Evaluation of 2002 USGS National Seismic Hazard Assessment
Final Report

National Seismic Hazard Mapping Project
U.S. Geological Survey

Introduction

We have developed and tested a new set of seismic hazard and deaggregation computer programs and applied them to the 29 sites of RG 1.165. The primary code uses a Monte Carlo method to sample through a series of logic trees that vary key parameters in the hazard calculation. This master code calls the hazard calculation codes that were used to make the 2002 USGS national seismic hazard maps. The new codes provide the mean seismic hazard curve and seismic hazard curves of selected fractiles, such as 15th, 50th, and 85th. Two other new codes we have developed for this project deaggregate the results in two forms. One code deaggregates the overall mean or median hazard curves. The other code finds the median value in each magnitude-distance bin as specified in the procedure in RG 1.165.

In the following report, we describe the logic trees used in the new procedure and show plots of the results. The logic trees are an expansion of those used in the 2002 national seismic hazard maps. We wanted to make sure that the added logic nodes did not substantially bias the mean hazard curves derived from the Monte Carlo runs relative to the mean hazard curves derived from the more limited logic trees used for the national maps. We also wanted the logic trees to produce an estimate of the uncertainties that was larger for regions that have had little historic seismicity but may have the potential for producing significant earthquakes, such as the Gulf Coast and the upper Midwest. These requirements led to extensive testing of alternative logic trees during the course of this project.

All of the hazard curves in this study were derived for a hard-rock site condition. The capping of median ground motions and the truncations in the ground motion distributions were the same as those used in the 2002 national maps (see Frankel et al., 2002).

Description of Logic Trees

Figures 1-3 show the logic trees used in the final calculations. For all sources, we used a logic tree of attenuaiton relations, with weighting identical to that used in the 2002 national maps. For the New Madrid source (Fig. 1) we used logic trees on characteristic magnitude, location of pseudofaults, recurrence time, and pseudofault endpoints. The first three of these are identical to those used in the 2002 maps. The recurrence time is characterized by a log-normal distribution with a variability determined from an analysis of the observed times of past large New Madrid earthquakes (1811-12, about 1450, and 900 A.D.). The logic tree on characteristic magnitude ranges from M7.3 to M8.0, essentially encompassing the range of magnitudes determined for the largest 1811-12 event by analyses of the isoseismal data by various investigators.

The logic tree for the Charleston, South Carolina source is depicted in Figure 2. The two areal source zones are specified in the node "1886 rupture model." These are the same areal zones used in the 2002 maps. The recurrence interval is specified as a log-normal distribution, with the same variability as that determined for New Madrid. The logic tree for the characteristic magnitude is the same as that used in the 2002 maps.

The logic tree for the spatially-smoothed seismicity required the most analysis for this work. To quantify the variability of the seismicity rate from the catalog, we re-sampled the catalog for each run. Originally we had also considered magnitude variability by re-assigning the magnitudes. However, we found that this biased the hazard lwo in areas dominated by the background source zones. Since the a-value for these zones were determined from the rates of magnitude 3.0 and above, varying the magnitudes preferentially reduced the rates, since our original catalog did not include events below magnitude 3.0. Thus the mean rate was underestimated from the re-assigned magnitudes. We removed the magnitude re-assignment from the code. Another question arose about varying the locations of the earthquakes in the catalog, after re-sampling. We found that this tended to reduce the hazard in areas with spatial clusters of earthquakes. We think that these clusters are real and their effects should not be diluted by varying the epicentral locations. Furthermore, the spatial smoothing tends to account for uncertainties in earthquake locations. Therefore, we did not vary the locations in the catalog.

The logic tree included a node for varying the Mmax used for the extended margin and craton Mmax zones. This Mmax is applied to all calculations involving the historic seismicity. We decided on using +- 0.2 magnitude units for the variation in the mbMax. We wanted a symmetrical variation so as not to introduce bias with respect to the 2002 values, but we did not want to use unrealistically large values of Mmax.

The key to the logic tree for the smoothed seismicity is the node for seismicity model. The 2002 maps use four models to capture the epistemic uncertainty. These models are: 1) M >=3.0 since 1924, 2) M>= 4.0 since 1860, 3) M>=5.0 since 1700, and 4) background zones. In particular, we felt that the occurrence of magnitude 5 and above earthquakes provided an important guide to where we might expect future moderate or large earthquakes. The separation of the seismicity into three models based on a minimum magnitude was designed to give the occurrence of larger events more importance in the hazard assessment. Using one model with a maximum likelihood method would treat the M4 and M5 events the same as the M3 events for the a-value estimation, although they would influence the b-value calculation. We did not vary the regional b-value of 0.95 used in the calculations, which is the same as that used in the 2002 maps (for Charlevoix area we used b of 0.76). We found that the catalog resampling provided reasonable variations in seismicity rates of moderate earthquakes, without the need for varying the regional b-value.

The 2002 maps employed adaptive weighting of the four seismicity models. This procedure was implemented so as not to lower the hazard in high seismicity areas by including the background zones. For each cell used in the seismicity-rate grid, the historic seismicity rate was compared to that of the background zone. The historic seismicity rate was determined from the weighted average of the seismicity rates from models 1-3. When the historic seismicity rate exceeded the background zone rate, only the historic seismicity rate was used for that cell (weights of 0.5, 0.25, 0.25 for models 1-3, respectively). In that case, the background zone was not considered. When the background zone rate was higher than the historic-seismicity rate for that cell, then the background zone was included in the mean seismicity rate for that cell, such that the weights were 0.4, 0.2, 0.2, 0.2 for models 1-4, respectively. The problem with this adaptive weighting is that it slightly violates the observed total rate of M3+ earthquakes by about 11%. This adaptive weighting is also described in the documentation for the 1996 maps (Frankel et al., 1996).

It is problematic applying the adaptive weighting scheme to the Monte Carlo simulations. We want to have a uniform procedure applied to all sites. We do not want to apply a three-model draw for some sites and a four-model draw for others. Even this approach would not be the same as that used in the national maps, since the number of models used varied with different seismicity-rate cells. We concluded that it was better to sacrifice total consistency with the national hazard map procedure in order to quantify the epistemic uncertainty in a uniform manner for all sites. The results of this decision are described in the next section.

Note that the M5 model actually consists of 0.8 wt for M5's and 0.1 wt for M3 and 0.1 weight for M4 models. We included the other two models with low weights so that drawing the M5 model would not lower the hazard to near zero at sites at large distances from M5's but close to M3's and M4's. This point will be discussed later with respect to individual sites. Without including the M3 and M4 models, sites such as North Anna that are lacking in nearby M5+ earthquakes would have many runs with very low hazard values that, we think, bias the uncertainty distribution to very low values of hazard. Using a large smoothing distance of 150 km raises the lower fractiles generated by distinct M5 earthquakes, but still results in highly asymmetric distributions. The use of the M3 and M4 models with small weights when the M5 model is drawn reduces this asymmetry.

The remaining node of the logic tree for the smoothed seismicity is that for the mblg to moment magnitude (Mw) conversion. This conversion is used in three places. First is the conversion from mblg to Mw when the attenuation relations are evaluated, since most of them are in terms of Mw. The code integrates the hazard using a truncated Gutenberg-Richter distribution for the incremental rates using mblg, the magnitude assigned in the catalog. The second procedure that uses the Mw to mblg conversion is when mblgmax is determined from Mwmax, since the maximum magnitudes are specified in terms of Mw. Finally, the Mw is used to calculate fault lengths used in the randomly-oriented strikes of the finite faults used in the hazard calculation. As in the

2002 maps, this node consists of the conversion formulas from Boore and Atkinson (1987) and Johnston (1996).

Logic trees were also used for the Meers and Cheraw faults. The recurrence times and characteristic magnitudes were varied for these faults. The variability of recurrence time was the same as that used for the New Madrid source. For the Cheraw fault, there is a node with branches for the characteristic and truncated Gutenberg-Richter recurrence models, as was used for the national maps. These are given equal weights, as in the national maps.

Results at Selected Sites

For most sites the mean seismic hazard curves derived from the Monte Carlo method are very similar to those derived from the national map model. In most cases, the probabilistic ground motions derived from the Monte Carlo mean hazard curve are within 15% of those from the mean curves used in the national seismic hazard maps, for an annual probability of 1×10^{-5}. For each site we used 200 runs to determine the mean, median, 15th and 85th percentile hazard curves. The hazard curves for the 200 runs were ranked at each ground motion level used in the calculations to determine the percentile hazard curves. We found that using greater than 200 runs did not significantly change the results.

Figure 4 shows the results for 10 Hz spectral acceleration (Sa) for Three Mile Island. This example shows very good agreement between the Monte Carlo mean curve and the mean from the national maps (Frankel et al, 2002). The median and mean curves from the Monte Carlo method are very similar in this case. In general the distribution of uncertainty is asymmetrical on a log probability log ground motion plot. That is, the ratio of the 50th/15th ground motions at 1×10^{-5} annual probability, for example, is much greater than the ratio of the 85th/50th ground motions at that probability. After some testing, we concluded that this asymmetry is largely caused by the M5+ seismicity model. In areas lacking magnitude 5 earthquakes in the catalog, the M5 models will produce very low hazard. This trend was mitigated by the large smoothing distances of 150 km used in some of the models and in the use of the M3 and M4 models with small weights in the draw for the M5 model.

The 1 Hz hazard curves (Figure 5) for Three Mile Island show a larger uncertainty than the 10 Hz curves. This is largely due to the greater epistemic uncertainty in the attneuaiton relations at 1 Hz compared to 10 Hz. The use of the double-corner model of Atkinson and Boore (1995) increase the epistemic uncertainty at 1 Hz. The Monte Carlo mean curve is slightly lower than that of the 2002 maps. We found that this is often true for sites in areas with significant historic seismicity. Part of this is due to the inclusion of the background zone in the Monte Carlo simulation as opposed to the adaptive weighting that does not use the background zone for the national maps for cells with substantial historic seismicity. At 1 Hz, the median curve in Figure 6 is significantly below the mean curve.

Figure 6 depicts the hazard curves for the South Texas site, an area removed from significant historic seismicity where the background zone hazard dominates. In this example the mean hazard curve from the Monte Carlo method is slightly higher than that from the 2002 maps. We found that this is generally true for sites where the background zones dominate. This is likely caused by the M3,4,5 models that had large smoothing distances and therefore, contribute to the hazard even at sites far from the epicenters.

The large uncertainty for the Texas site, compared to that at Three Mile Island, is largely caused by the contrast between the hazard from the background zone and the historic seismicity near the site. This contrast is greater for South Texas, since it is distant from clusters of historic earthquakes.

The hazard curves at Watts Bar for 10 Hz S.A. (Figure 7) show trends common to sites near seismically-active areas, in this case the eastern Tennessee seismic zone. The Monte Carlo mean is somewhat below the 2002 map mean, largely because of the inclusion of the background zone in the Monte Carlo mean. The median hazard curve is very similar to the mean hazard curve, as with Three Mile Island. The hazard curve uncertainty is again asymmetric with the larger uncertainty occurring for fractiles less than the median.

The site with the largest discrepancy between the Monte Carlo mean hazard curve and that derived from the 2002 map is Byron in northern Illinois (Figure 8). At 10 Hz S.A., the Monte Carlo mean at 10^{-5} is 76% of the map mean. The map mean slightly exceeds the 85^{th} fractile curve for probabilities from about 10^{-3} to 10^{-6}. This area of northern Illinois has had moderate earthquake activity in the past. In 1909 a mblg 5.0 earthquake occurred in this area. Three magnitude 4 earthquakes have also occurred in this region since 1860. The hazard curves may be sensitive to the re-sampling of the catalog, which can remove the single mblg 5.0 event for some re-sampled catalogs. Interstingly, the mean hazard curves for Braidwood and LaSalle are closer to the map means than those for Byron. Braidwood and LaSalle are closer to the 1909 event than Byron. We are currently analyzing the deaggregations for Byron to isolate the cause of the difference with the map mean hazard curves.

We have produced deaggregations for each site based on the procedure of RG1.165. A description of the deaggregations is provided in the Appendix after the figures. This Appendix also describes the filenaming procedure of the hazard curve and deaggregation files.

Figure 9 shows two types of deaggregations for Three Mile Island, for 5 Hz S.A. and an annual frequency of exceedance of 10^{-5}. The plot on the left shows the median hazard for each magnitude-distance bin, similar to the procedure of RG1.165, but only for 5 Hz. The plot on the right is the deaggregation for the median hazard curve. There are significant differences between the two deaggregations, although they both are dominated by earthquakes within 50 km of the site. The median hazard curve deaggregation lacks the contribution from close-in M5.0-5.5 events apparent in the

RG1.165 deaggregation. The median hazard curve deaggregation also has a greater contribution from M ≥7.0 earthquakes than the RG1.165 deaggregation.

These differences are mitigated when averaging the results from 5 and 10 Hz, as specified in RG1.165. The left plot on Figure 10 shows the deaggregation at Three Mile Island based on the full RG1.165 procedure, averaging the 5 and 10 Hz hazard values (annual frequency of exceedance of 10^{-5}). The right plot in Figure 10 displays the deaggregation of the median hazard curve, again averaging the 5 and 10 Hz hazard values. The two deaggregations are quite similar, although the RG1.165 procedure results in somewhat more contribution from earthquakes greater than 100 km from the site and less contribution from earthquakes with M6-7 within 20 km of the site. Averaging the two frequencies results in more similarity between the RG1.165 deaggreagtion and the median hazard curve deaggregation, compared to only using 5 Hz (Figure 9). Both deaggregations illustrate that the hazard to Three Mile Island at 5-10 Hz with an annual frequency of exceedance of 10^{-5} is dominated by earthquakes within 50 km of the site.

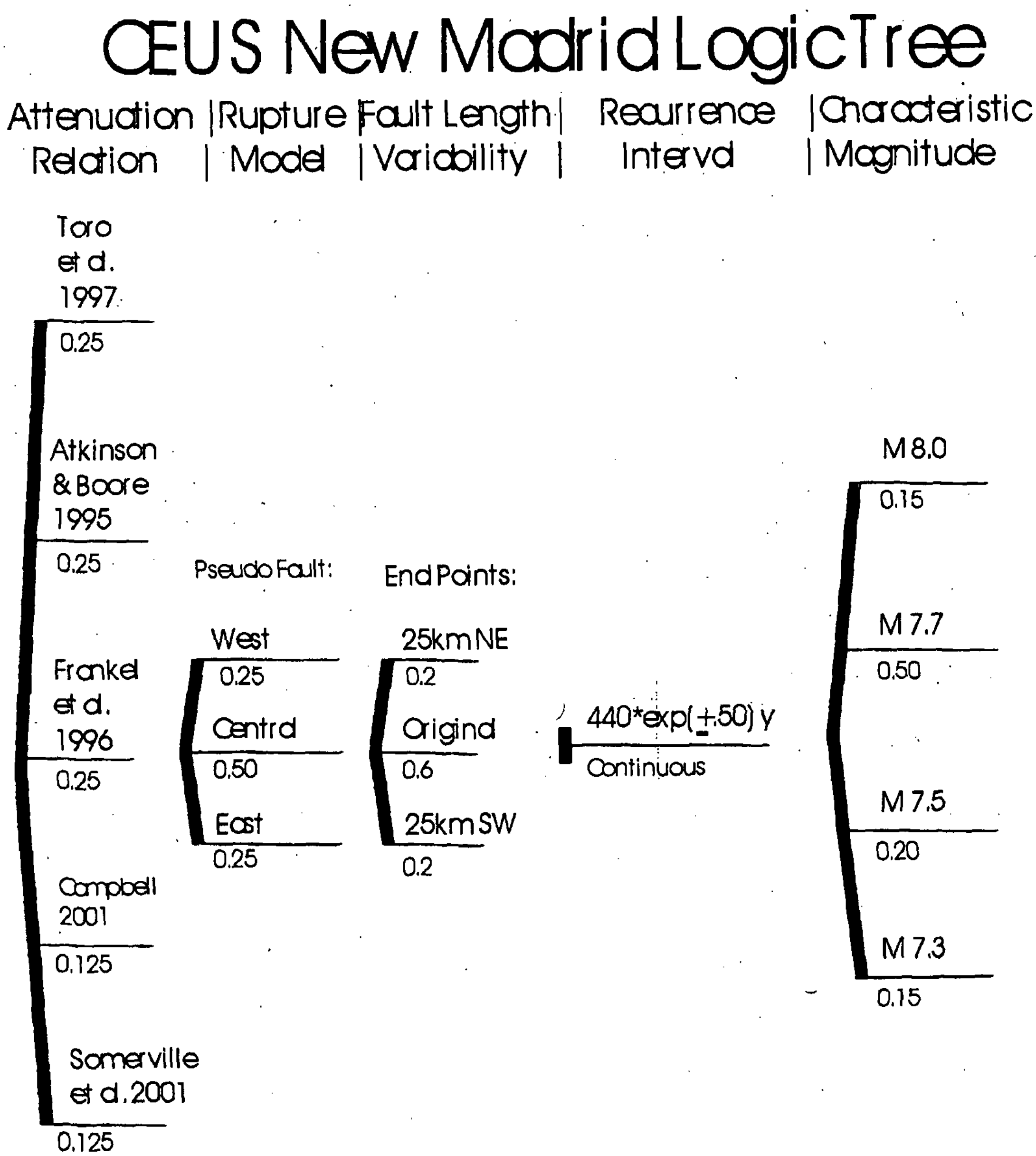


Figure 1. Logic tree used for New Madrid characteristic sources.

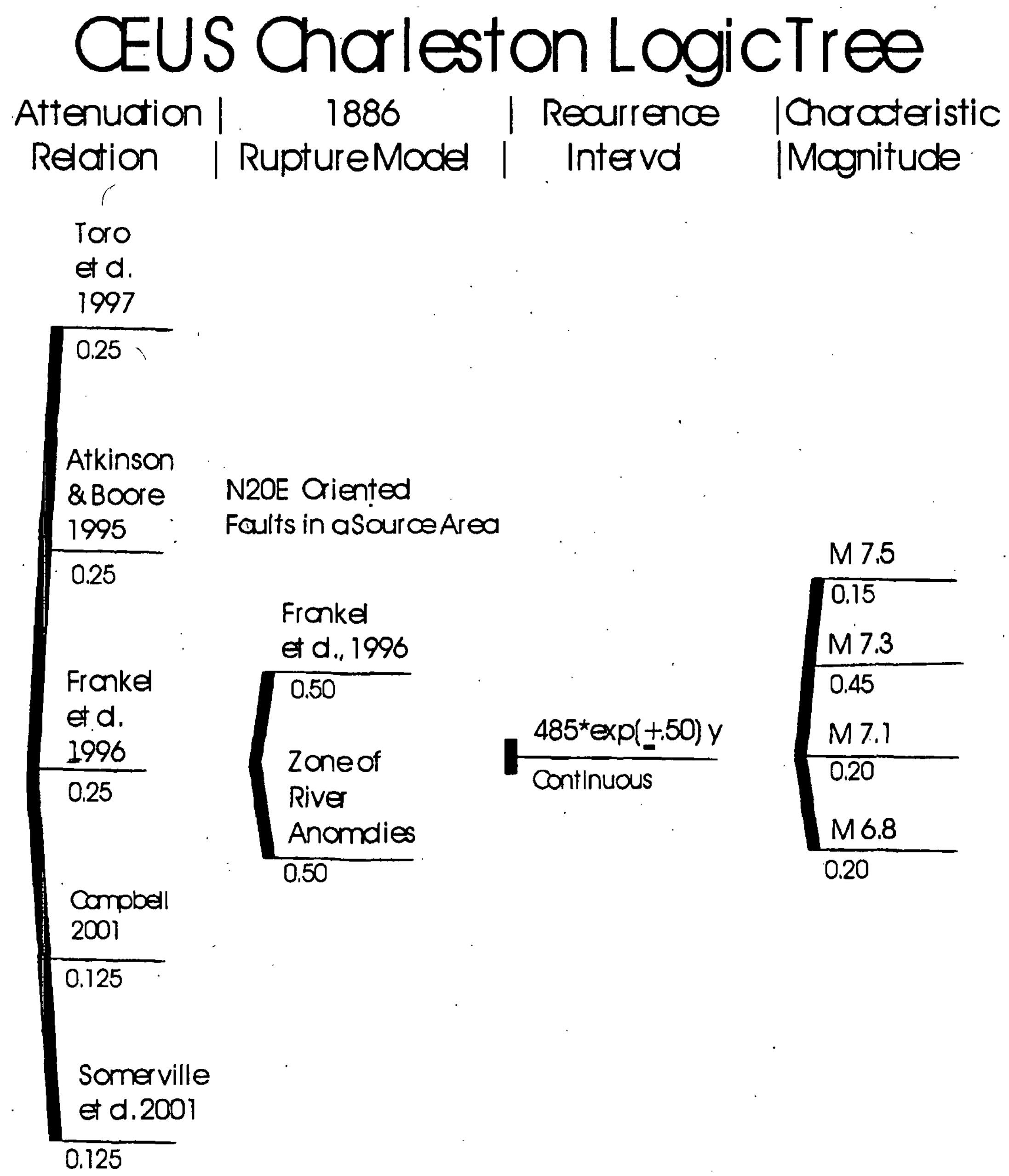


Figure 2. Logic tree used for Charleston, SC characteristic source.

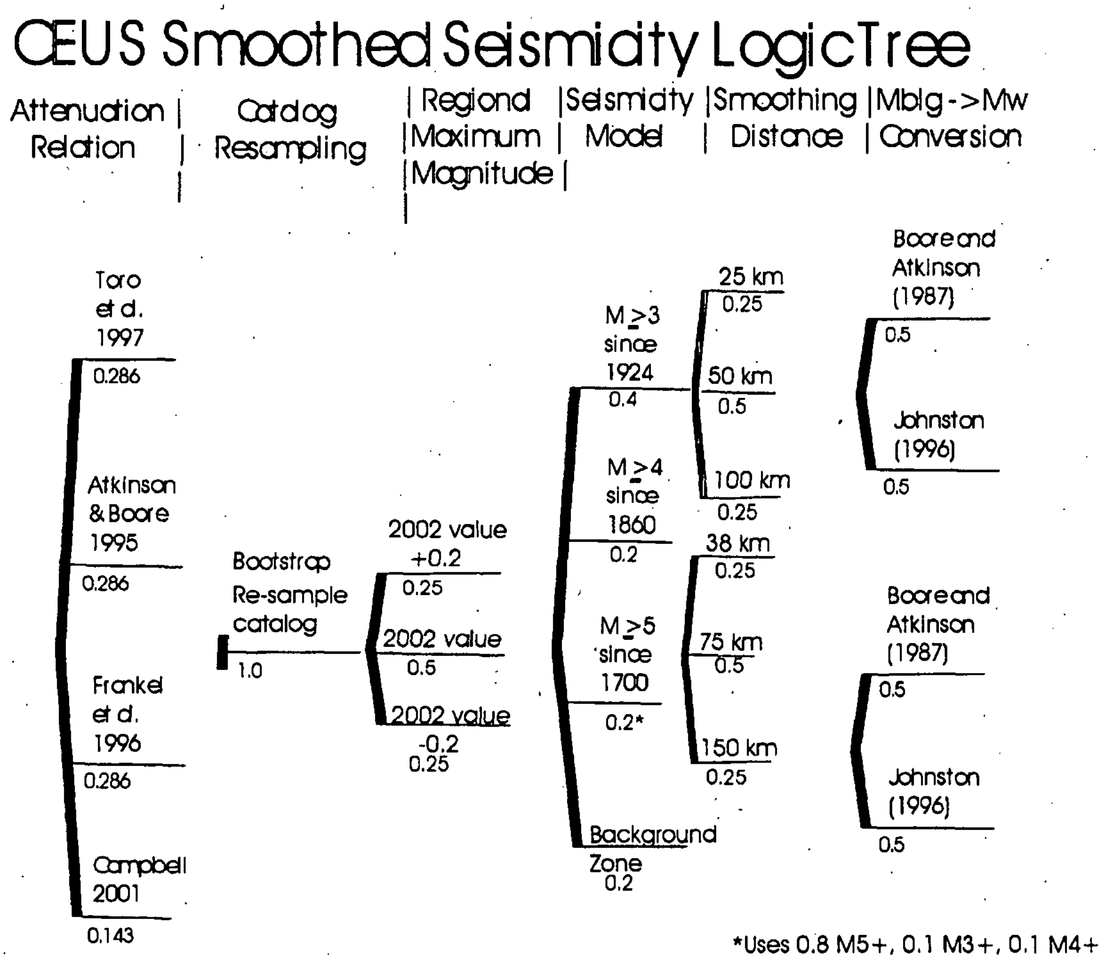


Figure 3. Logic tree used for spatially-smoothed seismicity.

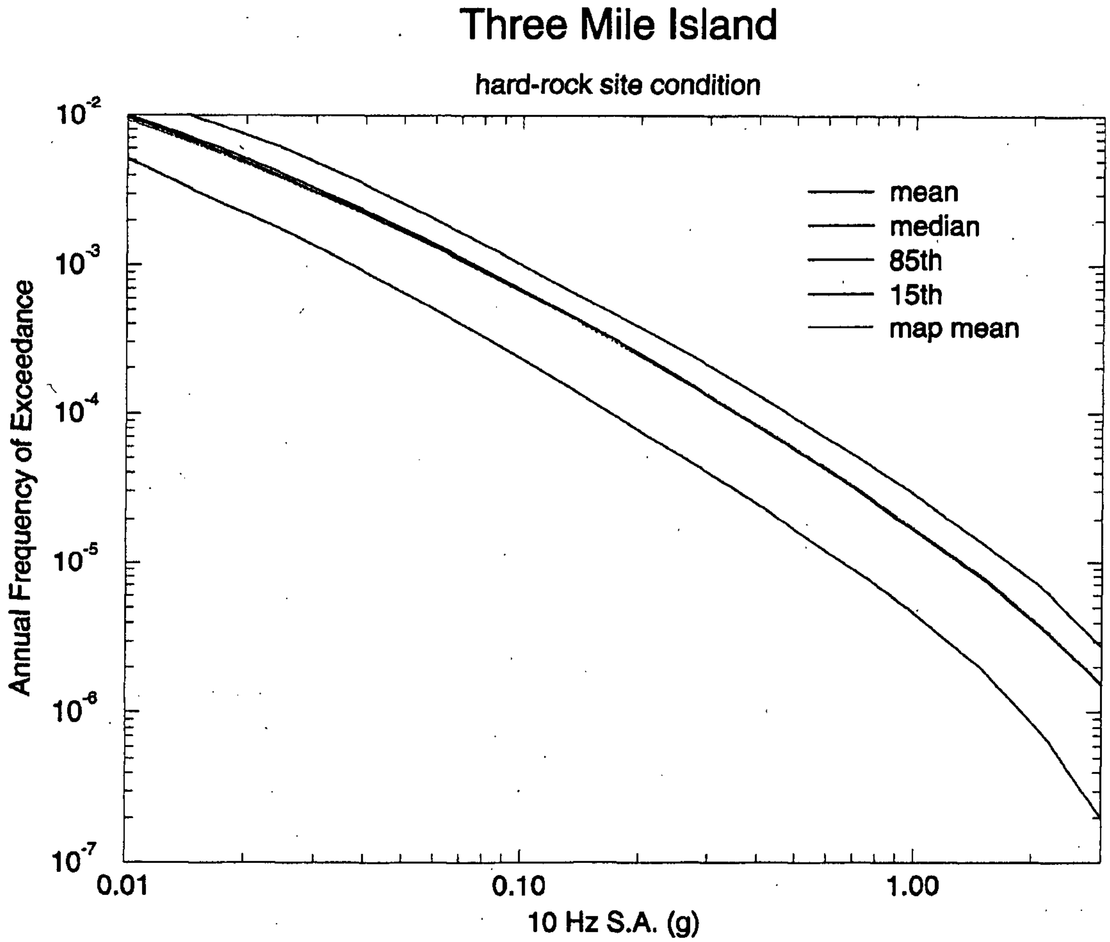


Figure 4. Hazard curves for 10 Hz S.A. at Three Mile Island. "Map mean" refers to the mean hazard curve derived from the procedure used in the 2002 national seismic hazard maps.

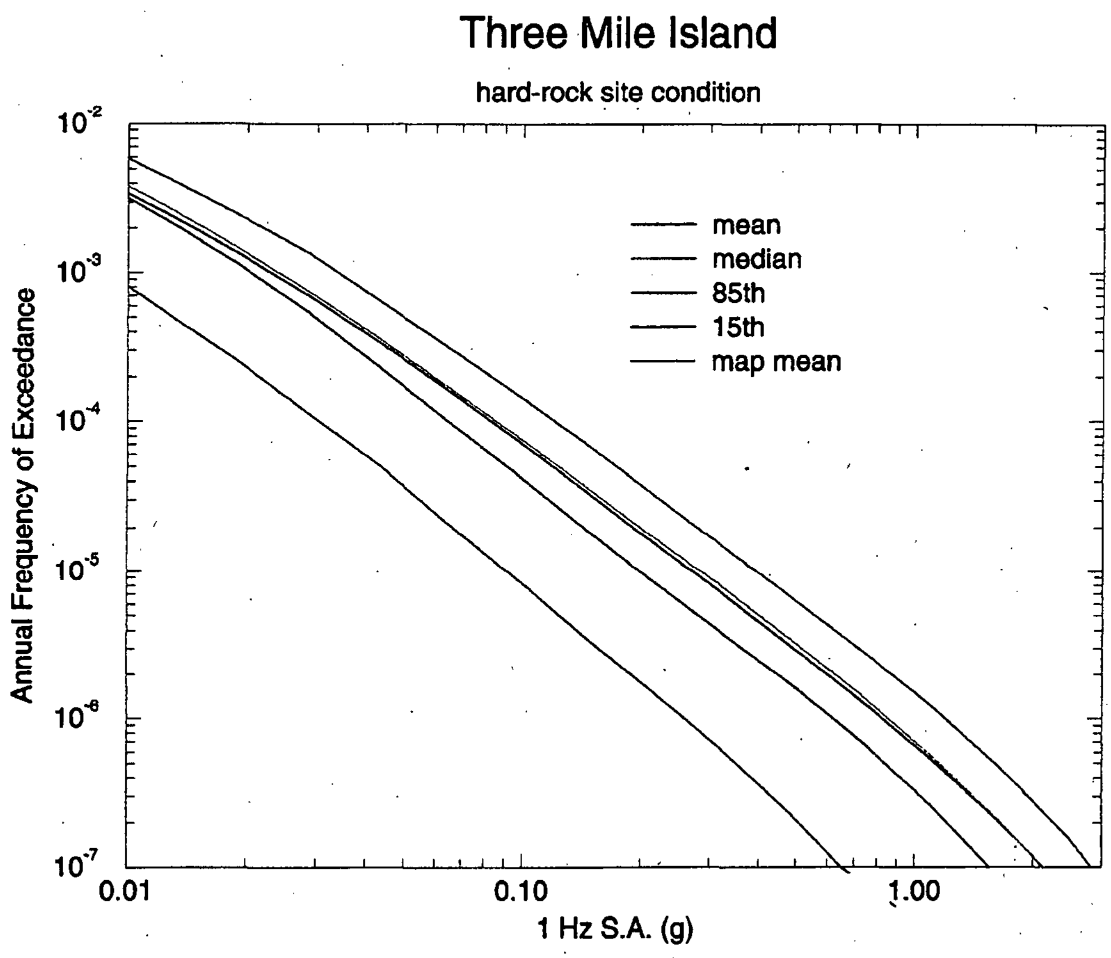


Figure 5. Hazard curves for 1 Hz S.A. at Three Mile Island.

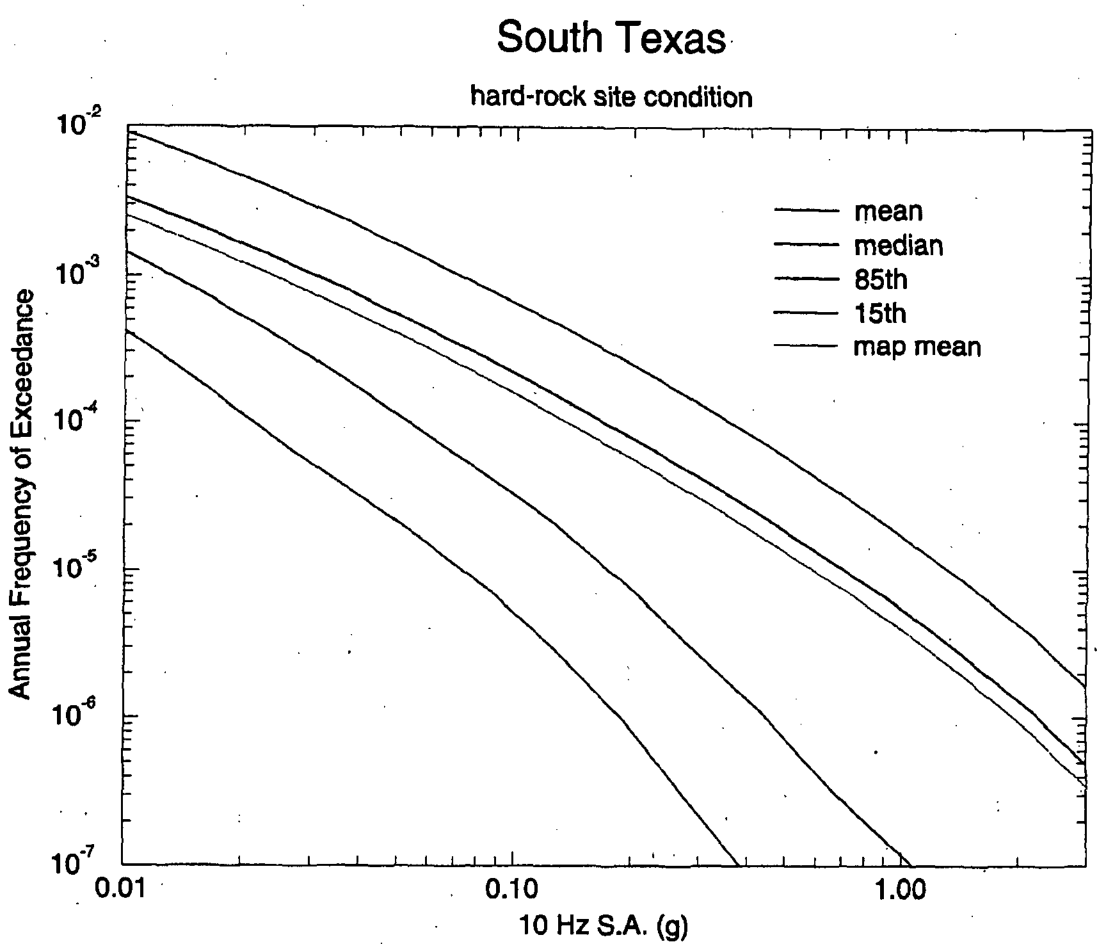


Figure 6. Hazard curves for 10 Hz S.A. at the South Texas site.

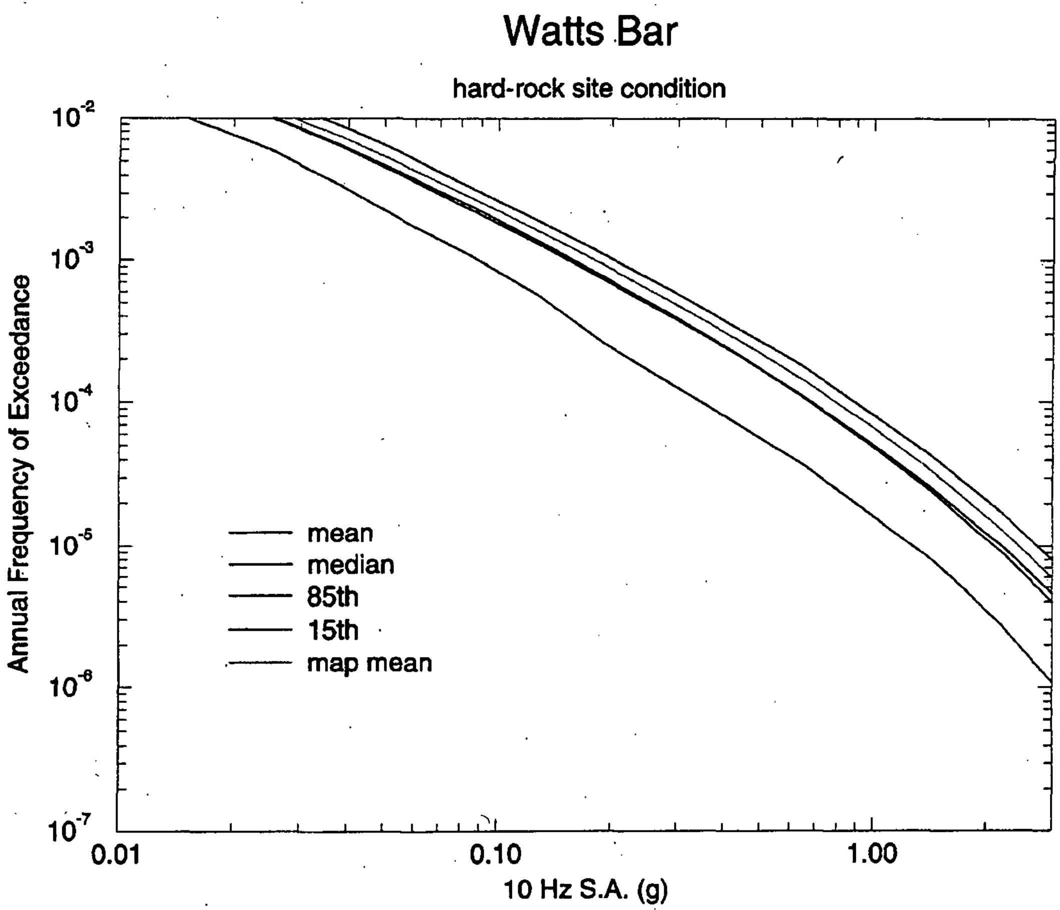


Figure 7. Hazard curves for 10 Hz S.A. at Watts Bar.

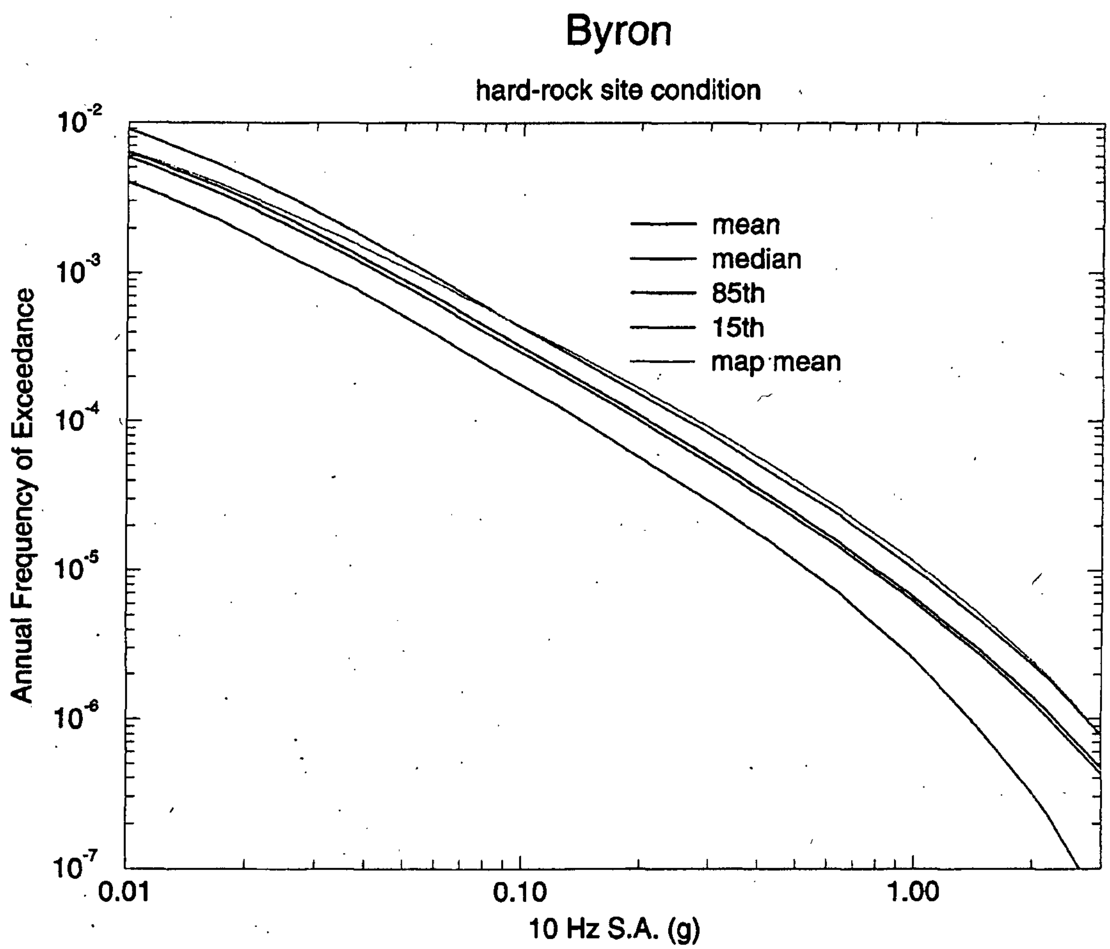


Figure 8. Hazard curves for 10 Hz S.A. at Byron.

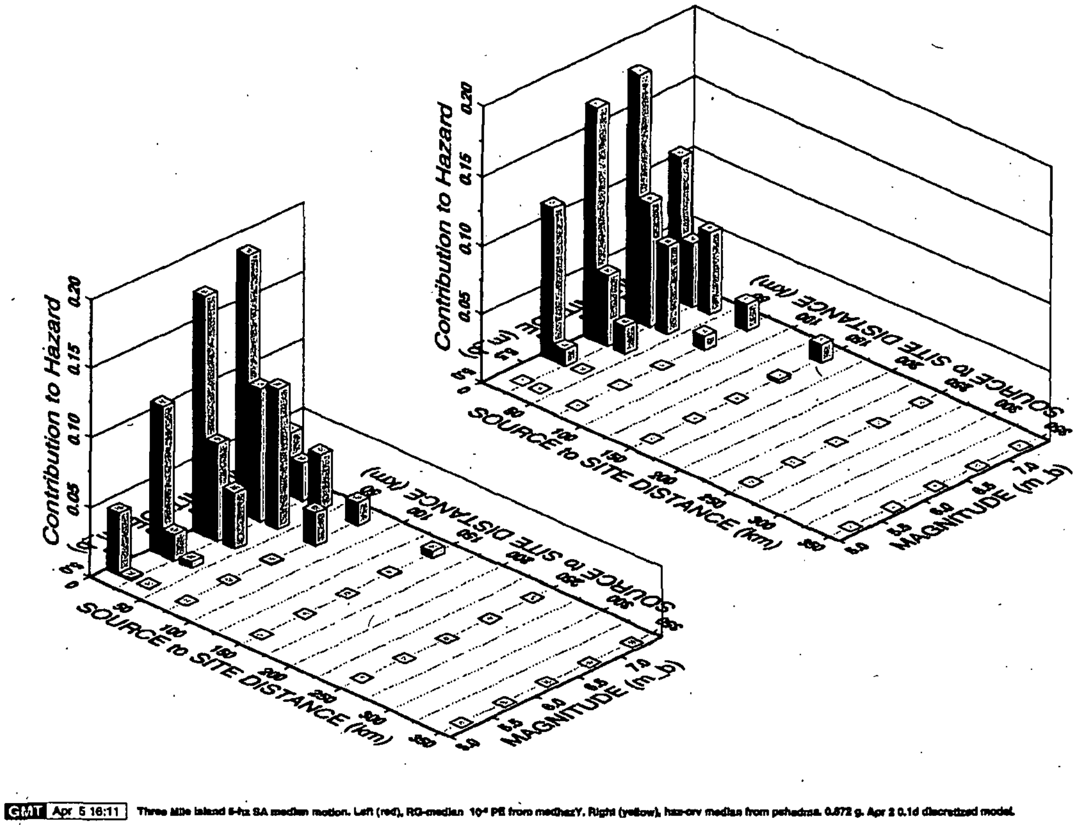
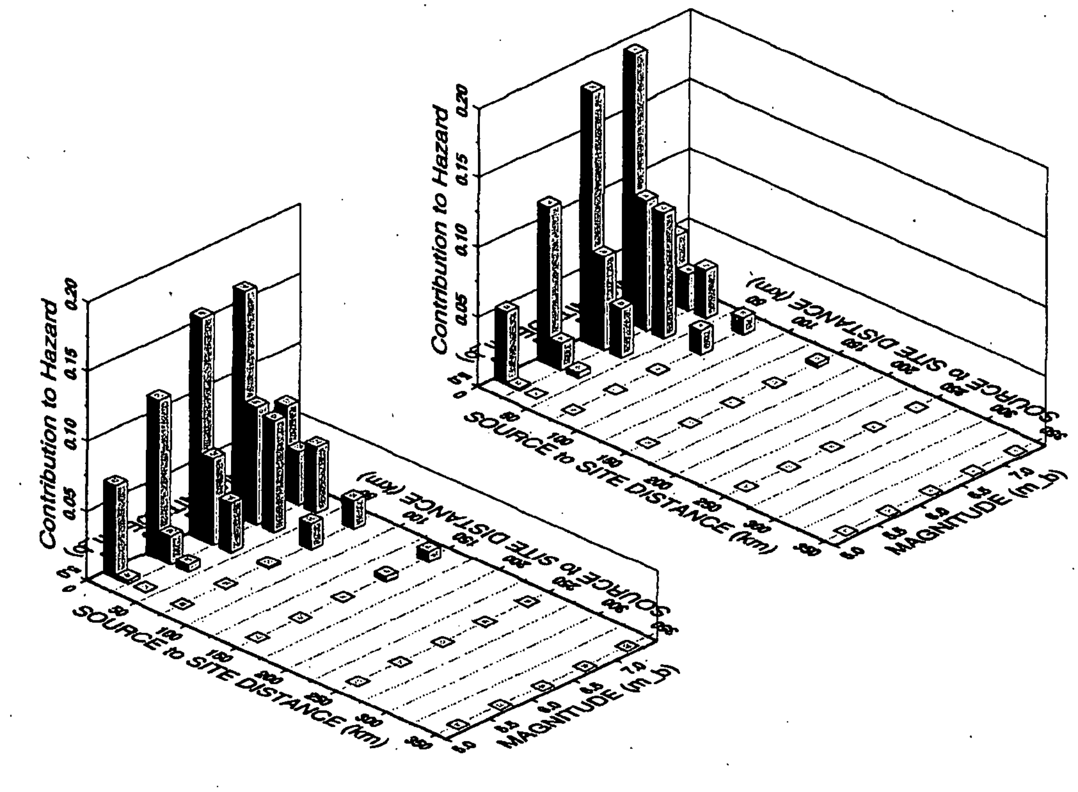


Figure 9. Deaggregations for Three Mile Island for 5 Hz S.A. at 10^{-5} annual frequency of exceedance. Deaggregation at left shows the median hazard for each magnitude-distance bin, following the RG1.165 procedure, but using a single frequency of spectral response. Plot at right is the deaggregation of the median hazard curve.



GMT Apr 5 15:22 Three Mile Island Combined 5.0 & 10.0-Hz SA. Left (red), MEAN motion 1.092 g 10^-5 PE from medhazY. Right (yellow), MEDIAN 1.093 g. Apr 2 0.1d discretized model.

Figure 10. Deaggregations for Three Mile Island from the combined 5 and 10 Hz S.A.values. Plot on the left is for mean hazard curve. Deaggregation on right is based on the RG1.165 procedure of taking the median of each magnitude-distance bin.

Appendix: Description of Deaggregation and Hazard Curve Files in Directory NRC29A (also on CD-ROM supplied to NRC)

The subdirectories in NRC29A contain logic-tree seismic hazard analysis for the 29 CEUS sites. The analysis in NRC29A uses a 0.1 degree gridding of background seismicity (NEW Apr 5 2004).in latitude and longitude.

The subdirectories also contain files with names that end in .4per.medhaz. These files are the deaggregated hazard according to NRC RG 1.165 Appendix C for 4 periods of Spectral Acceleration, 0.1 s, 0.2 s, 0.4 s, and 1.0 s, respectively. The last part of each of these contains the deaggregations for the "high-frequency" (5&10-Hz combined) and "low-frequency" (1 & 2.5 Hz combined) hazard. The last part also includes the analysis for 10**-4 and 10**-5 PE for both the mean and median. The "controlling earthquake" is listed at the end of each of these deaggs,

and the conditional controllng earthquake - for sources > 100 km away – is also listed if contribution is greater than 4.5% of total. You can get a controlling earthquake for each of the PEs. Everything should be annotated so that you should not have difficulty understanding what frequency, PE, and so on any given table corresponds to.

A second set of RG1.165 deaggregated periods is also available in all instances. These files have names ending 2per.medhaz. The 2 periods in question are 0.00 (PGA) and 0.04 (25-Hz), respectively. The file ending "medhaz" refers to median hazard but you will see that the mean hazard deaggregation is also contained in these files. You may also see that the mean hazard is the same as that contained in the files with names mapunc.mean.*.hzdt, where * is the period flag. Period flags are a10=1 sec SA, a04= 0.4-s SA, a02= 0.2-s SA, and a01= 0.1-s SA. Also, s04=0.04 s SA, and pga = PGA, horizontal peak ground acceleration.

Normalizations: For individual spectral periods, the deaggregation bin rates are "mean annual rates of exceedance" for the source(s) in the bin. When considering the mean, the bin rates sum to a good approximation of the target rate, which is either 10**-4 or 10**-5. When considering the median, RG1.165 Appendix C requires the median source contribution in each (magnitude,distance) bin. These conditional medians may add up to a quantity either greater than or less than the target rate, usually less than the target rate. The summed bin rates are printed at the end of each table for the median.

For combined spectral periods, the RG1.165 Appendix C procedure states that the binned contributions should be normalized in a specific manner. This normalization has been performed for the low- and high-frequency tables. Therefore, the sum of the contributions is 1.0 for the combined spectral period deagg. tables.

Note on File Names: The file names for the hazard curves are typically mapunc.ext1.ext2.hzdt. Ext1 gives the information on the fractile. The file name must be "decoded" somewhat. For example, ext1 = plus70cl means that the central part of the distribution contains 70% of the curves, and the plus means we are reporting the upper endpoint. Thus, the plus70cl curve is the same as the 85% fractile curve. Similarly minus70cl is the lower endpoint of this distribution, and is the same as the 15% fractile. Cl is a "confidence limit" band.

Note on Hazard-Curve Smoothing: The fractile curves have been smoothed somewhat. That is, the central curve is averaged with a few nearest neighbors on each side. When the number of Monte Carlo runs is 200, 10 curves on each side are included in the smoothing process. However, the deaggregated median bins are not smoothed in this manner. When the number of Monte Carlo runs is 200, the only averaging is between the 100th and 101st ordered contributions. We did not find any need to smooth the median deaggregation bin contents, nor does RG1.165 indicate any need to do so.

Units of motion in all files are g (1 g=980 cm/s/s).
For all files, PGA is denoted with 0.00-s period

For questions on interpreting deaggregation file contents, contact:
Steve Harmsen, USGS, 303 273 8567.
Email address: harmsen@usgs.gov

NRC FORM 8C
(7-94)
NRCMD 3.57

COVER SHEET FOR CORRESPONDENCE

USE THIS COVER SHEET TO PROTECT ORIGINALS OF MULTI-PAGE CORRESPONDENCE