Accepted for publication in Earthquake Spectra 22.09.2020

Capturing Epistemic Uncertainty in Site Response

Adrian Rodriguez-Marek,[a,*)] M.EERI, Julian J Bommer,[b)] M.EERI
Robert R Youngs,[c)] M.EERI, Maria J. Crespo,[d)] M.EERI, Peter J Stafford,[b)] M.EERI
and Mahdi Bahrampouri,[a)] M.EERI

The incorporation of local amplification factors determined through site response analyses has become standard practice in site-specific probabilistic seismic hazard analysis (PSHA). Another indispensable feature of the current state-of-practice in site-specific PSHA is the identification and quantification of all epistemic uncertainties that influence the final hazard estimates. Consequently, logic trees are constructed not only for seismic source characteristics and ground-motion models (GMMs) but also for the site amplification factors, the latter generally characterized by branches for alternative shear-wave velocity (V_S) profiles. However, in the same way that branch weights on alternative GMMs can give rise to unintentionally narrow distributions of predicted ground-motion amplitudes, the distribution of amplification factors obtained from a small number of weighted V_S profiles will often be quite narrow at some oscillator frequencies. We propose an alternative approach to capturing epistemic uncertainty in site response in order to avoid such unintentionally constricted distributions of amplification factors using more complete logic-trees for site response analyses. Nodes are included for all the factors that influence the calculated amplification factors, which may include shallow V_S profiles, deeper V_S profiles, depth of impedance contrasts, low-strain soil damping, and choice of modulus reduction and damping curves. Site response analyses are then executed for all branch combinations to generate a large number

[a)] Civil and Env. Eng. Dept., Virginia Tech, Blacksburg, VA 24061, USA. Email: adriarnm@vt.edu
[b)] Civil and Env. Eng. Dept., Imperial College London, South Kensington campus, London SW7 2AZ, UK
[c)] Wood Environment and Infrastructure Solutions, Oakland, CA 94612, USA
[d)] Principia Ingenieros Consultores SA, Velázquez 94, Madrid 28006, Spain. Also with Universidad Politécnica de Madrid – E.T.S.I. Minas y Energía, Ríos Rosas, 21, 28003 Madrid, Spain.

of frequency-dependent amplification factors. Finally, these are re-sampled as a discrete distribution with enough branches to capture the underlying distribution of amplification factors (AFs). While this approach improves the representation of epistemic uncertainty in the dynamic site response characteristics, modeling uncertainty in the AFs is not automatically captured in this way, for which reason it is also proposed that a minimum level of epistemic uncertainty should be imposed on the final distribution.

INTRODUCTION

Probabilistic seismic hazard analysis (PSHA) has become the standard of practice for determining site-specific earthquake ground-motion characteristics to be considered in the design of major infrastructure and safety-critical facilities. PSHA offers several advantages over deterministic assessments of design ground motions, including the effective integration of all sources of aleatory variability associated with earthquake occurrence patterns and the estimation of ground-motion amplitudes for a given earthquake scenario (i.e., magnitude-distance combination). The state-of-practice in PSHA now also includes the use of logic-trees to capture the epistemic uncertainty in different elements of the seismic source characterization (SSC) and ground-motion characterization (GMC) models.

The GMC model consists of multiple ground-motion models (GMMs) for the median and standard deviation. A key element of the GMMs in any site-specific PSHA is the influence of local site amplification effects resulting from the structure and dynamic properties of the near-surface layers at the site. Site amplification factors are routinely included in the GMMs, characterizing the frequency-dependent amplification factors as a function of proxy parameters such as V_{S30} (the time-averaged V_S over the uppermost 30 m) and Z_X (the depth at which a V_S of X is reached, typical values of X being 1.0 or 2.5 km/s). Such factors yield generic amplification factors (AFs) that will reflect an average site response behavior over the stations and recordings used in the derivation of the GMM. Since sites with the same V_{S30} may have both very different layering in the top 30 meters and entirely different V_S profiles below 30 m, site-specific AFs may be appreciably different from the generic AFs in GMMs (e.g., Papaspiliou et al., 2012). Consequently, it is generally not considered appropriate to model local amplification characteristics using GMMs in site-specific seismic hazard studies, especially since the dynamic site characteristics are among the only features of the GMM that

can be determined without requiring the occurrence of new earthquakes in the site region. Site-specific PSHA studies therefore routinely include *in situ* measurements of the V_S profiles and include AFs determined from site response analyses (SRAs) based on these profiles. Procedures for characterizing sites and for incorporating the AFs are now well established and practical implementations have been documented (e.g., Bazzurro and Cornell, 2004a; Bazzurro and Cornell, 2004b; Rodriguez-Marek et al., 2014; Stewart et al., 2014; Lessi-Cheimariou et al., 2019; Tromans et al., 2019). However, the procedures for ensuring that the epistemic uncertainty in the AFs is captured effectively are still evolving, and the purpose of this paper is to highlight pitfalls that are encountered in current practice and to propose procedures that can help to avoid these problems.

The paper begins with a discussion of epistemic uncertainty in PSHA and the use of logic-trees as the ubiquitous tool to incorporate epistemic uncertainty into hazard calculations, with a particular focus on how practice has evolved in recent years to ensure that logic-trees meet their intended objectives. This is followed by a discussion of how the use of logic-trees has evolved in site response analyses. We then present the pitfalls in current approaches to constructing logic-trees for site response analyses and propose alternative procedures designed to avoid these issues and to better capture the epistemic uncertainty. In the penultimate section of the paper we discuss considerations of model error and introduce an approach to account for it within the proposed methodology. The paper closes with a brief discussion of the implications for PSHA practice.

EPISTEMIC UNCERTAINTY AND LOGIC TREES IN PSHA

As noted in the Introduction, the original motivation underlying the development of PSHA was to treat the location and magnitude of future earthquakes as random variables (Cornell, 1968). The residuals in the ground-motion prediction model for any given magnitude-distance combination was subsequently added as another random variable in the hazard integral (Bommer and Abrahamson, 2006; McGuire, 2008).

As the practice of PSHA evolved, it was also recognized that there are several decisions made in the construction of both the SSC and GMC models that are subject to appreciable uncertainty as a result of incomplete or ambiguous information. In most cases, this uncertainty is related to the fact that different interpretations, all scientifically viable, can be drawn from the available data, and also to the need to extrapolate beyond the limits of the data to cover all

of the scenarios that must be considered in the hazard integrals. Being related to lack of knowledge, this uncertainty is classified as epistemic. More than 35 years ago, logic-trees were adopted as a tool to represent epistemic uncertainty and facilitate its incorporation into PSHA calculations. The first implementation of a PSHA logic-tree by Kulkarni et al. (1984) only considered uncertainties in the SSC model, but in subsequent studies nodes were also added to logic-trees for the uncertainties in the GMC models.

The purpose of a logic-tree has been defined as capturing the center, the body, and the range of technically defensible interpretations of the available data, methods and models, which is generally abbreviated as capturing the CBR of TDI (USNRC, 2018). For GMC logic-trees, in particular, practitioners came to the realization that while weights are generally assigned to alternative GMMs, the objective is to define the center, body and range of possible ground-motion amplitudes for different earthquake scenarios. The relationship between the weights assigned to different models and the resulting distribution is not easily visualized and the distributions may be found not to reflect the actual level of knowledge regarding predicted ground motions for the region and site in question. When a small number of GMMs populate the GMC logic-tree, there can be magnitude-distance combinations for which there is almost no spread among the predicted accelerations, but at the same time using a large number of GMMs—which will often have been derived from similar or overlapping databases—the resulting distribution can often be quite narrow (e.g., Bommer, 2012). This realization has prompted a move to the construction of GMC logic-trees using what is referred to as the backbone approach, in which the branches are populated with scaled or adjusted versions of a single GMM (Atkinson et al., 2014) which in turn has led to re-consideration of the criteria for selecting GMMs (Bommer and Stafford, 2020). In this way, the relationship between the logic-tree branch weights and the resulting distribution of predicted ground-motion amplitudes becomes much more transparent, and unintentionally narrow distributions are avoided.

The difference between how the logic-tree is constructed and the actual intended outcome also applies to SSC models. Weights may be assigned to alternative source zone boundaries, recurrence models and estimates of maximum magnitude, but the hazard estimates are ultimately driven by the distribution of recurrence rates of earthquakes of different magnitudes and the distribution of distances from the site at which they occur. Stromeyer and Grünthal (2015) proposed a procedure for re-sampling the recurrence models developed in PSHA studies, in a manner that could be compared with backbone GMM approach.

As is explained below, the purpose of this paper is to extend the backbone approach—which focuses attention of the final distributions of variables that drive the hazard estimates rather than distributions of the component models—to site response analyses conducted for PSHA.

ALEATORY VARIABILITY AND EPISTEMIC UNCERTAINTY IN SITE RESPONSE

Opinions differ regarding the importance of separating aleatory variability and epistemic uncertainty in PSHA, and it is recognized that in terms of calculating the mean hazard, the distinction is irrelevant provided the total uncertainty is correctly quantified. However, when the fractiles of the hazard are also required—which is often the case in the seismic assessment of critical structures such as nuclear power plants—then the distinction is important, and the spread of the hazard fractiles should reflect the degree of epistemic uncertainty (while the aleatory variability influences the shape of mean and fractile hazard curves). The correct separation of aleatory variability and epistemic uncertainty is also important because the latter can, at least in theory, be reduced through the acquisition of additional data and therefore identifying the reducible component of uncertainty can inform decisions regarding new data collection.

In ground-motion modelling, the understanding of how the total uncertainty can be partitioned has evolved considerably. The standard deviation of the logarithmic residuals in a GMM, usually referred to as sigma, is treated as aleatory variability, although it may be more appropriately referred to as apparent aleatory variability with respect to the model. Epistemic uncertainty in a GMC model can be identified both in the choice of median GMMs and in the associated sigmas. An important development, which has been parallel to advances in the incorporation of site response into PSHA, is that sigma values have generally included systematic as well as random effects. Specifically, at any given location, there are repeatable site amplification effects, which can be modeled explicitly. When local site effects are modeled explicitly, the station-to-station component of the within-event variability must be removed from sigma, resulting in what is referred to as single-station or partially-nonergodic sigma (Atkinson, 2006; Rodriguez-Marek et al., 2013). The site-specific amplification effects could in theory be perfectly constrained by recordings of a very large number of earthquakes—of different magnitudes and located at different distance and azimuths from the site—at the site,

but in practice such situations very rarely exist, especially at the proposed locations for new safety-critical structures. In a few cases, existing ground motion data at or near the site can help constrain the linear component of site response (Stewart et al. 2019). However, in most cases the site-specific amplification effects are modeled through site response analyses. If the removal of the site-to-site component of variability from the ergodic sigma values is to be justified, then the epistemic uncertainty in the estimated site amplification effects must be included in the hazard estimates.

For many years, site response calculations were performed modeling all of the variability and uncertainty in the V_S profiles as random variability (i.e., Rodriguez-Marek et al., 2014). A single best estimate V_S profile would be developed, usually based on *in situ* measurements, and a standard deviation on $\ln(V_S)$ defined (Figure 1b). The value of $\sigma_{\ln(Vs)}$ would ideally be calculated directly from multiple measurements at the site (Figure 1a), either at different locations and/or using different measurement techniques. The important observation from the perspective of this article is that in such an approach, all lateral variations and uncertainty in the V_S profile are treated as aleatory variability (Toro, 1995, Teague et al., 2018; Passeri et al., 2020). Using the best estimate profile and the assigned variability, randomized V_S profiles are generated—invoking appropriate rules for the layer-to-layer correlations—that may include randomization of the layer thicknesses as well (Figure 1c). Site response analyses are performed for all of the randomized profiles, enabling calculation of the site AFs and its associated variability (Figure 1d). In order to reduce the total number of calculations required, Bahrampouri et al. (2019) proposed a procedure to enable calculation of the contribution of the uncertainty in the modulus reduction and damping (MRD) functions to the overall uncertainty in the resulting AFs.

In recent years, advances have been made in separating epistemic uncertainty from aleatory variability in site response, for consistency with the partitioning of uncertainty in the other elements of site-specific PSHA studies. The distinction between the two is less straightforward than in other aspects of ground-motion modeling for several reasons, including the fact that while hazard is calculated at a single point, the motions that will affect the structure being designed will be influenced by the dynamic response of the V_S profiles at multiple points across the footprint of structure's foundation. Nonetheless, there is a rapidly growing consensus that uncertainty in measured V_S profiles is largely epistemic, especially when it reflects differences arising from the use of multiple measurement techniques [such as down-hole, cross-hole, PS

logging, Multichannel Analysis of Surface Waves (MASW), etc.] or when it reflects measurement errors within a single measurement technique (Passeri et al., 2019). The state-of-practice has consequently moved to including alternative V_S profiles within a logic-tree formulation in order to explicitly represent the epistemic uncertainty and carry it through to the AFs and the hazard estimates. The alternative profiles and their relative weightings can be inferred from in situ measurements and guidelines have been issued for establishing appropriate upper and lower branch profiles for cases where site data is limited or when V_S profiles are inferred rather than measured (EPRI, 2013). These profiles are still randomized for the site response calculations in order to account for aleatory (i.e. spatial) variability and to reduce the influence of resonances created by the artificially sharp impedance contrasts that result from a model with idealized horizontal layers with abrupt changes in velocity across layer boundaries, which could lead to bias in predicted amplification factors near the resonance period (Kaklamanos et al., 2020).

The separation of epistemic uncertainty and aleatory variability in site response models is a significant step forward in improving the state-of-practice in site-specific PSHA. However, as we discuss in the next section, in the same way that weights on alternative GMMs will often not result in the intended distribution of ground-motion amplitudes, weights on alternative V_S profiles may not yield the intended distribution of AFs.

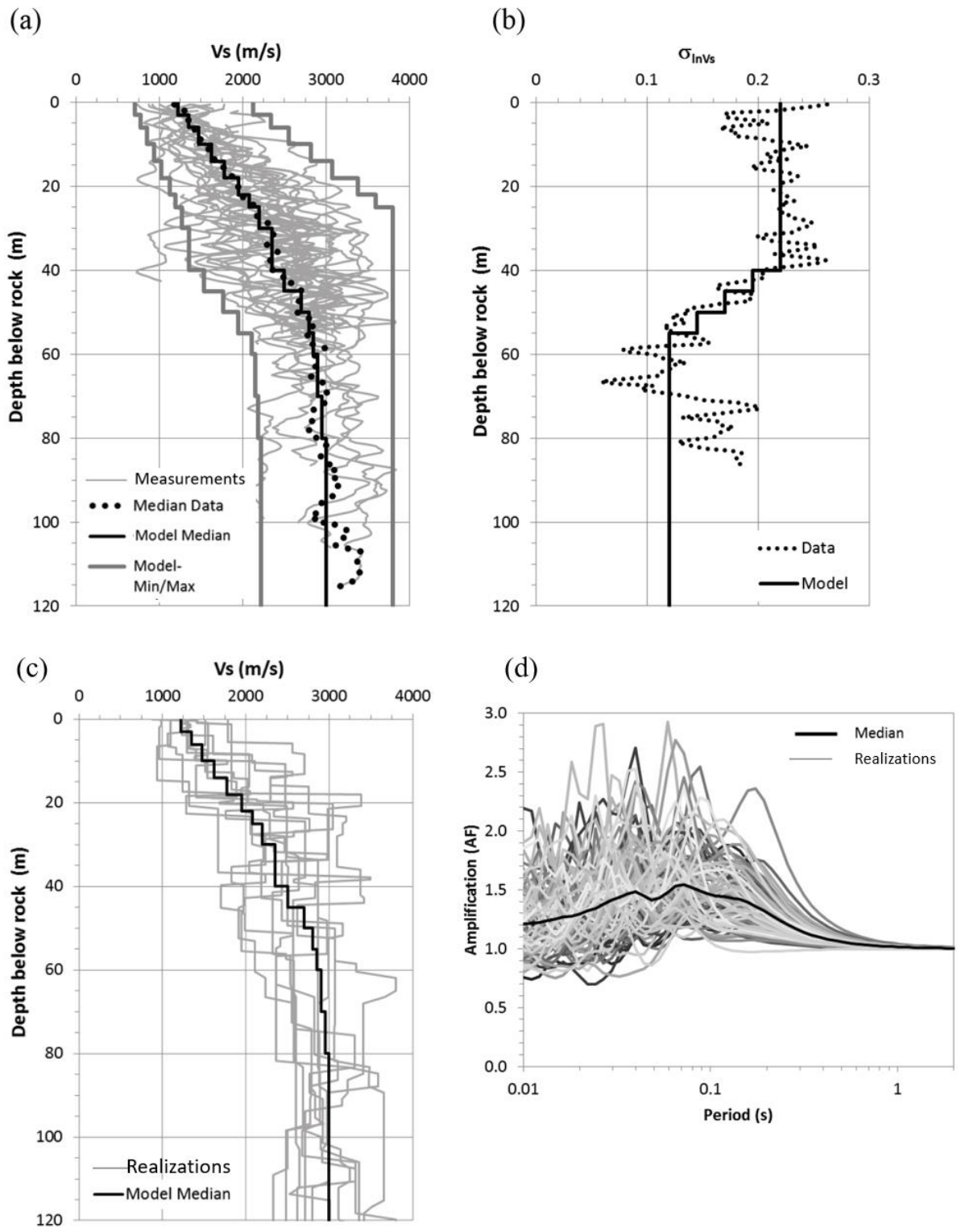


Figure 1. (a) Development of a best-estimate (median) V_S profile and bounding range from multiple measurements at a site, and (b) the associated variability in the profiles; (c) randomized V_S profiles generated for site response analyses, and (d) the resulting median AF from individual realizations of the V_S profiles computed using an equivalent linear site response analysis (from Rodriguez-Marek et al., 2014).

SITE RESPONSE LOGIC-TREES AND DISTRIBUTIONS OF AMPLIFICATION FACTORS

Figure 2 shows a fictitious, yet realistic, site profile and three alternative interpretations of the V_S profile, which are assigned weights reflecting the relative degree of confidence in each profile being the best representation of the site. The profile consists of an upper layer of stiff sand, with V_S increasing with depth, lying over rock which in turn overlies basement rock. The alternative profiles reflect uncertainties in the V_S, in the transition from the sand to the rock, and in the depth to basement rock. Figure 3 shows the AFs calculated from these three profiles using a linear visco-elastic formulation with the code Pysra (Kottke, 2019). Observe that at certain response frequencies, the range of values covered by the calculated AFs is very narrow. In particular, at certain frequencies (indicated with an arrow in Figure 3) the three AFs cross over, with the consequence that the epistemic uncertainty at these response frequencies would be close to zero.

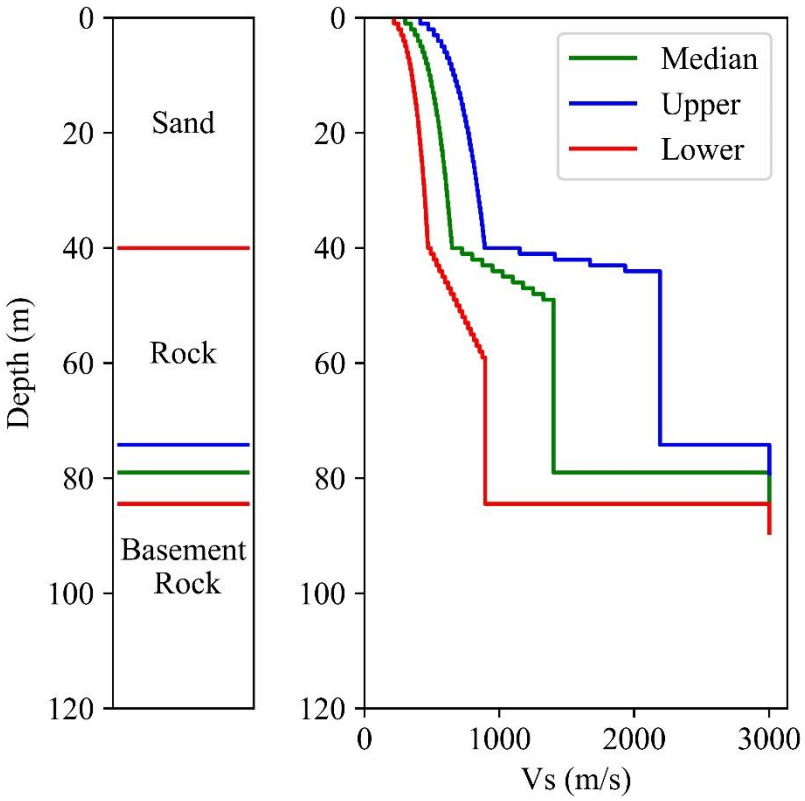


Figure 2. Stratigraphic profiles (left) and best-estimated and alternative upper- and lower-case V_S profiles (right) considered for this study.

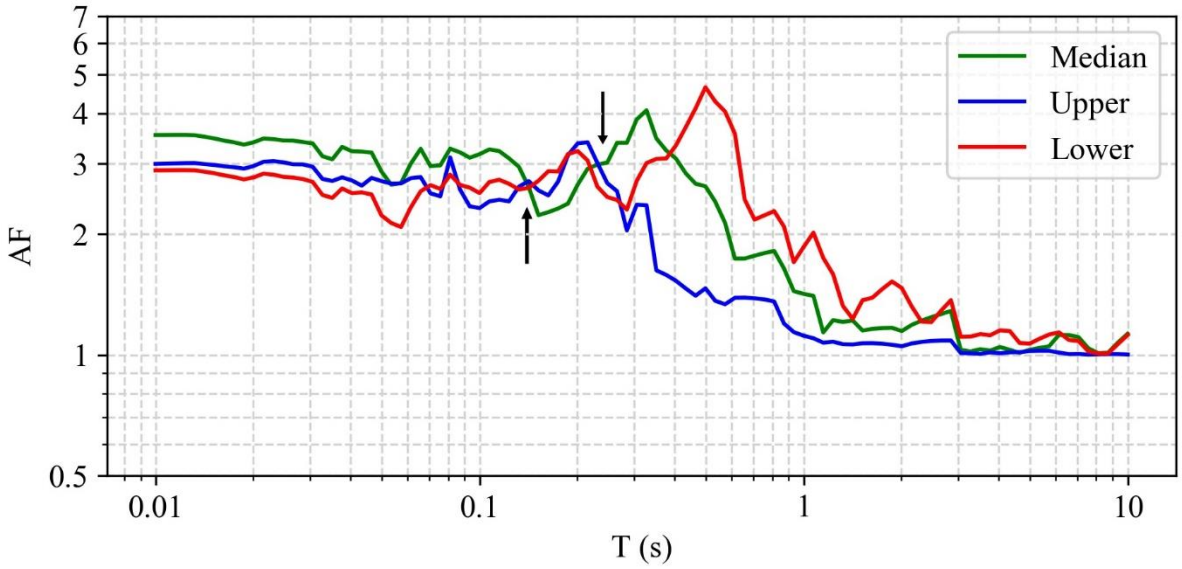


Figure 3. Linear amplification factors corresponding to the three profiles shown in Figure 2. The arrows point to oscillator periods where the three amplification factors cross over and the resulting epistemic uncertainty is low.

The crossing of the AFs in Figure 3 are symptomatic of the problem alluded to earlier: the representation of epistemic uncertainty in V_S profiles (or, more generally, in soil properties) does not map well into the desired spread of AFs. A potential solution to this problem would be to obtain a more refined representation of the epistemic uncertainty in the V_S profiles by capturing the different sources of uncertainties (i.e., V_S of the soil, V_S of the rock, depth to basement, etc.) via multiple profiles. The shortcoming of this approach is that the use of multiple profiles in a PSHA logic-tree results in a large increase in computational effort. The alternative approach proposed herein is to fully represent the uncertainty in site properties through a site response logic tree and then resample the AFs into a manageable number of branches. The paragraphs below illustrate this approach and discuss its advantages for applications in PSHA.

The different sources of uncertainty in the V_S profiles shown in Figure 2 can be represented in a site response logic-tree (Figure 4). The different sources of uncertainty in this logic tree include the uncertainty of V_S in the shallow profile and in the rock layer, uncertainty in the thickness of the rock layer, uncertainty in the thickness of the weathered zone, and uncertainty in the small-strain damping. To illustrate the flexibility of the method, an additional node of

the logic tree was used to represent the possibility of a velocity reversal at a depth of 35 m. Additional sources of uncertainty can include, among others, the V_S of the basement rock, the possible presence of a velocity gradient at the interface with the basement rock and, if large strains are expected, the undrained strength of the soil (Yee et al., 2013). For this example, the alternative choices in the profiles are selected rather arbitrarily. In practice, the different branches of each node of the logic tree would correspond to true epistemic uncertainties. For example, the node for *Soil* V_S can represent alternative measurements of V_S in the shallow profile or alternative base-case V_S profiles that are consistent with *in situ* measurements (Griffiths et al., 2016; Passeri et al., 2020), and the node for the small-strain damping can correspond to alternative estimates of the high-frequency attenuation parameter κ_0. Similarly, the weights on the logic tree in Figure 4 are assigned on simple bases: they represent either uninformed choice (equal weights) or discrete representations of a log-normal distribution. In an actual application, the weights in each branch should be selected to properly capture the center, body and range of defensible interpretations of the existing data. Additional nodes can be accommodated in cases where V_S values are inferred rather than directly measured or to represent uncertainties in other site properties, such as modulus reduction and damping versus strain curves.

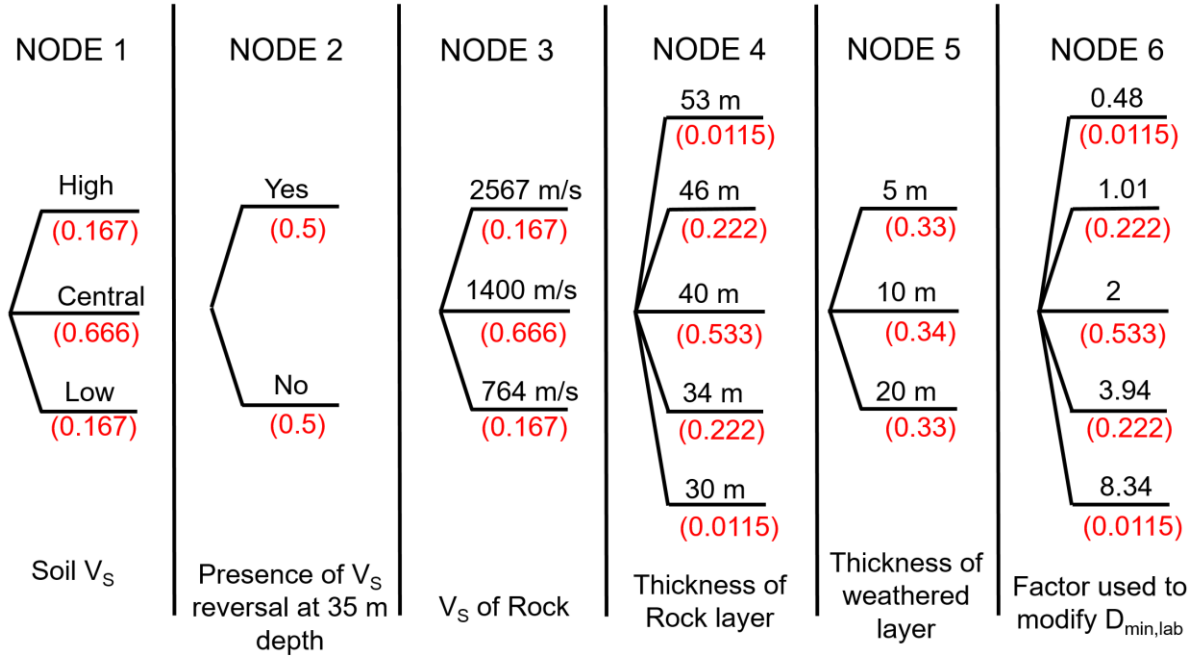


Figure 4. Site response logic tree representing the uncertainties in the profile shown in Figure 2.

To illustrate the application of the method, linear visco-elastic site response analyses were conducted for all of the profiles (a total of 1350 profiles) implied by the site response logic tree in Figure 4. The software Pysra (Kottke, 2019) was used for this purpose. The resulting AFs, plotted with light gray lines, are shown in Figure 5. The AFs corresponding to the profiles illustrated in Figure 2 are superposed on the figure. Not surprisingly, the range of AFs covered by the 1350 logic tree profiles is broader than the AFs captured by the three representative profiles in Figure 2. The corresponding standard deviations of the AFs (in natural log space) of the two set of profiles are also shown in Figure 5. We denote this as σ_{ep} to emphasize that these correspond to epistemic uncertainty. As noted earlier, the uncertainty in AFs that can be captured by a limited number of V_S profiles underestimates the more comprehensive estimate of the uncertainty that is properly captured by the logic tree approach.

The AFs captured by the logic tree are, collectively, a full representation of the epistemic uncertainty in site response. For application in a PSHA logic tree, this uncertainty has to be properly sampled. Within the proposed approach, the sampling can be easily achieved by selecting a certain number of AFs using the sampling described in Miller and Rice (1983). In this procedure, the AFs corresponding to a certain percentile are selected for each period. Each of the AFs is then associated with a weight that reproduces the weighted distribution of the AFs corresponding to the full logic tree. An example of the proposed sampling for five discrete samples is shown in Figure 6. The adequate number of branches needed to fully capture the distribution of AFs should be determined with a parametric study; preliminary results have shown that seven branches are generally sufficient. Alternatively, the uncertainty in AFs could be assumed to follow a certain model (i.e., log-normal) and the sampling can be based on the median and standard deviation of the AFs.

The proposed methodology has several advantages, the most obvious being that it allows for more complete site response logic-trees without a penalty on the computational effort in the PSHA. The more complete logic trees avoid the pitfall of under-representing epistemic uncertainty due to fortuitous crossings of the AFs for a limited number of profiles (e.g., Figure 2). More importantly, the analyst can more appropriately evaluate whether the full epistemic uncertainty in site response is properly captured by evaluating the uncertainty in the parameter that directly enters the hazard calculations (i.e., the AFs). Another important advantage of the proposed approach is that it allows for a proper accounting of minimum epistemic uncertainty,

as will be described in the next section. However, the proposed approach implies an additional effort on the part of the analyst to properly define the logic tree, and the computational cost associated with the site response analyses. Additional complexity in the definition of the logic tree can result if branches are mutually correlated. Similar complexities may arise in the definition of the logic tree for other components of the PSHA process (e.g., Bommer and Scherbaum, 2008).

The use of the site response logic-tree also allows for the types of sensitivity analyses that are common in the analysis of PSHA results. As an example, Figure 7 shows a tornado plot for the AFs for selected periods. The tornado plot shows, for each of the nodes in the site response logic, the mean conditioned on only one of the branches of each node being true. The size of each marker is proportional to the weight of the logic-tree branch. The nodes are sorted such that the node that results in the widest range of AFs is on top. Observe that for this example, the node for low-strain damping (D_{min}) has the highest influence for an oscillator period of 0.1 s, while the uncertainty in the shallow V_S has the highest impact for an oscillator period of 1.0 s. Such analyses can be used to identify additional site characterization that can result in the largest reduction in epistemic uncertainty. The plots can also indicate which nodes of the logic tree could be eliminated on the basis of which elements of uncertainty do not exert a marked influence on the hazard results. Other visualization tools, such as Sammons maps (Scherbaum et al., 2010) can also be used to better understand the resulting distribution of AFs from the proposed approach. The complexity of the site response logic tree can be reduced if the impact of parameter variability can be mapped in a straightforward manner to the uncertainty in the amplification factors (e.g., Bahrampouri et al., 2018).

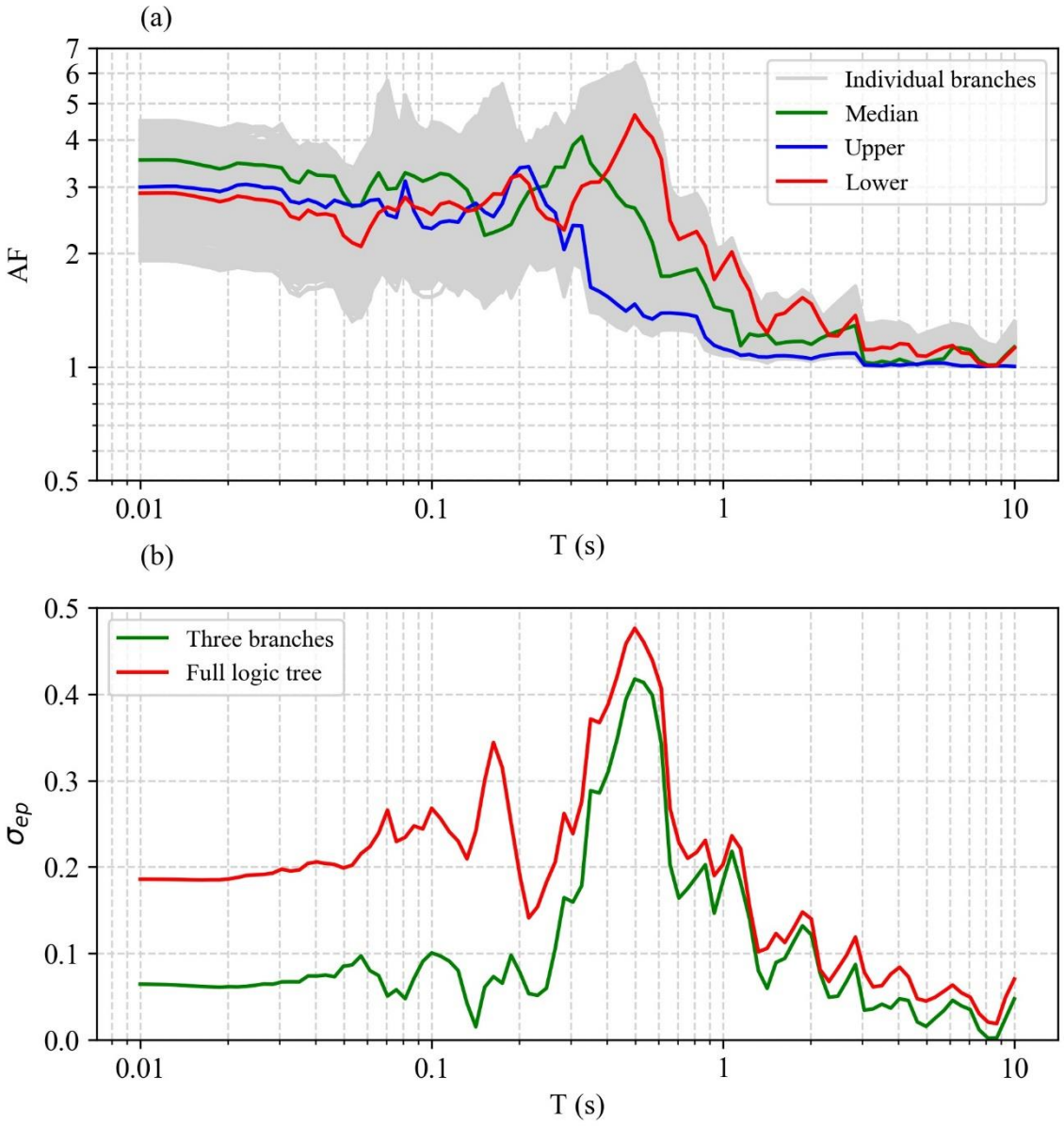


Figure 5. a) Amplification Factors for the profiles implicit in the logic tree in Figure 4 (light grey lines), as well as the AFs for the profiles in Figure 2 (colored lines). b) Corresponding standard deviation (in natural logarithm space) for the two sets of profiles in a).

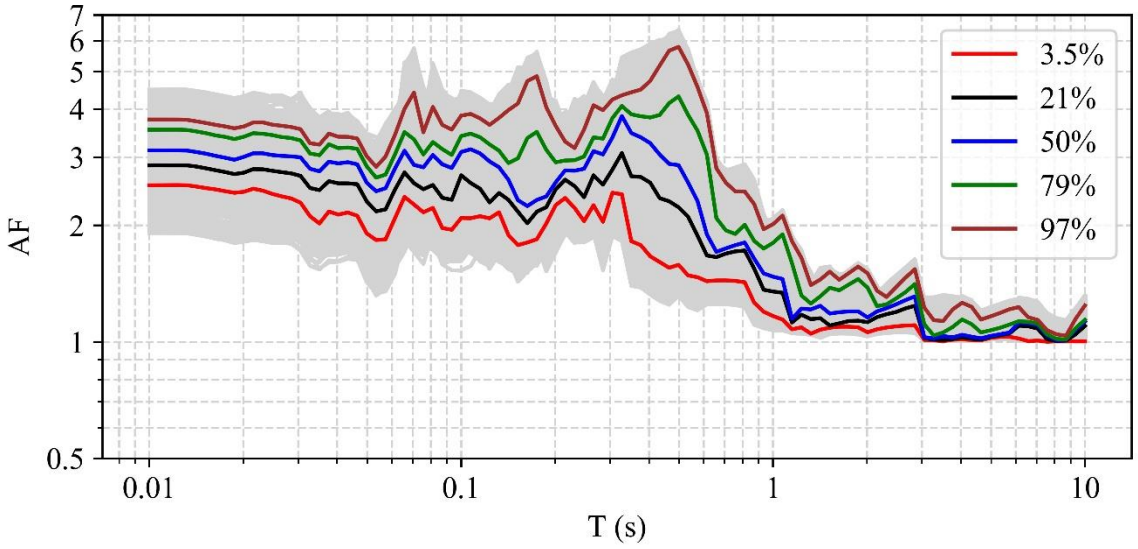
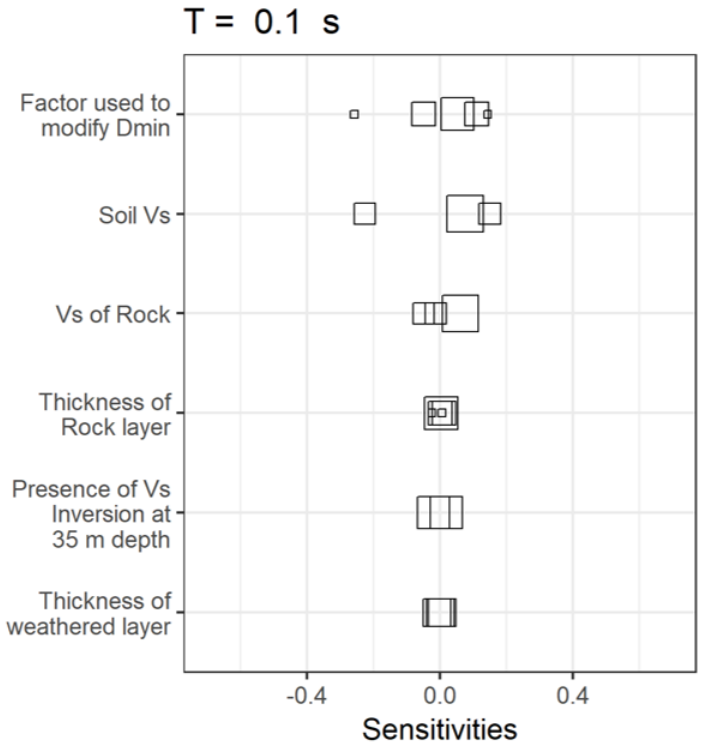


Figure 6. Amplification Factors for the profiles implicit in the logic tree in Figure 4 (light grey lines) and a statistical sample of five AF branches (colored lines). Following Miller and Rice (1983), these branches correspond to percentiles of 0.035, 0.212, 0.5, 0.788 and 0.965, with corresponding weights of 0.101, 0.244, 0.31, 0.244, 0.101.

The example shown in this publication is purposefully simple: it does not consider the potential for nonlinearity and it does not account for potential magnitude dependence of the amplification factors (Stafford et al., 2017). However, it is easy to expand the proposed approach by appropriately sampling the AFs for different magnitudes or different input motion intensities. Alternatively, nonlinearity can be captured by using the proposed approach to sample the parameters of a nonlinear model (i.e., a model that captures AF as a function of input motion intensity). The proposed approach is meant only for capturing epistemic uncertainty. To capture aleatory variability, each of the baseline profiles in the site response logic tree can be randomized (EPRI, 2013).

(a)



(b)

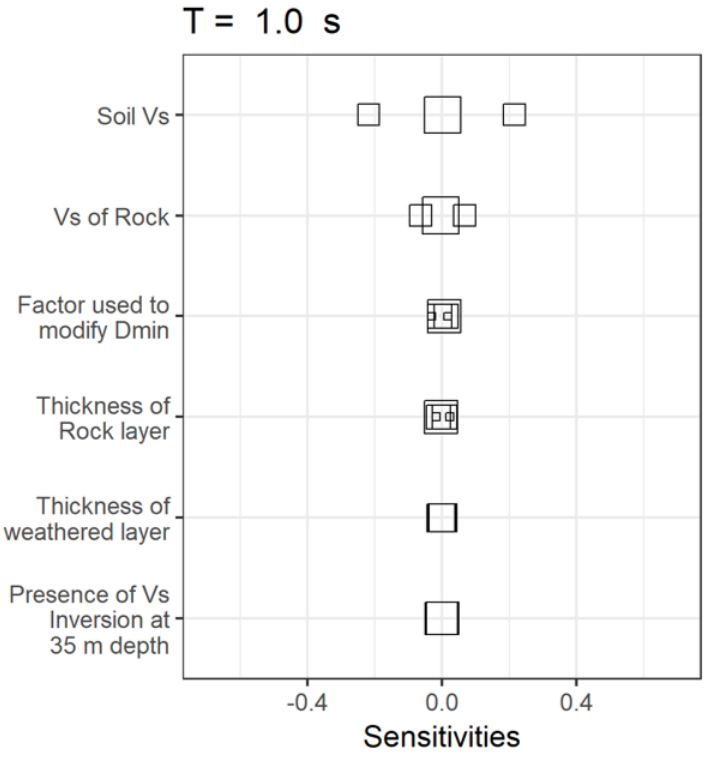


Figure 7. Tornado plots illustrating the sensitivity of AFs to the logic tree nodes for the logic tree in Figure 4. The x-axis corresponds to the deviation (in natural log space) from the median value of the full logic tree. The size of the square is proportional to the branch weight.

CONSIDERATION OF MODEL ERROR

The logic trees and resulting distributions of AFs discussed in the previous sections have only focused on parametric uncertainty. The resulting epistemic uncertainty in site response, as quantified by the σ_{AF} shown in Figure 5, tends to reduce to very low levels for long oscillator periods. This is a result of the assumptions of one-dimensional site response, which predicts that AFs reduce to unity for oscillator periods much longer than the site period. The values of σ_{AF} would seem to suggest that the predictions of site response are very accurate for long periods. One strong evidence to indicate that this is not the case comes from the KiK-net array. The KiK-net array has both surface and downhole instruments. The downhole instruments are generally placed at very stiff layers, generally with V_S values over 600 m/s. Rodriguez-Marek et al. (2011) performed a regression of KiK-net data where the uncertainty in the residuals was partitioned into repeatable site terms at the borehole and repeatable borehole-to-surface amplification terms at each station. If site response was entirely predictable for very stiff sites, it would be expected that there would be little site-to-site variability of the site term at the borehole level. This, however, is not the case, as the site-to-site variability for borehole records is relatively constant across all oscillator period (Rodriguez-Marek et al., 2011). This mismatch is a result of errors introduced by the assumptions implicit in models to compute site response. Other authors have tried to quantify directly the model error in one-dimensional site response analysis using downhole arrays (Afshari and Stewart, 2017; Stewart and Afshari, 2020) or multiple analyses across ground motion networks (Kaklamanos et al., 2013), and obtained similar errors for long oscillator periods.

The discussion in the previous paragraph points to the fact that estimates of site response have a certain amount of model error. We postulate that this error should constitute a lower bound to the epistemic uncertainty, or a *minimum epistemic uncertainty*, for the AFs. The approach proposed in this paper suggests a simple way in which this minimum epistemic uncertainty can be incorporated into PSHA. Figure 8 shows the epistemic uncertainty of the AFs along with an arbitrary level of minimum epistemic uncertainty (the quantification of the value of minimum epistemic uncertainty is outside the scope of this paper). Any time that the σ_{AF} falls below the level of minimum epistemic uncertainty, the σ_{AF} will take this larger value. The resulting σ_{AF} can then be used to widen the range of the branches of the AFs. One caveat to this approach is that the AFs should not fall outside of acceptable values due to physical considerations. In this particular case, the AFs are constrained to be larger than one for long

oscillator periods (Figure 8). We also note that an alternative approach is to assume that model error is statistically independent from the parametric uncertainty captured in the proposed approach (Stewart and Afshari 2020). In such a case, the model error must be added (via sum of variances) to the epistemic uncertainty computed by the proposed logic-tree approach.

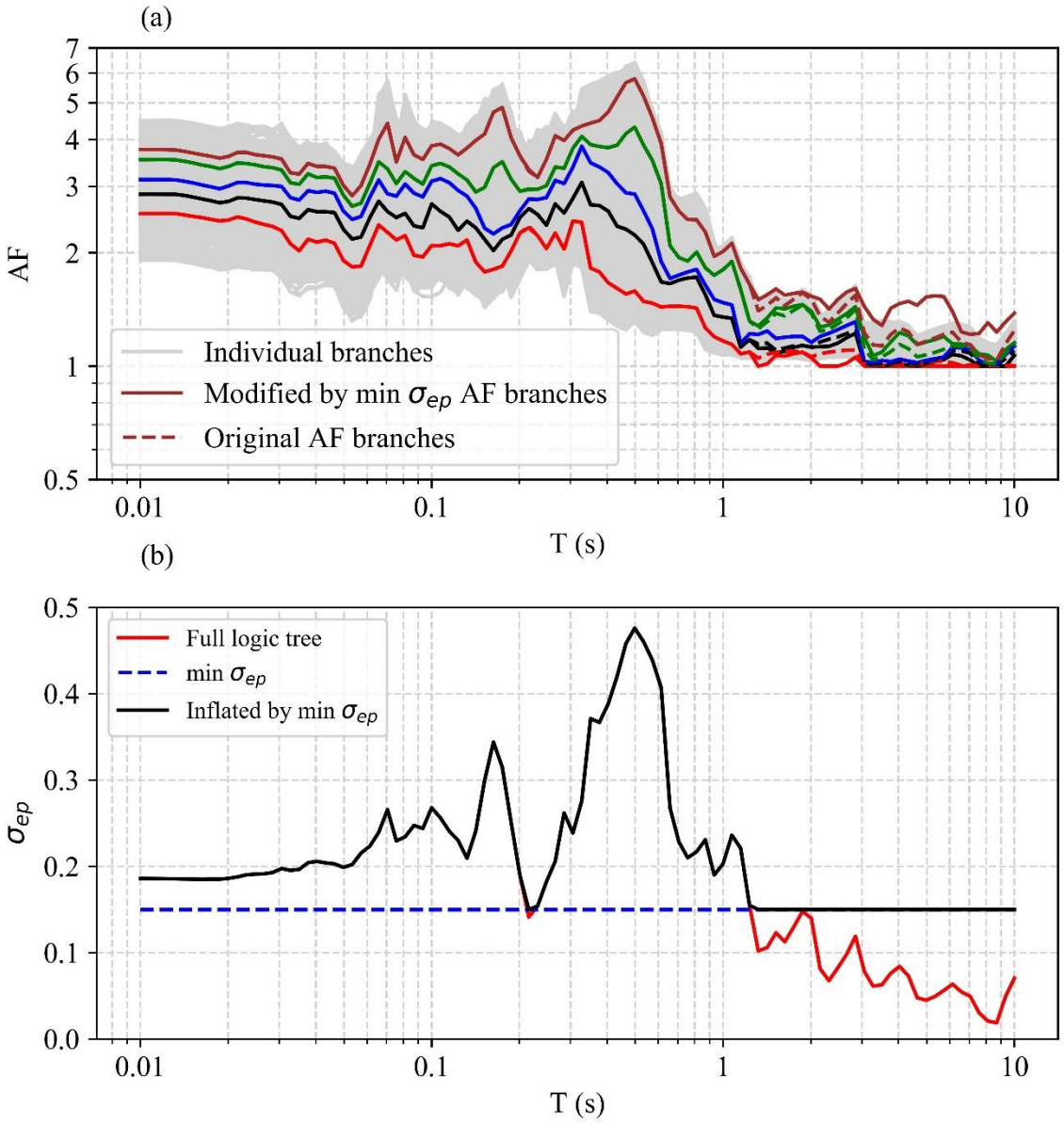


Figure 8. a) Branches of the AFs expanded to account for minimum epistemic uncertainty. b) standard deviation of AFs illustrating. The red line is the σ_{ep} that results from the full logic tree; the black line results from the imposition of a minimum level of epistemic uncertainty. The minimum level of epistemic uncertainty is set at 0.15.

DISCUSSION AND CONCLUSIONS

During the last three decades, and particularly in the last few years, the practice of PSHA has made important advances in terms of recognizing, quantifying and incorporating different sources of epistemic uncertainty. The distinctions between aleatory variability and epistemic uncertainty, together with the inclusion of the former within the hazard integral and the inclusion of the latter through logic-tree formulations, have been well established for some time in terms of SSC and GMC models, at least for rock sites. The separation of aleatory variability and epistemic uncertainty in site response analyses has evolved somewhat more slowly although it is now the state of practice to explicitly model epistemic uncertainty by including multiple V_S profiles (with randomizations of each profile). However, while it is now standard practice to develop logic trees to represent the epistemic uncertainty in most elements of both SSC and GMC models, attention has shifted from assigning weights to alternative models or parameter values to actually focusing more on the resulting distributions of variables that drive the hazard (source-to-site distances and recurrence rates of earthquakes of different magnitude in SSC models and response spectral accelerations in GMC models). In this paper, we have demonstrated that the same issues of unintendedly narrow distributions of amplitudes resulting from assigning weights to alternative models can occur in site response analysis as much as it can in prediction of rock-motion amplitudes. This can be avoided through the development of more complete logic trees to quantify all sources of uncertainty in a site response model and generating large numbers of alternative site profiles. While this creates the need for larger numbers of site response analyses, the computational burden for the PSHA calculations since the suite of AFs obtained from the full logic tree are resampled into a small number of discrete branches. This approach also brings two significant benefits, one being the insight into which elements of uncertainty exert greatest influence on the AFs at different response periods. This could influence decisions regarding additional data collection, targeting the period ranges of greatest importance to the site-specific project, which could be very different, for example, if this is a nuclear power plant, a large dam or a long-span bridge. The second benefit of the approach is that it provides a framework within which to impose a minimum level of epistemic uncertainty in order to avoid underestimating such uncertainty at longer periods as a result of simplified modeling approaches.

To move from current practice in site-specific PSHA, which generally now includes site response analyses with multiple V_S profiles, to the approach proposed in this paper is not an onerous step and we believe it is the next logical step in the development of PSHA practice.

ACKNOWLEDGMENTS

The ideas presented in this article first came about through discussions within the context of several SSHAC Level 3 PSHA studies in which the authors participated. The authors are grateful to the other participants in those projects, including members of the Technical Integration Teams and Participatory Peer Review Panels as well as Resource and Proponent Experts invited to participate in project workshops. At the risk of omitting individuals who contributed to these discussions, we would like to acknowledge the valuable contributions to these ideas through discussions with Norm Abrahamson and Jon Stewart. We are also very grateful to Jim Kaklamanos, Cliff Munson, Jon Stewart and an anonymous reviewer for constructive and insightful feedback on the manuscript that helped us to improve the presentation of our ideas.

DECLARATION OF CONFLICTING INTERESTS

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The ideas presented in this paper were originally conceived within the context of funded consultancy projects, but the development of this work and the generation of this article have been undertaken as independent efforts on the part of the authors without external funding.

REFERENCES

Afshari K and Stewart JP (2017) Implications of California vertical array data for the analysis of site response with 1D geotechnical modeling. Report to California Strong Motion Instrumentation Program. Civil and Environmental Engineering, University of California, Los Angeles.

Atkinson GM (2006) Single-station sigma. *Bulletin of the Seismological Society of America* 96(2): 446–455.

Atkinson GM, Bommer JJ and Abrahamson NA (2014) Alternative approaches to modeling epistemic uncertainty in ground motion in probabilistic seismic-hazard analysis. *Seismological Research Letters* 85(6): 1141-1144.

Bahrampouri M, Rodriguez-Marek A and & Bommer JJ (2019) Mapping the uncertainty in modulus reduction and damping curves into the uncertainty of site amplification factors. *Soil Dynamics & Earthquake Engineering* 126: Article 105091.

Bazzurro P and Cornell, CA (2004a) Ground-motion amplification in nonlinear soil sites with uncertain properties. *Bulletin of the Seismological Society of America* 94: 2090-2109.

Bazzurro P and Cornell, CA (2004b) Nonlinear soil-site effects in probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America* 94: 2110-2123.

Bommer JJ (2012) Challenges of building logic-trees for probabilistic seismic hazard analysis. *Earthquake Spectra* 28(4): 1723-1735.

Bommer JJ and Abrahamson NA (2006) Why do modern probabilistic seismic hazard analyses lead to increased hazard estimates? *Bulletin of Seismological Society of America* 96(6): 1967-1977.

Bommer JJ and Scherbaum F (2008) The use and misuse of logic trees in probabilistic seismic hazard analysis. *Earthquake Spectra* 24(4): 997-1009.

Bommer JJ and Stafford PJ (2020) Selecting ground-motion models for site-specific PSHA: adaptability vs applicability. *Bulletin of the Seismological Society of America*, in press.

Cornell CA (1968) Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 58(5): 1583-1606.

EPRI (2013) *Seismic Evaluation Guidance. Screening, Prioritization and Implementation Details (SPID) for the Resolution of Fukushima Near-Term Task Force Recommendation 2.1. Seismic*. EPRI Rep. No. 1025281, Palo Alto, California: Electrical Power Research Institute.

Griffiths SC, Cox BR, Rathje EM and Teague DP (2016) Surface-wave dispersion approach for evaluating statistical models that account for shear-wave velocity uncertainty. *Journal of Geotechnical and Geoenvironmental Engineering* 142(2016): 01016061.

Kaklamanos J, Bradley BA, Moolacattu AN and Picard BM (2020) Physical hypotheses for adjusting coarse profiles and improving 1D site-response estimation assessed at 10 KiK-net sites. *Bulletin of the Seismological Society of America* 110(3): 1338–1358.

Kaklamanos J, Bradley BA, Thompson EM and Baise LG (2013) Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data. *Bulletin of the Seismological Society of America* 103(3):1733-1749.

Kottke A (2019) Pysra: Site response analysis tool for Python, v0.4.5. DOI/10.5281/zenodo.3522104.

Kulkarni RB, Youngs RR and Coppersmith KJ (1984) Assessment of confidence intervals for results of seismic hazard analysis. *Proceedings of the Eighth World Conference on Earthquake Engineering*, San Francisco, vol. 1: 263-270.

Lessi-Cheimariou A, Tromans IJ, Rathje E and Robertson C (2019) Sensitivity of surface hazard to different factors and site response analysis approaches: a case study for a soft rock site. *Bulletin of Earthquake Engineering* 17(1): 73-96.

McGuire RK (2008) Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering & Structural Dynamics* 37: 329-338.

Miller AC and Rice TR (1983) Discrete approximations of probability distributions. *Management Science* 29(3): 352-362.

Papaspiliou M, Kontoe S and Bommer JJ (2012) An exploration of incorporating site response into PSHA – Part I: Issues related to site response analysis methods. *Soil Dynamics & Earthquake Engineering* 42: 302-315.

Passeri F, Foti S and Rodriguez-Marek A (2020) A new geostatistical model for the management of uncertainties in shear wave velocity profiles. *Soil Dynamics and Earthquake Engineering*, submitted for publication.

Passeri F, Foti S, Cox BR and Rodriguez-Marek A (2019) Influence of epistemic uncertainty in shear wave velocity on seismic ground response analyses. *Earthquake Spectra* 35(2): 929-954.

Rodriguez-Marek A, Cotton F, Abrahamson NA, Akkar S, Al Atik L, Edwards B, Montalva GA and Dawood HM (2013) A model for single-station standard deviation using data from various tectonic regions. *Bulletin of the Seismological Society of America* 103(6): 3149–3163.

Rodriguez-Marek A, Montalva GA., Cotton F and Bonilla F (2011) Analysis of single-station standard deviation using the KiK-net data. *Bulletin of the Seismological Society of America* 101(3): 1242-1258.

Rodriguez-Marek A, Rathje EM, Bommer JJ, Scherbaum F and Stafford PJ (2014) Application of single-station sigma and site response characterization in a probabilistic seismic hazard analysis for a new nuclear site. *Bulletin of the Seismological Society of America* 104(4): 1601-1619.

Scherbaum F, Kuehn NM, Ohrnberger M and Koehler A (2010) Exploring the proximity of ground-motion models using high-dimensional visualization techniques. *Earthquake Spectra* 26(4): 1117-1138.

Stafford PJ, Rodriguez-Marek A, Edwards B, Kruiver PP and Bommer JJ (2017) Scenario dependence of linear site effect factors for short-period response spectral ordinates. *Bulletin of the Seismological Society of America* 107(6): 2859-2872.

Stewart JP and Afshari K (2020) Epistemic uncertainty in site response as derived from one-dimensional ground response analyses. *Journal of Geotechnical and Geoenvironmental Engineering*, Accepted for publication.

Stewart JP, Afshari K, and Hashash YMA (2014) Guidelines for performing hazard-consistent one-dimensional ground response analysis for ground motion prediction, *PEER Report No. 2014/16*, Pacific Earthquake Engineering Research Center, UC Berkeley, CA.

Stewart JP, Wang P, Teague DP and Vecchietti A (2019) Applications of non-ergodic site response in ground motion modeling, Proceedings 7th International Conference on Earthquake Geotechnical Engineering, Rome Italy.

Stromeyer D and Grünthal G (2015) Capturing the uncertainty of seismic activity rates in probabilistic seismic-hazard assessments. *Bulletin of the Seismological Society of America* 105(2A): 580-589.

Teague DP, Cox BR and Rathje EM (2018) Measured vs. predicted site response at the Garner Valley Downhole Array considering shear wave velocity uncertainty from borehole and surface wave methods. *Soil Dynamics and Earthquake Engineering* 113: 339-355.

Toro, GR (1995) Probabilistic models of site velocity profiles for generic and site-specific ground-motion amplification studies, *Technical Report No. 779574*, Brookhaven National Laboratory, Upton, New York, 147 pp.

Tromans IJ, Aldama-Bustos G, Douglas J, Lessi-Cheimariou A, Hunt S, Davi M, Musson R, Garrard G, Strasser F and Robertson C (2019) Probabilistic seismic hazard assessment for a new-build nuclear power plant in the UK. *Bulletin of Earthquake Engineering* 17(1): 1-36.

USNRC (2018) *Updated implementation guidelines for SSHAC hazard studies*. NUREG-2213, Washington DC: US Nuclear Regulatory Commission.

Yee E, Stewart JP and Tokimatsu K (2013) Elastic and large-strain nonlinear seismic site response from analysis of vertical array recordings. *Journal of Geotechnical and Geoenvironmental Engineering* 139: 1789-1801.