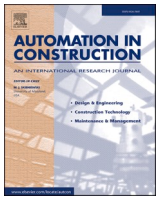
Automation in Construction 156 (2023) 105137

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon



Iterative hierarchical clustering algorithm for automated operational modal analysis

A. Romanazzi *, D. Scocciolini, M. Savoia, N. Buratti

DICAM, University of Bologna, Italy

ARTICLE INFO

Keywords:
Iterative hierarchical clustering analysis
Variance-based sensitivity analysis
Sobol's indexes
SSI covariance driven
Structural health monitoring
Machine learning

ABSTRACT

Recent developments in sensors and data processing made the structural health monitoring (SHM) sector expanding to big-data field, particularly when continuous long-term strategies are implemented. Nevertheless, main shortcomings are due to the identification and extraction of modal features. In fact, although machine learning methods have been implemented to automate modal identification processes, intense user interaction and time-consuming procedures are still required, limiting the extensive use of these techniques. In order to provide a fully automated procedure capable of identifying and extracting modal properties from covariance driven SSI analyses, an innovative and flexible algorithm for Iterative Hierarchical Clustering Analysis (IHCA) is proposed. To evaluate the stability and robustness of the IHCA method, a Variance-Based Global sensitivity Analysis (VBGA) was performed considering a numerical and experimental case study. The outcomes demonstrated that the IHCA is stable in clustering the physical structural modes and selecting the most representative modal features.

1. Introduction

In recent decades, structural health monitoring is gathering relevance due to its role in supporting decision making for the maintenance of civil infrastructures and architectural heritage [1–4]. In addition, recent advances in technology and accessibility of sensors and recording systems fostered the application of dynamic identification, in particular the installation and operation of continuous monitoring systems. Such incremented application turned the dynamic identification into big-data analysis problem [7–9]. In fact, the identification process of a dynamic system consists of various steps, namely data collection, signal processing, structural dynamic identification and estimation of valid sets of modal parameters, with the latter being the most time-consuming phase which involves considerable amount of user interaction. With this regard, several data-driven techniques with the use of machine learning-based methods are proposed, as in [4–6]. Nevertheless, such approaches do not provide any information on the modal parameters of a system, i.e. frequency, mode shape, and damping coefficient. Therefore, the development of reliable and robust tools for the automatic modal analysis and modal parameter extraction in operational conditions is still fundamental for the consequent damage detection [10,11].

With regard to dynamic identification, several methods are available in literature, which can be distinguished in either frequency [12,13] or time domain approaches [14,15]. Among the various time domain approaches, a fast and robust algorithm is the CoVariance driven-Stochastic Subspace Identification (SSI-CoV) method, in which the response of the system to unmeasured operational loading is recorded and the dynamic parameters are extracted from output-only data assuming a stationary stochastic white noise input. In general, parametric system identification techniques, such as SSI methods, refer to the equation of motion with auxiliary state variables. However, noise in the SSI results can be observed, which typically derives from modelling inaccuracies, as the data cannot be assumed exactly in a stochastic state-space model, measurement noise due to electronic devices and sensors, computational noise and finite number of data [14]. Therefore, the separation of physical and spurious modes is a central step of the identification algorithm. The definition of the stabilization diagram is a general accepted procedure to remove part of such errors. It is based on the observation that in a very large number of modal identification problems, physical modes of the structure are characterized by approximately the same modal parameters over the model order, while the spurious modes appear scattered [16,17]. In this regard, the model

* Corresponding author.
E-mail addresses: antonio.romanazzi2@unibo.it (A. Romanazzi), daniele.scocciolini2@unibo.it (D. Scocciolini), marco.savoia@unibo.it (M. Savoia), nicola.buratti@unibo.it (N. Buratti).

https://doi.org/10.1016/j.autcon.2023.105137
Received 28 July 2023; Received in revised form 19 September 2023; Accepted 15 October 2023
Available online 21 October 2023

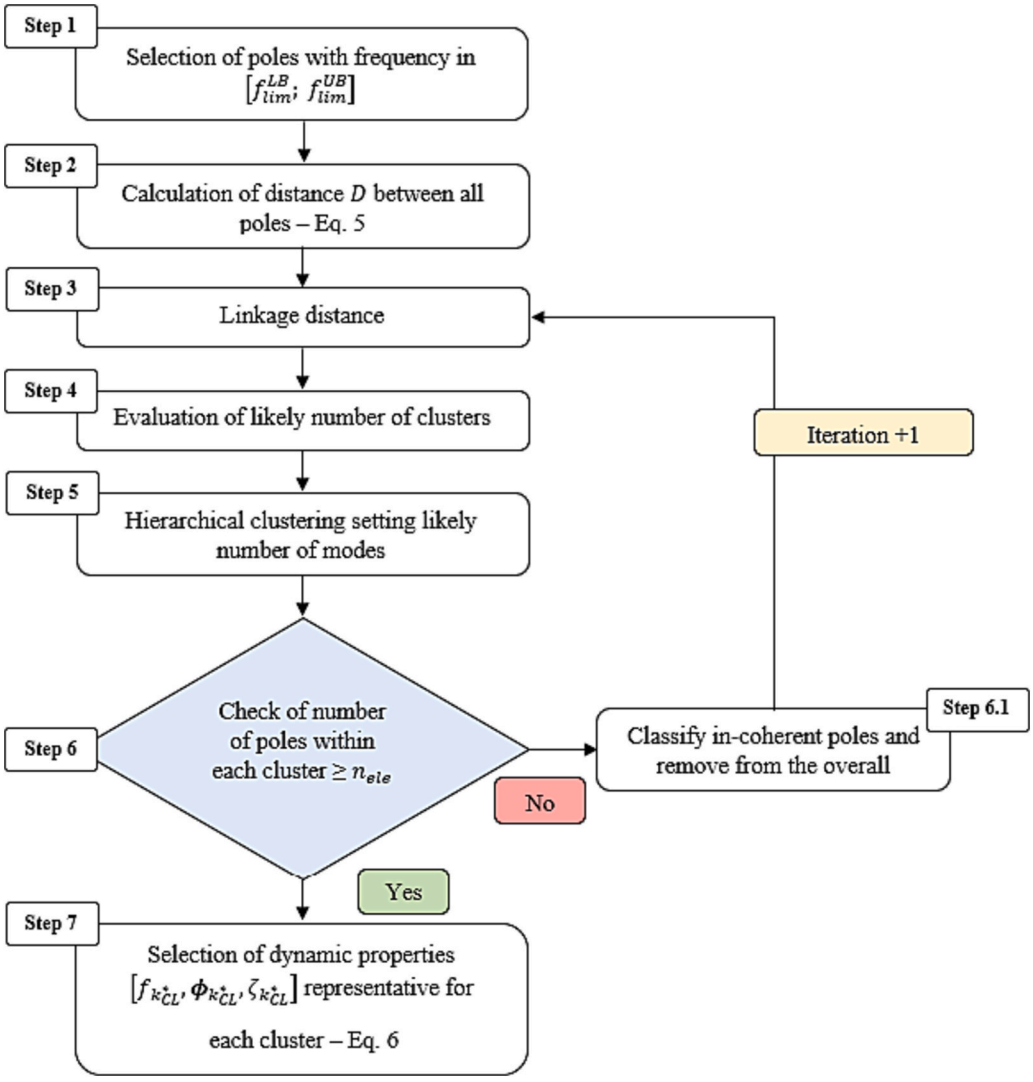


Fig. 1. Scheme of the Iterative Hierarchical Clustering Analysis (IHCA) algorithm.

order describes the modes of vibrations involved in the dynamic response for a certain excitation level [18]. Thus, the diagram allows to identify the poles in which frequencies, mode shapes and damping coefficients are consistent (stable) for increasing model orders and likely to be physical [19].

It is clear that the results of the stabilization diagrams depend on the criteria introduced to define the consistency of poles, on the used identification algorithm, on the values of the input parameters and on the Signal-to-Noise Ratio (SNR) of the recorded measurements. Hence, the interpretation of the stabilization diagram is often not straightforward and might require user interaction, which becomes a limitation when large datasets are analysed. In general, problems related to the analysis of the stabilization diagrams and the inference of modal parameters concern three main aspects, namely i) the individuation of groups of stable poles describing structural modes, ii) the definition of similarity based on which the poles are grouped, and iii) the selection of the most representative set of dynamic parameters for each group. In order to automate the interpretation of the stabilization diagrams and individuate the modes, several strategies based on cluster analysis in machine learning were proposed. Cluster analysis is a set of unsupervised learning methods to discover unknown groups of similar observations in data.

As a measure of similarity, several approaches were proposed to calculate the distance between poles in the stabilization diagrams. In [20,21], eigenfrequency difference and MAC value were considered as measure for similarity, while in [22,23] the eigenfrequency and damping ratio were used to define the distance. However, the damping coefficient is deemed unsuitable to separate modes due to scattering resulting from mathematical approximation in the analysis and because different modes might have the same damping ratio [1]. In [24], another distance measure was introduced based on the MAC value. In [25], further criteria for the definition of stable poles were introduced making the process more robust, however user interpretation of the stabilization diagram for the subsequent cluster analysis and selection of the representative modal parameters was required.

With regard to clustering algorithm, in [26], histogram analysis was implemented to separate the possible physical modes in the stabilization diagram. In such analysis, the frequency domain was divided into narrow bins in which the number of stable modes was counted; nevertheless, information on the mode shapes were not considered making the method not comprehensive and robust. In [27–31], K-means and fuzzy C-means non-hierarchical clustering were implemented as partitioning methods to divide the total set of modes into predefined number of clusters; in particular, the methods were used in single model order to separate physical modes from spurious modes. In [32], to determine the number of possible clusters in a Gaussian Mixture Model (GMM), the Dirichlet Process is used. However, four hyper-parameters are required which are manually defined on the basis of data observations.

Table 1
Mass matrix of the numerical benchmark.

M =	20	0	0	0	0	[ton]
	0	20	0	0	0	
	0	0	20	0	0	
	0	0	0	20	0	
	0	0	0	0	20	

The major limitation for non-hierarchical clustering procedures is the requirement to select the clusters seeds which might affect the individuation of clusters and, consequently, the reliability of the process. A more robust algorithm is the hierarchical clustering [33], in which objects are grouped in a succession of similarity based on measured distance which is updated at each step of the clustering process. In this way, since the actual number of clusters is not known a priori, a dendrogram in the form of tree-like linkage is obtained on the data similarity, which provides the global picture of each possible number of clusters in a multilevel hierarchy. However, at the final stage of the clustering process, a level or scale of clustering must be chosen to group the observations in the most appropriate way. With this regard, establishing a norm for cutting the dendrogram can be cumbersome, in particular when the results can be affected from approximation in the mathematical solution of the identification problem and noise in the signals.

In [33], to individuate the number of clusters in a dendrogram, a method to evaluate a cut-off distance is proposed, which is based on the statistical distribution of the distance between poles in consecutive model orders. Nevertheless, such method might be incapable of separating closed modes and result misleading, as discussed in [34]. In [34], the cut-off distance is manually tuned basing on a preliminary analytical solution of modal analysis; however, such approach might not be applied on big-data set and deceptive in case of complex structures, or uncertain conditions, e.g. boundary conditions, structural damage, material degradation, environmental influence.

Concerning the individuation of the number of groups, i.e. of physical modes, in [35], the number of clusters was assumed to be known a priori, although this is rarely the case. A more recent and successful application of hierarchical clustering is reported in [1,25], although the setting of a threshold distance to distinguish the clusters was introduced, making the algorithm not flexible on the different dataset under analysis.

Therefore, an automatic, flexible and robust algorithm for selection of modal parameters in a system identification analysis is still required for continuous SHM applications. In this framework, an automated and flexible Iterative Hierarchical Cluster Analysis (IHCA) is proposed in this paper, which considers the similarity between poles in terms of

Table 2
Stiffness matrix of the numerical benchmark.

K =	575,034.7	−227,812.5	0	0	0	[Nmm]
	−227,812.5	370,034.7	−142,222.2	0	0	
	0	−142,222.2	225,590.3	−83,368.1	0	
	0	0	−83,368.1	128,368.1	−45,000	
	0	0	0	−45,000	45,000	

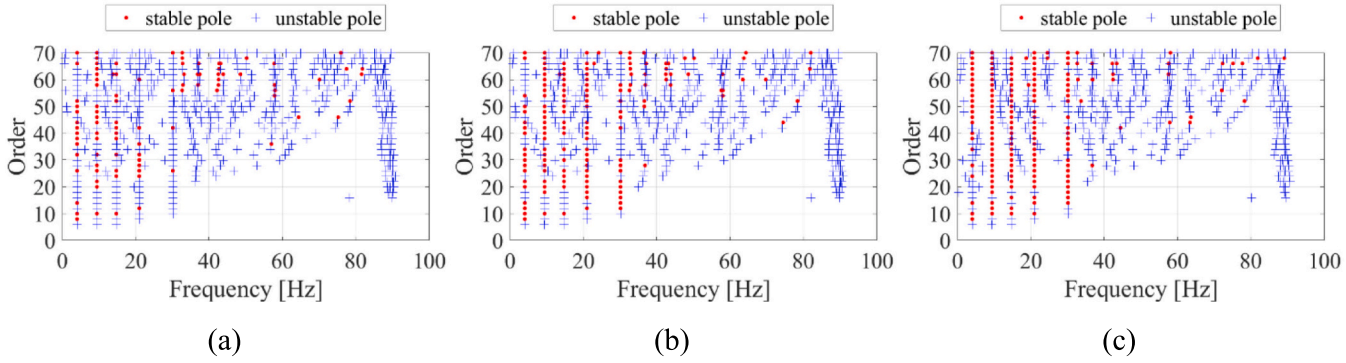


Fig. 2. Stabilization diagram of numerical benchmark for: a) SNR0, b) SNR25, and c) SNR50.

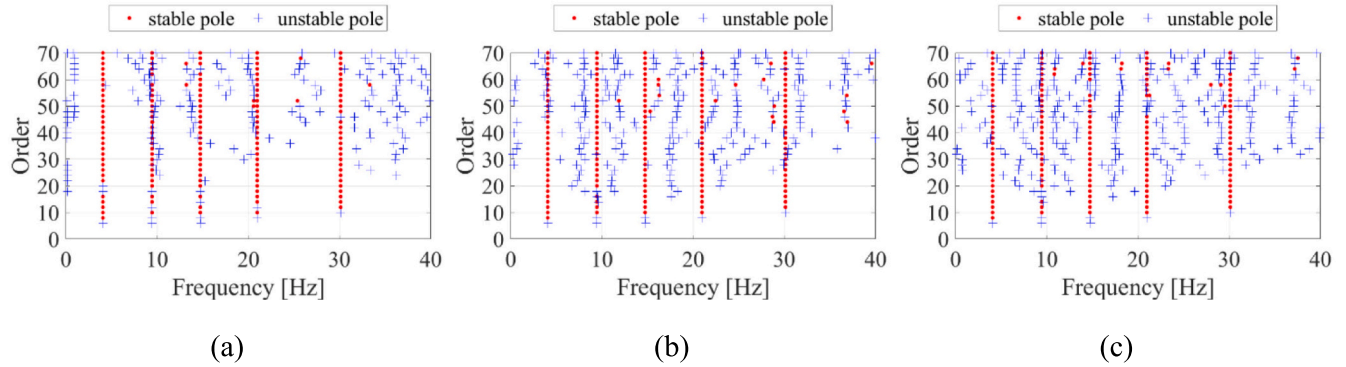


Fig. 3. Stabilization diagram of numerical benchmark considering $f_{lim}^{UB} = 40$ Hz for: a) SNR0, b) SNR25, and c) SNR50.

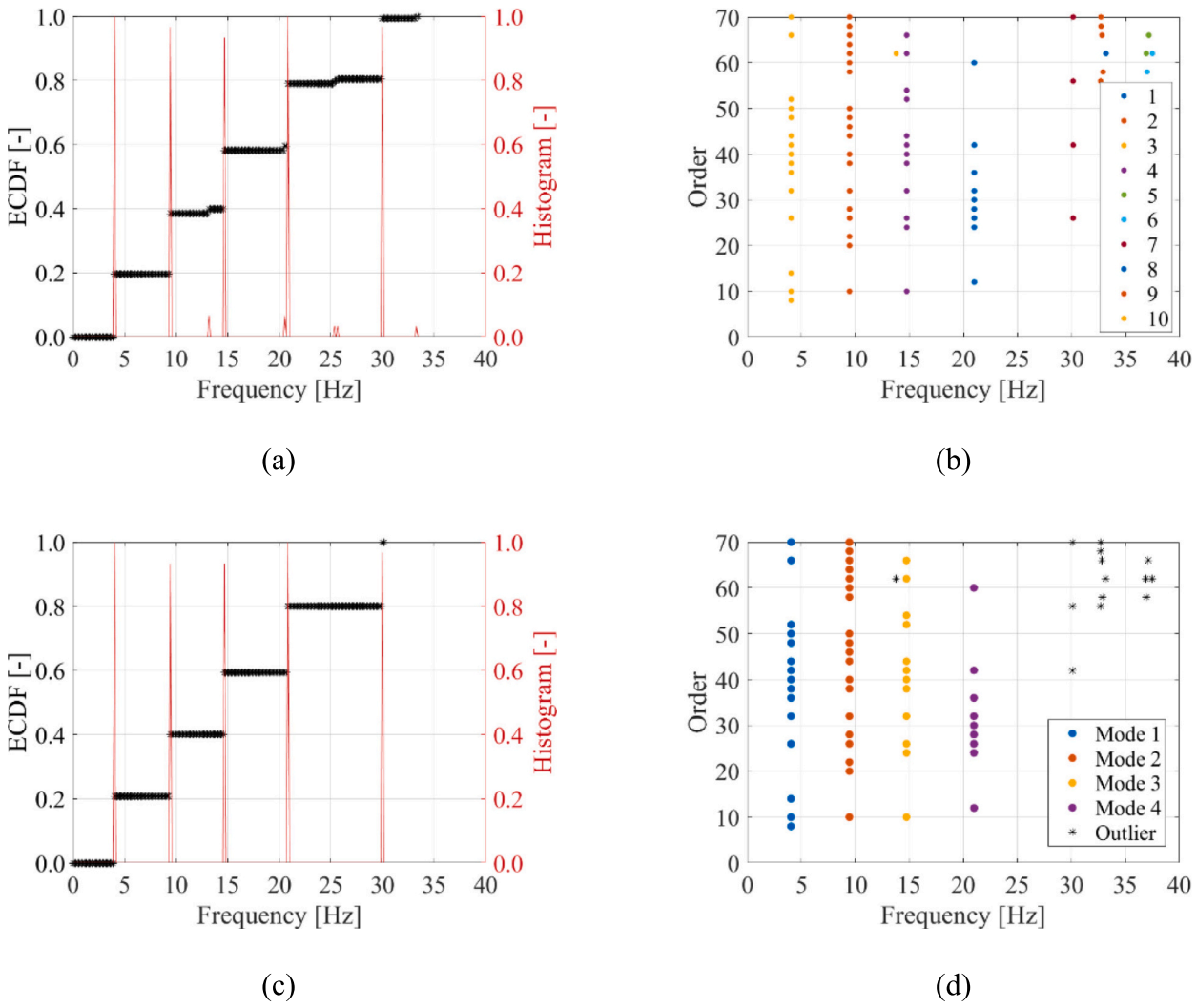


Fig. 4. Modal identification of the numerical case SNR0: a) CDF and PDF of the first iteration, b) identified cluster at the first iteration, c) CDF and PDF of the last iteration, and c) identified cluster at the last iteration.

frequency and MAC. In the IHCA, the number of clusters is not predefined, yet it is inferred through an iterative procedure; in addition, the requirement to initiate a cluster is related to the total dataset in the stabilization diagrams, making the algorithm flexible on the amount of information resulting from the modal analysis. At the final stage of the IHCA, the most representative set of modal parameters for each cluster is selected while introducing a weighting coefficient.

To assess the robustness of IHCA and validate the identification procedure, sensitivity analysis was performed on the IHCA algorithm applied to SSI-CoV results obtained for two different case studies, namely a numerical model and an experimental test. The numerical benchmark allows to evaluate the sensitivity of the IHCA against a numerical solution from modal analysis and to assess the effectiveness of the SSI-CoV method combined with the IHCA in a controlled environment; on the other hand, the experimental case study allows to consider the uncertainties of a real problem, related to randomness of signals, environmental effects and undetermined boundary conditions. Finally, the IHCA was applied on a large dataset of 720 records and the obtained results remarked that the IHCA is stable and robust in clustering and selecting the modal features and can be considered for automated modal identification for continuous SHM applications.

2. Iterative hierarchical cluster analysis

The algorithm of the IHCA is illustrated in Fig. 1. Provided the stability diagram from SSI-CoV analysis, all the stable poles are first detected with the concurrent complying of five requirements [25]. Specifically, two hard-criteria require that the mode shapes must be a complex and conjugate pair, and that the damping coefficient belongs in an interval which is in accordance with values observed in real cases, ζ_{max}, (Eq. (1)); and further four soft-criteria are based on the variation between the results of two consecutive increasing SSI-CoV orders, and consider the relative distance in frequency, $d_{f_{lim}}$, (Eq. (2)), relative distance in damping coefficient, $d_{\zeta_{lim}}$, (Eq. (3)) and similarity of mode shapes through MAC, MAC_{min}, (Eq. (4)). In case of civil buildings and in this investigation, the limit values of the requirements in Eq. (1) - Eq. (4) are set $\zeta_{max} = 0.05$, $d_{f_{lim}} = 0.01$, $d_{\zeta_{lim}} = 0.08$, and $MAC_{min} = 0.95$.

$$\zeta_n \leq \zeta_{max} \tag{1}$$

$$\frac{|f_n - f_{n+1}|}{f_n} \leq d_{f_{lim}} \tag{2}$$

$$\frac{|\zeta_n - \zeta_{n+1}|}{\zeta_n} \leq d_{\zeta_{lim}} \tag{3}$$

$$MAC_{n,n+1} \geq MAC_{min} \tag{4}$$

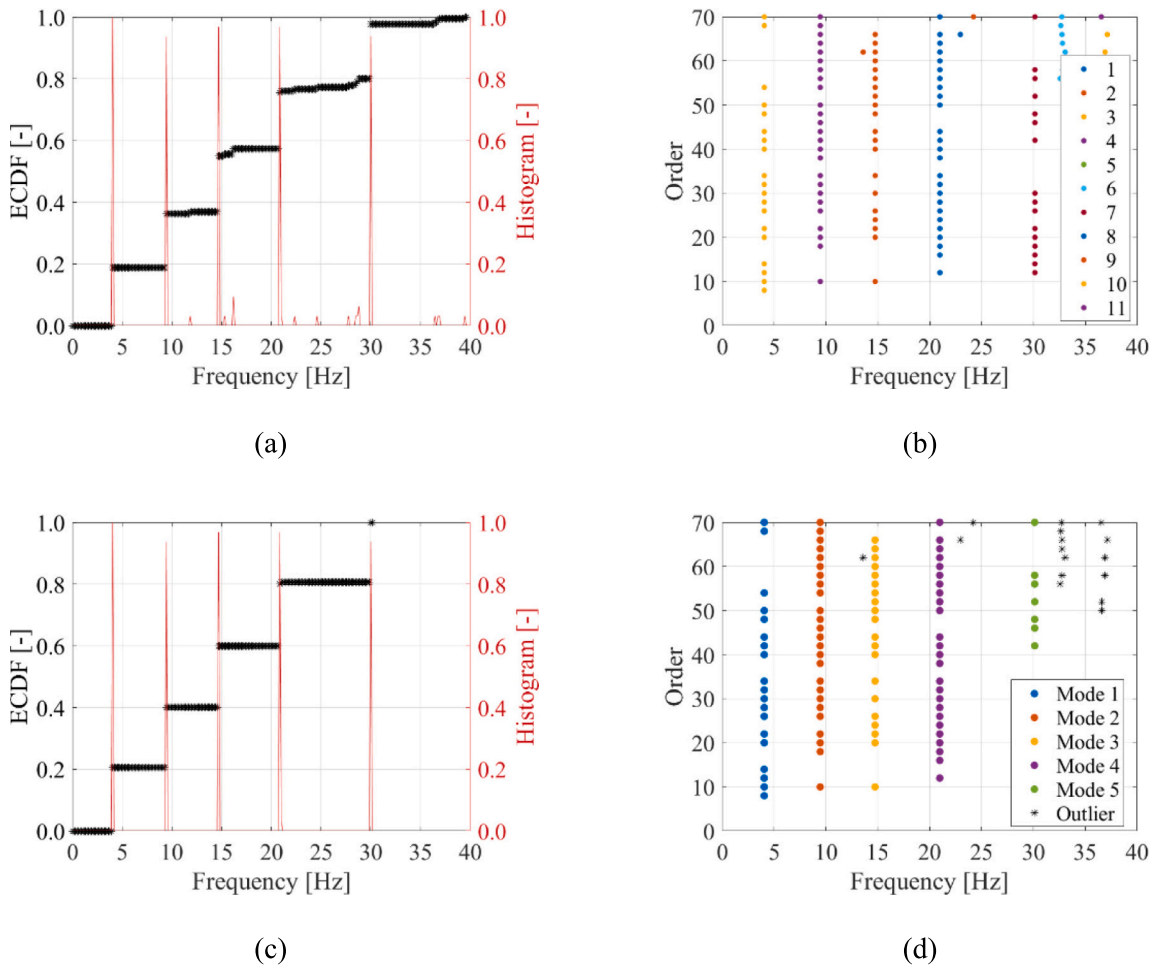


Fig. 5. Modal identification of the numerical case SNR25: a) CDF and PDF of the first iteration, b) identified cluster at the first iteration, c) CDF and PDF of the last iteration, and c) identified cluster at the last iteration.

Then, a range of frequencies of interest for the structural system $\left[f_{lim}^{LB}; f_{lim}^{UB}\right]$ is set and dataset is filtered accordingly (step 1 in Fig. 1). Afterwards, the similarity between each pair of poles $k-j$ in the filtered dataset is calculated using as distance measure a combination of frequency and MAC (step 2 in Fig. 1), as in Eq. (5):

$$D_{k-j} = \alpha \times \frac{|f_k - f_j|}{max(f_k, f_j)} + (1-\alpha) \times \left[1 - MAC_{k,j}\right] \tag{5}$$

where f_k and f_j are the frequencies of the k-th and j-th pole, respectively, and $MAC_{k,j}$ is the Modal Assurance Criterion computed considering the modal shapes of the aforementioned poles. In addition, a parameter α is introduced as weight coefficient in the domain [0; 1]. In this way, the calculation of the pair distance can be adjusted favouring frequency similarity ($\alpha > 0.5$), mode shape similarity ($\alpha < 0.5$), or equal frequency-modal-shape similarity ($\alpha = 0.5$).

Afterwards, based on the similarity defined by the introduced distance measure, the objects in the dataset are grouped following a binary hierarchical clustering tree-like algorithm (dendrogram) (step 3 in Fig. 1). The process of grouping considers the closest pair of observations at first; these are removed from the dataset to form a group which is considered as a new object. Subsequently, the distances between all the elements in the dataset are updated and the closest elements are grouped, proceeding to higher-order arrays. Among the various hierarchical algorithms, namely single linkage, complete linkage, average linkage, centroid method, and Ward's method [36], the latter is considered in the present investigation. The algorithm continues hierarchically until all the sub-groups are gathered in one group, which represents the entire dataset.

The clustering procedure returns a dendrogram, which represents a sequence of cluster combinations of the filtered poles of the stability diagram, yet the number of expected clusters needs to be set. The steps to define the likely number of clusters are the follows (step 4 in Fig. 1): i) the domain of frequency of interest is first discretized in intervals with amplitude Δf, ii) the Empirical Cumulative Distribution Function (ECDF) and the corresponding histogram are computed for the dataset, and iii) the local peaks in the histogram are counted. The local peaks in the histogram are individuated by checking that the value at the increment Δf_m is higher than the values at the increments Δf_{m-1} and Δf_{m+1}. Subsequently, the number of the local peaks in the histogram is assumed as the likely number of clusters in the dataset, that in turn represents the likely number of identified modes (step 5 in Fig. 1).

However, in such approach, an isolated instance might result in a local peak and consequently be counted as sub-group. To overcome such inconvenience, a threshold on the minimum number of entries in each cluster is established as requirement (n_{ele}), which is conveniently set as a percentage of the total number of data points in the entire dataset (step 6 in Fig. 1). In this way, clusters with a number of objects lower than the set requirement are not considered, and those observation are labelled as outliers in the analysis and removed from the overall dataset. In such case (step 6.1 in Fig. 1), the process is iterated from step 1 with the new dataset filtered from the previous outlier poles. Hence, the entire process is iterated until the requirement of minimum number of objects within each cluster is satisfied and the final clusters are individuated.

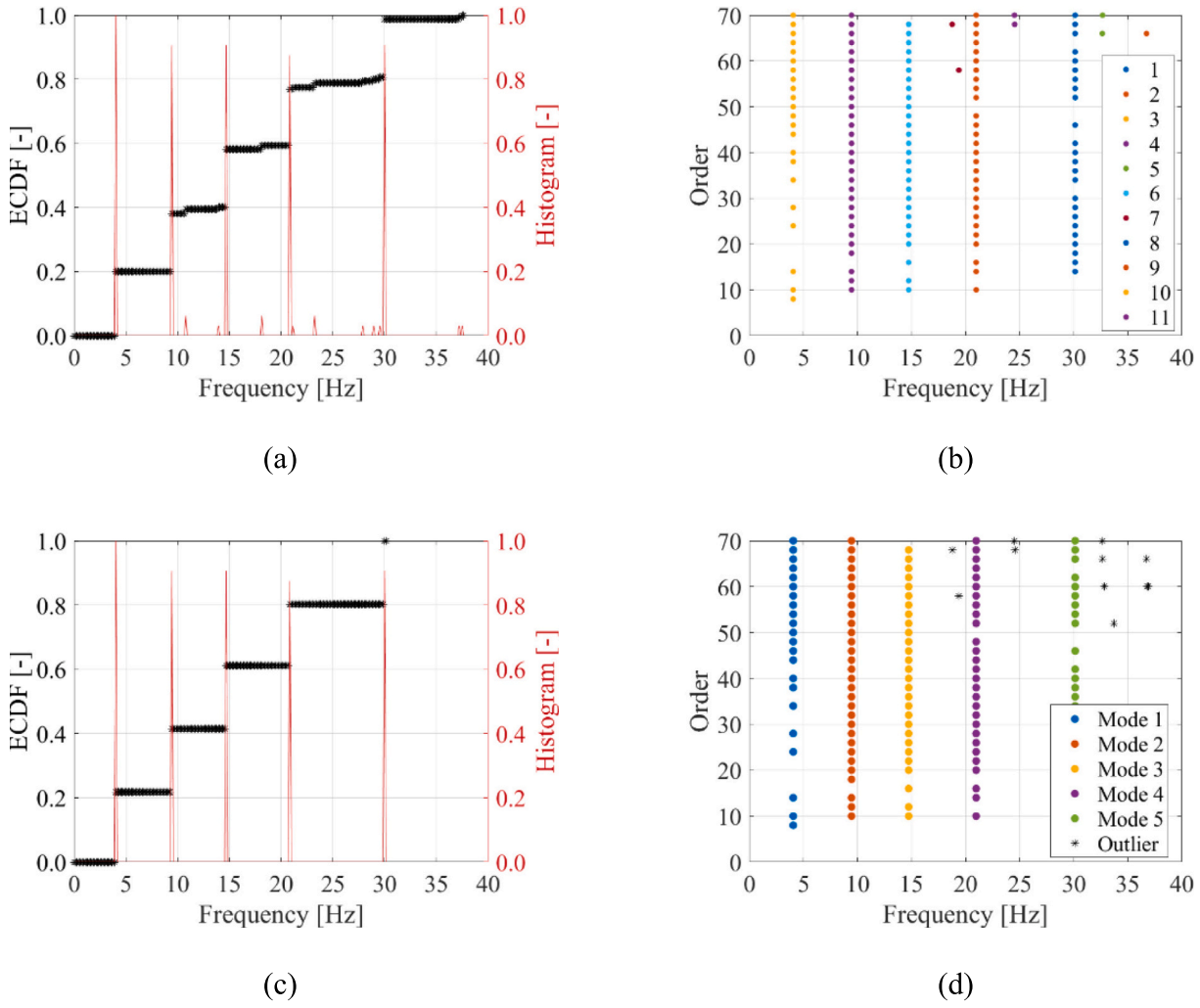


Fig. 6. Modal identification of the numerical case SNR50: a) CDF and PDF of the first iteration, b) identified cluster at the first iteration, c) CDF and PDF of the last iteration, and c) identified cluster at the last iteration.

Table 3
Comparison between modal properties of numerical benchmark and modal identification by means of IHCA and SSI-CoV.

	Mode 1 (f_{num} = 4.07 Hz)	Mode 2 (f_{num} = 9.47 Hz)	Mode 3 (f_{num} = 14.79 Hz)	Mode 4 (f_{num} = 21.11 Hz)	Mode 5 (f_{num} = 30.51 Hz)
SNR0	$MAC = 0.99$ $f_{SSI-CoV} = 4.07$ Hz $f_{Err\%} = 0.0$ %	$MAC = 1.00$ $f_{SSI-CoV} = 9.46$ Hz $f_{Err\%} = 0.1$%	$MAC = 1.00$ $f_{SSI-CoV} = 14.74$ Hz $f_{Err\%} = 0.3$ %	$MAC = 0.99$ $f_{SSI-CoV} = 20.98$ Hz $f_{Err\%} = 0.6$ %	$MAC = 0.92$ $f_{SSI-CoV} = 30.14$ Hz $f_{Err\%} = 1.2$ %
SNR25	$MAC = 1.00$ $f_{SSI-CoV} = 4.07$ Hz $f_{Err\%} = 0.0$ %	$MAC = 1.00$ $f_{SSI-CoV} = 9.46$ Hz $f_{Err\%} = 0.1$%	$MAC = 1.00$ $f_{SSI-CoV} = 14.74$ Hz $f_{Err\%} = 0.3$ %	$MAC = 0.99$ $f_{SSI-CoV} = 20.98$ Hz $f_{Err\%} = 0.6$ %	$MAC = 0.91$ $f_{SSI-CoV} = 30.14$ Hz $f_{Err\%} = 1.2$ %
SNR50	$MAC = 1.00$ $f_{SSI-CoV} = 4.07$ Hz $f_{Err\%} = 0.0$ %	$MAC = 1.00$ $f_{SSI-CoV} = 9.46$ Hz $f_{Err\%} = 0.1$%	$MAC = 1.00$ $f_{SSI-CoV} = 14.74$ Hz $f_{Err\%} = 0.3$ %	$MAC = 0.99$ $f_{SSI-CoV} = 20.98$ Hz $f_{Err\%} = 0.6$ %	$MAC = 0.91$ $f_{SSI-CoV} = 30.14$ Hz $f_{Err\%} = 1.2$ %

Finally, a set of dynamic properties $\left[f_{k^*_{CL}}, \boldsymbol{\phi}_{k^*_{CL}}, \zeta_{k^*_{CL}}\right]$, namely frequency, mode shapes and damping coefficient, is selected for each cluster (step 7 in Fig. 1). With this regard, the selected properties are those of the mode of vibration of the pole k^*_{CL} within the cluster that minimize Eq. (6):

$$k^*_{CL} = \underset{k\in[1,\ldots,N_{CL}]}{argmin}\left\{\beta \times \frac{|\widehat{f}_{CL} - f_k|}{\widehat{f}_{CL}} + (1-\beta)\times\left[1 - \overline{MAC}_{CL}(\boldsymbol{\phi}_k)\right]\right\} \quad (6)$$

where N_{CL} are the poles within the cluster, $\widehat{f}_{CL}$ is the median value of the frequency in the cluster, f_k is the frequency of the *k-th* mode of vibration in the cluster and $\overline{MAC}_{CL}(\boldsymbol{\phi}_k)$ is the mean value of MACs between the *k-th* mode and all the modes of the cluster. In addition, the parameter β is defined in the domain [0; 1] to favouring significance in the distance measure either in terms of frequency similarity ($\beta > 0.5$) or mode shape similarity ($\beta < 0.5$), or equal frequency-mode shape similarity ($\beta = 0.5$).

3. Variance-based global sensitivity analysis

Once that the IHCA was formulated, a Sensitivity Analysis (SA) was performed on the algorithm to assess its stability as following described.

(a) (b) (c)

Fig. 7. Masonry building mock-up: a) north-east façade, b) south-west façade, and c) south-east façade.

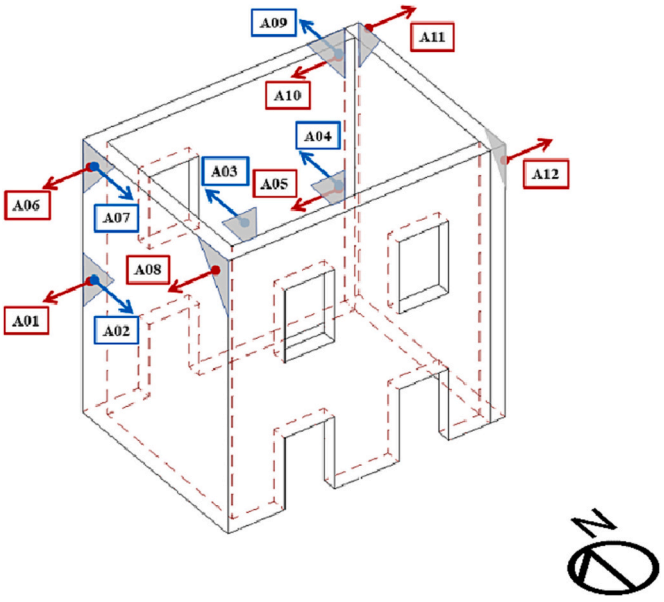


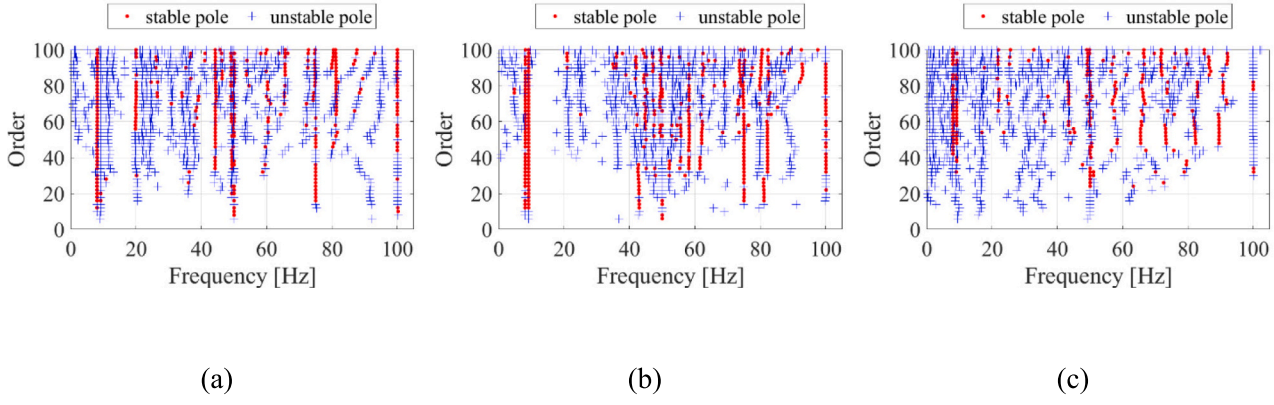
Fig. 8. Recording system setup.

SA aims at understanding how the uncertainty in the output of a model can be ascribed to different sources of uncertainty in the model input factors [37,38]. Therefore, the main application of SA is to identify the relative importance of each input factor considering its contribution to the output variance. Such analysis is defined as factor prioritization [39]. In addition, SA is an integral part of uncertainty quantification related to the estimation of uncertainty propagating through complex models. In this framework, SA supports decreasing the output variance by identifying noninfluential factors which are then fixed without a remarkable loss of information. This procedure is referred to as factor fixing [39,40,41].

Generally, SA methods can be divided in local (LSA) and global (GSA) SA. In LSA, the influence of the variation of inputs on the output is evaluated at a specific point in the parameter domain, and is typically performed through partial derivatives on individual parameters, while fixing all other factors. Consequently, the interaction among the inputs

Table 4
Variance-based sensitivity analysis on the DI01, DI02, and DI03, cases study, with $N = 10{,}000$, k = 3, and considering the number of identified modes as observed output of the model.

Parameter	Domain	DI01		DI02		DI03	
		S	S_t	S	S_t	S	S_t
Δf	[0.1; 0.3]	0.007	0.048	0.013	0.045	0.006	0.034
n_{ele}	[3; 7]	0.081	0.123	0.109	0.144	0.077	0.107
α	[0; 1]	0.034	0.076	0.024	0.062	0.003	0.029



(a) (b) (c)

Fig. 9. Stabilization diagram of numerical benchmark for: a) DI01, b) DI02, and c) DI03 recorded accelerations.

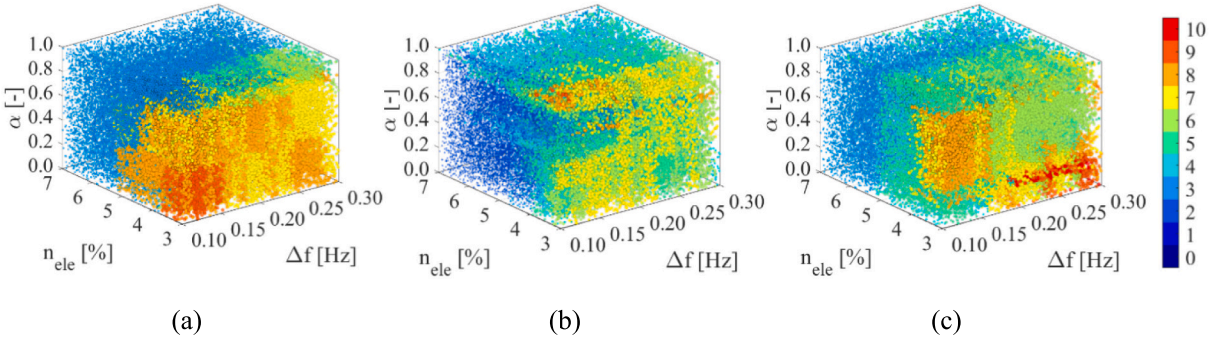


Fig. 10. Scatterplot of the individuated number of clusters (in colormap), as function of Δf, n_{ele}, and α, for: a) DI01, b) DI02, and c) DI03.

Table 5
Variance-based sensitivity analysis on the DI01, DI02, and DI03, cases study $N = 10{,}000$, k = 2, and considering the number of identified modes as observed output of the model.

Parameter	Domain	DI01		DI02		DI03	
		S	S_t	S	S_t	S	S_t
n_{ele}	[3; 7]	0.168	0.234	0.136	0.162	0.105	0.131
α	[0; 1]	0.035	0.104	0.024	0.048	0.029	0.057

is neglected in LSA and the sensitivity measures are not comprehensive. Whereas, in GSA, the output uncertainty is analysed over the entire variation domain of factors, therefore, providing more information on the model sensitivity to the interaction among the inputs [38,42]. Among the various method of GSA, variance-based methods (VBSA) are commonly used to assess the sensitivity of a model to inputs in terms of a reduction in the variance of the observed output [39].

Given a model in a general form $Y = f(X_1, X_2, \ldots, X_i)$, where the input factors X_i are independent, the sensitivity indices are related to the decomposition of the variance of Y into terms due to each X_i taken singularly, as well as into terms due to the joint effects of more than one factor [43,44].

In particular, the input factors in the present VBSA were the upper boundary f_{lim}^{UB}, the parameter α, the frequency step Δf, and the minimum size of the cluster n_{ele}, while considering the number of identified clusters as output, in the form of $n_{CL} = f\left(f_{lim}^{UB}, \alpha, \Delta f, n_{ele}\right)$. In addition, a further VBSA was performed on the effect of parameter β on the selection of the frequency associated to each cluster, as $f_{k_{CL}^*} = f(\beta)$.

The VBSA provides a framework whereby the variance of Y can be related to different sources of uncertainty in the inputs [40], as in the Sobol's approach in which two main sensitivity indexes quantify the variance of the output due to the uncertainty of the inputs, namely the first-order index (S^i) and the total effect (S_t^i) [45]. The first-order index (S^i) represents the direct contribution of the parameter X_i on the total variance $V(Y)$ of the observed output, which can be interpreted as the expected reduction in the total variance $V(Y)$ when X_i is fixed to a constant value [40]. The first-order index (S^i) are typically considered for factor prioritization. The total effect (S_t^i) is a consequential sensitivity metric which complements the first-order effect, and measures the effect of the parameter X_i in combination with all the other factors. The total effect is suit for factor fixing where insignificant inputs are set to a given value over their range of uncertainty [40]. The Sobol's first-order index (S^i) and total effect (S_t^i) are hereinafter calculated as [46]:

$$S^i = \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i))}{V(Y)} \tag{7}$$

$$S_t^i = \frac{E_{X_{\sim i}}(V_{X_i}(Y|X_i))}{V(Y)} = 1 - \frac{V_{X_{\sim i}}(E_{X_i}(Y|X_{\sim i}))}{V(Y)} \tag{8}$$

4. Validation of the IHCA method

To evaluate the stability of the proposed IHCA approach, variance-based global sensitivity analysis (VBSA) was performed, and the Sobol's indexes were calculated. The parameters which might affect the results of the IHCA, and thus to be investigated, were: *i*) the frequency that limits the range of interest $\left[f_{lim}^{LB}; f_{lim}^{UB}\right]$ (step 1), in particular, the lower boundary f_{lim}^{LB} was set to 0 Hz while varying the upper boundary f_{lim}^{UB}; *ii*) the parameter α for the calculation of the pair distance in terms of frequencies and mode shapes (step 2); *iii*) the frequency step Δf assumed to evaluate the number of peaks (step 4); *iv*) the minimum number of samples n_{ele} required to consider a cluster as valid (step 6); *v*) the parameter β to calculate the pair distance within each cluster in terms of frequency and modes for the identification of the most representative set of dynamic parameters (step 9). It is remarked that, in a process of

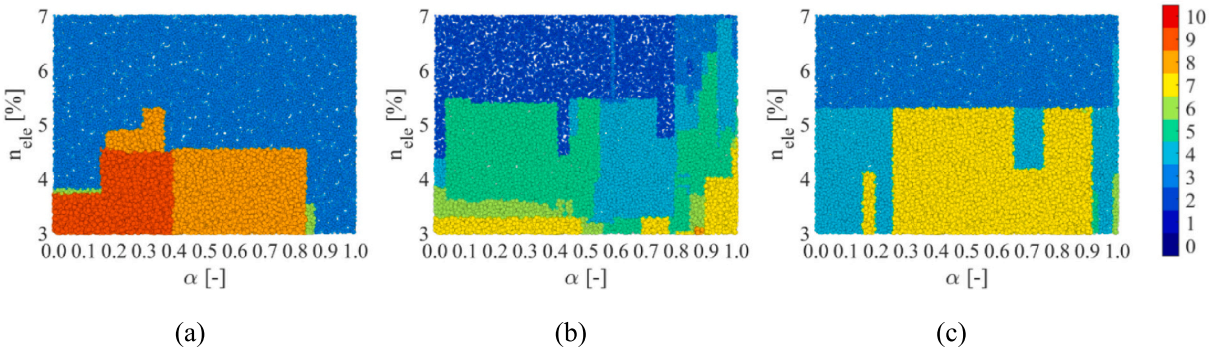


Fig. 11. Scatterplot of the individuated number of clusters (in colormap), as function of n_{ele} and α, for: a) DI01, b) DI02, and c) DI03.

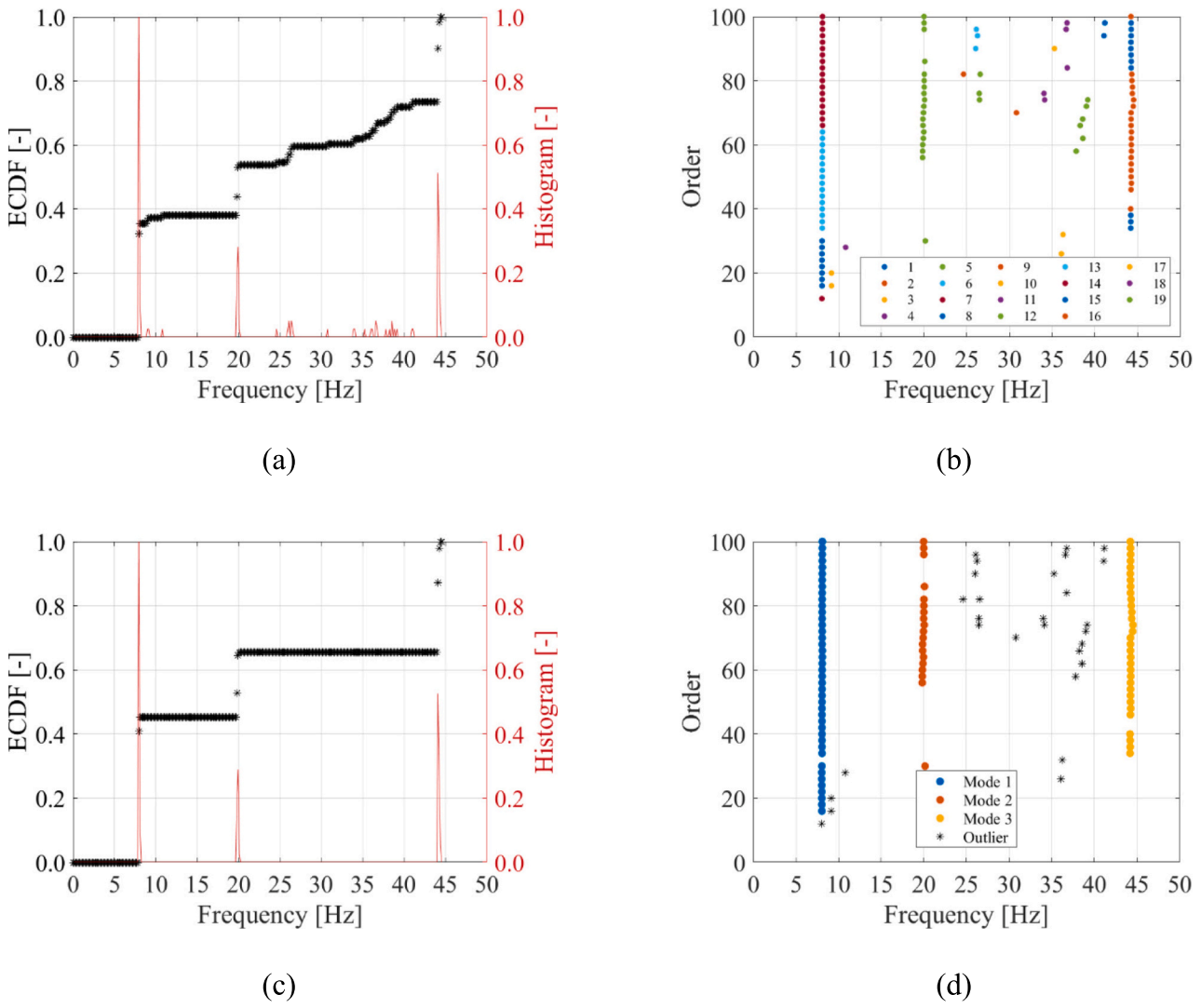


Fig. 12. Calibration of the experimental case DI01 with and IHCA parameters f_{lim}^{UB} = 45 Hz, Δf = 0.15 Hz, n_{ele} = 6% of total number of stable poles, α = 0.85 and β = 0.5: a) EDF and derivative of the first iteration, b) identified cluster at the first iteration, c) EDF and derivative of the last iteration, and c) identified cluster at the last iteration.

dynamic identification, the proposed IHCA approach considers as inputs data that were already defined as stable poles, as discussed in Section 2.

4.1. Case study: numerical model

The first case-study is the numerical model of a shear-type frame with five levels and four bays. At each level a mass of 20 tons is applied. The length of the beams is 5.00 m, and each storey is 3.00 m high with tapered columns, namely 50 × 50 cm^2, 45 × 45 cm^2, 40 × 40 cm^2, 35 × 35 cm^2, 30 × 30 cm^2. The Young's modulus of the material is 30 GPa. Rayleigh damping was considered, setting the damping ratio for the first and third mode to 0.1%. The mass matrix and the stiffness matrix of the numerical benchmark are reported in Table 1 and Table 2.

To simulate ambient vibrations, white noise acceleration with a duration of 600 s, with zero mean and unit standard deviation was applied at the base of the frame. The differential equations of motion were solved with the Newmark method with a time step dt = 0.002 s (500 Hz) and the response in terms of acceleration at each degree of freedom was recorded. In addition, in order to simulate measurement errors, white noise was added to recorded data; in particular, three different levels of noise were simulated considering increasing level of Signal-to-Noise Ratio (SNR), namely 0 dB, 25 dB and 50 dB; accordingly, the numerical simulations are hereinafter referred to as SNR0, SNR25, SNR50.

Therefore, SSI-CoV analysis was performed on the acceleration time histories obtained by numerical simulation. At first, the signals were filtered with a sixth-order pass-band filter in the range [0.5; 100] Hz, while the minimum order (n_{min}) and maximum order (n_{max}) of the SSI-CoV analysis were set to 4 and 70, respectively, with an order step of 2. According to the suggestion in [47], the time lag point (i) to be considered is a function of the expected fundamental period of the structure (T_0) and the sampling frequency (f_s), as $i \geq f_s T_0$. Being T_0 about 4 Hz, the number of instants considered for the time lag is 130. The stabilization diagrams for SNR0, SNR25, SNR50 numerical benchmarks are reported in Fig. 2.

Subsequently, the VBSA was conducted on IHCA for the SNR0, SNR25, SNR50 cases, where the observed output (Y) was the number of clusters representing the number of modes. The setting of the global sensitivity analysis focused on the factor prioritization and factor fixing. The total number of simulations was set to N = 10,000, while initially i = 4 parameters (f_{lim}^{UB}, Δf, n_{ele} and α) were investigated, namely in the domain f_{lim}^{UB}: [40; 60], Δf : [0.1; 0.3], n_{ele}: [3; 7], and α: [0; 1]. The parameter β was not considered since it influences the selection of the most representative dynamic properties once that the clusters are

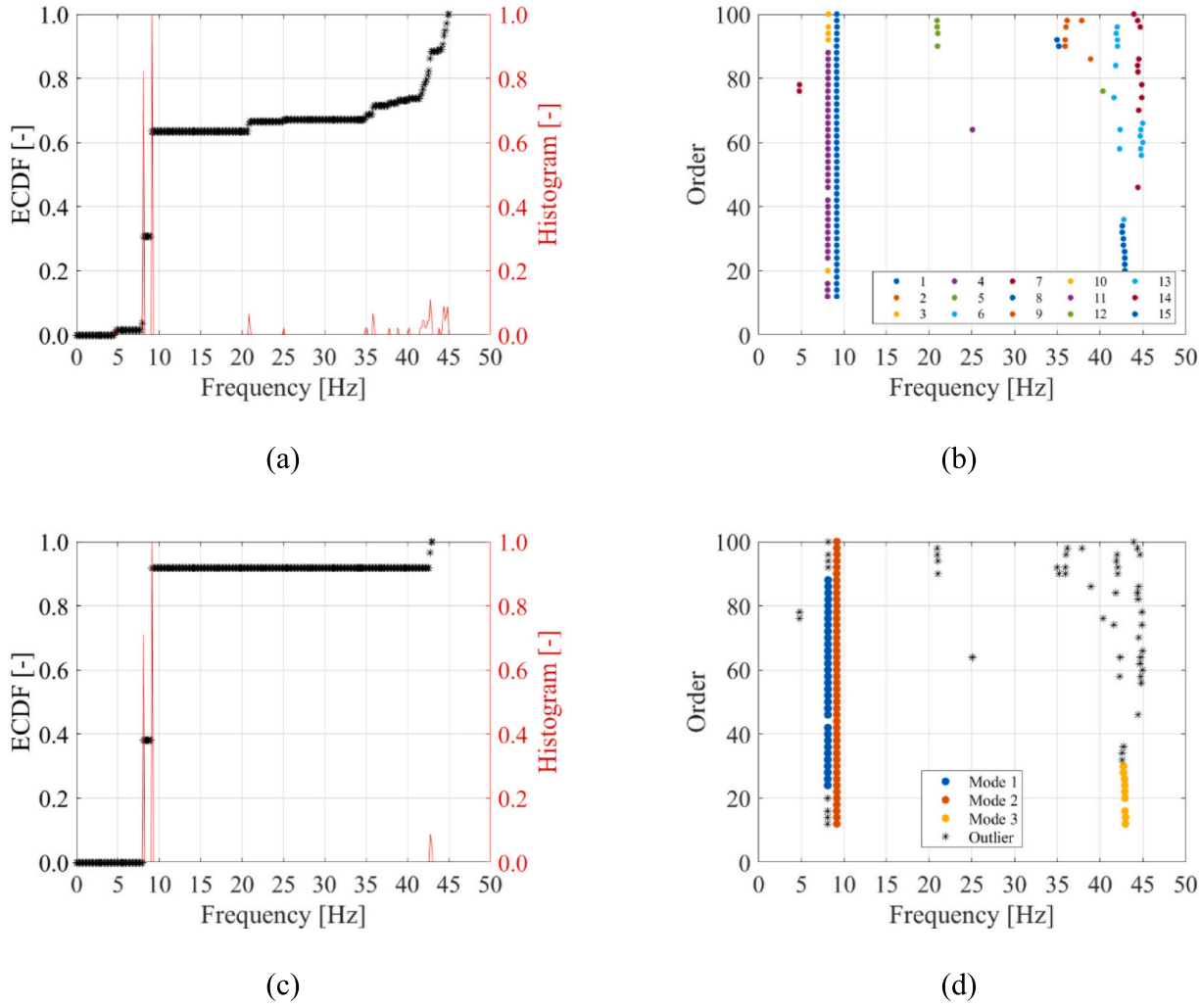


Fig. 13. Calibration of the experimental case DI02 with and IHCA parameters $f_{lim}^{UB} = 45$ Hz, $\Delta f = 0.15$ Hz, $n_{ele} = 6\%$ of total number of stable poles, $\alpha = 0.85$ and $\beta = 0.5$: a) EDF and derivative of the first iteration, b) identified cluster at the first iteration, c) EDF and derivative of the last iteration, and c) identified cluster at the last iteration.

grouped, and thus does not affect the individuation of the clusters. The values of the parameters for each simulation were generated as quasi-random low-discrepancy sequence points following Sobol's sequence [48,49]. Therefore, the first-order (S^i) and the total effect (S_t^i) sensitivity indexes were calculated according to Eq. (7) and Eq. (8).

As a result, first-order index (S^i) and the total effect (S_t^i) resulted null for all the parameters and for all the cases, except for the parameter n_{ele} and α which influenced the variance of the number of clusters only for the SNR50 case study. In fact, for the case SNR50, the first-order index and the total effect of the factor n_{ele} resulted 0.001 ($S^{n_{ele}}$) and 0.001 ($S_t^{n_{ele}}$), respectively. Moreover, the first-order index of the factor α resulted null, while the associated total effect 0.001 (S_t^{α}).

In order to define the influence of the parameters n_{ele} and α on the determination of the number of clusters, a further VBSA was conducted on the cases SNR50 prioritizing the parameters n_{ele} and α, while fixing $f_{lim}^{UB} = 40$ Hz and $\Delta f = 0.15$ Hz. Consequently, the first-order and the total effect indexes of the parameter n_{ele} and α resulted null in the entire domain investigated, confirming the stability and robustness of IHCA despite the level of SNR.

It is remarked that the VBSA provides insights on the output variance rather than the accuracy of the output. Aiming at validating the effectiveness of the algorithm, the IHCA was calibrated with the use of the VBSA against the numerical results of the modal analysis. At first, further VBSA was performed considering as output (Y) the frequency identified as the most representative for each cluster. With this regard, the factor β was investigated in the domain [0; 1], while the remaining parameters were selected in the domain previously investigated, whereby the variance of the model in terms of individuated clusters is null; thus, fixing $f_{lim}^{UB} = 40$ Hz, $\Delta f = 0.15$ Hz, $n_{ele} = 6\%$ and $\alpha = 0.5$. As a result, in all the case-studies, and considering all the identified modes, the model was not sensitive to parameter β in the entire domain [0; 1], meaning that the selection of the dynamic parameters is stable and robust once that the clusters are identified.

Subsequently, to validate the proposed IHCA for SSI-CoV, the modal parameters resulting from the numerical analysis were compared with the modal parameters identified in SNR0, SNR25, and SNR50. In this regard, based on the VBSA outcomes, the parameters of the IHCA were set as $f_{lim}^{UB} = 40$ Hz, $\Delta f = 0.15$ Hz, $n_{ele} = 6\%$ of total number of stable poles, $\alpha = 0.5$ and $\beta = 0.5$. Therefore, the stabilization diagrams resulting from SSI-CoV analysis and considering $f_{lim}^{UB} = 40$ Hz are presented in Fig. 3.

The steps of the IHCA for the cases SNR0, SNR25, and SNR50, are illustrated in Fig. 4, Fig. 5, and Fig. 6, respectively. After identifying the number of peaks (Fig. 4a, Fig. 5a, and Fig. 6a), as described for the step 4, the stable poles were grouped in the equivalent number of clusters and according to the distance as calculated in Eq. (5) (Fig. 4b, Fig. 5b,

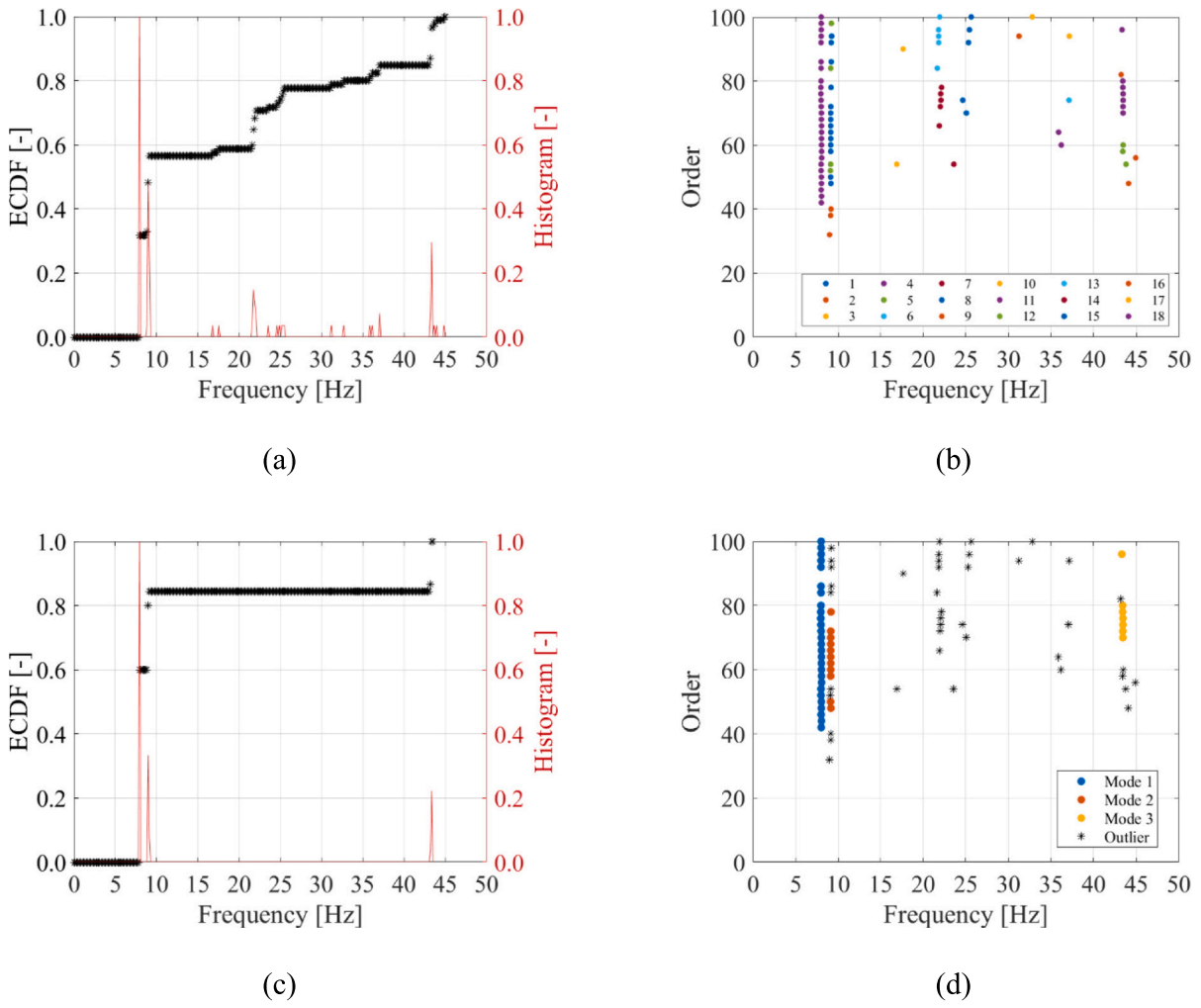


Fig. 14. Calibration of the experimental case DI03 with and IHCA parameters $f_{lim}^{UB} = 45$ Hz, $\Delta f = 0.15$ Hz, $n_{ele} = 6\%$ of total number of stable poles, $\alpha = 0.85$ and $\beta = 0.5$: a) EDF and derivative of the first iteration, b) identified cluster at the first iteration, c) EDF and derivative of the last iteration, and c) identified cluster at the last iteration.

and Fig. 6b). Therefore, the requirement of minimum number of objects for cluster was checked, as in the step 6. In general, clusters which did not met the set requirement (n_{ele}) were found at the first step of the iterative process. Consequently, the poles of those clusters were labelled as outlier and removed from the overall dataset. The IHCA continued with the second iteration, and the number of peaks was re-calculated (Fig. 4c, Fig. 5c, and Fig. 6c). Once that all the identified clusters met the requirement of minimum number of samples, the iterative process stopped (Fig. 4d, Fig. 5d, and Fig. 6d). The most representative dynamic properties were selected for each cluster according to Eq. (6), as described for the step 7.

Finally, to validate the effectiveness of SSI-CoV and IHCA in identifying the correct modes, the dynamic properties selected at the stage 7 were compared against the numerical solution in terms of MAC, frequency and relative error (see Table 3). The results show consistency between the numerical solution and the system identification, as the relative error of frequency ($f_{Err\%}$) does not exceed 1.2% for a consistency in terms of MAC equal to 0.92 in case of Mode 5.

4.2. Case study: vibrations of a masonry building

The second case study was a series of laboratory tests on a two-storey masonry building mock-up with plan size of 3.72*X*3.50 m^2 and total height of 4.67 *m*. The structure was not regular in plan and regular in elevation, with one opening on the north façade (Fig. 7a) and two openings on the south façade (Fig. 7b and Fig. 7c), at each level. The openings were 0.85 *m* wide and 1.25 *m* high, with timber lintels with size 0.25*X*0.10*X*1.05 m^3. The masonry walls were flemish bond brickwork with clay bricks and lime mortar with total thickness of 0.25 *m*. The first level was a polycentric vault 0.125 *m* thick with a timber floor of double plank. Four steel S235 tie-rods ϕ18 were insert across the vault with two square anchor keys with dimensions 0.40*X*0.40 m^2 and thickness of 0.025 *m*. The second level was a timber floor consisting of softwood beams with cross section 0.05*X*0.25 m^2 and spacing 1.00 *m*, and single plank. Further details can be found in [50].

The dynamic monitoring setup consisted of twelve piezoelectric accelerometers (model PCB 393B12, 0.15 to 1000 Hz frequency range, 10,000 mV/g sensitivity, 8 μg resolution) with sampling rate of 500 Hz and a total duration of each signal of 600 s (Fig. 8). Five accelerometers were placed on the west façade to monitor the in-plane (A02 and A07) and out-of-plane (A01, A06 and A08) modes of the wall; two accelerometers were set on the east façade to record out-of-plane modes (A11 and A12). Moreover, five accelerometers were placed on the north façade to detect out-of-plane (A03, A04 and A09) and in-plane (A05 and A10) modes.

SSI-CoV analysis was performed on the recorded accelerations, produced by ambient vibrations. At first, the signals were filtered with a sixth-order pass-band filter in the range [0.5; 100] Hz, while the minimum order (n_{min}) and maximum order (n_{max}) of the SSI-CoV analysis

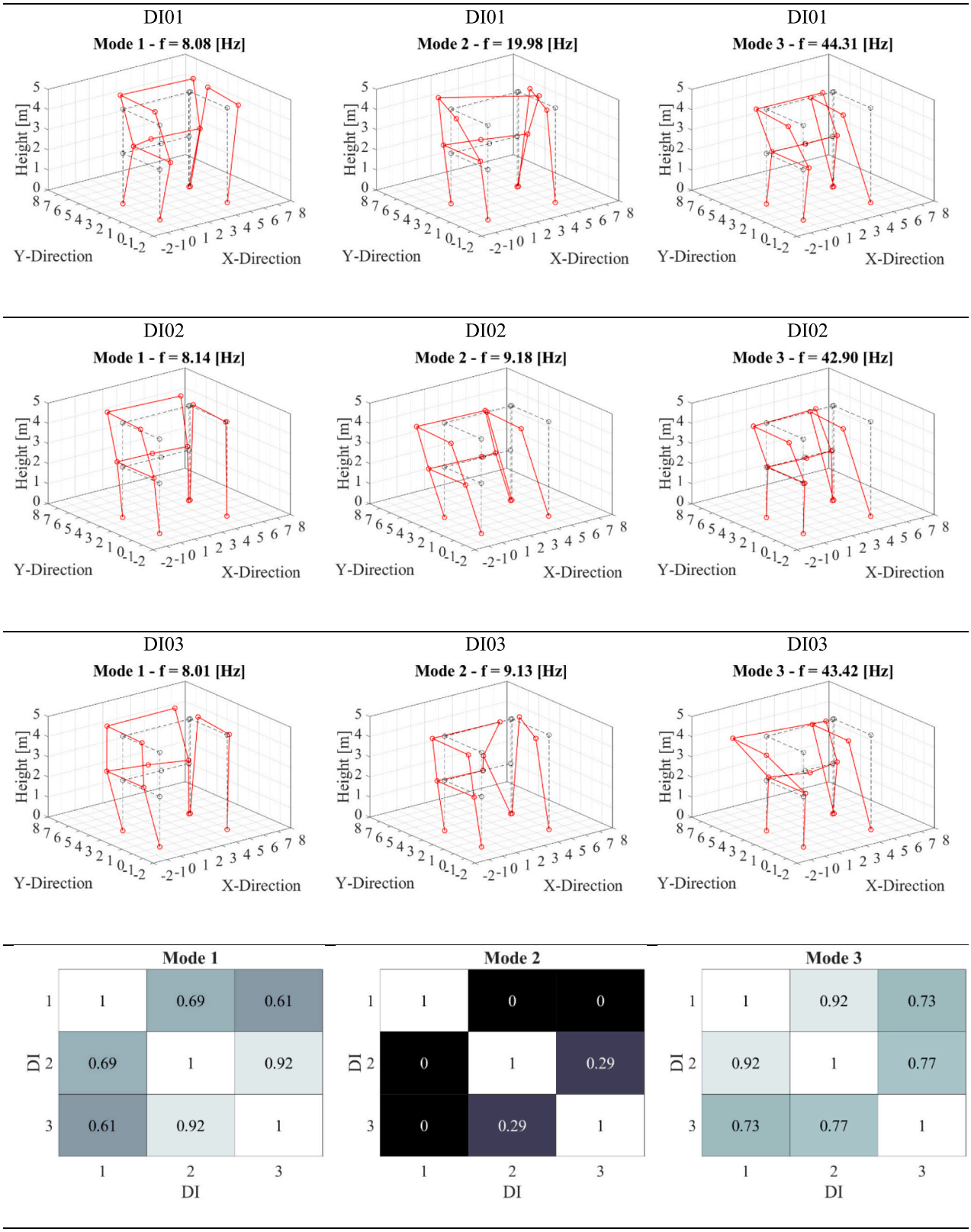


Fig. 15. Frequency and mode shapes identified, and MAC between the case DI01, D02, and DI03.

were set to 4 and 100, respectively, with an order step of 2. The number of instants considered for the time lag is 70. The stable poles in the SSI-CoV outcomes were detected following the procedure previously described in Section 2. For the sake of brevity, the dynamic identification and stability diagram of three different sets of 600 s recorded accelerations are here reported and referred to as DI01, DI02, and DI03 (Fig. 9).

As in the numerical benchmark, the global sensitivity analysis focused on the factor prioritization and factor fixing, and the observed output (Y) was the number of clusters. The value of f_{lim}^{UB} was fixed to 45

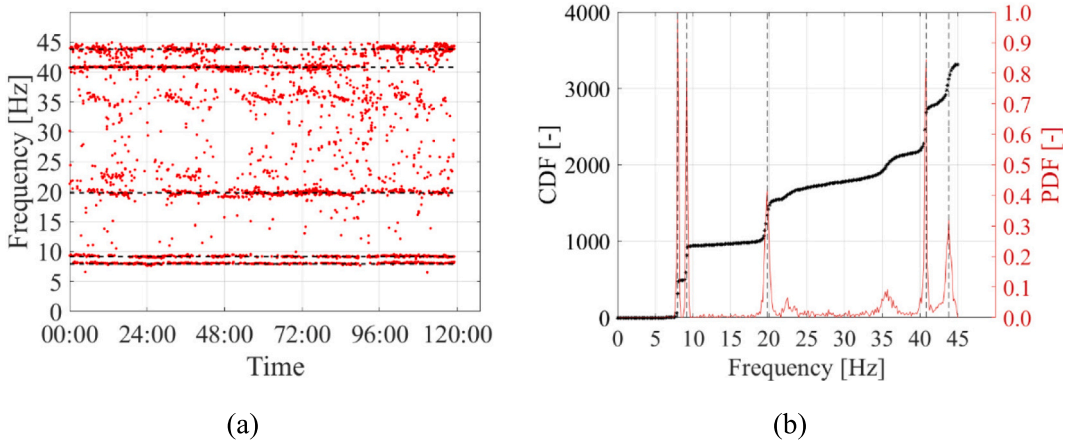


Fig. 16. SSI-CoV and IHCA analysis performed on data recorded during five days: a) identified frequencies in each time window, and b) density of identified frequencies.

Hz based on previous investigations on the dynamics of the building, and β was discarded because it does not affect the number of clusters. Thus, initially $i = 3$ parameters (Δf, n_{ele}, and α) were investigated. The total number of simulations was set to $N = 10{,}000$. The values of the parameters for each simulation were generated as quasi-random low-discrepancy sequence points following Sobol's sequence [48,49]. Therefore, the first-order (S^i) and the total effect (S_t^i) sensitivity indexes were calculated according to Eq. (7) and Eq. (8). Table 4 reports, for each parameter, the domain and the sensitivity indexes of the VBSA performed on the DI01, DI02, and DI03 recorded acceleration cases. In addition, Fig. 10 represents the number of clusters in colormap as function of the variables Δf, n_{ele}, and α. The parameter n_{ele} resulted the dominant factor influencing the variance in terms of number of modes, as the first-order ($S^{n_{ele}}$) and the total effect ($S_t^{n_{ele}}$) indexes described, respectively, the 8.1% and 12.3% of the variance in case of DI01, the 10.9% and 14.4% of the variance in case of DI02, and the 7.7% and 10.7% of the variance in case of DI03. Moreover, the factor α affected the variance of the observation, in particular in combination with parameters Δf and n_{ele}, as reported by the related first-order order (S^{α}) and the total effect (S_t^{α}) indexes. Finally, lower influence was observed for the parameter Δf.

Consequently, to reduce the output variance and considering the results of the VBSA, the factor Δf was fixed while prioritizing further sensitivity analysis on the parameters n_{ele} and α. Based on the stabilization diagrams, the Δf value to individuate the peaks, and hence the likely number of clusters, was selected equal to 0.15 Hz, as larger Δf value might merge two separate modes (see Fig. 9b). Therefore, a further sensitivity analysis was performed on total number of simulations of $N = 10{,}000$ and number of parameters $i = 2$ (n_{ele} and α), and the results are reported in Table 5 and Fig. 11. As summarized in Table 5, for all the DI cases, the variance in the number of identified modes is mainly ascribed to the factor n_{ele}, as the first-order ($S^{n_{ele}}$) and the total effect ($S_t^{n_{ele}}$), resulted above 10.5% and 13.1%, respectively. It should be noted that the results of the IHCA are consequent to the definition of stable poles according to the criteria previously mentioned, which can be further sources of variance for the observed number of clusters.

Aiming at comparing the identified modes between DI01, DI02 and DI03, and validating the HICA for the experimental case study, the variance in the number of clusters was reduced by fixing the factors n_{lim} and α to the value of 6% and 0.85, respectively. In fact, as observed in Fig. 11, in the neighbourhood of the fixed factors, the number of modes is equal to the different cases. Afterwards, VBSA was performed by observing, for each individuated mode, the selected most representative frequency as function of the factor β in the domain [0; 1]. The first-order and total effect sensitivity indexes resulted substantially null for all case-studies and mode, meaning that the variable β did not influence the variance of the selected frequency. Therefore, as already observed in the numerical benchmark, the selection of the dynamic parameters is stable and robust once that the clusters are identified.

Thereafter, the IHCA was calibrated for the experimental dataset based on the results of the VBSA. Thereby, the parameters of the IHCA were set as $f_{lim}^{UB} = 45$ Hz, $\Delta f = 0.15$ Hz, $n_{ele} = 6\%$ of total number of stable poles, $\alpha = 0.85$ and $\beta = 0.5$. Since the IHCA performed various iterations in order to satisfy the n_{lim} requirement, only the first iteration and the convergence steps are illustrated in Fig. 12, Fig. 13, and Fig. 14, for the cases DI01, DI02, and DI03, respectively. Fig. 12a, Fig. 13a, and Fig. 14a report the number of identified peaks (step 4 of IHCA), which represent the likely number of clusters as grouped in Fig. 12b, Fig. 13b, and Fig. 14b. Since the required minimum number of objects (n_{ele}) forming each cluster was not met, as described for the step 6, the iterative process continued removing those clusters from the entire dataset (step 6.1 of IHCA). In general, the number of iterations performed to complete the cluster analysis was two for DI01 and DI03, and four in case of DI02, while the peaks and the clusters are represented in Fig. 12c, Fig. 13c, and Fig. 14c, and Fig. 12d, Fig. 13d, and Fig. 14d, respectively.

Finally, to validate the effectiveness of SSI-CoV and IHCA, the frequencies and the mode shapes identified for DI01, DI02, and DI03 were analysed, see Fig. 15. As a result, the first frequency (f_1) was similar in all the cases in the range $f_1 = [8.01; 8.14]$ Hz, which corresponded to the first flexural mode in Y direction. The frequency associated to the second mode (f_2) was comparable between the cases DI02 and DI03, resulting $f_2 = 9.18$ Hz and $f_2 = 9.13$ Hz, respectively, which corresponded to the first flexural mode in X direction, and the third frequency (f_3) resulted similar in all the cases in the range $f_3 = [42.90; 44.31]$ Hz, which corresponded to the second flexural mode in X direction. A distinct value of f_2 was obtained for the case DI01, namely $f_2 = 19.98$ Hz, which corresponded to the torsional mode. Considering the MAC calculated for each mode shape between the three cases, similarity was found for the Mode 1 between DI02 and DI03 (MAC = 0.92), and for Mode 3 between DI01 and DI02 (MAC = 0.92). Nevertheless, low correspondence can be observed for the Mode 1 between DI01 and DI02 (MAC = 0.69), and DI03 (MAC = 0.61), and for Mode 3 between DI03 and DI01 (MAC = 0.73), and DI02 (MAC = 0.77). However, no similarity was found in any case for the mode associated to the second frequency (MAC < 0.29), as showed in Fig. 15. Such result is explained by the fact that the second frequency identified in DI01 corresponds to a distinct mode which was not identified from DI02 and DI03. In fact, as the IHCA is applied on SSI-CoV results, the recorded signals and the criteria to individuate the stable poles might affect the results of the singular dynamic identification, which can be otherwise regulated by a large number of analyses

along the time.

In order to verify this circumstance, additional SSI-CoV and IHCA analyses were performed on data continuously recorded during five days, which corresponded to about 720 records with a duration of 10 min. The setting of SSI-CoV and IHCA was the same as for the previous cases study DI01, DI02, and DI03. As a result, Fig. 16a reports the frequencies identified for each time window, while Fig. 16b summarises the EDF of detected frequencies. It is observed that, performing the automated dynamic identification through SSI-CoV and IHCA on a big database, the potential incongruence in the results is adjusted by the large dataset. In fact, the identified peaks in the EDF of 720 records are $f_1 = 7.95$ Hz, $f_2 = 9.15$ Hz, $f_3 = 19.8$ Hz, $f_4 = 40.80$ Hz, and $f_5 = 43.80$ Hz, and thus the frequencies determined from the case DI01, DI02, and DI03, results to be reliable for the experimental case study, remarking the stability and robustness of the IHCA.

5. Conclusion

An innovative hierarchical cluster analysis based on an iterative process is proposed. The method is fully automated and aims at supporting the selection of dynamic properties resulting from SSI-CoV analysis. To evaluate the method and determine which parameters of IHCA might influence the dynamic identification results, a variance-based global sensitivity analysis (VBGA) was performed on numerical benchmark considering further white noise in signals with different amplitude, and on experimental data.

The outcomes of the numerical simulations report that the identification of clusters is sensitive to the α factor to calculate the distance for clustering in terms of frequency and MAC. Nevertheless, the variance observing the frequency selected once that the clusters are formed is null, meaning that the IHCA is stable despite the amplitude of the white noise. In addition, the comparison between a numerical modal analysis and the dynamic identification from SSI-CoV with the use of the IHCA demonstrated the reliability of the IHCA method.

With regard to the experimental case study, the outcome in terms of number of identified modes is sensitive to the various parameters of the IHCA. It was observed that the output variance can be reduced by conveniently fixing the minimum number of elements required to initialise a cluster (n_{ele}), and that the weighting coefficient α to measure the distance between stable poles might require a proper selection depending on the case study. With regard to the selected frequencies, the sensitivity of IHCA to the β parameters resulted null; therefore, the IHCA is stable and robust in selecting the representative frequency once that the clusters are identified. Finally, a comparison between the results of the various dynamic identification remarks that in general similar frequencies and mode shapes were found among the various dataset, although dissimilarity was found for the Mode 2 identified in DI02. Such difference can be ascribed in particular to the recorded signals and the requirements for individuating stable poles, which can be regulated by considering different criteria or increasing the number of SSI-CoV and IHCA analyses to larger dataset. In fact, applying the automatic mode selection by IHCA on a dataset of 720 SSI-CoV, results demonstrated that the identified mode in DI01, DI02, and DI03, were all reliable. Therefore, the proposed IHCA is considered stable and robust.

Statement and declarations

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Author contribution

Romanazzi A. – Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Scocciolini D.** – Numerical simulation, Data curation, Writing - original draft. **Savoia M.** – Funding acquisition, Methodology, Supervision, Writing - review. **Buratti N.** – Funding acquisition, Methodology, Project administration, Supervision, Writing - review.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgment

This work is part of the "DETECT-AGING - Degradation Effects on sTructural safEty of Cultural heriTAGe constructions through simulation and health monitoring" project, funded by the Italian Ministry of University and Research through PRIN: Progetti di Ricerca di Rilevante Interesse Nazionale – Call 2017 – Prot. 201747Y73L. Acknowledgments are addressed to the CIRI Laboratory of Buildings and Structures of the University of Bologna, to Eng. Giorgio Virgulto and to the RELUIS Consortium (DPC-RELUIS 22-24 research agreement, WP6-Monitoring and satellite data).

References

[1] H.P. Chen, Y.Q. Ni, Structural Health Monitoring of Large Civil Engineering Structures, John Wiley & Sons Ltd, Hoboken, New Jersey, United States, 2018. ISBN: 9781119166634.
[2] F. Ubertini, G. Comanducci, N. Cavalagli, Vibration-based structural health monitoring of a historic bell-tower using output-only measurements and multivariate statistical analysis, Struct. Health Monit. 15 (4) (2016) 438–457, https://doi.org/10.1177/14759217166439.
[3] L.F. Ramos, L. Marques, P.B. Lourenço, G. De Roeck, A. Campos-Costa, J. Roque, Monitoring historical masonry structures with operational modal analysis: two case studies, Mech. Syst. Signal Process. 24 (2010) 1291–1305, https://doi.org/10.1016/j.ymssp.2010.01.011.
[4] Z. He, W. Li, H. Salehi, H. Zhang, H. Zhou, P. Jiao, Integrated structural health monitoring in bridge engineering, Autom. Constr. 136 (2022), 104168, https://doi.org/10.1016/j.autcon.2022.104168.
[5] F.N. Catbas, M. Malekzadeh, A machine learning-based algorithm for processing massive data collected from the mechanical components of movable bridges, Autom. Constr. 72 (3) (2016) 269–278, https://doi.org/10.1016/j.autcon.2016.02.008.
[6] T. Peng, M. Nogal, J.R. Casas, J. Turmo, Planning low-error SHM strategy by constrained observability method, Autom. Constr. 127 (2021), 103707, https://doi.org/10.1016/j.autcon.2021.103707.
[7] E. García-Macías, F. Ubertini, MOVA/MOSS: Two integrated software solutions for comprehensive structural health monitoring of structures, Mech. Syst. Signal Process. 143 (2020), 106830, https://doi.org/10.1016/j.ymssp.2020.106830.
[8] G. Toh, J. Park, Review of vibration-based structural health monitoring using deep learning, Appl. Sci. 10 (2020) 1680, https://doi.org/10.3390/app10051680.
[9] F. Magalhães, A. Cunha, E. Caetano, Vibration based structural health monitoring of an arch bridge: from automated OMA to damage detection, Mech. Syst. Signal Process. 28 (2012) 212–228, https://doi.org/10.1016/j.ymssp.2011.06.011.
[10] A. Entezami, H. Sarmadi, B. Behkamal, S. Mariani, Big data analytics and structural health monitoring: a statistical pattern recognition-based approach, Sensors 20 (2020) 2328, https://doi.org/10.3390/s20082328.
[11] F. Magalhaes, A. Cunha, E. Caetano, Online automatic identification of the modal parameters of a long span arch bridge, Mech. Syst. Signal Process. 23 (2009) 316–329, https://doi.org/10.1016/j.ymssp.2008.05.003.
[12] C.X. Qu, Y.F. Liu, T.H. Yi, H.N. Li, Structural damping ratio identification through iterative frequency domain decomposition, J. Struct. Eng. 149 (5) (2023) 04023042, https://doi.org/10.1061/JSENDH.STENG-11837.
[13] C.X. Qu, T.H. Yi, H.N. Li, Mode identification by eigensystem realization algorithm through virtual frequency response function, Struct. Control. Health Monit. 26 (2019) 10, https://doi.org/10.1002/stc.2429.
[14] B. Peeters, G. De Roeck, Stochastic system identification for operational modal analysis: a review, J. Dyn. Syst. Meas. Control. 123 (2001) 659–667, https://doi.org/10.1115/1.1410370.
[15] R. Brincker, C.E. Ventura, Introduction to Operational Modal Analysis, John Wiley & Sons Ltd, Hoboken, New Jersey, United States, 2015. ISBN: 9781118535141.
[16] L. Garibaldi, S. Marchesiello, E. Bonisoli, Identification and up-dating over the Z24 benchmark, Mech. Syst. Signal Process. 17 (2003) 153–161, https://doi.org/10.1006/mssp.2002.1553.

[17] J.B. Bodeux, J.C. Golinval, Application of ARMAV model to the identification and damage detection of mechanical and civil engineering structures, Smart Mater. Struct. 10 (2001) 479–489, https://doi.org/10.1088/0964-1726/10/3/309.
[18] C. Rainieri, G. Fabbrocino, Operational Modal Analysis of Civil Engineering Structures. An Introduction and Guide for Applications, Springer, New York, Heidelberg, Dordrecht, London, 2014. ISBN 978-1-4939-0767-0.
[19] F. Ubertini, C. Gentile, A.L. Materazzi, Automated modal identification in operational conditions and its application to bridges, Eng. Struct. 46 (2013) 264–278, https://doi.org/10.1016/j.engstruct.2012.07.031.
[20] J.-N. Juang, R.S. Pappa, An eigensystem realization algorithm for modal parameter identification and model reduction, J. Guid. Control. Dyn. 8 (1985) 620–627, https://doi.org/10.2514/3.20031.
[21] R.S. Pappa, G.H. James III, D.C. Zimmerman, Autonomous Modal Identification of the Space Shuttle Tail Rudder, Report NASA TM-112866,, National Aeronautics and Space Administration, 1997, https://ntrs.nasa.gov/api/citations/19970025582/downloads/19970025582.pdf.
[22] S. Chauhan, D. Tcherniak, Clustering approaches to automatic modal parameter estimation, in: Proceedings of the 27th International Modal Analysis Conference, Orlando, Florida, 2009, pp. 1072–1085. ISBN: 9781605609614.
[23] I. Goethals, B. Vanluyten, B. De Moor, Reliable spurious mode rejection using self learning algorithms, in: Proceedings of International Conference on Noise and Vibration Engineering, Leuven, Belgium, 2004. ISBN: 9073802822.
[24] R.J. Allemang, D.L. Brown, A.W. Phillips, Survey of modal techniques applicable to autonomous/semiautonomous parameter identification, in: Proceedings of International Conference on Noise and Vibration Engineering, Leuven, Belgium, 2010. ISBN: 9789073802872.
[25] E. Reynders, J. Houbrechts, G. De Roeck, Fully automated (operational) modal analysis, Mech. Syst. Signal Process. 29 (2012) 228–250, https://doi.org/10.1016/j.ymssp.2012.01.007.
[26] M. Scionti, J. Lanslots, I. Goethals, A. Vecchio, H. Van der Auweraer, B. Peeters, B. De Moor, Tools to improve detection of structural changes from in-flight flutter data, in: Proceedings of the of the 8th International Conference on Recent Advances in Structural Dynamics, Southampton, United Kingdom, 2003. ISBN: 0854327894.
[27] M. Scionti, J.P. Lanslots, Stabilization diagrams: pole identification using fuzzy clustering techniques, Adv. Eng. Softw. 36 (2005) 768–779, https://doi.org/10.1016/j.advengsoft.2005.03.029.
[28] S. Vanlanduit, P. Verboven, P. Guillaume, J. Schoukens, An automatic frequency domain modal parameter estimation algorithm, J. Sound Vib. 265 (2003) 647–661, https://doi.org/10.1016/S0022-460X(02)01461-X.
[29] P. Verboven, E. Parloo, P. Guillaume, M. Van Overmeire, Autonomous structural health monitoring - part 1: modal parameter estimation and tracking, Mech. Syst. Signal Process. 16 (2002) 637–657, https://doi.org/10.1006/mssp.2002.1492.
[30] P.-É. Charbonnel, Fuzzy-driven strategy for fully automated modal analysis: application to the SMART2013 shaking-table test campaign, Mech. Syst. Signal Process. 152 (2021), 107388, https://doi.org/10.1016/j.ymssp.2020.107388.
[31] J. Li, T. Bao, C.E. Ventura, An automated operational modal analysis algorithm and its application to concrete dams, Mech. Syst. Signal Process. 168 (2022), 108707, https://doi.org/10.1016/j.ymssp.2021.108707.
[32] J. Zeng, Z. Hu, Automated operational modal analysis using variational Gaussian mixture model, Eng. Struct. 273 (2022), 115139, https://doi.org/10.1016/j.engstruct.2022.115139.
[33] G. Zini, M. Betti, G. Bartoli, A quality-based automated procedure for operational modal analysis, Mech. Syst. Signal Process. 164 (2022), 108173, https://doi.org/10.1016/j.ymssp.2021.108173.
[34] E. Tomassini, E. García-Macías, E. Reynders, F. Ubertini, Model-assisted clustering for automated operational modal analysis of partially continuous multi-span bridges, Mech. Syst. Signal Process. 200 (2023), 110587, https://doi.org/10.1016/j.ymssp.2023.110587.
[35] P. Verboven, B. Cauberghe, E. Parloo, S. Vanlanduit, P. Guillaume, User-assisting tools for a fast frequency domain modal parameter estimation method, Mech. Syst. Signal Process. 18 (2004) 759–780, https://doi.org/10.1016/S0888-3270(03)00053-0.
[36] J. Hair, R. Anderson, R. Tatham, W. Black, Multivariate Data Analysis, 7th edition, Prentice Hall, New Jersey, 2010. ISBN: 0138132631.
[37] A. Saltelli, I.M. Sobol, About the use of rank transformation in sensitivity analysis of model output, Reliab. Eng. Syst. Saf. 50 (1995) 225–239, https://doi.org/10.1016/0951-8320(95)00099-2.
[38] A. Saltelli, S. Tarantola, F. Campolongo, Sensitivity analysis as an ingredient of modeling, Stat. Sci. 15 (2000) 377–395. http://www.jstor.org/stable/2676831.
[39] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis, The Primer. John Wiley & Sons Ltd, Hoboken, New Jersey, United States, 2008. ISBN: 9780470725184.
[40] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models, John Wiley & Sons Ltd, Hoboken, New Jersey, United States, 2004. ISBN: 978-0-470-87094-5.
[41] A. Saltelli, M. Saisana, Settings and methods for global sensitivity analysis – a short guide, Proc. Appl. Math. Mech. 7 (2007) 2140013–2140014, https://doi.org/10.1002/pamm.200700986.
[42] A. Saltelli, K. Chan, M. Scott, Sensitivity Analysis, Probability and Statistics Series, John Wiley & Sons Ltd, Hoboken, New Jersey, United States, 2000. ISBN: 978-0-470-74382-9.
[43] H. Rabitz, Global sensitivity analysis for systems with independent and/or correlated inputs, Procedia Soc. Behav. Sci. 2 (2010) 7587–7589, https://doi.org/10.1016/j.sbspro.2010.05.131.
[44] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun. 145 (2002) 280–297, https://doi.org/10.1016/S0010-4655(02)00280-1.
[45] I.M. Sobol', Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul. 55 (2001) 271–280, https://doi.org/10.1016/S0378-4754(00)00270-6.
[46] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Comput. Phys. Commun. 181 (2010) 259–270, https://doi.org/10.1016/j.cpc.2009.09.018.
[47] E. Reynders, G. De Roeck, Reference-based combined deterministic–stochastic subspace identification for experimental and operational modal analysis, Mech. Syst. Signal Process. 22 (3) (2008) 617–637, https://doi.org/10.1016/j.ymssp.2007.09.004.
[48] I.M. Sobol, On the distribution of points in a cube and the approximate evaluation of integrals, USSR Comput. Math. Math. Phys. 7 (1967) 86–112, https://doi.org/10.1016/0041-5553(67)90144-9.
[49] P. Bratley, B.L. Fox, Algorithm 659 Implementing Sobol's quasirandom sequence generator, Trans. Math. Softw. 14 (1988) 88–100, https://doi.org/10.1145/42288.214372.
[50] G. Virgulto, A. Meoni, A. D'Alessandro, N. Buratti, F. Ubertini, The Detect-Aging prototype building, a benchmark for structural health monitoring of masonry structures, Int. J. Masonry Res. Innov. (2023) (Accepted for publication).