QUANTIFYING THE EFFECTS OF EPISTEMIC UNCERTAINTY ON THE SEISMIC RISK OF A BUILDING INVENTORY

P. Kalakonas [(1)], V. Silva [(2)], A. Mouyiannou [(3)], A. Rao [(4)]

[(1)] *PhD student at IUSS - School for Advanced Studies, petros.kalakonas@iusspavia.it*
[(2)] *Risk Coordinator at Global Earthquake Model Foundation, vitor.silva@globalquakemodel.org*
[(3)] *Researcher at Partner Re, amaryllis.mouyiannou@partnerre.com*
[(4)] *Senior Risk Engineer at Global Earthquake Model Foundation, anirudh.rao@globalquakemodel.org*

Abstract

Probabilistic seismic risk models can have a wide range of purposes. National seismic risk assessments can support risk management, mitigation and reduction measures, which may include the creation of post-disaster emergency plans, informing decision-makers in cost-benefit analyses, and enforcement of building codes in regions in high risk. Furthermore, earthquake catastrophe models can be used across multiple segments of the (re)insurance industry to derive adequate pricing for (re)insurance contracts and transferring risk. Such models are affected by a large variability due to the uncertainties associated with each component. These sources of uncertainty arise from the input of various parameters that define the seismicity of a region, the ground shaking intensity, the seismic vulnerability and exposure characteristics of a building stock. The uncertainty related to the lack or incomplete scientific knowledge of a process is termed epistemic, and therefore it is considered to be reducible. The present study aims to quantify the effects of such sources of uncertainties on the estimated losses of a spatially distributed building inventory. As a case study, the residential building stock of Guatemala was selected, as this region is characterized by relatively high seismic risk. Various alternative assumptions in each component (hazard, exposure, vulnerability) were used, and their impact was assessed through sensitivity analyses using the OpenQuake-engine.

Keywords: risk modelling; uncertainty; loss assessment; sensitivity analysis; seismic hazard; OpenQuake.

1. Introduction

The end use of earthquake loss estimation models is multifold, including risk reduction and mitigation planning, emergency management, and risk management in the (re)insurance industry. There are various uncertainties related to the input parameters in each component of such models and therefore, decisions taken based on their results are affected. As demonstrated by [1], the results of a loss estimation model that does not properly incorporate all the sources of uncertainty can be misleading and lead to an underestimation of the actual risk. The seismic risk estimates of building portfolios carry a high degree of variability, as different assumptions during the modelling process can lead to relatively different results (e.g. [2],[3]). Nevertheless, modern catastrophe loss models tend to oversimplify risk estimates, leading to an insufficient treatment and quantification of the underlying uncertainties (e.g. [4]).

In general, uncertainties in risk modelling can be classified into aleatory and epistemic, according to their nature and physical basis. Epistemic uncertainty in earthquake risk assessment refers to the lack or insufficient knowledge of the various phenomena regarding the earthquake generation, propagation of seismic waves and the response of the built environment under ground shaking. The present study intends to quantify the relative effect of epistemic uncertainties arising from the different components of a seismic risk model. To this end, the OpenQuake-engine ([5],[6]) was used to perform probabilistic seismic risk analyses, due to its versatile capabilities and open-source nature.

2. Case Study

Guatemala is located in one of the most earthquake prone regions in the world, which has witnessed several destructive earthquakes in the past decades (e.g. [7],[8]). Some examples include the 1972 Managua (M_w 6.3) and 1976 Guatemala (M_w 7.5) earthquakes, which caused more than 10,000 and 20,000 fatalities, respectively (e.g.[9]). Guatemala City receives special attention since it is located in an area affected by a high level of seismic hazard (e.g. [10]) due to its vicinity to the subduction zone and crustal active faults (e.g. Motagua fault). Moreover, the fact that the first seismic design provisions in Guatemala were enforced in 1996 (which reflects the vulnerable characteristics of the building inventory) contributes to the significant level of seismic risk (e.g.[11]).

In the present study, as exposure model, the residential building inventory of Guatemala was adopted from[12]. In that study, an exposure model for each country in Central America was developed based on the national census databases, World Housing Encyclopedia reports, and local expert judgement. It is worth mentioning that in the case of Guatemala, the exposure model was developed according to the information from the national census survey conducted in 2002, because the results from the 2018 census were not available at the time this study was conducted. The authors used the GEM Building Taxonomy ([13]) to classify the building stock uniformly and associate each building class to a representative vulnerability function. In Table 1, the distribution of construction material of the building inventory at the national level and in the country capital (Guatemala City) are presented. It should be noted that approximately 12-13% of the total building stock is located in Guatemala City. The aforementioned census data for the country are publicly available at the municipality level, and thus the exposure model was created at this spatial resolution. All the buildings within each municipality are aggregated at its geometric centroid.

Table 1- Construction material distribution of the building inventory in Guatemala: at national level and in Guatemala City.

Material of Construction	National Level (%)	Guatemala City (%)
Wood	25.0	7.0
Adobe	26.0	5.0
Reinforced Concrete	1.0	3.0
Confined Masonry	19.5	41.0
Reinforced Masonry	7.0	8.0
Unreinforced Masonry	19.5	31.0
Other	2.0	5.0

Subsequently, the probabilistic seismic hazard assessment (PSHA) model was adopted from [9, 14]. By integrating tectonic and geological data, the researchers distinguished three types of seismogenic zones for Central America, related to focal depths: crustal seismicity ($h < 25$ km), subduction interface (25 km $< h <$ 60 km) and subduction intra-slab ($h > 60$ km). Moreover, they developed two seismic source models (seismic zonations) for the region with different degrees of detail. The *Regional* zonation, which distinguishes large scale area sources (seismogenic zones) that capture the primary seismic tectonic features and seismicity of the region, and the *National* zonation which is an increase of scale differentiating zones within each country and avoiding discontinuities at the national borders. As for the local intensity estimation, the hazard model uses the following ground motion prediction equations (GMPEs) for each of the tectonic regimes: Crustal (shallow) seismicity: [15] and [16]; subduction interface [17], and subduction intra-slab: [17] and [16]. There are four possible realizations and the researchers assigned equal weights (0.25) to all of them.

The vulnerability component was adopted from [18] due to its compatibility with the GEM taxonomy and building classes used by [12]. The seismic vulnerability model followed an analytical methodology which incorporates building-to-building and record-to-record variability through the generation of a large set of single-degree-of-freedom systems and the use of numerous ground motion records (GMRs), respectively. In particular, 150 capacity curves were generated for each building class using a Monte Carlo simulation and 300 GMRs were selected from subduction and active shallow earthquakes.

3. Considered Epistemic Uncertainties

As mentioned in the previous section, the definition of the building classes of the residential building inventory was based on an integration of information from the housing census database of Guatemala and local expert judgment. Therefore, the grouping of the building stock into a number of distinct building classes is affected by epistemic uncertainty, as it relies on the availability and detail of exposure data. The residential exposure model component was derived from [12] in collaboration with local experts and supported by online surveys regarding the construction practices and materials in the Central American countries. The results for Guatemala suggested two alternative versions of the exposure model in terms of construction material distribution at the national level, as shown in Table 2. These distributions were used to derive two alternative exposure models.

Table 2- Alternative versions of the exposure model in terms of the distribution of the main type of construction material in Guatemala.

Material of Construction	Version 1	Version 2
Wood	25%	25%
Adobe	26%	20%
Reinforced Concrete	5%	1%
Confined Masonry	16%	25%
Reinforced Masonry	16%	2%
Unreinforced Masonry	10%	25%
Other	2%	5%
Total	100%	100%

Inevitably, the spatial resolution of an exposure model is bound by the information regarding the location of the assets. For example, the adopted model was developed at the municipality level, as the available data were aggregated at such level. However, this is an important source of epistemic uncertainty that might affect the risk estimates. As [19] illustrated, this aggregation and relocation of the buildings leads to a misrepresentation of the distance between assets and the seismic sources, and the implicit correlation in the ground motion for all buildings at a given location. In the present study, the exposure model was spatially disaggregated into evenly spaced grids to assess the impact of such source of uncertainty. The methodology used by [20] was followed for this purpose, in which night-time lights are used to re-allocate the assets within each administrative region. The nighttime lights of Guatemala (extracted from [21]) are shown in Fig. 1, where Guatemala City is located in the area with the highest concentration of light. The exposure model was disaggregated to five different resolutions spanning from 480 to 30 arc-seconds. Indicatively, the original exposure model consists of 334 exposure locations (i.e. centroids of municipalities), while the 480 and 30 arc-second resolution models resulted into 314 and 30,842 locations, respectively.

Fig. 1- Night-time lights of Guatemala.

It is quite evident that the majority of the decisions based on the available data during the development of a PSHA model are also subjected to epistemic uncertainties. For example, these may involve the definition of the tectonic regimes, the development of an earthquake catalogue, the identification of crustal faults, the estimation of rate of occurrence of certain seismic events and the estimation of the maximum magnitude (M_w) of each source. Different assumptions or additional information during such tasks might influence the shape and parameters of the derived seismic sources. Taking these into account, the national and regional zonation were considered as two alternative source models. The effect of maximum magnitude of the seismic sources was also assessed, where the theoretical maximum and minimum M_{max} for each area source of the national zonation from [9] were used. Furthermore, the selection of the GMPEs representative for the seismicity of the respective region may have a significant impact in the loss estimates of building portfolios (e.g. [2]). Thus, the impact of the selected GMPEs in the logic tree was also explored.

In order to have an insight in the differences on the seismic hazard estimates between the two zonations, prior to the loss estimation, seismic hazard curves were computed using the OpenQuake-engine. A site model for the country was derived using Vs_{30} proxy values obtained from the USGS Vs_{30} server, following the slope topography methodology developed by [22]. The hazard maps in terms of peak ground acceleration (PGA), for the 9.5% probability of exceedance (PoE) in 50 years (i.e. 500-year return period) are presented in Fig. 2.

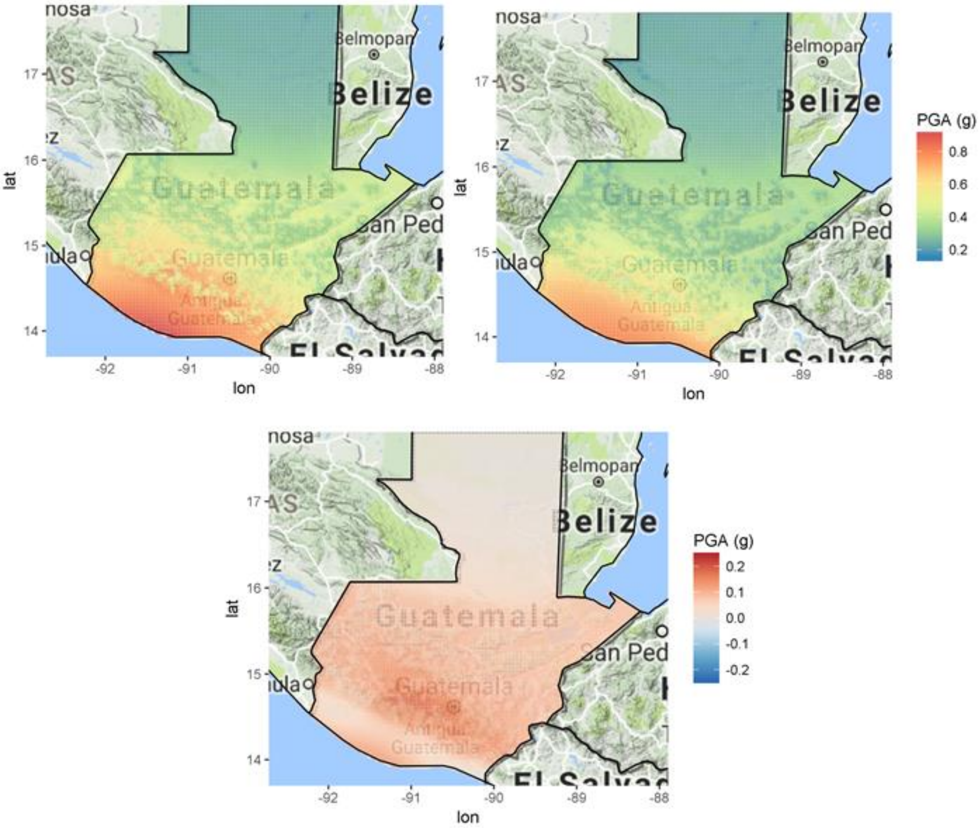

Fig. 2- *Top*: Seismic hazard maps of Guatemala for the 9.5% PoE in 50 years on soil conditions, using National (*left*) and Regional (*right*) zonation. *Bottom*: Difference between the two zonations (*National – Regional*).

In general, vulnerability functions for a given building class provide the expected loss ratio (LR) cost conditioned on a set of intensity measure levels (IML). Such functions can be derived analytically (e.g.[23]), empirically (e.g.[24]), and by employing expert judgement (e.g.[25]). Nevertheless, regardless of the derivation methodology, vulnerability functions are also typically characterized by large uncertainty (e.g.[26]), principally due to four main sources (e.g.[27]): 1) record-to-record variability, 2) building-to-building variability, 3) uncertainty in the damage criteria, and 4) uncertainty in the damage to loss model. Moreover, the definition of the loss ratios per intensity measure can follow a deterministic (i.e. mean LR value) or probabilistic approach. In this study, the uncertainty around the mean loss ratios arising from the propagation of the building-to-building and record-to-record variability was modelled using beta and lognormal distributions. For this purpose, the methodology proposed by [27] was applied at the adopted vulnerability functions, where a standard deviation (σ_{LR}) is calculated based on the expected LR and conditioned on the IML. Additionally, the effect of the consideration of correlation in the vulnerability of buildings among the same class was also assessed. However, due to limitations in the loss modelling approach, correlation was only considered for the case of lognormal distribution.

4. Results and Discussion

The effects of the selected uncertainties on the risk estimates of Guatemala and Guatemala City are presented in this section. To this end, the commonly applied technique of sensitivity analysis was followed (e.g.[2],[28]), and a base risk model was defined using the default modelling option for each component of the adopted models. Hence, the base country and city models were defined considering the construction material distribution and spatial aggregation (municipality level) developed by [12], the national seismic zonation, the expected M_{max}, the mean seismic hazard among the four realizations, and the vulnerability functions derived by [18]. Subsequently, event-based risk analyses were carried out in the OpenQuake-engine using a stochastic event set with a length of 250,000 years. The results are presented in terms of exceedance probability curves, average annualized losses (AAL), and histograms of annual losses. The loss metrics are expressed as loss ratio (i.e. loss/replacement cost) instead of monetary values. The loss exceedance curves considering the different hazard and exposure parameters are shown in Fig. 3 and Fig. 4, respectively, and the tornado (sensitivity) plot for AAL is presented in Fig. 5. In the latter figure, the variation from the base model is shown by altering one parameter at a time, while the base value is indicated by the vertical black line.

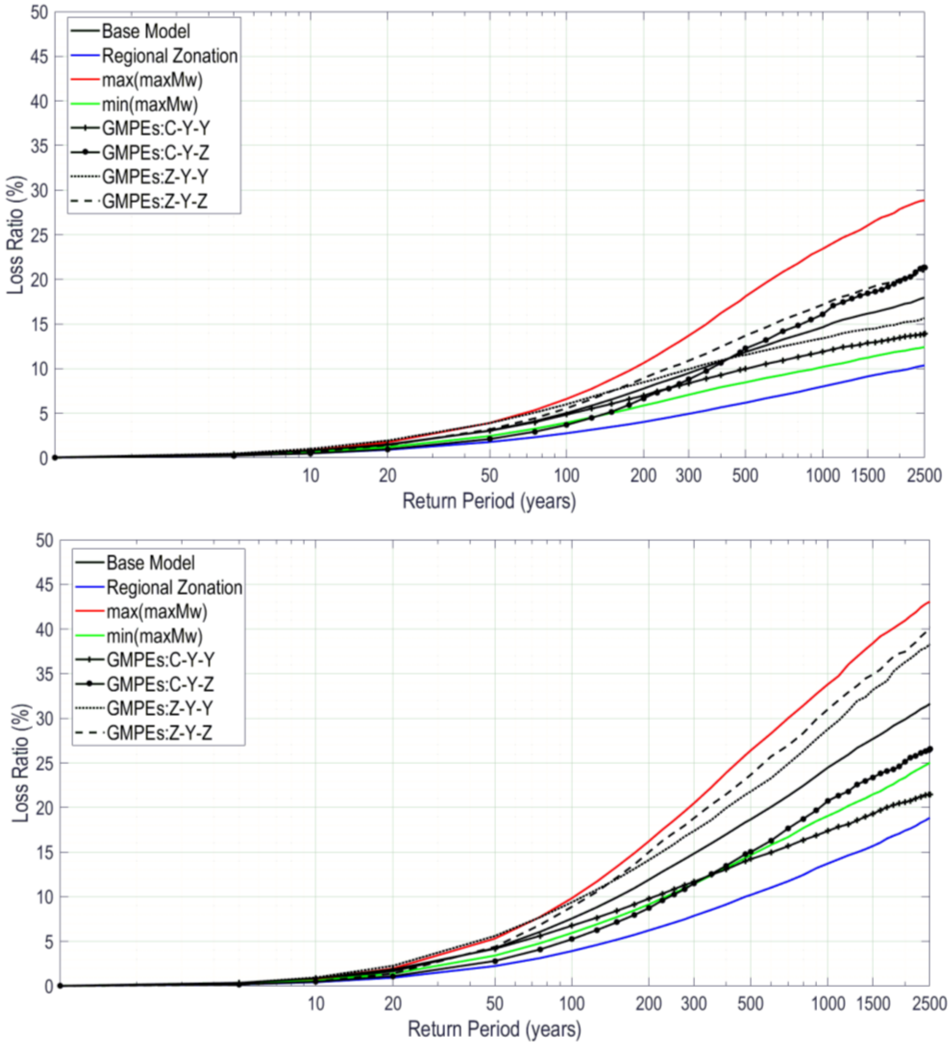

Fig. 3- Guatemala (*top*) and Guatemala City (*bottom*) loss exceedance curves derived from varying the hazard parameters.

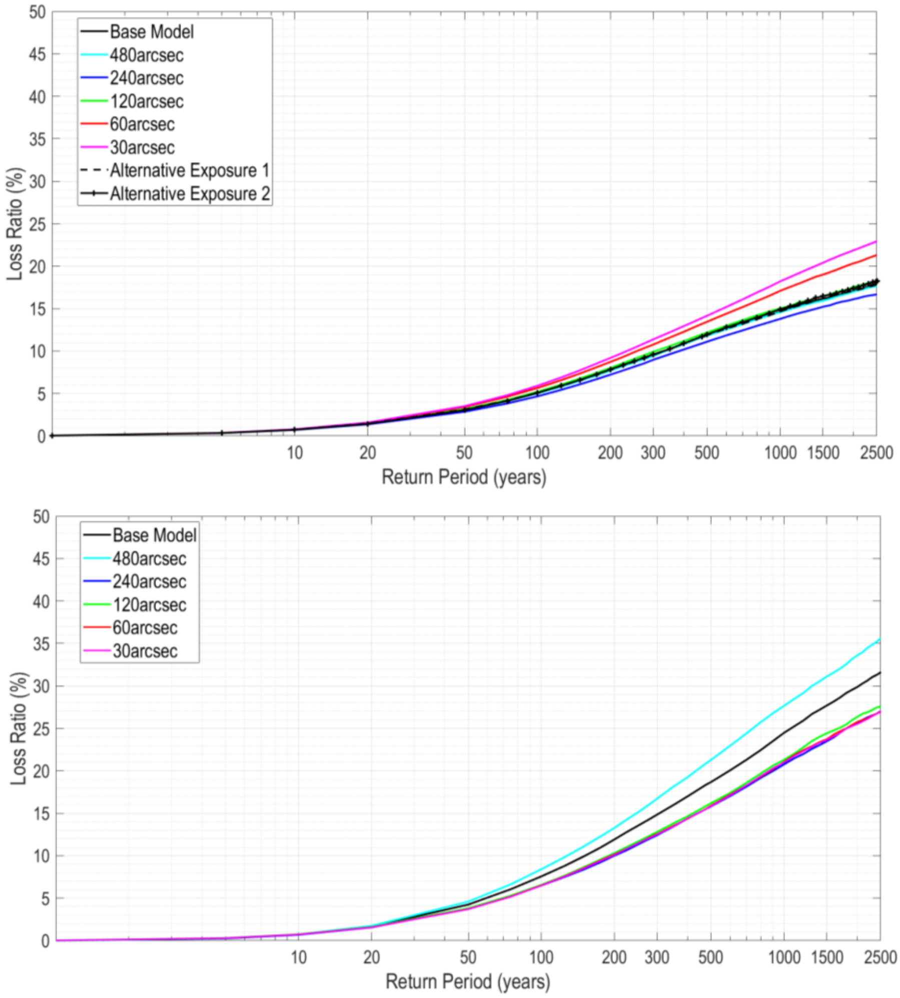

Fig. 4- Guatemala (*top*) and Guatemala City (*bottom*) loss exceedance curves derived from varying the exposure parameters.

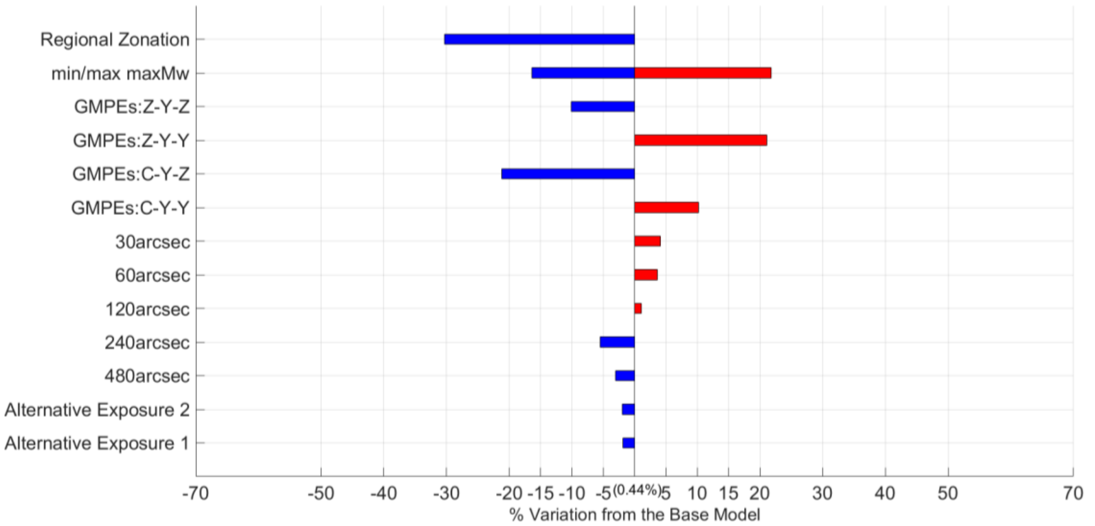

Fig. 5- Tornado diagram for the country's AAL estimate.

Amongst all the considered epistemic uncertainties, the regional zonation model led to the lowest predicted losses and AAL. Its impact appears to be similar and constant over the entire range of return periods on both Guatemala and Guatemala City. On the other hand, the maximum M_{max} parameter resulted in the highest losses, especially for high return periods where the losses are approximately 60% higher than the base model's (expected M_{max}) prediction. This trend is linked to the appearance of large rare events from the subduction zone for such return periods. The impact of the maximum M_{max} on the city's losses is also considerable, although it is not magnified at long return periods due to the smaller contribution of the large events from the subduction zone compared to the local shallow crustal events ([9]). On the contrary, the effect of the minimum M_{max} is constant throughout the return periods, leading to around 20% lower values.

The result at the national level from the individual branches of the GMPEs logic tree suggest that the contribution from the shallow crustal seismicity in the loss estimates at short return periods (i.e. <300 years) is greater than the subduction zone, and vice versa. Particularly, the selection of the GMPE for the associated tectonic type governs the predicted losses from each branch at the respective return period range. Furthermore, the sensitivity of the AAL indicates that the selection of the GMPE for the subduction intra-slab events dominates the results in all cases. The latter finding was expected due to the fact that an earthquake from the subduction zone is more likely to affect a wider region of the country. Nonetheless, for the case of Guatemala City, the losses for all return periods are strongly influenced by the shallow crustal seismicity, as the selection of the GMPEs for this tectonic regime governs the variability of the predicted losses from the individual branches.

The effect of the spatial resolution of the country's exposure model indicates that going towards finer resolutions (i.e. 60 and 30 arc seconds) results in higher losses and AAL. Such results also suggest that the aggregation of the building stock at the municipality level or generally at cruder spatial resolutions causes a shift of the assets to regions of lower seismic hazard (i.e. increases the site-to-source distances). It is important to note that these findings should not be generalized to other cases or regions, because as demonstrated by [29], the impact of spatial resolution of an exposure model on the loss metrics is strongly dependent on the spatial distribution of the seismic hazard and soil conditions. As for Guatemala City, it is evident that the models with finer spatial resolution than 480 arc-seconds present essentially the same losses in all return periods. A law of diminishing returns is observed in going to very fine resolutions, as sufficient loss convergence can be achieved at lower resolutions, a trend which was also observed and documented by [30]. In that study, it is also concluded to adopting very high spatial resolution models might not improve significantly the accuracy of the risk estimates to justify the additional computational effort. Finally, negligible differences in the loss estimates of the alternative versions of the exposure model were observed.

The results from the quantification of the impact of the vulnerability uncertainty on the loss estimates are summarized in Fig. 6. The results are presented at national level only, as quite similar trends where observed for the case of Guatemala City. Fig. 6 presents the histogram of the annual loss ratios considering the three modelling approaches. The results indicate similar average annual loss ratios (AALR) amongst all cases. Although such result is expected, as the mean loss ratios of the vulnerability functions in all three cases are the same, one might expect a higher standard deviation (σ_{ALR}) when modelling the uncertainty using a beta or lognormal distributions. However, in seismic risk analyses of building portfolios across large regions an averaging effect is observed. As also demonstrated in [27], the consideration of vulnerability uncertainty for a given event will lead to losses below the mean in some areas and above the mean in others, resulting in aggregated losses (i.e. sum of the losses across the entire portfolio) equivalent to the case in which deterministic vulnerability functions (i.e. no uncertainty) were used. The exception occurs when vulnerability correlation is considered within buildings of the same building class. In this case, it is possible that the majority of all the losses will be above or below the mean, thus leading to scenarios where very low or very high aggregated losses occur. This is shown in the histogram of Fig. 7, where annual loss ratios above 30% only occur for the full correlation case, leading to a standard deviation 50% higher.

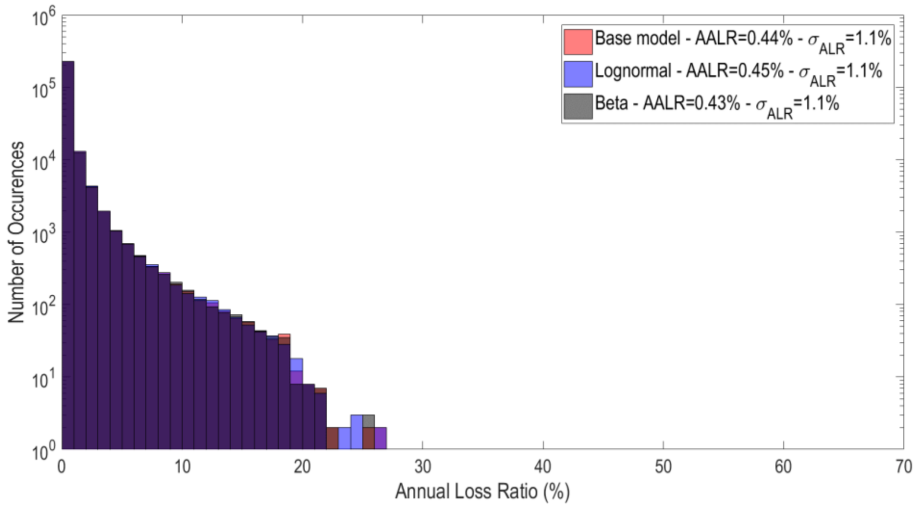

Fig. 6- Histogram of the annual loss ratios for Guatemala following three vulnerability modelling approaches.

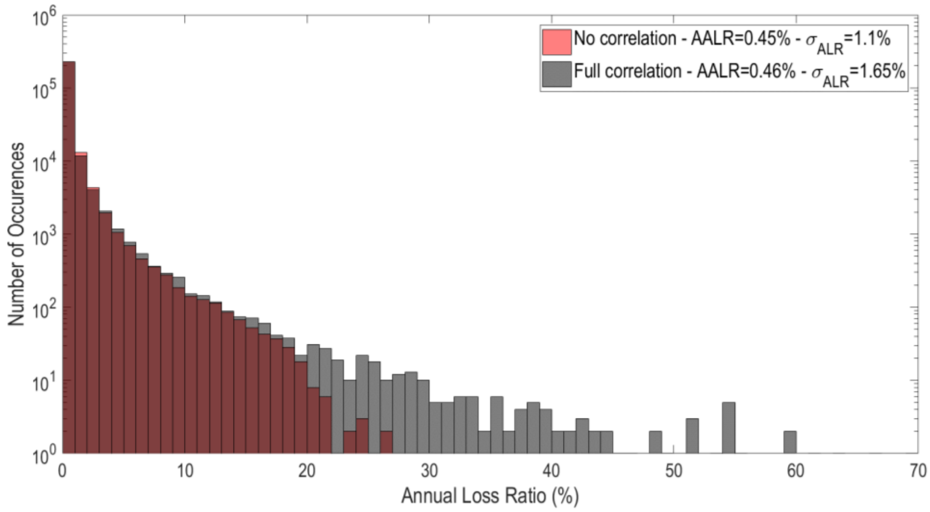

Fig. 7- Histogram of the annual loss ratios for Guatemala using lognormal distributions to model the uncertainty in the vulnerability, with and without full correlation.

5. Final Remarks

This study assessed the effects of a number of epistemic uncertainties on the earthquake loss estimation of a building inventory, explored with alternative versions of the model components of exposure, hazard and vulnerability implemented in the OpenQuake-engine. The risk estimates illustrated a high sensitivity on the seismic source model, the maximum moment magnitude of the seismic sources, and the selection of the GMPEs and associated weights. For the latter, a relative difference of up to 50% was observed in the predicted loss metrics between two branches of the logic tree.

The influence of spatial resolution of the exposure model was also explored, and it was found that for risk assessments at a city level, coarse resolutions should be refined in order to avoid biased loss estimates. For larger scale risk assessments, the estimated losses presented a strong dependency on the spatial distribution of seismic hazard. Regarding the consideration of the uncertainty in the vulnerability functions using probabilistic distributions, it was shown that the vulnerability correlation between assets of the same class dominates the losses. In summary, the results emphasize the significant impact of various sources of epistemic uncertainty on seismic risk assessment. Consequently, uncertainties in the input variables of a seismic risk model should be incorporated in the analyses and reflected in the risk estimates in order to support robust decision-making.

6. References

[1] Bazzurro P, Luco N (2007): Effects of Different Sources of Uncertainty and Correlation on Earthquake-Generated Losses. *Australian Journal of Civil Engineering*, **4**.

[2] Crowley H, Bommer JJ, Pinho R, Bird J (2005): *The impact of epistemic uncertainty on an earthquake loss model. Earthquake Engineering & Structural Dynamics*, **34**, 1653-1685.

[3] Silva V (2018): Critical Issues on Probabilistic Earthquake Loss Assessment. *Journal of Earthquake Engineering*, **22** (9), 1683-1709.

[4] Taylor P (2015): Calculating financial loss from catastrophes. in *SECED*, Cambridge, United Kingdom.

[5] Pagani M, Monelli D, Weatherill G, Danciu L, Crowley H, Silva V, Henshaw P, Butler L, Nastasi M, Panzeri L, Simionato M, Vigano D (2014): OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*, **85**, 692-702.

[6] Silva V, Crowley H, Pagani M, Monelli D, Pinho R (2014): Development of the OpenQuake engine, the Global Earthquake Model's open-source software for seism ic risk assessment. *Natural Hazards*, **72** (3), 1409-1427.

[7] Güendel F, Bungum H (1995): Earthquakes and Seismic Hazards in Central America. *Seismological Research Letters*, **66** (5), 19-25.

[8] Alvarado GE, Benito B, Staller A, Climent A, Camacho E, Rojas W, Marroquin G, Molina E, Talavera JE, Martinez-Cuevas S, Lindholm C (2017): The new Central American seismic hazard zonation: Mutual consensus based on up to day seismotectonic framework. *Tectonophysics*, **721**, 462-476.

[9] Benito MB, Lindholm C, Camacho E , Climent A, Molina E, Rojas W, Escobar JJ, Talavera E, Alvarado GE, Torres Y (2012): A New Evaluation of Seismic Hazard for the Central America Region. *Bulletin of the Seismological Society of America*, **102** (2), 504-523.

[10] Villagran M, Lindholm C, Dahle A, Cowan H, Bungum H (1996): Seismic hazard assessment for Guatemala City. *Natural Hazards*, **14** (2), 189-205.

[11] Silva, V., et al., (2020): Development of a global seismic risk model. *Earthquake Spectra*, DOI: 10.1177/8755293019899953

[12] Calderon A, S.V., Avilez M, Mendez R, Castillo R, Gil J.C, Lopez A (2019): Towards a Uniform Earthquake Loss Model across Central America. *Earthquake Spectra*.

[13]Brzev S, Scawthorn C, Charleson AW, Allen L, Greene M, Jaiswal K, Silva V (2013): *GEM Building Taxonomy Version 2.0*. Vol. GEM Technical Report 2013-02 V1.0.0. 2013, Pavia, IT: GEM Foundation. 188.

[14] Molina E, Marroquin G, Escobar JJ, Talavera E, Climent A, Rojas W, Camacho E, Benito B, Lindhom C (2008): Proyecto RESIS II Evaluación de la Amenaza Sísmica en Centro América. *NORSAR*.

[15] Climent ÁWT, Ciudad Real M, Strauch W, Villagran M, Dahle A, Bungum H (1994): Spectral strong motion attenuation in Central America. *NORSAR Technical Report*, **2-17**, 46.

[16] Zhao, J.X., et al., (2006): Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period. *Bulletin of the Seismological Society of America*, **96** (3), 898-913.

[17] Youngs RR, Chiou SJ, Silva WJ, Humphrey JR (1997): Strong Ground Motion Attenuation Relationships for Subduction Zone Earthquakes. *Seismological Research Letters*, **68** (1), 58-73.

[18] Martins L, Silva V (2018) : A global database of vulnerability models for seismic risk assessment. in *Proceedings of the 16th European Conference on Earthquake Engineering*, Thessaloniki, Greece

[19] Bazzurro P, Park J (2007): The effects of portfolio manipulation on earthquake portfolio loss estimates. in *Proceedings of the 10th International Conference of Application of Statistics and Probability in Civil Engineering (ICASP10)*, Tokyo, Japan.

[20] Dabbeek J, Silva V (2020): Modeling the residential building stock in the Middle East for multi-hazard risk assessment. Natural Hazards,. **100**, 781–810

[21] Román, M.O., et al., (2018): NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment*, **210**, 113-143.

[22] Wald D, Allen T (2007): Topographic Slope as a Proxy for Seismic Site Conditions and Amplification. *Bulletin of the Seismological Society of America*, **97**, 1379-1395.

[23] Yepes-Estrada C, Silva V, Rossetto T, D'Ayala, Ioannou I, Meslem A, Crowley H (2016): The Global Earthquake Model Physical Vulnerability Database. *Earthquake Spectra*, **32** (4), 2567-2585.

[24] Colombi M, Borzi B, Crowley H, Onida M, Meroni F, Pinho R (2008): Deriving vulnerability curves using Italian earthquake damage data. *Bulletin of Earthquake Engineering*, **6** (3), 485-504.

[25] Jaiswal KS, Aspinall WP, Perkins D, Wald D, Porter KA (2012): Use of expert judgment elicitation to estimate seismic vulnerability of selected building types. in *Proceedings of the 15th World Conference on Earthquake Engineering*, Lisbon, Portugal.

[26] Porter K (2010): Cracking an Open Safe: Uncertainty in HAZUS-Based Seismic Vulnerability Functions. *Earthquake Spectra*, **26** (3), 893-900.

[27] Silva V (2019): Uncertainty and correlation in seismic vulnerability functions of building classes. *Earthquake Spectra*, **35** (4), 1515-1539.

[28] Tyagunov S, Pittore M, Wieland M, Parolai S, Bindi D, Fleming K, Zschau J (2014): Uncertainty and sensitivity analyses in seismic risk assessments on the example of Cologne, Germany. *Natural Hazards and Earth System* Sciences, **14** (6), 1625-1640.

[29] Dabbeek J (2017): The effect of spatial resolution in exposure models for seismic loss estimation. *Master's Thesis*, Instituto Universitario di Studi Superiori di Pavia, Italy.

[30] Bal IE, Bommer JJ, Staffor PJ, Crowley H, Pinho R (2010): The influence of geographical resolution of urban exposure data in an earthquake loss model for Istanbul. *Earthquake Spectra*, **26** (3), 619-634.

