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

*R: "We thank the two anonymous reviewers for their constructive and detailed comments on our manuscript. We have made most of the changes they recommend. In the following sections we respond to each of their comments in turn. In addition, we have made some editorial changes based on our own internal review."*

### Reviewer C

**Summary:**

The article provides quantitative comparisons of epistemic uncertainty from 30 global PSHAs completed for a range of SSHAC levels. The results in the article provide a benchmark against which site-specific PSHAs can be compared. For example, if the epistemic uncertainty in a site-specific PSHA is significantly lower than the average from these 30 studies, this article provides a basis for reviewers to request justification for less uncertainty or a more thorough evaluation of uncertainty. This article is thus an important contribution to the professional literature. Furthermore, publishing this article in an open-access journal like Seismica assures the data set and results are accessible to more practitioners and reviewers, which I believe will improve the state-of-practice in incorporating and evaluating epistemic uncertainty in PSHA.

The article is generally suitable for publication with minor revisions. There are two major technical comments that I believe should be addressed prior to publication. Additionally, addressing the minor technical comments and presentation/editorial comments will improve the article. While the article is generally well written and clear, there is a propensity of isolated 2-3 sentence paragraphs lacking topic sentences throughout Section 4 that make the paper difficult to follow occasionally.

*R: "We appreciate the reviewer's comments, which certainly contribute towards an improved paper. We provide our responses below, indicating actions taken where needed, to each individual comment."*

**Major technical comments:**

1. A major shortcoming of this article is that comparisons between active and stable crustal settings are not provided. Providing benchmarks for SCRs based on the large data set herein would be valuable to practitioners. Furthermore, if no difference is found between active and stable regions (which is implied in the Discussion without basis), then that differs from earlier results (e.g., Douglas et al, 2014) and is important to directly demonstrate (similar to the surprising finding that uncertainty does not vary with SSHAC level). The statement that assigning active/stable labels is not feasible because it is too subjective (Line 486) is not accurate; while there is some ambiguity, there is precedent in the professional literature, and ultimately there is no issue if the assignment basis is clearly defined in the article.

*R: "A column has been added to Table 1 classifying the sites as being in active regions, stable regions, or near the boundary between them. The statement stating that it is not possible to classify sites has been removed. Additionally, Section 4.5 has been updated to discuss the variation of uncertainty by tectonic region rather than by ground-motion intensity. The original discussion on the uncertainty variation with ground-motion intensity was not retained due to the* *limitations associated with such an assessment, as highlighted by the various reviewer comments on this topic."*

2. Please provide an example of how $\sigma_{haz}$ is computed, such as in a short supplement. For example, I could not reproduce the $\sigma_{haz}$ (PGA) results for the first row in the supplemental CSV file. Reproducibility is important because the key feature of this article is that hazard analysts should be comparing the epistemic uncertainty in their PSHAs to the results herein. Please also state the units of $\sigma_{haz}$ (e.g., ln units).

*R: "A footnote has been added to the definition of $\sigma_{haz}$ to state how this metric has been computed."*

**Minor technical comments:**

3. The Idaho National Laboratory (INL) SSHAC Level 3 provides results for rock and soil for the MFC site (see Section 15.5.1 for fractiles for surface condition). Please comment on why the surface condition was not evaluated for INL (e.g., Figure 5).

*R: "Thanks for alerting us to this oversight. We were not aware of these data when we drew the figures. We believe that adding the results for surface conditions to the figures would make them more cluttered and would add only little value. Therefore, we have decided not to include them."*

4. It appears the statistics are computed with respect to AFE throughout the article, but it would be good to explicitly state this (perhaps near Line 96).

*R: "A sentence has been added to clarify this near the beginning of Section 2."*

5. Line ~265: The distribution of $\sigma_{haz}$ across ARPs is surprising. For example, the standard deviation of $\sigma_{haz}$ decreases at longer ARPs (e.g., 100k year, 0.2 s and 1.0 s). Typically, the mean hazard curve drifts to higher fractiles at longer ARPs, which suggests the spread across the fractiles is wider at longer ARPs. This pattern should manifest in $\sigma_{haz}$. Please provide additional discussion on this.

*R: "Although the mean hazard curve drifts to higher fractiles at longer ARPs, the widths of the fractiles in log space along the ground-motion intensity axis (x axis), which is what $\sigma_{haz}$ is measuring, does not significantly change. This can be seen from plots of hazard curves on log-log axes (e.g. Figure 10 of Villani et al., 2020). A sentence has been added to this section to discuss this."*

6. Line ~270: Are the ground motion models applicable to the very high $V_{S30}$ values for the rock sites (Table 1), or do the various studies incorporate additional/large epistemic uncertainty to account for this? If so, is this a factor in why $\sigma_{haz}$ is not lower for rock sites?

*R: "Overall, ground-motion models are poorly constrained to hard rock conditions, i.e. VS30 > ~1,000 m/s; however, this is typically addressed in PSHAs for nuclear facilities through the application of host-to-target adjustment factors (e.g. shallow crustal VS profile and the high frequency attenuation – kappa-0) and their associated epistemic uncertainties. Nevertheless, this does not explain why $\sigma_{haz}$ is not consistently higher at the bedrock horizon. The initial expectation was that uncertainties in the site response analysis will be additive to those observed at bedrock level; hence, uncertainty levels at the target horizon were expected to be systematically higher. As discussed in the text, the reasons for this behaviour are unclear at this* *stage and additional research would be required. Such work is beyond the objectives of the current study."*

7. Lines 312-314: Please provide a reference for the statement that higher uncertainties are associated with lower intensity. In my experience, uncertainties are wider at longer ARPs (and therefore higher intensities).

*R: "This sentence has been deleted, and the entire section updated as discussed in our response to Comment #1."*

8. Line 317: ARP is not a suitable proxy for ground motion intensity when stable and active settings are mixed. Stable regions have lower activity rates; this pushes the hazard curve down, leading to lower intensities at a given ARP compared to the same curve with a higher rate. Figure 5 shows a spread of 1.5 to almost 2 orders of magnitude along the x-axis.

*R: "This sentence has been deleted, and the entire section updated as discussed in our response to Comment #1."*

**Editorial comments:**

9. Line 48: is this really a problem? Are methods, results, etc. "blindly" applied in SSHAC studies? An alternative wording could be considered, especially for the non-technical (public) summary, as this could be taken out-of-context by public/political nuclear power interest groups.

*R: "We have modified this sentence to avoid misunderstanding."*

10. It would be very helpful to readers to move the bullet points from Line 185 up, to the end of Line 95. The discussion in Lines 96+ assumes some knowledge of the metrics.

*R: "The various metrics are defined as they are introduced in Section 2 (Previous Studies). Therefore, we do not agree that moving these bullet points, which are a useful reminder to the reader, would improve clarity."*

11. The authors could consider adding a table comparing the data set sizes (and perhaps other parameters, such as AFEs, IMs, etc.) among previous studies and the current study near Line 135. A key feature of the present work is the size of the data set, and a table is a nice way to highlight this.

*R: "Key information on previous studies is included in Section 2 when the studies are introduced. Therefore, we do not think that including an additional table is necessary, especially given the length of the manuscript."*

12. In Section 4.1 (Lines ~195), a discussion on comparisons between R and Unc is missing. The authors could consider adding a correlation matrix to supplement the discussion. It could be qualitative (e.g., low/medium/high) to be broadly applicable across the range of comparisons.

*R: "As Unc and $\sigma_{haz}$ are highly correlated (coefficient of determination = 0.97), the correlation between R and Unc will be like that already shown in the bottom left-hand corner of Figure 2. A sentence to mention this has been added to the text."*

13. Lines 258 to 259: are these needed? Appears redundant with Lines 250-257.

*R: "This paragraph has been revised to remove repetition."*

14. Lines 267 to 281: The authors in initially state there is no difference between reference horizons, then state otherwise in Line 274. Please clarify the writing.

*R: "This paragraph has been revised to avoid confusion concerning the Beznau results."*

15. Figure 5: Please note in the figure caption that sites #a and #b denote hard/soft.

*R: "We have added this information to the caption."*

16. Figure 5 is difficult to read; please consider replacing or adding a figure that shows the hard-to-soft $\sigma_{haz}$ ratio to support the statements at Lines ~270.

*R: "Figure 5 has been updated to shown target-to-bedrock ratios which makes easier to observe when $\sigma_{haz}$ at the target horizon is lower than at the bedrock level (i.e. target-to-bedrock ratio <1.0). Text in this section has been amended accordingly."*

17. Please consider renaming Section 4.4 to "Comparisons between intensity measures" for consistency with the other subsections.

*R: "We have renamed this section."*

18. Line 536, typo: "... aid the hazard analyst *in prioritising* inputs ..."

*R: "We have revised this sentence."*

19. Line 655: The link to the PEGASOS project does not work.

*R: "We have removed the links and email address in the PEGASOS and PEGASOS Refinement Project references."*

### Reviewer F

This study provides a valuable summary of the range of epistemic uncertainty from 30 SSHAC-level assessments. The summary provides a benchmark for epistemic uncertainty to compare the reasonableness of future studies. Assessing correlations across different uncertainty metrics helps identify reasonable metrics.

**Main comments:**

- The key issue is that epistemic uncertainty is not an absolute measure; rather, it is defined for a particular level of simplification, which depends on the modeled and unmodeled behavior. A citation by Liou and Abrahamson (2025), "Framework for aleatory variability and epistemic uncertainty for the ground-motion characterization based on the level of simplification," supports this.

*R: "Liou and Abrahamson (2025) focus particularly on the distinction between the epistemic uncertainty and aleatory variability in ground-motion model, which is one of the elements of the PSHA. In this study, we focused on the epistemic uncertainty captured by the entire PSHA which includes the uncertainty model in both the seismic source characterisation and the ground motion characterisation models. We have added two sentences clarifying this at the start of Section 2."*

- Information is missing on the level of simplification for each SSHAC study. A table is needed to describe the methods of each study (e.g., regionally adjusted ergodic GMM, partially non-ergodic single station sigma) because the term “site-specific” is vague and used across different types of studies.

*R: “We considered the level of effort to produce the recommended table is beyond the scope of this study as it would require a detailed understanding of the models and assumptions behind the development of the various models inputting into the individual PSHAs. Additionally, in many cases not enough information is available in the public domain to complete the table (e.g. in many cases only a journal or conference paper is available where only hazard results are summarised). We have added some sentences to the end of Section 3 on this.”*

- The epistemic uncertainty should be calculated for studies of the same level of simplification with similar aleatory variability.

*R: “We assume the reviewer is referring to the levels of simplification of the ground-motion models as defined by Liou and Abrahamson (2025), i.e. ergodic, partially ergodic or nonergodic. As mentioned in our previous response, this distinction is not possible given that information is not available for many of the studies; however, it is expected that all studies used either ergodic or partially ergodic GMMs. As discussed by Liou and Abrahamson (2025), the use of nonergodic models is not yet standard practice. To our knowledge, nonergodic GMMs have not yet being implemented as part of PSHAs for nuclear facilities. We have added some sentences to the end of Section 3 on this.”*

- Information is needed on how the epistemic uncertainty was calculated for each study.

*R: “As part of the current study, uncertainties were calculated based on the spread of the hazard curves using the alternative metrics as discussed in Section 2 – Previous studies, and summarised in Section 4 – Results. However, we believe the reviewer is referring to how the epistemic uncertainty was assessed within each of the studies considered in this paper. As discussed in Section 2, it is not yet common practice in PSHA to explicitly quantify the level of epistemic uncertainty captured by the study. This is the motivation for the current and similar previous studies.”*

- A few sentences are needed to clarify that: Conceptually, areas with less data should have greater epistemic uncertainty. Often, epistemic uncertainty is underestimated for areas with less data. The underestimation can result from comparing models of different levels of simplification within a single logic tree, leading to inconsistent epistemic uncertainty. There is a tendency to use more complex models with additional physical effects in areas with more data. This leads to an unequal comparison of epistemic uncertainty between the two areas because the models used differ in their levels of simplification. There is a need for a minimum level of epistemic uncertainty in areas with less data, which is reduced with more data.

*R: “This is now discussed in the updated Section 4.5.”*

- If, after the models are correctly separated and epistemic uncertainty is calculated for models at one level of simplification, the results still show less epistemic uncertainty in areas with less data, a note should be added that this supports the idea that epistemic uncertainty is underestimated in areas with less data.

*R: “As discussed in our response to previous comments, it is not possible to classify studies by the level of simplification as not enough information is available.”*

- Adding a few lines to the introduction about the purpose of identifying correlations among uncertainty measures will better define the study's purpose. I agree that sigma haz is the most direct; however, the point of other epistemic uncertainty metrics is that they provide possibilities for less time-intensive checks.

*R: "Some sentences have been added to the introduction to define the study's purpose better."*

- The fractile of the mean hazard is another metric to check that can strengthen the study and provide helpful comparisons across studies.

*R: "Although this would potentially be a useful metric it has not been previously proposed for use in this context and it requires many fractiles to have been reported for it to be accurately computed. Finally, the manuscript is already long without considering an additional metric. A paragraph has been added at the end of Section 4.1 to mention this metric and a sentence in the conclusions added to mention its consideration in a future study."*

Recommendation: Resubmit for Review

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

*R: "We thank the two anonymous reviewers for their additional comments on the updated manuscript. We have addressed their comments as discussed below."*

### Reviewer A

**Summary:**

The authors have addressed the major technical comments and most of the minor comments from my initial review. I have two remaining comments:

(1) Initial comment: "The Idaho National Laboratory (INL) SSHAC Level 3 provides results for rock and soil for the MFC site (see Section 15.5.1 for fractiles for surface condition). Please comment on why the surface condition was not evaluated for INL (e.g., Figure 5)."

> Response: "Thanks for alerting us to this oversight. We were not aware of these data when we drew the figures. We believe that adding the results for surface conditions to the figures would make them more cluttered and would add only little value. Therefore, we have decided not to include them."

I disagree that including the MFC study would add little value and clutter Figure 5. The authors emphasize that the lower σhaz values for soil sites is unexpected, so adding another site to the evaluation would indeed add value. Additionally, there are only six data points on Figure 5 (3 rock, 3 soil), whereas the other figures have more than a dozen data points, so the concern for clutter is weak.

Furthermore, I would like to gently note that one of the authors of this article is listed as a peer reviewer for a BSSA publication on the INL MFC host-to-target adjustments (Boore et al., 2022; DOI 10.1785/0120220056), so its exclusion in this study remains surprising. At minimum, a footnote should be added to Table 1 indicating that surface condition results for the MFC site where not evaluated; however, I encourage the authors to perform the evaluation.

*R: "Results for the INL SSHAC Level 3 study at the surface level have now been included in the study. This resulted on an update of all figures in the manuscript, except Figure 2, as well as Table 1. Target-to-bedrock ratios for this study are consistent with ratios derived for other studies. This is now discussed in the main text."*

(2) As for why the fractile spread (σhaz) is smaller for soil versus rock sites (e.g., lines 612-631 in track changes file), it might be that rock site response remains elastic, whereas soil response is nonlinear. At high frequencies, soil nonlinearity attenuates/suppresses strong motions, so there is less variability (low end doesn't really change, but high end is reduced). Rock sites remain essentially elastic, so there is no/minimal damping to suppress strong motions, leading to more variability. In other words, there is less variability in the transfer function at high frequencies for soil sites than rock sites. Including an evaluation of the INL MFC site would be valuable to see if the results follow this pattern.

*R: "The discussion on this topic in Section 4.3, immediately above Figure 4, has been expanded to incorporate the reviewer's view."*

## Reviewer B

For both the source characterization and ground motion models, the same concept applies that the size of the epistemic uncertainty depends on the level of simplification. The paper should acknowledge that there's no absolute value for epistemic uncertainty, as its size depends on the model's complexity. It's not a constant for all studies. For example, including renewal models in the seismic source characterization would increase the epistemic uncertainty.

*R: "The discussion in Section 5, Rows 503-506, has been expanded to indicate that there is no 'right' or 'absolute' level of uncertainty for a given study, and that its size depends on various project-specific factors, as suggested by the reviewer."*

Epistemic uncertainty tends to grow over time moving toward more complicated models. It's valuable to state the range of epistemic uncertainty; however, if there's a new study with more complicated models, then the range of epistemic uncertainty from simpler studies should not give a target range of the epistemic uncertainty.

*R: "Although there is logic on the reviewer's statement, particularly given the recent tendency to explicitly model epistemic uncertainty, previously treated as part of the aleatory variability, in more complex models (e.g. through the adoption of single-station sigma models), this trend is not supported by the current data. As discussed in Section 4.6, there is presently no clear or discernible increase, or decrease, of $\sigma_{haz}$ with the date of the study. Moreover, when considering sites for which multiple studies are available, the expected behaviour is that epistemic uncertainty should reduce over time as additional data becomes available and uncertainties become better constrained. No changes to the manuscript were made as result of this comment."*

Nonergodic GMMs have recently been implemented on engineering projects.

*R: "Text in Section 3, Rows 203-204, was updated to indicate this."*

In high-end seismic hazard studies, hazard fractiles are quantified. It is common practice to include logic trees and show the range of the hazard.

*R: "This comment appears to relate to our recommendation that* 'the total epistemic uncertainty should be routinely estimated and reported in the PSHA documentation'. *Although we agree with the reviewer that the calculation of hazard fractiles, along with the reporting of tornado plots illustrating the range of hazard across the individual nodes in the logic tree, is common practice, our recommendation stands. The estimation of the 'total' uncertainty is not yet routinely undertaken, nor is the comparison of this total uncertainty with estimates from other studies. No changes to the manuscript were made as result of this comment."*