Numerical Methods in Civil Engineering

Journal Homepage: https://nmce.kntu.ac.ir/

Parametric Study of Stochastic Subspace Algorithms in Modal Analysis of Moment-Resistant Frames

Mehran Pourgholi* and Saeid Mahdavi**

ARTICLE INFO

RESEARCH PAPER

Article history:
Received:
February 2022
Revised:
December 2023
Accepted:
May 2024

Keywords:
Hankel matrix,
Stochastic subspace,
Inverse problem,
Sensitivity analysis,
System identification

Abstract:

The Finite Element Model (FEM) derived from the design drawings may not precisely depict the behavior of the actual structure. This is due to various factors, such as construction variations, uncertainties in boundary conditions, discrepancies in material properties, inaccuracies in FEM discretization, uncertainties in external excitations, and more. Hence, this paper proposes a process that employs stochastic subspace identification (SSI) to estimate the modal parameters of the structural system with minimal user-defined parameters from ambient vibration data, which are then used to update the FEM. First, the optimal dimensions of the matrix with minimal noise errors are determined by analyzing the condition number of the Hankel matrix. Next, the models are filtered to remove modes caused by numerical instabilities resulting from over-determination in the system. Finally, selecting structural modes involves utilizing Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering and confirming the complexity of the shape modes. The algorithm was tested on a numerical model of a 2D concrete frame and used to analyze ambient vibration data from a 6-story building. The first five modes of the residential building with irregular plans were extracted successfully. Thus, the first two modes of the structure have a difference of less than 15%, and the other three modes have a 95% agreement with the results of the updated finite element model. It is important to note that the initial FEM did not accurately represent seismic behavior due to the used concrete strength.

1. Introduction

Reinforced concrete (RC)buildings are a common type of residential structure. However, the models used to analyze their seismic behavior need improvement due to infill wall effects, foundation compliance, and building material quality[1]. Thus, updating the numerical models of these structures based on experimental evaluation of vibration modes is essential. Ambient vibration tests have become famous for extracting these modal parameters, and many studies have explored output-only techniques for identifying the characteristics of RC buildings[2]. One notable study conducted by Mirteheri et al.[3] used environmental vibration tests to determine the natural frequency of 40 concrete and steel structures. The findings of this study uncovered significant discrepancies between numerical analysis and codes and standards. Another study by CLIK et al. used the analytical signal method to analyze an RC building that was subsequently reinforced with cast-in-place reinforced concrete infill shear walls near the North Anatolian fault. This method can quickly and accurately identify mode shapes, frequencies, and damping ratios from forced vibration testing[4].

One of the most popular output-only methods is the stochastic subspace method (SSI)[5]. The SSI approach estimates a linear, time-invariant system model from output-only data. This technique relies on projecting a matrix of recent data onto the subspace formed by the rows of a matrix containing the minor current data[6]. One of the primary reasons behind the popularity of SSI is its ability to offer high computational robustness and efficiency, along with accurate estimations[7]. Brownjohn conducted vibration monitoring on a 280-meter office tower and investigated ambient vibrations in a smaller building using SSI methods. This research has shown that the SSI procedure is a powerful tool

**Corresponding author: Assistant Professor, Department of Civil Engineering, Sarab Branch, Islamic Azad University, Sarab, Iran. Email: pourgholi@iausa.ac.ir*
***Department of Civil Engineering, Sarab Branch, Islamic Azad University, Sarab, Iran. Email: saied.mahdavi123@gmail.com*

https://doi.org/10.52547/NMCE.718.1

that works well for ambient and forced vibration testing, with some limitations[8]. Babak Moaveni et al. extracted the modal characteristics of a full-scale 7-story reinforced concrete building slice simulated on a seismic table with several output-only and input-output methods using SSI algorithms. Based on the results, the modal parameters identified in various ways are consistent. This suggests that the estimated modal parameters are likely a close approximation of the actual modal parameters of the building specimen[9].

However, one of the main challenges with SSI methods is their reliance on user-defined parameters, such as the number of block rows and columns in the Hankel matrix, the system's order, and the validation process for representative quantities[10, 11]. Priori et al. reduced the modal characteristic estimation error in SSI data by analyzing the noises in the past and future subspaces of the Hankel matrix. They presented the asymmetric dimensions of the matrix for this purpose. The method suggested has been effectively executed on seismic data obtained from a 4-story steel frame structure in a laboratory and environmental data collected from San Camillo de Lellis Hospital in Italy[12]. Li et al. conducted a parametric study on the SSI-COV method for concrete dams using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering. Through this study, they could extract the optimum user-defined parameters, such as the number of block rows and columns of the Hankel matrix. The proposed method was tested and validated using two finite element models of the Dagangshan dam and environmental data from the Xiluodu dam[13].

This research aims to determine the optimum size for the Hankel's matrix to identify the dynamic characteristics of RC building through process inversion. First, the optimal dimensions of the matrix with minimal noise errors are determined by analyzing the condition number of the Hankel matrix. Next, the models are filtered to remove modes caused by numerical instabilities resulting from over-determination in the system. Finally, selecting structural modes involves utilizing DBSCAN clustering and confirming the shape modes' complexity. The algorithm was tested on a numerical model of a 2D concrete frame and used to analyze ambient vibration data from a 6-story building. The resulting modal characteristics were used to update the building's numerical model.

2. Methodology

The structural modal properties, including modal frequencies, damping ratios, and shape modes, are accurately determined using SI techniques in our proposed strategy. However, the accuracy of the estimated system is mainly dependent on user-defined parameters, such as the dimensions of the Hankel matrix and the order of the predicted system. Critical structural modes can be omitted due to insufficient system information by a Hankel matrix with small dimensions. Conversely, a large matrix can increase computational costs and lead to estimated biases of physical poles through splitting phenomena and mathematical modes. Additionally, the order of the system is crucial in system identification, as numerical instability can be caused, and mathematical poles can be resulted in the model by overestimating the order. Therefore, a crucial aspect of SI algorithms is correctly selecting user-defined parameters.

This research presents a four-step method to ensure the reliability of identified systems. The steps are displayed in Figure 1 and include: 1) Using weight matrices based on CCA and CVA to extract state matrices via the SI algorithm. 2) Extracting the optimal dimensions of the Hankel matrix to avoid the magnification of noise effects in estimated systems. 3) Employ clustering analysis to eliminate noise modes and filter out mathematical modes, thereby mitigating the impact of numerical instabilities in estimated systems. 4) Selecting representative modal quantities by clustering the result of damping ratios and verifying their proportionality assumption.

2.1 Stochastic Subspace Identification

Innovation state space formulation works for all linear and time-invariant modal identification techniques: [14]:

$$x_{k+1} = \mathrm{A}x_k + \omega_k \quad (1.a)$$

$$y_k = \mathrm{C}x_k + \vartheta_k \quad (1.b)$$

data, $\mathrm{A} \in \mathbb{R}^{n\times n}$ and $\mathrm{C} \in \mathbb{R}^{m\times n}$ are the system matrices, and where $x \in \mathbb{R}^n$ is a discrete stochastic vector sampled in Δt intervals, $y \in \mathbb{R}^m$ denotes a vector of measured or output $\omega_k \in \mathbb{R}^n$ and $\vartheta_k \in \mathbb{R}^m$ represent the zero-mean white noise with the following covariance matrices [15]:

$$E\left\{\begin{bmatrix}\omega_k\\ \vartheta_k\end{bmatrix}\begin{bmatrix}\omega_k{}^T & \vartheta_k{}^T\end{bmatrix}\right\} = \begin{bmatrix}\mathrm{Q} & \mathrm{S}\\ \mathrm{S}^T & \mathrm{R}\end{bmatrix}\delta_{pq} \quad (2)$$

where $E\{\dots\}$ shows the mathematical expectation, δ_{pq} is the Kronecker delta, $\mathrm{Q} \in \mathbb{R}^{n\times n}$ denotes a definite non-negative matrix, and $\mathrm{R} \in \mathbb{R}^{\mathrm{m}\times\mathrm{m}}$ represents a definite positive matrix. If $y_t \in \mathbb{R}^{m\times s}$ are the measured data in the ambient vibration test, the data space can be defined as a Hankel matrix [16]:

$$\mathrm{H} = \left[\begin{array}{cccc} y_0 & y_1 & \cdots & y_{N-1} \\ y_1 & y_2 & \cdots & y_N \\ \vdots & \vdots & \ddots & \vdots \\ y_{k-1} & y_k & \cdots & y_{k+N-1} \\ \hline y_k & y_{k+1} & \cdots & y_{k+N} \\ y_{k+1} & y_{k+2} & \cdots & y_{k+N+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{2k-1} & y_{2k} & \cdots & y_{N+2k-2} \end{array}\right] = \left(\frac{\mathrm{Y}_p \in \mathbb{R}^{k.m\times j}}{\mathrm{Y}_f \in \mathbb{R}^{k.m\times j}}\right) \in \mathbb{R}^{i\times j} \quad (3)$$

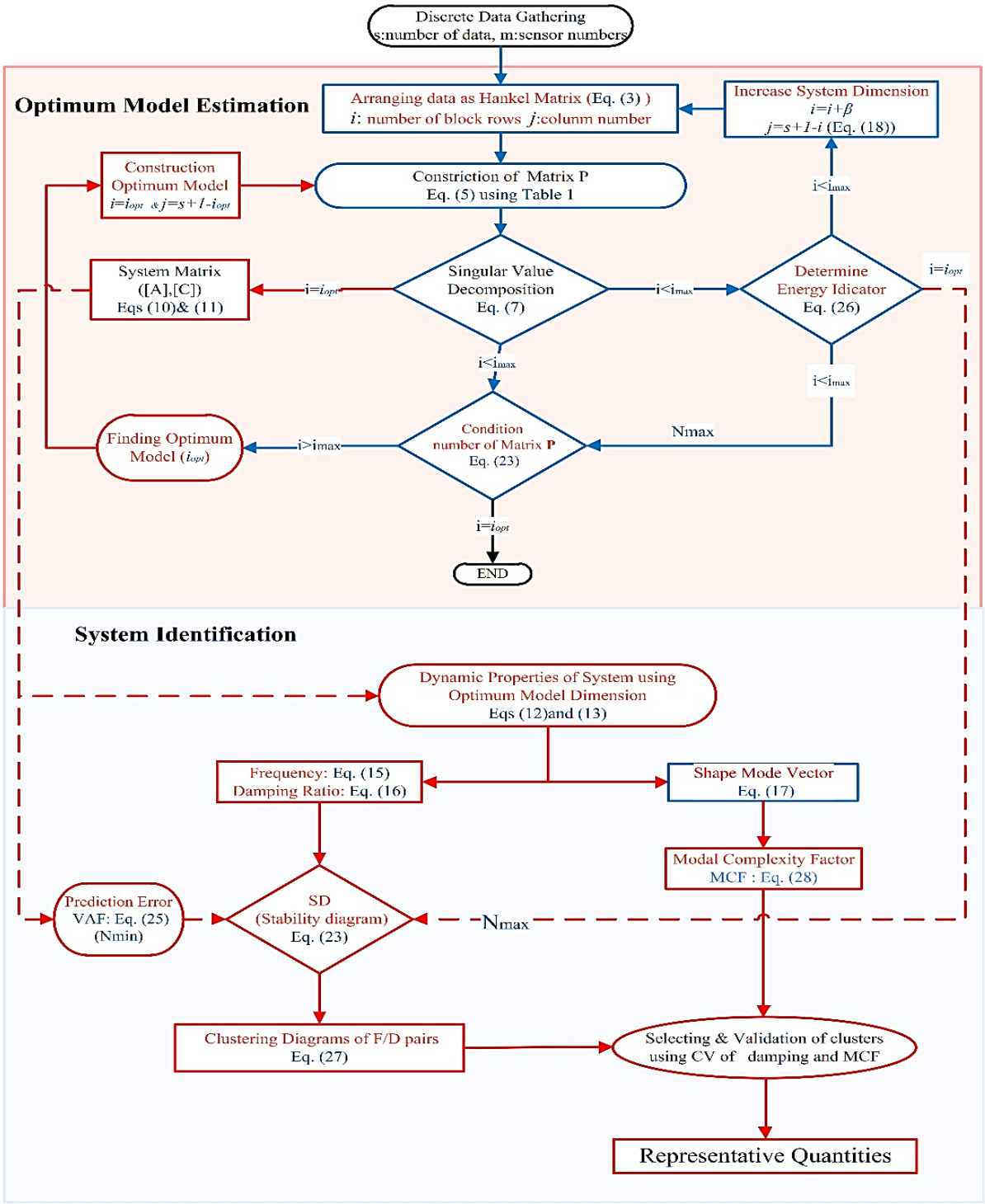


Fig. 1: Flowchart of propped Method

By selecting moment k as the present, Y_p is defined as the past subspace while Y_f stands for the future subspace. The prediction space ($\hat{X} \in \mathbb{R}^{k.m \times N}$) can be calculated by the conditional mean of future space over the past space [17].

$$\hat{X} = \hat{E}\{Y_f | Y_p\} = HT_{-}^{-1} Y_p \quad (4)$$

where $T_{-}$and T_{+} are the covariance of past and future matrices, respectively. To reduce the computational noise and better detect the true modes, Equation5 is multiplied by the weight matrices of W_1 and W_2.

$$P = W_1 \hat{X} W_2 \quad (5)$$

where $P \in \mathbb{R}^{k.m \times k.m}$, $W_1 \in \mathbb{R}^{k.m \times k.m}$, and $W_2 \in \mathbb{R}^{N \times k.m}$. Depending on the applied functions for W_1 and W_2, different algorithms can be defined for stochastic subspace methods. This study employs two algorithms, namely Canonical Correlation Analysis (CCA) and Conventional Variable Analysis (CVA), to identify the system. SSI-CVA helps detect modes with significant energy differences, making it ideal for identifying weak excitation modes. The SSI-CCA

extracts orthonormal basis vectors from the data space to find the optimum space. This technique estimates mode ranges that are highly correlated with the measured data and minimize the impact of noise on the identified system. Refer to Table 1 for the weighting matrices associated with both methods.

According to the Kalman filter, the matrix of **P** can be decomposed into the observability matrix ($\mathbf{O}_k$) and the state vector ($\hat{x}_k$) as follows:

$$\mathrm{P} = \begin{bmatrix} \mathrm{C} \\ \mathrm{CA} \\ \vdots \\ \mathrm{CA}^{k-1} \end{bmatrix} [\hat{x}_k \quad \hat{x}_{k+1} \quad \dots \quad \hat{x}_{k+N-1}] = \mathrm{O}_k \hat{x}_k \tag{6}$$

Table1: Weight matrices of stochastic subspace algorithms

Method	$\mathbf{W}_1$	$\mathbf{W}_2$
CVA	$\mathrm{T}_{+}^{-1/2}$	$\mathrm{I}_{N\times km}$
CCA	$\mathrm{T}_{-}^{-1/2}$	$\mathrm{T}_{+}^{-T/2}$

For this purpose, using the singular values decomposition method, we have:

$$\mathrm{P} = \mathrm{U}\Sigma\mathrm{V}^T \approx \hat{\mathrm{U}}\hat{\Sigma}\hat{\mathrm{V}}^T \tag{7}$$

where $\mathrm{U} \in \mathbb{R}^{k.m\times k.m}$ and $\mathrm{V} \in \mathbb{R}^{k.m\times k.m}$ are orthogonal matrices ($\mathrm{UU}^T = \mathrm{I}_{k.m}, \mathrm{VV}^T = \mathrm{I}_{k.m}$), and $\Sigma \in \mathbb{R}^{k.m\times k.m}$ represents the diagonal matrices of the eigenvalues. $\hat{\Sigma} \in \mathbb{R}^{n\times n}$ can be obtained by ignoring the small values of Σ so that the dimension of the system (n) will be equivalent to the order of $\hat{\Sigma}$. Finally, by comparing Eqs. (8) and (9), the state vector can be obtained as follows:

$$\mathrm{O}_k = \hat{\mathrm{U}}\hat{\Sigma}^{1/2} \tag{8}$$

$$\hat{x}_k = \hat{\Sigma}^{1/2}\hat{V}^T \tag{9}$$

The system matrix, A, can be calculated using Equation (2.8) by removing the first block on the left and the last block on the right as follows:

$$\mathrm{O}_k(2{:}k)\mathrm{A} = \mathrm{O}_k(1{:}k-1) \tag{10}$$

The observation matrix, C, is also the first block of the observability matrix:

$$\mathrm{C} = \mathrm{O}_k(1{:}1) \tag{11}$$

The modal analysis begins by eigenvalues decomposition of the system matrix as follows [18]:

$$\mathrm{A} = \psi[\mu_i]\psi^{-1} \tag{12}$$

$$\lambda_i^c = \frac{ln(\mu_i)}{\Delta t} \tag{13}$$

where ψ and μ_i are the vectors of the eigenvalues and the singular values of the system matrix, respectively and λ_i^c represents the continuous-time poles of the system matrix.

The continuous-time poles are of the complex form containing the frequency (f_i) and damping ratio (ξ_i) of the system as follow:

$$\omega_i = |\lambda_i^c| \tag{14}$$

$$f_i = \frac{\omega_i}{2\pi} \tag{15}$$

$$\xi_i = \frac{Re\{\lambda_i^c\}}{|\lambda_i^c|} \tag{16}$$

Finally, the mode shapes at the sensor installation location can be obtained from Equation (17):

$$\Phi = \mathrm{C}\Psi \tag{17}$$

2.2. Optimum dimension of Hankel Matrix

The effective parameters of the Hankel matrix include the number of block rows (*i*) and its columns (*j*) which are related to each other as follows [19]:

$$j + i - 1 \le s \tag{18}$$

where s is the sampled data count. Hankel matrix has two sub-matrices: past and future[13]. According to stochastic subspace theory, the dimensions of future and past submatrices must be greater or equal to the order of the system [5]:

$$k.m \ge n \tag{19}$$

To ensure the coverage of all system frequencies, especially in high ductility structures, in addition to Equation (19), the following conditions are proposed for the minimum dimension of the system [20, 21]:

$$k \ge \beta \tag{20}$$

where $\beta = \frac{f_s}{2f_o}$ is the lower bound of parameter *k*, f_o and f_s is the lowest frequency and the data sampling frequency, respectively.

In practice due to the limited length of the gathered data, QR [22] or LQ [23] decomposition is used to estimate the Hankel matrix, thus, the maximum value of parameter *j* can be obtained as:

$$j \le 2km \tag{21}$$

Comparing Eqs (18) and (21), the upper bound of parameter *k*, can be determined by:

$$k \le \frac{s+1}{2(m+1)} \tag{22}$$

After determining the range of parameter k, the number of block rows of the Hankel matrix is determined from the relationship$i = 2k$. Also, the parameter *j* is typically chosen according to Equation (18), once the value of i is selected. Large i values improve system information handling but increase computational effort. However, they can cause inaccuracies and numerical hypothetical modes in the presence of measurement noise. The occurrence of numerical error in the identified systems can be tracked by the transmission of the matrix P from invertible to non-invertible. For this, the condition number of matrix P, which is the ratio between the maximum (σ_1) and minimum (σ_m) singular values of matrix P can be used:

$$\bar{k}_m = \frac{|\sigma_1|}{|\sigma_m|} \tag{23}$$

The number of singular values between the maximum absolute value and the minimum acceptable singular value indicates the amount of available information, so selecting the appropriate number of block rows and corresponding condition number can ensure numerical stability. Conducting

a sensitivity analysis on the number of block rows i helps determine the optimal value for numerical stability [24].

2.3. System order estimation

As previously mentioned, determining the system's actual order is crucial in system identification. This is accomplished using two tools:

a) Stabilization diagram (SD): Real-order (n) extraction is one of the most important parts of system identification. In practice, uncertainties due to environmental noise and measurement, as well as the improper selection of the dimensions of the Hankel matrix, introduce unstable and computational poles in the system. To detect the real poles, all the dynamic characteristics of the poles, including frequency and damping percentage, are extracted from the order of n_{mim} to n_{max}. The pole satisfying the following conditions is considered stable and plotted in a SD [25]:

$$\frac{f_i^n - f_i^{n-1}}{f_i^{n-1}} < 0.01 \quad , \quad \frac{\xi_i^n - \xi_i^{n-1}}{\xi_i^{n-1}} < 0.05 \tag{24}$$

Therefore, limiting model order n within the range of $[n_{min} \quad n_{max}]$ is essential to generate a clear and high-quality stabilization diagram. To find the n_{min}, the concept of system identification, which is based on fitting of an assumed mathematical model to the measurements in order to extract the unknown model parameters, is used.

The $n_{min}$is determined by adapting the measured data to an assumed mathematical model. For this purpose, the variance accounting for (VAF) criterion was used [26]:

$$VAF = Max\left(0, \left(1 - \frac{\frac{1}{N}\sum_{k=1}^{N}(y_k - \widehat{y_k})^2}{\frac{1}{N}\sum_{k=1}^{N}(y_k)^2}\right) \times 100\%\right) \tag{25}$$

where y_k and $\widehat{y_k}$ are the measured data and estimated values, respectively. The closer this criterion to 100, the lower the prediction error and the more accurate the model. The order at which VAF begins converging provides an estimate of the number of the real poles of the systemic which can be used as an estimation of n_{min}in SD.

When the number of orders in the system gets too high, the matrix P's condition number increases towards infinity. This makes it impossible to invert, leading to substantial errors in the solution due to numerical inaccuracies and magnified errors, and is known as an ill-conditioned matrix. However, this issue can be controlled by only keeping the most significant singular values. As a result, the maximum order of the system (n_{max}) in the SD is restricted to the rank of P[18, 27]. To determine n_{max}, one method is to examine the $\bar{k}_m$ diagram and locate its inflection point. Nonetheless, in practical engineering, finding this inflection point can be challenging. The inflection point is hard for finding for practice engineering. Therefore, Ref. [20]'s energy indicator is employed to calculate n_{max}. Definition of energy indicator:

$$EI = \frac{\sum_{i=1}^{m}\sigma_i^2}{\sum_{i=1}^{no}\sigma_i^2} \tag{26}$$

where σ_i is the i-th singular value of the matrix Hn. Nmax is determined as the lowest integer such that $I > 99$ %.

b) Clustering analysis of the frequency and damping ratio (F/D pairs): To ensure accuracy and eliminate spurious modes, an outlier removal process may be conducted during or after the clustering stage. During the clustering stage, the stable modes are categorized into sets (clusters) that only have modal parameter estimates for the same structural mode. In this research, the DBSCAN procedure was utilized for clustering. The DBSCAN algorithm relies on two critical parameters, namely $Minpts$, and ε. $Minpts$ Determines the minimum number of points needed in a cluster, while and ε sets the maximum radius for a neighborhood around a point. When dealing with larger datasets or those with high noise levels, it is recommended to use larger values $of\ Minpts$ to create more significant clusters. On the other hand, selecting an improper value for and ε can lead to a significant portion of data being left unclustered. To calculate Minpts, an empirical relationship using $ln(n_p)$ is used, where n_p represents the total number of stable poles. As for ε, it is determined by the knee point of the k-distance diagram, where $k = Minpts$. Based on Trunchi's studies for buildings, using $\varepsilon = 0.01$in research is suggested.

2.4. Selection of modal representative quantities

The coefficient of variation (CV) δ of the clusters identified by DBSCAN is used to select the representative modal properties [28, 29]:

$$\delta = \frac{\sum_{J=1}^{Z}\frac{\sigma_J}{\mu_J}}{Z} \tag{27}$$

where Z represents the total number of identified clusters; σ_J and μ_J are the variance and mean of the J-th cluster, respectively.

It is important to note that damping ratios have a higher CV than other modal characteristics. Therefore, more emphasis will be placed on the damping ratios in choosing representative quantities. Choosing modes with lower CV damping is generally preferred when comparing clusters within the same frequency range.

After selecting representative clusters, the proportionality assumption of damping is used to validate them, which is acceptable for buildings. To assess this assumption, all shape mode components are plotted in a complex plot, where a horizontal plot represents a real-value mode, and a vertical plot represents a complex mode. The complexity of modes is then determined based on the modal complexity factor (MCF) criterion [30]:

$$MCF_r = 1 - \frac{(S_{xx} - S_{yy}) + 4S_{yy}^2}{(S_{xx} + S_{yy})^2} \tag{28}$$

$$\begin{cases} S_{xx} = Re\{\psi_r\}^T Re\{\psi_r\} \\ S_{yy} = Im\{\psi_r\}^T Im\{\psi_r\} \\ S_{xy} = Re\{\psi_r\}^T Im\{\psi_r\} \end{cases}$$

where $Re\{\psi_r\}$ and $Im\{\psi_r\}$ are the real and imaginary parts of the modal vector, respectively. MCF value is between 0 and 1; for real-value modes, this value should be close to zero.

3. Validation of proposed method

To test the effectiveness of the suggested approach, we constructed a 2D concrete frame with two close and weak modes (regular and irregular mass and stiffness distribution, respectively) subjected to the Electro earthquake's acceleration. The equation that governs the dynamics of moving systems is differential and can be written as[31, 32]:

$$\ddot{U}(t) + 2\omega_i\xi_i\dot{U}(t) + \omega_i^2 U(t) = -\begin{Bmatrix} m_1 \\ m_2 \\ m_3 \end{Bmatrix} L_i\ddot{u}_g(t) \quad (29)$$

where $\ddot{u}_g(t)$ is recorded acceleration of ground motion during El-Centro earthquake, M and K are the mass and stiffness matrices of the structure, respectively, defined as follows:

$$\mathrm{M} = m\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathrm{K} = k\begin{bmatrix} 4 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \quad (30)$$

where m=10 ton and k=1200 KN/m. Also, Cauchy damping has been adopted for both models. The exact values of modal characteristics for the considered case studies are presented in Table 2.

Using the Modal analysis technique in the Simulink MATLAB module, the 3-DOF system was modeled and analyzed (Fig2). The Fourth-order Runge-Kutta method was used to determine the damped response of the system. The time increment was modified to match the recording interval of 60 seconds and the sampling rate of 100 Hz. For example, the acceleration record of mass two can be found in Figure 3.

According to the frequency of the first mode (1.856 Hz) and the length of the data (6000), the minimum dimension is chosen as $\beta \approx 27$ based on Equation (20). From Equation (23), the maximum system dimension is $k_{max} \approx 750$. First, the

Table 2: The modal features of 3-DOF systems determined by the numerical model.

Mode No.	Frequency (Hz)	Damping (%)
1	1.856	5.00
2	3.190	3.00
3	3.711	2.00

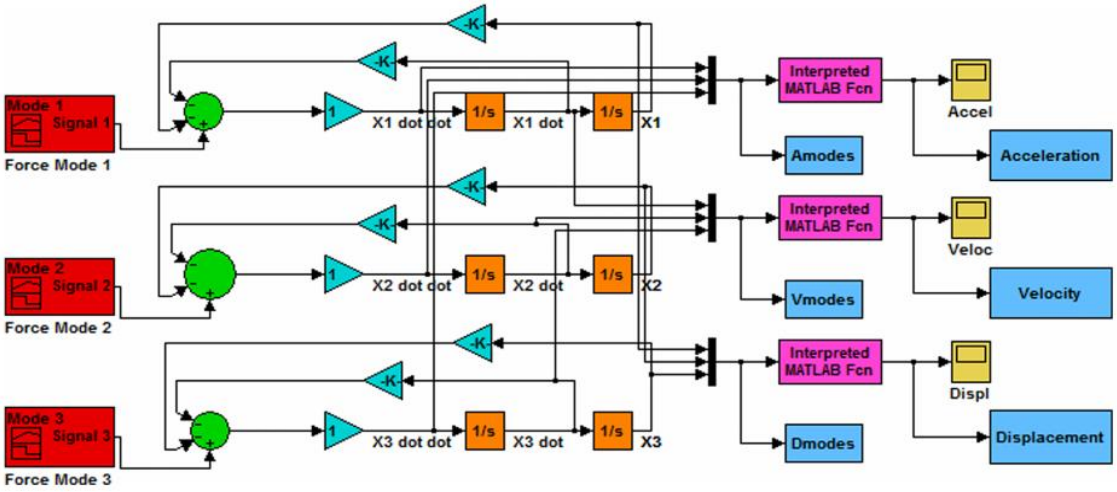


Fig. 2: Simulink model of the analytical 3DOF system[32].

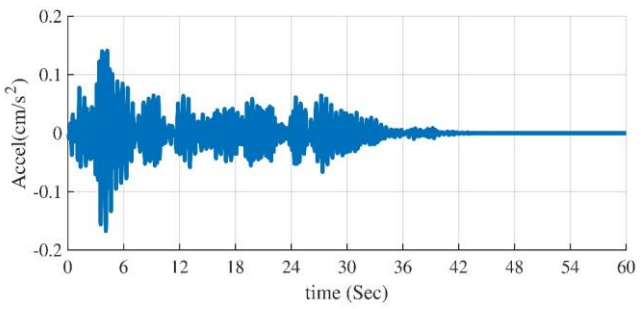


Fig. 3: Acceleration record of system 3DOF at mass 2 location under El-Centro excitation.

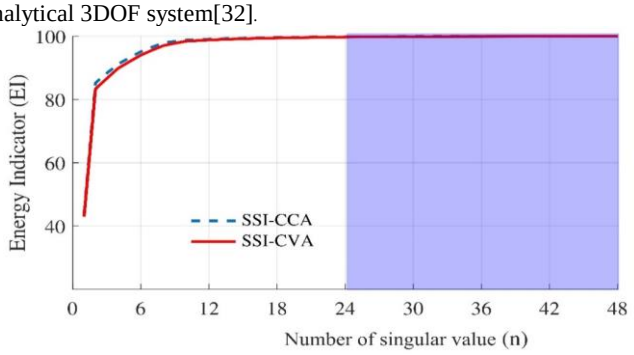


Fig. 4: Variations of energy indicator vs. the number of singular values of the matrix P.

maximum order (N_{max}^{i}) for each value of k from β to 15β is calculated. For example, Figure 4 illustrates the extraction of $N_{max}$for the 5β in SSI-CCA and SSI-CVA.

Once the N_{max}^{i} value has been determined, the $\bar{k}_m$ values are evaluated for each and plotted in Figure 5. From the fifth cycle onwards, it is clear that there is minimal variation in $\bar{k}_m$ for both methods. This indicates that the numerical stability of the models has been achieved, and any errors caused by noise have been minimized. Therefore, the modal characteristics for $i = 5\beta$ are extracted in both methods.

The first step to identifying the system is to check the predicted systems' estimation error with the VAF criterion. Figure 6 demonstrates that both methods reached 100% convergence from the sixth order (N_{min}). Using EI in Figure 5, N_{max} for the fifth cycle was obtained for both methods, which was 24.

SD has been plotted to the 60th order (highlighted in red) to examine the impact of numerical instabilities. In Figs 7 and 8, both methods identified three modal frequencies when $n < N_{max}$. However, SSI-CVA had more than 5% tolerance damping ratios due to noise errors, resulting in only stable modal frequencies. To conduct a thorough examination of this matter, a clustering analysis of the frequency and damping ratio (F/D) pair is utilized.

Comparing SSI-CVA to SSI-CCA, the first modal damping ratio was 1.94% and 3.86%, respectively, with an estimation error of over double in SSI-CCA. SSI-CCA also had a 28% and 7% estimation error for the second and third damping

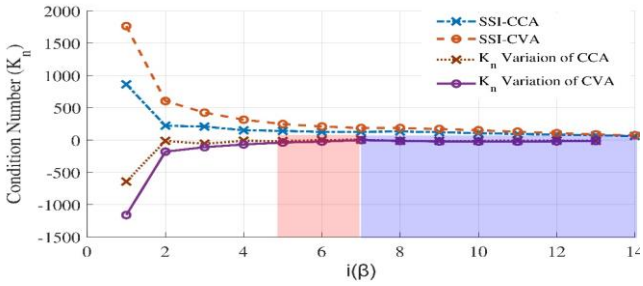
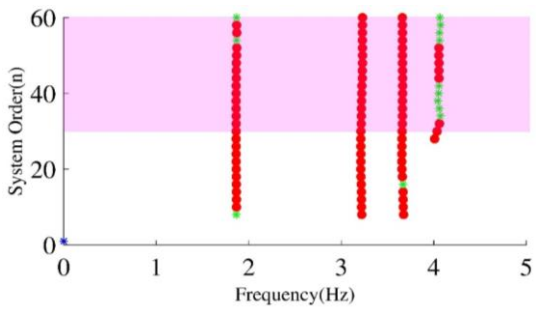


Fig. 5: Condition number vs. *i* of analytical 3-DoF system for El-Centro case.



(a)

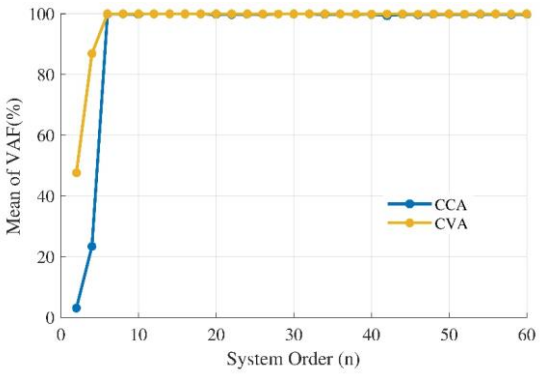
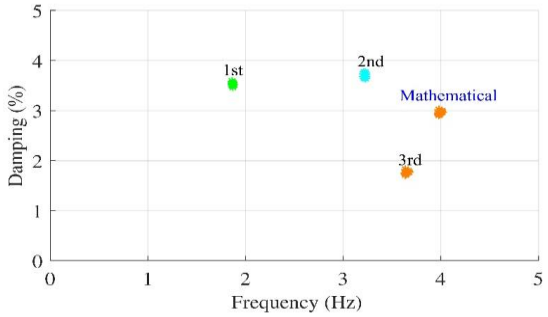


Fig. 6: Variability of average VAF vs. System order for 3-DoF system

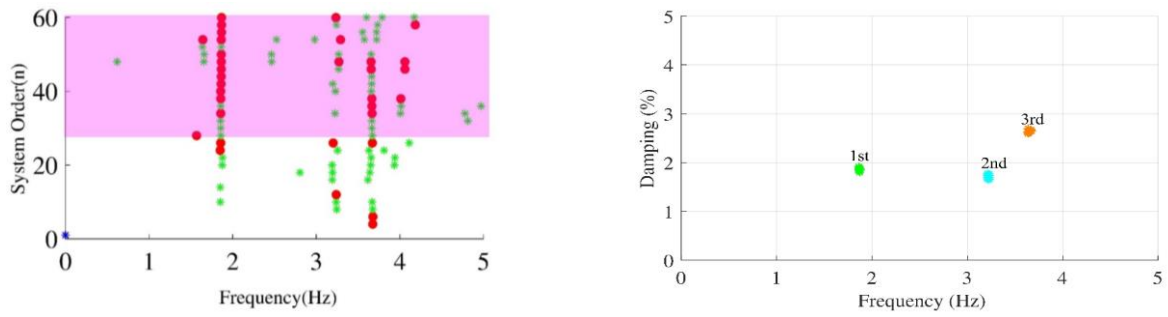
-ratios, while SSI-CVA had a 39% estimation error for both modes. Figs 7b and 8b illustrate these findings.

The CV of damping rations and the complexity of the shape modes are reanalyzed for a more thorough investigation into the reasons for errors in estimating damping ratios. According to Figure 9, the CV for damping ratios in SSI-CCA is not significant. However, in SSI-CVA, this value ranges between 4% and 8%. The increase in these values indicates a greater dispersion of damping ratios in CVA, leading to an increase in estimation errors. Although the CV in SSI-CCA is minimal, the damping ratio estimation error is notably high, particularly in the first mode. Analyzing the mode shapes' complexity depicted in Figure 10 can provide insight into this problem. The first mode has the highest level of complexity in both approaches, with CVA and CCA having 27% and 13.4%, respectively. Also, the third mode of CCA corresponds to the lowest MCF value, and its estimation error is minimal. So, using MCF and CV criteria together can be reliable for validating identified modal characteristics. If the value of *n* exceeds N_{max}, the SSI-CCA exhibits a modal frequency of 4.01 Hz in SD, as depicted in Figure 7. This mode's CV is considerably higher than other modes, and its complexity exceeds 40%, indicating that it is non-structural. Essentially, this mode results from numerical noise amplification due to the over-determination and ill-conditioning of the matrix P. Table 3 displays a summary of the system identification results



(b)

Fig. 7: Parameter identification results of 3-DoF system from El-Centro earthquake excitation based on SSI-CCA (a) Stability diagram (Red circle in the diagrams represent fully-stabilized solutions, green stars indicate only Frequency stable solutions.) and (b) Clustering diagrams of F/D pairs for $k = 5\beta$.



(a) (b)

Fig. 8: Parameter identification results of 3-DoF system from El-Centro earthquake excitation based on SSI-CVA. (a) Stability diagram (Red circle in the diagrams represent fully-stabilized solutions, green stars indicate only Frequency stable solutions.) and (b) Clustering diagrams of F/D pairs for 3-DoF for $k = 5\beta$.

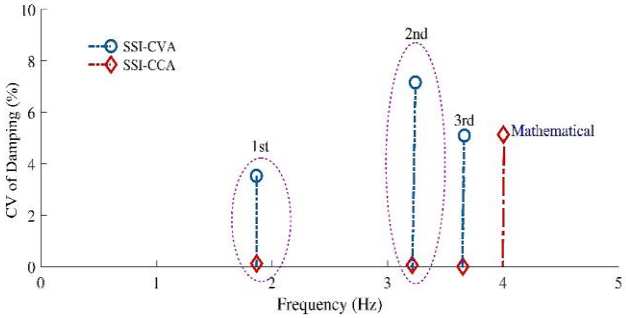


Fig. 9: CV of damping of 3-DoF system from El-Centro earthquake excitation for $k = 5\beta$.

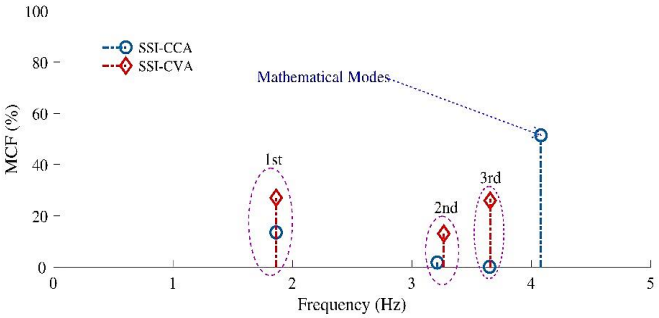


Fig. 10: Modal complexity factor of 3-DoF system from El-Centro earthquake excitation for $k = 5\beta$.

Table 3: Results of system identification of 3-DoF system for $i = 5\,\beta$.

Mode	Frequency (Hz)		Damping (%)		MCF (%)		CV (%)	
No.	CCA	CVA	CCA	CVA	CCA	CVA	CCA	CVA
1	1.863	1.864	3.56	1.94	13.52	27.03	0.18	3.53
2	3.212	3.267	3.86	1.81	1.75	13.01	0.06	7.98
3	3.65	3.655	1.809	2.77	0.08	0.22	0.00	5.14

4. In-situ test

4.1. Description of the structure

In this research, a residential building in Ardabil City, Iran, has undergone environmental vibration testing (Figure 11). This reinforced concrete building, one of Iran's most common structures, is 12 x 9.6 square meters. As shown in Figure 12, the building's plan is irregular due to its columns and roof slab arrangement. The height of the building from the ground level is 19.90 meters. The story height is 3.20 meters, except for the first floor, which measures 3 meters. The building's flooring system consists of joists and blocks covered by a layer of lightweight concrete designed to embed and conceal the pipes. The interior partition walls are constructed using hollow clay bricks, while the exterior walls comprise a layer of hollow clay bricks inside and stone outside. The building began construction in 2021 and is expected to be completed by winter 2023.

The FEM of the foundation and building was developed using the computer program SAP 2000 version 14.2.2. This program can perform linear, non-linear, static, and dynamic analyses of three-dimensional structural modeling [9]. The aim was to utilize the program to determine the fundamental

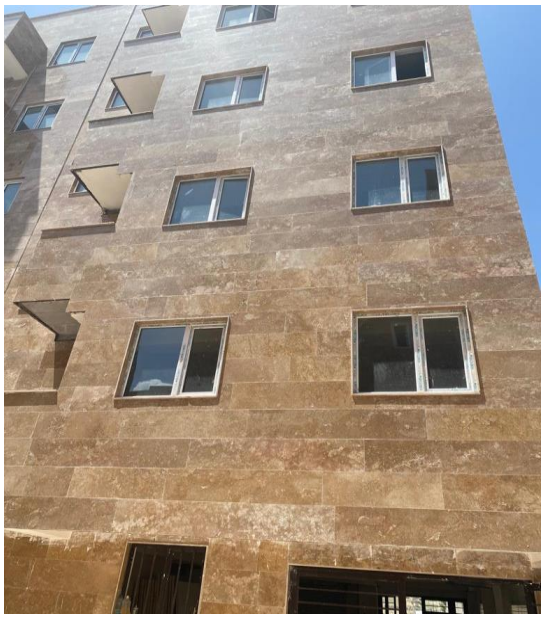


Fig. 11: View of tested six-story building.

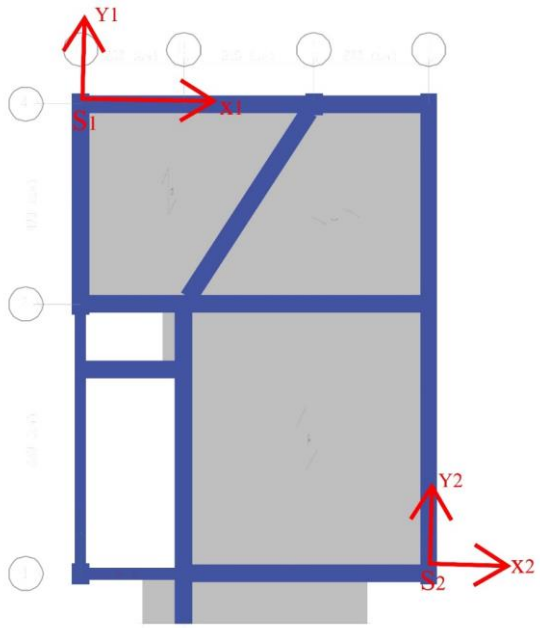


Fig. 12: Plan of Second Floor with Sensor Placement for tasted six-story building.

frequencies and corresponding mode shapes of the structure based on its physical properties. The FEM model comprises 5221 shell elements and 693 frame elements. The properties of the reinforced concrete structure, assumed to be homogeneous and isotropic, are listed in Table 4 for the structural simulation. The building was designed by the American Institute concrete design code (ACI318-99). Figure 13 displays the structure's shape mods and corresponding first five modal frequencies.

4.2. Ambient vibration tests

On June 14, 2023, a sunny and calm day, an environmental vibration test was conducted on the building. The irregularity of the building's plan necessitated a specific arrangement of the sensors. So, to effectively extract the torsional modes of the structure, 12 accelerometers were used (Figure 14). The first through fifth floors were equipped with two accelerometers diagonally placed in the N-W and S-E directions, while the foundation level and sixth floor had one sensor in the N-W direction (Figure 12). After installing all the sensors, the test was conducted for 25 minutes. A truck mixer enhanced the structure's vibration and excited higher modes.

Table 4: Material properties of reinforced concrete

Material properties	Concrete	Steel
Modulus of Elasticity (*Gpa*)	27	200
Compressive Strength (*Mpa*)	25	-
Yield stress (*Mpa*)	-	240
Poisson's ratio (%)	0.2	0.3

Fig. 14: Ambient vibration instruments.

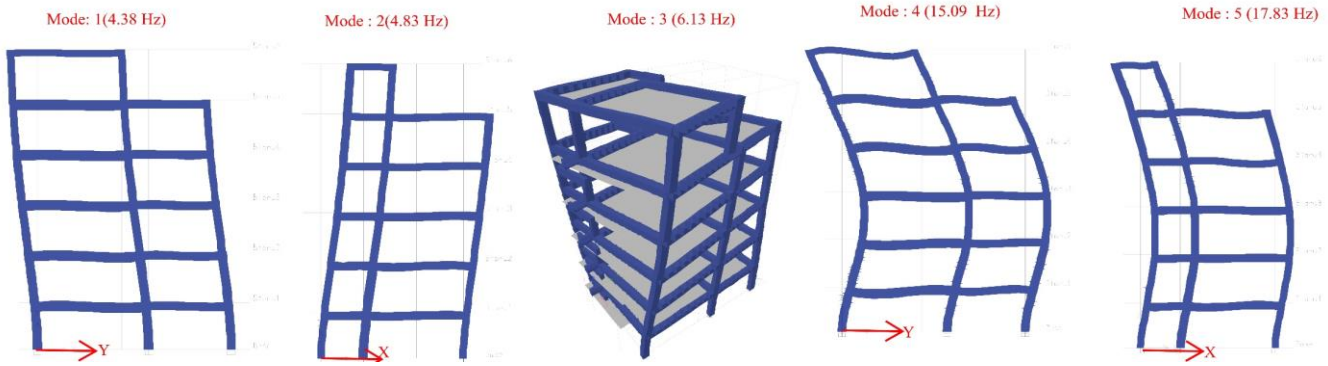


Fig. 13: 3D representation of the first five shape modes of tasted six-story building.

It kept moving back and forth near the building. Additionally, the design of multiple speed bumps near the structure resulted in more robust vibrations. Accelerometers simultaneously record data in three directions: North-South, West-East, and vertical, with a sampling time interval of 0.005 seconds. Based on Figure 15, the maximum response during the 9–11-minute period was recorded (highlighted in red) and was used to identify the structural system. The system identification process was simultaneously conducted for the N-S and W-E directions.

Based on the numerical model findings, the weighting matrices for SSI-CCA were selected to identify the building system with different values of the parameter i. Finite element analysis reveals that the lowest frequency, f_0, is 4.38 Hz. Therefore, the minimum k value guarantees that one cycle of the lowest frequency is 23, as per Equation (20). To select the desired values of parameter k, in the first step, N_{max} is evaluated for each k value from β to 15β using EI, and then the corresponding $\bar{k}_m$values are plotted in Figure 16. The diagram includes variations of $\bar{k}_m$ versus i to

provide a more detailed analysis. When the k value is increased, it becomes evident that the inversion problem transitions from ill-conditioned to well-posed. Additionally, the plot shows that for$k > 6\beta$ (denoted by the blue area), the changes of $\bar{k}_m$ are minimized, indicating that a high-quality model should be achieved. It is essential to remember that when the i parameter increases too much after the convergence of $\bar{k}_m$, it leads to an increased truncated error of SVD. To ensure high-quality models, it is necessary to check the initial i values related to the convergence of $\bar{k}_m$. In order to achieve the desired model, the proposed method extracts the modal characteristics of the system equivalent to $k = 6\beta, 7\beta$ and 8β. A method for working with $k = 7\beta$ will now be presented based on nearly identical outcomes.

After the CCA algorithm was conducted, the VAF diagram was created to display the estimated error of the identified system's response to the measured responses. As shown in Figure 17, 86% convergence was reached by the VAF from order 42 (displayed in blue).

According the FEM results, it was determined that the initial five modes of the structure were present within the frequency range of [0-25] Hz. Consequently, the SD was plotted within this range to identify and extract stable poles. The N_{min} in SD is set to 42 based on VAF, and the N_{max} is set to 118, according to EI. It should be noted that N_{max} was varied from 118 to 160 (marked in red) to show the numerical instability effects of the identified models.

To distinguish the structural modes from the spurious ones, the DBSCAN cluster analysis was implemented. Figure 19 Figure 18 displays the natural frequencies identified by the SD, with the n range varying from 42 to 160. The frequencies identified varied based on the model order *n*.

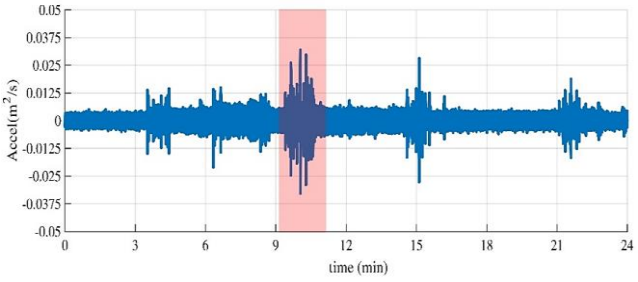


Fig. 15: Acceleration recorded during the ambient vibration test of the six-story building excited by a track mixer.

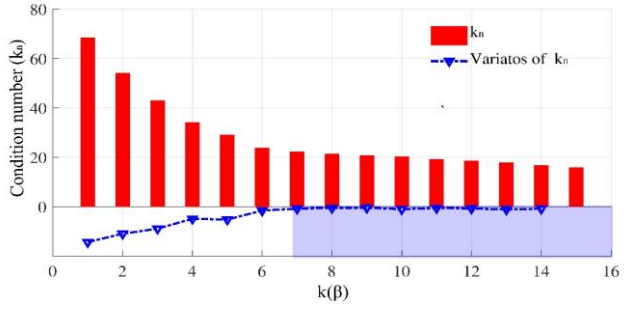


Fig. 16: Condition number vs. *k* of three-dimensional building for ambient vibration case.

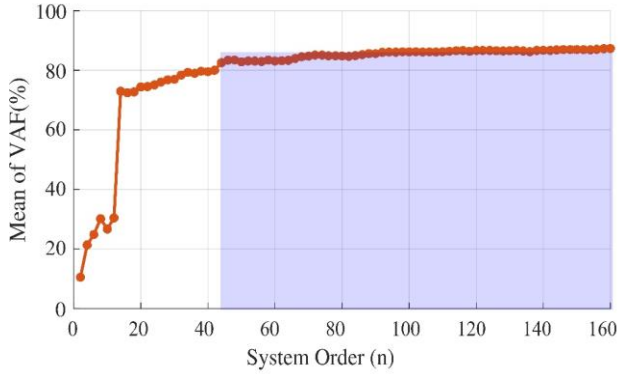


Fig. 17: Variability of average VAF vs. System order for three-dimensional building from ambient vibration.

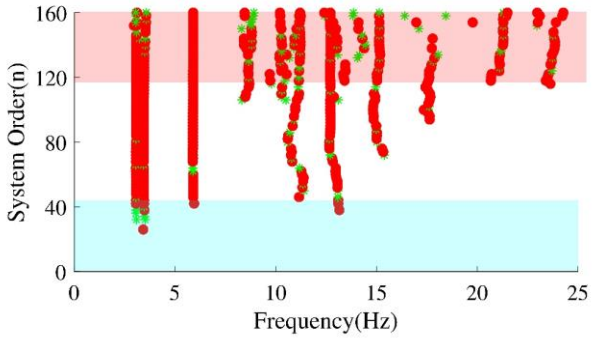


Fig18: Stability diagram (Red circle in the diagrams represent fully-stabilized solutions, green stars indicate only Frequency stable solutions.) of three-dimensional building from ambient vibration for $k = 7\beta$ in SSI-CCA.

When n was less than 118, seven modes were identified. However, when n exceeded 118, an additional four modes were observed in the SD, which can be attributed to the splitting phenomenon. Further research will be studied in the following.

To distinguish the structural modes from the spurious ones, the DBSCAN cluster analysis is implemented. Figure 19 displays the clustering diagrams for the frequency and damping ratio (F/D pairs). For $n < 118$, seven modes were identified, but the frequencies 10.73 and 17.63 Hz are not grouped into clusters based on their damping ratios. So, these frequencies are considered outlier. The damping ratios' CV and mode shapes' complexity were analyzed to validate and determine the structural characteristics. The identified clusters' CV values are shown in Figure 20, which indicates that all five structure clusters have a CV of less than 0.5%. Moreover, the first three modes have zero complexity, which confirms that they are the actual values. Then, the complexity of the frequency of 12.43 and 15.22 Hz is also less than 20%, significantly lower than the complexity of the four modes in the out-of-rank region ($n > 118$). It has been observed that the MCF of the modal shapes identified in n>118 is over 50%. This indicates that these modes may be classified as non-physical due to the numerical uncertainty of the predicted model. In Figs 13 and 21, we can see the

shape modes obtained from FEM and CCA, respectively. The first two modes correspond to the transition modes in the X and Y directions, while the third mode is torsion, which the proposed algorithm accurately extracted. Table 5 displays the system identification results summarized.

As seen in Table 5, the first two fundamental frequencies of the structure differ by more than 30% from the FEM values. This discrepancy is primarily due to the quality of materials, mainly concrete, implementation difficulties, and errors in estimating modal specifications. We developed a FEM model of the structure and foundation system for spectral dynamic analysis to address this issue. Upon reviewing the technical documents of the building, we discovered that the quality of the concrete was a significant concern. Initially, its compressive strength was assumed to be 25 MPa. However, after examining the average resistance of other samples, we found that the average compressive strength was between 20 and 21 MPa. Additionally, the construction site is located near Shorabil Lake, where the soil is saturated, unlike other areas of the city, which must be noted in the geotechnical report. After making changes in the FEM, as shown in Figure 23 and Table 5, the error in estimating two fundamental frequencies of the structure decreased from 35% and 30% to 16% and 12%, respectively. Moreover, three other modal frequencies of the building matched with the results of environmental vibration tests with minimal error. The design engineers and building supervisors were provided a summary of the analysis results to facilitate their preparations.

Table 5: Identified modal properties of tasted six-story building for optimum case of $\boldsymbol{i = 7\beta}$.

SSI-CCA		Finite Element	
Frequency (Hz)	Damping (%)	Frequency (Hz) Initial	Updated
3.23	2.59	4.38 *(35%)*	3.75 *(16%)*
3.64	3.89	4.83 *(32%)*	4.14 *(12%)*
5.78	2.25	6.13 *(6%)*	5.47 *(5%)*
12.43	2.53	15.09 *(21%)*	12.96 *(4%)*
15.22	2.00	17.83 *(17%)*	15.33 *(0%)*

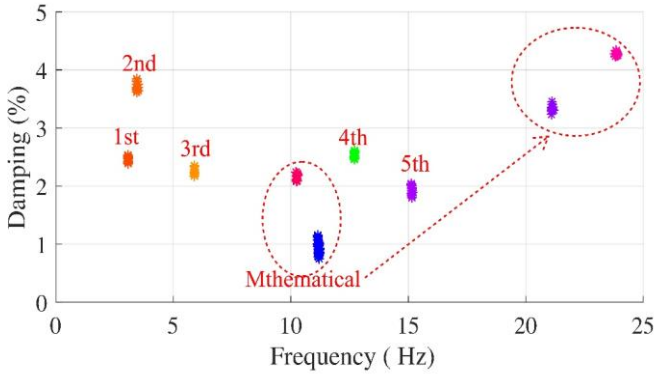


Fig. 19: Clustering diagrams of F/D pairs for three-dimensional building from ambient vibration for $k = 7\beta$ in SSI-CCA

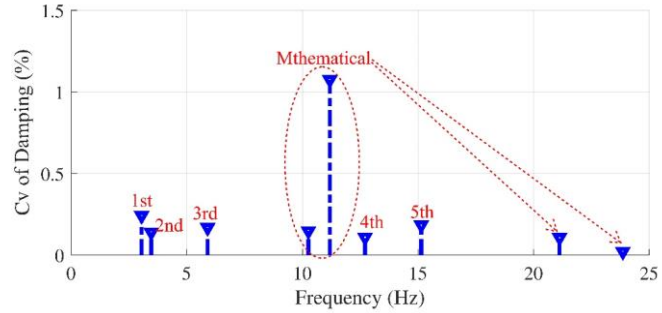


Fig. 20: CV of damping of three-dimensional building from ambient vibration for $k = 5\beta$.

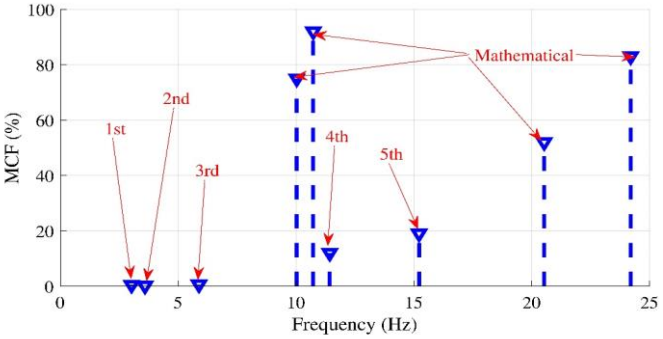


Fig. 21: MCF of three-dimensional building from ambient vibration for $k = 7\beta$.

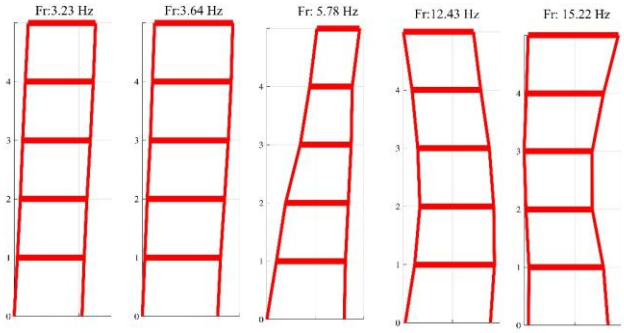


Fig. 22: First Five Shape Modes Identified through SSI-CCA for tasted Six-Story Building

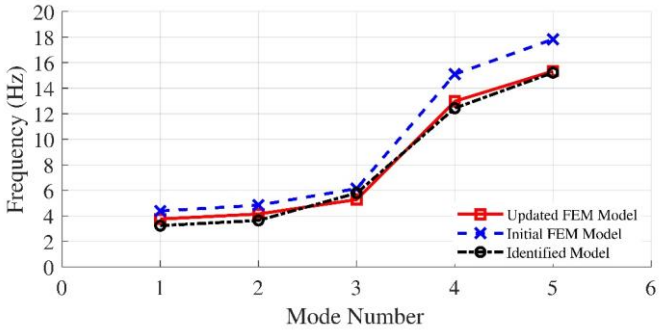


Fig. 23: Comparison of Modal Frequencies from SSI-CCA and Numerical Models.

8. Conclusion

This paper introduces a process based SSI that allows for the estimation of a structural system's modal parameters with minimal input from the user. The ambient vibration data collected from an irregularly planned RC residential building

was tested by our method. The identification phase of the process can be initiated by providing the required parameter ranges. During the initial identification stage, SSI utilizes two algorithms: SSI-CVA and SSI-CCA. These algorithms rely on two crucial parameters for accuracy - the number of block rows and system order. To optimize the number of block rows, an inverse process is applied. At the same time, the maximum and minimum order of the system is determined through the Hankel matrix rank and system estimation error control, respectively. The identified optimal system is analyzed using DBSCAN clustering to extract the characteristics of the structural model. Afterward, representative quantities are selected by examining the complexity of the modal shapes of the clustered characteristics.

Increasing the value of i leads to more stable frequency/damping ratio pairs, resulting in an increased number of cycles for each mode. However, an excessive increase in i raises the calculation costs and increases the SVD trashed error. The presented inversion process is designed to prevent the amplification of the effects caused by ill-conditioning of the systems while increasing i to provide the desired dimensions. The proposed method for validating the structural modal characteristics has succeeded in both algorithms. In the numerical model of the two-dimensional concrete frame, the modal appeared due to the overdetermination of the system was well distinguished by the lack of clustering. Additionally, the MCF criterion effectively revealed the high uncertainty of each of the mathematical modes clustered by examining the complexity of their mode shape. The Energy indication criterion filtered out the effects of numerical instabilities from the systems.

The SSI-CCA algorithm proved more effective than SSI-CVA, resulting in better outcomes. Therefore, it was utilized to identify the 6-story residential building system. Despite the noise and computational modes, the proposed approach successfully identified the first five modal frequencies of the structure, with the first three modal shapes being real-valued. Based on the real values of the first two modes, it was discovered that they differed by over 35% from the modal characteristics of FEM. Technical documents were reviewed, and FEM was updated using the compressive strength of the concrete and the soil saturation. This led to a decrease in the difference between the first two modes to less than 15%, while the frequencies of the other three modes matched over 95% with the updated FEM output. The findings indicate that the suggested approach is versatile and can be implemented with any identification method, not just the two algorithms examined in this research.

References

[1] H.S. Ulusoy, M.Q. Feng, P.J. Fanning, System identification of a building from multiple seismic records, Earthquake Engineering & Structural Dynamics, 40(6) (2011) 661-674.

[2] M. Inel, H.B. Ozmen, B.T. Cayci, Determination of period of RC buildings by the ambient vibration method, Advances in Civil Engineering, 2019 (2019) 1-10.

[3] M. Mirtaheri, F. Salehi, Ambient vibration testing of existing buildings: Experimental, numerical and code provisions, Advances in Mechanical Engineering, 10(4) (2018) 1687814018772718.

[**4]** O.C. Celik, H.P. Gülkan, Processing forced vibration test records of structural systems using the analytic signal, Journal of Vibration and Control, 27(19-20) (2021) 2253-2267.

[5] P. Van Overschee, B. De Moor, N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems, Automatica, 30(1) (1994) 75-93.

[6] E.P. Reynders, Uncertainty quantification in data-driven stochastic subspace identification, Mechanical Systems and Signal Processing, 151 (2021) 107338.

[7] E. Reynders, System identification methods for (operational) modal analysis: review and comparison, Archives of Computational Methods in Engineering, 19 (2012) 51-124.

[8] J.M. Brownjohn, Ambient vibration studies for system identification of tall buildings, Earthquake engineering & structural dynamics, 32(1) (2003) 71-95.

[9] B. Moaveni, X. He, J.P. Conte, J.I. Restrepo, M. Panagiotou, System identification study of a 7-story full-scale building slice tested on the UCSD-NEES shake table, Journal of Structural Engineering, 137(6) (2011) 705-717.

[10] E. Tronci, M. De Angelis, R. Betti, V. Altomare, Multi-stage semi-automated methodology for modal parameters estimation adopting parametric system identification algorithms, Mechanical Systems and Signal Processing, 165 (2022) 108317.

[11] M. Pourgholi, R. Tarinejad, M.E. Khabir, M. Mohammadzadeh Gilarlue, System identification of Karun IV Dam using balanced stochastic subspace algorithm considering the uncertainty of results, Journal of Vibration and Control, (2022) 10775463221133591.

[12] C. Priori, M. De Angelis, R. Betti, On the selection of user-defined parameters in data-driven stochastic subspace identification, Mechanical Systems and Signal Processing, 100 (2018) 501-523.

[13] S. Li, J.-T. Wang, A.-Y. Jin, G.-H. Luo, Parametric analysis of SSI algorithm in modal identification of high arch dams, Soil Dynamics and Earthquake Engineering, 129 (2020) 105929.

[14] T. Katayama, Subspace-Based System Identification-A View from Realization Theory, Systems, Control and Information Engineers, 41 (1997) 380-387.

[15] E.J. Hannan, M. Deistler, The statistical theory of linear systems, SIAM, 2012.

[16] T. Katayama, H. Kawauchi, G. Picci, Subspace identification of closed loop systems by stochastic realization, in: CD-ROM Preprints 15th IFAC World Congress, Barcelona, 2002.

[17] M. Verhaegen, P. Dewilde, Subspace model identification part 2. Analysis of the elementary output-error state-space model identification algorithm, International journal of control, 56(5) (1992) 1211-1241.

[18] E. Reynders, G. De Roeck, Reference-based combined deterministic–stochastic subspace identification for experimental and operational modal analysis, Mechanical Systems and Signal Processing, 22(3) (2008) 617-637.

[19] P. Van Overschee, B.L. De Moor, Subspace identification for linear systems: theory, implementation, applications, Kluwer academic publishers Dordrecht, 1996.

[20] A. Deraemaeker, E. Reynders, G. De Roeck, J. Kullaa, Vibration-based structural health monitoring using output-only measurements under changing environment, Mechanical systems and signal processing, 22(1) (2008) 34-56.

[21] F. Magalhaes, A. Cunha, E. Caetano, Online automatic identification of the modal parameters of a long span arch bridge, Mechanical Systems and Signal Processing, 23(2) (2009) 316-329.

[22] S. Kim, M. Vanderploeg, QR decomposition for state space representation of constrained mechanical dynamic systems, (1986).

[23] H. Tanaka, T. Katayama, A stochastic realization in a Hilbert space based on "LQ decomposition" with application to subspace identification, in: 13th IFAC Symposium on System Identification (SYSID 2003), 2003, pp. 899-904.

[24] J.C. Santamarina, D. Fratta, Discrete signals and inverse problems, An Introduction for Engineers and Scientists. UK: Wiley & Sons, (2005).

[25] J.-H. Yi, C.-B. Yun, Comparative study on modal identification methods using output-only information, Structural Engineering and Mechanics, 17(3-4) (2004) 445-466.

[26] M. Verhaegen, V. Verdult, Filtering and system identification: a least squares approach, Cambridge university press, United Kingdom, Cambridge, 2007.

[27] C. Rainieri, G. Fabbrocino, Influence of model order and number of block rows on accuracy and precision of modal parameter estimates in stochastic subspace identification, International Journal of Lifecycle Performance Engineering 10, 1(4) (2014) 317-334.

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

[29] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in: kdd, Munchen, Germany, 1996, pp. 226-231.

[30] P. Andersen, Identification of civil engineering structures using vector ARMA models, Aalborg University, Aalborg, Denmark, 1997.

[31] R. Tarinejad, M. Pourgholi, Processing of Ambient Vibration Results using Stochastic Subspace Identification based on Canonical Correlation Analysis, Modares Mechanical Engineering, 15(7) (2015).

[32] M. Damadi pour, R. Tarinejad, System identification of a concrete arch dam and calibration of its finite element model with emphasis on nonuniform ground motion, Tabriz, Tabriz, 2012.