



A Rapid Estimation Method for Post-earthquake Building Losses

Dengke Zhao¹ · Zifa Wang^{1,2} · Jianming Wang² · Dongliang Wei² · Yang Zhou² · Zhaoyan Li¹

Accepted: 29 April 2023 / Published online: 24 May 2023
© The Author(s) 2023

Abstract

Rapid estimation of post-earthquake building damage and loss is very important in urgent response efforts. The current approach leaves much room for improvement in estimating ground motion and correctly incorporating the uncertainty and spatial correlation of the loss. This study proposed a new approach for rapidly estimating post-earthquake building loss with reasonable accuracy. The proposed method interpolates ground motion based on the observed ground motion using the Ground Motion Prediction Equation (GMPE) as the weight. It samples the building seismic loss quantile considering the spatial loss correlation that is expressed by Gaussian copula, and kriging is applied to reduce the dimension of direct sampling for estimation speed. The proposed approach was validated using three historical earthquake events in Japan with actual loss reports, and was then applied to predict the building loss amount for the March 2022 Fukushima Mw7.3 earthquake. The proposed method has high potential in future emergency efforts such as search, rescue, and evacuation planning.

Keywords Earthquake building loss estimation · Fukushima earthquake 2022 · Gaussian copula sampling · Japan · Spatial correlation of earthquake losses · Spatial interpolation of ground motion

1 Introduction

Large earthquakes that occur near exposure concentrations result in high human casualties and economic losses as in the case of the 2008 Wenchuan Earthquake in China (Wang 2008) and the 2011 Great East Japan Earthquake (Krausmann and Cruz 2013). The economic loss is mostly caused by damage of buildings. Human casualties and business interruptions are highly correlated with building damage while post-earthquake efforts are also significantly related to the damage state of buildings. Therefore, rapidly estimating building losses and their distribution with reasonable accuracy based on improved scientific approaches is very important in post-event search and rescue, evacuation planning, and other activities.

The current earthquake loss estimate methods can be roughly divided into four categories: (1) image-based remote sensing (Dell'Acqua and Gamba 2012; Zhang et al. 2021); (2) performance-based probabilistic (Zeng et al.

2016; Ghasemof et al. 2022); (3) machine learning-based (Kalakonas and Silva 2022; Stojadinović et al. 2022); and (4) structural vulnerability-based approaches (Pnevmatikos et al. 2020).

The remote sensing approach requires high-resolution images that are not readily available right after an earthquake, and the image cannot effectively exhibit the internal damage of buildings. The performance-based probabilistic approach can predict the structural response with high precision, but it is dependent on model and input parameters and requires huge computing resources. The machine learning-based method can predict the building damage with high accuracy using a large number of factors if there are enough data available for training, but due to the lack of training datasets the generalization capability of the machine learning method is limited to seismic regions and specific types of events. The structural vulnerability-based approach is the widely used method for post-earthquake loss estimates given its effectiveness and speed, especially for large-scale building blocks (Jaiswal et al. 2010). This approach requires the ground motion, the building inventory, and structural vulnerability, and building-level loss is calculated and then aggregated to yield the total loss for the earthquake (Hosseinpour et al. 2021).

✉ Zifa Wang
zifa@iem.ac.cn

¹ Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, China

² CEAKJ ADPRHexa Inc., Shaoguan 512000, China

In the structural vulnerability-based approach, the accuracy and reliability of the ground motion prediction is of vital importance. Many parameters have been used to represent the earthquake ground motion, and the intensity parameter has been gradually phased out due to its lack of physics basis and logistic reasoning (Calvi et al. 2006). The currently used parameters include Peak Ground Acceleration (PGA), Peak Ground Velocity (PGV), and Spectral Acceleration (SA), among others. The ground motion at building sites is usually predicted with Ground Motion Prediction Equations (GMPE), and is established based on regression analysis of observed ground motion records. Although GMPEs are often used to represent ground motion attenuation from the earthquake source to the site, they are statistical representations of the ground motion, thus their prediction accuracy is limited and the prediction usually deviates from the actual ground motion for a specific earthquake event (Rohmer et al. 2014). Therefore, a better approach with higher accuracy needs to be found to estimate the ground motion for an actual event.

Building earthquake loss estimate is a complex process that involves many components from earthquake source, wave propagation, site effect, to building vulnerability, and each component encompasses various degrees of uncertainty that have significant impact on quantifying the tail risk of the loss event (Jiang and Ye 2020). The approach by Crowley et al. (2004) or Yong et al. (2001) only provided the mean loss estimate, without considering the impact of uncertainty. The HAZUS approach (HAZUS-MH 2022) assumes that with given ground motion level the building damage distribution in each state conforms to the shape of a log-normal function. The discrete damage state is not sufficient in representing the complex damage status, and the log-normal function does not represent the actual damage distribution well (Zhao et al. 2021). The PAGER approach (Jaiswal and Wald 2011) uses a coefficient of standard deviation ζ for each country or region to describe the overall loss estimation uncertainty (1.947 for Japan), which is clearly not sufficient to reflect the complex processes of earthquake loss estimates in different countries and for different types of earthquakes. Due to the propagation and convolution of uncertainty in the loss estimation process, the loss amount at each building site is a random number, whose determination requires a mean, a standard deviation, and a distribution shape (Gómez Zapata et al. 2022). When the building losses are added together for all the buildings in the seismic impact area, there is also the need to consider the spatial correlation in the aggregation process (Zhou et al. 2022). Therefore, the post-earthquake loss estimate is an estimate built upon the convoluted probability correlation of multiple variables, which cannot be simply represented by a mean, a standard deviation, and its simple distribution. In order to obtain an estimate with better accuracy, a simulated approach with a huge number of

samples that consider various types of probability and their correlation is required.

To address these issues, this study proposed an improved approach for rapid post-earthquake building loss estimation. In the proposed approach, building site ground motion is spatially interpolated based on the observed ground motion using GMPE as the weighting function to improve the accuracy. In loss calculation, the combination of Gaussian copula, Monte Carlo simulation, and kriging is used to derive the huge sample set of loss quantiles considering various uncertainties and their correlation. The approach is validated against past historical reported losses, and is applied to predict the building losses for the March 2022 Fukushima earthquake.

2 Methodology

In the earthquake loss estimate approach based on structural vulnerability, the loss of the buildings can be simply expressed as (Pnevmatikos et al. 2020):

$$Loss = \sum_i IM \times Dr|IM \quad (1)$$

where IM is the ground motion intensity measure of building i , and $Dr|IM$ denotes the relative loss probability of building i under the excitation of IM . It can be seen in Eq. 1 that the estimation of earthquake loss includes mainly two efforts: calculating the ground motion intensity measures at the target building, and calculating the possible losses of the building based on IM . Therefore, advances in both aspects can ultimately be used to improve the accuracy of the seismic loss analysis. This section details our recent findings in ground motion estimation and loss calculation, and approaches that can improve computational efficiency while ensuring sufficient accuracy.

2.1 Cell Grid

Building loss estimates for a large earthquake often involve the number of buildings in the order of millions and, due to the speed required and the complexity of the task, a number of tasks need to be performed beforehand, such as collecting site condition information and geocoding of exposure buildings. For efficiency, these tasks need to be performed at cell grid level to maintain reusability (Rusanen et al. 1993). The properly defined grid is used to represent the site conditions for the buildings within the grid to reduce the amount of computing efforts. The ideal approach is to use a dot-style grid so that each building is of its own grid to achieve the best accuracy, but this approach is practically impossible because of the speed requirement and the computing

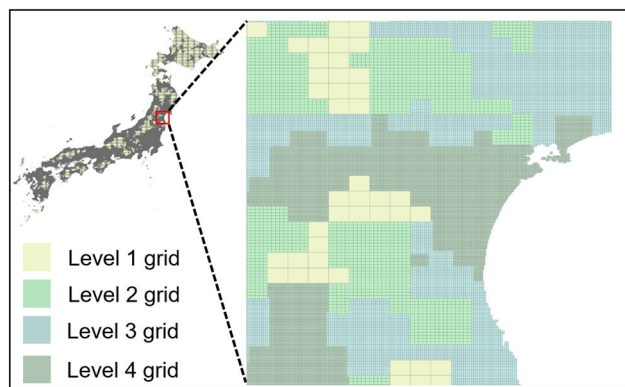


Fig. 1 A variable resolution grid for Japan. The grid size for level 1, 2, 3, and 4 is 5×5 km, 1×1 km, 0.5×0.5 km, and 0.25×0.25 km, respectively.

resource constraint. In this study, a variable resolution grid is defined based on the consideration of population density, gross domestic product (GDP), and seismicity level. Figure 1 shows the four-level resolution grid used for Japan in this study.

2.2 Spatial Interpolation of Ground Motion

The GMPE is a commonly used approach to predict the statistical mean ground motion at a given site, and its accuracy for any specific event is questionable since the approach is based on statistical regression. In a country like Japan where there are dense ground motion observation networks, the ground motion of sites close to the station can be better predicted using the observed ground motion records with proper implementation of spatial interpolation.

Spatial interpolation of ground motion needs to consider the impacting factors that affect ground motion, including earthquake source, path of seismic wave propagation, distance to earthquake source, and site conditions (Bora et al. 2015). In the case of an actual event, the earthquake source is the same, and the propagation path for neighboring sites are similar, so the variable factors are only the distance (epicentral or hypocentral distance) and site conditions. Therefore, the spatial interpolation only needs to consider the variation of distance and site conditions.

As shown in numerous previous studies (Hata et al. 2011), the ground motion attenuation with distance is very complex due to the complicated underground geological structure, and the site condition amplification has a strong nonlinear property, but the statistical effect of these two parameters is included in the GMPE. Therefore, to simplify the implementation process of ground motion spatial interpolation, the GMPE can be used as the weight to consider the impact of these two factors. The interpolation process for an example grid is shown in the following three steps.

Step 1 Selecting neighboring stations near the grid

For a given grid, the stations within 5 km are first identified, and if there are no matching stations, the search distance is increased to 10 km, 15 km, and 20 km, respectively until matching stations are found, which implies that multiple stations may be found to interpolate the ground motion for the defined grid. In Japan, there is at least one station within 20 km of a defined grid.

Step 2 Site condition modification

If the site condition of station i is different from the site of the grid center, modification needs to be performed to accommodate the change of site conditions. The GMPE is used to predict the ground motion at station i using the station site condition and the grid site condition. With the observed ground motion at station i , the site condition modified ground motion $Interp(Station_i, SC_{grid})$ can be expressed as:

$$Interp(Station_i, SC_{grid}) = Obs_i \times \frac{GMPE(Station_i, SC_{grid})}{GMPE(Station_i, SC_{station_i})} \quad (2)$$

where Obs_i is the actual record of station i , $GMPE(P_1, P_2)$ denotes the ground motion at location P_1 with site condition P_2 using GMPE, and similarly, $Interp(P_1, P_2)$ denotes the interpolated ground motion at location P_1 with site condition P_2 . According to the above definition, the difference between $GMPE(Station_i, SC_{grid})$ and $GMPE(Station_i, SC_{station_i})$ lies in the difference of the site conditions, so the ratio of the two can be used to correct the relative site amplification.

Step 3 Distance modification

When site condition modification is completed, the site conditions of the matching stations are the same as the defined grid, so the ratio of the predicted ground motion at the grid to the predicted ground motion at the station, $GMPE(Grid, SC_{grid}) / GMPE(Station_i, SC_{grid})$, can describe the influence of the earthquake focal distance, and when it is combined with $Interp(Station_i, SC_{grid})$ as in Step 2, the ground motion interpolation result $Interp(Grid, SC_{grid})_i$ of station i at the grid can be calculated as described in Eq. 4.

When the number of matching stations is more than 1, the $Interp(Grid, SC_{grid})_i$ obtained for different stations may differ, so it is necessary to consider the influence from each matching station. The predicted ground motion at each station under grid site conditions $GMPE(Station_i, SC_{grid})$ can be directly used as the weight W_i :

$$W_i = \frac{GMPE(Station_i, SC_{grid})}{\sum_i GMPE(Station_i, SC_{grid})} \quad (3)$$

Finally, the spatially interpolated ground motion $Interp(Grid, SC_{grid})$ at the grid is aggregated as in the following using the above weight:

$$\begin{aligned} \text{Interp}(\text{Grid}, SC_{\text{grid}}) &= \sum_i W_i \times \text{Interp}(\text{Grid}, SC_{\text{grid}})_i \\ &= \sum_i W_i \times \text{Interp}(\text{Station}_i, SC_{\text{grid}}) \times \frac{\text{GMPE}(\text{Grid}, SC_{\text{grid}})}{\text{GMPE}(\text{Station}_i, SC_{\text{grid}})} \end{aligned} \quad (4)$$

where W_i is the weight of station i , $\text{Interp}(\text{Grid}, SC_{\text{grid}})_i$ represents the ground motion interpolation result achieved based on station i , and it contains modification from two factors: site condition modification and distance modification, as explained in Steps 2 and 3, respectively.

The GMPE itself has substantial impact on the modification, therefore selecting applicable GMPE is very important. If rock site GMPE is chosen as the weight function, the site modification can be performed by adjusting ground motion directly instead of executing Step 2 (Paolucci et al. 2021).

2.3 Earthquake Loss Calculation Method

Earthquake loss distribution has to be considered in the loss estimate, and based on the analysis of a large amount of detailed earthquake loss data, it was found that the distribution is better represented by a Beta function (Zhao et al. 2021). Similar conclusions were reached by Hu et al. (2007) and Lallemand and Kiremidjian (2015) using earthquake loss data from the 1998 Ninglang earthquake in China and the 2010 Haiti earthquake, respectively. Assuming that the loss at the grid level is independent, Monte Carlo simulation can be used to sample the loss at the grid level directly, but due to the non-additive nature of the Beta function, the loss aggregation process must incorporate assumptions for loss spatial correlation.

Because of the lack of a large amount of actual detailed earthquake loss data, previous earthquake loss estimation mainly reflects the correlation of building losses indirectly by incorporating the spatial correlation of ground motion (Hu et al. 2022; Zhou et al. 2022). The closer the spatial distance of two grids (or buildings), the stronger the correlation of the ground motion intensity they are subjected to, and the more correlated the losses are. Researchers have found that the spatial correlation of ground motion differs in different regions and for different earthquakes in the same region (Hu et al. 2022), and that local site conditions and source parameters can also affect the ground motion correlations (Abbasnejadfar et al. 2020). In addition, Garakaninezhad and Bastami (2017) found that the isotropic assumption of spatial correlation of ground motion may be unreasonable, and it is complicated to accurately model the spatial correlation of ground motion.

In our previous study, a spatial correlation model of earthquake losses had been developed based on a large number of actual detailed building-level earthquake loss

data in Japan, and the model was validated using actual loss data of the 2010 Canterbury and 2011 Christchurch earthquakes in New Zealand (Zhou et al. 2022). Different from other studies, we did not model the spatial correlation of ground motion because we have access to a large number of the actual detailed earthquake loss data. Instead, we proposed a spatial correlation model for the earthquake loss, which is a function of the spatial distance between two grids (buildings) as in Eq. 5 (Zhou et al. 2022):

$$R = \text{Exp}(-0.02524 \times \Delta h) \quad (5)$$

where R is the correlation coefficient of damage ratio (Dr), Dr represents the ratio of building repair cost to replacement cost, which is a continuous number between 0 and 1 to better represent the building damage, and Δh is a scalar representing the spatial distance between two buildings (or grids).

Equation 5 indicates that exposed units spaced at shorter distances have higher correlation in earthquake losses, and the loss correlation decreases exponentially with increasing distance, which is similar to the decay pattern of the ground motion correlation (Abbasnejadfar et al. 2021), indirectly validating the formula format proposed in Eq. 5. Since Eq. 5 is directly derived from the actual data, it also contains correlation information for ground motion and for vulnerability as well as other factors, and there is no need to simulate the correlation through the convoluted process of integrating correlation information from ground motion, vulnerability, and other factors. Therefore, the correlation model based on actual earthquake loss is a more straightforward, simpler approach than other approaches to modeling the earthquake loss correlation.

In the earthquake loss aggregation process, spatial correlation affects the uncertainty estimation, which in turn impacts the distribution of the estimate. If we assume that the loss estimate at the grid level is a random variable, the estimate for all the grids involved will be a multi-variable set of (L_1, L_2, \dots, L_k) , where k is the number of grids. For any given L_i , its marginal distribution is assumed to be a Beta distribution. If the joint distribution for the multi-variable set can be constructed, the total loss can be integrated using the convolution of the joint distribution. For describing the correlation in the joint distribution Gaussian copula is introduced as in the following.

2.3.1 Copula Theory and Its Application

Copula function is an approach used for correlation in multi-random variables (Hochrainer-Stigler et al. 2018). For the multi-variable set described above, the joint probability density function can be expressed as:

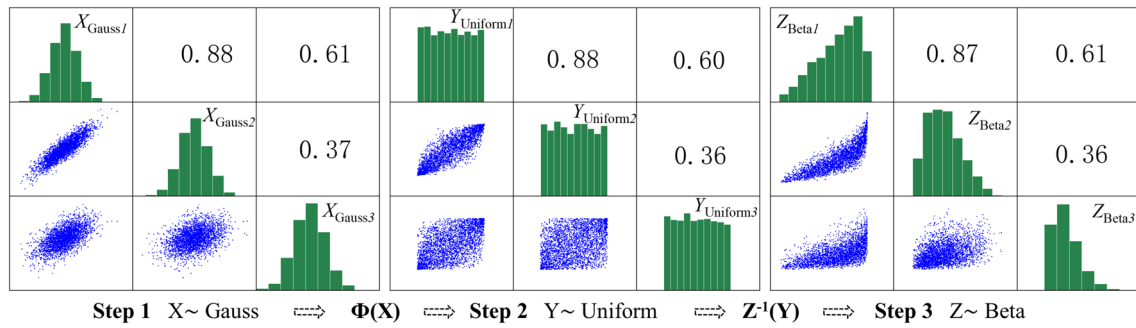


Fig. 2 Random loss number sampling based on Gaussian copula

$$\begin{aligned}
 f(l_1, l_2, \dots, l_k) &= \frac{\partial^k H(l_1, l_2, \dots, l_k)}{\partial l_1 \partial l_2 \dots \partial l_k} = \frac{\partial^k C(F(l_1), F(l_2), \dots, F(l_k))}{\partial l_1 \partial l_2 \dots \partial l_k} \\
 &= \frac{\partial^k C(u_1, u_2, \dots, u_k)}{\partial u_1 \partial u_2 \dots \partial u_k} \cdot \frac{\partial F_1(l_1)}{\partial l_1} \cdot \frac{\partial F_2(l_2)}{\partial l_2} \cdot \dots \cdot \frac{\partial F_k(l_k)}{\partial l_k} \\
 &= c(F(l_1), F(l_1) \dots, F(l_k)) \cdot f_1(l_1) f_2(l_2) \cdot \dots \cdot f_k(l_k)
 \end{aligned}
 \tag{6}$$

where $H(l_1, l_2, \dots, l_k)$ is the joint distribution function for multi-random variable (L_1, L_2, \dots, L_k) , and C is the copula function with marginal distributions F_1, F_2, \dots, F_k . Equation 6 shows that a joint probability density function can be represented by the product of copula function and marginal probability density distribution function, thus a joint function can be separated into the marginal distribution and correlation function. The randomness of the variables is expressed by their marginal distribution, and the correlation is expressed by copula function. There are many types of copula functions, and the Gaussian copula function was selected in this study due to its simplicity in computation and implementation (Goda and Tesfamariam 2015).

2.3.2 Seismic Loss Sampling Process

Loss integration over the joint probability distribution often requires a number of simplification assumptions for the correlation and integration, and the integration process is not only time- and resource-consuming but also lacks good accuracy. To address these issues, a sampling approach is proposed to estimate the random loss number from the joint distribution, which is explained as follows.

A random variable set $\mathbf{Z}_{\text{Beta}} = (Z_{\text{Beta}1}, \dots, Z_{\text{Beta}N})$ that conforms to marginal Beta distribution with a correlation matrix of \mathbf{R} can be constructed to represent seismic loss, where N is the number of grids within the seismic impact area, $Z_{\text{Beta}i}$ is the random loss number at grid i , and $\mathbf{R} \in [0,1]^{N \times N}$, R_{ij} represents the loss spatial correlation between grid i and grid j , and it can be established using Eq. 5, so \mathbf{R} satisfies $R_{ij} = R_{ji}$, and $R_{ii} = 1$. Three steps are needed to complete the sampling of the random variable set \mathbf{Z}_{Beta} .

Step 1 An N -dimension random variable set $\mathbf{X}_{\text{Gauss}} = (X_{\text{Gauss}1}, \dots, X_{\text{Gauss}N})$ with correlation matrix \mathbf{R} can be generated using Gaussian copula. Due to the linear property of Gaussian functions, the generation process can be realized by combining the Monte Carlo sampling of the Gaussian variable and Cholesky decomposition:

$$\mathbf{R} = \mathbf{A} \cdot \mathbf{A}^T \tag{7}$$

$$(X_{\text{Gauss}1}, \dots, X_{\text{Gauss}N}) = (I_1, \dots, I_N) \cdot \mathbf{A} \tag{8}$$

where \mathbf{A} is triangular matrix of \mathbf{R} after Cholesky decomposition, and (I_1, I_2, \dots, I_N) is the Gaussian independent variable by Monte Carlo sampling.

Step 2 Calculating the Gaussian distribution function $\Phi(\cdot)$ of $\mathbf{X}_{\text{Gauss}}$ will result in N -dimension random variable set $\mathbf{Y}_{\text{Uniform}} = (Y_{\text{Uniform}1}, \dots, Y_{\text{Uniform}N})$, where $Y_{\text{Uniform}i}$ is a uniform distribution in $[0,1]$ according to the probability integral transform approach as described in Angus (1994).

Step 3 Utilizing the quantile transform approach (Gatti 2004), the inverse function $Z_{\text{Beta}}^{-1}(\cdot)$ of the marginal distribution function of the uniform distribution $\mathbf{Y}_{\text{Uniform}}$ can be derived, thus resulting in the loss random number set of $(Z_{\text{Beta}1}, \dots, Z_{\text{Beta}N})$ that conforms to both the correlation matrix and the marginal distribution.

Figure 2 shows the sampling process for three grid cells. As shown in Fig. 2, the correlation matrix for the random variable has not changed. In addition, in Step 3, the marginal distribution $Z_{\text{Beta}i}$ for each grid is determined by the mean and variance of the building loss corresponding to the ground motion in the grid, and its shape changes as the Beta function parameters α and β change, demonstrating the variation for different ground motion and different types of buildings.

When the number of grids is small or in the order of thousands, the sampling process as described above can be directly applied. Because computing time increases exponentially with the increase of the grid number, the sampling has to be performed at limited grids, and kriging is used

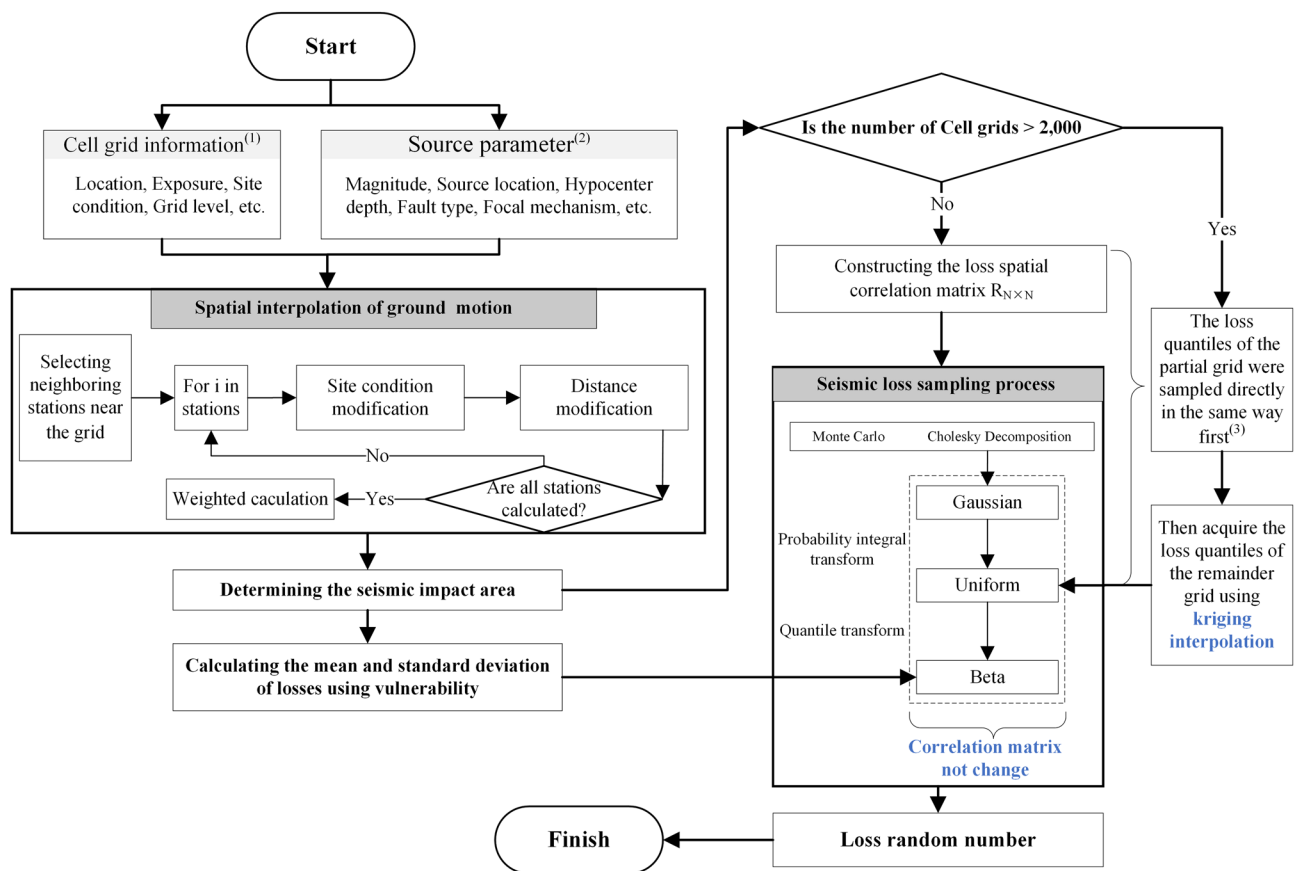


Fig. 3 A framework for post-earthquake building loss estimation based on observed ground motion considering various types of uncertainty and spatial loss correlation

to interpolate the loss quantile for other grids that are not sampled. The kriging approach has been verified by Wang (2020) to be effective in generating loss quantiles for earthquake loss estimate.

2.4 General Framework

To further explain the loss calculation process described above, a framework chart is provided in Fig. 3.

Additional explanations of Fig. 3 are:

- (1) Cell grid is used to improve the computational speed, and in practice grid information should be prepared in advance, as explained in Sect. 2.1.
- (2) Source parameters are determined by the GMPE model used in interpolating the ground motion.
- (3) Additional information on the recommended number of grids for direct sampling can be found in Zhou et al. (2022).

3 Method Validation

The objective of this section is to verify the correctness of the proposed method in the implementation and the validity of the earthquake loss estimation. Losses to Zenkyoren¹ insurance policies from three historical earthquakes in Japan—the 2011 Great East Japan Earthquake, February 2021 Fukushima Earthquake, and May 2021 Fukushima Earthquake—were used to validate the proposed method.

3.1 Data Preparation

To validate the proposed approach the exposure dataset of Zenkyoren is used. The exposure is an inventory of more than 3 million buildings, which are geocoded to the variable

¹ Zenkyoren (Japan) is also known as the National Federation of Agricultural Cooperative Mutual Insurance and provides building endowment insurance, which insures buildings owned by the policyholder against damage caused by fire, earthquake, or other natural hazard-related disasters.

Fig. 4 Building inventory distribution for the Zenkyoren insurance dataset. RC reinforced concrete

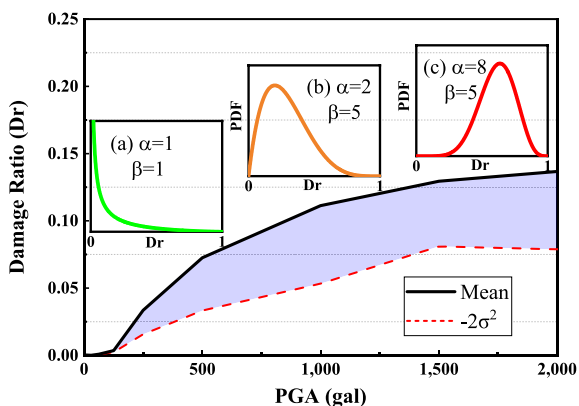
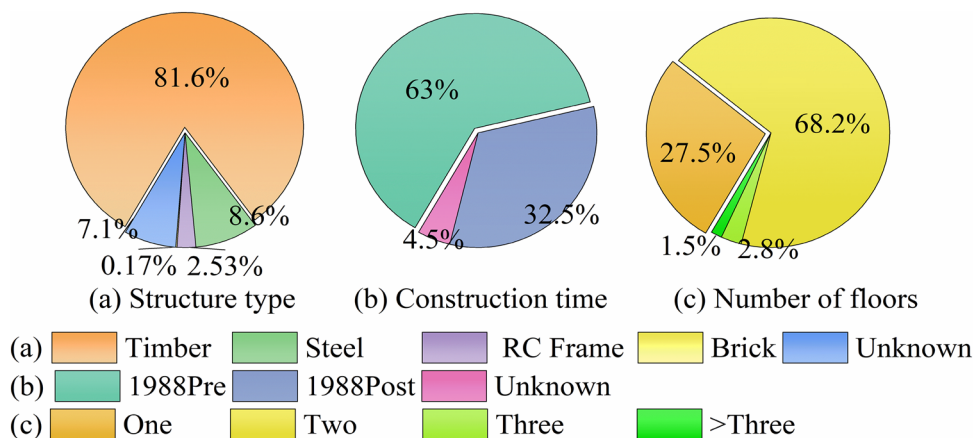


Fig. 5 Schematic vulnerability for a typical type of building with varying uncertainty distribution. PDF Probability density function; PGA Peak ground acceleration

grids explained above. The distribution of the building inventory is shown in Fig. 4, and the site condition of V_s30 (the average S-wave velocity of the top 30 m of the strata) is assembled according to the dataset developed by J-SHIS (Japan Seismic Hazard Information Station) at 250 m resolution (Imamura and Furuta 2015).

The structural vulnerability represents the relationship between ground motion and damage ratio (Armas et al. 2017) and is derived based on the detailed loss data from the 2011 Great East Japan Earthquake. Figure 5 shows the schematic vulnerability of a typical building type, and the uncertainty of loss distribution is also shown for three different levels with varying Beta function parameters. For this study, PGA was chosen as the ground motion parameter, and Dr was used as the building damage parameter.

Based on the observed ground motion data from KiK-net and K-NET, Zhao et al. (2016) proposed a GMPE for the subduction type of earthquakes in Japan considering the source mechanism and fault type. Since Zhao et al. (2016) used the Japanese dataset and its effectivity was also proven

by Tong et al. (2022), it was selected for this study. The formula of Zhao et al. (2016) is as follows:

$$\ln(y) = f_{mints} + g_{int} \cdot \ln(r) + g_{intsL} \cdot \ln(x + 200) + e_{ints} \cdot x + e_{int}^V \cdot x^V + g_{int} + \ln(A) \tag{9}$$

where the items with r or x are distance related, and A is the site amplification item including linear and nonlinear effects.

3.2 Validation of the Method

The first validation is to make sure that the approach is properly implemented. Since the vulnerability is mostly based on the data from the 2011 Great East Japan Earthquake, the estimate for the same event was performed, and the result was USD 6.08 billion (based on the exchange rate in March 2011), while the actual reported building loss for Zenkyoren was about USD 6.70 billion. Although there is a slight difference, the estimate is appropriate given the large uncertainty in the estimation process, verifying the correctness in implementation.

Further validation was performed for two other events. In February and May 2021, two earthquakes occurred near the east coast of Honshu Island, with magnitudes of 7.1 and 6.8, respectively (GIAJ 2021; Wikipedia 2021). The building loss estimation was performed for these two events, and the results are shown in Figs. 6 and 7 with a total building mean loss of USD 0.757 and 0.117 billion (based on the exchange rate in February 2021). Zenkyoren reported that the corresponding loss was USD 1.016 and 0.122 billion, respectively (Zenkyoren 2022), which included loss from other perils such as landslides and fire. Historical statistics show that the building loss by shaking was about 80% of the total loss (Bird and Bommer 2004). The estimate by this study for the February 2021 event was 74.5% of the total loss, and for the May 2021 event it was 96.2%. Overall, the loss estimates for the two earthquake events are acceptable considering the large uncertainty in the estimation process.

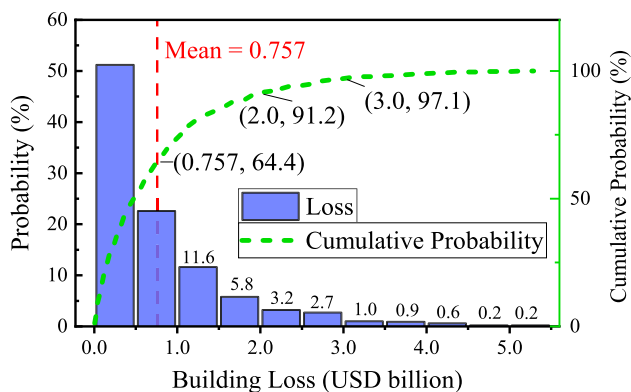


Fig. 6 Building loss estimate for the February 2021 earthquake event in Japan

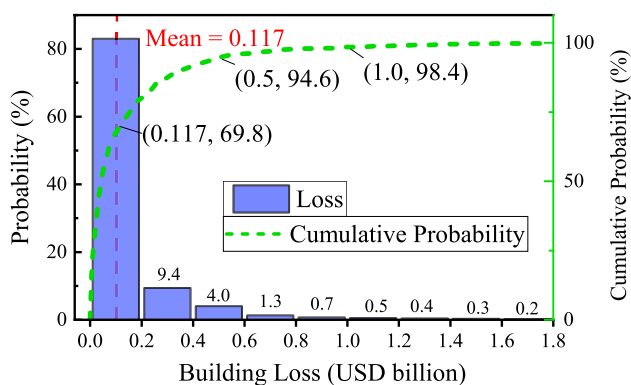


Fig. 7 Building loss estimate for the May 2021 earthquake event in Japan

Recent studies have shown that ignoring the spatial correlation of ground motion could overestimate the rare losses and underestimate frequent losses (Abbasnejadfar et al. 2020). From the loss estimation results for the two earthquakes in 2021, we slightly underestimated the loss for the February event and overestimated the loss for the May event, indicating no bias in our estimation. It should be noted that the higher estimate for the May event was probably caused by two factors—the weak buildings were already damaged or cleared out by the February event, and there was probably retrofitting for weak buildings after the February event.

The cumulative probability distribution curves in Figs. 6 and 7 show that the cumulative probability densities (CPD) at the mean values for the February and May events are different—64.4% and 69.8%, respectively. This is different from PAGER, whose standard deviation and distribution are fixed for all earthquakes, and the CPD at the mean loss estimates by PAGER is always 50% for all cases (Jaiswal and Wald 2011). This kind of difference is caused by the fact that the proposed approach is a simulated estimate and the PAGER approach is a statistical one.

The probability distribution of the two earthquake loss estimates is positively skewed with a long tail on the right side of the distribution in Figs. 6 and 7, which conforms to the statistical characteristics of catastrophe risk (Lane and Mahul 2008). The maximum loss for the February event is USD 5.24 billion, and the maximum loss for the May event is USD 1.8 billion, corresponding to about 6.9 times and 14.9 times the average loss of the two events, respectively, indicating that the possibility of extreme loss can be obtained in the proposed method. In contrast, the maximum loss by PAGER is infinite, and although the right long tail of PAGER’s loss distribution becomes thinner and thinner with the increase of the loss (Jaiswal and Wald 2011), it can extend indefinitely in theory. It is well known that in risk management and catastrophe insurance, the capital required for extreme losses with a small probability of occurrence is increasing in geometric order. Therefore, compared to PAGER, the proposed loss estimation method is more realistic in terms of precision and application in risk management.

Even though the probability distributions of both earthquake loss estimates are positively skewed, the coefficients of variation and skewness of the two distributions are different—1.12, 1.98, and 1.79, 3.67, respectively. In fact, the probability distributions of overall loss are jointly regulated by the marginal probability density of the loss at the grid level and the spatial loss correlation within the seismic impact area. Therefore, the probability distribution characteristics of the loss estimation results are unique for each individual earthquake event.

The above analysis demonstrates that the proposed method can estimate earthquake losses reliably, both in terms of mean and probability distributions. The loss estimation result is an integration of the spatial interpolation method in Sect. 2.2 and the loss calculation in Sect. 2.3 according to the general framework in Fig. 3. Through the comparison between estimated building loss and the actual earthquake building loss for three different earthquake events, the validity of the proposed approach is verified and the proposed approach can be used to predict future Japanese earthquake building loss.

4 Application

On 16 March 2022, an earthquake with a magnitude of 7.3 occurred offshore Fukushima with a depth of 60 km (USGS 2022). This event occurred near the Pacific plate boundary and was of the subduction type. The Japan Earthquake Research Committee provided more information on the fault model and source parameters (JERC 2022). It was the most damaging earthquake in the last three years, and caused derailment of the Shinkansen high-speed train, 4 deaths and 225 injuries, and left more than 2 million houses without

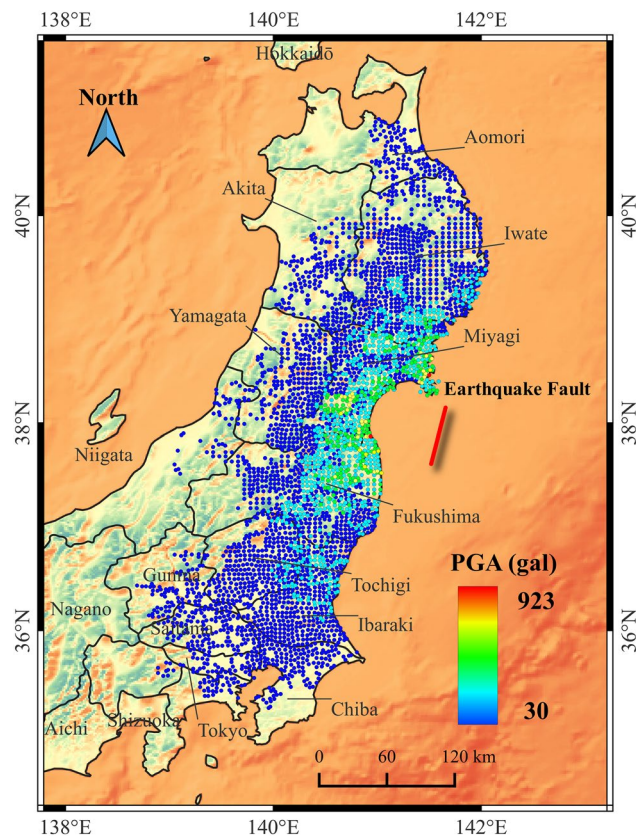


Fig. 8 Interpolated peak ground acceleration (PGA) within the earthquake impact area for the March 2022 Fukushima earthquake event in Japan

electricity (Wikipedia 2022). Zenkyoren had not published its loss report as of the time this study was completed, and the proposed approach in this study is used to predict the building losses with discussions on the approach and the results.

The spatial interpolation approach was used to estimate the PGA at all grids, and grids with a PGA level above 30 gal were used to define the seismic impact area. The PGA distribution within the impact area is shown in Fig. 8. Based on the PGA interpolated for the impacted grids, loss quantile sampling was performed 2000 times for all the grids to ensure good accuracy in uncertainty prediction and performance speed. The total building loss was aggregated for each sample. The distribution of the sampled results is shown in Fig. 9, indicating that the estimate has a 58.1% chance in USD 0–1 billion and a 27.1% chance in USD 1–2 billion (based on the exchange rate in March 2022). The estimate has a long tail in the distribution, representing high probability of small and medium loss and low probability of big loss. From the loss probability accumulation, the probability of the estimate below USD 1.325 (mean), 3, and 6 billion is 63.5%, 90.4%, and 98.5%, respectively. The mean estimate is USD 1.325 billion.

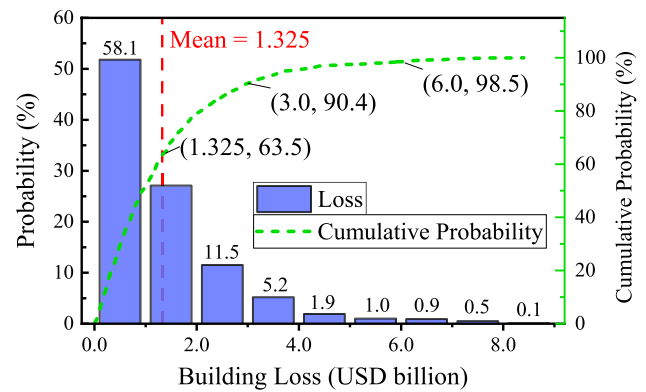


Fig. 9 Building loss estimate distribution for the March 2022 Fukushima earthquake event in Japan

5 Discussion

This section briefly discusses the results obtained in the application section. Figure 10 shows the comparison of PGA by GMPE, at observation stations and interpolated for site class II ($300 < V_s30 \leq 600$) and III ($200 < V_s30 \leq 300$). As shown in Fig. 10, when distance to fault is below 150 km, GMPE underpredicts PGA for both site classes II and III, but the interpolated PGA agrees well with station observed PGA, validating the argument that the proposed spatial interpolation approach has a higher degree of accuracy in predicting the ground motion than GMPE.

The station matching distribution for all grid cells is shown in Fig. 11, where 83.8% of the grids have only 1 or 2 matching stations for interpolation. For grids with only one matching station, 94.6% of the grids have a matching distance less than 10 km, indicating greater impact of distance than that from the number of matching stations. The farther the matching distance, the lower the correlation of ground motion, thus the lower the quality of the interpolation results. Conversely, the shorter the matching distance, the smaller the impact from other factors, thus the higher the accuracy of the interpolation results. A dense network of ground motion stations can help improve the reliability of the ground motion spatial interpolation results, therefore more monitoring stations are desirable in earthquake-prone areas to provide better ground motion information for future post-earthquake loss estimation and emergency management.

The weight by each station when there are only two matching stations is shown in Fig. 12, where the weight is pretty similar for both stations, and it is mostly between 0.4 and 0.6, and the reason is that the matching distance is mostly small (often less than 5 km). This shows that there is no uneven impact from a single station when there are two stations.

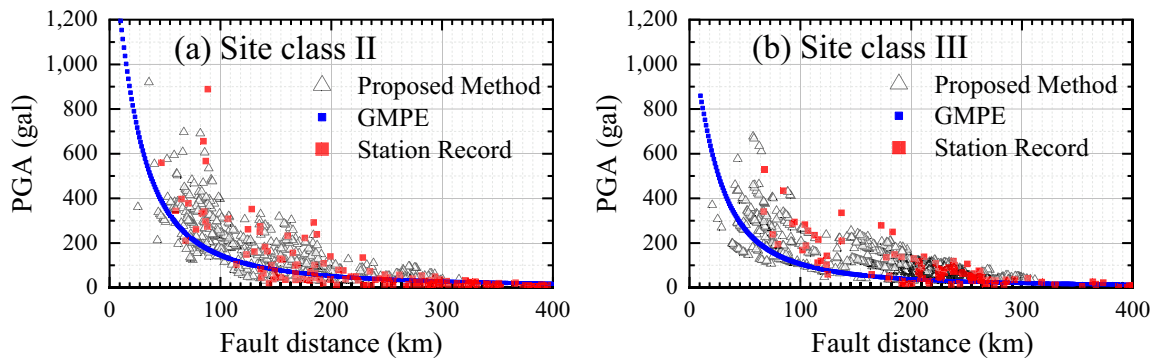


Fig. 10 Comparison of peak ground acceleration (PGA) by Ground Motion Prediction Equations (GMPE), observed and spatially interpolated for the March 2022 Fukushima earthquake event in Japan. **a** Site class II; **b** Site class III.

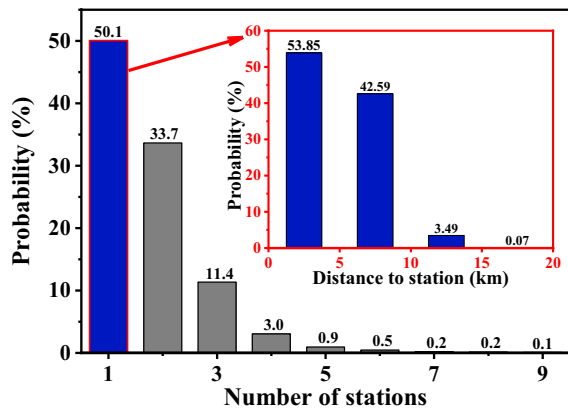


Fig. 11 The distribution of matching number of stations for the March 2022 Fukushima earthquake event in Japan

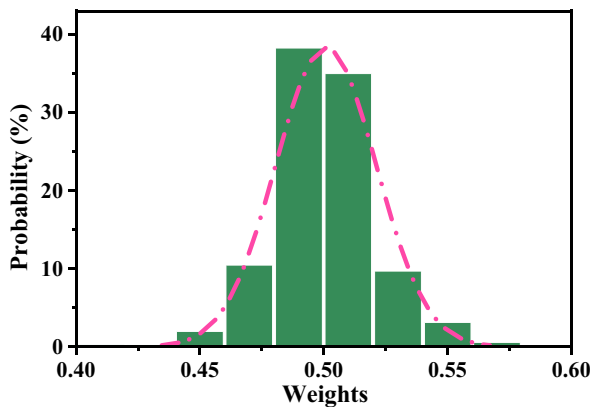


Fig. 12 Station weight distribution when there are two matching stations for the March 2022 Fukushima earthquake event in Japan

When the loss is aggregated at the prefecture level, the estimated mean and coefficient of variation (CV) are shown in Fig. 13. The loss is concentrated in Fukushima and Miyagi Prefectures, with a mean of USD 468.5 and 405.1 million,

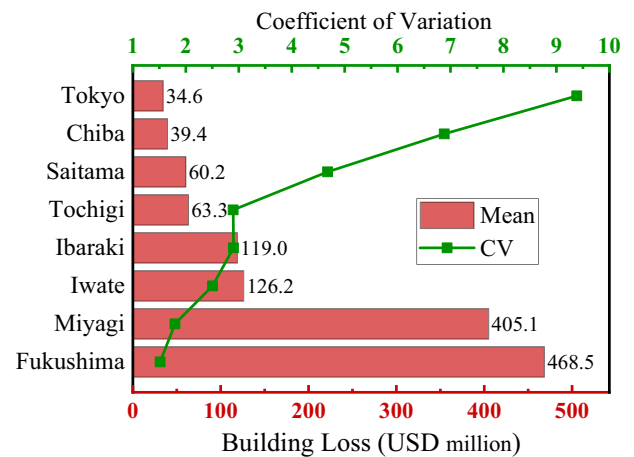


Fig. 13 Estimate mean and coefficient of variation (CV) at prefecture level for the March 2022 Fukushima earthquake event in Japan

respectively. This distribution of building loss estimates correlates well with the predicted PGA in Fig. 8, which is also observed in the three historical earthquakes (Sect. 3.2). The CV is much different for different prefectures, demonstrating the superiority of the proposed approach in representing the varying standard deviation in different regions over the PAGER approach that uses a constant to represent loss variation for any given country or region.

6 Conclusion

To address the issues in traditional loss estimate approaches, this study proposed an improved approach based on spatially interpolated ground motion and big data sampling that considers the spatial correlation of earthquake losses. The 2011 Great East Japan Earthquake was used to verify the implementation, and the February and May 2021 earthquake events were used to validate the appropriateness of the proposed approach. The proposed method was then applied

to predict the March 2022 Fukushima earthquake building losses with detailed discussions.

Comparison of PGA between the observed, the interpolated, and the estimated by GMPE shows that the interpolated approach is much better than the GMPE in predicting PGA for all the grids in the earthquake impact area. The interpolation using the GMPE as the weight to modify the change effect of both focal distance and site condition has been proven to be effective in predicting earthquake ground motion.

The detailed building loss estimate through intensive sampling not only provides a distribution of loss at different levels, but also in different prefectures. The varying CV in different prefectures demonstrates the effectiveness of the proposed approach in identifying variation difference for different regions in the same earthquake event, which has not been implemented in other approaches.

The proposed approach has good generalization capability for different countries and regions, where the corresponding GMPE and vulnerability need to be replaced. Using the grid cell and kriging approach, dimension reduction has been achieved for speed enhancement, providing a fast estimation approach after an earthquake. The two core components proposed in this study, the spatial interpolation of ground motion and the spatially correlated loss quantile sampling, can be easily adapted in future risk quantification systems for applications in other disaster mitigation efforts.

Correlation is a key factor in random variable quantification and aggregation, and only spatial correlation was considered in this study. Correlation of different buildings within the same grid and between different types of loss such as content and business interruption will further impact the uncertainty of loss estimate distribution and will be a topic in future studies.

Acknowledgements This study was supported by the Scientific Research Fund of the Institute of Engineering Mechanics, China Earthquake Administration (Grant No. 2021B09) and the National Natural Science Foundation of China (Grant No. 51978634).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbasnejadfar, M., M. Bastami, and A. Fallah. 2020. Investigation of anisotropic spatial correlations of intra-event residuals of multiple earthquake intensity measures using latent dimensions method. *Geophysical Journal International* 222(2): 1449–1469.
- Abbasnejadfar, M., M. Bastami, A. Fallah, and A. Garakaninezhad. 2021. Analyzing the effect of anisotropic spatial correlations of earthquake intensity measures on the result of seismic risk and resilience assessment of the portfolio of buildings and infrastructure systems. *Bulletin of Earthquake Engineering* 19(14): 5791–5817.
- Angus, J.E. 1994. The probability integral transform and related results. *SIAM Review* 36(4): 652–654.
- Armas, I., D. Toma-Danila, R. Ionescu, and A. Gavris. 2017. Vulnerability to earthquake hazard: Bucharest case study, Romania. *International Journal of Disaster Risk Science* 8(2): 182–195.
- Bird, J.F., and J.J. Bommer. 2004. Earthquake losses due to ground failure. *Engineering Geology* 75(2): 147–179.
- Bora, S.S., F. Scherbaum, N. Kuehn, P. Stafford, and B. Edwards. 2015. Development of a response spectral ground-motion prediction equation (GMPE) for seismic-hazard analysis from empirical Fourier spectral and duration models. *Bulletin of the Seismological Society of America* 105(4): 2192–2218.
- Calvi, G.M., R. Pinho, G. Magenes, J.J. Bommer, L.F. Restrepo-Vélez, and H. Crowley. 2006. Development of seismic vulnerability assessment methodologies over the past 30 years. *ISET Journal of Earthquake Technology* 43(3): 75–104.
- Crowley, H., R. Pinho, and J.J. Bommer. 2004. A probabilistic displacement-based vulnerability assessment procedure for earthquake loss estimation. *Bulletin of Earthquake Engineering* 2(2): 173–219.
- Dell'Acqua, F., and P. Gamba. 2012. Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. *Proceedings of the IEEE* 100(10): 2876–2890.
- Garakaninezhad, A., and M. Bastami. 2017. A novel spatial correlation model based on anisotropy of earthquake ground-motion intensity. *Bulletin of the Seismological Society of America* 107(6): 2809–2820.
- Gatti, P.L. 2004. *Probability theory and mathematical statistics for engineers*. Boca Raton: CRC Press.
- Ghasemof, A., M. Mirtaheeri, and R.K. Mohammadi. 2022. Multi-objective optimization for probabilistic performance-based design of buildings using FEMA P-58 methodology. *Engineering Structures* 254: Article 113856.
- GIAJ (Geospatial Information Authority of Japan). 2021. Source fault model for the May 1, 2021 Miyagi-ken-oki earthquake. <https://www.gsi.go.jp/cais/topic20210501-index.html>. Accessed 3 Oct 2022.
- Goda, K., and S. Tesfamariam. 2015. Multi-variate seismic demand modelling using copulas: Application to non-ductile reinforced concrete frame in Victoria, Canada. *Structural Safety* 56: 39–51.
- Gómez Zapata, J.C., M. Pittore, F. Cotton, H. Lilienkamp, S. Shinde, P. Aguirre, and H. Santa María. 2022. Epistemic uncertainty of probabilistic building exposure compositions in scenario-based earthquake loss models. *Bulletin of Earthquake Engineering* 20: 2401–2438.
- Hata, Y., A. Nozu, and K. Ichii. 2011. A practical method to estimate strong ground motions after an earthquake, based on site amplification and phase characteristics. *Bulletin of the Seismological Society of America* 101(2): 688–700.
- HAZUS-MH (Hazards U.S. Multi-Hazard). 2022. *Multi-hazard loss estimation methodology: Hazus5.1 earthquake model technical manual*. Washington, DC: Federal Emergency Management Agency (FEMA).

- Hochrainer-Stigler, S., G. Pflug, U. Dieckmann, E. Rovenskaya, S. Thurner, S. Poledna, G. Boza, J. Linnerooth-Bayer, and A. Brännström. 2018. Integrating systemic risk and risk analysis using copulas. *International Journal of Disaster Risk Science* 9(4): 561–567.
- Hosseinpour, V., A. Saeidi, M.J. Nollet, and M. Nastev. 2021. Seismic loss estimation software: a comprehensive review of risk assessment steps, software development and limitations. *Engineering Structures* 232: Article 111866.
- Hu, S., B. Sun, and D. Wang. 2007. Approach to making empirical earthquake damage matrix. *Earthquake Engineering and Engineering Dynamics* 27(6): 46–50 (in Chinese).
- Hu, J., L. Hu, H. Zhang, C. Jin, Z. Wang, and Y. Ding. 2022. Uncertainty analysis and spatial correlation of ground motion in the Kanto Basin. *Japan. Journal of Earthquake Engineering* 26(13): 7056–7072.
- Imamura, S., and A. Furuta. 2015. A practical method of nation wide Vs30 mapping with 250m grids for developing countries by automated topographical classification and published data in Japan. In *Proceedings of the 12th SEGJ International Symposium*, 18–20 November 2015, Tokyo, Japan.
- Jaiswal, K., D. Wald, and K. Porter. 2010. A global building inventory for earthquake loss estimation and risk management. *Earthquake Spectra* 26(3): 731–748.
- Jaiswal, K., and D.J. Wald. 2011. Rapid estimation of the economic consequences of global earthquakes. Reston, VA: US Department of the Interior, US Geological Survey.
- JERC (Japan Earthquake Research Committee). 2022. Assessment of the March 16, 2022, earthquake off the coast of Fukushima Prefecture. https://www.static.jishin.go.jp/resource/monthly/2022/20220316_fukushima_1.pdf. Accessed 7 Oct 2022.
- Jiang, L., and J. Ye. 2020. Quantifying the effects of various uncertainties on seismic risk assessment of CFS structures. *Bulletin of Earthquake Engineering* 18(1): 241–272.
- Kalakonas, P., and V. Silva. 2022. Earthquake scenarios for building portfolios using artificial neural networks: Part II—damage and loss assessment. *Bulletin of Earthquake Engineering*. <https://doi.org/10.1007/s10518-022-01599-2>.
- Krausmann, E., and A.M. Cruz. 2013. Impact of the 11 March 2011, Great East Japan earthquake and tsunami on the chemical industry. *Natural Hazards* 67(2): 811–828.
- Lallemant, D., and A. Kiremidjian. 2015. A beta distribution model for characterizing earthquake damage state distribution. *Earthquake Spectra* 31(3): 1337–1352.
- Lane, M., and O. Mahul. 2008. Catastrophe risk pricing: An empirical analysis. World Bank Policy Research Working Paper No. 4765. Washington, DC: World Bank.
- Paolucci, R., M. Aimar, A. Ciancimino, M. Dotti, S. Foti, G. Lanzano, P. Mattevi, F. Pacor, and M. Vanini. 2021. Checking the site categorization criteria and amplification factors of the 2021 draft of Eurocode 8 Part 1–1. *Bulletin of Earthquake Engineering* 19(11): 4199–4234.
- Pnevmatikos, N., F. Konstandakopoulou, and N. Koumoutsos. 2020. Seismic vulnerability assessment and loss estimation in Cephalonia and Ithaca islands, Greece, due to earthquake events: A case study. *Soil Dynamics and Earthquake Engineering* 136: Article 106252.
- Rohmer, J., J. Douglas, D. Bertil, D. Monfort, and O. Sedan. 2014. Weighing the importance of model uncertainty against parameter uncertainty in earthquake loss assessments. *Soil Dynamics and Earthquake Engineering* 58: 1–9.
- Rusanen, J., A. Naukkarinen, and A. Colpaert. 1993. Square kilometer grid system: An efficient database in rural studies. *Geography Research Forum* 13: 129–138.
- Stojadinović, Z., M. Kovačević, D. Marinković, and B. Stojadinović. 2022. Rapid earthquake loss assessment based on machine learning and representative sampling. *Earthquake Spectra* 38(1): 152–177.
- Tong, W., Z. Wang, D. Zhao, and X. Zhang. 2022. Comparison of vulnerability based on detailed loss data from the March 11, 2011 Japan earthquake. *Earthquake Engineering and Engineering Dynamics* 42(3): 171–179 (in Chinese).
- USGS (U.S. Geological Survey). 2022. M7.3-57 km ENE of Namie, Japan 2022. https://en.wikipedia.org/wiki/2022_Fukushima_earthquake. Accessed 9 Oct 2022.
- Wang, Z. 2008. A preliminary report on the Great Wenchuan Earthquake. *Earthquake Engineering and Engineering Vibration* 7(2): 225–234.
- Wang, Z. 2020. Spatial correlation of earthquake loss and its impact on loss aggregation. In *Proceedings of the 17th World Conference on Earthquake Engineering*, Paper No. C000089, 27 September–2 October 2021, Sendai, Japan.
- Wikipedia. 2021. 2021 Fukushima earthquake. https://en.wikipedia.org/wiki/2021_Fukushima_earthquake. Accessed 15 Oct 2022.
- Wikipedia. 2022. 2022 Fukushima earthquake. https://en.wikipedia.org/wiki/2022_Fukushima_earthquake. Accessed 15 Oct 2022.
- Yong, C., Q.-F. Chen, and C. Ling. 2001. Vulnerability analysis in earthquake loss estimate. *Natural Hazards* 23(2): 349–364.
- Zeng, X., X. Lu, T.Y. Yang, and X. Xu. 2016. Application of the FEMA-P58 methodology for regional earthquake loss prediction. *Natural Hazards* 83(1): 177–192.
- Zenkyoren (Japan National Mutual Insurance Federation of Agricultural Cooperatives). 2022. 2022 Annual Report. https://www.ja-kyosai.or.jp/about/annual_report/pdf/2022annual.pdf. Accessed 21 Oct 2022.
- Zhang, Y., S. Zheng, L. Sun, L. Long, W. Yang, and L. Li. 2021. Developing GIS-based earthquake loss model: A case study of Baqiao District. *China. Bulletin of Earthquake Engineering* 19(5): 2045–2079.
- Zhao, J.X., X. Liang, F. Jiang, H. Xing, M. Zhu, R. Hou, Y. Zhang, and X. Lan et al. 2016. Ground-motion prediction equations for subduction interface earthquakes in Japan using site class and simple geometric attenuation functions. *Bulletin of the Seismological Society of America* 106(4): 1518–1534.
- Zhao, D., Z. Wang, Y. Liu, and W. Tong. 2021. Earthquake loss uncertainty based on detailed loss data in New Zealand. *Earthquake Engineering and Engineering Dynamics* 41(2): 84–95 (in Chinese).
- Zhou, Y., Z. Wang, L. Shi, and W. Tong. 2022. Spatial correlation of earthquake damage and its influence on earthquake loss distribution estimation. *World Earthquake Engineering* 38(2): 151–159 (in Chinese).