Automated uncertainty-based extraction of modal parameters from stabilization diagrams

Johann Priou, Szymon Gres, Matthieu Perrault, Laurent Guerineau, Michael Döhler

9th International Operational Modal Analysis Conference
iomac
July 3 - 6, 2022
Vancouver, Canada

AUTOMATED UNCERTAINTY-BASED EXTRACTION OF MODAL PARAMETERS FROM STABILIZATION DIAGRAMS

Johann Priou[1], *Szymon Greś*[2], *Matthieu Perrault*[3], *Laurent Guerineau*[3], *Michael Döhler*[1]

[1] Univ. Gustave Eiffel, Inria, COSYS/SII, I4S, Rennes, France, {johann.priou,michael.doehler}@inria.fr
[2] Institute of Structural Engineering (IBK), SMM team, ETH Zürich, Zürich, Switzerland
[3] Sercel, Carquefou, France

ABSTRACT
The interpretation of stabilization diagrams is a classical task in operational modal analysis, and has the goal to obtain the set of physical modal parameters from estimates at the different model orders of the diagram. The diagrams are contaminated by spurious modes that appear due to the unknown (non-white) ambient excitation and sensor noise, as well as possible over-modelling. Under the premise that spurious modes will vary and physical modes will remain quite constant at different model orders, the focus is to retrieve the physical modes that constitute the identified model, while rejecting the non-physical, spurious modes. Over the last decade, extensive research has been devoted for developing automated strategies facilitating their interpretation. To this end, the interpretation is in principle disconnected from the identification method and boils down to three stages i.e., clearing the diagram from the spurious mode estimates, aggregating the modal parameter estimates in modal alignments and the final parameter choice. Besides the point estimates of the modal parameters, also their confidence bounds are available with some identification methods, such as subspace identification. These uncertainties provide useful information for an automated interpretation of the stabilization diagrams. First, modes with high uncertainty are most likely non-physical modes. Second, the confidence bounds provide a natural threshold for the automated extraction of modal alignments, avoiding the requirement of a deterministic threshold regarding the allowable variation within an alignment. In this paper, a strategy is presented for the automated mode extraction considering their uncertainties, based on clustering a statistical distance measures between the modes. The relevance of the uncertainty consideration in the automated extraction will be demonstrated on vibration data from two bridges.

Keywords: Operational Modal Analysis, uncertainty quantification, stabilization diagram, automated interpretation, hierarchical clustering, subspace methods

1. INTRODUCTION

Operational modal analysis is a fundamental task in engineering practice. Assuming vibrating structures with a linear time-invariant behaviour, linear system identification techniques such as subspace-based system identification [1, 2] can be used, where the system matrices of a state-space model are estimated before retrieving the modal parameters from their eigenstructure. The model order of the system (corresponding to the number of modes in the data) is in general unknown, and moreover non-physical modes due to colored noise are usually present besides the physical structural modes of interest. Under the premise that spurious modes will vary and physical modes will remain quite constant at different model orders, the identification is repeated at different model orders using efficient algorithms [2] to obtain the stabilization diagram. Then the focus is to retrieve the physical modes that constitute the identified model, while rejecting the non-physical, spurious modes. With subspace-based methods, not only the modal parameter estimates but also their uncertainty bounds can be obtained [3–6], which are valuable information for clearing the diagrams of modes with high uncertainty and their further processing.

Over the last decade, extensive research has been devoted for developing automated strategies facilitating the interpretation of the stabilization diagrams, e.g. [7–9], where an overview is given in [10]. Such strategies usually comprise three stages i.e., clearing the diagram from the spurious mode estimates, aggregating the modal parameter estimates in modal alignments and the final parameter choice. A popular tool for both the clearing and the aggregating stages is clustering based on a distance measure between the modes of the stabilization diagram. A drawback of clustering in this context is that either the number of clusters i.e., the number of modes, is usually assumed to be known a priori, or that thresholds on the allowable distance of the elements within a cluster need to be defined.

The focus of this work is to alleviate these issues to achieve less dependence on user-defined thresholds. First, a distance measure for clustering is proposed that is based on the modal parameters *and* their uncertainties. In this way, the natural variation of the modes between different model orders due to the intrinsic uncertainties is considered. Thresholds for the distances can then be defined based on a chosen confidence level on statistical grounds, instead of defining them in absolute terms on allowable changes in the modal parameters that could be quite different for different noise levels of the data. Second, to further reduce dependence on the user-defined thresholds, an automatic post-processing of the clusters with a merging and a separation step is introduced based on the cluster distances, where too tight or too loose thresholds are alleviated for the final clusters.

This paper is organized as follows. In Section 2., the basics of the subspace-based system identification and uncertainty quantification are recalled. In Section 3., the clustering strategy is developed for the automated interpretation of the stabilization diagrams and applied to vibration data of two bridges in Section 4., before concluding the paper in Section 5..

2. SUBSPACE-BASED SYSTEM IDENTIFICATION AND UNCERTAINTY QUANTIFICATION

2.1. System identification

Assume that the vibration behavior of the investigated structure can be modelled by a linear time-invariant system, then its dynamics can be described by the discrete-time state space model

$$\left\{ \begin{array}{rcl} \mathbf{x}_{k+1} & = & \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k & = & \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \end{array} \right. , \tag{1}$$

where $\mathbf{A}$ is the state transition matrix, $\mathbf{C}$ is the output matrix, and k is the integer time step corresponding to the system at time $t = k\Delta t$, where Δt is the sampling rate. Vector $\mathbf{y}_k \in \mathbb{R}^r$ contains the measured outputs (such as accelerations, velocities, displacements, strains), and $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector. The state noise $\mathbf{w}_k \in \mathbb{R}^n$ is related to the unknown ambient excitation, and vector $\mathbf{v}_k \in \mathbb{R}^r$ is the output noise. The modal parameters are related to the eigenvalues and eigenvectors (λ_i, ϕ_i), $i = 1, \ldots, n$, of $\mathbf{A}$ and to

$\mathbf{C}$ by

$$\mu_i = \frac{\log(\lambda_i)}{\Delta t}, \quad f_i = \frac{|\mu_i|}{2\pi}, \quad \zeta_i = \frac{-\mathrm{Re}(\mu_i)}{|\mu_i|}, \quad \varphi_i = \mathbf{C}\phi_i, \tag{2}$$

where μ_i is an eigenvalue of the corresponding continuous-time system, f_i is the natural frequency, ζ_i is the damping ratio and φ_i is the mode shape at the output coordinates.

To estimate the system matrices $\mathbf{A}$ and $\mathbf{C}$ from the output data $\mathbf{y}_k$ of length N, $k = 1, \ldots, N$, and consequently the modal parameters in (2), the reference-based covariance-driven subspace algorithm [1, 2] is used. The output covariance estimates with respect to a subset of reference sensors or projection channels are computed as $\hat{\boldsymbol{R}}_i = \frac{1}{N}\sum_{k=1}^{N} \mathbf{y}_{k+i}\mathbf{y}_k^{(\mathrm{ref})T}$. Arranged in block Hankel format, their theoretical values satisfy the decomposition

$$\boldsymbol{\mathcal{H}}_{p+1,q} = \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 & \ldots & \boldsymbol{R}_q \\ \boldsymbol{R}_2 & \boldsymbol{R}_3 & \ldots & \boldsymbol{R}_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}_{p+1} & \boldsymbol{R}_{p+2} & \ldots & \boldsymbol{R}_{p+q} \end{bmatrix} = \boldsymbol{\mathcal{O}}_{p+1}\boldsymbol{\mathcal{C}}_q, \quad \text{where} \quad \boldsymbol{\mathcal{O}}_{p+1} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^p \end{bmatrix} \tag{3}$$

with the observability and stochastic controllability matrices $\boldsymbol{\mathcal{O}}_{p+1}$ and $\boldsymbol{\mathcal{C}}_q$, and where p, and q are time lag parameters. With the singular value decomposition (SVD)

$$\hat{\boldsymbol{\mathcal{H}}}_{p+1,q} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_0 \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_0^T \end{bmatrix},$$

an estimate $\hat{\boldsymbol{\mathcal{O}}}_{p+1} = \mathbf{U}_1\mathbf{S}_1^{1/2}$ of the observability matrix is obtained for the model order corresponding to the truncation of the SVD. Then, the output matrix $\mathbf{C}$ is estimated from the first block row of the observability matrix, and the state transition matrix is estimated from the shift-invariance property $\hat{\mathbf{A}} = (\hat{\boldsymbol{\mathcal{O}}}_{p+1}^{\uparrow})^{\dagger}\hat{\boldsymbol{\mathcal{O}}}_{p+1}^{\downarrow}$ where $\hat{\boldsymbol{\mathcal{O}}}_{p+1}^{\uparrow}$ and $\hat{\boldsymbol{\mathcal{O}}}_{p+1}^{\downarrow}$ are the observability matrix estimate without the last and first block row, respectively. Ultimately, the modal parameters are obtained from the eigenvalues and eigenvectors of $\hat{\mathbf{A}}$ and from $\hat{\mathbf{C}}$ as in (2).

2.2. Uncertainty quantification

The computation of the modal parameter covariance results from the propagation of the sample covariance on the Hankel matrix estimate $\hat{\boldsymbol{\mathcal{H}}}_{p+1,q}$ through all steps of the modal identification algorithm. This sample covariance reflect in particular the unknown inputs due to non-measurable excitation sources and the sensor noise, and they contribute in a non-trivial way to the covariance of the modal parameter estimates. The estimation of the sample covariance of the Hankel matrix is straightforward and is done by separating the dataset for $k = 1, \ldots, N$ into blocks. The propagation to the modal parameter estimates is then based on the delta method [11], where the analytical sensitivity matrices are obtained from perturbation theory [3, 5].

Let ΔX be a first-order perturbation of a matrix-valued variable X, like $X = \boldsymbol{\mathcal{H}}_{p+1,q}$. Then, for a function $Y = f(X)$ it holds $\mathrm{vec}(\Delta Y) \approx \mathcal{J}_{Y,X}\,\mathrm{vec}(\Delta X)$, where $\mathcal{J}_{Y,X}$ is the derivative that can be obtained by analytically perturbing the functional relationship between X and Y, and where $\mathrm{vec}(\cdot)$ denotes the column stacking vectorization operator. Subsequently, covariance expressions for the estimates satisfy

$$\mathrm{cov}(\mathrm{vec}(\hat{Y})) \approx \mathcal{J}_{Y,X}\mathrm{cov}(\mathrm{vec}(\hat{X}))\mathcal{J}_{Y,X}^T. \tag{4}$$

In this context, the modal frequencies and damping ratios satisfy

$$\Delta f_i \approx \mathcal{J}_{f_i,\boldsymbol{\mathcal{H}}}\mathrm{vec}(\Delta\boldsymbol{\mathcal{H}}), \quad \Delta\zeta_i \approx \mathcal{J}_{\zeta_i,\boldsymbol{\mathcal{H}}}\mathrm{vec}(\Delta\boldsymbol{\mathcal{H}}), \tag{5}$$

where the analytical sensitivities are derived in detail in [3, 4], and the modal parameter covariance follows from (4) based on the Hankel matrix sample covariance and estimates of the respective sensitivities.

3. AUTOMATED INTERPRETATION OF STABILIZATION DIAGRAMS

Clustering techniques are well-suited for the automated interpretation of stabilization diagrams [8, 12], where data (the modes in the stabilization diagram) are regrouped according to similar properties into clusters (the modal alignments) with respect to a chosen distance measure. If the distance between two modes is "small", they are grouped together in one cluster. The clusters should contain modes with stable properties over several model orders. This section will first define appropriate distance measures, and in a second part will describe the clustering process to obtain alignments.

3.1. Definition of distance

Amongst several distances for clustering [8], a basic distance is based on the frequency difference and the MAC between two modes i and j in the stabilization diagram, defined as [7]

$$D_{ij}^{c} = \frac{|f_i - f_j|}{\max(f_i, f_j)} + (1 - MAC_{ij}). \tag{6}$$

Here the absence of the damping ratio can be noticed, which is often not used for clustering due to its high variability. The distance is normed in the sense that both the (normed) frequency and the MAC are in the interval $[0,\ 1]$, leading to $D_{ij} \in [0,\ 2]$ where D_{ij} indicates a perfect match between the modes.

Despite its high variability, the damping ratio is an important information to be considered in the distance measure, which can be achieved by appropriate weighting in

$$D_{ij}^{n} = \alpha_1 \frac{|f_i - f_j|}{\max(f_i, f_j)} + \alpha_2 \frac{|\zeta_i - \zeta_j|}{\max(\zeta_i, \zeta_j)} + \alpha_3 (1 - MAC_{ij}), \tag{7}$$

with the weighting factors $\alpha_l \in [0,\ 1]$ with $\sum \alpha_l = 1$ the distance can be interpreted as a change in percent between two modes. In the classical distance (6) a 50% weighting is chosen by default. Here a weighting is chosen as $\alpha = [50\%,\ 5\%,\ 45\%]$ between the modal parameters, reflecting the high uncertainty of the damping ratio. The threshold of the distance for elements in the same cluster is chosen ad-hoc as $D_{\lim} = 0.9$, corresponding to 10% allowable variation in each of the weighted modal parameter components.

The normed distance (7) is based on the (deterministic) relative difference between the modal parameters. Since the modal parameter uncertainties are available from their estimation as outlined in Section 2.2., this uncertainty can be used as a natural weighting in the computation of the distance. Considering the uncertainties of the frequencies and damping ratios, the first part of the distance can be evaluated statistically by

$$S_{ij} = \begin{bmatrix} f_i - f_j \\ d_i - d_j \end{bmatrix}^T \Sigma_{ij}^{-1} \begin{bmatrix} f_i - f_j \\ d_i - d_j \end{bmatrix} \tag{8}$$

where

$$\Sigma_{ij} = \mathbf{E}\left(\begin{bmatrix} \Delta(f_i - f_j) \\ \Delta(d_i - d_j) \end{bmatrix} \begin{bmatrix} \Delta(f_i - f_j) \\ \Delta(d_i - d_j) \end{bmatrix}^T \right),$$

which can be evaluated based on Equations (4) and (5). Since the modal parameter estimates are asymptotically Gaussian distributed [5], the variable S_{ij} is asymptotically χ^2 distributed with two degrees of freedom. Its non-centrality parameter is different from zero if the theoretical values of the frequencies and damping ratios are different. Hence, a known value for a target percentile of the distribution is available for verifying the assumption that the theoretical values of the frequency and damping ratios are equal, thus belonging to the same cluster. When targeting a 99% percentile, the corresponding threshold obtained from the χ^2 distribution is around 9.2, so modes with $S_{ij} < 9.2$ can be considered to be a member of the same cluster. A criterion for the MAC can also be added on statistical grounds based

on [13]; however, for ease of presentation its deterministic format is kept in the following, defining the combined distance, also called statistical distance,

$$D_{ij}^{s} = S_{ij} + \tilde{\alpha}(1 - MAC_{ij}). \tag{9}$$

with an appropriate weighting $\tilde{\alpha}$. This weighting factor can be chosen as follows. Considering a weighting of 45% of the MAC to the distance measure (9) as in (7), and the remaining 55% to S_{ij} having a threshold of 9.2, a total threshold for D_{ij}^{s} in (9) is set to $D_{\mathrm{lim}}^{s} = 16$. Then, assuming a criterion for the allowable change of the MAC in the order of $MAC_{ij} > 0.9$, this corresponds to $\tilde{\alpha} = 70$.

3.2. Clustering

There are many methods for clustering in the literature; amongst the most known are k-means and hierarchical clustering [14] which are both in use for automated modal analysis. The k-means clustering requires the number of clusters as an input, which is hard to know beforehand in modal analysis. However, it is mostly used to distinguish spurious modes from physical ones [8, 15]. Hierarchical linkage clustering is based on a tree format where elements are clustered according to their distance, starting with the smallest one and stopping when a maximal distance D_{lim} is reached. In the first step, each element defines a cluster. In an iterative process, the two clusters with the shortest distance are combined into a larger cluster, where the distance between two clusters is defined in different ways. Setting a limit on the distances between the modes instead of defining the number of modes seems to be more natural for clustering the stabilization diagram; hence the hierarchical linkage clustering is chosen in the following.

3.2.1. Linkage

The linkage is the choice of how the elements (i.e., the modes) form a cluster. Two popular ways of linkage are complete and single. Single linkage regroups elements together by the minimum distance, i.e., the distance between two clusters is defined by the minimum of the distances between their respective elements when combining them. It can also be called nearest neighbor. On the other side, the complete linkage regroups the clusters by the maximum distance of their elements, also known as farthest neighbor. The direct consequence is that the complete linkage is more strict as the maximum distance of elements should be under a specified value, while the single linkage allows more flexibility. For the first clustering step, complete linkage is chosen to ensure that the distances between all cluster elements are below the respective threshold D_{lim} as defined previously. Resulting clusters with too few elements are deleted. In a second step, a post-processing of the retained clusters is carried out. To avoid modal alignments that are split over the model orders due to too low thresholds, the clusters are checked with the single linkage distance for merging. Furthermore, to avoid the contrary effect of over-agglomerated clusters with several modes at single model orders due to too high thresholds, a separation step is introduced.

3.2.2. Post-processing

As presented before, the clustering depends on the threshold D_{lim} which is an *a priori* choice to a greater or lesser extent – greater when considering deterministic distance measures and lesser when considering statistical distance measures. If the threshold is too low physical modes will be missing; if it is too high different modes will be agglomerated in the same cluster. For less impact of the choice of this parameter several correction steps have been developed.

Merge

The first correction step is merging. The complete linkage for the initial clustering is strict and may tends to divide single mode alignments into different clusters. To correct this, a near neighbor merging was implemented comparing close neighbors between them based on the single linkage approach. The *minimum* distance between two close clusters is computed, then if this distance is below the limit D_{lim} the two cluster are merged.

Separation

The second step is a separation. A global mode cannot have different estimates at the same model order, yet this can happen when the threshold is too big, or when noise modes with similar properties are nearby. The purpose is then to clean some outlier points that could have been mixed with a real mode, and to possibly separate two modes that are close. To clean these clusters, a reference "mono-order" subset is defined by the elements of the cluster with unique model orders, i.e. at the corresponding model orders there is only one mode estimate in the cluster. For the remaining model orders with multiple estimates the closest element to the reference sub-set is kept, and the others deleted if too close.

The final step is a size verification of the found cluster. Every cluster whose number of elements is under a stability criterion nb_{stab} is deleted. The criterion nb_{stab} is chosen as 20% of the maximum order computed. The modal parameters are then determined from the means of the cluster elements when using the classical normed distance, or from the weighted mean of the elements, where the weights are related to the modal parameter uncertainties at the different model orders [16].

Algorithm 1 Clustering

Require: Frequencies f, damping ratios ζ, mode shapes φ, model orders n, maximum distance D_{lim}, minimum number of modes per alignment nb_{stab}

Ensure: Clusters

Compute distance D_{ij} between every pair of elements (Equation (7) or (9))
Compute the clusters with complete linkage with respect to D_{lim}
Sort the clusters by frequency and delete every cluster whose size is below nb_{stab}
Merge
Compare the neighboring clusters and merge the closest if their single linkage distance is below D_{lim}
Separation
for Every cluster **do**
  while There are model orders with several modes **do**
    Compare the different modes at each order to the mode set at orders with single modes
    Keep the closest and delete the other
  end while
end for

4. APPLICATIONS

Vibration data of two bridges, namely S101 Bridge [17] and Z24 Bridge [18, 19] are used to apply the developed cluster algorithm 1 using either of the two distances defined in (7) and in (9), denoted by *normed distance* D^n and *statistical distance* D^s, respectively.

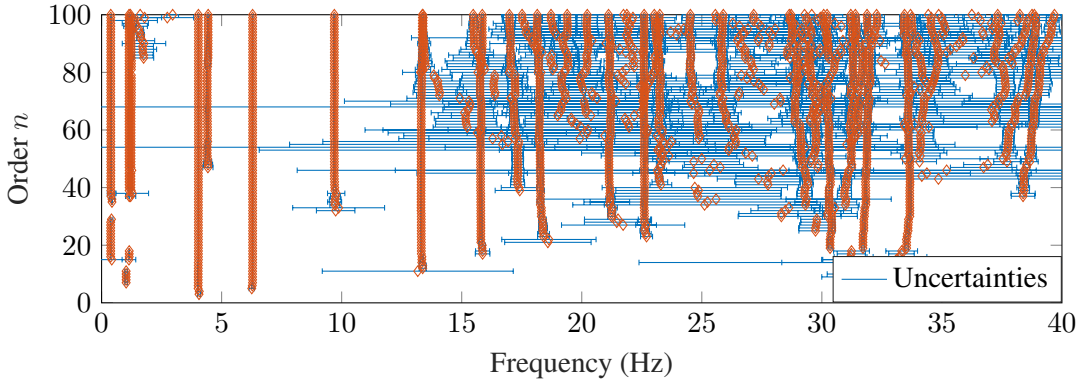

Figure 1: Full stabilization diagram of the S101 Bridge

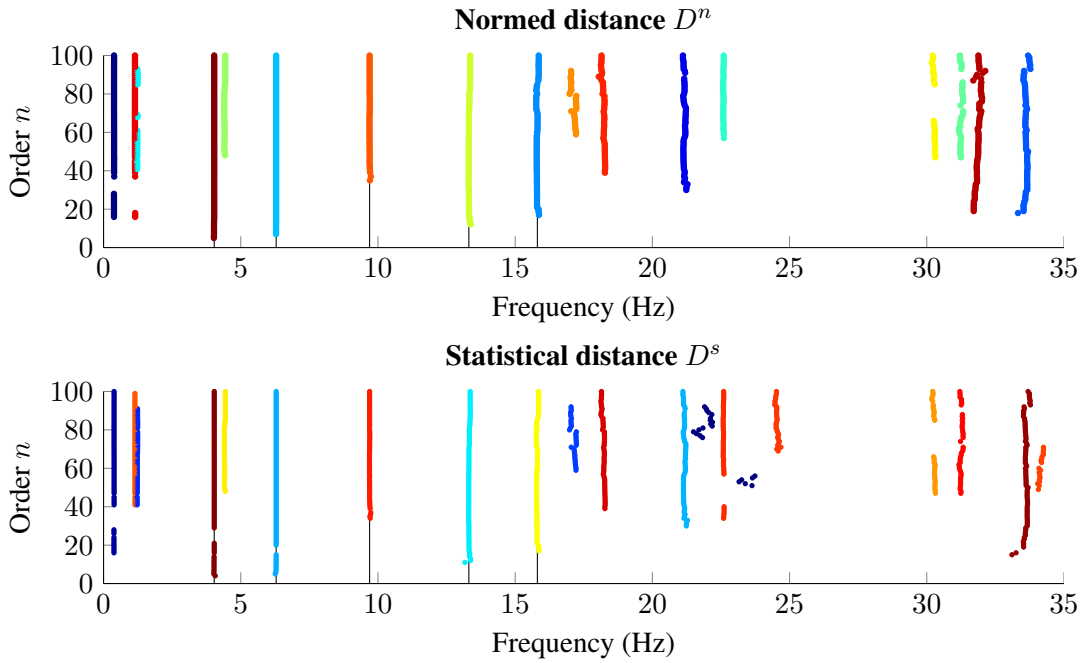

Figure 2: Hierarchical clustering from full stabilization diagram of the S101 Bridge (Figure 1) based on normed distance D^n (top) and statistical distance D^s (bottom)

4.1. S101 Bridge

One dataset of the reference state of S101 Bridge before damage [17] has been chosen for analysis, containing measurements from 45 sensors during 5.5 min. The respective stabilization diagram with the uncertainty bounds on the frequencies is shown in Figure 1, where the modal alignments are already quite clear.

In Figure 2, the clustering results with the developed algorithm are shown based on the classical normed distance (top) and the statistical distance (bottom). The five modes that were previously identified in [17] are shown as vertical lines in Figure 2. They are clearly present with both distances. This validates the use of statistical distance.

The uncertainty bounds of the modal parameters can also be used as hard criterion to delete very uncertain modes.With a threshold on the coefficient of variation of the frequencies (standard deviation of the frequency divided by the frequency), the stabilization diagram becomes clearer as shown in Figure 3 and seemingly easier to analyze. The clustering results based on the cleaned stabilization diagram are shown

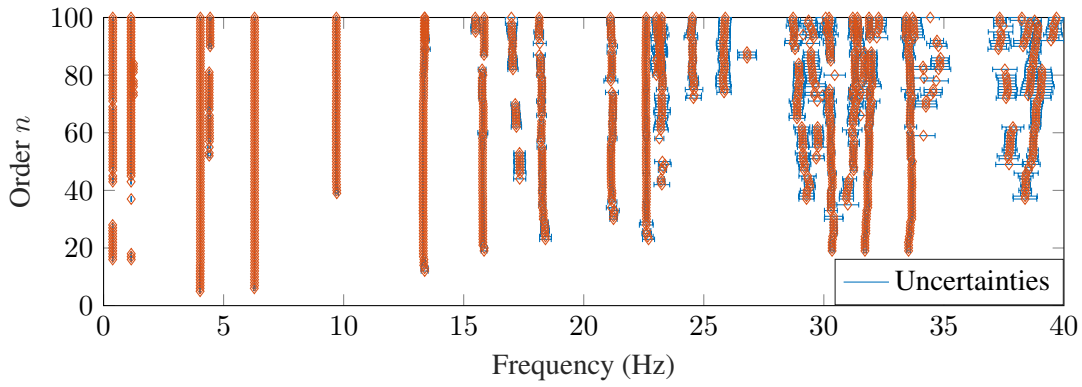

Figure 3: Cleaned stabilization diagram of the S101 Bridge with threshold on the coefficient of variation of frequencies of 1.5%

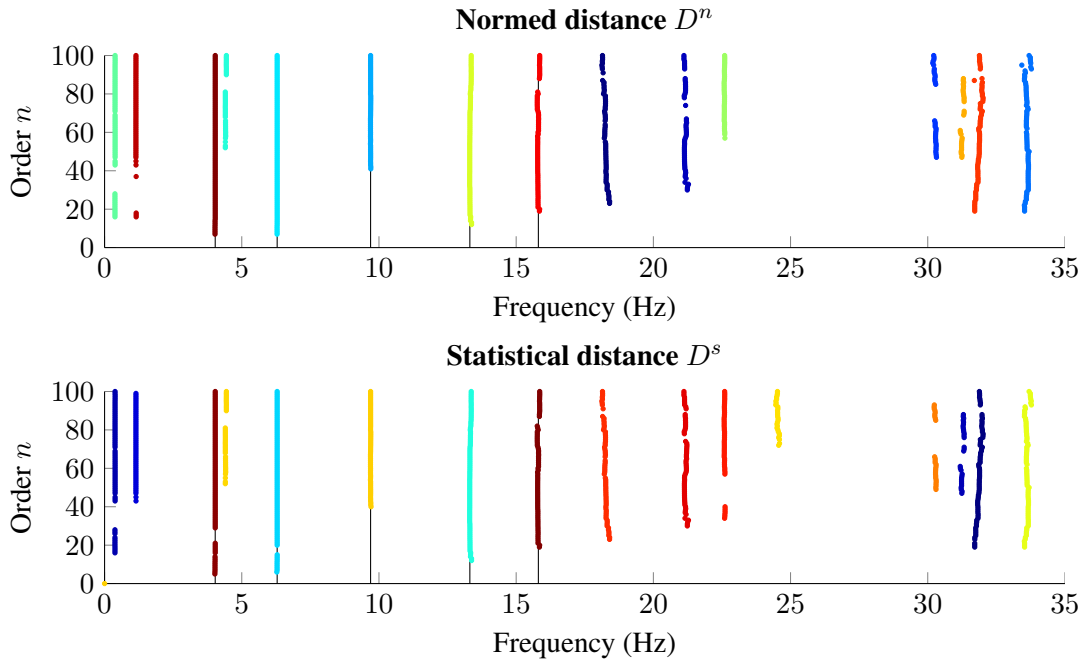

Figure 4: Hierarchical clustering from cleaned stabilization diagram of S101 Bridge (Figure 3) based on normed distance (top) and statistical distance (bottom)

in Figure 4, where they do not seem to be impacted by the application of the uncertainty threshold, while allowing as faster computation of the clustering.

The clustering algorithm is efficient to automatically retrieve stable alignments; however some alignments of spurious modes are also stable but tend to vary between different datasets. In order to retrieve the modes that are stable over different datasets, the same clustering technique can be applied on the set of modes that is obtained from each of the datasets in the spirit of "cross-validation" of the modes. Assuming only little change of the physical modes between the analyzed datasets, the same threshold for their distances as in the analysis of a single stabilization diagram can be used.

The results of this second clustering on the modal parameters obtained from 21 successive datasets in the reference state of S101 Bridge is shown in Figure 5 (left), where the five reference modes are

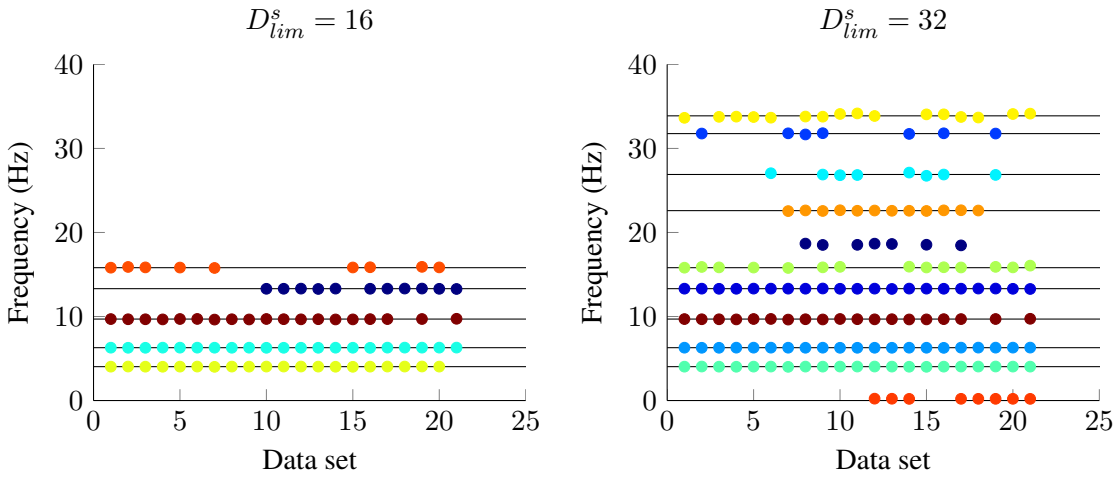

Figure 5: Hierarchical clustering of modes obtained from clustering of 21 stabilization diagrams using the statistical distance with the same threshold $D^s_{\text{lim}} = 16$ as in the clustering of the diagrams (left), and the doubled threshold (right)

well recognized and other modes with higher variability are dismissed. To account for higher mode variability between different datasets (as e.g. in tracking applications [12]), the threshold is usually increased. Results with the doubled threshold are shown in Figure 5 (right), where more stable modes are found. An in-depth analysis of these additional modes showed that the modes at 1 Hz and 18 Hz are most probably non-physical, while the other modes are indeed physical modes with lower excitation.

4.2. Z24 Bridge

The second example is the Z24 Bridge [18, 19], where one dataset of the multi-setup measurements has been used as in [4]. Measurements are more noisy, resulting in higher uncertainties of the modal parameters and more spurious modes, as shown in Figure 6, making the automated analysis more difficult. Nevertheless, the same thresholds as in the previous example are used for the clustering, and the impact of the developed steps for post-processing (merge and separate, see Algorithm 1) as well as the impact of the statistical distance measure are analyzed.

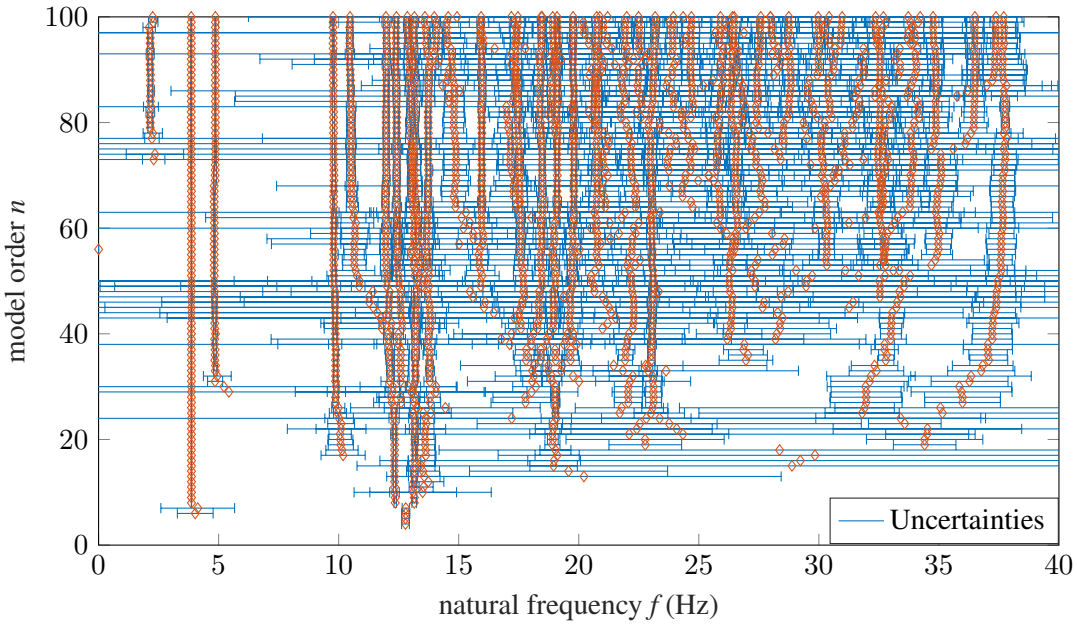

Figure 6: Full stabilization diagram of Z24 Bridge

In Figure 7 the result of the clustering is shown without any of the post-processing steps described in Section 3.2.2.. The black vertical lines represent the reference modes identified in [19]. Indeed, some of the alignments are split due to the more noisy data that would possibly require more relaxed thresholds. After the post-processing steps, the final clusters correspond well to the structural modes in Figure 8. However, there is one exception, namely the mode at 17 Hz that is very lowly excited and shows high variability. It is missing when using the classical normed distance; however the statistical distance considers its higher uncertainty intrinsically where it is well recognized. This shows that the statistical distance is more robust in a noisy situation.

It should be noted that it the effect of cleaning the diagram was tested by putting a threshold on the coefficient of variation of the frequencies, similarly as in Figures 2 and 3. First, since data is noisier, the threshold of 1.5% needed to be increased in order to keep modes with higher estimation uncertainty. Second, the results did not show improvement of the clustering, as already seen in Figure 3; moreover the mode at 17 Hz (which has a higher variation) tends to get lost if the threshold is not set very carefully. This shows that clustering strategy does not require a previous cleaning of the diagram, whereas such a cleaning may be helpful for a manual visual inspection of the diagram.

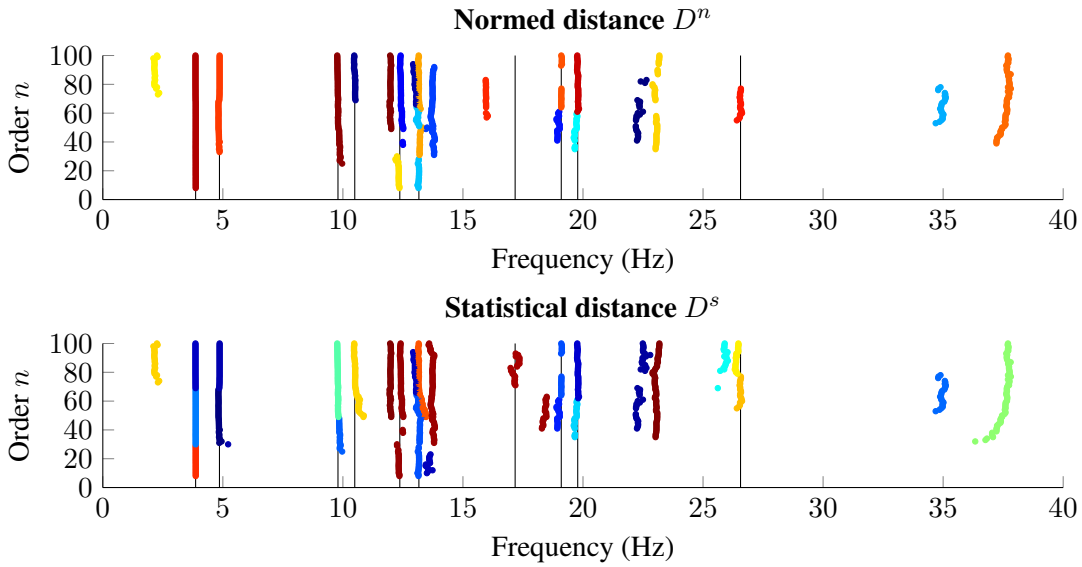

Figure 7: Hierarchical clustering without post-processing of the full stabilization diagram of Z24 Bridge based on normed distance (top) and statistical distance (bottom)

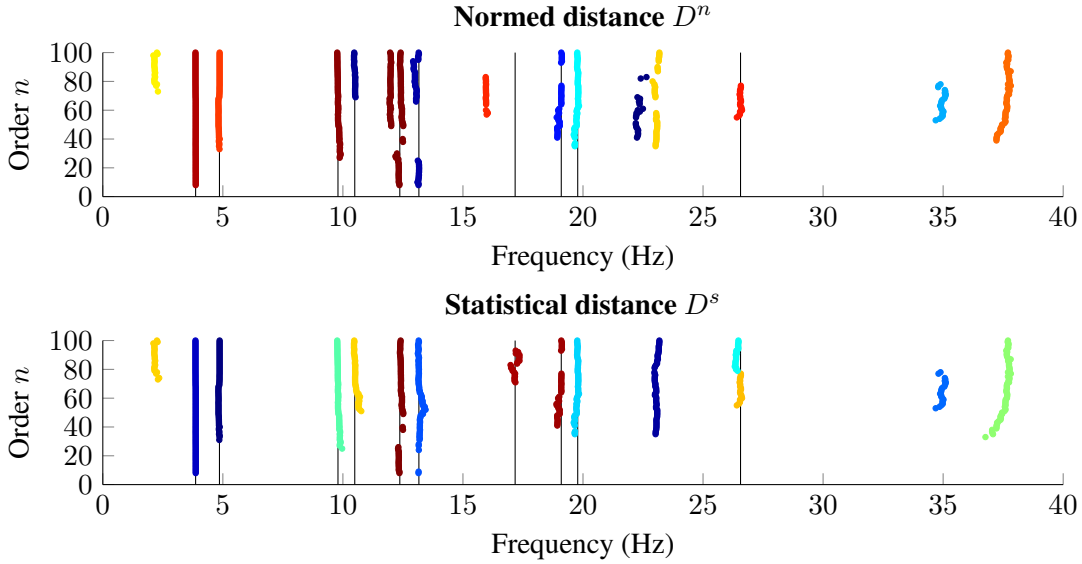

Figure 8: Hierarchical clustering after post-processing of the full stabilization diagram of Z24 Bridge based on normed distance (top) and statistical distance (bottom)

5. CONCLUSIONS

The proposed statistical distance for clustering of stabilization diagrams has been validated on the S101 Bridge data. Then an application on the Z24 Bridge showed that in difficult noise condition the statistical approach allows to find modal alignments even of very noisy modes due to its intrinsic consideration of the uncertainties, leading to a more robust algorithm. The examples also showed that clearing the diagrams with a hard criterion on modal uncertainty bounds may not be necessary, since the clustering algorithm takes them into account directly, and may even be counterproductive. Furthermore, it has been shown on the S101 Bridge data that the clustering approach is a means for cross-validation of modes from several datasets.

ACKNOWLEDGMENTS

The support from the ANR "France Relance" program is gratefully acknowledged.

REFERENCES

[1] Peeters, B., & De Roeck, G. (1999). Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing*, *13*(6), 855–878.

[2] Döhler, M., & Mevel, L. (2012). Fast multi-order computation of system matrices in subspace-based system identification. *Control Engineering Practice*, *20*(9), 882–894.

[3] 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.

[4] Döhler, M., & Mevel, L. (2013). Efficient multi-order uncertainty computation for stochastic subspace identification. *Mechanical Systems and Signal Processing*, *38*(2), 346–366.

[5] Mellinger, P., Döhler, M., & Mevel, L. (2016). Variance estimation of modal parameters from output-only and input/output subspace-based system identification. *Journal of Sound and Vibration*, *379*(100), 1–27.

[6] Greś, S., Riva, R., Süleyman, C. Y., Andersen, P., & Łuczak, M. (2022). Uncertainty quantification of modal parameter estimates obtained from subspace identification: An experimental validation on a laboratory test of a large-scale wind turbine blade. *Engineering Structures*, *256*, 114001.

[7] Magalhães, F., Cunha, Á., & Caetano, E. (2009). Online automatic identification of the modal parameters of a long span arch bridge. *Mechanical Systems and Signal Processing*, *23*(2), 316–329.

[8] Reynders, E., Houbrechts, J., & De Roeck, G. (2012). Fully automated (operational) modal analysis. *Mechanical Systems and Signal Processing*, *29*, 228–250.

[9] Yaghoubi, V., Vakilzadeh, M. K., & Abrahamsson, T. J. (2018). Automated modal parameter estimation using correlation analysis and bootstrap sampling. *Mechanical Systems and Signal Processing*, *100*, 289–310.

[10] Zini, G., Betti, M., & Bartoli, G. (2022). A quality-based automated procedure for operational modal analysis. *Mechanical Systems and Signal Processing*, *164*, 108173.

[11] Casella, G., & Berger, R. (2002). *Statistical inference*. Duxbury Press.

[12] Pereira, S., Magalhães, F., Gomes, J., & Cunha, A. (2022). Modal tracking under large environmental influence. *Journal of Civil Structural Health Monitoring*, *12*.

[13] Greś, S., Döhler, M., & Mevel, L. (2021). Uncertainty quantification of the Modal Assurance Criterion in Operational Modal Analysis. *Mechanical Systems and Signal Processing*, *152*, 107457.

[14] Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification* (2nd). John Wiley & Sons.

[15] Neu, E., Janser, F., Khatibi, A. A., & Orifici, A. C. (2017). Fully Automated Operational Modal Analysis using multi-stage clustering. *Mechanical Systems and Signal Processing*, *84*, 308–323.

[16] Döhler, M., Andersen, P., & Mevel, L. (2017). Variance computation of modal parameter estimates from UPC subspace identification. *7th International Operational Modal Analysis Conference*.

[17] Döhler, M., Hille, F., Mevel, L., & Rücker, W. (2014). Structural health monitoring with statistical methods during progressive damage test of S101 Bridge. *Engineering Structures*, *69*, 183–193.

[18] Reynders, E., & De Roeck, G. (2009). Continuous vibration monitoring and progressive damage testing on the Z24 bridge. *Encyclopedia of Structural Health Monitoring*.

[19] Döhler, M., Lam, X.-B., & Mevel, L. (2013). Uncertainty quantification for modal parameters from stochastic subspace identification on multi-setup measurements. *Mechanical Systems and Signal Processing*, *36*(2), 562–581.