Empirical fragility curves: The effect of uncertainty in ground motion intensity

Ioanna Ioannou[a*], Richard E. Chandler[b] and Tiziana Rossetto[a]

[a]Department of Civil, Environmental & Geomatic Engineering, University College London
[b]Department of Statistical Science , University College London.
[*]Corresponding author.

Abstract: Empirical fragility curves derived from large post-disaster databases with data aggregated at municipality-level, commonly make the assumption that the ground motion intensity level is known and is determined at the centroid of each municipality from a ground motion prediction equation. A flexible Bayesian framework is applied here to the 1980 Irpinia database to explore whether more complex statistical models that account for sources of uncertainty in the intensity can significantly change the shape of the fragility curves. Through this framework the effect of explicitly modelling the uncertainty in the intensity, the spatial correlation of its intra-event component and the uncertainty due to the scatter of the buildings in the municipality are investigated. The analyses showed that the results did not change substantively with increased model complexity or the choice of prior. Nonetheless, informed decisions should be based on the defensible modelling of the significant variability in the data between municipalities.

1 Introduction

Fragility is an important component of seismic risk and expresses the likelihood of damage sustained by buildings with certain characteristics in future earthquake events [1]. Fragility is commonly presented in terms of fragility curves, defined as the probability that a given building damage state will be reached or exceeded for a given ground motion intensity level, which is termed here X for simplicity. This study focuses on the empirical assessment of the fragility of buildings based on post-earthquake damage field data. The GEM Compendium database [2] identified 89 sets of empirical fragility curves for various building classes. Although two thirds of these sets determined X levels from the observed damage, the remaining and most recent curves estimate these levels from pre-selected Ground Motion Prediction Equations (GMPEs). These equations are linear statistical models associated with substantial error, which use ground motion records from past events to predict the X level at a specific site based on its local soil conditions and the characteristics of the seismic source and event. The importance of this error in the prediction of X on the seismic risk assessment of spatially distributed assets is well known (e.g., [3-8]). However, despite being extensively studied in the seismic hazard field, the error component of the GMPEs and spatial correlation of X levels has been largely ignored in the field of empirical fragility assessment of assets. Existing empirical fragility functions are typically based on large databases of observed building damage data that are often aggregated at municipality level. Such municipalities can cover a large geographical area, often with soils of significantly different properties. Yet the assumption is commonly made to represent the ground motion intensity across the municipality with a single best-estimate value of X, estimated at the centroid of the municipality. These values are estimated either by a GMPE [9] or by a ShakeMap [10]. In the latter case, the X best-estimates are determined by a GMPE whose outputs are calibrated with the earthquake ground motion records for the studied event at nearby stations. In 2014, a sensitivity study on fictitious data [11] showed that uncertainty in X can lead to significantly flatter empirical fragility curves. Moreover, the uncertainty in X levels has been explicitly taken into account in the construction of empirical fragility curves of elements of electrical stations affected by successive earthquakes by Straub & Der Kiureghian [12], who developed a Bayesian framework to model multiple sources of uncertainties. More recently, Yazgan [13-15] highlighted the importance of modelling both the uncertainty in X levels and the spatial correlation of the levels in nearby sites using a Bayesian framework. Their framework combined existing analytical fragility curves for 4-storey RC buildings with data from 516 individual buildings affected by the 1999 Düzce and 2003 Bingöl earthquakes.

Despite recent advances, it is not well understood whether the uncertainty in the X levels, their spatial correlation or the presence of ground motion records is important in the construction of empirical fragility curves using large databases, which are aggregated at municipality level. The present paper addresses this. A Bayesian framework is adopted here at it offers a natural and flexible way to assimilate multiple sources of uncertainty in a principled and transparent manner, as well as to combine available damage databases with prior knowledge accumulated over the past 40 years of research in the field of fragility assessment of buildings. Nonetheless, such framework requires a high degree of skill to implement. Therefore, the sensitivity analysis performed here aims to explore whether the explicit modelling of sources of uncertainty typically ignored in the literature leads to a significantly different shape of the fragility curves, and assesses whether a Bayesian framework is needed. In the empirical fragility assessment literature, fragility is typically expressed in terms of the best-estimate fragility curve ignoring the, often substantial, uncertainty in its shape. This uncertainty, however, needs to be quantified in a defensible way if the curves are to be considered useful as part of an informed risk assessment. In the adopted Bayesian framework, this uncertainty can be summarized using credible intervals. Therefore, the relative contribution of modelling different sources of uncertainty around the ground motion as well as the impact of the presence of ground motion records is assessed here by examining how they change the credible intervals of the fragility curves. Overall, five statistical models of increasing complexity are developed as part of this study. The models are then fitted to the well-studied 1980 Irpinia damage database and the constructed fragility curves are compared to existing fragility curves, which are also based on the 1980 Irpinia data.

2 The 1980 Irpinia Earthquake Building Damage Database

On 23rd November 1980, the Campania-Basilicata region was affected by a strong earthquake, with magnitude $M_w = 6.9$. Fig.1 highlights the 41 municipalities [16] for which post-earthquake damage data have been collected. In what follows, the building inventory, the damage scale used for the classification of the sustained damage and the ground motion intensity at each municipality are briefly presented.

Fig.1 Map of the Campania-Basilicata region affected by the 1980 Irpinia earthquake.

2.1 Intensity measure

The Campania – Basilicata region is located in Southern Italy along the Southern Appeninic chain, known for its high seismicity [17]. The 1980 event was the first strong event to occur since 1930. This moderately large event occurred at 19:34 local time on 23rd November 1980. The event was generated by a complex normal fault [18]. Fig.1 depicts the epicenter of the event and the projection of the fault to the surface. In an ideal world, each surveyed building should have a ground motion recording station installed so the actual ground motion intensity level at its known location could be recorded. In reality, there is a general lack of dense networks of recording stations and the buildings, especially in large databases such as the 1980 Irpinia damage database, are aggregated at municipality level. Both the lack of recordings and the data aggregation pose challenges as to how best to determine the actual ground motion intensity.

For the 1980 Irpinia event, the actual ground shaking caused by the main shock was recorded by 17 stations scattered in the Campania-Basilicata region as depicted in Fig.1 [19]. These stations are typically located in municipalities for which no damage data have been collected, with the exception of Arienzo (see Fig.1), where there is a nearby station, but this is not however included in the municipality. Given this, a popular approach is to determine the actual intensity measure level, x_j, at the geometrical centre of a given municipality, *j*, through a pre-selected GMPE. The use of the INGV ShakeMap is not deemed suitable in this study, as it is based on a larger scale geological map (1:100,000) than the one used in this study (1:50,000), it provides PGA estimated aggregated in bins of 0.4g and it does not differentiate between the different sources of uncertainty.

A typical GMPE is a function of the magnitude of the event, *M*, the source-to-point distance, R_j, the soil conditions, S_j, and the fault type, *F*. Nonetheless, the GMPE is not perfect and cannot capture the full variation in intensities, so that it is necessary to introduce a representation of the GMPE errors in order to fully reconcile the GMPE with the actual intensities. In modern GMPEs [20-22], the actual ground motion intensity is determined by explicitly accounting for an event-specific source of error and a spatially varying one. The event-specific error accounts for the fact that there will typically be unobserved features of an event that cause a GMPE to systematically over- or under- predict the actual intensity everywhere for a given event, although it may be unbiased "on average" over a large number of events. The spatially varying error accounts for the possibility that the precise parameterisation of the GMPE may not be appropriate for all events – for example, in a given event the decay rate of intensity with distance from the source may differ from that assumed in the GMPE due to differences of the wave path or local site conditions. In general, a GMPE determines the actual ground motion intensity level at the centroid of a municipality *j*, x_j, as:

$$\ln\left(x_j\right) = f\left(M, R_j, S_j, F\right) + \phi + \varepsilon_j \tag{1}$$

where ϕ is the event-specific error which is typically termed inter-event error; and ε_j which is the error spatially varying within a given event and is known as the intra-event error. Both of these variables are considered normally distributed, with mean equal to 0 and variance equal to σ^2_{inter} and σ^2_{intra}, respectively. Existing empirical fragility assessment studies, however, do not use Eq.(1) to determine the actual intensity levels. Instead, they ignore the two error components and they assign the estimated level of intensity at the centroid of each municipality, which can be obtained by rewriting Eq.(1) as:

$$\ln\left(\tilde{x}_j\right) = f\left(M, R_j, S_j, F\right) \tag{2}$$

The term 'estimated intensity' is used wherever the intensity level is determined by Eq.(2). The focus of the present study is to examine whether the explicit modelling of the actual, rather than the estimated, ground motion intensity leads to significantly different shape of fragility curves.

The identification of the most appropriate GMPE is not straightforward given the years of systematic research on Italian earthquakes, which produced a plethora of GMPEs. For this reason, three recent GMPEs [23-25], which explicitly model the two sources of error, are selected and their main characteristics are presented in Table 1. All three have used the ground motion records from the 1980 Irpinia earthquake and use the Peak Ground Acceleration (PGA) as a measure of ground motion intensity. PGA is adopted here as it is considered an efficient measure to predict the response of low-rise buildings with low ductility [26], which represent the majority of the building inventory in the Campania – Basilicata region affected by the earthquake, and it is a widely used measure of intensity in empirical fragility assessment studies [2].

Table 1: *Main characteristics of the three GMPEs adopted in this study.*

Study	Equation	PGA range in g

Bindi et al. [23]	$\log_{10}\left(x_j\right) = F_D + F_M + F_S + F_{sof} + \phi + \varepsilon_j$, where $\sigma_{inter} = 0.18, \sigma_{intra} = 0.28$	[0.01, 0.64]
Kotha et al. [24]	$\ln\left(x_j\right) = F_D + F_M + F_S + F_{sof} + \phi + \varepsilon_j$, where $\sigma_{inter} = 0.35, \sigma_{intra} = 0.57$	[0.01, 0.43]
Akkar et al. [25]	$\ln\left(x_j\right) = \ln\left[x_{REF}\left(M_w, R, F\right)\right] + \ln\left[\left(S, PGA_{REF}\right)\right] + \phi + \varepsilon_j$, where $\sigma_{inter} = 0.35, \sigma_{intra} = 0.62$	[0.01, 0.44]

In Fig.2, the PGA values from the 17 ground motion records are compared to the estimated PGA values at the location of the 17 stations based on the three GMPEs in the log-log scale. It can be seen that all three GMPEs provide PGA estimates with a sizeable uncertainty. Nonetheless, the Bindi et al. [23] GMPE appears to provide a better fit as the data points seem to be symmetrically scattered around the 45-degree line as opposed to the other two GMPEs [24, 25] which tend to overestimate the PGA values. Having established that Bindi et al. GMPE predicts the recorded PGA levels better than the other two, in section 3.1.1 it is further examined whether Bindi et al. [23] GMPE also fits the damage data better. In this study, the GMPE found to fit the damage data better than its alternatives would be selected to determine the PGA levels required for the fragility assessment.

Fig.2 Recorded (Rec. PGA) PGA values from 17 stations vs estimated (Est. PGA) PGA values for the same locations using the GMPEs proposed by Bindi et al. [23], Kothal et al. [24] and Akkar et al. [25] in log-log scale.

2.2 Building inventory

The database includes information from 29,661 buildings, which are considered to be a representative and unbiased sample of the total number of buildings located in the affected region. 89% of the surveyed buildings are masonry. Information regarding their vertical and horizontal construction materials is included in the database. In Table *2*, it can be seen that masonry buildings are built mainly using field stone (63%), and to a lesser degree hewn stone (32%). Brick masonry buildings appear to be the least common in the affected region. It can also be noted that 47% of the surveyed masonry buildings have wooden floors, followed, in decreasing frequency, by steel (30%) and RC (13%) floors. Masonry buildings constructed with vaults are the least common in the affected region. Reinforced concrete (RC) buildings are also present in the affected region. The RC buildings have been built either without a seismic design code or with an old seismic design code.

Table 2: *Classification of buildings to vulnerability classes according to their vertical and horizontal structure [16].*

Structure	**Vertical**			
Horizontal	*Field Stone*	*Hewn Stone*	*Brick Masonry*	*RC*
Vaults	A (1,532)	A (617)	A (16)	-
Wood	A (8,860)	A (3,294)	C (132)	-
Steel	B (5,216)	B (2,323)	C (468)	-
RC	C (855)	C (2,060)	C (601)	C (3,383)

If the buildings are classified into 13 classes according to their vertical and horizontal structure, this results in some classes having a very small number of buildings (see Table *2*). To avoid very small samples, the 13 building classes are reduced to the three vulnerability classes (i.e., A, B and C) following the re-classification approach found in Braga et al. [16] (see Table *2*). The vulnerability class ‘A’ is the best represented class in the database with overall 14,406 buildings. Class A is considered the most vulnerable class as it includes the worst quality masonry buildings. Poor quality masonry buildings are considered to have floors made of vaults or to have been constructed using field or hewn stone and wooden floors (see Table 2). The damage database also includes 7,816 class B buildings. This class includes the better quality field or hewn stone buildings, which have steel floors. Finally, Class C includes the least vulnerable buildings, having similar sample size with the latter class. The present study concentrates on the most vulnerable (Class A) and the least vulnerable (Class C) classes of buildings to present the results.

2.3 Damage scale

The damage sustained by the surveyed buildings has been classified into six discrete damage states according to the MSK-76 scale [27]. The six damage states characterize all possible levels of damage that a building can suffer, and are described briefly in Table 3.

Fig.3 shows the relationship between the estimated *PGA* level in each municipality and the percentages of Class A and C buildings sustaining damage state ds_i or above (i =1, 2, 5). Overall, damaged data from 14,406 Class A and of 7,439 C buildings aggregated in 41 data points. This is substantially larger than the minimum sample size required for the construction of meaningful fragility curves provided by the Global Earthquake Model, empirical vulnerability assessment guidelines [30]. According to the latter, at least 200 buildings are required to be aggregated in a minimum of 10 data points. Over 80% of the data points, (representing 70% of the building inventory), are clustered in the lower PGA levels, i.e., 0.024g to 0.30g. This scarcity of data for the higher intensity measure levels is not unusual. Moreover, the total number of recorded buildings in each municipality varies according to the building class: for example, there are no municipalities with fewer than 50 Class A buildings, but 35 municipalities with fewer than 50 Class C buildings.

Table 3: *Description of damage in each damage state [27].*

DS	**Description:**
ds_0	No damage
ds_1	Negligible or slight damage
ds_2	Moderate damage
ds_3	Substantial to heavy damage
ds_4	Very heavy damage
ds_5	Destruction

In Fig.3, the poor performance of class A buildings in the 1980 earthquake is highlighted, as in all 41 municipalities at least half of class A buildings sustained some level of damage (i.e., $DS \geq ds_1$) and 6% of the total class A buildings located in over 70% municipalities had been destroyed. Overall, each municipality appears to have recorded more damaged class A buildings with $DS \geq ds_i$ than class C buildings. In addition, the scatter in their data points appears to be wider than for class C buildings for higher damage states; indicating that there is greater uncertainty in the seismic performance of class A buildings for higher damage states. This is in line with observations in the literature regarding the high uncertainty in the seismic performance of low quality masonry buildings [28].

The better seismic performance of class C buildings is also depicted in Fig.3, where in most (71%) municipalities, less than half of the surveyed Class C buildings sustained some level of damage. The highest scatter in the data points can be seen for ds_1 and this substantially reduces for the higher damage states. A closer examination of the data shows that the highest percentage of buildings in any given municipality sustaining moderate damage or above (i.e., $DS \geq ds_2$) never exceeds 40%. and that in 80%

of the municipalities no Class C buildings have been destroyed (n.b. there are only 52 cases of collapsed Class C buildings). Therefore, very little information is available from which to construct fragility curves for the most severe damage states, for buildings with good seismic performance. This is a known issue in the empirical fragility assessment literature [29] and limits the ability of empirical approaches to determine the likelihood of collapse of buildings which perform well during earthquakes.

3 Empirical fragility assessment

The information regarding the damage, the intensity levels and building class are combined in order to empirically assess the fragility of class A and C buildings for three damage states (i.e, ds_1, ds_2 and ds_5). Two simplifications, are deemed necessary to substantially reduce the amount of time required to run each Bayesian analysis. Firstly, the fragility curves are constructed independently for each damage state instead of exploiting the ordinal nature of the damage scale. Secondly, the fragility curves are constructed independently for Class A and C.

Past studies commonly base their empirical fragility assessment of buildings on three main assumptions regarding the quality of the post-earthquake database. According to these assumptions, the database is considered representative of the seismic damage in the affected area, the misclassification error in assigning the ‘actual’ damage states to each building is considered insignificant and the uncertainty in the X level can be ignored [30]. In what follows, a reference model based on these assumptions is developed and fitted to the 1980 Irpinia data. Then, the importance of different sources of uncertainty in the ground motion intensity is examined by developing and fitting four models of increasing complexity. At each stage, the effect of accounting for additional information or uncertainties is assessed by examining the fitted fragility curves and their credible intervals. Finally, the most realistic fragility curves are constructed by combining the 1980 Irpinia damage data with prior information regarding the fragility of the Class A and C buildings.

3.1 Models M0-M1: Reference models

Many existing studies adopt a parametric statistical model, whose systematic component is typically expressed in terms of the cumulative lognormal distribution, and with random component following various assumptions that are often unrealistic and whose impact on the fragility is discussed in greater detail in [31]. The present study adopts a Generalised Linear Model (termed GLM) which is proposed by the Global Earthquake Model (GEM) Guidelines [30]. A GLM assumes that the number of buildings, y_{ij}, which sustained damage $DS \geq ds_i$ in municipality, j, follows a binomial distribution:

$$y_{ij} \sim Binomial\left(n_j, \pi_i\left(\tilde{x}_j\right)\right) \quad (3)$$

where $\pi_i\left(\tilde{x}_j\right)$ is the probability that a building located in municipality, j, will reach or exceed the ‘true’ damage state, ds_i, given estimated intensity level $\tilde{x}_j$; n_j is the total number of buildings of the examined building class in municipality, j. The binomial distribution is characterised by its mean:

$$\mu_{ij} = n_j \pi_i\left(\tilde{x}_j\right) \quad (4)$$

which is expressed here in terms of a probit model defined in terms of $\Phi(.)$, the cumulative distribution function of a standard normal distribution:

$$\Phi^{-1}\left[\pi_i\left(\tilde{x}_j\right)\right] = \eta_{ij} \quad (5)$$

where η_{ij} = is the linear predictor, which can be written in the form:

$$\eta_{ij} = \theta_{0i} + \theta_{1i} \ln\left(\tilde{x}_j\right) \quad (6)$$

Here, θ_{1i}, θ_{0i} are the two regression coefficients, representing the slope and the intercept, respectively, of the fragility curve corresponding to damage state ds_i. For the reference model, the ground motion intensity level, $\tilde{x}_j$, is estimated at the geographical centre of municipality j from a pre-selected GMPE, ignoring the two error components in Eq.(1). Throughout the remainder of the paper, the reference model fitted to the data using a maximum likelihood approach is referred to as M0; the same model fitted using a Bayesian approach is denoted M1. We next describe the two approaches, and highlight the differences between them.

3.1.1 M0 – The maximum likelihood approach

In a maximum likelihood approach, the regression coefficients and their standard error are determined from the log-likelihood function [32], as:

$$\boldsymbol{\theta} = \arg\max L(\boldsymbol{\theta}) = \arg\max \log\left\{\prod_{j=1}^{M}\left[\binom{n_j}{y_{ij}}\mu_{ij}^{y_{ij}}\left(1-\mu_{ij}\right)^{n_j-y_{ij}}\right]\right\} \quad (7)$$

where M is the total number of municipalities. It should be noted that for data aggregated at municipality level, the variability between municipalities is often greater than expected under the binomial distribution: a phenomenon known as "over-dispersion", which can be caused due to aggregation of non-homogeneous data in municipalities or the failure to account for other explanatory variables (for example, because their data are not available). Where over-dispersion occurs, standard errors for the regression coefficients will be underestimated, which in turn leads to underestimation of the uncertainty in the estimated fragility curves. A standard way to deal with over-dispersion in such situations is to carry out a "quasi-binomial" GLM fit [33], which provides an empirical adjustment to the standard errors to ensure that they correctly reflect the estimation uncertainty [34]. M0 is fitted to the data with estimated PGA levels from the three GMPEs. Due to the limited space in this study, the results are presented only for the GMPE that provides the best fit to the data. However, a sensitivity analysis is presented in the Appendix in which multiple GMPEs are used. To identify which GMPE fits the damage data best, the maximised log-likelihood for all fitted models are compared in Table 4. The Bindi et al [23] GMPE is found systematically to have the highest log-likelihood by a considerable margin: this means that the PGA levels obtained by this GMPE fit the data best.

Table 4: Maximized log-likelihood values for the three GMPEs.

GMPE	ds_1	ds_2	ds_5
Class A			
Bindi et al. [23]	**-498.2**	**-981.5**	**-559.7**
Kotha et al. [24]	-612.4	-1305.5	-966.2
Akkar et al. [25]	-589.1	-1249.1	-859.7
Class C			
Bindi et al. [23]	**-428.3**	**-269.3**	**-44.6**
Kotha et al. [24]	-552.7	-430.3	-98.0
Akkar et al. [25]	-561.1	-435.7	-95.0

Fig.3 depicts the fragility curves corresponding to ds_1, ds_2 and ds_5, obtained using maximum likelihood estimates of the regression coefficients and using the Bindi et al. [23] GMPE to determine the PGA levels at the geographical centre of each municipality. Within the range of estimated PGA levels corresponding to the municipalities for which there are available damage data, the best estimate curves represent a small part of the lognormal cumulative distribution function, rather than the full range from 0 to 1. Overall, the best-estimate fragility curves for Class A buildings appear to be systematically higher than their Class C counterparts, highlighting the poor seismic performance of the Class A buildings. The values of the regression coefficients are presented in Table *5*.

Table 5: *Regression coefficients estimates based on M0.*

DS	*Building Class*			
	A		*C*	
	θ_0	θ_1	θ_0	θ_1
ds_1	0.16	0.25	-1.51	0.27
ds_2	-1.23	0.29	-3.27	0.40
ds_5	-4.73	0.60	-6.14	0.67

In Fig.3, the 90% confidence intervals for the 'true' fragility curves are plotted accounting for the sampling variation in the damage data according to the fitted "quasi-binomial" model.

3.1.2 M1 – The Bayesian approach

In maximum likelihood estimation, uncertainties about unknown quantities are expressed using standard errors and confidence intervals. In Bayesian inference by contrast, such uncertainties are expressed by assigning probability distributions directly to those quantities. Using Bayes' theorem, the likelihood function is combined with prior knowledge regarding the probability distribution of the regression coefficients in order to obtain the posterior distribution of these coefficients:

$$p(\boldsymbol{\theta}|Y) \propto p(\boldsymbol{\theta}) L(\boldsymbol{\theta}) = p(\boldsymbol{\theta}) \prod_{j=1}^{M} \left[\binom{n_j}{y_{ij}} \mu_{ij}^{y_{ij}} \left(1-\mu_{ij}\right)^{n_j - y_{ij}} \right] \quad (8)$$

where $p(\boldsymbol{\theta})$ is the prior distribution of the vector of parameters $\boldsymbol{\theta}=[\theta_{0i}, \theta_{1i}]$; $L(\boldsymbol{\theta})$ is the likelihood function. In the present work, the posterior distribution of the regression coefficients, together with all other quantities of interest, is estimated by the Markov Chain Monte Carlo (MCMC) algorithm, using OpenBUGS [35]. All posterior distributions of interest are determined from three chains with 90,000 iterations each, ignoring the first 30,000 iterations. The simulations are fast: each lasts approximately 30 seconds. The convergence of the MCMC algorithm is assessed by the Gelman-Rubin diagnostic [36], known as Rhat statistic, which is found to be equal to 1 (to two significant digits) for all analyses, indicating the successful convergence of all chains.

The prior distribution provides a mechanism for incorporating additional information into the analysis to supplement the available post-earthquake data, if such information is available. For compatibility with model M0 however, initially we make no attempt to incorporate such information. Thus, the prior distribution for both regression coefficients (i.e., θ_{0i} or θ_{1i}) is taken to be normal, with zero mean and a very large variance:

$$\begin{aligned} \theta_{0i} &\sim Normal\left(\mu_{\theta_{0i}} = 0, \sigma^2_{\theta_{0i}} = 10^5\right) \\ \theta_{1i} &\sim Normal\left(\mu_{\theta_{1i}} = 0, \sigma^2_{\theta_{1i}} = 10^5\right) \end{aligned} \quad (9)$$

where $\mu_{\theta 0i}$, $\mu_{\theta 1i}$ is the prior mean of intercept θ_{0i} and slope θ_{1i}, respectively; $\sigma^2_{\theta_{0i}}$ and $\sigma^2_{\theta_{1i}}$ is their variance.

For model M0, a quasi-binomial fit was used to account for over-dispersion. This type of empirical adjustment is not possible in a Bayesian framework however, so an alternative approach must be sought. In model M1 therefore, over-dispersion is explicitly modelled by adding a random effect at municipality level to the linear predictor [37]. Equation (4) is therefore modified to:

$$\eta_{ij} = \theta_{0i} + \theta_{1i} \ln\left(\tilde{x}_j\right) + \xi_j \quad (10)$$

where the $\{\xi_j\}$ are random variables, realised independently for each municipality, which are considered normally distributed with:

$$\xi_j \sim Normal\left(\mu_{\xi_j} = 0, \tau^2\right) \qquad (11)$$

with τ^2 is the variance. The size of this variance determines the size of the over-dispersion [37]. For example, when τ^2=0, the model reduces to the binomial, otherwise the over-dispersion is taken into account. In OpenBUGS [35], the normal distribution is modelled in terms of its inverse variance instead of its variance. The inverse variance, termed precision, is here assigned a non-informative prior distribution in the form of a gamma distribution with mean 1 and variance equal to 100.

In Fig.3, the best estimate fragility curves obtained from the Bayesian approach are compared to their counterparts obtained by the maximum likelihood approach. The 90% credible intervals around the best estimates are also plotted. These credible intervals are much wider than the corresponding confidence intervals from the maximum likelihood method. This is predominantly due to the slightly different model formulations in the two methods. In model M0, a single fixed fragility curve is assumed for all municipalities and the scatter of observations is accounted for via an empirical adjustment for overdispersion: the confidence intervals for this model represent uncertainty about the single fixed curve. In model M1 however, the scatter of observations is attributed to municipality-specific fragility curves defined via the random effects in Eq.(10): the credible intervals here represent uncertainty about the fragility curve for an individual municipality. The differences are most notable for the curves corresponding to ds_2 for Class A and ds_1 for Class C, which are associated with the highest scatter in the data points. It should be noted that a smaller sample size would increase the uncertainty around the fragility curves leading to wider credible intervals. That would reduce the differences between the models.

Fig.3 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% confidence and the 90% credible intervals comparing M0 vs M1 are also presented.

3.2 Exploring the importance of the uncertainty in *X*

A sensitivity analysis is performed in order to examine whether the shape of the fragility curves is influenced by explicitly accounting for the error component in the GMPE, the spatial correlation in the intra-event component, the uncertainty due to the spread of the buildings in each municipality and the presence of ground motion records. To achieve that, four models (termed M2-M) are constructed here which explicitly account for the additional sources of uncertainty, the spatial structure of the data or the presence of the ground motion records. The models increase in complexity as depicted in Table 6.

Table 6: Summary of the five models used in this study.

Model	Over-dispersion	Uncertainty in X	Spatial Correlation in intra event component	Uncertainty due to the spread of buildings in each municipality	Presence of ground motion records.
M1	x	-	-	-	-
M2	x	x	-	-	-
M3	x	x	x	-	-
M4	x	x	x	x	-
M5	x	x	x	x	x

The fits of these models to the 1980 Irpinia data using the Bayesian approach are then compared to each other and to the reference model, M1. The fit is compared in terms of the best-estimate fragility curve as well as the 90% credible intervals. The conclusions were found to be the same if alternative credible intervals levels (e.g., 75% or 95%) were preferred. In order for the results to be directly comparable, non-informative prior distributions are assigned to the two regression coefficients: the effect of prior choice is addressed later. For these models, the posterior distributions of the probability that a damage state will be reached or exceeded by the buildings of a given class for a range of PGA levels are determined by three MCMC chains with 150,000 iterations each, ignoring the first 30,000 iterations. Each analysis lasts approximately 12 minutes. The convergence of the MCMC algorithm is assessed by the Rhat convergence statistic, which is found to be approximately equal to 1 for all analyses, indicating the successful convergence of the three chains.

3.2.1 M2 – Accounting for the error components in GMPE

A new model M2 is constructed here, which shares Eq.(2) to Eq.(6) with M1, but considers the actual ground motion intensity, x_j, at the center of municipality j by explicitly modelling the two error terms in Eq.(1). The inter-event error ϕ is common to all municipalities. As a first attempt to account for the intra-event errors $\{\varepsilon_j\}$, they are considered to be mutually independent in model M2 and to follow identical normal distributions:

$$\begin{aligned} \phi &\sim N\left(\mu_\phi, \sigma_\phi^2\right) \\ \varepsilon_j &\sim N\left(\mu_{\varepsilon_j}, \sigma_{\varepsilon_j}^2\right) \end{aligned} \qquad (12)$$

where μ_η, $\mu_{\varepsilon j}$ are the mean, which are considered equal to zero; $\sigma_\eta^2, \sigma_{\varepsilon_j}^2$ are the variance of the residuals. The average values of the variances are provided by the adopted GMPE and for this reason they are assigned informed prior distributions. Following standard practice, gamma distributions are assigned to the corresponding precision parameters, with mean and variance chosen to represent a judgement that 90% of the individual standard deviations are likely to lie within 20% of the provided values:

$$\begin{aligned} 1/\sigma_\phi^2 &\sim Gamma\left(mean = 1/0.18^2, Variance = 65.92\right) \\ 1/\sigma_{\varepsilon_j}^2 &\sim Gamma\left(mean = 1/0.28^2, Variance = 12.11\right) \end{aligned} \qquad (13)$$

For more information regarding the prior distributions and their parameters, the reader is referred to the comments in the file 'M5.R' uploaded as supplementary material accompanying this article.

In Fig.4, the best estimate fragility curves and their 90% credible intervals obtained by fitting the model M1 using the Bayesian approach are compared to their counterparts obtained by fitting M2. It can be seen that there is practically no difference in the best-estimate fragility curve between the two models. The credible intervals of M2 tend to be close and somewhat narrower than for M1. The differences appear to be notable but not significant for ds_5 for both building classes.

Fig.4 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best-estimate fragility curves and their corresponding 90% credible interval comparing M1 vs M2 are plotted.

3.2.2 M3 – Accounting for the error components in GMPE and spatial correlation

In model M2, the intra-event errors $\{\varepsilon_j\}$ were considered to be independent at each municipality. In reality however, if a GMPE over / underpredicts the ground motion intensity in one municipality for an event, it is conceivable that there will be a tendency to over / underpredict in neighbouring municipalities as well: the independence assumption is questionable, therefore. To allow for such dependence requires

a more complex statistical model, but it is also potentially beneficial because it allows the possibility for information to flow between municipalities: this can be particularly helpful when data from some municipalities are sparse, since the characteristics of the GMPE errors there can be inferred (with appropriate consideration of uncertainty, via the Bayesian approach to inference) from more data-rich locations nearby. Model M3 therefore removes the restriction that the intra-event errors are independent.
The significance of the intra-event spatial correlation in seismic risk assessment has generated an extensive literature aiming to determine the spatial correlation structure at given sites. Published studies exclusively consider that the natural logarithm of X levels in multiple sites follow a multivariate normal distribution [38]. The mean of this distribution is provided by Eq.(1) ignoring the two random effects, and the covariance matrix is determined as:

$$\Sigma = \sigma_{inter}^2 \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & & \vdots \\ \vdots & & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix} + \sigma_{intra}^2 \begin{bmatrix} 1 & \rho(h_{12}) & \cdots & \rho(h_{1m}) \\ & 1 & & \vdots \\ & & \ddots & \\ sym & & \cdots & 1 \end{bmatrix}$$

(14)

where $\rho(h_{jk})$ is the spatial correlation coefficient which is a function of the distance h_{jk} between the centroids of municipalities j and k, and is expected to decay with increasing distance. In a Bayesian context however, the use of correlation-based dependence models can be computationally expensive, because the intra-site correlation matrix must be inverted at each iteration in order to calculate the likelihood contribution to the posterior distribution [39]. With m municipalities in total, the computational cost of inverting a covariance matrix via any standard algorithm is roughly proportional to m^3: to do this at each MCMC iteration (we use 150,000 iterations per model for the results reported below, which is not an unusually high number, and is considered necessary in order to get a large effective sample size for the more complex models considered here) is simply not feasible for moderate or large numbers of locations. For this reason, MCMC-based Bayesian inference for spatial datasets is usually done using conditional autoregressive (CAR) representations of dependence, in which the distribution of the quantity of interest at each spatial location is specified conditional on the values at neighbouring locations. The CAR approach has gained wide acceptance in other fields (e.g., ecological studies and disease mapping [40]).

In M3, the intrinsic CAR model [41] (i.e., the simplest form of a CAR model) is used to capture the spatial correlation of the intra-event component of the uncertainty in the ground motion intensity. According to this model, the error ε_j in municipality j depends on the values of the errors in the neighboring municipalities, (defined here as those with which municipality j shares a border), and is modelled as a normal distribution:

$$\varepsilon_j \mid \boldsymbol{\varepsilon}_{i,i\neq j} \sim Normal\left(\frac{\sum w_{ji}\varepsilon_i}{\sum w_{ji}}, \frac{\sigma_{CAR}^2}{\sum w_{ji}} \right) \quad (15)$$

where the σ_{CAR}^2 is the unknown variance parameter; and w_{ji} are weights which account for the proximity of two municipalities, and can be written in the form:

$$w_{jk} = \begin{cases} 1 & \text{If } k \text{ and } j \text{ are neighbour municipalities} \\ 0 & \text{If } k \text{ and } j \text{ are not neighbour municipalities} \end{cases} \tag{16}$$

The variance of the CAR model, $\sigma_{CAR}^2 / \sum w_{ji}$ (see Eq.(15)) controls the variability of the effect in municipality *j* conditional on the effects of neighbour municipalities, and also depends on the number of these municipalities [42]. The intuition here is that the magnitude of the error in municipality *j* can be determined more precisely if the errors in a large number of neighbours are already known, than if there is little neighbourhood information. Given the lack of prior information regarding the variance of the spatially distributed effects, an uninformative prior distribution is assigned to the inverse of σ_{CAR}^2, in the form of a gamma distribution which is assumed to have mean equal to 1 and variance equal to 100. It should be noted that an informative prior based on existing spatial models (e.g., [38]) is not feasible as the CAR dependence structure cannot be directly related to a model for inter-station correlations.

In Fig.5, the best-estimate fragility curves corresponding to ds_1, ds_2 and ds_5 and their 90% credible intervals obtained by fitting M3 and M1 to the 1980 Irpinia data for class A and C are depicted. The best-estimate fragility curves for M3 appear to be identical to those for M1 . The differences between the credible intervals of the two models also appear to be negligible irrespective of damage state or class.

Fig.5 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% credible interval comparing M1 vs M3 are also plotted.

3.2.3 M4 – Accounting for the error components in GMPE, spatial correlation and uncertainty in IMLs due to the scatter of buildings in the municipality

In the empirical fragility assessment of aggregated post-disaster data, it is commonly assumed that the ground motion level is estimated at the geometrical centre of each municipality. For the Irpinia database, this means that municipalities with hundreds or thousands of buildings, are all assigned the same ground motion level. This is an unrealistic assumption as buildings can be far away from the geometrical centre of the municipality or be sited on soils with very different properties. The within-municipality variation is expected to decrease with distance from the fault (due to an exponential decay in ground motion with distance from the fault). Model M4 is designed to represent this variation in a simplified manner, and hence to test the sensitivity of the estimated fragility curves to the common assumption that all buildings within a municipality experience the same ground motion. To represent the variation, the area in every municipality is subdivided into a regular grid with cell size 1km x 1km. The ground motion intensity level at each centre is estimated by Eq.(2). The variation of these estimates is then calculated in order to determine an index of variation for that municipality. This approach essentially allows greater variation of the intra-event errors for those municipalities where the ground motion is more variable, thus making some allowance for the fact that the precise building locations are unknown.

In M4, the uncertainty due to the spatial distribution of the buildings in each municipality is modelled by adding a municipality-level random effect component at Eq.(1) in the form:

$$\ln\left(x_j\right) = f\left(M, R_j, S_j, F\right) + \phi + \varepsilon_j + w_j \tag{17}$$

where w_j is assumed to follow a normal distribution with:

$$w_j \sim N\left(\mu = 0, \mathrm{var}_{w_j}\right) \tag{18}$$

The inverse variance is also assigned an informative gamma distribution prior concentrated around the inverse sample variance of the ground motions across the municipality. It should be noted that unlike ε_j, w_j is not spatially correlated and for this reason, the two variables are modelled separately.

In Fig.6, the best estimate fragility curves and their 90% credible intervals for models M4 and M1 appear to be almost identical. Compared to the fit of M3 (see Fig.5), the addition of an extra random component in the estimation of the ground motion intensity in M4 leads to small and non-systematic differences. The fit of M4 appears to have a negligible effect on the width of the credible intervals for the curves corresponding to ds_1, ds_2 and ds_5 for Class A The width of the intervals corresponding to and ds_2 and ds_5 for Class C are reduced, whilst the width of the intervals corresponding to ds_1 for Class C remains unchanged.

Fig.6 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best-estimate fragility curves and their corresponding 90% credible interval comparing M1 vs M4 are also plotted.

3.2.4 M5 – Accounting for the error components in GMPE, spatial correlation, uncertainty in IMLs due to the scatter of buildings in the municipality and known IMLs

The models proposed so far have ignored the presence of observed ground motion records from the studied earthquake. This, however, is not realistic as there are often multiple records, which can be used to constrain the uncertainty in the X levels. M5 is the same as M4, the only difference between the two models is that in M5 the PGA level in the 17 municipalities for which there are ground motion records, is determined by these records. The PGA level at the remaining municipalities is estimated from the Bindi et al. [23] GMPE, by explicitly accounting for the two error components, the spatial correlation of the intra-event component and the uncertainty due to the spatial variation of the building inventory in each municipality.

In Fig.7, the best estimate fragility curves and their 90% credible intervals obtained by fitting M5 and M1 to the 1980 Irpinia data can be compared. The best-estimate curves and their 90% credible intervals appear to be identical for ds_1 and ds_2 for both building classes. The differences in the two models appear to be notable for ds_5, where the credible intervals appear to be reduced for M5. The fit of M5 is then compared to the fit of M4. The differences appear to be small, which is in line with similar observations made in the literature [7, 11] regarding the small impact of a few ground motion stations in the estimation of the likelihood of damage and highlight once more the need for a denser network of ground motion stations.

Fig.7 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% credible intervals comparing M1 vs M5 are also plotted.

3.3 Exploring the impact of informative priors

In models M1 to M5, the uninformative priors of the regression coefficients ignore the wealth of research regarding the fragility of the Italian building inventory. A search in the GEM compendium [43] identified twelve studies [9, 10, 16, 29, 44-50] that can be used to determine the fragility of Class A and C

Italian buildings. These include five empirical studies, six analytical and one heuristic fragility assessment study. It should be noted that the 1980 Irpinia earthquake is a well-studied event and its post-disaster data has been used by all the empirical fragility studies chosen [9, 10, 16, 29, 44]. Therefore, the 5 empirical fragility functions are excluded as sources of prior information regarding the regression coefficients, as they do not add new information regarding the shape of the fragility curves. Instead, the sets of fragility curves found in the six remaining analytical [45-49] and heuristic [50] studies are used here to determine the prior distributions.

In Fig.8, the plots of the slope against the intercept of the existing curves for the three damage states and the two building classes are depicted. Overall, the variation in the values of the slope and intercept is dominated by the variation between studies, which is substantially larger than the variation for individual studies that have constructed multiple fragility curves. It can also be noted that the intercept appears to be positively correlated to the slope. Nonetheless, this correlation can be attributed to the variation between studies. For individual studies which constructed multiple fragility curves there is little evidence of such correlation and for this reason, the correlation between the two coefficients is ignored.

In using this information to set informed priors for the two coefficients, the intercept is assigned a normal distribution as it can take both positive and negative values and the slope is considered to follow a gamma distribution, which accounts for the expectation that the increase in the ground motion intensity will increase the probability of a building to be damaged. The informative prior distributions for the two regression coefficients can be written as:

$$\begin{aligned} &\theta_{0i} \sim Normal\left(\mu_{\theta_{0i}}, \sigma^2_{\theta_{0i}}\right) \\ &\theta_{1i} \sim Gamma\left(\mu_{\theta_{1i}} \sigma^2_{\theta_{1i}}\right) \end{aligned} \qquad (19)$$

The mean and the variance of these prior distributions are determined from the existing fragility curves. In particular, the prior mean intercept and slope (i.e., $\mu_{\theta_{0i}}$ and $\mu_{\theta_{1i}}$) are set to the means of the corresponding values from the existing fragility curves. In setting the prior variances however, it is prudent to allow for a larger range of coefficient values than is present in the small number of existing fragility curves. The prior variances in equation (19are therefore set at 16 times the sample variance of the existing intercepts and slopes.

Fig.8 Plots of slopes against the intercept for existing fragility curves corresponding to 3 damage states used as priors for class A and C.

In Fig.9, the fragility curves obtained by fitting model M5 to the data using informative (M5-Info) and non-informative (M5-Uninfo) priors are plotted. The differences in the best-estimate curves as well as in their 90% credible intervals are not consistent but depend on the building class and damage state. Overall, the differences can be considered negligible or small with the exception of the collapse fragility curve (i.e., ds_5) for Class C buildings, which is associated with very small probabilities of collapse. For this case, the M5-Info leads to flatter best-estimate fragility curves and notably narrower credible intervals. These observations can be attributed to the relatively large sample sizes of Class A and C buildings in the 1980 Irpinia database, which ensure that the observations are highly informative and dominate any effect of the prior specification.

The existing fragility curves from which the priors were derived are also depicted in Fig.9. It can be noted that the analytical curves appear to be notably steeper than the best estimate fragility curves irrespective of whether informative or non-informative priors have been adopted. By contrast, the heuristic ones are more in line with the empirical ones, although this is not systematic. The systematic difference in steepness between the analytical and empirical curves could be attributed to the nature of the analytical studies which are typically based on a single well defined building exposed to known ground motion excitations. For this reason, they tend to result in steep fragility curves. By contrast, the empirical curves are based on large databases that include large variations in the performance and geometric characteristics of buildings in the same class as well as the variability in the ground motion excitation whose characteristics are generally not known.

Fig.9 Prior fragility curves corresponding to 3 damage states for class A and C as well as the best estimate fragility curves and their 90% credible intervals based on M5 assuming informative (M5-Info) and non-informative (M5-Uninfo) priors.

Fig.10 shows our estimated fragility curves based on M5-Info together with their counterparts based on M1 with uninformative priors. The differences in the credible intervals for the two models appear to be small, with the exception of the collapse fragility curves for Class C buildings. In this latter case, the model M5-info yields significantly narrower credible intervals than model M1. Fig.10 also depicts the curves constructed by the five [9, 10, 16, 44] other existing empirical studies that adopted the 1980 Irpinia damage database for class A and C buildings either on its own or among other databases from Italy or worldwide. Overall, the results show a significant variation in the fragility curves. This is most likely due to differences in data included in the studies (many studies include data sets from several earthquakes), how the data is manipulated for use in the construction of the fragility functions (e.g. whether the data is complete, biased etc.), differences in ground motion intensity estimation approach and due to differences in the statistical model used for their development. Such variations in empirical fragility functions developed from same/similar data is not uncommon, as seen in [51]. Interestingly, the variation is large both for Class A buildings, (which are known for their poor performance), and Class C buildings, (which are expected to perform better).

Fig.10 Plots of best-estimate fragility curves and their 90% credible intervals based on M5 with informative priors and M1 with uninformative priors. Fragility curves from five studies which also used the 1980 Irpinia damage database are also included.

3.4 Prediction of collapsed buildings

So far, this paper has focused on the construction of fragility curves. Despite the importance of such curves in the seismic risk assessment of the buildings inventory, decision makers often need answers to questions such as how many buildings are expected to collapse in a future event. The adopted Bayesian framework can be used to make these predictions, quantifying the uncertainty involved in the expected number of collapsed buildings. In general, predictions within as well as beyond the range of the data are possible. However, due to the limited information which produced fragility curves being only part of the lognormal cumulative distribution function, rather than its full range from 0 to 1, the credible intervals outside the provided range of data should be used for a qualitative assessment of the fragility. Fig.11 depicts the posterior distributions of the number of Class A and C buildings expected to collapse in three municipalities in a future repetition of the 1980 Irpinia earthquake. The predictions concern three municipalities (see Fig.1) associated with low, medium and high estimated PGA levels. It can be seen that there is higher uncertainty in predicting the collapsed Class A buildings for all three municipalities. By contrast, the predictions of collapsed Class C buildings are associated with smaller uncertainty, and the collapses of multiple buildings are expected only for the municipality very close to the epicentre.

Fig.11 Histogram of the posterior distribution of the number of buildings expected to be destroyed (i.e., to sustain damage $DS=ds_5$) in a future earthquake in three municipalities which are exected to be *affected by very low (0.03g), medium (0.09g) and high intensity (0.64g) levels (based on fitting the M1 model with uninfromative priors of regression coefficients).*

In the preceding sections, the differences between the models presented in this paper have been studied by visual inspecting the best-estimate fragility curves and the length of the corresponding 90% credible intervals. Notwithstanding the ease of intepretation of the results of such a qualititative approach, the strength of the conclusions is reinforced here by quantifying the differences between the models in predicting the number of buildings likely to suffer a given damage state. The focus is on the state of collapse, as the differences between the models were found to be more pronounced at this damage state. Table *7* depicts the relative change in the predictions of the number of Class A or C buildings likely to collapse in the municipality with the largest PGA value, based on models M2-M5 for uninformative prios and M5 for informative priors compared to model M1. The relative change in the best estimate and the length of the 90% credible intervals are both recorded. For this extreme case, the differences can be notable for Class A buildings for both the best estimate and the length of the credible intervals. The largest relative change is noted for the most complex M5 model with informative priors. For this model, the expected number of collapsed buildings is 14% lower than for M1 and the credible intervals are 20% narrower than for M1. The differences between the models are much higher for Class C buildings. In particular, for M5-info, 46% fewer buildings are expected to collapse than for M1 and the 90% credible intervlas appear 59% narrower than for M1. These observations highight that the more complex model has a substantial impact only for the probability of collapse of Class C buildings, where the probability of collapse is low.

Table 7: Predicted numbers of collapsed Class A and C buildings in Lioni, a municipality with a total of 356 Class A and 604 Class C buildings exposed to PGA=0.64g as well as the relative change of the predictions based on M2-M5 using uninformative priors and M5 using informative priors compared to M1.

Class	Model	Values Best Estimate (95%, 5%)	Interval length	Relative change compared to M1	
				Best Estimate	Interval length
A	M1	53 (159, 3)	156		
	M2	54 (139, 4)	135	-1%	13%
	M3	50 (146, 2)	144	6%	8%
	M4	50 (143, 3)	140	7%	10%
	M5	49 (131, 4)	127	9%	19%
	M5_Info	46 (127, 2)	125	14%	20%
C	M1	28 (116, 0)	116		
	M2	24 (95, 0)	95	13%	18%
	M3	24 (108, 0)	108	14%	7%
	M4	21 (83, 0)	83	24%	28%
	M5	21 (69, 0)	69	23%	41%
	M5_Info	15 (47, 0)	47	46%	59%

4 Conclusions

A Bayesian framework is used in this study in order to examine the importance of the uncertainty in the ground motion intensity in the shape of fragility curves based on post-disaster data aggregated at municipality level. The advantage of such framework is its ability to combine prior information regarding the shape of the fragility curves with post-disaster data and its flexibility in modelling additional sources of uncertainty. The novelty of the framework includes the use of a CAR model to model the spatial correlation in the intra-event component and the modelling of the uncertainty due the scatter of the buildings in the municipality.

The framework was applied to the 1980 Irpinia earthquake building damage database, which includes damage data from 21,845 Class A and C buildings aggregated in 41 municipalities and the ground motion intensity is estimated by a GMPE. Five models of increasing complexity were constructed in order to account for the uncertainty in the ground motion, the spatial correlation of its intra-event component and the uncertainty due to the scatter of the buildings in the municipality as well as to account for the known ground motion intensity records. The fit of these models to the data was compared to the fragility curves constructed by fitting a reference model using maximum likelihood analysis to determine the regression coefficients.

The analyses show that more complex models (e.g. M2-M5) yield almost identical results to those obtained from the reference model M1 that uses the estimated, instead of the actual, ground motion intensity and ignores the sources of uncertainty associated with the ground motion intensity or the presence of ground motion records. This suggests that the studies with aggregated post-disaster data did not introduce an error to the shape of the fragility curves by ignoring the uncertainty in the ground motion intensity. It was also noted that due to the large number of post-disaster data, the prior information was found to have a negligible impact on the shape of the fragility curves constructed here with the exception of the collapse fragility curves for Class C buildings, associated with low probability of exceedance. Finally, the analyses also highlight the need to appropriately model the significant over-dispersion in the data (i.e., the variation between the municipalities), which is typically ignored in the literature. Hence, the use of the reference model M1 can appropriately capture the uncertainty in the damage data and provide informed predictions regarding the number of buildings likely to suffer damage in a future earthquake. This has important implications for the seismic risk assessment of the building inventory as well as for decision makers interested in informed predictions of the vulnerability of the inventory in future events.

Acknowledgement

The time of Dr Ioanna Ioannou and Professor Tiziana Rossetto on this research work was supported by the European Research Council URBANWAVES grant (ERC Starting Grant 336084; awarded to Professor Tiziana Rossetto).

References

[1] I. Ioannou, T. Rossetto, Empirical Fragility, in: M. Beer, I.A. Kougioumtzoglou, E. Patelli, S.-K. Au (Eds.) Encyclopedia of Earthquake Engineering, Springer Berlin Heidelberg, Berlin, Heidelberg, 2015, pp. 976-986.
[2] Rossetto T, Ioannou I, Grant DN, Existing empirical fragility and vulnerability relationships: Compendium and guide for selection., in, GEM Foundation, Pavia, Italy, 2013.
[3] R. Lee, A.S. Kiremidjian, Uncertainty and Correlation for Loss Assessment of Spatially Distributed Systems, Earthquake Spectra, 23 (2007) 753-770, 10.1193/1.2791001.

[4] S.R. Kotha, P. Bazzurro, M. Pagani, Effects of Epistemic Uncertainty in Seismic Hazard Estimates on Building Portfolio Losses, Earthquake Spectra, 0 null, 10.1193/020515eqs020m.

[5] K. Goda, H.P. Hong, Estimation of Seismic Loss for Spatially Distributed Buildings, Earthquake Spectra, 24 (2008) 889-910, 10.1193/1.2983654.

[6] T.J. Liu, H.P. Hong, Application of Spatially Correlated and Coherent Records of Scenario Event to Estimate Seismic Loss of a Portfolio of Buildings, Earthquake Spectra, 31 (2015) 2047-2068, 10.1193/060513eqs143m.

[7] Crowley H, Stafford PJ, Bommer J, Can earthquake loss models be validated using field observations?, Journal of Earthquake Engineering, 12 (2008) 1078-1104, https://doi.org/10.1080/13632460802212923.

[8] V. Sokolov, F. Wenzel, Influence of spatial correlation of strong ground motion on uncertainty in earthquake loss estimation, Earthquake Engineering & Structural Dynamics, 40 (2011) 993-1009, 10.1002/eqe.1074.

[9] Rota M, Penna A, Strobbia CL, Processing Italian damage data to derive typological fragility curves Soil Dynamics and Earthquake Engineering, 28 (2008) 933-947, 10.1016/j.soildyn.2007.10.010.

[10] Ioannou I, Rossetto T, Grant DN, Use of regression analysis for the construction of empirical fragility curves, in: 15th World Conference on Earthquake Engineering, Lisbon, Portugal, 2012.

[11] Ioannou I, Douglas J, Rossetto T, Assessing the impact of ground-motion variability and uncertainty on empirical fragility curves, Soil Dynamics and Earthquake Engineering, 69 (2015) 83-92, http://dx.doi.org/10.1016/j.soildyn.2014.10.024.

[12] Straub D, Kiureghian Der A, Improved seismic fragility modeling from empirical data, Structural Safety, 30 (2008) 320-336, 10.1016/j.strusafe.2007.05.004.

[13] Yazgan U, A New Approach to Building Stock Vulnerability Modeling based on Bayesian Analysi, in: 15th World Conference on Earthquake Engineering, Lisbon, Portugal, 2012.

[14] Yazgan U, Empirical vulnerability modeling considering geospatial ground motion variability, in: ICOSSAR, New York, USA, 2013.

[15] Yazgan U, Empirical seismic fragility assessment with explicit modeling of spatial ground motion variability, Engineering Structures, 100 (2015) 479-489, http://dx.doi.org/10.1016/j.engstruct.2015.06.027.

[16] Braga F, Dolce M, Liberatore D, Southern Italy November 23, 1980 earthquake: a statistical study of damaged buildings and an ensuing review of the M.S.K.-76 scale, in: Geodynamics Project, CNR Publication 503, Rome, Italy, 1982.

[17] M. Lancieri, A. Zollo, Simulated shaking maps for the 1980 Irpinia earthquake, Ms 6.9: Insights on the observed damage distribution, Soil Dynamics and Earthquake Engineering, 29 (2009) 1208-1219, https://doi.org/10.1016/j.soildyn.2009.01.007.

[18] P. Bernard, A. Zollo, The Irpinia (Italy) 1980 earthquake: Detailed analysis of a complex normal faulting, Journal of Geophysical Research: Solid Earth, 94 (1989) 1631-1647, 10.1029/JB094iB02p01631.

[19] ITACA, http://itaca.mi.ingv.it/ItacaNet/, in.

[20] D.M. Boore, W.B. Joyner, T.E. Fumal, Equations for Estimating Horizontal Response Spectra and Peak Acceleration from Western North American Earthquakes: A Summary of Recent Work, Seismological Research Letters, 68 (1997) 128-153, 10.1785/gssrl.68.1.128.

[21] J. Douglas, Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates, Earth-Science Reviews, 61 (2003) 43-104, https://doi.org/10.1016/S0012-8252(02)00112-5.

[22] N. Abrahamson, G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I.M. Idriss, W. Silva, R. Youngs, Comparisons of the NGA Ground-Motion Relations, Earthquake Spectra, 24 (2008) 45-66, 10.1193/1.2924363.

[23] Bindi D, Pacor F, Luzi L, Puglia R, Massa M, Ameri G, Paolucci R, Ground motion prediction equations derived from the Italian strong motion database Bulletin of Earthquake Engineering, 8 (2011) 1209-1230, https://doi.org/10.1007/s10518-011-9313-z.
[24] Kotha SR, Bindi D, Cotton F, Partially non-ergodic region specific GMPE for Europe and Middle-East, Bulletin of Earthquake Engineering, 14 (2016) 1245-1263, 10.1007/s10518-016-9875-x.
[25] Akkar S, Sandıkkaya MA, Bommer J J, Empirical ground-motion models for point- and extended-source crustal earthquake scenarios in Europe and the Middle East, Bulletin of Earthquake Engineering, 12 (2014) 359-387, 10.1007/s10518-013-9461-4.
[26] Ioannou I, Bessason B, Kosmidis I, Bjarnason JÖ, Rossetto T, Empirical seismic vulnerability assessment of Icelandic buildings affected by the 2000 sequence of earthquakes, Bulletin of Earthquake Engineering, 16 (2018) 5875-5903, 10.1007/s10518-018-0413-x.
[27] Medvedev SV, Seismic intensity scale MSK-76, Publ Inst Geophys Pol Acad Sc, 117 (1977) 95-102.
[28] Maqsood T, Edwards M, Ioannou I, Kosmidis I, Rossetto T, Corby N, Seismic vulnerability functions for Australian buildings by using GEM empirical vulnerability assessment guidelines, Natural Hazards, 80 (2016) 1625-1650, 10.1007/s11069-015-2042-x.
[29] Rossetto T, Elnashai A, Derivation of vulnerbaility functions for European-type RC structures based on observational data, Engineering Structures, 25 (2003) 1241-1263, https://doi.org/10.1016/j.engstruct.2010.01.009.
[30] Rossetto T, Ioannou I, Grant DN, Maqsood T, Guidelines for empirical vulnerability assessment. , in, GEM Foundation, Pavia, Italy, 2014.
[31] Rossetto T, Ioannou I, Empirical Fragility and Vulnerability Assessment: Not Just a Regression, in: Risk Modeling for Hazards and Disasters, Elsevier, 2018, pp. 79-103.
[32] J.A. Rice, Mathematical statistics and data analysis, Cengage Learning, 2006.
[33] McCullagh P., Nelder J. A., Generalized linear models, Chapman and Hall, London, UK, 1989.
[34] Faraway JJ, Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models, CRC press, 2016.
[35] Lunn D.J., Thomas A., Best N., Spiegelhalter D., WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility, Statistics and Computing, 10 (2000) 325--337.
[36] A. Gelman, D.B. Rubin, Inference from iterative simulation using multiple sequences, Statistical science, 7 (1992) 457-472.
[37] Gelman A, Hill J, Data analysis using regression and multilevel/hierarchical models Cambridge University Press, Cambridge, UK, 2007.
[38] N. Jayaram, J.W. Baker, Statistical Tests of the Joint Distribution of Spectral Acceleration Values, Bulletin of the Seismological Society of America, 98 (2008) 2231-2243, 10.1785/0120070208.
[39] D. Higdon, A primer on space-time modeling from a Bayesian perspective, Monographs on Statistics and Applied Probability, 107 (2006) 217.
[40] D. Lee, A comparison of conditional autoregressive models used in Bayesian disease mapping, Spatial and spatio-temporal epidemiology, 2 (2011) 79-89, 10.1016/j.sste.2011.03.001.
[41] J. Besag, J. York, A. Mollié, Bayesian image restoration, with two applications in spatial statistics, Annals of the Institute of Statistical Mathematics, 43 (1991) 1-20, 10.1007/bf00116466.
[42] L. Bernardinelli, D. Clayton, C. Pascutto, C. Montomoli, M. Ghislandi, M. Songini, Bayesian analysis of space—time variation in disease risk, Statistics in Medicine, 14 (1995) 2433-2443, doi:10.1002/sim.4780142112.

[43] C. Yepes-Estrada, V. Silva, T. Rossetto, D. D'Ayala, I. Ioannou, A. Meslem, H. Crowley, The Global Earthquake Model Physical Vulnerability Database, Earthquake Spectra, 32 (2016) 2567-2585, 10.1193/011816eqs015dp.
[44] Coburn AW, Spence RJS, Earthquake protection, John Wiley and Sons, Chichester, UK, 2002.
[45] Syner-G, Fragility functions for common masonry building types in Europe, in, 2011.
[46] M. Rota, A. Penna, G. Magenes, A methodology for deriving analytical fragility curves for masonry buildings based on stochastic nonlinear analyses, Engineering Structures, 32 (2010) 1312-1323, https://doi.org/10.1016/j.engstruct.2010.01.009.
[47] SYNER-G, Fragility functions for common RC building types in Europe, in, 2011.
[48] B. Borzi, H. Crowley, R. Pinho, Simplified Pushover-Based Earthquake Loss Assessment (SP-BELA) Method for Masonry Buildings, International Journal of Architectural Heritage, 2 (2008) 353-376, 10.1080/15583050701828178.
[49] Minas S, Advancements in the Seismic Risk Assessment of Mid-rise Reinforced Concrete Buildings. , in: Civil, Engineering and Geomatic Engineering, UCL, London, UK, 2015.
[50] S. Lagomarsino, S. Giovinazzi, Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings, Bulletin of Earthquake Engineering, 4 (2006) 415-443, 10.1007/s10518-006-9024-z.
[51] Rossetto T, D' Ayala D, Ioannou I, Meslem A, Evaluation of Existing Fragility Curves, in: C.H. Pitilakis K, Kaynia AM, (Ed.) SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk, Springer, 2014, pp. 47-93.
[52] J.A. Hoeting, D. Madigan, A.E. Raftery, C.T. Volinsky, Bayesian model averaging: a tutorial, Statistical science, (1999) 382-401.
[53] R.E. Chandler, Exploiting strength, discounting weakness: combining information from multiple climate simulators, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371 (2013) 20120388.

Appendix – Ground motion intensity based on Multiple GMPEs

In the main article, the results were based on selecting a single GMPE (i.e. Bindi et al [23]) to estimate the ground motion intensities of interest which was found to fit the ground motion records for the 1980 Irpinai earthquake better and was found to fit the damage data substantially better than the two alternatives. A sensitivity analysis is performed here in which all three GMPEs are used to produce combined assessments of interest, under the simplifying assumption that each of them is equally credible a priori. Given the overwhelming differences between the fit of the three GMPEs, as measured by their log-likelihoods in Table 4, this sensitivity analysis can be regarded as extremely conservative in the sense that it encompasses a much broader range of scenarios than is supported by the available evidence.

The aim of this note is to determine how the combined information from the three GMPEs can be used within our analysis framework. Some notational changes are deemed necessary in Eq.(1) and Eq.(2) in the article because each GMPE has its own function $f(.)$, its own event-specific error φ and its own set of residual errors $\{\varepsilon_j\}$. Therefore, a subscript g is used to denote the values of these quantities from the gth GMPE. Eq.(1) thus becomes, for g=1,2,3:

$$\ln\left(x_j\right) = f_g\left(M, R_j, S_j, F\right) + \phi_g + \varepsilon_{jg} \quad \text{(A.1)}$$

and Eq. (2) becomes

$$\ln\left(\tilde{x}_{jg}\right) = f_g\left(M, R_j, S_j, F\right) \quad . \quad \text{(A.2)}$$

The absence of a g subscript on the left-hand side of Eq, (A.1) above is correct: x_j is the actual ground motion, which is the same regardless of which GMPE is used. Also in this equation, the quantities φ_g and ε_{jg} have their own GMPE-specific variances, $\sigma^2_{\text{inter}_g}$ and $\sigma^2_{\text{intra}_g}$ say.

According to equations (A.1) and (A.2), if the g^{th} GMPE is correct then $\ln x_j$ has a normal distribution with expected value $\ln \tilde{x}_{jg}$ and variance $\sigma^2_{\text{inter}_g} + \sigma^2_{\text{intra}_g}$. Moreover, under this GMPE, and under the assumption used in models M0 to M2 that the errors $\{\varepsilon_{\cdot\cdot}\}$ are uncorrelated between locations, the covariance between $\ln x_j$ and $\ln x_k$ is $\sigma^2_{\text{inter}_g}$. Formally, we can write $E(\ln x_j \mid g) = \ln \tilde{x}_{jg}$; $\text{Var}(\ln x_j \mid g) = \sigma^2_{\text{inter}_g} + \sigma^2_{\text{intra}_g}$; and $\text{Cov}(\ln x_j, \ln x_k \mid g) = \sigma^2_{\text{inter}_g}$, where the notation $\cdot\mid g$ denotes a property of a probability distribution conditional on the g^{th} GMPE being correct.

Now: assuming that all three GMPEs are considered equally credible a priori, then a probability of 1/3 is assigned to each of them. In this case, the laws of iterated expectation and variance are applied to combine the information from all three GMPEs. In particular, the expected actual ground motion intensity is:

$$E\left(\ln\left(x_j\right)\right) = E_g\left\{E\left(\ln\left(x_j\right) \mid g\right)\right\} = \sum_{g=1}^{3} E\left(\ln\left(x_j\right) \mid g\right) P(g) = \frac{1}{3}\sum_{g=1}^{3} \ln\left(\tilde{x}_{jg}\right) \qquad \text{(A.3)}$$

The variance and covariance are, respectively:

$$\begin{aligned} \text{Var}\left(\ln\left(x_j\right)\right) &= E_g\left\{\text{Var}\left(\ln\left(x_j\right) \mid g\right)\right\} + \text{Var}_g\left\{E\left(\ln\left(x_j\right) \mid g\right)\right\} = \\ &= \frac{1}{3}\sum_{g=1}^{3}\left(\sigma^2_{\text{inter}_g} + \sigma^2_{\text{intra}_g}\right) + \frac{1}{3}\sum_{g=1}^{3}\left(\ln\left(\tilde{x}_{jg}\right) - E\left(\ln\left(x_j\right)\right)\right)^2 \end{aligned} \qquad \text{(A.4)}$$

and

$$\begin{aligned} \text{Cov}\left(\ln\left(x_j\right), \ln\left(x_k\right)\right) &= E_g\left\{\text{Cov}\left(\ln\left(x_j\right), \ln\left(x_k\right) \mid g\right)\right\} + \text{Cov}_g\left\{E\left(\ln\left(x_j\right) \mid g\right), E\left(\ln\left(x_k\right) \mid g\right)\right\} = \\ &= \frac{1}{3}\sum_{g=1}^{3}\sigma^2_{\text{inter}_g} + \frac{1}{3}\sum_{g=1}^{3}\left[\ln\left(\tilde{x}_{jg}\right) - E\left(\ln\left(x_j\right)\right)\right]\left[\ln\left(\tilde{x}_{kg}\right) - E\left(\ln\left(x_k\right)\right)\right] \end{aligned}$$

(A.5)

To apply the framework of the main paper to this 'multi-GMPE' setting, it is necessary to identify 'aggregate' quantities $\sigma^2_{\text{inter.}}$ and $\sigma^2_{\text{intra.}}$ that can be interpreted as combined estimated of inter- and intra-event variation corresponding to quantities σ^2_{ϕ} and $\sigma^2_{\varepsilon_j}$ respectively in Eq. (12). In this case, the value of Eq.(A.4) must correspond to $\sigma^2_{\text{inter.}} + \sigma^2_{\text{intra.}}$ and Eq.(A.5) to $\sigma^2_{\text{inter.}}$. Unfortunately, the desired quantities do not exist in general because the values of (A.4) and (A.5) vary between sites (they depend on the values of j and k). Nonetheless, it is not necessary to be too precise about the values of $\sigma^2_{\text{inter.}}$ and $\sigma^2_{\text{intra.}}$. There are two reasons for this. First, the focus of this sensitivity analysis is on 'very conservative' assumptions as noted above, a rough approximation will suffice. Secondly the estimates of the inter- and intra- variations are used to determine the parameters of the respective prior distributions (see Eq.12) so that indicative values will suffice. In view of this, $\sigma^2_{\text{inter.}}$ is determined as the average value of expression.(A.5) over all sites, and $\sigma^2_{\text{inter.}} + \sigma^2_{\text{intra.}}$ as the average value of expression.(A.4) over all sites.

Using the mean, variance and covariance of the intensity levels based on the combined three GMPEs, the reference models M0-M1 (as defined in section 3.1) are fitted to the damage data and fragility curves corresponding to ds_1, ds_2 and ds_5 for Building Class A and C are constructed. Fig.A.1 depicts the best estimate fragility curves and the 90% credible intervals obtained by fitting M1 to damage data and intensity levels estimated by a single GMPE [23] and the combined GMPEs [23-25]. Overall, the intensity level at the centroid of each municipality is smaller when estimated by the combined GMPEs. This leads to the best-estimate fragility curves being systematically higher than the ones based only on Bindi et al. [23]. This suggests that the combined GMPEs leads to more conservative results with both examined building classes being more vulnerable to earthquakes than the curves based on the best-fitted GMPE. It can also be noted that the differences in the width of the credible intervals for the two tested assumptions are not systematic. They appear to vary according to the damage state and building class.

Fig.A.1 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% credible intervals for M1 for intensity levels based on the best-fitted Bindi et al. [23] GMPE and on combining the three selected GMPEs [23-25].

Table A.1: Summary of the three models used for this sensitivity study.

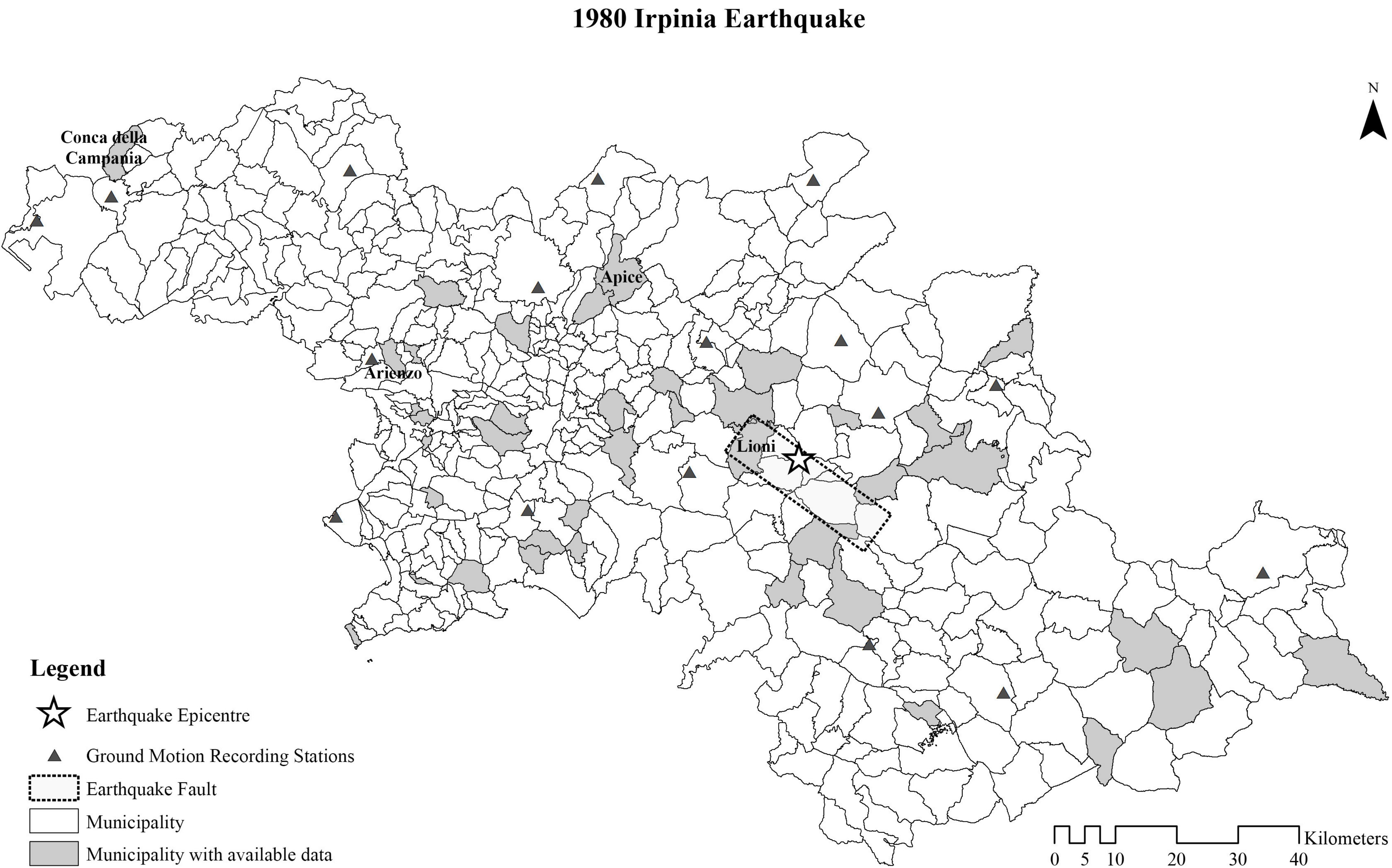
Model	Over-dispersion	Uncertainty in X	Uncertainty due to the spread of buildings in each municipality
M1	x	-	-
M2	x	x	-
M3	x	x	x

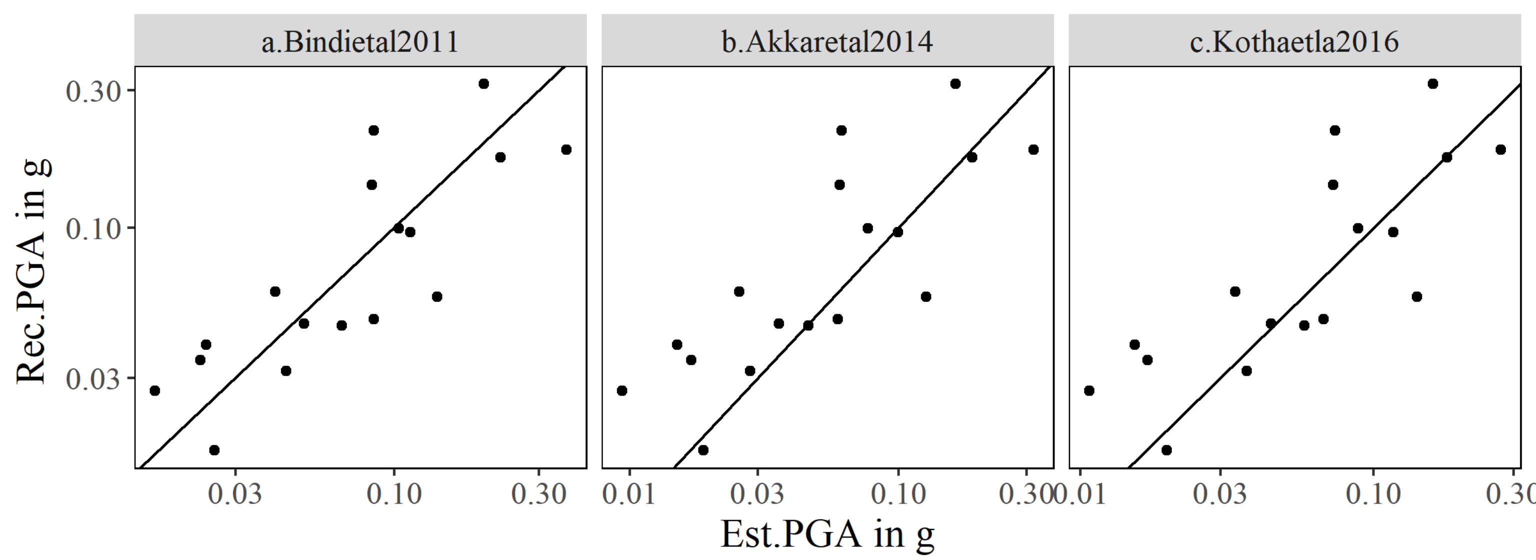
The inter-model variability on the shape of the fragility curves is examined next by explicitly accounting for the error component in the combined GMPEs as well as the uncertainty due to the spread of the buildings in each municipality. It should be noted that the most complex models used in this sensitivity analysis ignores the spatial correlation. In principle, it would be possible to relax the assumption that the intra-event errors are uncorrelated between locations for each GMPE and to account for the spatial correlation. Nonetheless, that would result in substantially more complicated expressions and, given the arguments in the preceding paragraph, it is not clear that the additional complexity is justified. A summary of the main characteristics of the three models used in this sensitivity analysis can be found in Tabel.A.1. A visual inspection of the fit on the reference model M1 with the more complex models M2 and M3* depicted in Fig.A.2 and Fig.A.3 show that neither the best-estimate fragility curves or their 90% credible intervals change significantly for M2 or M3 compared to M1.

The sensitivity analysis presented in this appendix highlighted that intensity levels based on the extremely conservative assumption which combined multiple GMPEs, leads to significant differences when compared to the curves based on the assumption of the best-fitted GMPE. It was also shown that with regard to the inter-model differences, in neither case do the best-estimate fragility curves or 90% credible intervals change significantly for M2 or M3 compared to M1 if the multiple GMPEs were used to estimate the intensity levels.

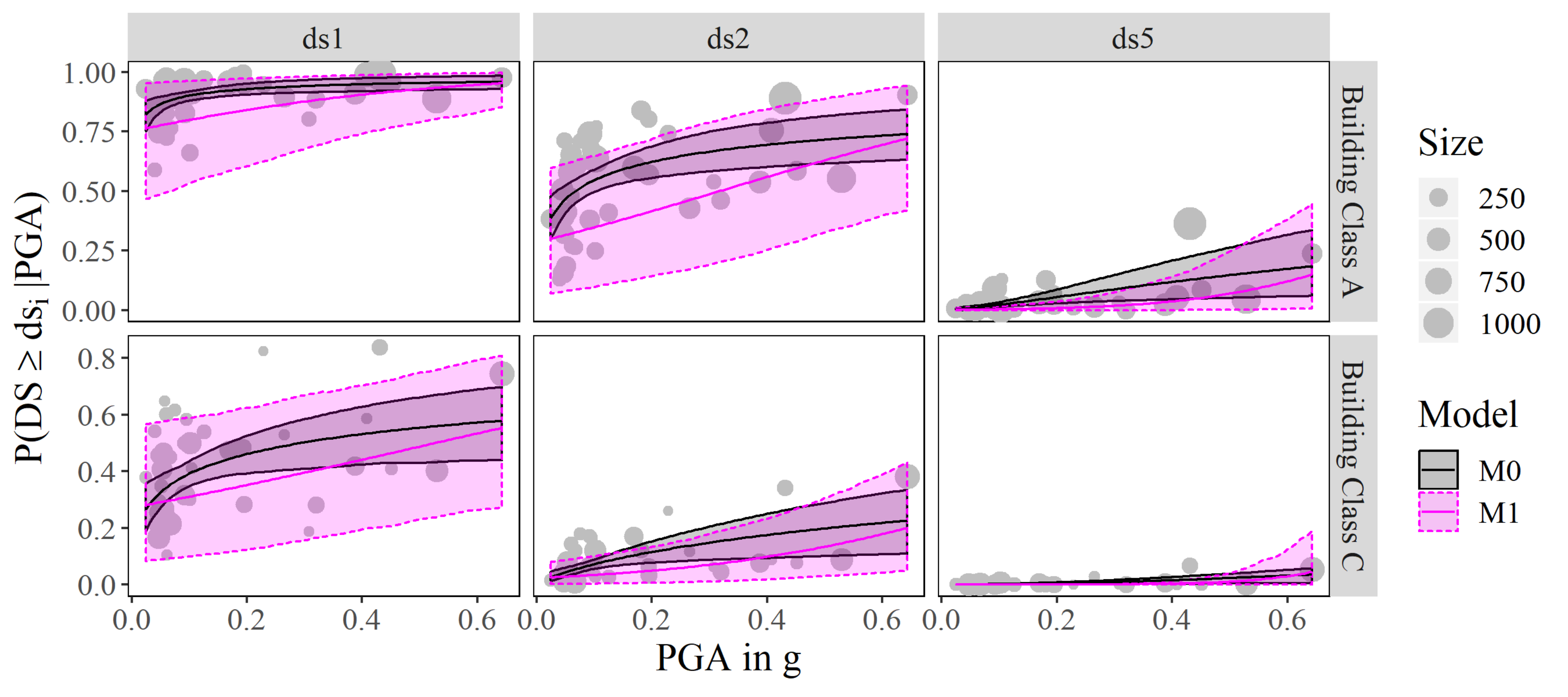
Fig.A.2 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The intensity levels are estimated by combining the three selected GMPEs. The size of each data point varies according to the total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% credible intervals comparing M1 vs M2 are also presented.

Fig.A.3 The data points represent the proportion of buildings of a given class (A or C) in each municipality, which sustained damage greater or equal to a given damage state ds_i. The intensity levels are estimated by combining the three selected GMPEs. The size of each data point varies according to the *total number of buildings of a given class located in a municipality. Best estimate fragility curves and their corresponding 90% credible intervals comparing M1 vs M3 are also presented.*

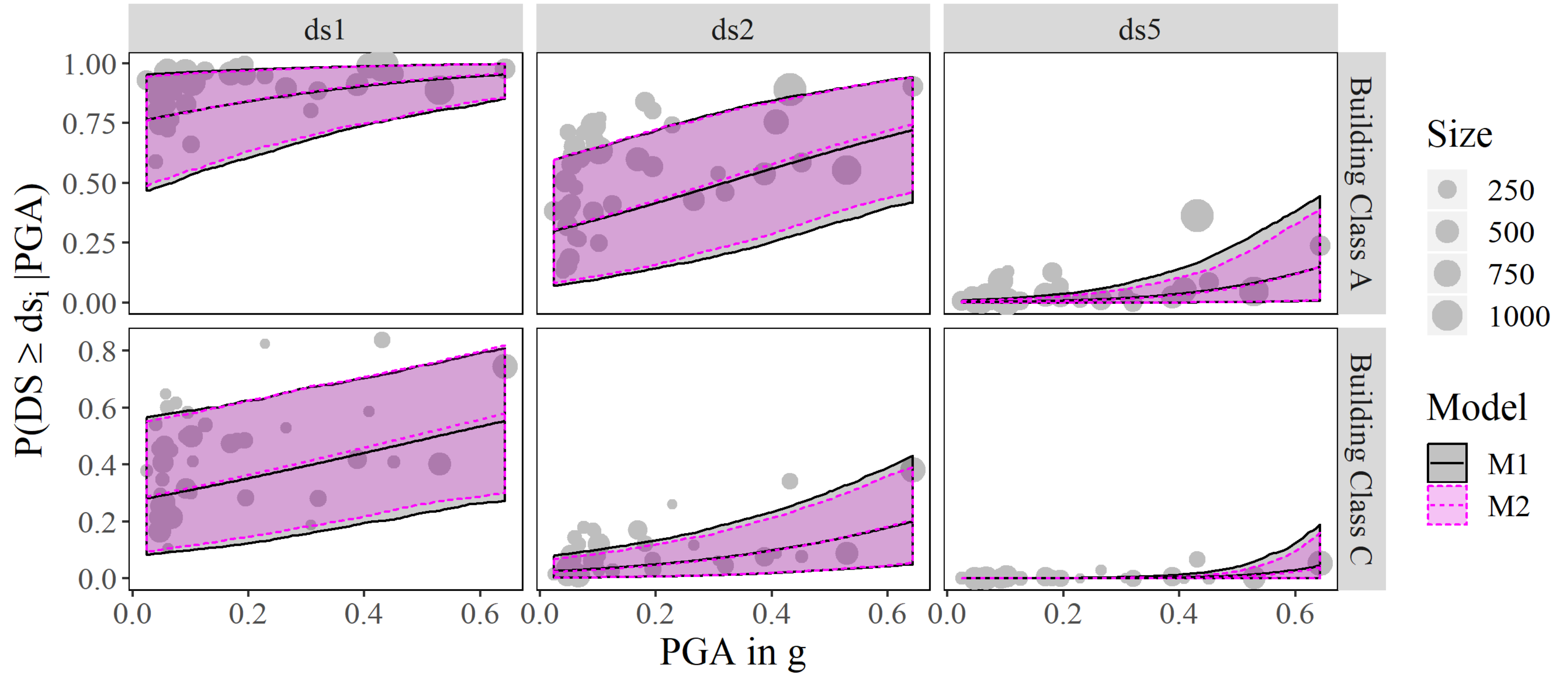

1980 Irpinia Earthquake
N
Conca della Campania
Apice
Arienzo
Lioni
Legend
Earthquake Epicentre
Ground Motion Recording Stations
Earthquake Fault
Municipality
Municipality with available data
0 5 10 20 30 40 Kilometers



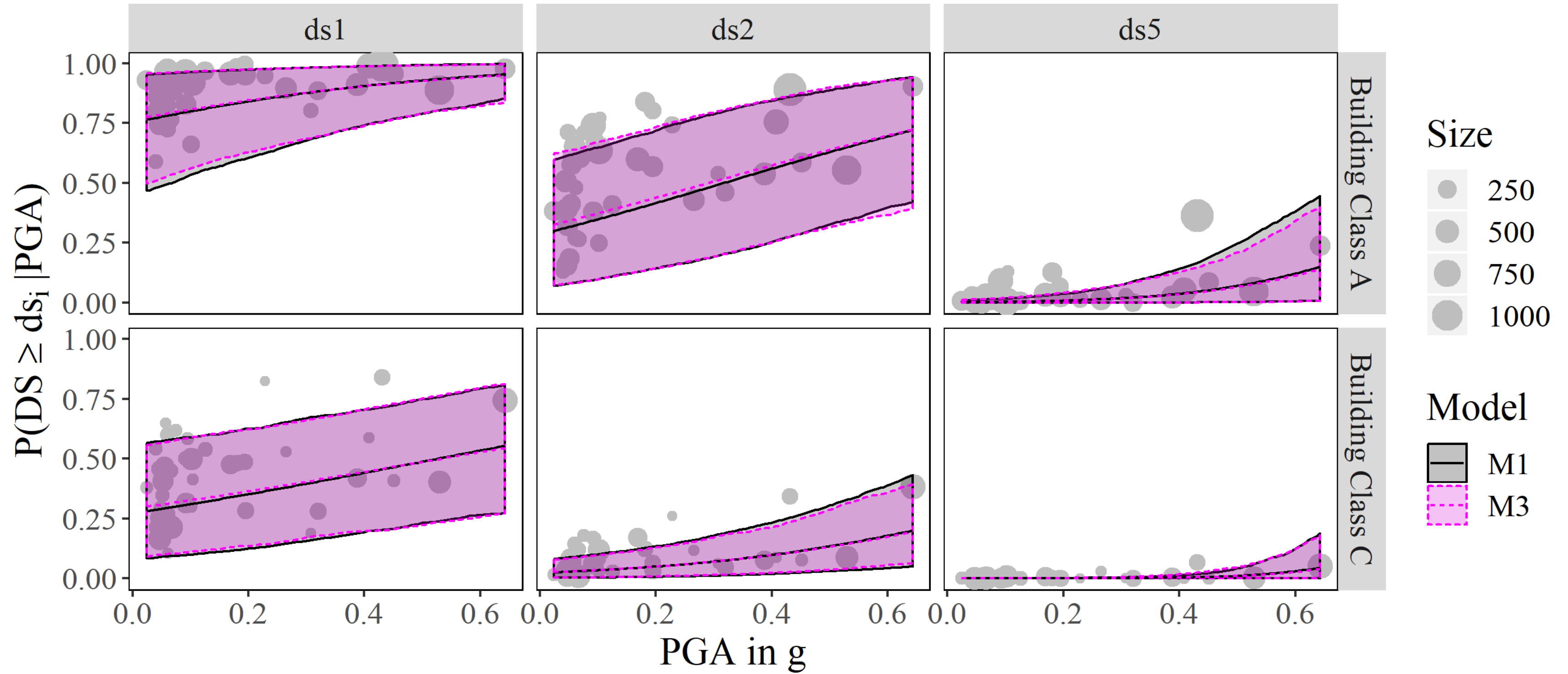
a.Bindietal2011
b.Akkaretal2014
c.Kothaetla2016
Rec.PGA in g
0.30
0.10
0.03
0.01
Est.PGA in g



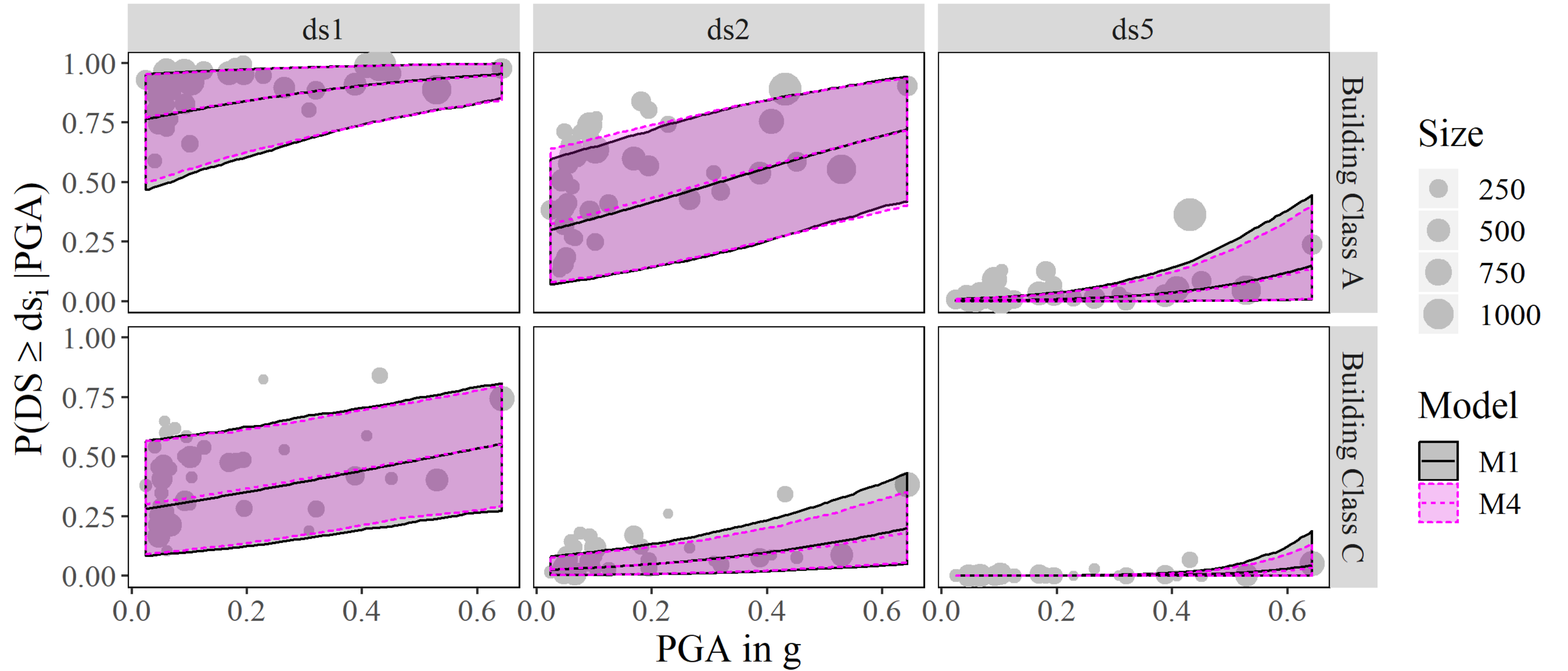
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i |PGA)
PGA in g
1.00
0.75
0.50
0.25
0.00
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
Size
250
500
750
1000
Model
M0
M1



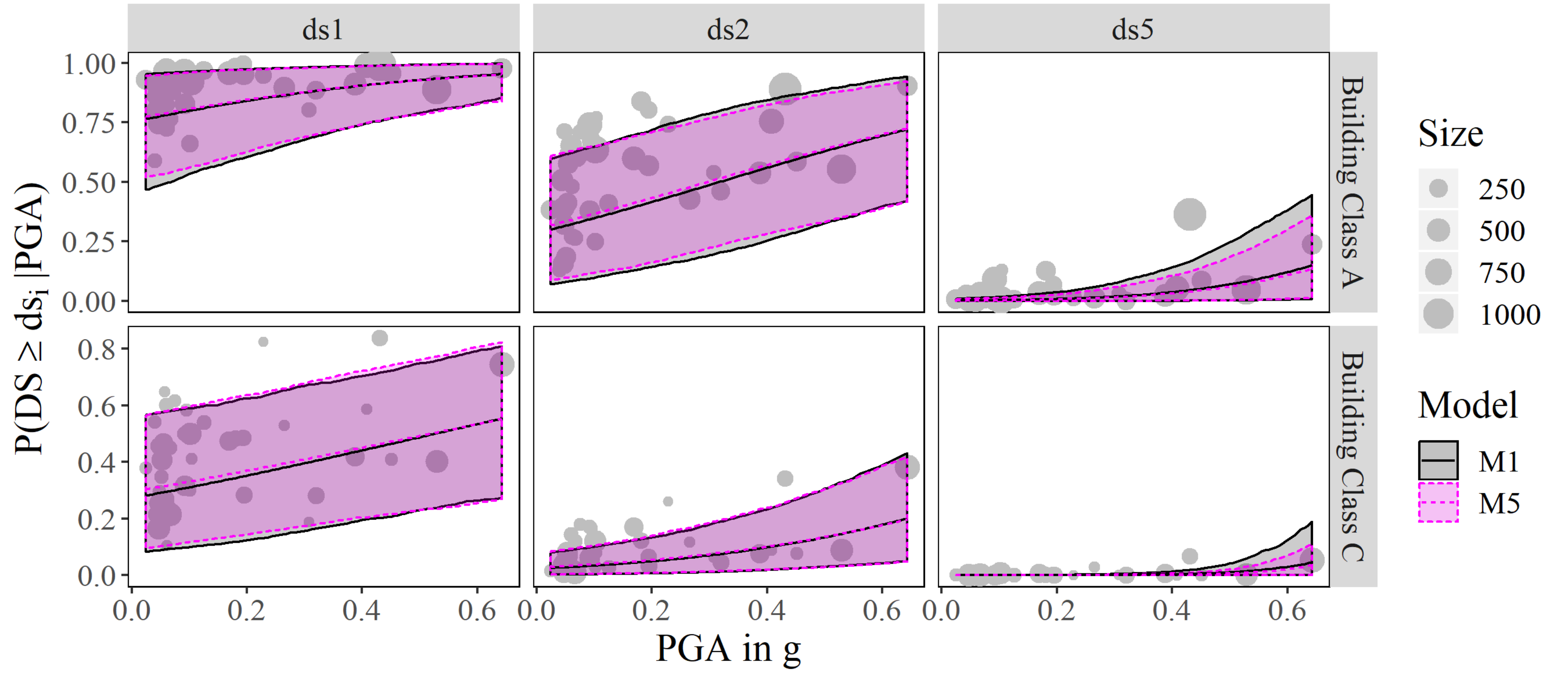
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i |PGA)
PGA in g
1.00
0.75
0.50
0.25
0.00
0.8
0.6
0.4
0.2
0.0
Size
250
500
750
1000
Model
M1
M2



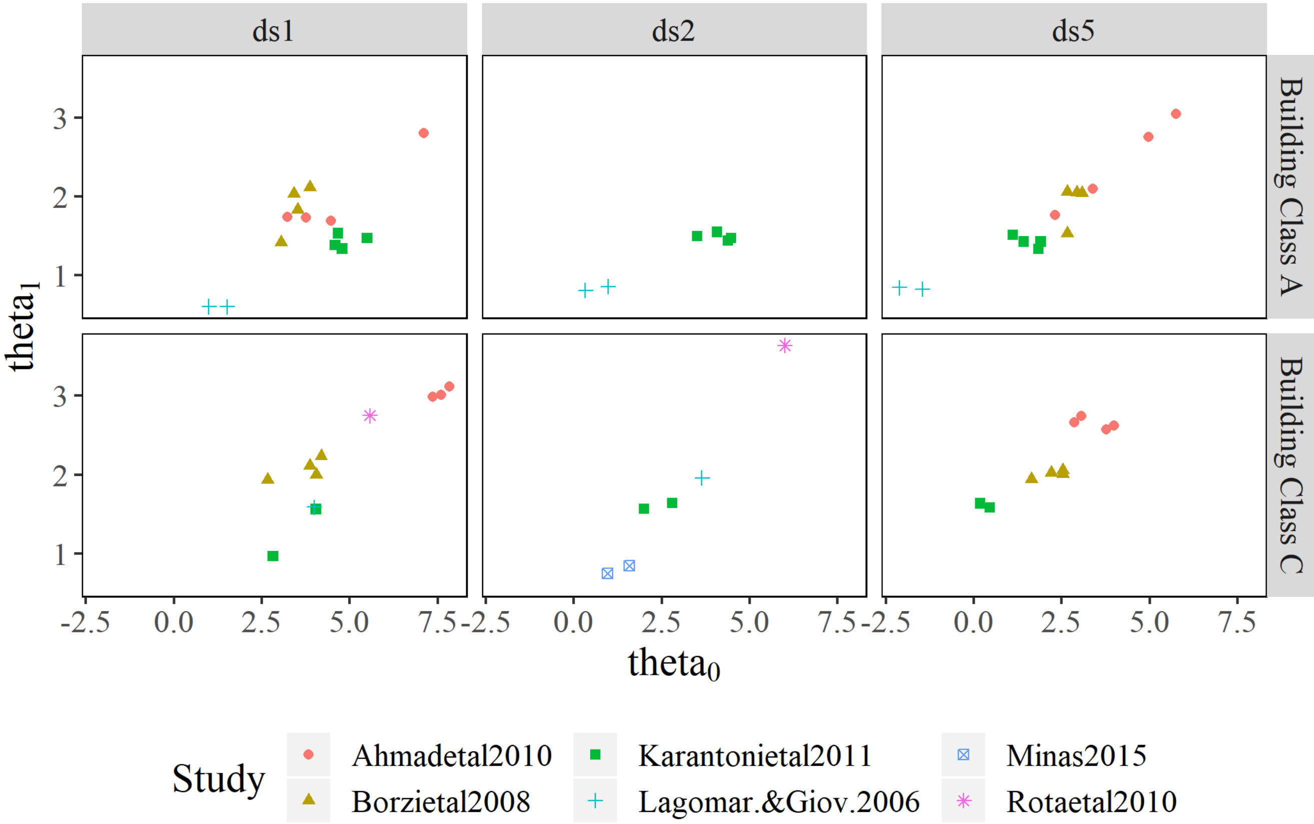
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i |PGA)
PGA in g
1.00
0.75
0.50
0.25
0.00
0.0
0.2
0.4
0.6
Size
250
500
750
1000
Model
M1
M3



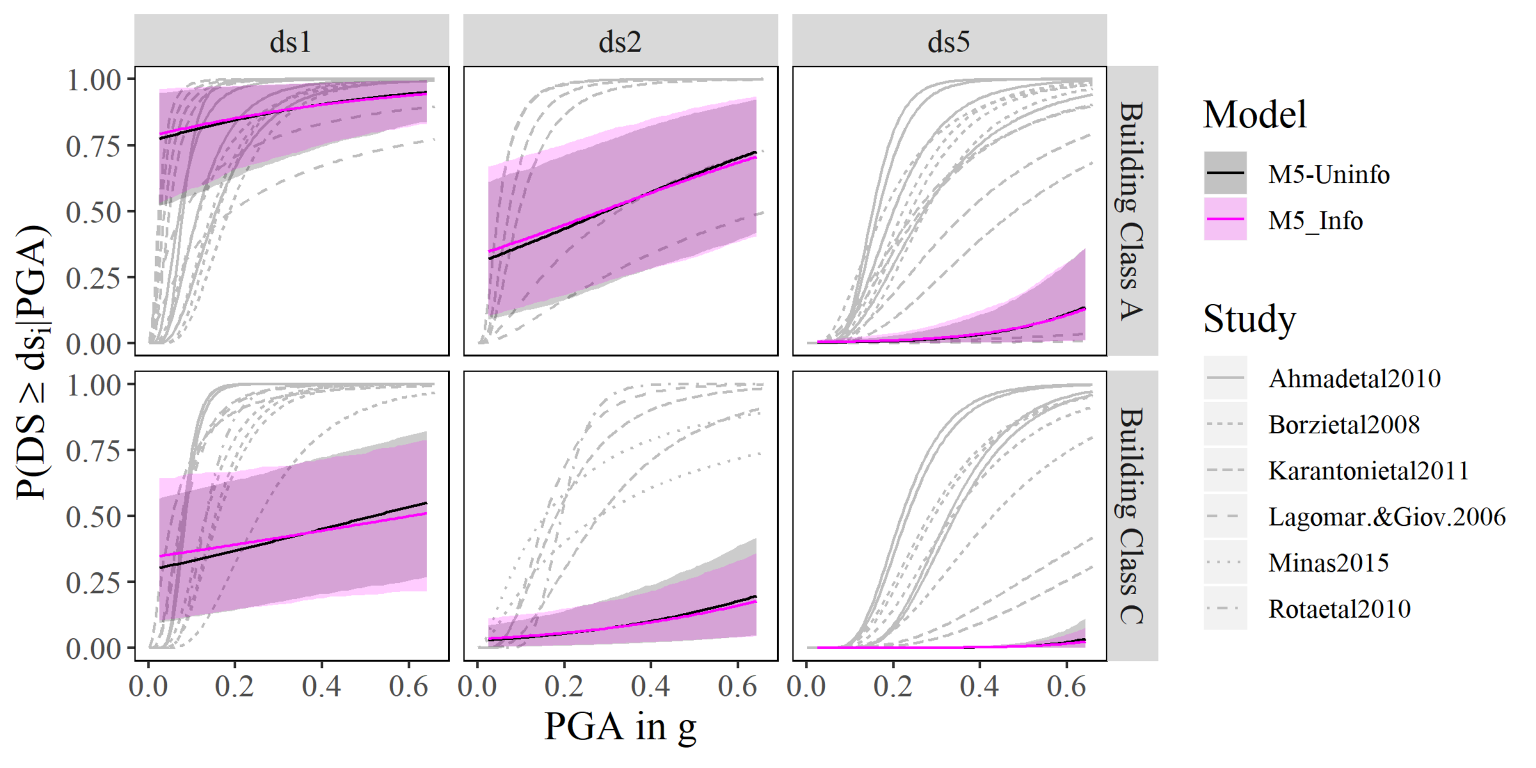
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i |PGA)
1.00
0.75
0.50
0.25
0.00
0.0
0.2
0.4
0.6
PGA in g
Size
250
500
750
1000
Model
M1
M4



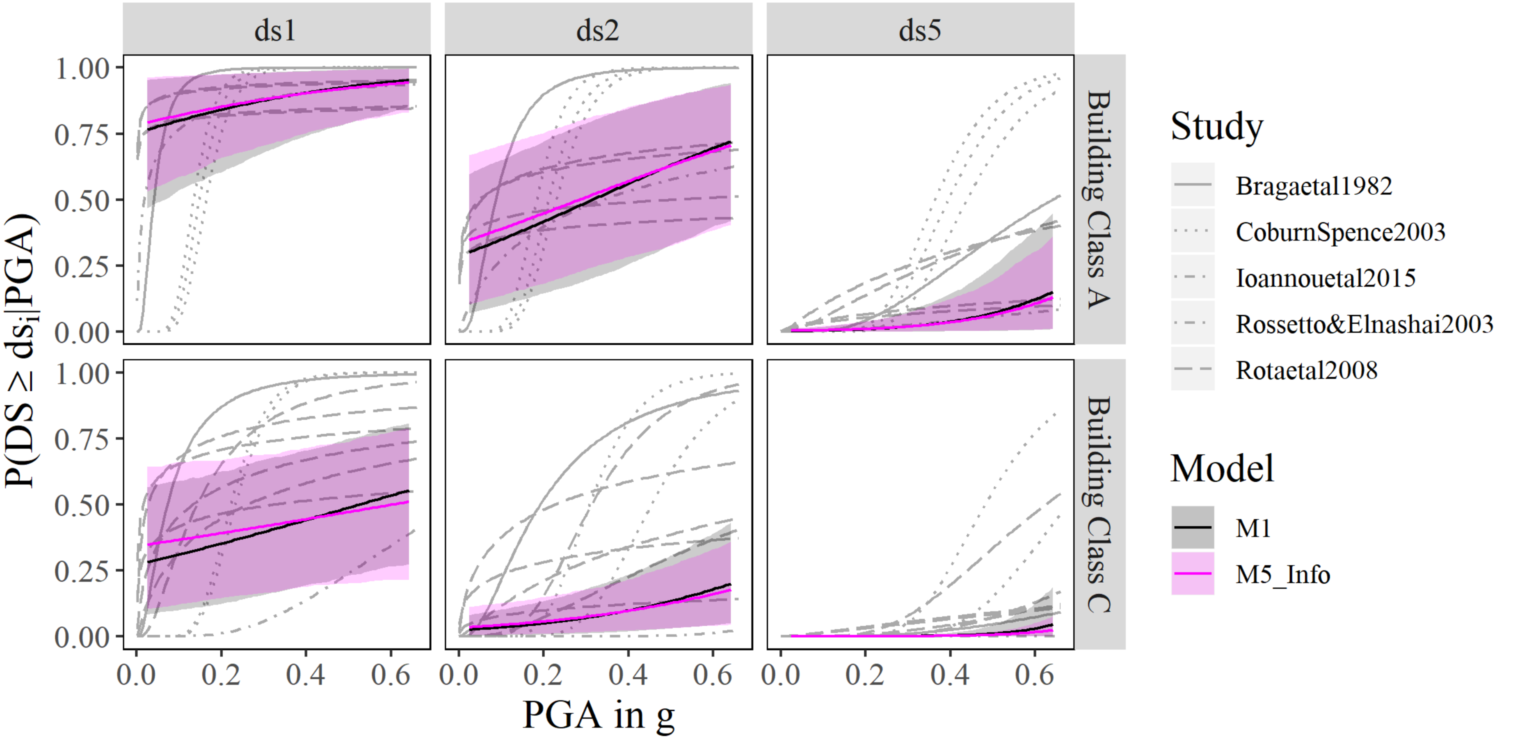
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i |PGA)
PGA in g
1.00
0.75
0.50
0.25
0.00
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
Size
250
500
750
1000
Model
M1
M5



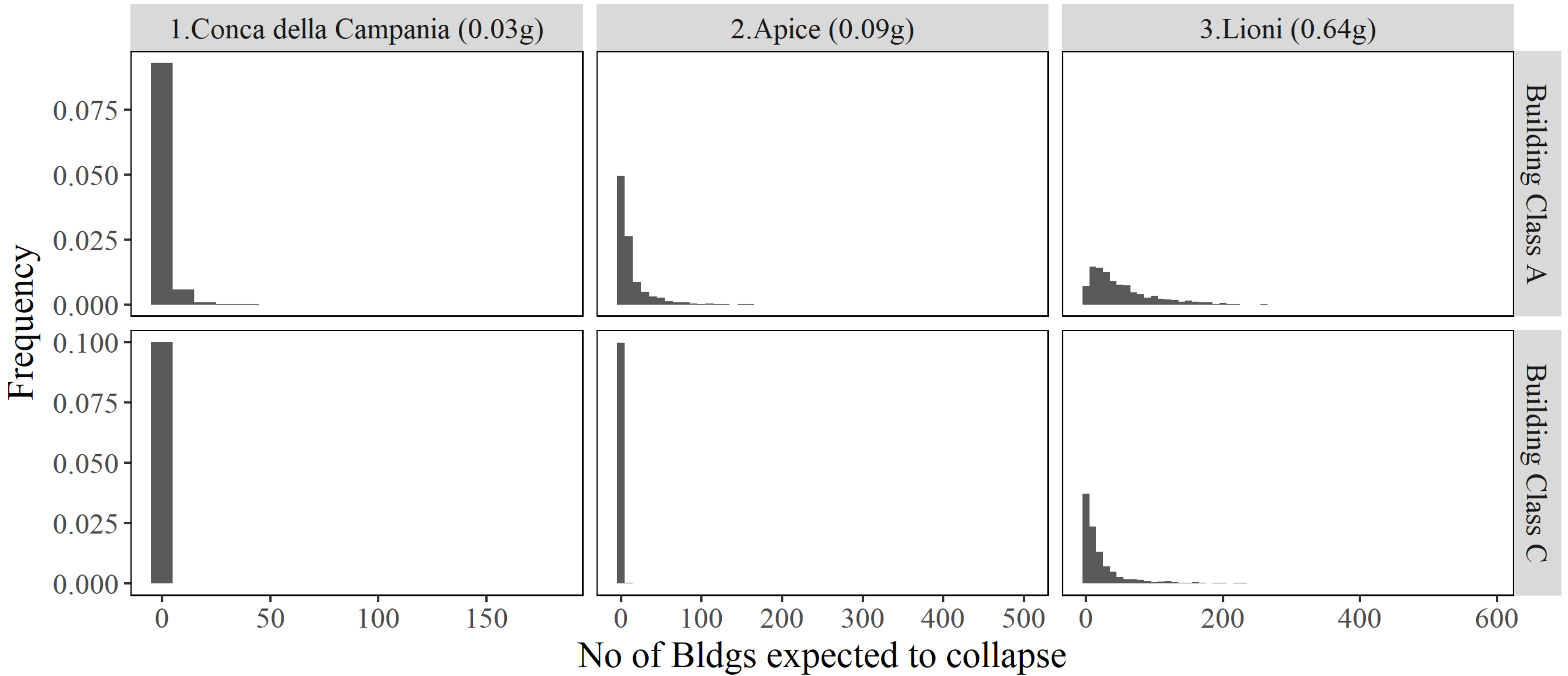
ds1
ds2
ds5
Building Class A
Building Class C
theta1
theta0
3
2
1
-2.5
0.0
2.5
5.0
7.5
Study
Ahmadetal2010
Borzietal2008
Karantonietal2011
Lagomar.&Giov.2006
Minas2015
Rotaetal2010



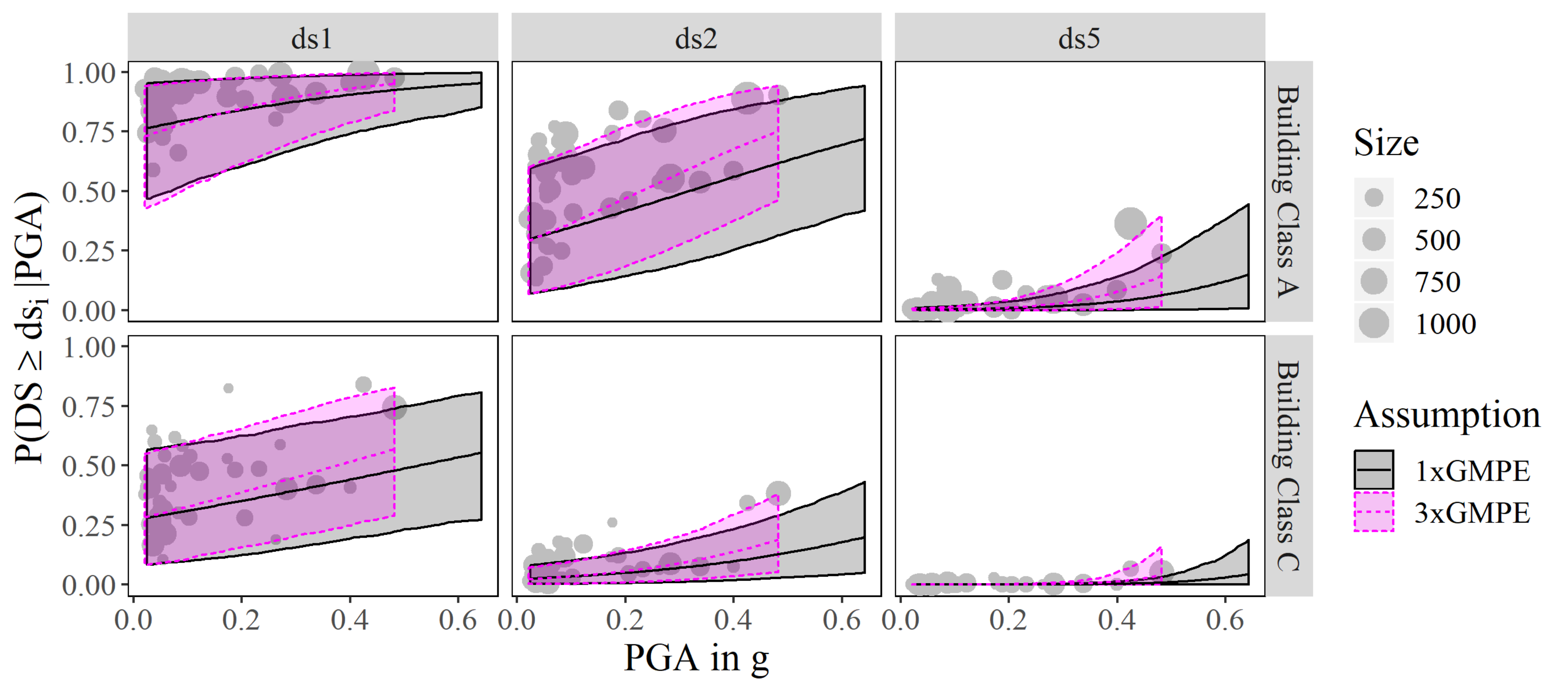
ds1
ds2
ds5
Building Class A
Building Class C
P(DS ≥ ds_i|PGA)
1.00
0.75
0.50
0.25
0.00
0.0
0.2
0.4
0.6
PGA in g
Model
M5-Uninfo
M5_Info
Study
Ahmadetal2010
Borzietal2008
Karantonietal2011
Lagomar.&Giov.2006
Minas2015
Rotaetal2010



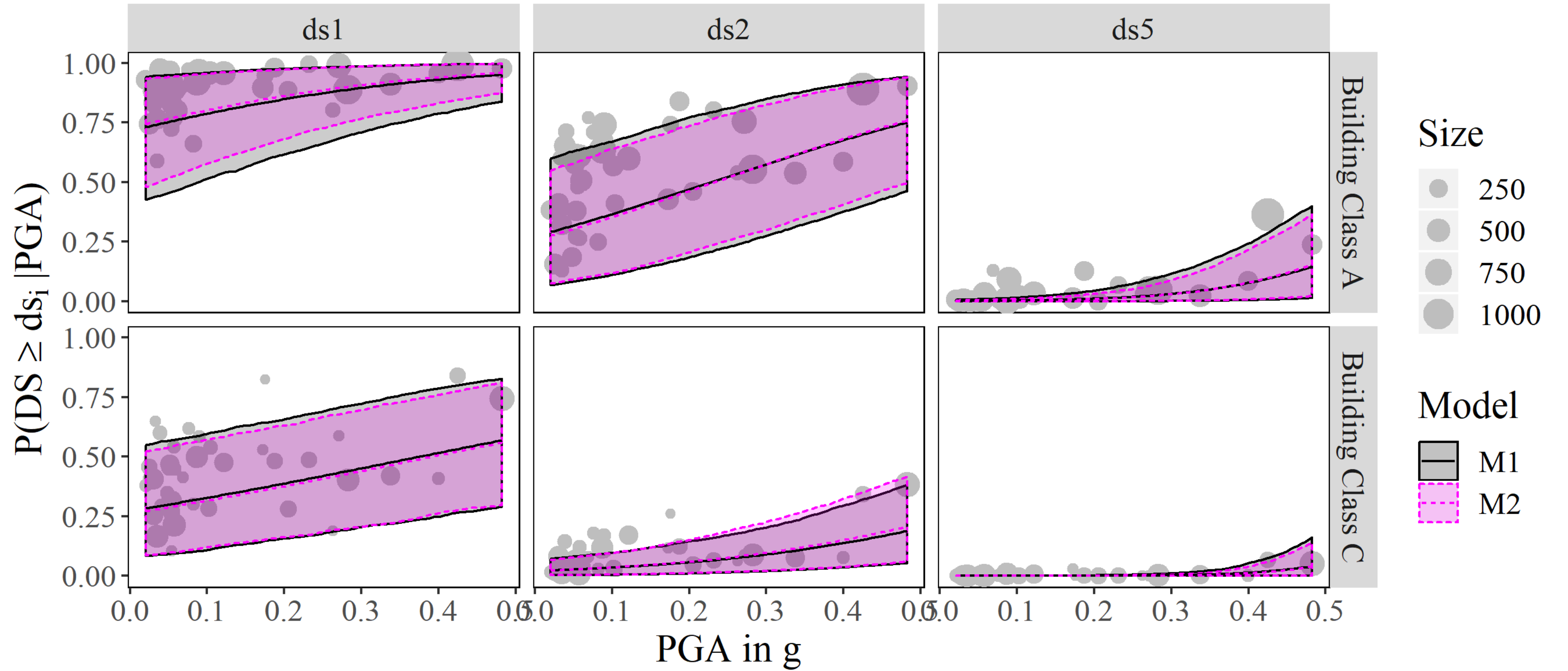
ds1
ds2
ds5
Building Class A
Building Class C
1.00
0.75
0.50
0.25
0.00
P(DS ≥ ds_i|PGA)
0.0
0.2
0.4
0.6
PGA in g
Study
Bragaetal1982
CoburnSpence2003
Ioannouetal2015
Rossetto&Elnashai2003
Rotaetal2008
Model
M1
M5_Info



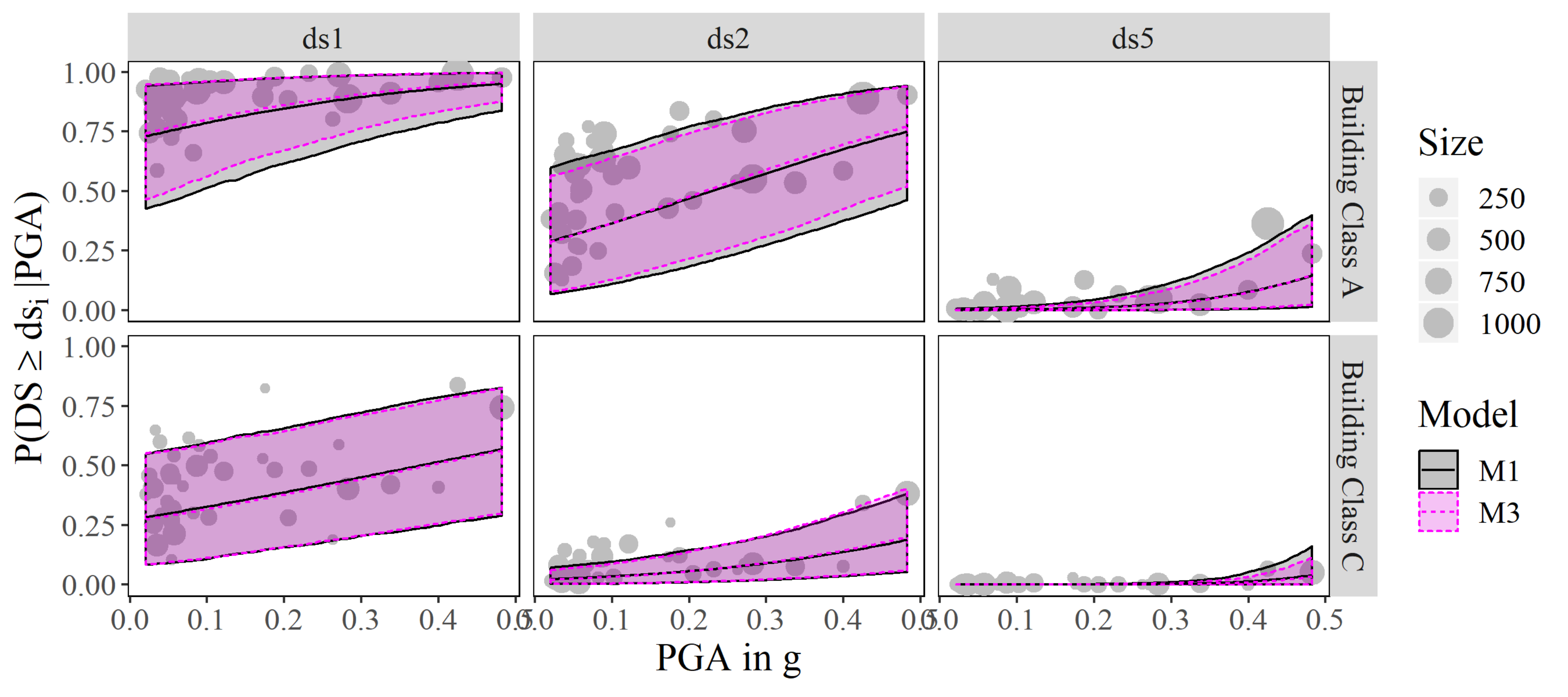
1.Conca della Campania (0.03g)
2.Apice (0.09g)
3.Lioni (0.64g)
Building Class A
Building Class C
Frequency
0.000
0.025
0.050
0.075
0.100
0
50
100
150
200
300
400
500
600
No of Bldgs expected to collapse



ds1
ds2
ds5
Building Class A
Building Class C
$P(DS \geq ds_i | PGA)$
PGA in g
1.00
0.75
0.50
0.25
0.00
0.0
0.2
0.4
0.6
Size
250
500
750
1000
Assumption
1xGMPE
3xGMPE



ds1
ds2
ds5
Building Class A
Building Class C
$P(DS \geq ds_i | PGA)$
PGA in g
1.00
0.75
0.50
0.25
0.00
0.0
0.1
0.2
0.3
0.4
0.5
Size
250
500
750
1000
Model
M1
M2



ds1
ds2
ds5
Building Class A
Building Class C
$P(DS \geq ds_i \,|PGA)$
PGA in g
1.00
0.75
0.50
0.25
0.00
0.0
0.1
0.2
0.3
0.4
0.5
Size
250
500
750
1000
Model
M1
M3