https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

Methods for evaluating the significance and importance of differences amongst probabilistic seismic hazard results for engineering and risk analyses: A review and insights

John Douglas[1], Helen Crowley[2], Vitor Silva[3], Warner Marzocchi[4], Laurentiu Danciu[5], Rui Pinho[6]

[1]Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, G1 1XJ, United Kingdom. https://orcid.org/0000-0003-3822-0060
[2]European Centre for Training and Research in Earthquake Engineering (EUCENTRE), Pavia, Italy. https://orcid.org/0000-0002-1216-1245
[3]Seismic Risk Team, Global Earthquake Model Foundation, Pavia, Italy. https://orcid.org/0000-0003-4408-3883
[4]Department of Earth, Environmental, and Resources Sciences, University of Naples, Federico II, 80126 Naples, Italy. https://orcid.org/0000-0002-9114-1516
[5]Swiss Seismological Service, ETH Zurich, Zurich, Switzerland. https://orcid.org/0000-0003-4086-8755
[6]Department of Civil Engineering and Architecture, University of Pavia, Via Adolfo Ferrata 5, 27100 Pavia, Italy. https://orcid.org/0000-0001-6767-9036

Correspondence to: John Douglas (john.douglas@strath.ac.uk)

Abstract. When new seismic hazard models are published it is natural to compare them to existing models for the same location. This type of comparison routinely indicates differences between the assessed hazards in the various models. The question that then arises is whether these differences are scientifically *significant*, given the large epistemic uncertainties inherent in all seismic hazard models, or practically *important*, given the use of hazard models as inputs to risk and engineering calculations. A difference that exceeds a given threshold could mean that building codes may need updating, risk models for insurance purposes may need to be revised, or emergency management procedures revisited. In the current literature there is little guidance on what constitutes a *significant* or *important* difference, which can lead to lengthy discussions amongst hazard analysts and end users. This study reviews proposals in the literature on this topic and examines how applicable these proposals are for several sites considering various seismic hazards models for each site, including the two European Seismic Hazard Models of 2013 and 2020. The implications of differences in hazard for risk and engineering purposes are also examined to understand how *important* such differences are for potential end users of seismic hazard models. Based on this, we discuss the relevance of such methods to determine the scientific *significance* and practical *importance* of differences between seismic hazard models and identify some open questions. Finally, we make some recommendations for the future.

1 Introduction

Probabilistic seismic hazard models provide assessments of the annual frequency of exceedance (AFE), or probabilities of exceedance over decades, of different levels of ground-motion intensity. Hundreds if not

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

thousands of such models have been derived since probabilistic seismic hazard assessment (PSHA) was pioneered in the late 1960s. These models are the basis of seismic design codes, risk evaluations for insurance and crisis management purposes, and, especially site-specific models, are used to assess the seismic safety of existing and planned critical infrastructure. Well-studied parts of the world (e.g., California, Italy, and New Zealand) have been the subject of dozens of PSHAs, all of which show differences in the assessed hazard for AFEs of engineering importance (generally 10^{-2} to 10^{-4}, but down to 10^{-8} for some highly sensitive infrastructure).

The differences in the assessed hazard between different generations of the models are often vigorously debated, particularly if the models are developed for national building codes. Given the importance of assessing whether differences between hazard results are of concern or expected given the large uncertainties inherent in PSHAs, it is surprising that there are only a handful of proposals in the literature discussing how these differences should be considered. The two aims of this article are the following: to review and test the proposed approaches about what is a *significant* or *important* difference between hazard results; and to discuss what could be considered when examining these differences for various uses of the hazard models.

Although the terms *significant* and *important* are often considered as synonyms, they have different meanings, at least within the context of this study. The term *significant* implies the application of statistical tests. Depending on the probabilistic framework adopted, the test seeks to detect if the differences are due to chance or if they describe changes that cannot be explained by epistemic uncertainties, or whether, given a set of data, one model performs remarkably better than the other. The term *important* has a more pragmatic meaning, suggesting that the differences (statistically *significant* or not) lead to consequences that may have a large economic or societal impact, or may even be unmanageable. We note that this terminology may create confusion to end users of hazard results. This is because a difference amongst results from different hazard models may not be *significant* from a statistical viewpoint, but the difference may be *important* for practical applications, such as for building codes where engineers may have to design or assess structures under different seismic loads depending on the model adopted.

It is noted that in the following calculations the exact values from the published hazard curves are used rather than contoured values. This is because contouring can mask or emphasise differences in computed hazard by either grouping a site in the same category in different models or separating them, solely depending on where the limits are drawn. For example, a site with a mean peak ground acceleration (PGA) at the 475-year return period in Hazard Model A of 0.19g, which subsequently changed to 0.21g in Hazard Model B, would either be in the same category if the limits were 0.15g, 0.25g and 0.35g, or different categories if the limits were 0.1g, 0.2g and 0.3g. Contouring can be useful in indicating that hazard cannot be determined precisely, but as this simple example shows, it can also unduly emphasise small differences. In addition, as shown by Table 3 of Douglas et al. (2013), contouring can have an impact on risk estimates (and other derived values) calculated using contoured results, for example within a seismic building code. As an aside, we encourage the use of consistent colours and contours through time and across organisations so that it is easier to visually compare hazard models for the same region without needing to redraw the maps from downloaded data. Within Europe, the European Facilities for Earthquake Hazard and Risk (EFEHR, http://www.efehr.org/) could be a vehicle for encouraging this harmonisation.

In this study we discuss differences in the hazard results as expressed by hazard curves for specific structural periods and ranges of AFE, and refrain from discussing differences in uniform hazard spectra or inputs to the hazard model. These types of differences are discussed in country-specific studies, e.g. Belvaux et al. (2014) for France, Kohrangi et al. (2018) for Iran, and Kalakonas et al. (2020) for Guatemala. Similarly, Beauval and Bard (2022) present informative plots (their Figures 4 and 5) showing how the assessed hazard for various locations in France varies depending on the study. Herein, methods proposed in the literature to evaluate whether a difference between hazard results is *significant* or *important* are summarised. In the following section, these methods are applied to some typical test cases and we highlight the different conclusions that would be reached by applying them. Next, the risk and engineering implications of differences between hazard results are discussed using various examples. The penultimate section discusses the results of these examples and describes areas for future research about how differences between hazard results could be managed. The article ends with some brief recommendations.

2 Methods proposed in the literature

It is relatively common to quantify the differences, often in percentage terms, in ground-motion levels for a given return period (or AFE) between hazard models (e.g., Belvaux et al., 2014; Tromans et al., 2019). These percentage differences, however, are often simply reported but their practical *importance* in terms of applications of the hazard model, or their statistical *significance* in terms of how accurately hazard can be assessed given the large epistemic uncertainties is rarely discussed. At least four methods to examine differences in pairs of seismic hazard models have been proposed, which are presented in this section. In the publicly available literature, there are few examples of these proposals being used beyond the initial application, however.

The first proposal is by McGuire (2012). He examines the differences between the hazard assessed by different experts or teams within site-specific seismic hazard assessments away from active plate boundaries (central and eastern North America and Switzerland). He uses this information to propose a lower bound for the precision to which seismic hazard can currently be assessed for well-studied sites within state-of-the-art studies. He concludes that if an alternative assessment changes the calculated mean AFE by less than 25% for ground motions corresponding to 10^{-4} annual frequency of exceedance and changes the mean AFE by less than 35% for ground motions corresponding to 10^{-6} annual frequency of exceedance, then, in the language used by the author, the change can be deemed not significant, even though the method is not based on a formal statistical test. This proposal could be considered as the simplest, as it uses the mean hazard curve directly and not any hazard fractiles or derived values. Therefore, it could be applied more easily and widely than other approaches. On the other hand, it is only applicable for low AFEs and site-specific studies and it is based on how precise hazard assessments can be from the point of view of available knowledge rather than the importance of differences from the point of view of end users of the hazard results.

The second proposal is by Malhotra (2014, 2015), which is based on Cohen (1977)'s effect size. This approach is based on the probability density functions (PDFs) derived from the mean hazard curves of the two studies being compared. The test statistic is the Cohen effect size, d, defined as:

$$d = \frac{\mu_1 - \mu_2}{\sqrt{0.5(\sigma_1^2 + \sigma_2^2)}} \tag{1}$$

where μ_1 and μ_2 are the means and σ_1 and σ_2 are the standard deviations of the two PDFs of the ground-motion levels. The criterion for a *large* difference, adopted by Malhotra (2014, 2015) from Cohen's original work, is d>0.8. Again, because this approach is based on mean hazard curves it can be applied when fractiles have not been computed or published. It does, however, need numerical values for full hazard curves, which are not always published (though digitised scans could be used). Because of the shape of the PDFs derived from hazard curves, the means characterise ground-motion levels of low engineering importance. Hence, a *large* effect size may not mean the hazard curves are greatly different at AFEs commonly used in applications.

The workshop presentation by Abrahamson (2017) proposes two independent approaches to assess the *robustness* (although the meaning of that term is not explained by the author) of a change in the hazard between the original and update hazard model. The change is considered *robust* by Abrahamson (2017) if the mean hazard for the target AFE is outside the 25th and 75th ground-motion fractiles of the updated hazard model. Although not being a formal statistical test (this point will be discussed later on), we interpret the term *robust* to refer to a difference worth being mentioned. This approach is simple to apply and does not require any calculations, although the 25th and 75th fractiles are not commonly computed (i.e., the 5th, 16th, 50th, 84th and 95th fractiles are more common). The second criterion can be described by the following inequality (rearranged from the original formulation):

$$\ln\left(\frac{\mathrm{IM}_{\mathrm{new}}}{\mathrm{IM}_{\mathrm{old}}}\right) - 0.5\sigma_{haz} > 0 \tag{2}$$

where $\mathrm{IM}_{\mathrm{new}}$ and $\mathrm{IM}_{\mathrm{old}}$ are the mean ground-motion levels at the target AFE and σ_{haz} is the standard deviation derived from the logarithms of the fractiles at the target AFE. Both these approaches have the advantage of explicitly considering the uncertainty in the assessed hazard through the fractiles. The second approach, however, requires the numerical values of the fractiles to have been published and further calculations. The benefit of this second approach over the method based solely on the relative locations of the original mean and updated 25th and 75th fractiles is not clear, especially since it may give misleading results for low hazard areas. Finally, it is not clear what conclusion should be drawn if the two approaches give contradictory results (i.e., one criterion says the change is *robust* whilst the other says it is not).

The methods of Malhotra (2014, 2015) and Abrahamson (2017) are based on the differences in ground-motion levels (e.g., PGAs) for a given return period (or AFE). This is how hazard results are often used in practice within building design codes or national hazard maps (e.g., what is the PGA for a return period of 475 years?). Hazard is assessed, however, the other way around, i.e., a hazard engine gives the AFE for a given ground-motion level. This is recognised by the proposal of McGuire (2012) who uses the AFEs for his criteria. This is also how hazard models could be used within risk evaluation through the convolution of the hazard and fragility/vulnerability curves. Therefore, it may be more appropriate among the four methods, although more difficult to visualise, to use the differences in AFE at a given ground-motion level (e.g., PGA of 0.1g), as has been done in Italy where the coefficient of variation of the hazard has been mapped (Meletti et al., 2021). Because of the steep slopes of hazard curves, differences in AFEs for a given ground-motion level can also be many times higher than differences in ground-motion level for a given AFE. Therefore, the same difference may appear greater when expressed in terms of AFEs than when expressed in terms of ground motion.

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

Although the term *significance* is used in some of the above-mentioned papers and some of the methods suggest a statistical test, we argue that a formal test on the statistical *significance* of the difference between two hazard outcomes requires the definition of the probabilistic framework adopted (Marzocchi and Jordan, 2017). For example, in the subjective interpretation of probability (Apostolakis, 1990; Vick, 2002), we consider only the mean hazard curves (here the fractiles do not have any meaning) and the statistical test to evaluate their differences requires a set of data to calculate the Bayes factor (Kass and Raftery, 1995); note that in this framework, the term *significance* does not refer to the *significance* level of a test, but to a different category of the Bayes factor [see section 3.2 in Kass and Raftery (1995)]. Conversely, fractiles have a formal meaning in the unified framework proposed by Marzocchi and Jordan (2014), where a meaningful statistical comparison between two hazard outcomes requires the inclusion of the distribution of fractiles from both models. All above-mentioned methods do not comply with these basic requirements to assess a general statistical *significance*. Hence, the value and interpretation of these techniques is primarily heuristic, and the best method to use depends on the context.

3 Results of applying the methods to some test cases

In this section the methods described in the previous section are applied to some European test cases: one site in Switzerland with results from five recent PSHAs; some example locations from two recent national hazard models for Italy; and a comparison of two European hazard models.

3.1 Beznau (Switzerland)

The first site used to test the methods described in the previous section is Beznau in northern Switzerland. This site is chosen as it was considered within two state-of-the-art site-specific PSHAs in 2004 (PEGASOS; NAGRA, 2004) and 2013 (PEGASOS Refinement Project, PRP; swissnuclear, 2013), as well as being covered by two European studies (European Seismic Hazard Model 2013, ESHM13 - Woessner et al., 2015; European Seismic Hazard Model 2020, ESHM20 - Danciu et al., 2021; 2022) and a national (Swiss Hazard Model, SUIhaz15 - Wiemer et al., 2015) study. Only PGA hazard curves for the 5th, 16th, 50th (median), 84th and 95th fractiles and the mean were considered. The hazard curves for PGA from the final reports of the PEGASOS and PEGASOS Refinement Projects were digitised, linearly interpolated in the logarithmic domain and then converted using the Poissonian assumption to probabilities of exceedance in 50 years. The hazard curves from the other three studies were obtained directly from the website of the European Facilities for Earthquake Hazard and Risk (Haslinger et al., 2022). These data were already in terms of probabilities of exceedance for 50 years but they were linearly interpolated in the logarithmic domain when required. The hazard results for ‘rock’ site conditions were considered. No attempt was made to adjust these values to a uniform definition of ‘rock’; any differences from this aspect are likely to be minimal. Finally, because of their common use for engineering purposes, two return periods: 475 years (10% probability of exceedance in 50 years) and 2475 years (2% probability of exceedance in 50 years) are considered.

All five studies find that Beznau is a low to moderate hazard site, with mean PGAs for a return period of 475 years between about 0.04 and 0.10g and between about 0.09g and 0.23g for a return period of 2475 years. The hazard results for this site from the five models are summarised in Figure 1. This figure allows the proposals of

Abrahamson (2017) to be assessed. Applying these to, for example, the pair of PEGASOS and PRP suggests that this change is *robust* as both the inequality is satisfied, and the old mean PGA is outside the 25th and 75th fractiles. On the other hazard, the change from SUIhaz15 to ESHM20 is not *robust* according to the same criteria. Based on a visual inspection of the spread of the hazard results this makes sense.

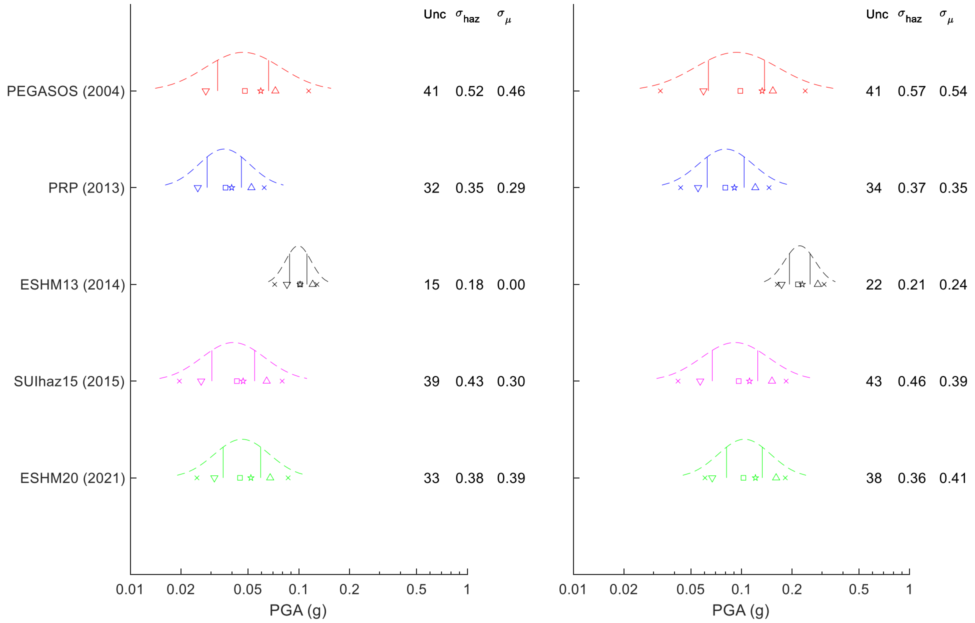

Figure 1: Hazard results for Beznau (Switzerland) from five studies for PGA and return periods of 475 (left) and 2475 years (right). Crosses correspond to 5th and 95th fractiles, downward triangle to 16th fractile, upward triangle to 84th fractile, square, vertical lines to the 25th and 75th fractile, square to the 50th (median) fractile and asterisk to the mean PGA. "Unc" is the measure of uncertainty of the hazard results computed by Douglas et al. (2014), σ_{haz} is the standard deviation of the lognormal distribution fitted to the fractiles and σ_{μ} is the standard deviation of the ground-motion model estimated using the approach proposed by Douglas (2018).

Assuming McGuire (2012)'s threshold of ±25% difference between AFEs at the ground-motion level corresponding to 10^{-4} (return period of 10,000 years) applies for these two return periods, this test can also be applied to these data. Considering the same pairs of hazard models as above, and using the McGuire terminology, this criterion leads to the conclusion that PEGASOS and PRP are again *significantly* different as the change in AFE is about 250% at the 475-year PGA of about 0.06g, whereas the SUIhaz15 and ESHM20 are not *significantly* different as the change in AFE is only about 10%.

Estimating the means and standard deviations of the PDFs from the hazard curves that are required by the approach of Malhotra (2014, 2015) leads to the following values of d (Cohen's effect size): 0.17 for the change between PEGASOS and PRP and 0.049 for the change between SUIhaz15 and ESHM20. Neither of these changes are, therefore, considered *large* according to this criterion (it is recalled that 0.8 is the threshold). The reason for the different conclusions for the pair PEGASOS and PRP compared with the other three proposals is because the standard deviations of the PDFs are large relative to the means. The only pairs where the change in hazard is considered *large* are those involving the ESHM13 model because the differences in the means of the PDFs are large relative to the size of the standard deviations.

3.2 Italy

Quite often, the proposal of a new national seismic hazard model (NSHM) triggers an intense debate on its validity and the differences with the existing model. For this reason, in this section we compare PGAs for a ‘rock’ site from the new NSHM for Italy (MPS19; Meletti et al., 2021) with those from the model that is currently used for the building code (MPS04; Stucchi et al., 2011).

MPS04 considers the larger horizontal acceleration component instead of the geometric mean of the two horizontal components as used in MPS19. Hence, for a more coherent comparison between the two models, we multiply the PGA of MPS19 by a constant factor of 1.15, which represents the average ratio between larger horizontal component and the geometric mean of the two horizontal components (Meletti and Marzocchi, 2019). Here, we compare the seismic hazard for four representative cities: Bologna, Florence, L’Aquila and Syracuse. In Table 1, we show the results of the four tests described in the previous sections. Note that the McGuire (2012) test aims at comparing hazard results at very low AFEs, which are not calculated for the Italian NSHM. Therefore, here we apply this method considering return periods of 475 and 2475 years (10% and 2% probabilities of exceedance in 50 years). As regards the first method proposed by Abrahamson (2017), we consider the 16th and 84th fractiles because they are the most commonly available.

The results indicate that each metric highlights different aspects of the differences in hazard models. In particular, the McGuire (2012) and Malhotra (2014, 2015) methods do not account for epistemic uncertainty because they only use the mean hazard curves. The two methods proposed by Abrahamson (2017) only account for the epistemic uncertainty of MPS19 and neglect the epistemic uncertainty of MPS04. This means that we are just checking if the mean hazard of MPS04 is compatible or not with MPS19, accounting for its epistemic uncertainty. The second Abrahamson (2017) method highlights only increases and not decreases in the NSHM (from MPS04 to MPS19). This is because an increase in the assessed seismic hazard has potentially a much larger impact on engineering practice than a decrease. For example, MPS19 implies a much lower seismic hazard for Syracuse: this difference is highlighted by the first method proposed by Abrahamson (2017), but not from the second one.

	McGuire percentage [percentage>25%]	Malhotra distance [d>0.8]	Abrahamson fractiles; [$\mathrm{PGA}_{\mathrm{MPS04}}$ outside 16th–84th $\mathrm{PGA}_{\mathrm{MPS19}}$]	Abrahamson inequality [>0]
Bologna		0.07		
2%in50	**42%**		0.28, 0.28–0.43	**0.25**
10%in50	20%		0.16, 0.13–0.20	**0.08**
Florence		0.04		
2%in50	15%		0.22, 0.19–0.36	-0.04
10%in50	11%		0.13, 0.09–0.17	-0.29
L'Aquila		0.2		
2%in50	1%		0.45, 0.34–0.53	-0.11
10%in50	9%		**0.26, 0.16–0.25**	-0.22
Syracuse		0.54		
2%in50	**42%**		**0.46, 0.16–0.37**	-0.72
10%in50	**47%**		**0.20, 0.08-0.14**	-0.75

Table 1: Results of the four methods applied to the two Italian models. The values in bold indicate a difference that meets the specific criterion.

3.2 Europe

We have extended our analysis to include five locations in the Euro-Mediterranean region: Bucharest (Romania), Izmit (Türkiye), Zagreb (Croatia), Lisbon (Portugal), and Syracuse (Italy), and compared the ESHM13 (Woessner et al., 2015) and ESHM20 (Danciu et al., 2021). Figure 2 illustrates side-by-side comparison of hazard curves for ESHM20 and ESHM13, including the mean and five fractiles of the former, alongside the latter's mean hazard curves. Upon a visual inspection of Figure 2, it is evident that the ESHM13 mean values for Bucharest and Izmit are lower than the ESHM20 values, while in Lisbon and Syracuse, the ESHM13 values are higher. As for Zagreb, the mean hazard curves of both models appear similar. Furthermore, Table 2 also includes examples of criteria for evaluating how different the two models are at specific locations. As shown in Figure 2, the differences between the two hazard models in Bucharest, Izmit, Lisbon, and Syracuse, are clearly indicated by all factors, whereas for Zagreb, the differences are negligible.

The percentage criterion applied either to AFE or PGA levels, indicate that for all cities, but Zagreb, exceeds the percentage criteria of 25% proposed by McGuire (2012). Note that the percentage change on PGA values of two models is most often used criteria to quantitatively assess the differences between the two models. Percentage difference maps between ESHM13 and ESHM20 are presented in Danciu et al. (2021, 2022). The Cohen's effect size (Malhotra (2014, 2015), *d*, indicates changes between the two models in the following order: Izmit, Lisbon, Bucharest, Syracuse and Zagreb. The negative value of *d* for Izmit ~ -0.83 exceeds the threshold criteria of 0.8, and the changes between the two models are described as *large* according to this criterion.

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

Furthermore, the out-of-range fractile (Abrahamson, 2017) of the ESHM13 PGA values is observed for Lisbon and Syracuse for both return periods as given in Table 2, and visually shown in Figure 2. Similarly, the left-hand side of equation 2 (Abrahamson (2017) is reported. Generally, we find: i) a negative value (i.e., the change is not *robust*) when ESHM13 is greater than ESHM20 or ii) a very small positive value (models are similar). The σ_{ESHM20} can be understood also as the distance between the quantiles, and in Figure 2 the spread between the quantiles appears to be larger for Izmit than for the other cities. The obtained value of σ_{ESHM20} is 0.66, which is double when compared with the values at the other locations (i.e., 0.3 to 0.38). It is worth mentioning, that such variations are caused by differences in modelling assumptions, seismogenic source models, and ground-motion models (Danciu et al., 2021; Danciu et al., 2022).

Among the quantitative factors and criteria considered to investigate differences between two models, the percentage difference is the most intuitive because it is simple to understand and widely used. It shall be noted that the percentage changes in AFE are higher than those computed from PGA values, and one should be aware that minor changes on small numbers often gives rise to large percentage changes.

The Cohen's effect size, *d*, might have the advantage that accounts for the sigma of both hazard models; though, computing the median and sigma for each model based on hazard curves is not straightforward. In this light, determining whether the mean or median fits within a given interquartile range may be a better tool for understanding the difference in the estimates; a value outside the uncertainty range will indicate either a reduction in the models' uncertainty range with time, and an increase or decrease in the mean/median hazard. Ideally, the user should be able to deduce the likely source of these changes from the models' documentation and data (Kohrangi et al. 2018).

Out of all the cities investigated, Izmit was identified by all criteria as portraying an *significant* change in hazard results. Notably, the hazard curves for Izmit depict a 50% increase with the new model, and this difference increases towards the lower probabilities (as illustrated in Figure 2), while the ESHM20 uncertainty range is the largest of all locations. Conversely, Lisbon displays a reduction in hazard values from ESHM20 to ESHM13, which can also be characterized as *significant*.

The following sections will examine the implications of seismic design and the seismic risk perspective.

	McGuire fractile [AFE percentage >25%]	McGuire fractile [PGA percentage >25%]	Malhotra distance [d>0.8]	Abrahamson fractiles		Abrahamson inequality [>0]
				PGA_{ESHM13}	16th–84th PGA_{ESHM20}	
Bucharest (RO)			-0.53			
2%in50	**-55%**	**-21%**		0.43	0.33–0.67	**0.08**
10%in50	**-43%**	**-20%**		0.24	0.19–0.39	**0.06**
Izmit (TR)			-2.05			
2%in50	**-80%**	**-47%**		0.68	0.39–1.38	**0.30**
10%in50	**-63%**	**-46%**		0.36	0.23–0.81	**0.29**
Zagreb (HR)			0.01			
2%in50	-13%	-2%		0.51	0.29–0.67	-0.16
10%in50	4%	1%		**0.25**	**0.15–0.31**	-0.19
Lisbon (PT)			0.64			
2%in50	**490%**	**82%**		0.51	0.15–0.36	-0.81
10%in50	**295%**	**80%**		**0.24**	**0.09–0.16**	-0.80
Syracuse (IT)			0.25			
2%in50	**-106%**	**-37%**		**0.51**	**0.23–0.47**	-0.53
10%in50	**-63%**	**-32%**		**0.25**	**0.12-0.25**	-0.38

Table 2: Results of the four methods applied to the two European models. The values in bold indicate a difference that meets the specific criterion.

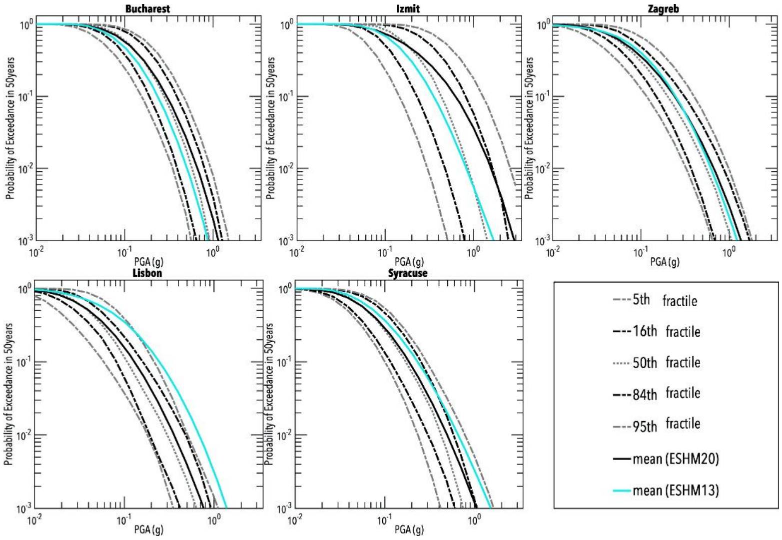

Figure 2: Hazard curves for ESHM20 and ESHM13 in Bucharest, Izmit, Zagreb, Lisbon and Syracuse. Mean and five fractiles of ESHM20 and mean values for ESHM13, respectively.

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

4 Importance of hazard differences for risk reduction and management

Risk reduction or management solutions can include the implementation of building codes, development of retrofitting campaigns, definition of post-disaster emergency plans, or the creation of financial instruments to transfer the risk from the public sector to the (re)insurance market. Methods to assess the impact of changes in seismic hazard on such applications have received limited attention in the open literature. Here we explore the impact of changes in hazard on various risk metrics, and from various viewpoints. We begin by illustrating ways in which the change in hazard might be important for the public and private sectors that are designing risk management applications. Then we look at the potential impact of changes in hazard on the application of building codes.

4.1 Impact of hazard changes on the design of risk management applications

The incorporation of risk results such as average annualised losses, losses for specific return periods (Silva, 2016) or the impact of specific events in risk reduction measures within the public sector has been demonstrated in several studies. For example, Dolce (2012) described how 1 billion euros were distributed throughout the Italian provinces in proportion to the average annual human losses of each region. The earthquake risk was calculated considering the seismic hazard model supported by the National Institute of Geophysics and Volcanology (INGV), and naturally a different distribution of risk across the country would have been obtained if another seismic hazard model had been used. Such modelling options would directly affect the available funds for each province to support risk reduction measures. Risk metrics are also used for urban planning or to assess the needs for temporary shelters in case of destructive events (e.g., Erdik and Durukal, 2008; Anhorn and Khazai, 2014).

To explore these applications, we calculated the expected loss ratios for three return periods (250, 475 and 2475 years) for an unreinforced masonry building. We used the seismic hazard curves for Beznau (see Figure 1) and the vulnerability model proposed by Martins and Silva (2020). We note that we did not consider uncertainty in the vulnerability model, and thus the hazard return period and loss ratio return period are the same. We purposely used a vulnerable building class as this type of construction is common in Europe (Crowley et al., 2020) and is of particular interest within the scope of risk reduction in the public sector. We note that we considered the 250-year return period (which is not typically provided as an output of hazard model studies), since losses for more frequent return periods are often needed for risk management. The median loss ratios and 16/84th fractiles for the three return periods are presented in Figure 3, along with a probability density function approximated by a Gaussian (normal) distribution. As expected, the loss ratios follow the same trend observed for the hazard results presented in Figure 1. However, we note that minor differences in the seismic hazard may lead to *important* differences in the risk. For example, the hazard proposed within the SUIhaz15 and ESHM20 projects for the 475-year return period differs by less than 5% (and indeed according to the previous methods they can be deemed as not *significantly* different), but the loss ratios produced with these models for the same return period differ by 18%. If we consider the hazard models that produce the minimum and maximum seismic hazard (PRP and ESHM13, respectively), the maximum PGA for the 475 years is 2.8 times greater than the minimum PGA, while the corresponding maximum loss ratio is 19.6 times larger than the minimum loss ratio.

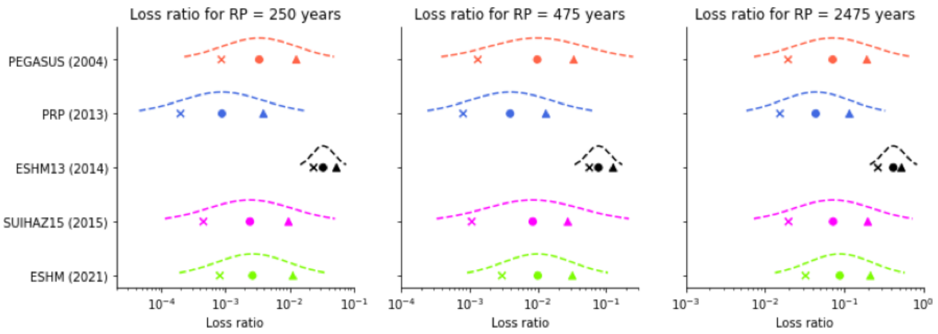

Figure 3: Loss ratio results for Beznau (Switzerland) storeys from five studies for PGA and return periods of 250 (left), 475 (centre) and 2475 years (right), and an unreinforced masonry building with 3 storeys. Crosses correspond to the 16th fractile, circles represent the median, and upward triangle correspond to the 84th fractile. The dashed lines represent an approximation of the probability density functions of the distribution of loss ratios.

Unlike what has been described previously for the seismic hazard component, there are no methods to evaluate how and when a variation in a seismic risk model (or seismic risk parameter) should be considered *significantly* different. This is partially due to the paucity of risk models in the public sector (in comparison with the hazard counterpart), as well as the fact that changes in the risk results can be caused not just by the seismic hazard, but also by variations in the exposure and vulnerability components. In the (re)insurance industry, there is a higher level of scrutiny between different versions of seismic risk models. Some changes might imply modifications in the design of risk transfer solutions, or the adjustment of existing products. For example, for the development of the Turkish Catastrophe Insurance Pool (TCIP), a probabilistic seismic risk model was developed for the country to understand the average annualised loss for each region, as well as the expected loss for specific return periods (Bommer et al. 2002). More recently, parametric insurance products (e.g., Goda 2021) have been proposed for specific parts of the world (e.g., Chile[1], Peru, Colombia and Mexico), which also rely on probabilistic seismic risk results to define the triggering thresholds. The criteria used by the insurance industry and regulators to decide whether a change in a seismic hazard model causes *important* changes in the resulting risk metrics are not publicly available, and usually depend on internal policies and the interests of the clients. From the experience of the authors working with partners from this sector, a change in the risk results of more than 10% requires a justification, while variations of more than 25% have resulted in the rejection of new risk models by national regulators.

To explore these applications, we have repeated the calculations presented previously for the unreinforced masonry building, but now considering a moderate code reinforced concrete building with 4 storeys (Martins and Silva 2020). The consideration of a different building class is prompted by two reasons: 1) vulnerable building classes are unlikely to be insured due to the high risk, and 2) it is important to understand whether the variations

---

[1] Chile CAT Bond - https://www.worldbank.org/en/news/press-release/2023/03/17/world-bank-executes-its-largest-single-country-catastrophe-bond-and-swap-transaction-to-provide-chile-630-million-in-fin

in the risk results are dependent on the building portfolio. The loss ratios for the three return periods are illustrated in Figure 4.

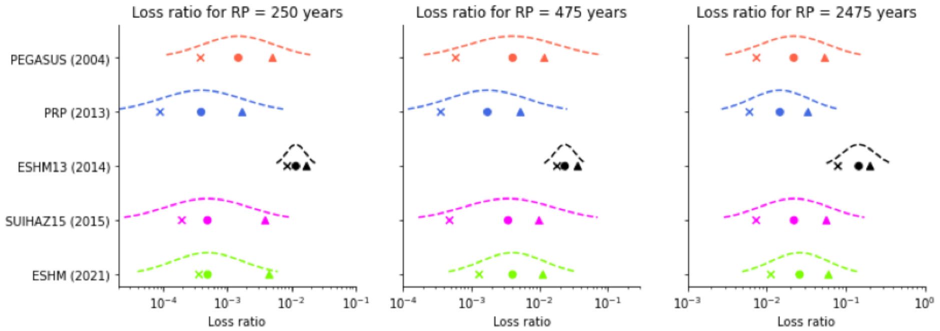

Figure 4: Loss ratio results for Beznau (Switzerland) storeys from five studies for PGA and return periods of 250 (left), 475 (centre) and 2475 years (right), and a reinforced concrete structure with 4 storeys. Crosses correspond to the 16th fractile, circles represent the median, and upward triangle correspond to the 84th fractile. The dashed lines represent an approximation of the probability density functions of the distribution of loss ratios.

Although these results seem to follow a similar trend to that shown previously in Figure 3 for the vulnerable building class, a lower impact was observed for the reinforced concrete structure. For example, the risk results produced with the PRP and ESHM13 seismic hazard previously led to differences by a factor of 19.6, while in this case this factor reduces to 12.4 (for the 475-year return period). Nonetheless, for the two cases of almost identical seismic hazard (i.e., SUIhaz15 and ESHM20), the differences in the risk metrics are above 10% for all return periods. This means that while the previously discussed criteria would deem both models not *significantly* different, most likely the differences in the risk results would trigger further investigation by the catastrophe risk modelling and insurance industries.

4.2 Impact of hazard changes on the application of building codes

Previous studies have looked at the impact of changes in seismic hazard on the design of buildings. For example, Gkimprixis et al. (2021) designed a 4-storey 3-bay reinforced concrete building to different levels of design peak ground acceleration (PGA) and showed that a variation of the design PGA values by as much as 60% in areas of high seismic hazard (in Italy) only affected the initial cost of construction by around 2%. The change in life cycle costs (i.e., the combination of the cost of construction and losses due to damage over the life of the structure) between hazard models was seen to be slightly higher, up to 7%. The largest impact was instead seen on the risk (in terms of the average annual frequency of collapse and the average annual loss), with differences of up to two orders of magnitude.

Given the limited impact that the change of hazard has been seen to have on the design and life cycle costs, in this section we focus on the impact that the change of hazard can have on existing structures, rather than new (yet to be built) structures. If changes in seismic hazard models are directly implemented as changes to seismic actions in seismic design/strengthening codes, this could lead to a large number of existing structures that would suddenly become no longer 'code compliant' from a life-safety viewpoint and would require seismic retrofitting. Some of these buildings might also have only recently been retrofitted (which is likely to be the case in countries such as Italy thanks to SismaBonus, the recently promoted national seismic strengthening strategy).

We propose a four-step method to evaluate whether a change in hazard could lead to a change in life-safety code-compliance, and we argue that if this occurs, such a change is *important* regardless of the actual amount by which the hazard has changed.

The steps of the method are as follows:

1. Design a given building class to modern design principles (e.g., Eurocode 8) for a number of different levels of PGA (or spectral acceleration) and produce collapse fragility functions for each design level.
2. For a given location, obtain the design seismic actions from a given hazard model (typically the mean PGA with a 10% probability of exceedance in 50 years, though it could be the median, if that is prescribed in the code, or it could refer to other return periods, and could also even account for importance factors), and select the fragility function for this design level.
3. Calculate the average annual probability of collapse (AAPC) by convolving the mean hazard curves[2] (from the design hazard model) with the collapse fragility function. Ensure that the target AAPC for code compliance (discussed further below) is met.
4. Repeat step 3 with the same fragility function, but with hazard curves from an alternative (revised) hazard model. Check if the target AAPC for code compliance is exceeded.

In recent years there has been an increase in studies that assess the fragility of buildings designed to modern seismic design principles for varying levels of seismic hazard, which provide useful input to the first step of the method. Examples include Gkimprixis et al. (2020) and Martins et al. (2018), who have produced fragility functions for mid-rise reinforced concrete structures designed to Eurocode 8 (CEN, 2004), and Suzuki and Iervolino (2021), who present fragility functions for residential reinforced concrete and masonry buildings and industrial steel and precast concrete frames designed to the Italian Norme Tecniche delle Costruzioni (NTC). Figure 5 shows the linear regression fit between the parameters of the lognormal fragility functions (namely the median and dispersion, denoted as theta and beta herein) and the design peak ground acceleration obtained with the numerical models from Martins et al. (2018).

---

[2] The hazard curves have been truncated to a return period of 100,000 years, and thus the probability of collapse is assumed to be 1 for all return periods above 100,000 years (see e.g., Suzuki and Iervolino, 2021). This effectively means that the lowest AAPC that can be computed is 1 in 100,000 or 10^{-5}.

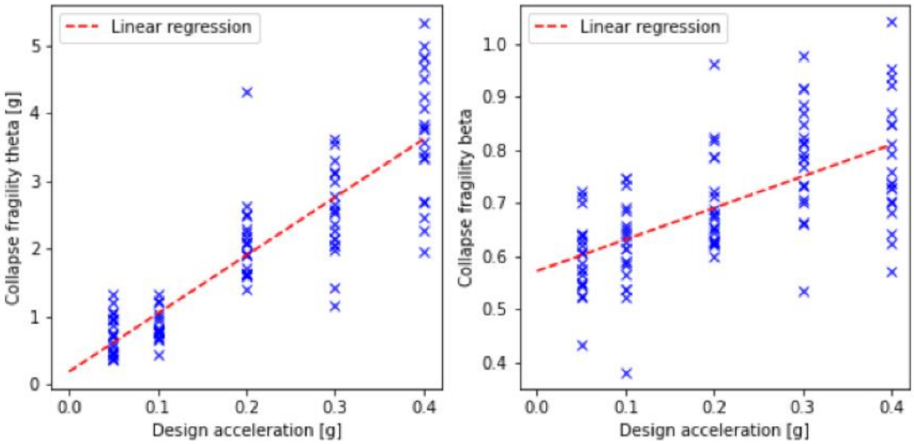

Figure 5: Linear regression between design PGA and the theta (PGA) and beta parameters of the lognormal fragility functions from Martins et al. (2018) for several 3 to 5 storey RC frames designed to EC8 site class A

The target AAPC for code compliance can be taken as 2 x 10^{-4} (which can also be expressed as a 1% probability of collapse in 50 years), following the recommendations introduced in the ongoing update to Eurocode 8 (in an Informative Annex), as well as in the ASCE Standard 7-10 (ASCE, 2010).

The hazard models used in the test cases above have been used to demonstrate this method. For Beznau, the design hazard has been obtained using the mean PGA (475-year return period) from the five hazard models presented in Figure 1. The AAPC has then been computed with the mean hazard curves from each of these studies. The results are presented in Table 3 where for a given design hazard model (say PRP, shown in the first column), the AAPC assessed according to each hazard model is shown. These results show that the target AAPC is respected for in all cases (i.e., is always less than 2 x 10^{-4}), regardless of the combination of design hazard and revised hazard. Previously it was shown that the change in hazard from the PEGASOS to PRP was *robust*, but these result show that the increase in mean hazard levels from the PRP model to the PEGASOS model does not affect the code compliance of the building class considered herein. Hence, from an engineering perspective, this change would not be *important* for the existing building stock.

Nevertheless, it is worth noting that even if life safety code-compliance is maintained, these results can give insight into the increase in the assessed risk due to a change in the hazard model. In the case of a PRP-designed structure which is then assessed with the ESHM20 model, the AAPC increases by a factor of 2. For the case of buildings designed to the PRP hazard and assessed with the ESHM13 code, the AAPC is seen to increase by almost an order of magnitude. This large impact of changing to the ESHM13 hazard is consistent with the findings on the *significance* of this hazard update, presented previously. On the other hand, for the buildings designed to the PEGASOS or SUIHaz15 hazard models, the AAPC would be almost unchanged if the hazard model were to be updated to the SUIHaz15 or ESHM20 models, respectively. This lack of *importance* in the change of hazard agrees with the findings of the methods applied to the hazard models previously.

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

Hazard model	Hazard model used in assessment (AAPC, × 10^{-5})				
used in design	PEGASOS	PRP	ESHM13	SUIHaz15	ESHM20
PEGASOS	2.81	1.61	10.7	2.81	2.84
PRP		2.05	15.5	3.81	3.95
ESHM13			3.41	1.38	1.35
SUIHaz15				3.10	3.16
ESHM20					3.02

Table 3: Values of AAPC for mid-rise RC buildings in Beznau designed according to five different hazard models and assessed using the same or subsequently developed hazard models.

The same exercise has been carried out using the European and Italian hazard curves presented in the previous section. In this case only two hazard models have been compared: ESHM13 and ESHM20 for Europe and MPS04 and MPS19, for Italy. The same example cities as before have been considered in both cases, with the results presented in Tables 4 and 5.

For the European hazard, there is only one city where the hazard change from ESHM13 to ESHM20 can be deemed *important*, shown in red in the table, as the AAPC exceeds the acceptable threshold due to the change in hazard. This would imply that buildings designed in Izmit due to the ESHM13 hazard would no longer respect the code according to the latest insights given by the ESHM20 model. This finding is perhaps not surprising, given that the design hazard PGA increased by 84% from ESHM13 to ESHM20. In all other cases the AAPC remains below the threshold, even though the AAPC doubles in Bucharest. Another interesting insight from this table is the comparison of the AAPC for buildings that are both designed and assessed with each hazard model. In Zagreb, the slight decrease in design hazard (together with the changes across the hazard curve) from ESHM13 to ESHM20 leads to an increase in the implicit risk of buildings designed to code. In Syracuse, despite a 30% decrease in the design hazard PGA level from ESHM13 to ESHM20, the implicit risk remains fairly constant (4.41×10^{-5} versus 4.18×10^{-5}).

Hazard model used in design	Hazard model used in assessment (AAPC, × 10^{-5})	
	ESHM13	ESHM20
Bucharest		
ESHM13	2.46	5.39
ESHM20	-	3.86
Izmit		
ESHM13	5.91	**24.3**
ESHM20	-	13.8
Zagreb		
ESHM13	3.77	4.78
ESHM20	-	4.82
Lisbon		
ESHM13	5.22	1.50
ESHM20	-	2.84
Syracuse		
ESHM13	4.41	2.43
ESHM20	-	4.18

Table 4: Values of AAPC for mid-rise RC buildings in designed according to two different European hazard models and assessed using the same or subsequently developed hazard model, for five example European cities. The value in bold is above the threshold of acceptable AAPC.

For Italy, the change in hazard from MPS04 to MPS19 would not be deemed *important* in any of the locations considered, as in all cases the change in hazard does not lead to an exceedance of the threshold AAPC. In Florence the AAPC increases three-fold, but still remains below the threshold. In Syracuse, despite a 35% reduction in the design PGA from MPS04 to MPS19, the implicit risk of buildings designed to each hazard model is seen to be very similar, which implies an increase in the hazard levels at other return periods that are influential on the AAPC. In Bologna the design PGA increases by around 20%, but the implicit risk is actually seen to be lower for buildings designed and assessed with MPS19. This is also the case in L'Aquila, where similar levels of design PGA are found in both hazard models, but the implicit risk of buildings designed and assessed with the later hazard model is lower.

Hazard model used in design	Hazard model used in assessment (AAPC, x10^{-5})	
	MPS04	MPS19
Bologna		
MPS04	4.81	5.91
MPS19		4.28
Florence		
MPS04	1.97	5.37
MPS19	-	3.70
L'Aquila		
MPS04	5.91	4.71
MPS19	-	4.63
Syracuse		
MPS04	5.04	2.86
MPS19	-	5.26

Table 5: Values of AAPC for mid-rise RC buildings in designed according to two different Italian hazard models, and assessed using the same or subsequently developed hazard model, for four example Italian cities.

5 Discussion

Changes in the assessed seismic hazard for a location over time can have varying implications in different contexts. These implications are discussed in this section.

The development of risk management measures often relies on risk metrics such as return period losses or average annual losses. The evaluation of the impact of variations in the seismic hazard indicated that small changes on the seismic hazard can lead to *important* differences on risk metrics. These findings indicate that updating seismic hazard models may render insurance products inadequate and can affect the distribution of public funds for risk reduction.

Seismic building codes imply an 'acceptable' level of risk for individual buildings. Sometimes this is explicitly stated (e.g., in the updated Eurocode 8 the annual probability of exceeding the near collapse limit state is 2×10^{-4}) or implicit. For the design and construction of new buildings, as shown above, changes in the hazard can affect the assessed annual probability of collapse (or other damage states). Therefore, if society does not want the population living in these buildings to be under a higher risk than previously thought (even if the risk still falls within 'acceptable' thresholds), the seismic actions in the building code would need to be adapted to reflect this change (higher seismic resistance in the case of higher hazard results and lower in the case of lower hazard). This change would then affect the life cycle costs (cost of construction and losses due to damage over the life of the structure) of the building, although if the changed hazard is considered at the design stage the effect on the life cycle costs would likely be minimal.

What change in the assessed risk to the users of the building or the life cycle costs is considered *important* is a question for society (government, regulators, builder's owners and the general public). This is a difficult question as there is uncertainty estimating this risk and, without employing risk-targeted seismic actions in the code (e.g., Luco et al., 2007), the risk will vary from one location to another. The risk may thus already exceed acceptable thresholds in some locations, especially those with higher levels of hazard (e.g., Silva et al., 2015, Iervolino and Pacifico, 2021), which should therefore be taken into account even before considering the impact that a change to the hazard could make. The answer to this question would then help determine the *importance* of changes in the assessed hazard. As a means of maintaining stability in building codes, it may be advisable that reductions in levels of hazard, from one generation of models to another, should not be implemented in the code for the design of new buildings and 'held in reserve' for the future.

For an existing building, if the reassessed hazard leads to an indication that the annual probability of collapse (or other risk measure) has decreased, then there is nothing that needs to be done. In contrast, in the case where the assessed hazard has increased from a previous generation then this may indicate that the building should be retrofitted to increase its capacity. If a structure has recently been constructed or retrofitted based on the previous hazard model, then the debate over whether additional capacity should be added could be intense. In the case of older unmodified buildings, in contrast, they may need improvements in any case due to the changes in design and construction standards from older building codes. Even without changes in the assessed hazard, seismic building codes have evolved greatly in the past fifty years, and so what to do with buildings constructed to older codes is a common problem. Countries such as Italy have recently promoted efforts to strengthen older structures (e.g., Sismabonus). If a decision has already been made to retrofit a structure through these programmes, then any increase in the hazard would be relatively easy, and cost relatively little more, to address at the same time.

It is not our role to provide thresholds to decide whether a change to the assessed hazard is important or not but there are actions that the construction industry could take to become more resilient to changes in seismic hazard modelling. Some ideas are discussed in the following paragraphs.

Firstly, policies to mitigate the impact of changes to the assessment of hazard and structural capacity (which is also continually being updated in building codes) are needed. Examples of such measures might be an explicit code-prescription and enforcement of time windows that define how long a building can remain non-code-compliant (with the time as a function of the level of increase of risk, often proxied by a seismic capacity/demand ratio), or enforcing a more realistic (less conservative) assessment of the capacity of recently designed/retrofitted buildings. For example, linear methods to assess the capacity of structures are admittedly conservative, and a building may be perceived as less likely to be code compliant when its seismic capacity has been evaluated using a linear structural analysis method as opposed to a nonlinear method. In addition, there are also many non-structural elements that contribute to the strength of a building (e.g., infill panels) that, again for the sake of simplifying structural analyses, are often overlooked in seismic assessment, whilst they may lend additional seismic resistance to the building.

Secondly, the design ground motions imposed by the seismic building code could be made more robust to changes in the hazard model. For example, the epistemic uncertainty in the assessed hazard could be better accounted for by using a higher fractile of the ground motion than using the expected (mean) value. This was suggested by

https://doi.org/10.5194/egusphere-2023-991
Preprint. Discussion started: 19 June 2023

McGuire et al. (2005), who note that designing for a higher level than implied by the mean hazard could be more cost-effective as it would avoid the cost of future retrofits in case of a revised hazard assessment. Using a higher fractile would mean that as epistemic uncertainties theoretically should decrease with new generations of hazard models the design ground motions would stay stable (or even decrease). As an example, in the UK, it is common to use the 84th fractile of the response spectral accelerations for 10^{-4} AFE rather than the mean (or median) as a conservative estimate of the hazard (ONR, 2022). A switch to using a higher fractile would, however, likely lead to large increases in the assessed ground motions within seismic building codes, which, although there would be the promise that they would decrease with time, would likely lead to difficulties in the short and medium term.

Thirdly, if the acceptable levels of risk or life cycle costs were explicitly stated as part of the building code, then conversations over whether a structure needs to be retrofitted (and consequently the importance of changes in the assessed hazard) would be easier. Defining these acceptable levels would, however, require considerable work and societal decisions.

Fourthly, inspiration could be sought from the Senior Seismic Hazard Analysis Committee (SSHAC) guidelines that are often used within nuclear-related seismic hazard projects, as these guidelines explicit refer to the need for longevity and stability in hazard results between different generations of hazard models (Budnitz et al., 1997). Recent SSHAC implementation documents have provided recommendations for when hazard studies should be updated (Ake et al., 2018).

Lastly, there should be a clear separation between the development of hazard models and their implementation for engineering purposes (Jordan et al., 2014). The inputs to a hazard model should capture our scientific understanding and our uncertainty without being constrained by what implications these inputs could have on changes in the ground motions for a given AFE and fractile. The implementation of the results of the hazard model within the seismic building code should account for the epistemic uncertainty in the results. Methods to assess the *significance* of the raw hazard results from different models will necessarily be different to those that assess the *importance* of differences in the final hazard results presented to end users (construction industry, infrastructure owners, insurance companies and general public). The large epistemic uncertainties captured in the raw hazard results mean that different models may be compatible, and the *significance* of differences are small, but when the hazard models are implemented for use by different communities there may be *important* implications of these differences.

6 Conclusions

This article discussed the vital topic of judging whether differences between results from different probabilistic seismic hazard assessments warrant further consideration and whether they are *important* in different contexts. The first conclusion is that there is no universal way of deciding whether a difference between hazard results is *important* or not. The approach used and the criteria adopted should depend on the application, e.g., whether it is differences between results from assessments used within seismic design codes or between assessments used for risk evaluations for the insurance industry or public risk mitigation campaigns.

When it comes to comparing just the hazard values, we note that the methods previously proposed are not adequate to evaluate formally the statistical *significance* of the difference between two models. Because seismic hazard can only be assessed to a relatively low level of precision due to the large epistemic uncertainties that are present in such assessments, even for well-studied regions, we think it is vital to consider these uncertainties when comparing hazard results. We would like to emphasise that two models that are not statistically *significantly* different in terms of the hazard values alone may actually lead to great differences when used by engineers for design, analysts when assessing risk or insurance companies when computing premiums. Different terminologies between groups and the ways that differences between hazard models are assessed can lead to lively discussions and confusion. We therefore advocate for the development of structured approaches to the evaluation of evolving hazard models, clearly driven by the intended applications of the models.

Author contribution

JD contributed the conceptualization of this study, led the development of the methodology, conducted the formal analysis for Section 3.1 and supervised the work. HC and RP contributed to the development of the methodology and conducted the formal analysis and writing for Section 4.2. VS contributed to the development of the methodology and conducted the formal analysis and writing for Section 4.1. WM contributed to the development of the methodology and conducted the formal analysis and writing or Section 3.2. LD contributed to the development of the methodology and conducted the formal analysis and writing for Section 3.3. JD led the writing and all authors contributed to the writing, review and editing.

Competing interests

Some authors are members of the editorial committee for this special issue of Natural Hazards and Earth System Sciences. The peer-review process was guided by an independent editor, and the authors have no other competing interests to declare.

Acknowledgements

We would like to thank the European Facilities for Earthquake Hazard and Risk for organising the webinar where the topics covered in this article were originally discussed. Also, we thank the developers of the seismic hazard models used in this analysis for providing access. We also gratefully acknowledge Dr Luis Martins for sharing the fragility functions and code that was used for the analyses in Section 4.2.

References

Abrahamson, N. (2017), Treatment of epistemic uncertainty in PSHA results, *PSHA workshop*, Lenzburg, Switzerland.

Ake, J., Munson, C., Stamatakos, J., Juckett, M., Coppersmith, K. and Bommer, J. (2018), Updated Implementation Guidelines for SSHAC Hazard Studies, NUREG-2213, Office of Nuclear Regulatory Research, Washington.

Anhorn, J. and Khazai, B. (2014), Open space suitability analysis for emergency shelter after an earthquake, *Natural Hazards and Earth System Sciences Discussions*, **1**(2), 4263-4297.

Apostolakis, G. (1990), The concept of probability in safety assessments of technological systems, *Science*, **250**(4986), 1359-1364, https://doi.org/10.1126/science.2255906.

ASCE (2010), Minimum Design Loads for Buildings and Other Structures (ASCE 7-10), American Society of Civil Engineers, Reston, VA.

Beauval, C. and Bard, P.-Y. (2022), History of probabilistic seismic hazard assessment studies and seismic zonations in mainland France, *Comptes Rendus de l'Académie des Sciences - Géosciences*, **353**(S1), 413-440, doi: 10.5802/crgeos.95

Belvaux, M., Douglas, J. and Ulrich, T. (2014), Comparisons between SHARE and current national seismic hazard maps for France, *Workshop: Results of the European Project: SHARE: Seismic Hazard Harmonization in Europe*, DGEB e.V.

Bommer J., Spence R., Erdik M., Tabuchi, S., Aydinoglu N., Booth E., Re D. and Pterken, D. (2002), Development of an Earthquake Loss Model for Turkish Catastrophe Insurance, *Journal of Seismology*, **6**, 431-446.

Budnitz, R. J., Apostolakis, G., Boore, D. M., Cluff, L. S., Coppersmith, K. J., Cornell, C. A. and Morris, P. A. (1997). Recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts, NUREG/CR-6372, two volumes, US Nuclear Regulatory Commission, Washington D.C.

CEN (2004) EN 1998-1:2004 Eurocode 8: design of structures for earthquake resistance—part 1: general rules, seismic actions and rules for buildings. European Committee for Standardization, Brussels

Cohen, J. (1977), Statistical Power Analysis for the Behavioral Sciences, Academic Press, ISBN: 978-0-12-179060-8, doi: 10.1016/C2013-0-10517-X.

Crowley H, Despotaki V, Rodrigues D, Silva V, Toma-Danila D, Riga E, Karatzetzou A, Zugic Z, Sousa L, Ozcebe S, Gamba P (2020). Exposure Model for European Seismic Risk Assessment. Earthquake Spectra. https://doi.org/10.1177/8755293020919429.

Danciu L., Nandan S., Reyes C., Basili R., Weatherill G., Beauval C., Rovida A., Vilanova S., Sesetyan K., Bard P-Y., Cotton F., Wiemer S., Giardini D. (2021) - The 2020 update of the European Seismic Hazard Model: Model Overview. EFEHR Technical Report 001, v1.0.0, https://doi.org/10.12686/a15.

Danciu, L. et al. (2022). The 2020 European Seismic Hazard Model: Milestones and Lessons Learned. In: Vacareanu, R., Ionescu, C. (eds) *Progresses in European Earthquake Engineering and Seismology. ECEES 2022*. Springer Proceedings in Earth and Environmental Sciences. Springer, Cham. doi: 10.1007/978-3-031-15104-0_1.

Dolce, M. (2012), The Italian National Seismic Prevention Program. *Proceedings of the 15th World Conference on Earthquake engineering*, Lisbon, Portugal.

Douglas, J. (2018), Capturing Geographically-Varying Uncertainty in Earthquake Ground Motion Models or What We Think We Know May Change, *Recent Advances in Earthquake Engineering in Europe, Geotechnical, Geological and Earthquake Engineering*, **46**, doi: 10.1007/978-3-319-75741-4_6

Douglas, J., Ulrich, T. and Negulescu, C. (2013), Risk-targeted seismic design maps for mainland France, *Natural Hazards*, **65**, 1099-2013, doi: 10.1007/s11069-012-0460-6.

Douglas, J., Ulrich, T., Bertil, D. and Rey, J. (2014), Comparison of the Ranges of Uncertainty Captured in Different Seismic-Hazard Studies, *Seismological Research Letters*, **85**(5), 977-985, doi: 10.1785/0220140084.

Gkimprixis, A., Tubaldi, E. and Douglas, J. (2020). Evaluating alternative approaches for the seismic design of structures, *Bulletin Earthquake Engineering*, **18**, 4331–4361. https://doi.org/10.1007/s10518-020-00858-4

Gkimprixis, A., Douglas, J. and Tubaldi, E. (2021). Seismic risk management through insurance and its sensitivity to uncertainty in the hazard model, *Natural Hazards*, **108**, 1629–1657, https://doi.org/10.1007/s11069-021-04748-z.

Goda, K. (2021), Multi-hazard parametric catastrophe bond trigger design for subduction earthquakes and tsunamis, *Earthquake Spectra*, **37**(3), 1827-1848, https://doi.org/10.1177/8755293020981974.

Haslinger, F., Basili, R., Bossu, R., Cauzzi, C., Cotton, F., Crowley, H., Custodio, S., Danciu, L., Locati, M., Michelini, A., Molinari, I., Ottemöller, L., Parolai, S. (2022), Coordinated and interoperable seismological data and product services in Europe: the EPOS thematic core service for seismology. *Annals of Geophysics*, **65**(2), DM213. https://doi.org/10.4401/ag-8767.

Iervolino, I. and Pacifico, A. (2021), Fatality rates implied by the Italian building code, *Earthquake Engineering and Structural Dynamics*, **50**(11), 3083-3089, doi: 10.1002/eqe.3472.

Jordan, T.H., Marzocchi, W., Michael, A. and Gerstenberger, M. (2014), Operational earthquake forecasting can enhance earthquake preparedness, *Seismological Research Letters*, **85**(5), 955-959.

Kalakonas, P., Silva, V., Mouyiannou, A. and Rao, A. (2020) Exploring the impact of epistemic uncertainty on a regional probabilistic seismic risk assessment model, *Natural Hazards*, **104**, 997–1020. https://doi.org/10.1007/s11069-020-04201-7.

Kass, R. E., and Raftery, A. E. (1995). Bayes factors, *Journal of the American Statistical Association*, 90(430), 773–795.

Kohrangi, M., Danciu, L. and Bazzurro, P. (2018), Comparison between outcomes of the 2014 Earthquake Hazard Model of the Middle East (EMME14) and national seismic design codes: The case of Iran, *Soil Dynamics and Earthquake Engineering*, **114**, 348-361, doi: 10.1016/j.soildyn.2018.07.022.

Luco, N., Ellingwood, B.R., Hamburger, R.O., Hooper, J.D., Kimball, J.K., Kircher, C.A. (2007) Risk-targeted versus current seismic design maps for the conterminous United States. In: SEAOC 2007 Convention Proceedings

Malhotra, P. K. (2014), Cost of Uncertainty in Seismic Hazard, *Technical report, StrongMotions Inc.*, Sharon, USA.

Malhotra, P. K. (2015), Myth of probabilistic seismic hazard analysis, *Structure*, July, 58-59.

Martins, L. and Silva, V. (2020), Development of a Fragility and Vulnerability Model for Global Seismic Risk Assessment, *Bulletin of Earthquake Engineering*, **19**, 6719-6745, https://doi.org/10.1007/s10518-020-00885-1.

Martins, L., Silva, V., Bazzurro, P. and Marques, M. (2018), Advances in the derivation of fragility functions for the development of risk-targeted hazard maps, *Engineering Structures*, **173**, 669-680, https://doi.org/10.1016/j.engstruct.2018.07.028.

Marzocchi, W. and Jordan, T. H. (2017), A unified probabilistic framework for seismic hazard analysis, *Bulletin of the Seismological Society of America*, **107**(6): 2738–2744. https://doi.org/10.1785/0120170008.

Marzocchi, W. and Jordan, T. H. (2014), Testing for ontological errors in probabilistic forecasting models of natural systems, *Proceedings of the National Academy of Sciences*, **111**(33), 11973-11978, https://doi.org/10.1073/pnas.1410183111.

McGuire, R. K. (2012), Precision of Seismic Hazard Evaluations in Central and Eastern North America, *Proceedings of the 15th World Conference on Earthquake Engineering*, Paper number 5501.

McGuire, R. K., Cornell, C. A. and Toro, G. R. (2005), The case for using mean seismic hazard, *Earthquake Spectra*, **21**(3), 879-886, https://doi.org/10.1193/1.1985447.

Meletti, C. and Marzocchi, W. (Eds.) (2019). Il modello di pericolosità sismica MPS19, Final Report, CPS-INGV, Roma, 168 + 2 Appendices (in Italian).

Meletti, C., W. Marzocchi, V. D'Amico, G. Lanzano, L. Luzi, F. Martinelli, B. Pace, A. Rovida, M. Taroni, F. Visini (2021). The new Italian seismic hazard model (MPS19). *Annals of Geophysics*, **64** (1), SE112

National Cooperative for the Disposal of Radioactive Waste (NAGRA) (2004), *Probabilistic seismic Hazard Analysis for Swiss Nuclear Power Plant sites (PEGASOS Project) prepared for Unterausschuss Kernenergie der Überlandwerke (UAK),* Final report Vols. 1/6, 2557 pp., to be obtained on request at swissnuclear by writing to info@swissnuclear.ch.

Office for Nuclear Regulation (ONR) (2022), External Hazards, Nuclear Safety Technical Assessment Guide, NS-TAST-GD-013 Issue 8.1.

Silva V, Crowley H, Bazzurro P (2015). Exploring Risk-targeted Hazard Maps for Europe. Earthquake Spectra, 32(2):1165-1186. doi.org/10.1193/112514eqs198m

Silva V (2017). Critical Issues on Probabilistic Earthquake Loss Assessment. Journal of Earthquake Engineering, 22:9, 1683-1709.

Stucchi, M., Meletti, C., Montaldo, V., Crowley, H., Calvi, G.M. and Boschi, E. (2011), Seismic hazard assessment (2003-2009) for the Italian building code, *Bulletin of the Seismological Society of America*, **101**, 1885-1911.

Suzuki, A. and Iervolino, I. (2021), Seismic fragility of code-conforming Italian buildings based on SDoF approximation, *Journal of Earthquake Engineering*, **25**(14), 2873-2907, doi: 10.1080/13632469.2019.1657989.

swissnuclear (2013), Probabilistic Seismic Hazard Analysis for Swiss Nuclear Power Plant Sites - PEGASOS Refinement Project, Volume 1-6, URL: http://swissnuclear.ch/de/downloads.html.

Tromans, I. J., Aldama-Bustos, G., Douglas, J., Lessi-Cheimariou, A., Hunt, S., Davi, M., Musson, R. M. W., Garrard, G., Strasser, F. O. and Robertson, C. (2019), Probabilistic seismic hazard assessment for a new-build nuclear power plant site in the UK, *Bulletin of Earthquake Engineering*, **17**(1), 1-26, doi: 10.1007/s10518-018-0441-6.

Vick, S. G. (2002), *Degrees of Belief: Subjective Probability and Engineering Judgment*, ASCE Press, Reston, VA. ISBN: 9780784405987.

Wiemer, S., Danciu, L., Edwards, B., Marti, M., Fäh, D., Hiemer, S., Wössner, J., Cauzzi, C., Kästli, P., Kremer, K. (2015), Seismic Hazard Model 2015 for Switzerland (SUIhaz2015), Swiss Seismological Service (SED) at ETH Zurich.

Woessner, J., Danciu L., Giardini, D. et al. The 2013 European Seismic Hazard Model: key components and results, *Bulletin of Earthquake Engineering*, **13**, 3553–3596. doi: 10.1007/s10518-015-9795-1.