



Philosophical aspects of probabilistic seismic hazard analysis (PSHA): a critical review

Luca Zanetti¹ · Daniele Chiffi² · Lorenza Petrini³ 

Received: 24 March 2022 / Accepted: 27 February 2023
© The Author(s) 2023

Abstract

The goal of this paper is to review and critically discuss the philosophical aspects of probabilistic seismic hazard analysis (PSHA). Given that estimates of seismic hazard are typically riddled with uncertainty, different epistemic values (related to the pursuit of scientific knowledge) compete in the selection of seismic hazard models, in a context influenced by non-epistemic values (related to practical goals and aims) as well. We first distinguish between the different types of uncertainty in PSHA. We claim that epistemic and non-epistemic considerations are closely related in the selection of the appropriate estimate of seismic hazard by the experts. Finally, we argue that the division of scientific responsibility among the experts can lead to responsibility gaps. This raises a problem for the ownership of the results (“no one’s model” problem) similar to the “problem of many hands” in the ethics of technology. We conclude with a plea for a close collaboration between philosophy and engineering.

Keywords Seismic risk · PSHA · Uncertainty · Logic trees · Epistemology of risk · Scientific responsibility

1 Introduction

The goal of this paper is to review and critically discuss the literature on the philosophical aspects of probabilistic seismic hazard analysis (PSHA). PSHA has received little attention in the philosophical literature, but its conceptual aspects have been object of continued interest in engineering (Marulanda et al. 2021; Foulser-Piggott et al. 2020; Baker et al.

✉ Lorenza Petrini
lorenza.petrini@polimi.it

Luca Zanetti
luca.zanetti@iusspavia.it

Daniele Chiffi
daniele.chiffi@polimi.it

¹ Scuola Universitaria Superiore IUSS Pavia, Pavia, Italy

² Department of Architecture and Urban Studies (DASU), Politecnico di Milano, Milan, Italy

³ Department of Civil and Environmental Engineering (DICA), Politecnico di Milano, Milan, Italy

2021). These conceptual aspects include the methodology of PSHA (Krinitzsky 1993a, b, 2002; Castaños and Lomnitz 2002; Mulargia et al. 2017), the assessment of seismic hazard models (e.g., Grandori 1991, 1993, 1994), and the scientific responsibility of the participants in the study, including the scientists that formulate the individual models, the experts that evaluate those models, and the team that performs the integration (SSHAC 1997). A clarification of the main issues from a philosophical point of view can contribute to a critical reflection on the foundations of the practice of PSHA.

Throughout the paper, we rely on a distinction between *epistemic values*, which are related to the pursuit of scientific knowledge, and *non-epistemic values*, which are related to practical goals and aims (Diekmann and Peterson 2013). Given that estimates of seismic hazard are typically riddled with uncertainty, different epistemic values compete in the selection of seismic hazard models, in a context influenced by non-epistemic values as well.

We will proceed as follows. In Sect. 2, we distinguish between the different types of uncertainty in seismic hazard analysis. In PSHA, seismic hazard is estimated as the probability of exceedance of a specified ground-motion intensity at a given site. These estimates are subject to two types of uncertainty. On the one hand, there are aleatoric uncertainties, which are due to the essential variability of seismic phenomena. On the other hand, there are epistemic uncertainties, which are due to insufficient understanding of these phenomena. Both these uncertainties are included in PSHA.

From the epistemic point of view, the main problem is that historical catalogs of seismic events, which often comprise only few centuries, are not sufficient for a complete validation of seismic hazard models, which would require thousands of years of data. Due to this problem, PSHA can lead to an *overestimation* of the hazard at a site due to mistakes in the interpretation and elaboration of data (Bommer and Abrahamson 2007), but it can also lead to an *underestimation* of seismic hazard, since “the ‘worst-case scenario’ is unlikely to have occurred yet” (Marulanda et al. 2021).

In Sect. 3, we compare different methodologies for the assessment of seismic hazard models. Given the practical impossibility to validate individual models, the estimates of seismic hazard are often based on ensembles. In an ensemble of models, the final estimate of seismic hazard corresponds to the weighted sum of the estimates produced by the models with the weights determined by the judgments of credibility assigned to those models by a group of experts. We raise some methodological challenges to the use of experts in PSHA. We then consider two alternative approaches. We claim that epistemic and non-epistemic considerations are closely related in the selection of the appropriate estimate of seismic hazard.

Finally, in Sect. 4 we critically discuss the issue of the scientific responsibility of the experts. Models ensembles allow to make decisions that take into account the judgments of the scientific community as a whole rather than those of a single expert, but it also makes it difficult to ascertain individual scientific responsibilities for the final result. We argue that the division of scientific responsibility among experts can lead to *responsibility gaps*, and so it raises a problem for the ownership of the results (“no one’s model” problem) similar to the “problem of many hands” in the philosophy of technology. We conclude with a plea for a close collaboration between the philosophy and sociology of science and earthquake engineering.

2 Probabilistic seismic hazard analysis

Probabilistic seismic hazard analysis (see Baker et al. 2021 for a recent presentation) is usually traced back to Cornell (1968). However, key elements of modern PSHA, including the explicit treatment of ground-motion prediction equations (see below), are due to Esteva (1969). Moreover, the Cornell's formulation is a special case of the Esteva's formulation (Alamilla et al. 2020); see McGuire 2007 for a historical overview.

PSHA estimates seismic hazard¹ as the probability of exceedance of a specified ground-motion intensity at a given site during a specified time interval (or its converse, the return period; for example, 10% in 50 year, corresponding to a return period of 475 years). The frequency of a seismic event corresponds to the annual rate of occurrence of an event that exceeds a given intensity level. Ground-motion acceleration can be measured in terms of peak ground acceleration (PGA), peak ground velocity, peak ground displacement, or response spectral acceleration; for convenience, in this paper we will always use PGA.

The first step in PSHA is to identify all possible sources of a seismic event that can affect the site of interest. An *earthquake rate model* (ERM) consists of a set of events characterized by the magnitude (M) of the earthquake, its location (usually represented as a point on a plane), and its rate of occurrence (for example, 10^{-4} /year). The average annual number of events with magnitude equal to or greater than some m in a given region is expressed by:

$$N_M(m) = \nu_0[1 - F_M(m)] \quad (1)$$

where ν_0 is the average total number of events per year with magnitude equal to or greater than a chosen threshold ("minimum earthquake") and F_M is the magnitude probability distribution function at the source of interest.

The PGA is calculated on the basis of the *ground-motion model* (GMM), or ground-motion attenuation relation, which is used to determine the peak ground-motion acceleration at the site of interest as a function of the magnitude of the event M and of the distance between the site and the source (calculated on the basis of the earthquake's location). A ground-motion attenuation relation has the general form²:

$$\ln \text{PGA} = cM - d \cdot \ln R + e \pm \epsilon, \quad (2)$$

where M is the magnitude of the seismic event, R is the distance from the source, c , d , and e are constants that characterize the site of interest, and ϵ is the deviation from the mean³; however, much more than just three parameters are considered today in the estimate of the ground-motion attenuation relation.

Finally, the frequency of *exceedance* at the site of a given $\text{PGA} = a$ is:

$$\nu(a) = \nu_0 P[\text{PGA} > a], \quad (3)$$

where $P[\text{PGA} > a]$ is calculated by integration over all relevant distances (determined by the ground-motion attenuation relation) and magnitudes (determined by the magnitude probability distribution F_M):

¹ In this paper, we will focus on seismic *hazard* (the probability that an adverse event occurs); similar considerations apply to vulnerability estimates (the probability of a loss was the adverse event to occur) in earthquake engineering.

² Cf. Hanks and Cornell (1994).

³ PGA is assumed to be lognormally distributed.

$$P[\text{PGA} > a] = \int_{M_{\min}}^{M_{\max}} \int_0^{r_{\max}} P[\text{PGA} > a \mid m, r] f_M(m) f_R(r) \, dm \, dr \quad (4)$$

where M_{\max} is the maximum intensity considered at that source, r_{\max} is the maximum distance between the source and the site, f_M is the probability density function of M , and f_R is the probability density function of R . If the site can be affected by n sources, the total hazard is determined by:

$$v(a) = \sum_{i=1}^n v_i(a) \quad (5)$$

The seismic hazard corresponds to the combination of all possible earthquakes that might affect the site of interest weighted by their annual frequencies. Notice that the seismic hazard is itself a mean value (see Eq. 4), therefore it also has an aleatory uncertainty that is not represented in the final estimate.⁴

Seismic hazard is an uncertain measure, and a characteristic feature of PSHA over the last 20 years has been the systematic treatment of uncertainties. The main types of uncertainties in PSHA are tabled in Fig. 1.

Aleatoric uncertainty⁵ is due to the essential randomness of seismic phenomena. It is uncertain where future earthquakes will occur (spatial uncertainty), when those earthquakes will occur (temporal uncertainty), and which level of ground motion they will produce (ground-motion uncertainty).⁶

The aleatoric variability at a site is represented by the shape of the hazard curve. A hazard curve plots PGAs (on the x -axis) and their frequencies of exceedance (on the y -axis). The characteristic shape of the hazard curve displays the fact that earthquakes that produce strong PGAs have longer return periods (bottom right of the curve) and earthquakes that produce small PGAs have shorter return periods (top left of the curve). A simplified hazard curve is depicted in Fig. 2.

Finally, aleatoric uncertainty is due to the stochastic rather than deterministic nature of seismic processes, and so, it does not decrease over time, even though, given enough time, all the values of the variables will eventually be sampled (Field 2001).

Epistemic uncertainty is due to limited knowledge of seismic phenomena (for example, incomplete dataset) and to insufficient understanding of the seismogenic processes (for example, differences between fault systems in different areas).

Epistemic uncertainty is represented by a *suite* or *bundle* (or family) of hazard curves. The spread of the family of hazard curves corresponds to how much the models diverge on their estimate of the seismic hazard at the site. For each point on the x -axis, the corresponding values on the y -axis represent the variance in the estimation of the frequency of events that produce a given ground-motion intensity level; *vice versa*, for each point on the y -axis, the values on the x -axis represent the variance in the estimation of the most intense event with a given frequency.

We can distinguish between two sorts of epistemic uncertainty. Model uncertainty, by contrast, is due to the inherent idealizations of the model (Field 2001). For example,

⁴ We are grateful to an anonymous reviewer for pointing this out.

⁵ The distinction between aleatoric and epistemic uncertainties was established in PSHA practice by the report of the US Senior Seismic Hazard Analysis Committee (SSHAC 1997).

⁶ See, e.g., Bazzurro and Luco (2005).

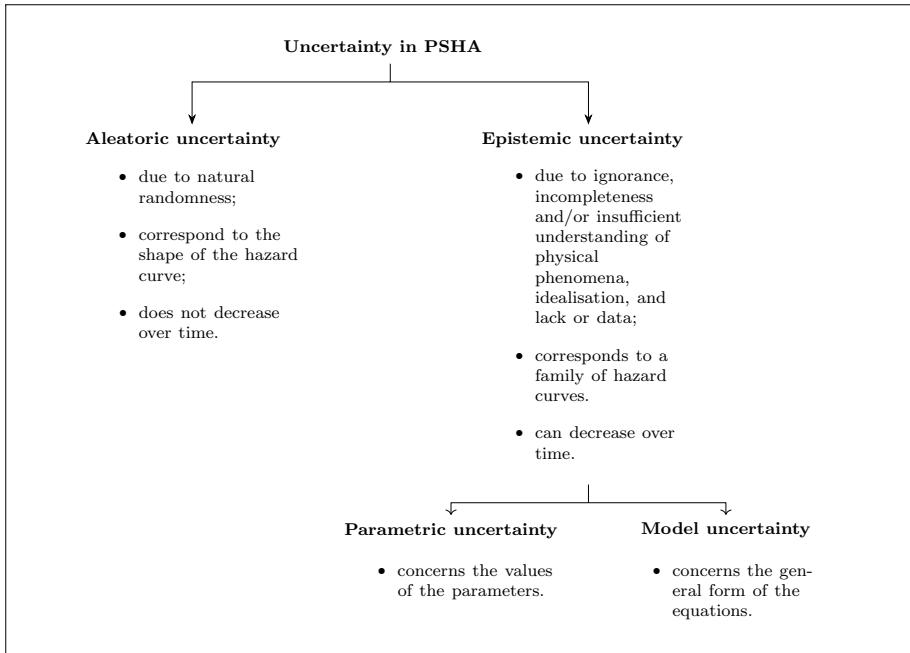


Fig. 1 A schema of different types of uncertainties in probabilistic seismic hazard analysis

multiple models of the same site may be available.⁷ Model uncertainty is due to both lack of data and incomplete understanding of the processes that generate seismic phenomena (Wang et al. 2003), given that historical catalogs are not enough for a statistical validation of seismic hazard models (Kijko 2011).

By contrast, parametric uncertainty concerns the value of the parameters of the ERM and of the GMM at the site of interest. An example is the Gutenberg–Richter equation.⁸ The *a*-value in the equation is related to the total number of earthquakes in the seismic zone, and the *b*-value corresponds to the relative ratio of small and large earthquakes, or alternatively, to the probability that a random seismic event at the source has magnitude *m*. Parametric uncertainty is due to the lack of data, given that historical catalogs are not enough to determine the exact values of the parameters.

⁷ For simplicity, we will only consider *probabilistic* models.

⁸ It is often assumed that earthquake magnitudes follow the Gutenberg–Richter distribution:

$$\log_{10} N(m) = a - bm$$

where N_M is the total number of earthquakes with magnitude *M*, and *a* and *b* are constants (“*a*-value” and “*b*-value,” respectively) that characterizes the source site.

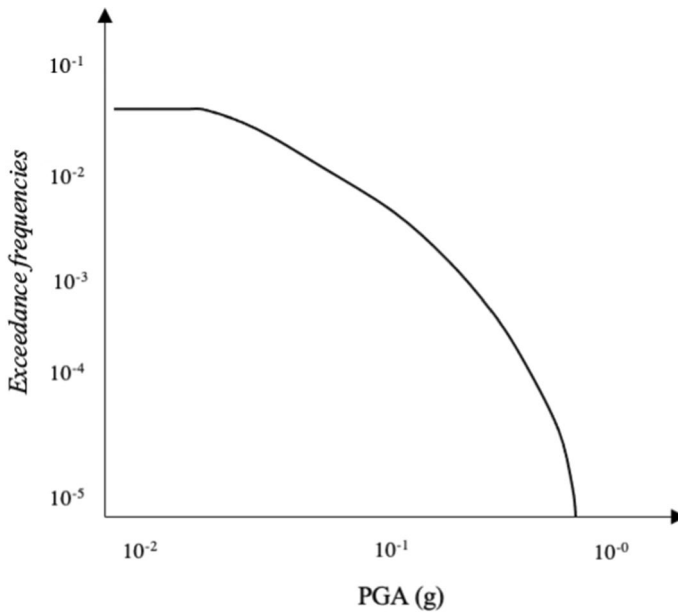


Fig. 2 A simplified hazard curve

We will now consider how these two types of epistemic uncertainties (model uncertainty and parametric uncertainty) are currently treated in PSHA.⁹

2.1 Logic trees

Logic trees (Kulkarni et al. 1984) have become the standard way of including epistemic uncertainties in PSHA.

Setting up a logic tree involves two stages. First, different ERMs and different GMMs are selected together with different values for their parameters (for example, different a -values and b -values in the Gutenberg–Richter equation). Each branch of the tree corresponds to an hazard curve.

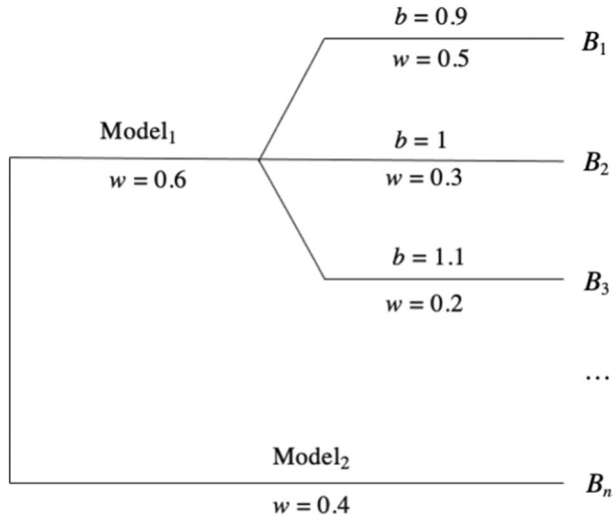
Second, a weight is assigned to each model. Weights are expressed as probabilities assigned to each model and values for its parameters; the weights of the branches departing from each node of the tree are required to sum up to one. Figure 3 depicts a simplified logic tree.

The mean seismic hazard¹⁰ outputted by a family of curves can be calculated by means of the total probability law as the sum of all exceedance probabilities estimated by each curve weighted by the weight of that curve:

⁹ For more philosophically oriented taxonomies of uncertainty, see Hansson (2022).

¹⁰ Bommer and Scherbaumb (2008) and Bommer (2012) argue in favor of the use of fractiles in the estimation of seismic hazard (for example, the 85th fractile or the 90th fractile); for a defense of the use of the mean hazard curve, see Musson (2005, 2012), and McGuire et al. (2005).

Fig. 3 A simplified logic tree; B_i = the i th branch of the logic tree; w = weight; b = model parameter



$$v(a) = \sum_n P(PGA > a | B_n)P(B_n), \tag{6}$$

where B_1, \dots, B_n each corresponds to a branch of the logic tree, $\{B_1, \dots, B_n\}$ is a partition of the probabilistic space, and $P(B_i) > 0$ for each B_i .

In order for the total probability law to be legitimately applied, the branches of the logic tree must be mutually exclusive and collectively exhaustive (MECE), that is, the models that are included in the logic tree should not overlap with each other (mutual exclusivity) and all the models should be included in the logic tree (collective exhaustivity).

As regards the first condition (mutual exclusivity), no two branches should correspond to the same model. This condition is easier to satisfy in the case of parametric uncertainty, since two branches arguably count as distinct if they correspond to different estimates of the value to the same parameter. In the case of model uncertainty, however, it can be hard to tell whether two models overlap with each other simply by looking at the form of the equations. In this case, exclusivity can be easier to satisfy if fewer and more diverse models are included. However, including less branches rather than more can undermine the exhaustivity of the tree; *vice versa*, including more branches increases the probability that branches are redundant, so that models are double-counted (Abrahamson and Bommer 2005).

The second condition (collective exhaustivity) can be interpreted in either of two ways that are not usually distinguished in the literature. A strong interpretation of collective exhaustivity is that all possible values must be included in the logic tree. A weak interpretation is, by contrast, that only values recognized by the scientific community (i.e., present in the scientific literature) are included in the logic tree.

The strong interpretation is more plausible in the case of parametric uncertainty: If it is uncertain what the true value of a parameter is, it is natural to think that the full range of possible values must be included (provided that those values are not unrealistic), not only those that have been considered in the literature. An example is the inclusion of all possible magnitudes between the minimum magnitude considered M_{min} and the maximum magnitude considered M_{max} (Bommer et al. (2004)).

The weak interpretation is more plausible in the case of model uncertainty: If it is uncertain which model is correct, it is natural to think that only the models that received attention in the scientific literature should be included. On this interpretation, a logic tree is exhaustive if it includes all available models that can plausibly be applied in the site of interest (Abrahamson and Bommer 2005; Bommer et al. 2010).

3 Models ensembles and credibility

In the face of epistemic uncertainty (lack of data, incompleteness, insufficient understanding, etc.), a complete validation of individual models may not be viable. Consequently, it may be practically impossible to determine which model is the correct one. On the opposite, many alternative models will often be available, and experts can disagree on which model is the most reliable.

However, decisions must be taken also in conditions of uncertainty. In particular, we need to estimate seismic hazard despite the insufficiency of data; these estimates must therefore take into account not only what we know (the available data), but also what we don't know (the epistemic uncertainty surrounding those data). This requires us to assess which models we should rely on among the available ones.

In this section, we compare three methodologies for the assessment of seismic hazard models. The first one is directly connected to the use of logic trees, with the weights determined by eliciting the judgments of the experts on the likelihood of the models. The second methodology selects a single model by a purely statistical procedure that allows to assess which model is more reliable for the estimation of a specific target quantity related to the seismic hazard at a given site. Finally, the third approach aims to assess the predictive power of the models using “experimental” data and can result either in the selection of single model or in the construction of an ensemble using a logic tree. We claim that in all these three approaches, epistemic and non-epistemic considerations are closely related to each other in the calculation of the appropriate estimate of seismic hazard.

3.1 The use of experts in PSHA

Probability theory is used twice in PSHA: first, to represent aleatoric uncertainty due to natural randomness (spatial uncertainty, time uncertainty, ground-motion uncertainty, etc.), and second, to represent epistemic uncertainty due to ignorance (model uncertainty and parametric uncertainty).

The weights in a logic trees are usually interpreted as degrees of confidence of the analyst that the model is correct (e.g., SSHAC 1997), the best one available (Bommer 2012), or the one that should be used (Musson 2012). For example, Scherbaum and Kuehn (2011) state that “the branch weights [are] subjective estimates for the degree-of-certainty or degree-of-belief ...that the corresponding model is the one that should be used.” In epistemology, subjective degrees of belief like the ones expressed by the experts are called *credences* (Huber and Schmidt-Petri 2008). In PSHA, these credences are elicited from the experts and the final weights are estimated by the integration of those judgments.

The systematic inclusion of experts' judgments in PSHA originated from the report of Senior Seismic Hazard Analysis Committee (SSHAC 1997). The SSHAC guidelines are

summarized by the US National Research Council Panel (NRC 1997); our presentation is based on Budnitz et al. (1998).

The SSHAC report distinguishes between the *proponents*, namely the scientist or team of scientists that formulate the models; the *evaluators*, namely the expert or group of experts that express the credibility judgments; and the *Technical Facilitator (TF) or Technical Facilitator/Integrator (TFI)*, namely the team that is in charge of the elicitation of the experts' judgments and of the production of the aggregated model.

The process consists of two main phases:¹¹ first, the *elicitation* of individual judgments or probabilities from the experts; and second, the *aggregation* of those judgments in the form of a "community distribution" (Budnitz et al. 1998). Scientists can participate in the process in two roles, both as proponents (of their own model) in the first phase and as evaluators (of others' models) in the second phase.¹²

In the second phase, the individual experts' judgments are integrated into a distribution that merges together the judgments expressed by the experts. It is important to notice that the aim of this phase is not that each expert believes in the same model or the same value for a variable or parameter of the model (Consensus 1), or that each expert believes in the same probability distribution for a random variable or model parameter (Consensus 2), but only that all experts agree that a particular composite probability distribution represents them as a group (Consensus 3), or, more weakly, and that all experts agree that distribution represents the overall scientific community (Consensus 4).

According to SSHAC (1997), consensus among experts can be achieved in five different ways.¹³ First and foremost, (1) experts might explicitly agree on a particular probability distribution. If not, then their judgments can be integrated either (2) by assigning the same weight to each judgment or (3) by assigning unequal weights. In this last case, a possible approach is (4) to assign quantitative but unequal weights (for example, if it is clear that a given expert is an outlier, but it still makes sense to assign to her judgment a numerical weight corresponding to how much that judgment is represented in the overall scientific community); a final approach would consist in (5) taking into account different judgments (for example, if the community ranks a model as more credible than another) but without assigning any definite numerical value to them. This last option is considered as the less desirable outcome.

This approach treats the judgments of the experts as evidence. This is clearly acknowledged, for example, by Hanks and Cornell (1994), who claim that "diverse expert opinion will become surrogate data," and by Budnitz et al. (1998), who claim that "although some of the key inputs to a PSHA can be determined reasonably well from observations or experiments, other key inputs require the judgment of experts." This has often been criticized on the grounds that it would make PSHA less scientific, supplying the shortage of hard data, namely historical catalogs of seismic events, with soft data, namely the estimates given by the experts (Krinitzsky 1993a, 2003; Castaños and Lomnitz 2002; Mulargia et al. 2017).

¹¹ See Coppersmith et al. (2010); SSHAC (1997) distinguishes between four phases, in which the integrator performs a literature review, evaluates the models, and estimates the community distribution (Phase 1); interacts with the proponents (Phase 2); fosters debate between proponents and experts (Phase 3); and puts together a panel of experts, elicits judgments, and infer the community distribution (Phase 4).

¹² An overview of the methods of elicitation and aggregation used in PSHA is given by Klügel (2011).

¹³ The SSHAC report considers, besides the cases mentioned below, also cases in which *unintentional* agreement/disagreement is reached among experts; for simplicity, we do not consider this distinction here.

By definition, the hazard value outputted by the logic tree is not the value estimated by the best model, which is the model that receives the highest score, but rather the sum of all the estimates of the models included in the tree weighted by the probability assigned to each branch. So it is possible that no single expert believes that the integrated model is “the one that should be used,” even though that model represents the judgements of the scientific community as a whole.

The use of experts in PSHA appears to be mainly motivated by non-epistemic values. The declared aim of the integration is to “represent the center, the body, and the range of technical interpretations that the larger informed technical community would have if they were to conduct the study” (SSHAC 1997). The fact that ensemble modeling in PSHA aims to “represent” the judgments of the experts indicates that epistemic values may also be involved: An ensemble of models may be praised because it accurately represents the uncertainty in the scientific community. However, the final goal of the aggregation is not to produce the most accurate model, but a model that best represents the judgments of the scientific community as a whole, and there may be practical reasons to rely on such model for public decisions. As stated, for instance, by Marzocchi and Zechar (2011), “one of the main goals of decision-makers is to minimize possible a posteriori critiques if a forecast model fails” (p. 446). Avoiding criticisms is easier if the decision-maker has considered all the available models and takes into consideration the judgments of all the experts.

3.2 Credibility

Grandori et al. (1998) move the following objection to the approach with the use of experts and logic trees:

This kind of procedure is much debatable. It must be noted, first, that the results obtained with n different models do not constitute a sample of a random variable of which mean value and variance could be estimated. Moreover, the dispersion of the n values depends on the subjective choice of possible models, and so does the mean value. Therefore, the procedure is formally not correct.

Starting from this criticism and from the consideration that it is not possible to statistically validate the estimates of seismic hazard produced by different models due to shortage of data, Grandori proposes a different approach to assess the models by means of a notion of credibility. He posed attention to the use of statistical tests for preferring one of two competing models in the estimation of a specific quantity (A) related to the seismic hazard at a given site (see also Lind 1996).

Referring in particular to the comparison of two earthquake rate models (F_1 and F_2) for the estimate of the peak ground acceleration with a specific return period (A), the fundamental tool for selection of one of the two is the evaluation of their foreseeable errors in estimating A , under a given hypothetical “true” distribution F^0 . It is assumed that the available catalog is a random sample S^0 drawn from F^0 and that the true value of A is obtained from the known procedure $Z: A^0 = Z(F^0)$.

The errors of F_1 under the hypothesis that the true magnitude distribution (F^0) has the same mathematical form as F_1 (i.e., F_1 is the right model) and under the alternative hypothesis that the model F_1 is not correct (F_2 is the right one) are calculated, using synthetic seismic catalogs generated from F^0 using Monte Carlo method. The same for F_2 .

The distribution of such errors is described by: (1) the mean value \hat{A}_m of 1000 independent estimations obtained from 1000 random samples S^0 drawn from F^0 with the same size

as the available catalog; (2) the standard deviation σ of the 1000 estimates of A_i ; and (3) an indicator Δ_i^0 called “credibility” of the model F_i with respect to F^0 :

$$\Delta_i^0 = P\{|\hat{A}_i - A^0| < kA^0\}$$

where \hat{A}_i is the estimator of A with the model F_i , and the parameter k defines a conventional limit.

The results of this analysis give a measure of the statistical uncertainty connected with the use of the model F_i (as right model) and of its robustness (as not correct model).

Finally, the relative credibility of one model with respect to the other is given by:

$$\Delta_{1,2}^0 = \Delta_1^0 - \Delta_2^0$$

The model 1 is more credible than the model 2 if $\Delta_{1,2}^0 > 0$, and equally or less credible otherwise. This procedure is akin to an empirical test (see 3.3), except that the data are drawn from a synthetic catalog rather than obtained directly by observation; this allows to consider samples that include strong earthquakes.

A model can have high credibility only when that same model is used as the conjectural truth but low credibility when another model is used, whereas another model can have high credibility in both cases (Grandori et al. 1998). Moreover, two models that give fairly different estimates for the same quantity A may have nonetheless similar credibility if their relative credibility indexes are similar in most cases; this suggests that two models can be “similar” besides the estimate of a particular value (Grandori et al. 2004). Moreover, it is worth noting that the credibility of a model is always relative to a specific quantity. Indeed, “when two models are proposed for the interpretation of reality, it may happen that [the first model] is more reliable than [the second model] in the estimate of PGA with a return period of 50 years, while for the estimate of PGA with a return period of 475 years [the second model] is more reliable than [the first one]” (Grandori et al. 1998).

3.3 Predictive power

Alternatively, the selection or preference for a model can be determined by the predictive power of that model in correctly estimating a specific quantity related to seismic hazard experimentally acquired (i.e., data observed in real word and not generated synthetically as proposed for credibility, see 3.2). This approach was already proposed by Esteva (1969) and has received an increasing attention in the last few years, becoming a conceptual alternative to the aggregation of experts’ judgments (Baker and Gupta 2016; Marzocchi and Jordan 2018; Secanell et al. 2018).

The approach is also called “Bayesian” (e.g., Esteva 1969; Secanell et al. 2018) because models are scored based on the conditional probability that the model is correct given the available data.¹⁴ Given an observation A consisting of recordings of the seismic activity at the site in a given interval of time, the probability that the model that corresponds to the i th branch of a logic tree is correct and is expressed by Bayes’ law:

¹⁴ Our presentation is based on Secanell et al. (2018).

$$P(B_i|A) = \frac{P(B_i) P(A|B_i)}{P(A)} \quad (7)$$

If the tree has n branches, the prior probability $P(B_i)$ can be set to:

$$P(B_i) = \frac{1}{n} \quad (8)$$

More often, $P(B_i)$ is set to the weights assigned to the model B_i by the experts that are then updated using “empirical” data. Finally, the probability of the observation A is expressed by the total probability theorem,

$$P(A) = \sum_{j=i}^n P(B_j) P(A|B_j), \quad (9)$$

as the sum of the probability of observing A if the model B_j is correct weighted by the probability that B_j is correct.

The posterior probability that a model is correct given some set of observations that can be used to score the predictive power of that model and are based purely on the performances of the models in predicting the observed phenomena (Scherbaum and Kuehn 2011). Notice that the scores represent therefore *degrees of confirmation* of the models with respect to the observations rather than to degrees of confidence of the experts.

The experimental approach can be used in two ways, either to select a single model or to construct an ensemble of models. In the first case, the model that is selected is the one that receives the highest score, and which is, therefore, the most successful in predicting the observations. In the second case, the hazard is determined as the mean value of an ensemble of models whose weights are determined by their predictive power.

Marzocchi and Zechar (2011) claim that scoring the predictive power of the models is “a scientific experiment in the traditional sense,” where the hypothesis corresponds to the forecasts of future seismic activity, and the experimental data are the yearly recordings against which those forecasts are tested.¹⁵ For this reason, they also argue that the experimental approach is more objective, since it provides an assessment of the models based on hard data rather than by integrating different estimates provided by the experts.

Unlike traditional experiments, however, (i) neither the experimenter has control over the initial conditions (i.e., the seismogenic process) nor he or she can intervene on that process during the experiment. Moreover, (ii) only a specific subclass of similar phenomena will likely be observed over a short period of time, namely small earthquakes, which are more frequent, rather than strong ones, that are rarer. It must therefore be assumed that the forecasting performances of the models with respect to those earthquakes carry over to their predictions of stronger events in order for the test to be reliable. However, given that strong earthquakes are rare, not only it is difficult to extrapolate robust probabilistic models but it is also difficult to determine whether large earthquakes have the similar distributions as small earthquakes.

Moreover, non-epistemic considerations are relevant to the experimental approach as well. In particular, practical aims and goals suggest in which situations the experimental

¹⁵ The example of the experimental approach that they provide is the Californian Regional Earthquake Likelihood Models Test (RELM), which produced an assessment of short-term earthquake forecast in the Bay area (Lee et al. 2011).

approach should be used. For example, this approach can be used to select a model that can be used to issue short-term forecasts of future seismic activity (i.e., small earthquakes) rather than when long return periods must be considered (e.g., long return periods must be considered (e.g., in the hazard assessment for nuclear power plant). Moreover, a dynamic update of the weights of the logic tree can be useful in “regions ...where large increases in the rate of earthquakes [are] triggered by anthropogenic activities” (Baker and Gupta 2016).

We can conclude that epistemic and non-epistemic considerations are therefore deeply intertwined in the assessment of seismic hazard model. Therefore, analyzing the epistemic aspects of PSHA leads to considering its non-epistemic aspects; we will now see that the converse is also true when we consider the non-epistemic aspects of PSHA, in particular, the *scientific responsibility* of the experts.

3.4 Discussion

We have considered three approaches to the assessment of seismic hazard models. These approaches differ of course from one another. The main difference is whether the approach selects a single model or builds an ensemble. The first approach aims to define an ensemble of models by integrating the judgments of the experts in a way that represents the overall epistemic uncertainty. By contrast, the second approach aims to determine which one of two competing models is the most credible one in estimating a given quantity. Finally, the third approach can be used either to select a single model, namely the model that receives the highest score with respect to some metric, or to update the weights of an ensemble of models on the basis of a posterior distribution calculated according to Bayes' Law.

As seen, the three approaches correspond to different meanings of “best model.” In the first approach, the best model is the one that reflects the whole body of judgments expressed by the experts. In the second approach, the best model is the model that has highest credibility index with respect to a certain quantity that must be estimated.¹⁶ Finally, in the third approach the best model is the one whose forecasting performances (with respect to a set of empirical data collected over a given period of time) receive the highest score.

The data that are used to assess the models (i.e., the inputs to PSHA) vary as well. The inputs to the first approach are both the historical catalogs and the judgments elicited from the experts that are used to determine the weights. The inputs to the second approach are historical catalogs alone, which are used to estimate the parameters of the true model F^0 from which synthetic catalogs are generated. Finally, the inputs to the third approach are both historical catalogs (that are used to formulate the models) and a set of independent data (that are used to assess the models).

The three approaches pursue, in a sense, different aims. Of course, the final aim is in all cases to estimate the seismic hazard at the site; however, those approaches differ on how this value is estimated. The aim of the first approach is to include epistemic uncertainties in the estimate of seismic hazard. The aim of the second approach is to be able to select on the basis of statistical tests the most suitable model. Finally, the aim of the third approach

¹⁶ Credibility explicitly targets a value that is significant in engineering (for example, a return period of 475 years), whereas both Credences and Confirmation produce an estimate of seismic hazard for any return period.

	Aim	Single model/ model ensemble	Inputs	Best model
Use of experts	incorporating epistemic uncertainty in the estimate of seismic hazard.	Ensemble	historical data; judgments of the experts	the model (ensemble) that represents the judgments of the scientific community
Credibility	selecting one model (the most credible) for the estimate of a given seismic hazard parameter.	Single model	historical data	the model with highest credibility index
Predictive power	scoring different models using new empirical data from the site.	Single model or ensemble	historical data; updated ground-motion recordings	the model (or ensemble) that receives the highest score.

Fig. 4 Synopsis of different approach for the assessment of probabilistic seismic hazard analysis

is scoring different models on the basis of empirical data that can be continuously updated. The three approaches are compared in Fig. 4.

Epistemic and non-epistemic considerations are therefore deeply intertwined in the assessment of seismic hazard models. In particular, practical aims determine which approach should be used. Each approach determines, in its turn, whether a single model or an ensemble of models should be used, how inputs are selected and interpreted, and which model is the best. Therefore, analyzing the epistemic aspects of PSHA leads to considering its non-epistemic aspects; we will now see that the converse is also true when we consider the non-epistemic aspects of PSHA, in particular, the *scientific responsibility* of the experts.

4 Scientific responsibility and the “no one’s model” problem

The philosophical literature, and especially Murphy and Gardoni (2011) and Doorn (2015), has focused on the moral and social responsibility related to seismic events.¹⁷ Given that those are random events that cannot be predicted (Geller 1997), there is arguably no responsibility for the occurrence of the event itself (namely for seismic hazards), even though that there can be responsibilities concerning the communication of scientific data (Allen 1976; Sol and Turan 2004) and the mitigation of seismic risk.

The literature in engineering has focused, by contrast, on the *intellectual responsibility of the scientist* (scientific responsibility). In the engineering literature, SSHAC (1997) provides indeed both a definition of intellectual responsibility and a criterion for attributing responsibility in PSHA.

¹⁷ van de Poel (2011) distinguishes between *active* responsibility, which consists in a proactive thinking about how a scientific result will be used in the future, and *passive* responsibility, namely backward-looking responsibility.

Intellectual responsibility is defined as “the responsibility not only for the accuracy and completeness of the results, but also for the process used to arrive at the results” (SSHAC 1997). The key term used in the report to describe scientific responsibility is “ownership.” For example, “it is absolutely necessary that there be a clear definition of ownership of the inputs into the PSHA, and hence ownership of the results of the PSHA”; this principle is important because “it assigns to an identified entity, the “owner,” clear intellectual or scientific responsibility for the conduct and results of a PSHA.” Specifically, scientific responsibility encompasses two levels of types of ownership: (1) ownership of the *input* of the PSHA and (2) ownership of its *results*. With respect to the standard approach based on the use of experts and the logic tree,

- (1) the inputs are defined in broad terms as “the composite distribution of the informed technical community” (SSHAC 1997). That includes not only the historical dataset and the models considered in the integration, but also the judgments elicited from the experts. The assumption of SSHAC (1997) is that “using the collective input of the *informed* technical community [is] the best, and most defensible, way of defining seismic hazard.”
- (2) the outputs of PSHA is the ensemble of models (aggregate distribution) and the corresponding mean hazard curve.¹⁸

The literature has discussed the different responsibilities, respectively, of the proponents of the individual models, of the expert evaluators, and of the team of integrators.

The scientists that propose the model have intellectual responsibility for it (Erto et al. 2016). The responsibility of the scientist includes both *intellectual integrity*, which requires that the scientist exercises his/her best professional judgment, does not fabricate data, etc., and also *due diligence*, which requires, for example, that the scientists “learn about the most recent advances in the field, often by direct contact with other experts,” update their credences in light of new data, etc. (SSHAC 1997). When a single model is selected, responsibility is also easy to ascertain, because it lies with the scientist that formulated the selected model.

However, scientific responsibility is neither a purely epistemic value nor a purely non-epistemic value, but rather a value that has both epistemic and non-epistemic facets. The proponents of the model are also responsible for the selection of the historical data that are considered for the formulation of the probabilistic model. These choices concern not only the inclusion of a particular historical catalog in the dataset, but also controversial issues such as the removal of foreshocks and aftershocks. The collection and elaboration of these data are among the epistemic facets of scientific responsibility, whereas scientific integrity, for example, belongs to the non-epistemic facets.

The responsibilities of the experts are briefly discussed by Klügel (2011). He lists five principles that experts should abide by. These principles are based on Cooke (1991). They concern (1) the *reproducibility* of results, (2) the *accountability* for the judgments that are expressed, (3) the *empirical control* of those judgments on the basis of the available dataset, (4) the *neutrality* of the procedure of aggregation, and (5) the *fairness* in giving each experts an equal say (Hanea et al. 2021). Applying these principles to PSHA is not easy;

¹⁸ Note that the very existence of such hazed curve may be a substantial assumption; see Anderson and Biasi (2016).

in particular, estimating the experts' performances (as proposed by Aspinall and Cooke (2011)) can be difficult in the context of seismic risk assessment due to the rarity of strong earthquakes and the shortage of historical data.

Finally, the integrator has responsibility for producing the overall distribution of weights to the branches and for the formulation of the final model. The integrators have always a choice on how to infer the "community distribution" from the judgements given by the experts, even if those judgements are expressed directly as probabilities.¹⁹ Moreover, the final estimate of seismic hazard is an aggregate result based on the amalgamation of different models and of different estimates of the values of their parameters; the shape of the final hazard curve will therefore depend not only on the models that are included in the ensemble, but also on how those models have been weighted by the experts.

This can create a problem for how responsibility is shared among the different participants in a PSHA. In SSHAC (1997), it is noticed, for example, that responsibility in PSHA is "typically diffused over a large group of experts, analysts and stakeholders in a nebulous way," and even "overly diffused" among them. The reason is that many PSHA studies²⁰ involve a very large number of participants; this makes it difficult to ascertain individual responsibilities for the final result.

This is similar to the so-called *problem of many hands* in the philosophy of technology (van de Poel et al. 2015): often, the complexities of an engineering project are such that "it is usually very difficult, if not impossible, to know who contributed to, or could have prevented a certain action, who knew or could have known what." In the case of seismic hazard analysis, in particular, the "many hands" are those of the proponents of the individual models, the evaluators of those models, and the team that performs the integration of the judgments elicited from the experts.

SSHAC (1997) solves the *problem of many hands* by attributing the overall scientific responsibility to the integrator. As stated by Budnitz et al. (1998), "final responsibility for the process of obtaining the aggregated product rests with the [integrator]." The report stresses also that the scientific responsibility of the integrator concerns both the inputs and of the results of the PSHA. In particular, the integrator selects the inputs (models and datasets), forms the panels of experts, elicits the individual judgments, infers the distribution of the judgments of the experts resulting from the elicitation, and produces the final estimate of seismic hazard. According to SSHAC (1997), the criterion allows to achieve "clear responsibility for the conduct of the study."

However, this can also lead to a *de-responsibilization* of individual scientists. In particular, the scientist who formulates a model could apply lower scientific standards to his/her work if he/she knows that the model will only contribute to the final hazard estimate as part of an ensemble, for which the scientist feels no responsibility. At the same time, the integrator could be tempted to include all available models, regardless of their accuracy, since the goal of the integrator is to represent the full spectrum of epistemic uncertainty, and he or she may feel in turn no responsibility for the accuracy of the proposed models.

¹⁹ The opinions of the experts can be elicited either as qualitative judgments about the models (how good the model is) that are then turned into probabilistic measures, or directly as probabilities; see Scherbaum and Kuehn (2011) and Runge et al. (2013).

²⁰ Some examples are the European SHARE project (Giardini et al. 2014), the National Seismic Hazard Model for New Zealand (e.g., Stirling et al. 2012) and the new Italian seismic hazard model (MPS19; see Meletti et al. 2021).

The proponents and the integrator have indeed different goals. The goal of the proponent is to produce a model that, according to his/her best judgment, is correct, given the available data; the hazard curve produced by the individual model should correspond to his/her best estimate of the “true” seismic hazard. By contrast, the goal of the integrator is to formulate the model that best represents the distribution of epistemic uncertainty in the scientific community. As highlighted in SSHAC (1997), “this does not necessarily mean that the ‘owner’ agrees with every particular input or result but that the owner feels confident that the PSHA has fulfilled the purpose of representing the larger technical community and can be defended in scientific and regulatory arenas.”

The criterion proposed by SSHAC seems therefore to conflate two different types of scientific responsibility: the responsibility for the hazard estimate itself (the final hazard curve considered as the result of PSHA) and the responsibility for the integration. Even though all scientific responsibility is attributed to the integrator, this latter cannot have full responsibility for the final model being the correct one.

With respect to the final hazard estimate, neither the individual proponent nor the integrator seem to have ownership of the final result. On the one hand, the proponents of the model can think that their best estimate is represented by the individual hazard curve and not by the ensemble of models. On the other, the integrator could take no responsibility for the final estimate, since that estimate is produced by an aggregation of different models, and the weights assigned to those models is not determined by scientists themselves. Responsibility for the scientific accuracy of the final result is still shared between the proponents, the experts, and the RFI; the final model is, in a relevant sense, *no one’s model*.

The “no one’s model problem” shows that scientific responsibility is still “overly diffused” (SSHAC 1997) even if the ownership of the final result is attributed to the integrator alone. The SSHAC criterion solves the problem of many hands at the price of a clear attribution of responsibility for the scientificity of the final model. This “no one’s model” problem has not been explored yet neither in the philosophical nor in the scientific literature. Finally, the interconnections between epistemic and non-epistemic aspects of PSHA suggest that this problem can be tackled only from a perspective that integrates philosophy and engineering.

5 Conclusions

PSHA estimates the seismic hazard at a site as the probability of exceedance of a given ground-motion intensity level in a given period of time. Such estimates are subject to two kinds of uncertainty. (i) Aleatoric uncertainty is due to the natural variability of seismic phenomena; for example, it is uncertain where future earthquakes will occur (spatial uncertainty), when they will occur (temporal uncertainty), and which level of ground motion they will produce (ground-motion uncertainty). (ii) Epistemic uncertainty, by contrast, is due to limited knowledge; in particular, it concerns the mathematical form of the seismic hazard models (model uncertainty) and the values of the model parameters (parametric uncertainty). Unlike the aleatoric uncertainty, epistemic uncertainty can decrease over time as more data are collected. However, historical data are not enough for a statistical validation of seismic hazard models.

We compared three approaches to the assessment of seismic hazard models. The first approach, which is based on the use of a logic tree, considers an ensemble of models weighted on the basis of the degree of confidence of the analyst. In the second approach,

a single model is selected by a purely statistical procedure that allows to assess the credibility of the model itself. Finally, the third approach assesses the predictive power of the models and can result either in the selection of single model or in the construction of an ensemble using a logic tree.

PSHA can create a “problem of many hands” due to the large number of people that commonly take part in a study, in particular individual model proponents, expert evaluators, and integrator team. This can make it difficult to ascertain individual responsibilities. As seen, SSHAC (1997) provides both a definition of scientific responsibility and a criterion for attributing responsibility within PSHA. Scientific responsibility is defined as “ownership” both for the inputs and the results; according to SSHAC (1997), the scientific responsibility for the final hazard curve lies with the integrator. We argued that scientific responsibility is still “overly diffused” in PSHA even if the ownership of the final result is attributed to the integrator alone; this raises a problem for the ownership of the results (“no one’s model” problem) that is yet unexplored in the literature. As we showed, epistemic considerations about PSHA are deeply intertwined with non-epistemic considerations; therefore, solving these problems required a close collaboration between geosciences and philosophy.

Author contributions All authors contributed equally to this work and read and approved the final manuscript.

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement. This research was partially funded by Next Generation EU, Piano Nazionale di Ripresa e Resilienza (PNRR), Ministry of University and Research: “RETURN. Multi-Risk Science for Resilient Communities Under a Changing Climate.”

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abrahamson NA, Bommer JJ (2005) Probability and uncertainty in seismic hazard analysis. *Earthq Spectra* 21(2):603–7
- Alamilla JL, Rodriguez JA, Vai R (2020) Unification of different approaches to probabilistic seismic hazard analysis. *Bull Seismol Soc Am* 110(6):2816–2827
- Allen CR (1976) Responsibilities in earthquake prediction. *Bull Seismol Soc Am* 66:2069–74
- Anderson J, Biasi G (2016) What is the basic assumption for probabilistic seismic hazard assessment? *Seismol Res Lett* 87:323–326
- Aspinall WP, Cooke RM (2011) Quantifying scientific uncertainty from expert judgement elicitation. In: Rougier J, Sparks S, Hill LJ (eds) *Risk and uncertainty assessment for natural hazards*. Cambridge University Press, Cambridge, pp 64–99

- Baker J, Gupta A (2016) Bayesian treatment of induced seismicity in probabilistic seismic-hazard analysis. *Bull Seismol Soc Am* 106:1–11
- Baker J, Bradley B, Stafford P (2021) *Seismic hazard and risk analysis*. Cambridge University Press, Cambridge
- Bazzurro P, Luco N (2005) Accounting for uncertainty and correlation in earthquake loss estimation. In: *Proceedings of 9th international conference on structural safety and reliability (ICOSSAR)*, Rome, Italy, June 19–23
- Bommer JJ (2012) Challenges of building logic trees for probabilistic seismic hazard analysis. *Earthq Spectra* 28(4):1723–1735
- Bommer J, Abrahamson NA (2007) Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bull Seismol Soc Am* 96:1967–1977
- Bommer JJ, Scherbaum F (2008) The use and misuse of logic trees in probabilistic seismic hazard analysis. *Earthq Spectra* 4(24):997–1009
- Bommer JJ, Abrahamson NA, Strasser FO, Pecker A, Bard P, Bungum H, Cotton F, Fäh D, Sabetta F, Scherbaum F, Studer J (2004) The challenge of defining upper bounds on earthquake ground motions. *Seismol Res Lett* 75(1):82–95
- Bommer JJ, Douglas J, Scherbaum F, Cotton F, Bungum H, Fäh D (2010) On the selection of ground-motion prediction equations for seismic hazard analysis. *Seismol Res Lett* 81(5):783–793
- Budnitz RJ, Apostolakis G, Boore DM, Cluff LS, Coppersmith KJ, Cornell CA, Morris PA (1998) Use of technical expert panels: applications to probabilistic seismic hazard analysis. *Risk Anal* 18(4):463–9
- Castaños H, Lomnitz C (2002) PSHA: Is it science? *Eng Geol* 66:315–317
- Cooke RM (1991) *Experts in uncertainty: opinion and subjective probability in science*. Clarendon Press, Oxford
- Coppersmith KJ, Bommer JJ, Kammerer AM, Ake JP (2010) Implementation guidance for SSHAC level 3 and 4 processes, PSMA 10, Seattle, WA
- Cornell AC (1968) Engineering seismic risk analysis. *Bull Seismol Soc Am* 58(5):1583–606
- Diekmann S, Peterson M (2013) The role of non-epistemic values in engineering models. *Sci Eng Ethics* 19(1):207–218
- Erto P, Giorgio M, Iervolino I (2016) About knowledge and responsibility in probabilistic seismic risk management. *Seismol Res Lett* 87(5):1161–6
- Esteva L (1969) Seismicity prediction: a Bayesian approach. In: *Proceedings of the fourth world conference on earthquake engineering*, Santiago, Chile
- Field EH (2001) Probabilistic seismic hazard analysis (PSHA): a primer (**unpublished manuscript**)
- Foulser-Piggott R, Bowman G, Hughes M (2020) A framework for understanding uncertainty in seismic risk assessment. *Risk Anal* 40(1):169–182
- Geller RJ (1997) Earthquake prediction: a critical review. *Geophys J Int* 131:425–50
- Giardini D, Woessner J, Danciu L (2014) Mapping Europe's seismic hazard. *Eos* 95(29):261–262
- Grandori G (1994) Falsification of hazard analysis. In: *Earthquake engineering—tenth world conference*, Balkema, pp 6623–6639
- Grandori G (1991) Paradigms and falsification in earthquakes engineering. *Meccanica* 26:17–21
- Grandori G (1993) A methodology for the falsification of local-seismic-hazard analysis. *Annali di Geofisica CCCV I(1):191–7*
- Grandori G, Gaugenti E, Tagliani A (1998) A proposal for comparing the reliabilities of alternative seismic hazard models. *J Seismol* 2:27–35
- Grandori G, Gaugenti E, Petrini L (2004) Earthquake catalogues and modelling strategies: a new testing procedure for the comparison between competing models. *J Seismol* 10:259–69
- Hanea AM, Nane GF, Bedford T, French S (2021) *Expert judgement in risk and decision analysis*. Springer, Dordrecht
- Hanks TC, Cornell A (1994) Probabilistic seismic hazard analysis: a beginner's guide (**unpublished manuscript**)
- Hansson SO (2022) Can uncertainty be quantified? *Perspect Sci* 30(2):210–236
- Huber F, Schmidt-Petri C (2008) *Degrees of belief*. Springer, Dordrecht
- Kijko A (2011) Seismic hazard. In: Gupta H (ed) *Encyclopedia of solid earth geophysics*. Springer, Dordrecht, pp 1107–20
- Klügel J (2011) Uncertainty analysis and expert judgment in seismic hazard analysis. *Pure Appl Geophys* 168:27–53
- Krinitzsky EL (1993a) Earthquake probability in engineering-part 1: the use and misuse of expert opinion. *Eng Geol* 33:257–288
- Krinitzsky EL (1993b) Earthquake probability in engineering-part 1: earthquake recurrence and limitations of Gutenberg–Richter *b*-values for the engineering of critical structures. *Eng Geol* 36:1–52

- Krinitzsky EL (2002) Epistematic and aleatory uncertainty: a new shtick for probabilistic seismic hazard analysis. *Environ Eng Geosci* 4(4):425–443
- Krinitzsky EL (2003) How to combine deterministic and probabilistic methods for assessing earthquake hazard. *Eng Geol* 70:157–63
- Kulkarni RB, Youngs RR, Coppersmith KJ (1984) Assessment of confidence intervals for results of seismic hazard analysis. *Proc Eighth World Conf Earthq Eng* 1:263–270
- Lind NC (1996) Validation of probabilistic models. *Civ Eng Syst* 13(3):175–183
- Marulanda MC, de la Llera JC, Bernal GA, Cardona OD (2021) Epistemic uncertainty in probabilistic estimates of seismic risk resulting from multiple hazard models. *Nat Hazards* 108:3203–3227
- Marzocchi W, Jordan TH (2018) Experimental concepts for testing probabilistic earthquake forecasting and seismic hazard models. *Geophys J Int* 215:780–798
- Marzocchi W, Zechar JD (2011) Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model. *Seismol Res Lett* 82(3):442–8
- McGuire RK (2007) Probabilistic seismic hazard analysis: early history. *Earthq Eng Struct Dyn* 37(3):329–338
- McGuire RK, Cornell CA, Toro GR (2005) The case for using mean seismic hazard. *Earthq Spectra* 21(3):879–886
- Meletti C, Marzocchi W, D'Amico V, Lanzano G, Luzi L, Martinelli F, Pace B, Rovida A, Taroni M, Visini F, The MPS19 Working Group (2021) The new Italian seismic hazard model (MPS19). *Ann Geophys.* <https://doi.org/10.4401/ag-8579>
- Mulargia F, Stark PB, Geller RJ (2017) Why is probabilistic seismic hazard analysis (PSHA) still used?. *Phys Earth Planet Inter* 264:63–75
- Murphy C, Gardoni P (2011) Evaluating the source of the risks associated with natural events. *Res Publica* 17:125–40
- Musson R (2005) Against fractiles. *Earthq Spectra* 21(3):887–891
- Musson R (2012) On the nature of logic trees in probabilistic seismic hazard assessment. *Earthq Spectra* 28:1291–1296
- NCR - National Research Council Panel on Seismic Hazard Evaluation (1997) Review of recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts. National Academy of Science, Washington, DC
- Neelke D (2015) The blind spot in risk ethics: managing natural hazards. *Risk Anal* 35(3):354–360
- Runge A, Scherbaum F, Curtis A, Riggelsen C (2013) An interactive tool for the elicitation of subjective probabilities in probabilistic seismic-hazard analysis. *Bull Seismol Soc Am* 103:2862–2874
- Scherbaum F, Kuehn NM (2011) Logic tree branch weights and probabilities: summing up to one is not enough. *Earthq Spectra* 27:1237–1251
- Secanell R, Martin C, Viallet E, Senfaute G (2018) A Bayesian methodology to update the probabilistic seismic hazard assessment. *Bull Earthq Eng* 16:2513–2527
- Sol A, Turan H (2004) The ethics of earthquake prediction. *Sci Eng Ethics* 10:655–66
- SSHAC—Senior Seismic Hazard Analysis Committee (1997) Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts. Report NUREG-CR-6372. U.S. Nuclear Regulatory Commission, Washington, DC
- Stirling M, McVerry G, Gerstenberger M, Litchfield N, Van Dissen R, Berryman K, Barnes P, Wallace L, Villamor P, Langridge R, Lamarche G, Nodder S, Reyners M, Bradley B, Rhoades D, Smith W, Nicol A, Pettinga J, Clark K, Jacobs K (2012) National seismic hazard model for New Zealand: 2010 update. *Bull Seismol Soc Am* 102(4):1514–1542
- van de Poel I (2011) *Ethics, technology, and engineering: an introduction*. Wiley-Blackwell, London
- van de Poel I, Royakkers L, Zwart SD (2015) *Moral responsibility and the problem of many hands*. Routledge, London
- Wang Z, Woolery EW, Shi B, Kiefer JD (2003) Communicating with uncertainty: a critical issue with probabilistic seismic hazard analysis. *Eos* 84(46):501, 506–508
- Ya-Ting L, Turcotte Donald L, Holliday James R, Sachs Michael K, Rundle John B, Chien-Chih C, Tiampo Kristy F (2011) Results of the regional earthquake likelihood models (RELM) test of earthquake forecasts in California. *PNAS* 108(40):16533–16538