} which can be partitioned into a past measurements matrix \mathbf{Y}_p and a future measurements matrix \mathbf{Y}_f ;
2. Calculate the orthogonal projection \mathbf{P} of the row space of \mathbf{Y}_f on the row space of \mathbf{Y}_p :

$$\mathbf{P} = \mathbf{Y}_f \mathbf{Y}_p^T \left(\mathbf{Y}_p \mathbf{Y}_p^T \right)^\dagger \mathbf{Y}_p \quad (4)$$

where \bullet^\dagger denotes the Moore-Penrose pseudo-inverse of the matrix \bullet ;

3. Compute the Singular Value Decomposition (SVD) of the projection weighted by matrices \mathbf{W}_1 and \mathbf{W}_2 :

$$\mathbf{W}_1 \mathbf{P} \mathbf{W}_2 = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (5)$$

4. Determine the system order n by inspecting the rank of \mathbf{S} (the number of non-zero singular values) and partition the SVD accordingly to obtain \mathbf{U}_1 and \mathbf{S}_1 :

$$\mathbf{U} \mathbf{S} \mathbf{V}^T = (\mathbf{U}_1 \quad \mathbf{U}_2) \begin{pmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} \quad (6)$$

5. Compute the extended observability matrix $\mathbf{\Gamma}$:

$$\mathbf{\Gamma} = \mathbf{W}_1^{-1} \mathbf{U}_1 \mathbf{S}_1^{1/2} \quad (7)$$

6. Determine, in a least square sense, the system (dynamic) matrix \mathbf{A} from $\mathbf{\Gamma}$ as:

$$\mathbf{A} = \underline{\mathbf{\Gamma}}^\dagger \bar{\mathbf{\Gamma}} \quad (8)$$

where $\underline{\mathbf{\Gamma}}$ denotes $\mathbf{\Gamma}$ without the last l (number of outputs) rows and $\bar{\mathbf{\Gamma}}$ denotes $\mathbf{\Gamma}$ without the first l rows.

7. Determine \mathbf{C} as the first l rows of $\mathbf{\Gamma}$.
8. Finally, knowing \mathbf{A} and \mathbf{C} , one can calculate the modal frequencies ω_i , the modal damping ratios ξ_i and the corresponding modal shapes ϕ_i . This calculation is made through the eigenvalue decomposition of matrix \mathbf{A} as the following statements show:

$$\lambda_i = \frac{\ln(\mu_i)}{\Delta t} \quad (9)$$

$$\omega_i = |\lambda_i| \text{ (rad / s)} \rightarrow f_i = \frac{|\lambda_i|}{2\pi} \text{ (Hz)} \quad (10)$$

$$\xi_i = \frac{-\text{Re}(\lambda_i)}{|\lambda_i|} \times 100 \text{ (\%)} \quad (11)$$

$$\phi_i = \mathbf{C} \mathbf{g}_i \quad (12)$$

where μ_i is the i^{th} discrete-time eigenvalue of \mathbf{A} and \mathbf{g}_i is its corresponding eigenvector. The eigenvalues of \mathbf{A} are also called *poles*. They are complex numbers and appear mostly in conjugated pairs. Only those poles which have a conjugated pair and have positive imaginary component are taken into account for the equations (9) to (12). Therefore, up to $n/2$ modes are expected to be estimated. The $\text{Re}(\bullet)$ is the symbol for the real part of \bullet and λ_i denotes the continuous-time eigenvalue of \mathbf{A} .

These 8 steps summarize the SSI-DATA routine which is, essentially, a fit of a state space model to the temporal output data by means of the geometric projection of the row space of the future measurements on the past measurements.

There are several valid sets of weighting matrices \mathbf{W}_1 and \mathbf{W}_2 to be used in the equations (5) and (7). They determine the state space basis in which the model will be identified. Special choices of these matrices correspond to algorithms described in the literature: Principal Component (PC), Unweighted Principal Component (UPC) and Canonical Variate Analysis (CVA). The PC algorithm is the one used in the applications shown in section 4.

It is necessary to point out that when one works with real structure response data, it is very hard to determine the system order through inspection of singular values like stated in the fourth step. This happens because there are no zeros at all, but an asymptotic smooth variation towards zero (\mathbf{S} is actually full rank). Thus, a threshold would be necessary to consider a singular value as being null. Therefore, instead of creating manners for defining this threshold, the system order n is readily given as an input to the algorithm, which solves the problem but creates the need for a further check whether this arbitrary order is able or not to accurately represent the actual data. One of the tools for this check is, exactly, the well-known stabilization diagram, where the modes estimates are plotted for various different system orders.

3 Automation methods for stabilization diagram interpretation

As previously mentioned, the mode estimates (usually presented in a stabilization diagram) are supposed to be analysed in order to get physical modes apart from the spurious (numerical) ones. The latter inevitably exist among the data, because they are a natural consequence of the parametric model attempt to better fit the response data.

Naturally, due to the large amount of collected data from online structural monitoring, a manual interpretation of modes estimates data is unfeasible. Thus, the obvious solution is the automation of such a process. For this purpose, in this paper, one proposes a novel approach, which was mainly inspired on the automation methodology proposed by Magalhães in his PhD thesis [3]. His method is considered as a benchmark.

Both methods are suitable for the analysis of the outputs generated by any parametric identification technique that produces modes estimates (frequencies, damping ratios and modal shapes) for several orders.

3.1 Benchmark methodology

In his approach, no pre-filter is applied to the modes estimates data, which means that the stabilization diagram would be still full of certainly spurious modes, i.e. like those with extremely high or negative damping ratios. This method is based on a hierarchical clustering algorithm.

Firstly, the dissimilarity (distance) measure between all pairs of estimated modes is calculated. The adopted 'metric' for calculation of such a distance is:

$$d_{i,j} = \left| \frac{f_i - f_j}{f_j} \right| + (1 - \text{MAC}_{i,j}) \quad (13)$$

where f_i and f_j are, respectively, the natural frequencies of the modes estimates i and j . The mode shape similarity between these two modes estimates is taken into account by the well-known Modal Assurance Criterion (MAC).

Hierarchical clustering algorithms differ in the way they compute the distance between two already formed clusters. In this methodology, the single-linkage is used, which means that the distance between two clusters will be considered as being equal to the smallest distance, also computed with equation (13), between objects inside these two clusters. This information allows constructing the hierarchical tree.

In the following step, one must define the tree's cut-level. This is usually dependent on the expected number of clusters. However, this number is not trivial to predict due to the unknown quantity of clusters grouping spurious modes. Thus, the alternative strategy is the use of a distance limit (d_{lim}). Such a threshold is the criterion to prone the branches of the hierarchical tree, generating clusters that are distant from each other more than that value. Hence, the lower is the limit, the higher is the number of resulting clusters.

At this point, one has each cluster representing a single mode. To decide whether the mode is spurious or not, the clusters are ranked according to their number of elements. Thus, the top n_m clusters with more elements are selected. Once physical modes are very consistent for models with different orders, their clusters present a much higher number of elements than the clusters that contain numerical estimates (which present a higher scatter between models of different orders). Therefore, it is expected to find that the n_m selected modes are indeed physical. In practice, it is common to realize that the number of physical modes n_m can be guessed, for instance, by a simple preliminary frequency domain analysis.

Then, the damping ratios are taken into account by means of an outlier analysis within each selected cluster. This internal filtering eliminates the modes estimates with extreme damping ratio values, which are out of the range defined by $\pm 2.698\sigma$ (standard deviation) that embraces 99.3% of the samples in a Gaussian distribution.

Finally, each selected cluster produces average values of the modal parameters (natural frequency, modal damping ratio and mode shape), which are the outputs of the methodology.

3.2 Proposed methodology

This method is also based on a hierarchical clustering algorithm. However, differently from the previous approach, one intends to avoid unnecessary computational efforts and aims to obtain better results from the clustering process. Thus, a pre-filter is applied on the modes estimates data, which removes all modes whose damping ratios are not between 0 and 15% (recommended for civil engineering structures). Besides

that, the user optionally may have the elimination of all modes with Modal Phase Collinearity (MPC) indicator [1, 6] lower than a specified m_{lim} ranging from 0 to 1. Finally, the user also may have removed all modes estimates out of a range of frequencies specified (much like a “band-pass” filter on the modes estimates data).

After having the modes estimates data processed by the pre-filter, a plot of the stabilization diagram would reveal that a significant amount of modes was removed, yielding a clearer aspect. Depending on the situation, and this is not rare, the removed modes may even represent something around 50% of initial data. These removed modes estimates are considered as certainly spurious (or undesired) and will not be processed by the subsequent clustering algorithm.

For the distance calculation between the modes estimates i and j , the novel ‘metric’ is used:

$$d_{i,j} = |f_i - f_j| + (1 - MAC_{i,j})c \quad (14)$$

where c is a constant arbitrarily set to 5 Hz (fair weight value, found after several simulations by the authors).

One can note that, oddly, $d_{i,j} \neq d_{j,i}$ is possible if the used metric is that one shown in equation (13), leading to an asymmetric dissimilarity matrix. On the other hand, this does not happen with the metric proposed in equation (14), which is perfectly symmetric.

Then, the hierarchical clustering is applied. However, differently from the benchmark methodology, the distance between two already formed clusters are measured according to a complete linkage, which means that such a distance is equivalent to the largest possible distance between two of their elements. With this information, one can construct the hierarchical tree that will be trimmed exactly like in the benchmark methodology: by a distance threshold d_{lim} .

Contrary to what the previous methodology does, the internal filter is here applied within the clusters before they are ranked according to their number of elements. This is done to penalize the clusters with high scatter (likely to be representing a numerical mode) before they could be ranked among the top n_m modes that will be selected as physical. Thereby, the internal filter removes outlier modes within the cluster considering, not only damping ratios, but also mode shapes. Mode estimates that have damping ratios out of the $\pm 2.698\sigma$ range and/or have a lower than 0.8 MAC related to the cluster average mode shape are eliminated.

Finally, one can carry out the clustering procedure according to their number of elements. The top n_m clusters (modes) with more elements are selected as being physical. Then, for the methodology output, each selected cluster has its average modal parameters (natural frequency, damping ratio and mode shape) calculated.

4 Results

4.1 Laboratory experiment

This section presents the experimental tests conducted at COPPE/UFRJ laboratory on a simply supported steel

beam depicted in Figure 1. This beam is 1.46 m long with rectangular cross-section (76.2 x 8.0 mm) and was instrumented with six piezoelectric accelerometers (PCB, 336C31). Data acquisition was carried out using Lynx ADS2002 equipment, which essentially is a conditioning/amplifier regulating system. This study considered random vibration tests (using a shaker). The random excitation was applied throughout the duration of the tests.

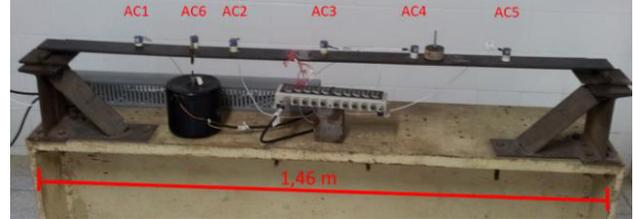


Figure 1. Instrumented steel beam.

The PC variant of SSI-DATA method was applied to the response series (six channels), leading to modes estimates of models with even order from 10 to 120. After a first quick manual judgment, the user-defined parameters were set up: $d_{lim}=0.01$ and $n_m=8$ (benchmark methodology); $d_{lim}=1\text{Hz}$ and $n_m=5$ (proposed methodology). The m_{lim} filter parameter was set to 0.9. Then, the different methodologies performed the automatic interpretation of such a data. Results of the processes, for both methods, can be checked in Figures 2 to 9.

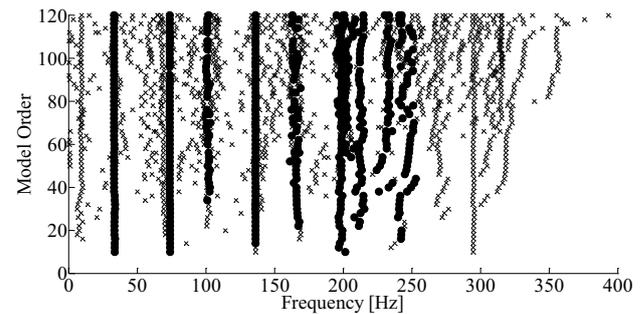


Figure 2. Stabilization diagram obtained by the benchmark methodology.

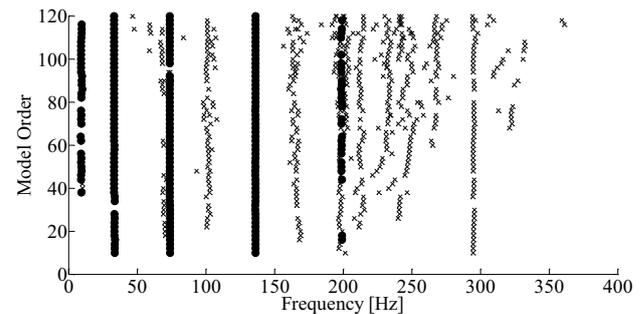


Figure 3. Stabilization diagram obtained by the proposed methodology.

Figures 2 and 3 depict the modes estimates considered physical (solid circles) by both approaches.

It is possible to see the clearer aspect of the stabilization diagram produced by the proposed methodology. Besides, one can note that the columns of modes estimates considered physical, are much more vertical (less scattered) for the proposed approach.

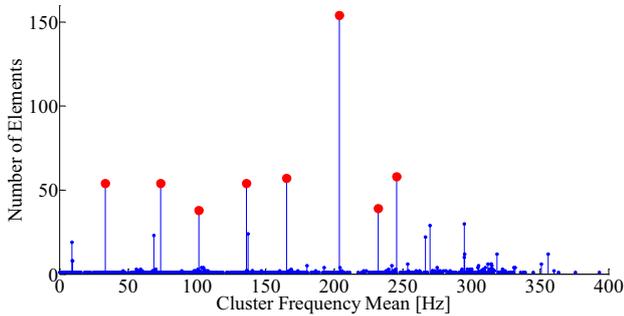


Figure 4. Number of elements per cluster. Benchmark method.

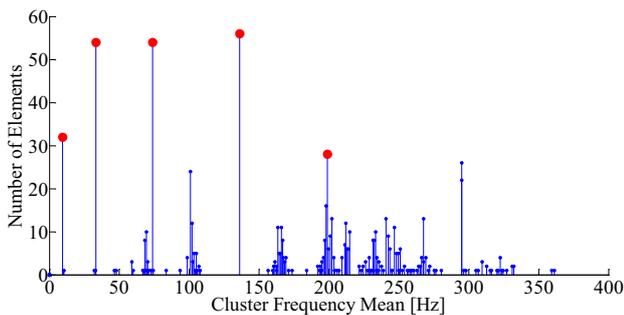


Figure 5. Number of elements per cluster. Proposed method.

The number of elements grouped into each cluster can be checked by means of Figures 4 and 5. The n_m selected modes are highlighted by the big circles on the top of the stem. One can note the high number of elements grouped, erroneously, into the cluster with frequency mean close to 200Hz in the case of the benchmark methodology (Figure 4).

Figures 6 and 7 show the modes estimates of the n_m selected clusters, scattered in function of damping ratio and natural frequency. Concerning the benchmark results (Figure 6), special attention should be paid to the high dispersion of modes estimates with frequencies higher than 150 Hz. This phenomenon does not happen with the proposed methodology, as can be seen on Figure 7.

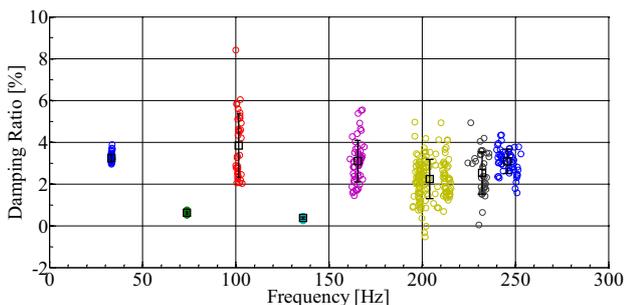


Figure 6. Frequency versus damping ratio. Benchmark method.

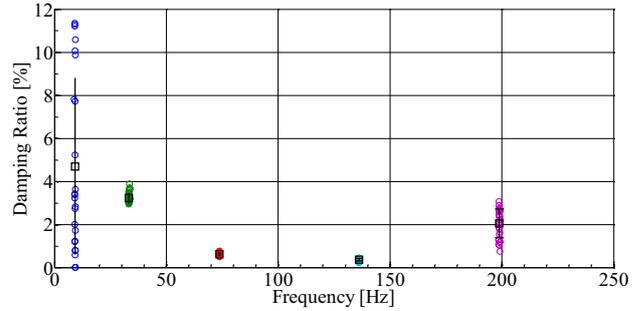


Figure 7. Frequency versus damping ratio. Proposed method.

The automated identification results can be checked by means of the Figures 8 e 9 for both methods. The small font texts inform about the frequency and damping ratio of each mode. ‘STD’ means standard deviation of the respective quantity.

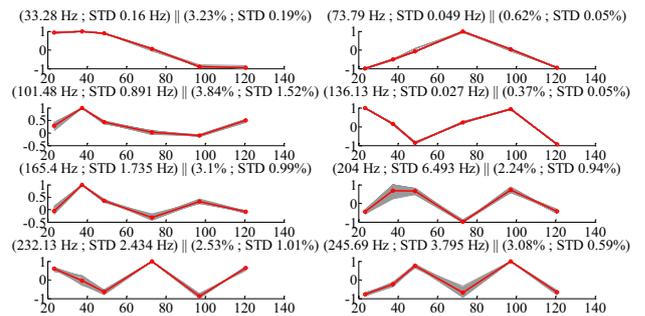


Figure 8. Mode shapes, natural frequencies and damping ratios for each selected cluster. Benchmark methodology.

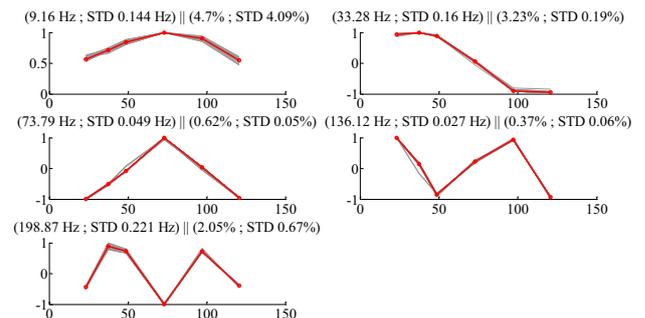


Figure 9. Mode shapes, natural frequencies and damping ratios for each selected cluster. Proposed methodology.

4.2 PI-57 Oise-Bridge

The PI-57 Bridge is a double-deck bridge located near the town of Senlis in France, crossing the Oise River and carrying the A1 motorway, which connects Paris to Lille (Figures 10). The bridge, a 116.50 m long, cast-in-place, post-tensioned segmental structure built in 1965, consists of three continuous spans of 18.00 m, 80.50 m, 18.00 m. The two lateral spans play the role of counterweights. Dynamic tests were performed under ambient excitation: the traffic was used as a source of excitation. Sixteen piezo-electric accelerometers (Bruel & Kjaer 4507B-005 with sensitivity 1 V/g, frequency range from 0.4 to 6000 Hz, maximum operational level 75g, temperature range

from -54 to 100°C) have been installed on the bridge deck (Lille/Paris – Figure 11). The data acquisition system was composed of two separate data acquisition systems. For the acceleration recording, a data programmable controller Gantner E-PAC DL was used and connected to an 8 GB USB flash drive. Data were transferred by a TCP/IP modem. Accelerations were filtered within the 0–30 Hz frequency range and sampling was set to 0.004 s (250 Hz) during 5 min. In order to make the data processing amenable, structural data were only recorded every 3 h over a 24-h time period and stored on a buffer hard disk. For the sake of simplicity, in this experiment, only the upper nine vertical accelerometers (in the plane view – Figure 11) had their signals processed. Thus, one can expect that the algorithms will identify torsional modes among the vertical bending modes.



Figure 10. PI-57 Bridge, France.

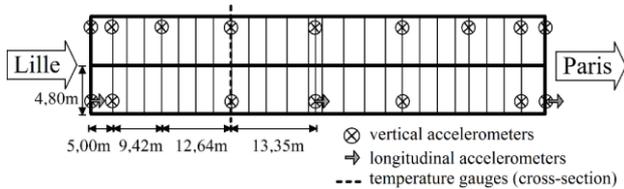


Figure 11. Plane view of the bridge with monitoring system.

Exactly like in the previous experiment, the SSI-DATA (PC) was applied to estimate structural modes considering both the benchmark methodology and the proposed methodology. However, differently from the previous test, in this application the first method had d_{lim} set to 0.1 and n_m set to 6. On the other hand, the proposed algorithm had its two parameters set to the very same values of the first experiment ($d_{lim}=1\text{Hz}$ and $n_m=5$) with one difference: here the proposed method also eliminated (through its pre-filter) all the modes estimates with frequencies out of the 0 to 50 Hz range.

The benchmark method detected, as requested, six modes, which can be seen in Figure 12 as the six vertical columns of solid circles.

Concerning the proposed method, the stabilization diagram, after filtering, can be analyzed in Figure 13. This diagram depicts, clearly, five vertical columns of physical modes estimates (solid circles). Because of the filters, much less computational effort was needed to achieve the results. One can note the clearer aspect of the diagram.

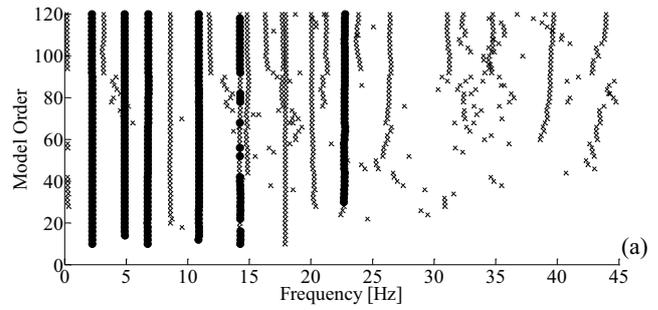


Figure 12. Stabilization diagram obtained by the benchmark methodology.

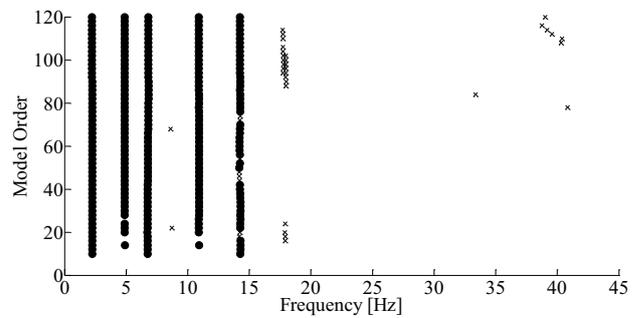


Figure 13. Stabilization diagram obtained by the proposed methodology.

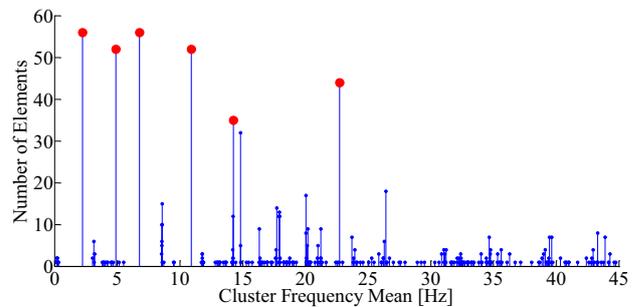


Figure 14. Number of elements per cluster. Benchmark method.

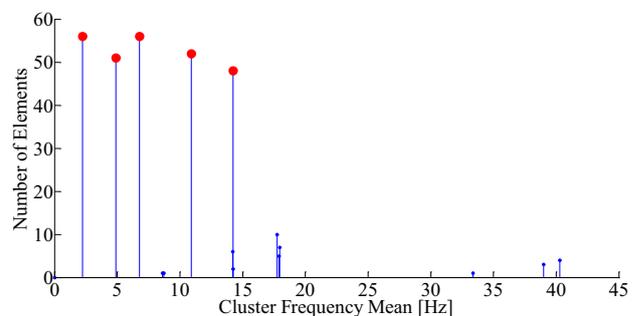


Figure 15. Number of elements per cluster. Proposed method.

The number of elements per cluster can be seen in Figures 14 and 15 for the benchmark and proposed methodologies, respectively. The big head stems means the cluster is selected as being physical.

A scatter plot (frequency versus damping ratio) of the selected modes estimates can be checked by means of Figure 16 and Figure 17 for both methodologies.

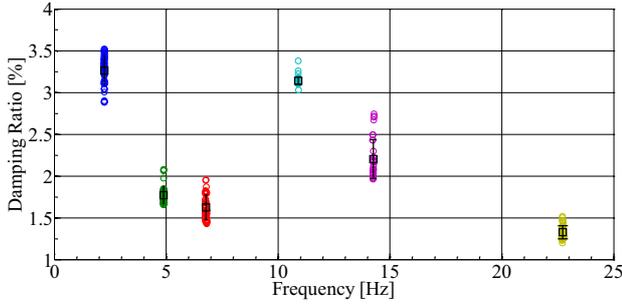


Figure 16. Frequency versus damping. Benchmark method.

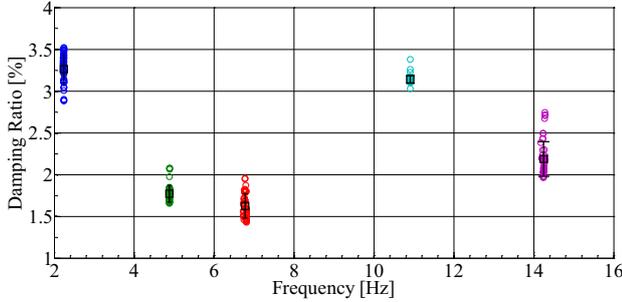


Figure 17. Frequency versus damping. Proposed method.

A summary of the automated identification results can be checked by means of the Figures 18 e 19 for both methods. The small font texts inform about the frequency and damping ratio of each mode. ‘STD’ means standard deviation of the respective quantity.

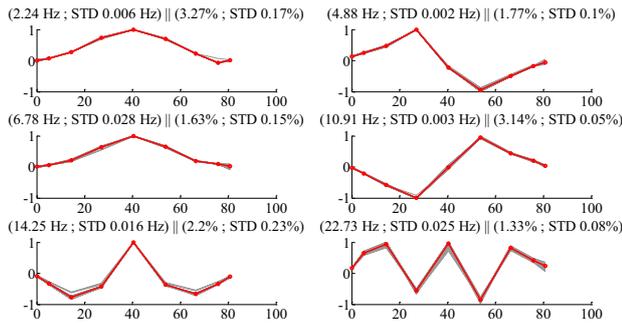


Figure 18. Mode shapes, natural frequencies and damping ratios for each selected cluster. Benchmark methodology.

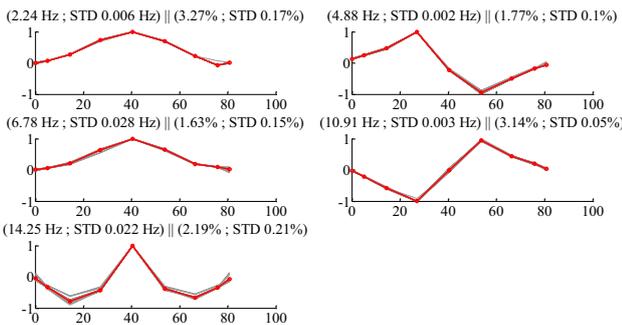


Figure 19. Mode shapes, natural frequencies and damping ratios for each selected cluster. Proposed methodology.

5 Final considerations and remarks

Considering the experimental applications presented in this paper, it is possible to note that, compared to the benchmark approach, the proposed methodology works with a much lower number of mode estimates due to the pre-filter application. As a result, a faster and more accurate clustering process is achieved. In the first test, for instance, 1015 (57.83%) of the 1755 initial modes estimates were automatically considered as certainly spurious and, therefore, were immediately removed. This fact can be noticed through the comparison between the full stabilization diagram in Figure 2 and the cleared one in Figure 3. In a similar way, concerning the second test, as can be seen in Figure 13, the pre-filter also deleted a large number of undesired noisy modes. More precisely, this quantity reached 81.39% (1347) of the initial 1655 modes estimates. These expressive numbers of removed modes are mainly due to the MPC limit, which was set to 0.9. Such a parameter must be used with caution [1].

The hierarchical clustering process over the modes estimates, or over the few remaining ones in case of the proposed method, deserves special attention. In the first experiment, by inspecting Figures 2 to 9, one may observe that the benchmark method was not able to identify the first vertical bending mode with frequency close to 9 Hz. However, in spite of the low energy excitation of this mode during the tests, the proposed method did find this mode. Several values for the distance threshold (d_{lim}) were tried on the benchmark method in order to induce the algorithm to ‘catch’ the first mode. However, the results were conclusive: as one increases the value of d_{lim} , before the low frequencies clusters could gain enough modes to be ranked among the top n_m , the higher frequencies clusters start to gather, in a too permissive way, a lot of modes (resulting in a fast growth of their stems in Figure 4). Consequently, a high scatter between modes belonging to a high frequency cluster can be observed in Figures 2, 6 and 8. Evidently, this excessive scattering is undesirable. A possible way of forcing the benchmark algorithm to catch the first mode is by keeping the same d_{lim} value and increasing the n_m parameter in an extreme conservative way. However, one can say that this solution is not elegant, since a lot of spurious clusters (modes) would be selected along with the physical ones.

The high frequency scattering phenomenon happens for two reasons. Firstly, the distance ‘metric’, which is written in equation (13), presents a dimensionless ratio of frequencies and, therefore, does not measure a distance between low frequency modes with the same rigor as it measures a distance between high frequency modes. Secondly, the single linkage criterion for measuring the distance between two already formed clusters promotes a decreasing in the heights of the hierarchical U-shaped branches, i.e. it makes easier to ‘gather’ two clusters when compared to the complete linkage criterion.

On the other hand, the novel methodology proposed in this paper is immune to such a phenomenon. That is due to the fact that the distance ‘metric’, stated in equation (14), does not contain any frequency ratio and, therefore, ensures that the distances between modes (with

either high or low frequency) will be expressed by an absolute value in Hz.

Another remarkable point is that the success of the algorithm proposed in this work is not as sensible to the threshold limit as the benchmark algorithm is. For instance, in both tests the user-defined parameters for the proposed algorithm did not need to be adjusted.

Both experiments showed that the peculiarities of the two examined automation methods produce significant differences at the final process output. However, at least for the studied cases, the proposed methodology, when compared to the benchmark, generates better results and, therefore, represents an advance towards an even more robust automatic modal identification method.

Acknowledgements

The authors would like to thank 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 the Federal Universities of Juiz de Fora (UFJF) and Ouro Preto (UFOP) for the financial support.

References

1. E. Reynders, *System Identification Methods for (operational) modal analysis: review and comparison*, CIMNE, Barcelona, (2012).
2. F. Magalhães, *Operational modal analysis for testing and monitoring of bridges and special structures*, PhD thesis by University of Porto, (2010).
3. A. Cabboi, *Automatic operational modal analysis: challenges and applications to historic structures and infrastructures*, PhD thesis for the Università degli Studi di Cagliari (2013).
4. F. Ubertini, C. Gentile, A. Materazzi, *Automated modal identification in operational conditions and its application to bridges*, *Engineering Structures* **46**: 264-278 (2013).
5. P. Van Overschee, B. De Moor, *Subspace identification for linear systems: theory, implementation and applications*, Kluwer Academic Publishers, (1996).
6. P. Vacher, B. Jacqier, A. Bucharles, *Extensions of the MAC criterion to complex modes*, International Conference on Noise and Vibration Engineering, ISMA (2010).