RESEARCH ARTICLE

Probabilistic investigation of error propagation in frequency domain decomposition-based operational modal analysis

Çağlayan Hızal [ORCID] | Engin Aktaş [ORCID]

Department of Civil Engineering, Izmir Institute of Technology, Urla, Turkey

Correspondence
Çağlayan Hızal, Department of Civil Engineering, Izmir Institute of Technology, Urla, Izmir, Turkey.
Email: caglayanhizal@iyte.edu.tr

Summary

Each operational modal analysis (OMA) technique may produce significant errors during the identification procedure due to the applied methodology, environmental/operational conditions, and instrumentation. Consequently, those errors can adversely affect the quality of identified parameters. In this context, this study aims at providing a comprehensive discussion on the propagation of predictions errors in the frequency domain OMA. To mitigate the prediction errors those considered to be induced by modeling and measurement errors, an extended formulation is presented based on a recently developed Modified Frequency and Spatial Domain Decomposition technique. A comprehensive investigation is presented for the probabilistic modeling of output power spectral density (PSD), considering prediction errors. Numerical and real data applications are conducted to show the effectiveness of the proposed methodology.

KEYWORDS

frequency domain identification, measurement noise, modeling errors, OMA, statistical errors

1 | INTRODUCTION

Output only modal identification techniques have an increasing popularity in the structural health monitoring (SHM) community since it efficiently works under ambient operational conditions without need of any artificial loading. In this context, the early-term implementations of output only techniques have generally dealt with the modal identification problem in the time domain. Natural Excitation and Eigensystem Realization Algorithm (NexT-ERA),[1] Ibrahim time domain identification,[2] or stochastic subspace identification (SSI)[3] can be count one of the best knowns of these time domain techniques. Among these, SSI techniques, which fundamentally undertake a stochastic process on the state space representation of the output response, are widely implemented since they present fast and efficient procedures. Various implementations of SSI techniques based on the forming of the Hankel matrix directly from the collected data (SSI-Data) or its covariance (SSI-Cov) are also available in the literature.[3–10] In the past two decades, more effective techniques have been developed based on Bayesian theory.[11,12] Bayesian time domain methods might be more successful in comparison to NExT-ERA and/or SSI techniques especially in case of noisy data. The computational cost, however, arises an important challenge for such kinds of methods.

In general sense, frequency-domain output only techniques can be considered more practical and effective compared to time-domain methods, due to the fact that they can focus on narrow bands in which a limited number

of modes are excited. From this aspect, probabilistic techniques such as, Bayesian Spectral Density Approach (BSDA)[13] and Bayesian Fast Fourier Transform Approach (BFFTA)[14,15] might be much more effective in comparison to conventional frequency domain techniques. All these methods estimate the modal parameters from a probabilistic model framed on either the fast Fourier transform (FFT) or power spectral density (PSD) matrix of the response data. During this procedure uncertainties stem from the modeling errors and/or measurement noise are also considered. Using this posterior uncertainty information, Bayesian operational modal analysis (BAYOMA) = techniques can also be integrated to two-stage finite element model updating and/or damage detection procedures.[16–20] Different variants of BAYOMA methods are available in the literature such as Bayesian Spectral Trace Approach (BSTA),[21] extended BFFTA for asynchronous vibration data[22] or Hierarchical Bayesian modal identification.[23] Despite their superiority, the most important disadvantage of Bayesian methods is that the computational effort can be significantly enhanced due to the increasing complexity under large uncertainties (e.g., weakly excited modes, large modeling error and channel noise). Therefore, some simplified versions of Bayesian methods are also available in literature especially for uncertainty quantification using the large signal-to-noise ratio (*snr*) asymptotic behavior of the response data.[24,25]

As one of the most classical frequency domain approaches, frequency domain decomposition (FDD) technique presents a fast and efficient tool for output only modal identification and it is widely referred to by the researchers during the past two decades.[26–28] Although the classical FDD method shows successful performance for natural frequency and mode shape vector identification, not achieving a similar success in modal damping estimation is an important issue. This issue has been partially addressed by a modified technique, called enhanced FDD (EFDD).[29] More complex and effective methods such as refined FDD (rFDD) are also available in the literature to surpass the limitations in the classical FDD implementation.[30–32] In rFDD, a special attention is given for the heavy modal damping and earthquake input motion problem which are solved by the implementation of an iterative procedure using various signal processing tools such as Welch's modified periodogram and Wiener Khinchin theorem.[33–36]

Frequency and spatial domain decomposition (FSDD) technique appears as one of the most successful modifications of the classical FDD approaches for more sensitive modal damping estimation.[37] Similar to the classical FDD, modal shape vector is estimated by singular value decomposition of the output PSD, at the peak response frequency. Then, an equivalent single degree of freedom (SDOF) response is filtered by pre and post multiplying the output PSD by the modal shape vector. Finally, the modal frequencies and damping ratios are determined by the solution of a least squares equation.

Motivated from the Bayesian approaches, a modified version of FSDD, which is named as modified frequency and spatial domain decomposition (MFSDD), has been recently developed.[38] Different from classical FDD techniques, the theoretical output PSD matrix is revisited from a probabilistic perspective, by inclusion of a prediction error term. Similar probabilistic approaches are also available in the literature which apply a maximum likelihood estimation on the frequency response functions (FRF) to extract the most probable modal parameters.[39–42] However, the major difference of MFSDD from the other maximum likelihood-based approaches is that the probabilistic expectation of the output PSD matrix is modeled using central complex Wishart distribution.

Assuming that the *snr* value is significantly large within a narrow band, the modal shape vector and prediction error term can be obtained in the closed form, similar to BFFTA. This assumption, however, may cause an additional modeling error, which may adversely influence the parameter identification quality. To see the propagation of the prediction (or identification) errors in MFSDD, an extended version is developed in this study. An extensive numerical example based on the channel noise, heavy modal damping, and earthquake input effects is conducted. Laboratory and real data examples which comprise large noise effects and closely spaced modes are also presented to see the performance of the extended MFSDD.

2 | THEORETICAL BACKGROUND OF FDD

In FDD, the theoretical value of the output PSD is estimated as a composition of the input PSD and the transfer function in the pole-residue form, which defines the relation between input force and output response. More specifically, the theoretical spectral density matrix of a structure, whose N number of degrees of freedom (DOF) response are measured, is given by[26]

$$\mathbf{G}_{yy}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{G}_{xx}(\omega_k)\mathbf{H}^*(\omega_k) \tag{1}$$

where ω_k = discrete excitation frequency at kth step, $\mathbf{G}_{yy}(\omega_k) = N \times N$ sized output PSD matrix, $\mathbf{G}_{xx}(\omega_k)$ = input PSD matrix, and $\mathbf{H}(\omega_k)$ represents the transfer function which is defined by

$$\mathbf{H}(\omega_k) = \sum_{n=1}^{N_m} \left[\frac{\mathbf{R}_n}{i\omega_k - \lambda_n} + \frac{\bar{\mathbf{R}}_n}{i\omega_k - \lambda_n^*} \right] \tag{2}$$

where $\mathbf{H}^*(\omega_k)$ stands for the complex conjugate and transpose of $\mathbf{H}(\omega_k)$, and $\bar{\mathbf{R}}_n$ represents the conjugate of $\mathbf{R}_n$. Here, both $\mathbf{G}_{xx}(\omega_k)$ and $\mathbf{H}(\omega_k)$ would be $N \times N$ sized matrices when all DOFs are considered as reference [26, 30]. In Equation (2), N_m = number of considered modes, $\mathbf{R}_n = \boldsymbol{\varphi}_n \boldsymbol{\Gamma}_n^*$, $\lambda_n = -\sigma_n + \omega_{dn}$, $\sigma_n = \xi_n \omega_n$, $\omega_{dn} = \omega_n \sqrt{(1-\xi_n^2)}$, and ω_n, ξ_n represent the nth mode natural frequency and modal damping ratio, respectively. In addition, $\boldsymbol{\varphi}_n = N \times 1$ sized modal shape vector and $\boldsymbol{\Gamma}_n = 1 \times N$ sized modal participation vector.

Under operational conditions such that the structure is excited by unknown and small amplitude input excitations, it can be assumed that these excitation forces are randomly distributed and uncorrelated variables, resulting in a stationary Gaussian process. Based on these assumptions, the input PSD can be modeled as a diagonal matrix such as $\mathbf{G}_{xx}(\omega_k) = G_{xx}\mathbf{I}_N$. In this modeling assumption, G_{xx} and $\mathbf{I}_N$ represent the spectral density of input force (corresponds to a scalar), and $N \times N$ sized identity matrix, respectively. Thus, the theoretical PSD matrix can be written as[37]

$$\mathbf{G}_{yy}(\omega_k) = \sum_{m=1}^{N_m} \sum_{n=1}^{N_m} \left[\left(\frac{\bar{\mathbf{R}}_m}{-i\omega_k - \lambda_m} + \frac{\mathbf{R}_m}{-i\omega_k - \lambda_m^*} \right) G_{xx} \left(\frac{\mathbf{R}_n^T}{i\omega_k - \lambda_n} + \frac{\mathbf{R}_n^*}{i\omega_k - \lambda_n^*} \right) \right] \tag{3}$$

where $\mathbf{R}_n^T$ represents the transpose of $\mathbf{R}_n$. After some arrangements and mathematical manipulations, the output PSD can be written as below.[30]

$$\mathbf{G}_{yy}(\omega_k) = \sum_{n=1}^{N_m} \left(\frac{\mathbf{A}_n}{i\omega_k - \lambda_n} - \frac{\mathbf{A}_n^*}{i\omega_k + \lambda_n^*} + \frac{\bar{\mathbf{A}}_n}{i\omega_k - \lambda_n^*} - \frac{\mathbf{A}_n^T}{i\omega_k + \lambda_n} \right) \tag{4}$$

Assuming small damping ratio (about 1%–2%), $\mathbf{A}_n$ can be approximately defined as[30]

$$\mathbf{A}_n = \sum_{n=1}^{N_m} \left[\frac{\mathbf{R}_n}{2(\xi_n \omega_n - i\omega_n)} + \frac{\bar{\mathbf{R}}_n}{2\xi_n \omega_n} \right] G_{xx} \mathbf{R}_n^T \cong \boldsymbol{\varphi}_n d_n \boldsymbol{\varphi}_n^* \tag{5}$$

where $d_n = (\boldsymbol{\Gamma}_n G_{xx} \boldsymbol{\Gamma}_n^*)/(2\sigma_n)$, corresponds to a real scalar. Substituting Equation (5) into (4) and after some arrangements, the output PSD matrix can be written as

$$\mathbf{G}_{yy}(\omega_k) = \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n \mathrm{Re}\left(\frac{2d_n}{i\omega_k - \lambda_n} \right) \boldsymbol{\varphi}_n^* \tag{6}$$

where "Re(.)" indicates the real part of a complex scalar.

3 | MODELING OF OUTPUT PSD CONSIDERING PREDICTION ERROR

Operational conditions (e.g., stationary input, adequately large measurement duration and sampling frequency) and considered dynamical properties in the theoretical formulation of FDD approaches have a crucial importance on the modal parameter estimation quality. In addition to these, measurement noise induced by the instrumented sensors

and/or data acquisition systems may cause a biased error during the analysis procedure. Considering all of these aspects, a statistical expected value can be defined for the theoretical output PSD as follows.[38]

$$E\left[\mathbf{G}_{yy}(\omega_k)\right] = \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n \mathrm{Re}\left(\frac{2d_n}{i\omega_k - \lambda_n}\right)\boldsymbol{\varphi}_n^* + \sum_{n=1}^{N_m} \delta_n \mathbf{I}_N + \varepsilon_m \mathbf{I}_N \tag{7}$$

where $E[\mathbf{G}_{yy}(\omega_k)]$ represents the statistical expectation of $\mathbf{G}_{yy}(\omega_k)$. In addition, δ_n and ε_m represent the modeling and measurement errors, respectively. Here, some leading reasons why such kinds of prediction errors arise can be classified as below.

Measurement errors:

- Channel noise
- Distortions in the sensor orientations
- Errors in the data transmission cables and/or data acquisition device

Modeling errors:

- Large damping ratio
- Non-stationary input data
- Non-orthogonal mode shapes

Identifying the measurement and modeling errors separately appears as a major challenge in OMA. Some probabilistic methods[21,38,43] are capable of providing a prediction error for the identification process. However, all these methods are more successful in extracting measurement errors rather than the modeling assumptions based one. Therefore, it might be more reasonable to define a prediction error term that covers both measurement and modeling errors, as follows.

$$\Delta_e = \sum_{n=1}^{N_m} \delta_n + \varepsilon_m \tag{8}$$

Substituting Equation (8) into (7) and after some mathematical manipulations yields:

$$E\left[\mathbf{G}_{yy}(\omega_k)\right] = \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n (c_n \alpha(\omega_k, \omega_n, \xi_n) + \Delta_e)\boldsymbol{\varphi}_n^* + \sum_{j=N_m+1}^{N} \Delta_e \mathbf{u}_j \mathbf{u}_j^* \tag{9}$$

where $c_n = \boldsymbol{\Gamma}_n G_{xx} \boldsymbol{\Gamma}_n^*$, $\alpha(\omega_k, \omega_n, \xi_n) = \mathrm{Re}(1/\sigma_n(i\omega_k - \lambda_n))$, and $\mathbf{U}\mathbf{U}^* = \mathbf{I}_N$. In addition, $\mathbf{U} = [\mathbf{u}_1, ..., \mathbf{u}_N]$ represents a unitary matrix that spans N number of orthogonal vectors. Consequently, the modal parameters including ω_n, ξ_n and $\boldsymbol{\varphi}_n$ as well as the prediction error, Δ_e, are estimated based on Equation (9). To show the relation between the channel noise, modeling errors and the quality of identified modal parameters, two similar methodologies are employed in this study, namely MFSDD and extended MFSDD, respectively. The first one has been recently introduced based on a modified theory of the classical FSDD technique, using a maximum likelihood estimation-based methodology. The latter one is an extended version of MFSDD which is developed to consider the effect of considerably large prediction errors.

4 | MFSDD METHODOLOGY FOR FREQUENCY DOMAIN OMA

The underlying theory of the MFSDD approach, which is developed by motivating from frequency domain BAYOMA techniques, are discussed in detail in this section. The major difference of MSDD from BAYOMA applications is that it is derived based on the classical frequency domain decomposition theory. Moreover, MFSDD does not require any posterior probability inference and uncertainty quantification in contrast to Bayesian approaches.

4.1 | Derivation of MFSDD

According to MFSDD, the expected output PSD is obtained by maximizing the following probability density function (pdf).[38]

$$p\left(\mathbf{G}_{yy}(\omega_k)\big|c_n,\omega_n,\xi_n,\Delta_e,\boldsymbol{\varphi}_n\right) = \prod_k \left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right|^{-1}\exp\left\{-tr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\mathbf{G}_{yy}(\omega_k)\right)\right\} \tag{10}$$

where "$tr(.)$" represents the trace of a matrix. The given pdf represents a central complex Wishart distribution with M degrees of freedom which must provide the condition such that $M > N$. To do so, FFTs of measured data can be separated into N_t/M non-overlapping segments, each spanning M number of independent sample PSDs. By making use of this, the set of averaged output PSDs are defined as $\left\{\bar{\mathbf{G}}_{yy}^s(\omega_{\bar{k}})\right\}_{s=1}^{N_t/M}$. Here, $N_t =$ the number of output PSD samples. In addition, $\bar{\mathbf{G}}_{yy}^s(\omega_{\bar{k}})$ and $\omega_{\bar{k}}$ represent the sample average of each segment and the average discrete frequency, which are defined by

$$\bar{\mathbf{G}}_{yy}^s(\omega_{\bar{k}}) = \frac{1}{M}\sum_{r=k}^{M+k}\mathbf{G}_{yy}(\omega_r); \quad \omega_{\bar{k}} = \frac{M(\bar{k}-1)}{\Delta t N_t}; \quad (\bar{k}=1,...,N_t/M) \tag{11}$$

where $\Delta t =$ sampling time. As the measurement duration increases ($N_t \rightarrow \infty$), the theoretical expected value of $\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)$ can be assumed as equivalent to that of $\mathbf{G}_{yy}(\omega_k)$. Consequently, $M\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)$ follows the following probability distribution.[13]

$$p\left(M\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)\right) = \kappa_0\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right|^{-M}\exp\left\{-Mtr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)\right)\right\} \tag{12}$$

where, $\kappa_0 = M^{N(M-N)}\left|\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)\right|^{(M-N)}\pi^{N(N-1)/2}/\prod_{j=1}^N(M-j)!$, represents a scalar for the normalization of the given pdf. Substituting Equation (11) into (12) yields

$$p\left(M\bar{\mathbf{G}}_{yy}^s(\bar{\omega}_k)\right) = \kappa_0\prod_{r=1}^M\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right|^{-1}\exp\left\{-tr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\mathbf{G}_{yy}(\omega_k)\right)\right\} \tag{13}$$

Since each segment is assumed as independent variables set, the expectations of discrete output PSD can be assembled as below.

$$p\left(\{\mathbf{G}_{yy}(\omega_k)\}_{N_t}\right) = \kappa_0\prod_{\bar{k}}\prod_{r=1}^M\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right|^{-1}\exp\left\{-tr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\mathbf{G}_{yy}(\omega_k)\right)\right\}$$
$$= \kappa_0\prod_k\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right|^{-1}\exp\left\{-tr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\mathbf{G}_{yy}(\omega_k)\right)\right\} \tag{14}$$

which is equivalent of Equation (10).

Selection of M and its influence on the accuracy of the assumed pdf have been widely investigated in the literature as regarding the theory of Bayesian PSD approach.[13,21,44] In general sense, choosing $M = 1$ does not affect the shape of the accepted pdf which is numerically equivalent to that obtained by Bayesian FFT method. On the other hand, a selection such as $M < N$ results in mathematically undefined pdf, since the scaling factor becomes an undefined number. This arises as an important issue in Bayesian analysis in which the modeling assumptions for the accepted pdf should be strictly obeyed. In maximum likelihood estimation, however, making such a selection might be acceptable in practice due to the fact that only the maximum likelihoods in the pdf are of interest.

In the original form of Bayesian PSD approach, it is assumed that there are M (>N) number of independent measurement sets, which seems not practical at all, for real life applications. From this aspect, as a more realistic situation, dividing a single set measurement into multiple segments might be much more useful. In such a case, however, the quality of identified modal parameters are negatively influenced by the larger values of M due to spectral leakage.[44] The

derivations presented above, however, mathematically show that segmenting the FFT ordinates instead of time-domain response gives the same solution with $M = 1$ case. Thus, it is unnecessary to divide time domain response data in the analysis procedure. This derivation can also produce modeling assumptions when N_t/M is not large enough. In this case, however, only κ_0 is affected by the selected M size, which is also irrelevant to the maximum likelihood estimation.

4.2 | Computational procedure

During the analysis process, employing a negative-logarithm likelihood function might be much more convenient instead of using Equation (10) directly. To do so, the negative logarithm likelihood function of Equation (10) is obtained as

$$
\begin{aligned}
L(c_n, \omega_n, \xi_n, \Delta_e, \boldsymbol{\varphi}_n) &= -\ln\left[p\left(\mathbf{G}_{yy}(\omega_k)\middle|c_n, \omega_n, \xi_n, \Delta_e, \boldsymbol{\varphi}_n\right)\right] \\
&= \ln\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right| + tr\left(E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1}\mathbf{G}_{yy}(\omega_k)\right)
\end{aligned}
\tag{15}
$$

Thus, most probable modal parameters can be estimated by gradient based minimization of Equation (15). Some important issues and computational difficulties such as calculation of inverse and determinant of $E[\mathbf{G}_{yy}(\omega_k)]$, however, might arise during this minimization process. In order to overcome those issues, the inverse and the determinant of $E[\mathbf{G}_{yy}(\omega_k)]$ can be analytically derived.[43,45] During this process, in a narrow resonant frequency band, MFSDD employs an assumption that $c_n\alpha(\omega_k, \omega_n, \xi_n) \gg \Delta_e$, which makes it possible to obtain a closed form solution for c_n, Δ_e and $\boldsymbol{\varphi}_n$.[38] This assumption, however, can produce an additional modeling error whose normalized value can be approximately represented by

$$
e_r = \frac{\Delta_e}{c_n\alpha(\omega_k, \omega_n, \xi_n)} \times 100(\%)
\tag{16}
$$

Here, one can be concluded the normalized error given by Equation (16) also corresponds to the modal noise-to-signal ratio (or inverse of modal *snr*).[45] Therefore, it is apparent that this assumption will give better results in case the spectral density level of the signal is considerably larger than the prediction error. In this context, Figure 1 presents the variation of normalized error with respect to the modal signal-to-noise ratio. For better illustration, the presented graph is separated into three regions according to the signal-to-noise ratio levels. In this classification, the first region (*snr* < 10) represents the unreasonable signal range in which the mode of interest is rather weakly excited and/or prediction error term is very large. Furthermore, the second region spans (10 < *snr* < 250) the reasonable range in which the signal-to-noise ratio remains within the acceptable limits. Finally, the third region (*snr* > 250) defines the ideal range that represents a very strong excitation signal and/or very small prediction error.
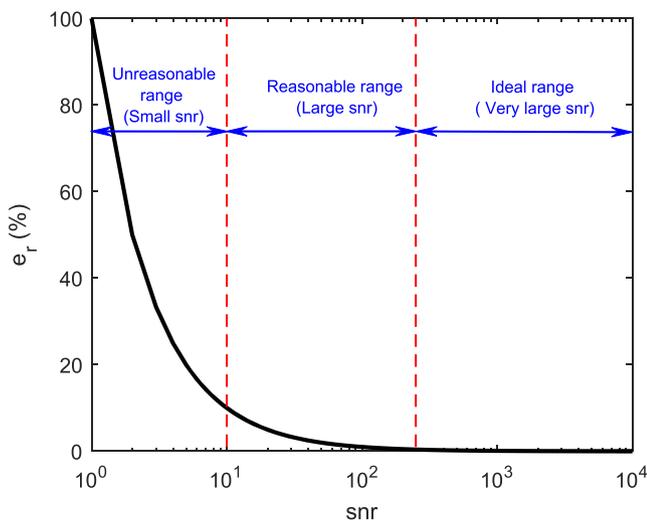

FIGURE 1 Variation of normalized error in expected output power spectral density (PSD) versus signal-to-noise ratio (*snr*)

5 | EXTENDED MFSDD

In the light of recent technological developments, it is possible to use highly sensitive accelerometer sensors and data acquisition systems which are widely produced for SHM applications. Therefore, it may not be possible to experience a *snr* level in unreasonable range unless there is an extreme problem such as cable error, sensor failure and/or large amount of modeling error (e.g., heavy damping, inadequate measurement duration). From this perspective, it should be clearly stated MFSDD provides better results for the reasonable and ideal signal levels, in comparison to classical FDD methods.[38] On the other hand, considering the exact theoretical spectral density matrix may significantly contribute to improve the modal parameter estimation even if it can lead to an increase in computational effort. To do so, an extended formulation is presented here for the defined unreasonable *snr* range.

In the extended MFSDD, similar to the calculations presented by Au[43], the inverse and determinant of the theoretical spectral density matrix are obtained as below.

$$E\left[\mathbf{G}_{yy}(\omega_k)\right]^{-1} = \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n [c_n \alpha(\omega_k, \omega_n, \xi_n) + \Delta_e]^{-1} \boldsymbol{\varphi}_n^* + \Delta_e^{-1} \left[\mathbf{I}_n - \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n \boldsymbol{\varphi}_n^*\right] \tag{17}$$

$$\left|E\left[\mathbf{G}_{yy}(\omega_k)\right]\right| = [c_n \alpha(\omega_k, \omega_n, \xi_n) + \Delta_e]\Delta_e^{(N-N_m)} \tag{18}$$

Substituting Equations (17) and (18) into (15), and after some mathematical manipulations, the following objective function is obtained.

$$J(c_n, \omega_n, \xi_n, \Delta_e, \boldsymbol{\varphi}_n) = N_\omega \ln(N - N_m) + \sum_k \ln[c_n \alpha(\omega_k, \omega_n, \xi_n) + \Delta_e]$$

$$+ \Delta_e^{-1} tr\left(\sum_k \mathbf{G}_{yy}(\omega_k)\right) - \boldsymbol{\varphi}_n^* \left(w_k \sum_k \mathbf{G}_{yy}(\omega_k)\right) \boldsymbol{\varphi}_n + \sum_{n=1}^{N_m} \beta_n \left(\boldsymbol{\varphi}_n^* \boldsymbol{\varphi}_n - 1\right) \tag{19}$$

where N_ω = number of FFT ordinates within the selected resonant band, $\beta = \left[\beta_1, ..., \beta_{N_m}\right]$ represents the set of Lagrange multipliers to satisfy the unitary assumption of theoretical spectral density matrix, and $w_k = \{\Delta_e(1+e_r)\}^{-1}$.

Minimizing of (19) in its present form might be a difficult task since the given objective function is highly non-linear and multi-dimensional. Therefore, a two-stage procedure can be adapted here to obtain a condensed objective function.[43,46] For this purpose, as a first step, the most probable mode shape vector can be obtained as

$$\nabla_{\boldsymbol{\varphi}_n} J = -\boldsymbol{\varphi}_n^* \left(\sum_k w_k \mathbf{G}_{yy}(\omega_k)\right) \boldsymbol{\varphi}_n + \boldsymbol{\varphi}_n^* \beta_n = 0 \Rightarrow \left(\sum_k w_k \mathbf{G}_{yy}(\omega_k)\right) \hat{\boldsymbol{\varphi}}_n = \beta_n \hat{\boldsymbol{\varphi}}_n \tag{20}$$

TABLE 1 Extended MFSDD algorithm

| *Initial Calculations* |
|---|
| *1. Detect the possible modes on SV or PSD spectrum by peak-picking* |
| *2. Select proper bandwidths for the detected possible modes* |
| ***Minimization process*** |
| *1. Determine the initial guess for modal frequency and damping ratio,* |
| *2. Determine initial guess of φ_n as the eigenvector of $\sum_k \mathbf{G}_{yy}(\omega_k)$* |
| *3. Determine initial guesses for c_n and Δ_e by using Equation (23)* |
| *4. Update the MPVs of c_n, ω_n, ξ_n and Δ_e by minimizing Equation (22)* |
| *5. Update the MPV of φ_n as the eigenvector of $\sum_k \hat{w}_k \mathbf{G}_{yy}(\omega_k)$* |
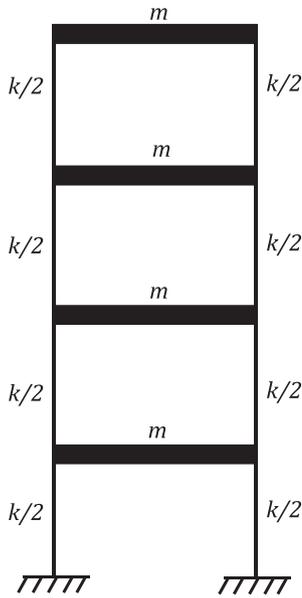
FIGURE 2 Schematic view of four-story shear frame

Interstory stiffness:

$k = 500 \; kN/m$

Story mass:

$m = 1000 \; kg$

Natural frequencies:

$f_1 = 1.236 \; Hz$

$f_2 = 3.559 \; Hz$

$f_3 = 5.452 \; Hz$

$f_4 = 6.688 \; Hz$

where "^" represents the most probable value (MPV). It can be deduced from the here that the most probable mode shape vector depends on the remaining parameters. Here, the MPV of nth mode shape vector can be parametrically obtained by solving the eigenvalue equation in Equation (20), which also indicates that β_n can be represented as a function of c_n, ω_n, ξ_n and Δ_e, at $\varphi_n = \hat{\varphi}_n$.[43,45,46]

$$\beta_n(c_n, \omega_n, \xi_n, \Delta_e) = \hat{\varphi}_n^* \left(\sum_k w_k \mathbf{G}_{yy}(\omega_k) \right) \hat{\varphi}_n \tag{21}$$

At the next step, a condensed objective function can be obtained for c_n, ω_n, ξ_n, and Δ_e at $\varphi_n = \hat{\varphi}_n$, by substituting Equation (21) into (19), as below.

$$J(c_n, \omega_n, \xi_n, \Delta_e) = N_\omega(N - N_m)\ln\Delta_e + \sum_k \ln[c_n \alpha(\omega_k, \omega_n, \xi_n) + \Delta_e] + \Delta_e^{-1} tr\left(\sum_k \mathbf{G}_{yy}(\omega_k) \right)$$
$$- \beta_n(c_n, \omega_n, \xi_n, \Delta_e) \tag{22}$$

Consequently, the most probable mode shape vector can be updated as the nth eigenvector of $\sum_k \hat{w}_k \mathbf{G}_{yy}(\omega_k)$, when eigenvalues are sorted in descending order. A pseudo code is presented in Table 1 to summarize the overall procedure. Here, as an initial guess, the possible modal frequencies can be detected on the SV or PSD spectrum by peak-picking. Then, a proper bandwidth selection is required for the detected modes. These bandwidths can be determined so that it covers the spectral bell curve on the spectrum plot provided that it does not exceed 20% of the modal frequency.[38] In the minimization phase, an initial guess about 1% can be selected for modal damping ratio.[45] If the measured structure possibly has heavy modal damping, a larger initial guess should be selected. In addition, an initial guess can be calculated for the mode shape vector as the eigenvector of $\sum_k \mathbf{G}_{yy}(\omega_k)$.[38,43,45] Finally, the initial guesses for c_n and Δ_e can be calculated by using the following formulae which are derived for MFSDD.[38]

$$c_{n0} = \sum_k \frac{\hat{\varphi}_n^* \mathbf{G}_{yy}(\omega_k)\hat{\varphi}_n}{N_\omega \alpha(\omega_k, \omega_n, \xi_n)}; \quad \hat{\Delta}_{e0} = \frac{\sum_k tr\left(\mathbf{G}_{yy}(\omega_k)\left[\mathbf{I}_N - \sum_{n=1}^{N_m} \varphi_n \varphi_n^* \right] \right)}{N_\omega(N - N_m)} \tag{23}$$
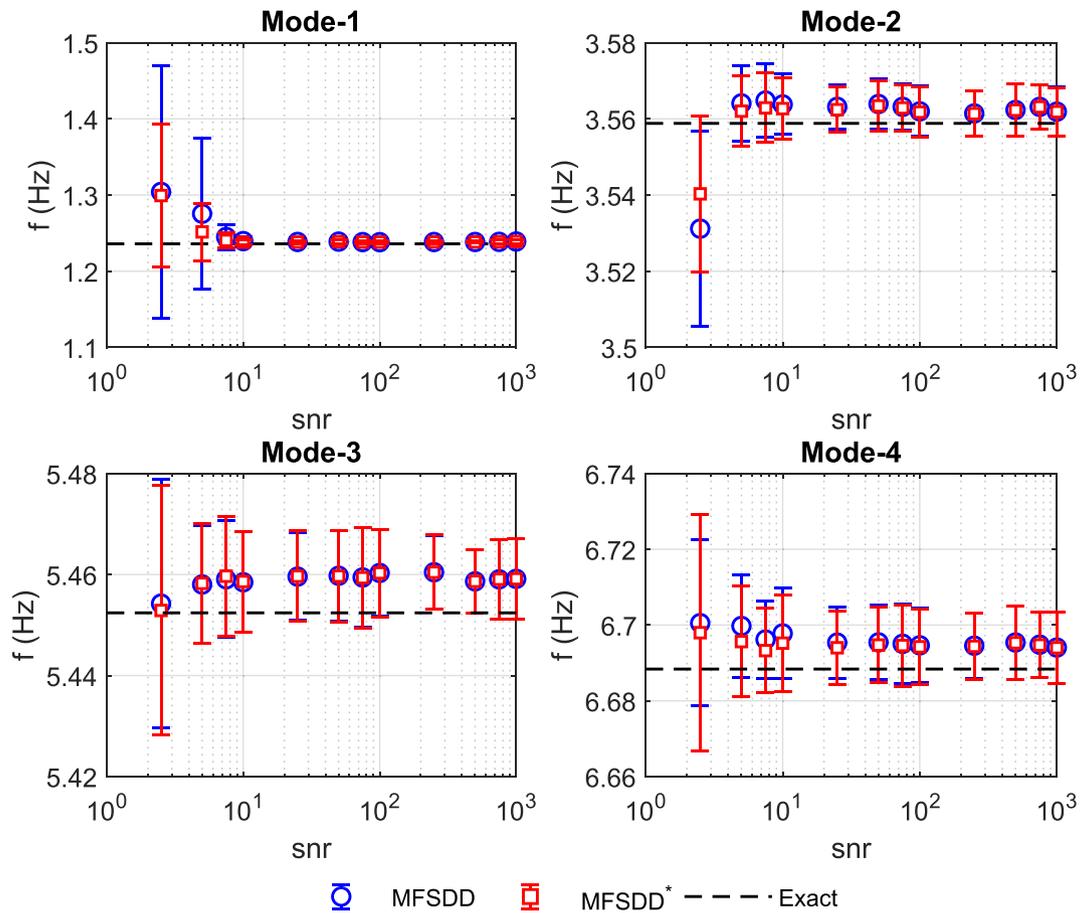
FIGURE 3 Variation of identified frequency values with respect to modal signal-to-noise ratio (*snr*)

6 | NUMERICAL ANALYSIS AND REAL DATA APPLICATIONS

In this section, a comprehensive investigation is undertaken to find out the effectiveness of MFSDD under some extreme cases and real operational conditions. To do so, a numerical example which investigates the effect of channel noise, heavy damping and earthquake loads is presented first. Second, a laboratory data example is investigated to see the effectiveness of the extended methodology under very low *snr* values. Then, a real data application which contains closely spaced and weakly excited modes is conducted.

6.1 | Channel noise, heavy modal damping, and earthquake load effects

To understand how the channel noise, heavy modal damping and earthquake loading influence the error propagation in identified modal parameters, an extensive numerical analysis is presented in this section. In this context, a four-story shear frame whose schematic representation is given in Figure 2 is numerically modeled. Then, the aforementioned effects are investigated separately based on the constructed numerical model. Finally, the combined effect of heavy modal damping and earthquake loading on the modal parameter estimation quality is assessed. A similar investigation which compares the effectiveness of MFSDD with respect to classical FSDD, SSI-Cov, and rFDD techniques can be found in Hızal.[38] Therefore, only MFSDD and its extended version (MFSDD*) are implemented in this presented numerical example from a more comprehensive perspective.
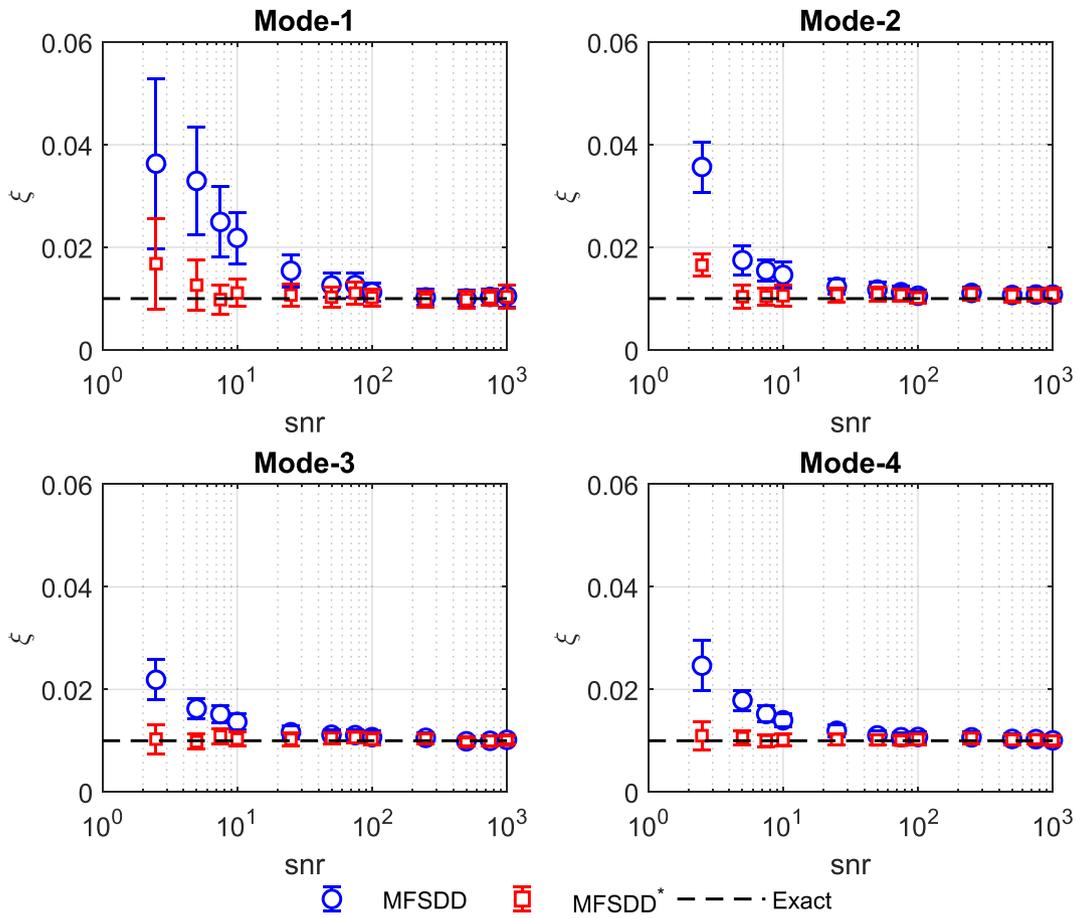
FIGURE 4 Variation of identified damping ratios with respect to modal signal-to-noise ratio (*snr*)

6.1.1 | Channel noise effect

To find out the pure channel noise effect on the modal identification results for the presented numerical model, a modal damping ratio is selected as 1% for all modes. Moreover, independent and identically distributed (*i.i.d.*) Gaussian forces are applied to floor levels of the structure so that the spectral density of the modal excitation will be 1 μg^2/Hz for all modes. The acceleration responses at each floor are simulated by the Newmark-beta integration scheme with 1/10000-s time sampling rate. Then, the simulated responses are contaminated by *i.i.d.* Gaussian numbers which represent the channel noise. Here, the considered channel noises are arranged so that their one-sided spectral density levels, S_n, will vary between 0.25 and 1000 μg^2/Hz. Thus, the resulting *snr* levels vary between 2.5 and 10,000. For each channel noise density level, totally 1000 trials have been simulated. The acceleration responses are acquired with 250-Hz sampling frequency and 5-min duration for each trial. Finally, the most probable values as well as the standard deviations of the modal parameters and prediction error are evaluated from those 1000 trials simulated for each channel noise density level.

Variations in the expected value of identified modal parameters as well as their standard deviations, with respect to the modal *snr* values are presented in Figure 3–6. The presented figures indicate that there is not a significant difference between the results obtained by MFSDD and MFSDD* for natural frequencies, modal shape vectors and prediction errors. Here, the most remarkable difference is observed in the results obtained for damping ratios at lower *snr* values. More specifically, the extended version of the MFSDD gives rather high-quality results for damping ratios under large channel noise effects (e.g., *snr* < 50). Moreover, it is also observed that the standard deviations of the expected values decrease for larger *snr*. However, an uncertainty still exists for the results obtained at larger *snr* values. In the other word, small channel noise and modeling errors do not guarantee that the uncertainty of the expected values are completely consumed.[45] Here, it should be noted that the given uncertainties are obtained by a frequentist perspective, which calculates the standard deviations obtained from the ensemble data sets. Instead, those uncertainties can also be
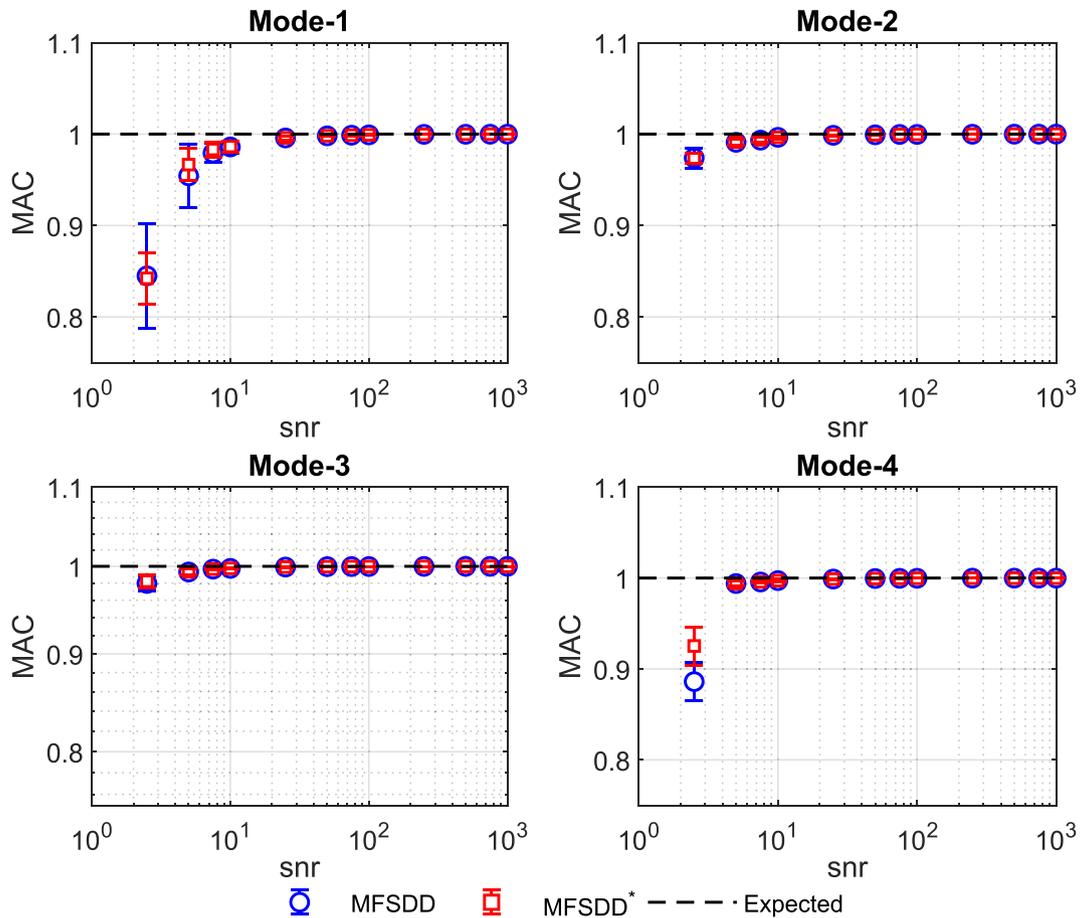
FIGURE 5   Variation of MAC values with respect to modal signal-to-noise ratio (*snr*)

estimated by a probabilistic perspective using a Bayesian approach, such as previous works.[13,14,43] On the other hand, it is known from the literature that the uncertainty information obtained by frequentist and probabilistic (Bayesian) perspectives equalize as the number of trials significantly increases.[47] Similarly, in case of uniform prior pdf, the expected values obtained by frequentist and Bayesian approaches become approximately the same for a large number of trials. Therefore, it can be deduced from here that the calculated expected values and standard deviations also reflect the probabilistic uncertainty information for the identified values.

Although a small modal damping ratio and Gaussian white noise input are considered for this numerical example, a modeling error may be inevitably observed, especially at the higher modes. This modeling error can be considered to stem from the contribution of unconsidered modes in the theoretical modeling of the single mode response. According to the results for prediction errors presented in Figure 6, one can observe that this modeling error term becomes more perceptible at higher modes. Here, Figure 6 presents the variation of the relative error ratio which is defined by Δ_e/S_n. For lower *snr* values, the modeling error term becomes negligibly small with respect to the channel noise, and therefore the relative error ratio approaches 1. As the channel noise decreases significantly (for *snr* > 100) modeling errors become more dominant on the identified prediction error term. The simulated results indicate that the modeling error term increases up to the pre-defined channel noise level (0.25 µg²/Hz) at fourth mode. This increase, however, does not influence the identification qualities significantly since it remains in very small levels with respect to the spectral density of modal response.

6.1.2 | Heavy damping effect

Neglecting the channel noise, the influence of heavy damping on the effectiveness of the applied methodology is investigated in this section. Here, modal damping ratios are arranged as 3%, 5%, 7%, and 9% for the considered numerical
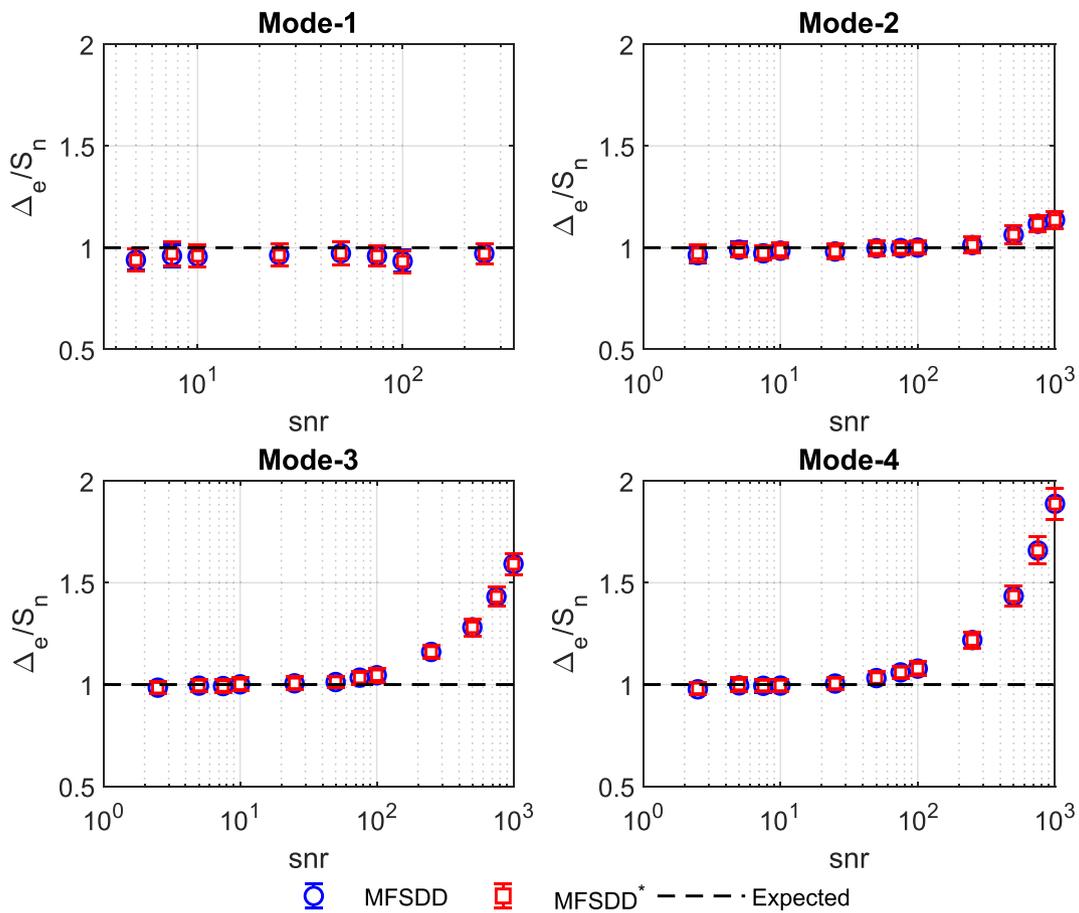
FIGURE 6 Variation of normalized prediction errors with respect to modal signal-to-noise ratio (*snr*)
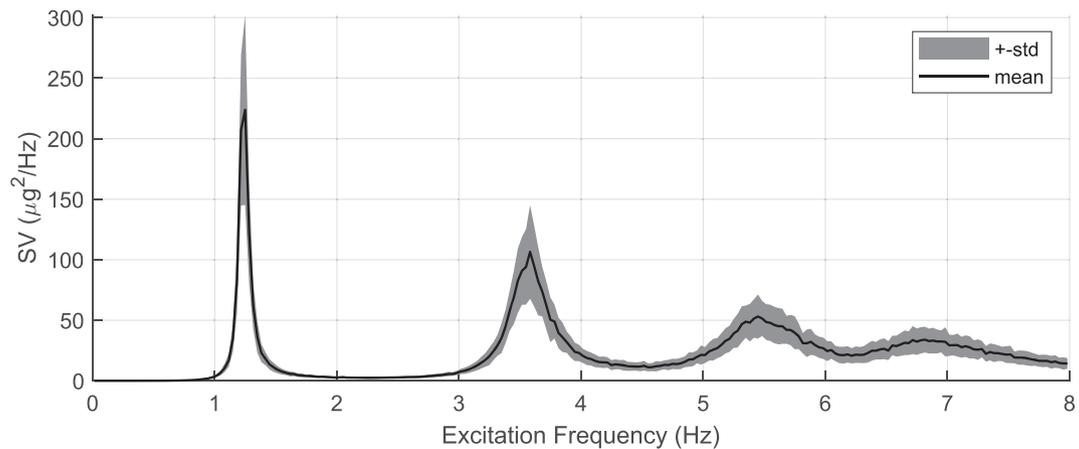

FIGURE 7 SV spectrum of the generated responses that represent heavy damped case (std: Standard deviation)

model, respectively. Totally 1000 trials have been simulated, each one is excited by Gaussian forces which make the spectral density of modal excitation 1 $\mu g^2/Hz$. For each trial, the acceleration responses at floor levels are recorded with 250-Hz sampling frequency and 5-min duration.

The SV spectrum that represents the simulated responses is presented in Figure 7. Here, the solid line indicates the expected value for the largest SVs calculated from the ensemble average of those obtained by the trials. Moreover, the standard deviations are also represented by the shaded area in the presented figure. One can deduce from here that the amplitude of the mean value shows a significant decrease at third and fourth modes, which amplifies the modeling
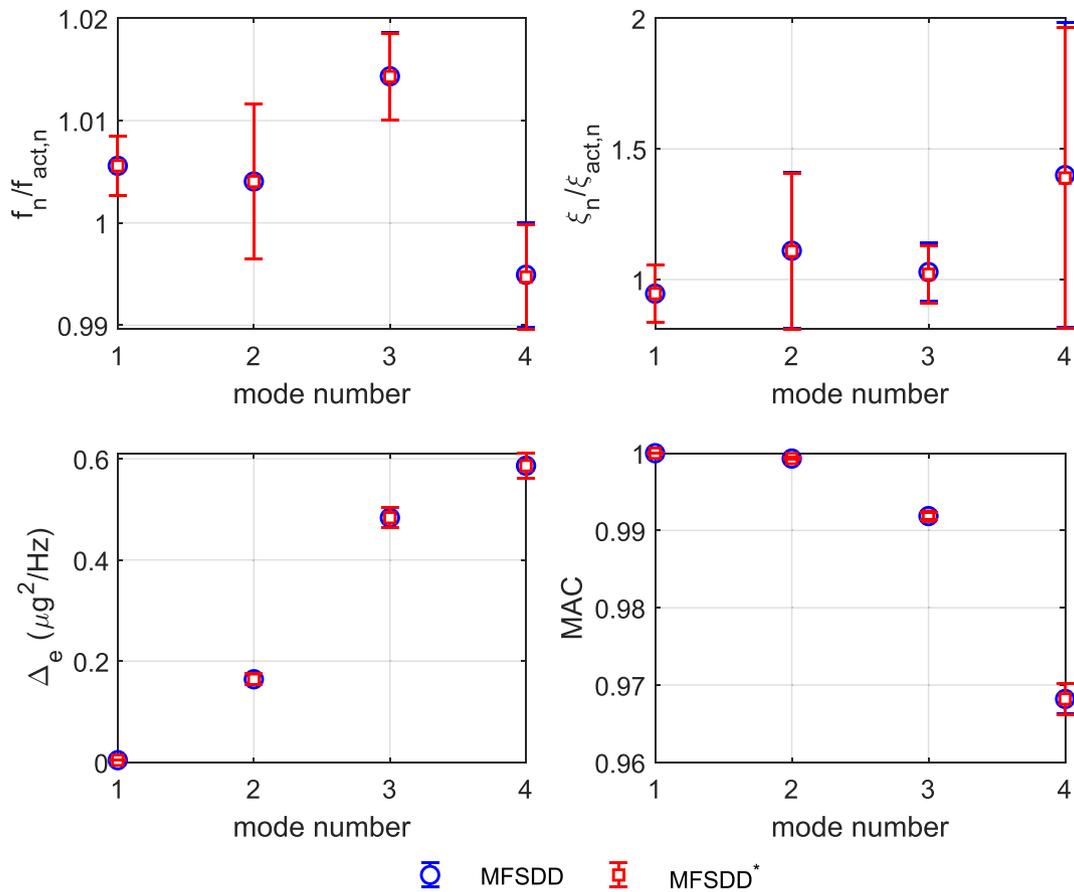
FIGURE 8 Variation of identified modal parameters with ±standard deviations

errors due to small modal damping assumption significantly. In addition, the deviations may also cause significant fluctuations among the simulated trials. Therefore, the standard deviations calculated for the modal parameters belonging to third and fourth modes can be expected larger than those calculated for the first two modes.

Expected value of the identified modal parameters as well as their standard deviations are presented in Figure 8. Here, the presented results for modal frequencies and damping ratios are normalized with respect to their actual values. At first view, it is apparently viewed that the results for MFSDD and MFSDD* perfectly match. In addition, the identified modal parameters show very small deviations from their actual values. Here, the maximum difference is observed in the third mode for modal frequencies (about 1.5%), and in the fourth mode for the damping ratio (about 40%) and the mode shape vector (about 3%), respectively. On the other hand, the identified standard deviations gradually increase at the larger modes. As it is mentioned above, it can be expected that the results obtained for the third and the fourth modes have larger standard deviations due to the large damping effect. This expectation is also compatible with the identified results. However, an unexpected increase is observed in the standard deviations of identified modal frequencies and damping ratios for the second mode. Similarly, the identified prediction errors gradually increase at higher modes, resulting in a modal *snr* about 25 for the fourth mode.

6.1.3 | Earthquake load effect

In this section, to illustrate the pure effect of the non-stationary input, the considered numerical model is investigated under earthquake excitations, considering 1% modal damping ratio, and neglecting the channel noise. For this purpose, totally eight earthquake records, whose auto-PSD spectra are given in Figure 9, are selected. Here, all the selected earthquakes have a sampling frequency rate of 200 Hz.

Identification results in terms of the expected modal parameters and corresponding standard deviations are presented in Figure 10. Here, the expected values reflect the ensemble average of the modal parameters obtained for the
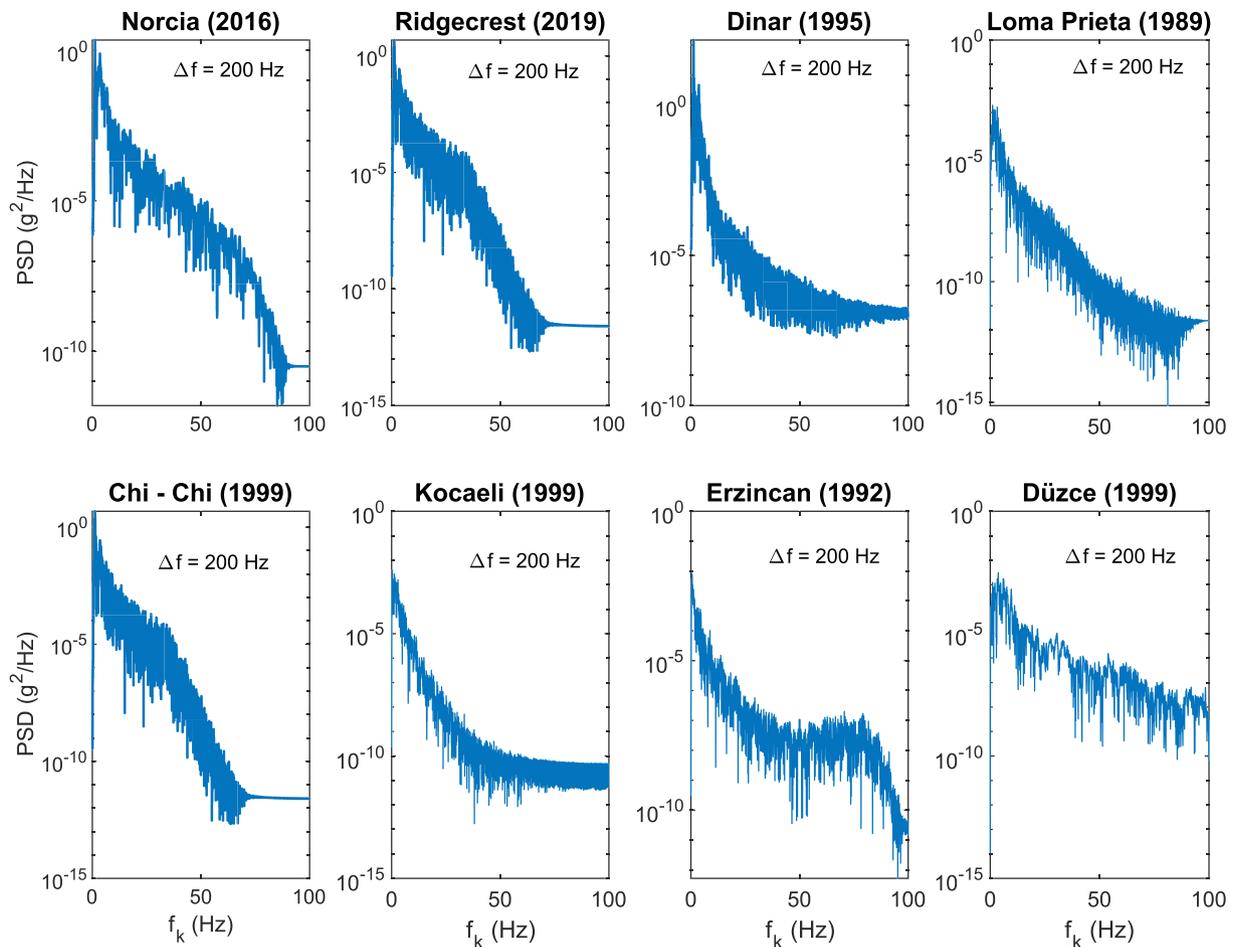
FIGURE 9 Power spectral density (PSD) spectra for the selected earthquake records (f_k: Excitation frequency, Δf: Sampling frequency)

selected records. Similar to the heavy modal damping case, the results by MFSDD* show better convergence to the actual values in comparison to MFSDD. The identified modal frequencies are compatible with the actual values and show a maximum deviation about 1%. For modal damping ratios, however, this deviation increases up to 40% for MFSDD while values by MFSDD* have a less divergence from the actual ones. On the other hand, the first three mode shape vectors show a good match with the actual values. Here, the minimum MAC value is obtained about 0.98 for the fourth modal shape vector, which can be considered quite reasonable for OMA applications.

Since the larger modes are weakly excited in earthquake induced motions, the prediction errors are expected to increase. In addition, the modeling errors may also increase in the prediction errors as it is observed from the previous examples. To the contrary, from Figure 10, one can observe that the identified prediction errors achieve a maximum at the third mode, and then show a decrease. This observation is also compatible with the result reported by Hızal[38]. Therefore, such a decrease can be induced by the loss of the quality in the identified prediction error term. This result indicates that if the signal quality of a mode significantly drops due to any reason (channel noise and/or modeling error) not only the identified modal frequencies, damping ratios and mode shape vectors but also the prediction error term can be adversely influenced. In the other word, inclusion of a prediction error term in the theoretical PSD does not guarantee that the possible prediction errors can be perfectly identified in case of weakly excited modes.

6.1.4 | Combined effect of heavy damping and earthquake load

To understand the influence of the combination of heavy damping and non-stationary input on the identification results, the four-story shear frame is investigated in this section considering those effects. Similar to the previous example, the previously used eight earthquake records are considered again. In addition, the modal damping ratios are
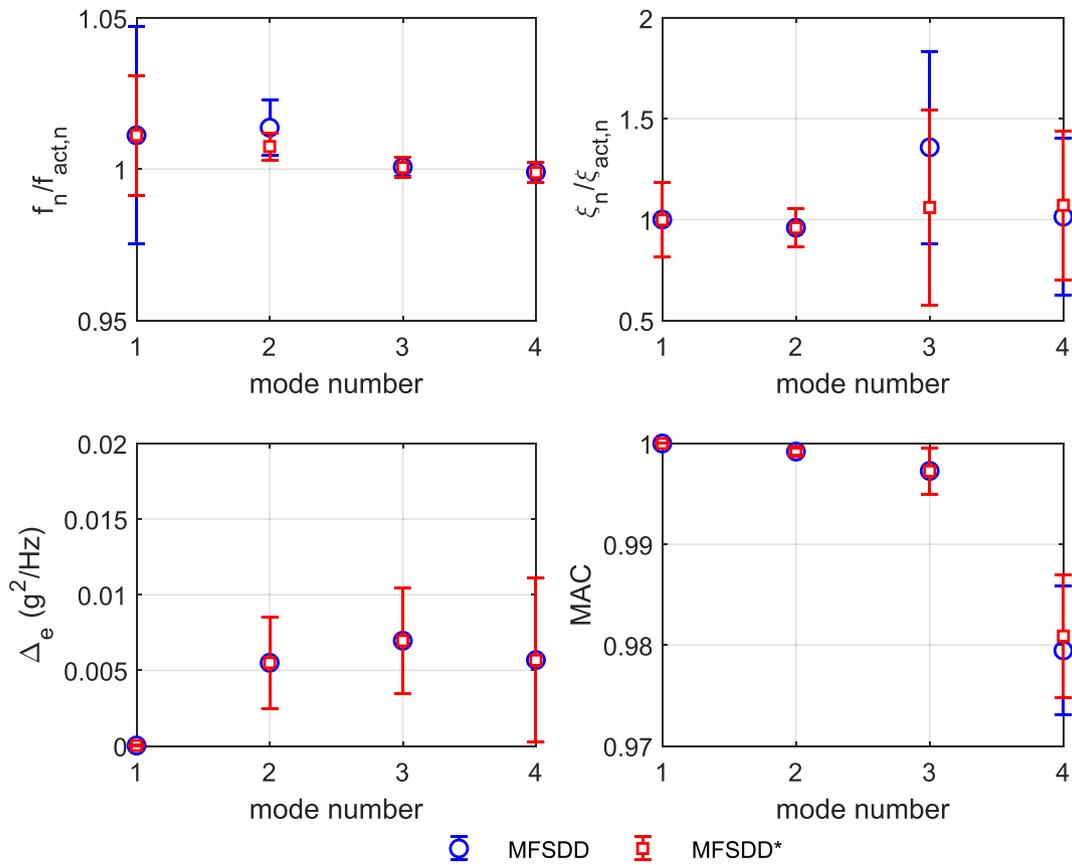
FIGURE 10 Variation of identified modal parameters with ±standard deviations

arranged as 3%, 5%, 7%, and 9%, as similar to the numerical example 6.1.2. In this context, the identification results are presented in Figure 11. In comparison to pure effects of heavy damping and earthquake loads, a dramatic decrease is observed in the quality of third and fourth modal shape vectors. Compared to the mode shapes, the modal frequencies and damping ratios show relatively better results. Here, the maximum relative deviations from the actual values are observed in the range of 0.1%–10% for the modal frequencies, and 1%–45% for damping ratios, respectively. Moreover, the prediction error follows a similar trend with results presented in Section 6.1.3. Considering all these aspects, one can conclude that the modal parameters are highly influenced by the combination of earthquake excitations and heavy modal damping. The decrease in the quality of identified results may neither be directly induced by the non-stationarity of input motion nor the heavy damping. Their combined effect, however, might be a much more challenging issue since the modeling errors increase ultimately. Additionally, the short data duration and small sampling frequency may also produce significant errors at the higher modes due to the increasing spectral leakage and aliasing effects.

6.2 | Laboratory data example

A 10-story laboratory shear frame structure is investigated in this section for the experimental validation of the extended methodology. This 10-story structure was previously used for the verification of various SHM applications such as Bayesian mode shape assembly and finite element model updating algorithms,[18,46] and pre-identification data merging strategies.[48,49] Therefore, reference values for modal parameters are also available for the considered structure to reveal the improvement in the modal identification quality by the proposed extended version.

Schematic view of the laboratory model and experimental system is shown in Figure 12. A more detailed description about the model and experimental system can also be found in the previous works.[18,46] In this study, a multi-setup measurement data which comprises significantly low *snr* values at the higher modes is used in the analysis. The corresponding multi-setup acceleration data was measured in the weak direction under small amplitude excitations,
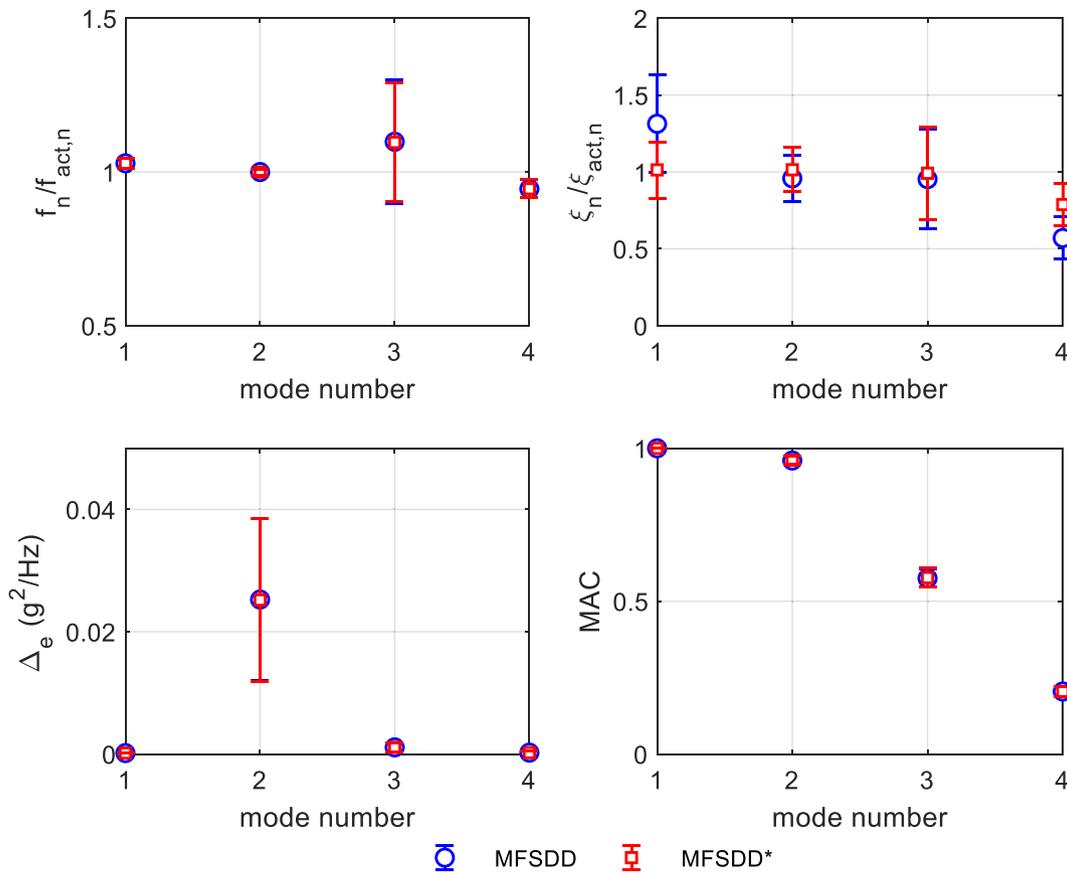
FIGURE 11 Variation of identified modal parameters with ±standard deviations

and has been originally investigated by Hızal et al.[46] for the experimental verification of a modified Bayesian mode shape assembly technique. The configuration of the measurement setups is presented in Table 2.

The SV spectra of the used multi-setup data is presented in Figure 12C. The possible modes and selected bandwidths are also highlighted in the presented figure. Here, the modal peak around the 15.50 Hz was not previously reported as a vibration mode of the 10-story model in the weak direction,[48] and therefore it is not considered in this study. The presented spectra reveal that the first five modes are well excited with neatly visible peaks. The larger modes, however, seem to be very poorly excited with the modal peaks which are not well perceptible. It can also be expected that these very poorly excited modes could negatively affect the modal identification quality. In this context, modal frequencies and damping ratios identified by MFSDD and MFSDD* as well as the reference values are presented in Table 3. These modal parameters are obtained as the sample mean of the values those identified at the individual setups. In addition, the modal shape vectors are assembled those obtained by the local setups using the Global Least Squares Approach.[50] The presented results for modal frequencies and damping ratios indicate that MFSDD* better converge to the reference values at the higher modes in comparison to MFSDD. The effectiveness of the MFSDD* is more apparently visible in the identified modal shape vectors which are presented in Figure 13. Especially at the seventh, eighth and ninth modes, whose peaks are nearly invisible in the SV spectrum, the MFSDD* shows significantly better performance.

6.3 | Real data example

A 58-story building whose schematic representation is presented in Figure 14 is investigated using the ambient and earthquake induced vibration records. The structural system of the considered building is composed of dual core reinforced concrete shear walls which are combined by an outrigger belt-truss.[51,52] A measurement system has been previously instrumented by the California Strong Motion Instrumentation Program (CSMIP) of the California Geological Survey (CGS).[51] The measurement system comprises totally 32 accelerometers deployed on 14 floors of the
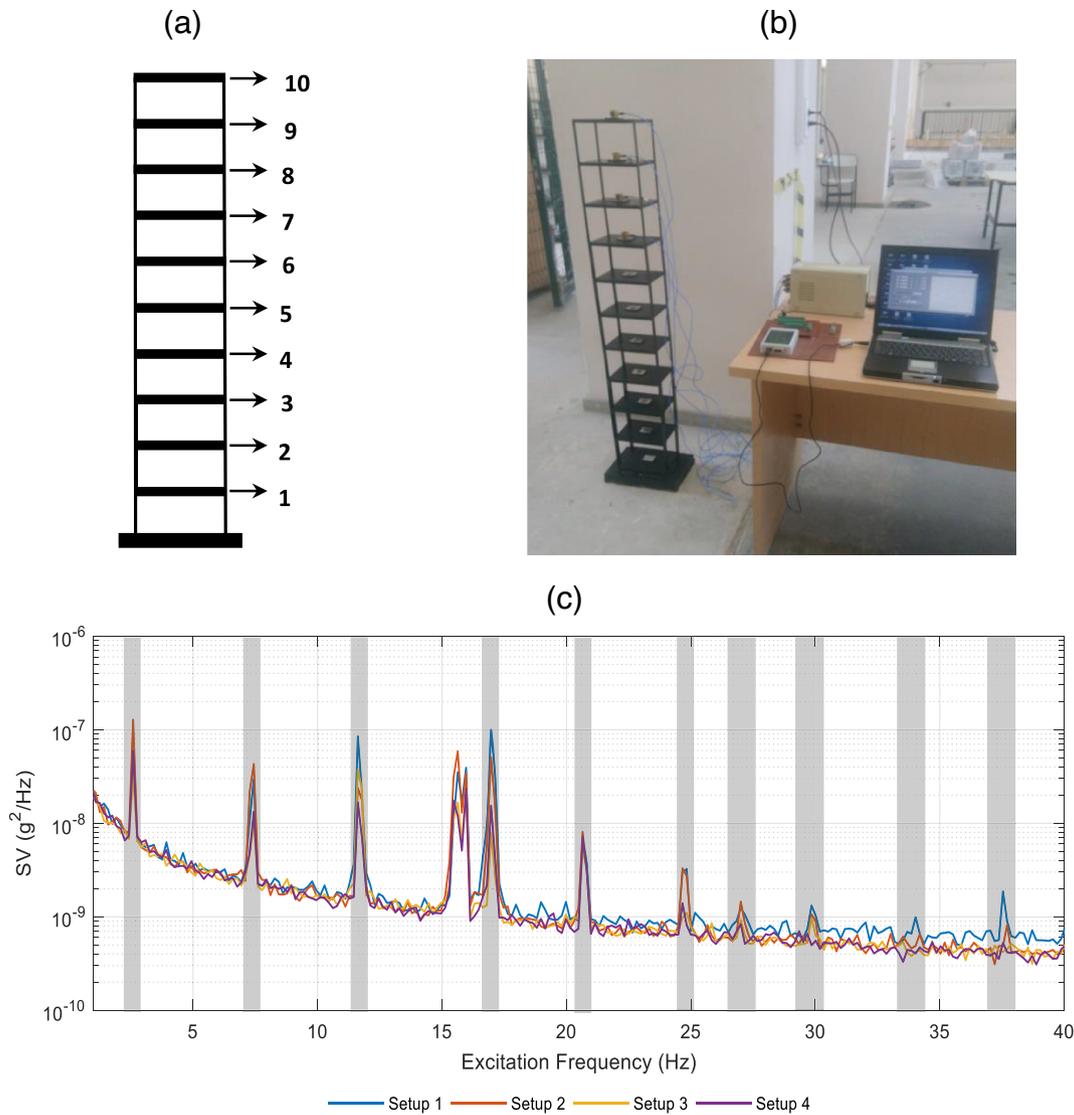
(a)

(b)

(c)

FIGURE 12 (A) Schematic view of the laboratory structure, (B) experimental system (reproduced from Hızal and Turan[18]), (C) SV spectrum obtained from the multi setup data

TABLE 2 Configuration of measurement setups

Setup number	Measured DOF
1	1, 2, 9, 10
2	3, 4, 9, 10
3	5, 6, 9, 10
4	7, 8, 9, 10

building.[51] The locations and directions of the instrumented accelerometer sensors are also presented in Figure 14A. In this context, the modal parameters of the considered structure are first investigated using ambient vibration data. Subsequently, the same procedure is undertaken using the measurement data obtained from the South Napa earthquake. All the considered measurement data can be accessed at the United States Geological Survey (USGS), Center for Engineering Strong Motion Data (CESMD).[53]

The SV spectra obtained for ambient measurement data is presented in Figure 14B. Here, the presented spectra show the first three SVs of $\mathbf{G}_{yy}(\omega_k)$. At first view, totally six modes namely first and second East–West (EW), North–South (NS) and torsional modes are visible, respectively. In addition, the selected bandwidths are also indicated

TABLE 3 Identified modal parameters and reference values

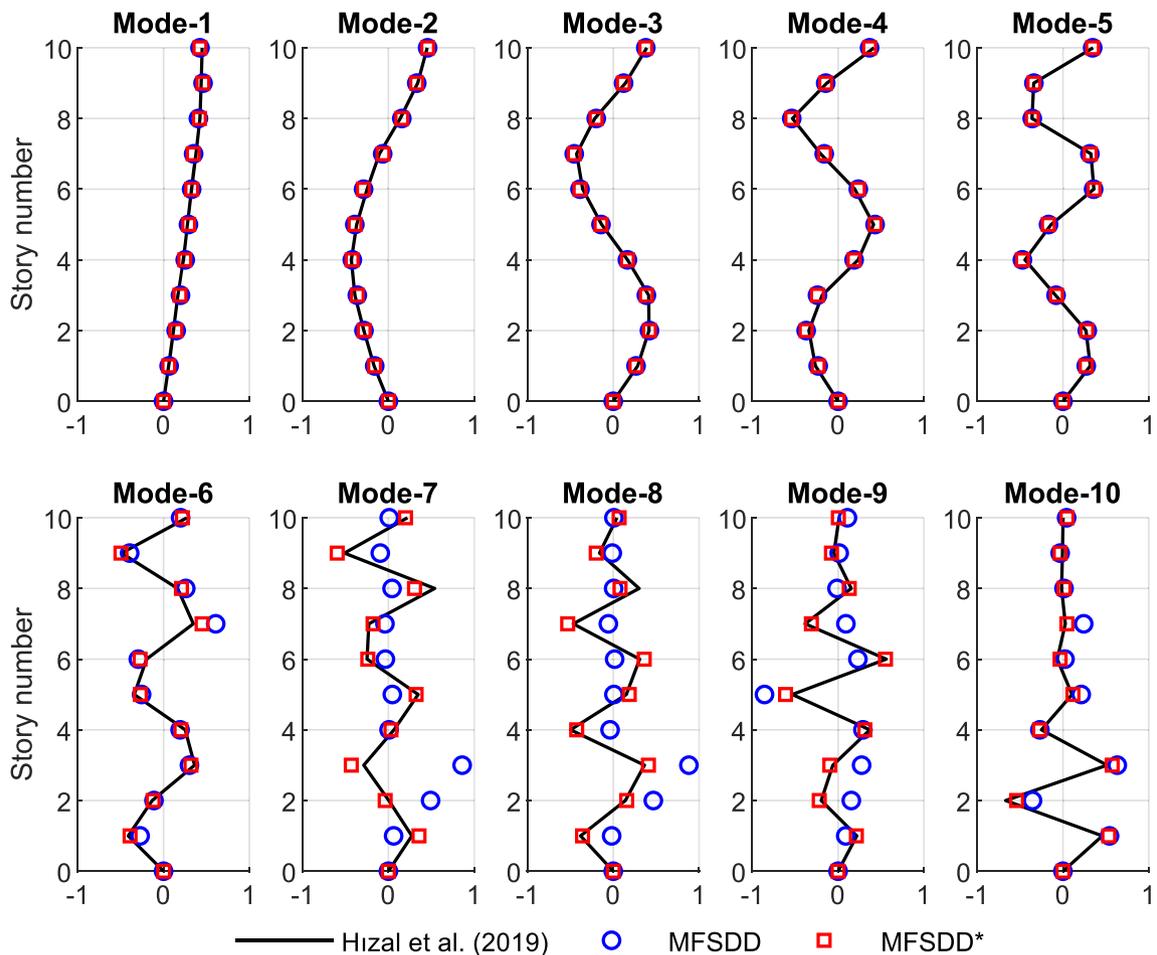| Mode number	f (Hz)			ξ (%)		
	Hızal et al.[46]	MFSDD	MFSDD*	Hızal et al.[46]	MFSDD	MFSDD*
1	2.59	2.61	2.62	0.34	0.17	0.22
2	7.32	7.38	7.37	0.26	0.21	0.19
3	11.65	11.70	11.69	0.23	0.23	0.15
4	16.96	17.02	17.02	0.16	0.27	0.21
5	20.65	20.71	20.71	0.14	0.25	0.15
6	24.69	24.73	24.73	0.16	0.62	0.20
7	26.94	27.03	27.04	0.18	0.67	0.19
8	29.85	29.69	29.74	0.16	0.36	0.16
9	33.19	32.72	33.36	0.16	0.93	0.18
10	37.47	38.01	37.66	0.18	0.48	0.14

FIGURE 13 Identified and reference mode shapes

in Figure 14B. In this context, the selected bandwidths span closely spaced modes at 0.26–0.30 and 1.10 Hz, which correspond to the first and second modes in the EW and NS directions, respectively.

The first six modal frequencies and damping ratios identified by using MFSDD and MFSDD* are presented in Table 4. For comparison purposes, BFFTA is also performed for the same problem. Here, BFFTA comprises two different algorithms for well separated and closely spaced modes.[54] The effectiveness of BFFTA algorithms has been
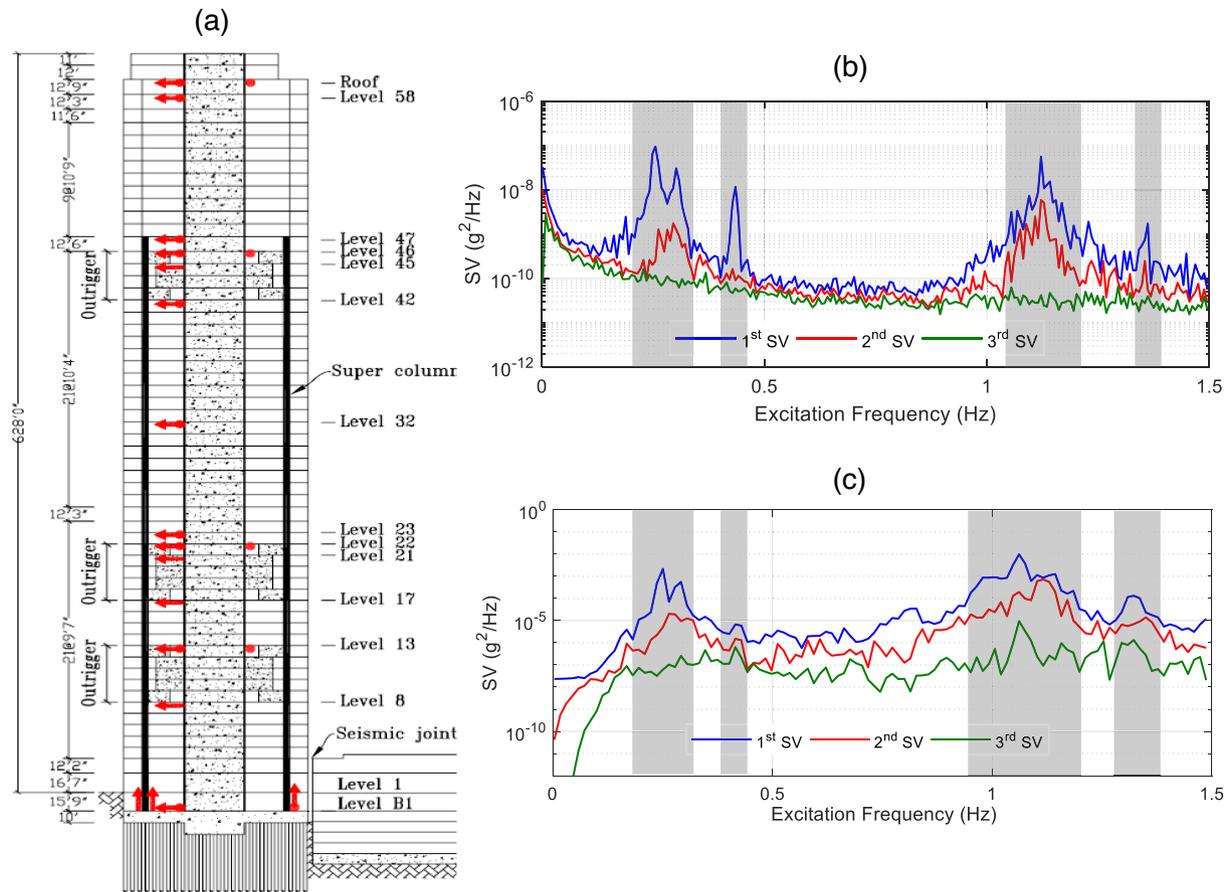
FIGURE 14 (A) Elevation view and sensor locations of the 58-story building (reproduced from Çelebi[51]). (B) SV spectrum for ambient data. (C) SV spectrum for South Napa earthquake data

TABLE 4 Identified modal parameters using ambient vibration data

Mode number	f (Hz)			ξ (%)		
	MFSDD	MFSDD*	BFFTA	MFSDD	MFSDD*	BFFTA
1 (EW)	0.26	0.26	0.26	2.01	1.75	1.91
2 (NS)	0.30	0.30	0.30	1.45	1.20	1.43
3 (Tors.)	0.43	0.43	0.43	0.79	0.79	0.47
4 (EW)	1.13	1.13	1.12	1.89	1.84	1.80
5 (NS)	1.13	1.13	1.13	1.13	1.03	1.05
6 (Tors.)	1.34	1.34	1.35	0.54	0.54	0.55

well-established in the literature.[54–56] Therefore, the results by BFFTA can be considered as reference values in this example. All the algorithms have been performed by using MATLAB[57] m-files developed by the first author. The presented results indicate that the considered algorithms provide rather similar solutions. This result also indicates that MFSDD and MFSDD* gives efficient solutions for closely spaced modes in case of those appearing in the perpendicular directions (e.g., in case one is NS and the other one is EW mode).

MAC values for the mode shapes identified by MFDD and MFSDD* with respect to those identified by BFFTA are calculated as {0,999, 1,000, 0.993, 1.000, 0.989} and {1.000, 1.000, 0.997, 1.000, 0.999}, respectively. In addition, 3D views of the mode shapes identified by MFSDD are shown in Figure 15.

Modal parameters of the 58-story building are investigated by using the acceleration response data recorded during the South Napa earthquake motion (24 August 2014, $M_w = 6.0$) as well.[51] The SV spectra of the recorded data are presented in Figure 14C. From the presented figure, it can be apparently observed that weakly excited modes appear in the
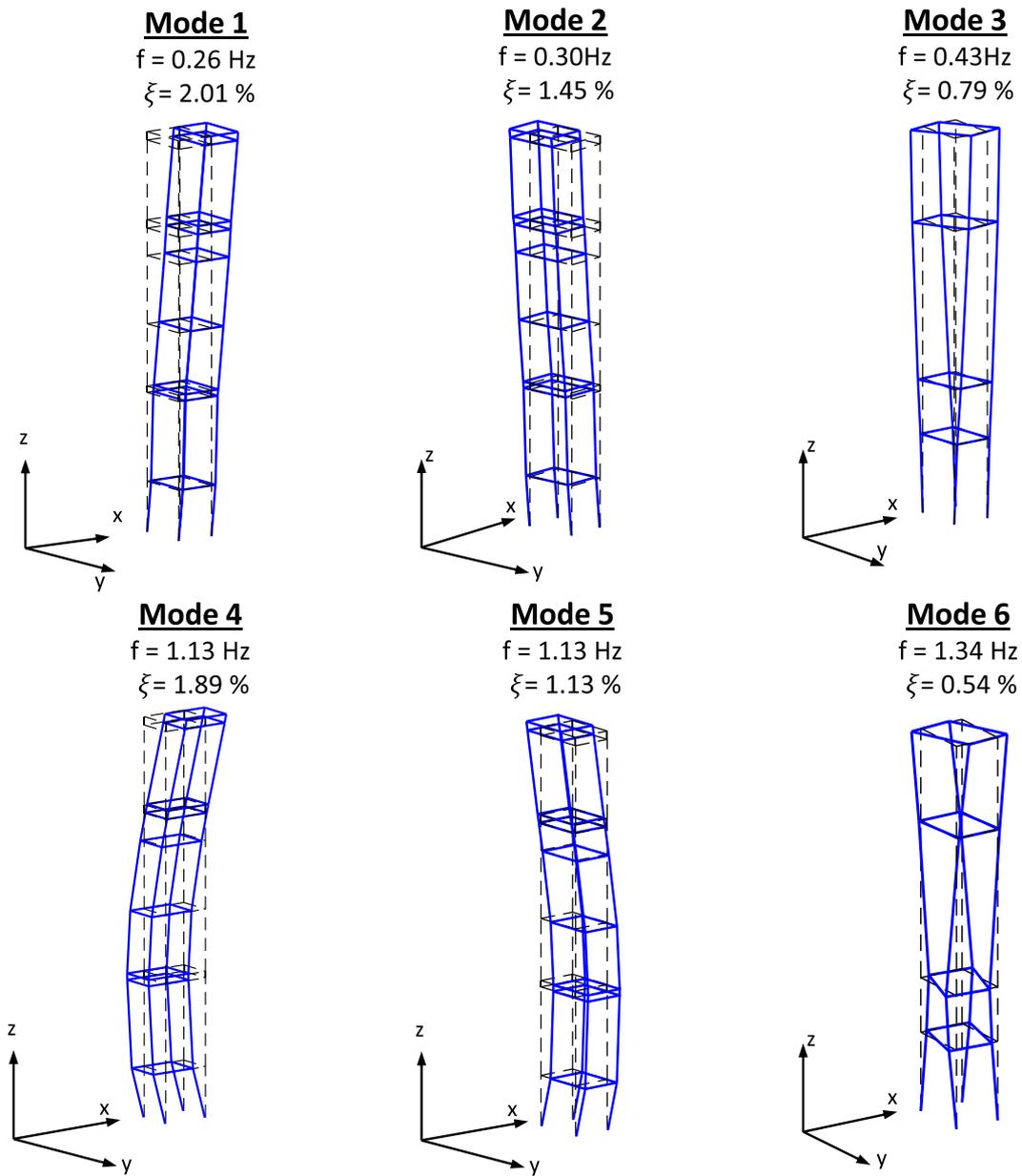
Mode 1
f = 0.26 Hz
ξ = 2.01 %

Mode 2
f = 0.30Hz
ξ = 1.45 %

Mode 3
f = 0.43Hz
ξ = 0.79 %

Mode 4
f = 1.13 Hz
ξ = 1.89 %

Mode 5
f = 1.13 Hz
ξ = 1.13 %

Mode 6
f = 1.34 Hz
ξ = 0.54 %

FIGURE 15 Identified 3D mode shapes

TABLE 5 Identified modal parameters using South Napa earthquake response data

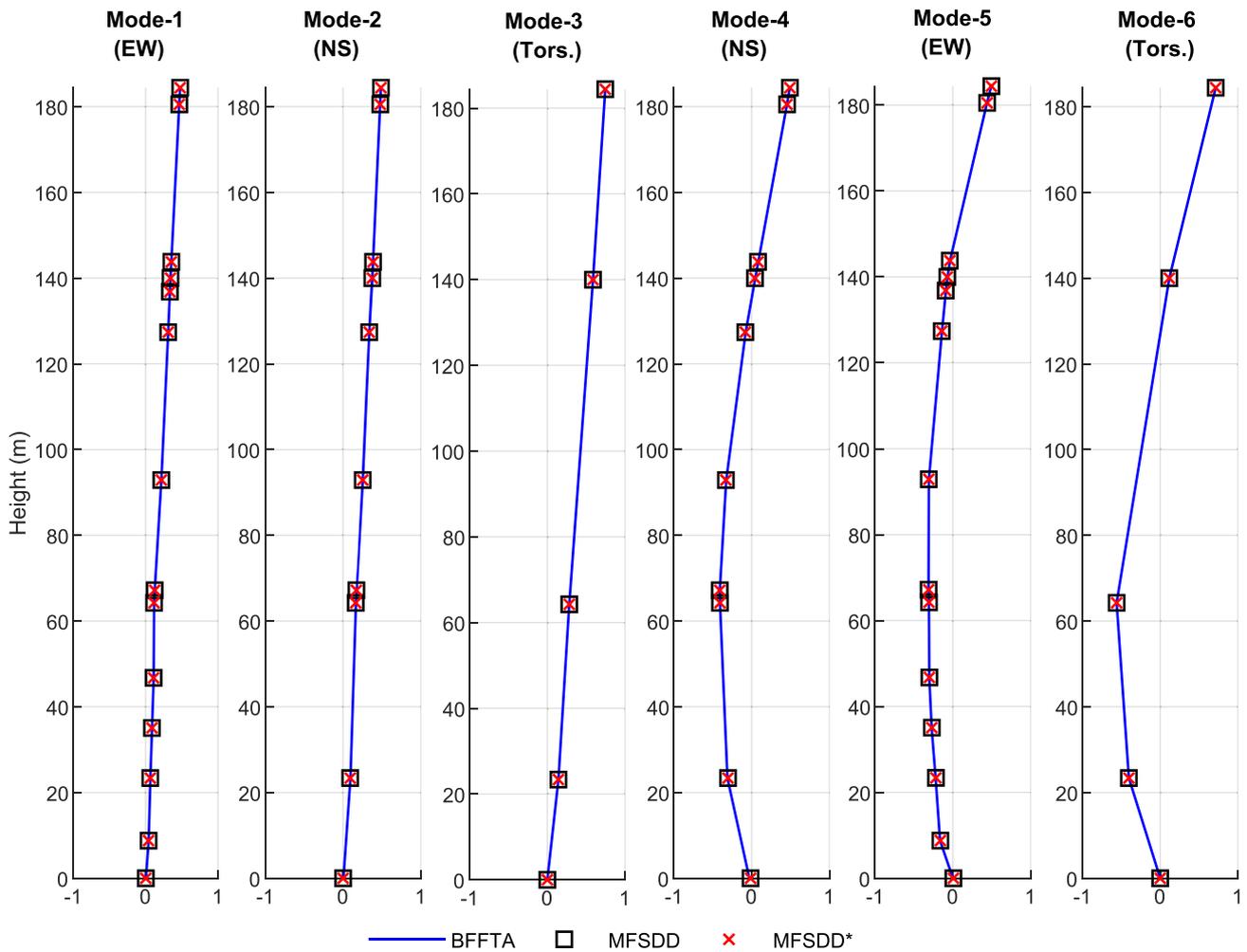| Mode number	f (Hz)				ξ (%)			
	MFSDD	MFSDD*	BFFTA	Çelebi[51]	MFSDD	MFSDD*	BFFTA	Çelebi[51]
1 (EW)	0.25	0.25	0.25	0.25	0.88	0.88	0.89	0.38
2 (NS)	0.29	0.29	0.29	0.28	1.68	1.67	1.62	0.35
3 (Tors.)	0.43	0.43	0.43	0.40	1.83	1.83	1.43	1.47
4 (NS)	1.07	1.07	1.07	1.06	1.99	1.97	0.97	0.46
5 (EW)	1.10	1.10	1.11	1.11	1.55	1.55	2.79	1.71
6 (Tors.)	1.33	1.33	1.35	1.31	1.99	1.97	2.35	0.62

FIGURE 16 2D mode shapes identified by using South Napa earthquake response data

torsional directions. The selected bandwidths are also highlighted in the presented spectra. Similarly, the selected bands comprise closely spaced modes in the NS and EW directions.

Identified modal parameters including natural frequencies and damping ratios are presented in Table 5. For comparison purposes, the results reported by Çelebi[51] are also given. Here, Çelebi[51] utilizes MATLAB System Identification which implements numerical algorithms for Subspace State Space System Identification (N4SID). The presented results indicate that MFSDD and BFFTA show fine agreement for modal frequencies and damping ratios. Although the modal frequencies reported by Çelebi[51] are quite similar with those obtained by MFSDD and BFFTA, a relatively large difference (about 400%) is observed in modal damping ratios. In addition, 2D mode shape vectors identified by MFSDD, MFSDD* and BFFTA are shown in Figure 16.

7 | CONCLUSIONS

To investigate the propagation of identification errors in FDD implementations, an extended probabilistic model is presented in this study. In this context, the prediction error term is included in the theoretical output PSD. Then, the theoretical expected value of the output PSD is updated by a maximum likelihood estimation using central complex Wishart distribution. The theoretical background of the applied methodology regarding the adaptation of Wishart distribution is discussed in detail. Finally, a comprehensive investigation is conducted considering several challenges such as channel noise, heavy modal damping, earthquake input and closely spaced modes, respectively. The fundamental results of the conducted study are summarized below.

- So far, the SHM applications regarding the central complex Wishart distributions, such as BSDA, requires multiple measurement sets due to preserve the mathematical stability of the considered pdf. Otherwise, the normalizing constant becomes a non-definite number which makes it impossible to incorporate the corresponding pdf in Bayesian applications. To do so, either multiple measurement sets are used, or a single set is separated to multiple segments. Both implementations comprise some drawbacks such as increase in the instrumentation/measurement effort and significant decrease in the identification quality due to the leakage effect. In this study however, it is mathematically shown that the central complex Wishart distribution can be used in case of a single measurement set without loss of identification quality.

- In the presence of channel noise, compared to MFSDD, the extended version, MFSDD*, provides significant improvement in the modal damping identification quality. For natural frequencies, modal shape vectors and prediction error terms, both MFSDD and MFSDD* show good performance especially in the case of $snr > 10$. Therefore, in case of noisy data, MFSDD* can be preferred to improve the identification quality in modal damping.

- According to the numerical analysis, no significant difference is observed between MFSDD and MFSDD* in case of the presence of model errors due to the heavy damping and earthquake input. A good performance is observed especially at the lower modes in the analysis conducted for pure heavy damping and earthquake input effects. However, a significant decrease is observed at the higher modes in case of the combined heavy modal damping and earthquake input effects. This indicates that the applied methodology is less successful in modeling error identification compared to channel noise. Since the identification of modeling errors is a rather difficult task in SHM application, the observed results make this conclusion reasonable.

- The conducted laboratory and real data examples indicate that MFSDD* show good performance in the presence of very poorly excited or closely spaced modes. Similar to numerical example, no significant difference is observed in identified modal parameters in case of earthquake input for real data application.

ORCID

Çağlayan Hızal https://orcid.org/0000-0002-9783-6511
Engin Aktaş https://orcid.org/0000-0002-5706-2101

REFERENCES

1. Caicedo JM. Practical guidelines for the natural excitation technique (NExT) and the eigensystem realization algorithm (ERA) for modal identification using ambient vibration. *Exp Tech*. 2011;35(4):52-58. https://doi.org/10.1111/j.1747-1567.2010.00643.x
2. Ibrahim SR. Fundementals of time domain modal identification. *Modal Anal Test*. 1999;241-250. https://doi.org/10.1007/978-94-011-4503-9_11
3. Van Overschee P, De Moor B. Subspace algorithms for the stochastic identification problem. *Automatica*. 1993;29(3):649-660.
4. Peeters B, De Roeck G. Reference-based stochastic subspace identification for output-only modal analysis. *Mech. Syst. Signal Process.* 1999;13(6):855-878. https://doi.org/10.1006/mssp.1999.1249
5. Fan G, Li J, Hao H. Improved automated operational modal identification of structures based on clustering. *Struct Control Heal Monit*. 2019;26(12):1-23. https://doi.org/10.1002/stc.2450
6. Wu W-H, Wang S-W, Chen C-C, Lai G. Application of stochastic subspace identification for stay cables with an alternative stabilization diagram and hierarchical sifting process. *Struct. Control Heal. Monit.* 2016;23(9):1194-1213. https://doi.org/10.1002/stc
7. Loh CH, Loh KJ, Sen Yang Y, Hsiung WY, Huang YT. Vibration-based system identification of wind turbine system. *Struct. Control Heal. Monit.* 2017;24(3):1-19. https://doi.org/10.1002/stc.1876
8. Carden EP, Mita A. Challenges in developing confidence intervals on modal parameters estimated for large civil infrastructure with stochastic subspace identification. *Struct. Control Heal. Monit.* 2011;18:53-78. https://doi.org/10.1002/stc.358
9. Loh C-H, Weng J-H, Chen C-H, Lu K-C. System identification of mid-story isolation building using both ambient and earthquake response data. *Struct. Control Heal. Monit.* 2013;20(2):139-155. https://doi.org/10.1002/stc.479
10. Mao JX, Wang H, Fu YG, Spencer BF. Automated modal identification using principal component and cluster analysis: application to a long-span cable-stayed bridge. *Struct Control Heal Monit*. 2019;26(10):1-20. https://doi.org/10.1002/stc.2430
11. Yang J-H, Lam H-F, Beck JL. Bayes-Mode-ID: a Bayesian modal-component-sampling method for operational modal analysis. *Eng Struct*. 2019;189:222-240. https://doi.org/10.1016/j.engstruct.2019.03.047
12. Yuen KV, Katafygiotis LS. Bayesian time-domain approach for modal updating using ambient data. *Probabilistic Eng Mech*. 2001;16(3):219-231. https://doi.org/10.1016/S0266-8920(01)00004-2
13. Katafygiotis LS, Yuen KV. Bayesian spectral density approach for modal updating using ambient data. *Earthq Eng Struct Dyn*. 2001;30(8):1103-1123. https://doi.org/10.1002/eqe.53
14. Yuen K, Katafygiotis LS. Bayesian fast Fourier transform approach for modal updating using ambient data. *Adv Struct Eng*. 2003;6(2):81-95. https://doi.org/10.1260/136943303769013183

15. Au SK, Zhang FL. Fast Bayesian ambient modal identification incorporating multiple setups. *J Eng Mech.* 2012;138(7):800-815. https://doi.org/10.1061/(asce)em.1943-7889.0000385

16. Zhang FL, Ni YC, Lam HF. Bayesian structural model updating using ambient vibration data collected by multiple setups. *Struct. Control Heal. Monit.* 2017;24(12):1-18. https://doi.org/10.1002/stc.2023

17. Zhang FL, Kim CW, Goi Y. Efficient Bayesian FFT method for damage detection using ambient vibration data with consideration of uncertainty. *Struct. Control Heal. Monit.* 2021;28(2):1-24. https://doi.org/10.1002/stc.2659

18. Hızal Ç, Turan G. A two-stage Bayesian algorithm for finite element model updating by using ambient response data from multiple measurement setups. *J Sound Vib.* 2020;469:115139. https://doi.org/10.1016/j.jsv.2019.115139

19. Au S-K, Ni Y-C. Fast Bayesian modal identification of structures using known single-input forced vibration data Siu-Kui. *Struct. Control Heal. Monit.* 2014;21(3):381-402. https://doi.org/10.1002/stc.1571

20. Zhang F-L, Au S-K, Lam H-F. Assessing uncertainty in operational modal analysis incorporating multiple setups using a Bayesian approach. *Struct. Control Heal. Monit.* 2015;22(3):395-416. https://doi.org/10.1002/stc.1679

21. Yan WJ, Katafygiotis LS. A two-stage fast Bayesian spectral density approach for ambient modal analysis. Part I: posterior most probable value and uncertainty. *Mech Syst Signal Process.* 2015;54:139-155. https://doi.org/10.1016/j.ymssp.2014.07.027

22. Zhu YC, Au SK. Spectral characteristics of asynchronous data in operational modal analysis. *Struct. Control Heal. Monit.* 2017;24(11):1-15. https://doi.org/10.1002/stc.1981

23. Sedehi O, Katafygiotis LS, Papadimitriou C. Hierarchical Bayesian operational modal analysis: theory and computations. *Mech. Syst. Signal Process.* 2020;140:106663. https://doi.org/10.1016/j.ymssp.2020.106663

24. Au SK. Uncertainty law in ambient modal identification—Part I: theory. *Mech. Syst. Signal Process.* 2014;48(1-2):15-33. https://doi.org/10.1016/j.ymssp.2013.07.016

25. Au SK. Uncertainty law in ambient modal identification—Part II: implication and field verification. *Mech. Syst. Signal Process.* 2014;48(1-2):34-48. https://doi.org/10.1016/j.ymssp.2013.07.017

26. Brincker R, Zhang L, Andersen P. Modal identification of output-only systems using frequency domain decomposition. *Smart Mater Struct.* 2001;10(3):441-445. https://doi.org/10.1088/0964-1726/10/3/303

27. Zimmerman AT, Lynch JP. Market-based frequency domain decomposition for automated mode shape estimation in wirwless sensor network. *Struct. Control Heal. Monit.* 2010;17(7):808-824. https://doi.org/10.1002/stc.415

28. Ozer E, Purasinghe R, Feng MQ. Multi-output modal identification of landmark suspension bridges with distributed smartphone data: Golden Gate Bridge. *Struct. Control Heal. Monit.* 2020;27(10):1-29. https://doi.org/10.1002/stc.2576

29. Brincker R, Ventura CE, Andersen P. Damping estimation by frequency domain decomposition. *Proc Int Modal Anal Conf-IMAC.* 2001;1:698-703.

30. Pioldi F, Ferrari R, Rizzi E. Output-only modal dynamic identification of frames by a refined FDD algorithm at seismic input and high damping. *Mech. Syst. Signal Process.* 2016;68–69:265-291. https://doi.org/10.1016/j.ymssp.2015.07.004

31. Pioldi F, Ferrari R, Rizzi E. Seismic FDD modal identification and monitoring of building properties from real strong-motion structural response signals. *Struct. Control Heal. Monit.* 2017;24(11):1-20. https://doi.org/10.1002/stc.1982

32. Pioldi F, Rizzi E. A refined Frequency Domain Decomposition tool for structural modal monitoring in earthquake engineering. *Earthq Eng Eng Vib.* 2017;16(3):627-648. https://doi.org/10.1007/s11803-017-0394-9

33. Pioldi F, Ferrari R, Rizzi E. Earthquake structural modal estimates of multi-storey frames by a refined Frequency Domain Decomposition algorithm. *J Vib Control.* 2017;23(13):2037-2063. https://doi.org/10.1177/1077546315608557

34. Pioldi F, Rizzi E. Earthquake-induced structural response output-only identification by two different Operational Modal Analysis techniques. *Earthq Eng Struct Dyn.* 2018;47(1):257-264. https://doi.org/10.1002/eqe.2947

35. Pioldi F, Rizzi E. Assessment of Frequency versus Time Domain enhanced technique for response-only modal dynamic identification under seismic excitation. *Bull Earthq Eng.* 2018;16(3):1547-1570. https://doi.org/10.1007/s10518-017-0259-7

36. Pioldi F, Salvi J, Rizzi E. Refined FDD modal dynamic identification from earthquake responses with Soil-Structure Interaction. *Int J Mech Sci.* 2017;127:47-61. https://doi.org/10.1016/j.ijmecsci.2016.10.032

37. Zhang L, Wang T, Tamura Y. A frequency-spatial domain decomposition (FSDD) method for operational modal analysis. *Mech Syst Signal Process.* 2010;24(5):1227-1239. https://doi.org/10.1016/j.ymssp.2009.10.024

38. Hızal Ç. Modified frequency and spatial domain decomposition method based on maximum likelihood estimation. *Eng Struct.* 2020;224:111007. https://doi.org/10.1016/j.engstruct.2020.111007

39. El-Kafafy M, Peeters B, Guillaume P, De Troyer T. Constrained maximum likelihood modal parameter identification applied to structural dynamics. *Mech. Syst. Signal Process.* 2016;72–73:567-589. https://doi.org/10.1016/j.ymssp.2015.10.030

40. El-Kafafy M, De Troyer T, Peeters B, Guillaume P. Fast maximum-likelihood identification of modal parameters with uncertainty intervals: a modal model-based formulation. *Mech. Syst. Signal Process.* 2013;37(1-2):422-439. https://doi.org/10.1016/j.ymssp.2013.01.013

41. El-Kafafy M, De Troyer T, Guillaume P. Fast maximum-likelihood identification of modal parameters with uncertainty intervals: a modal model formulation with enhanced residual term. *Mech. Syst. Signal Process.* 2014;48(1-2):49-66. https://doi.org/10.1016/j.ymssp.2014.02.011

42. El-Kafafy M, Guillaume P, Peeters B, Geluk T. The MLMM modal parameter estimation method: a new feature to maximize modal model robustness. *Mech. Syst. Signal Process.* 2019;120. https://doi.org/10.1016/j.ymssp.2018.10.015

43. Au S-K. Fast Bayesian FFT method for ambient modal identification with separated modes. *J Eng Mech.* 2011;137(3):214-226. https://doi.org/10.1061/(asce)em.1943-7889.0000213

44. Au SK. Insights on the Bayesian spectral density method for operational modal analysis. *Mech. Syst. Signal Process.* 2016;66–67:1-12. https://doi.org/10.1016/j.ymssp.2015.04.023

45. Au SK. *Operational Modal Analysis: Modeling, Bayesian Inference, Uncertainty Laws.* Springer; 2017 10.1007/978-981-10-4118-1.

46. Hızal Ç, Turan G, Aktaş E, Ceylan H. A mode shape assembly algorithm by using two stage Bayesian Fast Fourier Transform Approach. *Mech Syst Signal Process.* 2019;134:106328. https://doi.org/10.1016/j.ymssp.2019.106328

47. Au SK. Connecting Bayesian and frequentist quantification of parameter uncertainty in system identification. *Mech. Syst. Signal Process.* 2012;29:328-342. https://doi.org/10.1016/j.ymssp.2012.01.010

48. Ceylan H, Turan G, Hızal Ç. Pre-identification data merging for multiple setup measurements with roving references. *Exp Tech.* 2020;44(4):435-456. https://doi.org/10.1007/s40799-020-00365-w

49. Hızal Ç. Frequency domain data merging in operational modal analysis based on least squares approach. *Measurement.* 2021;170:108742. https://doi.org/10.1016/j.measurement.2020.108742

50. Au SK. Assembling mode shapes by least squares. *Mech. Syst. Signal Process.* 2011;25(1):163-179. https://doi.org/10.1016/j.ymssp.2010.08.002

51. Çelebi M. Responses of a 58 story RC dual core shear wall and outrigger frame building inferred from two earthquakes Responses of a 58 story RC dual core shear wall and outrigger frame building inferred from two earthquakes. *Earthq Spectra.* 2016;32(4):2449-2471.

52. Roorda DD, Rodrigues NJ. Design of the tallest reinforced concrete structure in California 58-story residential tower in San Francisco. In: Don A, Carlos V, David H, Marc H, eds. *Structures Congress 2008: Crossing Borders.* Vol. 314. ASCE; 2008:1-9. https://doi.org/10.1061/41016(314)85

53. United States Geological Survey, *Center of engineering strong motion data*, (2020). www.strongmotioncenter.org (accessed September 5, 2018).

54. Hızal Ç., Modal identification of structures by using Bayesian statistics, Ph.D. Dissertation, Izmir Institute of Technology, 2019.

55. Au SK. Fast Bayesian ambient modal identification in the frequency domain, Part II: posterior uncertainty. *Mech. Syst. Signal Process.* 2012;26:76-90. https://doi.org/10.1016/j.ymssp.2011.06.019

56. Au SK. Fast Bayesian ambient modal identification in the frequency domain, Part I: posterior most probable value. *Mech. Syst. Signal Process.* 2012;26:60-75. https://doi.org/10.1016/j.ymssp.2011.06.017

57. MATLAB. *2018a (Computer Software).* Natick, MA: MathWorks; 2018.