Proceedings of the Conference on Computational Engineering and Science JSCES
Vol. 29, Jun. 2024

Sequential Updating of Building Fragility Functions through City-Scale Seismic Simulations and Sensor Integration

Dongyang Tang [1)], Shuji Moriguchi [2)], Reika Nomura [3)], Jia Guo [4)], Kazuya Nojima [5)], Masaaki Sakuraba [6)], Susumu Ohno [7)] and Kenjiro Terada [8)]

[1)] Graduate student, (Civil and Environmental Engineering, Tohoku University, 980-0845, Sendai, Japan, E-mail: tang.dongyang.r3@dc.tohoku.ac.jp)
[2)] Associate Professor (International Research Institute of Disaster Science, Tohoku University, 980-0845, Sendai, Japan)
[3)] Assistant Professor (International Research Institute of Disaster Science, Tohoku University, 980-0845, Sendai, Japan)
[4)] Associate Professor (Division of Environmental Science and Technology, Kyoto University, 606-8502, Kyoto, Japan)
[5)] Visiting Associate Professor (Nippon Koei Co., Ltd.)
[6)] Visiting Professor (Nippon Koei Co., Ltd.)
[7)] Associate Professor (International Research Institute of Disaster Science, Tohoku University, 980-0845, Sendai, Japan)
[8)] Professor (International Research Institute of Disaster Science, Tohoku University, 980-0845, Sendai, Japan)

This study proposes a framework for assessing the seismic risk of buildings in large-scale urban areas. City-scale numerical simulations and sensor networks are integrated to update building-specific fragility functions. The proper orthogonal decomposition (POD) technique plays a crucial role in analyzing numerical simulation data. The POD-based approach helps to represent large-scale simulation results in a more compact and meaningful way, a crucial step in understanding seismic impacts. The framework also highlights the importance of sparse sensor distribution. Furthermore, this study uses cloud analysis and Bayesian updating to update the fragility functions of buildings based on sensor data. Fragility functions for all buildings are created based on cloud analysis using numerical simulation data, and then progressively refines them by incorporating sensor data through Bayesian inference. This dual-stage approach allows for a rapid risk assessment of all buildings in a target area, with continuous improvements to fragility functions by the sensor data.

Key Words : *Seismic Risk Assessment, Numerical simulation, Proper orthogonal decomposition, Sparse sensing, Fragility Function*

1. INTRODUCTION

This study proposes a framework for assessing the seismic risk of buildings in large-scale urban areas, which integrates city-scale numerical simulations and a sparse sensor network. The key aspect of this approach is the use of numerical simulation to create fragility functions for each building, and to identify efficient sensor distribution. Furthermore, the data obtained from the sensors are used to update the fragility functions, progressively enhancing their accuracy and providing a more precise assessment of seismic impacts on urban structures.

A critical part of this study is applying POD to analyze numerical simulation data. Proper Orthogonal Decomposition (POD) can identify principal components from numerical simulation data, enabling effective data decomposition and reconstruction. This approach helps represent large-scale simulation results more compactly and meaningfully, which is crucial for understanding seismic impacts. Additionally, the framework also highlights the importance of sparse sensor distribution. The study strategically places sensors to maximize data collection efficiency while minimizing the number of sensors. This optimization is vital for practical use in urban areas, where large sensor networks may be impractical.

The fragility functions for all buildings are created using simulation data based on the concept of the cloud analysis. Cloud analysis uses the linear regression in the logarithmic scale by least squares to establish the relationship between engineering demand parameter (EDP) and intensity measure(IM). Obtained fragility functions are then updated using the sensor data based on Bayesian inference. This updating makes the functions more accurate.

In this study, a trial calculation is finally conducted to verify the effectiveness of the proposed framework, focusing on Sendai City as the target area.

2. Numerical simulation

The following two methods were employed to calculate the propagation of seismic waves from the fault to the ground surface and the response of buildings on the ground.

(1) Stochastic Green's Function

The Stochastic Green's function method expands on the empirical Green's function method, which was initially introduced by Irikura[2]. The empirical Green's function method relies on observed records as Green's functions, presuming that the deep and shallow subsurface structures at the observation point are already integrated into the observed records. On the other hand, this method serves as an effective alternative when appropriate observation records cannot be obtained.

In the application of the Stochastic Green's function method by Dan and Sato[3], the fault surface is segmented into small sub-faults, and Boore's stochastic source model is taken into account for each sub-fault to compute the Green's functions. The deep subsurface structure is treated as a one-dimensional layered structure for ground response analysis. Random phase characteristics are attributed to this Green's function, and waveform synthesis is conducted by Irikura[2] to derive the seismic waveforms when the entire fault experiences rupture. This study utilizes the program provided by the National Research Institute for Earth Science and Disaster Resilience[4].

(2) Integrated Earthquake Simulator(IES)

IES is a program that is linked to a Geographic Information System (GIS) and incorporates earthquake motion simulation, structural response simulation, and response behavior simulation[5]. Wave propagation simulation: It outputs synthesized earthquake waves based on the fault mechanism. The propagation of waves passing through the crust is calculated, and the amplification of waves near the surface is calculated taking into account the non-linear characteristics of the 3-dimensional topographical effect and the shallow soil layer.
Structural response simulation: It calculates the response for all structures in the targeted area, including residential buildings, concrete infrastructure structures, geological structures, transportation networks, etc. It is necessary to choose an appropriate analysis method depending on the structure of the building.
Response behavior simulation: It is possible to analyze evacuation from building damage, crisis management, and restoration plans.
Finally, by modifying the fault model, multiple scenarios were generated, and using both the Stochastic Green's Function and IES, training and validation datasets were computed.

3. Sparse sensor distribution

The first part of this framework is a prediction model, based on proper orthogonal decompsition and sparse sensing, it can predict both seismic motion and structure response in a large area.
Proper Orthogonal Decomposition(POD) is mathematical operation that can extract modes from original data, allowing for mode decomposition based on the theory of singular value decomposition[6]. Let $\boldsymbol{x}_i$ represent the n-dimensional simulation result for a specific case i, and define the data $\boldsymbol{X}$ by arranging N cases in a row direction.

$$\boldsymbol{X} = \begin{bmatrix} | & & | \\ \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_N \\ | & & | \end{bmatrix}. \tag{1}$$

Let σ_j and v_j be the singular values and right singular vectors obtained, respectively, and let V be a matrix in which the eigenvectors are arranged in column direction. Using these, consider the singular value decomposition of $\mathbf{X}$ as follows:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \tag{2}$$

By retaining only r columns from the r singular values and their corresponding modes $\boldsymbol{U}$(left singular vector), we obtain an $n \times r$ matrix. Similarly, by retaining only r columns from the singular vector matrix $\boldsymbol{V}$, we obtain an $N \times r$ matrix. $\boldsymbol{x}$ as follows:

$$\boldsymbol{X} = \boldsymbol{U}_r\boldsymbol{\Sigma}_r\boldsymbol{V}_r^T = \sum_{k=1}^{r} \boldsymbol{u}_k(\sigma_k \boldsymbol{v}_k^T), \tag{3}$$

in which

$$\boldsymbol{x}_i = \sum_{j=1}^{N} (\sigma_j v_{ij}^T)\boldsymbol{u}_j = \sum_{j=1}^{N} z_{ij}\boldsymbol{u}_j. \tag{4}$$

Here σ_k represents the k-th singular value, and v_{ik} represents the i-th component of $\boldsymbol{V}$ in the k-th column. Thus, the data x_i for a specific case i is given by the sum of the products of the singular values, the i-th component of $\boldsymbol{U}$, and the i-th component of $\boldsymbol{V}$, for k ranging from 1 to r.

Then we consider the following linear system:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{U}\boldsymbol{z} = \boldsymbol{C}\boldsymbol{z}. \tag{5}$$

The equation describes a system where $\boldsymbol{y} \in \mathrm{R}^p$ is the observation vector, $\boldsymbol{H} \in \mathrm{R}^{p\times n}$ represents the sparse sensor location matrix, $\boldsymbol{U} \in \mathrm{R}^{n\times r}$ denotes the sensor candidate matrix, $z \in \mathrm{R}^r$ is the latent state vector, and $\boldsymbol{C} \in \mathrm{R}^{p\times r}$ is the measurement matrix, with $\boldsymbol{C} = \boldsymbol{H}\boldsymbol{U}$ being their relation. In the matrix $\boldsymbol{H}$, each row has a unity element indicating the sensor location, with all other elements being zero. When observations are subjected to uniform independent Gaussian noise, represented as $N(0, \sigma^2\boldsymbol{I})$, the estimated parameters $\hat{z}$ are derived applying a pseudo-inverse method as

$$\hat{\boldsymbol{z}} = \begin{cases} \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}\boldsymbol{y}, & p \le r, \\ (\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T\boldsymbol{y}, & p > r. \end{cases} \tag{6}$$
$$\text{subject to } \boldsymbol{C} \in \mathrm{R}^{p\times r},\ p, r \in \mathrm{N}.$$

A D-optimal design aims to minimize the determinant of the error covariance matrix. This can also be seen to minimize the determinant of the error covariance matrix.

$$\max f_D, \quad f_D = \begin{cases} \det\left(\boldsymbol{C}\boldsymbol{C}^{\boldsymbol{T}}\right) & (p \le r) \\ \det\left(\boldsymbol{C}^{\boldsymbol{T}}\boldsymbol{C}\right) & (p > r) \end{cases} \tag{7}$$

To maximaze the objective function, genetic algorithm is adopted in this study.

4. Fragility Function

Another key part of this framework is seismic assessment based on predicted data by sensor network, which contains construct initial fragility function and sequential update of such a function by bayesian inference.

(1) Cloud analysis

Fragility functions are derived from a structural assessment of the system (in the case of analytical form). In simpler terms, fragility can be defined as the susceptibility of a structure to collapse or being damaged. It is a continuous function showing the probability of exceeding a certain limit state (LS) for a specific level of ground motion intensity measure (IM)[8] as blow

$$Fragility = P[LS|IM = im] \tag{8}$$

Cloud analysis uses the linear regression in the logarithmic scale by least squares to establish the relationship between engineering demand parameter (EDP) and IM as follows:

$$\begin{gathered} E[\ln EDP \mid IM] = \ln \mu_d = \ln a + b \ln IM \\ \sigma_d = \sqrt{\sum_{j=1}^{N} (\ln EDP_j - \ln \mu_d)^2 / (N-2)} \end{gathered} \tag{9}$$

where IM, EDP$_j$ = EDP obtained from the j-th ground motion, a and b are regression coefficients; and N is number of ground motions. The fragility function is expressed as the damage probability that EDP exceeds the pre-defined value threshold for each limit state (LS) conditional on IM. This probability can be derived based on the above linear relationship between EDP and IM under the lognormal probability distribution[9] as

$$\begin{aligned} & P_f[EDP \geq LS \mid IM, \eta, \beta] \\ & = \Phi\left\{\frac{\ln(\mu_d) - \ln(LS)}{\sigma_d}\right\} = \Phi\left\{\frac{\ln(IM) - \ln(\eta)}{\beta}\right\}, \end{aligned} \tag{10}$$

where $\Phi(\cdot)$ is standard normal cumulative distribution function (CDF); η is median of the fragility function, i.e, $\ln(\eta) = [\ln(LS) - \ln(a)]/b$; and β is dispersion of the fragility function, i.e., $\beta = \sigma_d/b$. Note that Eq.(10) is a two parameter (η and β) fragility function given IM. When the fragility function is expressed with respect to ground motion intensity (denoted by x), such as peak ground acceleration (PGA), the mathematical form can be expressed as follows:

$$P_f = F_X(x; \mu, \sigma) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left[\frac{\ln(x) - \ln(\mu)}{\sigma\sqrt{2}}\right], \tag{11}$$

where erf is error function, $\sigma =$ is dispersion, i.e., and μ is natural logarithm of the median ground motion intensity.

Limit state refers to a specific level of damage or failure that is used to define fragility functions. In this case, the limit states of "moderate" was chosen from HAZUS[12] to develop the fragility functions.

(2) Bayesian updating

To update the fragility functions shown in Eq. (11) using the Bayesian framework under a seismic event, the random variables $\boldsymbol{\theta}$ need to be updated contains μ and σ; i.e., the natural logarithm of the meanground motion intensity and the dispersion[10], respectively. More specifically,

$$\begin{aligned} \theta_1 &= \mu \\ \theta_2 &= \sigma. \end{aligned} \tag{12}$$

We define a prior distribution for each parameter because the two are independent of each other, and the joint prior distribution is the product of the two prior distributions as

$$f'(\boldsymbol{\theta}) = P(\theta_1 \mid \mu_1, \sigma_1) \cdot P(\theta_2 \mid a_2, b_2). \tag{13}$$

On the other hand, the likelyhood function of fragility is shown as blow[11],

$$L(\varepsilon \mid \theta) = \prod_{i=1}^{n} [1 - F_X(x_i; \theta)]^{1-\varepsilon_i} F_X(x_i; \theta)^{\varepsilon_i} \tag{14}$$

where ε_i is expressed as a binary number. For each limit state, when the building collapse, then $\varepsilon_i = 1$, otherwise, $\varepsilon_i = 0$. x_i is observed intensity measure(PGA) for each building after an earthquake occured. Here, $F_X(x_i; \theta)$ is the fragility function given in Eq.(11). According to Bayesian theorem, the updated distribution of the parameters, called the posterior distribution (denoted by $f''(\boldsymbol{\theta})$), which combines the existing knowledge (denoted by $f'(\boldsymbol{\theta})$) and the newly obtain information (denoted by $L(\varepsilon \mid \theta)$), is given as follows:

$$f''(\boldsymbol{\theta}) \propto L(\boldsymbol{\varepsilon} \mid \boldsymbol{\theta}) f'(\boldsymbol{\theta}) \tag{15}$$

The posterior distribution is computed numerically by Markov chain Monte Carlo (MCMC) sampling, and then the marginal distributions of the two parameters θ_1, θ_2 are computed separately. Then the expectation for each marginal distribution is the desired updated parameter for fragility function and .

5. Seismic risk assessment in Sendai city

Based on Nagamachi-Rifu fault parameters and the recipe reported by Irikura (2011)[13], four more specific fault models were constructed by varying the original parameters of the Nagamachi-Rifu fault. The magnitudes for the training data set are 6.3, 7.0, and 7.5, while for the validation data set is 6.5, measured on the Japan Meteorological Agency magnitude scale, while the uncertainty of seismic wave, three different patterns of seismic wave are also considered for each magnitude.

Utilizing the Stochastic Green's Function method enabled the generation of seismic waves and the IES was used to acquire ground-level waveforms and building response. This approach yielded a data set that included peak ground acceleration and maximum inter-story drift angle. Following the principles of POD, spacial modes of peak ground acceleration and maximum inter-story drift angle

Fig. 1 Target area (32334 buildings)

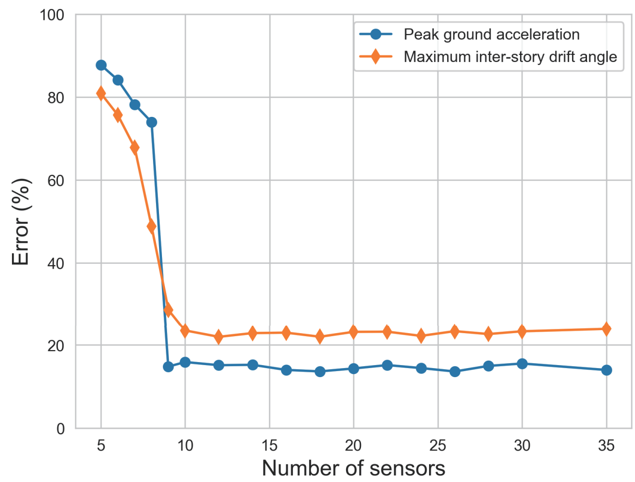

Fig. 2 Error of prediction

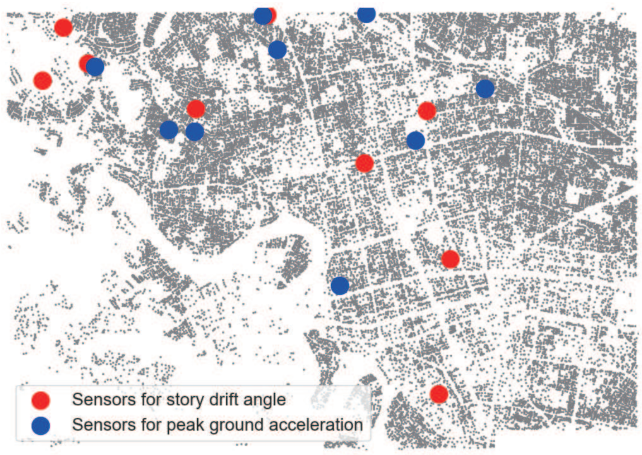

Fig. 3 Sensor distribution

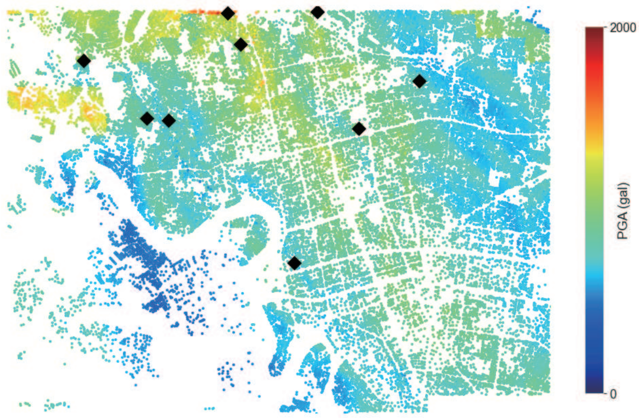

Fig. 4 Prediction

were obtained and allowed for the calculation of optimal sensor distribution. From Fig.2, it can be concluded that at least 9 sensors are needed for the estimation of the seismic motion and building responses. The obtained sensor distribution are shown in Fig3.
At the same time, the numerical simulation results are also used to construct the initial fragility function for each building type in Sendai city. As described in Eq.(10), both peak ground acceleration and maximum inter-story drift angle were used as intensity measures (IM) and engineering demand parameters (EDP), respectively. After categorizing the buildings according to their height categories, the initial fragility functions were calculated. Subsequently, the validation set was used as observational data. For each category, only a portion of this data was used as observation, them by the prediction model of this framework. For the validation case, at magnitude 6.3, by comparing the results obtained from the prediction with the theoretical numerical simulation, it can be seen that there is still an error of almost 20%. Then the predicted PGA map is used as observation, using the initial fragility function, according to Equation 10, a preliminary risk assessment can be obtained, as shown in Fig. 6, and then, according to the previously defined damage threshold, combined with the predicted observation PGA, an update can be made for the current fragility function, and then the risk assessment is performed again to obtain as shown in Fig. 7.

6. Conclusion

This study proposed novel risk assessment of buildings in urban area, and a trial calculation was conducted considering a part of Sendai city to examine the performance of the proposed framework. The obtained results suggest the utility of the proposed framework, and further research development is anticipated in the future.

However, using limit states provided by HAZUS (thresholds of maximum inter-story drift angle), the study categorized buildings based on floor and structural attributes and developed distinct fragility functions for each category. The reliability of these limit states is a matter of ongoing debate, especially for buildings in Japan, where adopting local standards might be more accurate. Moreover, due to insufficient data points for individual buildings, similar building types were combined to compute fragility functions. Future research should increase the data volume by altering fault parameters for more scenarios and considering greater uncertainties. This will not

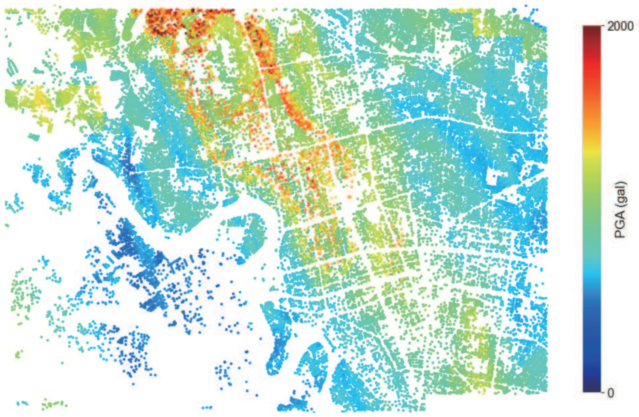

Fig. 5 Simulation

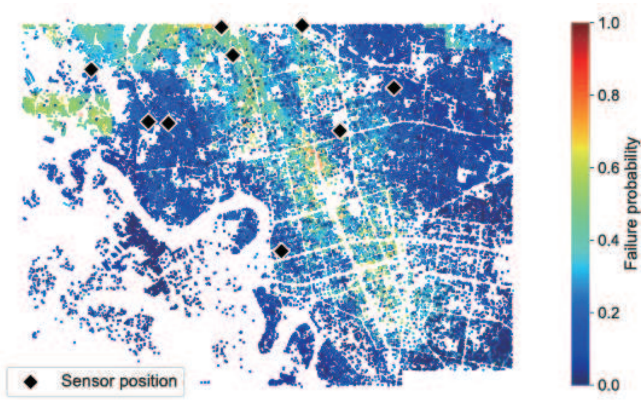

Fig. 6 Risk assessment (before update)

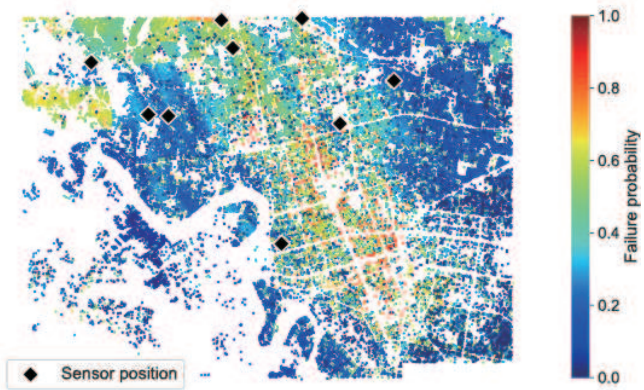

Fig. 7 Risk assessment (after update)

only address the data scarcity for individual buildings but also improve the sparse sensing models' accuracy, allowing for more precise predictions. Enhancing data volume is key to continuously updating fragility functions for each building using Bayesian inference, making them more representative of each building's unique characteristics.

I should be also noted that this study primarily utilizes PGA as the intensity measure for surface seismic motions. This measure, however, may not accurately reflect the impact of an earthquake, as it does not precisely represent the complexities of seismic waves. Future studies should employ more comprehensive and precise intensity measures that better capture the nuances of earthquake impacts on urban structures.

REFERENCES

[1] Irikura, K.: Prediction of strong acceleration motion using empirical Green's function, Proc. 7th Japan Earthq. Eng. Symp., vol. 151, 1986.

[2] Irikura, Kojiro: Prediction of strong acceleration motion using empirical Green's function, Proc. 7th Japan Earthq. Eng. Symp., vol. 151, pp. 151-156, 1986.

[3] Dan, K. and Sato, T.: Strong-motion prediction by semi-empirical method based on variable-slip rupture model of earthquake fault; J. Struct. Constr. Eng., AIJ, No.509, pp. 49-60, 1998

[4] National Research Institute for Earth Science and Disaster Resilience: J-SHIS Earthquake Hazard Station, 2011. Available from: https://www.j-shis.bosai.go.jp/.

[5] Hori, M.: Introduction to computational earthquake engineering, Imperial College Press, 2011.

[6] Liang, Y. C., et al.: Proper orthogonal decomposition and its applications—Part I: Theory, Journal of Sound and Vibration, Vol. 252, No. 3, pp. 527-544, 2002.

[7] Nakai, K., Yamada, et al.: Effect of Objective Function on Data-Driven Greedy Sparse Sensor Optimization, in IEEE Access, Vol. 9, pp. 46731-46743, 2022.

[8] Saouma, V.E., Hariri-Ardebili, M.A.: Fragility Functions, In Aging, Shaking, and Cracking of Infrastructures, Springer, Cham, 2021.

[9] Padgett, Jamie E., and DesRoches, Reginald., et al: Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios, Earthquake Engineering & Structural Dynamics, Vol. 37, No. 5, pp. 711-725, 2008.

[10] Li, Jian and Spencer, Billie F., et al.: Bayesian Updating of Fragility Functions Using Hybrid Simulation, Journal of Structural Engineering, Vol. 139, No. 7, pp. 1160-1171, 2013.

[11] Koutsourelakis, Phadeon-Stelios: Assessing structural vulnerability against earthquakes using multi-dimensional fragility surfaces: A Bayesian framework, Probabilistic Engineering Mechanics, Vol. 25, No. 1, pp. 49-60, Elsevier, 2010.

[12] FEMA: Hazus-MH 2.1 technical manual: Earthquake model, Federal Emergency Management Agency, Mitigation Division, Washington, DC, 2012.

[13] Irikura, K., Hiroe, M.: Recipe for Predicting Strong Ground Motion from Crustal Earthquake Scenarios, Pure and Applied Geophysics, Vol. 168, No. 1, pp. 85-104, 2011.