Journal of Physics: Conference Series

**PAPER** • **OPEN ACCESS**

# Intelligent Automatic Operational Modal Analysis: preliminary results

To cite this article: Marco Martino Rosso *et al* 2024 *J. Phys.: Conf. Ser.* **2647** 212005

Journal of Physics: Conference Series **2647** (2024) 212005 | doi:10.1088/1742-6596/2647/21/212005

# Intelligent Automatic Operational Modal Analysis: preliminary results

**Marco Martino Rosso[1,*], Angelo Aloisio[2], Giuseppe Carlo Marano[1], and Giuseppe Quaranta[3]**

[1,*]DISEG, Politecnico di Torino, Corso Duca Degli Abruzzi 24, Turin, Italy
[2]DICEAA, Universita' degli Studi dell'Aquila, Via Giovanni Gronchi 18, L'Aquila, Italy
[3]Department of Structural and Geotechnical Engineering, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy
[*]Corresponding Author

E-mail: `marco.rosso@polito.it`

**Abstract.** Within the structural health monitoring (SHM) field, consistent research efforts have been invested in developing automatic vibration-based indirect methodologies for inspecting existing heritage conditions. Current trends are mainly focused on output-only automatic operational modal analysis (AOMA), specifically throughout the stochastic subspace identification (SSI) technique among others. In the literature, a widespread workflow is implemented in a four-step solution: choice of the SSI control input parameters, computation of stabilization diagrams, stable poles' alignments detection, and their final clustering. However, the so far proposed solutions have not provided yet complete answers to some challenging and still open questions. For instance, an arbitrarily poor initial choice of the SSI control parameters may jeopardize the entire procedure. Therefore, in the current study, the authors present a novel intelligent-based AOMA framework in a machine learning perspective. Specifically, random-forest-driven Monte Carlo sampling of control parameters represents a promising intelligent way to automatically choose the proper SSI control parameters. Furthermore, the recurrent stable physical poles in the stabilization diagram among the Monte Carlo simulations deliver some special insights about mode shape confidence intervals. A numerical benchmark is herein analyzed illustrating some preliminary results and potentials of the proposed methodology.

## 1. Introduction

Nowadays existing infrastructure heritage is widely approaching its nominal life, and safety issues are occurring due to advanced degradation phenomena [1]. Therefore, in the last decades, a growing interest of civil engineers and the scientific community was focused to develop cutting-edge structural health monitoring (SHM) methodologies, thus providing more efficient and reliable methods for inspecting actual conditions of existing heritage [2, 3]. In particular, indirect and non-destructive testing became tremendously popular due to their rapidity, efficiency, and less-invasivity features [4]. Specifically, vibration-based dynamic identification strategies provide special insights into the dynamic modal properties of investigated structures, i.e. mode shapes, natural frequencies, and damping ratios, under operational and ambient vibration conditions. The commonly adopted approach is represented by the operational modal analysis (OMA) techniques, which include two typologies of methods: the parametric procedures on one side, and the non-parametric ones on the other side [1, 5, 6]. Among the others, the stochastic subspace identification (SSI) method represent a highly acknowledged and widely adopted time-domain parametric strategy. Based on the state-space identification, it configures in subspace projections computation (denoting the SSI-data variant) [6, 7] or via output covariance estimates (thus designated as SSI-cov) [6, 8].

The majority of the current research efforts within the SHM field are devoted to developing automatic operational modal analysis (AOMA) methods, especially suited for long-term dynamic monitoring procedures [9]. In the scientific literature [8, 10], a widespread workflow is often implemented according to four main steps. The foremost phase is related to a starting arbitrary choice of the SSI control input parameters, especially the time shift (number of block rows of the Hankel matrix) and the maximum order of the stabilization diagram [6]. Secondly, the next phase is the computation of the stabilization diagrams, collecting stables and spurious poles from the numerical solution for varying the model order. Each pole is characterized by specific modal parameters. In the third phase the stability check criteria are performed in terms of frequency, damping, and mode shape [6], preparatory for the identification of stable poles' alignments along the order axis. Finally, in the current literature approaches, the automation and innovative machine learning-based (ML) part is mainly based on clustering techniques for post-processing stabilization diagrams and automatically collecting stable poles' alignments. The arbitrary choice of SSI control parameters within the first phase plays a key role, since a poor choice of these parameters may jeopardize the quality of the entire identification process and the resulting modal properties. Notwithstanding many scholars analyzed this issue providing some suggestions for a proper choice of SSI input parameters, e.g. [11, 12], to the authors' knowledge, no fully automated SSI parameters' selection procedures have been proposed hitherto yet. Therefore, in the current manuscript, the authors present a novel intelligent AOMA (intelligent-AOMA) procedure aiming to provide an automatic Monte Carlo-based sampling of the input parameters as described in detail in the next section.

The current manuscript is organized as follows. In the next section, the proposed intelligent-AOMA method is illustrated. Finally, a preliminary numerical benchmark is analyzed based on the five degrees-of-freedom (DOF) example proposed in [5]. This benchmark permitted us to illustrate some preliminary results, showing the new advantages and potentialities of the proposed solution, i.e. enhancing the automation level of actual AOMA available methods and leveraging novel capabilities offered by innovative ML perspectives.

## 2. Description of the proposed AOMA method

Generally, the SSI-cov algorithm may provide the identification of spurious fictitious modes in the conventional stabilization diagram, which are completely unrelated to the real structural ones of actual interest. This issue is inherent to the SSI algorithm itself and it arises because the actual DOFs of a real-world structure are unknown, thus a conservative approach consists in an overestimation of the structural system order, however resulting in nonphysical poles next to the physical ones [6]. Moreover, the worthless SSI-cov's results may be further enhanced by the user's poor choice of the above-mentioned SSI-cov's governing parameters. Therefore, in [13, 14], the authors recently attempted to improve the SSI's accuracy in detecting the real structural modal parameters. They proposed a Monte Carlo-based stabilization diagram definition by considering two varying input parameters, i.e. the length of the signals' according to a shorter time window and the SSI-cov maximum model order. The time window extracts a portion of the entire signal with a specific time window length, symmetrically centered to a randomly generated time instant. In [13, 14], the authors demonstrated that executing a considerable number of Monte Carlo simulations of those parameters, and consequently of SSI-cov analyses, the spurious modal alignments were detected only occasionally whereas the actual physical structural modes were identified in virtually all the recurring analyses. Therefore, the true structural modes were supposed to remain basically unchanged among the SSI-cov analyses conducted for different

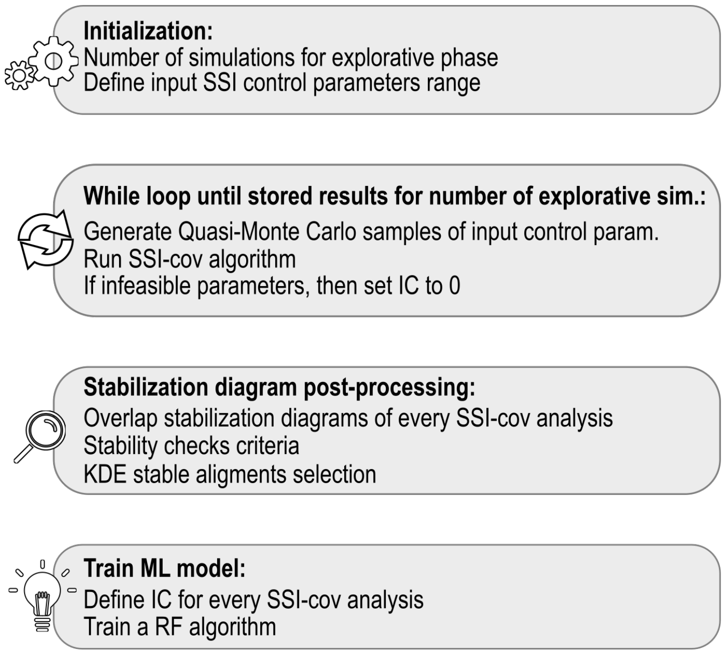


**Figure 1.** Flowchart of first step of intelligent AOMA.

small time windows of the same complete acquisition, centered at different time instants. Finally, the results post-processing of the simulated Monte Carlo-based stabilization diagrams permitted accurately discriminating between the recurrent true modal parameters of interest with respect to the spurious ones. Despite the potentialities, the method proposed in [13, 14] evidences two main limitations. The foremost is related to the arbitrary number of Monte Carlo simulations, suggested equal to 100 without any apparent reason or motivation due to any convergence criterion. The second limitation is related to the crucial choice of the time shift, still considered an arbitrarily user-defined parameter. A poor choice of time shift may strongly affect the reliability of the proposed AOMA framework. This also restricts the automatizing level of the AOMA, since the user must still perform a prior proper tuning of the critical SSI-cov input parameters. Furthermore, despite the final set of stabilization diagrams, another limitation is the lack of any uncertainty evaluation of the Monte Carlo-based results.

Therefore, in the current study, the authors aim to overcome all the above-mentioned issues by providing an intelligent-driven automatic OMA (denoted as intelligent-AOMA) to accomplish the basic requirements of the state-of-art AOMA approaches, and without relying on the commonly used ML-based clustering method. Inspired by the first attempts presented in [13, 14], the proposed intelligent-AOMA increased the level of automatization considering as well the time shift parameter into the Monte Carlo parameters' sampling, thus avoiding any careful pre-tuning user intervention. The proposed intelligent-AOMA has been conceptualized as a two main steps procedure. The first step illustrated in Figure 1 is denoted as the Monte Carlo exploration phase, whereas the second step depicted in Figure 2 is denoted as the intelligently-driven Monte Carlo phase.

With the proposed framework, the authors determined that the SSI-cov control parameters of main interest are the time shift, the time window length of a part of the entire signal, the time target in which symmetrically centering the time window, and the maximum model order. As depicted in Figure 1, in the first step of the proposed intelligent-AOMA, the algorithm may provide a quasi-Monte Carlo sampling [15] of these four governing parameters and evaluate the SSI-cov results. Since this first step is conceived as an exploration phase, the user is demanded

Journal of Physics: Conference Series **2647** (2024) 212005 doi:10.1088/1742-6596/2647/21/212005

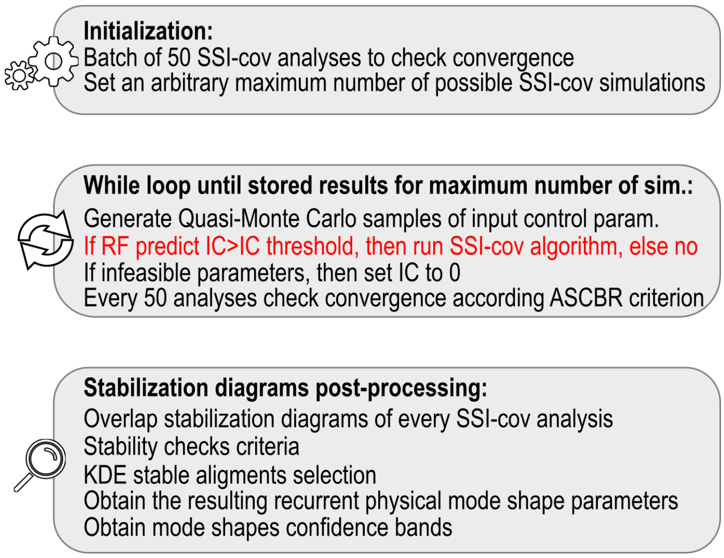


**Figure 2.** Flowchart of first step of intelligent AOMA.

to broadly define only the admissible range of the input governing parameters, which may be retrieved by suggestions provided in [11, 13]. The number of exploratory Monte Carlo simulation results may determine the quality of the next ML-based phase, therefore the authors adopted 100 simulations according to suggestions provided in [13]. The SSI-cov analysis is thus computed for a number of randomly sampled sets of input parameters. However, some unfeasible sets may happen, due to a wrong combination of time shift and model order parameters [6], or due to excessively long computational time (set in this case as 30 seconds). The code is therefore able to exclude those unfeasible sets and smartly invest computational resources in SSI-cov for feasible sets only. The exploration loop is stopped when reaching a number of useful results equal to the user's defined 100 quasi-Monte Carlo SSI-cov exploratory simulations. Stability check criteria retain only stable poles for each of the 100 generated stabilization diagrams. The resulting stabilization diagrams are finally overlapped in order to evaluate which stable poles' alignments are recurrent among the exploratory simulations and which aren't, according to [13]. To avoid traditional and computationally expensive clustering algorithms, the authors post-processed the overlapped stabilization diagram with the kernel density estimation (KDE) algorithm [16], providing a normalized KDE graph which exhibits highly sharp peaks only in natural frequencies associated with recurrent physical stable poles alignments. In summary, based on the normalized KDE, only those poles falling around the KDE graph's peaks within an automatic-defined retaining frequency band are selected. The selction of the peaks of interest in the normalized KDE graph was based on an automatic statistics-based criterion.

As illustrated in Figure 2, the second step of the proposed methodology relies on an ML-based method to intelligently drive the quasi-Monte Carlo sampling process toward the best promising sets of input parameters. The authors adopted a random forest (RF) model [17, 18] to predict if a new sampled set of input parameters is likely to produce useful results or not. To train the RF algorithm, the results of the first exploratory step have been post-processed to deliver an information content (IC) parameter associated with every Monte Carlo simulation. Specifically, the IC has been calculated as the ratio between the number of poles falling within the KDE-based frequency retaining bands and the total number of stable poles produced by every Monte Carlo SSI-cov simulation. For those unfeasible sets of parameters, the IC was set to zero. The RF algorithm was thus trained according to an information content threshold of 10%, attempting to

Journal of Physics: Conference Series **2647** (2024) 212005 doi:10.1088/1742-6596/2647/21/212005

exclude those sets of newly quasi-Monte Carlo sampled parameters which are likely to produce almost useless results with a waste of computational resources. The convergence analysis of the mode shape results is evaluated every 50 iterations according to the acceptable shifting convergence band rule (ASCBR) [19]. According to the generalized sample variance matrix, [20] of the mode shapes associated with the found natural frequency recurrent physical modes, the convergence criteria was set to a relative variation of the trace of the generalized sample variance matrix within $\pm 0.02$ for 50 iterations [19]. Once the convergence is reached, the intelligent-drive RF SSI-cov loop is stopped and some post-processing procedures are performed, similarly to the ones occurring in the first exploration step of the intelligent-AOMA method. Furthermore, since every cluster of stable poles' alignments contains poles coming from various choices of input parameters, the authors analyzed the confidence bands for epistemic uncertainties of the modal parameters, i.e. for natural frequencies, for damping ratios, and even for mode shapes.

## 3. Preliminary results on a benchmark problem

To validate the effectiveness of the proposed intelligent-AOMA methodology, the authors herein provided some preliminary results based on a literature benchmark provided in [5]. Specifically, the benchmark problem is related to a 5 DOF shear type system with lumped mass at each floor, whose implementation details are reported in [5]. This benchmark structure is excited by white noise simulating ambient vibration during operational conditions. The vibration response acceleration signals were monitored on each floor for an hour with a sampling frequency of 100 Hz. The theoretical expected natural frequencies and mode shapes are provided in Figure 3. Prior to the analysis with intelligent-AOMA, since all the natural frequencies are expected to fall in a frequency range less than 10Hz, the data were decimated with a decimation factor of 5 [5].

## 4. Results and discussion

Figure 4 illustrates the main results provided by the first exploratory step of the proposed intelligent-AOMA method retracing the entire procedure described in the previous sections. The exploratory quasi-Monte Carlo simulations delivered a quite populated overlapped stabilization diagram among the 100 SSI-cov analyses. However, the proposed KDE-based method showed an interesting alternative method with respect to standard and computationally more expensive ML-based clustering algorithms, able to reliably automatically detect those recurrent physical stable poles' alignments. The retaining frequency bands around the peaks were calibrated according to the automatically defined bandwidth defined by the KDE algorithm itself [16]. At the end of the first step of the proposed method, the IC was determined for all 100 exploratory simulations.

Considering the unfeasible sets of parameters as well, the entire IC dataset was employed to effectively train a predictive RF algorithm. Subsequently, in the second step of the proposed method, the RF permitted intelligently driving quasi-Monte Carlo sampling of input parameters attempting to avoid less informative sets and efficiently investing computational resources in the most promising parameters' sets. Thus after collecting 600 more actually informative SSI-cov

$$f_n = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} = \begin{bmatrix} 0.88995 \\ 2.59776 \\ 4.09511 \\ 5.2607 \\ 6.0001 \end{bmatrix} \text{[Hz]} \qquad \mathbf{\Phi} = [\phi_1 \ \phi_2 \ \phi_3 \ \phi_4 \ \phi_5] = \begin{bmatrix} 0.28463 & -0.763521 & 1 & 0.918986 & -0.5462 \\ 0.5462 & -1 & 0.28463 & -0.763521 & 0.918986 \\ 0.763521 & -0.5462 & -0.918986 & -0.28463 & -1 \\ 0.918986 & 0.28463 & -0.5462 & 10.763521 & \\ 1 & 0.918986 & 0.763521 & -0.5462 & -0.28463 \end{bmatrix}$$

**Figure 3.** Theoretical natural frequencies and mode shapes of the benchmark problem.

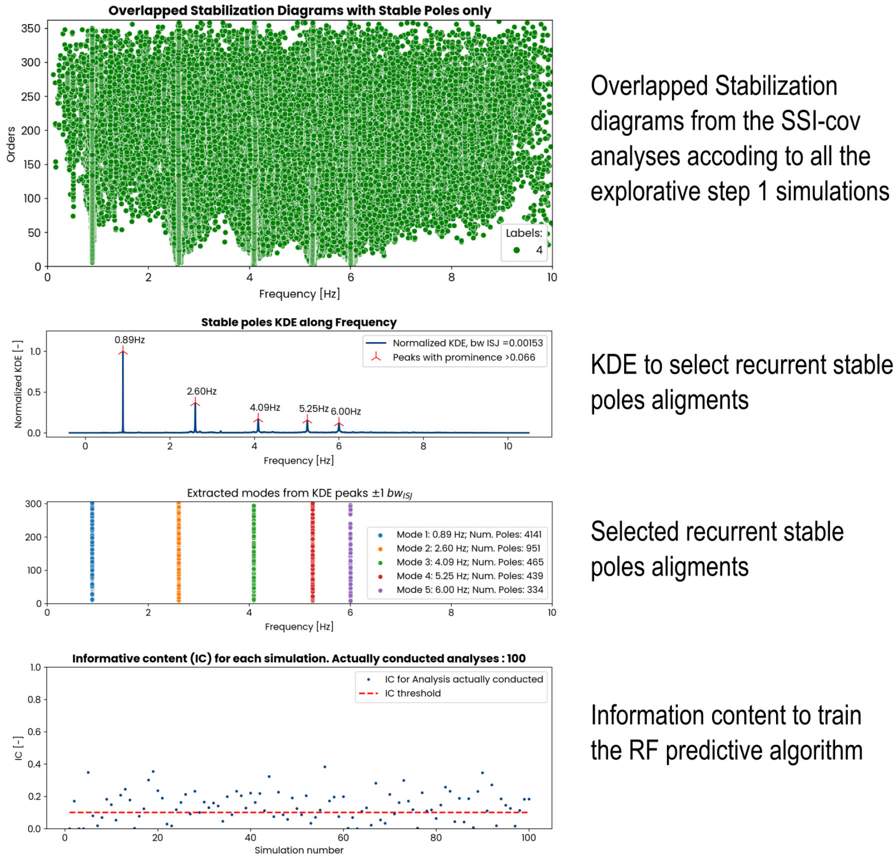


**Figure 4.** Benchmark results of step 1 of the proposed intelligent-AOMA.

results, the algorithm reached the convergence and exit from the quasi-Monte Carlo sampling loop. In total 2105 quasi-random samples of SSI-cov input parameters were considered, however only 600 were effectively computed with SSI-cov algorithms, demonstrating the efficiency of the proposed ML-based methodology. Therefore, considering the 700 effective results from the beginning, the overlapped stabilization diagram was post-processed with KDE to identify the final modal properties, as shown in Figure 5. Specifically, it was possible to demonstrate the great agreement with the natural frequencies of the five modes of interest with the expected theoretical results of Figure 3. Moreover, the intelligent-AOMA results provided uncertainties associated with the modal properties as shown in Table 1. In the Table, the symbol $\mu$ indicates the mean values of natural frequency $f_n$ and damping ratio $\xi_n$ for every mode, whereas $\sigma$ indicates their respective standard deviation. Despite the uncertainties related to the natural frequencies and damping ratios being very limited according to the Table, thus representing almost deterministic results, it is worth noting that the final mode shapes confidence bands around their respective average mode shape present a higher level of epistemic uncertainties, especially in the mode corresponding to 5.253 Hz. Therefore, from these results, it is evident how the various choice of parameters may provide quite uncertain results which propagate especially in the final mode shape of interest, and thus an automatic approach likewise the present one may avoid an arbitrary poor choice with deleterious effects.

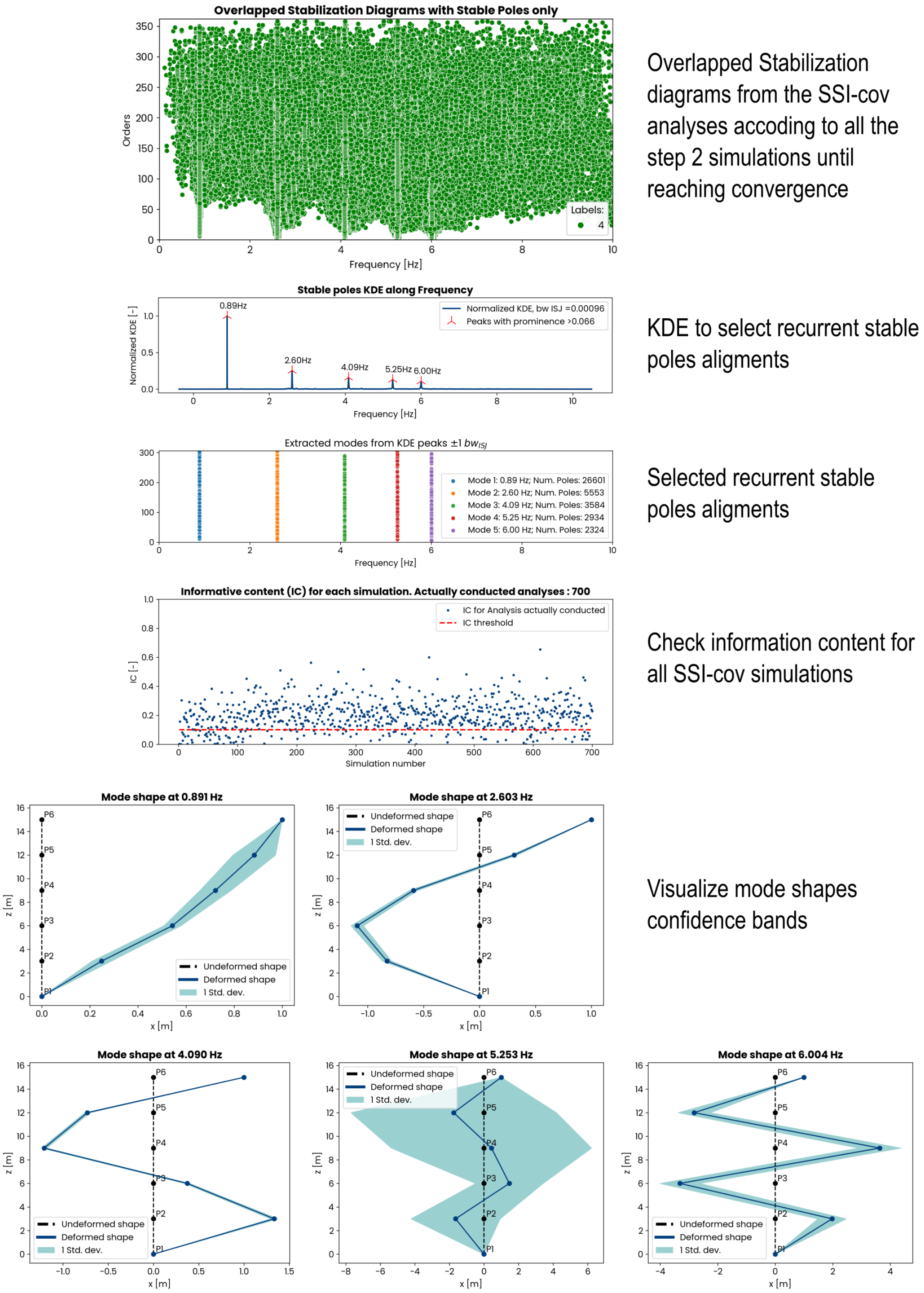


**Figure 5.** Benchmark results of step 2 of the proposed intelligent-AOMA.

## 5. Conclusions

In the current study, the authors presented an innovative machine learning-based automatic operational modal analysis framework conceived as a two main steps procedure attempting to overcome the main limitations of the commonly widespread four steps automatic operational

**Table 1.** Benchmark frequency and damping results.

| Theoretical | | intelligent-AOMA | | | |
|---|---|---|---|---|---|
| $f_n$ [Hz] | $\xi_n$ [%] | $\mu_{f_n}$ [Hz] | $\sigma_{f_n}$ [Hz] | $\mu_{\xi_n}$ [%] | $\sigma_{\xi_n}$ [%] |
| 0.8900 | 2.000 | 0.8906 | 0.00049 | 1.970 | 0.160 |
| 2.5978 | 2.000 | 2.6027 | 0.00064 | 1.990 | 0.213 |
| 4.0951 | 2.000 | 4.0897 | 0.00062 | 2.120 | 0.227 |
| 5.2607 | 2.000 | 5.2528 | 0.00062 | 2.131 | 0.237 |
| 6.0001 | 2.000 | 6.0040 | 0.00062 | 1.950 | 0.234 |

modal analyses framework in the existing literature. Inspired by [13, 14], the authors developed the currently proposed method based on quasi-Monte Carlo sampling of four critical parameters of the stochastic subspace identification algorithm. To validate the current method, the authors provide some preliminary insights for a literature numerical benchmark problem [5]. The main novelties of the current study may be summarized as follows:

- A first exploratory step of quasi-Monte Carlo stochastic subspace identification analyses is preparatory to train a machine learning random forest algorithm to intelligently guide the quasi-Monte Carlo parameters sampling in the second step;
- Stable poles' clusters on stabilization diagrams are selected throughout a KDE method in place of commonly widespread and more computationally expensive clustering techniques;
- The computational resources are smartly invested intelligently preferring those set of parameters expected to provide informative results with less computational efforts, and vice versa discarding the ones expected to provide useless results;
- The overall analysis of the overlapped stabilization diagram permitted investigation of the uncertainty propagation of the input parameters on the modal properties.

## Acknowledgments

This research has been supported by the ADDOPTML project: "ADDitively Manufactured OPTimized Structures by means of Machine Learning" (No: 101007595) belonging to the Marie Skłodowska-Curie Actions (MSCA) Research and Innovation Staff Exchange (RISE) H2020-MSCA-RISE-2020. Their support is highly acknowledged.

## References

[1] Aloisio A, Pasca D P, Di Battista L, Rosso M M, Cucuzza R, Marano G C and Alaggio R 2022 *Structures* vol 37 (Elsevier) pp 686–697
[2] Farrar C R and Worden K 2012 *Structural health monitoring: a machine learning perspective* (John Wiley & Sons)
[3] Marasco G, Rosso M M, Aiello S, Aloisio A, Cirrincione G, Chiaia B and Marano G C 2022 *International Conference on Engineering Applications of Neural Networks* (Springer) pp 165–176
[4] Melchiorre J, Manuello Bertetto A, Rosso M M and Marano G C 2023 *Sensors* **23** 693
[5] Pasca D P, Aloisio A, Rosso M M and Sotiropoulos S 2022 *SoftwareX* **20** 101216
[6] Rainieri C and Fabbrocino G 2014 *Springer, New York* **142** 143
[7] Cardoso R, Cury A and Barbosa F 2017 *Mechanical Systems and Signal Processing* **95** 24–41 cited By 44
[8] Reynders E, Houbrechts J and De Roeck G 2012 *Mechanical Systems and Signal Processing* **29** 228–250 cited By 311
[9] Alaggio R, Aloisio A, Antonacci E and Cirella R 2021 *Construction and Building Materials* **268** 121069
[10] Ubertini F, Gentile C and Materazzi A 2013 *Engineering Structures* **46** 264–278 cited By 190
[11] Rainieri C and Fabbrocino G 2014 *International Journal of Lifecycle Performance Engineering* **1** 317–334 cited By 34
[12] Zini G, Betti M and Bartoli G 2022 *Mechanical Systems and Signal Processing* **164** 108173
[13] Zhou K and Li Q S 2022 *Journal of Building Engineering* **52** 104373

[14] Zhou K, Li Q S and Han X L 2022 *Journal of Structural Engineering* **148** 04022066
[15] Owen A B 2017 *arXiv preprint arXiv:1706.02808*
[16] Gramacki A 2018 *Nonparametric kernel density estimation and its computational aspects* vol 37 (Springer)
[17] Raschka S and Mirjalili V 2019 *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2* (Packt Publishing Ltd)
[18] Géron A 2022 *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (" O'Reilly Media, Inc.")
[19] Ata M Y 2007 *Simulation Modelling Practice and Theory* **15** 237–246
[20] Ziegel E R 2003 *Technometrics* **45** 109