Manuscript accepted on 25 March 2023 for publication in Seismological Research Letters7

doi: 10.1785/022023037

Seismic Source Zones for Site-Specific Probabilistic Seismic Hazard Analysis: The Very Real Questions Raised by Virtual Fault Ruptures

Julian J Bommer[1*], Jon P Ake[2], and Clifford G Munson[3]

Abstract

Site-specific probabilistic seismic hazard analyses (PSHA) very often include areal source zones to represent diffuse seismicity that cannot be associated with known geological faults. Most modern ground-motion prediction models use distance metrics that are defined relative to the extended fault rupture rather than the epicenter or hypocenter. For these distances to be calculated correctly, virtual fault ruptures are generated, having dimensions consistent with the earthquake magnitude, within source zones when performing PSHA calculations. Although the generation of these virtual ruptures is necessary to achieve compatibility between the seismic source and ground motion models within the hazard calculations, the ruptures should, by definition, represent potentially realizable seismogenic structures within the crust. Frequently, algorithms for the generation of these virtual ruptures are generally embedded within the PSHA code as an interim calculation without generating any outputs to enable visualization of the location and extension of the resulting ruptures. Such visualizations can reveal features of these hypothetical ruptures that may challenge the assumptions underlying the definition of the source zone boundaries that separate and enclose distinct regions of diffuse seismicity, as well as raising questions regarding the recurrence parameters within each source, especially in terms of the assumed maximum magnitudes. Visualizing the virtual ruptures generated in PSHA calculations and ensuring their consistency with the criteria established, explicitly or otherwise, for the definition of seismic source zones, could lead to important improvements in the modeling of diffuse seismicity in PSHA. We propose that this visualization should become a standard step in any PSHA study that includes source zones of diffuse seismicity. In addition, the choice of strict or leaky source zone boundaries relative to these hypothetical ruptures should always be explained and justified rather than simply stated.

[*] Corresponding author. E: j.bommer@imperial.ac.uk

Introduction

All crustal earthquakes are associated with the rupture of geological faults. However, since it is generally impossible to associate all earthquakes with known faults, seismic source models for probabilistic seismic hazard analyses (PSHA) invariably include sources of diffuse seismicity in which future earthquake scenarios are not localized on mapped faults. To accommodate modern ground-motion models (GMMs), most of which employ distance metrics defined relative to the causative fault rupture, it has become standard practice in site-specific PSHA to generate virtual fault ruptures for each earthquake scenario. A virtual rupture is usually a plane with dimensions and geometry consistent with a geological fault rupture associated with an earthquake of specified magnitude but not associated with a mapped geological fault. Although these hypothetical ruptures resolve the potential inconsistency between the seismic source model and the GMMs, examination of the virtual fault ruptures can sometimes reveal apparent inconsistencies with the geologic criteria used to define the seismic sources.

In standard site-specific PSHA practice for critical facilities, algorithms are defined to generate the virtual fault ruptures, which are then used within the hazard engine to calculate the distances required by the GMMs. There could, however, be great value in making visualizations of the virtual fault ruptures generated within the PSHA calculations a routine component of constructing a seismic source model. The reason for this is that such visualizations would allow the hazard analyst to examine the locations and dimensions of the virtual ruptures and to determine whether they are consistent with the properties of the seismic source zone within which they have been generated, and also the properties of any other zone into which they are allowed to propagate. We suspect that in many cases such visualizations may challenge several of the assumptions made in defining the source zone. Although the initial response to such situations might be to consider modifying the algorithm for the virtual rupture generation, the more appropriate response will often be to reconsider the criteria that were established to define the seismic source zone in the first place.

We begin by presenting a brief overview of the criteria used to define seismic source zones in PSHA, particularly in terms of which characteristics of the seismicity are assumed uniform across a source zone and which may be allowed to vary spatially within the zone. We also discuss the issue of whether source zone boundaries are ‘strict’ or ‘leaky’, or in other words, whether virtual fault ruptures are allowed to cross the boundaries. This is followed by a brief discussion of virtual fault ruptures, starting with an explanation of why they are needed, and when they are unnecessary, and a summary of the procedures used for their generation within PSHA calculations. The paper then looks at the implications of virtual ruptures for seismic source models, and the questions that might be raised regarding consistency with the criteria that were applied to the specification of the seismic source zones. In the Discussions section we offer several suggestions for addressing, and in some cases resolving, the questions that are raised by the virtual ruptures, offering recommendations for making consistency between the criteria defining seismic source zones and the algorithms generating virtual ruptures a fundamental objective of any site-specific PSHA study.

Defining Seismic Source Zones in PSHA

Source zones (as polygons) or gridded seismicity, or a combination of these such as gridded seismicity within area source zones, are utilized in nearly all seismic hazard studies as a mechanism to capture the hazard at a site that arises from earthquakes that occur on faults that have no identifiable surface expression (at least not recognized at the time of the study). The expected characteristics of future seismicity within the source zone should be substantially uniform and distinct from regions outside of the zone (Baker et al., 2021). However, it should be noted that earthquake rates are often allowed to vary spatially within source zones to represent the observation that the general locations of future events are informed by the locations of past events, but those events are not an absolute predictor of the location of future events at the time scale of interest for most hazard studies (McGuire, 2004; Kafka, 2007; USNRC, 2012).

In defining source zone boundaries, the hazard analyst should be mindful that the ultimate purpose of the effort is to provide an assessment of the characteristics of future earthquakes in the region of interest. Several different factors (or criteria) have been utilized as the basis for defining seismic source zones in recent practice. Some of those factors include tectonic stress regime, style-of-faulting, strike and dip of ruptures, seismogenic thickness, fault rupture area, and the aspect ratio of ruptures (USNRC, 2012). Each of these factors is discussed briefly below. Maximum magnitude (Mmax) could be utilized as a factor but is often determined using source zone specific data after the source zones and their boundaries have been defined.

Tectonic stress regime has been included as a factor in identifying source zones to provide consistency with GMMs that often have alternative models based upon tectonic stress regimes (compressional versus extensional or stable continental versus active tectonic, for example). Both local and regional data on tectonic stress regime can be used in this assessment. Sense of slip/style of faulting should be specified for each zone. This information can be derived from existing focal mechanisms, stress indicators and possibly from geologic observations. This is also necessary for consistency with the GMMs. If future earthquakes are to be modelled as virtual ruptures, then a characterization of the expected strike and dip of future ruptures is also required. This is typically based on existing focal mechanism data for the region. Available geological information on significant fold and fault systems and structural grain in the area can also be utilized as confirmatory if that information is deemed to be consistent with the current tectonic stress regime. Seismogenic thickness must be identified for each potential zone as it is an important constraint on the geometry and size of future ruptures. This is often based upon a joint assessment of depth to some critical isotherm (determined from deep heat flow information) and an evaluation of a lower threshold of focal depth information. This is often determined from the 90^{th} or 95^{th} percentile of the reliable focal depths in the region of interest (frequently denoted as D90 or D95). A model that specifies a relationship between magnitude and fault-area is also required. This is often based on larger regional or tectonic regime specific empirical relationships. To produce a model for finite fault ruptures it is also necessary to specify a length-to-width relationship for future events. For larger

magnitudes, empirical relationships are typically used, for smaller magnitudes aspect ratios are often assumed (typically 1:1).

The final judgement in defining a source zone arises directly from the need to generate finite ruptures: a choice needs to be made whether the virtual fault ruptures can extend beyond the zone boundaries within which they are initiated or whether they are constrained to extend only within the source boundaries. If the virtual ruptures are truncated at the edge of the zone, then this is referred to as a strict boundary. If the rupture is allowed to cross the boundary and continue propagating into the adjacent source zone, then it is referred to as a leaky boundary. As will be discussed in the remainder of this paper, this choice can have many consequences and implications that warrant careful considerations.

There are no generalized rules for the dimensions of seismic source zones but there would often appear to be some correlation between the map size of the region under study and the sizes of the source zones, even though most studies acknowledge recommended zones of influence defined by a radius of ~300 km around the site. It has been observed that studies focused on smaller geographic regions sometimes produce a significant number of source zones of small dimensions (e.g., Tromans et al., 2019; Villani et al., 2020). Small source zones are sometimes justified on the basis of spatially constraining observed clusters of seismicity, whether or not a clear physical explanation can be provided for the localization of the earthquake activity. The same objective can be met using large seismic source zones but applying spatial smoothing of activity rates (and *b*-values as well) within the zone. Often a source model may capture both these scenarios (smaller or larger zones) with the preferred alternative receiving more weight.

The key issue in specifying the boundaries of a seismic source zone is to define which features of the seismicity are being modeled as uniform within the source, or indeed across multiple source zones if, for example, regional *b*-values or Mmax estimates are adopted.

The Need for Virtual Fault Ruptures in PSHA

The first generation of empirical GMMs generally used distance metrics defined relative to the epicenter or hypocenter, referred to as R_{epi} and R_{hyp}, respectively. Several model developers also noted, however, that these distance measures, based on representation of the earthquakes source as a point, did not reflect the dimensions of the sources of larger earthquakes and consequently could significantly overestimate the separation of the recording station and the source of seismic energy release. Donovan and Bernstein (1978) published the first GMMs that used distance metrics defined relative to the fault rupture. The landmark publications of Campbell (1981) and Joyner and Boore (1981) formally introduced the rupture distance, R_{rup}, and the Joyner-Boore distance, R_{JB}, the former being the distance to the closest point on the fault rupture, the latter the distance to the closest point on the projection of the rupture onto the Earth's surface. Over time, these metrics

defined relative to the extended fault rupture have been adopted in most GMMs such that the majority of current ground-motion models applicable to large earthquakes use either R_{JB} or R_{rup}.

Early computer programs for PSHA integrated over distance by discretizing the source zones into small areas or arcs that effectively treated the source of each earthquake as a point, which was fully consistent with the use of GMMs based on R_{epi} or R_{hyp}. The subsequent introduction of models using R_{JB} or R_{rup} created an inconsistency between the representation of the earthquake source in hazard integrations and in the GMM, the former treating each source as a point and the latter as an extended fault rupture. Because $R_{JB} \leq R_{epi}$ and $R_{rup} \leq R_{hyp}$, this inconsistency invariably results in underestimation of the seismic hazard (Bommer and Akkar, 2012; Monelli et al., 2014) since the point-source metrics will tend to place the source of each earthquake scenario farther away from the site than would be the case when using finite-rupture metrics.

There are three different options for resolving the inconsistency between hazard integrations that assume dimensionless earthquake sources and GMMs that describe the rupture dimensions of each earthquake: (1) revert to using GMMs based on R_{epi} or R_{hyp}; (2) apply conversions between different distance metrics; or (3) simulate virtual ruptures for each earthquake scenario to enable the calculation of R_{JB} or R_{hyp}. To address the first option, Bommer and Akkar (2012) recommended that GMPMs should be published as suites of models using both finite-rupture-based and point-based metrics (e.g., Akkar et al., 2014; Bindi et al., 2014). Bindi et al. (2017) echoed the recommendation, specifically adopting an R_{hyp}-based GMM for hazard calculations involving only areal source zones. Although this approach is internally consistent, it must be recognized that if larger magnitude earthquakes are significant contributors to the hazard estimates, the use of point sources for the scenarios is physically unrealistic. For site-specific PSHA studies for critical facilities, for which low annual exceedance frequencies are relevant and moderate-to-large magnitude earthquakes at relatively short distances will often be significant or even dominant contributors to the hazard, such an approximation is unlikely to be considered acceptable. An additional point that is often overlooked is that the aleatory variability associated with GMMs based on point-source metrics is inevitably larger than that obtained using finite-rupture metrics, and moreover the difference between the two is dependent on both magnitude and distance (e.g., EPRI, 2004; EPRI, 2013; Bommer et al., 2016; Thompson and Worden, 2018); sigma (standard deviation of residuals) values obtained from regression analysis will generally need modifying to capture the full extent of these dependencies.

The second option is to use GMMs based on R_{JB} or R_{rup} and to discretize the source zones in a manner that effectively treats each earthquake scenario as a point, and then apply conversions between the different metrics. EPRI (2004), EPRI (2013) and Scherbaum et al. (2004) derived such conversions by generating ground-motion fields around finite fault ruptures. The aleatory variability associated with these conversions, which depend on magnitude and distance, are large, and the propagation of this additional variability significantly inflates the sigma of the GMM and consequently the hazard estimates as well (Scherbaum et al., 2006). Thompson and Worden (2018) extended this approach for such applications as ShakeMap, where GMMs using R_{JB} or R_{rup} are deployed to calculate ground-motion fields for an earthquake scenario without knowledge of the fault rupture. Rather than

simulating or randomizing fault ruptures and site locations, Thompson and Worden (2018) numerically integrate the conversion factors from point-source-based to rupture-based distance metrics. Thompson and Worden (2018) performed the integrations to include variability in several parameters (dip, style-of-faulting, rupture area, depth to the top of rupture, and hypocenter location within the rupture plane).

Although both the options described above offer the convenience of continuing to model earthquakes sources within area source zones without needing to account for rupture dimensions, they are only approximate solutions to the issue of incompatibility. The third, and most rigorous, option is to generate virtual fault ruptures associated with each earthquake scenario considered in the hazard calculations, from which the distance from the earthquake source to the site can be calculated in accordance with the metric used in the GMM. This provides a more accurate and physically realistic resolution to the inconsistency, but it comes at the price of a considerable increase in the computational burden of the hazard integrations. However, the additional computational effort can be somewhat limited by not generating virtual ruptures in source zones that are remote from the site since it can be demonstrated that at larger distances the point-source approximation does not bias the hazard estimates (Bommer and Montaldo-Falero, 2020).

Although the focus herein is exclusively on site-specific PSHA for safety-critical facilities, we note in passing that virtual ruptures are also required for ShakeMaps when the earthquake scenarios are not associated with a mapped geological fault (e.g., Graves et al., 2011). Virtual faults are also relevant to regional risk assessments, since the orientation and length of the earthquake source will influence the number and location of buildings subjected to strong shaking for each earthquake scenario considered in the calculations.

The ruptures need to have dimensions consistent with the magnitude, as determined from empirical scaling relationships (e.g., Wells and Coppersmith, 1994; Leonard, 2014; Stafford, 2014; Thingbaijam et al., 2017). Generally, the rupture area will be fixed, which will lead to the rupture length depending on the fault dip and the defined seismogenic thickness of the crust for larger earthquakes. The length of the ruptures may also be bounded if the seismic source zone has strict rather than leaky boundaries (Figure 1). Other characteristics of the ruptures need to be either defined or randomized in the simulations, including the strike, dip, style-of-faulting, and depth-to-top-of-rupture, as well as the location of hypocenters down-dip and along-strike. Some of these features will, of course, be correlated, such as the style-of-faulting and the dip. The fault orientations may be constrained by the regional tectonics and the strike of known faults, which may be modeled as discrete individual seismic sources in addition to the area source zones. The procedures followed to generate virtual ruptures vary among PSHA codes (e.g., Hale et al., 2018) and may be one of the key factors leading to divergence in the results obtained from different software packages (Bommer et al., 2013). Campbell and Gupta (2018) provide an overview of some of the different choices that can be made in the simulation of virtual fault ruptures as well as an analysis of the sensitivity of hazard estimates to these choices. In essence, the faults are generated by fixing the hypocenter of an earthquake scenario and then generating rupture planes, with magnitude-consistent dimensions. Centered on the hypocenter, the rupture plane will then be rotated in strike and dip to sample the

ranges specified for these parameters, which may be 180° for strike unless there is constraint on preferred fault orientations, and for dip will be determined by the style-of-faulting.

Virtual Ruptures and Seismic Source Zones

For site-specific PSHA studies, the generation of virtual fault ruptures within areal source zones has become standard practice, albeit that the rules for generating these hypothetical earthquake sources vary considerably among hazard calculation codes. Virtual ruptures are an effective and rigorous solution to the inconsistency that would otherwise exist as a result of using GMMs based on finite-rupture R_{JB} or R_{rup}, but virtual fault ruptures also offer the opportunity to perform visualizations to reveal the implications for the seismicity associated with each source zone that might not otherwise be fully appreciated. However, in current practice, the virtual ruptures are generally simulated within the calculations as an interim step and are not included in the output; for simplified schemes that sample from distributions of distances calculated *a priori*, the actual nature of the implied fault ruptures can be even more obscure. In this section we discuss some of the issues related to virtual ruptures and the questions that these might raise regarding seismic source modeling.

Virtual Ruptures Should Reflect Potential Earthquakes on Physically Possible Faults

Although virtual ruptures are a tool of convenience to avoid an inconsistency between the seismic source model and the ground motion model in the hazard calculations, we would argue that they should always represent potentially realizable ruptures on currently unknown geological faults. This is the case because of the definition of the seismic source model that specified an earthquake scenario of a given magnitude with an epicenter at a particular location within the source zone. In a sense, the virtual ruptures can be viewed as illustrating the implications of the seismic source model by showing possible locations of the extended source of each of the defined earthquake scenarios.

As noted above, virtual ruptures are usually generated symmetrically about the scenario epicenter, which implies that the earthquakes are always associated with bilateral fault rupture, even though this is often not the case (e.g., Mai et al., 2005). If the convention of bilateral rupture is applied in all cases, epicenters located close to a strict (as opposed to leaky) source boundary will result in very short ruptures regardless of the earthquake magnitude unless one decides to simply abandon this convention and allow for non-bilateral ruptures (Figure 1). Similarly, for small source zones, the maximum rupture length will be limited by the shortest distance to a strict boundary (Figure 1). The largest magnitude that therefore can be accommodated when strict boundaries have been specified for a source zone is therefore constrained by the maximum dimensions of the zone, the rupture orientation, the seismogenic thickness of the crust, and the lowest permitted fault dip. If the largest Mmax value specified for a source zone is appreciably greater than the largest rupture that can be accommodated within the source boundaries, the implication is that the fault rupture would be associated with a very large displacement on the fault (since the seismic moment is the product of

the crustal rigidity, the rupture area, and the average fault slip). This would imply a very high stress drop earthquake, which was probably not envisaged in the source model – and hence a new inconsistency is created between the seismic source model and the ground motion model. This might not have a material impact on the hazard estimates (since stress drop is not usually included as an independent variable in GMMs) but following our premise that virtual ruptures should always correspond to potentially realizable earthquake sources, such an inconsistency should not be ignored.

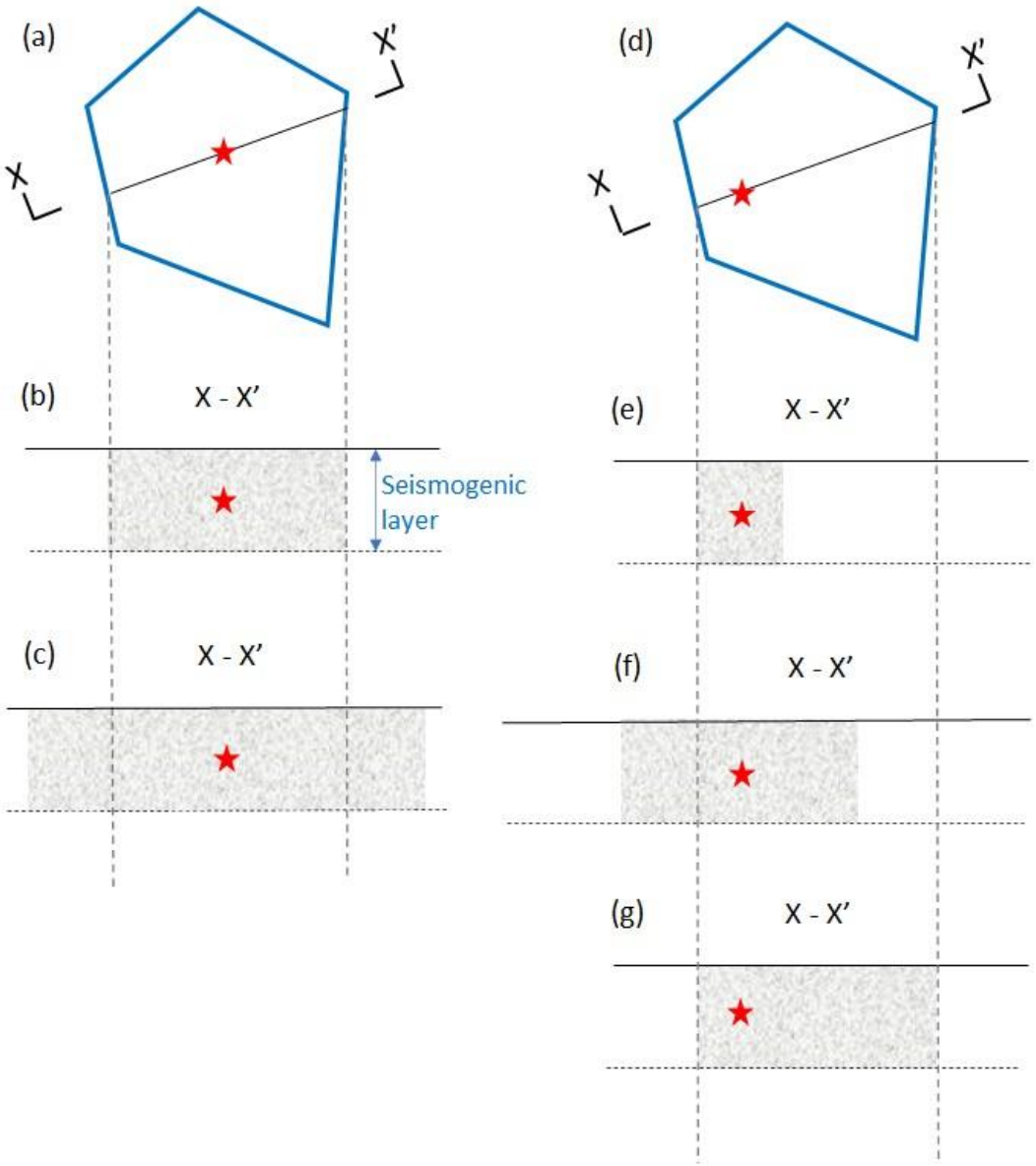


Figure 1. (a) Source zone and epicenter of earthquake scenario located in the center of the source, with X-X' indicating orientation of virtual rupture; (b) cross-section showing virtual rupture for largest Mmax when source has strict boundaries; (c) virtual rupture for Mmax with leaky source boundaries; (d) source zone and epicenter of earthquake scenario located towards the edge of the source; (e) maximum rupture size if hypocenter constrained to be at mid-point and strict source boundaries; (f) bilateral rupture if the boundaries are leaky; (g) maximum fault rupture with strict boundaries but virtual fault allowed to rupture asymmetrically.

Leaky Source Boundaries and Fault Rupture Scenarios

Larger Mmax values can be accommodated by allowing the virtual ruptures to extend beyond the source zone boundaries, which then need to be defined as leaky. This means that a significant proportion of the rupture may be located within another source zone. Invoking our implicit assumption that the faults should correspond to potentially real physical features, it must then be acknowledged that the same rupture could have initiated inside the adjacent source and propagated into the source currently under consideration. In other words, we create a situation whereby a possible fault traverses two source zones. Imagine now a case where a number of sub-parallel geological faults have been identified and are believed to be active but there is insufficient paleoseismological data to constrain their slip rates. In such a case, the mapped faults are likely to be accommodated within a seismic source zone rather than modeled as individual fault sources. However, it is difficult to imagine that a source zone boundary would be drawn across the fault traces; yet this is precisely what is implied by virtual fault ruptures that lie across two adjacent source zones.

There is another potentially serious consequence of leaky boundaries. Consider the case where the upper end of the Mmax distribution is large and the seismogenic thickness is relatively thin, hence the largest ruptures can assume great lengths that will extend far beyond the source limit if the boundaries are leaky. In extreme cases, it has been seen that for the largest Mmax values in remote sources that are significantly more active than the host zone (within which the site is located), when combined with the thinnest seismogenic crust, that fault ruptures have propagated to within very short distances of the site, leading to seemingly impossible magnitude-distance contributions when the hazard is disaggregated. Disaggregation to identify the hazard contributions from different magnitude-distance combinations, performed for each source zone separately, can usefully identify cases where virtual ruptures crossing leaky boundaries are approaching the site of interest. Suince such disaggregation is rarely performed, this situation may occur in many PSHA studies and pass undetected, especially if the host source zone is of comparable activity to neighboring source zones. The recommendation of Bommer and Montaldo-Falero (2020) to use point-source approximations rather than generating virtual ruptures in remote source zones avoids this issue to some extent. Another option is to specify orientations (strikes) for the virtual ruptures such that they would never propagate towards the site, but this would require geological or tectonic justification rather than being imposed solely to avoid this potential problem.

Figure 2 shows an example, from a real (but anonymized) project, of one possible set of virtual fault ruptures generated for a source zone having leaky boundaries. The fault orientations are unconstrained, the only criteria being imposed are that the earthquake epicenter is located within the source and the rupture dimensions are determined by the earthquake magnitude. While the figure shows that the ruptures do effectively sample the space within the source in terms of potential faults, there are a number of large ruptures for which only a very small fraction of the rupture length is actually within the source boundaries.

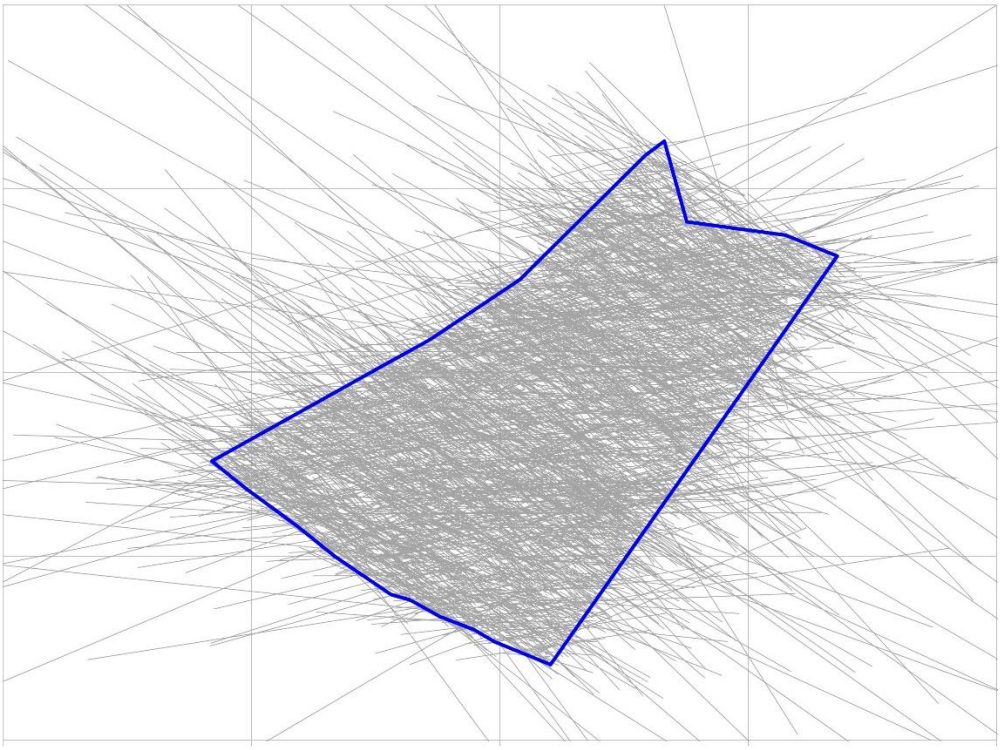


Figure 2. Example of virtual faults (grey lines) generated to represent the seismicity associated within a single seismic source zone (defined by the blue line); all the source boundaries are leaky, and the virtual ruptures are generated with random orientations. Note that several of the ruptures extend beyond the boundaries of the map area.

Virtual Ruptures and Maximum Magnitudes

One obvious way that virtual fault ruptures raise questions regarding seismic source zone characteristics is related to the largest magnitude earthquakes that are envisaged by the recurrence relations, specified as Mmax. For magnitudes greater than 7, for example, the virtual ruptures will have lengths of 50 km or greater, which implies that there is the possibility of a fault of at least this size present that has not been detected, despite the fact that the probability of such large crustal earthquakes not producing surface rupture is small (e.g., Youngs et al., 2003). Visualizing the corresponding ruptures for the largest earthquakes envisaged for a seismic source zone model and asking whether there is a possibility of a capable structure being present without leaving any signature on the landscape, may prompt reconsideration of the assigned Mmax values, although this needs to be balanced with the fact that large earthquakes can and do occur on previously unrecognized faults (e.g., the 1886 Charleston and 2001 Bhuj earthquakes). This is particularly the case when the seismic source model does include fault sources that serve to localize the largest earthquake scenarios. A great deal of effort has been invested to develop strategies for estimating Mmax for distributed seismicity, but it must also be recognized that it is a parameter that generally exerts a modest influence on seismic hazard estimates (Minson et al., 2021). Consequently, hazard modelers may often be comfortable assigning large values of Mmax that have very low probabilities

of ever being realized but visualization of the corresponding virtual fault ruptures might prompt re-evaluation of the upper end of the Mmax distribution.

The maximum magnitude also has obvious implications for the dimensions of source zones: if small source zones are defined, then the only options are to either limit the upper bound on Mmax such that the virtual ruptures can be accommodated within the source (perhaps relaxing the convention of specifying bilateral ruptures) or to make some, or all, of the boundaries leaky. The source zone depicted in Figure 2 is not particularly small by some standards – it has dimensions of about 300 km x 100 km – yet there are many virtual ruptures that extend beyond its geographical limits.

Discussion

The previous section highlights that visualizing the virtual fault ruptures can raise several questions regarding the bases for defining seismic source zones. A fundamental issue is clearly whether the source bounds are strict or leaky, a choice that only needs to be made in the first place because of the generation of virtual fault ruptures. We would propose that the decision on the choice of leaky versus strict boundaries should only be made after the full characterization of the zones is complete, including both the boundary locations and the Mmax distribution for earthquakes within the zone. The assessment of whether a particular source zone boundary should be leaky or strict should be most heavily influenced by the Mmax and the rupture orientation(s) in the adjacent zone. If, on the one hand, the Mmax and orientations of future ruptures differ significantly between adjacent zones, then a strict zone boundary seems most appropriate to reduce the potential for magnitudes and rupture orientations contributing to the hazard at a site that are inconsistent with the characteristics defined for the host zone. If, on the other hand, Mmax and rupture orientation distributions are similar between adjacent zones, then leaky boundaries may be justified.

In general, if relatively small source zones are defined, the zone dimensions will exert a controlling influence on Mmax if the boundaries are strict and do not permit ruptures to propagate across the edge of the zone. If the limited size of the source zone is technically justified, then the distribution of Mmax values should be carefully considered such that the ruptures do not need to extend outside the source zone. However, this might also require some flexibility regarding the rules of how the ruptures are generated, such as relaxing the (unnecessary) requirement for bilateral fault ruptures, especially for epicenters close to the source boundaries, such that the rupture extends preferentially towards the interior of the source zone.

This also raises consideration of whether the only two options are strict boundaries that the ruptures can never cross or leaky boundaries that are in fact effectively invisible to the virtual ruptures. Figure 3 shows a scenario for which these two end cases can lead to significant differences in the location and orientation of ruptures within and beyond the source zone and can also have a significant impact on the PSHA hazard curves for a site. Shown on the upper portion of Figure 3 are magnitude (**M**) 7 bilateral ruptures that are randomly oriented within a zone with leaky boundaries

(left) and strict boundaries (right). Although there are no **M**7 ruptures, approximately 35 km in length, that extend significantly beyond the source boundaries for the leaky case, many of the bilateral **M**8 ruptures, approximately 350 km in length, shown on the lower portion of Figure 3, extend well beyond the source boundaries into other zones and towards the site.

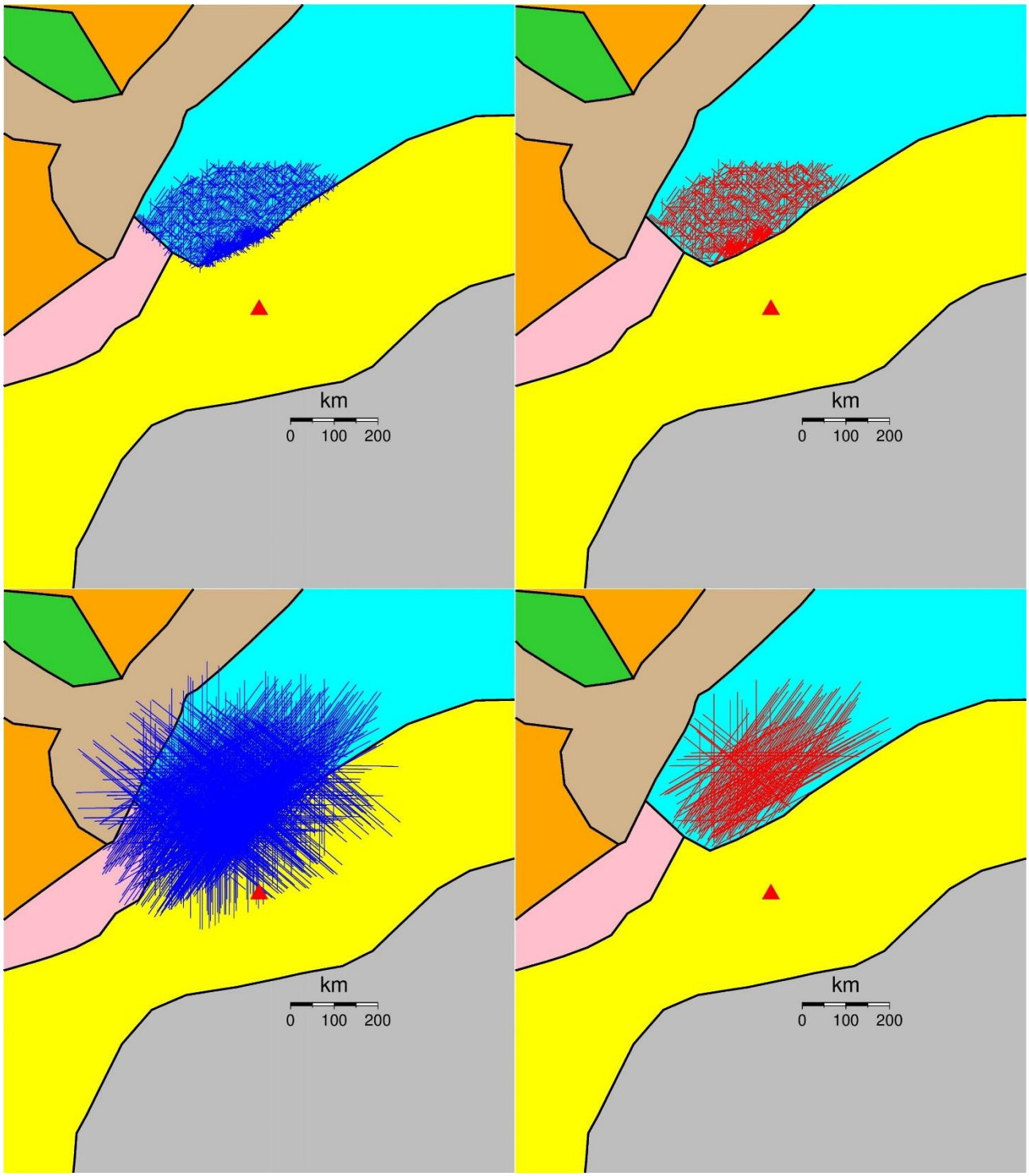


Figure 3. Examples of bilateral virtual fault ruptures generated for **M**7 (top) and **M**8 (bottom) earthquake scenarios in the light blue source zone when this is designated to have (left) leaky boundaries (blue ruptures) and (right) strict boundaries (red ruptures). The red triangle indicates the site for which the hazard is being calculated.

Figure 4 shows the number of bilateral ruptures as a function of magnitude for the two cases on the left and the mean 10 Hz hazard curves on the right. For the strict boundary case, the number of bilateral ruptures begins to fall off significantly for **M**>7 and the mean hazard curve for the strict boundary case at the site depicted by the red triangle is significantly lower than the curve for the leaky case. The visualization of this scenario raises several important questions with respect to the characteristics of smaller source zones and how virtual fault ruptures are modeled. First, with respect to the largest Mmax value for the zone, are 350-km in length **M**8 ruptures plausible for this region without surface expression of a geologic structure capable of generating such an event? Second, are there distinct differences between these source zones that warrant the development of these smaller zones? If there are distinct differences between these zones, then should the zone boundaries be leaky such that larger ruptures extend well beyond the zone boundaries? Finally, if the zone boundaries are modeled as being strict, then the spatial coverage and orientations for the larger ruptures, if assumed to be bilateral, may be limited as shown in the lower right corner of Figure 3.

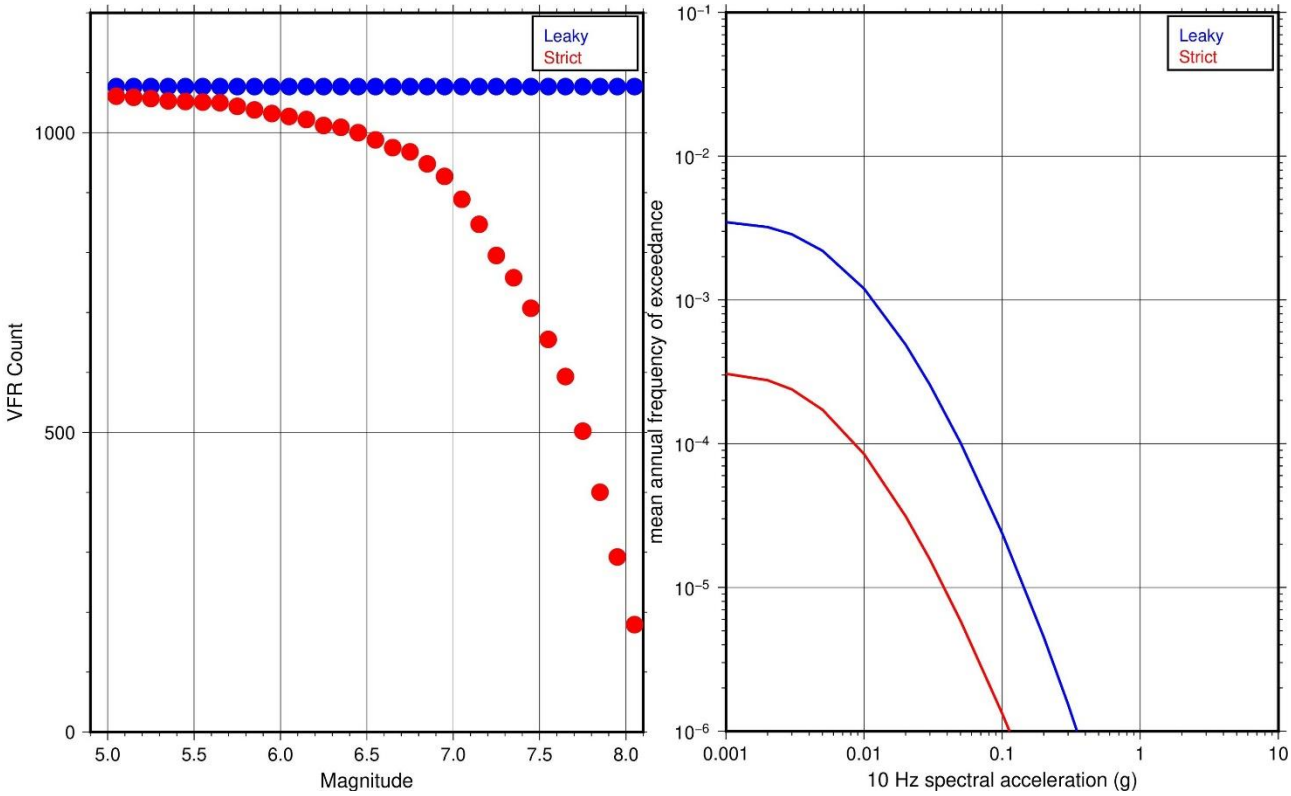


Figure 4. Virtual fault rupture counts (left) for the two scenarios (leaky and strict boundaries) depicted in Fig. 3 and hazard curves (right) in terms of PSA (10 Hz) for the site (red triangle) shown in Fig. 3 due to the two scenarios. The site is located on the Cape Cod peninsula approximately 150 km from the closest boundary of the light blue source zone depicted in Fig.3. The hazard curves were developed using the NGA-East GMM (Goulet et al., 2021) for the two rupture scenarios within the source zone, which has Mmax values ranging from **M** 6.1 to **M** 8.1 and gridded seismicity parameters estimated for each of the ¼° longitude by ¼° latitude cells within the zone.

There may be alternative options between the two end members of strict or leaky boundaries that would neither impose excessively restrictive constraints on the rupture dimensions nor permit the generation of virtual ruptures that lie predominantly outside the source boundaries. One option would be to specify a limit on how much of the virtual rupture can extend outside the source zone, which would be better expressed as a proportion of the total rupture length rather than a fixed distance. A possible basis for the permitted overshoot could be one standard deviation in the scaling relationship used to estimate the rupture dimensions. Such partially leaky or semi-permeable source boundaries would in a sense be reminiscent of the smoothed source boundaries, based on the standard errors in earthquake locations, proposed by Bender and Perkins (1982).

If all the source boundaries in a seismic source model are leaky, meaning that they are entirely transparent to fault ruptures, the following question needs to be asked: what exactly is being bounded within the source zones? Although this question may appear almost facetious, it is a challenge that should be addressed: if the response is activity rates, then in effect the model would be envisaging faults on which ruptures could initiate with different recurrence rates along its length (a situation that could occur perhaps but which should then be a deliberate modeling choice rather than an unintended consequence of a leaky source boundary). If all the source boundaries are defined to be leaky, then consideration should be given to merging zones into larger sources and then imposing spatially varying features of the seismicity within the larger zone (see, for example, Chapter 5 of USNRC, 2012, and Moschetti et al., 2015).

Figure 5 shows a flowchart of suggested steps that might be followed when generating seismic source zonations and visualizations of the associated virtual fault ruptures. This is not intended to be a prescriptive recipe for the process to follow but rather highlights the questions that could be considered with a view to achieving coherence between the objectives of the seismic source zones and the geological implications of the resulting virtual ruptures.

Conclusions

In this article we have highlighted several issues that arise from generation of virtual fault ruptures within seismic source zones that represent diffuse seismicity. Although the generation of these hypothetical ruptures is a laudable advance in PSHA practice to avoid a potential incompatibility between seismic source and ground motion models (or a backward step to using R_{epi} or R_{hyp} in GMMs), the ruptures should prompt reconsideration of the technical bases for defining seismic source zonation. We would propose that rather than viewing this as an inconvenience or an unwelcome complication, the virtual fault ruptures can be considered as diagnostic tools for understanding the physical implications for seismic source zone models. Whether the virtual ruptures are viewed as a burden or as an opportunity, it is essential that they are visualized by generating map views of realizations of the ruptures, as in Figures 2 and 3. This is best done for each source individually, but ideally on maps that show the other source boundaries and the location of the site of interest. We would propose that this visualization should become a standard step in any

PSHA study that includes source zones of diffuse seismicity. These plots should then be reviewed from the perspective of their consistency with the intended purpose of the source boundary configurations.

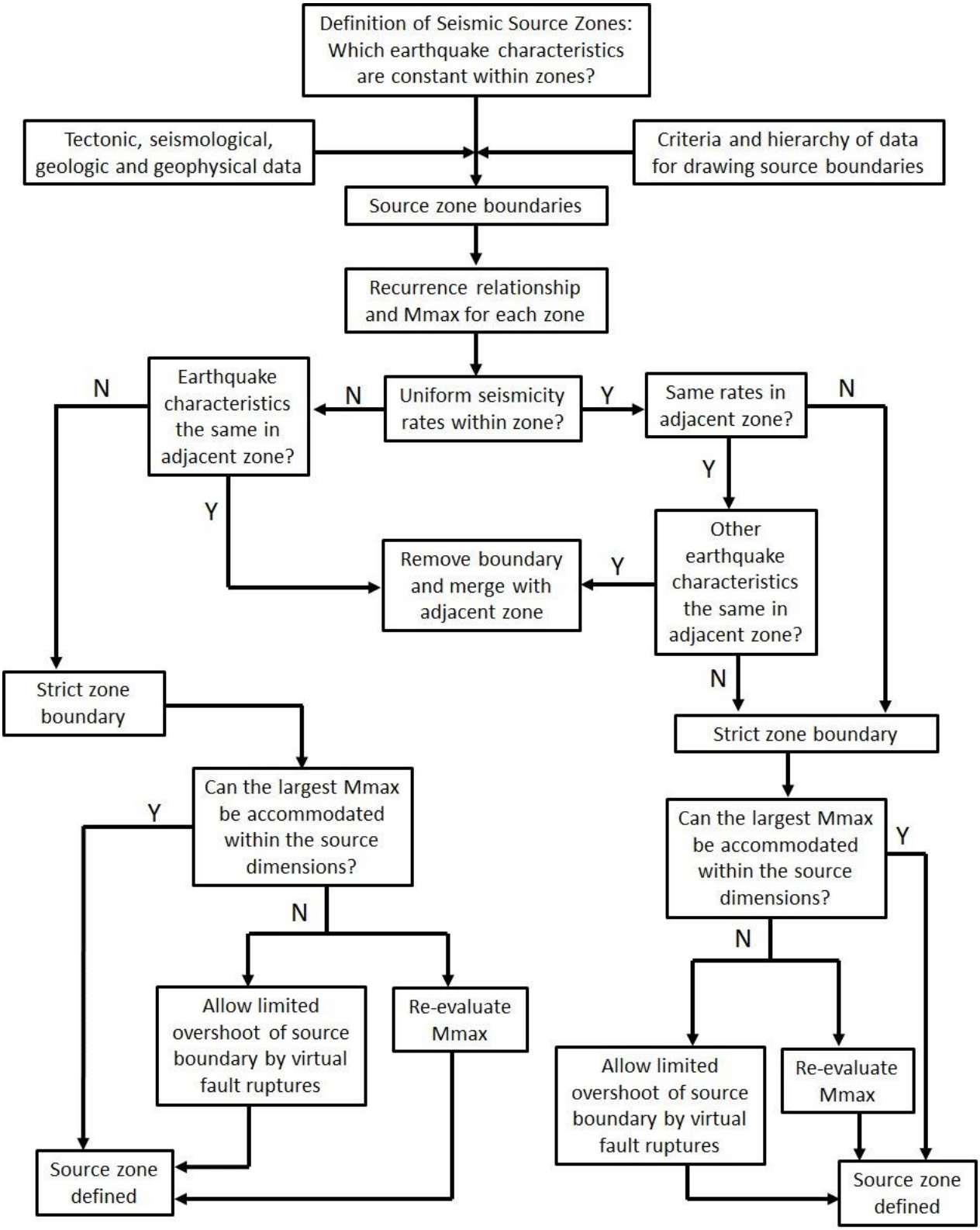


Figure 5. Flowchart for development of seismic source zones and selection of strict or leaky source boundaries.

A point to emphasize strongly is that the choice of strict or leaky source boundaries should always be explained and justified rather than simply stated. The technical bases for the nature of the boundaries should be documented and there should also be some discussion of the consequences of the choice and how these align with the criteria that were used to draw the source boundaries. As noted in the Discussion, there is no need to be constrained to generating virtual ruptures that are symmetrical with respect to the epicenter: if the rupture meets a source boundary, it should be able to propagate asymmetrically within the source zone. And there is also scope for considering alternatives between the end members of impenetrable and entirely permeable source zone boundaries.

Another recommendation we would make is that whenever there is a leaky boundary, hazard disaggregation to identify contributions from different magnitude-distance should be performed not only for the full hazard estimates but also for each individual source zone with leaky boundaries. This is particularly worthwhile if the leaky boundary could allow ruptures to propagate towards the site.

More than anything, we hope to have initiated a discussion within the seismic hazard modeling community and we hope that this will generate additional ideas for addressing the issues that have been highlighted herein. We believe that there may be real opportunities here to improve the state of practice in seismic source modeling for PSHA.

Data and Resources

No data was used in this paper. The seismic source zonation models used for Figures 1 and 3 have been deliberately anonymized since they are used for illustrative purposes only.

Declaration of Competing Interests

The authors have no competing interests but assume full ownership for the ideas expressed herein.

Acknowledgements

The authors are grateful to the participants of many SSHAC PSHA studies in which these issues have been raised and discussed, providing the stimulus for this paper, and generating interesting debates that have informed the ideas conveyed herein. Particular mention is due to Ken Campbell, Roland LaForge and Bob Youngs for insightful exchanges on the question of how virtual fault ruptures should be generated. We are very grateful to the Associate Editor, Professor Jack Baker, and an anonymous reviewer for constructive feedback that helped us to improve the manuscript.

For the third author, any opinions, findings, and conclusions expressed in this article are those of the author and do not necessarily reflect the views of the U.S. Nuclear Regulatory Commission.

References

Akkar, S., M. A. Sandıkkaya, and J. J. Bommer (2014). Empirical ground-motion models for point-and extended-source crustal earthquake scenarios in Europe and the Middle East. *Bulletin of Earthquake Engineering* **12**(1), 359-387.

Baker, J. W., B. A. Bradley, and P. J. Stafford (2021). *Seismic Hazard and Risk Analysis*. Cambridge University Press, Cambridge United Kingdom.

Bender, B., and D.M. Perkins (1982). *SEISRISK II, A computer program for seismic hazard estimation*. US Geological Survey Open-File Report 82-293.

Bindi, D., F. Cotton, S. R. Kotha, C. Bosse, D. Stromeyer, and G. Grünthal (2017). Application-driven ground motion prediction equation for seismic hazard assessments in non-cratonic moderate-seismicity areas. *Journal of Seismology* **21**(5), 1201-1218.

Bindi, D., M. Massa, L. Luzi, G. Ameri, F. Pacor, R. Puglia, and P. Augliera (2014). Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bulletin of Earthquake Engineering* **12**(1), 391-430.

Bommer, J. J., and S. Akkar (2012). Consistent source-to-site distance metrics in ground-motion prediction equations and seismic source models for PSHA. *Earthquake Spectra* **28**(1), 1-15.

Bommer, J. J., B. Dost, B. Edwards, P. J. Stafford, J. van Elk, D. Doornhof, and M. Ntinalexis (2016). Developing an application-specific ground-motion model for induced seismicity. *Bulletin of the Seismological Society of America* **106**(1), 158-173.

Bommer, J. J., and V. Montaldo Falero (2020). Virtual fault ruptures in area-source zones for PSHA: Are they always needed? *Seismological Research Letters* **91**(4), 2310-2319.

Bommer, J. J., F. O. Strasser, M. Pagani, and D. Monelli (2013). Quality assurance for logic-tree implementation in probabilistic seismic-hazard analysis for nuclear applications: A practical example. *Seismological Research Letters* **84**(6), 938-945.

Campbell, K. W. (1981). Near-source attenuation of peak horizontal acceleration. *Bulletin of the Seismological Society of America* **71**(6), 2039-2070.

Campbell, K. W., and N. Gupta (2018). Modeling diffuse seismicity in probabilistic seismic hazard analysis: treatment of virtual faults. *Earthquake Spectra* **34**(3), 1135-1154.

Donovan, N. C., and A. E. Bornstein (1978). Uncertainties in seismic risk procedures. *ASCE Journal of the Geotechnical Engineering Division* **104**(7), 869-887.

EPRI (2004). *CEUS ground motion project final report*. EPRI Report 1009684, Electric Power Research Institute, Palo Alto, California.

EPRI (2013). *EPRI (2004, 2006) ground-motion model (GMM) review project*. EPRI Report 3002000717, Electric Power Research Institute, Palo Alto, California.

Goulet, C. A., Y. Bozorgnia, N. Kuehn, L. Al Atik, R. R. Youngs, R. W. Graves, and G. M. Atkinson (2021). NGA-East ground-motion characterization model Part I: Summary of products and model development. *Earthquake Spectra* **37**(1_suppl), 1231-1282.

Graves, R., T. H. Jordan, S. Callaghan, E. Deelman, E. Field, G. Juve, C. Kesselman, P. Maechling, G. Mehta, K. Milner, and D. Okaya (2011). CyberShake: A physics-based seismic hazard model for southern California. *Pure and Applied Geophysics* **168**, 367-381.

Hale, C., N. Abrahamson, and Y. Bozorgnia (2018). *Probabilistic Seismic Hazard Analysis Code Verification*. PEER Report 2018/03, Pacific Earthquake Engineering Research Center, UC Berkeley.

Joyner, W. B., and D. M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the seismological Society of America* **71**(6), 2011-2038.

Kafka, A. L. (2007). Does seismicity delineate zones where future large earthquakes are likely to occur in intraplate environments? In: Stein, S. and Mazzotti, S. (editors), Continental Intraplate Earthquakes: Science, Hazard and Policy Issues, *Geological Society of America Special Publication* **425**, 35-48.

Leonard, M. (2014). Self-consistent earthquake fault-scaling relations: Update and extension to stable continental strike-slip faults. *Bulletin of the Seismological Society of America* **104**(6), 2953-2965.

Mai, P. M., P. Spudich, and J. Boatwright (2005). Hypocenter locations in finite-source rupture models. *Bulletin of the Seismological Society of America* **95**(3), 965-980.

McGuire, R. K. (2004). *Seismic Hazard and Risk Analysis*, EERI Monograph MNO-10, Earthquake Engineering Research Institute, Oakland, California.

Minson, S. E., A. S. Baltay, E. S. Cochran, S. K. McBride, and K. R. Milner (2021). Shaking is almost always a surprise: The earthquakes that produce significant ground motion. *Seismological Research Letters* **92**(1), 460-468.

Monelli, D., M. Pagani, G. Weatherill, L, Danciu, L., and J. Garcia (2014). Modeling distributed seismicity for probabilistic seismic-hazard analysis: Implementation and insights with the OpenQuake engine. *Bulletin of the Seismological Society of America* **104**(4), 1636-1649.

Moschetti, M. P., P. M. Powers, M. D. Petersen, O. S. Boyd, R. Chen, E. H. Field, A. D. Frankel, K. M. Haller, S. C. Harmsen, C. S. Mueller, R. L. Wheeler, and Y. Zeng Y (2015). Seismic source characterization for the 2014 update of the US national seismic hazard model. *Earthquake Spectra* **31**(1_suppl), S31-S57.

Scherbaum, F., J. J. Bommer, F. Cotton, H. Bungum, and F. Sabetta (2006). Ground-motion prediction in PSHA: a post-PEGASOS perspective. In *Proceedings of the First European Conference on Earthquake Engineering and Seismology*, Paper (No. 1312).

Scherbaum, F., J. Schmedes, J., and F. Cotton (2004). On the conversion of source-to-site distance measures for extended earthquake source models. *Bulletin of the Seismological Society of America* **94**(3), 1053-1069.

Stafford, P. J. (2014). Source-scaling relationships for the simulation of rupture geometry within probabilistic seismic-hazard analysis. *Bulletin of the Seismological Society of America* **104**(4), 1620-1635.

Thingbaijam, K. K. S., P. M. Mai, and K. Goda (2017). New empirical earthquake source-scaling laws. *Bulletin of the Seismological Society of America* **107**(5), 2225-2246.

Thompson, E. M., and C. B. Worden (2018). Estimating rupture distances without a rupture. *Bulletin of the Seismological Society of America* **108**(1), 371-379.

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

USNRC (2012). *Central and Eastern United States Seismic Source Characterization Project for Nuclear Facilities*. NUREG-2115, US Nuclear Regulatory Commission, Washington DC.

Villani, M., Z. Lubkowski , M. Free, R. M. Musson, B. Polidoro, R. McCully, A. Koskosidi , C. Oakman, T. Courtney, and M. Walsh (2020). A probabilistic seismic hazard assessment for Wylfa Newydd, a new nuclear site in the United Kingdom. *Bulletin of Earthquake Engineering* **18**(9), 4061-4089.

Wells, D. L., and K. J. Coppersmith (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the seismological Society of America* **84**(4), 974-1002.

Youngs, R.R., W. J. Arabasz, R. E. Anderson, A. R. Ramelli, J. P. Ake, D. B. Slemmons, J. P. McCalpin, D. I. Doser, C. J. Fridrich, F. H. Swan III, A. M. Rogers, J. C. Yount, L. W. Anderson, K. D. Smith, R. L. Bruhn,P. L. K. Knuepfer, R. B. Smith, C. M. dePolo, D. W. O'Leary, K. J. Coppersmith, S. K. Pezzopane, D. P. Schwartz, J. W. Whitney, S. S. Olig, and G. R. Toro (2003). A methodology for probabilistic fault displacement hazard analysis (PFDHA). *Earthquake Spectra* **19**(1), pp.191-219.

Author addresses:

Julian J Bommer
Department of Civil & Environmental Engineering,
Skempton Building
Imperial College London
South Kensington campus
London SW7 2AZ, UK
Email: j.bommer@imperial.ac.uk

Jon P Ake
Consulting Seismologist
2139 South Nile Street
Lakewood, CO 80228, USA
Email: jonake9@gmail.com

Clifford G Munson
U.S. Nuclear Regulatory Commission
1560 East 1850 South
St. George, UT 84790, USA
Email: Clifford.munson@nrc.gov