Received: 29 October 2022 | Revised: 1 August 2023 | Accepted: 2 August 2023

DOI: 10.1002/eqe.3993

SPECIAL ISSUE ARTICLE

WILEY

Toward multivariate fragility functions for seismic damage and loss estimation of high-rise buildings

Pouria Kourehpaz[1] | **Carlos Molina Hutt**[1] | **David Lallemant**[2,3]

[1]Department of Civil Engineering, University of British Columbia, Vancouver, BC, Canada

[2]Asian School of the Environment, Nanyang Technological University, Singapore, Singapore

[3]Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore

Correspondence
Carlos Molina Hutt
Department of Civil Engineering, University of British Columbia
Vancouver, BC, Canada
Email: carlos.molinahutt@civil.ubc.ca

Funding information
Mitacs Globalink Research Award; Canada's Natural Sciences and Engineering Research Council - Discovery Grant, Grant/Award Number: RGPIN-2019-04599; National Research Foundation, Singapore, Grant/Award Number: NRF-NRFF2018-06; AI Singapore.

Abstract

Data-driven models for seismic damage and loss assessment of buildings have become more common in recent years due to the availability of large repositories of recorded and synthetic ground motions coupled with structural response simulation data. This paper explores the benefits of using bivariate and multivariate fragility functions to estimate earthquake-induced damage and economic loss in high-rise buildings. The dataset used in this study encompasses 15,000 simulations of modern high-rise reinforced concrete shear wall buildings ranging from eight to 24 stories which are subjected to ground motion records at five different intensity levels. The proposed functions are conditioned on average spectral accelerations and ground motion significant duration. The results indicate that bivariate fragility functions improve damage state prediction success (Brier score) by 16%, and multivariate fragility functions by 24% relative to conventional univariate functions (standard of practice). To develop multivariate functions, nominal and ordinal probit regression models are fit to the dataset. While both models yield satisfactory predictive performance, ordinal functions can lead to a 15% reduction in misclassified collapse instances, that is, the minority class. Univariate functions tend to overestimate seismic losses at lower intensity levels while underestimating them at higher intensities. These loss estimates are significantly improved when bivariate or multivariate building fragility functions are used. Given the increase in the use of physics-based ground motion simulations and/or multi-variate ground motion models, from which multiple intensity measures can be extracted, a shift toward a more complex representation of fragility functions, for example, multivariate curves, is necessary and inevitable. The proposed functions are used to evaluate the performance of a portfolio of modern high-rise reinforced concrete shear wall buildings at four sites across the Seattle, Washington metropolitan area under a potential magnitude-9 Cascadia subduction zone earthquake scenario. The results indicate that the proposed functions can be beneficial in enhancing damage state predictions and loss estimates at a regional scale.

KEYWORDS
collapse misclassification, high-rise building, multivariate fragility function, probit regression model, seismic risk

1 | INTRODUCTION

The development of the Performance-Based Earthquake Engineering (PBEE) framework has enabled the estimation of earthquake-induced damage and loss of buildings considering various sources of uncertainty.[1,2] Although the PBEE framework was originally devised for individual buildings' performance assessments by aggregating component-level fragility information, it can also be applicable to building portfolio risk assessments. Within this framework, building-level fragility functions are used to define the probability of experiencing a certain level of damage conditioned on a ground motion intensity measure (IM). The damage state (DS) information can then be translated into various risk metrics (e.g., direct economic loss) by employing consequence functions. To derive the fragility functions from analytical and/or empirical data, various statistical methods have been used. In this regard, researchers typically follow parametric statistical procedures by fitting a Lognormal Cumulative Distribution Function (CDF) and using the maximum likelihood method to estimate the function parameters.[3–5] Baker[6] proposed more efficient fitting methods to estimate collapse fragility functions when there is a limited number of analytical data. Lallemant et al.[7] conducted a comprehensive study on the advantages and disadvantages of various statistical methods, including parametric, semi-parametric, and non-parametric methods. Noh et al.[8] focused exclusively on non-parametric methods and introduced various Gaussian kernel smoothing techniques to estimate seismic fragility curves. The fragility models that were predominantly used in previous studies were univariate functions developed primarily as a function of ground motion Spectral Acceleration (SA) at a specific period.

Although the performance-based approach was originally developed for earthquake engineering, researchers have applied it to other hazards such as flood[9] and tsunami.[10] In contrast with commonly-used seismic fragility functions, the studies on flood and tsunami hazards employed multivariate fragility functions to characterize damage and enhance the accuracy of risk assessments. For instance, Nofal et al.[9] proposed a multivariate flood damage model that includes flood depth and duration to characterize damage. Their results demonstrate that using multivariate fragility functions can significantly improve the accuracy of the flood loss calculation. Attary et al.[10] developed multivariate fragility functions to account for the interaction between earthquake SA and tsunami flow depth. Multivariate fragility functions in earthquake engineering have thus far primarily focused on component-level fragilities (e.g., reinforced concrete [RC] shear walls in Yazdi et al.[11]) or structure-specific fragilities (e.g., an unreinforced masonry building in Gehl et al.[12]), whereas building taxonomy-level fragilities are limited to functions conditioned on a unique ground motion IM. Using a single indicator to quantify earthquake-induced damage, particularly at the building taxonomy level, may lead to an unrealistic estimate of consequences, as it may not adequately account for the various drivers of damage. This can be particularly important in regional-scale assessments, which are intended to inform decision-making to enhance community resilience.[13] This issue is exacerbated for high-rise buildings that have unique dynamic properties. Thus, employing the same methodology to perform a seismic risk assessment of buildings with different heights may lead to an underestimation or overestimation of earthquake-induced economic losses. High-rise buildings also play a critical role in community resilience due to extended recovery times resulting from their large size and high occupancy levels.[14,15]

To address these shortcomings, this study explores the use of bivariate and multivariate seismic fragility models for modern high-rise RC shear wall buildings. The fragility models developed in this study leverage approximately 15,000 nonlinear simulation results obtained from a dataset based on archetypes developed by Marafi et al.[16] by considering six distinct design variations per building height. Each model was subjected to a suite of approximately 100 ground motion records at five distinct hazard levels ranging from 100 to 4975-year return periods. To extend and facilitate the application of multivariate fragility functions at a regional scale, (1) the availability of a sufficient amount of ground motion records, and (2) multivariate Ground Motion Models (GMMs) or vector-valued ground motion IMs is crucial. Firstly, recent advancements in high-resolution physics-based simulations for rare earthquakes have enabled the engineering community to supplement small repositories of empirical ground motion records with large datasets of simulated ground motions.[17] Secondly, multivariate GMMs have been proposed, for instance, by accounting for the correlation between SA at multiple periods.[18] Moreover, vector-valued ground motion IMs have been developed by modeling the joint probability distribution of correlated IMs.[19,20] Among various data-driven statistical methods, this study employs parametric fragility models in

order to simply represent the relationship between input and output parameters by means of a mathematical equation. To this end, a probit Generalized Linear Model (GLM) is used to represent seismic damage fragility function, which can even be more accessible to the engineering community. Furthermore, this paper examines the use of the multivariate GLM with a cumulative link model, that is, ordinal regression, to account for the ordinality of DS, which can offer a more informative representation of building damage and the ordinal characteristics of discrete DS.[21]

This article is organized as follows. Section 2 defines the characteristics of the input dataset as well as the criteria adopted for DS classification. Section 3 provides an overview and limitations of existing regional seismic risk assessment tools. Section 4 discusses the details of the fragility functions and the corresponding ground motion IMs considered as input variables for training the models. In addition, this section introduces performance metrics used to evaluate the prediction success of the various fragility models considered. In Section 5, the results are presented for DS predictions and economic losses under various ground motion shaking intensities. Finally, Section 6 showcases the application of the proposed fragility models in the context of a building portfolio seismic risk assessment in the Seattle, Washington metropolitan area under a plausible magnitude-9 Cascadia megathrust earthquake scenario.

2 | INPUT DATA

The dataset used in this study comprises 30 modern residential high-rise RC shear wall buildings with five distinct heights, namely, eight, 12, 16, 20 and 24 stories, designed by Marafi et al.[16] in Seattle, Washington State. The archetype buildings have floor dimensions of 30.5 m × 30.5 m (100 ft × 100 ft) above grade and 48.8 m × 48.8 m (160 ft × 160 ft) below grade. The height of each story is uniform across all archetypes and throughout the building, measuring 3.05 m (10 ft). The archetype with eight stories includes three basement levels, while all other archetypes have four basement levels. The lateral force-resisting system for all archetypes includes a symmetrical central core composed of two C-shaped walls that are coupled in one direction and uncoupled in the other. Marafi et al.[16] evaluated the nonlinear structural performance of these archetype buildings under suites of hazard-consistent ground motions, with records representative of crustal, intraslab, and interface earthquakes consistent with the 2018 National Seismic Hazard Model (NSHM)[22] for a downtown Seattle site (47.60° N, −122.30° W). The structural analyses were conducted following a multiple stripe analysis (MSA) procedure[23] considering five intensity stripes representing 100, 475, 975, 2475, and 4975-year return periods. To enable integration with the hazard data, the MSA employed the median SA across all orientations, that is, SA_{RotD50}. The target response spectrum employed was a conditional mean spectrum (CMS), conditioned on the target SA at the fundamental period of each archetype. The CMS was calculated by taking a weighted average of the ground motion prediction models corresponding to the applicable earthquake source characteristics (e.g., crustal, interface, intraslab) and site conditions. The weighting coefficients were determined based on the contribution of each seismic source to the hazard, which was derived from seismic hazard disaggregation results. At each hazard level, 100 ground motion records were selected and linearly scaled to match the target response spectrum mean and variance.[24] The proportion of crustal, intraslab, and interface records varied at each hazard level based on the source's contribution to the total hazard. Readers may refer to Marafi et al.[25] for further details.

For each building height, in addition to a reference archetype (REF), five enhanced design variations were included. While the REFs were designed to meet the minimum seismic design requirements of ASCE 7–16,[26] the enhanced designs included increased design seismic loads, reduced drift limits, and a combination of both. Marafi et al.[16] evaluated the seismic performance of all of the archetypes using two-dimensional (2D) nonlinear models. To develop these models, displacement-based beam-column elements with fiber sections were employed to account for the axial and flexural nonlinear response of the RC walls.[27] Readers may refer to Marafi et al.[25,28] for a detailed discussion of the modeling strategy.

Figure 1A shows the range of fundamental periods of the high-rise buildings considered in this study (the input dataset can be found in Electronic Supplement I). The figure displays the relative frequency of buildings with a specific period range. The number of buildings in each bin is divided by the total number of buildings to obtain the relative frequency. As seen in the figure, the fundamental period ranges from approximately 1 to 5 s, with the highest concertation of archetypes lying between a period of 2 and 3 s. Figure 1B illustrates the breakdown of the buildings' observed DS under different ground motion records categorized per their hazard levels (approximately 15,000 data points). The observed DS (DS0 to DS3), except collapse (DS4), are classified as a function of the story drift ratio thresholds defined in Federal Emergency Management Agency (FEMA) P-58[29] (see Table 1). The maximum story drift ratio along the building height is obtained

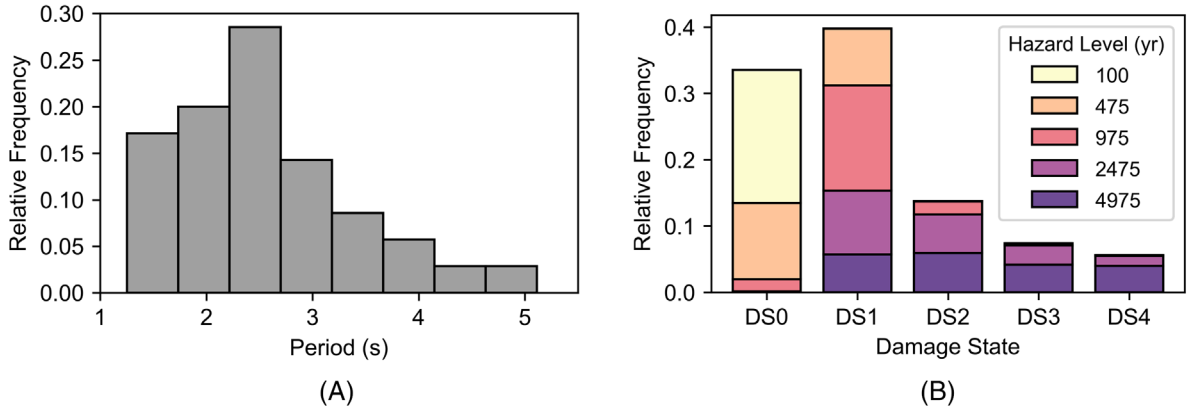


FIGURE 1 Distribution of building archetypes per (A) fundamental period of vibration, and (B) observed damage state (DS) per intensity level.

TABLE 1 Assumed damage state thresholds.

Story Drift Ratio (SDR)*	Damage State (DS)	Damage State Definition
SDR < 1%	DS0	Negligible (No Damage)
1% ≤ SDR < 2.6%	DS1	Slight
2.6% ≤ SDR < 3.6%	DS2	Moderate
3.6% ≤ SDR < 4.7%	DS3	Extensive
4.7% ≤ SDR	DS4	Complete (Collapse)

Note: *SDR threshold for all damage states, except collapse (DS4), is determined per FEMA P-58.[29] The collapse threshold is per Marafi et al.[16]

from nonlinear analysis simulations and this value is then compared against the specified thresholds to determine the anticipated DS. The threshold for collapse is determined based on story drift ratios that can lead to failure of the gravity system (i.e., slab-column punching shear failure). In modern RC shear wall buildings, failure of the lateral system is not generally expected, under the hazard levels considered, and gravity system failure is the governing collapse mode.[30] This failure mode can be associated with the slab-column rotation capacity which is dependent on the gravity shear ratio.[31,32] According to past experimental studies, for shear-reinforced slab-column connections with a gravity shear ratio ranging from 0.4 to 0.6, the median slab-column rotation capacity is 5.9%.[16,33–35] The relationship between the slab-column rotation and story drift ratios is linear supposing rigid-body rotations for walls and no shortening along the axial direction for the gravity system columns.[36] The archetype considered in this study had two symmetrical C-shaped central core walls, cantilevered in one direction, and uncoupled in the other, to resist lateral seismic forces. For these archetypes, the slab-column rotation value can be obtained when the story drift ratio is amplified by approximately a factor of 1.25. Therefore, structural collapse is assumed to occur if the story drift ratio exceeds 4.7%. At this threshold, considering drift amplification due to racking, the slab-column rotation demands reach values that testing has shown can lead to structural failure.

3 | POINT OF DEPARTURE

This study proposes more advanced tools/functions for use in existing regional seismic risk assessment frameworks that rely on existing univariate fragility functions for seismic damage characterization in building groups or taxonomies. A brief overview of these frameworks is provided below.

3.1 | Hazus

Hazus is a damage and loss estimation tool developed by the FEMA.[37] Hazus uses unique fragility functions for similar building types to obtain the probability of exceeding different damage states conditioned on a ground motion IM.

Buildings are grouped according to their occupancy type, seismic design level, structural system, material, and height. Hazus capacity curves for each building group that describe the lateral force (base shear) versus displacement response are developed by means of pushover analysis. Seismic performance is then evaluated using the capacity spectrum approach by overlaying capacity curves against response spectra. The Hazus fragility curves are then expressed by a lognormal CDF conditioned on spectral displacement. The median spectral displacements for different damage states are obtained using Hazus' unique DS drift thresholds as indicated in equ. 5-4 of the Hazus Technical Manual. The dispersion values are computed by accounting for the uncertainty associated with the capacity curve, the demand spectrum, and the median damage thresholds. For buildings that are components of utility and transportation systems, spectral displacements are converted to "equivalent" Peak Ground Accelerations (PGAs) for use in the fragility functions. While the fragility functions vary across structural systems and building heights, they are always characterized in terms of a single IM (e.g., spectral displacement, PGA). Hazus building groups are classified into different categories based on their height, where buildings with 1–3 stories are labeled as low-rise, 4–7 stories as mid-rise, and above 8 stories as high-rise. For instance, "C2H" represents high-rise RC shear wall buildings. To quantify earthquake-induced losses, vulnerability functions are derived using a probabilistic consequence model that generally assumes lognormal distribution for loss ratio per DS. The vulnerability functions describe the expected loss ratio conditioned on the ground motion IM as the fragility function.

3.2 | OpenQuake

OpenQuake, developed by the Global Earthquake Model (GEM) Foundation,[38] is a commonly used tool for regional seismic risk assessments. Similar to Hazus, this tool classifies buildings according to occupancy type, seismic design level, structural system, material, and height, where each class is referred to as a taxonomy.[39] OpenQuake supports empirical and analytical fragility functions for damage assessments. One advancement of OpenQuake relative to Hazus is the derivation of fragility curves using non-linear dynamic analysis methods. To achieve this by leveraging the capacity curves proposed by Hazus, equivalent Single-Degree-Of-Freedom (SDOF) oscillators are developed for each capacity curve. These SDOF models are then used in assessing seismic performance by means of non-linear time history analyses. The non-linear response of an SDOF system is used to determine the probability of exceeding a given DS under a ground motion intensity, which eventually leads to the derivation of building fragility functions.[40] Similar to Hazus, the OpenQuake fragility functions follow a lognormal CDF conditioned on a single IM for different taxonomies. For example, fragility functions developed by GEM for use in Canada's national seismic risk model, employ SA at 0.3, 0.6, and 1.0 s to characterize damage in low-, mid-, and high-rise buildings, respectively.[41] While it is important to acknowledge the differences between the behavior of buildings with different heights, characterizing earthquake-induced damage with a single indicator may be inadequate, particularly in high-rise buildings, for instance, due to fundamental periods of vibration well in excess of 1.0 s and contribution from higher mode effects.[42]

4 | METHODOLOGY

In this study, the PBEE framework is used to quantify earthquake-induced damage and losses in the context of regional seismic risk assessments with a focus on intensity-based and scenario-based assessments. In contrast with building-level PBEE, building-portfolio or regional-level PBEE merges the structural and damage analysis stages of the PBEE framework by means of fragility functions that describe the probability of damage conditioned on a ground motion IM. Equation (1) illustrates the annual exceedance rate of seismic loss (L) for a building portfolio.

$$\upsilon\ (L) = \sum_{i=1}^{N} \int_{D} \int_{IM} \mathrm{P}\left(L|D\right)\ \left|d\mathrm{P}\left(D|IM\right)\right|\ \left|d\lambda\left(IM\right)\right| \tag{1}$$

where $\lambda\ (IM)$ denotes the annual rate of exceedance of the ground motion IM, P($D|IM$) is the damage fragility function conditioned on a ground motion IM, P($L|D$) is the probability of loss given a certain level of damage (D), and N is the number of buildings in a portfolio. In the above equation, the IM is sufficient, that is, the seismic response and therefore, damage, are independent of earthquake magnitude and source-to-site distance.[43]

TABLE 2 Fragility functions definition and variables.

Model	Description	Intensity Measures	Regression Model
U_1	Conventional Univariate	$SA\ (T = 1\ \text{s})$	Nominal Probit
U^*_1	Structure-Specific Univariate	$SA\ (T = T_n\ \text{s})$	Nominal Probit
U_2	Enhanced Univariate	SA_{avg}	Nominal Probit
B_n	Bivariate	SA_{avg}, D_s	Nominal Probit
M_n	Multivariate	$SA^{SP}{}_{avg}, SA^{LP}{}_{avg}, D_s$	Nominal Probit
M_o	Multivariate	$SA^{SP}{}_{avg}, SA^{LP}{}_{avg}, D_s$	Ordinal Probit

Our work aims to enhance the damage analysis stage through the introduction of more advanced fragility functions. The risk metrics considered in this study are damage and loss conditioned on a given ground motion IM. For a given building, when evaluated under an earthquake scenario or pre-defined ground motion intensity level, the Expected Loss (EL), can be computed as shown in Equation (2).

$$\text{EL} = \sum_{i=1}^{n} \text{P}(L_i \mid D_i)\ \text{P}(D_i \mid IM) \tag{2}$$

where $\text{P}(D_i \mid IM)$ is the probability of experiencing DS i conditioned on a ground motion IM. $\text{P}(L_i \mid D_i)$ is the consequence function used to translate DS probabilities to losses. EL is then calculated by aggregating the loss and damage results across all DS, five in total ($n = 5$) for the buildings considered in this study. In contrast with the typical PBEE framework, which uses scalar IM quantities to describe the fragility models, in this study, we use IM vectors as explained in Section 4.1. We use a range of fragility models to compare their performance in relation to baseline models that employ scalar IM quantities.

4.1 | Fragility models

The proposed fragility functions advance the existing risk assessment frameworks by utilizing bivariate and multivariate functions (i.e., fragility surfaces) for seismic damage assessments of buildings. Table 2 summarizes the fragility models employed in this study. The conventional univariate function is denoted by U_1, which is commonly used in the existing seismic risk models for damage assessment of high-rise buildings for example, OpenQuake. U_1^* is the structure-specific univariate fragility function derived at the fundamental period (T_n) of a given structure. U_2 is an enhanced univariate function that considers a more robust IM, that is, SA_{avg}.[44] This IM corresponds to the geometric mean between $SA\ (T = 0\ \text{s})$ to $SA\ (T = 5\ \text{s})$ as depicted in Equation (3).

$$SA_{avg} = \left(\prod_{i=1}^{10} SA\ (T_i) \right)^{\frac{1}{10}} \tag{3}$$

where $T_1 = 0$ s, $T_2 = 0.1$ s, $T_3 = 0.2$ s, $T_4 = 0.3$ s, $T_5 = 0.75$ s, $T_6 = 1$ s, $T_7 = 2$ s, $T_8 = 3$ s, $T_9 = 4$ s, and $T_{10} = 5$ s. As seen in Equation (3), T_i values correspond to the fundamental period of vibration between 0 to 5 s, and they are consistent with the periods for which SA are provided in the NSHM, that is, the 2018 NSHM.

In addition to univariate fragility functions, bivariate and multivariate (three variables) probit regression models are chosen due to their simplicity and consistency with the existing form of building fragility functions. The probit models are basically GLMs with a standard normal CDF as the link function. These probit regression models employ a limited number of IMs (maximum three) that can be easily accessible to facilitate their integration into regional risk assessment frameworks. By contrast, other studies employ IMs for which GMMs are not widely available (e.g., Sousa et al.[45]) or utilize complex predictive models (e.g., Kiani et al.,[46] Kalakonas and Silva[47]) that would require a drastic change to the well-established regional risk assessment framework, thus hindering widespread adoption.

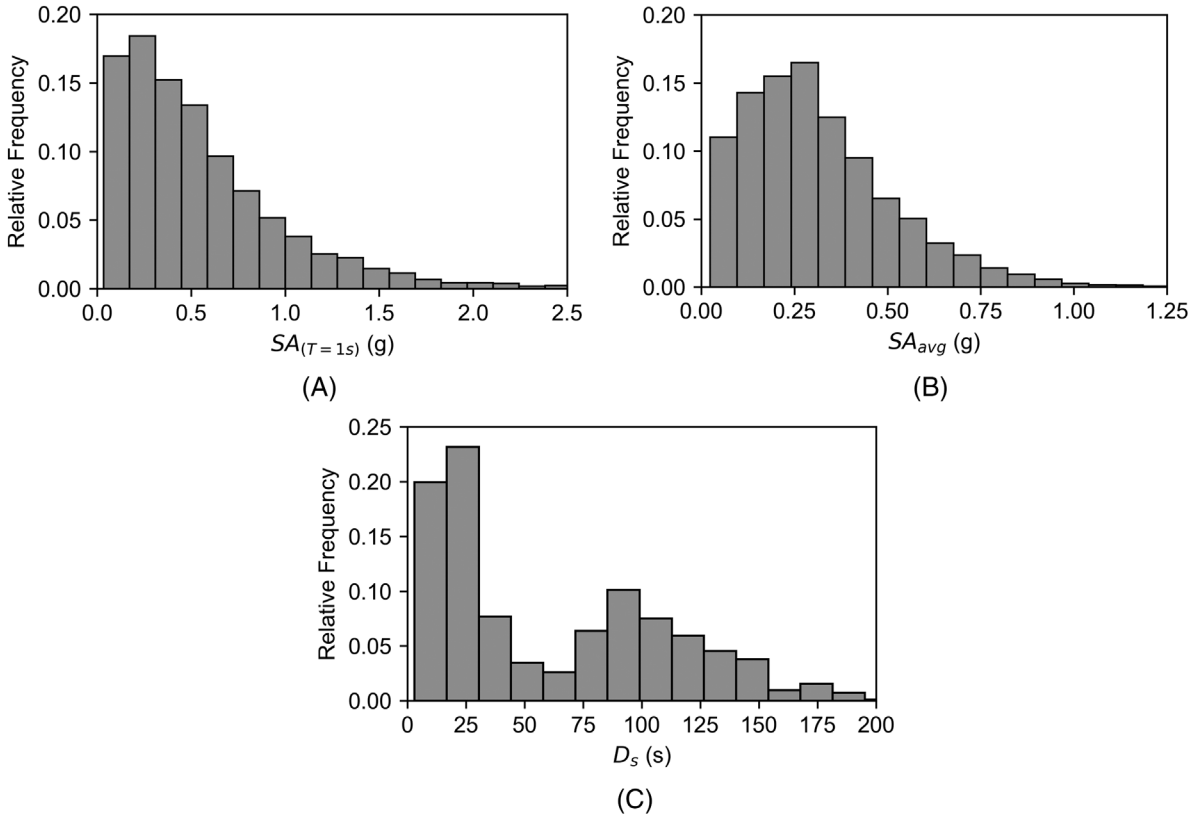


FIGURE 2 Distribution of input variables: (A) spectral acceleration at the fundamental period of 1 s ($SA_{(T=1s)}$), (B) average spectral acceleration (SA_{avg}), and (C) significant duration (D_s).

For the bivariate fragility function (B_n), significant duration (D_s) is considered, in addition to SA_{avg}, because long-duration ground motions can have a significant impact on the structural deformation capacity, damage level, and consequently collapse behavior.[48–54] Figure 2 illustrates the distribution of input variables considered in the univariate and bivariate models (U_1, U_2, and B_n). As seen in the figure, the dataset contains a large number of long-duration ground motion records, primarily to characterize shaking from subduction earthquakes. Note that in both univariate and bivariate models, the logarithm of SA, for example, SA ($T = 1$ s), and SA_{avg}, is considered as it is a common assumption in establishing the traditional form of seismic fragility curves.[55]

In the case of multivariate functions, the original SA_{avg} is split into two ranges: short-period SA_{avg} or $SA^{SP}{}_{avg}$ and long-period SA_{avg} or $SA^{LP}{}_{avg}$. $SA^{SP}{}_{avg}$ represents to average SA for periods ranging from 0 to 1 s and $SA^{LP}{}_{avg}$ for periods ranging from 2 to 5 s. The SA_{avg} split permits accounting for the spectral shape effects in the short and long period range distinctively, which is particularly important in more damaging ground motion records as they contain greater spectral ordinates at long periods.[56] Through this split, we can essentially obtain a spectral shape-based IM in the long-period range that can be crucial in characterizing the nonlinear behavior of the structures.[57] In the short-period range, we can potentially account for high-rise building damage due to higher dynamic mode effects as the higher mode periods are less than 1 s for the archetype buildings considered in this study.

Besides nominal probit regression models, a multivariate ordinal regression model is employed to investigate key differences in damage and loss estimates. Nominal and ordinal regression models are two of the most common Generalized Linear Models (GLMs). Nominal regression uses a probit link function and ordinal regression uses a cumulative probit link function to transform linear regression into a non-linear model.[58] These models are considered latent variable models that are formulated to represent observed variables in terms of variables that are not directly observed. The main purpose of generating latent, that is, hypothetical, variables is to provide a more precise explanation of phenomena, that is, outputs, in relation to observed input variables.[59] In this study, the latent variable is damage which is a function of ground motion IMs.

Using the nominal probit model, P_n, the probability of exceeding a DS $j \in \{1, 2, 3, 4\}$ can be computed as:

$$P_n\ (D \geq d_j) = \Phi\left(\beta_{0_j} + \sum_{i=1}^{m} x_i \beta_{i_j}\right) \tag{4}$$

where $\Phi()$ is the standard normal CDF or the probit link function, x_i is the IM or log-transformed IM (depending on the model), m is the number of IMs used in the model, β_0 and β_i are regression coefficients for a DS j obtained from the Maximum Likelihood Estimation (MLE). The values of the regression coefficients can be found in Electronic Supplement II. For example, using this formulation for the univariate fragility function U_2, where $x_1 = \log(SA_{\text{avg}})$, leads to a lognormal CDF fragility curve with the median of β_{0_k}/β_{1_k} and the dispersion of $1/\beta_{1_k}$ for DS k.

Using the ordinal probit model, P_o, the probability of exceeding a DS j can be computed as:

$$P_o\ (D \geq d_j) = \Phi\left(\xi_j + \sum_{i=1}^{m} x_i \beta_i\right) \tag{5}$$

Where ξ_j denotes the cut-off points or threshold between DS, which are obtained through the MLE procedure. The cut-off points are constrained to be arranged in ascending order, that is, the cut-off point for a higher DS must be greater than that for a lower DS. β_i are regression coefficients attained using the MLE method. The values of the cut-off points and regression coefficients can be found in Electronic Supplement II. In contrast with the nominal model, these coefficients do not vary across DS. The nominal and ordinal probit GLMs are readily accessible on standard statistical software packages. In this study, we have used a Python module, namely, statsmodels[60] to conduct regression analyses.

The principal difference between nominal and ordinal probit fragility functions relates to how they treat damage states in performing regression analyses. In model nominal models, the damage states are treated as distinct categories (mutually exclusive). In ordinal models, damage states are ordered categories and are fitted simultaneously.[21] To this end, in nominal models, the β values need to be estimated separately for each DS, whereas in ordinal models β values are estimated only once considering the entire dataset in the ordinal models, which leads to a shared β across damage states. In other words, the ordinal models tend to achieve the desired predictive power by using fewer predictor variables, that is, a parsimonious model.[58] Another advantageous characteristic of ordinal models is the capability to generate non-crossing fragility curves, which is often a problem in the nominal approach.[7]

4.2 | Prediction success evaluation

To evaluate the prediction results against observed values, that is, detailed nonlinear analysis simulation results, different metrics are employed. A brief overview of each metric is presented next.

4.2.1 | Brier score

Brier score (BS) is a widely used metric for evaluating the accuracy of probabilistic forecasts against realized specific outcomes.[61] It is commonly used for instance to test the accuracy of a weather prediction model that describes the probability of rain, against a history of rain/no-rain outcomes. Unlike most other evaluation metrics, it, therefore, accounts both for the prediction and the certainty of the prediction. The BS measures the mean squared error between predictions and observed outcomes. This metric can take values from 0 to 1, where a BS of 0 implies that the predicted probability matches the actual outcome, that is, optimal prediction accuracy. Therefore, a lower BS reflects better predictive performance. In this study, we report the average BS across damage states for the predictive models listed in Table 2.

4.2.2 | Top-k accuracy score

Top-k accuracy score is a generalized form of the accuracy score, which is used in multi-label classification problems. The accuracy score ($k = 1$) is computed as the ratio of correct predictions to the total observed data points. When $k = 2$, the top-k accuracy score, that is, top-2 accuracy, returns the fraction of predicted labels within one class difference

relative to the actual or observed labels. In this study, predictive models provide the probability of being in various damage states for each observed instance, that is, outputs are non-binary. Since these performance metrics are only applicable to classification models, the predicted label is taken as the most likely DS for each scenario when reporting the top-2 accuracy.

4.2.3 | Collapse misclassification

Collapse misclassification rate is another metric used to evaluate fragility models' performance as the seismic collapse risk can play a critical role in guiding policy decisions for enhanced resilience in urban regions.[62] This metric is calculated as the average of the predicted probability of no collapse across observed collapse cases. For instance, [0.0, 0.0, 0.1, 0.2, 0.7] is an output array obtained from a fragility model. The components of this array correspond to the probability of being in different damage states (starting from DS0 to DS4). If this case represents an actual collapse instance for a building, the probability of collapse misclassification will be 0.3.

5 | ASSESSMENT RESULTS

In this study, we use 80% of the data for training the fragility models and the remaining 20% as the test set to evaluate the models' predictive performance. When splitting the dataset, the proportions of damage states are preserved across the training and test sets, that is, a stratified split is adopted. To alleviate overfitting issues and bias in estimates, *k*-fold cross-validation[63] is carried out using different partitions of the dataset in five iterations, ($k = 5$).

Using the performance metrics defined earlier, we investigate the predictive power of various fragility models for seismic damage and loss estimations in the subsequent sections.

5.1 | Comparison of fragility models

A comparison of fragility models using different prediction success metrics is illustrated in Figure 3. On every subplot, in addition to the mean performance score value, the minimum and maximum values considering all five cross-validation folds are shown with the error bars. As seen in the figure, using enhanced univariate fragility functions (U_2) instead of conventional univariate functions (U_1) can improve the accuracy of DS predictions by reducing the Brier score by 10%. Using bivariate (B_n) and multivariate (M_n) functions can further reduce the Brier score by 16% and 24%, relative to U_1, respectively.

Figure 3 reveals that the highest prediction success, expressed in terms of Brier score, accuracy, and top-2 accuracy, is achieved when multivariate nominal fragility functions are used. While multivariate ordinal functions (M_o) are slightly less effective than the multivariate nominal functions (M_n) in damage predictions on the entire dataset, they can be beneficial in detecting the observed collapse cases, leading to a 15% reduction in the collapse misclassification rate relative to M_n (see Figure 3B). It should be noted that the large collapse misclassification rate within the fragility models is attributed to the comparison of predicted DS probabilities against the binary vectors (0 or 1 entries) of observed collapse. However, the comparative analysis of collapse misclassification rate across fragility models is still appropriate and insightful.

As outlined in the Methodology section, compared to ordinal regression models, the nominal models have more unknown parameters, which offer a higher degree of freedom. As a result, these models achieve higher accuracy, top-2 accuracy, and Brier score relative to ordinal models. On the other hand, ordinal models have higher parsimony and fit to all damage states concurrently. These models acknowledge the overall structure of the dataset and dependencies between DS. These features provide extra insights into the minority class, that is, collapse DS (refer back to Figure 1B), which consequently leads to a lower percentage of collapse misclassification as seen in Figure 3B.

Seismic vulnerability functions are developed by leveraging DS predictions obtained from the fragility functions and the consequence model proposed by Martins and Silva.[64] This consequence model assumes the mean loss ratio of 0.05, 0.2, 0.6, and 1.0 for a building experiencing the damage state DS1, DS2, DS3, and DS4, respectively. Figure 4 summarizes the expected loss per hazard level using different fragility models. A comparison of expected loss prediction results against the observed losses, reveals that the univariate functions (U_1 and U_2) tend to overestimate losses at low hazard levels (e.g., 475-year) and underestimate them at high hazard levels (e.g., 4975-year). At the 100-year hazard intensity, the observed losses are nearly zero. Although all fragility models tend to overestimate losses at this hazard level, the U_1 model shows

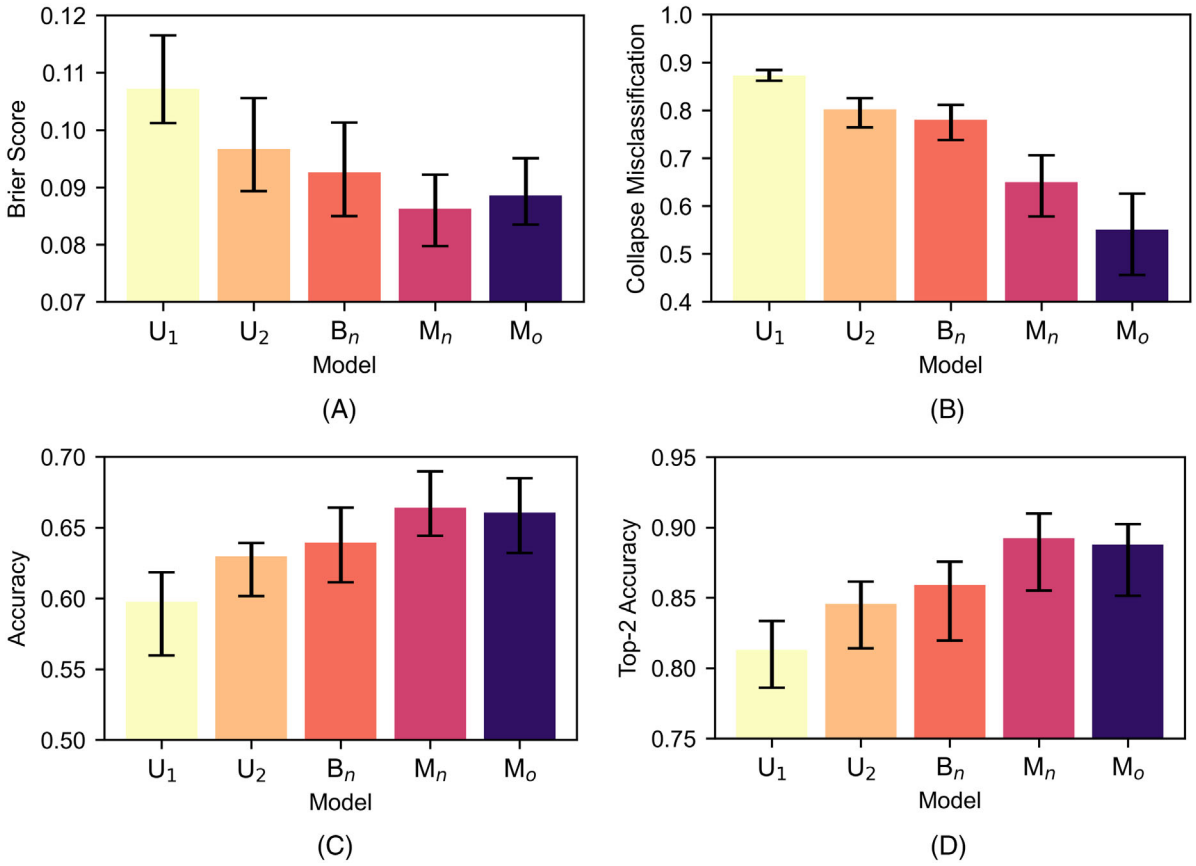
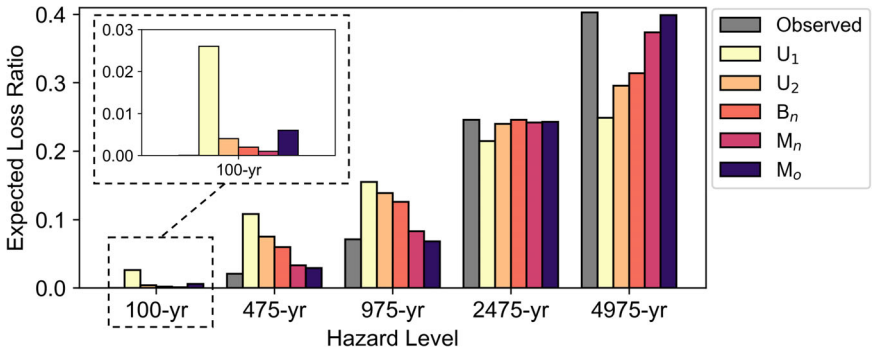


FIGURE 3 Prediction success in terms of (A) Brier score, (B) collapse misclassification, (C) accuracy, and (D) top-2 accuracy for damage fragility models on the test set.

FIGURE 4 Observed and predicted loss ratios using different fragility models for a range of hazard levels.



a more significant deviation, with predicted losses reaching approximately 3%. The main reason behind this significant overestimation of conventional univariate models in relation to multivariate models can be attributed to the DS probability distributions associated with the univariate functions which tend to significantly inflate the probability of experiencing the adjacent DS. For instance, for all observed DS0 damage instances, the univariate fragility functions predict on average around a 30% chance of experiencing DS1 or DS2. This issue is alleviated by using bivariate and multivariate fragility models, where this probability is reduced to 22% and 18%, respectively. A similar issue arises at high intensity levels, where conventional univariate models tend to underestimate losses and probabilities of collapse because, for observed collapse (i.e., DS4) instances, the univariate fragility functions predict considerable probabilities of experiencing DS3. Both nominal and ordinal models achieve significant improvement in the accuracy of loss estimates at all hazard levels. Under the highest ground motion shaking intensity, that is, 4975-year, the multivariate ordinal model provides the most accurate loss values (39.9% vs. 40.3% in observed losses), which can be attributed to a lower collapse misclassification (refer back to Figure 3B).

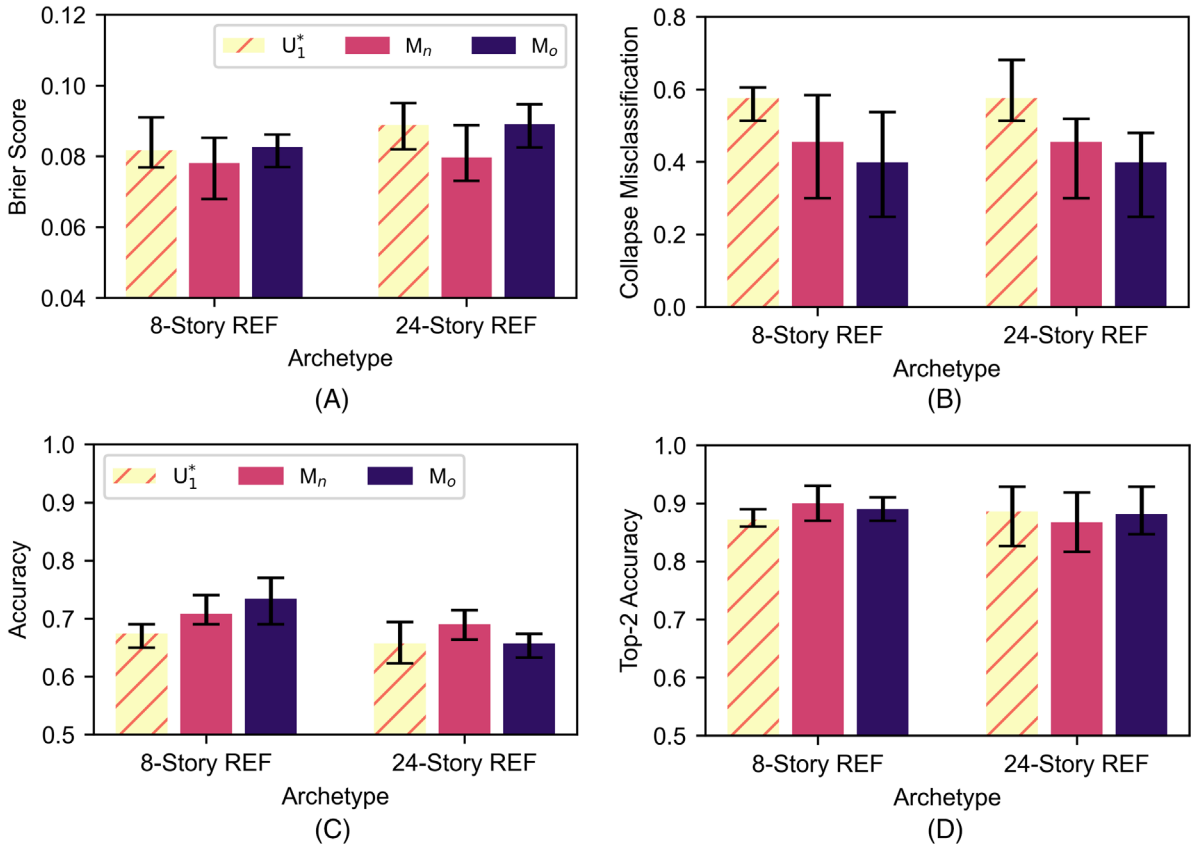


FIGURE 5 Prediction success for the structure-specific univariate versus multivariate fragility functions in terms of (A) Brier score, (B) collapse misclassification, (C) accuracy, and (D) top-2 accuracy.

5.2 | Structure-specific fragility functions

The primary purpose of this section is to explore the potential advantages of using the proposed multivariate fragility functions for an individual building, as opposed to lumping all the data together to represent a building taxonomy. To this end, we develop structure-specific fragility functions for a single height and design strategy to examine the effectiveness of multivariate fragilities against univariate fragility functions that utilize the SA, conditioned on the building's fundamental period, as the ground motion IM. To illustrate this comparison, fragility functions are exclusively developed for the shortest and tallest buildings, that is, 8-story and 24-story REF archetypes. The univariate fragility functions (U^*_1) are conditioned on SA at a period of 2 and 5 s, which correspond to the fundamental period of the 8-story and 24-story buildings, respectively. Figure 5 illustrates a comparison of prediction success for the multivariate nominal and ordinal fragilities with respect to the univariate fragilities. Figure 5 indicates that compared to the U^*_1 fragility model, utilizing multivariate fragility models leads to a slight improvement in the accuracy of predictions in terms of Brier score (a lower value denotes an improvement), accuracy, and top-2 accuracy for both buildings. Nevertheless, the primary advantage of using the proposed multivariate fragilities can be attributed to a significant reduction in the misclassification of collapse instances. As seen in Figure 5B, using the ordinal multivariate model, M_o, decreases the collapse misclassification rate on average by approximately 35% relative to the U^*_1 fragility model. These results suggest that multivariate fragility functions offer a more effective way of estimating seismic damage and loss in buildings in comparison to structure-specific univariate fragility functions.

5.3 | Impact of ground motion duration on seismic risk

This section explores the influence of significant duration on collapse risk and earthquake-induced losses. Figure 6 depicts that ground motion duration can have a considerable impact on collapse risk and seismic losses, which is not captured by

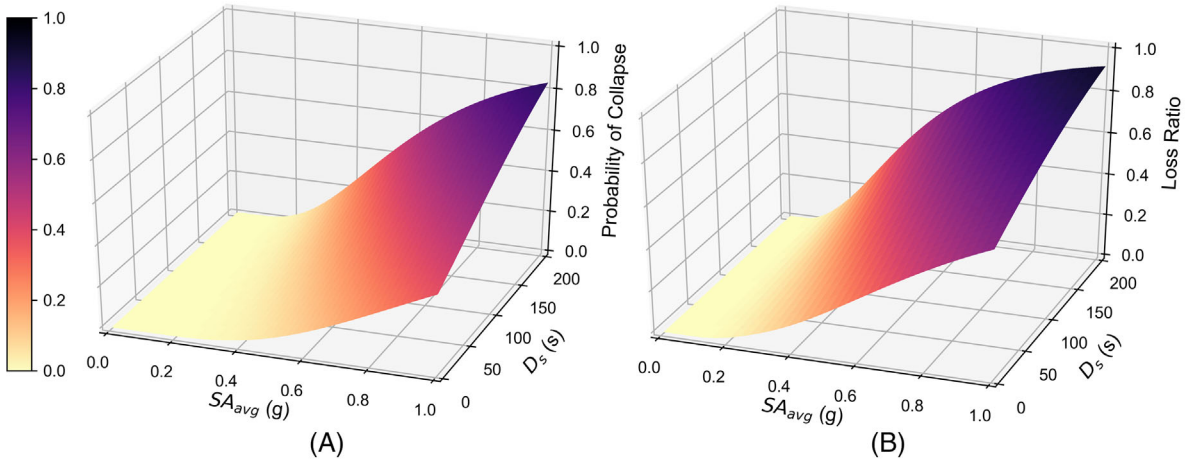


FIGURE 6 Bivariate model (B_n): (A) collapse fragility surface and (B) vulnerability surface.

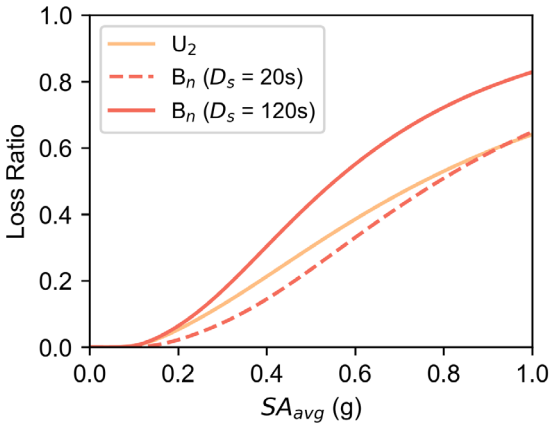


FIGURE 7 Vulnerability curves for univariate (U_2) and bivariate (B_n) models conditioned on different significant durations (D_s).

using univariate functions that are developed when buildings with different characteristics are lumped together under a specific building group or taxonomy. This impact is more evident in high earthquake intensities (i.e., SA_{avg} greater than 0.5 g). For instance, integrating the D_s into fragility models for a SA_{avg} of 1.0 g results in a 30% increase in collapse risk. A similar trend is observed for expected losses and the maximum value is achieved for the highest values of D_s and SA_{avg} over the variable grid. It should be noted that the D_s and SA_{avg} are weakly correlated IMs and thus any potential combination can be feasible.

The differences between bivariate (conditioned on a given significant duration) and univariate vulnerability functions are displayed in Figure 7. The bivariate vulnerability curves are generated for two earthquake scenarios with short and long durations, assuming a significant duration of 20 and 120 s, which represent two plausible scenarios (refer back to Figure 2C). As observed in the figure, under high earthquake intensities, for example, SA_{avg} greater than 0.8 g, with a D_s of 20 s, using bivariate and univariate models leads to similar loss ratios. However, when D_s is 120 s, the gap between bivariate and univariate vulnerability functions is substantial. For instance, assuming a SA_{avg} of 1.0 g, the expected losses are increased from approximately 0.6 to 0.8 in the event of a long-duration earthquake. This observation exemplifies the shortcoming of univariate fragility functions in recognizing differences between the impacts of short- and long-duration earthquake ground motions on the expected earthquake-included losses of buildings. It is important to highlight that, compared to ground motion duration, SA has a greater influence on expected earthquake-induced damage and losses. In this regard, for a specific earthquake scenario or intensity level, if ground motion records are selected based on seismic hazard disaggregation with careful consideration of earthquake magnitude and source-to-site distance, SA can to some extent contain certain ground-motion characteristics related to duration that may influence seismic risk.[65]

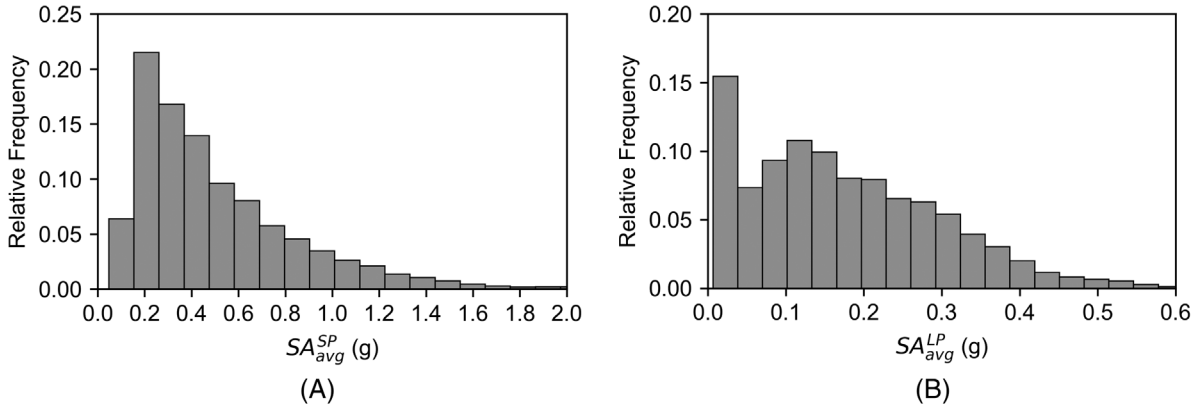


FIGURE 8 Distribution of (A) short-period average spectral acceleration (SA^{SP}_{avg}) and (B) long-period average spectral acceleration (SA^{LP}_{avg}) across the dataset.

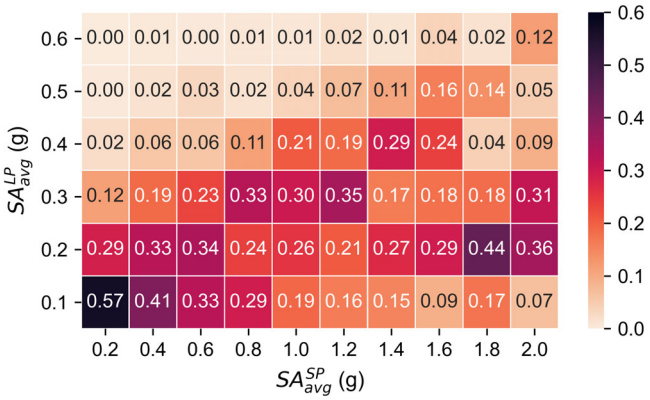


FIGURE 9 Conditional probability distribution of long-period average spectral acceleration (SA^{LP}_{avg}) given short-period average spectral acceleration (SA^{SP}_{avg}) from 0.2-2.0 g.

5.4 | Collapse risk assessment: Ordinal versus nominal probit regression models

As discussed earlier, multivariate ordinal functions can be a reliable choice for detecting structural collapse cases. In this section, we evaluate building collapse prediction performance using multivariate ordinal and nominal probit regression models. Figure 8 shows that the SA^{SP}_{avg} varies from 0.0 to 2.0 g, and SA^{LP}_{avg} from 0.0 to 0.6 g within the dataset used in this study. Figure 9 is a probability heatmap demonstrating conditional probabilities of SA^{LP}_{avg} given SA^{SP}_{avg} which is obtained using the kernel density estimation technique.[66] In this case, each of the vertical columns represents conditional probabilities for all possible values of SA^{LP}_{avg} for a fixed value of $SA^{SP}_{avg} = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ in units of g. As a result, the sum of the probabilities in each vertical column is one. For example, there is a 35% chance of observing SA^{LP}_{avg} of 0.3 g, if SA^{SP}_{avg} is equal to 1.2 g.

Figure 10 depicts nominal and ordinal collapse probabilities as a function of SA^{LP}_{avg} conditioned on two specified values of SA^{SP}_{avg} and D_s to represent: (A) a scenario with low ground motion intensity and short duration (dashed lines in Figure 10), (B) a scenario with high ground motion intensity and long duration (solid lines in Figure 10). In scenario A, there is a 92% chance that SA^{LP}_{avg} is less than 0.4 g (refer to Figure 9). In this case ($SA^{LP}_{avg} < 0.4$ g), nominal functions predict slightly higher collapse probabilities relative to ordinal functions. On the other hand, in scenario B, ordinal functions result in significantly larger collapse probabilities for all possible values of SA^{LP}_{avg}. For instance, there is a considerable chance of observing a SA^{LP}_{avg} of 0.4 g (~29%). In this case ($SA^{LP}_{avg} = 0.4$ g), the probability of collapse estimated by the ordinal fragility function is 70% as opposed to the 40% estimated by the nominal function. Considering a lower collapse misclassification rate of ordinal functions, the results suggest the nominal fragility functions tend to underestimate the probability of collapse, particularly under high-intensity and long-duration earthquakes.

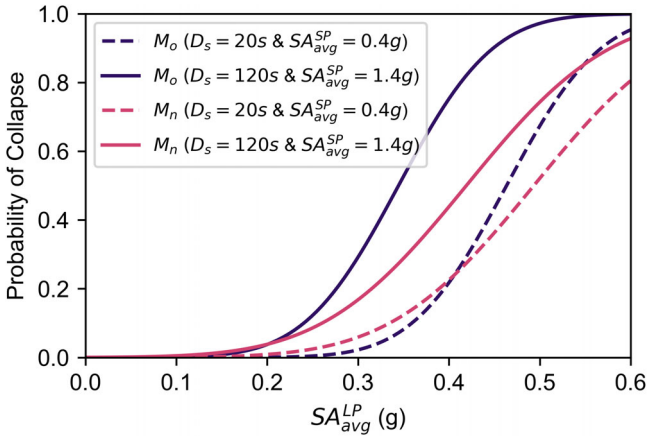


FIGURE 10 Multivariate nominal (M_n) and ordinal (M_o) collapse fragilities for different scenarios of ground motion intensity and significant duration.

6 | BUILDING PORTFOLIO RISK ASSESSMENT

In this section, the application of the proposed fragility models is presented for a building portfolio risk assessment. To this end, a portfolio of 30 modern high-rise RC shear wall buildings ranging from eight to 24 stories (assumed to be concentrated equally at all sites) is assessed. The building portfolio consists of all archetype buildings considered for training the fragility models as outlined in previous sections. This study uses four sites to conduct regional risk assessments, namely, Seattle, North Seattle, Bellevue, and Tacoma (see Figure 11A), consistent with the work of Kourehpaz and Molina Hutt.[67] The archetype buildings were subjected to a simulated M9 Cascadia Subduction Zone (CSZ) earthquake produced by Frankel et al.[68] The simulated M9 ground motions were obtained through 3D finite-difference simulations for periods greater than 1 s, while for periods less than 1 s, stochastic synthetics were used. Outside of deep sedimentary basins, simulated SAs at all periods were consistent with estimates derived from the BC Hydro GMM.[69] By contrast, at deep sedimentary basin sites, SA exceeded GMM predictions at periods greater than 1 s because these deep basin effects were explicitly captured in the finite-difference simulations.[70]

Here, the building portfolio is subjected to one of the most damaging scenarios of the M9 CSZ earthquake. This scenario is chosen such that the SA at a period of nearly 2.7 s, which is the dominant period within the dataset (refer back to Figure 1A), corresponds to the 95th percentile SA of the simulated M9 CSZ earthquake ground motions. Figure 11B depicts the 5%-damped geomean response spectra across the sites of interest for the M9 earthquake scenario considered. As seen in the figure, the seismic hazard can vary considerably across the region. This is attributed to the underlying deep sedimentary basins, which can amplify the SA at periods greater than 1 s.[28,71–73] The amplification of SA along with the long duration and more damaging spectral shape highlights the unique characteristics of the simulated M9 ground motions, which is the main motivation for their use in a regional risk assessment.

To quantify the impacts of the Cascadia megathrust earthquake scenario across the region, the probability of collapse and the expected loss ratios are estimated. The results are generated using the proposed probit regression models, which are trained by excluding simulated M9 CSZ earthquakes. The predictions are compared against observed outcomes derived from nonlinear analysis simulation results. Expected collapse probabilities and loss ratios across all buildings on each site are provided in Figure 12A, and 12B, respectively. The results suggest that the univariate models far underestimate collapse probabilities and seismic losses. For instance, U_1 predicts the overall expected loss ratio of 0.11 against the 0.34 observed loss ratio. Among the models analyzed in this study, multivariate fragility functions demonstrate the highest efficiency in estimating seismic damage and losses. For the case-study region, the multivariate fragility models (both nominal and ordinal) predict the overall probability of collapse to be approximately 14% versus the 12% observed collapse probability. Furthermore, using multivariate fragility models leads to the overall expected loss ratio of 0.34 across the region, which is equivalent to the observed loss ratio. The relationship between collapse risk and seismic loss predictions in relation to observed values varies across sites. For instance, North Seattle is the most critical site wherein the majority of buildings experience DS3 and DS4 damage states. At this site, while none of the observed DS3 cases are considered collapse

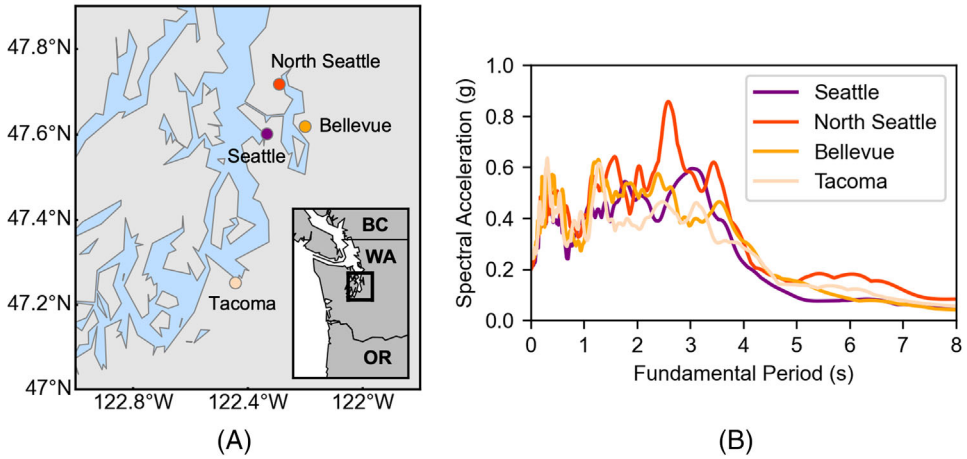


FIGURE 11 (A) Stations across the Seattle metropolitan area considered in the building portfolio seismic risk assessment, and corresponding (B) ground motion spectra for the simulated M9 Cascadia megathrust earthquake scenario.

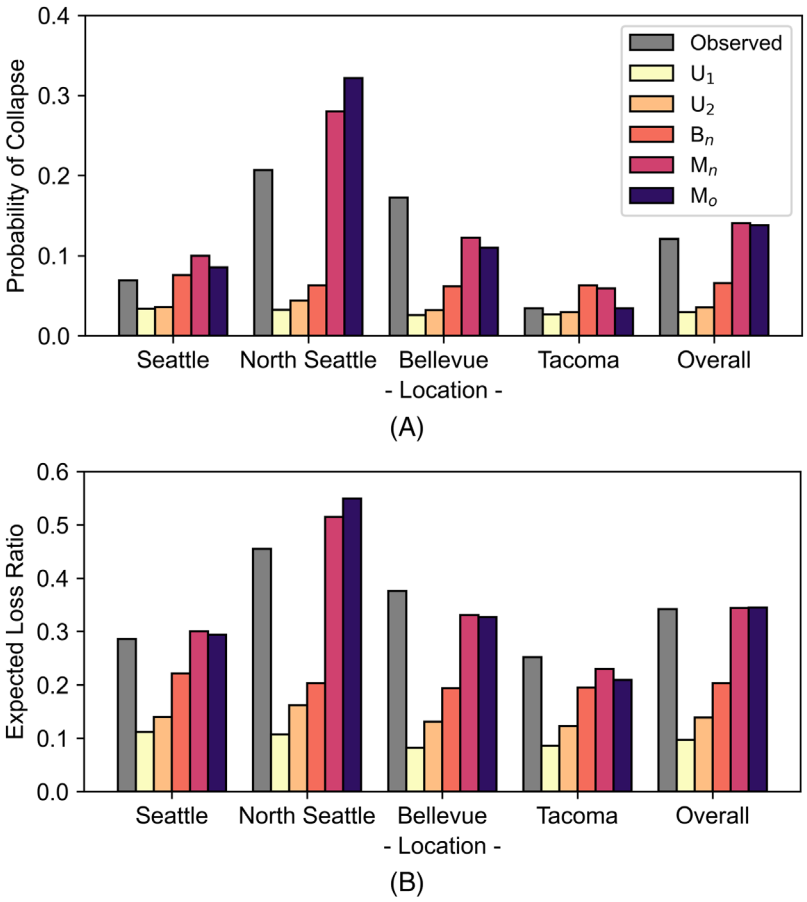


FIGURE 12 Observed and predicted (A) probability of collapse, and (B) expected loss ratios using different fragility models across the Seattle metropolitan area under an M9 CSZ earthquake scenario.

realizations, model predictions always have some small probabilities of collapse (i.e., DS4) even if the most likely DS coincides with the "observed" DS (i.e., DS3). Although nominal and ordinal functions yield decent performance in predicting the collapse probability and seismic losses, ordinal collapse fragilities can be particularly beneficial in reducing the collapse misclassification rate. For instance, at the most critical site, that is, North Seattle, using ordinal multivariate functions lowers collapse misclassification probability by 20% relative to nominal multivariate functions. A detailed calculation package for the building portfolio risk assessment can be found in Electronic Supplement III.

7 | DISCUSSION AND CONCLUSIONS

This study investigates the application of bivariate and multivariate building-level fragility and vulnerability functions for seismic damage and loss assessment. The study leverages a dataset of modern high-rise reinforced concrete shear wall buildings in Seattle, Washington subjected to earthquake ground motion records representing low to high hazard levels consistent with the 2018 US national seismic hazard model. The efficacy of fragility models is evaluated in terms of Brier score, accuracy score, top-2 accuracy score, and collapse misclassification rate by comparing the damage state predictions against observations derived from nonlinear response history analysis results. The earthquake-induced economic losses are also computed by converting the damage state results via an appropriate consequence function. Finally, the proposed fragility functions are implemented at a regional scale to quantify the risk of damage and losses to a portfolio of reinforced concrete shear wall buildings under a megathrust Cascadia subduction zone earthquake. The key findings can be summarized as follows:

- Using univariate building fragility functions (U_1 and U_2) leads to an overestimation of expected earthquake-induced economic losses at low-hazard levels, and an underestimation (up to 40%) at high-hazard levels relative to the expected observed losses obtained by leveraging nonlinear structural analysis results. The seismic loss prediction performance is significantly improved by employing multivariate fragility functions (even with only three variables) in comparison to univariate functions. The use of multivariate functions leads to predicted and observed values yielding almost the same results (within 1%) under high earthquake intensities.
- Among the models analyzed in this study, the multivariate nominal fragility function (M_n) is the most efficient model for predicting seismic damage state as it improves the Brier score by 24% relative to the conventional univariate fragility model (U_1). Although the multivariate ordinal and nominal models yield similar prediction success, using the multivariate ordinal model leads to a reduction of 15% in collapse misclassification rate relative to the nominal model.
- An application of the multivariate seismic fragility functions to a building-portfolio risk assessment (i.e., 30 buildings across four sites in the Seattle metropolitan area under a potential magnitude-9 earthquake scenario) results in an expected loss ratio of approximately 0.34 and a collapse probability of 14% across the portfolio of buildings relative to observed values of 0.34 and 12%, respectively. These results are a significant improvement over estimates obtained from traditional univariate functions, where the predicted loss ratio and collapse probability across the portfolio are approximately 0.11 and 4%, respectively.

In relation to the observed seismic damage and loss results, using univariate building fragility functions significantly biases (overestimates or underestimates) expected damage and loss predictions. This is attributed to the inability of the univariate model to capture (1) ground motion characteristics associated with short and long-duration shaking, and (2) damaging spectral shapes. The use of multivariate fragility functions that account for these phenomena shows improved seismic damage and loss prediction performance. Although multivariate models improve loss estimates at low hazard levels, that is, 100-year, in relation to univariate models, they still provide a small non-zero probability of observing damage under very low shaking intensities because these models provide a continuous probability distribution of damage.

The main advantage of ordinal models is their favorable performance in the reduction of collapse misclassification rate. While nominal models treat all data below collapse as the same, ordinal models distinguish near collapse from negligible damage realizations in constructing the collapse fragility curve. Thus, ordinal models make effective use of information in damage data and impose tighter constraints on fragility curves. This is particularly important for collapse since it typically represents the minority class.

This article highlights the potential benefits of using multivariate fragility functions based on the probit regression models for seismic damage and loss assessment of high-rise buildings. Given the increased prevalence of physics-based ground motion simulations (from which multiple intensity measures can be extracted) coupled with multi-variate ground motion models, it seems that developing and using multivariate fragility functions is often feasible and can significantly improve damage and loss predictions. Although the proposed functions can offer an efficient and simple tool for regional seismic risk assessments, this study does not attempt to perform a rigorous seismic risk assessment that involves integrating the fragility surfaces with the hazard maps that are developed by modeling the joint probability distribution of correlated intensity measures. Rather, it concentrates solely on quantifying the metrics of interest (i.e., seismic damage and loss) under a limited number of scenario- and intensity-based assessments. While this study has focused on high-rise modern reinforced concrete shear wall buildings, the method is generic and can be extended to buildings with different structural materials, systems, heights, and construction eras.

ACKNOWLEDGMENTS

The authors would like to thank Nasser Marafi (Risk Management Solutions) for sharing the structural nonlinear simulation results of archetype buildings used in this paper. The authors would also like to thank the Mitacs Globalink program and AI Singapore for their financial support to facilitate the research collaboration between the University of British Columbia and the Nanyang Technological University. This research was funded by Canada's Natural Sciences and Engineering Research Council under Discovery. RGPIN-2019-04599. Partial funding support was also provided by the National Research Foundation, Singapore under the NRF-NRFF2018-06 award.

DATA AVAILABILITY STATEMENT

The data compiled as a companion to this paper can be found by locating the paper in the publications list found on the following web page (https://www.carlosmolinahutt.com/publications) and accessing the supplied "Electronic Supplement" link. The electronic supplement includes the following:

(I): Input data spreadsheet.
(II): Fragility models: fitting procedure and parameter estimates.
(III): Building portfolio risk assessment.

ORCID

Pouria Kourehpaz https://orcid.org/0000-0003-1080-7089
Carlos Molina Hutt https://orcid.org/0000-0003-2116-1201
David Lallemant https://orcid.org/0000-0001-5759-9972

REFERENCES

1. Cornell C, Krawinkler H. *Progress and Challenges in Seismic Performance Assessment*. PEER Center News; 2000. https://apps.peer.berkeley.edu/news/2000spring/performance.html. Available at. accessed 24 Sept 2022.
2. Moehle J, Deierlein GG. A framework methodology for performance-based earthquake engineering. In: *Proceedings of the 13th World Conference on Earthquake Engineering (13WCEE)*. WCEE; Paper No. 679 2004.
3. Shinozuka M, Feng MQ, Lee J, Naganuma T. Statistical analysis of fragility curves. *J Eng Mech*. 2000;126(12):1224-1231.
4. Baker JW. *Vector-Valued Ground Motion Intensity Measures for Probabilistic Seismic Demand Analysis*. Ph.D. Thesis. Department of Civil and Environmental Engineering, Stanford University; 2005:347.
5. Straub D, Der Kiureghian A. Improved seismic fragility modeling from empirical data. *Struct Saf*. 2008;30(4):320-336.
6. Baker JW. Efficient analytical fragility function fitting using dynamic structural analysis. *Earthq Spectra*. 2015;31(1):579-599.
7. Lallemant D, Kiremidjian A, Burton H. Statistical procedures for developing earthquake damage fragility curves. *Earthq Eng Struct Dyn*. 2015;44(9):1373-1389.
8. Noh HY, Lallemant D, Kiremidjian AS. Development of empirical and analytical fragility functions using kernel smoothing methods. *Earthq Eng Struct Dyn*. 2015;44(8):1163-1180.
9. Nofal OM, van de Lindt JW, Do TQ. Multi-variate and single-variable flood fragility and loss approaches for buildings. *Reliab Eng Syst Saf*. 2020;202:106971.
10. Attary N, Van De Lindt JW, Barbosa AR, Cox DT, Unnikrishnan VU. Performance-based tsunami engineering for risk assessment of structures subjected to multi-hazards: tsunami following earthquake. *J Earthq Eng*. 2021;25(10):2065-2084.
11. Yazdi AJ, Haukaas T, Yang T, Gardoni P. Multivariate fragility models for earthquake engineering. *Earthq Spectra*. 2016;32(1):441-461.
12. Gehl P, Seyedi DM, Douglas J. Vector-valued fragility functions for seismic risk evaluation. *Bull Earthq Eng*. 2013;11(2):365-384.
13. Koliou M, Van de Lindt JW, McAllister TP, Ellingwood BR, Dillard M, Cutler H. State of the research in community resilience: progress and challenges. *Sustain Resilient Infrastruct*. 2020;5(3):131-151.
14. Molina Hutt C, Hulsey AM, Kakoty P, et al. Toward functional recovery performance in the seismic design of modern tall buildings. *Earthq Spectra*. 2022;38(1):283-309.
15. Molina Hutt C, Vahanvaty T, Kourehpaz P. An analytical framework to assess earthquake-induced downtime and model recovery of buildings. *Earthq Spectra*. 2022;38(2):1283-1320.
16. Marafi NA, Makdisi AJ, Berman JW, MO E. Design strategies to achieve target collapse risks for reinforced concrete wall buildings in sedimentary basins. *EarthqSpectra*. 2020;36(3):1038-1073.
17. Petrone F, Abrahamson N, McCallen D, Pitarka A, Rodgers A. Engineering evaluation of the EQSIM simulated ground-motion database: the San Francisco bay area region. *Earthq Eng Struct Dyn*. 2021;50(15):3939-3961.
18. Loth C, Baker JW. A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthq Eng Struct Dyn*. 2013;42(3):397-417.
19. Iervolino I, Giorgio M, Galasso C, Manfredi G. Conditional hazard maps for secondary intensity measures. *Bull Seismol Soc Am*. 2010;100(6):3312-3319.
20. Kohrangi M, Bazzurro P, Vamvatsikos D. Vector and scalar IMs in structural response estimation, part I: hazard analysis. *Earthq Spectra*. 2016;32(3):1507-1524.

21. Nguyen M, Lallemant D. Order matters: the benefits of ordinal fragility curves for damage and loss estimation. *Risk Anal*. 2022;42(5):1136-1148.
22. Petersen MD, Shumway AM, Powers PM, et al. The 2018 update of the US national seismic hazard model: overview of model and implications. *Earthq Spectra*. 2020;36(1):5-41.
23. Jalayer F, Cornell C. Alternative non-linear demand estimation methods for probability-based seismic assessments. *Earthq Eng Struct Dyn*. 2009;38(8):951-972.
24. Jayaram N, Lin T, Baker JW. A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance. *Earthq spectra*. 2011;27(3):797-815.
25. Marafi NA, Ahmed KA, Lehman DE, Lowes LN. Variability in seismic collapse probabilities of solid-and coupled-wall buildings. *J Struct Eng*. 2019;145(6):04019047.
26. *ASCE. Minimum Design Loads for Buildings and Other Structures*. ASCE/SEI 7–16; 2016.
27. Pugh JS, Lowes LN, Lehman DE. Nonlinear line-element modeling of flexural reinforced concrete walls. *Eng Struct*. 2015;104:174-192.
28. Marafi NA, Makdisi AJ, MO E, Berman JW. Impacts of an M9 Cascadia subduction zone earthquake and Seattle basin on performance of RC core wall buildings. *J Struct Eng*. 2020;146(2):04019201.
29. FEMA. Seismic Performance Assessment of Bui*ldings, FEMA P-58*. FEMA; 2012.
30. Kourehpaz P, Molina Hutt C, Marafi NA, Berman JW, Eberhard MO. Estimating economic losses of midrise reinforced concrete shear wall buildings in sedimentary basins by combining empirical and simulated seismic hazard characterizations. *Earthq Eng Struct Dyn*. 2021;50(1):26-42.
31. Hueste MBD, Browning J, Lepage A, Wallace JW. Seismic design criteria for slab-column connections. *ACI Structural J*. 2007;104(4):448.
32. Hueste MBD, Kang THK, Robertson IN. Lateral drift limits for structural concrete slab-column connections including shear reinforcement effects. In: *Proceedings of the ASCE Structures Congress*. ASCE; 2009.
33. Dilger W, Cao H. Behaviour of slab-column connections under reversed cyclic loading. In: *Proceedings of the 2nd International Conference of High-Rise Buildings*. Atlantis Press; 1991.
34. Dilger W, Brown S. Earthquake resistance of slab-column connections. *In Vol. 60 of Festschrift Professor Dr. Hugo BachmannZum*. Institut fur Baustatik unf Konstruktion; 1995:22-27.
35. Megally S, Ghali A. Punching shear design of earthquake-resistant slab-column connections. *Struct J*. 2000;97(5):720-730.
36. Charney FA. Wind drift serviceability limit state design of multistory buildings. *J Wind Eng Ind Aerodyn*. 1990;36:203-212.
37. FEMA. Hazus–MH 4.2*: Earthquake Model Technical Manual*. FEMA; 2020.
38. Silva V, Crowley H, Pagani M, Monelli D, Pinho R. Development of the OpenQuake engine, the global earthquake model's open-source software for seismic risk assessment. *Nat Hazards*. 2014;72(3):1409-1427.
39. Brzev S, Scawthorn C, Charleson AW, et al. *GEM Building Taxonomy (Version 2.0)*. GEM Foundation; 2013.
40. Silva V, Casotto C, Rao A, Villar M, Crowley H, Vamvatsikos D. OpenQuake risk modeller's toolkit—user guide. *Global Earthquake Model (GEM) Technical Report*. GEM; 2015.
41. Hobbs T, Journeay M, Anirudh R, Martins L, et al. *Scientific Basis of Canada's First Public National Seismic Risk Model*. Geological Survey of Canada; 2022.
42. Pejovic JR, Serdar NN, Pejovic RR. Optimal intensity measures for probabilistic seismic demand models of RC high-rise buildings. *Earthq Struct*. 2017;13(3):221-230.
43. Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthq Spectra*. 2007;23(2):357-392.
44. Eads L, Miranda E, Lignos DG. Average spectral acceleration as an intensity measure for collapse risk assessment. *Earthq Eng Struct Dyn*. 2015;44(12):2057-2073.
45. Sousa L, Silva A, Marques M, Crowley H, Pinho R. Including multiple IMTs in the development of fragility functions for earthquake loss estimation. In: *Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management*. ASCE; 2014.
46. Kiani J, Camp C, Pezeshk S. On the application of machine learning techniques to derive seismic fragility curves. *Comput Struct*. 2019;218:108-122.
47. Kalakonas P, Silva V. Seismic vulnerability modelling of building portfolios using artificial neural networks. *Earthq Eng Struct Dyn*. 2022;51(2):310-327.
48. Bhanu V, Chandramohan R, Sullivan TJ. Influence of ground motion duration on the dynamic deformation capacity of reinforced concrete frame structures. *Earthq Spectra*. 2021;37(4):2622-2637.
49. Bommer JJ, Magenes G, Hancock J, Penazzo P. The influence of strong-motion duration on the seismic response of masonry structures. *Bull Earthq Eng*. 2004;2:1-26.
50. Chandramohan R, Baker JW, Deierlein GG. Quantifying the influence of ground motion duration on structural collapse capacity using spectrally equivalent records. *Earthq Spectra*. 2016;32(2):927-950.
51. Fairhurst M, Bebamzadeh A, Ventura CE. Effect of ground motion duration on reinforced concrete shear wall buildings. *Earthq Spectra*. 2019;35(1):311-331.
52. Hancock J, Bommer JJ. A state-of-knowledge review of the influence of strong-motion duration on structural damage. *Earthq Spectra*. 2006;22(3):827-845.
53. Iervolino I, Manfredi G, Cosenza E. Ground motion duration effects on nonlinear seismic response. *Earthq Eng Struct Dyn*. 2006;35(1):21-38.
54. Raghunandan M, Liel AB. Effect of ground motion duration on earthquake-induced structural collapse. *Struct Saf*. 2013;41:119-133.

55. Porter K, Kennedy R, Bachman R. Creating fragility functions for performance-based earthquake engineering. *Earthq Spectra*. 2007;23(2):471-489.
56. Eads L, Miranda E, Lignos D. Spectral shape metrics and structural collapse potential. *Earthq Eng Struct Dyn*. 2016;45(10):1643-1659.
57. Bojórquez E, Iervolino I. Spectral shape proxies and nonlinear structural response. *Soil Dyn Earthq Eng*. 2011;31(7):996-1008.
58. Agresti A. *Categorical Data Analysis*. John Wiley & Sons; 2003.
59. Everitt B. *An Introduction to Latent Variable Models*. Chapman and Hall; 1984.
60. Seabold S, Statsmodels PJ, *Econometric and Statistical Modeling with Python*. In: *Proceedings of the 9th Python in science conference*. SciPy; 2010: 57-61.
61. Brier GW. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*. 1950;78(1):1-3.
62. Bijelic N, Lin T, Deierlein G. Classification algorithms for collapse prediction of tall buildings and regional risk estimation utilizing SCEC cybershake simulations. In: *Proceedings of the 13th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP)*. ICASP; 2019.
63. Cawley GC, Talbot NL. On over-fitting in model selection and subsequent selection bias in performance evaluation. *J Mach Learn Res*. 2010;11:2079-2107.
64. Martins L, Silva V. Development of a fragility and vulnerability model for global seismic risk analyses. *Bull Earthq Eng*. 2021;19(15):6719-6745.
65. Iervolino I, Manfredi G. A review of ground motion record selection strategies for dynamic structural analysis. In: *Modern Testing Techniques for Structural Systems*. Springer; 2008.
66. Scott DW. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons; 1992.
67. Kourehpaz P, Molina Hutt C. Machine learning for enhanced regional seismic risk assessments. *J Struct Eng*. 2022;148(9):04022126.
68. Frankel A, Wirth E, Marafi N, Vidale J, Stephenson W. Broadband synthetic seismograms for magnitude 9 earthquakes on the Cascadia megathrust based on 3D simulations and stochastic synthetics, part 1: methodology and overall results methodology and overall results. *Bull Seismol Soc Am*. 2018;108(5A):2347-2369.
69. Abrahamson NA, Gregor N, Addo K. BC hydro ground motion prediction equations for subduction earthquakes. *Earthq Spectra*. 2016;32:23-44.
70. Kakoty P, Molina Hutt C, Ghofrani H, Molnar S. Spectral acceleration basin amplification factors for interface Cascadia subduction zone earthquakes in Canada's 2020 national seismic hazard model. *Earthq Spectra*. 2023;39(2):1166-1188.
71. Kakoty P, Dyaga SM, Molina Hutt C. Impacts of simulated M9 Cascadia subduction zone earthquakes considering amplifications due to the Georgia sedimentary basin on reinforced concrete shear wall buildings. *Earthq Eng Struct Dyn*. 2021;50(1):237-256.
72. Eksir Monfared A, Molina Hutt C, Kakoty P, Kourehpaz P, Centeno J. Effects of the Georgia sedimentary basin on the response of modern tall RC shear-wall buildings to M9 Cascadia subduction zone earthquakes. *J Struct Eng*. 2021;147(8):05021003.
73. Molina Hutt C, Zahedimazandarani S, Marafi NA, Berman JW, Eberhard MO. Collapse risk of pre-Northridge tall steel moment-resisting frames in the Seattle basin during large-magnitude subduction earthquakes. *Eng Struct*. 2021;244:112751.

How to cite this article: Kourehpaz P, Molina Hutt C, Lallemant D. Toward multivariate fragility functions for seismic damage and loss estimation of high-rise buildings. *Earthquake Engng Struct Dyn*. 2023;52:4164–4182. https://doi.org/10.1002/eqe.3993