Machine-learning-driven automatic application of the stochastic subspace identification method

Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 64 (2024) 507–514

Structural Integrity Procedia

www.elsevier.com/locate/procedia

SMAR 2024 – 7th International Conference on Smart Monitoring, Assessment and Rehabilitation of Civil Structures

Machine-learning-driven automatic application of the stochastic subspace identification method

Marco Martino Rosso[a,*], Angelo Aloisio[b], Giuseppe Carlo Marano[a], Giuseppe Quaranta[c]

[a]*Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Corso Duca Degli Abruzzi 24, 10128 Turin, Italy*
[b]*Civil, Environmental and Architectural Engineering Department, Universita' degli Studi dell'Aquila, Via Giovanni Gronchi 18, L'Aquila, Italy*
[c]*Department of Structural and Geotechnical Engineering, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy*

Abstract

Vibration-based operational modal analysis (OMA) methods have been proven effective in identifying dynamic properties of existing structures and infrastructures under operational conditions. Nevertheless, the provision and installation of continuous monitoring systems for long-term structural health monitoring (SHM) purposes potentially applicable to the entire infrastructure networks or to the regional scale of existing vulnerable building heritage require significant economic planning efforts. Nowadays research trends are oriented toward developing effective automatic OMA (AOMA) methods for setting up novel and efficient long-term SHM solutions. The current study illustrates a new recent paradigm for the automatic output-only modal identification of linear structures under ambient vibrations called intelligent automatic operational modal analysis (i-AOMA). The proposed approach relies on the covariance-based stochastic subspace identification (SSI-cov) algorithm and effectively integrates a machine learning intelligent core, i.e. a random forest (RF) classifier, in a conceptually two steps procedure, i.e. an explorative phase and an intelligently-driven phase. The i-AOMA procedure provided a new framework that requires a minimum intervention to the user and is potentially able to deliver uncertainty measures of the modal parameters' estimates based on the explored SSI-cov control parameters. An application on a shear-type RC frame building typical of existing heritage in Italy is herein discussed and reported.

Keywords: Machine learning; Operational modal analysis; Random Forest; Stochastic subspace identification

* Corresponding author. Tel.: +39-011-090-4907.
E-mail address: marco.rosso@polito.it

10.1016/j.prostr.2024.09.295

1. Introduction

Monitoring the health status of existing structures and infrastructures over time is an essential and extremely topical issue worldwide. Severe economic and life losses are usually associated with structural failures, especially when strategic constructions are involved. Among the different causes of bridge failure, natural phenomena represent the preponderant source, principally floods and scour, see, e.g., Deng et al. (2016) and Fig. 1 inspired by Di Prisco et al. (2019), even though human factors are another decisive aspect affecting the bridges' remaining life and safety levels. The destructive impact of natural hazards is often enhanced by maintenance negligence and/or inadequacy of old structures. In addition, collapses that happened in recent years have particularly shaken public opinion, rekindling the general interest in investing in the Structural Civil Engineering sector, especially for risk mitigation strategies and promoting smart and innovative solutions to ensure and preserve the safety levels of our existing heritage, as reported in Deng et al. (2016).

Focusing on the Italian scenario, due to its varied and widespread orographic and hydrographic features, there is one of the most complex transportation networks in the world. On the other hand, concerning the Italian buildings existing heritage, nowadays it appears in a generalized old conservation state. Before the 1920s, the most used building typologies for residential purposes were exclusively masonry structures. Starting around the 1920s, load-bearing masonry began to be replaced even more often by reinforced concrete (RC) frames, and afterward leading to the RC frames' predominance for new constructions after the 1970s. It is worth underlining that following the 15th Italian National Institute of Statistics (ISTAT) general census of 2011, the prevalent structural typology of existing building heritage is nowadays still represented by ancient masonry structures, followed by old RC frame buildings often designed under gravity static loads only. However, focusing on natural hazards, earthquakes remain a significant concern for the resilience of the built environment, especially in seismic areas likewise in Italy. The vulnerability assessment of existing building heritage still poses significant challenges attributable to large uncertainties related to unknown material properties, lack of information about structural details, and undocumented previous structural interventions. Operational modal analysis (OMA) based on output-only operational vibration tests has demonstrated especially attractive to support the development and validation of numerical models employed for seismic assessment and retrofitting because it enables the collection of relevant experimental data in a short time while minimizing interference with the structure.

All the so far debated social and economic aspects motivate the growing and significant demand for the scientific community to develop effective and innovative smart structural health monitoring (SHM) solutions to be implemented also in historical buildings and infrastructures, bestowing them smart features to increase their conservation level ensuring enough safety levels. As discussed by Kanda et al. (2021), in Japan, due to the severe seismicity, the SHM solutions deployment on existing buildings already started in the 1950s, but it sharply rose and widespread in the last two decades, afterward the 1995 Kobe earthquake. In detail, before the 2011 Tohoky earthquake, only 150 buildings were equipped with an SHM system, and the number sharply rose to 500 in 2016. In 2018, it was estimated that about 850 buildings were equipped with an SHM system, often installed voluntarily by owners in the private sector, counting about 700 out of the estimated 850, see Kanda et al. (2021).

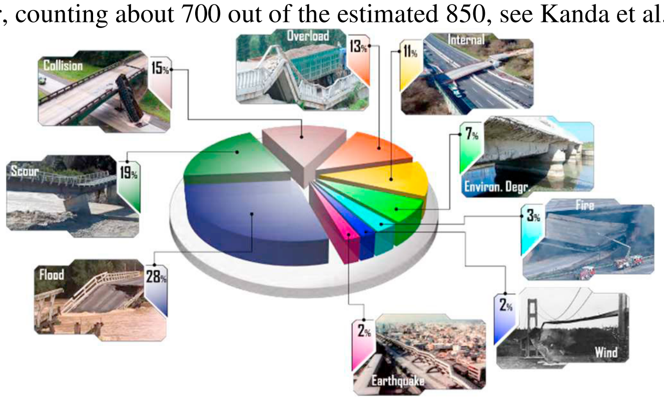


Fig. 1. Bridge failure causes pie chart inspired by Deng et al. (2016) and Di Prisco et al. (2019).

The growing and widespread adoption of SHM requires efficient solutions that aim at automatizing the extraction of the modal parameters (viz., natural frequencies, mode shapes, and damping ratios) from the recorded dynamic response of the structures. This need originated the development of some strategies able to facilitate the identification of the modal parameters under free or ambient vibrations, in such a way as to mitigate the influence of analyst's decisions on the whole elaboration process, see e.g. Magalhães et al (2012). Within this framework, as reported in Rainieri and Fabbrocino (2014), the stochastic subspace identification algorithm in its covariance-based formulation (SSI-cov) is often considered for the automatic operational modal analysis (AOMA) of linear structures subjected to ambient vibrations. Moreover, the nowadays novel and effective integration of machine learning (ML) driven methods is particularly attractive to overcome limitations of traditional existing AOMA methods.

In the current document, a new recently proposed AOMA approach is presented, which is devoted to identifying the modal characteristics of structures from ambient vibrations named Intelligent Operational Modal Analysis (i-AOMA), for further details please refer to Rosso et al. (2023). The method utilizes the SSI-cov algorithm for modal identification and is conceptually divided into two main steps, detailed in the next section 2. In sections 3 and 4, the proposed approach is validated numerically to demonstrate its potential for future seismic assessment and retrofitting applications of an existing RC frame structure case study located in Northern Italy.

2. Intelligent Automatic Operational Modal Analysis framework

As stated in Rainieri and Fabbrocino (2014), after the execution of the SSI-cov algorithm, the modal parameters of interest of the structure under investigation are derived from a graphical representation denoted as the stabilization diagram. This graph reports the natural frequency solutions of the identification algorithm versus the model order in which these solutions have been obtained. Nevertheless, due to noise in monitored vibration response data, and even due to a conservative model order overestimation method normally adopted in SSI-cov, spurious fictitious poles appear along with the physical-based solutions. However, considering various SSI-cov analyses with different sets of control parameters (model order and block row parameter), Zhou et al. (2022) stated that these spurious poles appear only occasionally, whilst actual physically-based ones are recursively appearing during the various analyses. Starting from this main observation, the authors in Rosso et al. (2023) formulated a new AOMA method that effectively integrates the automatic learning capabilities of machine learning (ML) methods, named intelligent AOMA, or briefly i-AOMA, and the Python code is freely available at the following GitHub repository https://github.com/marco-rosso-m/i-AOMA .

The i-AOMA methodology can be described by splitting the entire procedure into two main steps. As illustrated in Fig. 2, step 1 aims to make an initial exploration of SSI-cov control parameters. Specifically, this exploration is conducted with a quasi-Monte-Carlo sampling based on the Halton sampling scheme. The i-AOMA automatically defines the admissible intervals in which sampling the four main control parameters herein considered, i.e. being the model order, the block row parameter, the time window in which slicing the acceleration response, and the time instant in which centering the slicing time window. The automation level of the software is further enhanced in the i-AOMA because those sets of non-admissible control parameters (see Rainieri and Fabbrocino, 2014) or those sets that lead to SSI-cov computational time overpassing 30 seconds are excluded and labeled as unfeasible. Therefore, only minimum user intervention is required in this first step to basically define the number of useful simulations s to be collected in step 1. This is fundamental, since the exploration phase will serve as a database to train the intelligent core of the i-AOMA which will directly control the intelligent sampling in step 2 of the algorithm. Therefore, after collecting s useful SSI-cov results, the stability checks are performed (see Rainieri and Fabbrocino, 2014), and, retaining only fully stable poles, the s stabilization diagrams are overlapped and processed altogether. The nonparametric kernel density estimation (KDE) based on the automatic FFT-KDE implementation with the improved Sheather–Jones (ISJ) algorithm is employed to process the overlapped stabilization diagrams because of its advantageous automation level and efficiency, preferred rather than other traditional clustering methods (see Rosso et al, 2023). The resulting normalized KDE graph exhibits highly sharp peaks only in those natural frequencies associated with the most recurrent stable poles alignments, and the poles belonging to these alignments are distilled considering a retaining band calibrated according to the ISJ-based FFT-KDE estimated bandwidth parameter jointly with a statistical-based prominence approach to select only peaks of interest and excluding possible noisy ones. Finally, an information content (IC) is thus determined and associated with every SSI-cov control parameters set for distinguishing between informative and non-informative SSI-cov simulations. The IC has been simply calculated as the ratio between the number of poles falling within the KDE-based frequency retaining bands and the total number of stable poles of the stabilization diagram associated with that specific control parameters set. Indeed, all the control parameters sets jointly with their relative IC values configured a labeled database to train the intelligent core of the i-AOMA, i.e. a random forest (RF) classifier.

As illustrated in Fig. 3, step 2 of the i-AOMA methodology relies on an RF intelligently driven quasi-Monte Carlo sampling process. The control parameters sets which are predicted to be informative by the RF are thus adopted for SSI-cov evaluation, whereas the other ones are immediately excluded, thus saving significant computational resources and time. The intelligently-driven sampling continues until convergence criteria are reached. In i-AOMA, the convergence criteria have been evaluated for a batch of analyses, e.g. every 50 SSI-cov useful results, and formulated according to a limited relative variation of the trace of the total sampling variance of the mode shapes within 2%. This convergence rule is also known in the literature as the acceptable shifting convergence band rule (ASCBR). Once the convergence is reached, all the useful SSI-cov fully stable poles are again overlapped in a comprehensive stabilization diagram, again post-processed with the FFT-KDE, thus deriving the final stable alignments of interest. All these SSI-cov useful results permit evaluating statistical metrics associated with the modal parameters (natural frequencies, damping ratios, and mode shapes) due to the uncertainty propagation of the various control parameters sets.

```
Define s                                                   Successful runs of the SSI-cov for training the RF algorithm
Define n_max, j_min, j_max, i_min, i_max                   Reasonable control parameter bounds
Generate quasi-random samples of the control parameters
while s successful runs of SSI-cov are not completed do
        try
                @check the execution time < 30 s           Admissible elaboration time
                Perform SSI-cov                            Compute the SD
                Normalize mode shapes
        except
                Set IC equal to zero                       Unfeasible set of control parameters
end while
Overlap all the SDs                                        Detect possibly stable poles
Check the poles stability
Perform KDE                                                Perform FFT-KDE with ISJ algorithm
Recognize certanly stable poles from the normalized KDE
Calculate IC
Binarize IC according to a given threshold
Training RF
```

Fig. 2. i-AOMA step 1 pseudocode.

```
Set the batch size b                                       Convergence check every b runs of the SSI-cov
Set a large number s_max                                   Reasonable control parameter bounds
Generate quasi-random samples of the control parameters
while s_max successful runs of SSI-cov are not completed & statistical convergence is not achieved do
        try
                if Samples of the control parameters are classified as feasible by RF algorithm then
                @check the execution time < 30 s           Admissible elaboration time
                Perform SSI-cov                            Compute the SD
                Normalize mode shapes
                        if Number of runs is multiple of b then
                                Overlap all the SDs
                                Check the oples stability
                                Perform KDE
                        end if
                else
                        Set IC equal to zero
                end if
        except
                Set IC equal to zero
end while
Overlap all the SDs
Check the poles stability                                  Detect possibly stable poles
Perform KDE                                                Perform FFT-KDE with ISJ algorithm
Select certanly stable poles from the normalized KDE
Modal parameters and corresponding confidence level
```

Fig. 3. i-AOMA step 2 pseudocode.

3. Numerical vibration response simulation of a monitored reinforced concrete existing building case study

Case study herein under investigation is a RC existing building located in Northern Italy. As illustrated in Fig. 4, the building is four-story, regular in plan and height, characterized by an inter-story height of 3.20 m (4.00 m for the first floor), with plan dimensions of 24.90 m x 13.90 m. RC columns have constant cross section of 55 x 30 cm for the entire elevation of the building. Beams' cross-sections at 1st floor are 55x30 cm, whereas at 2nd and 3rd floors are 50x30 cm, and at last 4th level are 40x30 cm. As illustrated in Fig. 5, due to its planar and elevation regularity, the building behavior is shear-type diaphragmatic, thus allowing to describe its dynamics with a 3D lumped mass multiple degrees of freedom (MDOF) system. Therefore, a point mass is concentrated in the centre of the mass of each floor, and it is associated to three generalized DOFs each, viz. two translations in X (**u**) and Y (**v**) directions respectively, and a rotation ($\boldsymbol{\theta}$) around the vertical axis. Indeed, this building is described by 12 DOFs, i.e. 4 translations for each principal direction and 4 rotations. The mass and stiffness matrices of this RC building have been determined, and the damping matrix was computed by assuming a constant damping ratio equal to 2% for every mode. 12 natural frequencies (1.23 Hz, 1.28 Hz, 2.01 Hz, 3.53 Hz, 3.69 Hz, 5.63 Hz, 6.00 Hz, 6.10 Hz, 7.34 Hz, 8.22 Hz, 9.54 Hz, and 12.95 Hz) have been found by solving the eigenvalue problem for this 3D lumped mass system. Afterward, in order to simulate a SHM system collecting ambient vibration responses, the following stochastic state space (SSS) representation was formulated (see Rainieri and Fabbrocino, 2014), i.e. accounting for a random white noise dynamic excitation applied at the ground level DOFs in the process noise term $\boldsymbol{w}(t)$:

$$\dot{\boldsymbol{x}}(t+dt) = \boldsymbol{A}\,\boldsymbol{x}(t) + \boldsymbol{w}(t) \tag{1}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\,\boldsymbol{x}(t) + \boldsymbol{v}(t) \tag{2}$$

Eq. (1) are called state equations, whereas Eq. (2) are the observation equations. The symbol **x**(t) indicates the state vector which encompasses the state variables, i.e. the displacements and velocities of the system, whereas $\dot{\boldsymbol{x}}$(t+dt) denotes the first derivative with respect to the time of the next-state prediction, and the symbol A indicates the state transition matrix. The symbol **y**(t) refers to the acceleration vibration responses, observed at any time instant due to the random dynamic excitation of the process noise term **w**(t). The acceleration responses were further contaminated by an additive zero-mean Gaussian white noise called measurement noise process **v**(t). Eventually, the symbol **C** denotes the output influence matrix. The SSS model numerically simulated a monitoring system at 200 Hz which recorded 1 hour of acceleration ambient vibration response sampled using one biaxial accelerometer placed on every floor of this building (8 acceleration histories encoded in **y**(t)).

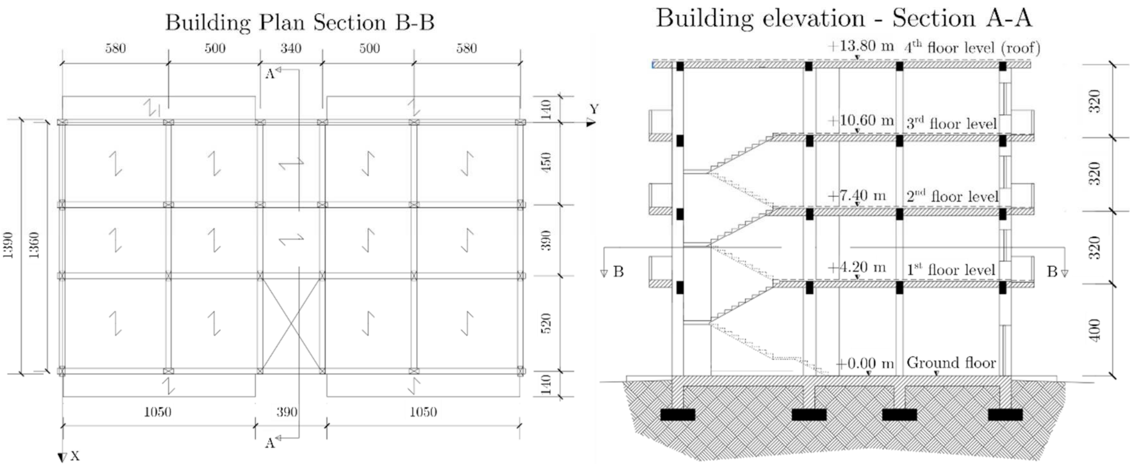


Fig. 4. Shear-type existing RC frame case study: plan view and lateral sectional view.

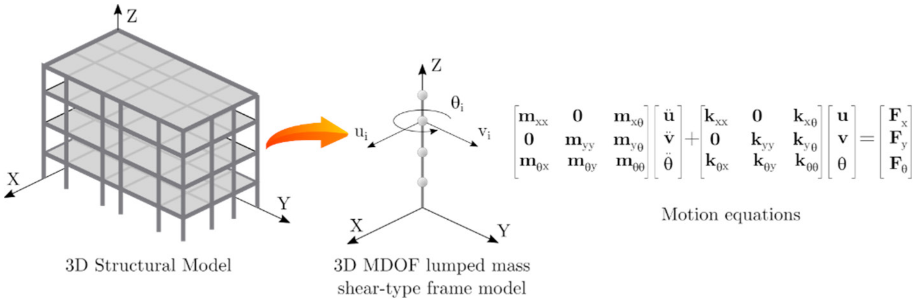


Fig. 5. Equivalent shear-type lumped mass 3D model of existing RC frame for the numerical validation of i-AOMA.

4. i-AOMA dynamic identification results and discussion

The numerical acceleration response data have been decimated with a factor equal to 5, thus restricting the frequency realizations domain within the Nyquist frequency of 20 Hz. The singular value decomposition of the power spectral density evidenced that the collected signals are informative since their peaks evidenced the natural frequencies of the system, except for the first mode at 1.23 Hz which appeared obscured by the spectral noise (see Fig. 6). Despite the SSI-cov control parameters sampling intervals have been automatically computed by the i-AOMA throughout the relationship illustrated in Rosso et al (2023), to further enhance the exploration step 1, the maximum model order has been limited to half of the theoretical possible value, i.e. to 680. Totally, 308 control parameters samplings have been carried out in the i-AOMA step 1 for collecting 200 user-defined successful SSI-cov analyses (success rate of 64.9 %). Consequently, to the stability checks, the stabilization diagram obtained by overlapping all the fully stable poles has been analyzed through the FFT-KDE algorithm delivering a bandwidth parameter of 0.00217 Hz. Therefore, 11 natural frequencies out of the theoretical 12 ones, and their stable poles' alignments, have been selected around the peaks of the normalized KDE, i.e. those peaks overcoming the statistical-based prominence value of 0.093.

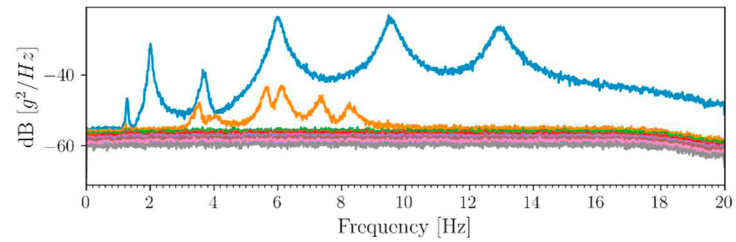
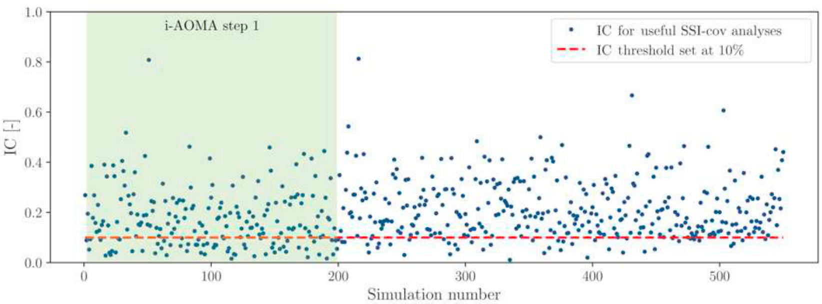


Fig. 6. Singular value decomposition of the power spectral density, typical of the frequency domain decomposition OMA method.



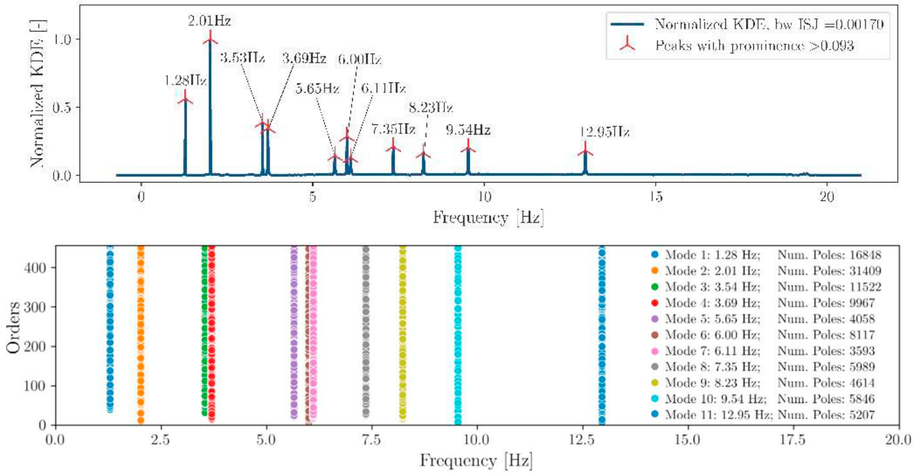


Fig. 7. i-AOMA step 2 results. From the top to the bottom: IC graph, FFT-KDE normalized graph, stable poles' alignments retained.

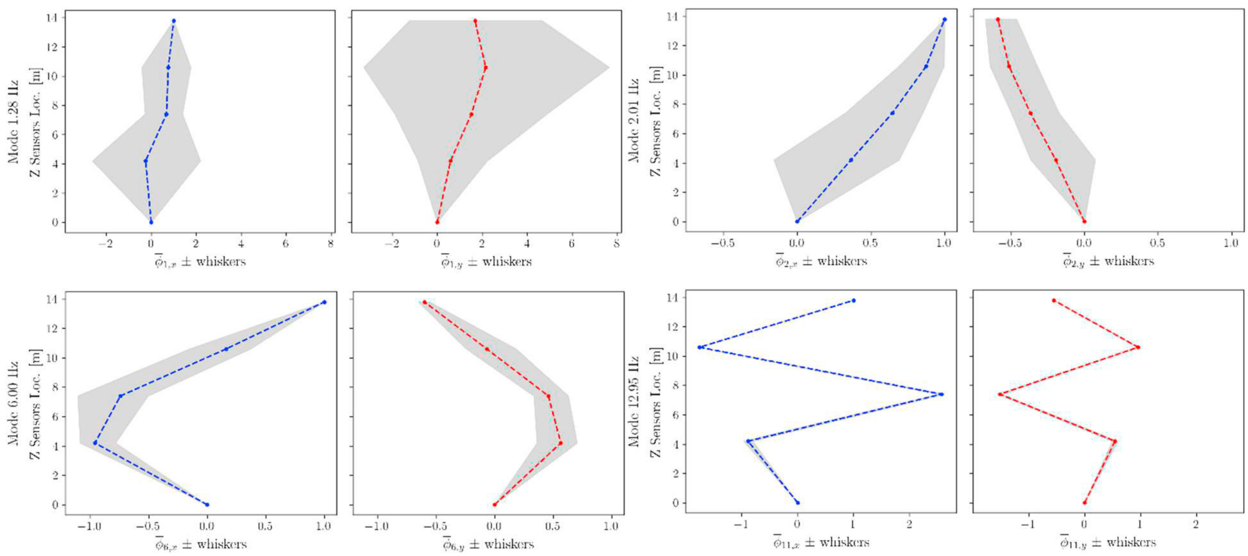


Fig. 8. i-AOMA sampled control parameters uncertainty propagation on mode shapes at modes No. 1, 2, 6, and 11.

Subsequently, the IC values have been computed to train the i-AOMA intelligent core RF classifier. In the i-AOMA step 2, new quasi-random samples of the control parameters have been generated, but only those predicted to be informative by the RF were actually employed to execute further SSI-cov analyses. Finally, further 350 useful results obtained by generating other 842 control parameters (RF filtering rate of 41.6 %), the ASCBR convergence criterion was met. The final convergence checks have been performed and overlapped stabilization diagrams have been processed by the KDE algorithm. A bandwidth parameter of 0.00170 Hz has been found, thus selecting the same 11 natural frequencies of step 1, as reported in Fig. 7. The uncertainties in the adoption of various control parameters set propagates to the final modal results. In detail, the 9 founded natural frequencies are 1.28 Hz, 2.01 Hz, 3.53 Hz, 3.69 Hz, 5.65 Hz, 6 Hz, 6.11 Hz, 7.35 Hz, 8.23 Hz, 9.54 Hz, and 12.95 Hz, all delivering the same standard deviation of about 0.001 Hz. Besides the first mode at 1.23 Hz which was lost due to excessive measurement additive noise, the obtained modal results are consistent with the modal analysis of the 3D lumped mass system. The respective estimates of damping ratios are in perfect agreement with the imposed 2 % at every mode: 2.15 ± 0.25%, 1.9 ± 0.21%, 1.96 ± 0.17%, 1.95 ± 0.24%, 1.94 ± 0.3%, 2.03 ± 0.22%, 1.92 ± 0.37%, 2.05 ± 0.18%, 1.9 ± 0.28%, 1.99 ± 0.26%, and 1.95 ± 0.29%. In conclusion, Fig. 8 depicts the control parameters uncertainties propagation on mode shapes in terms of median and boxplot whiskers, similarly to Rosso et al (2023).

5. Conclusions and remarks

Within the automatic operational modal analysis (AOMA) systems scenario for output-only vibration analysis, especially useful for continuous structural health monitoring (SHM), in the current contribution, the intelligent automatic operational modal analysis (i-AOMA) has been illustrated. This novel method attempted to overcome the arbitrary choice of the SSI-cov control parameters, permitting the exploration of various sets in reasonable ranges via a quasi-Monte Carlo sampling scheme. Moreover, the machine learning (ML) part has been effectively integrated within the proposed framework to save the computational burden traditionally associated with a Monte Carlo scheme to guess the quality of the modal results associated with a specific set of SSI-cov control parameters. Furthermore, all the stabilization diagrams associated with the various SSI-cov analyses are overlapped and comprehensively processed in one step using the efficient and automatic version of the nonparametric kernel density estimation (KDE) algorithm rather than a traditional clustering technique. In summary, the proposed i-AOMA framework has been formulated to increase the actual automation level of the existing AOMA methods, requiring a minimum intervention for the user to only setup the procedure the first time, and leveraging the AI and ML learning process, the system is able to autonomously recursively execute analysis automatically choosing the SSI-cov control parameters afterward this initial training phase. The effectiveness of the proposed i-AOMA approach has been herein validated on a numerical benchmark case referred to a typical archetype of existing RC frame building, which belong to existing seismic vulnerable heritage. The accurate modal identification, despite simulating using a single bi-axial accelerometer sensor for each floor, demonstrated the actual potentials of using the proposed i-AOMA procedure also for developing numerical models for seismic assessments purposes within earthquake engineering field.

Acknowledgements

Marco Martino Rosso, Giuseppe Carlo Marano and Giuseppe Quaranta acknowledge the support received through the project “Artificial Intelligence for SUstainable seismic risk reduction of STructures (AI-SUST)” (project code: 2022LEFKHS) funded by European Union – NextGenerationEU through the PRIN 2022 program of the Italian Ministry of University and Research (MUR) (D. D. n. 104, 02-02-2022). This work reflects only the authors’ views and opinions whereas the MUR cannot be considered responsible for them.

References

Deng, L., W. Wang, and Y. Yu. “State-of-the-art review on the causes and mechanisms of bridge collapse.” In: Journal of Performance of Constructed Facilities 30.2 (2016).

Di Prisco, M. et al. “Critical infrastructures in Italy: State of the art, case studies, rational approaches to select the intervention priorities.” In: Proceedings of the fib Symposium 2019: Concrete-Innovations in Materials, Design and Structures. International Federation for Structural Concrete. (2019).

Kanda, K., M. Nakashima, Y. Suzuki, and S. Ogasawara. ““q-NAVI”: A case of market-based implementation of structural health monitoring in Japan.” In: Earthquake Spectra 37.1 (2021).

Magalhães, F., Á. Cunha, and E. Caetano. “Vibration based structural health monitoring of an arch bridge: From automated OMA to damage detection.” In: Mechanical Systems and signal processing 28 (2012).

Mimmi, L. M. “Italy in Front of the Challenge of Infrastructure Maintenance: Existing Issues and Promising Responses.” In: Public Works Management & Policy (2023).

Rainieri, C. and G. Fabbrocino. Operational modal analysis of civil engineering structures. Vol. 142. Springer, (2014).

Rosso, Marco Martino, et al. "Intelligent automatic operational modal analysis." Mechanical Systems and Signal Processing 201 (2023).

Zhou, Kang, Qiu-Sheng Li, and Xu-Liang Han. "Modal identification of civil structures via stochastic subspace algorithm with Monte Carlo–based stabilization diagram." Journal of Structural Engineering 148.6 (2022).