Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Epistemic uncertainty of probabilistic building exposure compositions in scenario-based earthquake loss models

Juan Camilo Gomez-Zapata (jcgomez@gfz-potsdam.de)
Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam
https://orcid.org/0000-0001-7919-5549

Massimiliano Pittore
Institute for Earth Observation, EURAC Research

Fabrice Cotton
Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam

Henning Lilienkamp
Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam

Simantini Shinde
Helmholtz-Zentrum Potsdam GFZ: Deutsches Geoforschungszentrum Potsdam

Paula Aguirre
National Research Center for Integrated Natural Disaster Management (CIGIDEN)

Hernan Santa Maria
National Research Center for Integrated Natural Disaster Management (CIGIDEN)

Original Article

Keywords: epistemic uncertainty, faceted taxonomy, probabilistic exposure modelling, earthquake scenario, earthquake loss models, cross-correlated ground motion fields

Posted Date: February 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-178120/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Bulletin of Earthquake Engineering on January 20th, 2022. See the published version at https://doi.org/10.1007/s10518-021-01312-9.

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Epistemic uncertainty of probabilistic building exposure compositions in scenario-based earthquake loss models

Juan Camilo Gomez-Zapata[1, 2], Massimiliano Pittore[1, 3], Fabrice Cotton[1, 2], Henning Lilienkamp[1, 2], Simantini Shinde[1], Paula Aguirre[4, 5], Hernán Santa Maria[4, 5]

Abstract

In seismic risk assessment the sources of uncertainty associated to building exposure modelling have not received as much attention as other components related to hazard and vulnerability. We are introducing that the degree of knowledge of a building portfolio can be described within a Bayesian probabilistic approach to acknowledge the epistemic uncertainty of its composition (i.e. proportions per building class). We are investigating the impact of such uncertainty on earthquake loss models through a novel approach based on an exposure-oriented logic tree arrangement and scenario-based seismic risk simulations. We have found that the building class reconnaissance either from prior assumptions by desktop studies with aggregated data (top-down approach) or from building-by-building data collection (bottom-up approach), plays a fundamental role in statistical modelling of exposure. We successively show that the selection of the basic set of building classes is a major contribution to the uncertainties in the loss estimations given their dependence on specific fragility functions. If the set of fragility functions handle multiple spectral periods, cross-correlated ground motion fields are required for the vulnerability assessment which ultimately control the variability of the losses. When the exposure composition is poorly known due to, for instance, simplifications and assumptions over aggregated data, its absolute contribution to the loss uncertainties has been found to be larger than the one imposed by the spatially distributed cross-correlated ground motion fields. This work invites to redesign desktop exposure studies, while it also highlights the importance of a standardized iterative approach to exposure data collection and modelling.

Keywords: epistemic uncertainty; faceted taxonomy; probabilistic exposure modelling; earthquake scenario; earthquake loss models; cross-correlated ground motion fields.

Juan Camilo Gomez-Zapata
jcgomez@gfz-potsdam.de

[1] Helmholtzcentre Seismic Hazard and Risk Dynamics, Helmholtz Centre Potsdam – German Research Centre for Geosciences GFZ, Telegrafenberg, 14467 Potsdam, Germany
[2] Institute for Geosciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany
[3] Institute for Earth Observation, EURAC Research, Viale Druso 1, 39100, Bolzano, Italy
[4] National Research Center for Integrated Natural Disaster Management (CIDIGEN), ANID/FONDAP/15110017, Vicuña Mackenna 4860, Santiago 7820436, Chile
[5] Department of Structural and Geotechnical Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile
[6] Institute for Mathematical and Computational Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile

1. Introduction

Epistemic uncertainties stem from the incomplete knowledge of the actual problem and its parameters or simply from the, often unavoidable, modelling and methodology errors (e.g. Vamvatsikos et al. 2010). The performance of earthquake loss models for large-scale residential buildings portfolios under the influence of such epistemic uncertainties related to the lack of data availability in the exposure composition is the central aspect of this work. Exposure refers to the number, type and monetary value of elements (e.g. buildings) that are under threat from natural hazards and subjected to potential loss (e.g. UNISDR, 2009) and, together with the hazard and vulnerability components contributes to most of quantitative risk assessment applications. In these study the degree of knowledge of the hazard and exposure components plays a fundamental role since their associated uncertainties are propagated to the final loss estimates. Therefore, accurate estimations of the expected spatially distributed hazard intensities (i.e. PGA) for an earthquake scenario, together with a consistent classification of the building stock into suitable building vulnerability classes will minimize the variance of the final loss estimates over an area of interest. Being able to track and disaggregate the influence of the hazard and exposure is key for decision making, urban planning, and finance. In the latter, the smaller the variation about the mean loss values, the lower the risk is perceived (Wesson and Perkins, 2001).

While the hazard component has received significant research attention in seismic risk models, only recently the exposure uncertainty has been identified as an area that would particularly benefit from increased research attention (Crowley 2014; Corbane et al. 2017). Under that framework there are several epistemic uncertainty components that need further exploration such as the basic reconnaissance of the building class while gathering their individual attributes, as well a basic exploration in the loss outcomes if more than a single set of building classes is used.

In exposure modelling the buildings are classified in vulnerability-classes which ultimately describe their differential susceptibility to damage. The vulnerability-class definition is, therefore, linking the hazard intensities to the expected damage based on a clear reconnaissance of the building's structural and non-structural characteristics (e.g. Calvi et al. 2006). Conventionally, the relative uncertainty associated with the definition of building classes and their proportions have been assumed to contribute much less to the final loss estimates than the aleatory components of the risk processing chain (i.e. ground motion variability in seismic hazard). That typical underlying assumption has led to the general practice of assuming that the building exposure data collection is not as worthwhile compared to more detailed explorations required by the hazard component (Crowley and Bommer, 2006). This practice has led to some community- accepted assumptions such as supposing fixed proportions over aggregated data (i.e. census-based desktop studies), without exploring their underlying uncertainties. Therefore, the reconnaissance of a selected set of taxonomic attributes and their incorporation within a statistical building exposure model while investigating the uncertainty in the class assignment and their effects in the loss estimates is a pathway worth to be explored in the seismic risk framework.

Having in mind the aforementioned motivation, this work describes how the epistemic uncertainty associated with the building exposure model definition highly contributes to the variance in the loss estimates in earthquake scenarios. In the first place, the limitations in the current state of the art practices in building- exposure modelling for seismic risk are described. Subsequently, a method to explore the epistemic uncertainty of the building exposure models in earthquake loss models is proposed. As a premise, in this work the building exposure model is visualized under the scope of the compositional theory within a Bayesian framework in three stages: (1) statistical modelling from data collection of building's attributes and the configuration of synthetic building portfolios; (2) a novel method to obtain the probabilistic compatibility-levels between two sets of building classifications (schemes). And lastly, (3) the novel proposal of a logic-tree arrangement that propagates and disaggregates factors contributing to uncertainty within the exposure model up to the loss estimates while addressing the spatial cross-correlation of the seismic ground motion in an earthquake scenario. These steps are exemplified through a set of building exposure models in Valparaiso (Chile) to investigate their associated losses in the realistic framework of a subduction earthquake scenario.

2. Current state of the art in building- exposure modelling for seismic risk assessment

Few initial efforts to explore the associated uncertainties in the building exposure composition have been carried out. Among them, Crowley et al. (2005) performed a sensitivity analysis by comparing the variability of the expected damage grades varying the building compositions in terms of the storey ranges over a relatively homogeneous portfolio (the same reinforced concrete building type with height ranges between 1-10 storeys). This process implicitly accounted for the resignation of the fragility function while changing their proportions. It was shown the great impact imposed by a grouping procedure considering the number of storeys as the unique difference between classes. Under a similar framework, if a wider set of building attributes within a portfolio are considered, the variability in the expected damage and losses still remains unexplored up to the current date. This is partly because the associated complexity in the building classification would increase and lead to a more extensive set of classes in comparison with the available set of fragility functions (Haas, 2018; Martins and Silva, 2020). Recently, further contributions about the epistemic uncertainties in regional exposure models have been explored in Kalakonas et al. (2020). The authors observed negligible differences loss estimates when alternative exposure models were compared in a probabilistic risk assessment.

In classic exposure models for seismic risk only some basic building attributes have been used to classify a building stock (e.g. material of the lateral load resistance system (LLRS), height, and material technology). This has led to the consideration of the spatial variation of individual taxonomic attributes as random and smaller than the uncertainty induced by grouping different individual buildings into a single class (Crowley and Pinho, 2004). This simplification has been used to represent the epistemic uncertainty associated with the lack of building reconnaissance as aleatory uncertainty (Crowley et al. 2005). Therefore, it has been acknowledged that detailed inspections to collect attributes of all the buildings in a study area would allow to address this uncertainty as epistemic. Unfortunately, considering the extent and the very same evolution of the built environment, a full- enumeration of the assets could be a highly time and resource-intensive task, often simply unfeasible (Pittore et al. 2017). However, if only a sample of building structures within the entire stock is inspected, the epistemic uncertainly associated with the class assignment should be accounted for. Thus, considering the presence of certain building attributes as aleatory within the exposure model should not be longer valid mostly when they might be seismic-vulnerability drivers, such as foundation type; size and the shape of the vertical structure (e.g. Dell'Acqua, et al. 2013) or structural irregularities (e.g. Martínez-Cuevas et al. 2017).

Large-scale residential exposure models (i.e. nation, region-wide, or commune/ district level) are often based on information derived from census and are aggregated over specific geographical administrative units. Buildings are grouped into categories with expected similar performance when subjected to ground shaking. These categories describe vulnerability classes to a natural hazard and are described by a set of mutually exclusive, collectively exhaustive (MECE) building classes. Hereafter we will call "*scheme*" such a set of building classes. Every building class within a given scheme can be interpreted as a *risk-oriented building exposure taxonomy* which can be further decomposed into facets or attributes within a *faceted taxonomy* (e.g. GEM v.2.0, Brzev et al. 2013). This has been described in Pittore et al. (2018) and has been also noticed in Pavić et al. (2020).

Some of the most common *risk-oriented taxonomies* include: European Macroseismic Scale 1998 (EMS-98, Grünthal, 1998); HAZUS (FEMA, 2003); PAGER-STR (Jaiswal et al. 2010). Given the lack of local models they have been applied outside their original geographical scope. This is the case of HAZUS, originally designed for the USA (specific hazard intensities and local building types), that, due to the similarities between seismic-code compliance with other countries, it has been used to classify building stocks and estimate losses in other geographical contexts (e.g. Aguirre et al. 2018 in Chile). Similar practices have been reported using the European EMS-98 risk-oriented taxonomy in Central Asia (e.g. Bindi et al. 2011; Pilz et al. 2013).

Faceted taxonomies, in contrast, provide an exhaustive and structured set of mutually exclusive and well-described attributes. These taxonomies allow to describe both individual structures and homogenous populations in a standard way, and are largely independent on specific fragility or vulnerability models. The most widely used and well-established one is the GEM Building Taxonomy (GEM v.2.0, Brzev et al. 2013). It has been adapted for a multi-hazard-risk initiative (GED4ALL, Silva et al. 2018) and to classify structures with special occupancies such as schools to assess their seismic vulnerability in the Global Library of School Infrastructure project (GLOSI in D'Ayala et al. 2020) led by the Global Program for Safer Schools (GPSS) of the World Bank.

Regardless of the type of taxonomy (either risk-oriented or faceted), there are two conventional methods for exposure modelling of large scale spatially distributed buildings: (1) *a top-down* approach: the analysis of aggregated data (e.g. census data) through expert elicitation and (2) *a bottom-up* approach using individual observations. These two approaches classify the building stock addressing a double expert elicitation:

(1) To classify the building inventory into assumed building classes in a given study area.
(2) To obtain the building exposure composition (i.e. proportions in every building class) through the use of exposure mapping- schemes.

2.1. Top-down approach: building class from the analysis of aggregated data

Recently, the implementation of the GEM *"mapping-schemes"* for the analysis of aggregated data has been outlined in Yepes-Estrada et al. (2017) and further implemented in the European Exposure model (Crowley et al. 2020) and in the Global seismic risk model (Silva et al. 2020). These mapping-schemes classify a building stock through desktop studies and expert elicitation into earthquake vulnerability classes. Each class is described by selected attributes from the GEM V2.0 faceted taxonomy. They rely on the available regionally aggregated data (e.g. region-specific census) while addressing socioeconomic characteristics for *dwellings* and not at the building level (Crowley, 2014). Since the few valuable attributes to physical vulnerability assessment refer to the building façade predominant material and occupancy type, these mapping-schemas have been customized to include other attributes by defining covariate relations between census descriptors and expected proportions per building class (e.g. Dabbeek and Silva, 2020; Acevedo et al. 2020) to ultimately use a single set of building classes to represent a building stock. Therefore, the variation of taxonomic building attributes is still being addressed as random within an aleatory uncertainty framework instead of a reducible epistemic uncertainty. Some drawbacks of this approach have been described in Villar-Vega and Silva, (2017). Moreover, building exposure models derived from purely top-down desktop studies neglect the temporal evolution of the exposure data. Since census data are neither standard across regions nor in time (possible changing data formats), once the mapping-schema is used the resultant exposure model would remain static until a new census information is generated over the same area.

2.2. Bottom-up approach: individual building observations

When the composition of the portfolio is expected to be heterogeneous, data collection of attributes over a selected sample of individual buildings is required to constrain and validate the underlying assumptions imposed by a top-down vision. For that purpose, freely available data products from volunteering mapping initiatives such as OpenStreetMap (OSM) may offer some attributes (occupancy or footprint shape). However, due to the lack of standardized data formats, other vulnerability driver's attributes are still required to describe more robust schemes. Thus, efficient data collection with harmonized data- formats is still a pending aspect to be solved by volunteered mapping. This harmonization issue has been addressed through the use of standards while including building taxonomic attributes. This is the case of FEMA-154 (FEMA 154, 2002); the SASPARM 2.0 project (Grigoratos et al. 2016); CARTIS (e.g. Zuccaro and Cacace 2015; Polese et al. 2020); some initiatives for data collection of post-earthquake damage such as AeDES (Baggio et al. 2007; Nicodemo et al. 2020); and focus on historical buildings (e.g.

Jiménez, et al. 2018). Moreover it has been used to collect features in the field or remotely using the IDCT tool (e.g. Jordan et al. 2014; Santa María et al. 2017) or the RRVS one (e.g. Wieland et al. 2014; Pittore and Wieland 2013) for a continuous and dynamic data assimilation for building exposure modelling in terms of a faceted taxonomy keeping harmonized data standards.

2.3. Dynamic building exposure based on data collection and statistical modelling

As a way to combine top-down and bottom-up approaches in a consistent modelling framework, a dynamic building exposure modelling with statistical nature has been recently suggested by (Pittore et al. 2020). They combine the use of risk-oriented schemes and building-by-building data collection (using a faceted taxonomy). The portfolio's vulnerability classes are defined in a top-down manner, while the expected frequency of the related classes can are further constrained through a bottom-up approach by integrating data collection from surveys, or even from building-features extraction through satellite imagery analyses (e.g. Torres et al. 2019, Liuzzi et al. 2019). To the best of authors' knowledge, these statistical models have not yet been exploited to investigate the epistemic uncertainties in the building exposure composition and its impact in the loss estimates. A detailed exploration of the epistemic uncertainty carried by statistical building exposure models and its propagation until scenario-based loss estimates is introduced hereafter.

3. Methodology

3.1. Probabilistic - exposure models: a Bayesian formulation

The building portfolio configuration is conceptualized within the *compositional theory* within a fully probabilistic Bayesian framework as initially suggested by (Pittore et al. 2020). First we will introduce the concept of the likelihood function, followed by the assumptions on the prior and the posterior distributions in this context. This formulation considers risk-oriented schemes that contain a finite set of building class and their associated fragility functions. The resulting exposure model highlights its statistical nature related to its dynamic spatiotemporal evolution.

3.1.1. The definition of the likelihood function: The intra-scheme compatibility levels

A suitable scheme containing k risk-oriented building classes has been selected for the area of interest, where we assume some data has also been collected through surveying (evidence). We assume that the probability of observing a given distribution of $\boldsymbol{n} = \{n_1, \dots, n_k\}$, $\sum_k n_i = N$ building types, where n_i is the number of specimens of building type i, follows a multinomial probability distribution. As proposed in (Pittore et al. 2020), the probability of observing $\boldsymbol{n}$ conditional on $\boldsymbol{\theta}$ is given by:

$$p(\boldsymbol{n}|\boldsymbol{\theta}) = Mul\,(\boldsymbol{n}|\theta)\frac{N!}{\prod_{i=1}^{k} n_i!}\prod_{i=1}^{k}\theta_k^{n_i} \qquad \text{Eq. 1}$$

With k typologies or building class, whose frequencies are represented by the proportion $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_k\}$, $\theta_i > 0\ \forall i$ and $\sum_k \theta_{\mathrm{k}} = 1$. We assume that the set of observations given their proportions is the likelihood distribution of the Bayesian formulation as explained hereafter.

Every building class within a considered scheme is translated into the attribute values offered by a faceted building taxonomy. This emerges naturally considering the bottom-up data collection of individual attributes. Triangular fuzzy values are assigned through expert criteria to score the compatibility degree between the observed attribute values and the every building class as formulated in Pittore et al. (2018) to constrain the actual proportion of every class within

the exposure model. Subsequently, the data collection over individual buildings either on the field or through remote visual inspections are used in the class assignment. Every attribute type *j* has an associated numerical weight, w_j that acknowledges their relevance in vulnerability assessment as well as their ability of reconnaissance during the survey. By evaluating the compatibility degree between the observed building attributes and the building class, a transparent assignment of the most likely class within a fully probabilistic framework is achieved. This formulation contains several ingredients subjected to epistemic uncertainty (i.e. selection of the fuzzy values; weights; classes; fuzzy scores). We have only considered a sensitivity on the weights.

3.1.2. Prior and posterior distributions

Some of the concepts described in (Pittore et al. 2020) are recalled hereafter. The expected proportion θ_{ob_i} for every building class are treated as Dirichlet-distributed random variables (see Eq. 2)

$$Dir(\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k} \theta_i^{\alpha_i - 1} \qquad \text{Eq. 2}$$

Where $\alpha = \{\alpha_1, \dots, \alpha_k\}$, $\alpha_i > 0 \ \forall i$ with $\alpha_0 = \sum_{i=1}^{k} \alpha_i$ being called concentration factor. The Dirichlet hyper parameter α_i is factorized as a product of a proportion (θ_k) and a common concentration factor such as: $\alpha_i = (\theta_k)(\alpha_o)$ where α_0 increases the virtual counts for the category k. By Bayes theorem, and since the prior Dirichlet is the conjugate prior for the multinomial likelihood term, we describe the posterior probability distribution of θ in terms of the likelihood and prior:

$$p(\boldsymbol{\theta}|\boldsymbol{n}) \propto p(\boldsymbol{n}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) \qquad \text{Eq. 3}$$

When the number of observations increases the probability estimate is dominated by the multinomial likelihood which allows the definition of expert-based priors that are increasingly superseded by real data captured during surveys ((Pittore et al. 2020). Being the Dirichlet a probability distribution over probabilities, for every set of hyper-parameters, the distribution is sampled. Each sample provides a different composition within synthetic building portfolios. Larger values of α_0 ($\alpha_0 \sim 50$) give similar compositions to the considered posterior. Smaller ones ($\alpha_0 \sim 1$) result in sparser distributions (Hastie et al. 2015) and thus is interpreted as divergent opinions respect to the portfolio composition of the posterior distribution.

3.1.3. Logic tree construction; synthetic building portfolio construction and allocation

The prior and likelihood terms are further characterised to obtain customized posterior distributions. This arrangement describes a logic tree as a tool to explore the epistemic uncertainty in the building exposure as composition data. With this formulation we keep the statistical nature of the exposure modelling overcoming a unique top-down vision of the portfolio composition. It has five complexity levels explained as follows.

(1) The selection of the building class scheme (group of building classes)
(2) The selection of the numerical weight, w_j (per attribute type *j*) that score and rank the relevance of every attribute type in vulnerability assessment and their ability of reconnaissance during the survey. The set of w_j is called 'weight arrangement" (W.A)
(3) The definition of a prior distribution that describes the initial guess about the composition of the building portfolio in the form of a Dirichlet distribution. This composition describes the representability of every building class in the area and is driven by data collection, expert criteria or aggregated data (e.g. official census).

(4) The selection of the hyper-parameter α_0 (concentration factor(s)) of the conjugate posterior Dirichlet distribution obtained from Eq. 3. This selection acknowledges the degree of trust in the former assumptions. α_0 ~ 1.0 means low information content and hence low knowledge of the portfolio composition whilst α_0 ~ 50 represents a higher knowledge.
(5) The selection of the number of samples (m) to be obtained for each concentration factor of the posterior Dirichlet distribution in order to get the stochastic compositions within synthetic building portfolios. α_0 ~ 1 means divergent opinions across m, whilst α_0 ~ 50 emulates an almost unanimous consensus on the portfolio composition across m.

In order to spatially distribute every synthetic building portfolio, a dasymetric disaggregation from population counts is followed as proposed in (Pittore et al. 2020). This approach is a suitable for the purpose of a differential spatial allocation of the synthetic building portfolios whose composition is being reconfigured with every sample. Population counts reported in any aggregated data sources (e.g. LandScan; WorldPop; GPWv4) can be used. This procedure is similar to the method reported in Dell'Acqua, et al. (2013).

It should be noted that the former steps regard the use of a single group of building classes (scheme). However, if more than one scheme are found to be suitable to describe the building stock in a given area for seismic risk, besides ensuring the availability of associated fragility functions, some basic exposure metrics for every building class such as their proportions, average nigh-time residents and replacement cost have to be obtained for this alternative exposure classification. This is not a trivial task since this information might be only available in terms of one reference scheme. For that purpose we propose to obtain these metrics for other scheme through the formulation of a probabilistic inter-scheme compatibility matrix. This method is described hereafter.

3.2. Probabilistic inter-scheme compatibility matrix

Besides the classes themselves, other exposure descriptors such as the proportions within every building class; their replacement costs and number of inhabitants may be desired under more two reference schemes that represent the same building portfolio. Subjective relations between the building class contained in two different schemes have been proposed between HAZUS and EMS-98 (Hancilar et al. 2010) for ELER, the Earthquake Loss Estimation Routine (by RISK-EU). However, in order to exploit the decomposition of a risk-oriented building class into a faceted taxonomy, we propose a probabilistic representation of the building class to express the exposure descriptors as a conditional probability for two references, the source scheme (T_k^A) and the target- scheme (T_j^B) composed by k and j building classes respectively. We obtain their compatibility degree as $p(T_k^A|T_j^B)$ in a matrix. The first step is to define every building type in terms of a probabilistic description of basic taxonomic features ($\{F\}_m$) as expressed in Eq. 4.

$$\sum_m p(T_k^A|\{F\}_m) \qquad \text{Eq. 4}$$

A straightforward application of the total probability theorem and a probabilistic description of the building type in taxonomic features allow us to present Eq. 5.

$$p(T_k^A|T_j^B) = \sum_m p(T_k^A|\{F\}_m \cap T_j^B)p(\{F\}_m|T_j^B) \qquad \text{Eq. 5}$$

Since we assume that the representations of a building within the two considered schemes are conditionally independent given the information on the individual taxonomical features we can describe the source scheme (T_k^A) to be modelled in terms of the taxonomic attributes that also compose the target- scheme (T_j^B): $T_k^A \vDash T_j^B|\{F\}_m$, the former equation can be expressed as a product, as expressed in Eq. 6.

$$p\left(T_k^A \middle| T_j^B\right) = \sum_m p(T_k^A|\{F\}_m)p(\{F\}_m|T_j^B) \;\; since\; T_k^A \perp\!\!\!\perp T_j^B|\{F\}_m \qquad \text{Eq. 6}$$

We obtain a probabilistic representation of the compatibility degree across the two considered building classes in an alternative Bayesian formulation as presented in Eq. 7.

$$p(T_k^A|T_j^B) = \sum_m p(T_k^A|\{F\}_m)p(T_j^B|\{F\}_m)\frac{p(\{F\}_m)}{p(T_j^B)} \qquad \text{Eq. 7}$$

Synthetic surveys based in the possible combinations of attributes that may describe every building class are input to solve the compatibility scores and are integrated through the selection of the weight arrangement of the scheme that, as proposed Pittore, et al. (2018), rank the commonly considered attribute values (e.g. material of the LLRS, number of storeys). Using this matrix we obtain the missing exposure descriptors of the target scheme (e.g. average nigh-time residents and replacement cost of every class) from the most compatible class of the source scheme. Once we have the residents and prior compositions of the alternative building portfolios, we can perform the formerly described dasymetric disaggregation procedure and estimate losses. This approach highlights the advantages of representing the building classes by a faceted taxonomy.

3.3. Scenario-based earthquake-rupture and ground motion distribution

An earthquake scenario that replicate a historical event is selected for the construction of a seismic rupture and the simulation of spatially distributed ground motion from suitable GMPE (ground motion prediction equations). Cross-correlated ground motions are generated for the spectral-periods that serve the fragility functions as intensity measures (IM) to and its use is transversal to all of the logic tree levels explained in title 3.1.3.

3.4. Epistemic uncertainty exploration of the building exposure definition in earthquake loss models

The impact of the building exposure composition in earthquake loss models is investigated by comparing the differences in the losses obtained from every synthetic building portfolio from the logic-tree proposed (title 3.1.3). To complement the vulnerability analysis, we have selected a consequence model that includes the total replacement cost for the building class and their loss ratios for every damage state. 1,000 ground motion simulations are computed for the considered earthquake rupture to address its aleatory uncertainty (Silva, 2016). For every realisation the ground motion cross-correlation of the intra-event residuals is also estimated. This provides a robust framework to propagate and explore the exposure uncertainty within the seismic risk processing chain.

4. Application

4.1. Context of the study area: Valparaiso, Chile

The study area comprises the communes of Valparaiso and Viña del Mar (see Figure 1). Hereafter, for simplicity, both communes will be called ‘Valparaiso’. It is the second largest Chilean urban centre. Its port is the main container and passenger port in Chile and vital for the country’s economy. As described by Indirli et al. (2011), Valparaiso shows a very heterogeneous building inventory in which its historic district was declared World Heritage Site by UNESCO in 2003 after recognizing its diverse urban layout and architecture (Jiménez, et al. 2018).

The Central Chile area, and in particular Valparaiso, have been hit by powerful historical earthquakes. One of the few

with historical descriptions is the 1906 earthquake, with an inferred Mw 8.0 - 8.2 (Carvajal et al. 2017), which caused widespread damage (Montessus de Ballore, 1914). In 1985 a Mw 7.8 event with an epicentre located just 120 km west of the city, destroyed 70.000 houses and damaged an additional 140.000 dwellings, leaving 950.000 persons homeless, and caused losses of about $1.8 billion (Comte et al. 1986). The 2010 Mw 8.8 earthquake caused structural damage to some buildings in Viña del Mar (de la Llera et al. 2017) and impacted the labour market recovery and overall economy (Jiménez Martínez et al. 2020). Furthermore, recent seismic activity has been noticeable during the 2017 Mw 6.9 which was triggered by a slow slip event described in Ruiz et al. (2017). It is notable the input of the MARVASTO project (Indirli et al. 2011) that addressed the construction of earthquake scenarios to obtain the expected seismic ground motion. However, to best authors' knowledge, no scenario-based for seismic risk over building portfolios in Valparaiso has been reported in the scientific literature.

4.2. Probabilistic exposure model construction for Valparaiso

4.2.1. The definition of the likelihood function: The intra-scheme compatibility levels

Two earthquake-oriented building schemes, namely SARA and HAZUS, have been considered to represent the building portfolio in Valparaiso. Both schemes have been already proposed for exposure modelling at the third administrative division "commune" in Chile in former studies. SARA constitutes an effort to harmonize and define all the building types in the South-American Andes region (GEM, 2014) largely based in the World House Encyclopaedia reports (WHE, 2014) through expert judgment and has let us to infer 17 building class for Valparaiso combining the storey ranges when it was possible. HAZUS addresses 11 ones according to Aguirre et al. (2018). Every building class within both schemes has an associated fragility function reported in Villar-Vega et al. (2017) and FEMA (2012) respectively. SARA considers three intensity measures (IM) (i.e. Spectral acceleration at the structural yield period), whilst HAZUS considers PGA as the only IM Furthermore, SARA implies the assumption that the construction practices for residential buildings in Chile does not include steel types nor reinforced concrete framed structures, and notably only considers wall-structures. Moreover, it only considers buildings up to 19 storeys. This is in contrast with Jiménez, et al. (2018) and Aguirre et al. (2018) that both report higher rise and steel-type building classes for the study area.

Every building class- within the two schemes- has been decomposed into attribute types and values contained in the GEM v.2.0 taxonomy in order to assign their corresponding fuzzy compatibility levels through expert elicitation as proposed in Pittore et al. (2018). Their graphical representation is depicted in Figure 2. Data-collection was surveyed in order to test the actual plausibility of the schemes in the study area and to construct the customized likelihood distributions. 604 randomly distributed buildings in the urban area of Valparaiso (Figure 1) were inspected by local structural engineers from the Chilean Research Centre for Integrated Disaster Risk Management (CIGIDEN) while collecting their attribute values in terms of the GEM v.2.0 taxonomy through the RRVS web-platform (Haas et al. 2016).

In the original formulation of Pittore et al. (2018), a set of numerical weights to rank differently the importance of the considered attribute types is needed. This set jointly must sum up to 1.0 and acknowledge the relevance of the attribute value for the seismic performance of the building as well as the availability of reconnaissance during the particular survey. This selection is based on expert elicitation and carries a further epistemic uncertainty. In SARA and HAZUS, four common attribute types have been selected and two different sets of weight arrangements, WA-1 and W.A-2 reported in in Table 2 assign a differential importance to the attributes, being the largest differences between the LLRS ductility and storey range. Every weight arrangement leads to different survey compositions as shown in Figure 3 and Figure 4 mostly due to the fact that the LLRS ductility was, in most of the cases, not identified during the surveys, leading to the identification of either unreinforced structures (W.A-1) or reinforced structures (W.A-2).

Table 1. Building classes in terms of the SARA scheme proposed for Valparaiso with the respective intensity measure of their associated fragility functions (as reported in Villar- Vega et al. 2017); together with the prior proportions per building class; average night-time residents (Res. /bdg) and replacement cost (Repl. Cost (USD/bdg)) as reported in Yepes-Estrada et al. 2017.

SARA building class	Description	IM	Prior Prop	Res. /bdg	Repl. Cost (USD/bdg)
CR-PC-LWAL-H1-3	RC with shear wall, low rise	PGA	0.005	18	360000
CR-LWAL-DNO-H1-3	RC with shear wall, low rise (non-ductile)	PGA	0.032	14	288000
CR-LWAL-DNO-H4-7	RC with shear wall, medium rise (non-ductile)	S.a at 1.0s	0.010	54	1080000
CR-LWAL-DUC-H1-3	RC with shear wall, low rise (ducitle)	PGA	0.011	15	336000
CR-LWAL-DUC-H4-7	RC with shear wall, medium rise (ductile)	S.a at 1.0s	0.006	54	1260000
CR-LWAL-DUC-H8-19	RC with shear wall, high rise (ductile)	S.a at 1.0s	0.002	173	4032000
ER-ETR-H1-2	Rammed earth, low rise	PGA	0.029	4	43750
MCF-DNO-H1-3	Masonry confined, low rise (non-ductile)	PGA	0.152	5	94500
MCF-DUC-H1-3	Masonry confined, low rise (ductile)	PGA	0.034	14	288000
MR-DNO-H1-3	Reinforced masonry, low rise (non-ductile)	PGA	0.029	18	360000
MR-DUC-H1-3	Reinforced masonry, low rise (ductile)	PGA	0.012	18	360000
MUR-ADO-H1-2	Unreinforced masonry with adobe, low rise	PGA	0.111	4	43750
MUR-STDRE-H1-2	Unreinforced masonry with stone, low rise	PGA	0.006	5	43750
MUR-H1-3	Unreinforced masonry, low rise	PGA	0.060	6	52500
UNK	Unknown	S.a at 0.3s	0.108	4	35000
W-WLI-H1-3	Light wood, low rise	S.a at 0.3s	0.273	5	108000
W-WS-H1-2	Solid wood, low rise	S.a at 0.3s	0.121	4	43750
			∑ 1.0		

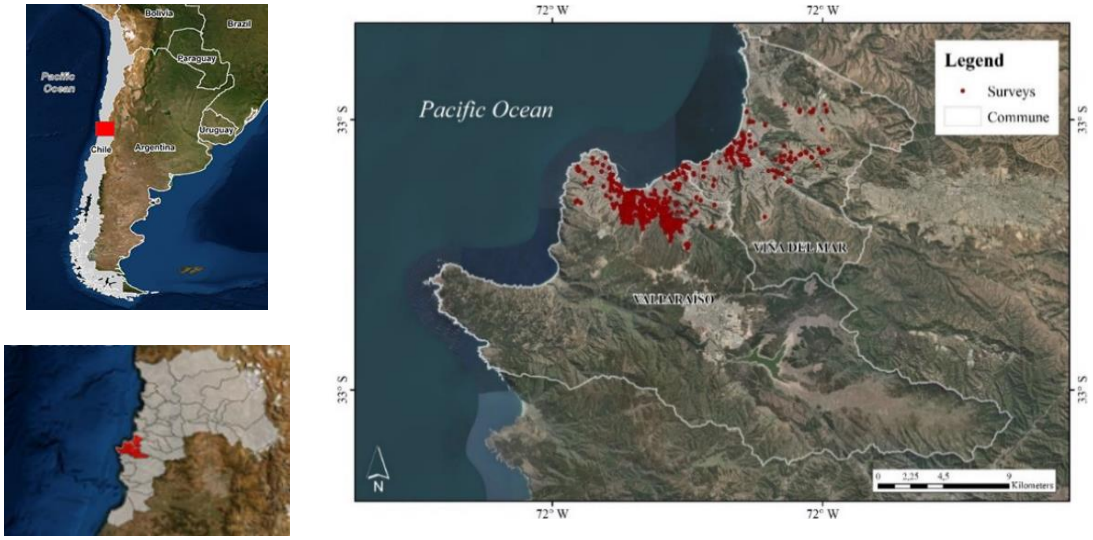


Figure 1. Location of the surveyed buildings within the urban area of the communes of Valparaiso and Viña del Mar within Valparaiso Region and Chile.

Table 2. Weight arrangement for SARA and HAZUS schemes.

Taxonomic Attribute	W.A-1	W.A-2
Material type	0.40	0.30
Material technology	0.10	0.10
LLRS	0.15	0.20
LLRS ductility	0.25	0.10
Storey range	0.10	0.30

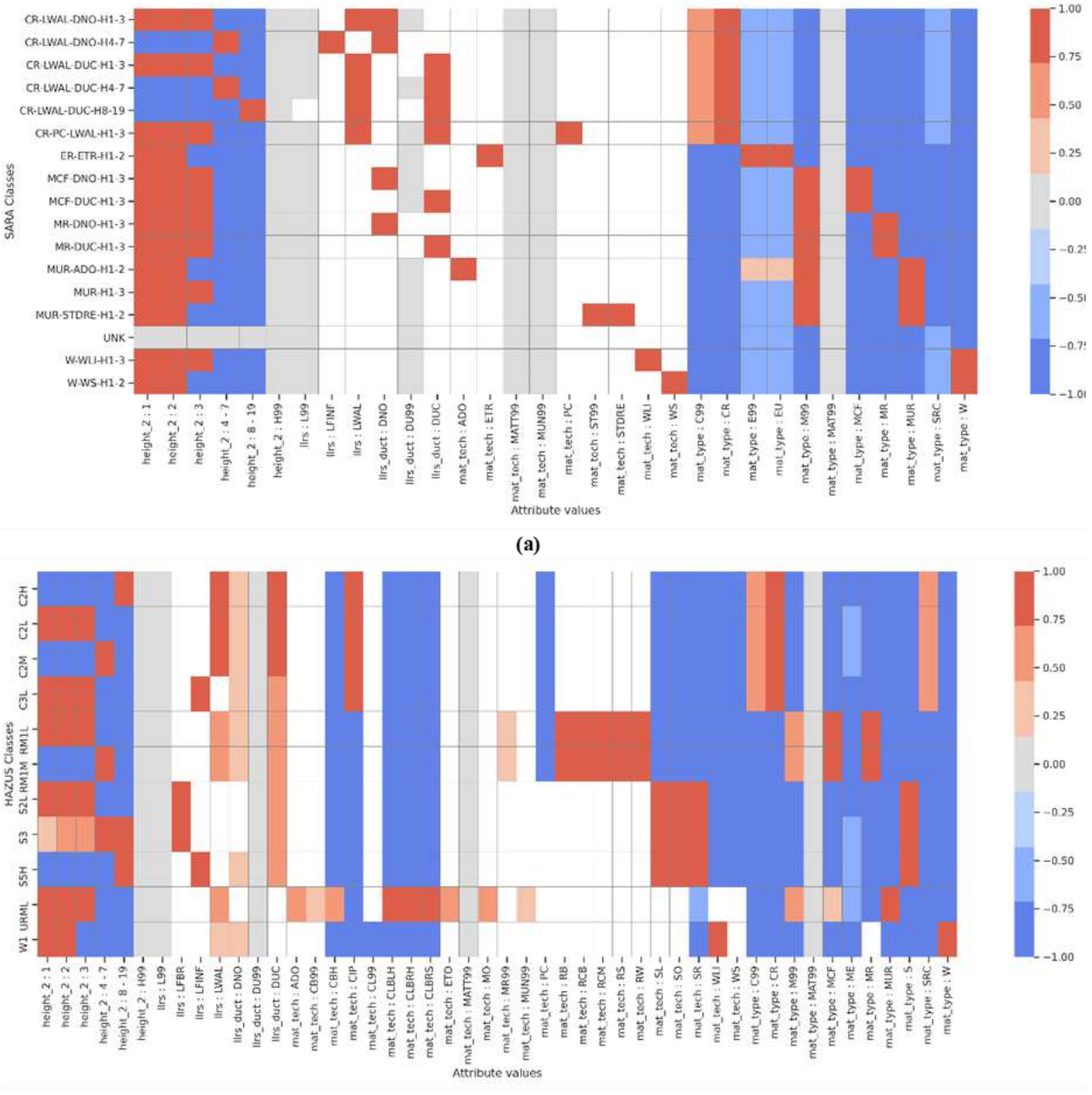


Figure 2. Graphical representation of the (a) SARA and (b) HAZUS building classes scheme definition. The colour encodes the compatibility value with extremes in red and blue that represent high and low compatibilities respectively. Grey: neutral and white indicates no explicit compatibility value has been assigned.

The two sets of weights reported in Table 2 are used to obtain the inter-scheme compatibility matrices in Figure 5 following the method in title 3.2.

	(a)	(b)	(c)
SARA	W-WLI-H1	CR-LWAL-DUC-H8-19	CR-LWAL-DUC-H4-7
HAZUS	W1	C2H	RM1M

	(d)	(e)	(f)	(g)	(h)
SARA	UNK	MUR-STDRE-H1-2 *(W.A-1)*; MUR-H1-3 *(W.A-2)*	UNK	MCF-DNO-H1-3 *(W.A-1)*; MCF-DUC-H1-3 *(W.A.2)*	MR-DNO-H1-3 *(W.A-1)*; MR-DUC-H1-3 *(W.A.2)*
HAZUS	C2H	URML	S3	RM1L	RM1L

Figure 3. Pictures of some buildings' façade surveyed in Valparaiso. Their classification into the SARA and HAZUS schemes is displayed considering the two weight arrangements in Table 2. No distinction is made when both weights led to the same class.

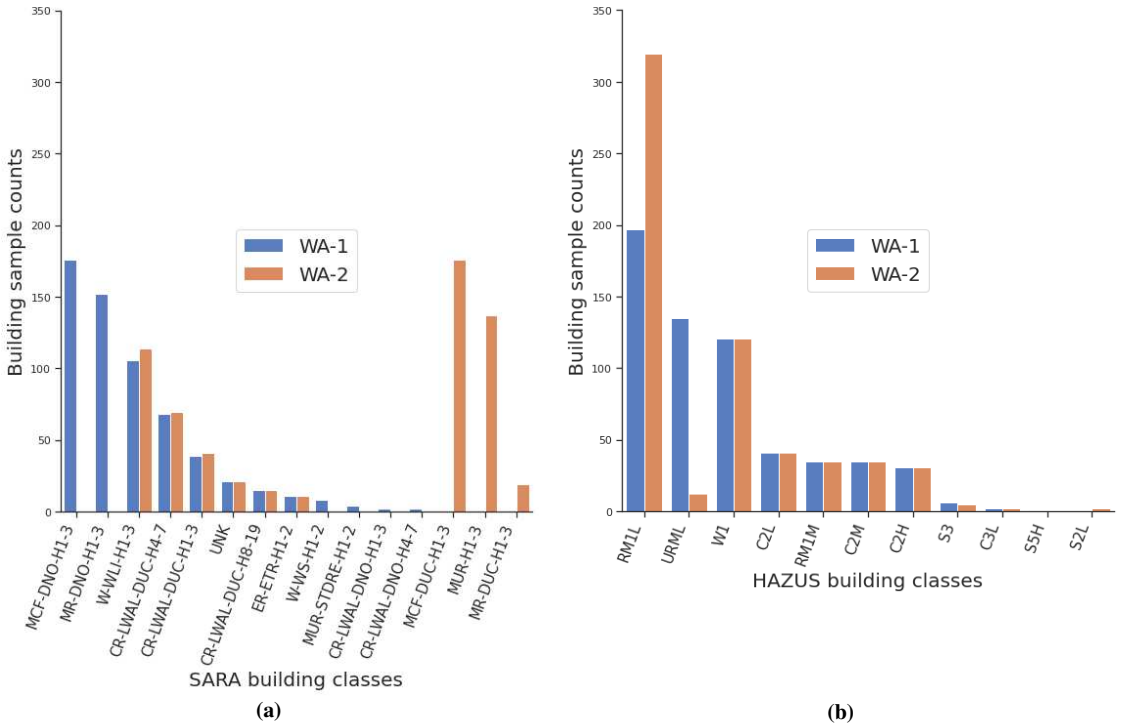


Figure 4. Distributions of the most likely building class of the survey from the compatibility levels between the observed building attributes and the classes of every scheme: (a) SARA and (b) HAZUS for two weight arrangements (WA-1, WA-2) in Table 2.

4.2.2. Prior and posterior distributions

It has been assumed that the definition of expert-based priors has been increasingly superseded by the real data being captured during surveys. Priors have been only considered as: (1) informative prior if the portfolio composition is

assumed from expert elicitation (GEM, 2014) and (2) uninformative prior if the portfolio has identical proportions per class. Since the informative priors have been only assigned to the SARA building classes, and we know the average night-time occupancy, and replacement costs only for that reference scheme, the method to obtain probabilistic inter-scheme compatibility matrices is followed to obtain these quantities for the HAZUS scheme. This is first done through the generation of the possible combinations of attributes values per scheme (see horizontal axis of Figure 2) leading to 71.280 combinations.

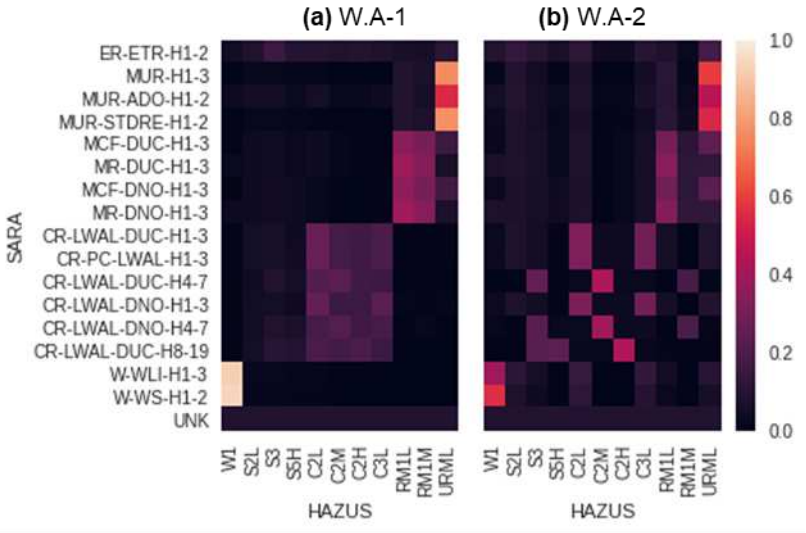


Figure 5. Inter-scheme compatibility matrix for SARA (source) and HAZUS (target) in Valparaiso through the weight arrangements: W.A-1 (a) and W.A-2 (b) reported in Table 2.

Table 3. Building classes in terms of the HAZUS scheme proposed for Valparaiso with the respective intensity measure of their associated fragility functions (as reported in FEMA 2012); together with the prior proportions per building class; obtained from the inter-scheme compatibility matrices shown in Figure 5. Average night-time residents (Res /bdg.) and replacement cost (Repl. Cost (USD/bdg.) have been assumed to be the same as the highest compatibility score respect to a class of SARA. In the description H means the range of number of storeys.

			W.A-1			**W.A-2**		
HAZUS building class	**Description**	**IM**	**Prior prop.**	**Res. /bdg.**	**Repl. Cost (USD/bdg)**	**Prior prop.**	**Res. /bdg.**	**Repl. Cost (USD/bdg.)**
W1	Wood,light frame < 5000 sq. ft.	PGA	0.384	4	43750	0.204	4	43750
S2L	Steel braced frame, H:1,3	PGA	0.039	15	336000	0.091	15	336000
S3	Steel light frame, any H.	PGA	0.044	54	1260000	0.061	54	1260000
S5H	Steel frame, unreinforced masonry infill walls, high rise	PGA	0.033	173	4032000	0.026	173	4032000
C2L	RC shear walls, H:1,3	PGA	0.046	18	336000	0.114	18	336000
C2M	RC shear walls, H:4,7	PGA	0.028	54	1260000	0.022	54	1260000
C2H	RC shear walls, high rise	PGA	0.027	173	4032000	0.021	173	4032000
C3L	RC frame buildings, unreinforced masonry infill walls, H:1,3	PGA	0.033	14	288000	0.100	15	336000
RM1L	Reinforced masonry walls; wood or metal deck diaphragms, H:1,3	PGA	0.112	18	360000	0.120	18	420000
RM1M	Reinforced masonry walls; wood or metal deck diaphragms, H:4,7	PGA	0.094	18	360000	0.046	54	1080000
URML	Unreinforced masonry bearing walls, H:1,3	PGA	0.160	5	43750	0.196	6	52500
			∑ 1.0			**∑ 1.0**		

We have obtained two sets informative prior for HAZUS (see Table 3). The replacement costs and night-time residents per HAZUS class have been assumed to be the same as the SARA class for which the largest compatibility value was obtained. Identical values of replacement costs and residents are obtained for both W.A-1 and W.A-2, except for the classes C3L, RM1L, RM1M and URML (Table 3) for which different highest compatibility values respect to the SARA classes were obtained. Once we have the prior and likelihood terms for every building class in both schemes, the posterior distributions are constructed following Eq. 3. They are shown in Figure 6 where the proportions of the four posteriors sums up to 1.0. We note that W.A-1 associates to the HAZUS class URML the largest compatibility (77%) with MUR-STDRE-H1-2 in SARA, while the W.A-2 does it for a 0.60% with MUR-H-1-3. The latter leads to a larger replacement cost. The presence of the URML class is high in the surveys in Valparaiso (Figure 4) and due to the large informative prior value, its presence in the posteriors is also large, as observed for W-2 in Figure 6). Since the assignment of priors through expert elicitation has been at the commune level (GEM, 2014), the posteriors obtained are up-scaled to represent the building composition in Valparaiso.

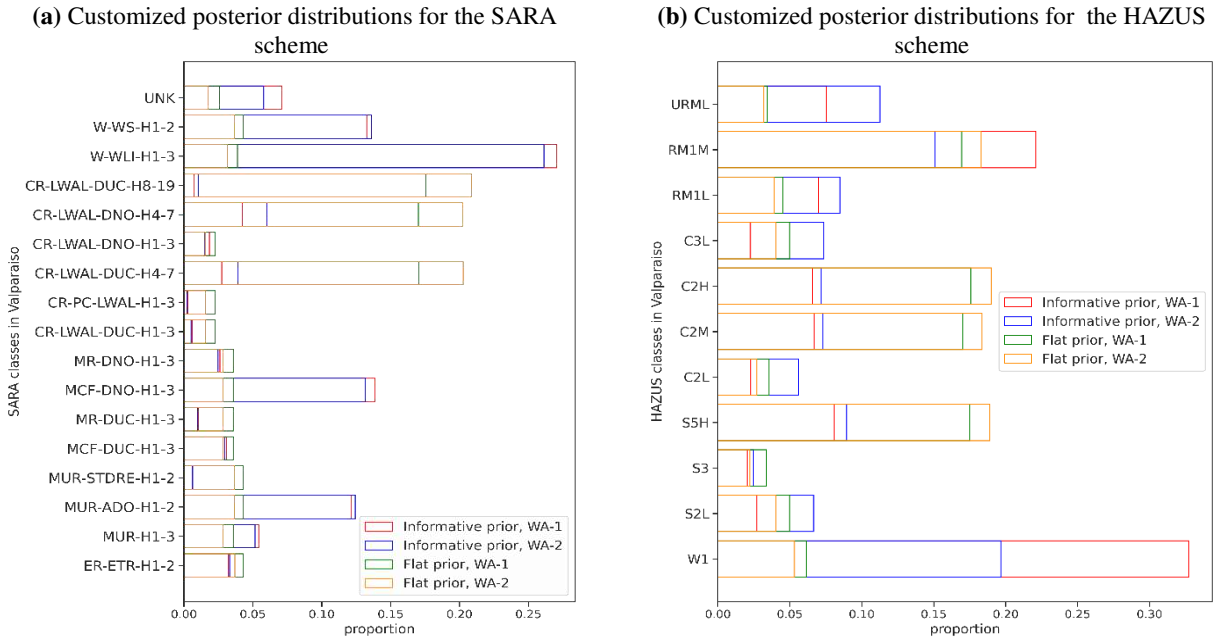


Figure 6. Posterior distributions obtained for the (a) SARA and (b) HAZUS schemes for the 604 inspected building surveyed in Valparaiso while considering different weight arrangements (WA) and flat and informative priors.

4.2.3. Logic tree construction; synthetic building portfolio construction and allocation

The four posterior Dirichlet distributions obtained for every scheme are depicted in Figure 6. Each of them are sampled varying three selected parametrized concentration factors as follows: α_{0_1}= 1.0; α_{0_2}= 15 and α_{0_3}= 50 to generate 300 synthetic building portfolios in every case. α_{0_1} describes three different degree of trust in the assumptions that lead to the construction of the posterior distribution: very low, moderate, very high respectively. They can also be interpreted as 300 different criteria (e.g. a pool of virtual experts) regarding the portfolio composition: very divergent opinions; moderately similar; and very similar compositions respect to the proportions in the customized posteriors. Twelve scenarios in terms of the exposure composition are investigated per every considered scheme. A customized logic tree with five levels is ultimately constructed as shown in Figure 7.

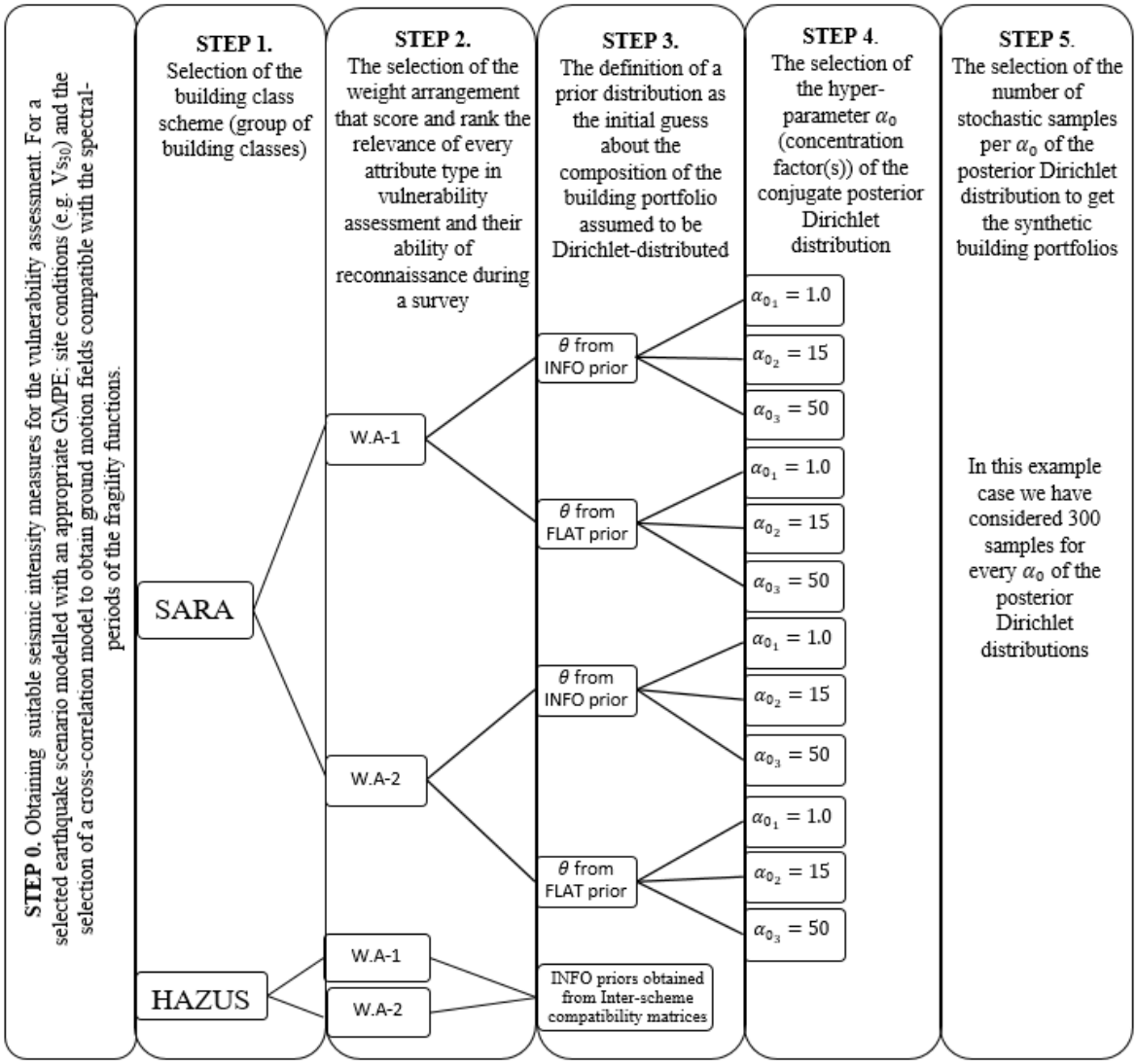


Figure 7. Logic tree with five levels constructed explore the impact of the building exposure composition modelling in Valparaiso. The twelve branches result from every considered scheme. The entire one is shown only for SARA. The same procedure has been carried out for HAZUS.

The gridded population product GPWv4 with a 30 arc-second grid resolution model and population projections for 2020 (CIESIN, 2018) was used to perform a dasymetric disaggregation based on population as described in (Pittore et al. 2020). With this process we can spatially locate all the synthetic buildings portfolios obtained from every case displayed in Figure 7 using the occupancy (residents per building class) of the SARA (Table 1) and HAZUS schemes (Table 3). Despite that the population is the same, all the synthetic portfolios have varying proportions per building class, each with different occupancies that result in very different estimates of the total number of buildings expected in the location. The corresponding building count values to the single observation assuming equally composed portfolios and with the expert criteria defined by GEM, (2014) for SARA and using the compatibility matrix for HAZUS are shown on the first column of Figure 8. Since the number of residents in every HAZUS building class is different for WA-1 and WA-2, the building counts vary accordingly. However, from the top-down approach, the building counts in Valparaiso are almost identical regardless the scheme implemented. The plots in the other four columns display the associated variabilities in the estimations of the total number of buildings for the synthetic portfolios constructed with the 300 samples.

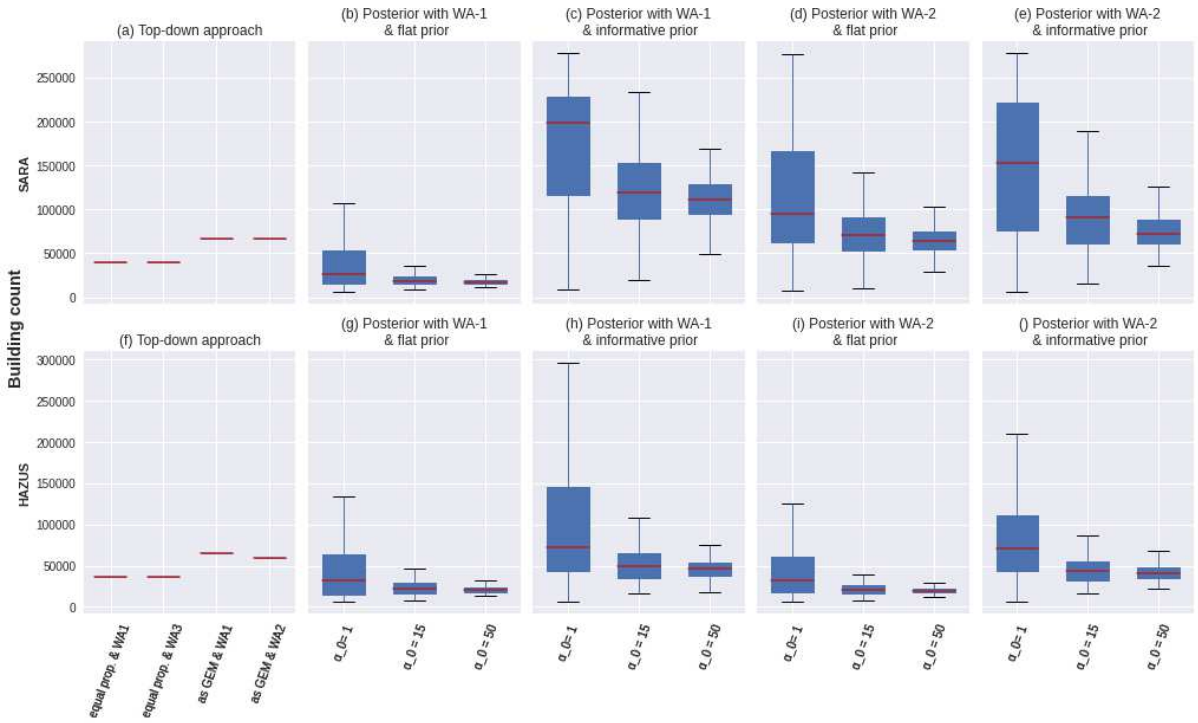


Figure 8. Dependency between the variance of the building counts (from population extraction and occupancy) and the portfolio composition. The first column comprises a single value in which the weight arrangement WA did not impact SARA, contrary to what occurs in HAZUS. The distributions on the other columns corresponds to every case in the logic tree depicted in Figure 7.

Similarities between the estimates from the single top-down single approach and the respective mean building counts from every synthetic portfolio are more evident in the HAZUS scheme whilst larger differences appears in SARA that is more sensitive to the prior definition and weight arrangements. However, for both schemes these similarities increase when the WA-2 is addressed together with informative priors (Figure 8-e, j). On one hand, the selection of W.A-2 imposes a larger amount of observations of ductile buildings which have a larger number of residents associated (Table 1 and Table 3), and therefore there is a consequent reduction in the variance of their respective building counts distributions. On the other hand, when α_0 =50 is addressed, (the similarities between the means are reduced, but the value from the top-down approach stays within the range of associated variation. Instead of having a fixed building count value, we are defining a suitable range of values whose variation is consistent with prior assumptions and observations.

4.3. Scenario-based earthquake-rupture and ground motion distribution

The earthquake rupture with a Mw 8.2 similar to the 1906 Valparaiso event is used throughout this example. Given the lack of instrumentation during the occurrence of that earthquake, its exact location and other parameters are uncertain. Nevertheless, some basic ones have been used to generate a finite fault model making use of the OpenQuake Engine (Pagani et al. 2014) as displayed in Figure 10. In order to model the distribution of the resulting ground motion values, the suitable GMPE for the Chilean inter- plate subduction area reported in Montalva et al. (2017) has been considered. For the GMPE's "site" term, the Vs_{30} values from slope as proposed by the USGS (Allen and Wald, 2007) have been used over the selected grid, and replaced when it was possible by the values obtained from a local seismic microzonation reported in Mendoza et al. (2018). The latter work performed in-situ seismic refraction and HVSR (Nakamura, 1989) instrumental campaigns in Valparaiso. The final Vs_{30} gridded values are displayed in Figure 9-b.

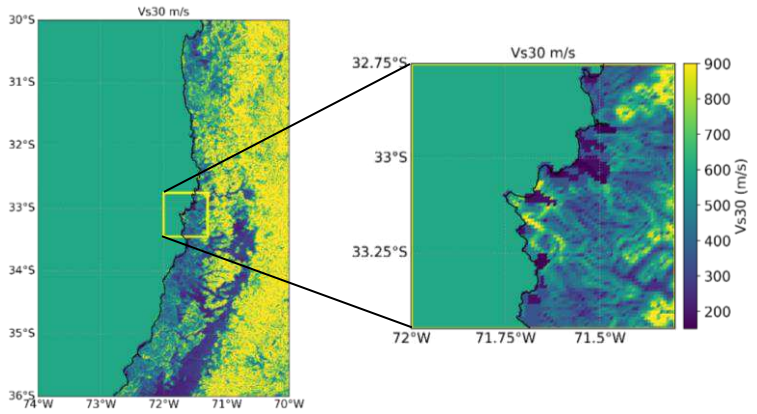


Figure 9. Distribution of the Vs_{30} values in Chile from slope as proposed by the USGS, and refined within the study area with the seismic microzonation reported in Mendoza et al. 2018.

Ground motion fields for pseudo-spectral acceleration values at the spectral periods: PGA; S.A(0.3 s); S.A(1.0 s) are selected and allow consistency with the required IM by the fragility functions of the SARA scheme as reported in Villar- Vega et al. (2017). The aleatory uncertainty in the simulated ground motions has been addressed with 1,000 realisations in two cases: with the Markhvida et al. (2018) cross-correlation model and without any ground motion correlation. The median values obtained from the Mw 8.2 scenario for one realisation over the three spectral periods is shown in Figure 10 with a detail of Valparaiso with cross-correlated ground motion fields.

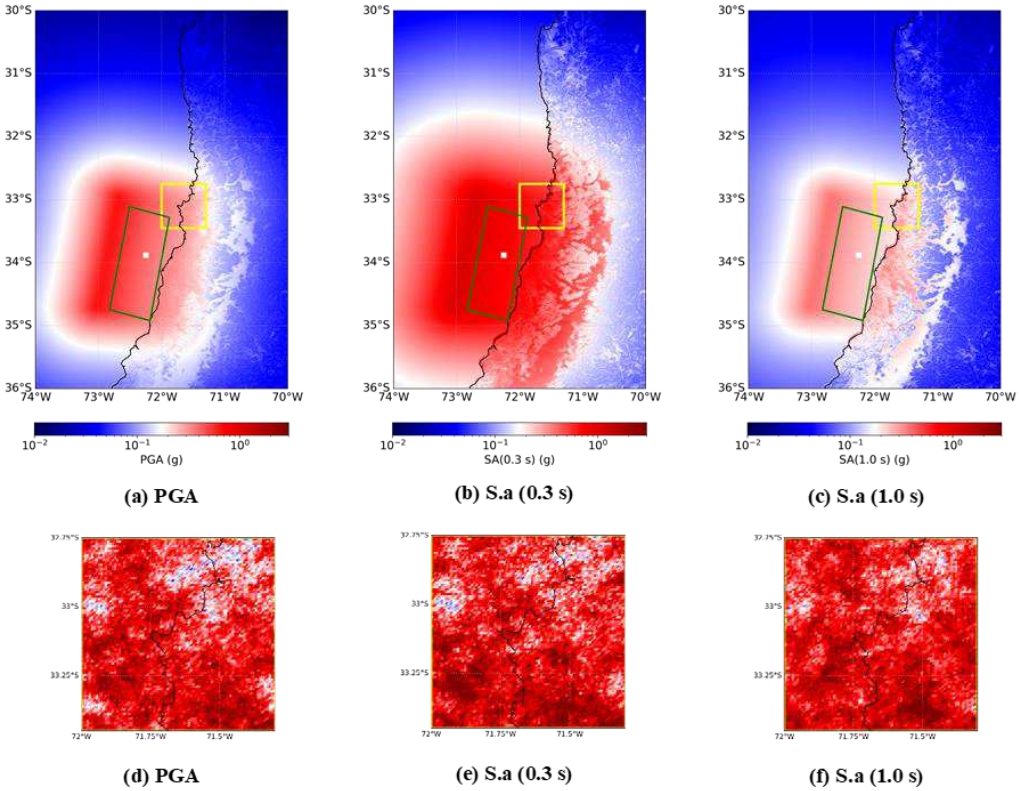


Figure 10. First row: distributed median values obtained from a single realisation of the Mw 8.2 earthquake rupture for three spectral periods (a) PGA, (b) S.a (0.3 s) and (c) S.a (1.0 s) using the Montalva et al. 2017 GMPE. The earthquake hypocentre is depicted by a white square. The rupture plane is displayed by a green rectangle. The study area is shown by a yellow square Second row: zoom of the study area with a single realisation of the cross-correlated seismic ground motion fields for their correspondent considered spectral periods using the Markhvida et al. 2018 model.

4.4. Epistemic uncertainty exploration of the building exposure definition in earthquake loss models

An initial exploration of the epistemic uncertainty in the building portfolio composition has been carried out computing the absolute losses for the selected earthquake scenario (Mw 8.2) considering the four sets of posterior distributions obtained for the SARA and HAZUS schemes as reported in Figure 6 and only for the mean values of building counts shown in Figure 8. These customized posteriors have embedded the use of the priors and likelihood terms constructed with the surveyed data. Their respective losses have been further compared with the losses from equally composed portfolios (same proportions per class) and with a portfolio composition as defined by the informative criteria through the joint expert elicitation at the Commune level (GEM, 2014) and the use of a top-down approach using the GEM mapping-schemas over census data (Yepes-Estrada et al. 2017). The loss estimates are obtained with the corresponding fragility function, the required ground motion field at their pseudo-spectral period (with and without correlation) and their economical consequence models. This first comparison is reported in Figure 11. The losses for the HAZUS scheme are calculated considering with two code compliances: low and moderate levels. It can be seen that the absolute losses obtained from SARA are drastically lower than the values obtained when any of the two sets of fragilities by code provisions of HAZUS is implemented.

Figure 11 shows that the overall variability in the losses is imposed by the ground motion cross-correlation model. However, the mean loss values in all the cases are controlled by the selection of the portfolio composition. The role of selecting the prior distribution has a greater impact than the weight- arrangement of the likelihood term. Comparatively lower loss values are obtained from the top-down and expert elicitation in both schemes (first row) than when evidence (i.e. surveyed buildings) are integrated into the posterior distributions, except in the cases when a flat prior is addressed in the moderate seismic code of HAZUS (plots “o” and “q” in Figure 11). The code-level in HAZUS seems to increase almost linearly the losses when comparing the second row with the respective cases of the third row of the figure.

Subsequently we present the results obtained from an exhaustive seismic risk assessment over the 300 synthetic portfolios obtained for each of the 12 customized posteriors per scheme as proposed in the logic tree presented in Figure 7. This is done for two independent schemes and 2.000 ground motion realisations, thus, we have performed $2 \times 12 \times 300 \times 2.000 = 14.400.000$ stochastic risk computations for a single seismic rupture. Loss exceedance curves (LEC) obtained from the earthquake scenario considering the SARA and HAZUS schemes are illustrated in Figure 12 and Figure 13 respectively. There are two sets of curves that have been included throughout all the plots: (1) a pair curves that represent the losses obtained from a unique composition according to the joint expert elicitation at the Commune level (GEM, 2014) and the use of a top-down approach using the GEM mapping-schemas over census data (Yepes-Estrada et al. 2017). They are coloured in purple when the seismic ground motion was modelled with a cross-correlation model and in red when it was not addressed; and (2) a pair of curves generated foreseeing the portfolio with equally composed proportions which are represented by black curves either they the cross correlated model has been included or not. These curves have a smoother shape and shorter initial plateau when correlation was addressed contrasting with sharper curves when correlation was not included. A third set of curves that represent each of the four correspondent customized posterior distributions (as reported in Figure 6). The same sets of curves are represented in every row by non-continuous white curves. Thus, 606 loss curves are displayed in every subplot.

A decision to normalize all of the LEC respect to the maximum value obtained without intra-event residuals’ cross-correlation has been taken in order to compare the differences of the loss outputs in percentages. This decision is supported in the work of Vamvatsikos et al. (2010) who argues that, due to the continuously evolving exposure in time and location, erroneous physical damage predictions can arise especially if the losses are absolute instead of normalised. Therefore, to compare the losses of the synthetic portfolios, the metric is: normalized number of losses. It should be noticed that the axis is not identical for SARA (Figure 12) and HAZUS (Figure 13) and the horizontal axis in every plot related to HAZUS starts at 0.7 to graphically highlight some differences.

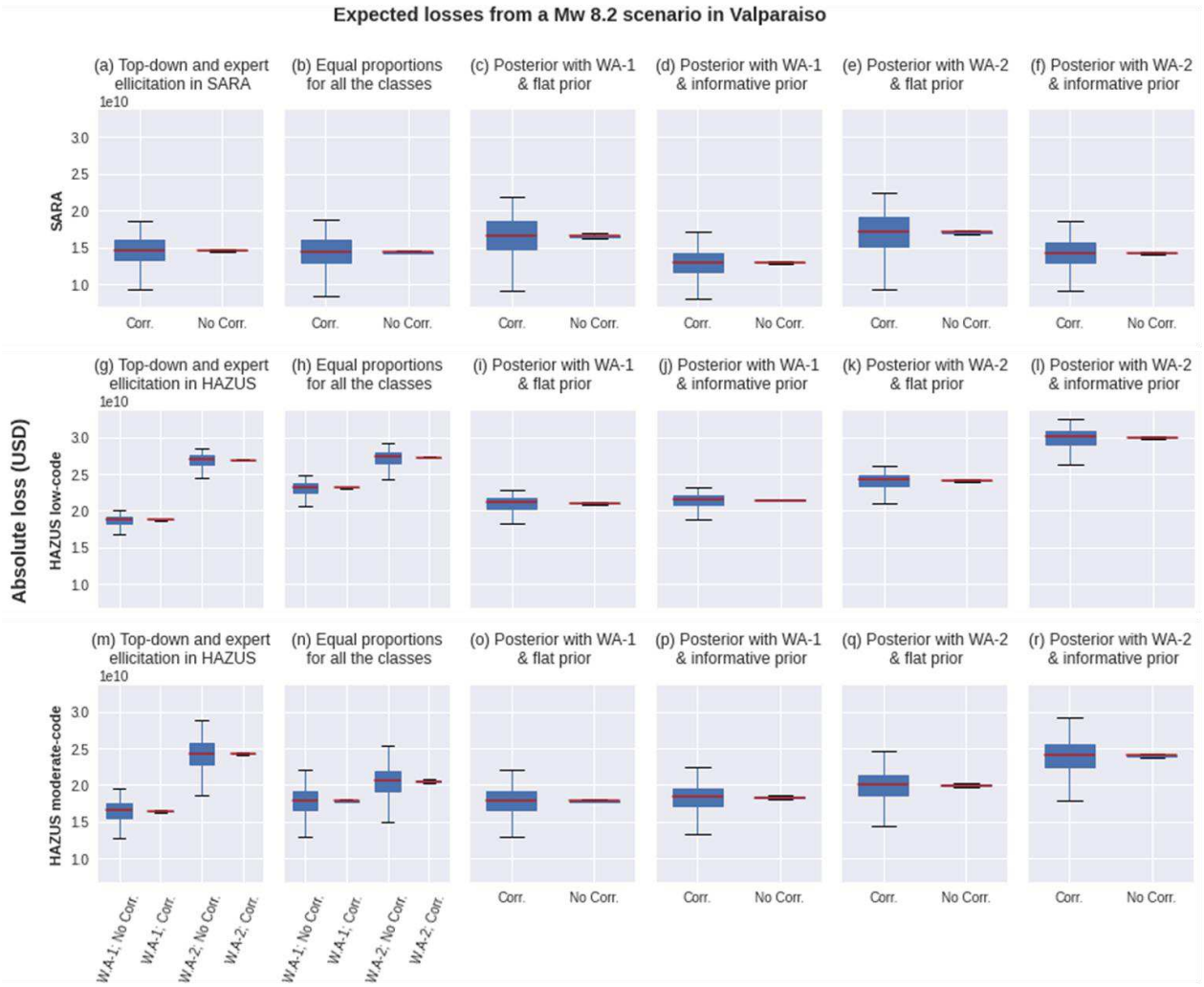


Figure 11. Comparison of the absolute loss (USD) considering SARA (first row) and HAZUS with varying code-compliance level (low code in second row; and moderate code in third row). The losses are obtained after the earthquake scenario Mw 8.2 in the residential building stock in Valparaiso using the cross-correlation model (Corr.) and without it (No Corr.). Different portfolio compositions are considered: the top-down vision and expert elicitation (first plot in every row); equally composed portfolios (second plot in every row) and the customized posterior distributions for both schemes as shown in Figure 6.

The 300 loss curves (per α_0) obtained from every synthetic portfolio are represented by blue curves when the correlation model was addressed and by yellow curves when it was not. Regardless the building portfolio composition, when the cross-correlation model is accounted, their corresponding LEC show a greater variability. This is not particularly surprising and it is aligned with the findings discussed in Crowley et al. (2008) related to the log-normal distribution of the seismic ground motion induced by the Montalva et al. (2017) GMPE. When the cross-correlation is considered in the HAZUS scheme, the variability in the normalized loss values is generally low, and only start to be perceptible, for greater probabilities, at significantly large values of the normalised metric (i.e. 0.7) as formerly explained. This lower variability is a contribution of the HAZUS fragility functions which account for the structural demand of a single ground motion intensity (PGA). These features were formerly noted in Luco and Cornell, (2007).

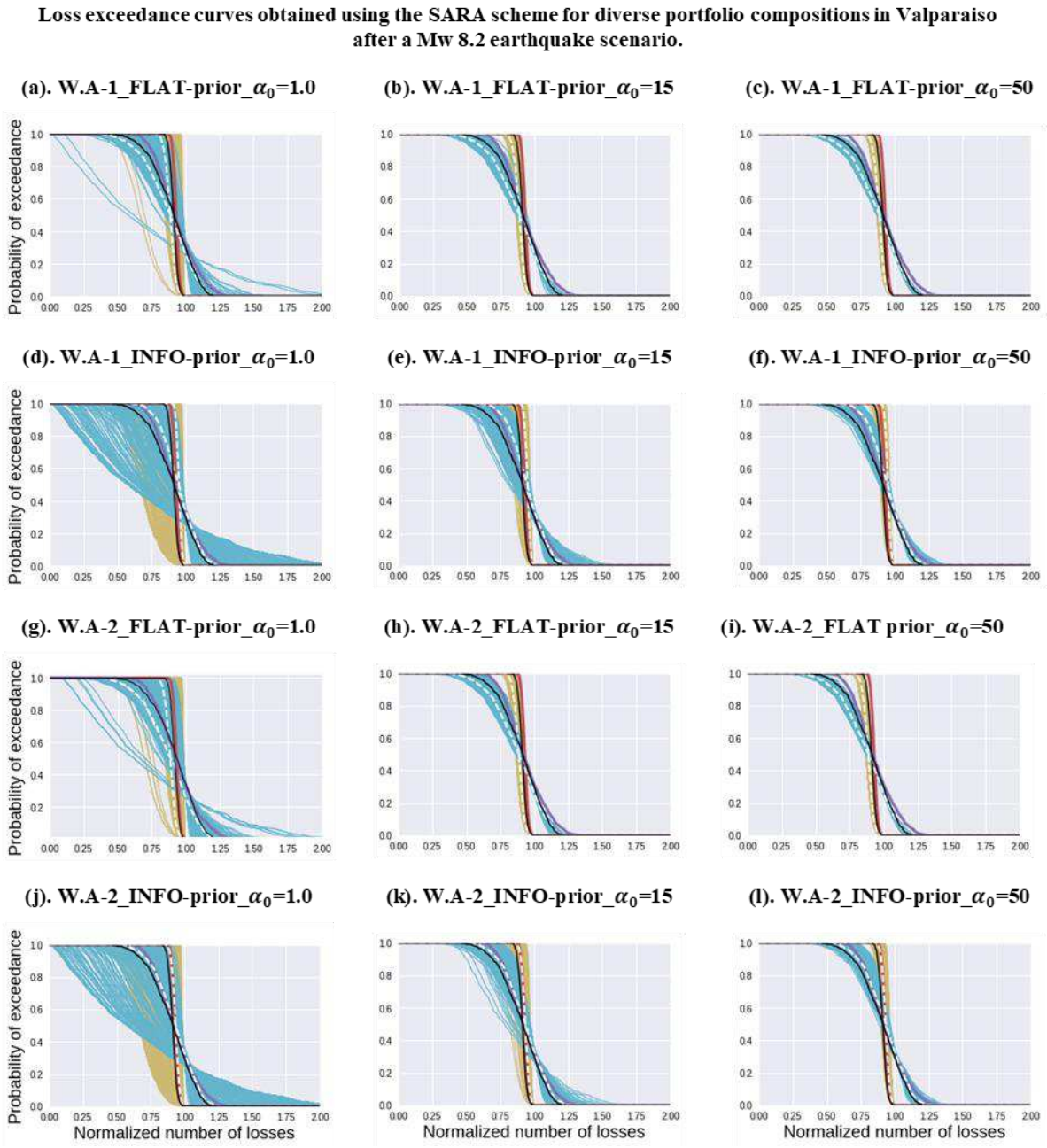


Figure 12. Normalized loss exceedance curves (LEC) for different exposure compositions as depicted in the logic tree (Figure 7) for the SARA scheme. Curves are displayed in blue if the ground motion cross-correlation model was addressed and in yellow with no ground motion cross-correlation. In every plot, the LEC from their respective customized posterior distributions are displayed in non- continours white curves. The single composition vision as informative priors (as defined by GEM) are depicted in purple when the ground motion cross-correlation model was addressed and in red with no ground motion cross-correlation. Black curves represent the losses portfolio whose composition is assumed to be have equal proportions for all of the classes. The black curve with a smoother shape is obtained while modeling cross-correlation in oposition to the sharper one when it was not.

The selection of the weights which score and integrate the taxonomic attributes types, and that together with the collected data evidence construct the likelihood term, do not show a relative large impact in the resultant LEC for the SARA scheme. This could be inferred from Figure 6 where the weight arrangement did not impose a large difference

as the prior definition did. However, these weights also impact the construction of the inter-scheme compatibility matrices (Figure 5) which are used to obtain the informative prior proportions for HAZUS to ultimately create the posteriors. Therefore, the weight arrangement selection also impacts the loss estimates. This is perceptible in the relocation in the proportions as shown in Figure 4 with diverse replacement costs for HAZUS (see Table 3) as well as the larger seismic vulnerability of the W.A-1 in comparison to W.A-2 that, due to the larger value assigned to the ductility level (see Table 2) leads to non- ductile building classes (Figure 4). For the W.A-1 driven classes, lower mean acceleration values (of their respective fragility functions) are needed to reach the same damage state (e.g. MCF-DNO-H1-3 vs. MCF-DUC-H1-3 in SARA; and URML vs. RM1L in HAZUS).

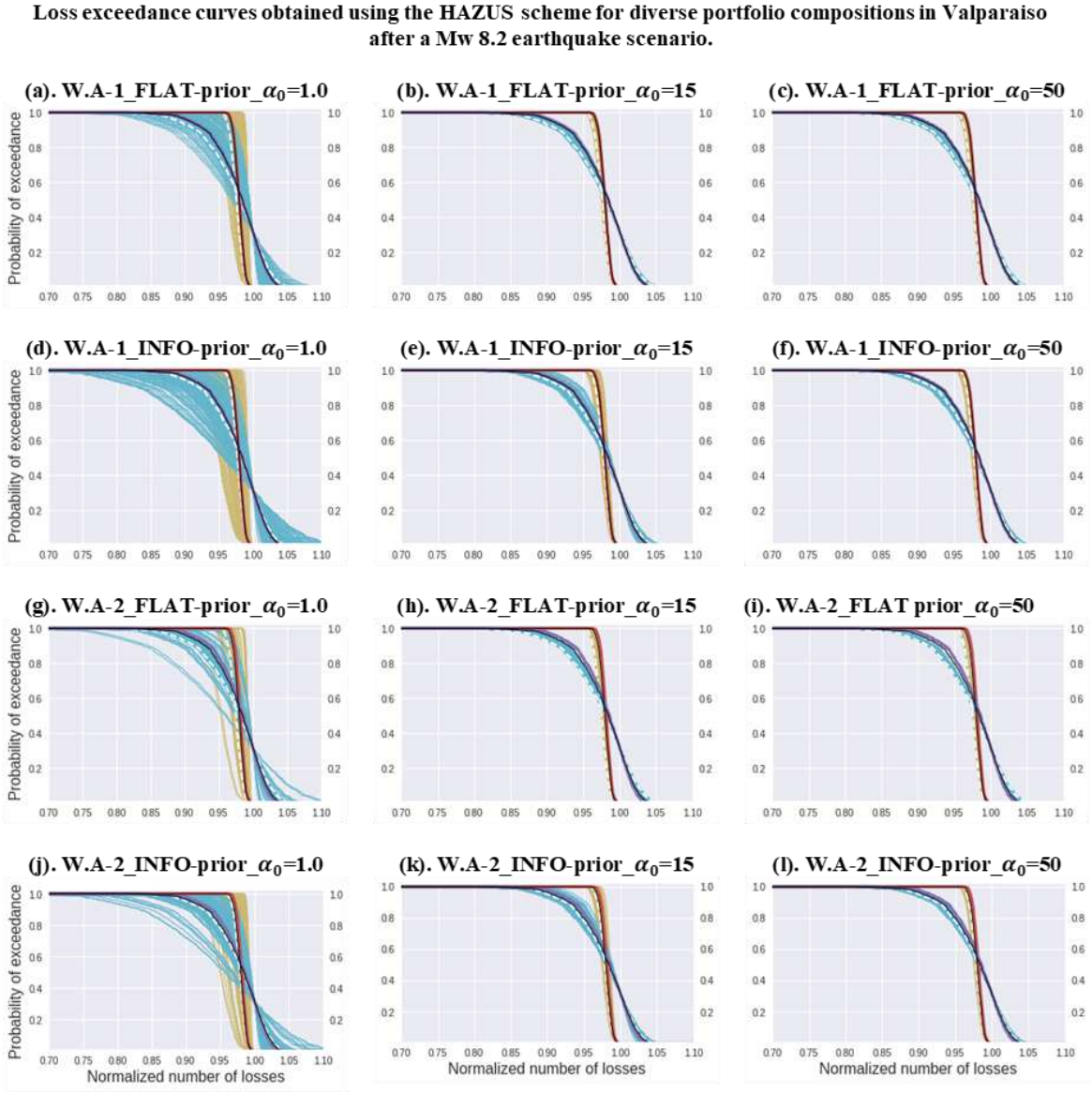


Figure 13. Normalized loss exceedance curves (LEC) for different exposure compositions as depicted in the logic tree (Figure 7) for the HAZUS scheme. Curves are displayed in blue if the ground motion cross-correlation model was addressed and in yellow with no ground motion cross-correlation. In every plot, the LEC from their respective customized posterior distributions are displayed in non- continours white curves. The single composition vision as informative priors (from the inter-scheme conversion matrices in Figure 5) are depicted in purple when the ground motion cross-correlation model was addressed and in red with no ground motion cross-correlation. Black curves represent the losses of a portfolio whose composition is assumed to be have equal proportions for the classes. The black curve with a smoother shape is obtained while modeling cross-correlation in oposition to the sharper one when it was not.

The plots in the first column of Figure 12 (SARA) and Figure 13 (HAZUS) show a greater dispersion (and therefore uncertainty in the results) when α_{0_1}=1.0 is addressed for a lower degree of trust in the portfolio compositions that emulates an "increased spread" around mean values, given by their respective posterior distribution which is strongly related to the prior through the conjugacy with the likelihood term. This variability decreases for $\alpha_{0_2} = 15$ (intermediate degree of trust) in the second column, being the lowest when $\alpha_{0_3} = 50$ is considered (third column) that emulates very similar compositions as the customized posterior distributions.

The loss curves obtained from the posteriors constructed with flat prior (non-continuous white curves in first and third rows) are in all the cases leading to lower values in comparison with their respective black curves (equally composed portfolios) and from the solely top-down vision using the GEM mapping-schemes (purple and red curves). This is due to the incorporation of evidence (surveys) in the construction of the likelihood term. The opposite is observed when informative priors are addressed. Posteriors created with flat priors tend to concentrate lower loss values at the lowest probabilities of exceedance (p.o.e) and greater values at the largest p.o.e, being more evident when $\alpha_0 = 1.0$. This is due the fact that flat priors largely impact the creation of the posterior as observed in Figure 6. When larger proportions are relocated to classes with relatively big replacement cost values, more losses are expected (Figure 11-c, d). This is for instance perceptible for the SARA classes CR-LWAL-DNO-H4-7 and CR-LWAL-DUC-H4-7, that despite their different seismic vulnerability, both of them have comparatively large replacement costs (Table 3).

In SARA the difference between the losses obtained with and without correlation (both non-continuous white curves) is around 25%. Nevertheless, when we take a look at the stochastically constructed portfolios, for a similar α_0 (e.g. $\alpha_{0_2} = 15$) we realise that, the largest difference between the loss curves with cross-correlation (blue ones) and without (yellow ones) is still 25% from the posterior constructed with a flat prior, but it reaches at least a 60% from the posterior with an informative prior. However, if we do a similar comparison over the HAZUS scheme, this difference is almost negligible. Therefore, we cannot state this is only due to the prior definition. The greater dispersion in SARA's LEC might be due to by the joint effect between the prior definition and the variability induced by the cross-correlation model for three spectral periods. In this sense, relatively homogenous building compositions obtained from posterior constructed with a flat prior, would require similar spectral periods and thus a lower variability in the losses would be expected because no cross-correlation model would be needed. This is consistent with the fact that no cross-correlation is required by HAZUS since its fragility functions only use PGA as IM. This is acknowledging that the spatial variation in the ground motion places a lower limit on the uncertainty in the loss estimates that cannot be entirely reduced as described in Bal et al. (2010), even when the composition of the building portfolio is well known (i.e. $\alpha_{0_3} = 50$).

5. Discussion

Suitable prior distributions for the building portfolio composition could be also obtained by drawing several realisations varying the concentration factor α_0 of its associated Dirichlet distribution. This would allow to acknowledge the degree of belief in the prior, as for instance done by a pool of experts with similar or diverse opinions for even smaller sectors within a large study area. This is aligned with O'Hagan et al. (2006) and Beven et al. (2018) who proposed to give the expert elicitation scientific rigour by providing empirical control (i.e. surveys) on how well the experts' opinions works. This would update the initial assumptions by the "mapping-schemas" (title 2.1) on the census co-variants of the top-down exposure modelling while exploiting emerging technologies such as remote sensing and image reconnaissance for the definition of prior distributions. In this framework, the expert-based priors will be increasingly superseded by real data captured during surveys. Hence, the posteriors will be dominated by the likelihood term, assigning a long-term importance to the data assimilation within statistical exposure models.

For the sake of simplicity in the illustration of the proposed approach we did consider only some of the uncertainties in the parameters employed. Number inhabitants and replacement cost per building class, for instance, should be also considered as random variables to create a more comprehensive probabilistic exposure model. This is particularly important for absolute losses that are significantly influenced (through the estimation of total number of buildings) by these parameters. This would overcome the drawback of having static exposure models which depend on the availability of the aggregated data (i.e. census). The combination of probabilistic exposure models with individual observations for a data-driven prior-definition is worth to explore.

The description of building classes in terms of taxonomic attribute types has been shown to be instrumental to identify the most likely building class in a scheme, and are used by the likelihood term and in the construction of the posterior distributions. This description is also input for the novel probabilistic inter-scheme conversion. This latter approach is useful to obtain exposure descriptors (i.e. number of buildings belonging to a certain class, their expected nigh-time residents and replacement cost) under another reference (target) scheme. However it is important to remark that, as both schemes should reflect local construction practices and are attached to fragility functions for local IM, these kind of matrices are only applicable only for typical residential buildings that compose large-scale exposure models.

We have found that the contribution to the overall variance to the losses induced by the ground motion cross-correlation model is lower than the one induced by the level of knowledge of the building portfolio composition. The selection the correlation/cross-correlation model has further impacts that have been not considered herein. Their associated epistemic uncertainty are discussed in Weatherill et al. (2015) and state of the art correlations models need to be also tested. Similarly, the cost ratios per damage state within the economic consequence model for vulnerability assessment is also subject to uncertainties (e.g. Martins et al. 2016; Corbane et al. 2017) that have been neither considered herein. To analyse their contribution, the logic tree proposed in this work could be suitably extended. However, it is worth to effectively rank them once post-damage buildings and macroseismic intensity estimations are available for specific study areas in order to validate the related underlying assumptions.

The results and comparisons have been presented with the aim of discussing the role of uncertainty in the exposure model definition, without arguing whether one particular scheme (SARA or HAZUS) is better suited for representing the exposure model or the expected losses in Valparaiso. The degree of variability obtained by exploring the underlying uncertainty and the resulting differences in the loss estimation are rather calling for caution when selecting "foreign" exposure and vulnerability models for loss estimates and urban planning. It highlights the importance of constructing local building exposure models together with local fragility functions designed from local strong ground motion records. This also acknowledges that besides the hazard- consistency of the seismic fragility functions (e.g. Kohrangi et al. 2017; Sousa et al. 2018), a clear building exposure definition of building classes based on local construction practices is also highly relevant in seismic risk practices.

6. Conclusions

An exploratory Bayesian framework to study the epistemic uncertainty in seismic risk loss estimates associated with the probabilistic nature of the building exposure model has been presented. A sensitivity analysis was performed based on a logic tree approach considering five modelling components : 1) the risk-oriented building class schemes (set of MECE classes); 2) the weights patterns to select and rank the relevance of single attribute types in estimating compatibility and assigning class labels; 3) the prior distribution describing the prior information about the composition of the building portfolio assumed to be Dirichlet-distributed; (4) the concentration factor α_0 of the posterior Dirichlet distribution.

The concentration factor α_0 of the Dirichlet distribution describes the extent of trust in the building portfolio initial assumptions and can also be interpreted as different opinions respect to the initial portfolio composition (e.g. very

similar ($\alpha_0 \sim 50$) or "increased spread" criteria ($\alpha_0 \sim 1$)). In this work we have sampled the posterior distribution that is itself strongly related to the prior through the conjugacy with multinomial. As a matter of fact, prior and posterior are two Dirichlet distribution which differ only by the empirical contribution from the frequency observed in the field, but the concentration factor comes from the prior since it is the factor that "weights" and adjust the prior proportions to the observed frequencies.

This work has confirmed the observations of former works regarding that the spatial variation in the ground motion places a lower limit on the uncertainty in the final loss estimates that cannot be entirely reduced even when the composition of the building portfolio is very well known. This variability increases when the building portfolio is heterogeneous and needs to be modelled using fragility functions for diverse spectral periods. We have shown that the contribution of the seismic cross-correlation model to the overall loss variance in scenario-based earthquakes is generally lower than that induced by the level of knowledge of the building portfolio composition that controls the mean loss values. This is shown for all the alternative portfolio compositions obtained through a probabilistic method that include prior assumptions and empirical evidence (i.e. building surveys).

This work has also highlighted that considering the presence of certain building attributes within a probabilistic exposure model as part of a reducible epistemic uncertainty framework, instead of purely aleatory, is better linking the attribute reconnaissance with the actual building classes and associated fragility functions for loss estimates. This is important when these attributes drive the seismic vulnerability of the building (e.g. LLRS ductility). This is emphasising that well-structured standards for data collection based on a faceted taxonomy would lead to better practices in the exposure model for seismic risk modelling. Expert-elicited models used in top-down approaches, even when carefully crafted, often provide only a partial perspective on the real composition of the building portfolio, and do not easily allow for a proper understanding and analytic modelling of the underlying epistemic and aleatory uncertainties. An exposure data collection based on vulnerability-independent standard taxonomies, combined with a sound probabilistic modelling is instrumental for more robust risk assessment practices.

Acknowledgments We would like to thank the structural engineers from CIGIDEN for having carried out the RRVS building surveys and to Yvonne Merino (PUC) and Patrick Aravena (DLR) for having supervised it. We thank Catalina Yepes (GEM) for having provided the SARA model. Thanks to Dr. Esteban Sáez (PUC) for providing the seismic microzonation of Valparaiso. Our gratitude to Dr. Graeme Wetherill, Nils Brinckmann, Matthias Rüster and Dr. Jörn Lauterjung (GFZ) for their advice and support.

Funding The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: J.C. Gomez-Zapata, M. Pittore and S. Shinde from the RIESGOS project, funded by the German Federal Ministry of Education and Research (BMBF) Grant No. 03G0876, as part of the funding programme CLIENT II – International Partnerships for Sustainable Innovations'. H. Lilienkamp has been funded by the Helmholtz Einstein International Berlin Research School in Data Science (HEIBRiDS). P. Aguirre and H. Santa Maria have been funded by the Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017 and ANID-FONDECYT 1191543.

Declarations The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

References

Acevedo, A.B., Yepes-Estrada, C., González, D., Silva, V., Mora, M., Arcila, M., Posada, G., 2020. Seismic risk assessment for the residential buildings of the major three cities in Colombia: Bogotá, Medellín, and Cali. Earthquake Spectra 8755293020942537. https://doi.org/10.1177/8755293020942537

Aguirre, P., Vásquez, J., Llera, J.C. de la, González, J., González, G., 2018. Earthquake damage assessment for deterministic scenarios in Iquique, Chile. Natural Hazards 92, 1433–1461.

Allen, T.I., Wald, D.J., 2007. Topographic Slope as a Proxy for Seismic Site-Conditions (VS30) and Amplification Around the Globe (Report No. 2007–1357), Open-File Report. https://doi.org/10.3133/ofr20071357

Baggio, C., Bernardini, A., Colozza, R., Corazza, L., Della Orsini, M., Di Pascuale, G., Dolce, M., Goretti, A., Martinelli, A., Orsini, G., Papa, F., Zuccaro, G., 2007. Field Manual for post-earthquake damage and safety assessment and short term countermeasures (AeDES), EUR 22868 EN-Joint Research Centre-Institute for the Protection and Security of the Citizen.

ed, EUR – Scientific and Technical Research series – ISSN 1018-559. : Office for Official Publications of the European Communities, Luxembourg.
Bal, I.E., Bommer, J.J., Stafford, P.J., Crowley, H., Pinho, R., 2010. The Influence of Geographical Resolution of Urban Exposure Data in an Earthquake Loss Model for Istanbul. Earthquake Spectra 26, 619–634. https://doi.org/10.1193/1.3459127
Beven, K.J., Aspinall, W.P., Bates, P.D., Borgomeo, E., Goda, K., Hall, J.W., Page, T., Phillips, J.C., Simpson, M., Smith, P.J., Wagener, T., Watson, M., 2018. Epistemic uncertainties and natural hazard risk assessment – Part 2: What should constitute good practice? Natural Hazards and Earth System Sciences 18, 2769–2783. https://doi.org/10.5194/nhess-18-2769-2018
Bindi, D., Mayfield, M., Parolai, S., Tyagunov, S., Begaliev, U.T., Abdrakhmatov, K., Moldobekov, B., Zschau, J., 2011. Towards an improved seismic risk scenario for Bishkek, Kyrgyz Republic. Soil Dynamics and Earthquake Engineering 31, 521–525. https://doi.org/10.1016/j.soildyn.2010.08.009
Brzev, S., Scawthor, C., Charleson, AW., Allen, L., Greene, M., Jaiswal, K., Silva, V., 2013. GEM building taxonomy version 2.0 (GEM technical report 2013-02, June). GEM Foundation, Pavia.
Calvi, G.M., Pinho, R., Magenes, G., Crowley, H., Bommer, J.J., Restrepo-Velez, L.F., 2006. Development of seismic vulnerability assessment methodologies over the past 30 years. ISET journal 483 of Earthquake Technology 43, 75–104.
Carvajal, M., Cisternas, M., Catalán, P.A., 2017. Source of the 1730 Chilean earthquake from historical records: Implications for the future tsunami hazard on the coast of Metropolitan Chile. Journal of Geophysical Research: Solid Earth 122, 3648–3660. https://doi.org/10.1002/2017JB014063
CIESIN, 2018. Documentation for the Gridded Population of theWorld, Version 4 (GPWv4), Revision 11 Data Sets, [WWW Document]. NASA Socioeconomic Data And Applications Center (SEDAC) Palisades, NY, USA. URL https://doi.org/10.7927/H45Q4T5F
Comte, D., Eisenberg, A., Lorca, E., Pardo, M., Ponce, L., Saragoni, R., Singh, S.K., Suarez, G., 1986. The 1985 Central Chile Earthquake: A Repeat of Previous Great Earthquakes in the Region? Science 233, 449. https://doi.org/10.1126/science.233.4762.449
Corbane, C., Hancilar, U., Ehrlich, D., De Groeve, T., 2017. Pan-European seismic risk assessment: a proof of concept using the Earthquake Loss Estimation Routine (ELER). Bulletin of Earthquake Engineering 15, 1057–1083. https://doi.org/10.1007/s10518-016-9993-5
Crowley, H., 2014. Earthquake Risk Assessment: Present Shortcomings and Future Directions, in: Ansal, A. (Ed.), Perspectives on European Earthquake Engineering and Seismology: Volume 1. Springer International Publishing, Cham, pp. 515–532. https://doi.org/10.1007/978-3-319-07118-3_16
Crowley, H., Bommer, J.J., 2006. Modelling Seismic Hazard in Earthquake Loss Models with Spatially Distributed Exposure. Bulletin of Earthquake Engineering 4, 275–275. https://doi.org/10.1007/s10518-006-9011-4
Crowley, H., Bommer, J.J., Pinho, R., Bird, J., 2005. The impact of epistemic uncertainty on an earthquake loss model. Earthquake Engineering & Structural Dynamics 34, 1653–1685. https://doi.org/10.1002/eqe.498
Crowley, H., Despotaki, V., Rodrigues, D., Silva, V., Toma-Danila, D., Riga, E., Karatzetzou, A., Fotopoulou, S., Zugic, Z., Sousa, L., Ozcebe, S., Gamba, P., 2020. Exposure model for European seismic risk assessment. Earthquake Spectra 8755293020919429. https://doi.org/10.1177/8755293020919429
Crowley, H., Pinho, R., 2004. Period-Height Relationship for Existing European Reinforced Concrete Buildings. Journal of Earthquake Engineering 8, 93–119. https://doi.org/10.1080/13632460409350522
Crowley, H., Stafford, P.J., Bommer, J.J., 2008. Can Earthquake Loss Models be Validated Using Field Observations? Journal of Earthquake Engineering 12, 1078–1104. https://doi.org/10.1080/13632460802212923
Dabbeek, J., Silva, V., 2020. Modeling the residential building stock in the Middle East for multi-hazard risk assessment. Natural Hazards 100, 781–810. https://doi.org/10.1007/s11069-019-03842-7
D'Ayala, D., Galasso, C., Nassirpour, A., Adhikari, R.K., Yamin, L., Fernandez, R., Lo, D., Garciano, L., Oreta, A., 2020. Resilient communities through safer schools. International Journal of Disaster Risk Reduction 45, 101446. https://doi.org/10.1016/j.ijdrr.2019.101446
de la Llera, J.C., Rivera, F., Mitrani-Reiser, J., Jünemann, R., Fortuño, C., Ríos, M., Hube, M., Santa María, H., Cienfuegos, R., 2017. Data collection after the 2010 Maule earthquake in Chile. Bulletin of Earthquake Engineering 15, 555–588. https://doi.org/10.1007/s10518-016-9918-3
Dell'Acqua, F., Gamba, P., Jaiswal, K., 2013. Spatial aspects of building and population exposure data and their implications for global earthquake exposure modeling. Natural Hazards 68, 1291–1309. https://doi.org/10.1007/s11069-012-0241-2
FEMA, 2012. Hazus mr-MH 2.1 user manual, earthquake model (Technical report, FEMA).
FEMA, 2003. Multi-hazard loss estimation methodology. Federal Emergency Management Agency, Washington.
FEMA 154, 2002. Rapid Visual Screening of Buildings for Potential Seismic Hazards, 2nd ed. APPLIED TECHNOLOGY COUNCIL 555 Twin Dolphin Drive, Suite 550 Redwood City, California 94065.
GEM, 2014. Report on the SARA Exposure and Vulnerability Workshop in Medellin, Colombia (Report produced in the context of the GEM South America integrated Risk Assessment (SARA) project No. Version 1.0-May 2014).
Grigoratos, I., Dabeek, J., Faravelli, M., Di Meo, A., Cerchiello, V., Borzi, B., Monteiro, R., Ceresa, P., 2016. Development of A Fragility and Exposure Model for Palestine – Application to The City of Nablus. Procedia Engineering 161, 2023–2029. https://doi.org/10.1016/j.proeng.2016.08.797
Grünthal, G., 1998. European Macroseismic Scale 1998, Centre Européen de Géodynamique et de Séismologie. ed. Luxembourg.

Haas, M., 2018. Towards time- and state-dependent seismic risk over urban scales. der Technischen Universität Berlin, Berlin.
Haas, M., Wieland, M., Pittore, M., 2016. DEMO: Remote Rapid Visual Screening (RRVS) [WWW Document]. URL https://vimeo.com/158600573
Hancilar, U., Tuzun, C., Yenidogan, C., Erdik, M., 2010. ELER software – a new tool for urban earthquake loss assessment. Natural Hazards and Earth System Sciences 10, 2677–2696. https://doi.org/10.5194/nhess-10-2677-2010
Hastie, D.I., Liverani, S., Richardson, S., 2015. Sampling from Dirichlet process mixture models with unknown concentration parameter: mixing issues in large data implementations. Statistics and Computing 25, 1023–1037. https://doi.org/10.1007/s11222-014-9471-3
Indirli, M., Razafindrakoto, H., Romanelli, F., Puglisi, C., Lanzoni, L., Milani, E., Munari, M., Apablaza, S., 2011. Hazard Evaluation in Valparaíso: the MAR VASTO Project. Pure and Applied Geophysics 168, 543–582. https://doi.org/10.1007/s00024-010-0164-3
Jaiswal, K., Wald, D., Porter, K., 2010. A Global Building Inventory for Earthquake Loss Estimation and Risk Management. Earthquake Spectra 26, 731–748. https://doi.org/10.1193/1.3450316
Jiménez, B., Pelà, L., Hurtado, M., 2018. Building survey forms for heterogeneous urban areas in seismically hazardous zones. Application to the historical center of Valparaíso, Chile. International Journal of Architectural Heritage 12, 1076–1111. https://doi.org/10.1080/15583058.2018.1503370
Jiménez Martínez, Maribel, Jiménez Martínez, Mónica, Romero-Jarén, R., 2020. How resilient is the labour market against natural disaster? Evaluating the effects from the 2010 earthquake in Chile. Natural Hazards 104, 1481–1533. https://doi.org/10.1007/s11069-020-04229-9
Jordan, C.J., Adlam, K., Lauri, K., Shelley, W., Bevington, J., 2014. User guide: Windows tool for field data collection and management (Technical Report 2014---04). pp., GEM Foundation, Pavia, Italy.
Kalakonas, P., Silva, V., Mouyiannou, A., Rao, A., 2020. Exploring the impact of epistemic uncertainty on a regional probabilistic seismic risk assessment model. Natural Hazards. https://doi.org/10.1007/s11069-020-04201-7
Kohrangi, M., Kotha, S.R., Bazzurro, P., 2018. Ground-motion models for average spectral acceleration in a period range: direct and indirect methods. Bulletin of Earthquake Engineering 16, 45–65. https://doi.org/10.1007/s10518-017-0216-5
Kohrangi, M., Vamvatsikos, D., Bazzurro, P., 2017. Site dependence and record selection schemes for building fragility and regional loss assessment. Earthquake Engineering & Structural Dynamics 46, 1625–1643. https://doi.org/10.1002/eqe.2873
Liuzzi, M., Aravena Pelizari, P., Geiß, C., Masi, A., Tramutoli, V., Taubenböck, H., 2019. A transferable remote sensing approach to classify building structural types for seismic risk analyses: the case of Val d'Agri area (Italy). Bulletin of Earthquake Engineering 17, 4825–4853. https://doi.org/10.1007/s10518-019-00648-7
Luco, N., Cornell, C.A., 2007. Structure-Specific Scalar Intensity Measures for Near-Source and Ordinary Earthquake Ground Motions. Earthquake Spectra 23, 357–392. https://doi.org/10.1193/1.2723158
Markhvida, M., Ceferino, L., Baker, J.W., 2018. Modeling spatially correlated spectral accelerations at multiple periods using principal component analysis and geostatistics. Earthquake Engineering & Structural Dynamics 47, 1107–1123. https://doi.org/10.1002/eqe.3007
Martínez-Cuevas, S., Benito, M.B., Cervera, J., Morillo, M.C., Luna, M., 2017. Urban modifiers of seismic vulnerability aimed at Urban Zoning Regulations. Bulletin of Earthquake Engineering 15, 4719–4750. https://doi.org/10.1007/s10518-017-0162-2
Martins, L., Silva, V., 2020. Development of a fragility and vulnerability model for global seismic risk analyses. Bulletin of Earthquake Engineering. https://doi.org/10.1007/s10518-020-00885-1
Martins, L., Silva, V., Marques, M., Crowley, H., Delgado, R., 2016. Development and assessment of damage-to-loss models for moment-frame reinforced concrete buildings. Earthquake Engineering & Structural Dynamics 45, 797–817. https://doi.org/10.1002/eqe.2687
Mendoza, L., Ayala, F., Fuentes, B., Soto, V., Sáez, E., Yañez, G., Montalva, Gonzalo., Gález, C., Sepúlveda, N., Lazo, I., Ruiz, J., 2018. Estimación cuantitativa de la amenaza sísmica en base a métodos geofísicos: aplicación a las localidades costeras del segmento los Vilos – San Antonio. Presented at the 50 Congreso SOCHIGE., Valparaiso, Chile.
Montalva, G.A., Bastías, N., Rodriguez-Marek, A., 2017. Ground-Motion Prediction Equation for the Chilean Subduction Zone. The Bulletin of the Seismological Society of America 107, 901–911. https://doi.org/10.1785/0120160221
Montessus de Ballore, F., 1914. Historia sísmica de los Andes Meridionales al sur del paralelo XVI, Cuarta parte. Imprenta Cervantes, Santiago, Chile.
Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface QR Railway Tech. Res. Inst., 30(1), 25-33.
Nicodemo, G., Pittore, M., Masi, A., Manfredi, V., 2020. Modelling exposure and vulnerability from post-earthquake survey data with risk-oriented taxonomies: AeDES form, GEM taxonomy and EMS-98 typologies. International Journal of Disaster Risk Reduction 50, 101894. https://doi.org/10.1016/j.ijdrr.2020.101894
O'Hagan, A., Buck, C.E., Daneshkhah, A., Eiser, J.R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T., 2006. Uncertain Judgements: Eliciting Expert Probabilities, Uncertain Judgements: Eliciting Expert Probabilities. John Wiley.
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. https://doi.org/10.1785/0220130087

Pavić, G., Hadzima-Nyarko, M., Bulajić, B., Jurković, Ž., 2020. Development of Seismic Vulnerability and Exposure Models—A Case Study of Croatia. Sustainability 12, 973. https://doi.org/10.3390/su12030973
Pilz, M., Bindi, D., Boxberger, T., Hakimov, F., Moldobekov, B., Murodkulov, S., Orunbaev, S., Pittore, M., Stankiewicz, J., Ullah, S., Verjee, F., Wieland, M., Yasunov, P., Parolai, S., 2013. First Steps toward a Reassessment of the Seismic Risk of the City of Dushanbe (Tajikistan). Seismological Research Letters 84, 1026–1038. https://doi.org/10.1785/0220130040
Pittore, M., Haas, M., Megalooikonomou, K.G., 2018. Risk-Oriented, Bottom-Up Modeling of Building Portfolios With Faceted Taxonomies. Frontiers in Built Environment 4, 41. https://doi.org/10.3389/fbuil.2018.00041
Pittore, M., Haas, M., Silva, V., 2020. Variable resolution probabilistic modeling of residential exposure and vulnerability for risk applications. Earthquake Spectra 36, 321–344. https://doi.org/10.1177/8755293020951582
Pittore, M., Wieland, M., 2013. Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. Natural Hazards 68, 115–145. https://doi.org/10.1007/s11069-012-0475-z
Pittore, M., Wieland, M., Fleming, K., 2017. Perspectives on global dynamic exposure modelling for geo-risk assessment. Natural Hazards 86, 7–30. https://doi.org/10.1007/s11069-016-2437-3
Polese, M., Di Ludovico, M., Gaetani d'Aragona, M., Prota, A., Manfredi, G., 2020. Regional vulnerability and risk assessment accounting for local building typologies. International Journal of Disaster Risk Reduction 43, 101400. https://doi.org/10.1016/j.ijdrr.2019.101400
Ruiz, S., Aden-Antoniow, F., Baez, J.C., Otarola, C., Potin, B., del Campo, F., Poli, P., Flores, C., Satriano, C., Leyton, F., Madariaga, R., Bernard, P., 2017. Nucleation Phase and Dynamic Inversion of the Mw 6.9 Valparaíso 2017 Earthquake in Central Chile. Geophysical Research Letters 44, 10,290-10,297. https://doi.org/10.1002/2017GL075675
Santa María, H., Hube, M.A., Rivera, F., Yepes-Estrada, C., Valcárcel, J.A., 2017. Development of national and local exposure models of residential structures in Chile. Natural Hazards 86, 55–79. https://doi.org/10.1007/s11069-016-2518-3
Silva, V., 2016. Critical Issues in Earthquake Scenario Loss Modeling. Journal of Earthquake Engineering 20, 1322–1341. https://doi.org/10.1080/13632469.2016.1138172
Silva, V., Amo-Oduro, D., Calderon, A., Costa, C., Dabbeek, J., Despotaki, V., Martins, L., Pagani, M., Rao, A., Simionato, M., Viganò, D., Yepes-Estrada, C., Acevedo, A., Crowley, H., Horspool, N., Jaiswal, K., Journeay, M., Pittore, M., 2020. Development of a global seismic risk model. Earthquake Spectra 8755293019899953. https://doi.org/10.1177/8755293019899953
Silva, V., Yepes-Estrada, C., Dabbeek, J., Martins, L., Brzev, S., 2018. GED4ALL: Global exposure database for multi-hazard risk analysis. Multi-hazard exposure taxonomy, GEM technical report 2018-01. GEM Foundation, Pavia.
Sousa, L., Marques, M., Silva, V., Graeme, W., 2018. Epistemic Uncertainty in Hazard and Fragility Modelling for Earthquake Engineering. Presented at the 16th European Conference on Earthquake Engineering, Thessaloniki, Greece.
Torres, Y., Arranz, J.J., Gaspar-Escribano, J.M., Haghi, A., Martínez-Cuevas, S., Benito, B., Ojeda, J.C., 2019. Integration of LiDAR and multispectral images for rapid exposure and earthquake vulnerability estimation. Application in Lorca, Spain. International Journal of Applied Earth Observation and Geoinformation 81, 161–175. https://doi.org/10.1016/j.jag.2019.05.015
UNISDR, 2009. UNISDR terminology on disaster risk reduction, United Nations International Strategy forDisaster Reductio, UNISDR-20-2009. Geneva.
Vamvatsikos, D., Panagopoulos, G., Kappos, A.J., Nigro, E., Rossetto, T., Lloyd, T.O., Stathopoulos, T., 2010. Structural Vulnerability Assessment under Natural Hazards: A review, in: Urban Habitat Constructions under Catastrophic Events, Chapter: 3-4. CRC Press. Editor: Mazzolani, F.M.
Villar-Vega, M., Silva, V., 2017. Assessment of earthquake damage considering the characteristics of past events in South America. Soil Dynamics and Earthquake Engineering 99, 86–96. https://doi.org/10.1016/j.soildyn.2017.05.004
Villar-Vega, M., Silva, V., Crowley, H., Yepes, C., Tarque, N., Acevedo, A.B., Hube, M.A., Gustavo, C.D., María, H.S., 2017. Development of a Fragility Model for the Residential Building Stock in South America. Earthquake Spectra 33, 581–604. https://doi.org/10.1193/010716EQS005M
Weatherill, G.A., Silva, V., Crowley, H., Bazzurro, P., 2015. Exploring the impact of spatial correlations and uncertainties for portfolio analysis in probabilistic seismic loss estimation. Bulletin of Earthquake Engineering 13, 957–981. https://doi.org/10.1007/s10518-015-9730-5
Wesson, R.L., Perkins, D.M., 2001. Spatial correlation of probabilistic earthquake ground motion and loss. Bulletin of the Seismological Society of America 91, 1498–1515. https://doi.org/10.1785/0120000284
WHE, 2014. World Housing Encyclopedia [WWW Document]. Housing reports. URL http://db.world-housing.net (accessed 5.15.20).
Wieland, M., Pittore, M., Parolai, S., Begaliev, U., Yasunov, P., Tyagunov, S., Moldobekov, B., Saidiy, S., Ilyasov, I., Abakanov, T., 2014. A Multiscale Exposure Model for Seismic Risk Assessment in Central Asia. Seismological Research Letters 86, 210–222. https://doi.org/10.1785/0220140130
Yepes-Estrada, C., Silva, V., Valcárcel, J., Acevedo, A.B., Tarque, N., Hube, M.A., Coronel, G., María, H.S., 2017. Modeling the Residential Building Inventory in South America for Seismic Risk Assessment. Earthquake Spectra 33, 299–322. https://doi.org/10.1193/101915eqs155dp
Zuccaro, G., Cacace, F., 2015. Seismic vulnerability assessment based on typological characteristics. The first level procedure "SAVE." Soil Dynamics and Earthquake Engineering 69, 262–269. https://doi.org/10.1016/j.soildyn.2014.11.003

Figures

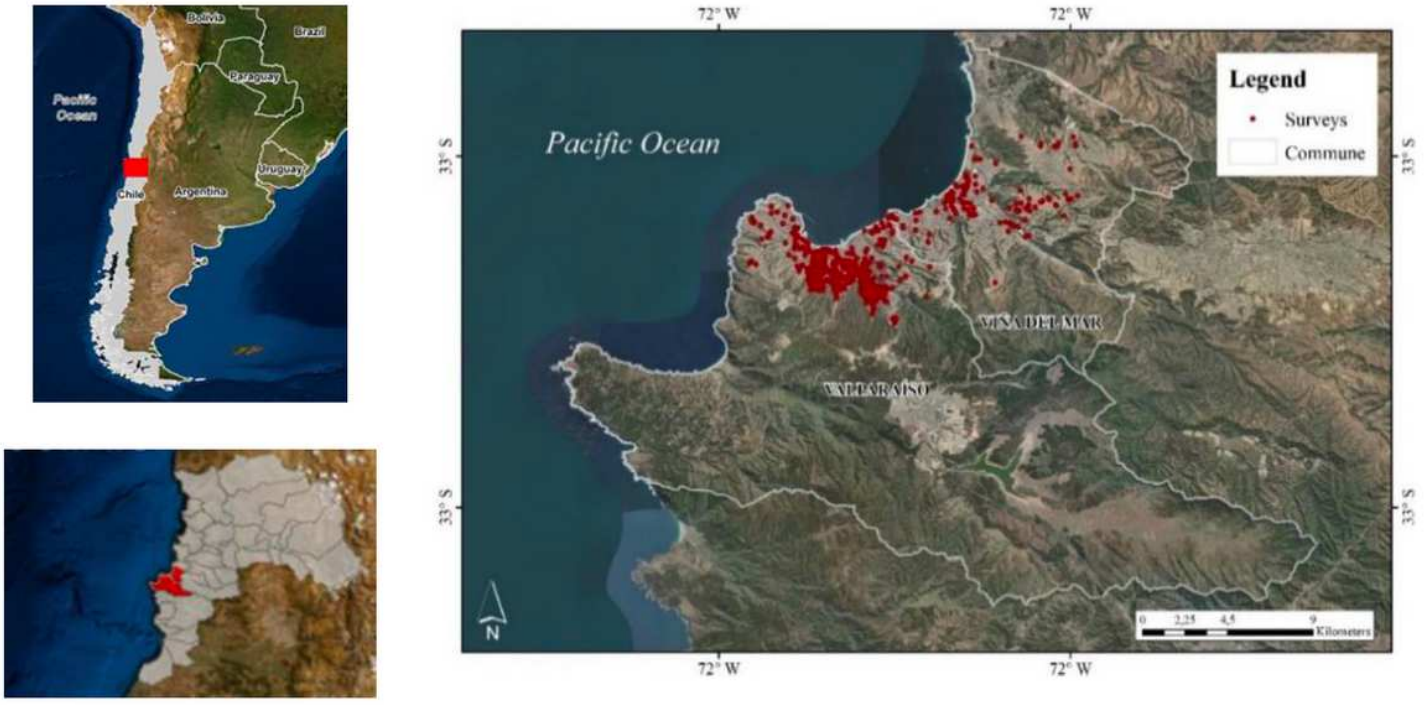


Figure 1

Location of the surveyed buildings within the urban area of the communes of Valparaiso and Viña del Mar within Valparaiso Region and Chile.

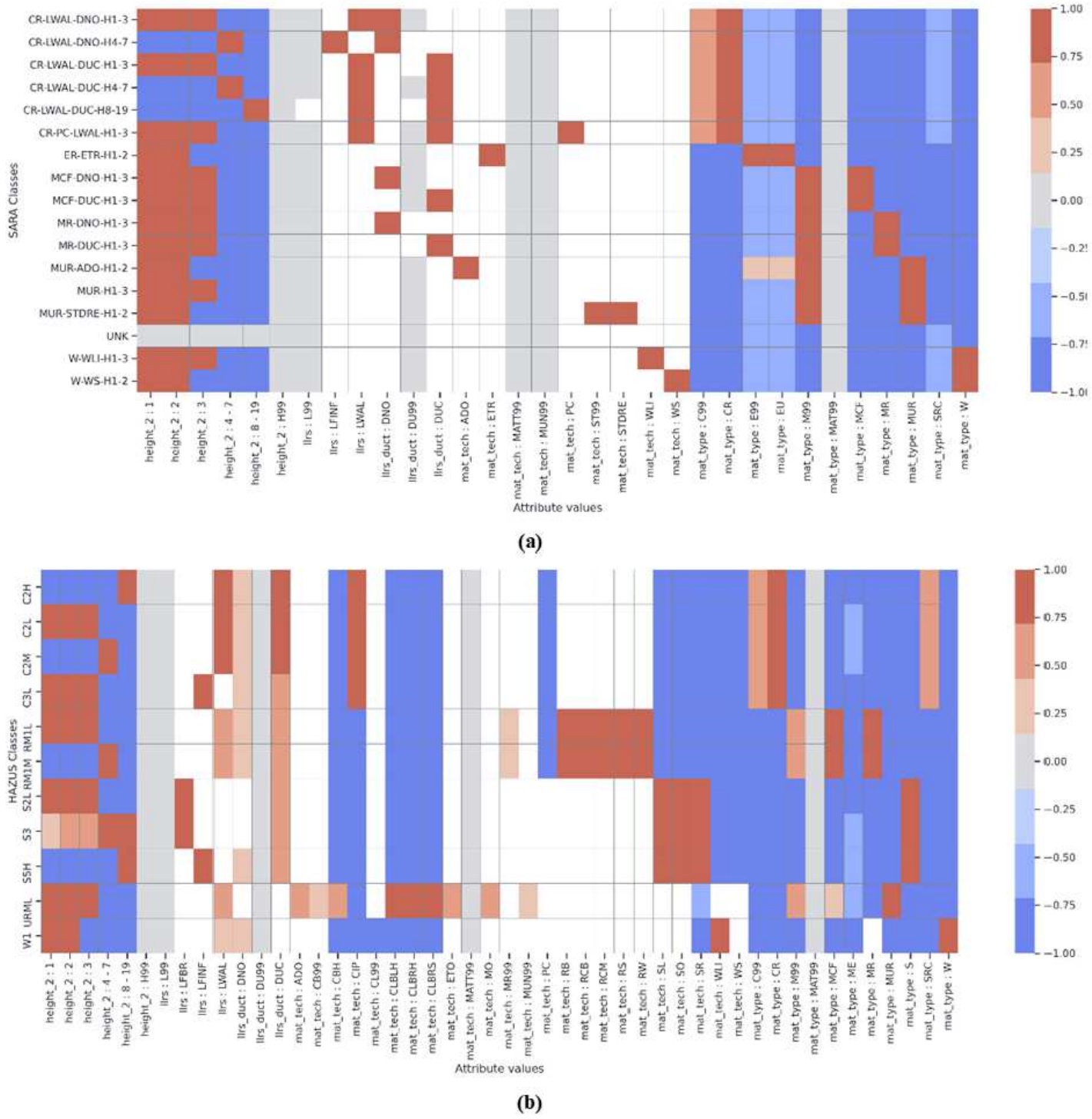


Figure 2

Graphical representation of the (a) SARA and (b) HAZUS building classes scheme definition. The colour encodes the compatibility value with extremes in red and blue that represent high and low compatibilities respectively. Grey: neutral and white indicates no explicit compatibility value has been assigned.

SARA	W-WLI-H1	CR-LWAL-DUC-H8-19	CR-LWAL-DUC-H4-7		
HAZUS	W1	C2H	RM1M		

SARA	UNK	MUR-STDRE-H1-2 (*W.A-1*); MUR-H1-3 (*W.A-2*)	UNK	MCF-DNO-H1-3 *(W.A-1);* MCF-DUC-H1-3 *(W.A.2)*	MR-DNO-H1-3 *(W.A-1);* MR-DUC-H1-3 *(W.A.2)*
HAZUS	C2H	URML	S3	RM1L	RM1L

Figure 3

Pictures of some buildings' façade surveyed in Valparaiso. Their classification into the SARA and HAZUS schemes is displayed considering the two weight arrangements in Table 2. No distinction is made when both weights led to the same class.

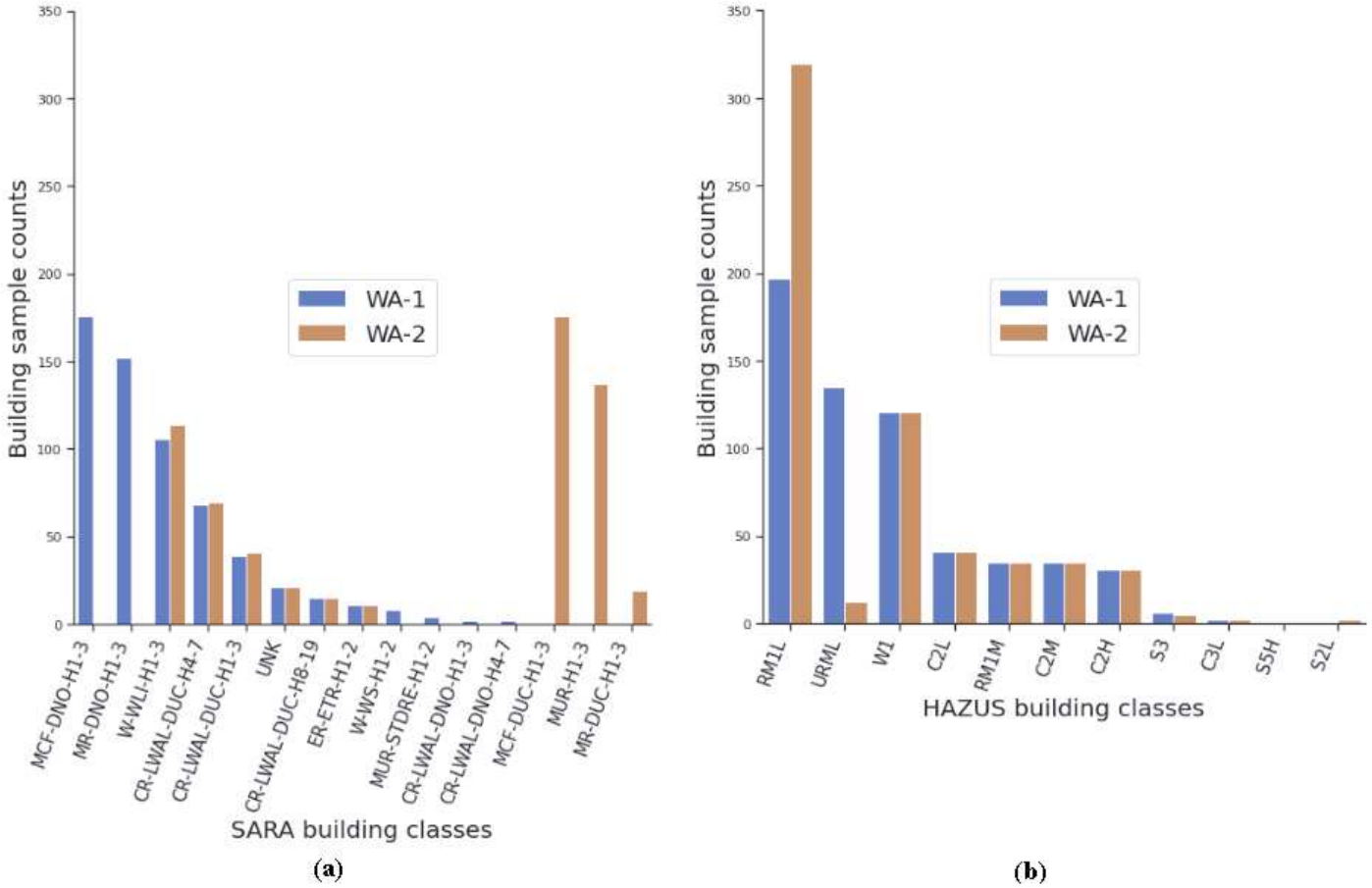


Figure 4

Distributions of the most likely building class of the survey from the compatibility levels between the observed building attributes and the classes of every scheme: (a) SARA and (b) HAZUS for two weight arrangements (WA-1, WA-2) in Table 2.

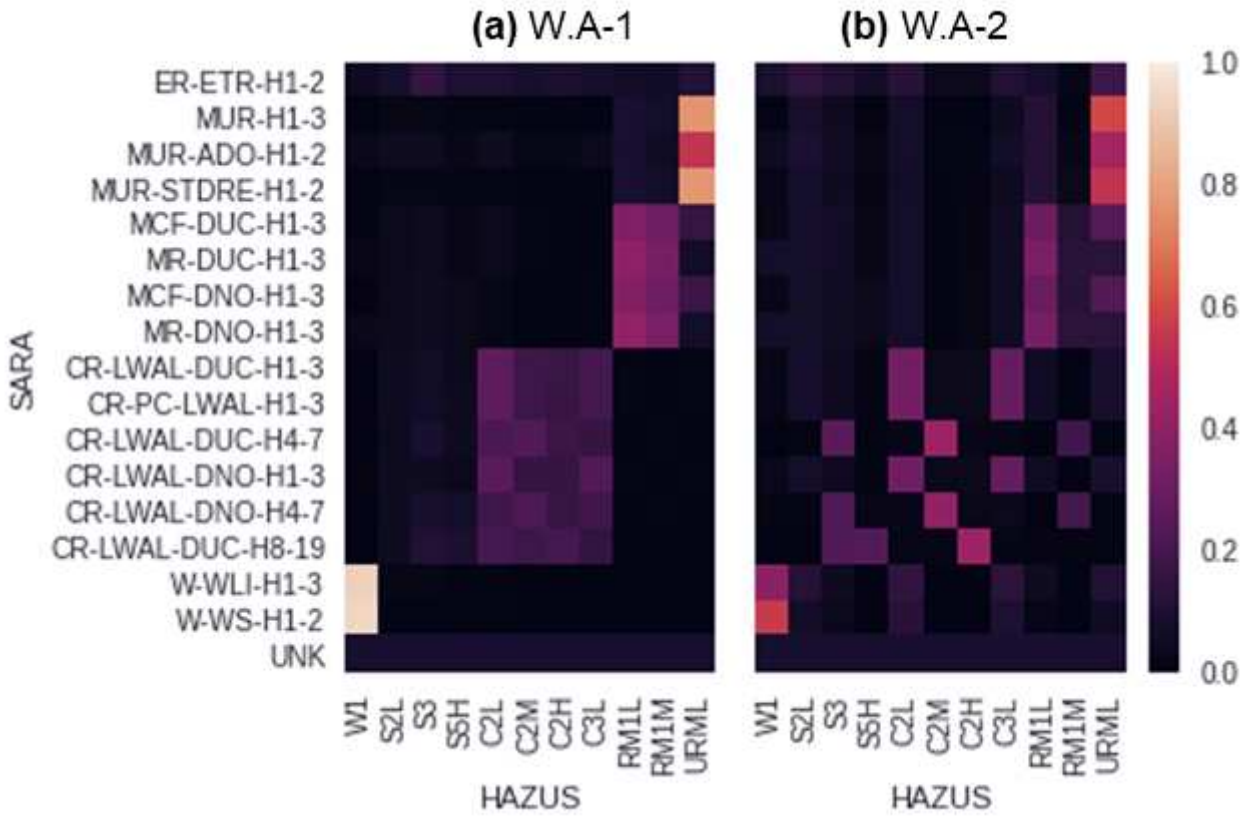


Figure 5

Inter-scheme compatibility matrix for SARA (source) and HAZUS (target) in Valparaiso through the weight arrangements: W.A-1 (a) and W.A-2 (b) reported in Table 2.

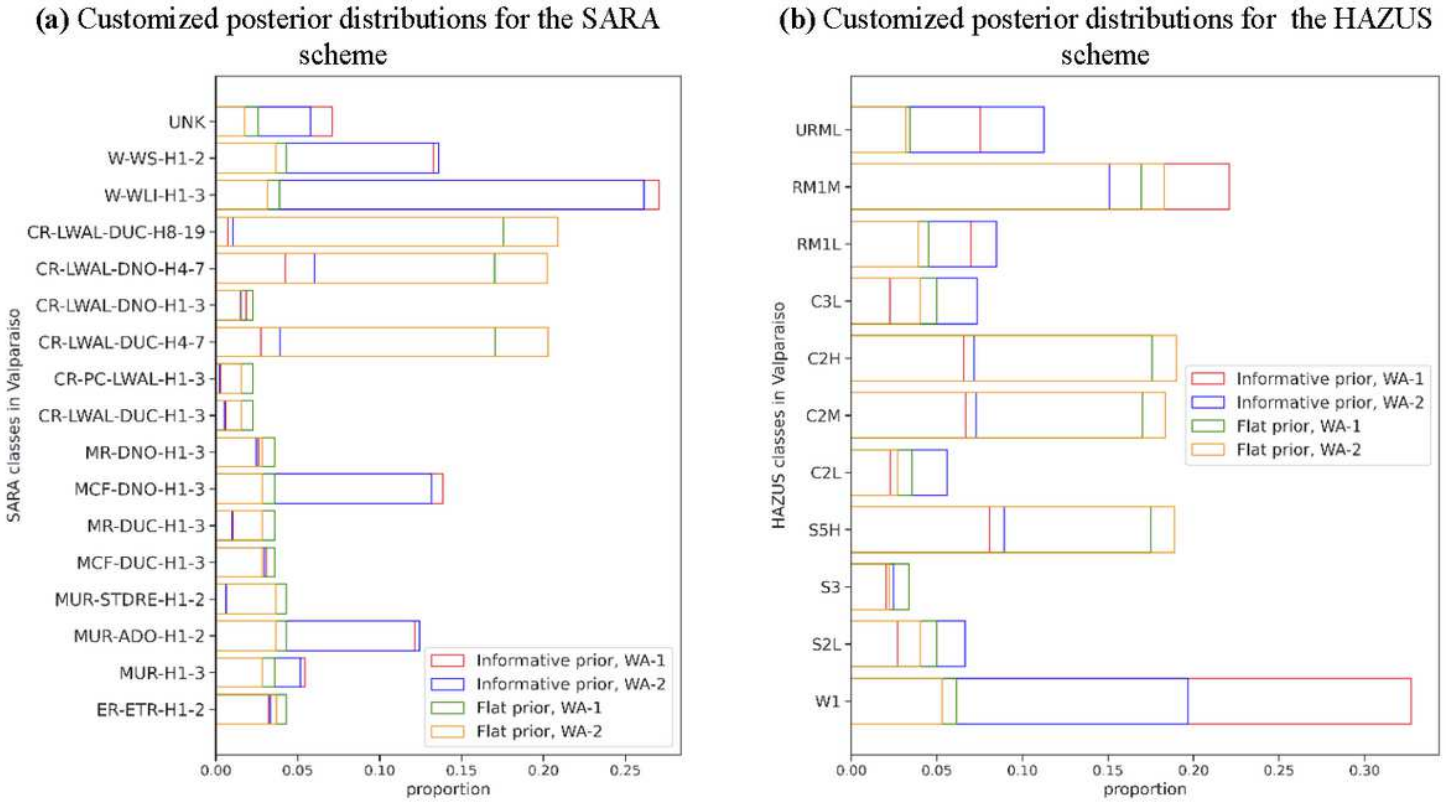
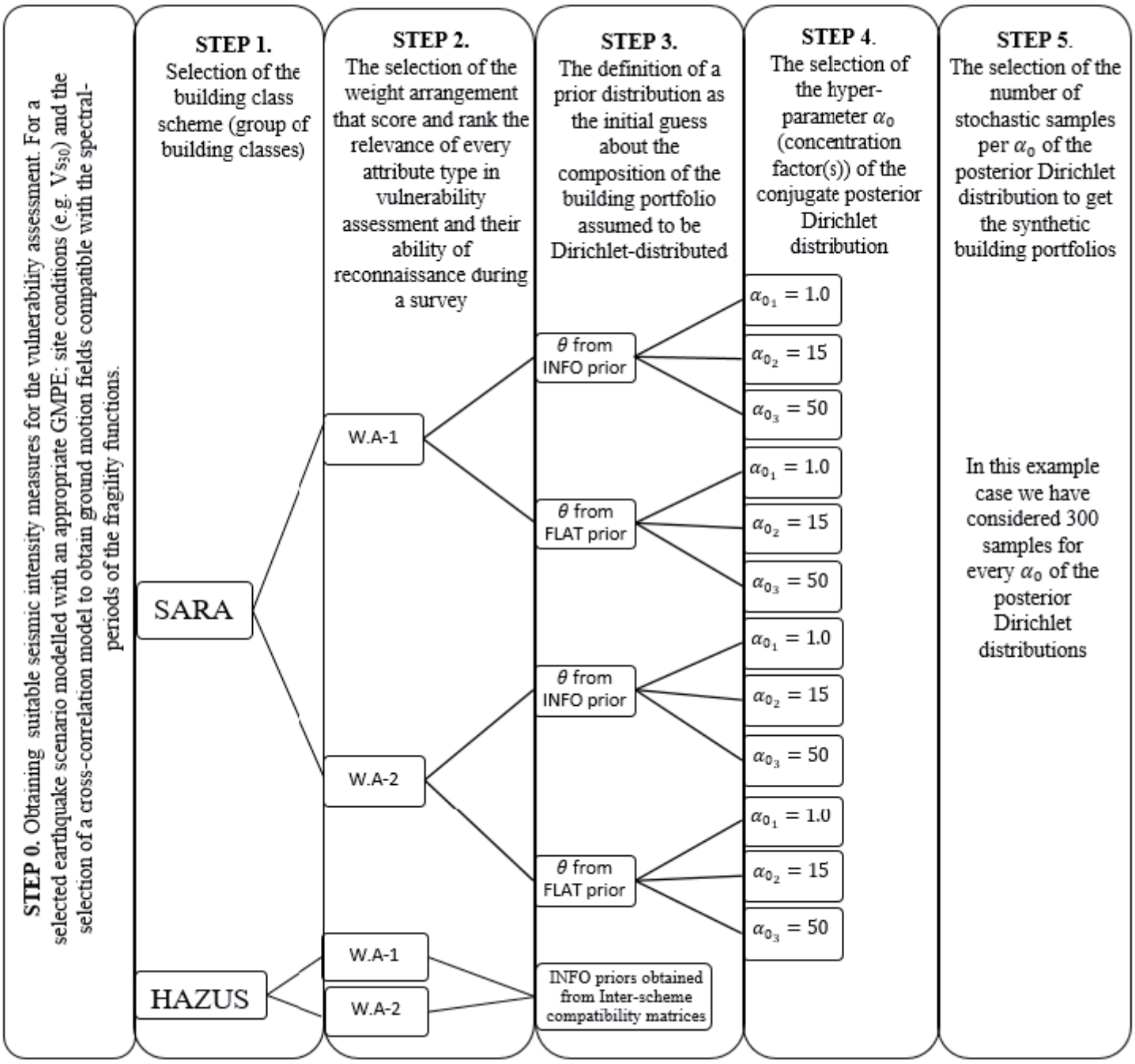


Figure 6

Posterior distributions obtained for the (a) SARA and (b) HAZUS schemes for the 604 inspected building surveyed in Valparaiso while considering different weight arrangements (WA) and flat and informative priors.

Figure 7



Logic tree with five levels constructed explore the impact of the building exposure composition modelling in Valparaiso. The twelve branches result from every considered scheme. The entire one is shown only for SARA. The same procedure has been carried out for HAZUS.

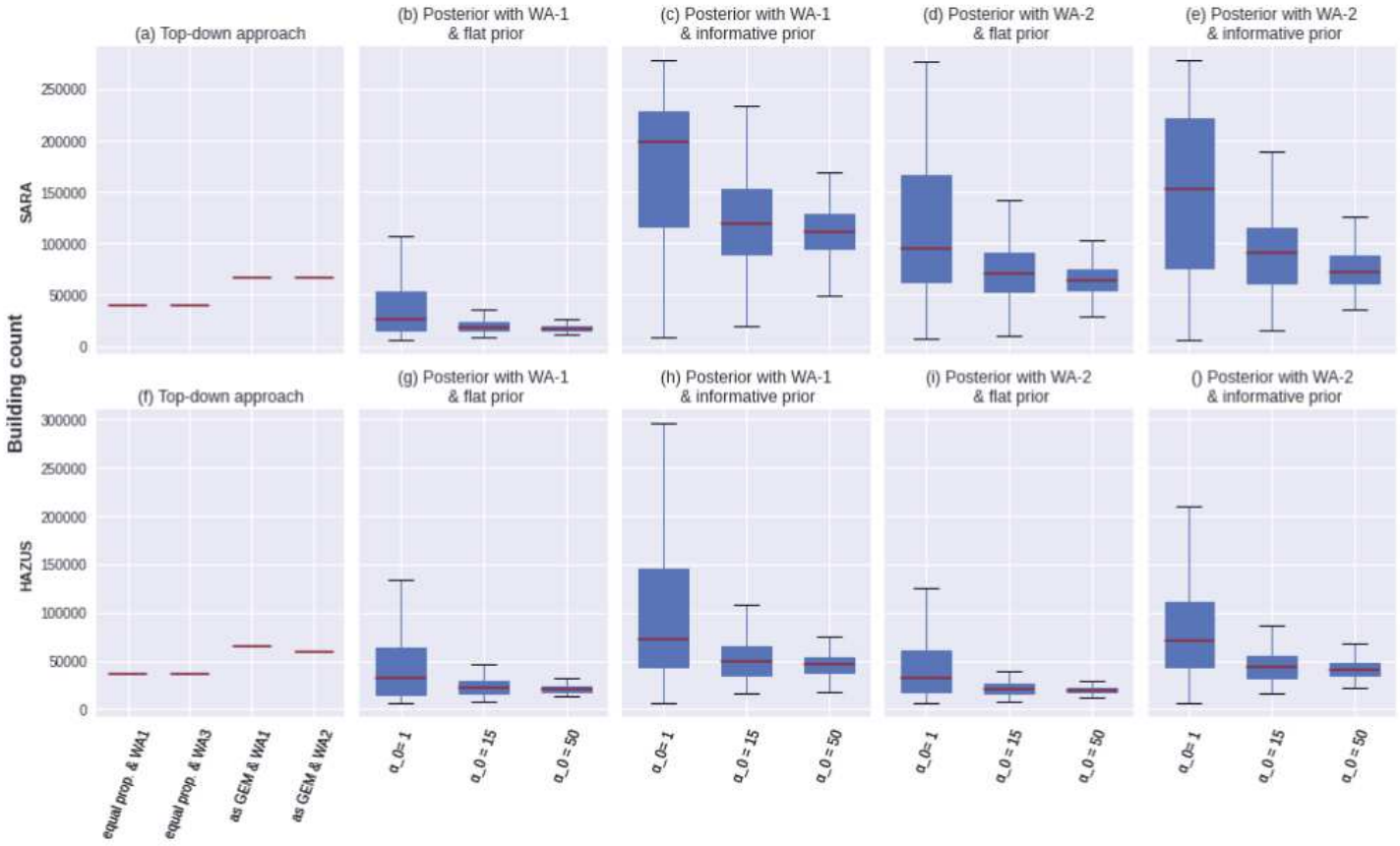


Figure 8

Dependency between the variance of the building counts (from population extraction and occupancy) and the portfolio composition. The first column comprises a single value in which the weight arrangement WA did not impact SARA, contrary to what occurs in HAZUS. The distributions on the other columns corresponds to every case in the logic tree depicted in Figure 7.

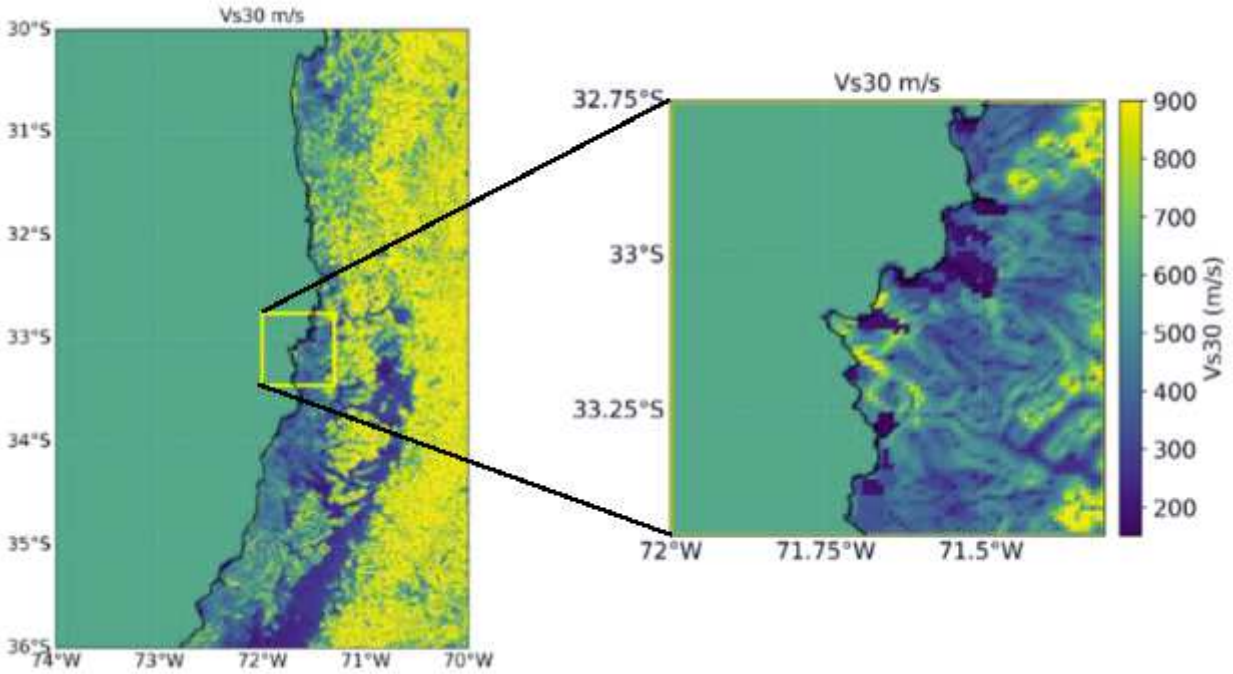


Figure 9

Distribution of the Vs30 values in Chile from slope as proposed by the USGS, and refined within the study area with the seismic microzonation reported in Mendoza et al. 2018.

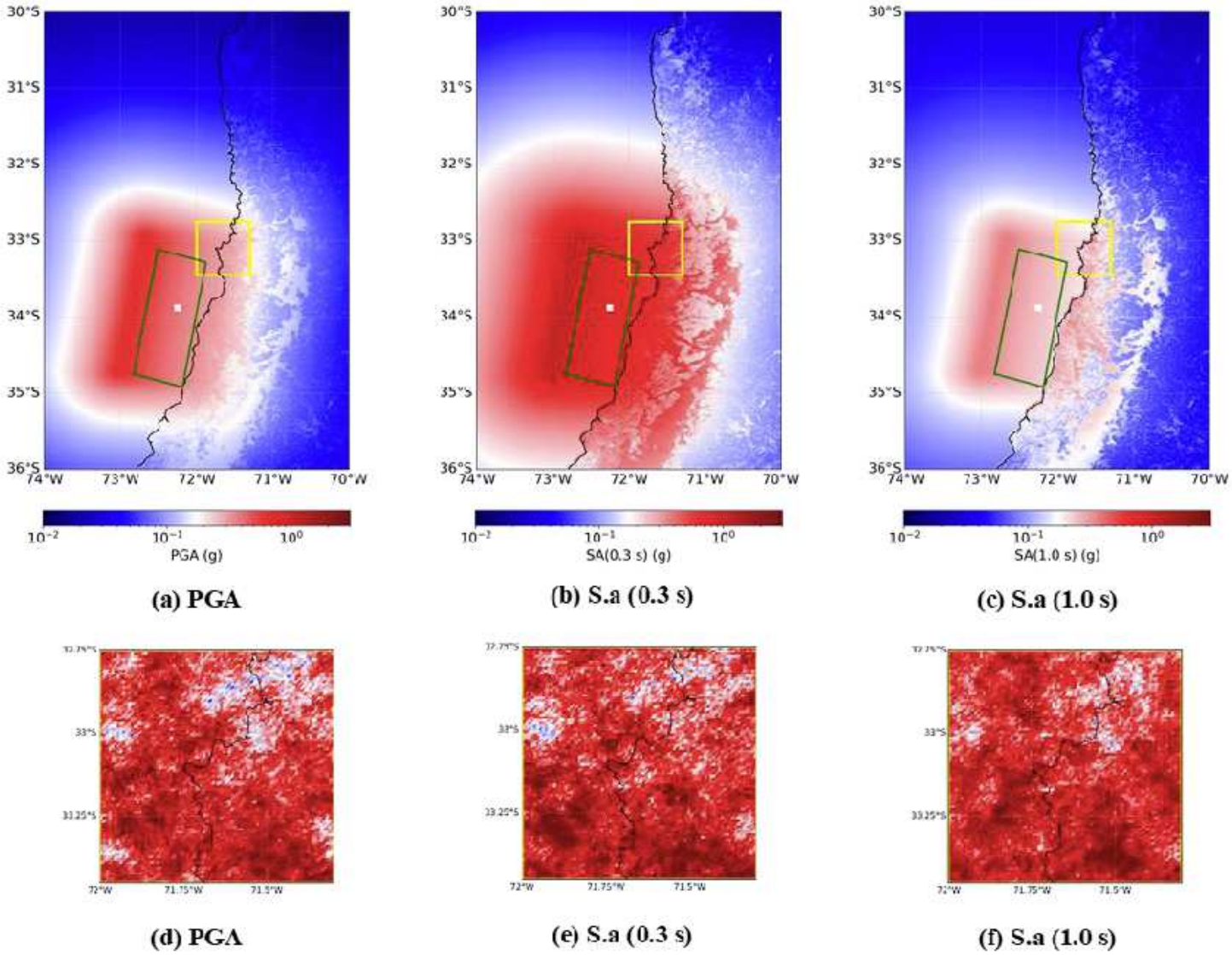


Figure 10

First row: distributed median values obtained from a single realisation of the Mw 8.2 earthquake rupture for three spectral periods (a) PGA, (b) S.a (0.3 s) and (c) S.a (1.0 s) using the Montalva et al. 2017 GMPE. The earthquake hypocentre is depicted by a white square. The rupture plane is displayed by a green rectangle. The study area is shown by a yellow square Second row: zoom of the study area with a single realisation of the cross-correlated seismic ground motion fields for their correspondent considered spectral periods using the Markhvida et al. 2018 model.

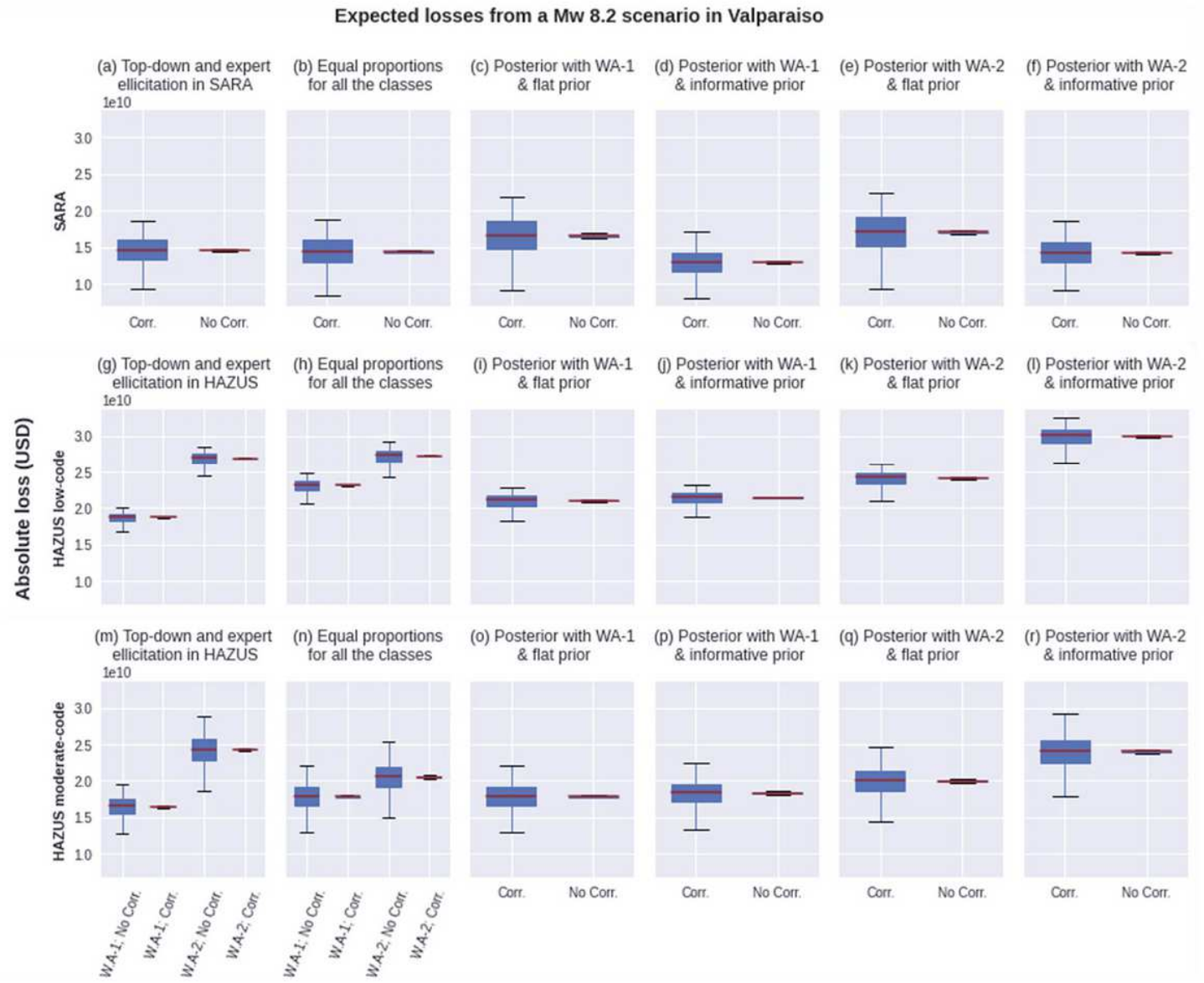


Figure 11

Comparison of the absolute loss (USD) considering SARA (first row) and HAZUS with varying code-compliance level (low code in second row; and moderate code in third row). The losses are obtained after the earthquake scenario Mw 8.2 in the residential building stock in Valparaiso using the cross-correlation model (Corr.) and without it (No Corr.). Different portfolio compositions are considered: the top-down vision and expert elicitation (first plot in every row); equally composed portfolios (second plot in every row) and the customized posterior distributions for both schemes as shown in Figure 6.

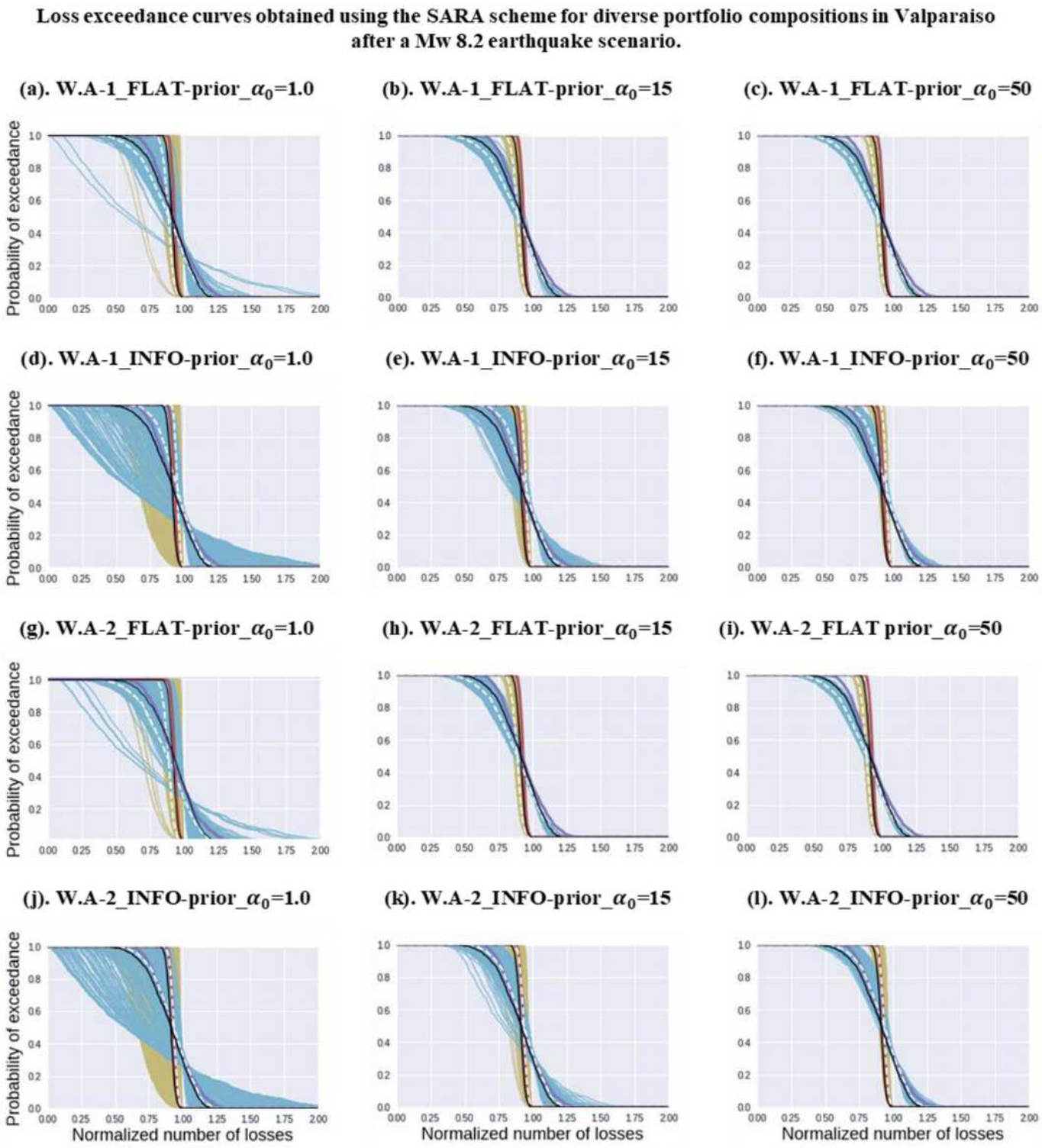


Figure 12

Normalized loss exceedance curves (LEC) for different exposure compositions as depicted in the logic tree (Figure 7) for the SARA scheme. Curves are displayed in blue if the ground motion cross-correlation model was addressed and in yellow with no ground motion cross-correlation. In every plot, the LEC from their respective customized posterior distributions are displayed in non- continours white curves. The single composition vision as informative priors (as defined by GEM) are depicted in purple when the

ground motion cross-correlation model was addressed and in red with no ground motion cross-correlation. Black curves represent the losses portfolio whose composition is assumed to be have equal proportions for all of the classes. The black curve with a smoother shape is obtained while modeling cross-correlation in oposition to the sharper one when it was not.

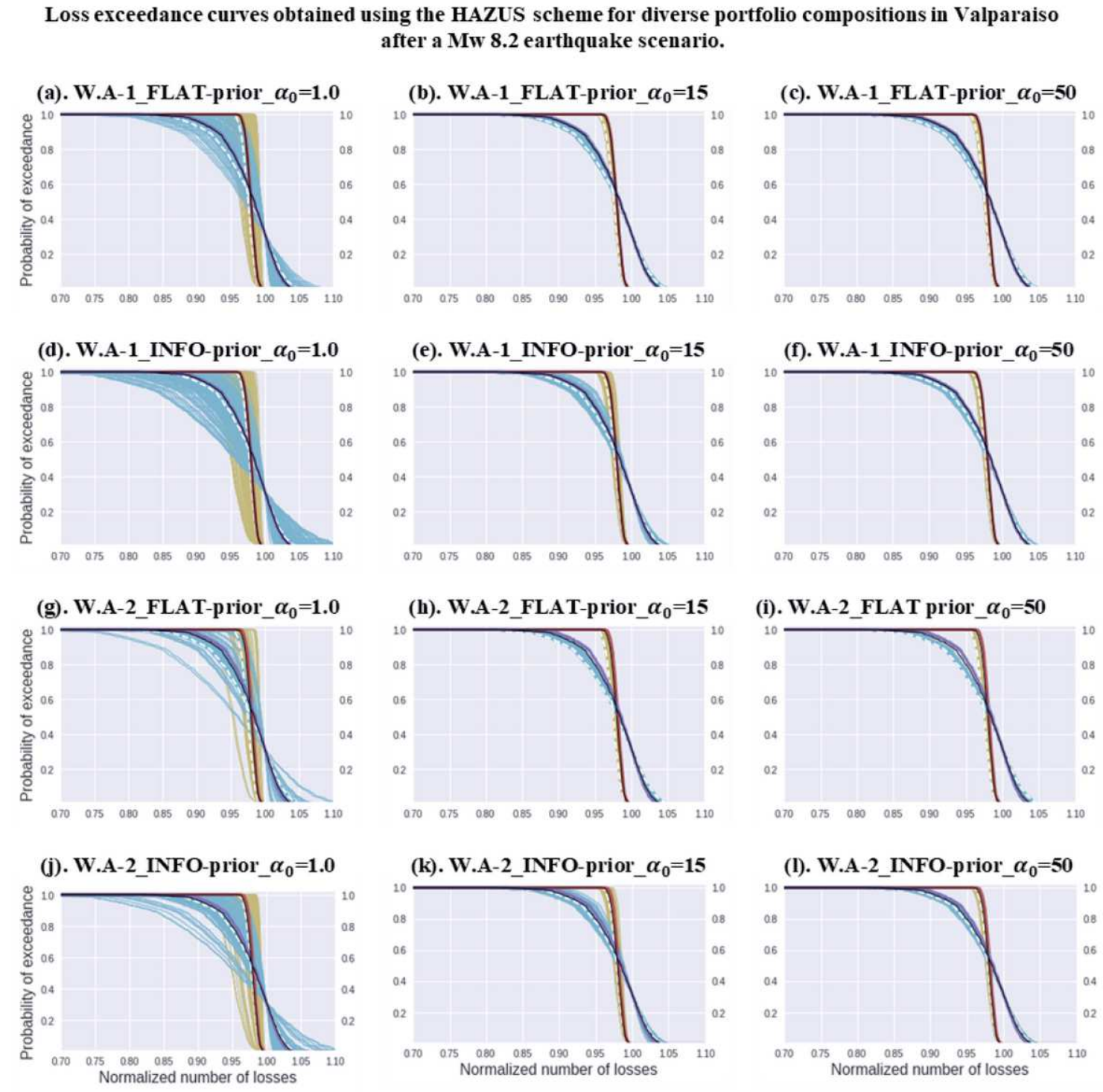


Figure 13

Normalized loss exceedance curves (LEC) for different exposure compositions as depicted in the logic tree (Figure 7) for the HAZUS scheme. Curves are displayed in blue if the ground motion cross-correlation

model was addressed and in yellow with no ground motion cross-correlation. In every plot, the LEC from their respective customized posterior distributions are displayed in non- continours white curves. The single composition vision as informative priors (from the inter-scheme conversion matrices in Figure 5) are depicted in purple when the ground motion cross-correlation model was addressed and in red with no ground motion cross-correlation. Black curves represent the losses of a portfolio whose composition is assumed to be have equal proportions for the classes. The black curve with a smoother shape is obtained while modeling cross-correlation in oposition to the sharper one when it was not.