USGS review of the Central and Eastern United States Source Characterization for Nuclear Facilities Draft Report (July 31, 2010 version)

August 30, 2010

The Central and Eastern United States Seismic Source Characterization for Nuclear Facilities Draft Report dated July 31, 2010 represents a comprehensive review and analysis of hazard data, models, and methodologies for the Central and Eastern United States. Earthquake source information and hazard methodologies and framework developed for this study will be very influential in future seismic hazard analyses for this region, including the U.S. National Seismic Hazard Maps that will be updated in 2013. The USGS National Seismic Hazard Mapping Project made a significant effort during the month of August, 2010 to review the input documents and to understand the differences between the model and the 2008 National Seismic Hazard models for the Central and Eastern U.S. Because this review period was very short given the amount of material that we received, it was difficult to comprehensively review all chapters with equal rigor. Nevertheless, we have compiled a list of general comments and editorial changes and provide these to the TI team in the hope that they will be useful feedback for consideration in finalizing the model and report. The review comments have been written by individual scientists and because of time constraints it was difficult to come to consensus on each point in the review. In addition, we were not able to consolidate our comments with those of our colleague Art Frankel who is located in Seattle, WA. Therefore, we have separated our USGS comments into two parts: the first is a compilation of comments by the group in Golden, CO and the second contains comments by Art Frankel in Seattle, WA.

PART 1: Review comments of the National Seismic Hazard Mapping Project Members at USGS in Golden, Colorado

By Mark Petersen, Russ Wheeler, Steve Harmsen, Chuck Mueller, and Dave Perkins

General comment: Many of the references cited in the text are not listed in the References of chapter 10.

Executive Summary

The executive summary provides a basic description of the master logic tree, RLME and distributed seismicity sources applied in the analysis, and a description of the results. One point that should be discussed is why this model differs significantly from previous models including the USGS and COLA models. This is important in gaining acceptance of the document.

Pages viii-ix of the Executive Summary contain lists of source zones that do not match the lists in the Table of Contents and text.

Page ix last sentence on page states that if an alternative assumption or parameter is used in a seismic hazard study and it potentially changes the hazard by less than 25% from ground motions with hazards in the range 10-4 to 10-6 , it is within the level of precision with which one can calculate seismic hazard. In the Chapter 9 comments we discuss how the typical differences in ground motions between the CUES SSC and the COLA and USGS models are much larger than 25% so we are not sure if this result is reasonable. It may be the case if the same users developed the model using the same parameters and methods a second time that they would have this level of precision.

Acronyms

SHmax is defined as “maximum horizontal shortening”. In contrast, the Zobacks and others use SHmax to mean maximum horizontal stress, compression, or principal stress, with no implication of shortening or other deformation. Until horizontal pressure solution, seismic or aseismic faulting, or other deformation occurs, the crust merely supports the SHmax. Has SHmax been used anywhere to mean deformation instead of stress?

Chapter 1 – Introduction

The Introduction provides a clear description of the purpose of the document, the approach used in developing a SSHAC level 3 analysis, and a description of how the TI team determined that they were accomplishing the goals of obtaining the center, body, and range of the informed technical community.

1.1.5 p 1-4, paragraph 1: Extra comma after "facilities"?

1.1.5 p 1-4, paragraph 1: You state that the national seismic-hazard maps (NSHM) focus on annual frequencies of exceedance (AFE) of 10^{-2} to 10^{-3}. The documentation of the 2008 NSHM points out that the NSHM include maps with AFE as low as $4x10^{-4}$, and 35 of the 53 map sheets show this AFE (http://earthquake.usgs.gov/hazards/products/conterminous/2008/maps/). Thus, AFEs of the NSHM approach the upper limit of the SSC AFEs that are intended for design purposes.

1.2.2 p 1-5: The Section heading doesn't look right?

1.4.1 p 1-8, paragraph 3: "E" should be italicized in "RLME"?

1.4.4.5 p 1-11, paragraph 4: Missing word or words in first sentence?

Chapter 2 – SSHAC Level 3 Process And Implementation

Chapter 2 provides a necessary discussion of the meetings and goals of the project.

2.1.1.2 p 2-11, paragraph 2: Unclear phrasing in the second half of the crucial second sentence. The difference between PEGASOS results and the older results were shown to be caused by "an appropriate treatment of the ground motion aleatory variability and an error in the calculations in the previous hazard study …." Was the treatment appropriate in the older study or in PEGASOS?

2.1.2.1 p 2-17, paragraph 1: Extra “like” in last sentence?

Table 2-2 Names duplicated, misformatted, etc.: Al-Shukri, Ravat, Baldwin, Mueller?

Chapter 3 – Earthquake Catalog

The earthquake catalog is one of the most important components of the hazard analysis for the CEUS. The CEUS SSC catalog is a major improvement over previous catalogs in that it incorporates more regional catalogs and has developed moment magnitude estimates. This catalog is one of the most important contributions of the CEUS SSC project to the hazard community. The USGS reviewed the catalog and made several suggested changes. However, it is difficult to assess whether or not these changes were implemented in the final catalog.

3.1.2 p 3-2, paragraph 3: EPRI (1988) missing in References (Chapter 10)? Is this reference accessible to most readers?

3.1.3 p 3-3, paragraph 1: Should be “an important”, not “and important”? Should be “seismologists”, not “seismologist”?

3.2.1 p 3-3, paragraph 3: Should be “Mueller”, not “Muller”

3.2.1 p 3-3, paragraph 4: The list of USGS source catalogs is not correct. For CEUS the USGS uses: NCEER91; Sanford’s New Mexico catalog; Stover & Coffman; Stover, Reagor & Algermissen (state-by state catalogs); PDE; and DNAG. If the first sentence of the paragraph is rewritten accordingly, then should sources like SUSN, ISC, and ANSS be added to the list of regional catalogs in Section 3.2.2? If you are still using ISC as a source catalog, has it been filtered to avoid the problems discussed in the previous USGS review? The next sentence is confusing: Do you mean the primary USGS purpose, or the primary purpose of including these as separate sources in CEUS-SSC even though they had already been included by USGS? (I’m guessing the latter, but it’s not clear.) The CEUS-SSC catalog-construction scheme is described here relative to what USGS did, and this seems to be causing some confusion. A more straightforward description/separation might be clearer.

3.2.2 p 3-4, paragraph 1: Should be "for each earthquake" or "for earthquakes" in last sentence?

3.2.3 p 3-4, paragraph 2: "Boatwright" misspelled? "McCulloch" misspelled?

3.2.4 p 3-4, paragraph 4: Provide reference (web?) for the NEIC Mining Catalog? Should be "Section 3.2.3", not "3.2.4"?

3.3.1.1 p 3-5, paragraph 5: Clarify the location of the added blue square? Is it +0.2 mag units?

3.3.1.3 p 3-6, paragraph 2: Should be "a coda", not "an coda"?

3.3.1.3 p 3-6, paragraph 2: Should be "albeit", not "abet"?

3.3.2.2 p 3-7, paragraph 2: Should be "reasonably", not "reasonable"?

3.3.3.1 p 3-7: Provide a definition/description of M_N?

3.3.3.1 p 3-7, paragraph 5: should be "particular", not "particulars"

3.3.3.1 p 3-7, paragraph 5: I think "sources" means catalogs, but the text after eqn 3.3.3-1 suggests that earthquake location also matters. So for an earthquake in southern Canada with Canadian-network M_N, both $Z_{CAN} = 1$ and $Z_{NE} = 1$, right? Clarify this?

3.3.3.2 p 3-8: Provide a definition/description of m_{bLg}? Should m_{bLg} be included in the acronym table?

3.3.3.3 p 3-8: Provide a definition/description of m_b? The text & eqn refer to m_{bLg}? Ordinate label is cropped in Fig 3.3.3-3?

3.3.3.4 p 3-8: Provide a definition/description of M_L? Discuss truncation level (Fig 3.3.3-4)? Ordinate label is cropped in Fig 3.3.3-4?

3.3.3.1–3.3.3.4: Simple offsets between **M** and M_N (Fig 3.3.3-1) and **M** and m_{bLg} (Fig 3.3.3-2) do not fit the smaller-magnitude data (~ magnitude 3) very well. M_N, m_{bLg}, m_b, and M_L will begin to saturate somewhere in the magnitude 6–7 range, and will fall below **M** at greater magnitudes; the simple offsets do not account for this.

3.3.3.5 p 3-9: Provide a definition/description of M_S? The text refers to "local magnitude M_L"? Discuss meaning of very small M_S? Ordinate label is cropped in Fig 3.3.3-5?

3.3.3.6 p 3-9: Provide a definition/description of M_C? Discuss truncation level (Fig 3.3.3-6)? Ordinate label is cropped in Fig 3.3.3-6?

3.3.3.7 p 3-9: Should be subscript "D" in heading? Provide a definition/description of M_D? Discuss truncation level (Fig 3.3.3-7)? Ordinate label is cropped in Fig 3.3.3-7?

3.3.3.8 p 3-9: Should be subscript "U" in heading?

3.3.4 p 3-10, paragraph 1: Should be "Sections", not "Section s"?

3.3.4 p 3-10: I suspect that the derivations and applications of the different magnitude measures **M** and **M*** will be confusing to many readers. The discussion in old EPRI is pretty arcane, and many readers will not have access to it anyway. I wonder if the discussion here should be expanded, maybe with an example? How should these magnitudes be used?

3.3.4 p 3-10: How was the sigma "nominal value of 0.1" for instrumental **M** determined? What is the behavior of $\text{sigma}_{E[M|X]}$? Is sigma_M in eqn 3.3.4-4 and eqn 3.3.4-5 and Appendix B the same thing as sigma_P in eqn 3.3.4-3?

3.4 p 3-11: The declustering analysis is interesting. About the EPRI method, you state: "If the rate of earthquakes is significantly higher than the background rate …, then earthquakes are removed until the rate becomes consistent with the background rate." Does this mean that a few earthquakes that would be clearly declared as aftershocks by, say, Gardner & Knopoff remain in the final catalog in order to match the background rate? In other words, is the declustered catalog not strictly a catalog of mainshocks (in, say, the G&K sense)? My recollection is that G&K try to envelope the southern California cluster durations and dimensions, not match averages; comparisons to the EPRI CEUS cluster durations look OK to me, but the G&K time windows look too small (even when multiplied by 1.5).

3.4 p 3-11, paragraph 5: Should be “European earthquakes”, not “European earthquake”?

3.4 p 3-11, paragraph 5: Should be “more than the average”, not “more the average”?

3.5 p 3-12: Again, the completeness analysis relies on the reader having knowledge of the old EPRI study. What is the rationale for the EPRI completeness regionalization: geology, demographics? The latitude-longitude-bound (blocky) zonation does a poor job of defining some demographic or natural boundaries (for example, along the Atlantic margin).

3.5 p 3-12, paragraph 7: Should be “2009”, not “2008”?

3.5 p 3-13: How should TE be used?

Chapter 4: Conceptual SSC framework

Chapter 4 discusses the three attributes that are needed for a conceptual SSC framework, a systematic approach to treat alternatives for spatial distributions of seismicity and seismic sources, an approach to identify applicable data for source models, and a methodology for identifying seismic sources. The master logic tree shows the weights of the major branches for the Mmax and seismotectonic zones. The USGS review team did not understand the impact of the spatial smoothing in the two models on the resultant hazard estimates. In Part 2 there is a more complete explanation of this question.

4.1: NEEDS FOR A CONCEPTUAL SSC FRAMEWORK

(1) Concerning section 4.1, the USGS and SSC utilize logic trees to greatly different degrees, and this difference may not be reducible. The appropriate standard of practice for the SSC is the SSHAC procedures, which require inclusion of the range of credible alternatives, which in turn forces the extensive use of logic trees. In contrast, the standard of practice for scientists can be summarized as striving toward using reproducible evidence and its implications to select the one best answer. The two approaches overlap, but they differ fundamentally in focus. Additionally, seismologists' understanding of the geological controls on the physics of rupture nucleation and propagation in the CEUS is rudimentary. The different standards and poor understanding lead USGS scientists to reject most branches of most logic trees, which lead to a few small trees. The USGS does not reject the concept of logic trees entirely. This is evidenced by our summing of weighted hazard from three different choices of lower magnitude limits, and by our representations of epistemic uncertainty in Mmax and in the sizes of large earthquakes at Charleston and New Madrid. However, we should not expect the SSC and USGS to agree much on the preferable number and sizes of logic trees or on some of the conclusions that are based on them.

4.4: MASTER LOGIC TREE

(2) Figure 4.4.1–1 appears to be incomplete. The figure and its discussion need revision and expansion to clarify the following matters. The new material would be most effective if it were near the beginning of section 4.4.1, with details later as they are now.

First, apparently there are four kinds of zones: RLME sources, seismotectonic zones, Mmax zones, and distributed seismicity zones. The last two are not clearly defined and distinguished from the others. Citations of maps showing the different kinds of zones would help greatly.

Second, Figure 4.4.1–1 shows two kinds of conceptual approaches, Mmax zones and seismotectonic zones. There are two source groups, or kinds of Mmax zones: distributed historical seismicity zones (dhs hereafter) and RLME source zones. What is a "source group"? The dhs zones are used as background source zones to the RLME source zones (p. viii). Do the dhs zones exist everywhere including within the RLME source zones, so that the two hazards are summed? Additionally, the label “Mmax zones” implies that both source groups deals primarily with Mmax. If so, then it would be helpful if the link between Mmax and each source group were stated.

Third, there are also two source groups of seismotectonic zones: seismotectonic zones and RLME zones. If a conceptual approach has the same name as a source group, each mention of either is likely to be confusing and potentially misleading. The dhs zones “exist in different forms depending on whether the Mmax zones or the seismotectonic branches of the logic tree are being followed” (p. 5–6). However, Figure 4.4.1–1 does not show a dhs source group for the seismotectonic conceptual approach. The quoted sentence implies that the two forms in which the dhs zones exist are similar enough that they should have similar but not identical names in the figure and throughout the text.

(3) The captions of Figures 4.4.1.2–2, 6.2–1 and 6.2–2 use “wide” where “narrow’ should be, and vice versa. It is advisable to check usage of these two adjectives throughout the report. Usage of Mmax could be similarly checked to make sure that the zone name cannot be confused with the largest-magnitude sense of the term. Finally, different versions of some of the zone names are used in different parts of the text. Examples are in the Executive Summary; there may be other examples elsewhere.

Chapter 5 : SSC Model: Overview and Methodology

Chapter 5 provides a description of the spatial and temporal distribution of earthquakes in the CEUS. The USGS team appreciated the efforts to systematically assess the reasonable maximum magnitudes using the updated SCR catalog for each superdomain and generating a catalog by Monte Carlo simulation. The USGS review team was not clear on details of how the prior and likelihood distributions were constructed. The prior distributions are based on the superdomains and assume that the Mmax distribution is spatially variable. The likelihood distributions assume the largest earthquake in the catalog but do not consider the historical earthquakes (e.g., Charleston). By neglecting the RLME type sources it seems that the model assumes that all RLME sources have been identified. The Mmax distribution seems reasonable if there is a valid reason to throw out the large historic earthquakes.

The penalized maximum –likelihood method for smoothing seismicity also seems like a reasonable model as far as we can understand it. However, we do not understand how this method results in such high rates of earthquake exceedances compared to the USGS and COLA models (as described in Chapter 8 commentary). We do not understand all of the details in the penalized maximum approach so it is difficult to assess this methodology. Implementation of the penalized maximum-likelihood approach would be difficult for most of the user community, therefore, we feel that if this approach is maintained in the model the software to calculate the smoothed seismicity should be made available so that further comparisons could be made.

(1) Pg 5-1 Second paragraph, second sentence should be shortened.

Some Positive Items

We applaud the decision not to do time dependence for Central and Eastern sources. While time dependence almost surely characterizes single isolated fault sources, the sources in places like New Madrid and Charleston, and likely other places, are multiple, and if independent, as a group the occurrences would appear Poissonian. If dependent (or "contagious") clustering is rather more likely.

We also appreciate the use of likelihood functions for rates, as it more clearly expresses the recurrence rate uncertainties, especially for large events, say, in New Madrid.

Discretization of rates is also helpful, inasmuch as rates carry through proportionately in exceedances. For other continuous parameters we expect that Latin Squares for more finally divided discretization is more suitable for exploring hazard uncertainties.

We think it helpful to recognize in 5.3.2.4 that more realistic estimates of the uncertainty would be obtained by using bootstrapping from synthetic catalogs derived from alternatives models (even though this was not done).

Maximum Magnitude

Observed maximum magnitude data should be expected to be strongly rate dependent. The larger the sample derived from a distribution, the larger should be the fractile of the maximum of the sample. A low maximum magnitude from a source with low rates should not be a candidate for a prior to be applied to a source with high rates. As an example, using the data from Table 5.2.1-1 for EC type superdomains, figure A shows a plot of observed maximum magnitude against number of events above magnitude M. There is a clear dependence of maximum magnitude on observed number.

To correct for this dependence by the adjustment procedure proposed using equation 5.2.1-2 and illustrated in figures 5.2.1-3, and 4 is certainly reasonable. Thus, of the points in figure A, observed maximum magnitudes with low numbers of events above 4.5, either transform to values in the high 5, low 6 range or give no evidence at all. This suggests in general, that priors for maximum magnitude should have minima in the vicinity of magnitude 6. The examples given in figure 5.2.1-7, and 8 display this desirable quality. (On the other hand, the priors of figures 5.2.1-1 and 2, therefore should be hypothetical and not a typical example.)

The two-priors model of section 5.2.1.1.1 seems to us to reflect rather strongly the difference on the rate of earthquake occurrence of the two regimes, even with bias correction. Because prior distributions should not be lower than largest we can expect with no known observed structures, we believe that minima of these distributions should give very low probabilities that true maxima can be lower than 5.8, and should give credible (though low) probabilities for the exceedance of 7.9. The model prior with the lower mean gives around 15 percent probability to the former and 0.003 to the latter. The composite-prior model, on the other hand, seems to us to satisfy our criterion, and be more correct for philosophical reasons, as well.

In practice, we notice an unfortunate consequence of the normal model for the prior. There is always some probability of a maximum magnitude in the tails, no matter how extreme. As an example of what happens, figures 7.4.2-2, and 3 show a characteristic double maximum in the posterior. This is coming from the interaction of a strongly spiked likelihood function with the small-magnitude tail of the prior at magnitudes below 5.8. If the lower tails of the priors were truncated below, say, 5.6, this bimodal behavior could not occur.

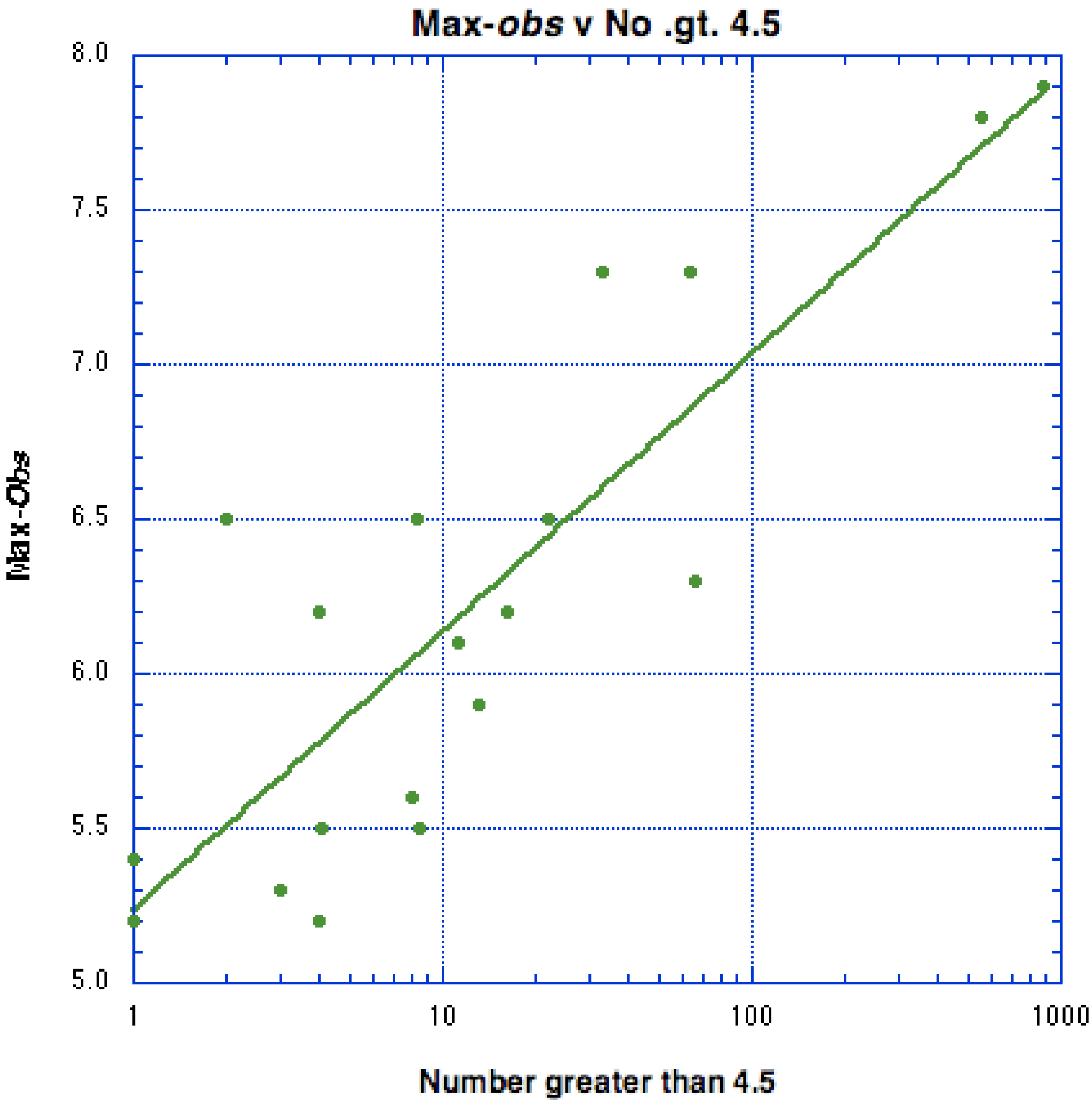


Figure A. Max-obs versus logarithm of the number of events greater than 4.5 from Table 5.2.1-1 for EC-labeled superdomains. Note the strong correlation.

Concluding General Remarks on Maximum Magnitude

Concerning section 5.2, there is a fundamental disagreement about Mmax between the USGS, on the one hand, and EPRI-SOG and SSC on the other. EPRI-SOG made a fundamental assumption that Mmax varies spatially in SCRs. SSC adopts the assumption and implements it with a Bayesian analysis. USGS cannot see any justification for the assumption. However, we also cannot disprove it. Accordingly, USGS sidesteps the matter and makes what it can of the sparse data contained in the global catalog of SCR earthquakes. Accordingly, SSC uses a more elaborate and mathematical methodology than does USGS, although not necessarily a more objective or reproducible one. EPRI-SOG's Mmax values tended to be significantly lower than the USGS values. SSC values appear to be closer to the USGS values. So far, no one proved or disproved either approach. We see no reason to expect that this fundamental disagreement will be resolved any time soon.

Spatially varying b-values.

Inasmuch as b values are almost as dependent on sample size as maximum magnitude values are, and because maximum likelihood b-values are also dependent on the maximum magnitude observed, we doubt that b-values should be allowed to vary spatially in most of the CEUS. In figures 5.3.2-16 and 17, we find that the patterns of changing b-values over the range 0.74 to 0.85 seems to reflect the methodology's allowance of variation of a-value or variation of observed maximum magnitude to affect the b-value determination (notably, in New England, oppositely between the two figures. We are not convinced that a statistical distinction within this small range can be justified by the relatively small number of historical events available.

Predictive spatial patterns from observed seismicity.

Previous work has shown that alternative subjective interpretation of source zones is a major source of hazard map variability from expert to expert. We commend the use of more objective kernel smoothing methods as a way of setting bounds on likely variability.

On the other hand, it is well known that randomly located points in areas show apparent spatial clustering analogous to temporal clustering in Poisson time processes. A predictive procedure that adapts to show detail at the scale of what may be random clustering is ill-advised, unless that clustering is also associated with geologic or seismogenic structure. To us, the rate variations shown in figures 5.3.2-16, 17 and 18 all appear overly detailed. Insufficiently broad spatial smoothing may account for the unanimity seen in alternative recurrence estimates seen in figure 5.3.2-20 through 26, for instance, which we find very unlikely.

For this reason, we continue to prefer using multiple-scale constant kernel smoothers with what are apparently larger scale than those of the SSC methodology.

On the other hand, the range in values obtained for Minnesota, with reduced M weights (Figure 5.3.2-29 seem to reflect likely uncertainties where there is so little historical information.

Recurrence determinations

Figures 5.3.2-20-23 worry us. Here rates are fit, by maximum likelihood, to data from magnitudes less than 3.75, presumably adjusted for incompleteness. Any uncertainty in the completeness correction produces additional uncertainty in already uncertain historical recurrence rates (and magnitudes). Maximum likelihood fitting strongly depends on the largest rates observed, and these must be the lowest magnitude. Reducing the M weights mitigates some of this, but we would think that predicted rates which are double the observed rates at magnitude 4 should be considered disqualified. Later in the report are shown hazard curves at some locations which are double the exceedance probability of those both of the USGS and of the [COLA] models.

Is there perhaps some maximum likelihood model that penalizes a fit that falls outside the error bounds of the observed data?

Specific Comments on the Text
Section 5.1.2. The sentence,, "With time and study the earthquake research community in the WUS has developed information that suggests that the assumption of an exponential distribution of magnitudes may not be appropriate for individual faults." is probably the basis for considering only characteristic sources for the RLMEs, rather than considering branching which includes the truncated exponential distribution with Mmin in the neighborhood of M6.5 and Mmax at Mchar.

The exclusion of models having truncated exponential source frequencies on faults such as the Cheraw fault in southeast Colorado has a strong effect on hazard. The above view is embodied in the California A-fault model, on which just big events are believed to occur. However, for Quaternary faults in all the rest of the western US, and for b-faults in California, the NSHMP hazard model continues to give the GR branches 1/3 to 1/2 weight, that is, weight on par with the characteristic source model. It looks like the current SSC has weighed in on identifying all faults in the CEUS as A-faults, in the California sense, even though not much is known about CEUS faults compared to California A-faults.

It is not clear that the methodology is considering the possibility that some CEUS faults may be more like b-faults, where small M events may be somewhat more likely than large M events. The effect on hazard is explored in the Harmsen review of the new EPRI Cheraw RLME, for example. When not much is known about seismogenesis on specific faults in the CEUS, it is hard to believe that smaller sources simply cannot occur, or can only occur at frequencies that make them immaterial to the seismic hazard when compared to larger events.

The use of a truncated GR distribution for Mmax and distributed zone seismicity where essentially nothing is known may also be considered questionable. Such an assumption essentially denies the possibility that other RLME source zones may exist, as yet undiscovered. The turning on and off of known RLME source zones should give the authors pause, that is, other zones may exist but may be at this moment in an out-of-cluster mode. Great benefit to society would be accrued by discovering such a zone, if such a discovery included proof that it will turn on in the next 50 years. The promulgation of models that essentially deny the existence of such zones may discourage the search for as yet undiscovered RLMEs.

Section 5.2.1. The second paragraph on p. 5–7 notes that the global SCR catalog contains earthquakes that occurred in RLME sources, whereas the Bayesian method is being applied only to CEUS sources that are not RLME sources. The suggestion is made that using global RLME sources might lead to overestimation of Mmax in the CEUS non-RLME sources. Why would Mmax be potentially overestimated?

Again concerning the second paragraph, RLME sources lack recognized geologic records of large cumulative Quaternary deformation. This implies that probably the RLME sources have not generated RLMEs throughout geologically long periods of time. Absent a clear preference for RLMEs to be in specific geologic settings, the likelihood that presently active RLMEs have geologically short lifetimes suggests that RLME sources turn on and off at unknown intervals. The SSC report lists ten RLME sources, two of which are in parts of the craton that are sparsely seismically active to aseismic. Five of the ten RLME sources were not known 5-10 years ago. Thus, it seems likely that more RLME sources exist in the CEUS but are presently inactive or undiscovered. It is seems likely that RLMEs can occur anywhere within the CEUS. Furthermore, prehistoric and historical surface ruptures are abundant and weakly clustered spatially in much of the Australian SCR. The diffuse clusters of Australian scarps could be regarded as RLME sources. This reasoning suggests that RLME sources worldwide may be appropriate estimators of the value of Mmax, but not its frequency, in the CEUS outside RLMEs.

Section 5.2.1.1. In the SSC Bayesian estimation of Mmax, a bias correction deals quantitatively with the likelihood that Mmax exceeds the largest observed magnitude. We are not aware of any published test of the bias correction. The choice of a test is not straightforward because the bias correction estimates the median of the range of Mmax instead of a single value. This difficulty does not reduce the need for a test to increase confidence in the bias correction. An interesting example test could be presented, say for the Wabash area, where there is a limited amount of historical seismicity, but some paleoestimates of large magnitudes.

Section 5.2.1.1. Pages 5–9 summarize the pooling of domains into superdomains according to shared values of geologic and seismological variables. C.A. Cornell, in proposing the pooling, listed eight variables that might be used, including those used in pages 5–9. None of Cornell's variables were justified or explained. Of the variables, only crustal type and crustal age have been observed to be associated spatially with SCR earthquakes of M7.0 or larger worldwide. What evidence supports use of any of the other variables? Statistical tests are unlikely to help because there are 246 ways of choosing two or more of the eight variables and 154 ways of choosing two to four of the variables. Given this many possibilities, it is probable that chance alone would cause a statistical test to identify spuriously several of the possible choices as significant.

Sections 5.2.1.1. Pages 5–10 through 5–12 present a multi-step process for identifying variables that are associated from step to step with an increasing ability to distinguish groups of superdomains according to their mean values of maximum observed magnitude. It is not clear from the text whether the analysis was done in an exploratory sense, without formal testing of hypotheses, or in the standard statistical way of formulating a testable hypothesis and the corresponding null hypothesis. If it were done in an exploratory sense, then probably it would be valid to use the p-values as descriptors, although some statisticians might disagree. However, tests of hypotheses and conclusions about significance would be invalid. In contrast, if the analysis were done in the usual way, then the hypotheses to be tested would have to be determined before examining the data. Otherwise the tests of significance would be invalid. Which was done should be stated clearly at the outset to avoid misinterpretation and confusion about validity. Additionally, most of the p-values are not stated. Providing them would help the reader follow the argument.

Other matters also could be stated more clearly. (a) In standard hypothesis testing, is it necessary to select a significance value, such as 0.05, before performing the test? (b) Was the null hypothesis rejected or not? (c) Other tests can give the probability that the alternative hypothesis is true; what were their results? (d) Some of the samples were so small that the t-test might not have had much power to detect differences that exist; then the true p-value may have been smaller than 0.14, which would weaken the conclusion. Did a test less sensitive to sample size give similar results to the t-test? (e) For small samples the t-test is sensitive to non-normal distributions. Did another test less sensitive to small sample size give similar results to the t-test?

The analysis involves a family of similar statistical tests. The similarity creates the problem of multiple comparisons, by which a single test with a significance level of 0.05, repeated n times in similar fashion, increases the probability of at least one spuriously significant result to approximately n(0.05). Therefore, to avoid one or more spuriously significant results, the significance level must be reduced. Determination of the exact amount of reduction is a complex problem, but the simplifying Bonferroni inequality suggests reduction to about 0.5/n, or in this case about 0.01. The reduction further weakens the interpretation of the p-value of 0.14. The other p-values in the analysis are not presented; were any p-values less than about 0.01 obtained at the end of the analysis?

Section 5.2.1.1.5. This section is insufficiently clear. It can be interpreted as saying that instead of using only the historical maximum magnitude, a distribution of maximum magnitudes is calculated by simulating 100,000 sets of magnitudes and taking the maxima, binning these to make a distribution, and then applying the bias correction procedures and developing a likelihood function for each 0.1-unit magnitude in the distribution, and getting a posterior for each, the posterior being weighted in accordance with the distribution. It seems to say, then, that observed maximum magnitude no longer plays a role, and it implies that the observed number of historical events alone controls the maximum magnitude distribution. Is this correct? Please make the section more specific.

Section 5.2.1.2. We applaud the effort to compare the Kijko with the Bayesian approach. The result seems to be that we can have confidence in the method when it more or less agrees with the Bayesian. Although a rule is given for assigning 0 weight to the Kijko approach, it seems that, in general, one would never be secure in adopting this approach when the number of observed events is moderate, without also attempting a Bayesian approach as a check. This suggests that a Bayesian approach should always be preferred. Is this a fair assessment? From your experience is it possible to provide other rules of thumb for using/not using the Kijko method?

Section 5.3.2.1.2. It is not completely clear to us from the *details* of how the MCMC procedure described here is related to the usual uses of MCMC, namely sampling from posteriors of distributions whose joint discription would be too complex, or exploring the joint distribution of parameters compatible with some objective criterion, like minimal square error of predicted vs observed data, as in Bayesian inversion. Please provide an *overview* of the procedure.

Section 5.3.2.3.3. Typo. Page 5-33, first paragraph, line 4, Figure 5.2.3-16 should be 5.3.2-16.

Section 5.3.2.3.3. Last paragraph. “Although there are no events within this area, the model predicts activity rates that are comparable to those obtained for Miami.” Thanks for pointing this out, but you have no comments as to whether this is good or bad. From a structural point, one should not expect similarity. From a statistical point, since, for Poisson occurrences, when 1 event is expected, the probability of observing 0 is the same as the probability of observing 1, the process may be respecting this.

Section 5.3.2.4. Typo. Page 5-33, last paragraph, last line, Figure 5.2.3-1 should be 5.3.2-1

Section 5.3.3.2. The last paragraph on p. 5–38 states that the coefficient of variation α of the interevent time t in CEUS RLME sources is estimated from data that had been used to estimate earthquake probabilities in the San Francisco Bay area. The paragraph explains that this was done because data are few from the CEUS RLMEs. Presumably most of the larger datasets used in the San Francisco Bay study are from active continental regions and plate boundaries. Most faults in those areas have slipped farther and faster than CEUS faults, so that the former are likely to be the more mature. What evidence or reasoning supports the use of active-area data to estimate α of CEUS RLME sources? Are CEUS data sufficient to tell whether or not they are consistent with α from active areas?

The same question arises in section 5.3.3.4 with respect to assuming that the aleatory uncertainty of the magnitude of an RLME is 0.25.

Section 5.3.3.3 discusses uncertainties in ages, some of which were represented as normal distributions. Laboratory dates are commonly given with standard deviations of the analytical uncertainty. However, the geological uncertainty of the age that the date constrains may commonly be larger, comprising uncertainties in the various structural and stratigraphic relations seen in a trench. The geologic uncertainties may have distributions of various shapes, some of which may be asymmetrical. How were these complexities incorporated?

Section 5.3.3.4. (See the comment at the end of the one on section 5.3.3.2.)

Chapter 6: SSC Model, Mmax Zones Branch

Chapter 6 describes one of the major branches of the master logic tree for the CEUS SSC model. This branch, however, is only given 20% weight.

6.1.1: CHARLEVOIX RLME SOURCE

(1) Figure 6.1.1–1, which shows the Charlevoix RLME source, shows the M of the 1663 earthquake as 3.71. Also, the M of the 1925 earthquake is actually 6.3, not 6.2. Bent's 1992 paper lists M6.2 but it also lists the moment as 3.1 x 10^{25}. The moment gives M6.3. She and I were not able to determine where the 6.2 came from. She advised me to always check M by calculating it from the moment.

6.1.2: CHARLESTON RLME SOURCE

(2) Section 6.1.2.1 uses the example of the Eastern California shear zone and its clear geomorphic expression to argue that the Charleston seismic zone, which lacks such an expression, may have a variable recurrence rate. The point is well taken, but the example is not the best. The Eastern California shear zone is in a dry climate. Additionally, SHmax orientations

indicate that the causal fault or faults of the Charleston seismic zone may be largely strike-slip. Is there a better example from a humid climate?

(3) Section 6.1.2.3.1 states that a northeast strike is assumed for future rupture zones in the Charleston Local source. Figure 6.1.2–4 shows the gravity and aeromagnetic expressions of the north edge of the broad South Georgia rift trending easterly across the northern part of the Local source. The rift contains the Mesozoic rocks that are widespread beneath the Charleston source zones. State geologic maps show that eastern seaboard Mesozoic basins have short faults of many orientations, but long faults parallel to the basin boundaries. Should not some weight be given to east-striking rupture zones? The same question applies to the Charleston Regional source of section 6.1.2.3.3.

(4) Section 6.1.2.3.2 explains that the Charleston Narrow source is truncated at the north end because the strength of the evidence for its existence decreases from south to north. Which specific features show that the zone should be truncated where shown in Figure 6.1.2–4?

6.1.3: CHERAW FAULT RLME SOURCE

(5) "Average and Maximum Displacement" in section 6.1.3, paragraph 1, reports displacement for the three ruptures as 1.5 m, 1.1–1.6 m, and 0.5–1.1 m, listing the oldest first. After noting that the second two ruptures might have been one, the discussion lists estimates of displacements differently as 1.1 m, 1.5 m, and 1.9 m. These sum to 4.5 m, whereas the other three displacement estimates sum to only 3.1–4.2 m. Please explain the origin of the three final displacement estimates.

(6) In section 6.1.3 "Rupture Area" states that magnitude was estimated from rupture length, fault dip, and fault width. However fault displacement is generally preferred by paleoseismologists because length can be severely censored by erosion. What effect would use of displacement have on the final magnitude distribution?

(7) Section 6.1.3.4, paragraph 6 gives recurrence intervals of 2–5 centuries for the out-of-cluster alternative. Is this an error?

6.1.4: MEERS FAULT RLME SOURCE

(8) Section 6.1.4.4 states that rupture length is preferred to displacement as a magnitude estimator because length requires fewer assumptions. Displacement is subject to uncertain interpretations of structural and stratigraphic relations in a trench. It is not clear how large these uncertainties might have been. However, scarp length is recognized as being subject to severe censoring by erosion over time. The ends of the 26-km section of the fault are constrained by the changes in geomorphic expressions into the two adjacent sections. The locations of these geomorphic changes may have uncertainties measured in kilometers. What are the uncertainties in displacement and scarp length, and how much magnitude uncertainty would each estimator produce?

(9) Also in section 6.1.4.4, the single-event displacements measured by Crone and Luza were carefully made and well constrained. They may have better accuracy and precision that Swan's measurements of displacement. If so, why were the better displacement measurements not used?

6.1.5: REELFOOT RIFT - NEW MADRID FAULT SYSTEM RLME SOURCE

(10) The fourth paragraph on p. 6–31 discusses alternative locations for the January 23, 1812 New Madrid earthquake. Not mentioned is the 2006 abstract of Cramer and others, which reports that rupture directivity can explain the evidence interpreted as supporting a location in White County, Illinois.

(11) Tables 6.1.5–1 through 6.1.5 –3 are missing.

(12) Paragraph 3 on p. 6–38 summarizes some observations of Tuttle and others. The paragraph omits their observation that the Bhuj and New Madrid liquefaction fields cover comparable areas. Area is a more robust estimator than is the distance to the single largest feature.

(13) Page 6–39 states that equal weights are to be given to the magnitude estimates of Bakun and Hopper, Hough and Page, and Johnston. Some members of the informed technical

community might down-weight the Hough and Johnston estimates because they have not been published for others to examine and critique. Additionally, it is not clear whether the low estimates around magnitude 7.0 are consistent with paleoseismic evidence. The most direct way to justify such low estimates would be to compare the sizes, abundance, and geographic extent of liquefaction features caused by an instrumental M 7.0 earthquake with the features attributed to one of the very large New Madrid earthquakes. Tuttle and others made such a comparison for the 2001 Bhuj earthquake of magnitude 7.6, and found the two liquefaction fields to be comparable. Equal weights on the high and low magnitude estimates, and especially on the low estimate, may be hard to defend.

6.1.9: WABASH VALLEY RLME SOURCE

(14) Figure 6.1.9–2 lists earthquake epicenters in the map explanation, but no map symbol is shown and no epicenters are on the map. The cited paper does not include a map or list of feature locations. The map does not show any paleoliquefaction features in Kentucky, although Appendixes C and D cite a report of sand dikes by Counts.

6.2.1: CRITERIA FOR DEFINITION OF BOUNDARY (of alternative Mmax zones)

(15) In the first paragraph of section 6.2.1, the most inboard of the three Atlantic domains includes the Mesozoic grabens and half grabens that are stated to underlie the continental shelf. However, the Mesozoic extensional basins extend also underlie the Coastal Plain and extend northwestward into the exposed Appalachian crust.

(16) The first paragraph of section 6.2.1.1 concludes that the ECC-AM, AHEX and NAP seismotectonic zones should be included in the Mesozoic Narrow Extended zone. The paragraph cites good evidence for the first two, but the sparse cited evidence for the NAP would be strengthened by citations of the 1989 paper and 1993 map and cross section by Stewart.

(17) The second paragraph on p. 6–64 includes the GMH seismotectonic zone in the Mesozoic Extended Zone on the grounds of Cretaceous volcanism that has been attributed to the passage of a hotspot. We know that the GMH has unusually abundant low- to moderate-

magnitude seismicity. Do we know anything else about the GMH that is pertinent to the occurrence or frequency of larger earthquakes? Adding the GMH to the Mesozoic Extended Zone requires evidence for Mesozoic extension within or below the seismogenic zone. Does the chemistry or mineralogy of the volcanic rocks indicate extension at melting depths? Are any extensional faults known? The hot spot hypothesis is consistent with Mesozoic extension, but consistency does not imply validity or relevance. In what way does the hotspot hypothesis contribute evidence of Mesozoic extension that penetrates well into or through the seismogenic zone?

6.4: RECURRENCE PARAMETERS

(18) Figures 6.4–1 through 6.4–6 show b-values of about 0.8, plus or minus about 0.03. It was worthwhile to demonstrate that variability is so small. Assuming a value of 0.8 everywhere would prune the logic tree. Would the hazard be much affected if a constant value of 0.8 were used? Why should it be used or not used?

CHAPTER 7: SSC MODEL, SEISMOTECTONIC ZONES BRANCH

Chapter 7 describes the heavily weighted (80%) branch of the master logic tree for sources. This chapter shows the depth of information that was reviewed in developing the CEUS SSC model and provides geologic and geophysical evidence for the source and its tectonic framework as well as evidence for activity. The analysis applies seismicity, magnetic, gravity, and paleoseismic data to assess these parameters.

The USGS team had trouble understanding the poor fits of the catalog compared to the realizations on several figures (e.g., 7.5.5-11 to 7.5.2-30). Is this the reason that the CEUS SSC model has such high rates of exceedance?

7.3.2: GREAT METEOR HOTSPOT SEISMOTECTONIC ZONE

(1) Section 7.3.2 attributes the Monteregian Hills alkalic igneous rocks of Cretaceous age to formation of a northwest-trending hotspot track in Cretaceous time. However, the alkalic

rocks align east-west. The geologic maps of Ontario and Quebec show several east-striking faults in, south of, and west of the outcrop area of the alkalic rocks. Elsewhere the report summarizes evidence for Cambrian and Mesozoic regional extension to form inner parts of plate-scale passive margins. Why should the west-trending alkalic rocks not be attributed to Cretaceous reactivation of earlier Mesozoic or Cambrian normal faults, instead of to a northwest-trending hotspot track?

(2) Section 7.3.2.4 explains why the seismogenic crust in the seismotectonic zone is modeled as 25–30 km thick. This is an unusually thick seismogenic layer for central and eastern North America. Is the unusually thick brittle crust expected in an area that has been uplifted and heated by a hotspot?

(3) Table D–7.3.3 lists four arguments by McHone against the hotspot hypothesis. Have McHone's arguments been addressed in the literature? If not, does other evidence support a separate Great Meteor Hotspot seismotectonic zone? If not, what would be the impact on hazard if the seismotectonic zone were deleted, leaving only its abundant seismicity?

7.3.4: PALEOZOIC EXTENDED SEISMOTECTONIC ZONE

(4) Section 7.3.4 cites Data Evaluation Table C–7.3.4, but Appendix C does not contain such a table.

(5) Section 7.3.4.1.2 refers to the Chain Lakes massif as "coastal". The massif is on the Maine-Quebec border, far inland. The same error occurs on page 7–43.

7.3.5: ILLINOIS BASIN EXTENDED BASEMENT SEISMOTECTONIC ZONE

(6) The first paragraph refers to "subsequent studies". What are they?

(7) Figure 7.3.5–1 shows an energy center on the Meramec River. The corresponding earthquake is assigned the same magnitude range as the Iona and Elnora earthquakes of the Illinois Basin Extended Basement seismotectonic zone. The first paragraph of section 7.3.5 explains that the prehistoric earthquakes are part of the reason for defining the Illinois Basin

Extended Basement Zone. Yet, the Meramec River energy center is excluded from the zone. Please explain the reason for the exclusion, and cite appropriate references.

(8) The beginning of section 7.3.5.1 lists five characteristics of the southern Illinois Basin that may influence Mmax and the properties of future earthquakes, and which support delineation of the Illinois Basin Extended Basement seismotectonic zone. The first characteristic is an unusual number of prehistoric and historical earthquakes. The prehistoric earthquakes are known because the southern Illinois Basin has been searched for paleoliquefaction features, whereas no other similarly large part of the Midcontinent has. The other four characteristics are likely results of the basin's having been much more intensely characterized by exploration for and production of petroleum and coal than most other parts of the Midcontinent. Thus, only the unusually abundant historical seismicity appears to distinguish the southern Illinois Basin from other basins throughout the Midcontinent. Yet in 1988 Coppersmith pointed out that seismicity probably migrates within the central and eastern U.S., at unknown intervals. Migration implies that some factor other than long-lived faults, folds, uplifts, and basins likely controls whether and where most seismicity concentrates. Therefore, the stated reasons that support the first sentence of section 7.3.5.1 may not be valid. Please justify the sentence.

(9) As stated in the text, the boundaries of the Illinois Basin Extended Basement Zone are highly uncertain because they are estimated from the geographic distributions of several geologic structural types, stratigraphic sequences, and potential-field anomalies, each of which has its own uncertainty. Section 7.3.5.4 explains how other kinds of uncertainty are incorporated into the hazard-computation model. Please explain how geographic uncertainty in the zone boundary is incorporated.

7.3.6: REELFOOT RIFT SEISMOTECTONIC ZONE

(10) Figure 7.3.6–1 shows geologic features in and around the Reelfoot Rift. The steeply dipping rift-bounding faults are dominantly normal, with minor contractional reactivation. Why are they shown with the standard map symbol that identifies shallowly dipping thrust faults? Do other figures contain the same error?

(11) The last sentence of the first paragraph on p. 7–41 misstates two of the three kinds of information from which Wheeler defined a boundary between the Reelfoot Rift and the Rough Creek Graben. First, across the boundary, the largest Cambrian normal slip switches from a fault on one side of the rift-graben system to another fault on the other side of the system. Thus, both of the faults with the largest slip and consequently the greatest likely width die out at the boundary. The discontinuity might limit propagation of large ruptures. The amounts of Cambrian normal slip are recorded by the thicknesses of Cambrian sandstones. Fault reactivation after the Cambrian rifting could have reversed part of all of the Cambrian slip, thereby affecting the dip direction of the basement/cover contact within the rift-graben system. However, the thicknesses of the Cambrian strata would have been preserved. Thus, the dip direction of the basement-cover contact is not pertinent to the location of the boundary. Second, the northeast limit of alkalic igneous rocks might indicate the northeast limit of extensional rift faults that are wide enough to sample alkalic melts that were generated by extension, and which are therefore wide enough to contain large rupture zones. In contrast, the northeastern limit of Mesozoic strata of the embayment is controlled largely by erosion and by a broad, regional downwarp southwest of the boundary between the rift and the graben. The distribution of Mesozoic strata is not pertinent. The following paragraph repeats the error with respect to Mesozoic strata, and adds an unclear statement about Mesozoic deformation.

(12) Section 7.3.6.4 appears to assign M 6 to both the 1843 and the1895 earthquakes. However, intensity VI, VII, and VIII were reported over much larger areas in 1895 than in 1843. There were more reports in 1895 because the region was more densely settled, but density of reports seems unlikely to explain the greater distance of reports from the epicenters. The highest intensity reports from both earthquakes were concentrated along major rivers, so it is not clear that greater amplification of the 1895 shaking might explain the larger reporting areas. What is the evidence that the two earthquakes had the same magnitude?

7.3.7: EXTENDED CONTINENTAL MARGIN - ATLANTIC MARGIN SEISMOTECTONIC ZONE

(13) The last paragraph on page 7–43 states that a topographic high would resist extension. By what mechanism would high topography resist extension, rather than encouraging it as in orogenic collapse?

(14) The first paragraph of section 7.3.7.1 reports a “highly significant” finding of a difference between rifted and unrifted crust, and “a statistical basis for separating Mesozoic and younger extended crust to establish a prior distribution of Mmax.” Pages 5–10 through 5–12 appear to present a contradicting statistical assessment of the separation between Mesozoic and younger extended crust, on the one hand, and other crust, on the other hand (although see one of the comments on section 5.2). Please resolve the apparent inconsistency.

(15) Section 7.3.7.1 lists three criteria for distinguishing the more seaward northern part of the Extended Continental Crust - Atlantic Margin Zone (ECC-AM) from the onland Northern Appalachian Zone (NAP). First, the multiple phases of NAP reactivation do not seem like a valid criterion because the reactivation is pre-Mesozoic and, therefore, it may be unlikely to have affected the seismic potential of any Mesozoic faults. Furthermore, at least part of the recognition of multiple phases might be caused by the better exposure in the NAP compared to that in the ECC-AM.

Second, the absence of exposed Mesozoic rift basins in the NAP does not appear to be a valid criterion, because the greater uplift there could have allowed erosion to destroy the sediments and sedimentary rocks that filled the basins, leaving only the normal faults whose movement created the basins. In fact, the 1991 cross sections of Stewart, which are cited in the reviewed report, show several interpreted examples of NAP faults whose basins presumably have been eroded. If the beheaded faults are of Mesozoic age, then Johnston’s observation would imply potentially similar Mmax in NAP and ECC-AM. Additionally, Figure 7.1–7 shows that the difference between NAP seismicity level and ECC-AM seismicity level is similar to the differences along trend within the ECC-AM. Finally, examination of state geologic maps from Connecticut to North Carolina shows that Mesozoic rift basins contain faults of all orientations, although the longer ones parallel the local basin trend. Thus, except for the largest earthquake ruptures, the difference in structural grain between the NAP and the ECC-AM does not seem likely to produce significantly different preferred orientations of strikes of Mesozoic faults. It seems to be difficult to argue for a difference in the numbers, styles, orientations, or sizes of Mesozoic faults between the NAP and the ECC-AM.

Third, absent such differences, it seems equally difficult to argue for a difference in seismicity characteristics or rupture characteristics between the NAP and the ECC-AM. In sum, a convincing argument for the existence of a distinct NAP seismotectonic zone has not been presented.

(16) The second sentence of section 7.3.7.3 refers to the Extended Continental Crust - Atlantic Margin Zone as "highly extended, transitional crust". Those words do not describe the onland central and southern Appalachians that comprise roughly half of the zone. The words are better applied to the Atlantic Highly Extended Crust of section 7.3.8.

7.3.8: ATLANTIC HIGHLY EXTENDED CRUST SEISMOTECTONIC ZONE

(17) Section 7.3.7.4 states that the strike expected of future ruptures is taken as the default distribution of Table 5.4–1. The distribution is strongly asymmetrical about N35°E. The strong asymmetry is inconsistent with the little that is known about the Atlantic Highly Extended Crust Zone (AHEX).

As the report summarizes, the AHEX is far offshore under the shelf edge, has crust that thins southeastward, is defined by linear potential-field anomalies, is sparsely seismically active, and is defined by the long, thin East Coast Magnetic Anomaly. The anomaly trends approximately N20°E south of New Jersey and about N60°E north of New Jersey. The anomaly is attributed to abundant rift-related volcanic rocks within the AHEX.

The most likely properties of any faults in the AHEX can be suggested from the generalized characteristics of the AHEX. Any continental crust is probably highly extended. Large extension tends to produce long faults that strike at high angles to the extension directions, as exemplified by the mapped border faults of Mesozoic grabens from Massachusetts to North Carolina. At least the older of the volcanic rocks may also contain extensional faults. The faults are likely to strike perpendicular to the direction of extension, which would be more or less parallel to the AHEX and approximately N20°–60°E. Emplacement of large amounts of volcanic and perhaps intrusive rocks might disrupt this simple fault geometry. These considerations argue against application of a strongly asymmetric default distribution of fault strikes (Table 5.4–1). If there is a preferred strike, it is likely to be about the same as the mode of the default distribution.

7.3.9: EXTENDED CONTINENTAL CRUST-GULF COAST SEISMOTECTONIC ZONE

(18) The third bullet in section 7.3.9.2.1 states that aeromagnetic anomalies are margin-perpendicular in the Extended Continental Crust-Gulf Coast Zone (ECC-GM). However, the 2002 Magnetic Map of North America shows margin-perpendicular aeromagnetic anomalies only in central Texas. Please cite whichever aeromagnetic map supports the bullet, and explain why it is preferred over the 2002 map.

(19) On page 7–55, second paragraph, the description of the northern edge of the ECC-GM includes the phrase “The source zone is extended to the Southern Arkansas fault zone ….” The preceding sentences indicate that the source zone extends north of the fault zone.

(20) Page 7–58 states that the exposures interpreted in the cited papers by Cox do not contain unequivocal evidence of large earthquakes, but gives no reason for this surprising conclusion. Given the recognized quality of Cox’s work and its publication in respected journals, it might be wise for the text to state in a sentence the reason for discounting Cox’s interpretation. The present reference to Appendix E is not sufficient.

(21) The last two lines of section 7.3.9.5 attribute a lack of “additional research conducted to support Quaternary offsets along these faults” to a 2005 paper by Wheeler. The paper cannot reflect any research that has been done since 2005 and it cannot reflect any research that was not found or not published. The citation is wrong and the quotation implies omniscience.

7.4: MAXIMUM MAGNITUDE DISTRIBUTIONS FOR SEISMOTECTONIC DISTRIBUTED SEISMICITY SOURCES

(22) The text explains how the double-peaked distributions in Figures 7.4.2–2 and 7.4.2–3 resulted from the procedures of chapter 5. However, if I understand the analysis correctly, the explanation does not answer the more fundamental question of whether a double peaked distribution is likely to seem reasonable to the informed technical community. Figure 7.4.2–2 implies that Mmax 5.0 is just as probable as Mmax 6.0 and Mmax 6.9. Does any substantial portion of the community think that Mmax can be as low as 5.0, when we have seen larger earthquakes occur in very unlikely places? A “yes” answer would require documentation. Additionally, one might ask whether it makes sense to conclude that Mmax 5.4 is less likely than Mmax 5.0 when the majority of the distribution is above M 6.0. However elegant the mathematics, it seems unlikely that Mmax 5.0 would be acceptable to a majority of the informed technical community at any probability.

7.5: RECURRENCE PARAMETERS

(23) The text introduces the figures but does not seem to draw any conclusions from them. For example, Figures 7.4.2–1 to 7.4.2–8 show that seismicity rates are highest where there are the most earthquakes, and that b does not vary much from 0.80. Section 7.5.2 explains the effects of different proportions of small and moderate earthquakes. What conclusions may be drawn from these observations, and how do the conclusions affect the analysis?

Chapter 8: Demonstration Hazard Calculations Using CEUS SSC Model

The purpose of Chapter 8 is to demonstrate the hazard calculations, show the major contributors to the hazard, show the sensitivity of parameters to the hazard, and compare the hazard with the USGS and COLA models. The CEUS SSC model rates are almost always higher, often by a factor of two or more, over a large range of ground motions. The slopes of the hazard curves are similar because they all assume the same ground motion prediction equations. It seems that if Mmax were the only consideration the USGS ground motions might be higher than the CEUS SSC ground motions because they typically have a higher mean value.

This higher rate of ground motions compared to earlier models is not clearly explained in the text. This higher hazard indicates that the CEUS SSC model predicts a rate of earthquakes that is considerably higher than the earthquake rate predicted in the USGS and COLA models. These higher rates can be seen in Chapter 5-7 figures (e.g., 6.4-7 to 6.4-16; 5.3.2-22, where the model over-predicts the historical rate. Has the earthquake catalog changed significantly? We examined the catalogs and find that the CEUS SSC catalog is higher than the USGS catalog but probably not by a factor of 2. Is the higher rate related to the decluster or smoothing parameters and methods (a and b-values)? Why doesn't the regression analysis fit the historical seismicity rates better in the figures shown in Figure 7 discussed previously? The USGS model uses a b-value of 0.95 for the CEUS. The b-values for the CEUS SSC model seem to be lower implying more large earthquakes.

Central Illinois Site: The CEUS SSC model is almost a factor of 2 higher than USGS/EPRI-SOG models. The major contributor is the IBEB (Illinois Basin) zone. The New Madrid (NMFS) RLME is most important at 1 s SA. However, background seismicity dominates at shorter periods. Why does the background hazard from CEUS SSC model give higher rates than were applied in the USGS and COLA models for short periods? At 1 s period the USGS and CEUS-SSC models are much more similar because the NMFS models are much more similar.

Chattanooga Site: The CEUS SSC model hazard is more than a factor of 2 higher in annual frequency of exceedance than the USGS and COLA models. At the Chattanooga site the ground motion hazard at e-3 to e-5 is more than a factor of 2 higher. Background sources contribute most to the hazard. However, the USGS ground motions are higher at 1 hz for exceedances of e-4 to e-6. This result is not explained in the text.

Houston Site: CEUS-SSC model hazard is dominated by GHEX (Gulf of Mexico) which is the zone that encompasses the site. Contributions from other background sources are much lower. Hazard is dominated by background sources at all periods (except for very low ground motions at 1 s SA). The SSC model indicates about a factor of 2 higher annual frequency of exceedance than the USGS model frequencies for short periods (10 hz and PGA) but is more similar at longer periods (1 hz). This is probably because NMFS is significant at 1 hz and the USGS and

CEUS-SSC models are more similar for NMFS. However, the differences are not explained in the text.

Jackson Site: For Jackson Site the NMFS is important at all frequencies. Therefore, the CEUS-SSC, COLA, and USGS models are quite similar for pga, 10 hz, and 1 hz.

Manchester Site: Similar to the other sites dominated by background hazard the CEUS SSC hazard at the Manchester site is considerably higher than the hazard for the USGS and COLA models. The deaggregation for the Manchester site at 10 hz is dominated by earthquakes with magnitudes less than 6.0 and distances less than 10 km. The CEUS SSC deaggregation for 10 hz at e-4 is similar to that produced by the USGS for pga at 4e-4. The higher rates for the Manchester Site should be explained in the text.

Savannah Site: For the CEUS SSC model the major contributors to the ground motion hazard are the Charleston RLME source and the $ACC_A M$ background source model. The CEUS-SSC, COLA, and USGS models are quite similar with the CEUS-SSC model showing a little higher ground motions for a large range of exceedances.

Topeka Site: The major contributor to the background source is MIDC-A which encompasses the site. The next important contributors are MIDC–B, MIDC-C, and MIDC-D. Background seismicity dominates the hazard at PGA and 10 hz and the NMFS dominates hazard at 1 hz. The hazard curves for the CEUS-SSC, COLA, and USGS and similar, especially at 1 hz. The hazard is typically higher for the CEUS-SSC model with rates almost a factor of two higher for a large range of ground motions. This discrepancy should be explained in the text.

CHAPTER 9: Use of the CEUS SSC Model in PSHA

This chapter is intended to provide information necessary in implementing the hazard model (contained in the Hazard Input Document). The main focus of these sections is on simplifying the logic tree to those branches that are most important and in understanding the level of precision associated with the hazard values. The main comment for this section is that it is hard to justify a +/- 25% uncertainty between experts when the results of Chapter 8 show a factor of 2 difference in rate of exceedance between the USGS and CEUS SSC models at many sites.

Section 9.1 gives an overview of the chapter and section 9.2 discusses the Hazard Input Document (HID) that is discussed in Appendix H. Section 9.3 discusses hazard sensitivity studies for different crustal thickness and rupture orientation (dip, direction) at the seven test sites. In general these parameters are shown to have little sensitivity in the hazard results with the exception of a source near the Illinois Basin. Therefore, a single branch of the logic tree is selected rather than maintaining the multiple branches. Sections 9.3.2 and 9.3.3 were not available and will be written later.

Section 9.4 discusses the level of precision associated with seismic hazard estimates. For example, how would the hazard vary if different experts made up the TI team. This section discusses three inputs of the hazard analysis: sources, ground motions, and site response. The sources involve source geometry, maximum magnitude, paleoseismic record length, and activity rate given a record. These results may change when NGA-East becomes available and may be specific for these test site and not apply to all sites. The authors make a case for using COV_{WT} of 0.5, 0.3, and 0.7. Would this estimate depend on the number of ground motion equations available? Results generally show that the annual frequencies of ground motion exceedance increase with ground motions, which seems reasonable. However, I do not understand why the COVs decrease in annual frequencies of exceedence greater than 1E-5 on Figure 9.4-53 and 9.4-57. The authors show at the Savannah , Chattanooga, and Columbia sites that the term “cl. Mean COV” is quite a bit different from the “wts COV”. This was not intuitive and I do not understand this result. Section 9.4.3 discusses the conclusions on the precision in seismic hazard estimates. They conclude that the ground motions can vary by about 25% for ground motions with hazards in the range 10^{-4} to 10^{-6}. Some members of our group have questioned whether or not experts would all obtain answers within 25% of one another, especially in light of the fact that in previous hazard models the difference between the COLA and USGS models have had factor of 2 differences at some sites. The differences in rate of exceedances and the ground motion hazard between the USGS and CEUS SSC model are sometimes more than a factor of two.

Editorial Comments: The figure numbers for 9.4-1 are mislabeled. The counting seems to have proceeded up to 9.4-71 even though the last number seems to be 40 greater than the number referenced in the text. Some of the notation in section 9.4 was a little ambiguous.

It took me a while to realize that COV_K was for the parameters GEOM, Mmax, RATE, and RECORD. COV_T is probability the same as SRSS. Later in the chapter there are references to COVHAZ and COV_{HAZ} wts COV, and COV_{WT} , cl mean COV and σ_{CL} and σ_H. I think that these terms need to be cleaned up to be more consistent with the equations and figures.

APPENDIX H: EPRI/DOE/NRC CEUS SEISMIC SOURCE CHARACTERIZATION PROJECT- Draft Final Seismic Source Model Hazard Input Document (HID) July 6, 2010

The intent of the HID is to give future users details on how to implement the CEUS-SSC model. It contains the logic tree structure that defines the frequency, locations, and sizes of future earthquakes in this region. The appendix describes how the zones are characterized. A description of why the TI team chose a particular equation, occurrence rate, magnitude, or source geometry, or references is not given in this section of the report. The logic tree does not seem to be trimmed according to Chapter 8 sensitivity studies. It would be helpful to have a logic tree in which only the important branches were included.

It would be difficult for most users to implement the variable a and b routines described in previous chapters. Therefore, the process is not open for most users to evaluate that methodology. It seems like the computer codes should be made available for these analyses. Alternatively, the TI team could release the output gridded data. However, this is not the best alternative since most users would not understand how these numbers were generated. A third alternative is for the TI team to revert to the smoothed seismicity kernel that is more intuitive to the user community.

Editorial comments: I can't see any differences in Figure 8 and 9. Page H-19 states that an objective approach is used to select the degree of smoothing. It would be very helpful to refer back in the text where this approach is described. I am confused why Tables Charleston HID-2 and HID-3 have identical RLME frequencies even though they both consider 4 events and one table is for a 2000 year period and the other is for a 5000 year period.

Steve Harmsen reproduced the 1000 branches of the logic tree for the Cheraw Fault from the HID information. His analysis is described below.

My primary purpose was to evaluate the information in Appendix H to determine whether the information in the document was sufficient for the USGS to replicate the draft SSC source model in its hazard calculations.

This was too difficult to fully evaluate in the time available. Instead, I concentrated on trying to replicate the RLME model for the Cheraw fault in southeast Colorado and to compare hazard from the EPRI model with that of the USGS in its 2008 NSHMP update. The Cheraw RLME is neither the most complex nor the most simple of those considered in Appendix H. It is a representative case for early evaluation.

The EPRI model for Cheraw fault has 5 major branches, labeled HID-2 through HID-6 in Figure Cheraw_HID-1, the seismic source logic tree. These branches were sufficiently explained that I was able to perform the calculations for hazard from the source descriptions that are given.

The EPRI model differs from the USGS model in some significant ways. Primarily, the EPRI model just considers characteristic events. In contrast, Gutenberg-Richter branching is given equal weight in the USGS model for this fault. This has a significant impact on hazard at the return times that are customarily considered in seismic resistant design for buildings, bridges, and so on. It has lesser significance at very long return times, such as 100,000 years. The EPRI model has branching based on use of recurrence intervals and slip rate. The USGS model just considers recurrence intervals. It turns out that the slip rate EPRI Cheraw RLME model produces about the same mean hazard as RI based hazard. Thus adding slip rate branches, at least those considered in Appendix H, does not have more than trivial impact on the mean hazard (although it may impact fractal hazard, not evaluated here). A relatively important difference is that EPRI considers in-cluster and out-of-cluster branching, whereas USGS assumes in-cluster. Another new feature of EPRI relative to USGS is the lengthening of the fault to 62 km on some branches, whereas USGS just uses one fault model. Variable dip is another feature, USGS model just has one dip for this source (variable dip is implemented in Basin and Range only in 2008 NSHMP update). Variable thickness of seismogenic crust is yet another detail that USGS has not previously considered at this source. I programmed in all of these branches when performing this review. For Cheraw, considering all branches, there are about 1000 source models in the new EPRI model (versus basically 2 in the USGS model, characteristic vs. GR).

Figure 1 below shows the mean hazard from the Cheraw source at Topeka KS, the only site in the EPRI set that is affected by this source. The figure shows that the USGS model is significantly higher at short return times, or lowest ground motions, but is lower at relatively long return times, because the smaller GR-based sources cease to produce significant rates of exceedance at the higher ground motion. This figure is comparable to Chapter 8 Topeka hazard Curve (TBD), but it needs to be kept in mind that the calculations here are for firm rock not hard rock, and that the GMPEs that I used are the set used by NSHMP, not the EPRI/SOG GMPEs used in the EPRI HID analysis.

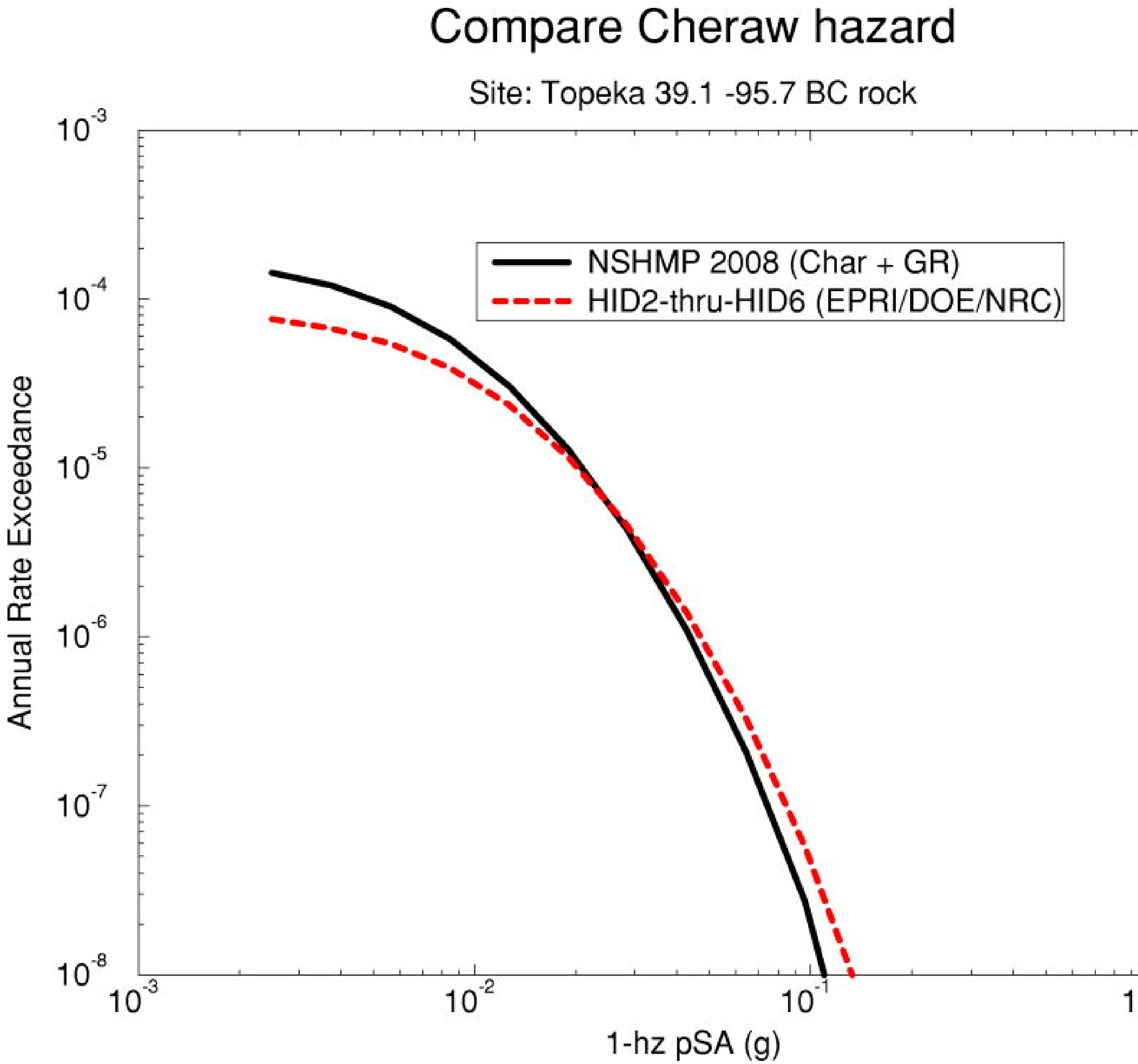


Figure 1. Topeka Kansas hazard curves for 1-s SA, 5% damping. USGS black, EPRI/DOE/NRC red.

Figure 2 shows 1-s hazard curves at a closer site, Colorado Springs, where this source is more relevant to hazard. In figure 2, besides comparing USGS with EPRI mean hazard curves, I also show the EPRI mean hazard from the combined HID-4 and HID-6 major branches. At Cheraw, these two branches include all slip-based hazard (versus recurrence-interval based). Figure 2 shows that slip-based hazard (green curve) is basically the same as total hazard (red curve), that is, slip based source recurrences say the same thing in the mean as the total logic tree. This is one of several examples where apparent diversity is really not significantly present, at least as it impacts mean hazard.

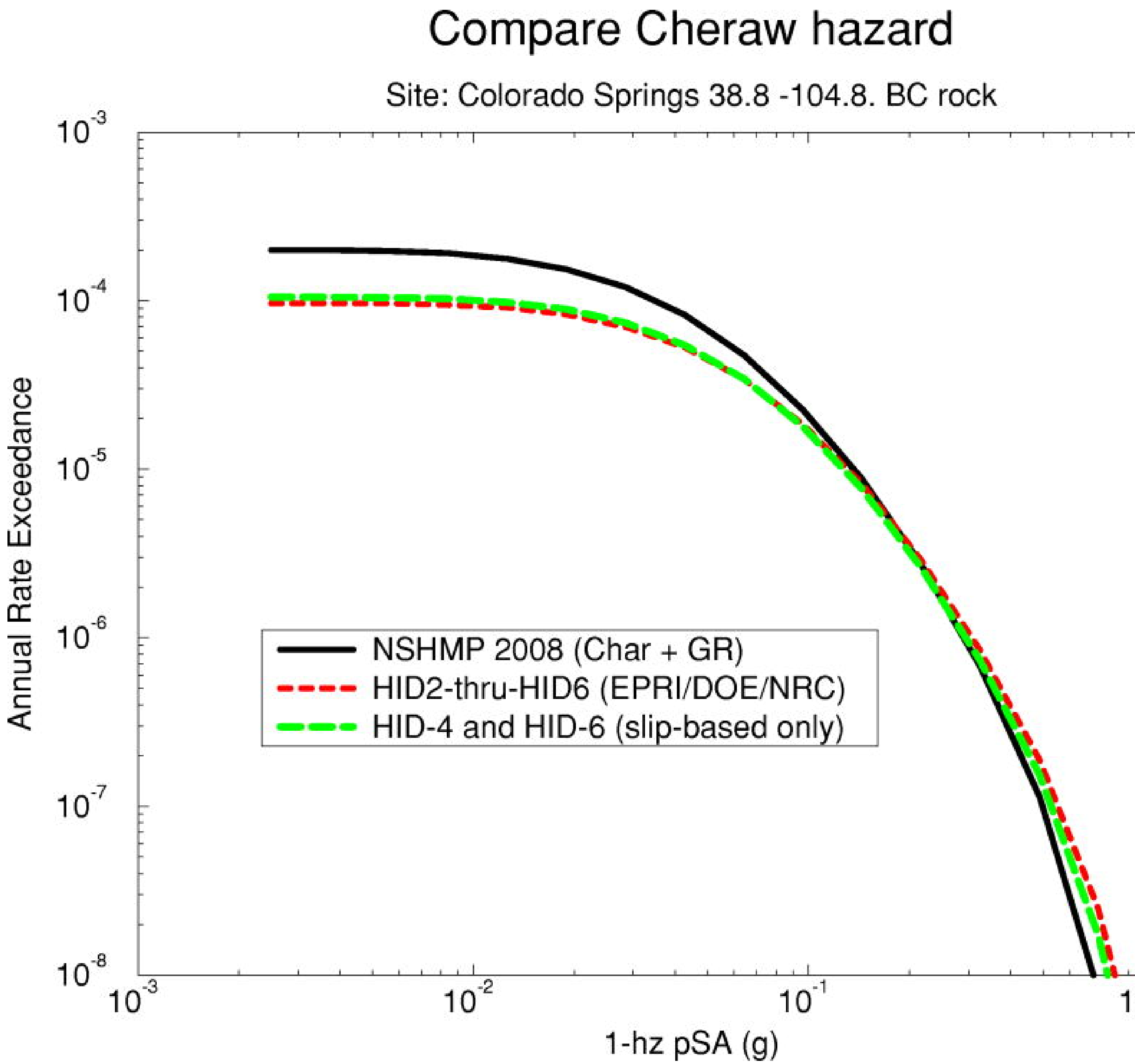


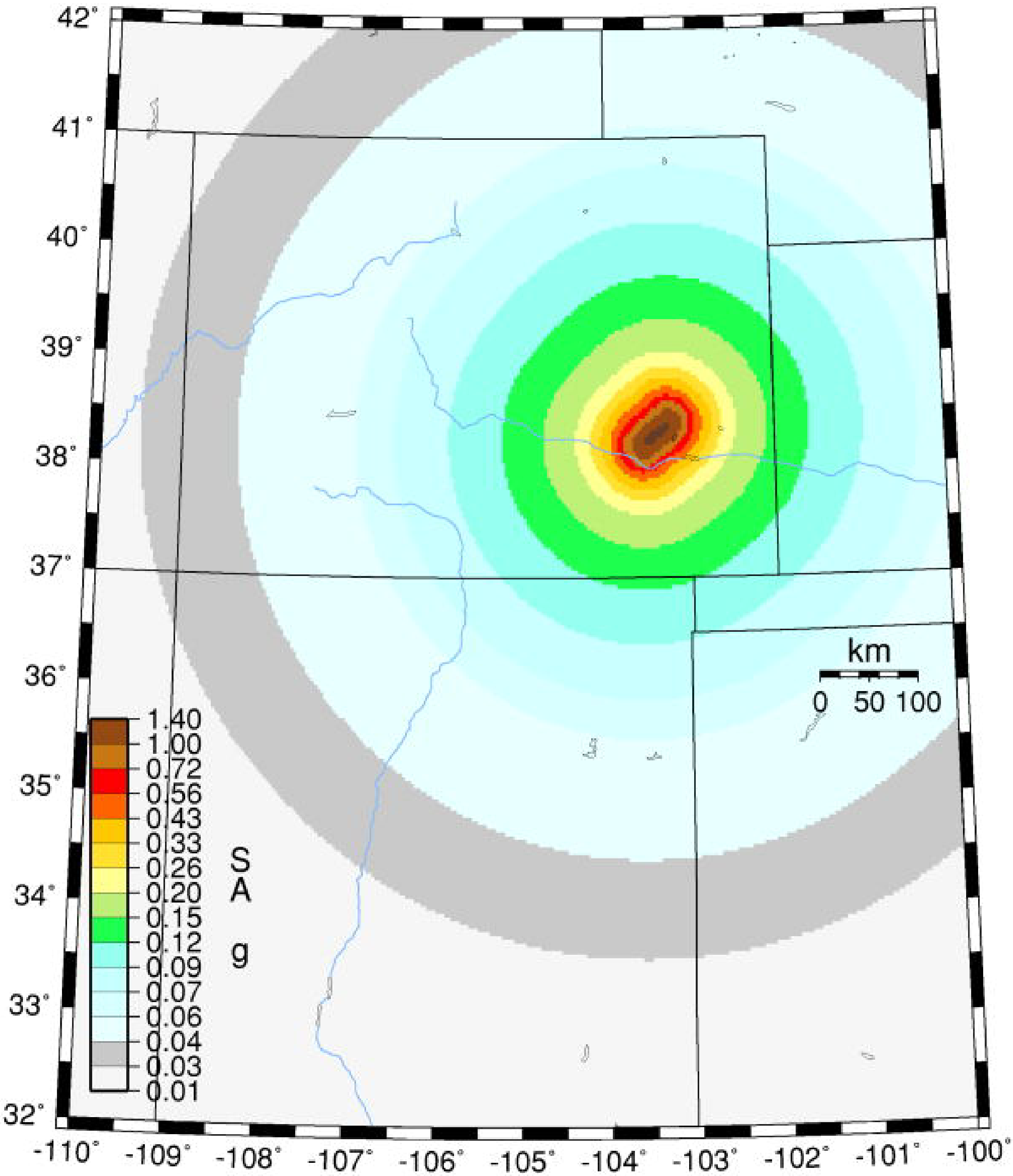
Figure 2. Hazard curves for 1-s SA at Colorado Springs comparing USGS model (black) , EPRI full model (red), and EPRI model restricted to just slip-based source rate estimates (green). These occur on the HID-4 (in-cluster) and HID-6 (out-of-cluster) branches, respectively.

From the USGS perspective, the difference between the black and colored curves might be of some significance and will probably require some high-level workshop participant decisions to change the model.

I also looked at the mapped hazard at 10^{-5} annual exceedance, not a typical value, but needed to get any interesting results when this source is considered in isolation as it was here.

Figures 3 and 4 are the USGS model and the EPRI/DOE/NRC model, respectively, for the uniform hazard at this low PE and for a uniform NEHRP B/C boundary rock site condition, for sites surrounding the fault. These figures do not differ much, although effects from the different fault geometries are evident at sites very near the fault. For example, the dark brown region is broader in Figure 4 than in fig. 3, probably because of the greater crustal thickness, on average, in the EPRI models, and the shallower fault dip in the main branches, 50˚ instead of NSHMP's 60˚ dip. These sites tend to be cattle ranches and watermelon farms for the most part (just a guess).

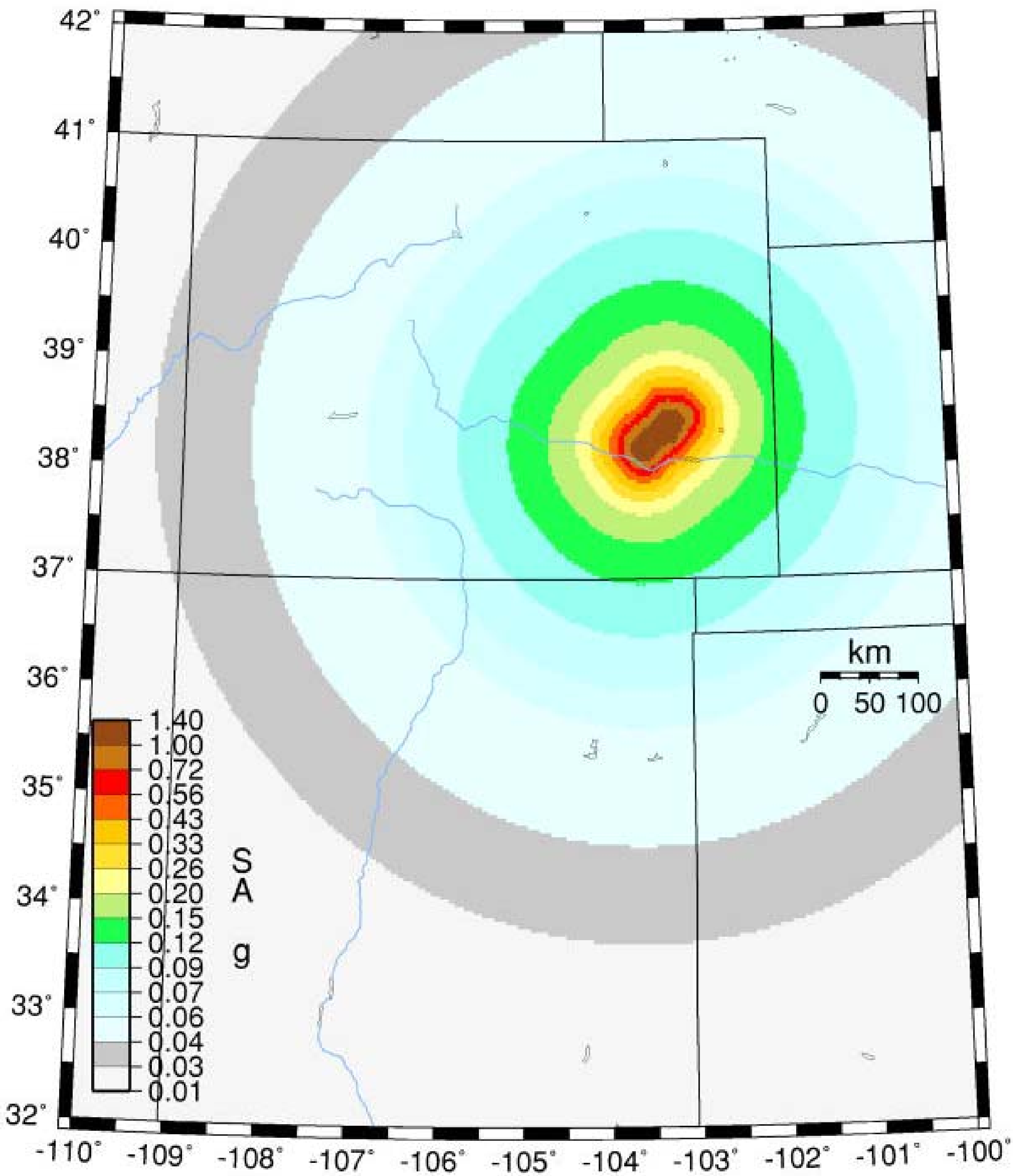
NSHMP 2008 1-Hz SA w/10**-5 rate



GMT 2010 Aug 26 12:07:37 Probabilistic Spectral Accel (SA) for NSHMP, USGS model. Site Vs30 760 m/s. 1 Hz 100,000 yr. Cheraw fault hazard GR+Char.

Figure 3. USGS/NSHMP model of hazard from the Cheraw fault in SE Colorado and vicinity.

EPRI/DOE/NRC 2010 1-Hz SA w/10**-5 rate



GMT 2010 Aug 27 10:03:26 Probabilistic Spectral Accel (SA) for Hid2-d6, Cheraw. Site Vs30 760 m/s. 1 Hz 100,000 yr. Cheraw fault hazard depth,dip,length

Figure 4. New EPRI model of hazard from Cheraw source at long return time (100,000 years).

Implementing Other RLME Sources

The logic-tree presentations for the New Madrid and Charleston source zones are somewhat more complex than that of the Cheraw zone, and only time will tell if we can implement them without significant problems. We already have source-clustering capability in the software which was used in the NMSZ source model calculations for the 2008 update, so this new feature should not be a problem. Time-dependent hazard branching should not be a *calculational* problem, in that it is treated as Poissonian with adjusted rate parameters. However, time-dependent renewal models have not been used in previous NSHMP final map products, so will probably require debate at workshops that include discussions of New Madrid, Charleston, and other RLMEs with time-dependent branches.

Other source logic trees tend to be simpler than that for the Cheraw RLME, other than the above two. I do not anticipate problems implementing these.

Mmax zones and distributed seismicity zones

One feature that NSHMP does not currently have software for is the tight (versus leaky) boundary. I have worked on this in our software but this could be an issue for further work. We do have capability of inputting and calculating with variable *b*-value, as well as variable *a*-value matrices, and multiple Mmax matrices, using current versions of NSHMP seismic-hazard software. Whether the penalized likelihood methodology used to estimated variable *a*- and *b*-parameters is acceptable is another topic, beyond the scope of this short article. The modeling of virtual variable-dip dip-slip sources along with virtual variable-dip strike-slip sources in these zones is another matter that may require software development. Current NSHMP software assumes vertically oriented virtual faults.

PART 2: Review comments from USGS Seattle: Art Frankel

This study represents a large amount of work and a significant attempt to engage a group of experts to discuss seismic hazard issues in the CEUS. That said, I think there are major problems in the methodology used in this draft that make its utilization for seismic hazard assessment questionable. From my experience convening several workshops on seismic hazard in the CEUS, the spatial detail in the seismicity rate and b-value determinations in this report would not be generally acceptable to most of the regional experts, at least the ones who attended our workshops and commented on our maps. The Mmax distributions used in this report would also not be acceptable to most of the experts who attended our workshops.

Does the SSHAC process include the dissemination of the draft report to the people who participated in the workshops and other experts to get more general feedback? It appears that the review of the draft report is really limited to a small group of people. Therefore, I think that this SSHAC level 3 exercise leaves something to be desired as a way to capture expert opinion in PSHA.

Obviously there was only limited time to review this large report. The following are my initial comments. I have not reviewed the site amplification portion of the report in Chapter 8 with the demonstration hazard calculations. I have not had sufficient time to thoroughly review the parameters in the RLME logic trees or to review chapter 9.

The Master Logic Tree

First of all, the master logic tree in Fig. 4.4.1-1 is not clear and may be inaccurate. The RLME is not an alternative branch of a logic tree. It is a separate, additive component of the overall hazard "model." For the other component of the model, there should be two branches of the logic tree for the seismicity-parameter calculations based on either 1) Mmax zones or 2) seismotectonic zones.

Figure 6.4-1 shows the seismicity parameters for the case of no seismotectonic zones or Mmax zones. I don't see such a branch in the master logic tree. If this case is used, then it needs to be a third branch in the master logic tree.

The high weight assigned to the seismotectonic zones is not consistent with the weighting favored by most of the participants of the USGS workshops.

Seismotectonic Source zones

This report gives high weight to the idea that seismotectonic source zones defined on the basis of similar timing of ancient faulting mean something to current seismicity rates and b-values. However, the authors of this study don't think these source zones imply uniform hazard throughout an individual zone. Then, what is the point of having these source zones for the seismicity rate calculation? Having these source zones often causes unphysical artifacts in the seismicity rates. Seismicity trends often cross the source zone boundaries (e.g., Ottawa-Massena trend and in southern New York state). Yet the seismicity rate calculation technique causes discontinuities of rate and b-values at zone boundaries.

The bottom line is: how much difference in the hazard does it make to use the seismotectonic zones versus the Mmax zones in the seismicity rate calculation? I suspect very little difference overall, for the full M weights. Figures 6.4-5 and 7.5.2-3, which show the mean results from the two approaches, have a lot of similarities. The spatial patterns of the seismicity rate are very similar. The seismicity rates are a bit lower in the Mmax zone case, but the b-value is lower too. So the hazard may be very similar between these two cases. It would be useful to see what the differences in the mean and median hazard curves are between these two cases, all other factors being equal.

It would be informative to have maps that show the mean expected rate of M>=5.5 and M>=6.0 earthquakes, the magnitudes that often dominate hazard in the CEUS (outside of RLME's).

I do think there is a need for some areal source zones in the CEUS, based on the classical notion of uniform hazard within a zone. These are needed, for example, in the case of a linear structure that has a cluster of seismicity in only one portion. The SSC study is deficient because it doesn't include these traditional areal source zones in certain areas.

I think the Eastern Tennessee (ETN) area requires a traditional areal source zone. I think the fine grained nature of the seismicity rates derived from the report's methodology is not justified as a guide to future seismicity. Also, by not having an ETN source zone, the hazard is underestimated in northeastern Alabama, the site of the Fort Payne earthquake and other recent earthquakes. It appears that this area is on the southwesterly extension of the ETN zone.

The Nemaha Ridge is another example of an area which may require an areal source zone. The M 4.5-5 earthquakes in the 1800's may be associated with the Nemaha Ridge structure. If so, one should connect these epicenters with an areal zone along the Ridge, so that the area between these earthquakes has higher hazard. This is not captured in a smoothed seismicity approach. I would think this source zone deserves some weight in the comprehensive approach advocated in this SSC. The report discusses some possible liquefaction features in the vicinity of the Nemaha Ridge, but then states on page 7-69 that the M5.2 1867 earthquake may "characterize the seismic source in this region." This statement needs elaboration. Could the M5.2 earthquake cause the observed features that may be from liquefaction? Sounds unlikely to me.

Spatial smoothing technique to get rate and b-value

The procedure for determining seismicity rates and b-values on a grid is extremely complicated, difficult for even an expert to understand, opaque, and not reproducible by others. I wonder if the complexity of the analysis technique is justified by the limited seismicity data and speculative nature of the source zones.

A key difficulty is that the programs used to make these seismicity rate and b-value grids have not been released and are not available to scientists outside the SSC project. Therefore, these results cannot be reproduced or checked by other investigators. Given that this method will provide the basis for hazard assessments for nuclear power plants, I don't think this is acceptable from a public policy standpoint.

I recognize that there may be some potential advantages of this complicated and turbid method, but I think it also has pitfalls which make its use problematic.

The report often uses the term “objective” to describe how the weights for the smoothing were derived for each source zone. I presume that this “objective” aspect is described on page 5-27, but the text is very hard to follow. It is unclear to me how the data are “objectively” constraining the smoothing parameters. I am doubtful that the data do constrain the smoothing parameters. If you’re just trying to get a best fit to the earthquake data, then having no smoothing would produce the lowest residuals.

It would be very helpful for the authors to provide more details about the penalty method, by showing how it works in some special cases, e.g. a cluster of events in one place and events spread out over an area. The use of the Laplacian is a mathematical construct and doesn’t provide insight on physical parameters, such as smoothing distances and the correlation function of the observed seismicity. This appears to be a mathematical exercise that provides little physical insight.

In the USGS maps, we used seismicity models with different low-magnitude cutoffs (M3, M4, M5). However, this was done with the opposite intention to that invoked in the SSC report for the reduced magnitude weights. We wanted to give higher weight to the M4 and M5 earthquakes, to emphasize the hazard for areas that had these larger earthquakes, but had relatively few M3 earthquakes in the catalog. Such areas include New York City and Nemaha Ridge. Our application of models with different magnitude cutoffs addresses possible local variations from a GR recurrence relation.

In contrast, the SSC report uses the reduced M weights to create maps with excessively smoothed seismicity rates, which increase the hazard at large distances from the observed M5 earthquakes. The reduced magnitude weight models produce seismicity rate maps that have odd changes within source zones that form almost linear boundaries (e.g., fig. 6.4-6, fig 7.5.2-6). The authors should explain these physically unreasonable patterns and also show the locations of only the M4 and M5 earthquakes that apparently drive these patterns.

It appears that the SSC method systematically overestimates the rate of M>=5.0 earthquakes in the CEUS compared to the earthquake catalog. This is illustrated in figures 6.4-7 through 6.4-16 (for the Mmax zones and in Chapter 7 for the seismotectonic zones). I think the observed rate of M>=5.0 in the CEUS should be a key constraint in a hazard model for the CEUS. We tried to adhere to this constraint for the USGS maps. The LLNL study in the 1990's also had a large over prediction of M>= 5.0 earthquakes.

It is important for the report to show the total mean recurrence rates from all the models and compare to the observed rates, for the entire CEUS. If there is a major discrepancy in the observed and predicted (mean) rates of M>=5.0, this should be explained and justified.

It is likely that the over prediction of the rate of M5.0 and larger earthquakes from the SSC model is a major reason why the SSC hazard estimates are usually higher than the USGS ones for the sites shown, although I'm not totally convinced that the USGS model was correctly implemented.

For St. Paul, the reduced M weight realizations have a mean rate of M>=5.0 of about 0.1 (Fig. 5.3.2-30). So we should expect a M>=5.0 earthquake near St. Paul about every 10 years, according to the mean of these results. Interesting, given that there hasn't been a M>=3.0 in this zone for at least 100 years, according to the plot in the report. I would hope that some of the models based on the occurrence of M5 earthquakes (the reduced M weight models) would give low hazard to a place that hasn't had a M5 historically. This does not seem to be the case and highlights, I believe, a flaw in the methodology.

The reduced M weight models have the most spatial smoothing and help to form the floor of hazard in low-hazard areas. So the weighting of the reduced M weight model (0.333) relative to the full M weight model (0.667) is important to establishing the floor of hazard. **The choice of this weighting is not objective, so the floor of hazard is also not objective.** This is illustrated in Figure 5.3.2-11 for Houston. The reduced M weight model dominates the hazard, so its weight will control the floor of hazard.

There is too much fine-grained structure in the seismicity rates determined in the active areas (e.g., fig 6.4-3). As time goes on, we are seeing more M3 earthquakes on the fringes of the active areas. This is to be expected from an ETAS type model of seismicity. Taking such a fine grained approach to the location of future earthquakes is not physically reasonable. It is likely that the areas within about 70 km of the cells with the highest rate values will be filled in by future earthquakes. Again, it is not adequately explained how the “objective” smoothing in this report is achieved.

It does not appear that the “objective” smoothing method takes into account the uncertainties in earthquake locations and the change of these uncertainties with time (older events more uncertain in location than more recent ones). It should.

Using the seismotectonic or Mmax source zones creates large, unphysical discontinuities in seismicity rates and b-values across the source zone boundaries (e.g., Fig 6.4-4 for Texas, fig 6.4-6 western PA; fig 7.5.2-3 Gulf of Mexico).

Fig 5.3.2-16 is enlightening because it shows the mean rates and b-values from the highest weight source-zone configuration. It’s not reasonable that in Fig. 5.3.2-16 the area off the coast of South Carolina (near the Charleston earthquake sources) has lower seismicity rate than the area offshore of Florida. This seems to be another unintended consequence of the “objective” smoothing method applied in a source zone context. This map also shows the artificial north-south division of seismicity rate in the Gulf of Mexico. Can we resolve the higher b-value found for the Oklahoma aulocogen?

How well-resolved are the spatial variations in b-value determined from this method? I would doubt that most of these differences are resolvable, given the number of earthquakes. This level of spatial detail is simply not justified by the data. See my comment in the first paragraph about the views of the experts on this subject in the USGS workshops.

Why is the b-value prior different from the b-values listed in Table 5.2.1-4?

Mmax

The driving sentiment behind the Mmax distributions in this report is that a region is innocent of having a high Mmax (say M7.5) until proven guilty (I am paraphrasing Bill Lettis). Even when a region is guilty of having large (M>7) earthquakes, somehow the defendant is let off on a technicality by the judge throwing out incriminating evidence.

I don't personally subscribe to the view that a region is innocent of having a high Mmax until proven guilty, and neither did the people attending our workshops. I think there needs to be a whole separate branch of the Mmax logic tree that represents this point of view. That way people can see this popular opinion directly in the total model. This also forces the report authors to explicitly assign a weight to the view that regions are guilty of high Mmax until proven innocent, rather than perhaps claiming that it is embedded in the result of the Bayesian procedure.

Why exclude the magnitudes of the RLME's from the Mmax determinations for the rest of the CEUS? In particular how can you not use the RLME magnitudes for the Mmax in the source zone where each RLME resides? For example, how can you have a modal Mmax of 5.6-5.8 for the NMESE_W zone (fig. 6.3.2.-5)? This is too low in my view and the view of our workshop participants. Note that the Meers and Cheraw faults are within this zone. Are the authors of the report telling us that these are unique places within this zone? Given the paucity of neotectonic research in this region, how can this be the highest probability logic tree branch? Why not at least use these RLME earthquakes in the Bayesian Mmax determination for the Mmax zone they are in?

Similarly, the MESE-N zone contains New Madrid and Charleston, yet has a modal Mmax of only 6.6. By the way, this is lower than the modal Mmax used for NMSE-N (6.8), which is in non-extended crust. This is counterintuitive. Note that the NMSE-N zone contains Cheraw and Meers, yet has a modal Mmax of 6.8.

MESE-W contains New Madrid and Charleston and has modal Mmax of only 7.1.

For the seismotectonic zones, the dissonance between the RLME's and the Mmax is sometimes profound. Note that the modal Mmax for the Oklahoma Aulocagen zone is 6.5, yet the Meers fault is within this zone. This is a major disconnect. Again the authors are assuming that all the RLME's in the OKA have probably been found. Another problem: modal Mmax of 6.4 in Reelfoot Rift source zone. This contradicts the Marianna paleoliquefaction data. Presumably this would have been the Mmax distribution used for the Marianna zone if this study had been done *before* the Marianna features were found. If so, the modal value would not have represented the magnitude of the Marianna earthquakes.

The SSC Mmax methodology is basically stating that we have probably found all the RLME's in the CEUS. That is likely to be wrong.

So what makes the CEUS tectonically different from the area of the M7.7 Bhuj, India earthquake or the M>7 paleoearthquakes found in Australia? Note that the aftershock zone of the M7.7 Bhuj earthquake was only about 50 km long. There are likely to be many faults of that length throughout the extended margin of the CEUS.

The Bayesian method implicitly assumes that there is some spatially variable nature to Mmax, for a given type of tectonic region, so that a prior distribution is formed based on a collection of observed "Mmax's" from different regions. This implies that the observation of a M7.7 earthquake in a tectonically-analogous region does not necessarily imply a best estimate Mmax of 7.7 for the region of interest. I think it does, within the uncertainty determination of the magnitude.

So the SSC does not consider the tectonic analog approach as implemented by the USGS. I strongly think this causes an underestimation of Mmax's in this study.

Conversion of Mn and mblg to moment magnitude

It has been problematic to develop a moment magnitude catalog for CEUS hazard studies, since there are relatively few determinations of moment magnitudes for mblg 2.5-3.5 earthquakes. The SSC uses a constant offset to go from mblg or mn determinations to M (moment mag). I think this may seriously underestimate the true rate of M>=3.0 earthquakes. Figure 3.3.3-1 shows that the constant offset does not fit the "approximate M" data for mn < 3.5 and figure 3.3.3-2 shows the same thing for mblg. I presume the term "approximate M" means the M determination came from the use of S-wave spectral levels, adjusted using the LS fit in Figure 3.3.1-1 and the figures that follow for Boatwright and Macheridas. The report notes that the constant offset fits the data well for mblg or mn > 3.5. However, it's my understanding that the seismicity rate calculations use earthquakes with M down to 3.0. So it is important to get the

M's correct for these small events. Why not use the locally-weighted LS fits in figures 3.3.3-1 and 3.3.3-2 to convert mn and mblg to M? The results for M3-3.5 would be substantially different than that found from the constant offset. At least this approach should be given some significant weight.

At larger mblg's there should be a crossover as the corner frequency of the earthquake gets around 1 Hz, say about magnitude 5.5-6. So there should not be a constant offset with all magnitudes applied for mn or mblg to get to M.

Including speculative RLME's

The existence of RLME's in the Commerce fault zone is uncertain, but this is not reflected in the logic tree.

Clustering

Allowing for the possibility that a source is out of a time cluster is another way to down weight the hazard from that source, for example, the 0.5 weight for clustering of the Marianna source. Given the long recurrence times for this source region, is a cluster model resolvable?

New Madrid

The report does a detailed assessment of the hazard in the New Madrid area. Separate sources are used for the Reelfoot fault, southern and northern NM faults, and the Bootheel lineament. However, I don't see the possibility that the 1811-12 earthquakes (and the 1450 and 900 A.D. quakes) occurred on none of these features, but on faults nearby. I think this is a possibility that needs to be included. Recent reflection surveys have found many possible faults in the New Madrid area, other than the ones inferred from the current microseismicity lineations.

Other

Is the Oklahoma Aulacogen treated as an RLME? It is shown in Fig. 6.1-1 where the RLME's are indicated. It is listed in Table 4-6 as an RLME source. Yet from the text it appears to be only treated as a seismotectonic zone.

How is the b-value and seismicity rate calculated for the continental slop source zone (AHEX)? There are almost no earthquakes in this zone.

I thought there was too much certainty implied in some of the geologic narratives for the seismotectonic source zones. Hypotheses are often presented as facts. For example, on page 7-22: the report states that “a mantle plume initiated Iapetan rifting along the Sutton Mountain triple junction…” Do we really know this is what happened? It’s a hypothesis.

I wonder if the USGS model was really implemented properly in some of the comparisons. For our Charleston source zones we aligned the strike of the faults and allowed them to extend outside of the source zones.