Journal of Vibration Engineering & Technologies (2024) 12:4313–4333
https://doi.org/10.1007/s42417-023-01121-1

ORIGINAL PAPER

An Automated Procedure for Continuous Dynamic Monitoring of Structures: Theory and Validation

Y. V. Ardila[1,2] · I. D. Gómez-Araújo[1,2] · J. D. Villalba-Morales[1,2]

Received: 25 November 2022 / Revised: 11 July 2023 / Accepted: 9 August 2023 / Published online: 9 September 2023
© The Author(s) 2023

Abstract

Introduction Stability diagrams are a helpful tool for operational modal analysis to obtain the physical modes of a structure. These modes can be defined by visualizing stable columns formed by consistently identified modes over a range of model orders. Extracting these modes manually becomes an obstacle if continuous identification of modal parameters is required.

Materials and methods In this paper, a procedure is configured to automatically interpret the stability diagrams constructed with the identification results of the SSI-COV/ref algorithm. This procedure is based on some methodologies found in the literature, which follow three stages. First, a stability diagram cleaning stage is defined where modal validation criteria and partitioning clustering algorithms are used to detect spurious modes. Second, a mode grouping stage based on a hierarchical clustering algorithm is implemented to form sets of modes that share similar modal information. Finally, a selection stage is applied to define representative modal parameters from the set of physical modes.

Results The proposed procedure is validated by simulating a beam-type structural model with ten degrees of freedom affected by ambient temperature functions. Natural frequencies computed for the DT140 and DT220 datasets collection with the frequency-domain decomposition method agree with the computed ones with the proposed procedure, presenting MAC coefficients higher than 0.97. A total of 192 datasets are simulated, and the acceleration responses are polluted with two noise levels, SRN = 40 [dB] and SNR = 20 [dB].

Conclusions For the analyzed beam, the modal tracking results showed that the procedure could perform continuous identification automatically. The variations in the natural frequencies are correlated to the variations in the ambient temperature functions.

Keywords Modal parameters · Operational modal analysis · Dynamic monitoring · Clustering · Modal identification

Introduction

The advantages that offer operational modal analysis (OMA) have led to its increasing study and application in the mechanical [1, 2], aerospace, [3–5], and civil [6, 7] fields. As OMA has been developed to identify modal parameters (natural frequencies, damping ratios, and mode shapes) from ambient vibration measurements, no equipment is required to excite the structures, and the tests can be conducted under operational conditions. Furthermore, applying this type of test makes it possible to obtain the entire system's dynamic characteristics and effectively identify repeated and closely space modes [8]. The purpose of identifying modal parameters from OMA is to obtain reliable information that can be used in different applications, such as dynamic structural modification, numerical model updating, optimal dynamic design, vibration control, and structural health monitoring (SHM). Although

I. D. Gomez-Araujo and J. D. Villalba-Morales have contributed equally to this work.

✉ J. D. Villalba-Morales
jesus.villalba@javeriana.edu.co

Y. V. Ardila
yeny.ardila@correo.uis.edu.co

I. D. Gómez-Araújo
ivan.araujo@unila.edu.br

[1] Deparment of Civil and Infrastructure Engineering, Federal University of Latin American Integration, Foz do Iguaçu, Parana, Brazil

[2] Departamento de Ingeniería Civil, Facultad de Ingeniería, Pontificia Universidad Javeriana, Bogota, Cundinamarca, Colombia

vibration-based SHM is becoming more and more popular, the implementation of such systems is not an easy task. These must contain the dynamic characterization of structures in real time to observe and detect changes in the temporal evolution of modal parameters to be used as damage indicators. In this way, continuous or long-term identification entails a significant challenge as it requires periodic analysis of large amounts of data, ensuring the reliability of the estimated modal information. Much human input makes this process inefficient, so automation of OMA techniques is essential.

Several numerical algorithms have been developed to perform OMA. Some of them, in the frequency domain, such as the frequency-domain decomposition (FDD) [9] or the least-squares complex frequency domain (LSCF) [10], and in the time domain, such as the stochastic subspace identification algorithms (SSI-DATA, SSI-COV). The SSI-COV algorithm has been successfully used to identify modal parameters of large civil structures [11–13]. One of the benefits of this algorithm is that it allows the simultaneous identification of multiple modes, which makes it quite efficient for continuous dynamic characterization. Notwithstanding, the main issue in using the SSI-COV is to select a proper model order. It is usually recommended to set this parameter by detecting a large gap between two successive singular values, but this gap is difficult to detect when using ambient vibration measurements. Different strategies have been developed to determine an optimal model order, and some of them can be consulted in [14–16]. However, the exact value of the model order will never be known, and applying the above strategies leads to a high probability of obtaining underestimated or overestimated values of this parameter. In the first scenario, not all the physical modes of the system would be identified. On the contrary, overestimating the order would result in noise absorption and thus in modal characteristics represented by both physical and spurious modes.

Stability diagrams are a useful tool to define the true modes of a structure. Thus, the SSI-COV algorithm can be applied to a range of model orders. Then, the identification results can be displayed on a frequency vs. model order plane. The physical modes are expected to be identified consistently in each order. Since the orders are over-specified beyond a certain order, identifying spurious modes is inevitable. However, unlike the physical modes, they are not consistently identified in each order and therefore appear in the stability diagram as scattered points. Therefore, vertical alignments or stable columns can be defined as the physical modes. Manually extracting these modes can be laborious and can easily lead to errors in the modal identification process. Then, estimating modal parameters without user interaction demands the automatic interpretation of the stability diagram.

The first challenge faced when interpreting stability diagrams is the detection of spurious modes. Commonly, the degree of physicality of the identified modes is evaluated through filters using hard validation criteria (HVC) and soft validation criteria (SVC) calculation. HVC are defined by delimiting domains with fixed thresholds and rigorous application of physical principles, while the calculation of SVC yields a range of values that require thresholds to determine the point at which modes are considered spurious. Conventional modal validation criteria include modal phase collinearity (MPC) [17], mean phase deviation (MPD) [18], modal transfer norm (MTN) [19], and modal coherence indicator (MCI)[20]. When spurious modes substantially contaminate a stability diagram, the application of these criteria may be insufficient, allowing many spurious modes to remain in the diagram, affecting the accuracy of the estimated modal parameters. Thus, some approaches for the automatic interpretation of the stability diagram have incorporated the calculation of the uncertainty in modal parameters that are used to distinguish physical and spurious modes. The modes whose modal parameters have the largest standard deviations are defined as spurious [21, 22].

Different clustering approaches based on partitioning methods and hierarchical algorithms have been applied for the automatic interpretation of the stability diagram. Partitioning methods such as K-means and Fuzzy C-means are fast in processing large amounts of data; however, one of the difficulties in conducting such algorithms is that a priori knowledge of the number of clusters is required. These algorithms have been used in different processes, such as classifying modes between physical and spurious [3, 10, 23], selecting physical modes [18], and separating close modes [24]. The hierarchical algorithm has been widely used to cluster modes with similar characteristics based on the relative differences between natural frequencies, damping rates, and MAC coefficients [25–27]. One of the main issues of this algorithm is the definition of a tree cut-off threshold, which establishes the number of clusters obtained. Furthermore, the user's definition of a static threshold in long-term monitoring can lead to identification errors, since a different stability diagram is obtained for each dataset.

In [18], a fully automated approach is proposed, which combines the capabilities of the hierarchical algorithm and k-means through the following stages: 1. Application of soft validation criteria and use of a k-means algorithm to detect spurious modes. 2. Automatic calculation of the cluster threshold based on the remaining modes and grouping the modes with similar characteristics employing a hierarchical algorithm. 3. Use a k-means algorithm to select clusters containing the physical modes and select a representative element from each cluster of physical modes. Later, in [28], an uncertainty criterion of natural frequencies is incorporated at the end of the first stage to improving the cleaning of spurious modes. In addition, an outlier detection procedure is introduced after selecting the clusters containing the physical modes to make the representative values more accurate. Both approaches were validated through automatic

operational modal analysis on bridges (the Z24 bridge and the Wetteren footbridge), showing high robustness in the estimated modal parameters.

In this work, a procedure based on the approaches proposed in [18, 28] is configured to assess its potential for long-term monitoring of structures. It is considered that ambient vibration measurements present low SNR conditions to account for operational conditions. The procedure is presented and validated by tracking the modal parameters of a beam-type structure with ten degrees of freedom. This paper is organized as follows. First, in Section "Numerical Simulation of an Experimental Test", the structural model used to validate the automated procedure is introduced; Section "Covariance-Driven Stochastic Identification (SSI-COV)" provides the background of the SSI-COV identification method and the stability diagram. Then, in Section "Automated Modal Parameter Estimation (MPE) Procedure", the automated procedure is presented, and in Section "Tracking of Modal Parameters", the results of the modal parameter tracking of the simulated structural model are shown.

Numerical Simulation of an Experimental Test

A simply supported beam is simulated to determine the performance of the automated procedure for structural monitoring. The beam has a length of 10 [m] and an uniform square cross-section of side 0.2 [m]. The numerical model of the structure is discretized into 11 finite elements of equal length. The Young's modulus in each element is $E = 1.787*10^9$ [N/m^2] and the material density $\rho = 12{,}816$ [Kg/m^3]. As shown in Fig. 1, this beam is also elastically restrained against translational and rotational movements at nodes 2 and 3. $k_T = 20{,}000*T$ [N/m] and $k_R = 40{,}000*T$ [N/m] are the translational and rotational spring constants, respectively, where T is a temperature function. The mass **M** and stiffness **K** matrices obtained from geometric and material properties are condensed into ten vertical degrees of freedom (DOF) to reduce the problem size. A damping matrix **C** proportional to the mass is also estimated. The ambient temperature functions used in this work are shown in Fig. 2.

The beam is excited with a white noise-type load on each vertical DOF. Then, the vibration responses are obtained using the Newmark integration method, [29] considering the matrices **M**, **K** and **C**, and parameters $\beta = 0.25$ and $\gamma = 0.5$ at a sampling frequency of 200 [Hz] and 15 [min] duration. Vertical acceleration measurements in the ten DOF of the beam are grouped into a dataset to be used in the modal identification process. A total of 192 measurement datasets are obtained from the simulation of the same quantity of white noise inputs. A signal with zero mean and variance equal to the noise intensity is added to each response to investigate the sensitivity of modal parameters to noise-type perturbation. A collection of 96 simulated datasets, DT140, were obtained considering the temperature function T_1 in Fig. 2a, with an SNR = 40 [dB]. The other 96 datasets depend on the temperature function T_2

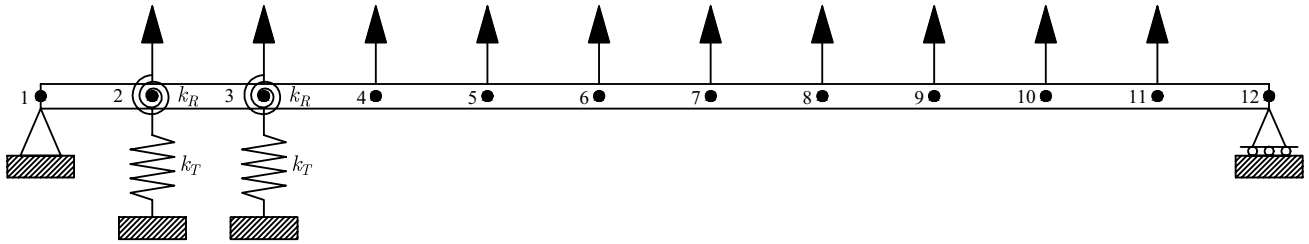


Fig. 1 Structural model

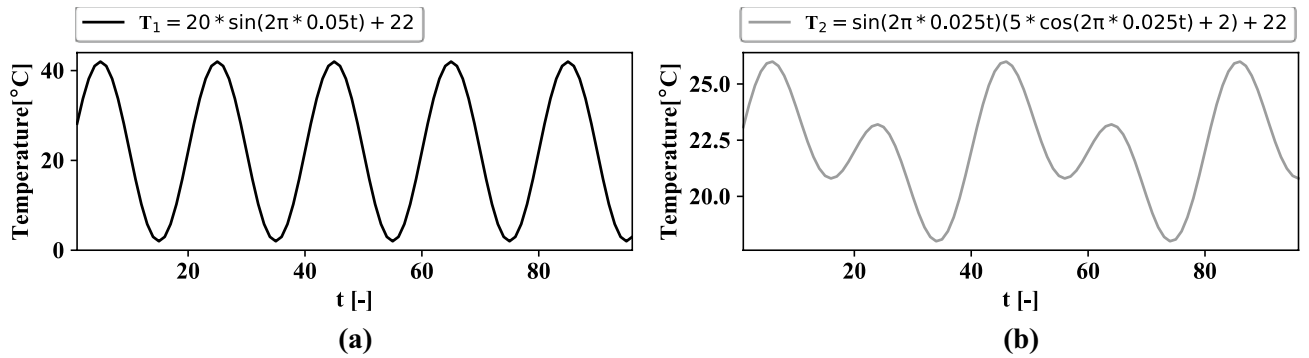


Fig. 2 Temperature functions

in Fig. 2b, with an SNR = 20 [dB]. That datasets' collection is named DT220.

The modes of the structure are in the frequency range [0–45] [Hz]. Thus, each time series was decimated by a factor of 2 to resample the data at a frequency of 100 [Hz]. To assess the performance of the automated procedure at each stage, the results are shown for the two collections of datasets, DT140 and DT220. Subsequently, the identification is conducted using the 192 simulated datasets, and the results of the modal tracking are presented.

Covariance-Driven Stochastic Identification (SSI-COV)

The SSI-COV algorithm can be used to identify modal parameters from covariance matrices of ambient vibration measurements. The reference-based generalization version of this algorithm was later proposed to reduce the influence of noise on the identification results. The SSI-COV/ref [30] algorithm establishes covariance functions $\mathbf{\Lambda}_l^{ref}$ for $l = 1, 2, ..., 2i-1$ between all outputs r and a limited set of reference outputs r_o to construct a Toeplitz matrix $\mathbf{T}_{1|i}^{ref} \in \mathbb{R}^{ir \times ir_o}$ and subsequently by singular value decomposition to estimate the observability matrix $\mathcal{O} \in \mathbb{R}^{ir \times n}$. From the observability matrix, it is possible to identify the state transition matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and the observation matrix $\mathbf{C} \in \mathbb{R}^{r \times n}$ of a discrete stochastic state-space model given by

$$\begin{aligned} x_{k+1} &= \mathbf{A}x_k + w_k, \\ y_k &= \mathbf{C}x_k + v_k, \end{aligned} \tag{1}$$

where k is the sampling instant, x_k the state vector, and y_k the measurements vector. w_k and v_k are stochastic processes with zero mean representing unknown inputs and measurement noise. i is the number of block rows of the Toeplitz matrix and n the model order with $n/2$ equal to the number of the system's degrees of freedom.

The modal parameters can be identified from the system matrices $\mathbf{A}$ and $\mathbf{C}$. Using the eigenproperties decomposition of matrix $\mathbf{A}$ given by $\mathbf{A}\psi_k = \lambda_k \psi_k$, it is obtained that

$$\psi_{ck} = \psi_k, \lambda_k = \frac{\ln(\lambda_k)}{\Delta t} \begin{cases} f_k = \frac{|\lambda_{ck}|}{2\pi}[Hz] \\ \xi_k = \frac{\Re(\lambda_{ck})}{|\lambda_{ck}|} * 100[\%], \\ \phi_k = C\psi_k \end{cases} \tag{2}$$

where ψ_k and λ_k are the discrete-time eigenvectors and eigenvalues of $\mathbf{A}$, respectively. ψ_{ck} and λ_{ck} are the equivalent continuous-time eigenproperties, Δt the sampling period, $\Re(\bullet)$ represents the real part of a complex, and f_k, ξ_k, and ϕ_k the identified natural frequencies, damping ratios, and mode shapes, respectively.

Stabilization Diagram

The stabilization diagram provides a graphical output of the identification results of the system. Figure 3a, b shows the poles identified for the datasets collections DT140 and DT220, respectively. A previous analysis of the identification algorithm on the DT140 showed that for model orders below $n = 40$, frequencies lower than 1 [Hz] were not being identified, so the order range for the identification was chosen from $n_{min} = 2$ to $n_{max} = 120$ in steps of 2. The number of block rows of the Toeplitz matrix was set in $i = 60$ and reference channels = (2, 3, 4, 5, 6, 7, 8, 9, 10, 11).

In both cases, well-defined vertical alignments representing physical modes are noticeable in the frequency range of 0–20 [Hz]. Manual extraction of these modes requires a lot of user interaction, as they are not given directly but require

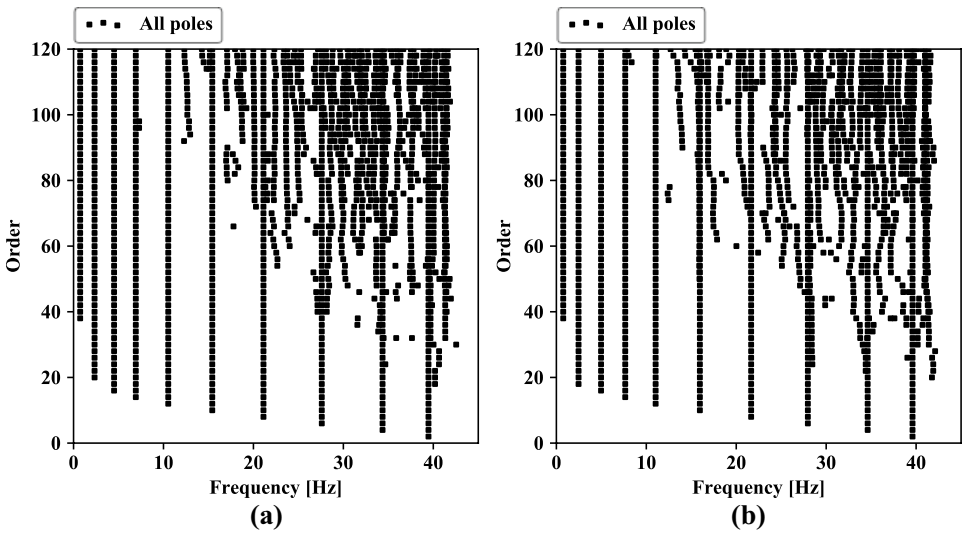


Fig. 3 SSI-COV/ref results. **a** DT140 dataset and **b** DT220 dataset

visual recognition. As the identification is performed for a wide range of model orders and beyond a certain order, these are overestimated, resulting in many candidate modes. For the DT140 dataset, 1455 modes were identified, while for DT220, 1558 modes. These modes are composed of genuine physical and spurious modes. The stability diagrams in Fig. 3 show that contamination by spurious modes is present for frequencies higher than 20 [Hz] and model orders higher than $n = 40$, which makes stable modes with higher frequencies difficult to extract manually as their visual perception is not reliable. The following section introduces the configured procedure to automatically estimate the physical modes from the stability diagram.

Automated Modal Parameter Estimation (MPE) Procedure

Overview

The configured procedure for automatically estimating modal parameters from the stabilization diagram is presented in Fig. 4. Initially, ambient vibration measurements are received as input data. Then, these signals are pre-processed by decimation to obtain a frequency band containing the frequencies of interest. Also, the averages of the signals are subtracted, since it is necessary to compute the output covariance matrices in the identification

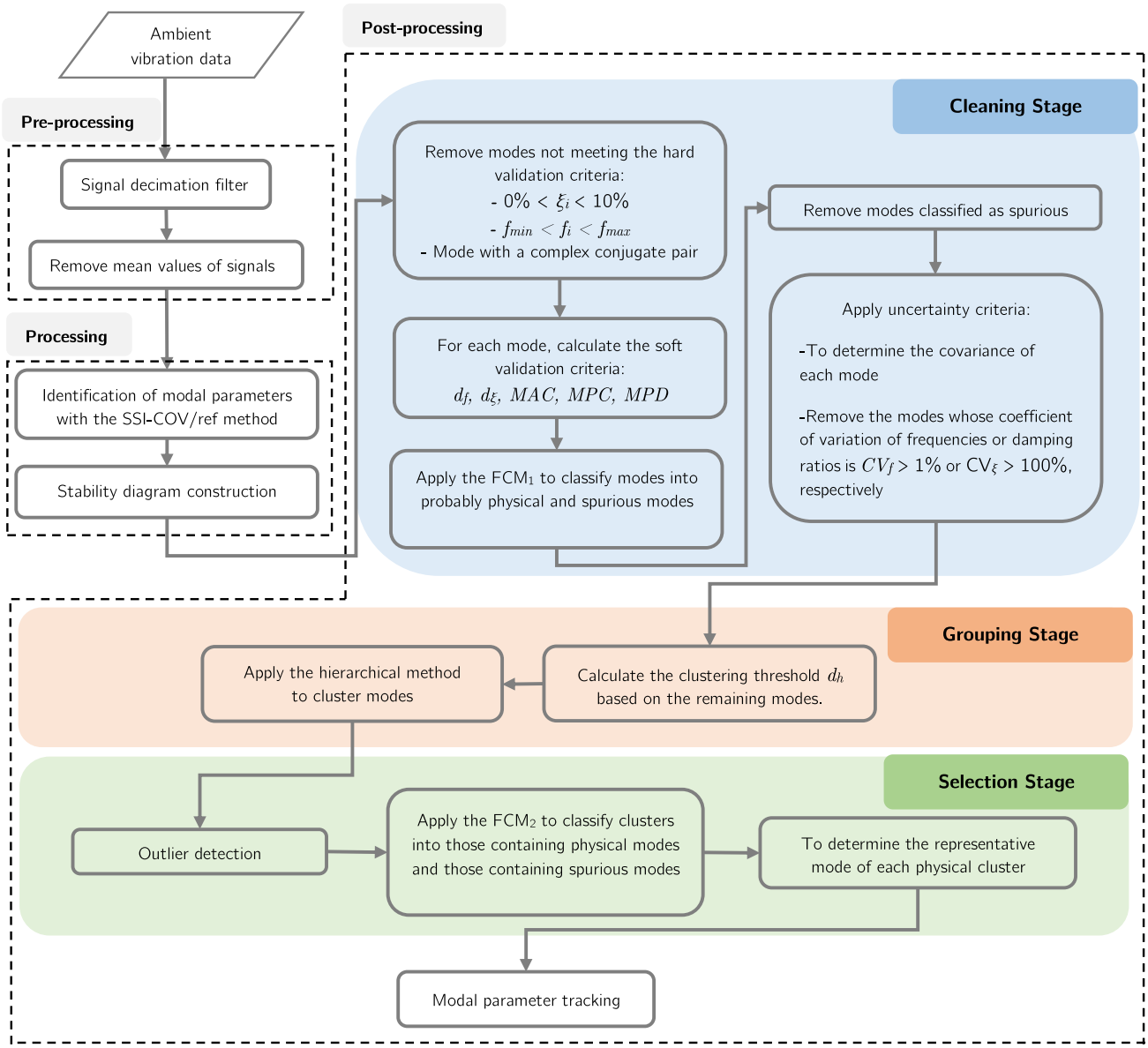


Fig. 4 Automatic MPE procedure flowchart

algorithm. Subsequently, the signals are processed by the SSI-COV/ref algorithm to obtain a substantial number of candidate modes. Next, the post-processing of the information includes the automated procedure for the estimation of modal parameters and modal tracking.

The automatic MPE procedure follows the stages proposed in [18], and the core of the procedure is conceptually similar to the one described in [28]; however, some variations are made. The process consists of three main stages: cleaning the stabilization diagram, grouping similar modes, and selecting the physical modes. Hard validation criteria (HVC) and soft validation criteria (SVC) are applied to detect spurious modes in the first stage. Relative differences between frequencies and damping ratios, as well as the MAC, are applied to define the similarity between two modes. The single-mode criteria MPC and MPD, which define the complexity of the modes, are also incorporated. Instead of hard clustering as used in [28], soft clustering via a 5D Fuzzy C-means algorithm is employed to separate the modes into spurious and probably physical ones. Subsequently, the spurious modes found are discarded. Finally, in the first stage of the process, in [28], an uncertainty criterion of the natural frequencies is used to detect more spurious modes. In this work, this criterion is also included for the damping ratios. These two uncertainty criteria are applied to the remaining modes to reclassify them into spurious and probably physical.

A hierarchical algorithm is used in the grouping stage to obtain clusters of the probably physical modes based on a distance matrix that depends on the relative difference between frequencies and MAC values. An automatically calculated threshold proposed in [18] is used to determine the number of final clusters. In the selection stage, a statistical procedure is applied to detect and remove outliers in each cluster obtained in the previous step. In [28], this procedure is performed at the end of the selection stage; however, in this work, it was found to be more advantageous to apply it at the beginning of this stage. Since the selection of clusters representing physical modes is performed based on the number of elements in each cluster using a fuzzy C-means algorithm, the detection of outliers at the beginning helps dissolve clusters containing spurious modes, thus, increasing the probability of selecting only genuinely physical modes. Finally, the procedure ends with selecting a representative mode for each cluster based on the distance between the modes within each cluster and the MPC indicator. In [28], representative values of each cluster are computed through the mean. However, since modal shapes and damping ratios are not identified with the same accuracy as natural frequencies, the mean of the cluster elements can lead to distorted values, which backfire if these parameters are used as similarity criteria in modal tracking. Each stage is explained in detail below.

Spurious Mode Elimination

Hard Validation Criteria

A set of candidate modes are identified at each model order $\left(n_{min}, n_{min+2}, ..., n_{max-2}, n_{max}\right)$. At each iteration, the modes are filtered by hard validation criteria to find some spurious modes. Only those modes that meet the following criteria will be considered in the next stage:

1. Modes with positive damping ratios whose value is less than 10[%].
2. Mode estimates within the frequency band $[f_{min}, f_{max}]$.
3. Modes with a complex conjugate pair.

Soft Validation Criteria

Spurious modes do not share modal information in common with other modes; therefore, a physical mode can be distinguished from a spurious one by the similarity with modes identified in other model orders. Also, single-mode criteria can be applied to distinguish between weak and powerful modes and real and complex modes, with weak and complex modes defined as spurious [31]. A list of soft validation criteria (SVC) that can be applied for mode classification is presented in [18]. Table 1 summarizes the SVC chosen for this work.

Assuming that i is a mode identified at model order n, a mode k identified at one of the model orders $2, ..., n-2$ is sought, such that

$$\{j\} = \arg\min \left(d_f\left(f_i, f_k\right) + 1 - MAC\left(\phi_i, \phi_k\right)\right) \quad for \quad k = 1, 2...n_p, \tag{3}$$

where n_p is the total number of modes identified with the model orders $2, ..., n-2$ that meet the HVC. Thus, the modes i and k represent a pair of closest modes. d_f and MAC measure the similarity between the two modes. d_f can be calculated as

Table 1 Soft validation criteria

Criterion	Value	Ideal physical	Ideal spurious
$\zeta^{(1)}$	d_f = Relative frequency difference	0	1
$\zeta^{(2)}$	d_ξ = Relative damping ratio difference	0	1
$\zeta^{(3)}$	MAC = Modal assurance criterion	1	0
$\zeta^{(4)}$	MPC = Modal phase collinearity	1	0
$\zeta^{(5)}$	MPD = Mean phase deviation	0	1

$$d_f(f_i,f_j) = \frac{|f_i - f_j|}{\max(f_i,f_j)}; \tag{4}$$

MAC is the *modal assurance criterion* [32], a dimensionless coefficient between 0 and 1 that provides the degree of correlation between two mode shapes ϕ_i and ϕ_j, defined as

$$MAC(\phi_i,\phi_j) = \frac{\left|\phi_i^*\phi_j\right|^2}{\left(\phi_i^*\phi_i\right)\left(\phi_j^*\phi_j\right)}, \tag{5}$$

where * is the complex conjugate transpose vector.

The Modal Phase Collinearity (MPC) [17] is used to quantify the spatial consistency of the identification results. In complex modes, the two eigenvalues of the variance–covariance matrix of the real and imaginary parts of the mode shape ϕ_i are approximately equal, while one of the eigenvalues of this matrix is zero in mode shapes with monophase behavior. The values of this indicator are dimensionless and lie between 0 for mode shapes with no collinearity and 1 for real modes. The MPC is obtained as follows: define the variance–covariance matrix of the real and imaginary parts of ϕ_i:

$$\mathbf{S} = \begin{bmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy} \end{bmatrix}, \tag{6}$$

with

$$\begin{aligned} S_{xx} &= \Re(\phi_i)^T \Re(\phi_i), \\ S_{yy} &= \Im(\phi_i)^T \Im(\phi_i), \\ S_{xy} &= \Re(\phi_i)^T \Im(\phi_i). \end{aligned} \tag{7}$$

The eigenvalues of matrix $\mathbf{S}$ are

$$\begin{aligned} \lambda_1 &= S_{xx} + \frac{S_{xy}\left(2\left(\mu_{MPC}^2+1\right)sin^2\left(\theta_{MPC}\right)-1\right)}{\mu_{MPC}}, \\ \lambda_2 &= S_{yy} - \frac{S_{xy}\left(2\left(\mu_{MPC}^2+1\right)sin^2\left(\theta_{MPC}\right)-1\right)}{\mu_{MPC}}, \end{aligned} \tag{8}$$

where

$$\mu_{MPC} = \frac{S_{yy} - S_{xx}}{2S_{xy}}, \tag{9}$$

$$\theta_{MPC} = arctan\left(\mu_{MPC} + sgn\left(S_{xy}\right)\sqrt{\mu_{MPC}^2+1}\right), \tag{10}$$

and the MPC of the mode shape ϕ_i is defined as

$$MPC(\phi_i) = \left[2\left(\frac{\lambda_1}{\lambda_1+\lambda_2} - 0.5\right)\right]^2. \tag{11}$$

Another indicator to measure the mode shape complexity is the Mean Phase Deviation (MPD). This indicator computes the mean phase angle deviation of the mode shape components with respect to a straight line that presents a special phase angle concerning the real axis. The phase angles are computed by minimizing the orthogonal distance of the modal shape components to the fit of this straight line [18]. The MPD is calculated first using a total least-squares fit of the mode shape in the complex plane

$$\mathbf{USV}^T = \left[\, \Re(\phi_i) \;\; \Im(\phi_i) \,\right], \tag{12}$$

and then a weighted sum of phase angle deviations

$$MPD(\phi_i) = \frac{\sum_{k=1}^{r} w_k arccos\left|\frac{\Re(\phi_{ik})V_{22} - \Im(\phi_{ik})V_{12}}{\sqrt{V_{12}^2+V_{22}^2}\phi_{ik}}\right|}{\sum_{k=1}^{r} w_k}; \sum_{k=1}^{r} w_k \neq 0, \tag{13}$$

where V_{12} and V_{22} are the elements (1, 2) and (2, 2) of the right singular vector $\mathbf{V}$ in Eq. 12, r is the number of components of the mode shape, and w_k are weighting factors, commonly equal to $w_k = |\phi_{ik}|$ to give elements with larger modal amplitude a stronger weight. For complex modes, the MPD should be close to 1, while for real modes, close to 0.

The criteria matrix $\mathbf{Z}$ of Eq. 14 is constructed by evaluating the SVC to all n_p modes remaining after applying the HVC, where the row i represents a vector that contains the SVC of mode i. In the case of criteria d_f, d_ξ, and *MPD* when values tend to 0, they represent a high level of physical relevance, same when *MAC* and *MPC* tend to 1. In contrast, when values of d_f, d_ξ and *MPD* are close to 1 and values of *MAC* and *MPC* are close to 0, they represent spurious modes

$$\mathbf{Z} = \begin{bmatrix} \zeta_1^{(1)} & \zeta_1^{(2)} & \zeta_1^{(3)} & \zeta_1^{(4)} & \zeta_1^{(5)} \\ \zeta_2^{(1)} & \zeta_2^{(2)} & \zeta_2^{(3)} & \zeta_2^{(4)} & \zeta_2^{(5)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \zeta_{n_p}^{(1)} & \zeta_{n_p}^{(2)} & \zeta_{n_p}^{(3)} & \zeta_{n_p}^{(4)} & \zeta_{n_p}^{(5)} \end{bmatrix}. \tag{14}$$

Fuzzy C-Means Clustering

An iterative fuzzy C-means clustering algorithm [33, 34] is employed to classify the modes into probably physical and spurious modes. This approach divides a dataset into C clusters or classes by minimizing the weighted sum of the Euclidean distance between the dataset and centroids of each class. Thus, for the case in question, the objective function of the FCM_1 can be written as

$$\{V,\eta\} = arg\,\min_{V,\eta} \sum_{j=1}^{n_p}\sum_{i=1}^{2} \eta_{ij}^{m} \left\| Z_j - V_i \right\|^2, \tag{15}$$

where Z is the matrix of soft validation criteria, η_{ij} is the degree of membership of the mode j to the class i (probably physical or spurious mode), V_i is the center of the classes and $m \in [1, \infty)$ is a weighting exponent that determines the fuzziness of the resulting clusters. $m = 2$ is chosen in this work. The steps to apply the algorithm are

1. Randomly initiate the cluster membership values η_{ij} in the range [0, 1].
2. Compute the center of the clusters as $V_i^k = \frac{\sum_{j=1}^{n_p} \left(\eta_{ij}^{k-1}\right)^m Z_j}{\sum_{j=1}^{n_p} \left(\eta_{ij}^{k-1}\right)^m}, i = 1, 2.$
3. Update the membership matrix as $\eta_{ij}^k = \left(\sum_{q=1}^{2} \left(\frac{\left\| Z_j - V_i^k \right\|}{\left\| Z_j - V_q^k \right\|} \right)^{\frac{2}{m-1}} \right)^{-1}, i = 1, 2; j = 1, 2, ..., n_p.$

   $\eta_{ij}^k = \left(\sum_{q=1}^{2} \left(\frac{\left\| Z_j - V_i^k \right\|}{\left\| Z_j - V_q^k \right\|} \right)^{\frac{2}{m-1}} \right)^{-1}, i = 1, 2; j = 1, 2, ..., n_p.$
4. If $\left\| \eta_{ij}^k - \eta_{ij}^{k-1} \right\| < \epsilon$ stop. Otherwise, do $k = k + 1$ and repeat steps 2–4.

As a result, the factor $0 < \eta_{ij} < 1$ indicates how much the modes $j = 1, 2, ..., n_p$ belong to each class $i = 1, 2$. According to Table 1, the vector $C = [0, 0, 1, 1, 0]$ represents the centroid of the ideal physical class. Supposing that V_1 is the closest cetroid to C; therefore, V_1 represents the centroid of the probably physical class. As $\sum_{j=1}^{2} \eta_{ij} = 1$, the modes with $\eta_{1j} < 0.5$ are considered as spurious and are discarded.

Figure 5 shows the mode classification for the DT140 and DT220 datasets' collections. As the optimization problem generates a five-dimensional space, Fig. 6 shows the result of the FCM_1 for DT140 in four projections of two-dimensional space. In all cases, the ordinate axis was chosen as d_ξ due to the large dispersion in the modal parameter ξ, which provides better visualization of the criteria. The plots also illustrate the initial centroids of the classes from Table 1 and the final centroids obtained by running the algorithm. It is observed that the values obtained for each criterion span a continuous space instead of concentrating on the initial centroids. This finding reaffirms the advantage of applying optimization to classify modes instead of carrying out a manual process based on non-optimal threshold values.

Figure 6e shows the membership value of each mode to the probably physical class. In this plot, it is notable that several modes are located very close to the threshold 0.5, which means that some spurious modes are being classified as probably physical and vice versa. However, the stability diagrams in Fig. 5a, b show that most of the spurious modes are not on the vertical alignments, so the latter case can be left out. Then, not all spurious modes are detected. For instance, in the DT140 dataset collection, 33.68% of the candidate modes were classified as spurious, while in DT220, 27.73%. This result is because the fuzzy C-means algorithm is suitable for normally distributed features, which is not the case for the validation criteria. In addition, when implementing the algorithm, soft validation criteria with a long range, such as d_ξ, MPC, and MPD, have more weight and dominate the results. These limitations of fuzzy C-means do not allow the detection of a significant number of spurious modes. To overcome this difficulty in [10, 31, 35, 36], it is proposed before

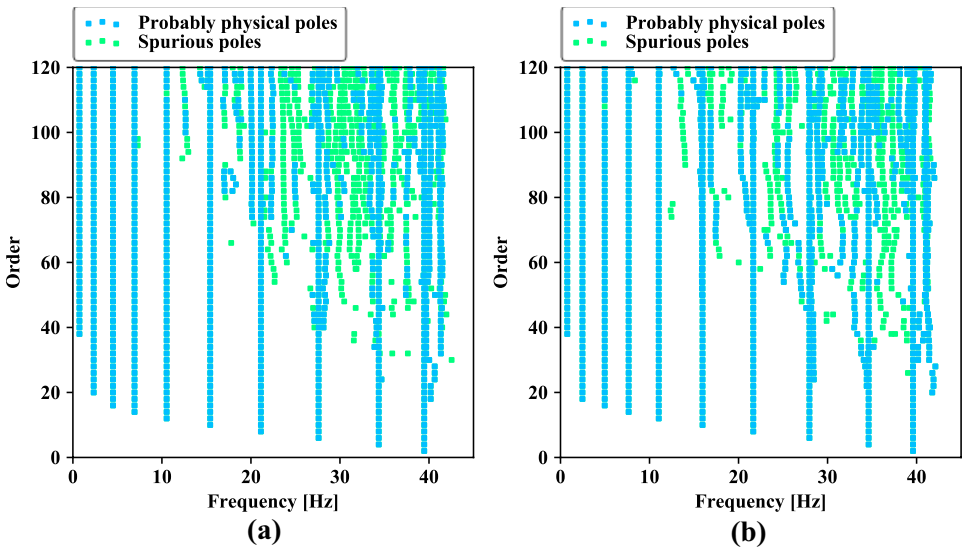


Fig. 5 Classification of modes. **a** DT140 dataset and **b** DT220 dataset

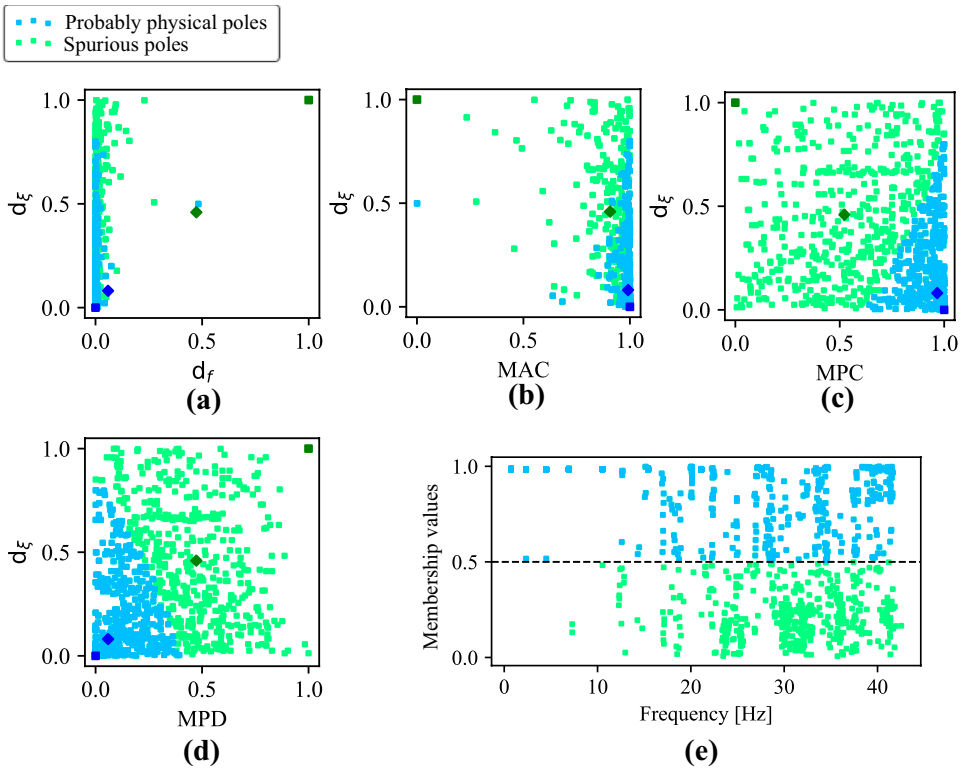


Fig. 6 Outcome of the fuzzy C-means. **a** Relative frequency difference vs relative damping ratio difference, **b** modal assurance criterion vs relative damping ratio difference, **c** modal phase collinearity vs relative damping ratio difference, **d** mean phase deviation vs relative damping ratio difference, **e** membership value of each mode. ■ and ◆ are the initial and final centroids, respectively, of the probably physical class and ■ and ◆ the initial and final centroids of the spurious class

conducting partitioning algorithms such as fuzzy C-means and K-means to normalize the variables and logarithmically transform those with the largest variation. In this paper, to improve the cleanliness of the stability diagram, more spurious modes are detected using uncertainty criteria for natural frequencies and damping ratios.

Uncertainty Criteria

All the identification process results obtained with algorithms either in the time domain or in the frequency domain are not absolutely accurate; they have a certain degree of uncertainty. The sources of error that lead to these uncertainties are the finite number of data, the loads being non-stationary, colored, and deterministic; the presence of nonlinearities in the system, noise in the measurements, and errors induced by the identification algorithm [21]. Different uncertainties impact the SSI-COV/ref method's stages during its application. At the beginning of the process, the uncertainty associated with the collected data affects the estimation of the covariance functions; then, the accumulated uncertainty propagates through each of the stages of the system identification and finally to the modal parameters. Thus, the last stage of the method inherits the uncertainties from the previous step.

Uncertainties in modal parameters can be quantified through statistical values such as variance or its square root, the standard deviation. To determine the uncertainty of a mode i, the estimation of the standard deviations of the frequency $\sigma(f_i)$, damping ratio $\sigma(\xi_i)$, and modal shape components $\sigma(\phi_{i,k})$ requires the availability of a representative large number of samples. Reynders et al. in [21] developed a procedure in which the measured time series to conduct the SSI-COV/ref are divided into n_b samples, and subsequently, the uncertainties are estimated based on the first-order sensitivity of the modal parameter estimates to perturbations of the measured output-only data. However, the direct implementation of this methodology requires high computational and memory costs, especially when a large sensor network and a high model order are used. Döhler and Mevel in [22] reformulated this procedure mathematically and proposed a more efficient one to determine the variances in multiple model orders as required by the stabilization diagram. Validations and

applications of this procedure can be found in [37, 38]. This procedure is briefly summarized below. The steps to compute the variance of modal parameters obtained from SSI-COV/ref algorithm in multiple model orders are as follows:

1. The procedure input is: Number of data blocks nb, the Toeplitz matrix $\mathbf{T}_{1|i}^{ref}$, observability matrix $\mathcal{O}$, system matrices $\mathbf{A}$ and $\mathbf{C}$, and values of $f_k^{(n)}$, $\xi_k^{(n)}$, $\phi_k^{(n)}$ for $k = 1, 2, ..., n_d$ identified at model orders $n = 2, 4, ..., n_{max}$.
2. Split the acceleration time series into n_b statistically independent blocks and compute the covariance of the Toeplitz matrix $\hat{\mathbf{\Sigma}}_{\mathbf{T}_{1|i}^{ref}}$. Then, compute the sensitivity $\mathcal{J}_{\mathcal{O},T_{1|i}^{ref}}$ of the observability matrix $\mathcal{O}$.
3. For each model order n, compute the sensitivity of the system matrices A and C given by $\mathcal{J}_A, \mathcal{O}$ and $\mathcal{J}_C, \mathcal{O}$, respectively. Then, estimate the covariance $\mathbf{\Sigma}_{A,C}$.
4. For each mode $k = 1, 2, ..., n_d$ identified at model order n, compute the sensitivity matrices $\mathcal{J}_{f,A}$, $\mathcal{J}_{\xi,A}$ and $\mathcal{J}_{\varphi_k,A,C}$, where φ_k is the scale mode shape ϕ_k. Finally, compute the covariance of frequencies f_k, damping ratios ξ_k, and mode shapes φ_k as: $cov\left(\begin{bmatrix} f_k \\ \xi_k \end{bmatrix}, \begin{bmatrix} f_k \\ \xi_k \end{bmatrix}\right)$, $cov\left(\begin{bmatrix} \Re(\varphi_k) \\ \Im(\varphi_k) \end{bmatrix}, \begin{bmatrix} \Re(\varphi_k) \\ \Im(\varphi_k) \end{bmatrix}\right)$.

This procedure was applied to the DT140 and DT220 datasets collections, with $n_b = 200$. Figure 7a, b shows the standard deviations ($\pm\sigma$) uncertainty bounds of frequencies as horizontal bars. It is noted that some standard deviations are so small that the bars are not visible. Figure 7c, d shows the standard deviations for the damping ratios associated with each pole. According to the color map, most scattered spots corresponding to spurious poles have a higher uncertainty in the computed damping ratios compared to the stable spots. It is proven that the standard deviation values for physical modes are smaller than for spurious modes [22, 39–41]. In this work, the remaining modes after applying the FCM_1 are again classified into probably physical and spurious modes, based on the coefficients of variation of natural frequencies and damping ratios. The coefficients of variation are obtained as the ratio between the standard deviations and the mean (frequency and damping ratio values). Modes with coefficients of variation of frequencies $CV_f > 1\%$ or coefficients of variation of damping ratios $CV_\xi > 100\%$ are classified as spurious and are eliminated. By applying these criteria, the stability diagrams in Fig. 7e, f were obtained. A total of 443 and 511 spurious modes were detected for the DT140 and DT220 datasets collections, respectively. This criterion has shown to be more effective in eliminating spurious modes than soft validation criteria [28, 42]. However, the computational cost of estimating the uncertainties for all candidate modes identified with the SSI-COV/ref is much larger than that employed in evaluating the soft validation criteria and subsequent classification by FCM_1. In this way, the application of SVC in a previous step is practical to detect some spurious modes. Thus, the computational time for uncertainty estimation is reduced, since it must be performed for a reduced set of modal parameters.

Mode Clustering

Calculation of Clustering Threshold

Calculating the clustering threshold is a crucial step in the automation process as it determines the quality of the final clustering results. A small threshold gives rise to a large variety of clusters where the genuine physical modes are not compact, while a large threshold leads to merging physical modes with spurious modes. Furthermore, selecting a static threshold in the context of long-term monitoring is not a good alternative, as it is not guaranteed to be appropriate for all datasets to be analyzed. Therefore, the approach proposed by Reynders et al. in [18] is adopted to determine the threshold in this work. Thus, the threshold is automatically calculated for each dataset analyzed through a statistical perspective according to

$$d_h = \mu_p + 2\sigma_p, \tag{16}$$

where μ_p and σ_p are the mean and standard deviation, respectively, of $\zeta^{(1)} + 1 - \zeta^{(3)}$ for the n_e probably physical modes. $\zeta^{(1)}$ and $\zeta^{(3)}$ represent the criteria d_f and MAC, respectively, according to Table 1.

Hierarchical Clustering Approach

An agglomerative hierarchical algorithm is used to group modes that share similarities. First, a matrix of distances between each pair of modes must be computed. The Euclidean distance between two elements has been commonly used to construct this matrix; however, regarding operational modal analysis, good results have been obtained by measuring the distance between two modes i and j based on the natural frequencies and mode shapes according to

$$\mathbf{D}_{i,j} = \frac{|f_i - f_j|}{\max |f_i, f_j|} + 1 - MAC\left(\phi_i, \phi_j\right). \tag{17}$$

The hierarchical algorithm works as follows

1. Calculate the distance matrix $\mathbf{D}$ using Eq. 17.
2. Each mode starts in its own cluster.

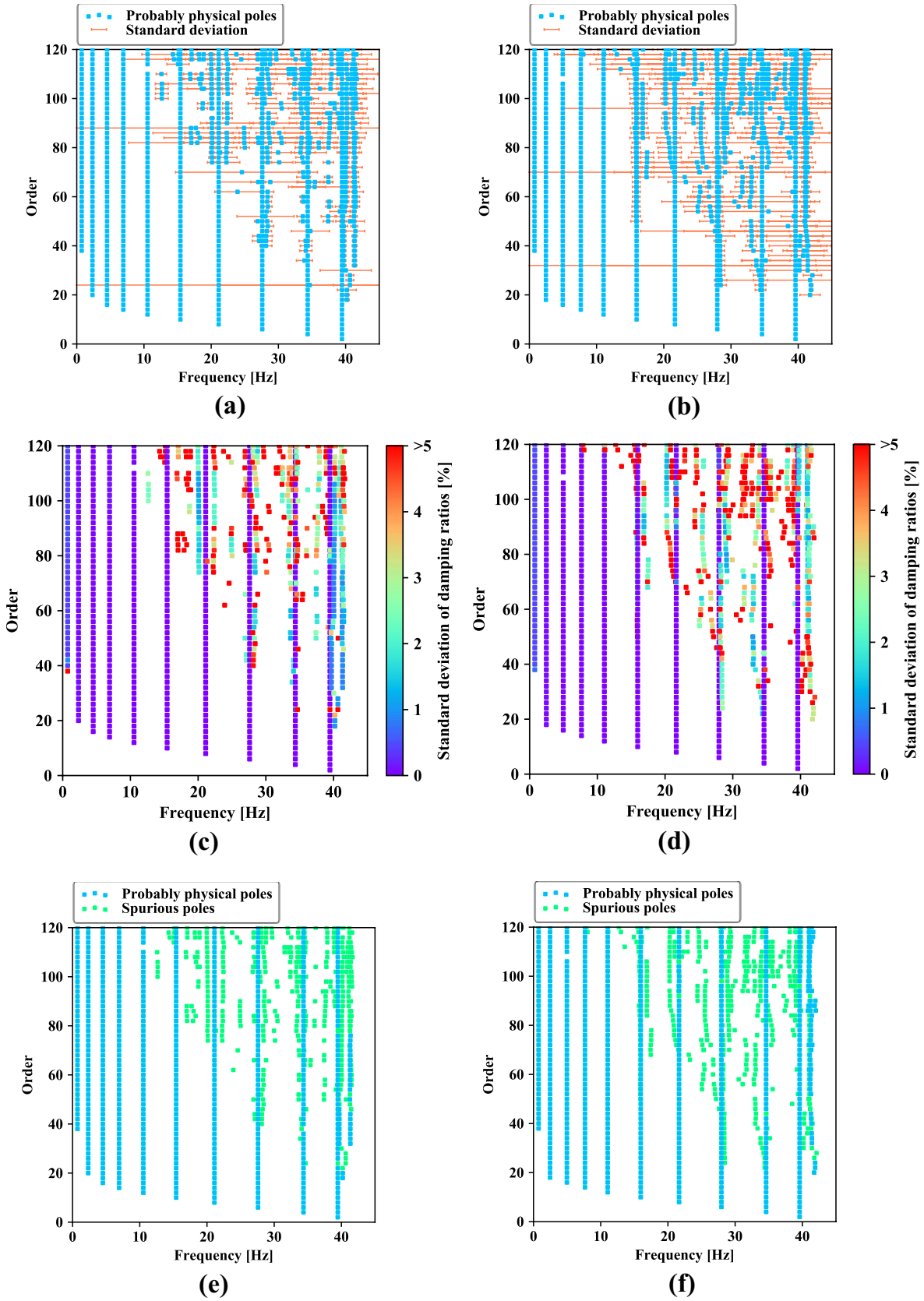


Fig. 7 Uncertainty criteria results. **a**, **b** Standard deviations of frequencies. **c**, **d** Standard deviations of damping ratios. **e**, **f** Second classification of spurious and probably physical modes. **a**, **c**, **e** DT140 dataset. **b**, **d**, **f** DT220 dataset

3. Merge the modes/clusters that are in close proximity and update matrix **D**. This step is performed until all modes are all included in a single cluster.
4. Use the threshold d_h in Eq. 16 to cut the hierarchical tree into clusters.

The distance matrix **D** is updated at each step using linkage methods as clusters are combined into larger ones. The most common are single linkage [25, 43], average linkage [18, 31], and complete linkage [26]. The distance between the closest elements of two clusters defines their overall distance in the simple linkage. This method is prone to chaining, since, at any step, two clusters are merged if two modes are sufficiently close, and the proximity between all the modes forming the two clusters is not taken into account. The complete linkage uses the distance of the farthest apart elements as the overall distance between two clusters. One of the disadvantages of this method is that it can attract outliers, resulting in a non-optimal merging. The average linkage, the method used in this work averages the distance between all pairs of elements, where each pair consists of one element from each cluster. Although this method is more computationally demanding, the resulting clusters tend to be relatively compact and separated.

Once the clusters are obtained, an inspection is performed to ensure that modes identified in the same model order are not located in the same cluster. Suppose they are located in the same cluster. In that case, the relative distance between the frequencies of the modes identified at the same model order and the mean of the frequencies of the modes in the cluster is computed. Only the mode whose relative distance is smaller is retained in the cluster.

The cut-off distances d_h were automatically estimated based on the probably physical modes for the two datasets. Thresholds of 0.171 and 0.152 were obtained for the DT140 and DT220 datasets collections, respectively. The fewer spurious modes retained in the previous stage, the more accurate the threshold calculation will be since spurious modes do not share common modal information with each other compared to physical modes. Without effective cleaning of the stability diagram, many spurious modes could be contained in groups of physical modes in the clustering process. For instance, Fig. 8a shows the results of the hierarchical clustering process for the probably physical modes of the DT140 dataset collection without applying the uncertainty criterion. For this scenario, a threshold of 0.223 was obtained, which led to the formation of 32 clusters. Spurious modes close to the physical modes were contained in the same cluster due to a large threshold. Figure 8b, on the contrary, shows the clustering results by applying the uncertainty criterion previously. A much lower threshold was obtained, resulting in tighter clusters defined by the stable columns. Additionally, in the stability diagram of Fig. 8b, it is observed that in

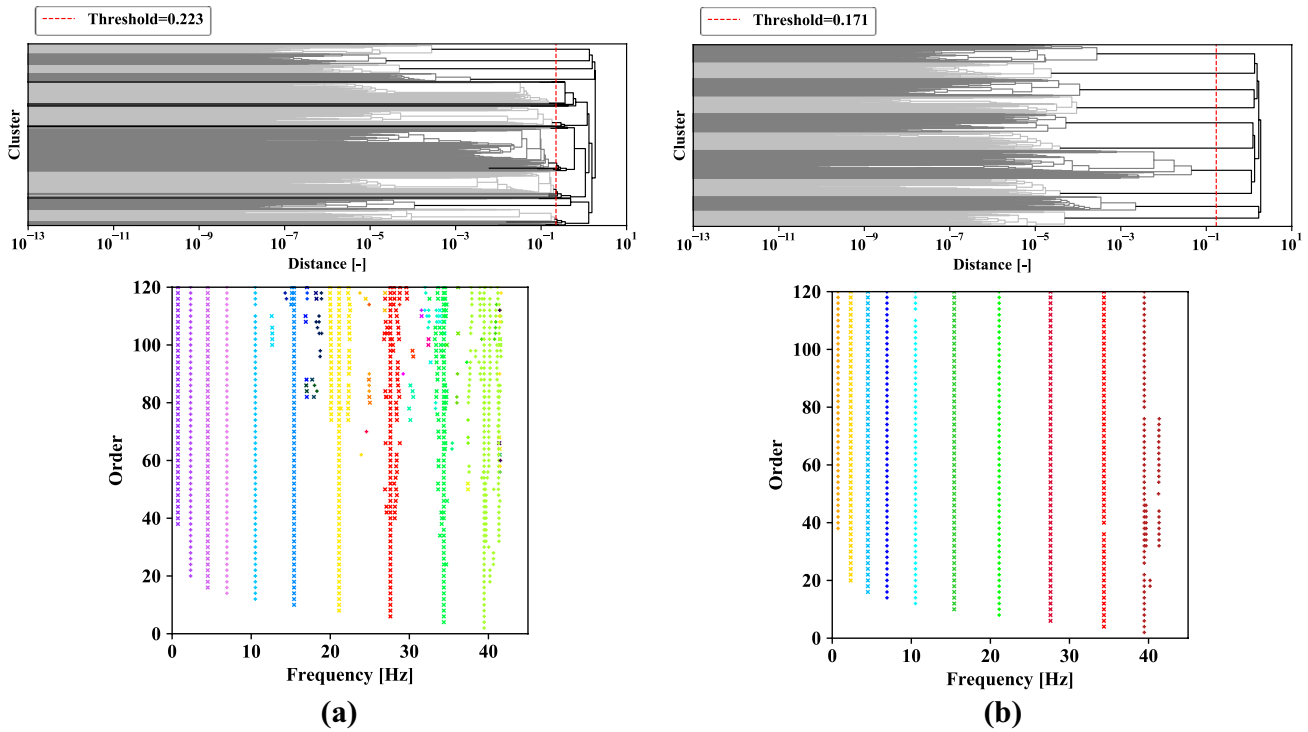


Fig. 8 Hierarchical clustering results. **a** Without uncertainty criteria. **b** With uncertainty criteria

the cluster containing the frequencies around 40 [Hz], some modes were identified in the same order of the model. This cluster was inspected and debugged, as explained above.

Physical Modal Results

Outlier Detection

Because the identification is performed for different model orders and the threshold used to cut the hierarchical tree is not 100% optimal, there is variability in the modal parameters within each cluster obtained in the previous stage. Outlier detection procedures based on minimum covariance determinant (MCD) [27], Thompson–Tau technique [31, 43], mean absolute deviation (MAD) [44], and box plots [28, 42, 45] have been implemented in different MPE automation approaches to reduce variability and safeguard the quality of modal information. In this work, box plots are implemented to find the outliers of the natural frequencies in each cluster. In addition to visually summarize the variation of long datasets, box plots can be used to find outliers without prior knowledge of the data distribution. The steps to estimate outliers using box plots are as follows:

1. For each cluster, compute the lower L_L and upper L_U limits as $L_L = Q_1 - 1.5 * IQR$; $L_U = Q_3 + 1.5 * IQR$, where Q_1 is the first quartile, Q_3 the third quartile, and IQR the interquartile range.
2. Find outliers of natural frequencies.
3. Remove the natural frequencies, damping ratios, and mode shapes associated with the outliers found.
4. Repeat steps 1–3 until no outliers are found.

The box plots for the DT140 and DT220 datasets collections are presented in Figs. 9 and 10, respectively. In most modes, extreme values were found, as seen in the black boxes. This indicates that the calculated cluster thresholds are not precise enough to obtain tighter clusters. The refined results are shown in the blue boxes, whose clusters have no outliers. In most modes, the means values do not coincide with those of the medians, which indicates that not all modes have normal distributions. Next, the clusters containing the physical modes are selected from the refined clusters.

Selection of the Physical Mode Sets

The clusters obtained in the clustering phase can be divided into two classes: those containing physical modes and those containing spurious modes. Physical modes are consistently identified from a certain model order, while spurious modes do not appear with such consistency. Therefore, clusters containing physical modes are expected to contain more

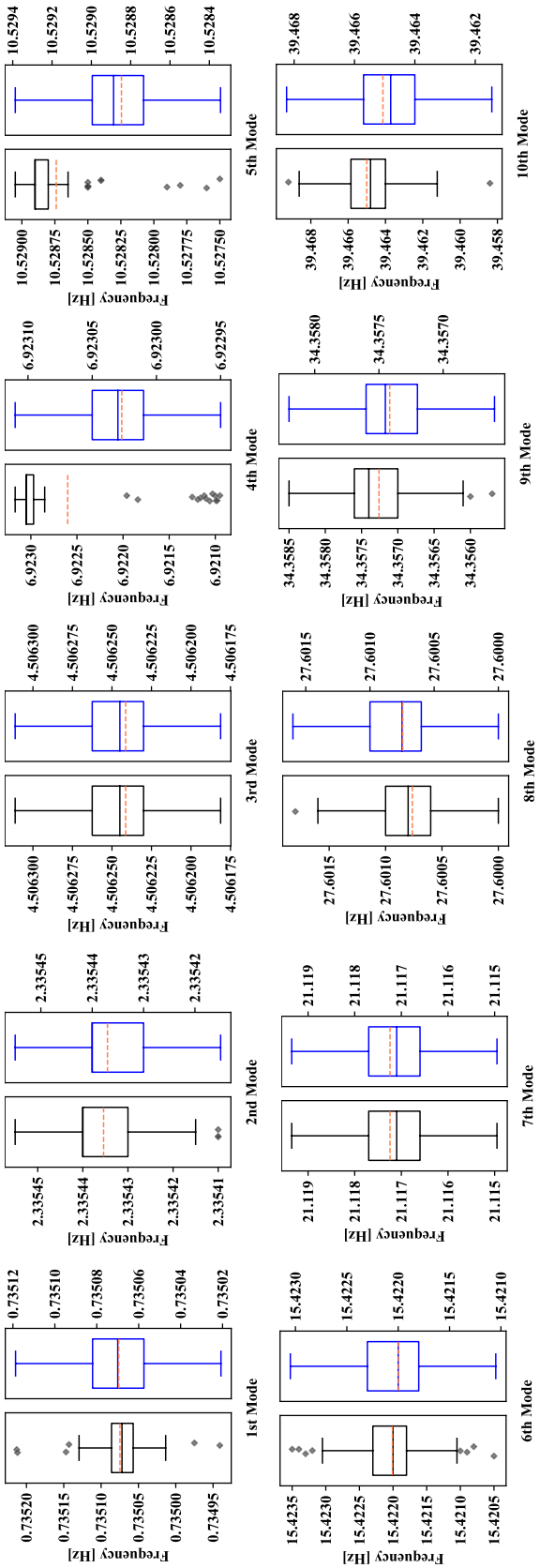


Fig. 9 Box plots of natural frequencies of DT140 dataset. Black boxes: before outlier detection. Blue boxes: after outlier detection

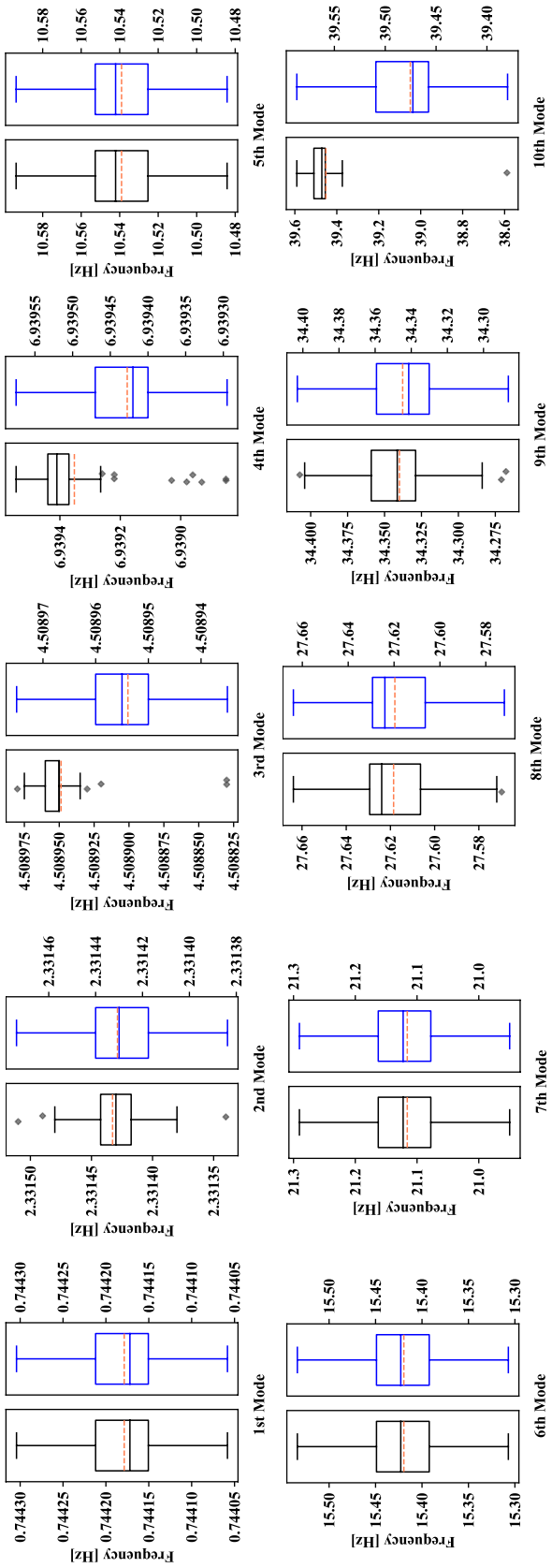


Fig. 10 Box plots of natural frequencies of DT220 dataset. Black boxes: before outlier detection. Blue boxes: after outlier detection

elements than those containing spurious modes. A fuzzy C-means clustering algorithm FCM_2 is used to select the modes that contain the physical modes. Here, the objective function to minimize is

$$\{V,\eta\} = arg\,\min_{V,\eta} \sum_{j=1}^{n_c}\sum_{i=1}^{2} \eta_{ij}^{m}\left(\mathcal{L}_j - V_i\right)^2, \tag{18}$$

where $\mathcal{L}_j$ is the number of elements of the cluster j, n_c the number of clusters, η_{ij} the degree of membership of the cluster j to the class i (physical os spurious), and V_i the center of classes and m the weighting exponent set to 2. The steps of the algorithm are as follows:

1. Randomly initiate the cluster membership values η_{ij} in the range [0, 1].
2. Compute the center of the clusters as $V_i^k = \frac{\sum_{j=1}^{n_c}\left(\eta_{ij}^{k-1}\right)^m \mathcal{L}_j}{\sum_{j=1}^{n_c}\left(\eta_{ij}^{k-1}\right)^m}, i = 1, 2.$
3. Update the membership matrix as $\eta_{ij}^k = \left(\sum_{q=1}^{2}\left(\frac{\mathcal{L}_j - V_i^k}{\mathcal{L}_j - V_q^k}\right)^{\frac{2}{m-1}}\right)^{-1}, i = 1, 2; j = 1, 2, ..., n_c.$
4. If $\left\|\eta_{ij}^k - \eta_{ij}^{k-1}\right\| < \epsilon$ stop. Otherwise, do $k = k + 1$ and repeat steps 2–4.

As a result of the application of the hierarchical algorithm, there may be the case of obtaining as output-only clusters with physical modes. For this reason, before conducting the FCM_2, a set of empty clusters equal to the number of clusters containing more than one-fifth of the number of elements of the cluster with the most elements is added [18]. Supposing that η_{1j} are the obtained membership degrees of clusters to the physical class, then clusters with $\eta_{1j} > 0.5$ are selected as clusters containing physical modes.

Representative Physical Modes

Finally, the automatic MPE selects the representative mode for each cluster or stable column in the stabilization diagram. In this paper, the following strategy is proposed. By applying the hierarchical algorithm, a distance matrix $\mathbf{D} \in \mathbb{R}^{n_e \times n_e}$ was constructed, and then, $\mathbf{D}_j \in \mathbb{R}^{\mathcal{L}_j x \mathcal{L}_j}$ is a subset of this matrix containing the distances of the elements of cluster j. The mode k for which Eq. 19 is minimized is selected as the representative mode of cluster j. This way, the mode closest to all modes, with a high *MPC* value, is selected as representative

$$\{k\} = arg\,\min_{l}\left[\left(\sum_{i=1}^{\mathcal{L}_j}\mathbf{D}_{il}^{j}\right) + 1 - MPC_l \,for\, l = 1, 2, ..., \mathcal{L}_j\right]. \tag{19}$$

The FCM_2 algorithm was applied to select the clusters containing the physical modes for the DT140 and DT220 datasets' collections. As a result, ten clusters were obtained for each set, the same number resulting from the hierarchical clustering process. Based on the theoretical modes of the simulated structural model, ten modes were expected to be identified in each dataset; therefore, the strategy of adding empty clusters proposed in [18] gives optimal results. According to the above approach, a representative mode was extracted from each cluster. In Fig. 11, the estimated stable modes for the DT140 and DT220 datasets collections are plotted over the singular value lines of the two datasets.

The DT140 and DT220 datasets' collections were also analyzed with the frequency-domain decomposition (FDD) algorithm [9]. Modal parameters were extracted using an automated local peak search procedure. Ten modes were identified for the DT140 dataset collection, while for the DT220, the automated procedure could not detect the peak of the first mode. The results obtained by the FDD and the modal information associated with each representative mode obtained automatically by the procedure configured in this work are presented in Table 2. The natural frequencies identified by the two methods are very similar, and MAC coefficients exceed 0.97, indicating that the modes estimated by the two methods are in agreement.

Table 2 also presents the MPC and MPD modal criteria and the coefficients of variation of the natural frequencies and damping ratios for each representative mode. The MPC and MPD values are very close to the values defined for an ideal physical mode. It is also noteworthy that the coefficients of variation satisfactorily meet the uncertainty criteria presented in Section 5. The coefficients of variation of the modes identified with the DT140 dataset collection are lower than those obtained with DT220. This result is consistent insofar as the noise on the response of the DT220 dataset collection is higher than that of DT140.

Tracking of Modal Parameters

The automated MPE procedure was applied to the 192 simulated datasets. The modal parameters identified with the DT140 dataset collection served as reference modes for tracking. Thus, in each analysis, the identified modes were compared with the reference modes to associate them with the same physical mode. In this work, this comparison was based on natural frequencies and mode shapes. Each new set of modal parameters was associated with a reference mode if the relative difference between the frequencies was lower than 5% and the MAC between the mode shapes and the

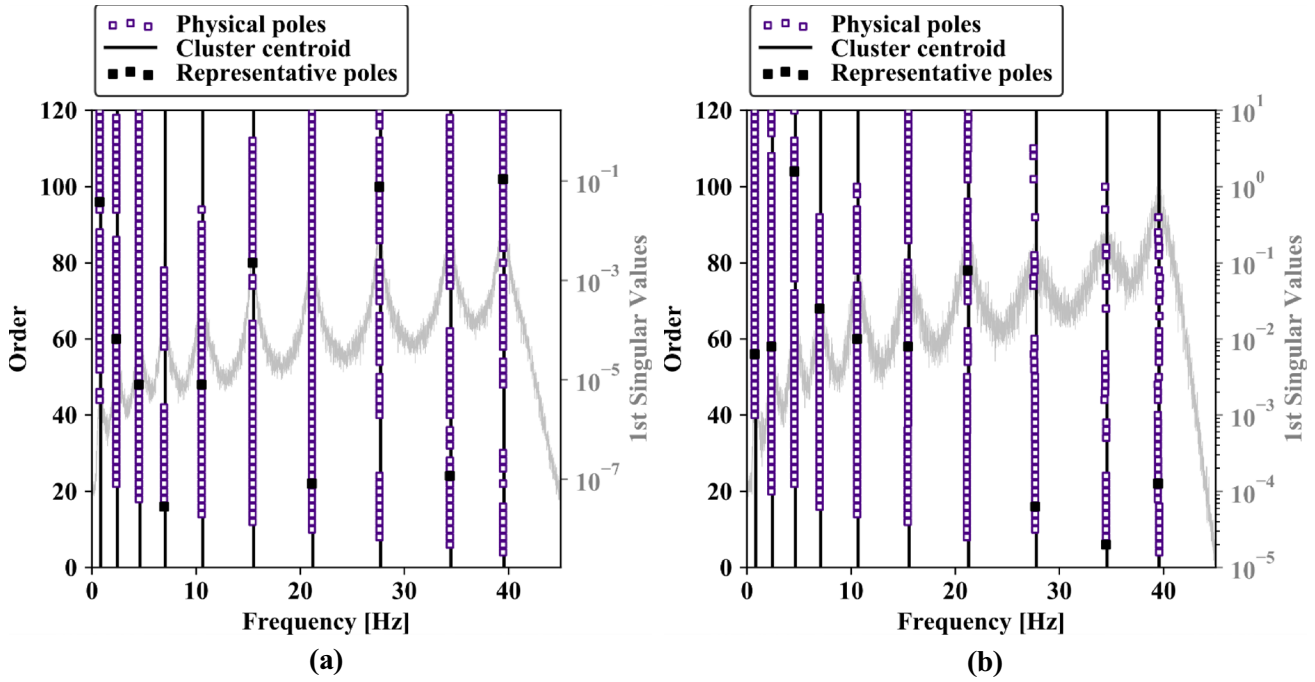


Fig. 11 Physical modes. Stable modes from dataset **a** DT140 and **b** DT220

Table 2 Identified modal parameters from FDD and the automated MPE procedure

	Mode	FDD		Automated MPE						
		f[Hz]	ξ[%]	f[Hz]	ξ[%]	MAC	MPC	MPD	CV_f[%]	CV_ξ[%]
DT140	1st	0.726	3.180	0.735	5.906	0.99	0.99	0.0035	0.34	7.10
	2nd	2.331	1.450	2.335	1.702	0.98	1.00	0.0040	0.10	3.89
	3rd	4.486	0.910	4.506	0.849	0.98	1.00	0.0018	0.05	3.83
	4th	6.927	0.490	6.923	0.556	0.99	0.98	0.0013	0.03	3.39
	5th	10.511	0.340	10.529	0.382	0.98	1.00	0.0004	0.02	3.55
	6th	15.424	0.250	15.422	0.249	0.97	1.00	0.0006	0.02	7.20
	7th	21.118	0.150	21.117	0.179	0.98	1.00	0.0003	0.01	4.04
	8th	27.606	0.080	27.601	0.118	0.99	0.97	0.0001	0.01	7.08
	9th	34.358	0.070	34.357	0.103	0.99	0.99	0.0001	0.31	22.63
	10th	39.466	0.070	39.465	0.078	0.98	0.98	0.0002	0.02	5.87
DT220	1st	–	–	0.744	6.487	–	0.98	0.0073	0.35	7.46
	2nd	2.301	1.290	2.331	1.628	0.98	0.99	0.0067	0.23	4.42
	3rd	4.500	0.560	4.509	0.746	0.99	0.99	0.0029	0.11	3.89
	4th	6.947	0.490	6.939	0.478	0.98	0.99	0.0037	0.14	3.75
	5th	10.542	0.270	10.553	0.349	0.99	1.00	0.0046	0.16	4.30
	6th	15.456	0.200	15.424	0.227	0.97	0.98	0.0090	0.49	8.86
	7th	21.170	0.110	21.124	0.136	0.98	0.99	0.0048	0.48	9.60
	8th	27.695	0.080	27.628	0.137	0.99	1.00	0.0001	0.20	13.50
	9th	34.328	0.080	34.343	0.094	0.99	0.99	0.0043	0.32	51.80
	10th	39.447	0.060	39.470	0.072	0.97	0.99	0.0012	0.12	28.76

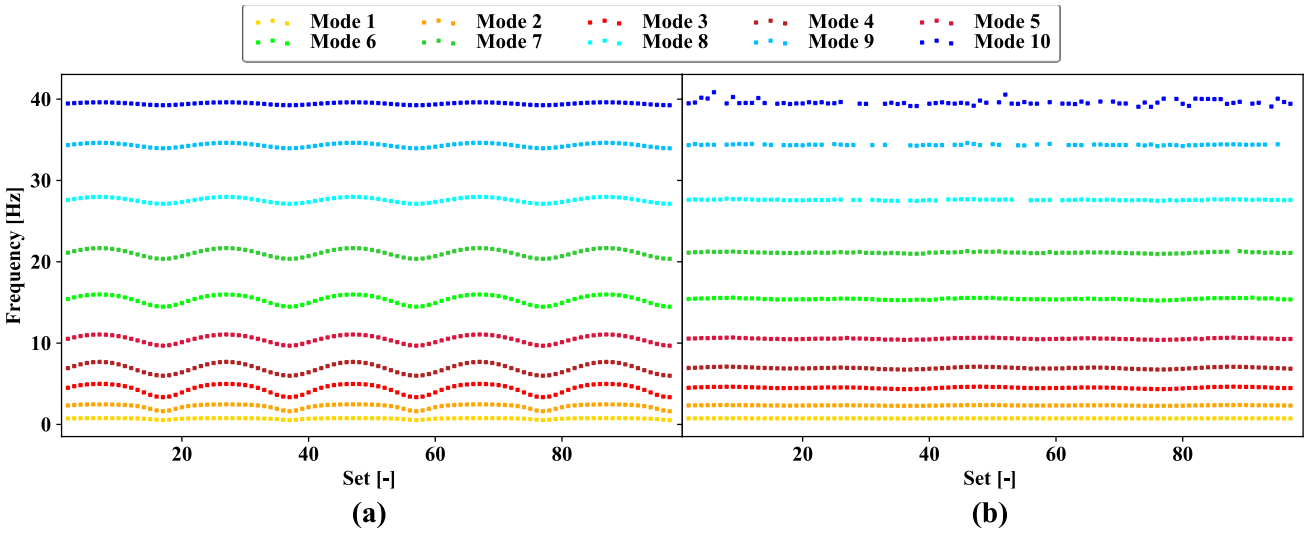


Fig. 12 Tracking of natural frequencies. **a** Datasets influenced by T_1 and SNR = 40 [dB]. **b** Datasets influenced by T_2 and SNR = 20 [dB]

mode shapes of any of the reference modes was higher than 80%.

Figure 12 shows the tracking of the natural frequencies. It can be observed that the modes identified in each analysis were correctly linked to the reference modes. Under the SNR conditions applied to the vibration responses, the procedure showed a good mode detection capability. Figure 12a shows the frequencies of the datasets influenced by the temperature function T_1 and a noise level of SNR = 40 [dB]. The modes estimated from these datasets had an identification rate of 100% for all modes. In contrast, for the other 96 datasets perturbed with a noise level of SNR = 20 [dB] and influenced by the temperature function T_2, the modes obtained showed identification rates of 100% only for the first six modes and 98%, 93%, 89%, and 91% for modes 7, 8, 9, and 10, respectively. One way to verify the consistency of the identified parameters is to compare the seasonal evolution of the natural frequencies with those of the temperature functions. Figure 12a shows that the natural frequency waves coincide with those of the T_1 function in Fig. 2a. In contrast, the waveform trend is not visible for datasets influenced by the T_2 function, as shown in Fig. 12b, since the amplitude of the T_2 function is lower than that of T_1. Figure 13 shows the temperature function T_2 vs. natural frequencies plots for each mode, revealing a positive linear correlation between the two variables, thus evidencing that the identified frequencies are coherent concerning the expected temporal evolution.

Figure 14 shows the box plots of the damping ratios identified for each mode in the two proposed cases. It can be observed that the obtained estimates are reasonably consistent as the means of damping ratios decrease with increasing frequencies, which is characteristic of mass-proportional damping. In addition, as expected, the obtained damping ratios are more sensitive to perturbations in the measured data than the natural frequencies. In Fig. 14a, variations of the damping ratios estimated from the datasets contaminated with a noise equal to SNR = 40 [dB] are observed. A significant increase in the variations of this modal parameter is noticeable in Fig. 14b, as the noise level increases to SNR = 20 [dB]. In Fig. 14b, the box plots' lower limits, estimated as $L_L = Q_1 - 1.5 * IQR$, yield negative values for modes 2, 3, and 4. For the simulated structural model, damping ratios between 0% and 10% are expected, as specified in Section "Hard Validation Criteria". However, these results are consistent with the fact that the modes identified from the SSI-COV/ref are not constrained to be stable, leading to the possible occurrence of negative damping ratios [21].

Conclusions

This paper configured a procedure for automatically estimating the modal parameters from the stability diagram constructed with the SSI-COV/ref identification results. This procedure is based on the approaches proposed in [18, 28]. Using five soft validation criteria in conjunction with uncertainty criteria for natural frequencies and damping rates allowed the detection of a substantial number of spurious

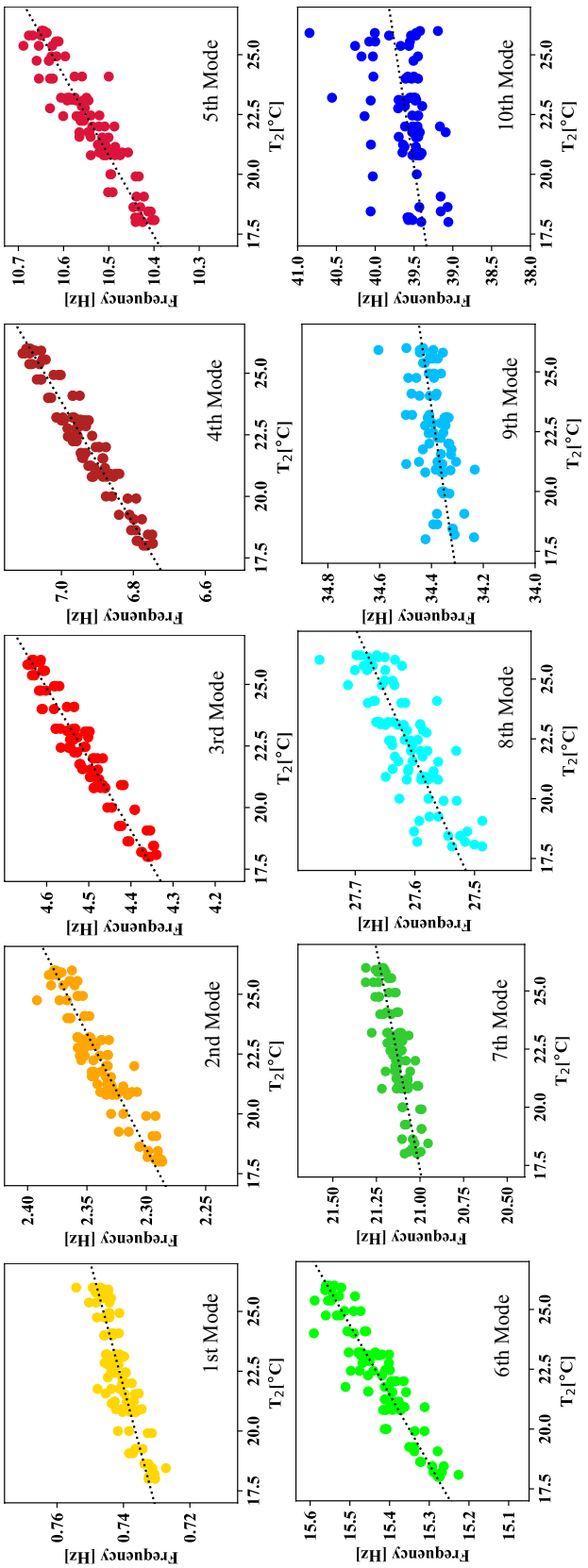


Fig. 13 Temperature function T_2 vs. natural frequencies

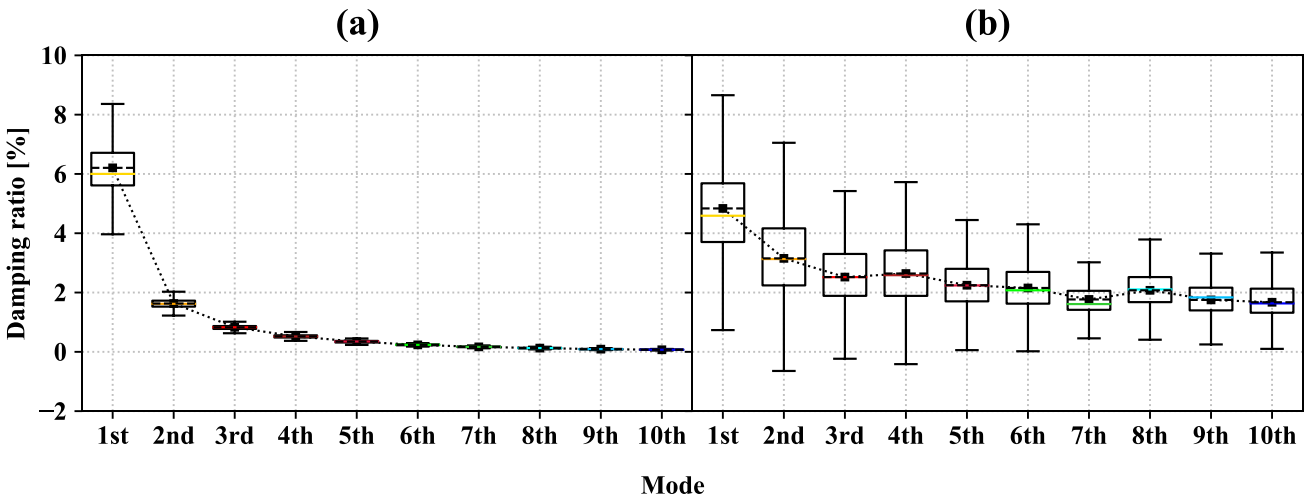


Fig. 14 Box plots of damping ratios. **a** Datasets influenced by T_1 and SNR = 40 [dB]. **b** Datasets influenced by T_2 and SNR = 20 [dB]

modes, which was reflected in the accuracy of the identified modal parameters.

The acceleration responses of a beam-type structural model whose stiffness is intentionally affected by ambient temperature functions were simulated to validate the procedure. The acceleration responses were contaminated with noise to assess the modal parameters' sensitivity. Natural frequencies and modal shapes identified for two datasets collections, one perturbed with an SNR = 40 [dB] and one with SNR = 20 [dB], were shown to agree with those estimated through the frequency-domain identification algorithm (FDD).

Continuous identification using 192 simulated datasets yielded identification rates greater than 89% for those sets perturbed with an SNR=20. The temporal evolution observed in the identified natural frequencies showed coherence, since they are correlated with the seasonal waves of the ambient temperature functions affecting the beam. The identified damping ratios are consistent with the mass-proportional damping behavior as it was simulated. This procedure requires only the input parameters of the SSI-COV/ref algorithm, which makes it suitable for long-term monitoring.

Funding Open Access funding provided by Colombia Consortium. The authors would like to thank UNILA (University of Latin American Integration) for financial support.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Amer M, Wallaschek J, Seume JR (2022) Operational modal analysis of an axial compressor rotor and casing system for the online identification of a digital twin. Applied Mechanics 3(1):244–258. https://doi.org/10.3390/applmech3010017
2. Devriendt C, Magalhães F, Weijtjens W, Sitter GD, Cunha Á, Guillaume P (2014) Structural health monitoring of offshore wind turbines using automated operational modal analysis. Struct Health Monit 13(6):644–659. https://doi.org/10.1177/1475921714556568
3. Neu E, Janser F, Khatibi AA, Braun C, Orifici AC (2016) Operational modal analysis of a wing excited by transonic flow. Aerosp Sci Technol 49:73–79. https://doi.org/10.1016/j.ast.2015.11.032
4. Jelicic G, Schwochow J, Govers Y, Sinske J, Buchbach R, Springer J (2017). Online monitoring of aircraft modal parameters during flight test based on permanent output-only modal analysis. https://doi.org/10.2514/6.2017-1825
5. Kocan C (2020) A comparative study on in-flight modal identification of an aircraft using time- and frequency-domain techniques. J Vib Control 26(21–22):1920–1934. https://doi.org/10.1177/1077546320908457
6. Xiong C, Lu H, Zhu J (2017) Operational modal analysis of bridge structures with data from gnss/accelerometer measurements. Sensors 17(3). https://doi.org/10.3390/s17030436

7. Ramos LF, Marques L, Lourenço PB, De Roeck G, Campos-Costa A, Roque J (2010) Monitoring historical masonry structures with operational modal analysis: Two case studies. Mechanical Systems and Signal Processing 24(5):1291–1305. https://doi.org/10.1016/j.ymssp.2010.01.011. Special Issue: Operational Modal Analysis
8. Zhang L, Brincker R (2005) An overview of operational modal analysis: major development and issues. In: Proceedings of the 1st International Operational Modal Analysis Conference, April 26-27, 2005, Copenhagen, Denmark, pp. 179–190. Aalborg Universitet
9. Brincker R, Zhang L, Andersen P, et al. (2000) Modal identification from ambient responses using frequency domain decomposition. In: Proceedings of the 18th International Modal Analysis Conference (IMAC), vol. 1, pp. 625–630. San Antonio, TX, USA
10. Vanlanduit S, Verboven P, Guillaume P, Schoukens J (2003) An automatic frequency domain modal parameter estimation algorithm. J Sound Vib 265(3):647–661. https://doi.org/10.1016/S0022-460X(02)01461-X
11. Peeters B, De Roeck G (2001) One-year monitoring of the z24-bridge: environmental effects versus damage events. Earthquake engineering & structural dynamics 30(2):149–171
12. Brownjohn J, Magalhaes F, Caetano E, Cunha A (2010) Ambient vibration re-testing and operational modal analysis of the humber bridge. Eng Struct 32(8):2003–2018
13. Liu Y-C, Loh C-H, Ni Y-Q (2013) Stochastic subspace identification for output-only modal analysis: application to super high-rise tower under abnormal loading condition. Earthquake Engineering & Structural Dynamics 42(4):477–498. https://doi.org/10.1002/eqe.2223
14. Qin S, Kang J, Wang Q (2016) Operational modal analysis based on subspace algorithm with an improved stabilization diagram method. Shock and Vibration 2016. https://doi.org/10.1155/2016/7598965
15. Hu S-LJ, Bao X, Li H (2010) Model order determination and noise removal for modal parameter estimation. Mech Syst Signal Process 24(6):1605–1620. https://doi.org/10.1016/j.ymssp.2010.01.005
16. Teng J, Tang D-H, Zhang X, Hu W-H, Said S, Rohrmann RG (2019) Automated modal analysis for tracking structural change during construction and operation phases. Sensors 19(4):927. https://doi.org/10.3390/s19040927
17. Pappa RS, Elliott KB, Schenk A (1993) Consistent-mode indicator for the eigensystem realization algorithm. J Guid Control Dyn 16(5):852–858. https://doi.org/10.1016/j.ymssp.2010.01.011
18. Reynders E, Houbrechts J, De Roeck G (2012) Fully automated (operational) modal analysis. Mech Syst Signal Process 29:228–250. https://doi.org/10.1016/j.ymssp.2012.01.007
19. Reynders E, De Roeck G (2008) Reference-based combined deterministic-stochastic subspace identification for experimental and operational modal analysis. Mech Syst Signal Process 22(3):617–637. https://doi.org/10.1016/j.ymssp.2007.09.004
20. Lardies J, Minh-Ngi T (2011) Modal parameter identification of stay cables from output-only measurements. Mech Syst Signal Process 25(1):133–150. https://doi.org/10.1016/j.ymssp.2010.05.020
21. Reynders E, Pintelon R, De Roeck G (2008) Uncertainty bounds on modal parameters obtained from stochastic subspace identification. Mechanical Systems and Signal Processing 22(4):948–969. Special Issue: Crack Effects in Rotordynamics
22. Döhler M, Mevel L (2013) Efficient multi-order uncertainty computation for stochastic subspace identification. Mech Syst Signal Process 38(2):346–366. https://doi.org/10.1016/j.ymssp.2013.01.012
23. Verboven P, Parloo E, Guillaume P, Van Overmeire M (2002) Autonomous structural health monitoring-part i: modal parameter estimation and tracking. Mech Syst Signal Process 16(4):637–657
24. Goethals I, Vanluyten B, De Moor B (2004) Reliable spurious mode rejection using self learning algorithms. In: Proceedings of the International Conference on Noise and Vibration Engineering (ISMA 2004), Leuven, Belgium, pp. 991–1003
25. Magalhães F, Cunha A, Caetano E (2009) Online automatic identification of the modal parameters of a long span arch bridge. Mech Syst Signal Process 23(2):316–329. https://doi.org/10.1016/j.ymssp.2008.05.003
26. Cardoso R, Cury A, Barbosa F (2017) A robust methodology for modal parameters estimation applied to shm. Mech Syst Signal Process 95:24–41. https://doi.org/10.1016/j.ymssp.2017.03.021
27. Zeng J, Hoon Kim Y (2021) A two-stage framework for automated operational modal identification. Structure and Infrastructure Engineering, 1–20. https://doi.org/10.1080/15732479.2021.1919151
28. Wu G, He M, Liang P, Ye C, Xu Y (2020) Automated modal identification based on improved clustering method. Mathematical Problems in Engineering 2020. https://doi.org/10.1155/2020/5698609
29. Lindfield G, Penny J (2019) Chapter 5 - solution of differential equations. In: Lindfield G, Penny J (eds.) Numerical Methods (Fourth Edition), Fourth edition edn., pp. 239–299. Academic Press. https://doi.org/10.1016/B978-0-12-812256-3.00014-2
30. Peeters B, De Roeck G (1999) Reference-based stochastic subspace identification for output-only modal analysis. Mech Syst Signal Process 13(6):855–878. https://doi.org/10.1006/mssp.1999.1249
31. Neu E, Janser F, Khatibi AA, Orifici AC (2017) Fully automated operational modal analysis using multi-stage clustering. Mech Syst Signal Process 84:308–323. https://doi.org/10.1016/j.ymssp.2016.07.031
32. Allemang RJ, Brown DL (1982) A correlation coefficient for modal vector analysis. Proceedings of the 1st International Modal Analysis Conference, Orlando, FL
33. Dunn JC (1973) A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters
34. Bezdek JC, Ehrlich R, Full W (1984) Fcm: The fuzzy c-means clustering algorithm. Comput Geosci 10(2):191–203. https://doi.org/10.1016/0098-3004(84)90020-7
35. Charbonnel PÉ (2021) Fuzzy-driven strategy for fully automated modal analysis: Application to the smart2013 shaking-table test campaign. Mech Syst Signal Process 152:107388. https://doi.org/10.1016/j.ymssp.2020.107388
36. Yang X-M, Yi T-H, Qu C-X, Li H-N, Liu H (2019) Automated eigensystem realization algorithm for operational modal identification of bridge structures. J Aerosp Eng 32(2):04018148. https://doi.org/10.1061/(ASCE)AS.1943-5525.0000984
37. Reynders E, Maes K, Lombaert G, De Roeck G (2016) Uncertainty quantification in operational modal analysis with stochastic subspace identification: Validation and applications. Mech Syst Signal Process 66–67:13–30. https://doi.org/10.1016/j.ymssp.2015.04.018
38. Ardila Y, Gomez I, Villalba J, Aracayo L (2021) Validation of the uncertainty bounds on modal parameters identified with the ssi-cov method. Latin American Journal of Solids and Structures 18. https://doi.org/10.1590/1679-78256725
39. Verboven P, Par E, Guillaume P, Van Overmeire M (2001) Autonomous modal parameter estimation based on a statistical

frequency domain maximum likelihood approach. Proceedings of the International Modal Analysis Conference - IMAC 2
40. Nord TS, Petersen ØW, Hendrikse H (2019) Stochastic subspace identification of modal parameters during ice-structure interaction. Phil Trans R Soc A 377(2155):20190030. https://doi.org/10.1098/rsta.2019.0030
41. Pereira S, Reynders E, Magalhães F, Cunha Á, Gomes JP (2020) The role of modal parameters uncertainty estimation in automated modal identification, modal tracking and data normalization. Eng Struct 224:111208. https://doi.org/10.1016/j.engstruct.2020.111208
42. He M, Liang P, Li J, Zhang Y, Liu Y (2021) Fully automated precise operational modal identification. Eng Struct 234:111988. https://doi.org/10.1016/j.engstruct.2021.111988
43. Ye X, Huang P, Pan C, Mei L (2021) Innovative stabilization diagram for automated structural modal identification based on era and hierarchical cluster analysis. J Civ Struct Heal Monit 11(5):1355–1373. https://doi.org/10.1007/s13349-021-00514-8
44. Civera M, Mugnaini V, Zanotti Fragonara L (2022) Machine learning-based automatic operational modal analysis: A structural health monitoring application to masonry arch bridges. Structural Control and Health Monitoring, 3028. https://doi.org/10.1002/stc.3028
45. Marrongelli G, Magalhães F, Cunha Á (2017) Automated operational modal analysis of an arch bridge considering the influence of the parametric methods inputs. Procedia engineering 199:2172–2177. https://doi.org/10.1016/j.proeng.2017.09.170

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.