WILEY | Hindawi

Research Article

A Development of Seismic Fragility Curve of Box-Type Squat Shear Walls Using Data-Driven Surrogate Model Techniques

Sangwoo Lee [ID],[1] **Shinyoung Kwag** [ID],[2] **and Bu-seog Ju** [ID][1]

[1]Department of Civil Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do, Republic of Korea 17104
[2]Department of Civil and Environmental Engineering, Hanbat National University, Daejeon, Republic of Korea 34158

Correspondence should be addressed to Shinyoung Kwag; skwag@hanbat.ac.kr and Bu-seog Ju; bju2@khu.ac.kr

Recent experience with the strongest earthquakes greater than magnitude 5.4 in Korea leads to public interest in the safety and resilience of critical infrastructures. There are many nuclear power plants near the epicenters of the earthquakes. Nuclear power plants are essential infrastructures that provide stable and enough energy for human life, and simultaneously, control systems for the safety and security of nuclear power plants are critical due to the risk of nuclear accidents on public health and the environment. The nuclear area uses probabilistic risk assessment to estimate the risk of structures, systems, and components in nuclear power plants, and the evaluation of the fragility curve is a key process for probabilistic risk assessment. The challenges of a seismic fragility analysis lie in estimating the influence of various uncertainties in material, geometry, and earthquake and improving the existing fragility analysis methods, which require time-consuming nonlinear time history analysis. Thus, this paper conducts a multivariate seismic fragility analysis using surrogate models for reinforced concrete squat shear walls and proposes a simplified closed form equation for a chosen surrogate seismic demand model. The surrogate models are trained and validated by several approaches: response surface method, support vector machine, Gaussian process regression, and neural network. In addition, a correlation analysis is used to evaluate the relative importance of the variables to the seismic demand to simplify the surrogate model further. Finally, simplified surrogate models based on the importance of the variables are proposed as closed form of polynomials, and the performance of these models on the fragility analysis is evaluated.

1. Introduction

The Great East Japan Earthquake severely damaged infrastructures such as telecommunication facilities, power plants, bridges, and buildings. On the day of the earthquake, 8.5 million households lost electricity, 1.5 million were still powerless on March 13, and 0.3 million had no power supply even a week after the earthquake. Due to the lack of power, communication service outages were widespread [1]. Also, Korea has recently experienced the strongest earthquakes greater than magnitude 5.4 in Gyeongju and Pohang, which are beyond-design earthquakes, as shown in Figure 1. They are the largest earthquakes since the beginning of the instrumental earthquake records in 1978. These earthquakes caused life and property damage: Gyeongju and Pohang earthquakes resulted in 111 victims, 23 injured

people, and 5,368 damaged properties and 135 injured people and 57,000 damaged properties, respectively [2, 3]. In addition, Korea has the highest density of NPPs, and many NPPs in Korea are operated on the country's east coast near the epicenters of the earthquakes, as shown in Figure 1. In particular, the Kori NPP is located in Busan metropolitan city, which is the second largest city with a population of 3.5 million, and the Wolseong NPP in Gyeongju City is located within around 20 km of Ulsan metropolitan city with a population of over 1.1 million. As a result, Korean society pays strong attention to the safety of nuclear power plants (NPPs). A probabilistic seismic risk assessment of the structures, components, and systems in an NPP is important to estimate the seismic fragility of them and the overall risk of the NPP. The overall risk is estimated by the convolution of hazard and fragility curves as the probability of occurrence of
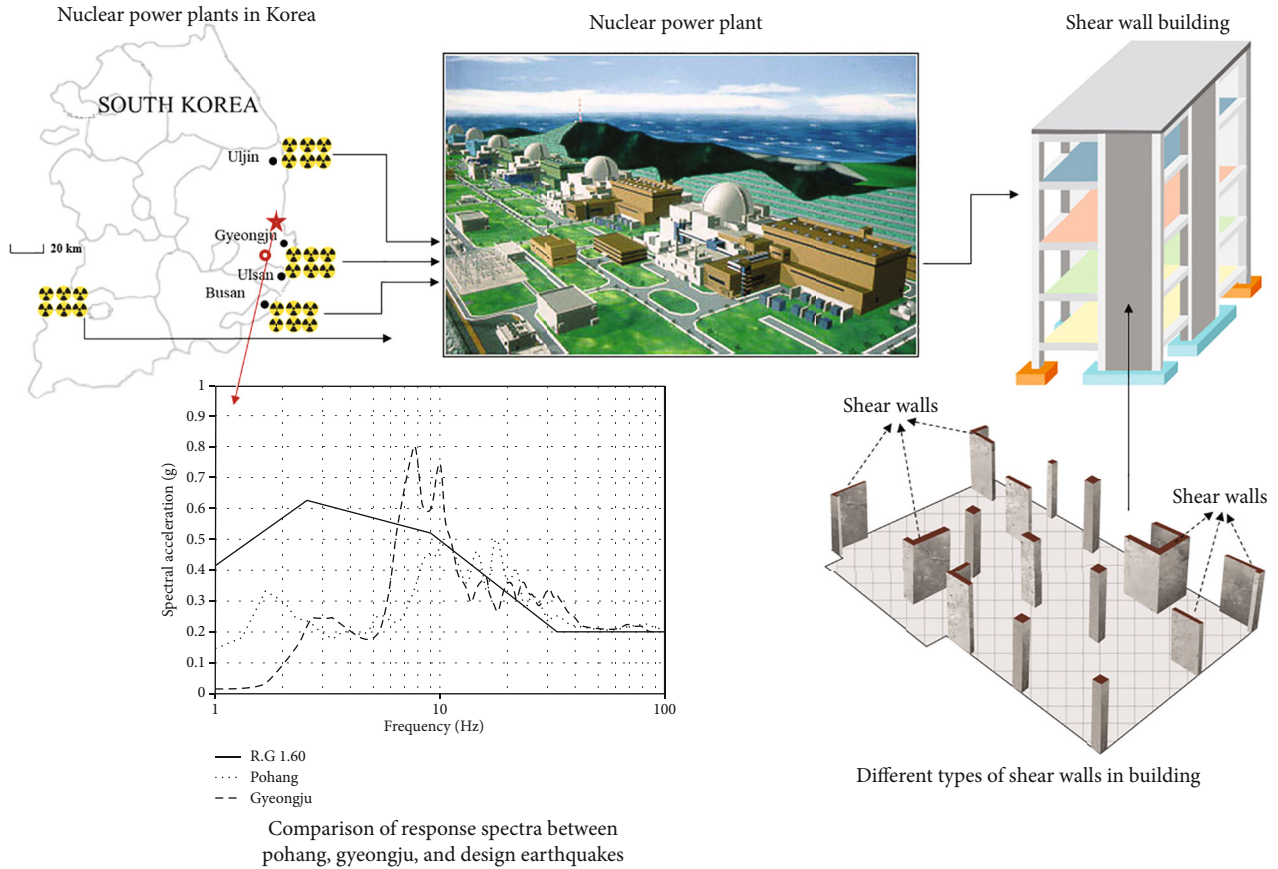
Figure 1: Different type of shear walls in NPPs and response spectra of Gyeongju (O) and Pohang (⋆) earthquakes.

each consequence; therefore, an accurate evaluation of the risk is necessary to define and quantify the levels of structural safety and resilience.

Reinforced concrete (RC) squat shear walls are the most typical structural systems used in NPPs because they effectively resist lateral forces like earthquakes and winds. The NPPs consist of various types of squat shear walls, as shown in Figure 1. The squat shear wall has a height-to-length ratio of less than two, resulting in shear-dominant lateral behaviors with small lateral deformation [4]. In the seismic risk assessment, calculating a fragility curve defined as the conditional probability of a failure exceeding a specific capacity for a given ground motion intensity measure (IM) is a key process. The fragility analysis must include all sources of uncertainties such as ground motions, the geometry, and material properties of structures. However, there is a lack of studies on the seismic fragility analysis of the squat shear wall. Many earlier studies [5–10] mainly focused on experiment and simulation works for investigating and predicting the nonlinear behavior of squat shear walls under monotonic and cyclic loading. Furthermore, while Syed and Gupta [11] conducted a fragility analysis for an RC squat shear wall with uncertainty in ground motions, uncertainties in the geometry of the shear wall were not included.

Cloud analysis (CA) and multiple stripe analysis (MSA) are the conventional methods used to develop seismic fragility curves in many studies [11–16]. Both analyses employ nonlinear time history analysis (NTHA) to collect the seismic demands. In the cloud analysis, a single-parameter demand model is estimated by linear regression between demands and unscaled ground motions in a lognormal space. In the MSA, scaled ground motions to IMs of interest are used, and the probability of a failure at a given IM level is calculated as a ratio of the number of samples in which the seismic demand obtained from NTHA exceeds the capacity to the total number of samples ($p_{f,IM} = N_{f,IM}/N$). However, recent fragility analysis studies have shown the need for a multiparameter demand model using surrogate models because the conventional single-parameter demand model has limitations [17–19]. In addition, new NTHAs are repeatedly necessary to reestimate a fragility curve due to changing or updating the input variables; however, NTHA is a very time-consuming process. Several researchers have attempted to generate multiparameter surrogate models for a seismic demand model [16–19]. However, all studies have focused on highway bridges, and few studies have considered RC squat shear walls. The seismic demand characteristics of RC squat shear walls are quite different from those of the highway bridge piers, decks, and bearings. Accordingly, it demands that a seismic demand model be efficiently developed for the fragility of the RC squat shear wall. Also, the current multiparameter surrogate model studies for highway bridges have developed black-box surrogate models (i.e., implicit functions) with a sole focus on the model accuracy; thus, such developed models cannot be accessed
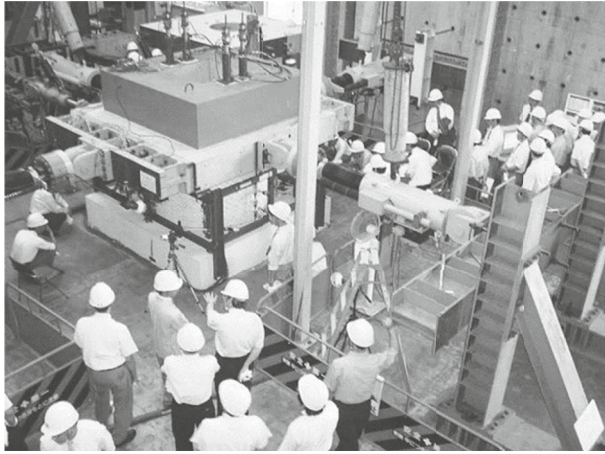
(a) NUPEC experiment [8]
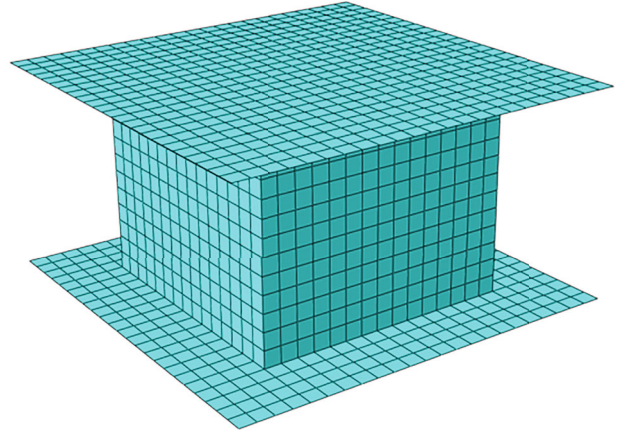
(b) Finite element model

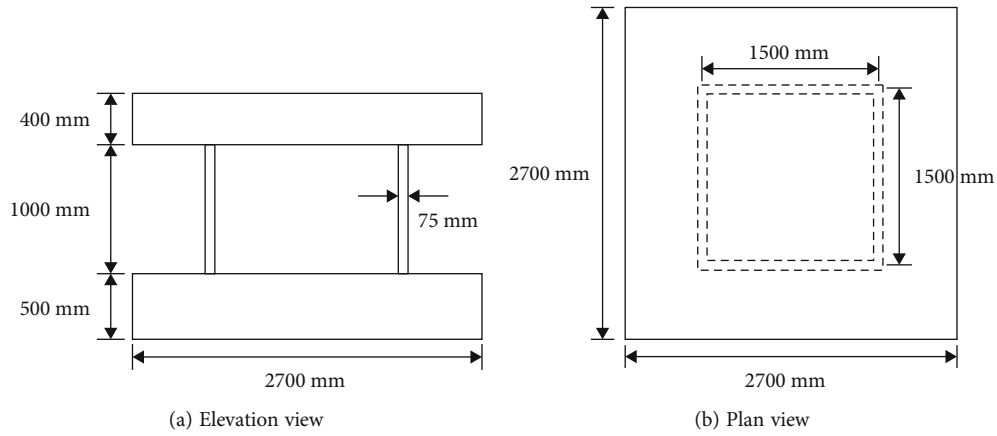FIGURE 2: Experimental test and finite element model of the box-type shear wall system.

(a) Elevation view

(b) Plan view

FIGURE 3: Dimensions of the box-type shear wall system.

TABLE 1: Material properties of shear walls.

	Young's modulus (E_o, MPa)	Compressive and tensile strength (f_c' and f_t, MPa)	Yield strength (f_y, MPa)
Concrete	30,700	41.3 and 1.74	—
Rebar	200,000	—	375

easily by other engineers and researchers, which hinders utilizing such models for design and safety assessment in a comprehensible manner.

The purposes of this study are to (1) develop multiparameter surrogate models of a box-type RC squat shear wall for the ultimate shear forces and (2) determine a best-fitting surrogate model. These models consider the geometric uncertainty of the shear walls as well as the material and ground motion uncertainties. Surrogate models are developed using four different techniques: second-order response surface method (RSM), support vector machine (SVM), Gaussian process regression (GPR), and neural Network (NN). (3) The fragility curves of the box-type shear walls obtained from the surrogate models are accurately estimated

TABLE 2: Concrete damage plasticity parameters.

Dilation angle	Eccentricity	f_{b0}/f_{c0}	K_c	w_c	w_t
30	0.1	1.16	0.67	0.5	0.55

and compared with the ones obtained from the traditional approaches mentioned above. (4) Through a correlation analysis, the relative importance of the variables to the seismic demand is evaluated and reduced parameter surrogate models are suggested. As a result, the multiparameter surrogate models help to efficiently generate the fragility curve for the box-type RC shear wall and to update the seismic fragility curve due to the variation of the input variables without time-consuming simulations.

(a) Load-deformation backbone curves

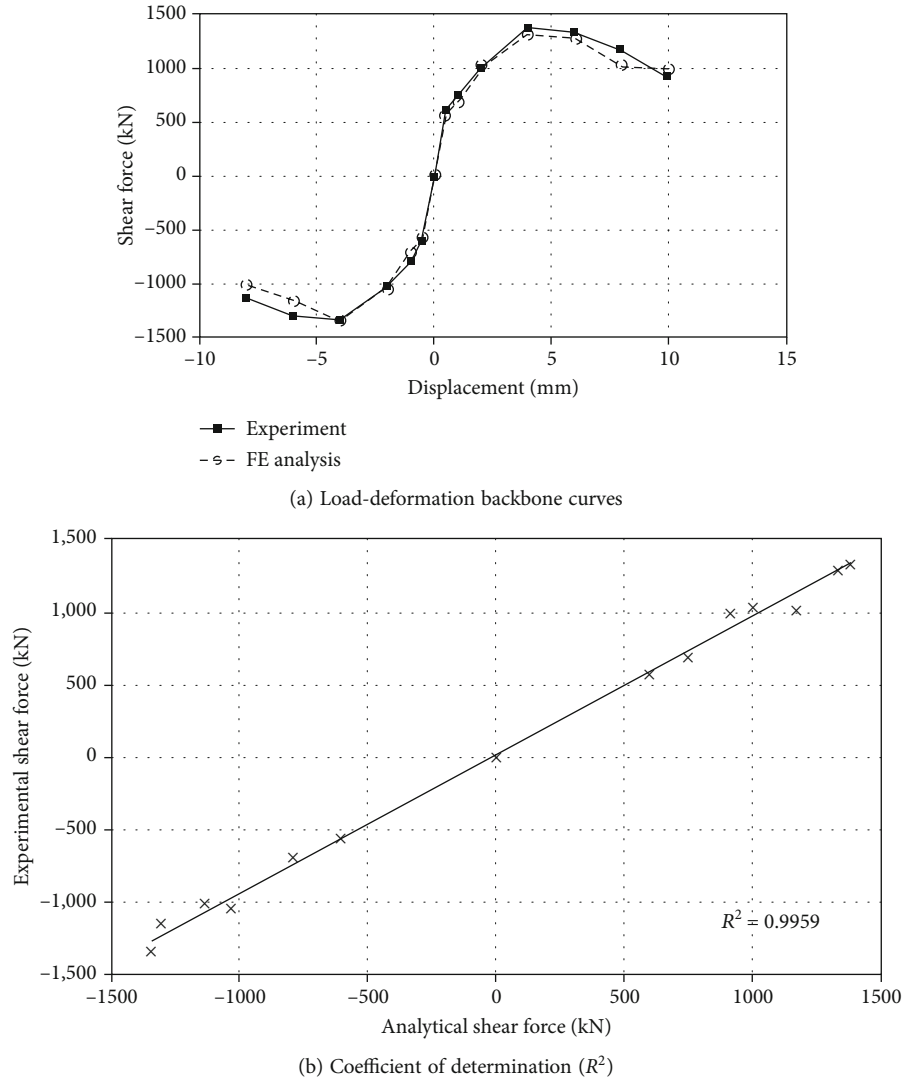(b) Coefficient of determination (R^2)

FIGURE 4: Comparison between experimental and analytical results.

TABLE 3: Uncertainties in the material and geometry included in the shear wall model.

Material	Distribution	Mean (μ)	Coefficient of variation	Reference
f'_c (MPa)	Normal	41.3	0.09	[11]
f_y (MPa)	Normal	475	0.08	[15]
Geometry	Distribution	Lower limit	Upper limit	Reference
L (mm)	Uniform	600	3000	[36]
H/L	Uniform	0.25	2	[36]
L/t	Uniform	7.5	27.5	[36]
P (N)	Uniform	$0.022A_w f'_c$	$0.182A_w f'_c$	[36]
ρ_h	Uniform	0.00125	0.0161	[36]
ρ_v	Uniform	0.00125	0.0287	[36]

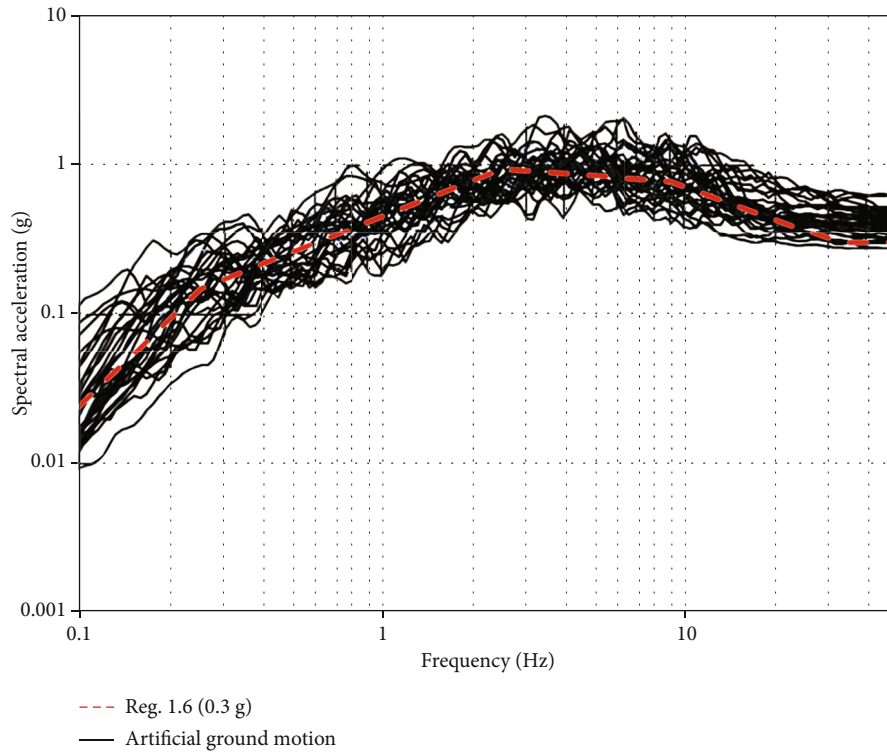A_w: gross area of the web cross section (mm^2).

FIGURE 5: Response spectra of artificial ground motions compared with Regulatory Guide 1.60 response spectra scaled 0.3 g.
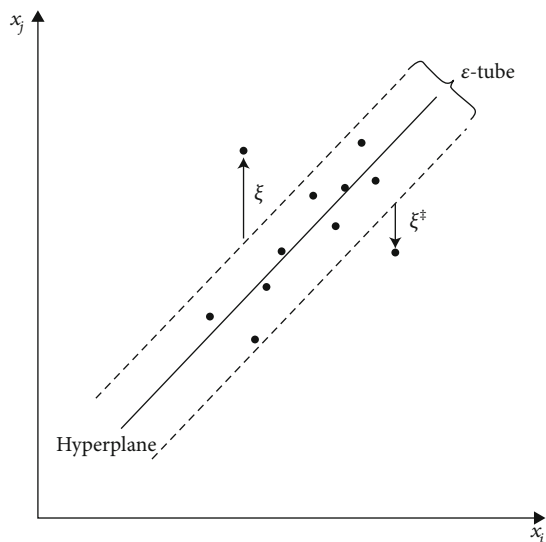

FIGURE 6: Principle of one-dimensional support vector regression model.

2. Finite Element Model of Box-Type Shear Wall for Surrogate Model and Fragility Analysis

For training surrogate models and conducting fragility analysis for the shear walls, developing a finite element model and running NTHAs is essential. This study uses a finite element model of box-type RC shear walls, which was developed and validated with the experimental data [20–24]. The validated finite element model is analyzed considering uncertainties, and corresponding analytical results are used to develop surrogate models and conduct a fragility analysis in the following sections. An experiment on the box-type RC shear walls was conducted by the Nuclear Power Engineering Corporation (NUPEC) to assess the ultimate seismic capacity of the shear walls, as shown in Figure 2(a) [8]. The shear walls consisted of upper slab, lower slab, and four rectangular shear walls with an aspect ratio of 0.67. The box-type shear walls were subjected to multiaxial cyclic loading, and vertical preloading was applied to the upper slab, resulting in a typical value of compressive stress at the bottom of shear walls for the lower story of a building in NPPs. The rectangular shear walls had a double layer of D6 rebar with a rebar ratio of 1.2% in both the horizontal and vertical directions. Figure 3 shows the dimensions of the box-type shear wall system.

The finite element model was developed consisting of a box wall, loading slab, and bottom slab using ABAQUS. An analysis of the shear wall under cyclic loading was performed through the ABAQUS implicit solver. The developed model is shown in Figure 2(b) [20–22]. The walls were modeled by a 4-node shell element with reduced integration (S4R) and a mesh size of 100 mm, assuming that the rebar is perfectly bonded with the concrete. In the experimental observations, there was no damage to the upper and lower slabs, so an elastic isotropic concrete model was applied to the slabs. The bottom slab is fixed at the ground, and multiaxial cyclic loading is applied to the loading slab. To

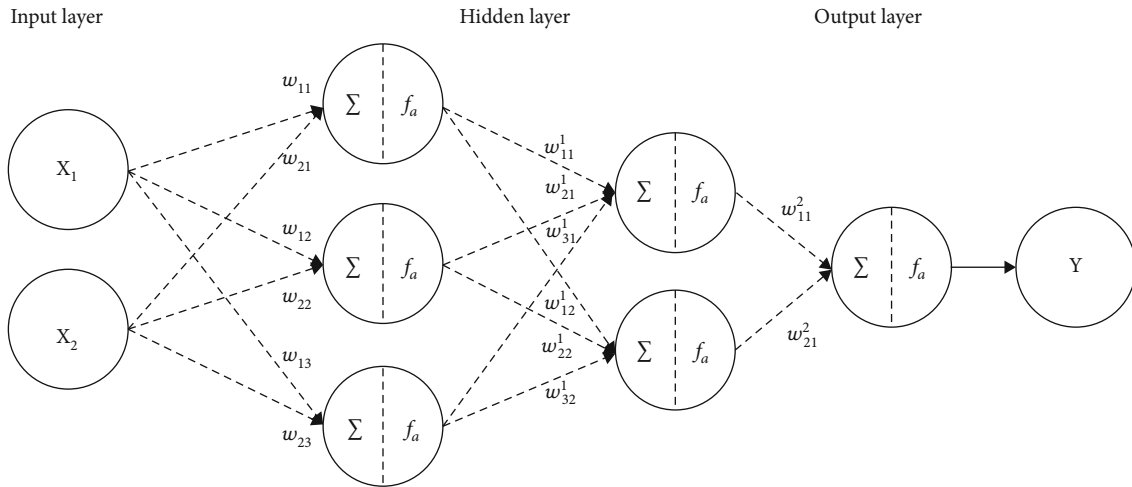FIGURE 7: Principle of neural networks with two hidden layers.
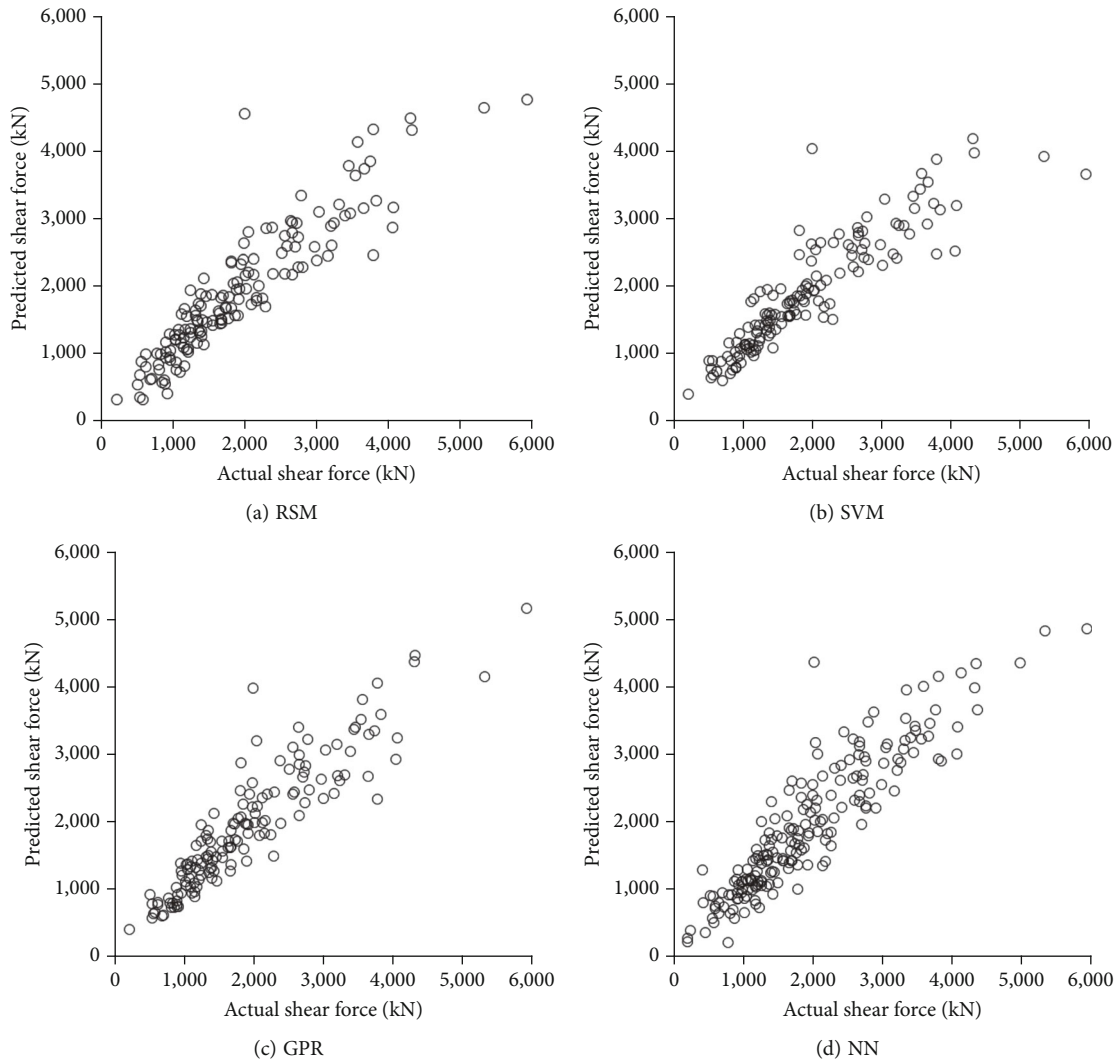

(a) RSM

(b) SVM

(c) GPR

(d) NN

FIGURE 8: Comparison between predicted and actual ultimate shear force of box-type shear walls for training dataset.
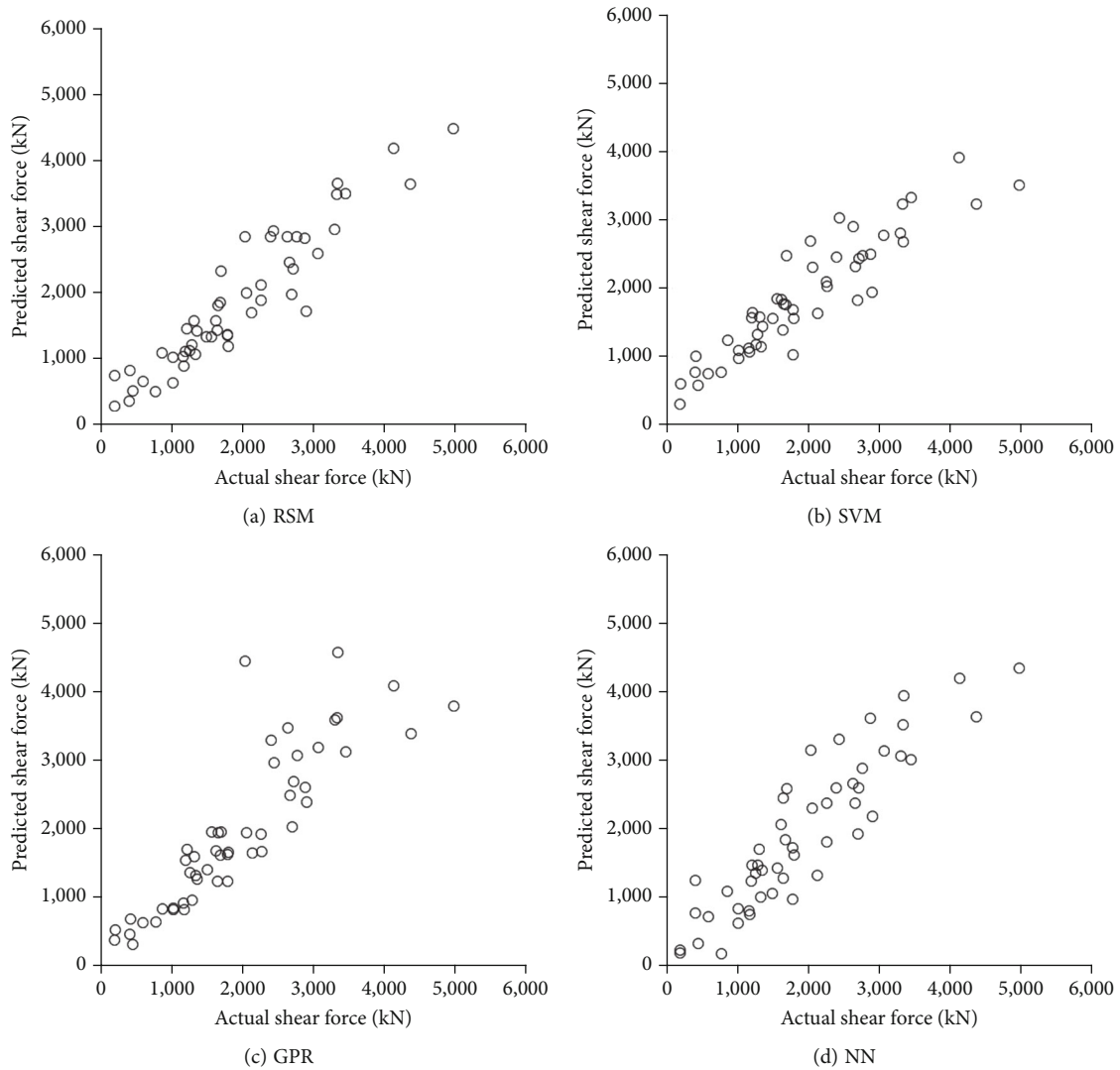
Figure 9: Comparison between predicted and actual ultimate shear force of box-type shear walls for testing dataset.

represent the nonlinear behavior of the system, a concrete damage plasticity model proposed by Lubliner et al. [23] and Lee and Fenves [24] was applied to the rectangular shear walls. It has been used in many studies on the nonlinear behavior of concrete structures [20–22, 25–32]. In addition, the uniaxial concrete tension stiffening model and compression hardening model that were proposed by Maekawa and Okamura [33] and Izumo [34], respectively, were used in this study. Tables 1 and 2 summarize the material properties of the shear wall and parameters of the concrete damage plasticity model [20], where f_{b0} and f_{c0} are the biaxial compressive yield stress and the uniaxial compressive yield stress, K_c is the ratio of the second invariant of the tensile meridian to that of the compressive meridian at initial yield for any given value of the first stress invariant, and w_c and w_t represent the compression and tension recovery factors.

Figure 4(a) shows that the experimentally obtained load-deformation backbone curve coincides well with the corresponding curve obtained from the finite element analysis, particularly at the ultimate (maximum) shear strength which needs to be measured in this study. In addition, the coefficient of determination (R^2) between the experimental and analytical shear forces represents 0.9959, as presented in Figure 4(b); therefore, this finite element model can be used to collect ultimate shear forces for developing surrogate demand models and for conducting a fragility analysis in the following sections.

3. Uncertainty in Material, Geometry, and Ground Motion

All sources of uncertainties must be considered in a fragility analysis. This study considers uncertainties in the material, the geometry of the box-type shear walls, and earthquakes. The uncertainties in the material properties and geometry are presented in Table 3. f_c' and f_y are included as the

TABLE 4: R^2 and RMSE values of surrogate models.

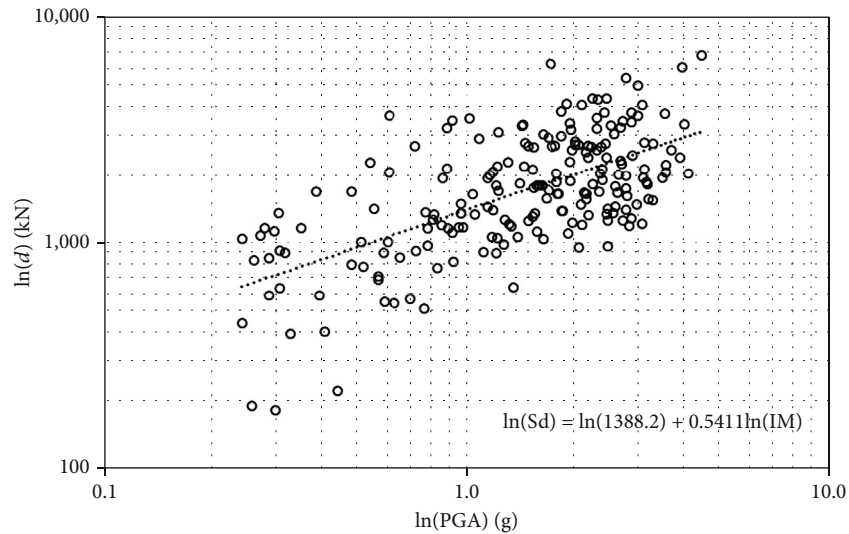|	RSM	SVM	GPR	NN
Training set				
R^2	0.86	0.80	0.86	0.86
RMSE	426	514	422	423
MAE	298	304	294	303
VAF (%)	62.8	55.4	63.1	63.1
Test set				
R^2	0.88	0.82	0.74	0.81
RMSE	378	456	553	472
MAE	294	343	373	376
VAF (%)	65.7	58.3	49.0	56.4
Average				
R^2	0.87	0.81	0.80	0.84
RMSE	402	485	488	448
MAE	296	324	334	340
VAF (%)	64.25	56.85	56.05	59.75

$$\ln(Sd) = \ln(1388.2) + 0.5411\ln(IM)$$

FIGURE 10: Cloud analysis of shear walls: ultimate shear forces and corresponding PGA levels.
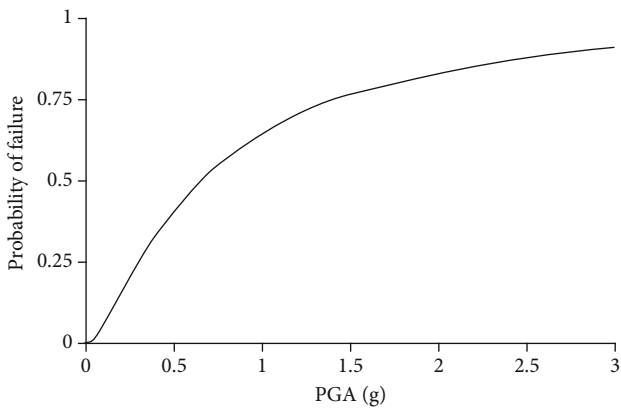
FIGURE 11: Fragility curve for the box-type shear walls using CA.

material uncertainty. The material properties of concrete, such as f'_c, E_o, and f_t, are correlated in nature, thereby being considered by the following equations [35]:

$$E_o = 3,320\sqrt{f'_c} + 6,900, \tag{1}$$

$$f_t = 0.33\left(f'_c\right)^{1/2}. \tag{2}$$

The degree of uncertainties in steel-related strength parameters is much less compared to the corresponding strength parameters in concrete [11]; thus, in this study, f_y for the uncertainty in the steel is only considered. The geometry uncertainties, such as the aspect ratio of the shear wall (H/L), ratio of length to thickness (L/t), vertical preloading

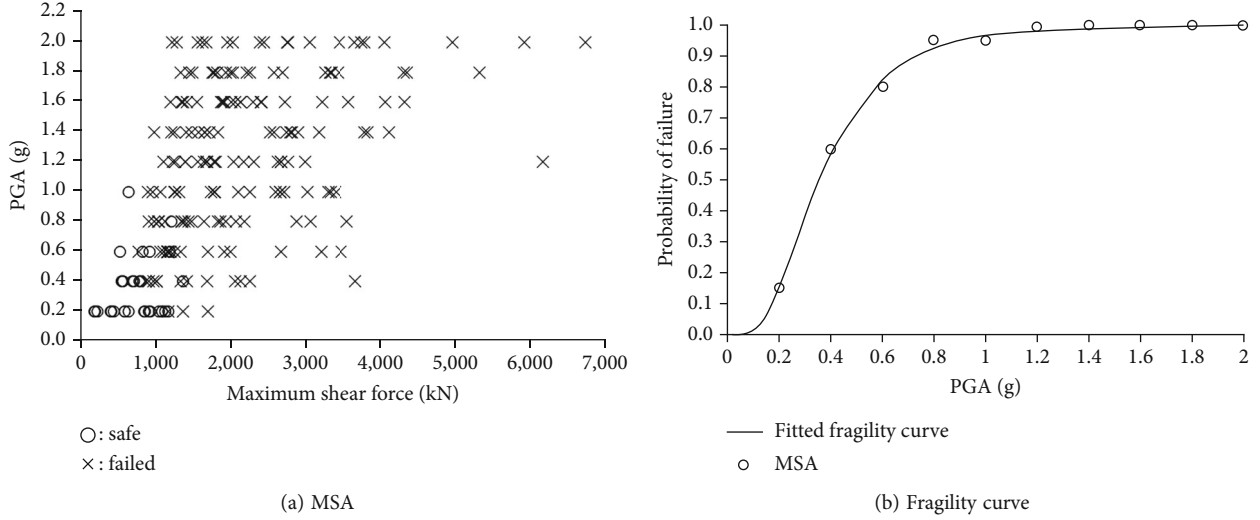FIGURE 12: Fragility curve for the box-type shear wall using MSA.

(P), and horizontal and vertical rebar ratio ρ_h and ρ_v, are considered, and the correlations between the parameters are not considered because of the inherent variability in structural geometry across different structural types and purposes. The geometric uncertainties were determined by research from the Multidisciplinary Center for Earthquake Engineering Research [36]. In the research, a database was assembled comprising data from experiments with 434 squat walls to improve the current state of knowledge on squat wall response and develop improved empirical equations for ultimate shear strength for shear walls [36]. The database for the 150 rectangular wall experiments is used in this study to consider uncertainties in the geometry of the rectangular shear walls. All input variables, including material and dimensional properties, would be randomly extracted based on the defined distribution in Table 3. Furthermore, 20 pairs of artificial ground motions that were generated in a previous study [37] are selected to account for the uncertainty in earthquakes. The artificial ground motions consist of two horizontal and a vertical earthquake time history, and they are compatible with a design response spectrum anchored to 0.3 g for NPPs based on Regulatory Guide 1.60 [38], as shown in Figure 5. The entire ground motions are scaled to 10 different peak ground acceleration (PGA): 0.2 g, 0.4 g, 0.6 g, 0.8 g, 1.0 g, 1.2 g, 1.4 g, 1.6 g, 1.8 g, and 2.0 g for generating surrogate models and estimating fragility curves. 200 sets of random variables for the material and geometry of the boxed-type shear walls are generated by the Latin hypercube sampling technique, and the sets are randomly matched with the selected ground motions. As a result, 200 NTHAs are performed to obtain the ultimate shear forces of the shear walls for given input variables.

4. Surrogate Models for Seismic Demand Model

Probabilistic risk assessments provide realistic and subjective risk index of structures that describes losses as a function of hazard intensity and the probability of occurrence of each consequence. In the probabilistic risk assessment, the fragil-

ity curve is used to evaluate seismic performance on structures and estimate the probability of a specific failure given earthquake intensity. In addition, the fragility curve could estimate potential damage to structures during an earthquake. The conventional fragility analysis employs NTHA to collect the seismic demands. However, recent fragility analysis studies have shown the need for a multiparameter demand model using surrogate models because the existing method requires repetitive NTHAs to reestimate a fragility curve when the input parameters are changed or updated. On the other hand, the pretrained surrogate models help reduce a lot of computational time caused by NTHAs for reestimating the fragility curves and identifying the impact of each parameter on the fragility curve. Thus, this study uses common machine learning algorithms, such as the response surface method, support vector machine, Gaussian process regression, and neural network, to develop surrogate models for the seismic demand of box-type shear walls.

4.1. Response Surface Method (RSM). RSM consists of a group of mathematical and statistical techniques used to develop an adequate functional relationship between a response of interest and several input variables [39]. The relationship could be approximated with a linear polynomial function and second-order polynomial function, as follows [39]:

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \sum_{j=1, i>j}^{n} \beta_i x_i x_j, \tag{3}$$

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \sum_{j=1, i>j}^{n} \beta_i x_i x_j + \sum_{i=1}^{n} \beta_i x_i^2, \tag{4}$$

where y, x, and β represent the predicted responses, input variables, and regression coefficients. The coefficients are estimated using the least square method, which minimizes the gap between the response surface (polynomial function) and seismic demands obtained from the NTHAs.
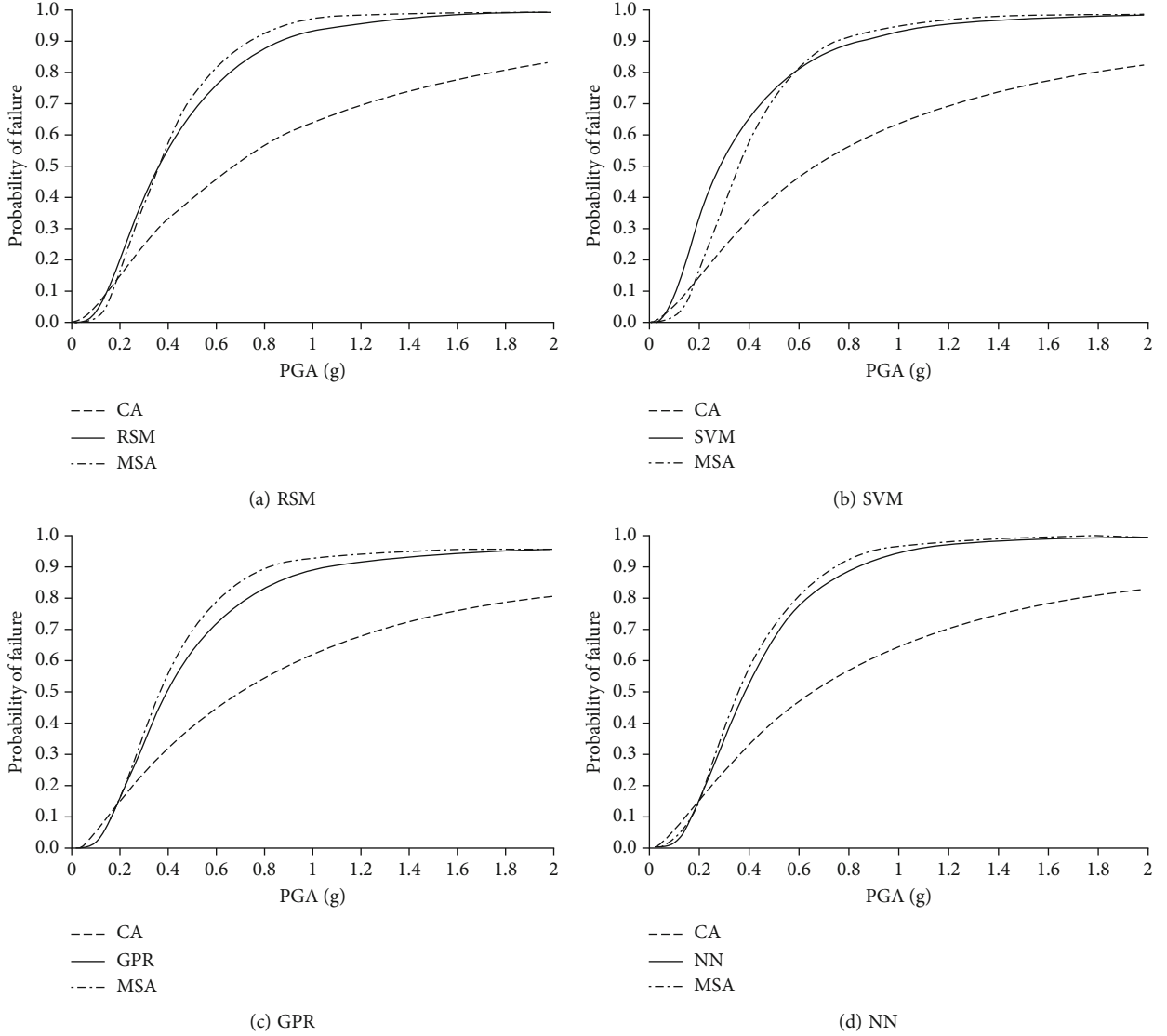
(a) RSM

(b) SVM

(c) GPR

(d) NN

FIGURE 13: Comparison between fragility curves using surrogate models and traditional approaches.

4.2. Support Vector Machine (SVM). SVM is a learning method to define a hyperplane for data classification and regression [18, 40–42]. In the regression case, the goal of the SVM is to define the hyperplane close to as many of the data points as possible [40]. The objective function of the SVM is to choose a hyperplane with a small norm while simultaneously minimizing the sum of the distances from the data points outside of ε-tube to the hyperplane, and the linear function is written as follows [40]:

$$
\begin{aligned}
\text{minimize}: \quad & \frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^{n} \left(\xi_i + \xi_i^* \right) \\
\text{subject to} \quad & \mathbf{w}^T \mathbf{x} + b - y \leq \varepsilon + \xi_i \\
& y - \mathbf{w}^T \mathbf{x} - b \leq \varepsilon + \xi_i^* \\
& \xi_i, \xi_i^* \geq 0,
\end{aligned}
\tag{5}
$$

where $\mathbf{w}$, $\mathbf{x}$, and b is the n-dimensional orthogonal vector to the hyperplane, n-dimensional input vector, and scalar. C is the positive penalty parameter, and slack variables (ξ_i and ξ_i^*) represent the deviation of data points from the ε-tube. In the SVM, the error of the data points within the ε-tube is ignored. Figure 6 shows a schematic of the support vector regression model.

When it is difficult to linearly define a hyperplane in low-dimensional space, a kernel function is used, and it implicitly transforms the input variables into high-dimensional space so that a linear hyperplane is provided in the high-dimensional space. A function for multiple input variables is Gaussian basis function with standard deviation (σ) as follows [40]:

$$
k\left(x_i, x_j \right) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma} \right).
\tag{6}
$$

4.3. Gaussian Process Regression (GPR).

GPR is a nonparametric Bayesian regression combined with the properties of Gaussian processes [43–45]. Its goal is to obtain the distribution of predicted responses at given input variables, in contrast to a representative regression, which would fit actual responses. In GPR, it is assumed that a function $f(x)$ is distributed as a Gaussian process, which is a probability distribution over functions that fit the data as follows [43]:

$$f(x) \sim \mathrm{GP}\Big(m(x), k\big(x, x'\big)\Big), \tag{7}$$

where $m(x)$ is the mean and $k(x, x')$ represents the covariance functions between each pair in x (input variables). In this study, the squared exponential kernel function is used as written in the following [43]:

$$k\big(x, x'\big) = 2\sigma_\mathrm{f}^2 \exp\left(-\frac{\|x - x'\|^2}{2\sigma_\mathrm{m}^2}\right), \tag{8}$$

where σ_f and σ_m are hyperparameters. The kernel function represents the similarity of data, which means that similar input variables produce similar outputs. The multivariate Gaussian distribution of f and f' is expressed as follows [43]:

$$\begin{bmatrix} f \\ f' \end{bmatrix} \sim N\left(\begin{bmatrix} m \\ m' \end{bmatrix} \begin{bmatrix} K(X, X) & K\big(X, X'\big) \\ K\big(X', X\big) & K\big(X', X'\big) \end{bmatrix}\right), \tag{9}$$

where f and f' are output functions obtained from observed and expected data. The conditional distribution over f' given f represents a posterior distribution given new data and is written as follows [43]:

$$\begin{aligned} f'|f \sim N\Big(&m' + K\big(X', X\big) K(X, X)^{-1}(f - m), K\big(X', X'\big) \\ &- K\big(X', X\big) K(X, X)^{-1} K\big(X, X'\big)\Big). \end{aligned} \tag{10}$$

4.4. Neural Networks (NN).

NN is one of the widely used machine learning techniques inspired by the structure of the human brain [41, 44, 46, 47]. The NN consists of an input layer, an output layer, and different hidden layers between the input and output layers, as shown in Figure 7. Each layer is connected and has neurons (nodes) that are assigned numerical values. Nodes in the input layer collect input data and transmit the data multiplied by the corresponding weight value to the hidden layer. The sum of the weight data from the input layer is saved at each node in the hidden layer, and an activation function that exists in the hidden layers determines the value of the data to be transmitted to a subsequent hidden layer or output layer. The function plays a significant role in defining nonlinear relationships between each node. Finally, the output layer produces the seismic demand for given input variables. Commonly, the backpropagation process is applied for training the model to minimize the error between actual and predicted
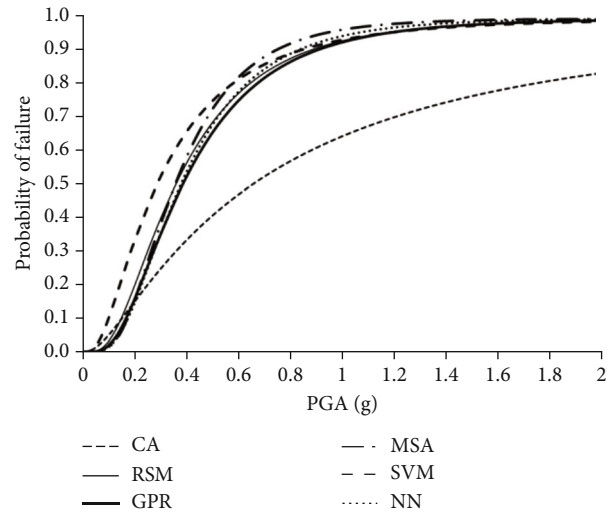

Figure 14: Fragility curves using surrogate models and traditional approaches.

outputs by modifying the weight and bias values. In this process, the model is retrained based on the error in the output. If the error does not satisfy a predefined accuracy threshold, it returns back to the input layer to adjust the weight and bias, reestimates the output, and repeats the process until the desired accuracy is achieved.

5. Fragility Analysis of Box-Type RC Squat Shear Wall

5.1. Evaluation for the Accuracy of Surrogate Models.

Surrogate models using various techniques, such as RSM, SVM, GPR, and NN, are generated for predicting the ultimate shear forces of box-type shear walls with different input data. The accuracy of the generated surrogate models is evaluated, and a best-fitting model is identified. To avoid overfitting in the surrogate models, the input data of the shear walls are randomly divided into a training set and a test set. The ratio of the testing to the training data is adopted as 0.333 in the current study. The training set is used to generate surrogate models, and the test set is used to evaluate the performance of the surrogate models. The accuracy of the surrogate models is evaluated with the following performance indices: coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and variance accounted for (VAF). The predicted and actual ultimate shear forces are compared in Figures 8 and 9. Table 4 presents the performance indices for surrogate seismic demand models. The surrogate models show the highest prediction of 88% and the lowest prediction of 74% in terms of R^2. Results from the GPR- and NN-based model show overfitting, whereas results from the RSM- and SVM-based model present slight underfitting. For selecting the best fitting model, average values of training and test set at each performance index are calculated as presented in Table 4, and as a result, it is found that the RSM-based surrogate model produces the best performance.

TABLE 5: Estimated parameters of lognormal fragility curves.

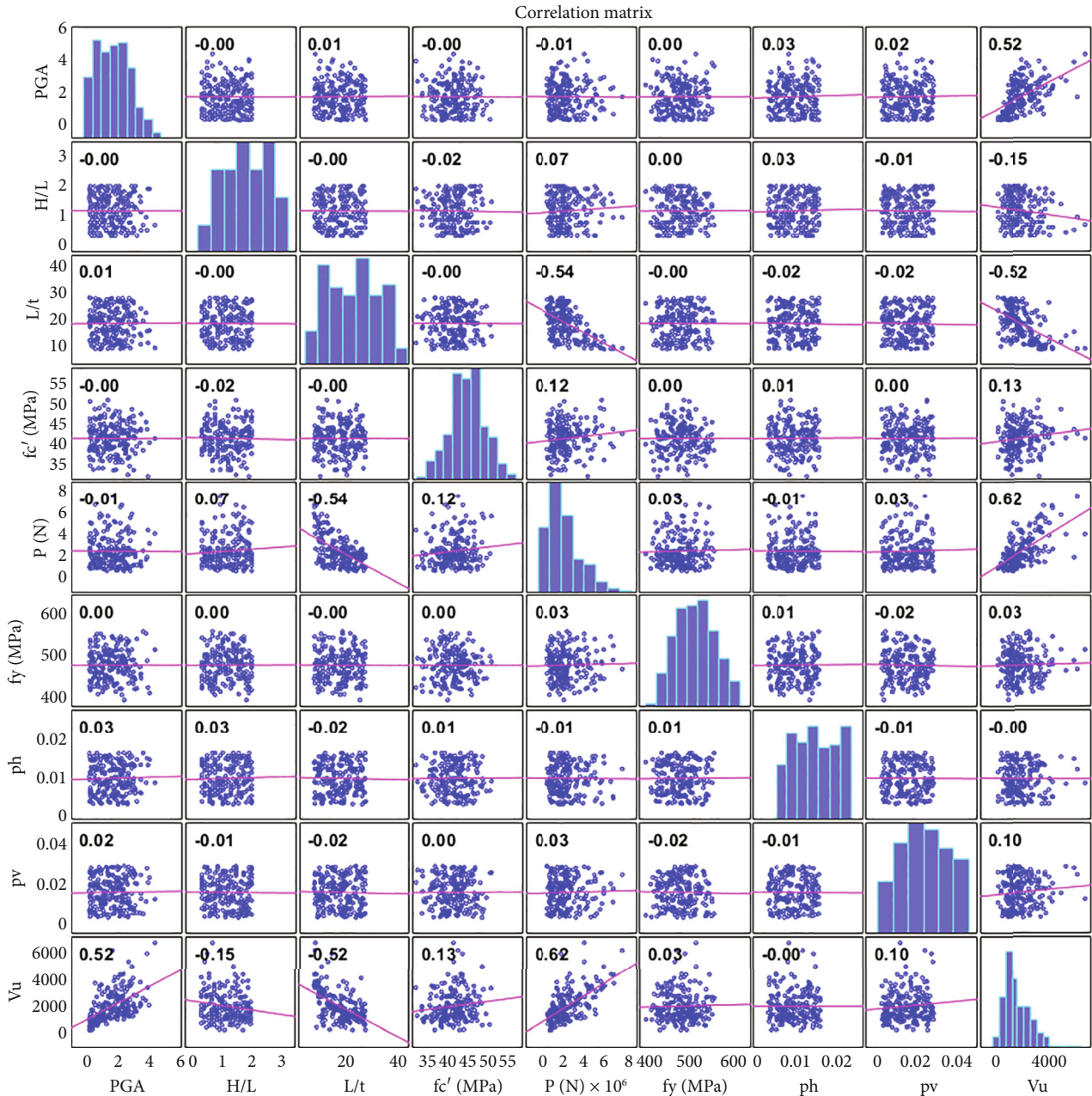|	CA	MSA	RSM	SVM	GPR	NN
Median (θ (g))	0.657	0.359	0.361	0.288	0.388	0.381
Dispersion (β)	1.146	0.566	0.688	0.826	0.653	0.605

FIGURE 15: Correlation matrix of the variables.

5.2. Fragility Curve Using Traditional Approach

5.2.1. Cloud Analysis (CA). A seismic fragility curve represents the conditional probability of a failure that a seismic demand exceeds a specific capacity for a given IM, and it is generally estimated in conjunction with the NTHA of a finite element model. IM may include spectral acceleration or peak ground acceleration (PGA). In this study, PGA is used as IM, considering consistency with previous studies [11, 21, 22] for fragility analysis of shear walls. A lognormal cumulative distribution for a specific IM often defines the fragility function. Cornell et al. [48] suggested the conditional probability of a failure as follows:
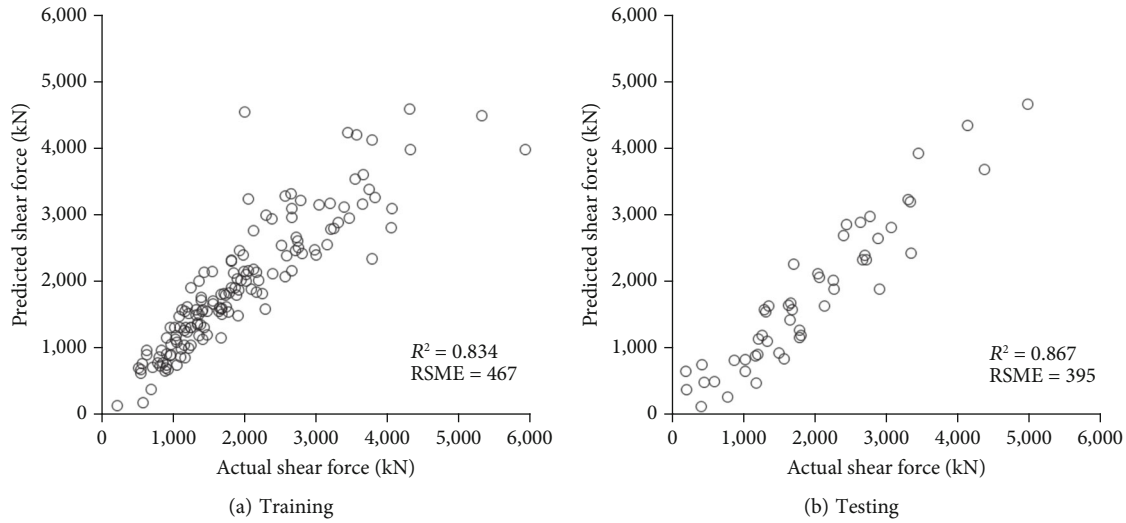
FIGURE 16: Comparison between predicted and actual ultimate shear force of box-type shear walls with five-parameter model.

$$P_f[\text{demand} > \text{capacity}|\text{IM}] = \Phi\left(\frac{\ln(S_d/S_c)}{\sqrt{\beta_{d|\text{IM}}^2 + \beta_c^2}}\right), \qquad (11)$$

where $\Phi()$ is the standard normal cumulative distribution function, S_d and S_c are the median of the demand and capacity, $\beta_{d|\text{IM}}$ is the dispersion of demand given in the IM, and β_c is the dispersion of the capacity. The S_d has a linear regression in the logarithmic space, and it can be expressed as follows:

$$\ln(S_d) = \ln(a) + b\ln(\text{IM}), \qquad (12)$$

where a and b are regression coefficients obtained from NTHA. The $\beta_{d|\text{IM}}$ is calculated as follows:

$$\beta_{d|\text{IM}} = \sqrt{\frac{\sum_{i=1}^{N}(\ln d_i - \ln(a\text{IM}^b))}{N-2}}, \qquad (13)$$

where N is the number of simulations and d_i is the seismic demand obtained from ith NTHA at a given IM. Figure 10 shows the seismic demands and IM plot obtained from the NTHAs of the FE models of the shear walls. It includes the assumed linear regression for the median of seismic demand, and the linear regression model shows a low prediction of 40% ($R^2 = 0.404$). The capacities of 200 shear walls (S_c) are calculated based on ACI 349 design code [49] to estimate the fragility curve of the shear walls. Finally, the fragility curve through the CA is estimated in conjunction with Equations (11)–(13), and Figure 11 presents the results.

5.2.2. Multiple Stripe Analysis (MSA). In the traditional MSA, the probability of a failure at a specific IM is estimated in conjunction with NTHA's results and is expressed as a ratio of the number of failures to the total number of simulations; therefore, the probability of a failure for the box-type shear walls is estimated at 10 different PGAs in this study.

However, the fragility curve cannot be estimated directly with these discrete data, and the maximum likelihood estimation (MLE) [50] is thereby used to estimate the fragility function. Figure 12 presents the fitted fragility curve by MLE compared with the MSA results at the different PGAs, and the conditional probability of a failure is written as follows:

$$P_f[\text{demand} > \text{capacity}|\text{IM}] = \Phi\left(\frac{\ln(x/\theta)}{\beta}\right), \qquad (14)$$

where θ is the median of IM and β is the standard deviation of $\ln(\text{IM})$.

5.3. Fragility Curve Using Surrogate Models. The fragility curves of the box-type shear walls using generated surrogate models are estimated in this section. To estimate the fragility curves, seismic demands at each PGA level obtained from the surrogate models are compared with the capacity of each shear wall model. The probability of a failure at the specific PGAs ($p_{f,\text{IM}} = N_{f,\text{PGA}}/N$) is calculated, and the fragility curve is estimated by MLE. Figures 13 and 14 present a comparison of the fragility curves between the traditional approaches and each surrogate model, and Table 5 summarizes the estimates of the fragility curves.

The fragility curve using CA represents the most overestimated fragility values, and the difference between the estimates of the fragility curves represents more than 50% compared with the MSA. It shows that the assumption of the CA that the median of the demands has a linear regression in the logarithmic space leads to unrealistic and overestimated results. Thus, the CA is not appropriate for the fragility analysis considering uncertainties in extensive input variables such as material, geometry, and earthquake. On the other hand, the RSM-based surrogate model produces the most similar fragility curve to the MSA. Also, the GPR and NN show errors within 10% regarding the probability of a
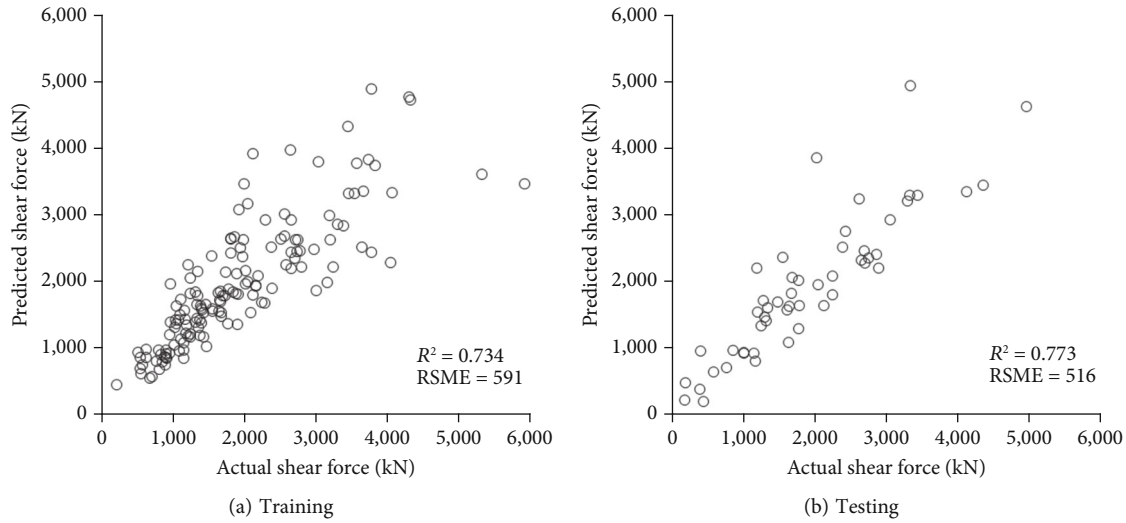
(a) Training

(b) Testing

FIGURE 17: Comparison between predicted and actual ultimate shear force of box-type shear walls with three-parameter model.

failure at each PGA. However, the SVM-based surrogate model results in greatly underestimated fragility curves, and the median value is less than 25% compared with the MSA. The dispersion of the fragility curves generated by surrogate models tends to be slightly larger than that generated by MSA; however, the overall trend of the fragility curve is quite similar to MSA's result, as shown in Figure 14. As a result, the RSM-based surrogate model is the best-fitted seismic demand model for box-type shear walls.

6. Reduced Multiparameter Surrogate Model Using RSM

While the RSM-based surrogate model is generated using the eight variables in the previous section, the model is simplified through a correlation analysis, which measures the correlation or dependence between variables, as presented in this section. Figure 15 shows the correlation between the variables. It is found that while the seismic demand (V_u) is rarely affected by f_y, ρ_h, and ρ_v, five other variables (PGA, f_c', H/L, L/t, and P) have quite an effect on it. Moreover, three of the five variables (PGA, L/t, and P) have the greatest influences on the seismic demand of the shear walls. To simplify the generated surrogate model, two cases are considered in this study: in case 1, the f_y, ρ_h, and ρ_v are eliminated, and only five variables are used as the input parameters. Case 2 is that only the three most important variables are used as the input. As a result, the two surrogate models are trained by RSM with different input variables (case 1 and case 2) and the accuracy of the surrogate models is presented in Figures 16 and 17. The RSM-based surrogate model with five parameters shows around 80% accuracy and has little difference from the original eight-parameter model; however, the three-parameter model is less accurate than the other model, at around 70%. In terms of the fragility curve, the results estimated by the reduced parameter surrogate models are shown in Figure 18 and compared with the
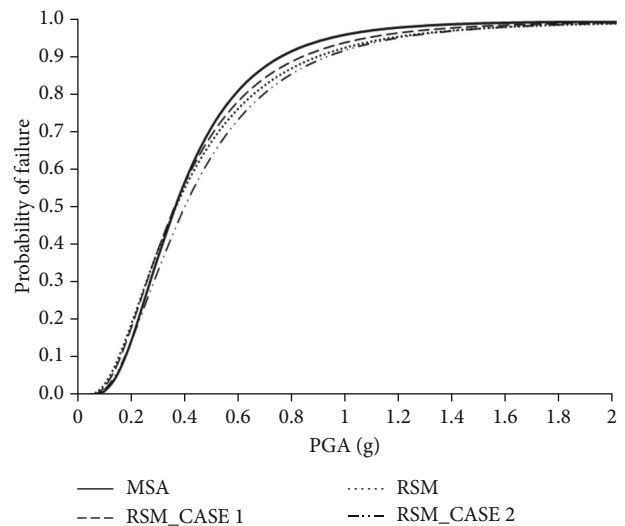

FIGURE 18: Comparison between fragility curves using RSM surrogate models and MSA.

original RSM-based model and MSA. The median of a fragility curve from the five-parameter model represents the closest value of MSA and RSM original, and the three-parameter model provides an overestimated median. Nevertheless, the three-parameter model still provides a better result than the CA. The estimated parameters of lognormal fragility curves are summarized in Table 6.

Although there is a slight difference in a fragility curve between full parameters and reduced parameters, the advantage of the RSM-based surrogate model with reduced parameters is that the model could be derived as a simple form due to the assumption of the polynomial regression. As a result, it is easily used with little understanding of surrogate models. Therefore, the seismic demand and fragility curves can be conveniently estimated based on the closed forms originating from Equations (3) and (4). The

TABLE 6: Estimated parameters of lognormal fragility curves for reduced parameter surrogate models.

	MSA	RSM original	RSM case 1	RSM case 2
Median (μ (g))	0.359	0.361	0.354	0.393
Dispersion (β)	0.566	0.688	0.651	0.656

five-parameter and three-parameter models are written as follows:

$$
\begin{aligned}
\text{Max shear force (kN)} = {} & -97.9(\text{PGA})^2 - 33.3(\text{PGA})\left(\frac{L}{t}\right) \\
& - 1.39 \times 10^{-5}(\text{PGA})(P) \\
& - 497(\text{PGA})\left(\frac{H}{L}\right) + 14.9(\text{PGA})\left(f_c'\right) \\
& + 1548(\text{PGA}) + 3.30\left(\frac{L}{t}\right)^2 - 9.7 \\
& \times 10^{-6}\left(\frac{L}{t}\right)(P) - 45.7\left(\frac{L}{t}\right)\left(\frac{H}{L}\right) \\
& + 5.14\left(\frac{L}{t}\right)\left(f_c'\right) - 261\left(\frac{L}{t}\right) \\
& + 2.07 \times 10^{-12}(P)^2 - 3.0 \\
& \times 10^{-4}(P)\left(\frac{H}{L}\right) + 1.49 \\
& \times 10^{-5}(P)\left(f_c'\right) + 2.14 \\
& \times 10^{4}(P) + 8.17\left(\frac{H}{L}\right)^2 \\
& + 44.2\left(\frac{H}{L}\right)\left(f_c'\right) + 15.5\left(\frac{H}{L}\right) \\
& - 2.23\left(f_c'\right)^2 + 9.33\left(f_c'\right) + 1823,
\end{aligned}
\tag{15}
$$

$$
\begin{aligned}
\text{Max shear force (kN)} = {} & -35.3(\text{PGA})^2 - 29.1(\text{PGA})\left(\frac{L}{t}\right) \\
& - 2.05 \times 10^{-5}(\text{PGA})(P) \\
& + 1356(\text{PGA}) + 4.18\left(\frac{L}{t}\right)^2 \\
& + 3.13 \times 10^{-6}\left(\frac{L}{t}\right)(P) - 155\left(\frac{L}{t}\right) \\
& + 2.44 \times 10^{-12}(P)^2 + 3.44 \\
& \times 10^{-4}(P) + 1122.
\end{aligned}
\tag{16}
$$

7. Conclusion

While existing fragility methods to reestimate the fragility curves of structures require repetitive NTHAs, surrogate model-based fragility methods are expected to reduce a lot of computational time and conveniently reestimate the fragility curve. This study generates multiparameter surrogate models of a box-type RC squat shear wall for a fragility analysis. The summary and conclusion of this study are the following:

(1) By comparing four performance indices, the RSM-based surrogate model is selected as the best-fitting demand model of the box-type shear walls and has 87% of R^2

(2) The RSM-based surrogate model produces the most similar fragility curve to the MSA. In addition, the GPR- and NN-based models lead to errors within 10% in terms of the probability of a failure at each PGA. On the other hand, the SVM-based surrogate model results in quite underestimated fragility curves, and the median value represents less than 25% compared with the MSA

(3) Based on a correlation analysis, simplified surrogate models are expressed as closed forms of polynomials. The simplified surrogate model with five parameters yields a fragility curve similar to the results of the MSA and RSM-based surrogate models. However, the simplified surrogate model with three parameters produces an overestimated fragility curve

(4) The suggested surrogate model and the reduced parameter closed forms help to reestimate and update the seismic shear force and fragility of the box-type RC squat shear walls without the time-consuming NTHAs when the change in the input variables of the shear walls occurs (i.e., degradation of material properties and discrepancy in the variables between design and construction stages)

While the surrogate models with multiparameters for estimating the seismic demand and fragility are generated in this study, the characteristic of earthquakes is only defined by PGA. Further studies are needed to consider various characteristics of the earthquakes (i.e., spectral acceleration at a period of 1.0 s and at the fundamental frequency of structures and duration of strong earthquake). In addition, the accuracy of the surrogate models lies around 87%; therefore, efforts to improve the performance of the surrogate models are needed in further studies.

Data Availability

Data are available on request.

Additional Points

Highlights. (i) This study presents surrogate models for a shear wall seismic demand. (ii) The seismic demand data for the shear wall are collected based on an experimentally

validated finite element model. (iii) The surrogate models are developed using several machine learning methods. (iv) The reduced parameter surrogate model is further proposed. (v) The efficient and accurate closed form equation is finally devised.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] M. Kobayashi, "Experience of infrastructure damage caused by the Great East Japan Earthquake and countermeasures against future disasters," *IEEE Communications Magazine*, vol. 52, no. 3, pp. 23–29, 2014.

[2] K. K. Lee, *Final Report of the Korean Government Commission on Relations between the 2017 Pohang Earthquake and EGS Project*, The Geological Society of Korea, 2019.

[3] A. S. Nur and C. W. Lee, "Damage proxy map (DPM) of the 2016 Gyeongju and 2017 Pohang earthquakes using Sentinel-1 imagery," *Korean Journal of Remote Sensing*, vol. 37, no. 1, pp. 13–22, 2021.

[4] S. M. Barbachyn, R. D. Devine, A. P. Thrall, and Y. C. Kurama, "Behavior of nuclear RC shear walls designed for similar lateral strengths using normal-strength versus high-strength materials," *Journal of Structural Engineering*, vol. 146, no. 11, article 04020252, 2020.

[5] N. Ile and J. M. Reynouard, "Nonlinear analysis of reinforced concrete shear wall under earthquake loading," *Journal of Earthquake Engineering*, vol. 4, no. 2, pp. 183–213, 2000.

[6] P. A. Hidalgo, C. A. Ledezma, and R. M. Jordan, "Seismic behavior of squat reinforced concrete shear walls," *Earthquake Spectra*, vol. 18, no. 2, pp. 287–308, 2002.

[7] D. Palermo, F. J. Vecchio, and H. Solanki, "Behavior of three-dimensional reinforced concrete shear walls," *ACI Structural Journal*, vol. 99, no. 1, pp. 81–89, 2002.

[8] J. Xu, J. Nie, J. Braverman, and C. Hofmayer, *Assessment of analysis methods for seismic shear wall capacity using JNES/NUPEC multi-axial cyclic and shaking table test data*, NUREG/CR-6925, 2007.

[9] J. J. Wang, C. Liu, X. Nie, R. Ding, and Y. J. Zhu, "Biaxial constitutive models for simulation of low-aspect-ratio reinforced concrete shear walls," *Journal of Engineering Mechanics*, vol. 148, no. 2, article 04021157, 2022.

[10] F. Wei, H. Chen, and Y. Xie, "Experimental study on seismic behavior of reinforced concrete shear walls with low shear span ratio," *Journal of Building Engineering*, vol. 45, article 103602, 2022.

[11] S. Syed and A. Gupta, "Seismic fragility of RC shear walls in nuclear power plant part 2: influence of uncertainty in material parameters on fragility of concrete shear walls," *Nuclear Engineering and Design*, vol. 295, pp. 587–596, 2015.

[12] E. Brunesi, R. Nascimbene, F. Parisi, and N. Augenti, "Progressive collapse fragility of reinforced concrete framed structures through incremental dynamic analysis," *Engineering Structures*, vol. 104, pp. 65–79, 2015.

[13] T. T. Tran, A. T. Cao, T. H. X. Nguyen, and D. Kim, "Fragility assessment for electric cabinet in nuclear power plant using response surface methodology," *Nuclear Engineering and Technology*, vol. 51, no. 3, pp. 894–903, 2019.

[14] M. Montazeri, G. Ghodrati Amiri, and P. Namiranian, "Seismic fragility and cost-benefit analysis of a conventional bridge with retrofit implements," *Soil Dynamics and Earthquake Engineering*, vol. 141, article 106456, 2021.

[15] S. Kameshwar and J. E. Padgett, "Multi-hazard risk assessment of highway bridges subjected to earthquake and hurricane hazards," *Engineering Structures*, vol. 78, pp. 154–166, 2014.

[16] S. Mangalathu and J. S. Jeon, "Stripe-based fragility analysis of multispan concrete bridge classes using machine learning techniques," *Earthquake Engineering & Structural Dynamics*, vol. 48, no. 11, pp. 1238–1255, 2019.

[17] S. Mangalathu, J. S. Jeon, and R. DesRoches, "Critical uncertainty parameters influencing seismic performance of bridges using Lasso regression," *Earthquake Engineering & Structural Dynamics*, vol. 47, no. 3, pp. 784–801, 2018.

[18] Y. Yan, H. Huang, and L. Sun, "Multivariate structural seismic fragility analysis and comparative study based on moment estimation surrogate model and Gaussian copula function," *Engineering Structures*, vol. 262, article 114324, 2022.

[19] J. Ghosh, J. E. Padgett, and L. Dueñas-Osorio, "Surrogate modeling and failure surface visualization for efficient seismic vulnerability assessment of highway bridges," *Probabilistic Engineering Mechanics*, vol. 34, pp. 189–199, 2013.

[20] S. Lee, A. Gupta, and G. T. Proestos, "Performance-based characterization and quantification of uncertainty in damage plasticity model for seismic fragility assessment of concrete structures," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 9, no. 1, article 04022060, 2023.

[21] S. Lee, S. Kwag, and B. S. Ju, "On efficient seismic fragility assessment using sequential Bayesian inference and truncation scheme: a case study of shear wall structure," *Computers & Structures*, vol. 289, article 107150, 2023.

[22] S. Lee, S. Kwag, and B. S. Ju, "Improved ensemble machine learning framework for seismic fragility analysis of concrete shear wall system," *Computers and Concrete*, vol. 32, no. 3, pp. 313–326, 2023.

[23] J. Lubliner, J. Oliver, S. Oller, and E. Oñate, "A plastic-damage model for concrete," *International Journal of Solids and Structures*, vol. 25, no. 3, pp. 299–326, 1989.

[24] J. Lee and G. L. Fenves, "Plastic-damage model for cyclic loading of concrete structures," *Journal of Engineering Mechanics*, vol. 124, no. 8, pp. 892–900, 1998.

[25] S. Syed and A. Gupta, "Seismic fragility of RC shear walls in nuclear power plant part 1: characterization of uncertainty in concrete constitutive model," *Nuclear Engineering and Design*, vol. 295, pp. 576–586, 2015.

[26] H. L. Minh, S. Khatir, M. A. Wahab, and T. Cuong-Le, "A concrete damage plasticity model for predicting the effects of compressive high-strength concrete under static and dynamic loads," *Journal of Building Engineering*, vol. 44, article 103239, 2021.

[27] Y. Sümer and M. Aktaş, "Defining parameters for concrete damage plasticity model," *Challenge Journal of Structural Mechanics*, vol. 1, no. 3, pp. 149–155, 2015.

[28] B. Alfarah, F. López-Almansa, and S. Oller, "New methodology for calculating damage variables evolution in plastic damage model for RC structures," *Engineering Structures*, vol. 132, pp. 70–86, 2017.

[29] T. Puddicome, *Finite Element Analysis of Reinforced Concrete and Steel Fiber Reinforced Concrete Slabs in Punching Shear*, Doctoral dissertation, Memorial University of Newfoundland, 2018.

[30] M. A. Najafgholipour, S. M. Dehghan, A. Dooshabi, and A. Niroomandi, "Finite element analysis of reinforced concrete beam-column connections with governing joint shear failure mode," *Latin American Journal of Solids and Structures*, vol. 14, no. 7, pp. 1200–1225, 2017.

[31] B. S. Ju and H. Y. Son, "Reevaluation of failure criteria location and novel improvement of 1/4 PCCV high fidelity simulation model under material uncertainty quantifications," *Nuclear Engineering and Technology*, vol. 55, no. 9, pp. 3493–3505, 2023.

[32] B. S. Ju, H. Son, S. Kwag, and S. Lee, "Statistical approach of performance-based uncertainty quantification of prestressed concrete containment structures for internal pressure capacity," *Nuclear Engineering and Design*, vol. 414, p. 112654, 2023.

[33] K. Maekawa and H. Okamura, "The deformational behavior and constitutive equation of concrete using the elasto-plastic and fracture model," *Journal of the Faculty of Engineering University of Tokyo, Series B*, vol. 37, no. 2, pp. 253–328, 1983.

[34] J. Izumo, "Analytical models for RC panel elements subjected to in-plane forces," *Concrete Library of JSCE*, vol. 12, pp. 155–181, 1989.

[35] M. P. Collins and D. Mitchell, *Prestressed Concrete Structures*, Response Publications, 1997.

[36] C. K. Gulec, *Performance-Based Assessment and Design of Squat Reinforced Concrete Shear Walls*, State University of New York at Buffalo, 2009.

[37] S. Eem, I. K. Choi, S. L. Cha, and S. Kwag, "Seismic response correlation coefficient for the structures, systems and components of the Korean nuclear power plant for seismic probabilistic safety assessment," *Annals of Nuclear Energy*, vol. 150, article 107759, 2021.

[38] US Nuclear Regulatory Commission, "Design response spectra for seismic design of nuclear power plants," *Regulatory Guide*, vol. 1, no. 60, 2014.

[39] A. I. Khuri and S. Mukhopadhyay, "Response surface methodology," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 2, pp. 128–149, 2010.

[40] K. Luu, K. Ricanek, T. D. Bui, and C. Y. Suen, "Age estimation using active appearance models and support vector machine regression," in *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, Washington, DC, USA, September 2009.

[41] L. Li, W. Zheng, and Y. Wang, "Prediction of moment redistribution in statically indeterminate reinforced concrete structures using artificial neural network and support vector regression," *Applied Sciences*, vol. 9, no. 1, p. 28, 2019.

[42] U. Aich and S. Banerjee, "Modeling of EDM responses by support vector machine regression with parameters selected by particle swarm optimization," *Applied Mathematical Modelling*, vol. 38, no. 11-12, pp. 2800–2818, 2014.

[43] R. Gentile and C. Galasso, "Gaussian process regression for seismic fragility assessment of building portfolios," *Structural Safety*, vol. 87, article 101980, 2020.

[44] P. G. Asteris, A. D. Skentou, A. Bardhan, P. Samui, and K. Pilakoutas, "Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models," *Cement and Concrete Research*, vol. 145, article 106449, 2021.

[45] Y. Pang, X. Zhou, W. He, J. Zhong, and O. Hui, "Uniform design–based Gaussian process regression for data-driven rapid fragility assessment of bridges," *Journal of Structural Engineering*, vol. 147, no. 4, article 04021008, 2021.

[46] W. S. Sarle, "Neural networks and statistical models," in *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Dallas, 1994.

[47] J. I. Kim and D. K. Kim, "Application of neural networks for estimation of concrete strength," *KSCE Journal of Civil Engineering*, vol. 6, no. 4, pp. 429–438, 2002.

[48] C. A. Cornell, F. Jalayer, R. O. Hamburger, and D. A. Foutch, "Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines," *Journal of Structural Engineering*, vol. 128, no. 4, pp. 526–533, 2002.

[49] American Concrete Institute (ACI), *Code requirements for nuclear safety-related concrete structures and commentary*, ACI 349-13, Farmington Hills, 2013.

[50] J. W. Baker, "Efficient analytical fragility function fitting using dynamic structural analysis," *Earthquake Spectra*, vol. 31, no. 1, pp. 579–599, 2015.