

# Comparison of epistemic uncertainty captured in probabilistic seismic hazard assessments for critical infrastructure

Guillermo Aldama-Bustos \*<sup>1</sup>, John Douglas <sup>2</sup>

<sup>1</sup>Jacobs, Cottons Centre, Cottons Lane, London SE1 2QG, UK, <sup>2</sup>Frazer-Nash Consultancy, 180 West George Street, Glasgow, G2 2NR, UK

**Author contributions:** *Conceptualization:* John Douglas, Guillermo Aldama-Bustos. *Data curation:* John Douglas, Guillermo Aldama-Bustos. *Formal Analysis:* Guillermo Aldama-Bustos, John Douglas. *Visualization:* Guillermo Aldama-Bustos, John Douglas. *Writing – Original draft:* Guillermo Aldama-Bustos, John Douglas. *Writing – Revised draft:* Guillermo Aldama-Bustos, John Douglas. *Writing – Revised 2nd draft:* Guillermo Aldama-Bustos, John Douglas.

**Abstract** Site-specific probabilistic seismic hazard assessments (PSHAs) are fundamental for designing and evaluating critical infrastructure, e.g. nuclear power plants, but involve significant epistemic uncertainties. This study compares epistemic uncertainties from thirty site-specific PSHAs, half following the Senior Seismic Hazard Analysis Committee (SSHAC) guidelines. Uncertainties were estimated for three intensity measures (IMs) at four annual frequencies of exceedances (AFEs), and for some sites, at bedrock and target horizons. Four uncertainty metrics were used: mean-to-median ratio (R); log-ratio of the 84<sup>th</sup> and 16<sup>th</sup> fractiles (Unc); standard deviation of the fractiles assuming a lognormal distribution ( $\sigma_{\text{haz}}$ ); and standard deviation of a R-based relationship ( $\sigma_{\mu}$ ). Results show strong correlations amongst Unc,  $\sigma_{\text{haz}}$  and  $\sigma_{\mu}$ . R correlates well with  $\sigma_{\mu}$  but only moderately with the others, likely due to fractile asymmetry. Unexpectedly, uncertainties are constant across AFEs but vary across IMs. SSHAC level has limited impact on uncertainty levels, though non-SSHAC studies show greater dispersion. As expected, uncertainties in stable regions are generally higher than in active regions. Results provide a benchmark for assessing the appropriateness of uncertainty levels in PSHAs; however, the risk of ‘anchoring’ to previous studies must be considered. Routine estimation and documentation of total epistemic uncertainty are recommended, especially for critical infrastructure.

Production Editor:  
Gareth Funning  
Handling Editor:  
Pablo Heresi  
Copy & Layout Editor:  
Théa Ragon

Received:  
August 9, 2025  
Accepted:  
April 30, 2026  
Published:  
May 26, 2026

**Non-technical summary** We study how uncertainty is measured and understood in assessments of how often strong earthquake shaking will occur at a location. These assessments help guide the design of important structures like nuclear power plants. Uncertainty in such assessments comes from gaps in our knowledge—what is called epistemic uncertainty—which can make it harder to decide how strong an earthquake a structure should be designed to withstand. We review thirty such assessments from different parts of the world. Our analysis focuses on uncertainty for measures of ground shaking intensity, for different frequencies of earthquake shaking, from rare to extremely rare, and in some cases at both the ground surface and the underlying rock. We use four different metrics to compare the level of uncertainty captured in each study. Our findings show that most of these metrics are closely related, and in many cases, the amount of uncertainty is similar for different earthquake frequencies but that it varies with the measure of ground shaking intensity. Following industry-standard guidelines does not always result in more consistent uncertainty. We recommend that uncertainty be calculated and reported in all assessments, especially for projects where safety is critical, and suggest using these observations as guidance on typical levels of uncertainty.

## 1 Introduction

Probabilistic seismic hazard assessments (PSHAs) are conducted to estimate the earthquake ground motions that may occur at a given location (e.g. a site of critical infrastructure) over a specified period (e.g. the infrastructure’s design life). Because of the infrequency of large earthquakes in a particular region and limited observational data (e.g. strong-motion records from the local area and shear-wave velocity measurements at the site), PSHA results are commonly associated with significant epistemic uncertainties. The design and assessment of the infrastructure need to account for these uncertainties, e.g., by using the mean, or the 84<sup>th</sup> frac-

tile, hazard curve. Larger uncertainties generally lead to larger assessed ground motions, leading to a more expensive design, a more complicated safety case or the acceptance of a higher level of seismic risk. Consequently, it is important that uncertainties are appropriately captured and that they are neither overestimated, leading to over-design, or under-estimated, leading to insufficient safety margins.

The purpose of this article is to assess the level of epistemic uncertainty captured in thirty site-specific PSHAs, and to explore the correlations amongst the uncertainty levels and characteristics of the PSHAs. A review of previous studies with similar scopes is provided next along with the novelties of the present work. This is followed by a summary of the database of PSHA re-

\*Corresponding author: guillermo.aldama-bustos@jacobs.com

sults that was compiled for this article and the results of analysing this database. Various uncertainty metrics have been proposed previously and hence exploring correlations amongst these metrics allows identification of the most appropriate metric(s). Some metrics are more complicated to compute, and hence quantifying correlations allows simpler metrics to be used to obtain a good understanding of the epistemic uncertainty captured in a particular PSHA. The article ends with a discussion of the findings, a summary of the conclusions reached, and gives some recommendations.

## 2 Previous Studies

Interest in correctly capturing epistemic uncertainty within PSHAs has grown rapidly over the past decades. Kulkarni et al. (1984) pioneered the use of logic trees to allow uncertainties to be transparently captured within hazard calculations. This approach has been almost universally adopted and guidelines for their use have been developed (e.g. Bommer and Scherbaum, 2008; Scherbaum and Kuehn, 2011). It should be noted that the distinction between epistemic uncertainty (due to lack of knowledge) and aleatory variability (due to randomness) is not strict and depends on the level of simplification in the components of the hazard model (e.g. Liou and Abrahamson, 2025). This study focuses on the epistemic uncertainty that is manifest in the hazard results.

The PSHAs conducted for nuclear power plants (NPPs) in the central and eastern USA by the Electric Power Research Institute (EPRI) and the Lawrence Livermore National Laboratory (LLNL) in the late 1980s and early 1990s highlighted that teams conducting PSHAs for a particular site can have widely differing views on the level of epistemic uncertainty that is appropriate. Cornell (1992) analysed these two studies and found that, although the median hazard curves for a particular site were broadly similar, the wider hazard fractiles (i.e. higher uncertainty) in LLNL study resulted in significantly higher mean hazard estimates. Based on his findings, Cornell (1992) proposed an approach for estimating the mean hazard in studies with poorly constrained uncertainty levels. To do this the median hazard curve is multiplied by an estimate of the mean-to-median ratio derived from studies with better constrained levels of uncertainty.

In response to the findings of Cornell (1992), the Senior Seismic Hazard Analysis Committee (SSHAC) developed guidelines for PSHAs (Budnitz et al., 1997). These guidelines provide a formal approach for capturing uncertainties within the inputs to the PSHA. They define four SSHAC levels, ranging from Level 1 (a simple PSHA conducted by a small team) to Level 4 (a complex PSHA conducted by many interacting teams of experts). Updates to this approach have been published in U.S. Nuclear Regulatory Commission (2018) and the SSHAC guidelines have been followed in dozens of projects since the late 1990s.

It is not yet standard practice to explicitly quantify the level of epistemic uncertainty captured in the outputs of PSHAs (e.g. hazard curves). Nevertheless, in the past

two decades, several studies have provided insights on this topic and various metrics have been proposed for this purpose. These metrics are based on various estimates of intensity measures (IMs) for particular annual frequencies of exceedance (AFE) or their reciprocal (return periods,  $T_r$ ) reported in PSHAs.

Douglas et al. (2014) computed:  $100 \log (PGA_{84}/PGA_{16})$  and  $100 \log [SA(1.0s)_{84}/SA(1.0s)_{16}]$ , where the subscript refers to the hazard curve for a particular fractile (either 16<sup>th</sup> or 84<sup>th</sup>), PGA is the peak ground acceleration, and SA(1.0s) is the spectral acceleration at a period of 1.0s for a given AFE (either 1/475 or 1/2,475 in that study). They demonstrated that PSHAs conducted at a higher SSHAC level (Level 3 or 4) had high and consistent values of this metric (around 40), indicating wider fractiles, and that some regional and national PSHAs had relatively low values (around 20), indicating narrow fractiles.

By rearranging an equation of Toro (2006)<sup>1</sup>, Douglas (2018, his Table 2) estimated  $\sigma_\mu$ , which represents the epistemic uncertainty in the ground motion (in terms of natural logarithms) if all the uncertainty in PSHA results arises from the ground-motion model. This equation relies on the mean-to-median ratio (R) of the IM of interest for a given AFE and the slope of the hazard curve in log-log space at that AFE. Douglas (2018) showed that there was a close correlation between the metric used by Douglas et al. (2014) and  $\sigma_\mu$ , with  $\sigma_\mu$  values of about 0.5 for SSHAC Level 3 or 4 studies and values of about 0.2 for regional PSHAs (again for AFEs of 1/475 and 1/2475). Aldama-Bustos et al. (2023) used  $\sigma_\mu$  to assess the epistemic uncertainty captured at the reference velocity horizon in the final hazard calculations for the Sizewell C site in the UK. They provide plots of  $\sigma_\mu$  against response spectral period and AFE, revealing considerable dispersion (roughly 0.2 to 0.6), particularly with respect to spectral period, largely attributed to the ground-motion model used.

Recently, Douglas et al. (2024, their Figure 2) compared the outputs of five studies for a single Swiss site (Beznau), for AFEs of 1/475 and 1/2,475, using the metrics considered in Douglas et al. (2014), Douglas (2018) and a third metric,  $\sigma_{\text{haz}}$ . The latter represents the standard deviation of a lognormal distribution fitted to the hazard fractiles (5<sup>th</sup>, 16<sup>th</sup>, 50<sup>th</sup>, 84<sup>th</sup> and 95<sup>th</sup>)<sup>2</sup>. Douglas et al. (2024) found a close correlation amongst the measures with values of  $\sigma_{\text{haz}}$  of 0.35-0.57 for the two SSHAC Level 4 studies. All three metrics for the two national studies and one regional study for this site are comparable to those from the SSHAC Level 4 studies.

Returning to the approach of Cornell (1992), Valentini and Beltran (2024) examine R for a given AFE for 33 SSHAC Level 3 studies. They only examined the AFEs of  $10^{-4}$ ,  $10^{-5}$  and  $10^{-6}$  as their interest is in the AFEs of main concern for nuclear-related facilities. They found

<sup>1</sup>This equation was previously given in Electric Power Research Institute (2004) and appears to be based on an equation in Bender (1984).

<sup>2</sup> $\sigma_{\text{haz}}$  is computed by calculating the number of standard deviations from the mean (i.e. the standard, or Z, score) for each of the fractiles assuming a normal distribution (e.g. 5<sup>th</sup> fractile corresponds to a Z score of -1.64) and then undertaking a linear regression between these Z scores and natural logarithms of the IM at each of the available fractiles;  $\sigma_{\text{haz}}$  equals the slope of this line.

reasonable consistency in this ratio amongst the studies they examined and a variation with spectral period, with higher values at shorter periods ( $\sim 0.1$ s), likely due to uncertainty in  $\kappa$  (a measure of high-frequency attenuation, [Anderson and Hough, 1984](#)), and longer periods ( $\sim 1$ s), likely due to uncertainty in near-surface site amplification.

The aim of the current article is to extend the findings from previous studies by a deeper examination of a large and more diverse database. This database includes both SSHAC-compliant PSHAs, performed at various levels (Levels 1 to 4), and non-SSHAC compliant studies. The analysis also considers four AFEs, three IMs and sites with data for alternative reference horizons (bedrock and surface). Additionally, uncertainty levels were assessed using the four metrics mentioned above. The breadth and heterogeneity of the database, combined with the use of different metrics, allow observations to be made on how uncertainty varies with SSHAC level, AFE, IM, tectonic classification (stable v active) and date of publication.

### 3 Database

We chose PSHAs for this study using the following criteria. Firstly, the results of the study needed to be publicly available, either in a journal/conference article or within a technical report, in a form that could be readily used. Many results were presented in the form of hazard curves or spectra that we digitised and interpolated (if required), which leads to negligible inaccuracies in the computed metrics. The remaining results were taken from tables and hence these are not affected by digitisation inaccuracies.

Secondly, only studies that were specifically developed for the design or safety assessment of engineering infrastructure are included. In all but one case (for an offshore oil platform) this infrastructure is a NPP or another nuclear-related facility. This criterion excludes PSHAs that were conducted for research purposes, regional hazard mapping or for national building code development. This criterion implies that we only included site-specific studies, which again excludes national or regional studies, although some studies present results for a handful of locations (e.g. PEGASOS, which presents results for four NPP sites in Switzerland).

Thirdly, we only included PSHAs for which at least the mean and two or more fractiles were reported. Generally, the fractiles reported in PSHAs are some combination of the: 5<sup>th</sup>, 16<sup>th</sup> (alternatively 15<sup>th</sup>), 50<sup>th</sup> (median), 84<sup>th</sup> (alternatively 85<sup>th</sup>), and 95<sup>th</sup> fractiles. This requirement allows the four uncertainty metrics considered here to be computed.

Fourthly, only studies presenting results for AFEs between 1/475 and 1/100,000 were selected. As one of the uncertainty metrics uses the slope of the hazard curve as an input variable, only studies for which this slope could be estimated were included (this excludes studies that only present results for a single AFE from consideration).

Finally, only studies presenting results for at least one of the considered IMs: PGA (or SA at a sufficiently

high frequency to be assumed to correspond to PGA), SA(0.2s) and SA(1.0s). Some studies only provide results for a single IM (often PGA) or for periods other than 0.2s or 1.0s, meaning that the number of studies included on each figure in the following sections varies with IM.

Overall, thirty PSHAs are included in the database. Most studies date from the last 15 years but there are a handful of PSHAs from the early-to-mid 1990s (pre-SSHAC). The geographical coverage of the PSHAs is relatively broad with studies from Europe (the UK, Slovenia, Finland and Switzerland), eastern and western USA, eastern Canada, Brazil, South Africa, and Japan. While compiling the database we noted the SSHAC level at which the study was conducted. SSHAC levels were only assigned when it is explicitly stated in the PSHAs documentation that the SSHAC guidelines were followed. The distribution with respect to SSHAC level is: None (16 studies), Level 1 (1 study), Level 2 (3 studies), Level 3 (7 studies) and Level 4 (3 studies). Note that one of the studies in the database ([Stamatikos et al., 2024](#)) is classified as a SSHAC Enhanced Level 2 but considered here as SSHAC Level 2 for convenience.

For 24 PSHAs only the rock results are included (either because no results for other ground conditions are available or because fractiles for other ground conditions were not available). Nevertheless, for six studies results for two reference horizons are included, which allows an investigation of site conditions on uncertainties. Table 1 summarizes the PSHAs included. Figure 1 indicates the locations of the sites on a global seismic hazard map, which shows that most sites are in areas of low-to-moderate hazard but that there are a handful of sites from high hazard areas. The sites are classified as either within an active shallow crustal region (ASCR) (8 sites) or a stable continental region (SCR) (12 sites) using the maps of [Schulte and Mooney \(2005\)](#). Three sites close to the boundary between ASCR and SCR are indicated in Table 1.

Using the classification of [Liou and Abrahamson \(2025\)](#), studies before about 2010 would have relied on the ergodic assumption whereas more recent studies would have generally used the partial non-ergodic approach. To our knowledge, fully non-ergodic models are not yet commonly used in practice. Although there have been a few recent applications in nuclear-related projects ([IAEA, 2025](#)), their use remains primarily confined to the research domain. Providing more details on the 30 hazard models, or classifying them according to the approach used for developing their ground-motion model, is beyond the scope of this study and would require more information on the study than is often in the public domain.

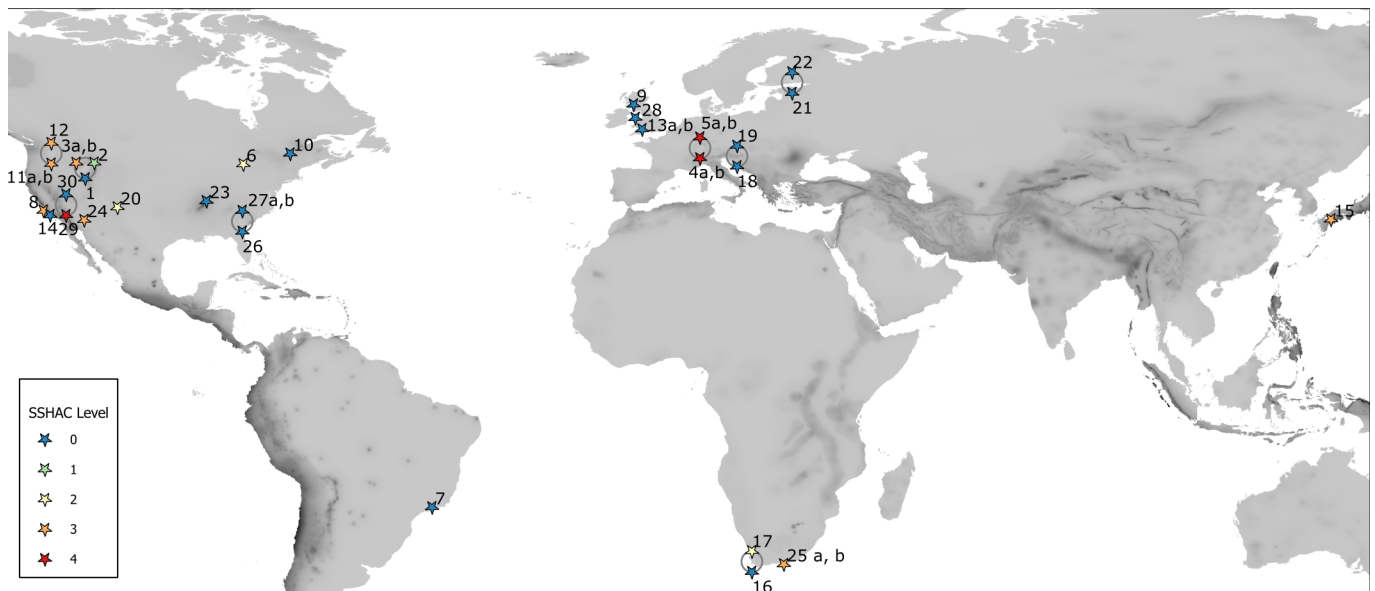
### 4 Results

As a reminder, the four metrics considered in this article are:

- $\sigma_{\text{haz}}$ , the standard deviation of the lognormal distribution fit to the hazard results for the available fractiles (at least 16<sup>th</sup>, 50<sup>th</sup> and 84<sup>th</sup>, but often also 5<sup>th</sup> and 95<sup>th</sup>);

No.	Site	Reference	SSHAC Level	Site class or VS30 (m/s)	ASCR or SCR
1	Argonne National Laboratory West/ Idaho National Laboratory (Materials and Fuels Complex - MFC), USA	INL (1996)	0	Rock (1,030m/s) Surface (607m/s, INL, 2022)	Near boundary
2		INL (2016)	1		
3 a, b		INL (2022)	3		
4 a, b	Beznau (Nuclear Power Plant), Switzerland	NAGRA (2004)	4	Rock [2,000m/s for NAGRA (2004) and 1,800m/s for Swissnuclear (2013)]. Soil [unknown for NAGRA (2004) and 516m/s for Swissnuclear (2013)]	Near boundary
5 a, b		Swissnuclear (2013)	4		
6	Bruce (Deep Geologic Repository), Canada	AMEC-Geomatrix (2011)	2	Hard rock (2,830m/s)	SCR
7	Angra dos Reis (Central Nuclear Almirante Álvaro Alberto), Brazil	de Almeida et al. (2019)	0	Hard rock (2,800m/s)	SCR
8	Diablo Canyon (Nuclear Power Plant), USA	Pacific Electric and Gas Company (2015)	3	Rock (760m/s)	ASCR
9	Faslane (HMNB Clyde), UK	Nielsen (2023)	0	Hard rock (2,100m/s)	SCR
10	Gentilly 2 (Nuclear Power Plant), Canada	Weston Geophysical Corpo- ration (1996)	0	Hard rock (3,000m/s)	SCR
11 a, b	Hanford (Columbia Generating Station), USA	Energy Northwest (2015)	3	Hard rock (3,000m/s) Soil (361m/s)	ASCR
12		PNNL (2014)	3	Hard rock (3,000m/s)	
13 a, b	Hinkley Point C (Nuclear Power Plant), UK	Tromans et al. (2019)	0	Rock (1,100m/s) Soft rock (802m/s)	SCR
14	Houchin Platform, USA	Smith and Murray (1999)	0	Stiff soil	ASCR
15	Ikata (Nuclear Power Plant), Japan	Fujiwara et al. (2024)	3	Hard rock (2,600m/s)	ASCR
16	Koeberg (Nuclear Power Plant), South Africa	Stamatakos et al. (2022)	0	Rock (733m/s)	SCR
17		Stamatakos et al. (2024)	2		
18	Krsko (Nuclear Power Plant), Slovenia	Lapajne and Fajfar (1995)	0	Deep alluvium	ASCR
19		Ursic et al. (2005)	0		
20	Los Alamos National Laboratory (Chemistry and Metallurgy Research Replacement Facility), USA	Wong et al. (2007)	2	Soil (280m/s)	Near boundary
21	Loviisa (Nuclear Power Plant), Finland	Varpasuo (2008)	0	Hard rock (2,800m/s)	SCR
22		Visakorpi (2022)	0		
23	Paducah (Gaseous Diffusion Plant), USA	Toro and McGuire (1991)	0	Rock	SCR
24	Palo Verde (Nuclear Power Plant), USA	Arizona Public Service Company (2015)	3	Soil (335m/s)	ASCR
25 a, b	Thyspunt (Nuclear Power Plant), South Africa	Bommer et al. (2015)	3	Hard rock (3,000m/s) Rock (1,680m/s)	SCR
26	Vogtle (Nuclear Power Plant), USA	U.S. Nuclear Regulatory Commission (1994)	0	Rock	SCR
27 a, b		Southern Nuclear Operat- ing Company (2014)	0	Hard rock (2,800m/s) Soil (300m/s)	
28	Wylfa B (Nuclear Power Plant), UK	Villani et al. (2020)	0	Hard rock (3,000m/s)	SCR
29	Yucca Mountain (Nuclear Waste Repository), USA	Wong et al. (1996)	0	Hard rock (1,900m/s)	ASCR
30		Stepp et al. (2001)	4		

**Table 1** PSHAs included in this study. Some references cited above provide summaries of confidential reports dating from several years before. The classification into ASCR or SCR is based on the maps of Schulte and Mooney (2005).



**Figure 1** Location of the sites included in the database. Note: When a single study is available for a given site, this is denoted by a star, when there are multiple studies for a single site the site location corresponds to the centroid of the circle. Project numbers correspond to project number in Table 1.

- $Unc = 100 \log [IM_{84}/IM_{16}]$ , where IM is the intensity measure of interest (PGA or SA) and the subscript indicates the fractile (for a few studies the 85<sup>th</sup> and 15<sup>th</sup> fractiles are used instead; this has a negligible impact on the results);
- $R = IM_{mean}/IM_{50}$ ; and
- $\sigma_{\mu} = [2 \ln(R)/k]^{1/2}$ , where k is the slope of the hazard curve in log-log space at the AFE of interest

These metrics are computed for the four AFEs considered (i.e.  $2.1 \times 10^{-3}$ ,  $4 \times 10^{-4}$ ,  $10^{-4}$  and  $10^{-5}$ ) and the three IMs [PGA, SA(0.2s) and SA(1.0s)]. Not all studies provide results for all these AFEs or all these IMs. No interpolation/extrapolation is attempted to complete the database for these gaps.

An additional metric that could have been included is the fractile corresponding to the mean ground motion, i.e. X in the equation  $IM_X = IM_{mean}$  (e.g., Adams and Halchuk, 2019). This was not considered in this study as it has not been widely proposed to compare uncertainties captured in PSHAs and would require reporting multiple fractiles to compute accurately.

#### 4.1 Correlations Amongst Metrics

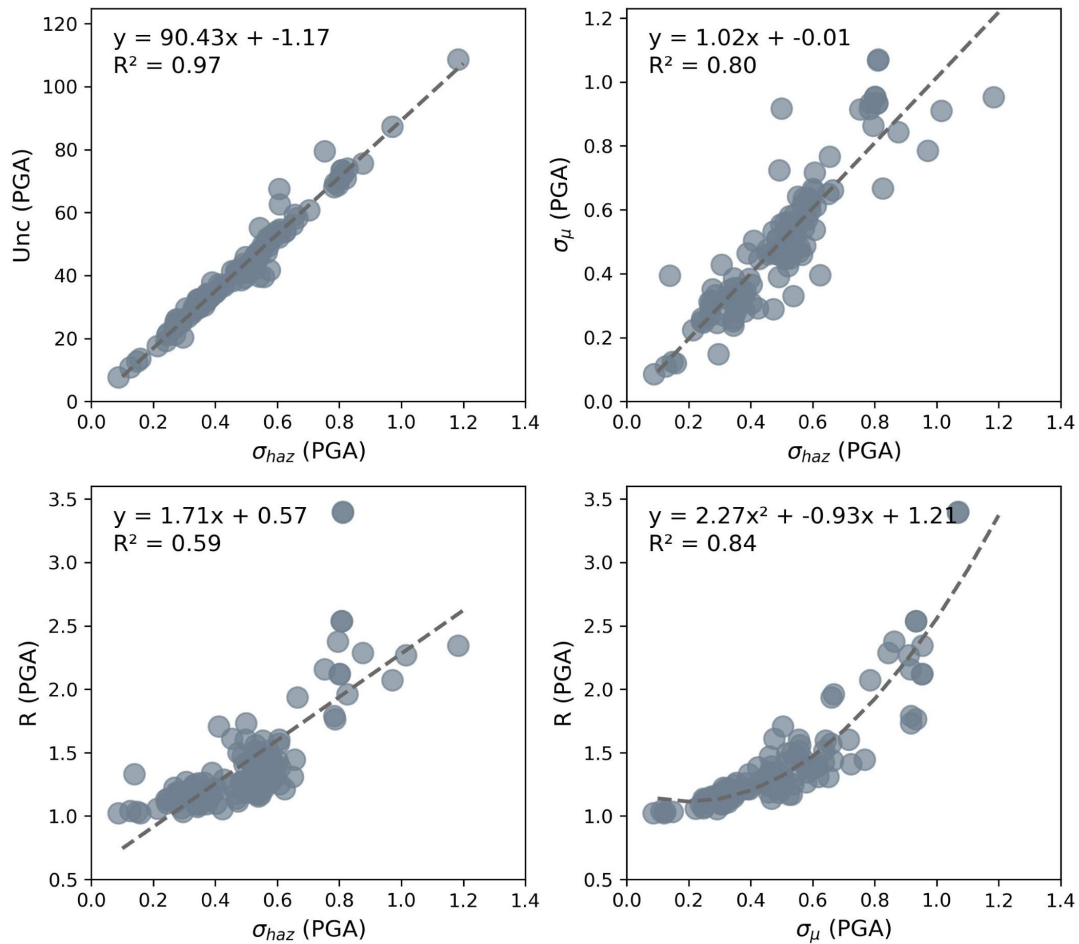
Figure 2 shows that Unc,  $\sigma_{\mu}$  and  $\sigma_{haz}$  for PGA are strongly positively correlated. Similar correlations are observed for SA (0.2s) and SA (1.0s). This high correlation is expected given the definitions of the metrics.  $\sigma_{haz}$  and Unc exhibit an almost perfect correlation (coefficient of determination = 0.97), this is because hazard fractiles generally follow a lognormal distribution; and in this case  $Unc = 86.9 \sigma_{haz}$ . By contrast, R and  $\sigma_{haz}$  show only a moderate correlation, suggesting that R may be a poor metric for quantifying uncertainty as, mathematically,  $\sigma_{haz}$  is the most direct metric for measuring uncertainty in PSHA results. Since  $\sigma_{haz}$  and Unc are almost

perfectly correlated, the correlation between R and Unc will also be weak. R and  $\sigma_{\mu}$  are highly correlated, primarily because the slope of the hazard curve used to calculate  $\sigma_{\mu}$  varies within a tight range of 1.5 to 3.5 for the AFEs of interest here. As a result,  $\sigma_{\mu}$  is mainly controlled by R (also see Table 1 of Douglas, 2018). From the definition of  $\sigma_{\mu}$ , there is a quadratic dependency between  $\ln(R)$  and  $\sigma_{\mu}$ , which for a small range of R implies a quadratic dependency between R and  $\sigma_{\mu}$  (see bottom right plot of Figure 2).

Due to the high correlations amongst the uncertainty metrics, only a single uncertainty measure,  $\sigma_{haz}$ , is used in the subsequent figures to investigate how uncertainty varies with characteristics of each study. These characteristics are: SSHAC level, AFE, tectonic classification (ASCR v SCR) and year of publication. Figures were generated for all three IMs considered in this study. Where data are available, results for different ground conditions are also presented.  $\sigma_{haz}$  was selected as the representative uncertainty metric because it can be estimated for all PSHAs that present two or more fractiles, regardless of which fractiles they are available, and since it has a clear mathematical meaning. Results for other metrics are presented where relevant, particularly to demonstrate their strengths or limitations.

#### 4.2 Comparisons by SSHAC Level

Figure 3 shows  $\sigma_{haz}$  values, grouped by SSHAC levels, for return periods (Tr) of 475, 2,475, 10,000 and 100,000 years (i.e.  $2.1 \times 10^{-3}$ ,  $4 \times 10^{-4}$ ,  $10^{-4}$  and  $10^{-5}$  AFE) and the IMs of interest. For convenience, non-SSHAC studies are shown as SSHAC Level 0. The mean (indicated by the red dot) and standard deviation (represented by the red bar) were calculated only for the SSHAC levels with enough data to support meaningful statistical estimates. Note that a small amount of jitter has been applied on the x-axis in figures presented here and following sec-



**Figure 2** Correlation plots for PGA amongst the uncertainty metrics considered in this study.

tions. This is done to improve visualisation of overlapping data points. As a result, a data point, for example, for the 475-yr Tr, may appear slightly offset along the x-axis.

Overall, there is no clear pattern across SSHAC levels (including non-SSHAC studies). Nevertheless, when comparing SSHAC Level 0 and 3 studies specifically, SSHAC Level 3 studies tend to exhibit a slightly higher mean  $\sigma_{\text{haz}}$  values but lower standard deviations, particularly at SA(1.0s). This trend seems to imply a more consistent and potentially better-constrained characterisation of uncertainty in SSHAC Level 3 studies compared with SSHAC Level 0 studies, although the difference is marginal in most cases. The reduced variability observed in SSHAC Level 3 studies, particularly at SA(1.0s), may be attributed to methodological choices, such as the use of backbone ground motion models, which are known to produce less dispersion at long periods than multi-GMPE models (e.g., Aldama-Bustos et al., 2023). This is normally the result of multi-GMPE logic trees including models for different tectonic regions (e.g. SCR and ASCR) along with other potential factors such as the need to extrapolate to periods for which coefficients are not provided and the consideration of models that are not particularly well constrained at long periods.

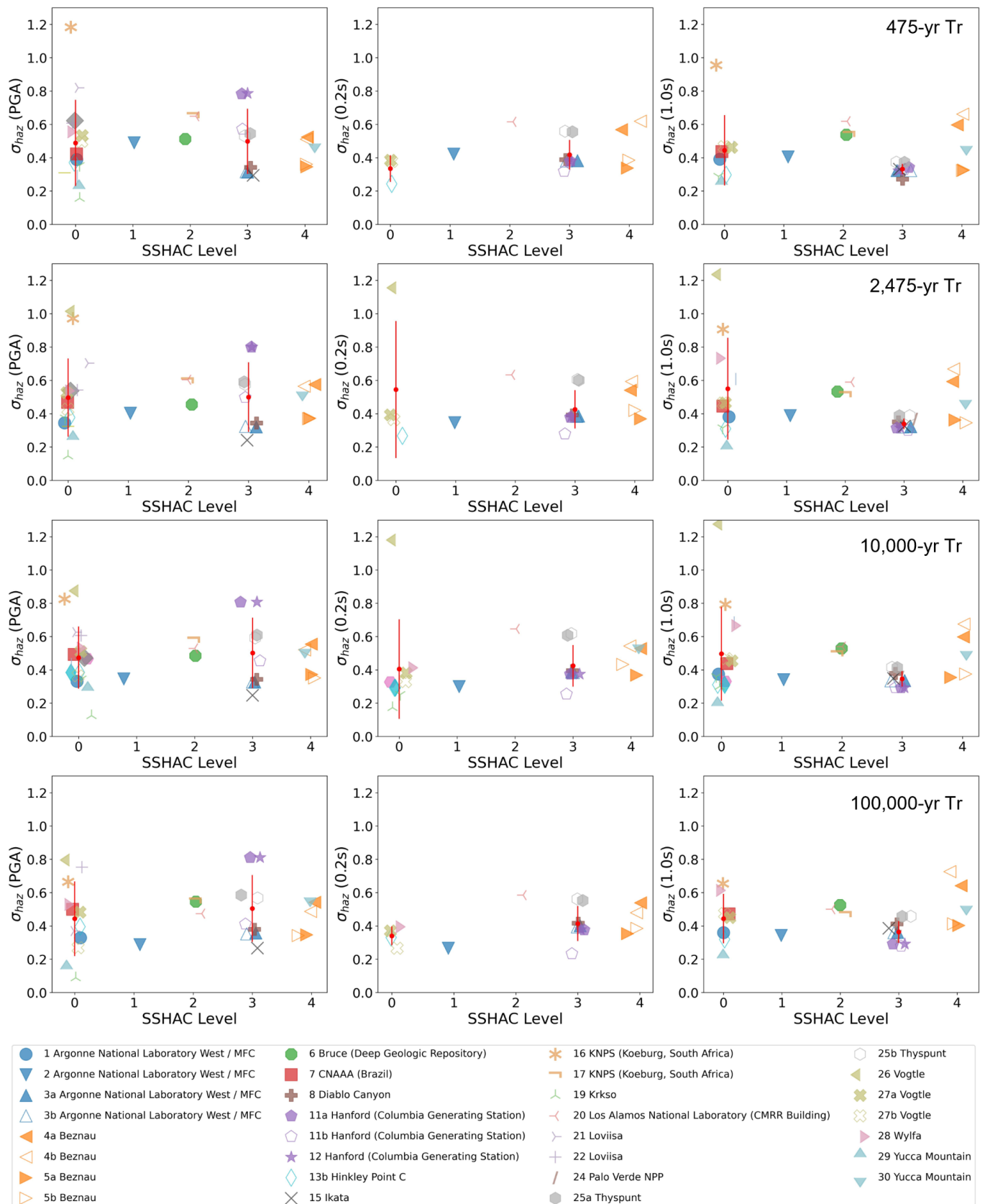
In contrast, SSHAC Level 0 studies show a wider spread in  $\sigma_{\text{haz}}$  values, with some notable outliers, specifically ‘16 KNPS’ and ‘26 Vogtle’, which if excluded would

decrease the standard deviation of  $\sigma_{\text{haz}}$ . This adjustment would bring the standard deviation of the SSHAC Level 0 studies closer to those of the Level 3 studies. The characteristics and implications of these outlier studies are discussed in more detail later in this section.

### 4.3 Comparisons by Return Period

The variability of  $\sigma_{\text{haz}}$  across return period, and IMs of interest, is shown in Figure 4, along with the corresponding mean (red dot) and standard deviation (red bar) for the four return periods considered in this study. Across these return periods,  $\sigma_{\text{haz}}$  does not show clear trends in terms of variations in mean and standard deviation. Like the analysis by SSHAC Level, the ‘16 KNPS’ and ‘26 Vogtle’ studies are clear outliers, significantly influencing the overall statistics, particularly the standard deviation at specific return periods, which vary depending on the IM. Excluding these studies from the statistical analysis would result in more uniform estimates across return periods. It is important to note that no values are available for SA(0.2s) for ‘16 KNPS’, and for ‘26 Vogtle’ values are only available for 2,475-yr and 10,000-yr Tr.

This finding contrasts with earlier studies (e.g. Douglas et al., 2014), which suggested a slight reduction in uncertainty at longer return periods due to the dominance of better-constrained magnitude-distance scenarios. The observed constant trend may reflect

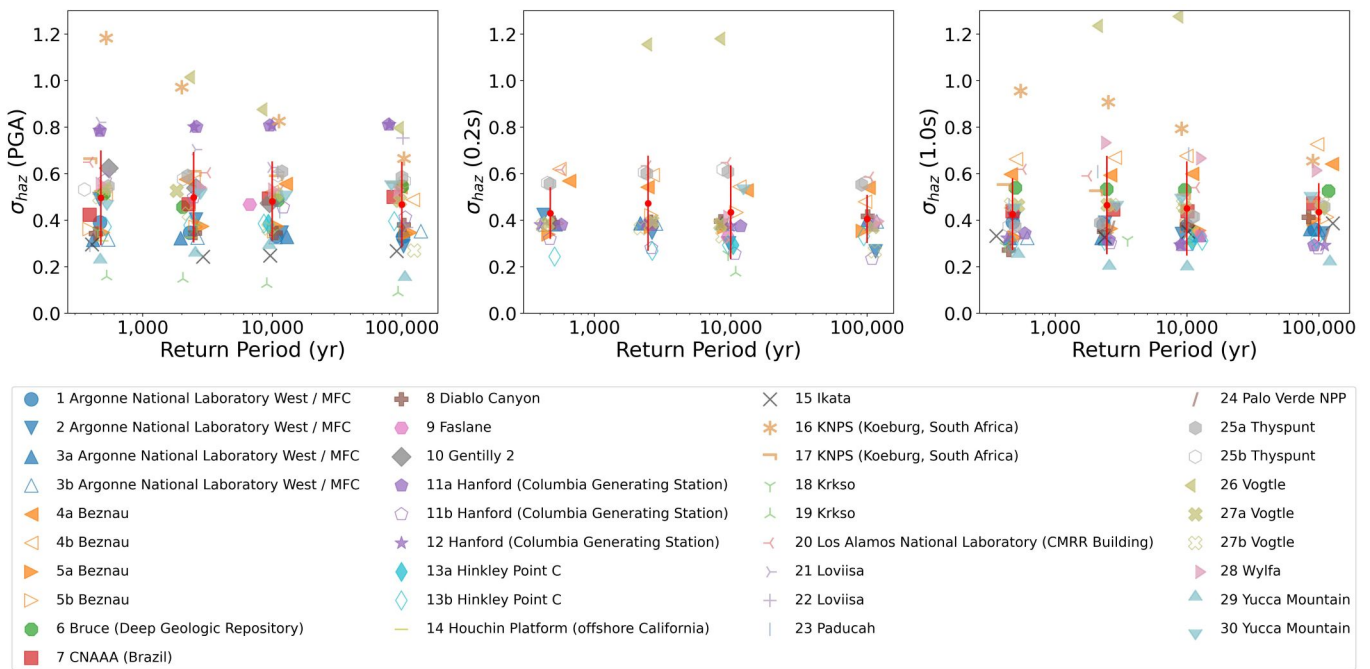


**Figure 3**  $\sigma_{\text{haz}}$  values for alternative SSHAC levels. Note: Red dots show the mean value while the red bar shows the  $\pm 1$  standard deviation. Dots and bars are only provided for SSHAC levels with enough data to provide meaningful results (0, i.e. not SSHAC-compliant, and 3).

the diversity in the tectonic settings, modelling approaches and approaches to capture epistemic uncertainties across the studies analysed.

It should be noted that the mean hazard curve often

tracks across hazard curves for higher fractiles so that at longer return periods the mean hazard curve corresponds to higher fractiles (e.g. the 84<sup>th</sup>) but at short return periods it is close to the median (50<sup>th</sup>) (e.g. Fig-



**Figure 4** Variation of  $\sigma_{\text{haz}}$  with return period.

ure 10 of Villani et al., 2020). This behaviour is not well captured by the  $\sigma_{\text{haz}}$  metric that measures the widths of the fractiles not the location of the mean.

Using a subset of the studies for which data are available for two reference horizons, typically ‘bedrock’ and ‘target’ horizons, target-to-bedrock  $\sigma_{\text{haz}}$  ratios were calculated and presented in Figure 5 (see Table 1 for the  $V_{S30}$  values from each study at each horizon). Notably, the uncertainty at the target horizon is lower than at bedrock (i.e. target-to-bedrock ratio < 1.0) for some studies. This is unexpected, as initial expectation was that uncertainties in the site response analysis, affected by uncertainties in the shear-wave velocity profile of the shallow deposits, would be additive to those observed at bedrock level.

A clear and significant reduction in  $\sigma_{\text{haz}}$  for PGA is observed for sites where the target horizon corresponds to soil ground conditions (i.e.  $V_{S30} < 600$  m/s), such as the Beznau, Hanford and Vogtle sites; although the reduction is less pronounced for the Beznau site. For Hinkley Point C and Thyspunt sites, where the target horizon is classified as rock ( $V_{S30} > 800$  m/s), along with the MFC site (INL, 2022) with a  $V_{S30}$  at the target horizon of 607 m/s, the reduction is minimal or null. A similar pattern, though with smaller differences, is observed for SA(0.2s) and SA(1.0s), except for the Beznau site, where both studies (4a,b: NAGRA, 2004; 5a,b: Swissnuclear, 2013) report generally higher  $\sigma_{\text{haz}}$  values at the target horizon compared to bedrock (i.e. target-to-bedrock ratios > 1.0).

The reason for the reduction in  $\sigma_{\text{haz}}$  values for PGA and SA(0.2s) between bedrock and target horizons is not entirely clear. The observed decrease in the target-to-bedrock ratios with increasing return period seems to suggest that this may be related, at least in part, to the overall de-amplification of the ground motions and non-linearity at longer return periods, which would influ-

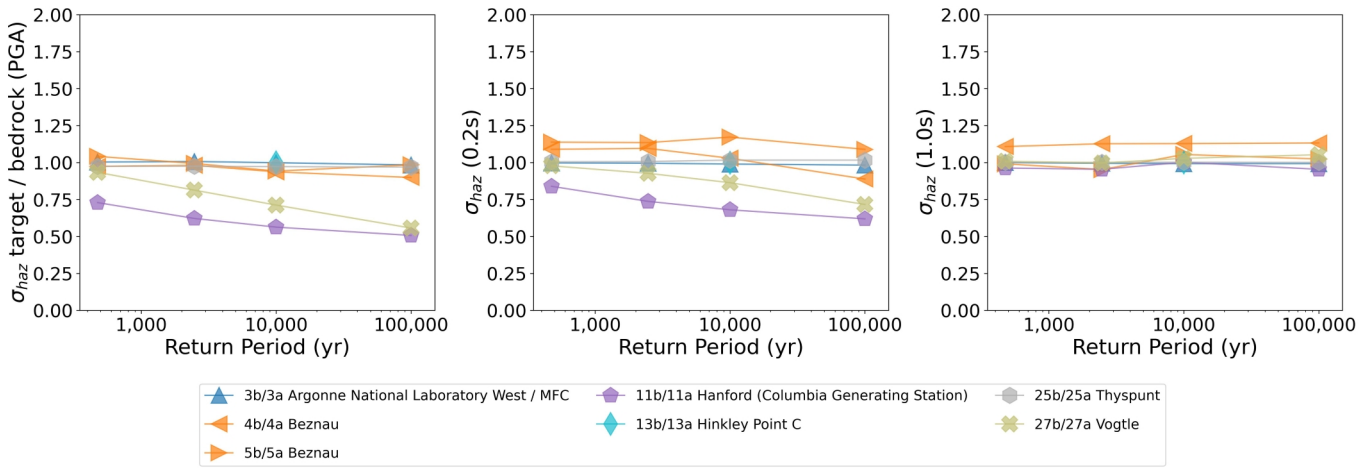
ence the slope of the hazard curves and, consequently, the uncertainty estimates. In other words, while ground motions at bedrock are expected to remain in the elastic range, site response for soil horizons is often affected by non-linearity. Particularly at high frequencies, soil non-linearity seems to strongly attenuate ground motions at the higher fractiles compared with the lower ones, leading to lower uncertainty (i.e. lower fractiles remain largely unchanged, whereas higher fractiles are diminished).

#### 4.4 Comparisons Amongst Intensity Measures

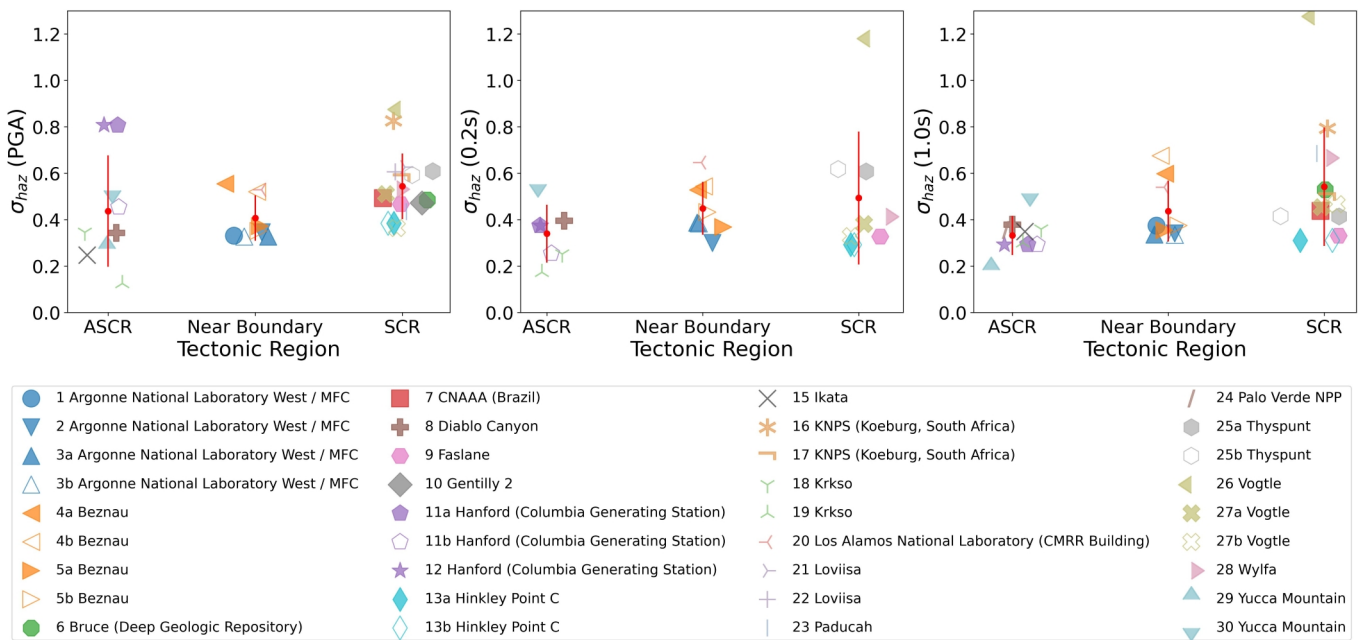
From Figure 3 and Figure 4, it is clear that PGA consistently shows the highest mean  $\sigma_{\text{haz}}$ , indicating larger uncertainties at short spectral periods. In contrast, lower mean  $\sigma_{\text{haz}}$  values are normally observed for SA (0.2s), with a slight increase for SA (1.0s). Differences in the mean values across all IMs are relatively small, in particular between SA (0.2s) and SA (1.0s), but consistent. These observations are in line with those in Valentini and Beltran (2024).

Regarding the variability in the results, represented by the estimated standard deviation, when grouped by SSHAC level (Figure 3), it is difficult to draw definite conclusions as these are strongly affected by the amount of data available and particularly outliers. Nevertheless, it is noteworthy that for SA(1.0s), SSHAC Level 3 studies tend to cluster closely, suggesting lower variability. As previously discussed, this may be due to methodological choices in SSHAC Level 3 studies such as the use of backbone ground-motion models.

When the data are grouped by return period (Figure 4), the resulting standard deviations provide a more representative view of the variability across IMs as they are less influenced by methodological differences amongst studies. Figure 4 shows that the standard



**Figure 5** Target to bedrock  $\sigma_{\text{haz}}$  ratios across return periods.



**Figure 6** Variation of  $\sigma_{\text{haz}}$  with tectonic region for the 10,000-yr return period.

deviations are relatively uniform across IMs, with the highest values clearly influenced by outliers. Excluding these outliers from the statistics, particularly the ‘16 KNPS’ and ‘26 Vogtle’ studies, would result in more consistent standard deviations across IMs.

#### 4.5 Comparisons by Tectonic Region

Figure 6 shows  $\sigma_{\text{haz}}$  values plotted against tectonic region for the three IMs at the 10,000-yr  $T_r$ , using the classification provided in Table 1. In addition to the values for the individual studies, the mean (red dot) and standard deviation (red bar) for each tectonic region are also shown. Results for the other return periods of interest are omitted here, as they show similar trends.

Uncertainty levels in PSHAs for ASCRs tend to be lower than those for SCRs, while studies classified as ‘near boundary’ fall between these two categories. This pattern aligns with expectations, given the more limited data typically available for SCRs compared with ASCRs.

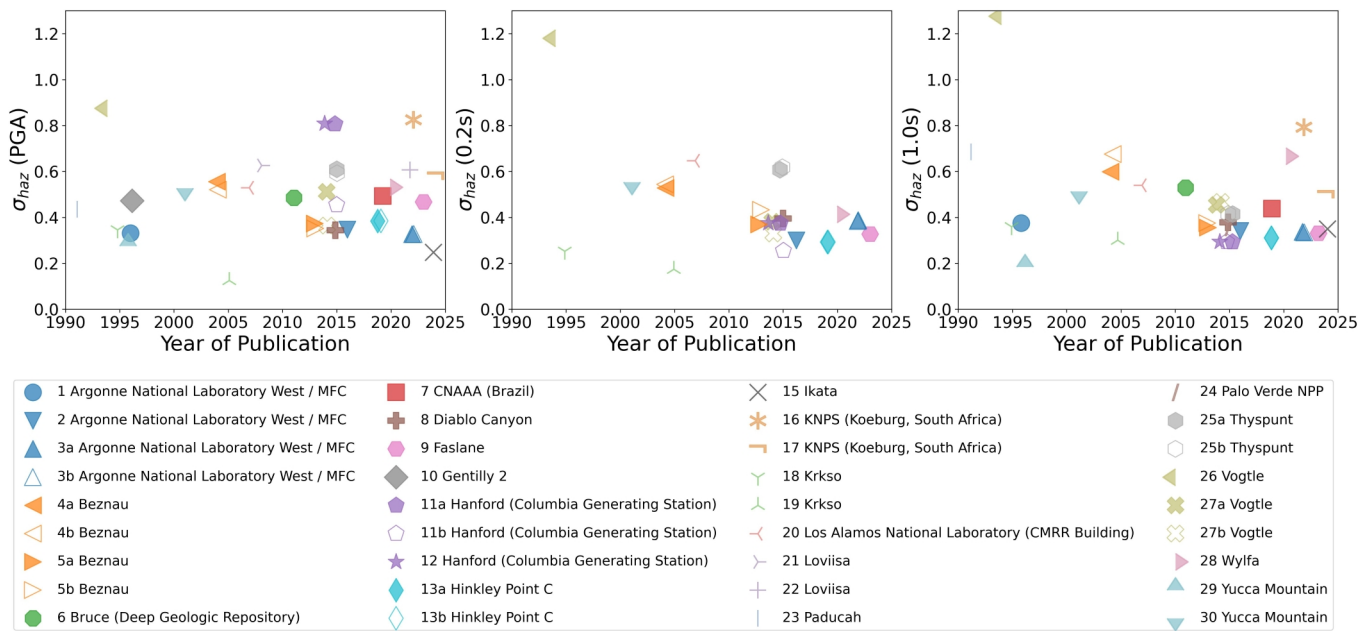
The spread of the uncertainty estimates, indicated by

the standard deviation (red bar) in Figure 6, also appears to be greater for SCR sites. This is particularly evident for SA(0.2s) and SA(1.0s), which, again, is consistent with expectations. However, for these two IMs, the larger standard deviations are partially influenced by the results from ‘26 Vogtle’. Similarly, for PGA, the wider standard deviation observed for the ASCR sites seems to be driven by the higher uncertainty levels associated with ‘11a Hanford’ and ‘12 Hanford’.

#### 4.6 Comparisons by Publication Date

The variation of  $\sigma_{\text{haz}}$  with the date of publication of the PSHA results, for the 10,000-yr  $T_r$ , is presented in Figure 7. Similar trends are observed for the other return periods of interest; therefore, they are not shown here. Note that the stated year of publication corresponds to the date of the data source we used and may not reflect the completion date of the PSHA study or the first publication of its results.

Overall, no clear or discernible pattern is observed,



**Figure 7** Variation of  $\sigma_{\text{haz}}$  with date of publication of PSHA results for the 10,000-year return period.

with  $\sigma_{\text{haz}}$  for all IMs being relatively constant with date. A slight decrease may be noted for SA (0.2s); however, this reduction is minimal and could be attributed to limited data availability, particularly the absence of available results for this IM from the '16 KNPS' study, which exhibits atypically high uncertainties. Notably, the largest outliers (i.e. the highest and lowest  $\sigma_{\text{haz}}$  values) correspond to older studies, specifically '26 Vogtle', a pre-SSHAC (i.e. before 1997) study, and '19 Krsko' which is also a non-SSHAC study. From 2005 onward the spread of  $\sigma_{\text{haz}}$  remains relatively constant. The only exception in this period (post 2005) is the '16 KNPS' study which shows atypically high uncertainty levels at shorter return periods (i.e. 475-yr and 2,475-yr Tr) as previously observed in Figure 3 and Figure 4.

Figure 8 shows the variation of  $\sigma_{\text{haz}}$  with publication date for the 10,000-yr Tr, considering only sites for which one or more studies are available. In most cases, there is a clear reduction in epistemic uncertainty with time. This was expected as successive studies for the same site typically build upon the knowledge, data and experience from previous assessments, thereby reducing epistemic uncertainty.

Nevertheless, there are a couple noticeable exceptions. For the Hanford site,  $\sigma_{\text{haz}}$  values for the bedrock horizon from both the '11a Hanford (GCS)' and '12 Hanford (CGS)' studies are nearly identical. It is important to note that no data are available for the target horizon from the '12 Hanford (CGS)' study. For the Yucca Mountain site an increase in the epistemic uncertainty is observed between the '29 Yucca Mountain' and the '30 Yucca Mountain' studies. This increase is likely due to a more comprehensive characterisation of the epistemic uncertainty in the '30 Yucca Mountain' study, which followed the SSHAC Level 4 process, in contrast to the earlier '29 Yucca Mountain' study, which predates the SSHAC guidelines.

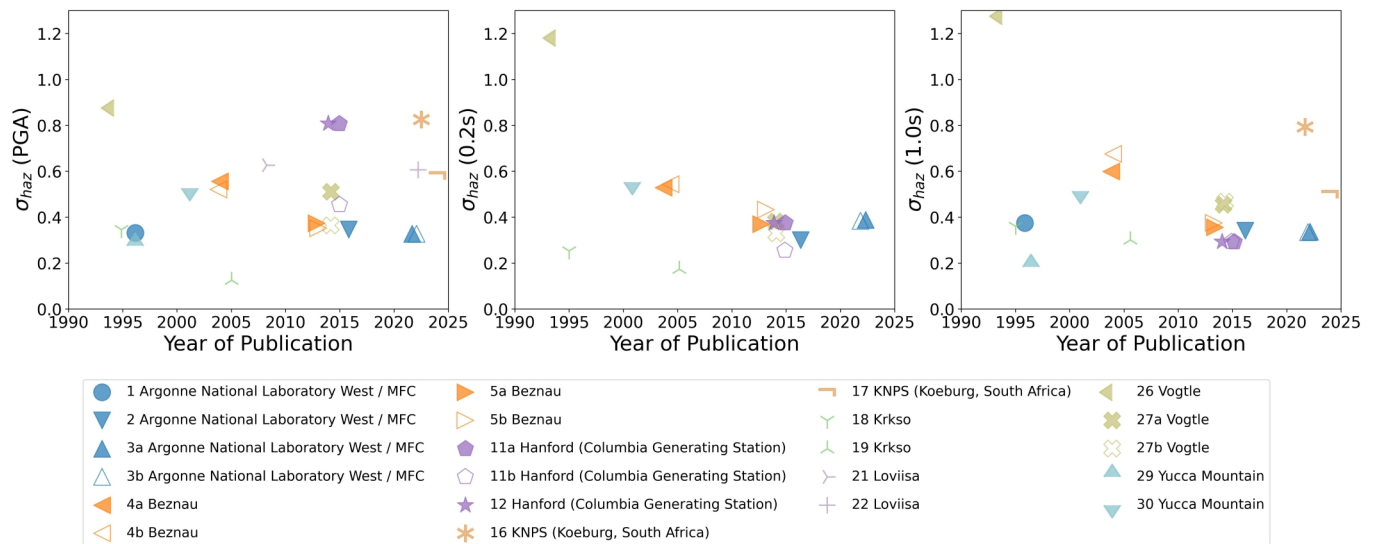
### 4.7 Comparisons Amongst Uncertainty Metrics

As discussed above, there is a strong positive correlation between  $\sigma_{\text{haz}}$ , Unc and  $\sigma_{\mu}$ . However, the correlation between  $\sigma_{\text{haz}}$  and R is only moderate and clearly non-linear. Given these strong correlations, the findings presented in previous sections for  $\sigma_{\text{haz}}$  are also applicable to Unc and  $\sigma_{\mu}$ , and to a lesser degree, to R. Figure 9 presents the variation of the four uncertainty metrics with return period.

As expected, the behaviour of the uncertainty estimates based on  $\sigma_{\text{haz}}$  and Unc is nearly identical. The uncertainty estimates using  $\sigma_{\mu}$  also show a similar trend, though the correlation is somewhat weaker. Notable differences are, however, observed when uncertainty is quantified using the R metric. Most prominently, the majority of studies cluster around  $R = 1.2$ , while a few studies show R values that are two or three times higher.

The change in the trend of the uncertainty estimates for some studies is worth noting. The behaviour for the Hanford site is probably the most evident. For both the '11a,b Hanford (CGS)' and '12 Hanford (CGS)' studies, uncertainty estimates remain relatively constant with return period for  $\sigma_{\text{haz}}$ , Unc and  $\sigma_{\mu}$  but show a pronounced increase with return period when using R. A comparable situation is observed for the '16 KNPS' study, particularly for SA(1.0s).

These findings reinforce earlier observations that R is a poor metric for quantifying uncertainty, as suggested by its moderate and non-linear correlation with  $\sigma_{\text{haz}}$ . The R metric is strongly affected by outliers in the PSHA results, which is expected given that R is more indicative of the skewness of the distribution of the hazard results rather than of uncertainty. Despite its direct correlation with  $\sigma_{\mu}$ , the latter appears less affected by the skewness of the data, likely due to the consideration of the slope of the hazard curve in log-log space (k) within its derivation.



**Figure 8** Variation of  $\sigma_{\text{haz}}$  with date of publication of PSHA results for the 10,000-yr return period and selected studies.

Our findings are in line with those from the previous studies discussed above. For example, Figure 9 shows that  $\text{Unc}$  is generally between about 30 and 50 with some studies leading to values as low as 20 and as high as 60 (excluding a handful of outliers). These ranges match those reported by Douglas et al. (2014). The values of  $\sigma_{\mu}$  are similar to the values reported in Douglas (2018) and Douglas et al. (2024), with most values clustered between 0.3 and 0.6. The values of  $\sigma_{\text{haz}}$  are also between 0.3 and 0.6, which is close to the range reported in Douglas et al. (2024). Finally, our values of  $R$  generally fall within a tight range from 1.0 to 1.5 (excluding a handful of outliers), which matches the range of Valentini and Beltran (2024).

#### 4.8 Studies With Atypical Uncertainty Levels

As discussed above, the four uncertainty metrics computed from the hazard results of the thirty PSHAs considered in this study generally fall within relatively tight limits. These limits have weak to no correlation with various characteristics (SSHAC level, date, site conditions and return period) and only weak correlations with spectral period, with slightly higher values observed for PGA and SA(1.0s), and lower values for SA(0.2s), and tectonic classification. A handful of studies fall outside these limits, often appearing as outliers across all metrics.

The clearest high outlier is ‘26 Vogtle’ PSHA conducted by LLNL, as reported in U.S. Nuclear Regulatory Commission (1994). Considering the results for the 10,000-yr Tr,  $\sigma_{\text{haz}}$  for this study is about 0.9 for PGA and around 1.2 for both SA(0.2s) and SA(1.0s), which is roughly double the average for this metric. As noted earlier, the PSHAs conducted by LLNL were one of the triggers for developing the SSHAC guidelines because the uncertainties captured by those studies were much higher than those from the EPRI PSHAs conducted at the same time. Therefore, it is unsurprising that this study is an outlier with respect to subsequent PSHAs, including the 2014 ‘27a,b Vogtle’ study for the same site.

Other high outliers are the studies ‘11 a Hanford

(CGS)’, ‘12 Hanford (CGS)’ for PGA, and ‘16 KNPS’ for both PGA and SA(1.0s). The uncertainty metrics for these studies are in line with other studies for SA(0.2s), and in the case of ‘11 a Hanford (CGS)’ and ‘12 Hanford (CGS)’, also for SA(1.0s). The  $\sigma_{\text{haz}}$  values for ‘11 a Hanford (CGS)’ and ‘12 Hanford (CGS)’ are about 0.8 across all return periods, whereas for ‘16 KNPS’,  $\sigma_{\text{haz}}$  varies from about 1.2 at the 475-yr Tr (more than double the average for this metric) to approximately 0.7 at the 100,000-yr Tr. A potential reason for the high values for ‘11 a Hanford (CGS)’ and ‘12 Hanford (CGS)’ could be high uncertainties in estimating very hard rock ( $V_{S30} = 3,000\text{m/s}$ ) site response, in particular kappa. The value of  $\sigma_{\text{haz}}$  is lower (about 0.45) in ‘11 b Hanford (CGS)’ for softer surface deposits ( $V_{S30} = 361\text{m/s}$ ). The reason for the large uncertainties and strong return period dependency in ‘16 KNPS’ is unknown, but it could be related to the difficulties in estimating seismic hazard in a relatively low seismicity area with limited data (South Africa). The uncertainties from ‘17 KNPS’, which was published two years later based on an Enhanced SSHAC Level 2 study, are lower and more in line with other studies, including ‘25 a Thyspunt’, which is a SSHAC Level 3 PSHA for another site in South Africa.

There are fewer outliers at the low end. The clearest is ‘19 Krsko’, a study for a NPP site in Slovenia, where  $\sigma_{\text{haz}}$  is around 0.1 for PGA and about 0.2 for SA(0.2s) at the 10,000-yr Tr. This site is within an ASCR where the earthquake catalogue and ground-motion databases are relatively large, and for which published GMPEs using data from similar tectonics were available. Another study with relatively low uncertainty metrics is ‘29 Yucca Mountain’, particularly for SA(1.0s), where  $\sigma_{\text{haz}}$  is around 0.2 for the 10,000-yr Tr. The reason for this low value is unknown, but this was a preliminary PSHA for the SSHAC Level 4 study ‘30 Yucca Mountain’, which led to uncertainty metrics within typical ranges, e.g. for SA(1.0s) the value of  $\sigma_{\text{haz}}$  is around 0.5 for the 10,000-yr Tr.

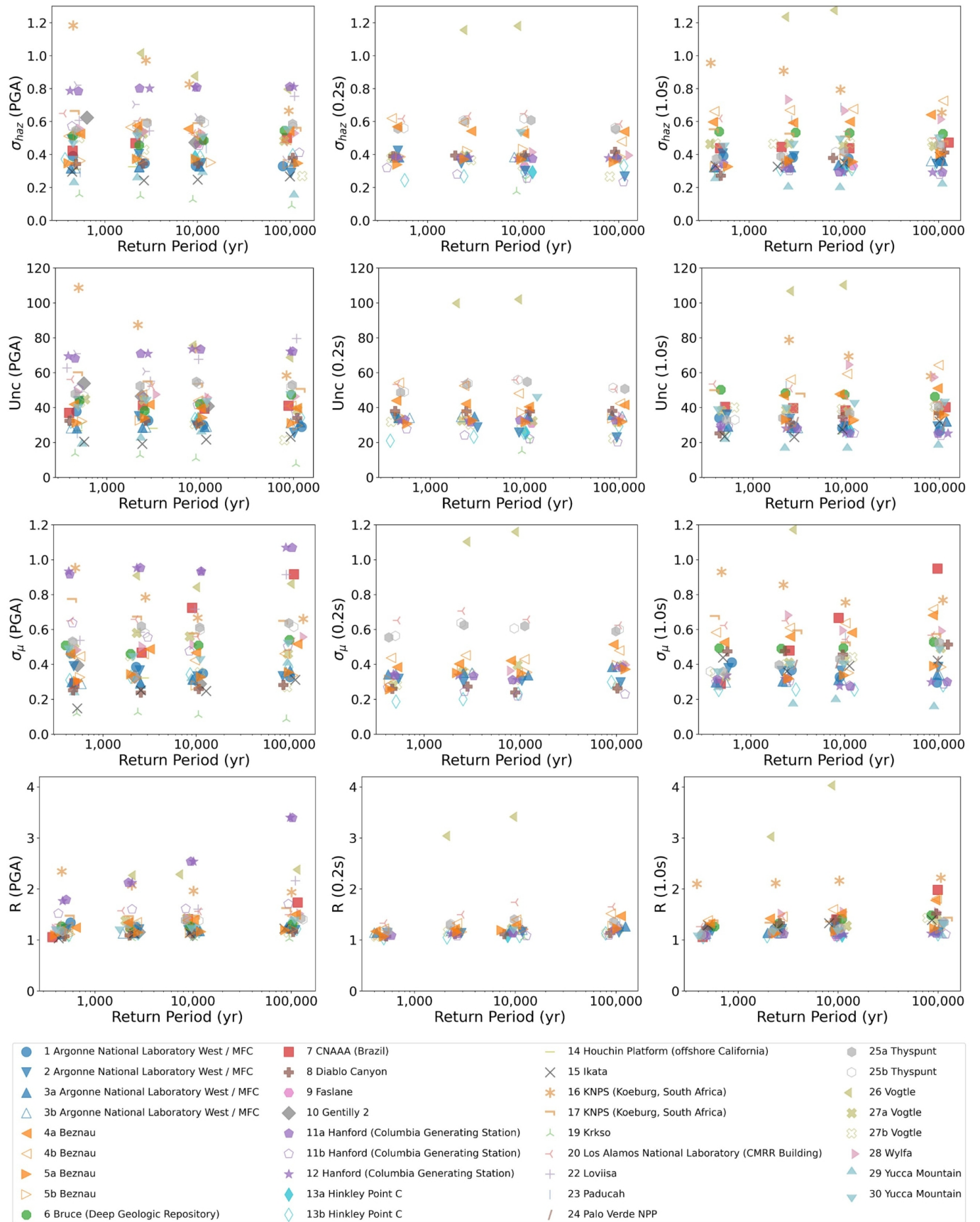


Figure 9 Variation of  $\sigma_{haz}$ , Unc,  $\sigma_{\mu}$  and R with return period.

## 5 Discussion

As previously discussed, the understanding and quantification of the epistemic uncertainty in PSHA studies is of particular importance, specifically for safety-critical infrastructure. However, evaluating whether the level of epistemic uncertainty captured in a PSHA is adequate or not is not a trivial task as it implies, by definition, the quantification of the unknown.

The first step towards this objective is to define a metric to measure in a consistent manner the uncertainty level. In this study, four alternative uncertainty metrics were examined. Amongst these metrics,  $\sigma_{\text{haz}}$  is preferred as it is mathematically the most direct measure of uncertainty in PSHA.  $\text{Unc}$  and  $\sigma_{\mu}$  are strong alternatives given their strong correlation with  $\sigma_{\text{haz}}$ , resulting in comparable results. Of these two,  $\text{Unc}$  has the advantage of requiring less data for its computation and stronger correlation with  $\sigma_{\text{haz}}$ .  $R$  was found to be a relatively poor metric, as it is highly sensitive to outliers in the PSHA logic tree and is, therefore, influenced by the skewness of the distribution of the hazard results.

Once a metric has been selected, the second step would be to assess whether the uncertainty level is 'adequate'. A way of judging this is through the comparison with the level of epistemic uncertainty captured in the PSHA with that observed in studies for other sites, ideally with similar ground conditions and tectonic regime. Although, based on the results of this study, the latter seem to have a clear but limited influence on the overall uncertainty levels. The observations presented above provide a reference baseline for this purpose. If the uncertainty metrics for a particular PSHA were considerably higher or lower than those from comparable studies, this could indicate that additional work is required to understand the source of the differences. For instance, a lower level of uncertainty may be justifiable if extensive, high-quality geotechnical data are available, allowing for well-constrained inputs to the site-response analyses or that show the site response is relatively simple. Similarly, high uncertainty may be reasonable if, for example, the available earthquake catalogue is short, and the events have highly uncertain magnitudes and locations. Nevertheless, although having an established baseline against which to compare uncertainty levels provides a useful check on the PSHA, it is important to be cautious of the risk of 'anchoring', that is, giving disproportionate weight to results from previous studies, and to ensure that such comparisons do not override site-specific considerations. It is important to emphasize that there is no 'right' or 'absolute' level of epistemic uncertainty for a particular study. As discussed above, this depends on multiple factors, including the complexity of the hazard models, the quality and quantity of the available data, and other study-specific considerations.

The various metrics discussed in this study can be computed easily from hazard curves and could therefore be routinely computed as part of the hazard calculations. This baseline could then be used to determine if changing the logic tree weights or branches 'move the needle' in terms of uncertainty. This information would

allow the hazard analyst to focus on the most influential components of the model and, where relevant, undertake targeted data collection to reduce the uncertainty in the most effective manner. This step is undertaken in some studies, but it is unusual to use quantitative metrics to guide this.

Several of our observations challenge conventional expectations and, in some cases, appear counterintuitive. Notably, the SSHAC level seems to have little influence on the level of uncertainty captured in PSHA studies, particularly for post-2005 studies. This may be the result of regulatory requirements that impose strict requirements on non-SSHAC studies for critical infrastructure, along with some influence from the SSHAC guidelines on the methodology applied for the delivery of at least some of these studies. Some studies may not be SSHAC compliant but still incorporate elements of the SSHAC framework (e.g., [Aldama-Bustos et al., 2023](#)).

Another noteworthy finding is that uncertainty levels remain almost constant across AFEs. The expectation was that uncertainty would increase with decreasing AFE, given that lower probability scenarios are associated with larger uncertainties. We did find, as expected, that studies conducted for sites in ASCRs, which tend to have higher design ground-motion intensities, generally show lower uncertainty levels compared to those in SCRs, which is likely due to varying data availability (e.g. the length and completeness of the earthquake catalogue and availability of ground-motion records). In contrast, when examining the results of [Valentini and Beltran \(2024\)](#), specifically their Figure 5, there is no clear difference in the results for their SCR sites and those from more active regions (e.g. Hanford, Diablo Canyon and Ikata).

Arguably the most revealing finding from this study is the negligible difference in the epistemic uncertainty between the bedrock and target horizon when the target horizon can be classified as a rock site, but a reduction, in some cases a substantial reduction, in the epistemic uncertainty for PGA and SA(0.2s), when the target horizon corresponds to soil. These results are counterintuitive as the incorporation of the epistemic uncertainty from the site response analysis was expected to increase the overall uncertainty, not reduce it. This is an area for further research and discussion to better understand the underlying mechanism and implications for future PSHAs.

## 6 Conclusions and Recommendations

This study has quantified and compared the epistemic uncertainty level captured in thirty site-specific probabilistic seismic hazard assessments, all but one for nuclear-related facilities. The database covers a broad geographic distribution, a wide range of SSHAC levels, from non-SSHAC compliant studies to SSHAC Level 4 studies, and a variety of ground conditions. This diversity allowed a comprehensive examination of the epistemic uncertainty estimated using four alternative metrics and their behaviour across SSHAC levels, return periods, ground-motion intensities, spectral periods, publication dates and site conditions. The study builds

upon and expands the scope of previous studies, with overall findings in good agreement with earlier work.

Amongst the four metrics considered in this study, the standard deviation computed from hazard fractiles ( $\sigma_{\text{haz}}$ ) is the preferred option as it is the most direct and mathematically robust measure of uncertainty. The ratio of 84<sup>th</sup> to 16<sup>th</sup> fractiles (Unc) and a standard deviation estimated assuming all uncertainty is due to the ground-motion model ( $\sigma_{\mu}$ ) are also considered strong alternatives due to their good correlation with  $\sigma_{\text{haz}}$ , with Unc having the advantage of requiring less data and a simpler calculation. In contrast, the ratio of the mean to median ground motions (R) is less reliable as an uncertainty measure, being strongly affected by outliers, and therefore skewness, in the logic tree outputs. A potential alternative metric is the fractile corresponding to the mean hazard, which was not considered here.

Many of our findings challenge conventional expectations. Most notably the SSHAC level, or lack of formal compliance with the SSHAC framework, seems to have limited influence on the level of epistemic uncertainty, particularly when considering studies conducted after 2005. This may be the result of strict regulatory requirements for nuclear facilities and the broader influence of the SSHAC principles, even in studies not formally designated as SSHAC-compliant.

Similarly, the reduction in the epistemic uncertainty between the bedrock and target horizons, when the ground conditions at the target horizon can be classified as soil is also noteworthy, particularly for PGA and short spectral periods. This is against the expectation that site response analysis would increase the overall epistemic uncertainty. This is expected to be related mainly to the attenuation of the high frequency motions, along with the non-linearity of the amplification factors at high intensities / long return periods, in soil profiles, which influences the slope of the hazard curves. This is an area that needs further investigation, particularly regarding its implications for the interpretation of PSHA results.

A key recommendation is that the total epistemic uncertainty should be routinely estimated and reported in the PSHA documentation in the same manner as disaggregated results, which have become a standard output in PSHA studies. The results from this study can provide a useful benchmark for assessing whether the level of epistemic uncertainty in a PSHA is appropriate. This may be particularly useful for PSHAs for infrastructure projects subject to less strict regulatory oversight than that for nuclear facilities. Such assessments can be used to re-evaluate the logic tree and the assigned weights and would assist the hazard analyst in prioritising inputs that contribute most significantly to the spread (i.e. uncertainty) in the hazard results. However, care must be taken to avoid ‘anchoring’ to previous studies, particularly when there are sound technical reasons to expect higher or lower uncertainty levels than the average.

## Acknowledgements

We would like to thank clients who allow the results of the hazard studies they sponsor to be published; this transparency enabled this study. We thank two anonymous

reviewers for their careful and useful suggestions on previous versions of this manuscript.

## Data and Code Availability

The database containing the seismic hazard results used in this study, along with the associated uncertainty values for all four metrics considered, is available at: <https://figshare.com/s/4fbf432b17495c4744ae>.

## Competing interests

The authors have no competing interests

## References

- Adams, J. and Halchuk, S. Uncertainty spread in the 5th Generation seismic hazard results used in NBCC2015. In *12th Canadian Conference on Earthquake Engineering*, Quebec City, Chateau Frontenac, Canada, 2019.
- Aldama-Bustos, G., Douglas, J., Strasser, F. O., Daví, M., and MacGregor, A. Methods for assessing the epistemic uncertainty captured in ground-motion models. *Bulletin of Earthquake Engineering*, 21(1):1–26, 2023. doi: 10.1007/s10518-022-01515-8.
- AMEC-Geomatrix. Seismic hazard assessment, OPG’s geologic repository for low and intermediate level waste, NWMO DGR-TR-2011-20, revision R000, 2011.
- Anderson, J. and Hough, S. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America*, 74(5):1969–1993, 1984.
- Arizona Public Service Company. Seismic hazard and screening report for Palo Verde Nuclear Generating Station Units 1, 2 and 3, Revision 0, Arizona, USA, 2015.
- Bender, B. Incorporating acceleration variability into seismic hazard analysis. *Bulletin of the Seismological Society of America*, 74(4):1451–1462, 1984.
- Bommer, J. J. and Scherbaum, F. The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis. *Earthquake Spectra*, 24(4):997–1009, Nov. 2008. doi: 10.1193/1.2977755.
- Bommer, J. J., Coppersmith, K. J., Coppersmith, R. T., Hanson, K. L., Mangongolo, A., Neveling, J., Rathje, E. M., Rodriguez-Marek, A., Scherbaum, F., Shelembe, R., Stafford, P. J., and Strasser, F. O. A SSHAC Level 3 Probabilistic Seismic Hazard Analysis for a New-Build Nuclear Site in South Africa. *Earthquake Spectra*, 31(2):661–698, May 2015. doi: 10.1193/060913eqs145m.
- Budnitz, R., Apostolakis, G., Boore, D., Cluff, L., Coppersmith, K., Cornell, C., and Morris, P. Recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts. *NUREG/CR-6372, two volumes, U.S. Nuclear Regulatory Commission*, Apr. 1997. doi: 10.2172/479072.
- Cornell, C. *An interim recommendation for DOE use of the LLNL and EPRI hazard curves. Appendix D in DOE standard DOE-STD-1024-92, Guidelines for use of probabilistic seismic hazard curves.* Department of Energy Sites for Department of Energy Facilities, Washington, D.C., 1992.
- de Almeida, A. A. D., Assumpção, M., Bommer, J. J., Drouet, S., Riccomini, C., and Prates, C. L. M. Probabilistic seismic hazard analysis for a nuclear power plant site in southeast Brazil. *Journal of Seismology*, 23(1):1–23, 2019. doi: 10.1007/s10950-018-9755-8.
- Douglas, J. *Capturing Geographically-Varying Uncertainty in Earthquake Ground Motion Models or What We Think We Know May*

- Change, page 153–181. Springer International Publishing, 2018. doi: 10.1007/978-3-319-75741-4\_6.
- Douglas, J., Ulrich, T., Bertil, D., and Rey, J. Comparison of the Ranges of Uncertainty Captured in Different Seismic-Hazard Studies. *Seismological Research Letters*, 85(5):977–985, 2014. doi: 10.1785/0220140084.
- Douglas, J., Crowley, H., Silva, V., Marzocchi, W., Danciu, L., and Pinho, R. Methods for evaluating the significance and importance of differences amongst probabilistic seismic hazard results for engineering and risk analyses: a review and insights. *Bulletin of Earthquake Engineering*, 22(6):2769–2796, Mar. 2024. doi: 10.1007/s10518-024-01896-y.
- Electric Power Research Institute. CEUS ground motion project final report. Technical report 1009684, EPRI Palo Alto, CA; Dominion Energy, Glen Allen, VA; Entergy Nuclear, Jackson, MS; and Exelon Generation Company, Kennett Square, PA, 2004.
- Energy Northwest. Seismic hazard and screening report for the Columbia Generating Station. Technical report, Columbia Generating Station, Richland, Washington, USA, 2015.
- Fujiwara, H., Ebisawa, K., Kagawa, T., Si, H., Furumura, T., Miyake, H., Morikawa, N., Shiota, T., Ogawa, H., Matsusaki, S., Miyakoshi, J., Sakai, T., and Kameda, H. Development of ground motion characterization model at the Ikata site based on guidelines for SSHAC level 3. *Journal of Japan Association for Earthquake Engineering*, 24(2), 2024. doi: 10.5610/jaee.24.2\_1.
- IAEA. *Non-ergodic Ground Motion Models for Site Specific Seismic Hazard Assessment at Nuclear Installation Sites, IAEA TECDOC series*. International Atomic Energy Agency, May 2025. doi: 10.61092/iaea.9f3l-atpr.
- INL. Site-specific probabilistic seismic hazard analyses for the Idaho National Engineering Laboratory. Final report INEL-95/0536, volume 1, Woodward-Clyde Federal Services, Geomatrix Consultants and Pacific Engineering and Analysis, 1996.
- INL. SSHAC Level 1 probabilistic seismic hazard analysis for the Idaho National Laboratory. Revision INL/EXT-15-36682, Idaho National Laboratory, Idaho Falls, Idaho., 2016.
- INL. *Idaho National Laboratory sitewide SSHAC Level 3 probabilistic seismic hazard analysis, INL/RPT-22-70233*. Idaho National Laboratory, Idaho Falls, Idaho, 2022.
- Kulkarni, R., Youngs, R., and Coppersmith, K. Assessment of confidence intervals for results of seismic hazard analysis. In *Proceedings of the Eighth World Conference on Earthquake Engineering*, volume 1, page 263–270, San Francisco, USA, 1984.
- Lapajne, J. K. and Fajfar, P. Seismic hazard assessment of an existing NPP in Slovenia. In *Transactions of the 13th International Conference on Structural Mechanics in Reactor Technology (SMiRT 13)*, M043-1, 797-802, Porto Alegre, Brazil, 1995.
- Liou, I. and Abrahamson, N. Framework for Aleatory Variability and Epistemic Uncertainty for the Ground-Motion Characterization Based on the Level of Simplification. *Bulletin of the Seismological Society of America*, 115(1):296–314, Dec. 2025. doi: 10.1785/0120240141.
- NAGRA. Probabilistic seismic hazard analysis for Swiss nuclear power plant sites (PEGASOS Project), prepared for Unterausschuss Kernenergie der Überlandwerke (UAK) by the Nationale Genossenschaft für die Lagerung radioaktiver Abfälle (NAGRA), Final report. 1(6):2557, 2004.
- Nielsen, A. Updating the seismic design guidance for HMNB Clyde. In *SECED 2023 Conference: Earthquake Engineering and Dynamics for a Sustainable Future, 14-15 September*, Cambridge, UK, 2023.
- Pacific Electric and Gas Company. Seismic hazard and screening report, Diablo Canyon Power Plant, Units 1 and 2, PGandE Letter DCL-15-035, Diablo Canyon Power Plant. Technical report, Avila Beach, California, USA, 2015.
- PNNL. Hanford sitewide probabilistic seismic hazard analysis. In *Pacific Northwest National Laboratory, prepared for the U.S. Pacific Northwest National Laboratory. In Department of Energy, under Contract DE-AC05-76RL01830, and Energy Northwest, PNNL-23361*. 2014.
- Scherbaum, F. and Kuehn, N. M. Logic Tree Branch Weights and Probabilities: Summing up to One is not Enough. *Earthquake Spectra*, 27(4):1237–1251, Nov. 2011. doi: 10.1193/1.3652744.
- Schulte, S. M. and Mooney, W. D. An updated global earthquake catalogue for stable continental regions: reassessing the correlation with ancient rifts. *Geophysical Journal International*, 161(3):707–721, 2005. doi: 10.1111/j.1365-246x.2005.02554.x.
- Smith, C. and Murray, R. Development of a probabilistic seismic hazard analysis for offshore facilities in the Santa Barbara Channel. In *Proceedings of the 31st U.S.-Japan Joint Panel on Wind and Seismic Effects*, page 140–155, 1999.
- Southern Nuclear Operating Company. Seismic hazard and screening report - Vogtle Electric Generating Plant - Units 1 and 2, Enclosure 1 to SNC Letter, NL-14-0344. Technical report, Birmingham, Alabama, 2014.
- Stamatikos, J., Watson-Lamprey, J., Cawthra, H., Claassen, D., Coppersmith, R., Johnson, C., Largent, M., Manzunzu, B., Midzi, V., Mulabisana, T., Murphy, D., Rathje, E., and Wooddell, K. Baseline PSHA for the Duynefontyn site and the Koeberg Nuclear Power Station. Report number: CGS-2022-0009, Council for Geoscience, South Africa, 2022.
- Stamatikos, J., Watson-Lamprey, J., Cawthra, H., Claassen, D., Coppersmith, R., Johnson, C., Midzi, V., Montaldo Falero, V., Mulabisana, T., Rathje, E., Stafford, P., Ulmer, K., Largent, M., Williams, T., and Shaw, K. Enhanced SSHAC Level 2 probabilistic seismic hazard analysis for the Duynefontyn nuclear site, Western Cape Province, South Africa. Report number: 2024-0001, Council for Geoscience, South Africa, 2024.
- Stepp, J. C., Wong, I., Whitney, J., Quittmeyer, R., Abrahamson, N., Toro, G., Youngs, R., Coppersmith, K., Savy, J., and Sullivan, T. Probabilistic Seismic Hazard Analyses for Ground Motions and Fault Displacement at Yucca Mountain, Nevada. *Earthquake Spectra*, 17(1):113–151, Feb. 2001. doi: 10.1193/1.1586169.
- Swissnuclear. Probabilistic seismic hazard analysis for Swiss nuclear power plant sites—PEGASOS Refinement Project, 2013.
- Toro, G. and McGuire, R. Evaluation of seismic hazard at the Paducah gaseous diffusion plant. In *Proceedings of the Third DOE Natural Phenomena Hazards Mitigation Conference*, page 157–165, 1991.
- Toro, G. R. The effects of ground-motion uncertainty on seismic hazard results: examples and approximate results, 2006. doi: 10.13140/RG.2.1.1322.2007.
- Tromans, I. J., Aldama-Bustos, G., Douglas, J., Lessi-Cheimariou, A., Hunt, S., Daví, M., Musson, R. M. W., Garrard, G., Strasser, F. O., and Robertson, C. Probabilistic seismic hazard assessment for a new-build nuclear power plant site in the UK. *Bulletin of Earthquake Engineering*, 17(1):1–36, 2019. doi: 10.1007/s10518-018-0441-6.
- Ursic, M., Stritar, A., Vojnovic, D., and Mühleisen, A. Evolution of the NPP Krško safety assessment due to seismic hazard. In *Proceedings of the Nuclear Energy for New Europe International Conference, 5-8 September*, Bled, Slovenia, 2005.
- U.S. Nuclear Regulatory Commission. Revised Livermore seismic hazard estimates for sixty-nine nuclear power plant sites east of the Rocky Mountains. Final report, NUREG-1488, Washington, DC, 1994.

- U.S. Nuclear Regulatory Commission. Updated Implementation Guidelines for SSHAC Hazard Studies. Report NUREG-2213, 2018.
- Valentini, A. and Beltran, F. Mean-to-median spectral acceleration ratios in Senior Seismic Hazard Analysis Committee Level 3 probabilistic seismic hazard analyses: An analysis of reported results. *Earthquake Spectra*, 40(1):469–485, Feb. 2024. doi: 10.1177/87552930231217003.
- Varpasuo, P. The update of the seismic probabilistic site hazard assessment for Loviisa NPP in Finland. In *Proceedings of the 14th World Conference on Earthquake Engineering, 12-17 October*, Beijing, China, 2008.
- Villani, M., Lubkowski, Z., Free, M., Musson, R. M. W., Polidoro, B., McCully, R., Koskosidi, A., Oakman, C., Courtney, T., and Walsh, M. A probabilistic seismic hazard assessment for Wylfa Newydd, a new nuclear site in the United Kingdom. *Bulletin of Earthquake Engineering*, 18(9):4061–4089, May 2020. doi: 10.1007/s10518-020-00862-8.
- Visakorpi, V. *Sensitivity analysis of a seismic hazard assessment for a Finnish nuclear power plant*. Master of Science in Technology, School of Science, Aalto University, 2022.
- Weston Geophysical Corporation. Probabilistic seismic hazard assessment Gentilly 2. AECB project 2.311.1, Atomic Energy Control Board, Canada, 1996.
- Wong, I., Pezzopane, S., Abrahamson, N., Green, R., Sun, J., and Quittmeyer, R. A preliminary assessment of earthquake ground shaking hazard at Yucca Mountain, Nevada and implications to the Las Vegas region. Conference on Seismic Hazard in the Las Vegas Region, Las Vegas, USA, 1996.
- Wong, I., Silva, W., Olig, S., Dober, M., Gregor, N., Gardner, J., Lewis, C., Terra, F., Zachariassen, J., Stokoe, K., Thomas, P., and Upadhyaya, S. Update of the probabilistic seismic hazard analysis and development of seismic design ground motions at the Los Alamos National Laboratory. Final report, URS Corporation, Oakland, USA, 2007. Job No. 24342433.

The article *Comparison of epistemic uncertainty captured in probabilistic seismic hazard assessments for critical infrastructure* © 2026 by Guillermo Aldama-Bustos is licensed under CC BY 4.0.