A framework for understanding uncertainty in seismic risk assessment

Foulser-Piggott, Roxane; Bowman, Gary; Hughes, Martin

Published in:
Risk Analysis

DOI:
10.1111/risa.12919

A Framework for Understanding Uncertainty in Seismic Risk Assessment

Roxane Foulser-Piggott[1], Gary Bowman[2*], and Martin Hughes[3]

[1] School of Mathematics and Physics, University of Queensland, Brisbane, Queensland, Australia

[2] Bond Business School, Bond University, Gold Coast, Queensland, Australia

[3] Cambridge Architectural Research, Cambridge, UK

[*] Correspondence to Gary Bowman, Bond Business School, Bond University, 14 University Drive, Robina, Queensland, 4226; gbowman@bond.edu.au

(there are no acknowledgements)

This is the peer reviewed version of the following article: Foulser-Piggott, R., Bowman, G., & Hughes, M. (2020). A framework for understanding uncertainty in seismic risk assessment. Risk Analysis, 40(1), 169-182, which has been published in final form at https://doi.org/10.1111/risa.12919.

ABSTRACT

A better understanding of the uncertainty that exists in models used for seismic risk assessment is critical to improving risk-based decisions pertaining to earthquake safety. Current models estimating the probability of collapse of a building do not consider comprehensively the nature and impact of uncertainty. This paper presents a model framework to enhance seismic risk assessment and thus give decision-makers a fuller understanding of the nature and limitations of the estimates. This can help ensure risks are not over or under estimated and the value of acquiring accurate data is appreciated fully. The methodology presented provides a novel treatment of uncertainties in input variables, their propagation through the model and their effect on the results. The study presents ranges of possible annual collapse probabilities for different case studies on buildings in different parts of the world, exposed to different levels of seismicity and with different vulnerabilities. A global sensitivity analysis was conducted to determine the significance of uncertain variables. Two key outcomes are (1) that the uncertainty in ground-motion conversion equations has the largest effect on the uncertainty in the calculation of annual collapse probability; and (2) the vulnerability of a building appears to have an effect on the range of annual collapse probabilities produced, i.e. the level of uncertainty in the estimate of annual collapse probability, with less vulnerable buildings having a smaller uncertainty.

Keywords: Aleatory and epistemic uncertainty; risk assessment; seismic hazard; sensitivity analysis; annual probability of collapse.

1. INTRODUCTION

Uncertainty plays a critical role in the analysis and modeling of both human and natural disasters [1-3], and has significant impact on risk-based decisions [4-7]. Building on recent studies on uncertainty analysis [8-10], and specifically in an engineering context [11-13], this paper utilizes research on seismic risk assessment to present a novel modeling framework to enrich and enhance the treatment of uncertainty.

Building collapse is the predominant cause of death and injury in an Earthquake[14-16]. Building collapse occurs when ground motion shakes the building beyond the capacities it is designed and built to withstand. Seismic risk can be defined as a probability of human losses given a likely future earthquake or the probability that the built environment will be damaged given such an earthquake. These probabilities usually represent a level of loss or damage that is equaled or exceeded over some time period; in this study, the seismic risk corresponds to the annual probability of collapse of a given structure.

Taking risk as a function of hazard and vulnerability, in order to calculate seismic risk, the seismic hazard at the building location and the vulnerability of the building must be quantified. The seismic hazard is calculated as the frequency that a ground motion amplitude is greater than a specified value. In this study the primary seismic hazard is ground-motion with a 10% probability of exceedance in 50 years. Assuming a Poisson process for ground motion occurrences, this corresponds to a ground-motion with a 475-year return period. The ground-motion measure used is S_1, spectral acceleration at a 1 second reference period. The effect of soil conditions at individual sites on S_1 is accounted for by using an amplification factor.

The vulnerability of a building is a relationship for estimating the probability of collapse for an individual property, given a certain level of ground motion. The vulnerability of buildings is dependent on their principal structural system (e.g. steel frame or unreinforced masonry) as well as factors influencing the vulnerability of individual buildings (e.g. quality of construction). In this study, the global building stock is classified into 6 primary classes which define expected vulnerability levels of buildings.

In this study, the seismic hazard and vulnerability are brought together to form a seismic risk calculation model which can be used to calculate the seismic risk to an individual building. This procedure is routinely conducted, with varying levels of sophistication, by risk management companies, insurance companies, governments and researchers. Since Cornell's [17] original methodology, a number of other commonly used methodologies have been developed [18, 19]. This paper explores the errors in the results when using these methodologies attributable to the approximations and assumptions in the calculations. The paper classifies these uncertainties as aleatory or epistemic. Uncertainties are classified as epistemic if they are reducible (e.g. through further data gathering or through refining models) and as aleatory if there is no possibility of reducing them. This classification of uncertainties within a model is an important stage: firstly, it helps distinguish which uncertainties may be reduced; secondly, epistemic uncertainties may introduce dependence among random events, which may not be identified if uncertainties are not correctly classified [20].

Uncertainty should be recognized, classified and its impact considered if models are to be validly used to support informed decision-making. Previous studies, such as Ellingwood et al. [21] and Liel et al. [10] discuss the importance of quantifying and communicating the uncertainty in seismic risk assessment, using the example of risk assessment of steel moment frames and collapse risk of RC frames respectively. However, in commonly used seismic hazard assessment tools, such as HAZUS, the uncertainties are rarely quantified either at the input variable level or at the level of attaching an overall uncertainty to a seismic risk assessment. HAZUS's deterministic model is extended to incorporate a subset of the recognized uncertainties in the estimates of seismic hazard and building vulnerability. The input variable uncertainties are propagated through the model using Monte Carlo simulation, resulting in a distribution of estimates for the annual probability of collapse.

Another widely used framework, which supports the propagation of uncertainty in calculating earthquake based performance, is the PEER framework equation (http://peer.berkeley.edu). In this framework, probabilistic functions link earthquake intensity measures with engineering demand parameters. This is a highly technical and comprehensive methodology that requires detailed building construction characteristics (e.g. detailed design information) and behaviours to be collected and as such its utility in situations where limited information is available and uncertainties can only be broadly quantified is limited. In general, there are a lack of empirical studies in this area that provide a straightforward way to characterise and quantify uncertainty at each stage of the risk assessment process. There are many frameworks posed but few that are practical in the way in which uncertainty is incorporated or which allow the separation of uncertainties into the epistemic and aleatoric components. The paper's particular focus is on allowing users of this framework to quantify uncertainties at each stage of the process.

The PEER framework also gave rise to numerous papers which address the problem of uncertainties at different stages in the risk assessment process. For example, Deierlin[22] and Cornell[23] look at the quantification of uncertainties and their sources, whereas Porter and Beck [24] and Hamburger[25] examine the propagation of uncertainties through the process and the impact on the final results for performance based design. Although directly relevant to performance based earthquake engineering and, like the PEER framework, dependent on detailed technical information on building design, they provide useful context to help situate this study's contribution. In the aforementioned papers, calculated performance may be expressed in a variety of metrics, including average annual loss, expected loss at a specified hazard level, or probable maximum loss, as best suits decision makers. The effects of aleatory and epistemic uncertainty are directly accounted for in these performance calculations, however these cannot easily be separated.

The paper introduces a novel seismic risk model framework and through the use of individual case studies investigates uncertainties in seismic hazard and building vulnerability; how the uncertainties in these variables can be quantified; and how these uncertainties can be applied in the risk model. This model framework also allows investigation of the subsequent impact of these uncertainties on estimates for annual probability of collapse, and, through sensitivity analysis, identifies the parameters most responsible for the observed range of outputs.

The paper will present the seismic risk calculation model, using global case studies to consider specific hazards and vulnerabilities. It will then focus on the quantification of uncertainties, before using case studies from six different countries around the world representing a range of economic development, seismic activity, building quality, and availability of data.

2. SEISMIC RISK CALCULATION MODEL

The model developed in this study calculates the seismic risk to a single building, represented by the annual collapse probability, i.e. the percentage of buildings of that particular type that would collapse in any given year. For the building location, the variables in this calculation are:

- The seismic hazard, characterized by the 475 or 2475 year return period spectral acceleration at a time period of 1 second (475yr or 2475yr S_1).
- The soil type or site class, characterized by the shear-wave velocity in the top 30m of the site (V_{s30}).

- The vulnerability of the building, a building class from A to E is used, defined using EMS-98 [26]. Where A is the most vulnerable and E the least vulnerable.

The model takes the location-specific seismic hazard data, modifies it for local soil conditions, extrapolates the single return period value to approximate a seismic hazard curve covering multiple return periods, converts the intensity value to Modified Mercalli Intensity (MMI) as required for the collapse probability calculation, and combines this hazard data with building-specific vulnerability data to estimate the annual probability of collapse; for a given building in a given location, this is described in the following section and in Figure 1. The data sources, parameters and assumptions used in this paper are consistent with those used as part of standard industry practice for calculating seismic hazard assessment. As such, the results of the study, including the uncertainties obtained and the method of calculation, will be relevant to the majority of seismic hazard assessments conducted.

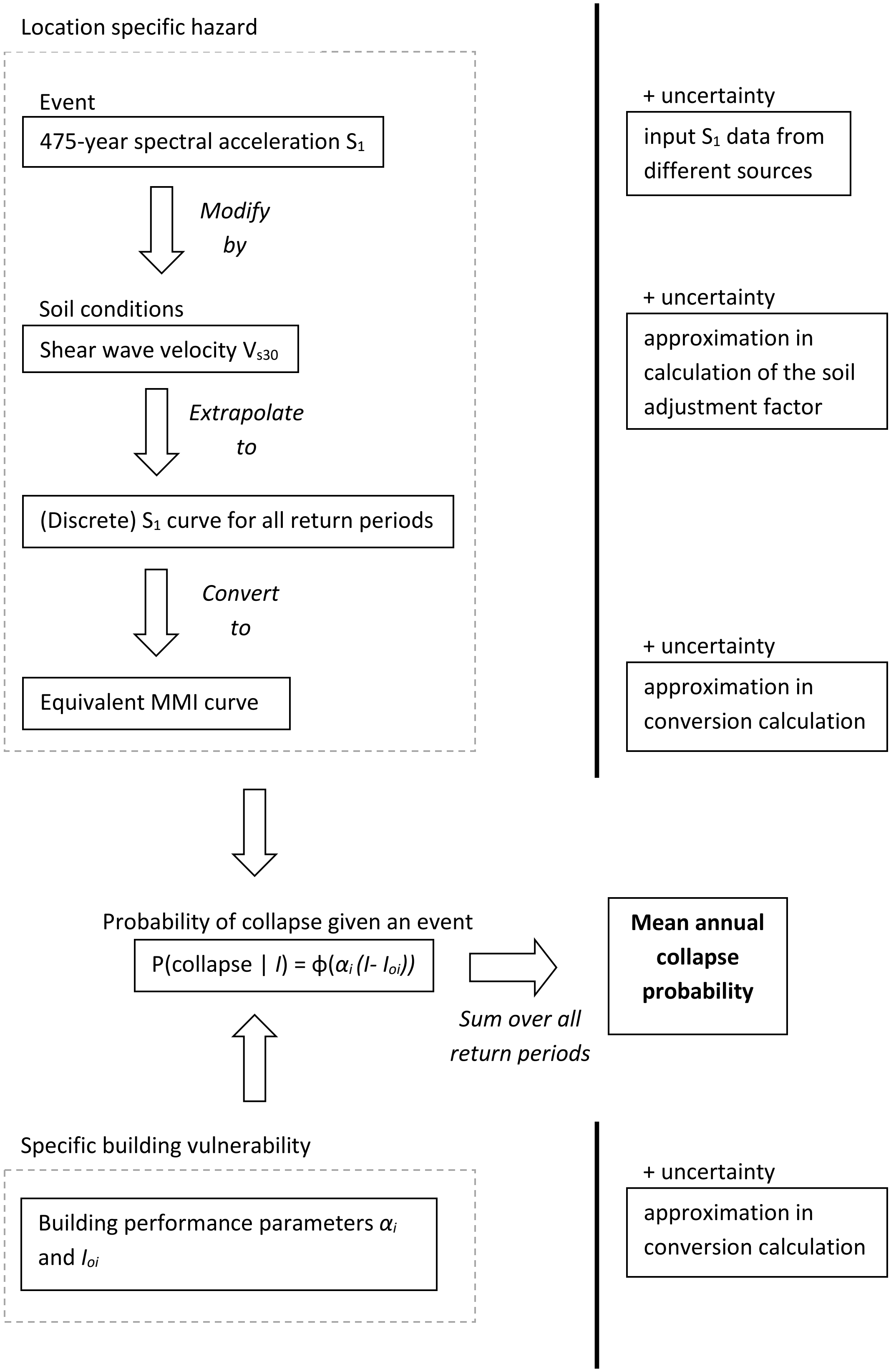
2.1. Site Specific Seismic Hazard

At a building location, the S_1 value on rock (475 or 2475 year return period) can be obtained using sources included in the USGS worldwide seismic "DesignMaps" Web Application [27]. This application provides earthquake ground-shaking parameters worldwide that are needed for seismic design of structures using the International Building Code (IBC) and similar standards, e.g. ASCE/SEI 7 Standard, U.S. Department of Defence Unified Facilities Criteria (DoD UFC). The sources included on the USGS "DesignMaps" website are: GSHAP (Worldwide), 2009/2006 IBC (U.S. & Territories), U.S. DoD UFC 3-310-01 (Worldwide) as well as a number of location specific reports.

In order to provide a meaningful estimate of seismic hazard at an individual property location, it is necessary to gather information on the site conditions, which can be obtained from USGS Global Vs30 maps [28] and existing micro-zoning studies available for some cities. This information is used to obtain ground-motion amplification factors (F) for different soil types, using the equations presented in Choi et al. [29]. These are then applied to the prediction of ground-shaking hazard on bedrock to provide amplified values of S_1 and MMI. F is estimated here using Choi and Stewart [29], who give tables of values for parameters c, V_{ref} and b, n and ε (which are represented in terms of their standard deviations, τ and σ respectively):

$$\ln(F) = c ln\left(\frac{V_{S30}}{V_{ref}}\right) + b ln\left(\frac{S_1/g}{0.1}\right) + n + \varepsilon \qquad (1)$$

Calculating the mean annual frequency of collapse requires the determination of the seismic hazard curve - that is spectral acceleration values over all return periods, rather than just the single 475 year (amplified) S_1 value. This will also allow data sources based on different return periods to be used in the same analysis, if required [30]. A key point here is that the aleatory variability gives the shape of the hazard curve and the epistemic uncertainty results in alternative hazard curves. For risk assessment, the selection of a design ground motion involves selecting a return period and a hazard curve. These two decisions can be interpreted as deciding what level of safety is required (the return period) and how sure one wants to be that this level of safety is being achieved (the hazard curve); the latter is a decision regarding epistemic uncertainty [31].



Location specific hazard
Event
475-year spectral acceleration S_1
Modify by
Soil conditions
Shear wave velocity V_{s30}
Extrapolate to
(Discrete) S_1 curve for all return periods
Convert to
Equivalent MMI curve
+ uncertainty
input S_1 data from different sources
+ uncertainty
approximation in calculation of the soil adjustment factor
+ uncertainty
approximation in conversion calculation
Probability of collapse given an event
P(collapse | I) = $\phi(\alpha_i (I - I_{oi}))$
Sum over all return periods
Mean annual collapse probability
Specific building vulnerability
Building performance parameters α_i and I_{oi}
+ uncertainty
approximation in conversion calculation

Figure 1. Seismic risk calculation model combining seismic hazard and building vulnerability. The deterministic form of the model (left) can be augmented by the introduction of uncertainty in input variables (right). The resulting mean annual collapse probability then evolves from a single point estimate to a distribution of estimates.

In this study, a curve is generated with a shape which gives a high probability of exceedance for low intensity value and low probability of exceedance for high intensity values, and values that appear reasonable when extrapolated from cited 475 year values. The curve use an exponential (quasi Weibull) formulation [32] and is expressed in terms of the mean annual frequency of exceedance of the ground motion S_1; mean annual frequency of exceedance is the reciprocal of the return period:

$$P(exceedance) = \left(e^{-\lambda . {S_1}^{\kappa}}\right) \qquad (2)$$

Here λ is a *scaling* parameter. For a defined value of the *shape* parameter κ, λ can be calculated from a single known S_1 value. The equation is then applied to all return periods. Example S1 and corresponding MMI curves (where k = 0.45) are shown in Figure 2, from Equations 2, 3a and 3b.

A shape parameter of approximately 0.5 is commonly used, corresponding to a multiplicative factor of 2 to convert 475 year to 2475 year return periods[30]. In order to determine a valid shape parameter for this study, simulations have been used to generate a wide variety of input data scenarios, and then considered the impact across these scenarios of varying k; k values between 0.25 and 0.7 were considered in incremental steps of 0.05. An assessment of the ranges of results across multiple scenarios suggest that low k values generate a number of negative or unrealistically low MMI values, as well a unrealistically high S1 values. High k values generate excessively narrow ranges of MMI values, suggesting a limitation on the ability to adequately account for ranges of S1 values. The value of k = 0.45 was selected to generate realistic MMI results and this value approximately equates to a factor of 1.7 between 475 and 2475 year return periods.

In order to align with the building vulnerability annual collapse probability calculation that will be introduced in section 2.2, the final stage in the hazard calculation requires the S_1 hazard curve to be converted into MMI. Here the S_1 values are converted using the Worden et al. [(33)] equations to calculate MMI for the site of interest:

$$\mathrm{MMI} = c_1 + c_2 \log(Y) \ \text{ for } \ \log(Y) \leq t_1 \qquad (3a)$$

$$\mathrm{MMI} = c_3 + c_4 \log(Y) \ \text{ for } \ \log(Y) > t_1 \qquad (3b)$$

Where $Y = S_a = S_1$: $c_1 = 2.5, c_2 = 1.51, c_3 = 0.20, c_4 = 2.90, t_1 = 1.65, \sigma_{\log\,(\mathrm{Sa})} = 0.51$ (the standard deviation associated with the conversion equation).

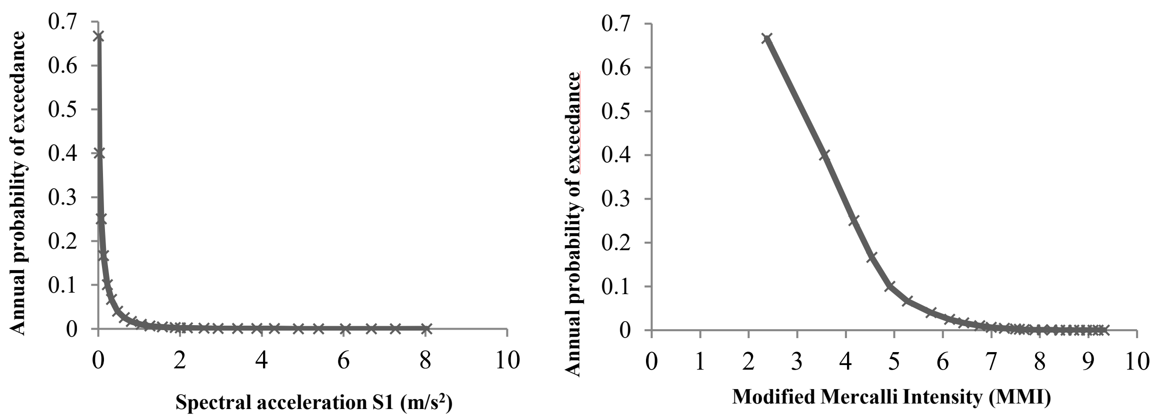


Figure 2: Example seismic hazard curves. Using Equation 2 the single S_1 value is extrapolated over all return periods. The corresponding MMI curve from Equation 3 is shown for completeness.

2.2. Building Vulnerability Assessment

Vulnerability relationships are used to estimate the probability of collapse for an individual property, given a certain level of ground motion. The proposed form of the vulnerability curve as a function of MMI is a cumulative normal distribution. This is equivalent to a lognormal distribution against spectral acceleration values, a form of relationship which is very widely used in damage assessment [19, 18, 34]. For each vulnerability class the general form of the relationship to estimate the collapse probability, given ground motion intensity MMI is:

$$P(\text{collapse}|I) = \Phi\left(\alpha_i(I - I_{oi})\right) \quad (4)$$

Where I is the MMI intensity at the location, Φ is the standard normal cumulative distribution function, and for each primary class a set of parameters α_i, I_{oi}, representing the performance of that building type globally, is determined.

The classification of the global building stock used in this study starts with a classification by principal structural system into 6 primary classes. These are shown in Table I. These types are close to the primary classes defined in EMS-98 [26], and have been found in numerous studies to represent a sensible compromise between the need for distinctions to be made on the basis of known seismic performance, and the need for simplicity, given the limited amount of empirical damage data available. Spence et al. [35] discuss that modifications to these basic classes can also be made based on regional (4 regions) and specific characteristics of the building.

2.2.1. *Determination of vulnerability parameters from empirical data*

The scope of the data used in this analysis is shown in Table I. Using these data, the parameters αi, I_{oi}, were calculated for each class as described in Spence and Foulser-Piggott [35].

Table I. Classification of global building stock, characteristics of damage database, and empirical vulnerability parameters.

Class	Description	Countries	Events	Range of MMI intensity	Total buildings	Vulnerability parameters for mean collapse rate	
						α	I_0 (ICLASS)
A	Weak masonry	8	12	4.7 to 8.1	42683	0.7	9.1
B	Load-bearing masonry, unreinforced	11	16	5.75 to 10	36758	0.7	10.7
C	Structural masonry; pre-code RC frame	13	23	4.5 to 10	79417	0.7	11.4
D1	Moderate code RC frame; RC shear wall	7	9	6 to 9.5	11285	0.7	12.0
D2	Timber frame	3	9	4.9 to 9.0	19110	0.5	12.6
E	Steel frame; high code RC	4	6	6.0 to 9.2	7104	0.5	14.2

For future studies, it may be appropriate to modify the value I_0 to take into account the level of seismicity and code compliance of a region where a building is located as well as characteristics of the building that can result in an increase or decrease in its vulnerability, e.g. height, construction quality and configuration. With these modifiers, the collapse rate for any given building can be determined using Equation 4, where I_{oi} is calculated using Equation 5:

$$I_{oi} = ICLASS + CM + BM1+BM2+BM3 \qquad (5)$$

Where *ICLASS* is the value of I_0 specific to the class of building (Table I), *CM* is the country modifier, *BM1* is the building modifier for height, *BM2* is the building modifier for quality of construction and *BM3* is the building modifier for earthquake resistant configuration. Expert judgement is used in this study in the determination of the building modifiers appropriate to any individual building. The values obtained are derived from previous studies based on the CEQID database [34, 36].

2.2.2. *Calculation of annual collapse probability*

The annual collapse probability calculation follows the methodology proposed by McGuire [18]. If both hazard and vulnerability are expressed as a function of the same ground motion amplitude, x, then annual failure probability is determined by integrating the collapse probability curve over the seismic hazard curve:

$$P_f = \int_0^\infty P_{f|x}(x) . \frac{\mathrm{d}\,\gamma(\mathrm{x})}{\mathrm{dx}} .\mathrm{dx} \tag{6}$$

Here $\gamma(\mathrm{x})$ is the hazard curve and the vulnerability curve $P_{f|x}(x)$ is equivalent to $P(\mathrm{collapse}|I)$ in Equation 4.

Equation 6 is typically solved using numerical integration [37]. Here a range of return periods between 1.5 years and 100,000 years is considered, divided into a discrete number of approximately equal logarithmic intervals. The contribution to annual collapse probability is then taken as the product of the expected annual probability of the hazard being within the limits of the step (that is the mean annual frequency of collapse within each discrete range of return periods) and the collapse probability $P(\text{collapse}|I)$ associated with this hazard value. Total annual collapse probability is the sum of the contribution from all steps.

3. QUANTIFICATION OF UNCERTAINTIES

Uncertainty and the impact it may have can be considered through uncertainty analysis and sensitivity analysis [38, 6, 1, 10, 39]. Uncertainty analysis identifies the level of uncertainty in model outputs. Sensitivity analysis identifies those variables most responsible for the observed range of outputs [40]. Application of such analyses helps to explore the relationship between model inputs and outputs and use unexpected results to expose issues [38]; this should improve model testing and lead to an increased understanding of the model and its limitations in a decision-making context.

The uncertainty analysis presented here is not intended to be exhaustive. Rather, it serves to investigate the impact of a subset of the recognized uncertainties on model outputs, as well as the process of quantifying and applying these uncertainties. Uncertainty is a subset of model input data only—specifically uncertainties in the estimates of seismic hazard and building vulnerability. Uncertainty analysis requires that the nature of the uncertainties in the modeling process be determined and the effects of these uncertainties combined simultaneously. Here uncertainty is incorporated into the model using Monte Carlo simulation, whereby the model is repeatedly run, each time randomly sampling from defined input distributions of values for each of the uncertain input variables. For a specific scenario the outcome of the analysis is multiple estimates of annual probability of collapse, for an individual building in a specific location.

The uncertainty in the prediction of seismic hazard is both epistemic and aleatory, where the former component arises due to lack of knowledge and understanding (e.g. seismic hazard, site conditions and conversion equations) and may eventually be better understood and therefore reduced. The presence of this uncertainty is clear when one considers the numerous different sources of seismic hazard data[41, 42] and the differences in the seismic hazard estimates that these provide. The same applies to local site conditions as there are also numerous estimates of site conditions and amplification factors. The second type of uncertainty is described as aleatory, which is a natural (random) variability which cannot be reduced.

Although the total uncertainty in the seismic hazard assessment is made up of an aleatory and epistemic component, the aim of this study is not to attempt to classify or distinguish between these types of uncertainty. Instead, the aim is to identify the uncertainty in the results of seismic hazard predictions, by generating a distribution of the resulting annual collapse rates. After generating this distribution of estimates for the annual probability of collapse, a global sensitivity analysis will subsequently be applied, to identify which variables contribute to the observed range of outputs.

Four specific sources of uncertainty are considered here: spectral acceleration; site conditions; the conversion of spectral acceleration to MMI; and vulnerability relationships. The first three sources are used in the seismic hazard calculation, and can be subject to uncertainty in many different forms. Spectral acceleration is a measure of ground motion and depends on the accuracy of the estimates of seismic hazard data; uncertainty can originate from differences in data sources, or at the level of individual data sources for the seismic hazard. Site conditions represent the site and specification of soil types; data on local soil conditions may be subject to uncertainty in: the definition of site classes from different sources, for example, the USGS Vs30 map defines the site as rock (site class B, 1500 > Vs30 > 760) or a local site survey defines the site as very dense soil and soft rock (site class C, 760 > Vs30 > 360); a difference in amplification factors from different sources (e.g. Choi and Stewart [29], Stewart et al. [43], and Borcherdt [44]); difference in the uncertainty in amplification factor estimates specified by different authors. The third source is the conversion from S_1 to MMI, where the uncertainty comes from different conversion models, such as Worden et al. [(33)] or Tselentis et al. [(45)], or there may be explicit uncertainty in the conversion process. The final source of uncertainty is in the vulnerability relationships (annual collapse probabilities), where uncertainty exists in the mean collapse ratios.

A single source of uncertainty is considered for each of the first three seismic hazard elements, as detailed below, as well as uncertainty in the building hazard information due to variations in observed collapse rates. The methods of uncertainty propagation are also described.

The first three sources of uncertainty are considered to be independent, i.e. the uncertainty in the seismic hazard estimate, the effect of site conditions, and in the conversion equations are not correlated. This can be considered a reasonable and conservative assumption and these uncertainties are therefore represented separately here. The fourth source of uncertainty, in the calculation of mean collapse ratio, is not assumed independent as the distribution of values is based on the first three input variables.

The assumption of independence of variables has important implications for the calculation of overall uncertainty. If uncertainties are assumed independent but are actually correlated, this can lead to double counting of uncertainties which results in an overestimation of uncertainty. Conversely, where uncertainties are assumed not to be independent but are actually independent, uncertainty can be underestimated resulting in a non-conservative estimation of the overall uncertainty.

3.1. Seismic Hazard Data

There are multiple sources of spectral acceleration data. The framework presented in this paper bases the calculations on the assumption that there are three sources of seismic hazard data to be taken into account in the calculations (e.g. GSHAP, USGS-OFR and UFC). In each Monte Carlo run a single value is sampled from a discrete distribution based on weightings associated with each data source. The weighting indicates relative reliability of the available sources used in the analysis; the weightings sum to 1[46]. As indicated in Musson[46], using expert judged weighting is standard industry practice when using logic trees to deal with different seismic hazard estimates assigned by different experts (which are subjective by their nature). In this study, the weightings are assigned using expert judgement, considering the following factors:

- Vintage of data: modern data have a higher weight.
- Methodology of the study on which the hazard estimate is based: peer-reviewed studies have a higher weight.
- Area over which the study has taken place: studies which involve a more detailed investigation of smaller areas and are as such more detailed, have a higher weight.
- Validation of study results: where multiple studies exist in a certain region, the results of the studies can be compared and any study which can be validated would have a higher weight.

3.2. Site Conditions

A soil amplification factor is applied to the ground-motion S_1 input value in order to represent the influence of local soil conditions on the seismic hazard estimate at that location. Equation 7 is used to calculate the amplification factor F, where ln(F) has standard deviation calculated using Equation 8 [29]:

$$F = c\ln\left(\frac{V_{s30}}{V_{ref}}\right) + b\ln\left(\frac{S_1/g}{0.1}\right) \quad (7)$$

$$\sigma_{total} = \sqrt{\sigma^2 + \mathrm{r}^2} \quad (8)$$

In this study, ln(F) is assumed uncertain and distributed normally here, with mean given by Equation (7) and standard deviation by Equation (8). In each Monte Carlo run a value for ln(F) is sampled and the original S_1 value is then modified by multiplication by F. The calculated adjustment factor is only valid for S_1 in the range 0.02-0.8g, so for values outside this range the adjustment factor has been calculated here on the basis that S_1 is either at the lower or upper end of this range, as appropriate.

3.3. Conversion Equations

For the calculation of the collapse probability the S_1 values must be converted to MMI. (In this study, MMI is defined as a continuous variable [(33)], and the MMI conversion is based on Worden et al. [(33)], Equation 3, where it is assumed that the calculated MMI value is distributed normally with standard deviation 0.8. Therefore, uncertainty is applied to each MMI value on this basis: in each Monte Carlo run a single value is sampled from the standard normal distribution to determine an uncertain adjustment based on the number of standard deviations from the mean, then for each MMI value (across multiple return periods) the MMI value is adjusted by adding the sampled number of standard deviations. MMI is bounded within the range 1 to 12, therefore any calculated values outside this range are truncated to 1 or 12 as appropriate.

3.4. Vulnerability Relationships

A wide variation in collapse rates is observed in the Cambridge Earthquake Impact Database (CEQID)[47]. In particular, a large number of the individual surveys were found to have a zero collapse rate, while a small number had collapse rates very much greater than the mean. This made the use of a standard statistical distribution inappropriate; instead, a probability distribution based on point values was determined using data from CEQID. This probability distribution depends on the mean estimated collapse rate (MCR), and not on the building class. The resulting proposed distribution is shown in Table II. Due to the limited amount of data available in the dataset, these distributions are relevant only to this study. The model allows any appropriate distribution to be used to describe the uncertainty in collapse rates.

Table II. Proposed discrete distribution of MCR values.

Central value in range (* MCR)		0	0.5	1	2	8.5
Proportion of samples in range	MCR<0.02	0.5	0.15	0.1	0.2	0.05
Proportion of samples in range	0.02<MCR<0.1	0.375	0.125	0.175	0.275	0.025
Proportion of samples in range	MCR>0.1	0.250	0.100	0.25	0.350	0

The probability of collapse of a building is determined by the ground motion intensity and the vulnerability of the building, Equation 4. The collapse probability P(collapse|MMI) or Mean Collapse Ratio (MCR), is calculated for each MMI value for each return period. To account for the uncertainty in the MCR a value is sampled from the discrete distribution presented in Table II to determine a factor by which the calculated MCR value is multiplied. There are 5 MCR weighting factors and in each Monte Carlo run a single value is sampled from this discrete distribution based on associated weightings. The weightings vary according to the value of the originally calculated MCR however there is a single MCR value for each return period. Therefore the sampling of the discrete distribution has been anchored to the 475 year return period value only. The sampled MCR multiplication factor is then applied to all originally calculated MCR values to give adjusted MCRs. There is also a truncation in the Mean MCR calculation as the probability of collapse cannot be greater than 100%. Within this procedure, it can be recognized that the choice of MCR value distribution is based on the three uncertain input variables listed previously. This uncertainty cannot therefore be assumed independent.

4. CASE STUDIES: RESULTS OF THE UNCERTAINTY ANALYSIS

The model is used to calculate annual collapse probabilities for buildings in six different parts of the world, each with different hazard, site and vulnerability characteristics for which data was collected in previous projects. The data shown in Table III for V_{s30} and S_1 was obtained using the USGS website[28, 27]; building class data was collected by the author on earthquake reconnaissance field missions and the remaining parameters assigned based on building class.

Code	Vs30	S_1 (m/s2)	Building class	alpha	I0	CM	BM1	BM2	BM3	ITOT
ARMENIA	180 - 360	4.4	B	0.7	10.7	-1.3	0	0	0	9.4
BANGLADESH	<180	3.6	C	0.7	11.4	-1.7	-0.3	1.4	-0.5	10.3
PHILIPPINES	180 - 360	4	D1	0.7	12	-1.7	0	0	0	10.3
UZBEKISTAN	180 - 360	4	B	0.7	10.7	-1.3	0	-0.5	0	8.9
INDONESIA	<180	3.2	D1	0.7	12	-1.7	0	1.4	0	11.7
JAPAN	<180	4.4	E	0.5	14.2	0	0	0	0	14.2

Table III. Building information for case studies

For each case, two sets of results were obtained. The first are 'deterministic' results obtained by running the model using single point estimate baseline values – that is with no uncertainty in any variables. The second are 'stochastic' results obtained using the Microsoft Excel add-on software @RISK. Monte Carlo simulation was undertaken on the six case studies in Table III, performing 10,000 sample runs for each case study, and using Latin hypercube sampling from the defined uncertain input distributions for the four parameters.

Results are given for each case study in Table IV. The deterministic model value for annual collapse probability is given, along with the mean, standard deviation, median, 5% and 95% confidence intervals from the stochastic analysis. Figure 3 shows the cumulative distribution functions (CDF) obtained from the stochastic analysis for the six cases (simulations 1 to 6). The x-axis of the CDF gives the range of possible values of annual collapse probability (ACP) for each case; the 95% confidence line is also indicated. The probability density function (PDF) for Uzbekistan is shown in Figure 4, to give a fuller indication of the shape of the output range of uncertain estimates; the shape is similar for all case studies.

Code	**Deterministic**	**Mean**	**Standard deviation**	**Median**	**5%**	**95%**
ARMENIA	0.528%	1.045%	1.837%	0.319%	0.000%	4.620%
BANGLADESH	0.156%	0.390%	0.822%	0.072%	0.000%	1.825%
PHILIPPINES	0.129%	0.344%	0.755%	0.054%	0.000%	1.630%
UZBEKISTAN	0.855%	1.543%	2.505%	0.540%	0.000%	6.556%
INDONESIA	0.013%	0.056%	0.175%	0.001%	0.000%	0.313%
JAPAN	0.004%	0.013%	0.055%	0.000%	0.000%	0.068%

Table IV. Deterministic and stochastic analysis model results for annual collapse probability for the six case studies.

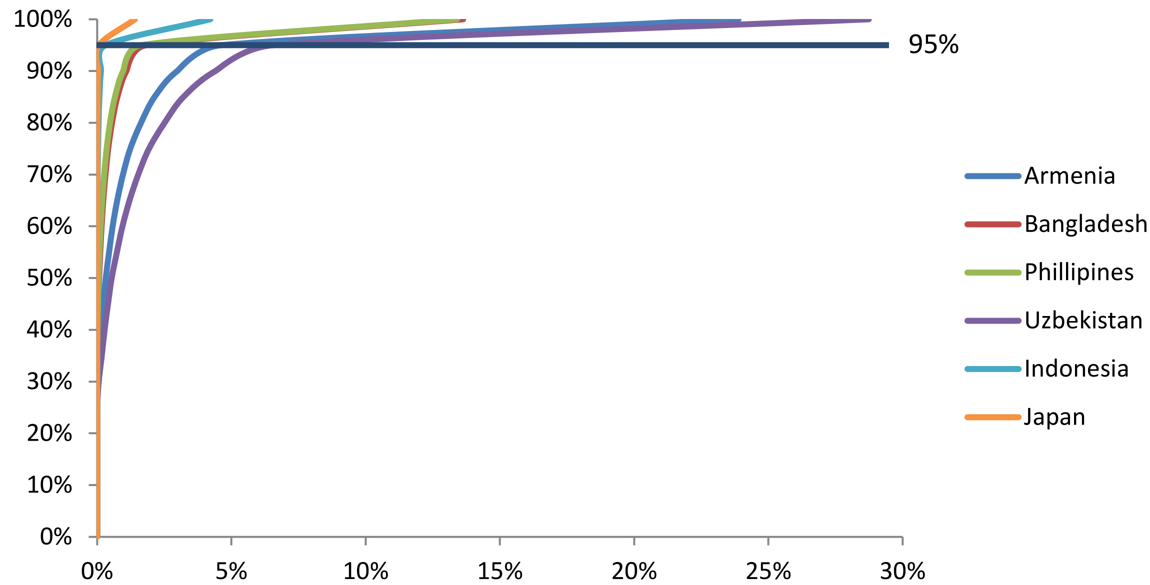


Figure 3. Cumulative distribution function for the annual collapse probability results from the stochastic analysis for the six case study buildings; the 95% confidence line is also shown.

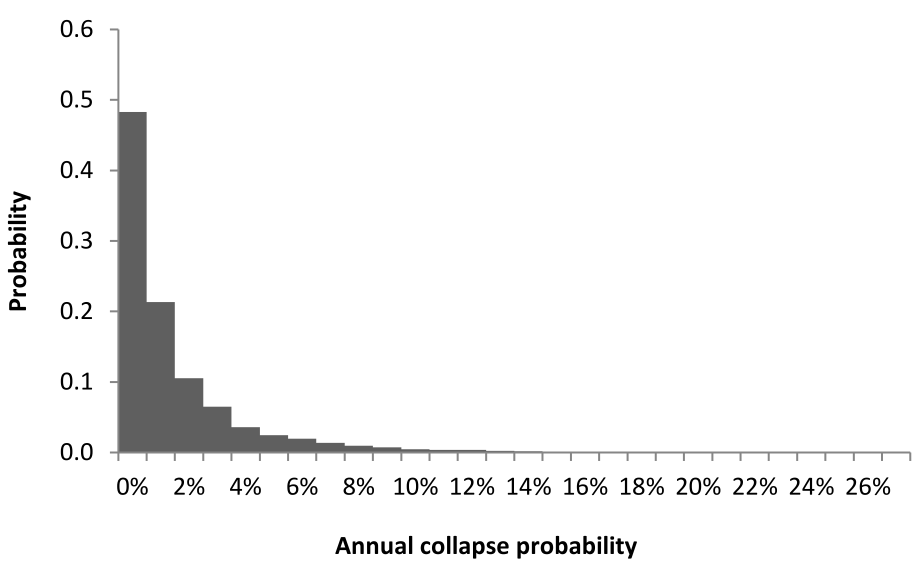


Figure 4. Probability density function for the annual collapse probability results for the stochastic analysis on the Uzbekistan case study building; the shape of the output distribution is consistent for all case studies.

As can be determined from Figures 3 and 4 and Table IV, in all case studies the distribution is left skewed: the majority of values are small values of the annual collapse probability and the median is always substantially lower than the mean. Secondly, although outputs are non-normal and are Exponential-like, Chi-squared tests show that the Exponential distribution is not a good fit to the data for any of the case studies. Thirdly, Japan has the smallest range of annual collapse probabilities with a 90% confidence interval only spanning 0 to 0.07% and a maximum value of 1.7%. This may be due to the building being low vulnerability and located in a region with high quality of construction; the likely range of annual collapse probabilities is small, as building behavior is relatively consistent across well-designed buildings. Finally, Uzbekistan has the largest range of annual collapse probabilities, with a 90% confidence interval from 0 to 6.6% and a maximum value of 28.7%. Conversely to the results for Japan, this may be due to the building being high vulnerability and located in a region with poor quality of construction; the likely range of annual collapse probabilities is very high. The two key factors which give rise to the variation in collapse probability between countries are: level of building vulnerability and quality of construction. This indicates that in regions where buildings have a low level of vulnerability and are well constructed, there is lower uncertainty regarding how that building will behave in an earthquake. This is an intuitive result as we expect to have a higher degree of certainty of building performance in areas with well-constructed, low vulnerability buildings as there is less variation in construction standards leading to unexpected collapses.

The results show that when communicating the results of a calculation of annual collapse probability analysis, the mean and standard deviation can be referenced, as long as the context of the distribution is clear and confidence interval values are referenced. Displaying the CDF of results is helpful to understand the range of results and the difference between mean and median values. It is also clear that it is not adequate to give the deterministic model result as a single point estimate of the annual collapse probability as the 95% confidence value is of the order of ten times the single point estimate. Additionally, giving the results of the stochastic analysis as the mean of the distribution is also not adequate as this may significantly overestimate the true annual probability of collapse of the structure.

The practical implication of these results is that representing a distribution as a median (or mean) value may cause a significant underestimation of risks. For example, if an insurance company is seeking a value to assign as maximum potential exposure to charge a premium over, it would be worthwhile considering a value of the distribution more in line with a 95% confidence level. If there is a large disparity between the 95% confidence level and the median value, the company may consider collecting more data on the particular situation thus potentially reducing the uncertainty and refining the distribution.

5. SENSITIVITY ANALYSIS

Having identified uncertainties in the inputs through Monte Carlo simulation, the aim of the sensitivity analysis is to reveal which variables are the major contributors to the overall uncertainties.

Here a global technique is used, Elementary Effects (EE) [48, 49] with a radial design [40, 50], is used to achieve this. The EE technique is an extension of a one-at-a-time (OAT) approach, starting from multiple baseline points in the input parameter space, varying a single parameter, and measuring the impact on the output mean annual collapse probability due to that single parameter variation. The input uncertainties are taken into account using Latin hypercube sampling. For each of the four uncertain parameters (for each case study) a single EE output indicates the sensitivity of the model to variations in that parameter; here all results are scaled to be equivalent to the impact due to a 1% quantile variation in the input uncertainty. Repeating the analysis multiple times produces multiple EE values for each parameter (for each case study). The average EE value for each parameter indicates the sensitivity of the model to that uncertain input whilst the variation in the sampled EEs indicates the impact of starting from multiple baseline positions and changing the parameter values by varying amounts – that is the impact due to both interactions between the parameters and nonlinearities. For the six case studies, Figures 5 and 6 show the mean μ and standard deviation σ of the resultant Elementary Effects on the mean ACP results respectively, for each input variables and each case study.

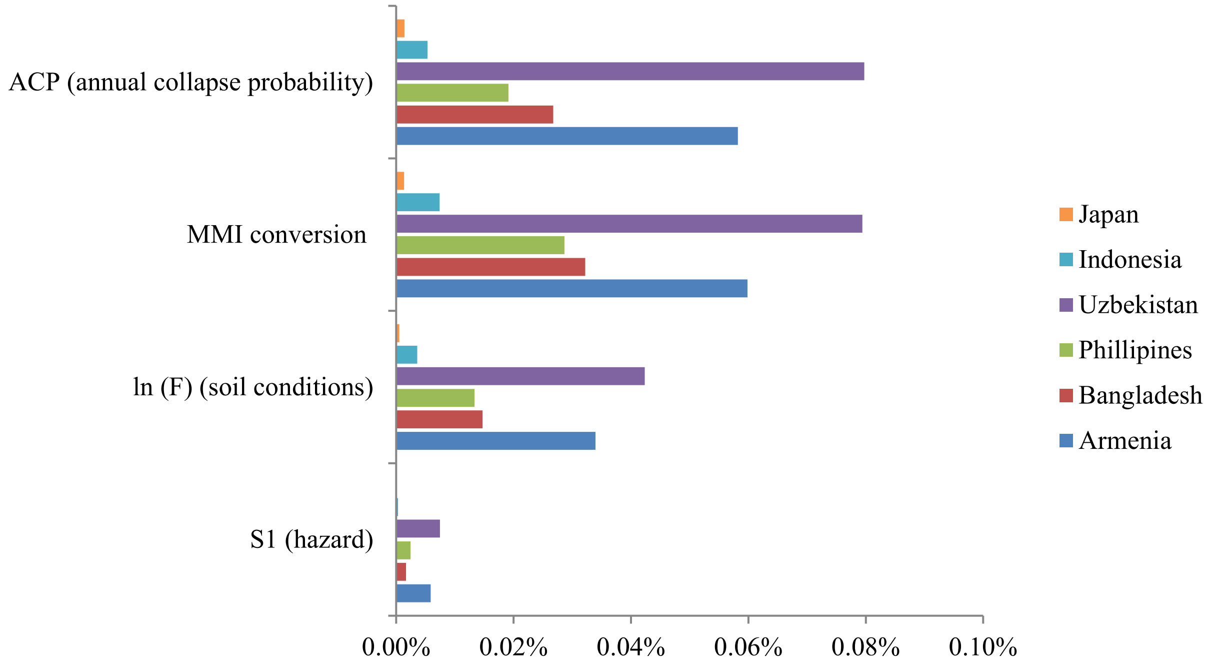


Figure 5. Results of the sensitivity analysis: mean μ of the Elementary Effects on the mean ACP results, for each input variable and each case study; impact due to a 1% quantile variation in the input uncertainty.

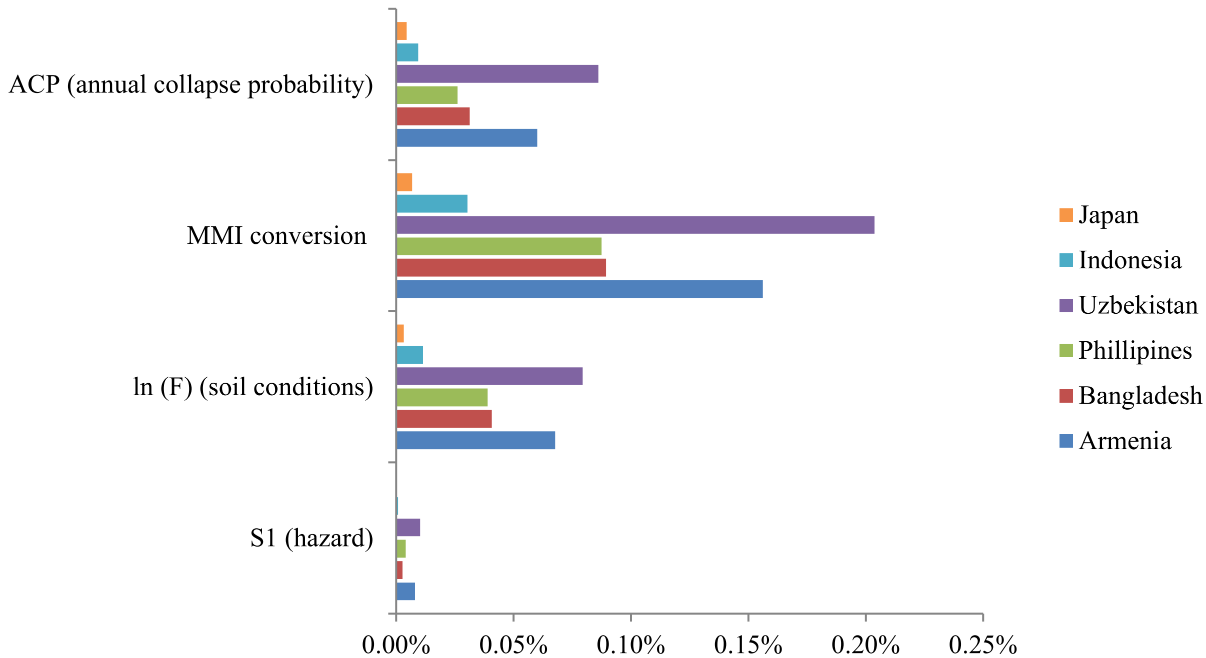


Figure 6. Results of the sensitivity analysis: standard deviation σ of the Elementary Effects on the mean ACP results, for each input variable and each case study; impact due to a 1% quantile variation in the input uncertainty.

The μ results (Figure 5) indicate that ACP is most sensitive to both the MMI conversion and the MCR factors across all of the case studies. However, the σ results (Figure 6) further indicate that there is considerably more variability in the response of the model to variations in the MMI conversion, due to nonlinearities in the response of the model associated with this parameter and/or interactions with the other three uncertain parameters. The analysis also shows that the levels of uncertainty considered for S_1 do not appear to be significant contributors to the uncertainty.

In comparing Uzbekistan and Japan (see Figures 3 and 4), the results suggest that for a building with high vulnerability and a poor standard of construction, the uncertainty resulting from a single input can be considerable greater than the total uncertainty (from all uncertain input variables) for a well-constructed building with low vulnerability. This is further supported by the results of the sensitivity analysis. Figures 5 and 6 show that the vulnerability of the building and the standard of construction are key global sensitivity variables, i.e. these variables determine the range of ACP values for different cases.

There are two main implications of these results. The first is that the result of the uncertainty in the conversion of S_1 to MMI is significant and therefore the development or use of vulnerability relationships in terms of S_1 or other ground-motion parameters is warranted and may act to reduce the overall uncertainty in results. However, 'hidden' uncertainties in the hazard value provided should also be considered as it is possible that conversions have taken place in order to provide a value of S_1. The second is that analysts should focus on gathering accurate information on the building vulnerability characteristics because the standard of construction is a key variable in determining the range of ACP values. This is an important result for earthquake risk assessment, as high value assets or buildings (which are more likely to be insured) if well-constructed have a much better constrained value of annual collapse probability (i.e. a lower uncertainty).

The analysis in this section is based on the assumption that the estimates of uncertainty are reasonable. In the context of this study, the uncertainty estimates used are based on the authors' best judgment of the uncertainty or published values of uncertainty in the input variables and can therefore be considered reasonable. In future studies, however, more information may be available and these estimates may therefore change. Additionally, assumptions made, such as the independence of variables and that variable uncertainty is normally distributed may be challenged.

6. CONCLUSION

This paper has introduced a model framework for estimating the annual probability of collapse for a given building in given locations and quantifying the uncertainty in this estimate. The methodology is based on previous risk assessment works however the way in which uncertainties in input variables are considered, the way their propagation through the model is investigated and their effect on the results demonstrated is novel. This is an important development for seismic risk assessment: understanding the nature of uncertainty within a model can be used to better inform decision makers in terms of data collection (in terms of its needs, prioritization, and limitations) and overall risk assessment.

The communication of the uncertainties in the results has been discussed, particularly in relation to acquisition of more data, the development of future vulnerability relations and the inadequacy of providing single point estimates from a risk analysis. The investigation into the different uncertainties affecting the calculation of the annual probability of collapse for buildings has also considered sources of uncertainty that could affect the single point estimates generated by a deterministic model.

In this paper, four input variable uncertainties have been considered: spectral acceleration, site conditions, MMI conversion and vulnerability relationships. Global sensitivity analysis was conducted to determine the significance of each of the four uncertain variables. The outcome of this analysis is that the uncertainty on the MMI conversion and the MCR factors has the largest effect on the calculation of annual collapse probability for all case studies. For the case of Uzbekistan and Armenia, the uncertainties have the largest effect on the annual collapse probability results; these are also the case studies for which the mean and range of the stochastic analyses are largest. The results of the global analysis are significantly more variable for the MMI conversion parameter.

As quantifying and communicating uncertainty is to some extent subjective and the aims of the risk analysis will likely vary from study to study, a number of issues may be explored in future studies. The consideration of uncertainty in different variables and assumptions about independence of these variables will be of benefit, particularly examining parameters within the building vulnerability calculation such as height, structural system etc., the magnitude of uncertainty in site class (and possible inclusion of this uncertainty by using a three-point estimate of V_{s30} as inputs to the model), and inclusion of uncertainty in the shape parameter. Further investigation into improved quantification of uncertainty in input variables could be developed to help, for example, refine uncertainty in the collapse probability be refined. Finally, this study, and all studies into seismic risk assessment would benefit from research into whether uncertainties are truly epistemic or aleatory and whether altering characterization of uncertainties affects the results, for example whether there is epistemic uncertainty in the shape of the hazard curve which is assumed to be aleatory.

The model framework developed in this paper can be used in its current form when conducting risk analyses for buildings or portfolios of buildings to better understand the likely variation in values of annual collapse probability or risk of collapse. This can help ensure that risks are not over or under estimated and the value of acquiring good quality, accurate data is appreciated. This extra dimension to seismic risk assessment (and all risk assessment) can help inform risk-based decisions and ensure decision-makers are more aware of the potential and limits of risk assessments.

REFERENCES

1. Linkov I, Burmistrov D. Model uncertainty and choices made by modelers: Lessons learned from the international atomic energy agency model intercomparisons. Risk Analysis, 2003; 23 (6):1297-308.

2. Borgonovo E. Measuring uncertainty importance: Investigation and comparison of alternative approaches. Risk Analysis, 2006; 26 (5):1349-61.

3. Goda K, Ren J. Assessment of seismic loss dependence using copula. Risk Analysis, 2010; 30 (7):1076-91.

4. Bier VM, Lin S-W. On the treatment of uncertainty and variability in making decisions about risk. Risk Analysis, 2013; 33 (10):1899-907.

5. Tesfamariam S, Sadiq R, Najjaran H. Decision making under uncertainty—an example for seismic risk management. Risk Analysis, 2010; 30 (1):78-94.

6. Paté-Cornell ME. Risk and uncertainty analysis in government safety decisions. Risk Analysis, 2002; 22.

7. Thompson KM. Variability and uncertainty meet risk management and risk communication. Risk Analysis, 2002; 22 (3):647-54.

8. Arunraj NS, Mandal S, Maiti J. Modeling uncertainty in risk assessment: An integrated approach with fuzzy set theory and monte carlo simulation. Accident Analysis & Prevention, 2013; 55:242-55.

9. Khanzode VV, Maiti J, Ray PK. A methodology for evaluation and monitoring of recurring hazards in underground coal mining. Safety Science, 2011; 49 (8–9):1172-9.

10. Liel AB, Haselton CB, Deierlein GG et al. Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings. Structural Safety, 2009; 31 (2):197-211.

11. Karimi I, Hüllermeier E, Meskouris K. A fuzzy-probabilistic earthquake risk assessment system. Soft Computing, 2006; 11 (3):229-38.

12. Wang Y-M, Elhag TMS. A fuzzy group decision making approach for bridge risk assessment. Computers & Industrial Engineering, 2007; 53 (1):137-48.

13. Chauhan SS, Bowles DS. Dam safety risk assessment with uncertainty analysis.; 2003. (Proceedings of the australian committee on large dams risk workshop).

14. Coburn AW, Spence R, Pomonis A. Factors determining human casualty level in earthquakes: Mortality prediction in building collapse. In Tenth World Conference on Earthquake Engineering. Spain; 1992. 5989-94 p.

15. Spence R, So E, Scawthorn C. Human casualties in earthquakes: Progress in modelling and mitigation: Springer Netherlands; 2011.

16. Yamazaki F, Nishimura A, Ueno Y. Estimation of human casualties due to urban earthquakes. In Eleventh World Conference on Earthquake Engineering. Mexico; 1996.

17. Cornell CA. Engineering seismic risk analysis. Bulletin of the Seismological Society of America, 1968; 58 (5):1583-606.

18. McGuire RK. Seismic hazard and risk analysis. Oakland, CA: Earthquake Engineering Research Institure; 2004.

19. HAZUS. Earthquake loss estimation methodology. Washington, DC: National Institute of Building for the Federal Emergency Managaement Agency.

20. Kiureghian AD, Ditlevsen O. Aleatory or epistemic? Does it matter? Structural Safety, 2009; 31 (2):105-12.

21. Ellingwood BR, Kinali K. Quantifying and communicating uncertainty in seismic risk assessment. Structural Safety, 2009; 31 (2):179-87.

22. Deierlein GG. Overview of a comprehensive framework for earthquake performance assessment. In Performance-based seismic design concepts and implementation. Bled, Slovenia; 2004.

23. Cornell CA. Hazard, ground motions, and probabilistic assessments for pbsd. In Performance-based seismic design concepts and implementation. Bled, Slovenia; 2004.

24. Porter KA, Beck JL. Simplified pbee to estimate economic seismic risk for buildings. In Performance-based seismic design concepts and implementation. Bled, Slovenia; 2004.

25. Hamburger RO. Development of next-generation performance-based seismic design guidelines. In Performance-based seismic design concepts and implementation. Bled, Slovenia; 2004.

26. Grunthal G, editor European macroseismic scale 1998 Luxembourg: Conseil de l'Europe; 1998.

27. USGS. Worldwide designmaps app documentation [online] 2012.

28. USGS. Global vs30 map server [online] 2012.

29. Choi Y, Stewart JP. Nonlinear site amplification as function of 30 m shear wave velocity. Earthquake Spectra, 2005; 21 (1):1-30.

30. Lubkowski ZA, Aluisi B. Deriving ss and s1 parameters from pga maps. In 15th World Conference in Earthquake Engineering Lisbon; 2012.

31. Abrahamson NA, Bommer JJ. Probability and uncertainty in seismic hazard analysis. Earthquake Spectra, 2005; 21 (2):603-7.

32. Booth E. Assessment of seismic hazard. In: The Arup Journal. Series Assessment of seismic hazard. London, UK: Ove Arup & Partners; 1984.

33. Worden CB, Gerstenberger MC, Rhoades DA et al. Probabilistic relationships between ground‐motion parameters and modified mercalli intensity in california. Bulletin of the Seismological Society of America, 2012; 102 (1):204-21.

34. Spence R, So E, Jenny S et al. The global earthquake vulnerability estimation system (geves): An approach for earthquake risk assessment for insurance applications. Bulletin of Earthquake Engineering, 2008; 6 (3):463-83.

35. Spence R, Foulser-Piggott R. The use of macroseismic intensity as a basis for empirical vulnerability assessment. In: VEESD. Series The use of macroseismic intensity as a basis for empirical vulnerability assessment. Vienna, Austria; 2013.

36. Booth E, Spence R, Curtin D et al. Seismic risk management of an internatinal portfolio of buildings. In: 14th World Conference in Earthquake Engineering Conference Proceedings (14WCEE). Series Seismic risk management of an internatinal portfolio of buildings. Beijing, China; 2008.

37. Eads L, Miranda E, Krawinkler H et al. An efficient method for estimating the collapse risk of structures in seismic regions. Risk Analysis, 2013; 42 (1):25-41.

38. Saltelli A, Tarantola S, Campolongo F et al. Sensitivity analysis in practice. A guide to assessing scientific models. London: John Wiley and Sons; 2004.

39. Paté-Cornell ME. Uncertainties in risk analysis: Six levels of treatment. Reliability Engineering and System Safety, 1996; 54:95-111.

40. Saltelli A, Annoni P. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software, 2010; 25 (12):1508-17.

41. GSHAP. Global seismic hazard assessment program [online] 1999.

42. SHARE. Seismic hazard harmonization in europe [online] 2012.

43. Stewart JP, Liu AH, Choi Y. Amplification factors for spectral acceleration in tectonically active regions. Bulletin of the Seismological Society of America, 2003; 93 (1):332-52.

44. Borcherdt RD. Empirical evidence for site coefficients in building code provisions. Earthquake Spectra, 2002; 18 (2):189-217.

45. Tselentis G-A, Danciu L. Empirical relationships between modified mercalli intensity and engineering ground-motion parameters in greece. Bulletin of the Seismological Society of America, 2008; 98 (4):1863-75.

46. Musson R. On the nature of logic trees in probabilistic seismic hazard assessment. Earthquake Spectra, 2012; 28 (3):1291-6.

47. CEQID. Cambridge earthquak impact database 2016.

48. Campolongo F, Cariboni J. Sensitivity analysis: How to detect important factors in large models. Luxembourg: European Commission: Directorate General Joint Research Centre EUR 22798 EN.

49. Campolongo F, Cariboni J, Saltelli A. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software, 2007; 22:1509-18.

50. Saltelli A, Annoni P, Azzini I et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications, 2010; 181 (2):259-70.