

## Graphical Abstract

**Using experimental modal analysis updating finite element model with two-steps optimization: A case study of the building at geothermal power plant in Taufkirchen, Munich**

Wei-Teng Kao<sup>a</sup>, Francesca Taddei<sup>a</sup>, Gerhard Mueller<sup>a</sup>

## Highlights

**Using experimental modal analysis updating finite element model with two-steps optimization: A case study of the building at geothermal power plant in Taufkirchen, Munich**

Wei-Teng Kao<sup>a</sup>, Francesca Taddei<sup>a</sup>, Gerhard Mueller<sup>a</sup>

- <https://www.sciencedirect.com/journal/learning-and-instruction/publish/guide-for-authors>
- Research highlight 2

# Using experimental modal analysis updating finite element model with two-steps optimization: A case study of the building at geothermal power plant in Taufkirchen, Munich

Wei-Teng Kao<sup>a</sup>, Francesca Taddei<sup>a</sup>, Gerhard Mueller<sup>a</sup>

<sup>a</sup>*Chair of Structural Mechanics, Technical University of Munich, Arcisstr.  
21, Munich, 80333, , Germany*

---

## Abstract

This paper presents finite element model updating (FEMU) for low-rise buildings to enhance the accuracy of FE models in representing real-world systems, which is one of the crucial processes for seismic safety assessment in civil engineering.

The proposed FEMU approach employs a two-step optimization process. In the first step, Gaussian process regression (GPR) models the relationship between mode frequencies and building parameters. Global sensitivity analysis, Sobol indices, identifies significant parameters. The significant parameters are optimized using genetic algorithms (GA) and L-BFGS-B to minimize errors of mode frequencies between the GPR model and experimental data. In the second step, Bayesian optimization (BO) refines non-significant parameters while keeping significant parameters fixed, minimizing the normalized FRF difference between simulation and experiment.

For the test, experimental modal analysis (EMA) was used to extract modal parameters and FRFs, while automated SSI-COV, an operational modal analysis (OMA) method, validated the EMA results.

*Keywords:* Finite Element Model Update, Experimental Modal Analysis, Operational Modal Analysis, Response Surface Model, Surrogate Model, Optimization

---

## 1. Introduction

Munich Data Science Institute (MDSI) and the Chair of Structural Mechanics at Technical University Munich (TUM) are cooperating on the seismic safety assessment of micro-seismic events generated from deep geothermal plants in the south of Munich. To understand the dynamic behavior of the low-rise building under micro-seismic events, it is crucial to build a representable finite element (FE) model for further seismic assessment. The investigating building is the control room for the geothermal plants in Taufkriechen, Munich, which is the closest low-rise building to the geothermal plants.

Finite element model update (FEMU) methods are developed (Ereiz et al., 2022) to build an FE model and calibrate it based on the actual behavior of the structure from dynamic structural testing. There are diverse FEMU methods for calibrating the parameters of a building. Most of the updating methods rely on an optimization process. The first type is derivative-free optimization. In Tran-Ngoc et al. (2018) works, Genetic Algorithm (GA) and particle swarm optimization (PSO) are used to update the FE model of a bridge. Additionally, derivative-free optimization is utilized in several literature of FEMU (Mosquera et al., 2012; Srinivas et al., 2011; Luong et al., 2017), which can be treated as one of the effective tools in this field. The second type is related to the gradient-based optimization method. One of the common FEMU methods, named Sensitivity FEMU, belongs to this type that uses a sensitivity matrix to provide gradient information for iteratively FE updating (He et al., 2020; Jiang et al., 2014; Zhu et al., 2015; Ren and Chen, 2010).

To form the objective function for optimization, The measure dynamics responses of the system, i.e., Frequency response function (FRF) or modal parameters, are widely used. Pu et al. (2019) forms the objective function by the error between the simulated and the experimental FRFs. Furthermore, modal parameters such as natural frequencies, damping ratios, or mode shapes are also popular choices due to their concise and meaningful representation of the dynamic behavior of a structure. Compared to FRF, which is obtained by the experimental modal analysis (EMA) technique, e.g., hammer testing or shaker vibration testing, modal parameters can be gained by the operation modal analysis (OMA) technique also. This method is more applicable to large structures in the civil engineering field.

Updating FEM via modal parameter can be realized by building the sur-

rogate model that outputs the natural frequencies with building parameters as inputs first, followed by updating the building parameter and minimizing the difference between output result from the model and experiential data (Luczak et al., 2014; Mao et al., 2019). The updated natural frequencies and their Modal Assurance Criterion (MAC) are often used for validation.

This study performs FEMU by first preparing the FRFs and the modal parameters from the result of EMA, then building the Gaussian process regression (GPR) surrogate model for the first 2 natural frequencies in section 2 of the paper. Secondly, global sensitivity analysis (GSA) is conducted to determine the significant parameters. For optimization processes, in the first step, we search for the best fit for the first 2 modes by combining GA and Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS-B) based on significant building parameters. In the second step, Bayesian optimization (BO) is proposed to tune the non-significant building parameters to align FRFs from simulation to experiment. This two-step optimization is presented in section 3. Finally, the conclusion and the outlook will be drawn in section 4.

## 2. Building for investigation

The structure for investigation is the low-rise building at the geothermal site in Taufkirchen, Munich. It is a G+2 (ground floor with 2 additional floors) building where the slab is a right-trapezoid shape (figure 1). During the hammer test, the building was already operated. The control room on the first floor had already been placed with the control systems, equipment, and office. The second floor is the storage room with the unused furniture and construction leftovers. Figure 2 shows the environment inside the building for the experiment. Further descriptions on the geometry and dimensions of the building are introduced in the following:

- The height of each floor is around 3.37 meters; therefore, this G+2 building is a total of 10.11 meters. Since this value is an approximation, it is considered the parameter that requires updating.
- Although the target building appears to be connected to the neighboring structures in the figure 1, there are gaps between them, making it an independent structure with no physical contact with the adjacent buildings.

- There are no machines within the building, which might generate additional operating vibration affecting the experiment.
- The slab's geometry is right-trapezoid, which has a shorter base of 3 meters, a longer base of 5 meters, a height of 4 meters, and a slanted side of 6 meters.



Figure 1: Target building for investigation. The image shows the full view of the geothermal power plant. The target building is marked by a red box (IKAV).

### *2.1. Dynamics testing on slabs*

The hammer test was performed to identify natural frequencies and modal shapes of the building of the slab on the first and second floors. Compared to alternative methods, the main advantages of a hammer test are its cost-effectiveness, rapid setup time, and the ability to test an unlimited number of degrees of freedom simultaneously (Eiras et al., 2018). In our study, the sensor has 8 channels connecting the accelerometers for receiving the signal from the out-of-plane vibration in  $z$  direction. 4 accelerometers are mounted on each floor, which totally provides 8 degrees of freedom information in the sum for two floors. The vibration responses were obtained by impacting the slab of the building on the second floor with PCB® impact hammer (Model 086D50). 13 independent hammer tests were conducted with respect to three



Figure 2: The interior of the building. The left image is the control room on the first floor, and the right image is the storage place on the second floor.

different exciting locations. According to each exciting location, the tests are clustered into three groups labeled EX220, EX221, and EX22, respectively. Tests 1 to 6 belong to group EX220, 7 to 9 belong to group EX221, and the remaining tests, i.e., tests 10 to 13, are included in group EX22 (Table 1). Since the excitation point and measurement points of the testings inside each group are the same, the FRFs obtained from the group should share a similar pattern or be identical. Figure 3 shows the approximated position of the excitation and provides the location of the installation of each accelerometer for both floors.

Table 1: Experiment plan of hammer test.  $X_{z,2}$  indicates the height of the second floor. The coordinates of excitation points refer to figure 3.

Test case	Event label	Excitation point
Test 1-6	EX220	$X_{ex220} : \{7.19, 8.14, X_{z,2}\}$
Test 7-9	EX221	$X_{ex221} : \{12.78, 2.75, X_{z,2}\}$
Test 10-13	EX22	$X_{ex22} : \{9.30, 5.90, X_{z,2}\}$

## 2.2. Testing result

The time series of signals are collected by accelerometers. All the signals in each test are assumed to be synchronized. To extract modal information from testing results, the FRFs need to be formulated from the input and output measurement data. The FRFs can be directly computed as the ratio of the output and input signals transferred in the frequency domain via the

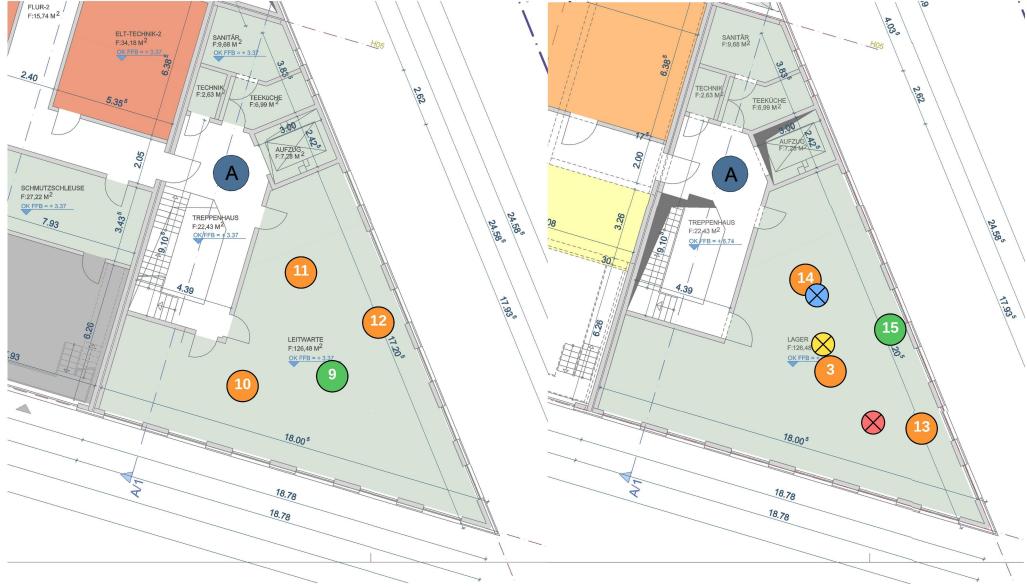


Figure 3: The approximate locations of the measurement and excitation points. The orange dots are the location of uniaxial accelerometers for  $z$  direction, and the green dots are triaxial accelerometers. The number on them corresponds to the ID of the channels. The blue, yellow, and red dots with the cross symbol are the excitation points corresponding to the event: EX220, EX22, EX221, respectively.

Fourier transforms. However, in practice, it is achieved by determining the FRF as the ratio of the cross-spectrum between the input and output signals to the auto-power spectrum due to its noise reduction feature (Mao et al., 2019). The FRF is formed in equation (1).

$$H(f) = \frac{G_{xy}(f)}{G_{xx}(f)} \quad (1)$$

, where  $H(f)$  represents the FRF in acceleration and  $f$  denoted the frequency.  $G_{xy}(f)$  and  $G_{xx}(f)$  are the Cross-spectral density (CSD) between input  $x(t)$  and output  $y(t)$  in the time domain, and Auto-spectral density (PSD) of the input  $x(t)$ , respectively. The input  $x(t)$  and output  $y(t)$  are acceleration, meaning the resulting FRF from equation (1) will also be an acceleration in the frequency domain. To compare with the FRFs in displacement generated from harmonic analysis in the section 3, the FRFs in equation (1) are

transferred to displacement in the frequency domain for convenience. The FRF in displacement  $H_d(f)$  is computed by following equation (2).

$$H_d(f) = \frac{1}{-(2\pi f)^2} H(f) \quad (2)$$

Practically,  $G_{xy}(f)$  and  $G_{xx}(f)$  are estimated over multiple averages to against noise. They are formulated in equation (3) and equation (4).

$$G_{xy}(f) = \frac{1}{N} \sum_{k=1}^N X_k^*(f) Y_k(f) \quad (3)$$

$$G_{xx}(f) = \frac{1}{N} \sum_{k=1}^N |X_k(f)|^2 \quad (4)$$

where  $N$  is number of frequency series segments,  $X_k(f)$  is Fourier transform of the  $k^{th}$  segment of input  $x(t)$  and  $Y_k(f)$  is Fourier transform of the  $k^{th}$  segment of output  $y(t)$ .  $*$  denotes complex conjugate.

Furthermore, using CSD and PSD allows us to compute the coherence function  $C_{xy}(f)$ . This function quantifies the degree of the output signal that the input signal can explain.  $C_{xy}(f)$  is defined as equation (5).  $G_{yy}$  represents the PSD of the output  $y(t)$ .

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx} G_{yy}} \quad (5)$$

Based on the definitions of the PSD and the CSD, the coherence function will equal to 1 if there is no measurement noise and the system is linear. In contrast, 0 will be computed when signals are totally uncorrelated.

The modal parameters were obtained from FRFs using the Peak Amplitude Method (Ewins, 2009), including peak identification, damping ratio estimation, and mode shapes extraction. In this study, the natural frequencies are determined using the pick-peak method developed by Liutkus (2015), and the damping ratios are obtained by the half-power bandwidth method using equation (6).

$$\xi = \frac{f_2 - f_1}{2f_n} \quad (6)$$

that  $f_1, f_2$  are the frequencies on either side of the peak where the amplitude is  $\frac{1}{\sqrt{2}}$  of the amplitude at natural frequency  $f_n$ . Once the natural frequencies are detected, the mode shapes with direction can be obtained from the

imaginary part of FRFs (Avitabile, 2001). To compare with the pick-peak method in order to validate the selected modal parameters. One of the Operating modal analysis (OMA) methods naming automated Covariance driven Stochastic Subspace Identification (automated SSI-COV), which is developed by Cheynet (2020) is applied. SSI-COV is the time domain OMA method, which has the advantages of high parameter estimation accuracy and high computational efficiency compared to other OMA methods (Reynders et al., 2016). The SSI-COV method identifies the modal parameters by first forming the stochastic discrete state space model, and the input  $X(t)$  is assumed as a white noise (Magalhães et al., 2009). The discrete state space model is defined by the following equation (8).

$$x_{k+1} = A \cdot x_k + w_k \quad (7)$$

$$y_k = C \cdot x_k + v_k \quad (8)$$

where  $x_k$  is the discrete-time state vector at time instant  $k$  and  $y_k$  is a vector of sampled outputs.  $A, C$  denote the discrete state matrix and discrete output matrix, receptively, and  $w_k, v_k$  are vectors representing the noise. At the beginning of the algorithm, the output covariance  $R_i$  is formed as equation (9).  $N$  is the data samples and  $i$  is time lag.

$$R_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-i} y_{k+i} y_k^T \quad (9)$$

Secondly, Toeplitz matrix, the output covariance matrix  $T_i$ , is obtained from equation (10).

$$T_i = \begin{bmatrix} R_i & R_{i-1} & \cdots & R_1 \\ R_{i+1} & R_i & \cdots & R_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_{2i-1} & R_{2i-2} & \cdots & R_i \end{bmatrix} \quad (10)$$

$T_i$  can be decomposed into product of observability matrix  $\mathbf{O}_i$  and controllability matrix  $\mathbf{\Gamma}_i$  as equation (11),following the instruction in Qin et al. (2016).

$$T_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} [A^{i-1}G, A^{i-2}G, \cdots, G] = \mathbf{O}_i \mathbf{\Gamma}_i \quad (11)$$

After that, the Singular Value Decomposition (SVD) is performed on  $T_i$  to extract the dominant modes and reduce noise as equation (12)

$$T_i = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (12)$$

where  $\mathbf{U}$  is singular matrix with diagonal values, and  $\mathbf{V}, \mathbf{D}$  are orthogonal matrices. The system matrices  $A, C$  can be solved by solving equation (11) and equation (12) (Qin et al., 2016). Finally, the modal parameters can be obtained by eigenvalue decomposition on matrix A by equation (13).

$$A = \Psi \Lambda \Psi^{-1} = \sum_{k=1}^n \psi_k \lambda_k \psi_k. \quad (13)$$

$\lambda_k$  is the eigenvalue of discrete time system and  $\psi_k$  is the eigenvector of  $A$  (Qin et al., 2016). The natural frequencies  $f_k$ , damping ratio  $\xi_k$  and mode shapes  $\varphi_k$  are computed by equation (14), where eigenvalue in continuous time system,  $\lambda_{ck}$ , equal to  $\frac{\ln(\lambda_k)}{\Delta t}$  and  $\Delta t$  represents discrete time step.

$$f_k = \frac{|\lambda_{ck}|}{2\pi}; \quad \xi_k = \frac{-100 \operatorname{Re}(\lambda_{ck})}{|\lambda_{ck}|}; \quad \varphi_k = C\psi_k \quad (14)$$

For model order selection, there are several criteria such as natural frequency variation  $\Delta f_i$ , modal damping coefficient variation  $\Delta \xi_i$ , MAC (Qin et al., 2016; Magalhães et al., 2009). The MAC is a value used to measure the similarity between two mode shapes, defined as equation (15).

$$\text{MAC}(\phi_1, \phi_2) = \frac{|\phi_1^T \phi_2|^2}{(\phi_1^T \phi_1) \cdot (\phi_2^T \phi_2)} \quad (15)$$

where  $\phi_1, \phi_2$  are the mode shape vectors for comparison, and MAC ranges as  $0 \leq \text{MAC} \leq 1$ . MAC value equals 1, indicating that the two mode shapes are identical. On the other hand, MAC = 0 shows no similarity between the two modes.

In our study, the following limits are defined for system identification:  $\Delta f_i \leq 1\%$ ,  $\Delta \xi_i \leq 2\%$ , and  $\text{MAC} \geq 0.99$ . Additionally, the automatic identification feature purposed by Magalhães et al. (2009) is applied, which uses the Hierarchical clustering (HC) technique (Ogier, 2025) to group the similar estimated modes. The max distance criterion is set as 0.2 Hz.

Figure 4 plots the detected natural frequencies and estimated damping ratios for each test. The common global modes crossing each test are approximately 12 and 18.5 Hz. This result can also be observed by the SSI-COV

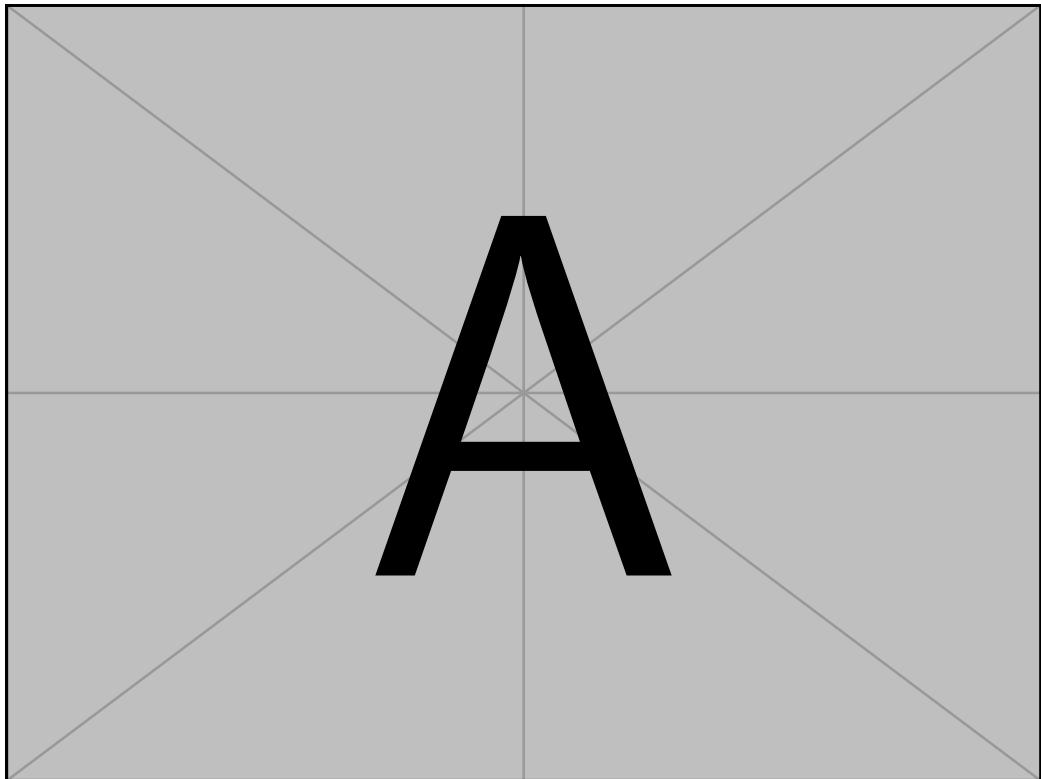


Figure 4: Modal parameters estimating by peak amplitude method from all tests. The left sub-figure shows the damping ratios against detected natural frequencies, and the right sub-figure demonstrates the detected natural frequencies from the  $i$  test.

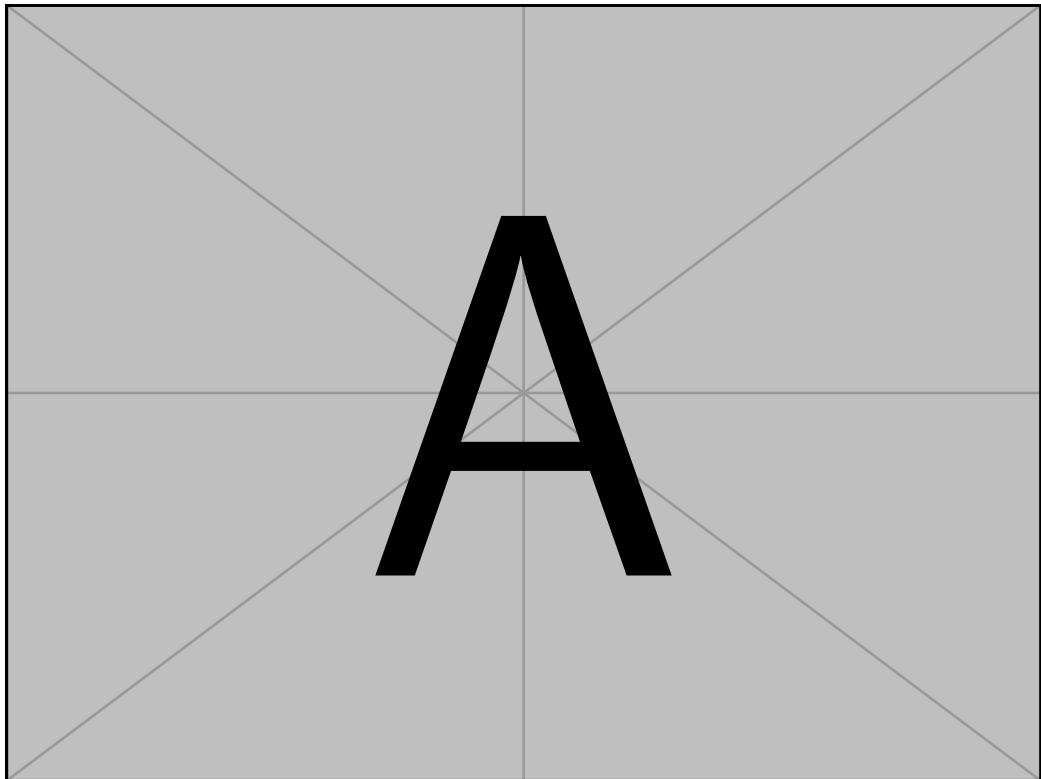


Figure 5: Modal parameters estimating by SSI-COV method from all tests. The left sub-figure shows the damping ratios against detected natural frequencies, and the right sub-figure demonstrates the detected natural frequencies from the  $i$  test.

analysis (figure 5). Figure 6 shows the real and imaginary parts of FRFs of test 1, which belong to the event EX220. The imaginary parts of FRFs at each natural frequency are extracted and normalized to form the mode shape vectors. In this study, only the modes at 12 and 18.5 Hz are investigated and used for further FEM updating. The mode shapes are plotted in figure 7.

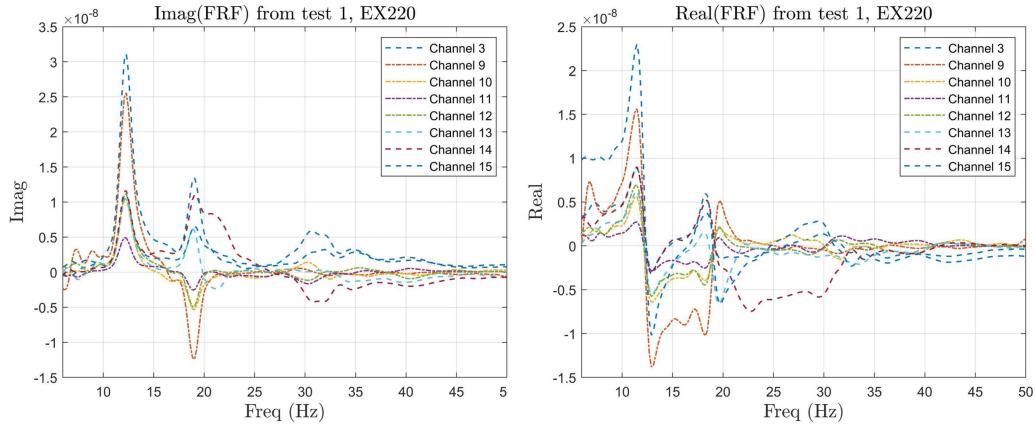


Figure 6: The FRFs measuring in test 1, EX220

### 3. Updating FE model to testing results

A two-step optimization technique for calibrating the FE model with experimental results is proposed. In the optimization process, the natural frequencies of the model will be first calibrated. This requires the surrogate model for each targeted natural frequency.

After building the surrogate models, global sensitivity analysis, e.g., Sobol's indices, is applied to separate the building parameters into significant and non-significant groups. For updating significant building parameters, the GA and L-BFGS-B are applied sequentially to minimize the error of natural frequencies between the surrogate model and experiment. Next, it updates the non-significant parameters while the significant parameters are fixed by GP optimization. This step minimizes the difference of normalized FRFs between simulation and experiment without heavily varying the natural frequencies.

#### 3.1. Finite element model

In this section, the dynamic characteristics of the building will be determined by the finite element model. Several components are assumed to

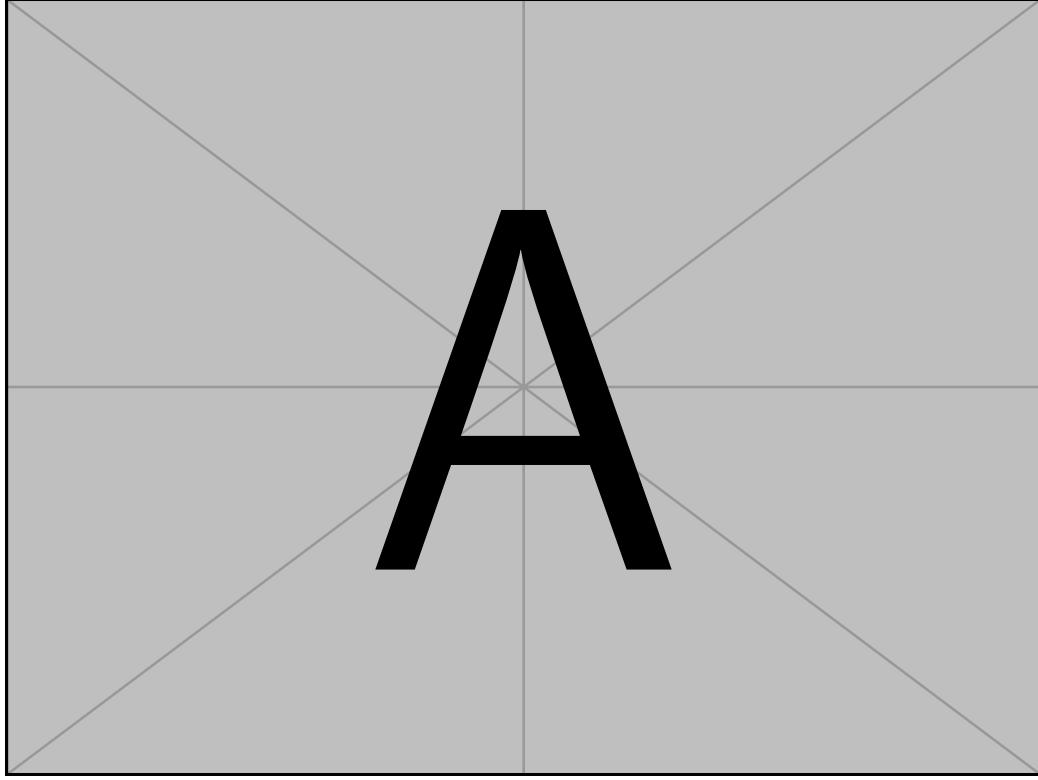


Figure 7: 2 mode shapes extracted. Only the points plotted as red dots are the mode shapes extracted from the measurement. The mesh and silhouette of each slab are for demonstration.

assemble the building, each with different material and geometric properties:

- The building has components: slabs, corridors, grounds, exterior walls, interior walls, and stairs. The stair is assumed to have the same property as the slab. The FE model is presented in figure 8.
- The shell element (SHELL181<sup>1</sup>) and beam element (BEAM189<sup>1</sup>) are used for modeling the building. The shell element is defined by thickness  $t$ , Young's modulus  $E$ , Poisson's ratio  $\nu$ , and density  $\rho$ . Additionally, the cross-section of the beam is square.
- The soil-structure interaction is considered using a frequency-independent

---

<sup>1</sup>Element defined by ANSYS

lumped parameter model (LPM) proposed by Wolf (1994). It is implemented by MATRIX27<sup>1</sup> and MASS21<sup>1</sup> element. The LPM model is formulated by the shear-wave velocity  $v_s$ , Poisson's ratio of soil  $\nu_s$ , and density of soil  $\rho_s$ . The implementation and setting of boundary conditions are referred to Kao and Pfleger (2023).

- The constant modal damping ratio  $\xi_b$  is assumed.
- Totally, there are 26 parameters for material properties of building  $\theta_{bld}$  and 3 parameters for LPM modeling  $\theta_{ssi}$  including in building parameters  $\theta_b$  for updating. Their initial values for the FE model are defined in table 2.

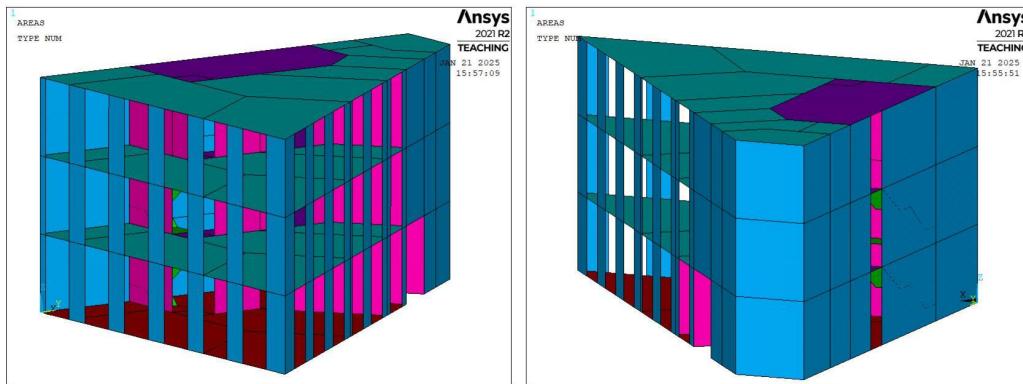


Figure 8: FE model of the target building. The green, purple, brown, blue, pink, and light green areas represent the slab, corridor, ground, exterior wall, interior wall, and stair, respectively. The LPM model, connected with the ground area, is not shown here.

In this study, the PyMAPDL package is used. PyMAPDL is a Python API developed by Kaszynski (2021) that offers access to the Mechanical Ansys Parametric Design Language (MAPDL) solver and its generated simulation data.

### 3.2. Surrogate modeling

The surrogate model is typically a non-intrusive technique used to describe the relationship between output and given input, which can be treated as a mathematical approximation or an alternative representation of a more complex or computationally expensive model or system. It is well used in

Table 2: Initial values of building parameter  $\theta_b$ 

Building parameters $\theta_{bld}$							
	Thickness [m]	Young's modulus [N/m <sup>2</sup> ]	Poisson's ratio [-]	density kg/m <sup>3</sup>			
Slab	$t_{sl}$	0.468	$E_{sl}$	20E9	$\nu_{sl}$	0.27	$\rho_{sl}$
Corridor	$t_c$	0.8	$E_c$	20E9	$\nu_c$	0.27	$\rho_c$
Ground	$t_g$	0.05	$E_g$	20E9	$\nu_g$	0.27	$\rho_g$
Exterior walls	$t_{ew}$	0.2	$E_{ew}$	20E9	$\nu_{ew}$	0.27	$\rho_{ew}$
Interior walls	$t_{iw}$	0.2	$E_{iw}$	20E9	$\nu_{iw}$	0.27	$\rho_{iw}$
Stair	$t_{st}$	0.468	$E_{st}$	20E9	$\nu_{st}$	0.27	$\rho_{st}$
Height per floor [m]				$H_b = 3.37$			
Damping ratio [-]				$\xi_b = 0.045$			
Soil-structure-interaction parameters $\theta_{ssi}$							
Shear-wave velocity [m/s]				$v_s = 350$			
Poisson's ratio [-]				$\nu_s = 0.33$			
Density [kg/m <sup>3</sup> ]				$\rho_s = 1400$			

optimization to reduce computational costs while maintaining reasonable accuracy. This study selects the GPR model to describe the dominant mode frequency.

GPR is a non-parametric surrogate model that can model arbitrarily complex systems if enough data is provided (Schulz et al., 2018). Furthermore, GP-based algorithms are particularly effective when data is scarce because they provide a principled way of combining prior knowledge with observed data (Williams and Rasmussen, 2006). These features make it appealing for this study, which simulation is typically time-consuming. Moreover, although the GPR is reported to suffer from *curse of dimensionality*, The surrogate model with 19 inputs should be practically realizable (Frazier, 2018).

For training the GPR model, the training and validation data are prepared by following steps:

1. It is assumed that  $t_j, E_j, \rho_j$  for  $j^{th}$  components and  $H_b, \xi_b$  are variable while the  $\nu_j$  is fixed. Considering the assumption that the stair has the same properties as the slab, there are totally 19 parameters for updating.
2. For Design of Experiments (DOE), 1000 normalized samples are generated via the Latin Hypercube Sampling (LHS) method for each event following a uniform distribution. After that, the iso-probabilistic transform maps the uniform space to physical space using equation 16.

$$\theta_i = \mu_i + \frac{\mu_i(2U_i - 1)}{r}; U_i \sim \mathcal{U}(0, 1) \quad (16)$$

where  $\Theta = \{\theta_1, \theta_2, \dots, \theta_{19}\}^T$  is a vector of independent uniform variables and  $\theta_i \sim \mathcal{U}(a, b)$  is uniform variable of  $i^{th}$  building parameter.  $\mu_i$  is the

mean of  $X_i$  defined due to prior knowledge;  $r$  is the ratio controlling the range for parameter varying; the default is  $r = 2$ . The  $U_i$  is a standard uniform vector obtained from LHS.

3. Harmonic analysis is performed using the modal-superposition method (MSM). The advantages of the MSM method are that it increases computational efficiency (Shah et al., 1979) and provides natural frequencies without additional modal analysis. The excitation is the unit force loading applied on the locations shown in figure 3.
4. Extracting the natural mode frequencies by comparing the mode shapes with the experiment. The modes are extracted from each simulation if high amplitudes are detected at certain modes from FRFs and  $\text{MAC} \geq 0.5$ . Otherwise, the corresponding simulation result will be ignored. The FEM updating in this study doesn't aim to fit all of the modes. Instead, the focus is on matching the modes that have a dominating impact in experiments. According to section 2.1, two dominated modes reported at 12 and 18.5 Hz, respectively, are our target modes for matching.
5. Spiting the data to training and validation sets. The correlation between the two datasets is limited to under 20 percent to ensure independence.

Based on DOE, the response surfaces of each mode-frequency are computed using the GPR. The validation and errors of the GPR surrogate model for each event are shown. Figure 9 is the comparison of results from the GPR models and experiment results for 1<sup>st</sup> and 2<sup>nd</sup> dominated modes.

Global sensitivity analysis is performed since further optimization requires the separate updating of significant and non-significant parameters. One of the variance-based sensitivity analysis techniques, called Sobol indices, is introduced to quantify how much each input variable contributes to the output variance of a model. The formulation of First-order Sobol indices  $S_i$  is in the following equation 17

$$S_i = \frac{\text{Var}[E(Y|\theta_i)]}{\text{Var}(Y)} \quad (17)$$

where Y is the GPR model output.

Figure 10 demonstrates the result of first-order indices from all events;

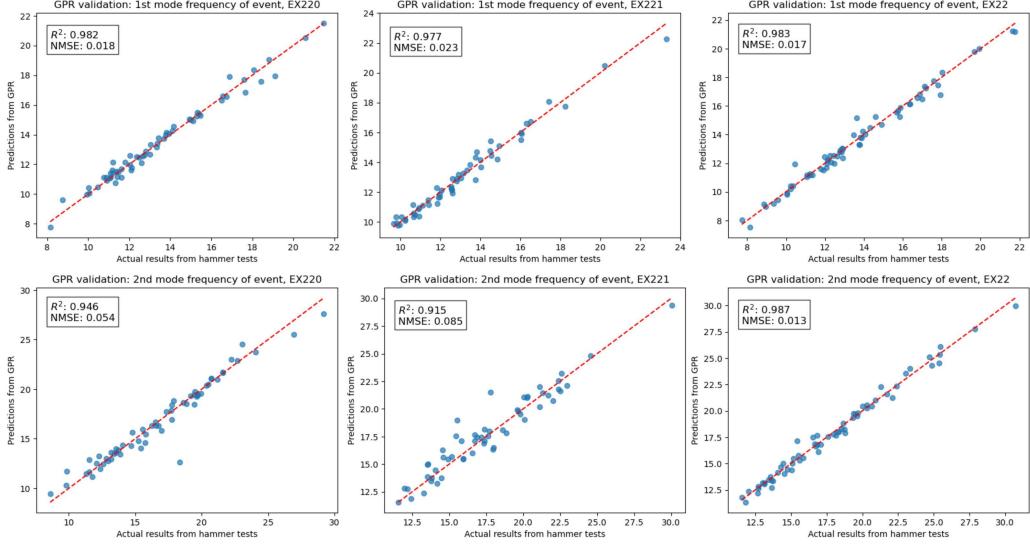


Figure 9: Validation of the GPR models for 1<sup>st</sup> and 2<sup>nd</sup> dominant mode frequency.

the significant parameters influencing each dominant mode frequency are:

$$\begin{aligned}\theta_{fn_1} &= \{t_{sl}, t_{ew}, t_{iw}, E_{sl}, E_{ew}, E_{iw}, \rho_{sl}, v_s, \rho_s, H_b\} \\ \theta_{fn_2} &= \{t_{sl}, t_{ew}, E_{sl}, \rho_{sl}, H_b\}\end{aligned}$$

$\theta_{fn_1}$  and  $\theta_{fn_2}$  denoted the high-sensitive parameters for the top two dominant mode frequencies. For convenience, we define high-sensitivity parameters  $\theta_{hs}$  and low-sensitivity parameters  $\theta_{ls}$  in equation (19) using union ( $\cup$ ) and complement ( $\sim$ ) symbol. Finally, 10 sensitive parameters are extracted.

$$\theta_{hs} = \theta_{fn_1} \cup \theta_{fn_2} \quad (18)$$

$$\theta_{ls} = \sim \theta_{hs} \quad (19)$$

The degree of sensitivity can also be observed from the response surface of the GPR model plotted in figure 11.

### 3.3. Optimization for aligning dominant mode frequencies

To update the  $\theta_{hs}$ , the proposed process applies GA and L-BFGS-B algorithms fitting the dominant mode frequencies. In this study, GA is applied first to search the potential combination of  $\theta_{hs}$ , followed by further calibration

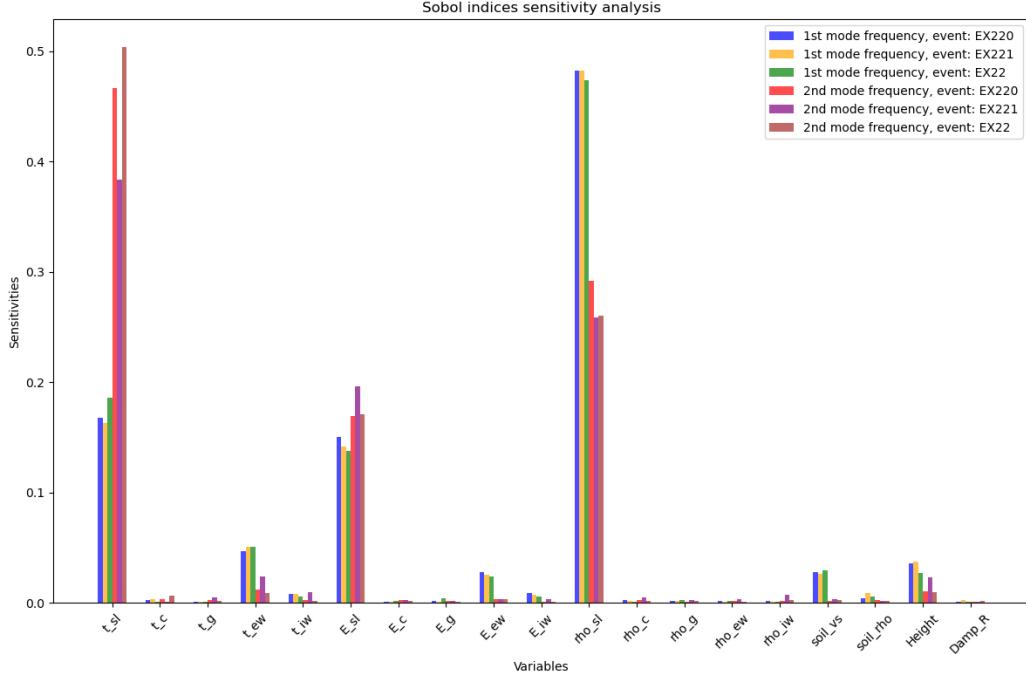


Figure 10: Sensitivity analysis on top two dominated mode-frequencies using first-order indices. The bars in blue, yellow, and orange present the significance of 1<sup>st</sup> mode frequency based on each variable, while the bars in red, purple, and brown are for the 2<sup>nd</sup> mode frequency. The results from all three events are compared.

via the L-BFGS-B algorithm. The objective of the optimization is aligning dominant mode frequencies between the GPR outputs and experiment results by the objective function defined in equation (20):

$$\hat{\theta}_{hs} = \arg \min_{\theta_{hs}} \sum_{m=1}^3 \sum_{k=1}^2 (w_k^m (fn_{k,gpr}^m(\theta_{hs}, \theta_{ls}) - fn_{k,exp}^m))^2 \quad (20)$$

where  $w = 100$  is the weight,  $\hat{\theta}_{hs}$  is the optimal solution,  $fn_{k,gpr}^m$  is the natural frequency return from the GPR model from input  $\{\theta_{hs}, \theta_{ls}\}$ .  $\theta_{hs}$  has 10 parameters in the range between 0 and 1, and  $\theta_{ls} = 0.5\mathbf{1}_9$ , that  $\mathbf{1}_d$  is a vector of one with size  $d$ . In this step,  $\theta_{ls}$  is fixed.  $fn_{exp}$  is the natural frequency from experiment. The index  $k$  denotes  $k^{th}$  mode frequency and index  $m$  represents result of  $m^{th}$  event.

The selected optimization algorithms are briefly described: GA is a population based approach for solving complex optimization problems. Population

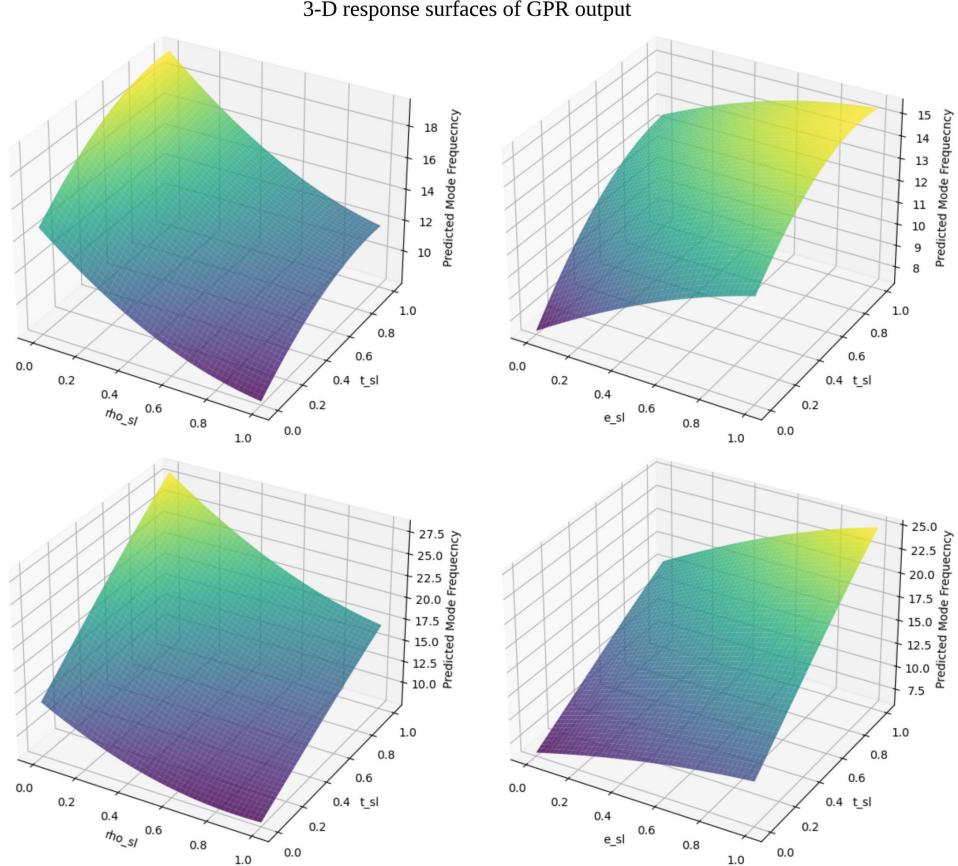


Figure 11: 3-D response surfaces of GPR output. The surfaces on the top of the figure show predicted 1<sup>st</sup> mode frequency in the z-axis against  $\rho_{sl}$ ,  $t_{sl}$ , and  $E_{sl}$ . The surfaces at the bottom of the figure demonstrate the response surfaces for predicted 2<sup>nd</sup> mode frequency

based algorithm enhances global search capabilities, which helps us search the possible combination of  $\theta_{hs}$  without trapping in local optima. Additionally, GA is suitable for optimization problems in under-determined systems (Holland, 1992; Coello, 2007), which fits our current case with greater unknowns in the model than known experimental results. The GA involves selection, crossover, and mutation operation. The purpose of each operation is:

1. Selection: Choosing individuals from the current population with better fitness ensures that optimal solutions are more likely to contribute

to the next generation.

2. Crossover: Combining information from two parent solutions to generate new offspring, which potentially inherits the strengths of both parents.
3. Mutation: Introducing random changes into an individual to maintain genetic diversity for global searching and prevent premature convergence in local optima.

The theoretical detail of GA is introduced in Mitchell (1998). This study implements the GA using the package Distributed Evolutionary Algorithms in Python (DEAP), developed by Fortin et al. (2012).

Furthermore, the L-BFGS-B algorithm is introduced. L-BFGS-B is an extension of the limited-memory BFGS (L-BFGS) algorithm that overcomes the limitation for only unconstrained problems. It can deal with bounds on the variables (Zhu et al., 1997). It is in the family of the quasi-Newton method, which requires a gradient of the objective function for optimization. The gradient information in this study could be obtained with the help of the GPR model, which can be treated as the smooth response surface due to its kernel. One main concern of L-BFGS-B is its sensitivity of initial guess since it relies on local gradient information. As a remedy, a population-based approach, e.g., GA, is applied beforehand to provide candidates with an initial guess, which is assumed to be the nearest of the global minima. This hybrid method is also mentioned in (Dalvand and Hajarian, 2020; Talbi, 2009; Renders and Flasse, 1996). The SciPy Python package developed by Virtanen et al. (2020) is used for implementation.

The FRFs obtained from the FE model after update with input  $\hat{\theta}_{hs}$  are shown in figure 12. The error compared to experimental mode frequencies and the MAC values after the update are also demonstrated in the table.

Figure 12 shows that only the dominant mode frequencies at approximately 12 and 18.5 Hz match, while the FRFs at higher frequencies range still do not align well. This observation is due to the objective function in equation (20), which is formed by the GPR models without considering the whole FRFs from the FE model. Figure 13 shows the mode shapes of two dominant modes: the 4<sup>th</sup> and 10<sup>th</sup> modes located at 12.1684 Hz and 18.7840 Hz, respectively. For mode at 12.1684 Hz, both slabs on the first and second floors deform in the positive  $z$ -direction, opposite the hammer impact. For mode at 18.7840 Hz, the first floor's slab deforms along the direction of impact while the slab of the second floor deforms opposite of the hammer

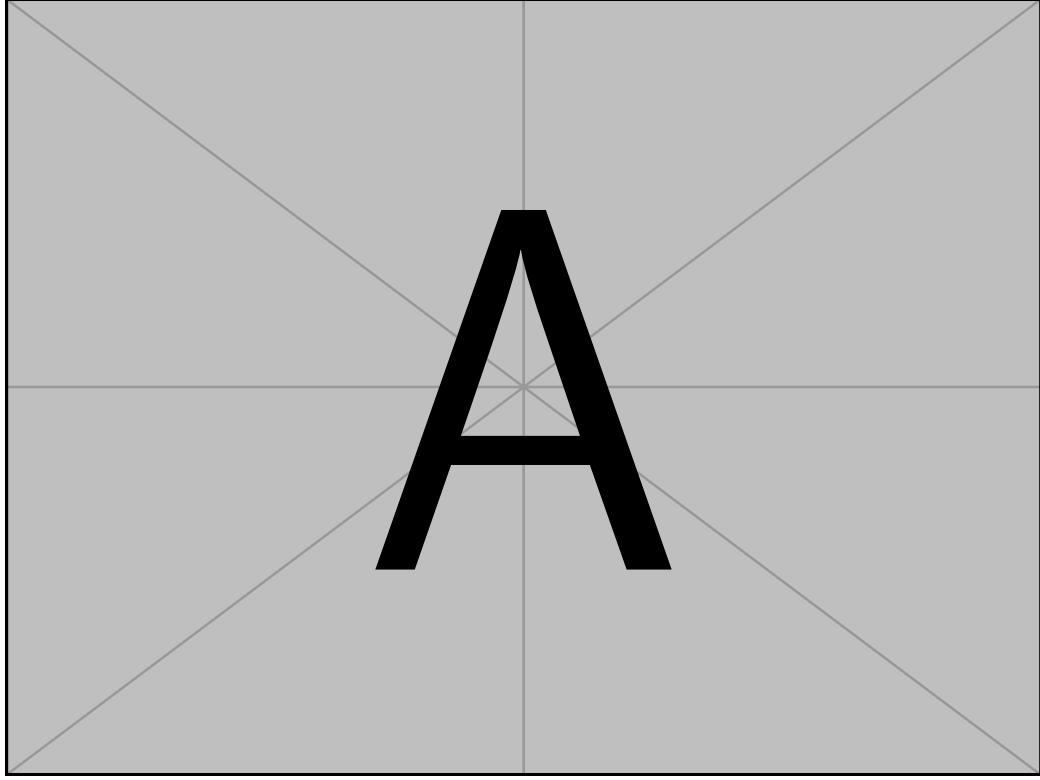


Figure 12: The FRFs of the FE model after aligning dominant mode frequencies from each event.

impact. This observation matches the FRFs obtained from the experiment in figure 6.

In this step, the FEM update based on mode frequencies and mode shapes is conducted. In the scenario that only mode parameters are available, e.g., OMA, this proposed step could help the user update the FE model with satisfied dynamics interpretation. The next step of optimization is introduced in section 3.4 to improve the accuracy of the FE model further. However, the FRFs obtained from the simulation and the experiment are required.

TABLE HERE

### 3.4. Fine-tuning of matching frequency response functions

This optimization step is called *fine-tuning* because it minimizes the difference of FRFs between simulation and experiment in the whole frequency span by updating  $\theta_{ls}$ . It is realized by minimizing the objective function in

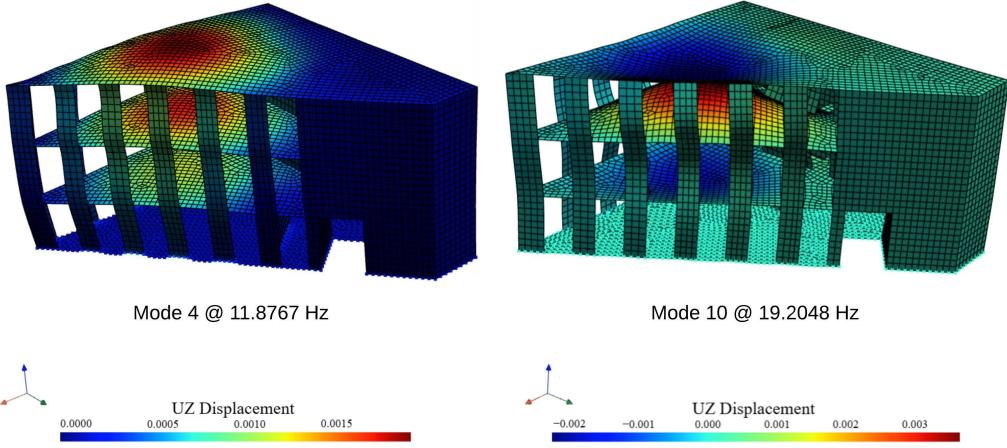


Figure 13: Mode shapes of two dominant modes after first-step FEM update

equation (21) using the BO algorithm.

$$\begin{aligned}\hat{\theta}_{ls} &= \arg \min_{\theta_{ls}} E_t(\hat{\theta}_{hs}, \theta_{ls}) \\ E_t &= E_{fn}(\hat{\theta}_{hs}, \theta_{ls}) + E_H(\hat{\theta}_{hs}, \theta_{ls})\end{aligned}\quad (21)$$

where  $E_t$  is total error, and the error  $E_{fn}$  from GPR model output for fine-tuning and the error  $E_H$  from FRFs are defined:

$$E_{fn}(\hat{\theta}_{hs}, \theta_{ls}) = \frac{\sum_{m=1}^3 \sum_{k=1}^2 (w(fn_{k,simu}^m(\hat{\theta}_{hs}, \theta_{ls}) - fn_{k,exp}^m))^2}{\sum_{m=1}^3 \sum_{k=1}^2 (fn_{k,exp}^m)^2}; w = 25 \quad (22)$$

$$E_H(\hat{\theta}_{hs}, \theta_{ls}) = \frac{\sum_{m=1}^3 \sum_{j=1}^8 (M \cdot (\hat{H}_{simu}^{m,j}(\hat{\theta}_{hs}, \theta_{ls}) - \hat{H}_{exp}^{m,j}))^2}{\sum_{m=1}^3 \sum_{j=1}^8 2(\hat{H}_{exp}^{m,j})^2} \quad (23)$$

where  $w$  is the weight,  $\hat{H}_{simu}(\hat{\theta}_{hs}, \theta_{ls})$  and  $\hat{H}_{exp}$  are normalized FRFs based on 8 channels in each event from simulation and experiment, respectively.  $fn_{simu}$  is the mode frequency extracted from each event.  $m$  denotes the  $m^{th}$  event,  $k^{th}$  represents the  $k^{th}$  mode frequency and  $j$  is the index of measurement points or channels. The existence of  $E_{fn}$  prevents the heavy shifting of mode frequencies from the optimized result obtained in section 3.3. For  $E_H$ , the mask  $M$  is applied to amplify the difference at specific frequency spans. The mask is defined in equation (24), where the scaling of each frequency span is

hyper-parameter.

$$M(f) = \begin{cases} 3 & \text{if } 14 \leq f < 17, \\ 2 & \text{if } 17 \leq f \leq 50, \\ 1 & \text{otherwise.} \end{cases} \quad (24)$$

Since the computation of  $E_H$  in objective function requires FE simulation, which is typically time-consuming, choosing the optimization method that can deal with small data or expensive evaluations is better. To fulfill this, BO is proposed. BO is a probabilistic model-based method that aims to search for the global optimum of an expensive objective function. Unlike conventional optimization methods requiring a large population or multiple iterations, BO balances exploration and exploitation by combining a surrogate model, e.g. Gaussian process, and an acquisition function. This study uses the Scikit-optimize Python package developed by Head et al. (2018). The bounds of each input are [-0.35, 1.35], and the same mapping as equation (16) is used. Since the optimization of  $E_H$  is based on simulation, not the GPR model, the searching range doesn't need to be identical to DOE. The initial guess of BO is set as  $\{\hat{\theta}_{hs}, \theta_{ls} = 0.5\mathbf{1}_9\}$ .

Figure 14 compares the FRFs before and after fine-tuning with 100 iterations. It is observed that difference between  $\hat{H}_{simu}(\hat{\theta}_{hs}, \hat{\theta}_{ls})$  and  $\hat{H}_{exp}(\hat{\theta}_{hs}, \hat{\theta}_{ls})$  after 20 Hz decrease, implying the increase of similarity of their dynamics characteristic.

#### 4. Conclusion

This study demonstrates FEMU through a two-step optimization process to calibrate the FE model using experimental FRFs and model parameters. The first step employs GA and L-BFGS-B algorithms to optimize significant building parameters  $\theta_{hs}$  on the GPR response surface, aligning mode frequencies. Fine-tuning in the second step utilizes the BO algorithm to minimize the difference between simulated and experimental FRFs.

Modal parameters, including natural frequencies and mode shapes, were extracted. Both EMA and OMA identified dominant frequencies of approximately 12 Hz and 18.5 Hz across all hammer tests. The agreement between EMA and OMA results suggests that FEM updating is feasible using only the OMA method without FRF data from the EMA method.

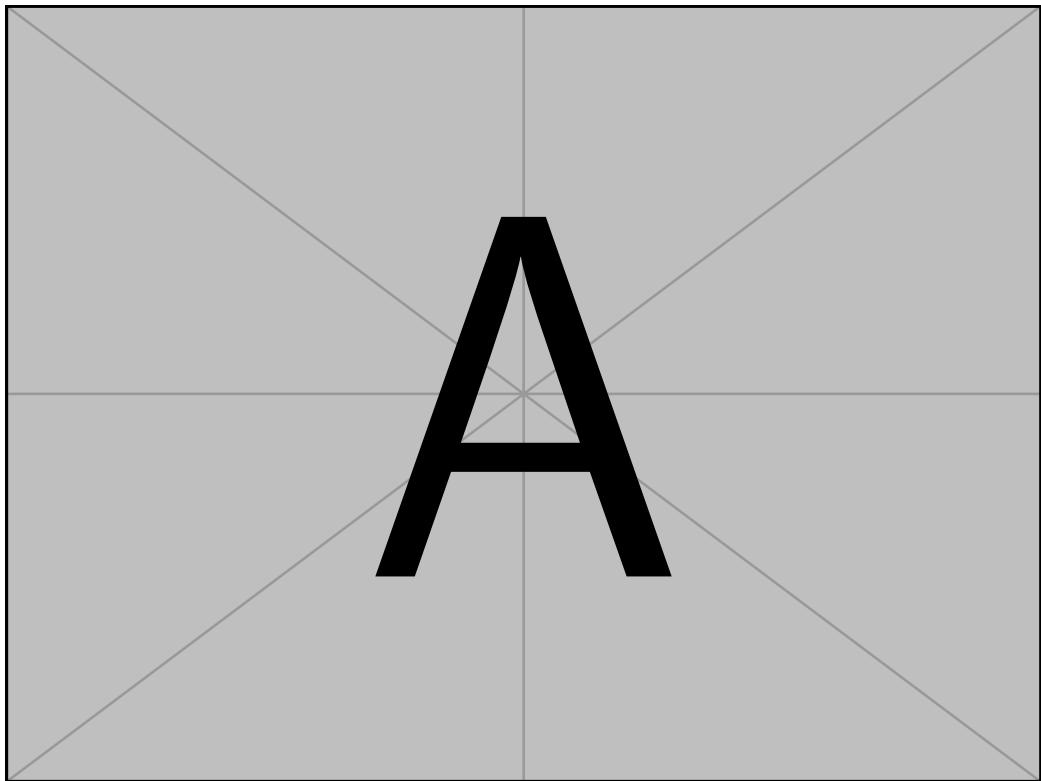


Figure 14: The FRFs before and after fine-tuning.  $E_t$  is the current total error, equaling the summation of  $E_{fn}$  and  $E_H$ .

In the first step, the GPR surrogate models are built. GA, followed by L-BFGS-B, ensures robust parameter estimation of  $\theta_{hs}$  with sufficient accuracy, calibrating modal frequencies from 10 Hz and 20 Hz to 12 Hz and 18.5 Hz, with MAC values of 0.7 and 0.6, respectively. It is noticed that the first optimization step focuses on modal parameters without requiring FRFs, making it applicable even when using only OMA data for collection. In the second step, BO optimization further refines the FE model, minimizing dynamic response errors  $E_t$  from 1.2 to 0.4 in the frequency range of 5–50 Hz and enhancing similarity above 20 Hz through iterative parameter updates.

The updated FE model outputs a dynamic response similar to the target low-rise building through optimization techniques. This enables more robust dynamic analysis and reliable seismic assessment following the model update.

#### 4.0.1. Mathematics

This is an example for the symbol  $\alpha$  tagged as inline mathematics.

$$f(x) = (x + a)(x + b) \quad (25)$$

$$f(x) = (x + a)(x + b)$$

$$f(x) = (x + a)(x + b) \quad (26)$$

$$= x^2 + (a + b)x + ab \quad (27)$$

$$\begin{aligned} f(x) &= (x + a)(x + b) \\ &= x^2 + (a + b)x + ab \end{aligned} \quad (28)$$

$$\begin{aligned} f(x) &= (x + a)(x + b) \\ &= x^2 + (a + b)x + ab \end{aligned}$$

$$\begin{aligned} f(x) &= (x + a)(x + b) \\ &= x^2 + (a + b)x + ab \end{aligned}$$

1	2	3
4	5	6
7	8	9

Table 3: Table Caption

## Appendix A. Example Appendix Section

### References

- Avitabile, P., 2001. Experimental modal analysis. *Sound and vibration* 35, 20–31.
- Cheynet, E., 2020. Operational modal analysis with automated ssi-cov algorithm. URL: <https://zenodo.org/record/3774061>, doi:10.5281/ZENODO.3774061.
- Coello, C.A.C., 2007. Evolutionary algorithms for solving multi-objective problems. Springer.
- Dalvand, Z., Hajarian, M., 2020. Solving generalized inverse eigenvalue problems via l-bfgs-b method. *Inverse Problems in Science and Engineering* 28, 1719–1746.
- Eiras, J.N., Payan, C., Rakotonarivo, S., Garnier, V., 2018. Experimental modal analysis and finite element model updating for structural health monitoring of reinforced concrete radioactive waste packages. *Construction and Building Materials* 180, 531–543.
- Ereiz, S., Duvnjak, I., Jiménez-Alonso, J.F., 2022. Review of finite element model updating methods for structural applications, in: *Structures*, Elsevier. pp. 684–723.
- Ewins, D.J., 2009. Modal testing: theory, practice and application. John Wiley & Sons.
- Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., Gagné, C., 2012. DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13, 2171–2175.
- Frazier, P.I., 2018. A tutorial on bayesian optimization. arXiv preprint arXiv:1807.02811 .

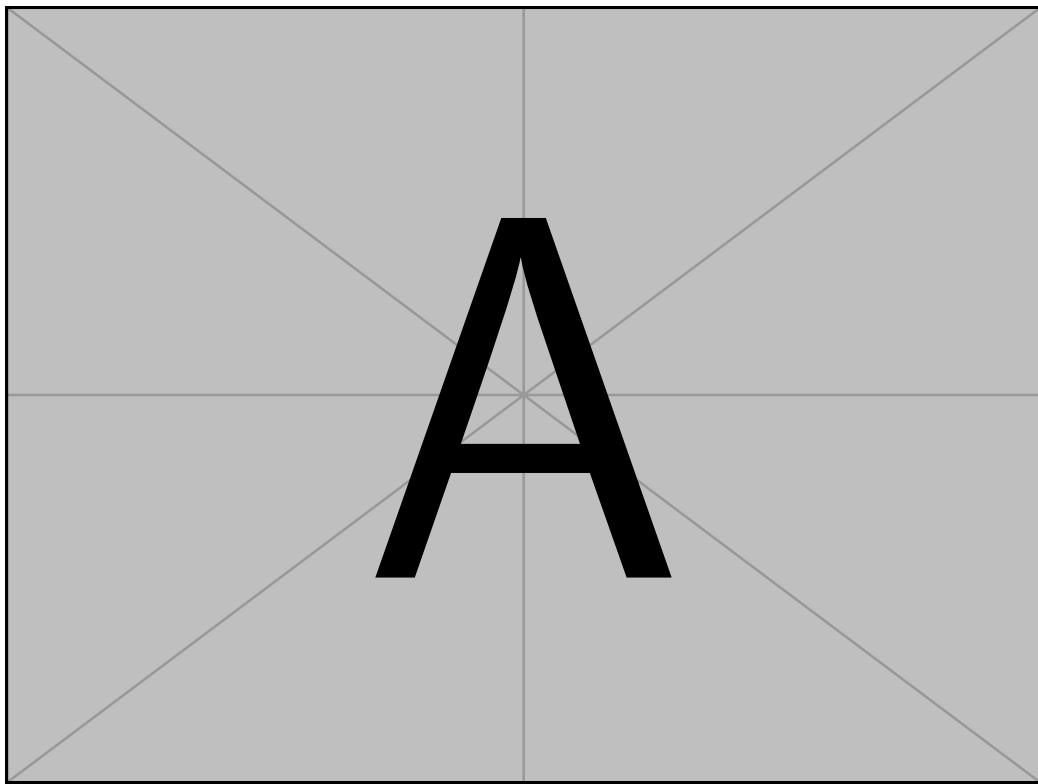


Figure 15: Figure Caption

He, L., Reynders, E., García-Palacios, J.H., Carlo Marano, G., Briseghella, B., De Roeck, G., 2020. Wireless-based identification and model updating of a skewed highway bridge for structural health monitoring. *Applied Sciences* 10, 2347.

Head, T., MechCoder, Louppe, G., Shcherbatyi, I., fcharras, Vinícius, Z., cm-malone, Schröder, C., nel215, Campos, N., Young, T., Cereda, S., Fan, T., rene rex, Shi, K.K., Schwabedal, J., carlosdanielcsantos, Hvass-Labs, Pak, M., SoManyUsernamesTaken, Callaway, F., Estève, L., Besson, L., Cherti, M., Pfannschmidt, K., Linzberger, F., Cauet, C., Gut, A., Mueller, A., Fabisch, A., 2018. scikit-optimize/scikit-optimize: v0.5.2. URL: <https://doi.org/10.5281/zenodo.1207017>, doi:10.5281/zenodo.1207017.

Holland, J.H., 1992. Adaptation in natural and artificial systems: an intro-

- ductory analysis with applications to biology, control, and artificial intelligence. MIT press.
- IKAV, . The existing geothermal power plants in munich. IKAV. URL: <https://www.ikav.com/assets>. [Photograph].
- Jiang, D., Zhang, P., Fei, Q., Wu, S., 2014. Comparative study of model updating methods using frequency response function data. Journal of Vibroengineering 16, 2305–2318.
- Kao, W.T., Pfleger, S., 2023. Frequency dependent lpm for soil-structure interaction under induced seismicity URL: <https://collab.dvb.bayern/pages/viewpage.action?pageId=71122813>.
- Kaszynski, A., 2021. pyansys: Pythonic interface to MAPDL. URL: <https://doi.org/10.5281/zenodo.4009466>, doi:10.5281/zenodo.4009466.
- Liutkus, A., 2015. Scale-space peak picking. Ph.D. thesis. Inria Nancy-Grand Est (Villers-lès-Nancy, France).
- Luczak, M., Manzato, S., Peeters, B., Branner, K., Berring, P., Kahsin, M., 2014. Updating finite element model of a wind turbine blade section using experimental modal analysis results. Shock and Vibration 2014, 684786.
- Luong, H.T., Zabel, V., Lorenz, W., Rohrmann, R.G., 2017. Vibration-based model updating and identification of multiple axial forces in truss structures. Procedia Engineering 188, 385–392.
- Magalhães, F., Cunha, A., Caetano, E., 2009. Online automatic identification of the modal parameters of a long span arch bridge. Mechanical Systems and Signal Processing 23, 316–329.
- Mao, Q., Mazzotti, M., DeVitis, J., Braley, J., Young, C., Sjoblom, K., Aktan, E., Moon, F., Bartoli, I., 2019. Structural condition assessment of a bridge pier: A case study using experimental modal analysis and finite element model updating. Structural Control and Health Monitoring 26, e2273.
- Mitchell, M., 1998. An introduction to genetic algorithms. MIT press.

- Mosquera, V., Smyth, A.W., Betti, R., 2012. Rapid evaluation and damage assessment of instrumented highway bridges. *Earthquake engineering & structural dynamics* 41, 755–774.
- Ogier, E., 2025. Hierarchical clustering. URL: <https://www.mathworks.com/matlabcentral/fileexchange/56844-hierarchical-clustering>.
- Pu, Q., Hong, Y., Chen, L., Yang, S., Xu, X., 2019. Model updating-based damage detection of a concrete beam utilizing experimental damped frequency response functions. *Advances in Structural Engineering* 22, 935–947.
- Qin, S., Kang, J., Wang, Q., 2016. Operational modal analysis based on subspace algorithm with an improved stabilization diagram method. *Shock and Vibration* 2016, 7598965.
- Ren, W.X., Chen, H.B., 2010. Finite element model updating in structural dynamics by using the response surface method. *Engineering structures* 32, 2455–2465.
- Renders, J.M., Flasse, S.P., 1996. Hybrid methods using genetic algorithms for global optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 26, 243–258.
- Reynders, E., Maes, K., Lombaert, G., De Roeck, G., 2016. Uncertainty quantification in operational modal analysis with stochastic subspace identification: Validation and applications. *Mechanical Systems and Signal Processing* 66, 13–30.
- Schulz, E., Speekenbrink, M., Krause, A., 2018. A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of mathematical psychology* 85, 1–16.
- Shah, V., Bohm, G., Nahavandi, A., 1979. Modal superposition method for computationally economical nonlinear structural analysis .
- Srinivas, V., Ramanjaneyulu, K., Jeyasehar, C.A., 2011. Multi-stage approach for structural damage identification using modal strain energy and evolutionary optimization techniques. *Structural Health Monitoring* 10, 219–230.

- Talbi, E., 2009. Metaheuristics: From design to implementation. John Wiley & Sons google schola 2, 268–308.
- Tran-Ngoc, H., Khatir, S., De Roeck, G., Bui-Tien, T., Nguyen-Ngoc, L., Abdel Wahab, M., 2018. Model updating for nam o bridge using particle swarm optimization algorithm and genetic algorithm. Sensors 18, 4131.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17, 261–272. doi:10.1038/s41592-019-0686-2.
- Williams, C.K., Rasmussen, C.E., 2006. Gaussian processes for machine learning. volume 2. MIT press Cambridge, MA.
- Wolf, J.P., 1994. Foundation vibration analysis using simple physical models. Pearson Education.
- Zhu, C., Byrd, R.H., Lu, P., Nocedal, J., 1997. Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. ACM Transactions on mathematical software (TOMS) 23, 550–560.
- Zhu, Q., Xu, Y.L., Xiao, X., 2015. Multiscale modeling and model updating of a cable-stayed bridge. i: Modeling and influence line analysis. Journal of Bridge Engineering 20, 04014112.