## Mean and Modal ε in the Deaggregation of Probabilistic Ground Motion

### by Stephen C. Harmsen

Abstract An important element of probabilistic seismic-hazard analysis (PSHA) is the incorporation of ground-motion uncertainty from the earthquake sources. The standard normal variate  $\varepsilon$  measures the difference between any specified spectral-acceleration level, or SA<sub>0</sub>, and the estimated median spectral acceleration from each probabilistic source. In this article, mean and modal values of  $\varepsilon$  for a specified SA<sub>0</sub> are defined and computed from all sources considered in the USGS 1996 PSHA maps. Contour maps of  $\varepsilon$  are presented for the conterminous United States for 1-, 0.3-, and 0.2-sec SA<sub>0</sub> and for peak horizontal acceleration, PGA<sub>0</sub> corresponding to a 2% probability of exceedance (PE) in 50 yr, or mean annual rate of exceedance, *r*, of 0.000404.

Mean and modal  $\varepsilon$  exhibit a wide variation geographically for any specified PE. Modal  $\varepsilon$  for the 2% in 50 yr PE exceeds 2 near the most active western California faults, is less than -1 near some less active faults of the western United States (principally in the Basin and Range), and may be less than 0 in areal fault zones of the central and eastern United States (CEUS). This geographic variation is useful for comparing probabilistic ground motions with ground motions from scenario earthquakes on dominating faults, often used in seismic-resistant provisions of building codes. An interactive seismic-hazard deaggregation menu item has been added to the USGS probabilistic seismic-hazard analysis Web site, *http://geohazards.cr.usgs.gov /eq/*, allowing visitors to compute mean and modal distance, magnitude, and  $\varepsilon$  corresponding to ground motions having mean return times from 250 to 5000 yr for any site in the United States.

#### Introduction

A standard product of probabilistic seismic-hazard analysis (PSHA) is a set of site-specific ground motions, or 5%damped-response spectral acceleration (SA<sub>0</sub>) corresponding to a fixed probability of exceedance (PE) in a 50-yr period, or exposure time. Structural engineers may use these SA<sub>0</sub> data and related products, such as Uniform Hazard Spectra (Leyendecker *et al.*, 2000) and site-specific deaggregations, when designing to earthquake-resistant provisions of building codes. Private- and public-sector planners consult maps of SA<sub>0</sub> and related PSHA products to support earthquake mitigation, preparedness, and management policy decisions.

Whenever a theoretical value, such as SA<sub>0</sub>, is translated into a set of practical decision rules, such as design to withstand some fraction of the uniform hazard-spectral ordinates, which are approximately the 2% in 50 yr SA<sub>0</sub> (Leyendecker *et al.*, 2000), it is of some practical interest to know how SA<sub>0</sub> is related to ground motion from potential sources. The deaggregation of source contributions (McGuire, 1995) informs us about the most likely magnitude, *M*, and distance, *R*, contributing in a mean annual sense, to the PE, given that M,R pair. Deaggregation also indicates the quantile (location) of SA<sub>0</sub> on the conditional SA distribution given M and R. The Greek letter epsilon ( $\varepsilon$ ) has been chosen to represent this quantile (McGuire, 1995). When we compute the PE of a specific ground motion at the site, then  $\varepsilon$  for each M,R is a specific number, which we designate  $\varepsilon_0$ .

According to the ground-motion prediction equations currently invoked in PSHA by the USGS and many other agencies, SA has a lognormal conditional distribution (Abrahamson and Shedlock, 1997). That is, sa = log (SA) from sources having a given M,R (and perhaps other parameters) that might be recorded at any site is approximately normally distributed, with mean  $\mu_A$  and standard deviation  $\sigma_A$ . Some caveats about the relationship between empirical SA distributions and the conditional ground-motion probability density functions (p.d.f.) used in PSHA are mentioned in the Appendix.

Here, the subscript A represents one of the groundmotion prediction models that is used in the PSHA.  $\mu_A(M,R)$ and  $\sigma_A(M,R)$  may exhibit dependence on geographic region, fault maturity, and fault style, for example, normal, strike slip, and reverse slip. Maximum likelihood values of  $\mu$  and  $\sigma$  differ from model to model because analysts use different physical/statistical models to explain the observations. Following McGuire (1995), we define

$$\varepsilon_{0A}(S) = (sa_0 - \mu_A) / \sigma_A, \qquad (1)$$

that is, the standardized value of  $sa_0 = log(SA_0)$  or the (signed) distance, in standard deviations, from  $sa_0$  to the logarithmic mean ground motion,  $\mu_A$ , for a specific source *S*, and attenuation model, *A*. Every source in a PSHA is defined with a mean recurrence rate, a magnitude, and a distance to site. In the western conterminous United States (WUS) a source is frequently a mapped fault. In the United States east of the Intermountain Seismic Belt, few sources can be associated with mapped faults, but are instead areas characterized by historical activity rates.

Uncertainty about some source properties, including stress drop and directivity, propagation properties, including regional variations in apparent attenuation, and site properties, including degree of nonlinearity of response as SA increases, is called epistemic uncertainty. Epistemic uncertainty is that part of the total uncertainty that promises to diminish as understanding of the Earth increases. Although epistemic uncertainty is not a central topic of this paper, its effect on estimates of  $\varepsilon_0$  is illustrated in a subsequent section. The remaining uncertainty is called aleatory (random) uncertainty. For the 1996 USGS PSHA maps, aleatory uncertainty in SA for a given *M*,*R* is equated to the strong-motion data regression's estimated total  $\sigma$  for a given attenuation model. Aleatory uncertainty in future earthquake locations and magnitudes are discussed in Frankel *et al.*, (1996).

PSHA methodology suggests the inclusion of several ground-motion prediction models to account for epistemic uncertainty in ground-motion prediction (SSHAC, 1997). The 1996 National Seismic Hazard Maps assign several analysts' ground-motion prediction equations equal weight in the PSHA (Frankel *et al.*, 1996), which implies that these equations are equally likely to be valid. Weights sum to 1. In practice, new models may be added and weights modified in future revisions.

#### Mean $\varepsilon$ ( $\overline{\varepsilon}_0$ and $\overline{\varepsilon}$ ) for a Specified PE

For the set of sources that contribute to a specified seismic hazard at the site, we define the mean or average  $\varepsilon_0$ ,

$$\overline{\varepsilon}_{0} = \left( \sum_{A} \sum_{S} \varepsilon_{0A} Wt(A) \lambda_{s} \operatorname{Pr}[\operatorname{SA} \geq \operatorname{SA}_{0} \mid S, \mu_{A}, \sigma_{A}] \right) / \left( \sum_{A} \sum_{S} Wt(A) \lambda_{s} \operatorname{Pr}[\operatorname{SA} \geq \operatorname{SA}_{0} \mid S, \mu_{A}, \sigma_{A}] \right)$$
(2)

Here, Wt(A) is the weight (>0) applied to terms computed with ground-motion prediction model A.  $\lambda_s$  is the estimated mean annual rate of occurrence of source S. Earthquake rate estimates are discussed in Petersen *et al.* (1996) for California sources, in Wesson *et al.* (1999) for Alaska sources, and in Frankel *et al.* (1996) for all other United States sources. The Pr[] factor is the conditional probability that earthquake *S* will produce a ground-motion exceedance, given the maximum likelihood parameters of the ground-motion uncertainty distribution for model *A*. In the 1996 USGS seismichazard calculations, and in later USGS PSHA products, this probability is computed by integrating the normal density function from sa<sub>0</sub> = log(SA<sub>0</sub>) to  $\infty$ ,

$$\Pr[SA \ge SA_0 \mid S, \mu_A, \sigma_A] = \frac{1}{\sigma_A \sqrt{2\pi}} \int_{sa_0}^{\infty} \exp\left(\frac{-(y - \mu_A)^2}{2\sigma_A^2}\right) dy.$$
(3)

With this definition, ground motion is not bounded. Note that equation (2) can be interpreted as if all of the PE is concentrated at  $SA = SA_0$ . This definition is therefore equivalent to the definition by McGuire (1995) of mean  $\varepsilon$ .

The denominator of equation (2) is the mean annual rate of exceedance of  $SA_0$ , *r*. Making the standard PSHA assumption that ground-motion exceedances are a Poisson (memoryless) process (Cornell, 1968), the mean annual rate of exceedance  $r = 4.04 \times 10^{-4}$  corresponds to the 2% PE in 50 yr, and  $r = 2.107 \times 10^{-3}$  corresponds to the 10% PE in 50 yr. A mean annual rate *r* is equivalent to a mean return time of 1/*r*. In many PSHA reports the various expressions mean return time, mean annual rate *r*, or X% PE in 50 yr are used interchangeably to describe the hazard, for independent Poisson random exceedances. The distinction between mean return time, for exceedances, and mean recurrence time, for a specific source, which can be a point source, a fault, or an area, is an important one.

The definition of  $\overline{\varepsilon}_0(r)$  in equation (2) is used in the  $\varepsilon$ -contour maps that follow. Summation is over the same sources and ground-motion prediction models as were used in the 1996 USGS seismic-hazard maps. In this report a bar over a PSHA statistic, such as  $\overline{M}$  or  $\overline{R}$ , implies a statistical mean value, using the same source and attenuation-model weighting scheme as in equation (2). Bazzurro and Cornell (1999) note that  $\overline{R}$  when computed with multiple attenuation-model distance metrics, as in equation (2), may be a questionable quantity. For example, any or all of the distance metrics discussed by Abrahamson and Shedlock (1997) might be invoked during a deaggregation of PGA, but the resulting source-to-site distance metric.

A natural definition of mean  $\varepsilon$ , conditional on SA  $\ge$  SA<sub>0</sub>, utilizes the normal distribution of ground-motion exceedances for a given *M*,*R*, and attenuation model *A*:

$$\overline{\varepsilon}_{0A} = E[\varepsilon \mid \varepsilon \ge \varepsilon_{0A}] = \frac{\int_{\varepsilon_{0A}}^{\infty} xn(x)dx}{\int_{\varepsilon_{0A}}^{\infty} n(x)dx}$$

where n(x) is the standard normal density function. If in equation (2)  $\overline{\varepsilon}_{0A}$  is substituted for  $\varepsilon_{0A}$ , the resulting left-handside is the conditional mean  $\varepsilon$  or  $\overline{\varepsilon}$ . We report  $\overline{\varepsilon}$  as well as  $\overline{\varepsilon}_0$  at our web site of interactive deaggregations, but do not consider it further here.

#### Modal-Event Epsilon ( $\hat{\varepsilon}_0$ ) and Modal Epsilon ( $\varepsilon^*$ ) for a Specified PE

Statistical means have some desirable as well as undesirable properties for specific applications. A desirable property is that the mean, at least as defined in equation (2), is invariant with respect to binning schemes. An undesirable property of the triple  $\overline{M}, \overline{R}, \overline{\varepsilon}_0$  is that it may be associated with an unlikely source for bimodal and multimodal deaggregated hazard distributions, which are common at both CEUS sites (Harmsen *et al.*, 1999) and western United States sites (Harmsen and Frankel, 2001). For applications that choose to consider only those sources that have a likely probability of future occurrence, an alternative to mean source statistics is needed. PSHA modal (most likely source) statistics are frequently suggested. For this article, hats on statistics indicate modal values consistent with McGuire (1995) and Chapman (1995).

To define  $\hat{\varepsilon}_0$  some background is needed. During a PSHA, information about each source is converted into a source-to-site distance,  $R_A$ , and moment magnitude, M. The summation over S is, in practice, a double summation over  $R_A$  and **M**. The subscript A is a reminder that distance to source is defined differently for different ground-motion prediction models, A. M also has an A-dependence for sites in the CEUS for the 1996 USGS seismic-hazard calculations (Harmsen et al., 1999, equations 5 and 6). When deaggregation is performed, seismic hazard is computed and summed for sources assigned to bins. Following McGuire (1995), a bin contains sources with magnitudes in the range  $\mathbf{M}_i \pm \Delta M/2$ , and distances in the range  $\mathbf{R}_i \pm \Delta R/2$ , where i and *j* are indices on magnitude and distance bin dimensions, respectively. For this article,  $\hat{\varepsilon}_0$  is computed from sources in the (M,R) bin with the largest mean annual rate of exceedance.

USGS deaggregations use regionally constant  $\Delta R$  and constant  $\Delta M$ . Magnitude lower and upper limits, and distance upper limit are site-dependent, and possibly application-dependent, quantities. Various source-binning schemes are encountered in applications. Some analysts recommend deaggregations with distance bins whose  $\Delta R$  is equally spaced in log *R* rather than *R*. Some applications require that we consider sources associated with known structures separately from randomly occurring sources. For sites in the Pacific Northwest, we keep deep and shallow random sources separate. At many of those sites, events that occur in the subducting Juan de Fuca plate present significant hazard in the USGS 10% and 2% in 50 yr PSHA maps (Harmsen and Frankel, 2000). Application-specific source-binning requirements can affect estimates of the modal, or most likely, magnitude, distance, and epsilon (Bazzurro and Cornell, 1999).

Once the source-binning scheme is defined, we define a modal-event  $\varepsilon$ ,  $\hat{\varepsilon}_0$  as the  $\overline{\varepsilon}$  computed for sources confined to the (M,R) bin having the greatest hazard or mean annual rate of SA exceedances. Let  $r_i$  = mean annual rate for sources in the *i*th bin, and let *j* be the index of the bin such that  $r_j = \max(r_i)$ . Then, we define

$$\hat{\varepsilon}_0 = \sum_A \sum_{S_j} \varepsilon_{0A} \lambda(S_j) \Pr\left[ \mathsf{SA} \ge \mathsf{SA}_0 \mid S_j, \mu_A, \sigma_A \right] Wt(A) / r_j,$$
(4)

where  $S_j$  refers to a source in the *j*th bin. This definition of  $\hat{\varepsilon}_0$  corresponds to the definitions of modal distance,  $\hat{R}$ , and modal magnitude,  $\hat{M}$  in Harmsen *et al.* (1999) and Harmsen and Frankel (2001) and is consistent with the modal event in Chapman (1995). The modal-event triple,  $(\hat{R}, \hat{M}, \hat{\varepsilon}_0)$  and the mean  $(\bar{R}, \bar{M}, \bar{\varepsilon}_0)$ , are reported in the interactive web-site seismic-hazard deaggregations available at the URL *http:// geohazards.cr.usgs.gov/eq/.* The decision to define modal statistics from the joint distribution of *M* and *R* is based on the argument that the most likely *M*, *R* pair represents a physically plausible event, that is, one of the sources (or binned set of similar sources) considered by the probabilistic model.

The physical meaning of equation (4) is the quantile (location relative to median ground motion) of SA<sub>0</sub> for the most likely (magnitude,distance) pair in the hazard deaggregation. Using equation (4), the target motion, SA<sub>0</sub>, is recovered from the triple  $(\hat{M}, \hat{R}, \hat{\varepsilon}_0)$  when just one attenuation model is used in the PSHA and just one source occupies the modal-event bin. If a structure is designed such that the 2% in 50 yr ground motion represents a collapse level (cf. Leyendecker *et al.*, 2000), and if the modal event (most-likely M, R pair) occurs, that collapse level is  $\hat{\varepsilon}_0 \sigma$  from the logarithm of the median ground motion from the modal event. In this sense,  $\hat{\varepsilon}_0$  is a useful statistic for thinking about probabilistic ground motions in a seismic-engineering context.

The modal-event  $\hat{\varepsilon}_0$  of equation (4) is distinct from another potentially useful modal  $\varepsilon$  obtained by deaggregating  $\varepsilon$  independently of R and  $\mathbf{M}$ . ( $\mathbf{M}, R, \varepsilon$ ) deaggregation is performed at our interactive deaggregation web site. The modal triple ( $\mathbf{M}^*, R^*, \varepsilon^*$ ) is the mean magnitude, mean distance, and epsilon-interval for sources in the ( $\mathbf{M}, R, \varepsilon$ ) bin having the greatest mean annual rate of exceedances, that is, the mode of the joint conditional distribution of binned  $M, R, \varepsilon$  (Bazzurro and Cornell, 1999). This triple is now reported at the interactive deaggregation web site;  $\varepsilon^*$  is reported as an interval (e.g.,  $1\sigma$  to  $2\sigma$ ) rather than a scalar. As stated by Bazzurro and Cornell (1999), ( $\mathbf{M}^*, R^*, \varepsilon^*$ ) can correspond to SA > SA<sub>0</sub>. For example, if  $\hat{\varepsilon}_0 < O$ , the triple ( $\hat{M}, \hat{R}, 0$ ) is more probable than ( $\hat{M}, \hat{R}, \hat{\varepsilon}_0$ ) but corresponds to a higher ground motion.

We generally find that  $(M^*, R^*) \cong (\hat{M}, \hat{R})$  for the 2% in 50 yr PE for United States sites when using the relatively coarse intervals,  $\Delta M = 0.5$ ,  $\Delta R = 10$  or 25 km, and  $\Delta \varepsilon =$ 1 used at our interactive web site. Some sites in the CEUS where these pairs differ are given in Harmsen *et al.* (1999). A WUS site where they can differ is discussed in a section that follows.

Because the definitions of modal statistics such as  $\hat{\varepsilon}_0$ sum over events, sensitivity to binning details should be considered in some applications. A calculation illustrating sensitivity of modal-event estimates to  $\Delta R$  is given below for a site near the 1992 Landers, California, mainshock. The triple  $(M^*, R^*, \varepsilon^*)$  is sensitive to the definition of the  $\varepsilon$  bin boundaries as well as those of R and  $\mathbf{M}$ . Using  $1\sigma$ -wide bins, we find that  $(M^*, R^*)$  can vary significantly as the  $\varepsilon$  bin boundaries are changed from  $k\sigma$  to  $(k + 1/2)\sigma$ ,  $k = 0, \pm 1, \pm 2$ , sometimes equaling  $(\hat{M}, \hat{R})$  sometimes not.

#### How Different Attenuation Models Affect Epsilon Estimates

In many studies, if two processes yield response values  $s_1$  and  $s_2$ , then the sample mean,  $0.5(s_1 + s_2)$ , is an estimate of the mean response. In PSHA, equations (2) and (4) indicate that if  $\varepsilon_{01}$  and  $\varepsilon_{02}$  are the epsilons for source *S* corresponding to attenuation models 1 and 2, respectively, then  $\overline{\varepsilon}_0$  is in general not equal to  $0.5(\varepsilon_{01} + \varepsilon_{02})$  even though the two attenuation models have equal weights, Wt(1) = Wt(2) = 0.5. Another weighting factor is important, the conditional PE given the occurrence of *S*. This factor is often considerably larger for one of the several attenuation models being used in the PSHA.

To illustrate how  $\overline{\varepsilon}_0$  and  $\hat{\varepsilon}_0$  depend on ground-motion prediction equations used in the USGS 1996 PSHA, a realistic  $\hat{\varepsilon}_0$  calculation is now performed. We consider the 0.2sec spectral acceleration from an M 8 New Madrid Seismic Zone (NMSZ) source at a CEUS rock site 200 km from that source. Suppose that the PSHA 2% in 50 yr SA<sub>0</sub> at that site is 0.6g (g = standard gravity  $\approx$  9.8 m/s<sup>2</sup>). Two attenuation models are used for 1996 USGS CEUS PSHA calculations (Frankel et al., 1996). According to the attenuation model of Frankel et al., or A1, the median 0.2-sec SA for an M 8 earthquake at 200 km is 0.307g, and according to the attenuation model of Toro et al. (1997), or A2, the median SA is 0.159g. For both models, the CEUS 0.2-sec logarithmic standard deviation,  $\sigma$ , is 0.751 (natural log units). For the 1996 seismic-hazard maps,  $Wt(A_i) = 0.5$ , i = 1, 2. Figure 1 illustrates the two ground-motion p.d.f.s for this example. The area of each of the curves to the right of the solid vertical line is the conditional probability of observing an exceedance of 0.6g at the site, given the occurrence of the M 8 source, for each attenuation model.  $\ln(0.6)$  is  $\ln(\text{median}) +$  $0.891\sigma$  for A<sub>1</sub>, and is ln(median) +  $1.766\sigma$  for A<sub>2</sub>, so that  $\varepsilon_1$  is 0.891 and  $\varepsilon_2$  is 1.766. Consulting a tabulation of the standard normal distribution upper tail, we determine that the weighted conditional probability is 0.5(0.1866 + 0.0387)= 0.1127. The mean annual rate of occurrence for a CEUS M 8 source is 0.001 (1000 yr mean recurrence) (Frankel et al., 1996). At the site under consideration, the M 8's contribution to the mean annual rate of SA exceedances is 0.001  $\times$  0.01127 = 0.0001127, or 27.9% of the total rate,



Figure 1. Ground-motion uncertainty p.d.f.s for an M 8 earthquake at distance 200 km from two ground-motion attenuation models used in the 1996 National Seismic Hazard maps. For each of these, the logarithmic  $\sigma$  is 0.751. The vertical line is located at  $\ln(SA_0) = -0.51$ , or  $SA_0 = 0.6g$ . PSHA computes exceedance probability, Pr[SA > SA\_0] = Pr[ln(SA) > ln(SA\_0)], which is the area of the curves to the right of the vertical line. Median motions are for a BC rock site, with average Vs = 760 m/sec in the top 30 m.

0.000404, associated with the 2% in 50 yr PE. Suppose this **M** 8 source is the modal source for the deaggregation. Therefore, according to equation (4),  $\hat{\varepsilon}_0 = (0.5 \times 0.891 \times 0.1866 + 0.5 \times 1.766 \times 0.0387) / 0.1127 = 1.037$ . Note that  $\hat{\varepsilon}_0$  is closer to  $\varepsilon_1$  (0.891) than to  $\varepsilon_2$  (1.766).

For the 1.0-sec SA, the attenuation functions  $A_1$  and  $A_2$  predict median ground motions of 0.134*g* and 0.080*g*, respectively, for this source and site. For 1-sec SA, the greater similarity between median estimated ground motions implies less variation in  $\varepsilon$  for  $A_1$  and  $A_2$ . Although this example illustrates effects on epsilon, the same recipe for averaging over attenuation functions is used to determine  $\overline{M}$ ,  $\hat{M}$ ,  $\overline{R}$ , and  $\hat{R}$ . Thus, all of these statistics are closer to the values associated with the highest-predicted-response attenuation function(s), when the  $Wt(A_i)$  are equal. Bender and Perkins (1993) note that the hazard curve that results from averaging over attenuation functions is closest to the curve for the function that predicts the highest response, and ask if that attenuation function should be given lower weight when averaging.

If the conditional probability of exceedance of  $SA_0$  given *S* and  $A_2$  is low, then *S* is very likely not the modal source when deaggregating hazard with  $Wt(A_2) = 1$ , even though *S* is the modal source when deaggregating with  $Wt(A_1) = 1$ . At many sites in the CEUS, the modal source using just the attenuation model of Frankel *et al.* is an **M** 8 NMSZ event at regional distances, while the modal source at that site is a small to moderate (**M** < 6) local earthquake when using the attenuation model of Toro *et al.* (1997). This

circumstance is likely to exist when the deaggregated seismic-hazard distribution shows a bi- or multimodal pattern, with the primary peak corresponding to NMSZ **M** 8, and the secondary peak corresponding to a local moderate earthquake. In these cases, applications often should work with (e.g., design to resist) both (several) of these  $M,R,\varepsilon$  triples. Beta earthquakes, in the terminology of McGuire (1995) are modal  $M,R,\varepsilon$  triples associated with each considered attenuation function.

# How Source-Recurrence Times Affect $\varepsilon$ for a Specified PE

We next discuss the relationship of  $\varepsilon$  to event occurrence rates or probabilities. If a probabilistic source *S* has mean annual rate  $\lambda$ , that is, mean recurrence time  $T = 1/\lambda$ , then the mean annual rate for which logarithmic ground motion, sa = log (SA), from that source exceeds the mean logarithmic motion,  $\mu$ , is 0.5 $\lambda$ . For any  $\varepsilon$ , and assuming an unbounded upper ground-motion limit, the annual rate,  $r_S$ , for which SA exceeds  $\mu_A + \varepsilon \sigma_A$  from that source is  $\lambda N(\varepsilon)$ , where  $N(\varepsilon)$  is area under the upper tail of the standard normal distribution,

$$r_{\rm S} = \frac{\lambda}{\sqrt{2\pi}} \int_{\varepsilon}^{\infty} {\rm e}^{-y^2/2} \, dy.$$
 (5)

The subscript A refers to a ground-motion attenuation function. If only one source contributes to the probabilistic hazard, then  $r_s = r$ , which implies that large  $\lambda$  is associated with large  $\varepsilon$ , and small  $\lambda$  is associated with small  $\varepsilon$  for a given PE. That is,  $\varepsilon$  increases with  $\lambda$  for a fixed probability of ground-motion exceedance. As *r* decreases, the  $\varepsilon_0$  associated with each source increases.

When many sources contribute to seismic hazard, SA<sub>0</sub> and  $\varepsilon$  values are affected by the source-site R distribution, by source M distribution, and by recurrence-time distributions. In this case, quantitative generalizations about the relation of  $\bar{\varepsilon}_0$  and  $\hat{\varepsilon}_0$  to source-recurrence times are more difficult. Qualitatively, we can say that shorter distances, larger magnitudes, and shorter recurrence times tend to increase  $\overline{\varepsilon}_0$ and  $\hat{\varepsilon}_0$  for a given PE. Thus, geographically, we expect to observe higher  $\varepsilon$  associated with a given PE at locations where larger, closer, and/or more frequent earthquakes are encountered in the PSHA. The geographic variation of  $\overline{\epsilon}_0$  and  $\hat{\varepsilon}_0$  in the United States computed in this article may be understood by thinking about the geographic variation of these source factors, along with attenuation-model influences. For example,  $\hat{\varepsilon}_0$  for sites close to the surface trace of the San Andreas fault (SAF) is determined by characteristic events (large magnitude, close distance, short recurrence times) on that fault, and is larger than  $\hat{\varepsilon}_0$  for sites further from the SAF, where the influence of longer recurrence time sources increases.

 $\hat{\varepsilon}_0$  may be used to calculate the mean annual rate of occurrence of the modal event (or modal-event set) from a

deaggregation analysis. If the ensemble of seismic sources yields an annual rate of exceedance r, and if the fraction of exceedances contributed by the modal bin is f, then the mean annual rate of occurrence of the modal event,  $\hat{\lambda}$ , is

$$\hat{\lambda} \approx fr / \int_{\hat{\varepsilon}_0}^{\infty} n(\varepsilon) d\varepsilon.$$
(6)

Here, the integral is the upper tail of the standard normal distribution, with lower integration limit  $\hat{\varepsilon}_0$  The equality is approximate when at least two ground-motion prediction models are used in the PSHA, and is exact when one ground-motion prediction equation is used, for example, when determining beta earthquake parameters (McGuire, 1995). Equation (6) is just equation (5) for binned data, solved for  $\lambda$ . For a given PE and a given modal-event *M*,*R* we see that relatively low SA<sub>0</sub>  $\Leftrightarrow$  low  $\hat{\varepsilon}_0 \Leftrightarrow$  long mean recurrence time of the modal event, that is, low  $\hat{\lambda}$ .

#### Modal Sources, Predominant Earthquakes, and $\varepsilon$

The distinction among source types can be important when we inquire about modal M,R, which by analogy to equation (4), are averages in the modal bin. Sources in the 1996 USGS PSHA (Franket et al., 1996) include characteristic earthquakes and smaller sources on faults with known strike, dip, and Quaternary slip. Other sources include randomly occurring earthquakes, often called background seismicity. Randomly occurring earthquakes include point sources for smaller magnitudes, and earthquakes on unidentified, and therefore random-strike, finite faults for M 6.5 and greater. USGS PSHA also considers areal source zones with a variety of source treatments. In a seismic-hazard deaggregation, binned sources may contain earthquakes on known faults and random earthquakes. At most sites in the WUS, the main exceptions being some Basin and Range sites, the annual exceedance rate contributed by mapped fault sources significantly exceeds the rate contributed by random seismicity in each M, R bin. Often there is no overlap of random-seismicity and known source or source-zone events. Non-zero hazard in an M 7.5-to-8 bin for a WUS site means only characteristic events on the SAF are contributing. Similarly, non-zero hazard in an M 7.5-to-8 bin for an eastern United States site implies that NMSZ characteristic earthquakes are the sole contributors, since random seismicity's  $M_{max}$  < 7.5 (Frankel *et. al.*, 1996). On the other hand, at many sites in the Basin and Range province, sources on faults having low Quaternary slip rates contribute less to the modal (M,R) bin than does random seismicity. In any given bin,  $\varepsilon_0$  for individual sources having larger magnitudes and shorter distances is lower than  $\varepsilon_0$  for smaller or more distant sources. Just as there is a geographically varying  $\hat{\mathcal{E}}_0$  corresponding to a fixed PE, at many United States sites there can be considerable variation in  $\ensuremath{\epsilon_{0A}}$  for individual sources and attenuation models contributing to  $\hat{\varepsilon}_0$  for that PE when using bin dimensions  $\Delta M = 0.5$  and so on.

In western California, where the spatial density of Quaternary faults is relatively high, it is sometimes difficult to associate the dominant hazard with any one fault system. For example, in Santa Barbara, California, each of four fault systems contributes more than 20% of the exceedances to the 10% in 50 yr SA<sub>0</sub> according to the California Division of Mines and Geology (CDMG) and USGS source models. For lower probabilities, contributions from the Mission Ridge–Arroyo Parida–Santa Ana fault system (FSI) tend to dominate North Channel Slope and other fault sources. This illustrates the general principle that as probabilistic ground motion increases, the relative contributions from the nearest sources increase. For the Santa Barbara site, with coordinates 34.423° N, 119.703° W,  $\varepsilon_0$  (averaged over attenuation models) from FSI sources for the 5-Hz  $SA_0$  is less than -0.1when considering the 10% in 50 yr PE, and is about 0.7 when considering the 2% in 50 yr PE.

In much of California, characteristic events on one or two fault systems often strongly dominate the hazard for the 10% PE in 50 yr probability. These predominant earthquakes can sometimes be determined by reducing deaggregation  $\Delta$ M and  $\Delta$ R to small values. Maps that show these predominant earthquakes are available from CDMG for several urban areas in California, for example the Hollywood quadrangle (CDMG, 1998). Rarely, however, do we find PSHA reports that focus equal attention on  $\varepsilon$  as on magnitude and distance, even though the question naturally arises, how does SA<sub>0</sub> differ from the median motion for the predominant, modal, or other distinguishing earthquake?

Some sources may be dominant under certain deaggregation rules, whereas others may be dominant under other rules. We have already seen examples of this, resulting from using different attenuation models. Individual source  $M,R,\varepsilon$ triples are briefly discussed in this article, sometimes for sources that do not happen to occupy the modal bin. It would seem prudent for decision makers to inform themselves about several hazardous sources, not just about the source that a PSHA analysis determines to be the most hazardous, for example,  $(\hat{M}, \hat{R}, \hat{\varepsilon}_0)$  or  $(M^*, R^*, \varepsilon^*)$ , given a somewhat arbitrary set of binning rules. By analogy, the medical community strives to immunize the population against many known diseases, not just the most likely one or few.

#### $\overline{\varepsilon}_0$ and $\hat{\varepsilon}_0$ for CEUS Seismic Hazard

Figure 2 is a set of contour maps of SA<sub>0</sub> and mean epsilon,  $\overline{\epsilon_0}$ , for the 2% PE in 50 yr case, for the CEUS. Here, CEUS just means the United States with longitude  $\leq 100^{\circ}$  W. The maps are for 1-, 0.3-, and 0.2-sec SA<sub>0</sub> and for peak ground acceleration (PGA<sub>0</sub>). SA<sub>0</sub> is discussed in the online documentation for the National Seismic Hazard Maps, and in Frankel *et al.* (1996). Historically active, large-magnitude sources in the CEUS include the **M** 8 NMSZ and **M** 7.3 Charleston, South Carolina, earthquakes. Recurrences of characteristic events in NMSZ and **M** 7.3 Charleston, South Caolina, dominate hazard in much of the CEUS for the 2% PE in 50 yr ground motion (Harmsen *et al.*, 1999). Figure 2 indicates that  $\overline{\varepsilon}_0$ , for the 2% PE in 50 yr motions, is between 0 and 1 almost everywhere in the CEUS. For the 2% in 50 yr exceedance probability,  $\overline{\varepsilon}_0$  has a spatial median of 0.40 for PGA, and  $\overline{\varepsilon}_0$  has a spatial median of about 0.76 for 1.0sec SA, in the CEUS. The spatial median is a summary statistic: for a randomly chosen location in the CEUS, the probability is 0.5 that  $\overline{\varepsilon}_0$  for that site will be greater than (less than) the spatial median of  $\overline{\varepsilon}_0$ . Other spatial medians are given in Table 1 for CEUS and for western United States epsilons.

 $\hat{\mathcal{E}}_0$  exhibits a greater range and variance in the CEUS than does  $\overline{\mathcal{E}}_0$ .  $\hat{\mathcal{E}}_0$  corresponding to short-period SA exhibits greater spatial variance than  $\hat{\mathcal{E}}_0$  for intermediate-period SA. Figure 3 shows  $\hat{M}$  on the left and  $\hat{\mathcal{E}}_0$  on the right for 1-, 0.3-, and 0.2-sec SA, and PGA associated with 2% PE in 50 yr.  $\Delta R = 25$  km in Figure 3. CEUS modal magnitude and distance are discussed further in Harmsen *et al.* (1999). Table 1 exhibits spatial medians of  $\hat{\mathcal{E}}_0$  for CEUS sites. The spatial medians of  $\hat{\mathcal{E}}_0$  are 0.2–0.3 less than those for  $\overline{\mathcal{E}}_0$ . For more than half of the CEUS, if an event having the modal M,R (for the 2% in 50 yr PE) were to occur, SA<sub>0</sub> would be exceeded with greater likelihood than if an event with the mean M,R, were to occur. For the CEUS sites of Figures 2 and 3, the spatial standard deviation of  $\hat{\mathcal{E}}_0$  is about twice that of  $\overline{\mathcal{E}}_0$ .

Figure 3 indicates that  $\hat{\varepsilon}_0$  is generally greater than 0.5 for the 1-sec SA in the CEUS and is generally greater than 0 for shorter-period SA or PGA, for the 2% in 50 yr PE.  $\hat{\varepsilon}_0 =$ 0 is a useful reference point because it represents median ground motion for the modal source. Locations where  $\hat{\varepsilon}_0$  for 0.2-sec SA and PGA generally occur in low-hazard regions like northern Minnesota.  $\hat{\varepsilon}_0 = 0$  is also found in a relatively high-hazard region, coastal South Carolina. This  $\hat{\varepsilon}_0$  distribution appears anomalous, as the mean recurrence time for an M 7.3 characteristic earthquake, which dominates the hazard, is 650 yr (Frankel et al., 1996). The result may be understood by remembering that the M 7.3 host fault location is unknown. To model source-location uncertainty, sources are distributed uniformly over a large area, such that the ensemble's average interevent time is 650 yr (Frankel et al., 1996). For M 7.3 sources confined to any one distance annulus, however, the mean recurrence time is much greater than 650 yr. Areally spreading the M 7.3 hazard may decrease SA<sub>0</sub> and  $\hat{\varepsilon}_0$  at many interior sites and at many sites near but exterior to this areal region, versus concentrating the hazard at a single fault at some preferred location in the region. Using  $\Delta R = 10$  km also decreases  $\hat{\varepsilon}_0$  from that shown in Figure 3 for sites in the areal source zone. Contributions from M 7.3 (and all other) sources in more distant annuli keep  $\overline{\varepsilon}_0 > 0$  in Figure 2, however.

For all frequencies of SA shown in Figure 3, geographic regions in the CEUS with  $1 < \hat{\varepsilon}_0 < 1.5$  are not uncommon. These are often transition areas where mean annual rate of exceedance of SA<sub>0</sub> from the distant NMSZ **M** 8 or Charleston **M** 7.3 source, while still dominant, is comparable to that from lower-magnitude or less-likely local and regional seis-



Figure 2. Left side, SA<sub>0</sub> and PGA<sub>0</sub> for PE = 2% in 50 yr, for sites in the CEUS. Values from National Seismic Hazards Maps (Frankel *et al.*, 1996). Right side, mean  $\varepsilon(\bar{\varepsilon}_0)$  for sites in the CEUS, corresponding to these ground motions. SA frequency increases down the page, 1.0-sec SA top, then 0.3-sec SA, 0.2-sec SA, then PGA.

Table 1Spatial Median Values of  $\bar{\epsilon}_0$  and  $\hat{\epsilon}_0$  for USGS Seismic-Hazard Sources for SA with 2% in 50 yrPE For CEUS,  $\Delta R = 25$  km, but for WUS,  $\Delta R = 10$  km.

Frequency (Hz) Statistic	1.0		3.33		5.0		PGA	
	$\bar{\varepsilon}_0$	$\hat{\varepsilon}_0$	$\bar{\varepsilon}_0$	$\hat{\varepsilon}_0$	$\bar{\varepsilon}_0$	$\hat{\varepsilon}_0$	$\bar{\varepsilon}_0$	$\hat{\varepsilon}_0$
CEUS	0.76	0.60	0.65	0.40	0.59	0.30	0.40	0.12
WUS	0.84	0.59	0.68	0.45	0.62	0.46	0.43	0.31

micity. In such regions,  $\overline{\epsilon}_0 < 1$ . Contributions from several sources increase the 2% in 50 yr PE motions. Crustal attenuation raises  $\hat{\epsilon}_0$  for these distant sources, while  $\overline{\epsilon}_0$  averages over distant and nearby sources, and is therefore less than  $\hat{\epsilon}_0$  in these regions.

#### $\bar{\boldsymbol{\varepsilon}}$ and $\hat{\boldsymbol{\varepsilon}}_0$ for Western United States Seismic Hazard

Figure 4 is a set of contour maps of SA<sub>0</sub> and  $\overline{\varepsilon}_0$  for the 2475-yr return accelerations, for 1-, 0.3-, and 0.2-sec SA and for PGA<sub>0</sub>, for the WUS. We see that there is far more geographic variation of  $\overline{\varepsilon}$  in the WUS than in the CEUS. This is because in the WUS, there are many active faults with mean source-recurrence times, *T*, on the order of a few hundred years, and many active faults with *T* on the order of 10,000 yr. As  $\lambda$  in equation (2) is the reciprocal of *T*, the conditional probability factors in equation (2) may be quite low (upper tail of ground-motion uncertainty distribution) for sites near sources with short (few hundred years) *T* when compared with those factors for sites near the dominant long-recurrence time sources, for a given *r*.

The data of Figure 4 indicate that  $\overline{\varepsilon}_0$  at some locations in the WUS is significantly less than 0 for the 2% PE in 50 yr ground motions. In other words, the 2% in 50 yr ground motions are less than the predicted medians for many sources at those sites. Negative  $\varepsilon$  occurs where dominating sources are associated with faults that have mean recurrence times greater than 5000 yr, quite typically in the Basin and Range. Few such faults have known historical activity. Wellknown examples in eastern California are Owens Valley (1872) and Landers (1992). The 16 October 1999, M 7.1 Hector Mine earthquake in southern California ruptured a fault that had previously exhibited no surface displacement in the Holocene (Scientists from the USGS, SCEC, and CDMG, 2000). Well-known examples in Nevada include Pleasant Valley (1915), Dixie Valley (1954) and Fairview Peak (1954) (Rogers *et al.*, 1991). Spatial medians of  $\overline{\varepsilon}_0$  in the WUS are quite similar to those for the CEUS (see Table 1).

WUS  $\hat{M}$  is contoured in Figure 5 on the left side, and modal epsilon,  $\hat{\varepsilon}_0$  is contoured on the right side, for the 2% in 50 yr PE. We attempt to separate sources on closely spaced faults into different bins using smaller  $\Delta R$  (10 km) in the WUS regional and site-specific seismic-hazard deaggregations. Modal magnitude and distance for the WUS are discussed further in Harmsen and Frankel (2000). Table 1 exhibits spatial medians of  $\hat{\varepsilon}_0$  for WUS sites. For the sites of Figures 4 and 5, the spatial standard deviation of  $\hat{\varepsilon}_0$  is about 2.5 times that of  $\bar{\varepsilon}_0$ .

 $\hat{\varepsilon}_0 > 1$  is frequently encountered in WUS deaggregations of the 2% in 50 yr (and less probable) ground-motion exceedances.  $\hat{\varepsilon}_0 > 1.5$  is common in western California, and  $\hat{\varepsilon}_0 > 2$  occurs at locations near the Salton Sea, mostly associated with Brawley seismic-zone sources. In the WUS, high  $\hat{\varepsilon}_0$  corresponds to short *T*, on the order of 1/5 to 1/10 the ground-motion return time, 2475 yr. Whereas sources with short recurrence times are relatively common in tectonically active western North America, they are absent from the CEUS, at least in the 1996 PSHA maps. Relatively high  $\hat{\varepsilon}_0$  at sites in the CEUS is a consequence of many distant and local sources affecting the hazard additively.

Modal sources with long recurrence times, on the order of five times the ground-motion return time, are also frequently encountered in many western states. Sites near these sources have low  $\hat{\varepsilon}_0$  Modal values like -0.5 are common, and sites with  $\hat{\varepsilon}_0 < -1$  can be found for the 2% in 50 yr motions. The most negative  $\hat{\varepsilon}_0$  in Figure 5 occurs near the Cheraw fault of eastern Colorado. Significant earthquakes on the Cheraw fault have a very long recurrence time; however, unlike the eastern California seismic zone, there is no high regional slip rate in eastern Colorado. Thus, neither gridded (random) seismicity nor other faults have much influence on SA<sub>0</sub> in the vicinity of the Cheraw fault.

Negative  $\varepsilon$  is sometimes associated with sources having earthquake recurrence times less than five times the probabilistic ground-motion return time. For a site at 116.6° W, 34.6° N, in the immediate vicinity of the Landers, California, fault system, which ruptured with an **M** 7.3 earthquake on 28 June 1992,  $\hat{M} \approx 6.78$  and  $\hat{\varepsilon}_0 \approx 0.5$  for the 1-sec SA. The second-largest hazard bin, with **M** from 7 to 7.5 and R < 5km, has nearly the same hazard (bimodal distribution) at most SA periods, and a major contributor to that bin is an **M** 7.3 characteristic earthquake, like the 1992 Landers mainshock, whose 1-sec SA  $\varepsilon_0 = -0.47$ . The mean recurrence time for characteristic events on the Landers fault is 5000 yr (Petersen *et al.*, 1996).

For this eastern California site, estimates of modal parameters are sensitive to the choice of  $\Delta R$ . The modal source has a higher magnitude with a significantly lower  $\varepsilon$  when  $\Delta R$  is reduced from 10 to 5 km. Table 2 shows the modal



Figure 3. Left side, Modal magnitude  $(\hat{M})$  for PE = 2% in 50 yr, for sites in the CEUS, from Harmsen *et al.* (1999). Right side, modal-event  $\varepsilon(\hat{\varepsilon}_0)$  for sites in the CEUS. Computations use  $\Delta R = 25$  km and  $\Delta M = 0.5$ . Map arrangement as in Figure 2.





Figure 4. Contour maps of SA<sub>0</sub> (Frankel *et al.*, 1996) and mean  $\varepsilon(\overline{\varepsilon}_0)$  arranged like those of Figure 2, for the WUS and PE = 2% in 50 yr.



Figure 5. Contour maps of modal M (from Harmsen and Frankel, 2001) and modalevent  $\hat{\varepsilon}_0$ , arranged like those of Figure 3, for the WUS and PE = 2% in 50 yr. Modal sources determined using  $\Delta R = 10$  km.

data when using these two  $\Delta R$  values. In both cases,  $\Delta M = 0.5$ .

We next illustrate how magnitude-bin interval may affect  $\hat{\varepsilon}$ . Here and in web-site deaggregations, we use  $\Delta M =$ 0.5. For a site in the northern Owens Valley, 36.51° N, 118.11° W, the 1-sec SA  $\hat{M} \cong 6.7$  and  $\hat{\varepsilon}_0 \cong 0.25$ ,  $\overline{\varepsilon}_0 = 0.55$ . For this location near the Independence fault (R < 1 km, mean recurrence time 5500 yr) (Petersen et al., 1996), the modal bin includes effects from a characteristic event on that fault.  $\hat{\varepsilon}_0$  is greater than 0, but the biggest contributor to the modal bin, an M 6.9 event on the Independence fault, has  $\varepsilon_0 < -0.3$ . An M 7.6 characteristic event on the Owens Valley fault (R = 7.5 km, recurrence time 4000 yr, similar to the March 1872 earthquake) (Petersen et al., 1996) is another significant PSHA source at this site, with 1-sec SA  $\varepsilon_0 = -0.24$ . Reduction of  $\Delta M$  to 0.2 or 0.1 makes one of these negative- $\varepsilon_0$  sources modal. Regardless of binning details,  $\varepsilon_0$  for important individual sources can be significantly less than either  $\overline{\varepsilon}_0$  or  $\hat{\varepsilon}_0$ .

The subduction of the Juan de Fuca plate dominates seismic hazard in much of the coastal Pacific Northwest. From Figure 5, we see that for many coastal sites in the Pacific Northwest, the primary (modal) source of seismic hazard is an **M** 8.3 or **M** 9 earthquake. **M** 8.3 and **M** 9 are the two subduction-source magnitudes considered in the 1996 hazard maps. These two Cascadia sources contribute to high  $\hat{\varepsilon}_0$  in a transition zone, that is, the zone where the two subduction sources contribute comparable groundmotion exceedances.

Although our knowledge of Quaternary fault locations in the Pacific Northwest is limited, at sites near known faults, such as the Seattle fault and the South Whidbey Island fault, the 2% in 50 yr SA exceedances are often dominated by motion on those faults (Harmsen and Frankel, 2000). For the Seattle fault, both characteristic and smaller-magnitude sources are considered in the 1996 PSHA. The characteristic event magnitude is 7.1, and the 2% in 50 yr 1-sec SA<sub>0</sub>, which is 0.524g, is lower than the median motion from this **M** 7.1 event for a site 1.6 km from the fault. For this site/source  $\varepsilon_0 = -0.32$ . The mean  $\varepsilon_0$  for lower-magnitude Seattle fault sources for this same 1-sec SA<sub>0</sub> is 0.11. In other words, the 2% in 50 yr probabilistic motion is approximately equal to the median motion from the modal source for many sites in downtown Seattle.

#### Which Mode Should I Use?

The modal-event  $\hat{M}$ ,  $\hat{R}$  can be different from, M\*,R\* and even if equal,  $\varepsilon^*$  can exceed  $\hat{\varepsilon}_0$ . Figure 6 is the USGS website deaggregation of 0.2-sec SA at a site near a potential national nuclear-waste repository at Yucca Mountain, Nevada. The deaggregation analysis for the 2% in 50 yr PE at this site yields  $(\hat{M}, \hat{R}, \hat{\varepsilon}_0) = (6.25, 7.3 \text{ km}, 0.1)$  and  $(M^*, R^*, \varepsilon^*) = (6.81, 32.8 \text{ km}, \varepsilon > 2)$ . The hatted triple corresponds to random local seismicity, whereas the asterisk triple corresponds to earthquakes on the Death Valley/Fur-

 Table 2

 One-Second SA for Site at 34.6° N, 116.6° W. %C is percent contribution to 2% in 50-yr hazard.

$\Delta R$ (km)	Ŕ	Ŵ	$\hat{\varepsilon}_0$	%C	
5	1.3	7.24	-0.40	28.0	
10	4.2	6.67	0.51	28.9	

nace Creek fault system. Adoption of one or the other of these (*M*,*R*) pairs as the scenario earthquake for 0.2-sec response could result in different design considerations. These distinct sources comprise the largest and second-largest hazard bins, with 17.2% and 13.5% contributions to the 0.2-sec SA exceedances, respectively. Altering the bin widths and/ or centers could change these estimates. One strategy for design that appears quite sensible is to include both of these scenario earthquakes.  $\hat{M}$ ,  $\hat{R}$  and M\*,R\* tend to be equal over a broad range of response-spectral periods at this site near Yucca Mountain. For PGA and 0.1-sec SA, local sources determine the two modes, whereas for 0.3- to 2.0-sec period SA, Death Valley sources determine the two modes. All example deaggregations are for the 2% in 50 yr PE.

The fact that  $(M^*, R^*, \varepsilon^*)$  tends toward the mode of the three-dimensional conditional p.d.f. of  $M, R, \varepsilon$  as the bin widths tend to 0 (Bazzurro and Cornell, 1999) does not provide much more than academic insight. In practice, USGS deaggregations always use substantial bin widths, and these are often required if we wish to associate significant probability with a binned event. When comparing the influence of random seismicity hazard with fault hazard, we may use R bins several hundred kilometers in width (Harmsen and Frankel, 2001). The Nisqually, Washington, earthquake of 28 February 2001 is an example of random seismicity hazard (no mapped fault is associated with a 52-km deep event) that yields a significant percentage of PSHA ground-motion exceedances for many sites in the Puget Sound area, especially at the 10% in 50 yr PE. We believe deep intraplate sources should be considered in seismic resistant design decisions for structures in parts of the Pacific Northwest. Large bin dimensions are useful in PSHA deaggregations when aleatory uncertainty in M,R is large.

#### Summary, Conclusions, and Recommendations

This article attempts to supplement the information given in similar previous articles that focused on PSHA mean and modal magnitude and distance in the conterminous United States. (Harmsen *et al.*, 1999; Harmsen and Frankel, 2001). Here, two mean  $\varepsilon$  and two modal  $\varepsilon$  are defined. Maps of  $\overline{\varepsilon}_0$  and  $\hat{\varepsilon}_0$  are presented for the 2475-yr-return ground motions.  $\varepsilon_0$  is the quantile that relates SA<sub>0</sub> to the median ground motion from significant seismic sources. The geographic variation of  $\overline{\varepsilon}_0$  and  $\hat{\varepsilon}_0$  for a given frequency of ground motion, and for a specified probability of exceedance, is large.

Negative  $\hat{\varepsilon}_0$  can be an indicator that a site is near a



Figure 6. Probabilistic seismic-hazard deaggregation for a site near Yucca Mountain, Nevada, for the 0.2-sec SA and the 2% PE in 50 yr. The front face of each column is colored to indicate the relative contribution to hazard by  $\varepsilon$ . Acceleration  $\ge 0.7293g$ ; annual exceedance rate  $= 0.401 \times 10^{-3}$ 

known source of potentially damaging earthquakes even though the 2% in 50 yr probabilistic ground motion is quite low. For the 2% in 50 yr SA<sub>0</sub>, regions and sites where  $\hat{\epsilon}_0 < 0$  are identified. Several locations in western states of the United States, principally Colorado, New Mexico, and Nevada, but also Texas, Washington, Oregon, and Arizona, have relatively low 2% PE in 50 yr SA<sub>0</sub> levels, yet they are near long recurrence-time active faults believed to be capable of **M** 7 + earthquakes. The Meers fault of southwest Oklahoma is an example of a similar CEUS seismic source.

The Landers, California, case discussed previously illustrates that to exclusively focus attention on the deaggregation's modal-event M,R pair may result in overlooking significant earthquake sources near the site. Seismologists often advise engineers to consider the contents of nonmodal but significantly contributing hazard bins (e.g., Cramer and Petersen, 1996). Geographic deaggregations to determine all faults with significant contributions (Bazzurro and Cornell, 1999) often emphasize contributions from such faults.

In southeastern California, because of the spatial density of mapped Quaternary faults, and the high rate of random or gridded seismicity, PSHA SA<sub>0</sub> values are much higher than in many other states, where long recurrence-time faults exist in relative isolation and where random seismicity hazard is of limited importance. The Cheraw fault of southeastern Colorado and the Meers fault of southwestern Oklahoma are examples of long-recurrence-time faults having  $\overline{\epsilon}_0 < 0$  and  $\hat{\epsilon}_0 < 0$ . The list of such sources grows in many states, as a result of advances in paleoseismology and geophysical disciplines that study low-strain-rate phenomena.

At sites near several active, short-recurrence-time faults in the WUS, the 2% PE in 50 yr  $SA_0$  is often a higher ground motion than that which current building codes consider for earthquake-resistant design (Leyendecker et al., 2000). Often, these are sites where  $\hat{\varepsilon}_0 > 2$ , or where SA<sub>0</sub> is two standard deviations above the median for the predominant source. Geographically, these are sites near some of the large western California faults, such as the San Andreas, Maacama, Hayward, San Jacinto, Brawley, Imperial, and others. Some cities containing locations with  $1.9 < \overline{\varepsilon}_0 < 2$  for the 2475-yr probabilistic SA<sub>0</sub> include San Bernardino, San Jose, and Sacramento, California, and Anchorage, Alaska.  $\hat{\varepsilon}_0$  and  $\overline{\varepsilon}_0$  both exceed 2.1 for a range of SA periods at Brawley, California. Because  $\hat{\varepsilon}_0 > 2$  implies a ground motion having less than 0.023 probability of being exceeded if the modal event occurs, the code decision made in western California to use as a design ground motion a lower value than the 2% in 50 yr  $SA_0$  is currently deemed to be adequate.

Indeed, the 1997 NEHRP Provisions (BSSC, 1998) and the International Building Code (IBC) (ICC, 2000) define a maximum considered earthquake (MCE) ground motion for sites near several large, historically active, short-recurrencetime faults. These MCE values are  $1.5 \times$  median motion for a deterministic earthquake for sites sufficiently near the active faults (Leyendecker *et al.*, 2000).

The developers of these building codes did not explicitly consider  $\varepsilon$  for significant probabilistic sources when deciding how to determine MCE values. However, they did believe that the 1.5 margin-of-safety factor also represented approximately  $1\sigma$  above the median (Leyendecker *et al.*, 2000) for a deterministic event. A more quantitative estimate of the distance of near-fault MCE values to median motion may be derived from the strong-ground-motion regression equations. For example, the 1-sec SA attenuation equations of Boore *et al.* (1997) tabulate  $\sigma_{\ln Y} = 0.61$ . This  $\sigma$  corresponds to  $\varepsilon \approx 0.66$ . The equation for determining  $\varepsilon$  is

#### $(1.5 \times \text{median motion from the earthquake}) = e^{\mu + \epsilon \sigma}$ (7)

with  $\sigma = 0.61$ . The probability that sa exceeds  $\mu + 0.66\sigma$  given the occurrence of the deterministic event is 0.25. Recognizing that ground motions near certain faults can be quite high, developers of the 1997 Uniform Building Code (UBC) (ICBO, 1997) incorporate a near-source factor to increase the design motion (Petersen *et al.*, 2000).

Public policy seeks an appropriate level of protection given the seismic hazard. A deterministic definition of protection is seismic design that should resist  $\mu + \varepsilon \sigma$ , for a specified  $\varepsilon$ , for the sources which contribute most to the hazard. The  $\varepsilon$ -choice involves a balance among economics, engineering experience, safety, and uncertainty. In the previous paragraph, it was shown that near some major faults, including several in southern California, IBC-2000 code in effect requires engineers to design structures for elastic response for sa  $\leq \mu$  and noncollapse for sa  $< \mu + \varepsilon \sigma$ , with  $\varepsilon \approx 2/3$  for the 1.0-sec ordinate of the design spectrum.

The probabilistic definition, in contrast, says protect against SA having a specified probability of exceedance. IBC-2000 adopts a probabilistically determined SA for its design value when the PSHA 2% in 50 yr level is less than deterministic-event median, for example, for sites near many long-recurrence-time faults of Nevada, Colorado, Texas, and many other states (Leyendecker *et al.*, 2000). These areas include the blue zones for  $\hat{\varepsilon}_0$  in Figures 3 and 5. This choice shows that IBC-2000 code considers that earthquakeresistant design to SA less than two-thirds the median predicted ground motion from the infrequent earthquake on such faults represents acceptable risk for structures near these faults, because the probability of damage is small.

Whitman (1995) proposes that a long-range goal (for the year 2077) of NEHRP might be the protection of 99.9% of buildings against life-threatening damage from any earthquake with a mean recurrence interval of 2000 yr. We determine the 2% in 50 yr  $\varepsilon_0$  for such an earthquake. First, we assume that other sources make a negligible contribution to the seismic hazard for buildings near this 2000-yr source. The 2% in 50 yr *r* is 0.000404, and Whitman's 2000-yr source has  $\lambda = 1/2000$ . Thus, from equation (5),  $\varepsilon_0 =$  -0.8705. In this case, IBC-2000 code's requirement to protect to two-thirds of this SA<sub>0</sub> would probably need to be strengthened to meet Whitman's proposed goal. Seismic provisions of this building code imply that the MCE or collapse ground motion is expected to be exceeded with probability 0.8 if this 2000-yr source occurs. Faults with, say, four to six Holocene events are rare in the 1996 USGS fault database (Frankel *et al.*, 1996) The Owyhee River fault system in southeast Oregon has a several-thousand-year recurrence time. For sites about 10 km from this fault system, the 2% in 50 yr PGA<sub>0</sub> can be less than 0.2g, and  $\hat{\varepsilon}_0$  can be less than -1.

Second, we consider sites where Whitman's long-term NEHRP proposal is perhaps already met by the IBC-2000 seismic provisions. Background seismicity can elevate the 2% in 50 yr SA<sub>0</sub> at the site near the 2000-yr fault so that  $\varepsilon_0$ can be greater than zero. In the 1996 USGS maps,  $PGA_0 >$ 0.5g near the long-recurrence time Mission fault in Montana, located near 114° W, 48° N. The mean rate for characteristic events is 2.44  $\times$  10<sup>-4</sup> although the rate associated with smaller Mission fault events is higher. For a site 2 km from this fault, whose characteristic event M is 7.4,  $\varepsilon_0 = 0.26$  for PGA<sub>0</sub>, whereas for a site 8 km from this fault,  $\varepsilon_0 = 0.84$ , using the attenuation function of Boore et al. (1997). Thus, the combined effect of fault hazard and random seismicity hazard at some sites in the United States yields 2% in 50 yr ground motion that if protected against, would also protect against median ground motion from the nearby characteristic earthquake whose recurrence time is several thousand years. However, in this instance, source-directivity effects and longer ground-shaking duration of the M 7.4 earthquake are likely to be separate design issues for many structures.

This article does not advocate a specific preference for probabilistic or deterministic approaches in seismic-resistant design—both suffer from limited but growing understanding of where the active faults are, how active they are, and what the pertinent features of the next challenging rupture are. This article suggests that the triple  $(\hat{M}, \hat{R}, \hat{\varepsilon}_0)$  does approximately recover the probabilistic ground motion, for example, the 2% in 50 yr motion. Therefore,  $(\hat{M}, \hat{R}, \hat{\varepsilon}_0)$  is useful for facilitating discussion on the appropriateness of SA<sub>0</sub> as a design value. By considering  $\varepsilon$  with other deaggregation parameters, principally M and R, some of the fog surrounding PSHA ground motions (Allen, 1995) is lifted.

#### A New Product Available on the Web

An examination of deaggregated M, R, and  $\varepsilon$  data at the site of interest should help engineers to understand the relation between the ground-motion level they are protecting against, and the probabilistic source *M*,*R*, and  $\varepsilon$  distribution associated with that ground-motion level. The USGS's National Seismic Hazards Mapping Project web site, *http:// geohazards.cr.usgs.gov/eq/*, now includes an interactive seismic-hazard deaggregation page. This page may be accessed by clicking the Interactive Deaggregations menu item under Seismic Hazard. The user can select one of a large range of probabilities (from 20% PE in 50 yr to 1% PE in 50 yr) for any location in the conterminous United States and Alaska, for SA periods in the 0.1- to 2.0-sec range, or for PGA. The request generates a deaggregation hazard analysis and produces a graph of deaggregated magnitude, distance, and  $\varepsilon$ . The analysis displays the site's mean and modal distance, magnitude, and  $\varepsilon$ . The first mode is for binned (*M*,*R*) and the second is for binned (*M*,*R*, $\varepsilon$ ). Site conditions are everywhere assumed to be firm rock—760 m/sec average shear-wave velocity in the top 30 m. This web page includes the option to generate a plot of geographically deaggregated seismic hazard as discussed by Harmsen and Frankel (2001), similar to that of Bazzurro and Cornell (1999).

#### Acknowledgments

Discussions with A. Frankel, R. LaForge, R. Wheeler, and others were helpful. E. V. Leyendecker suggested the development of an interactive seismic-hazard deaggregation web page. Technical reviews by P. Bazzurro, D. Perkins, M. Petersen, and C. P. Tsai are gratefully acknowledged. Wessel and Smith's GMT routines were used for Figures 2–5 and in much of our PSHA web site. Opinions expressed here are the author's and are not necessarily shared by other project members or by USGS management.

#### References

- Abrahamson, N. A., and K. M. Shedlock (1997). Overview, *Seism. Res. Lett.* 68, 9–23.
- Allen, C. R. (1995). Earthquake hazard assessment: has our approach been modified in the light of recent earthquakes? *Earthquake Spectra* 11, 357–366.
- Bazzurro, P., and C. A. Cornell (1999). Disaggregation of seismic hazard, Bull. Seism. Soc. Am. 89, 501–520.
- Bender, B., and D. Perkins (1993). Treatment of parameter uncertainty and variability for a single seismic hazard map, *Earthquake Spectra* 9, 165–194.
- Boore, D., W. Joyner, and T. Fumal (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work, *Seism. Res. Lett.* **70**, 128–153.
- Building Seismic Safety Council (BSSC) (1998). 1997 Edition NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, Part 1 (Provisions) and Part 2 (Commentary), FEMA 302/303, Washington, D.C.
- California Division of Mines and Geology (CDMG) (1998). Seismic hazard evaluation of the Hollywood 7.5 minute quadrangle, Los Angeles County, California, Department of Conservation OFR 98–17, Section III.
- Campbell, K. (1997). Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra, *Seism. Res. Lett.* 68, 154–180.
- Chapman, M. C. (1995). A probabilistic approach to ground-motion selection for engineering design, *Bull. Seism. Soc. Am.* 85, 937–942.
- Cornell, C. A. (1968). Engineering seismic risk analysis, Bull Seism Soc. Am. 58, 1583–1606.
- Cramer, C., and M. Petersen (1996). Predominant seismic source distance and magnitude maps for Los Angeles, Orange, and Ventura counties, California, *Bull. Seism. Soc. Am.* 86, 1645–1649.

- Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E. Leyendecker, N. Dickman, S. Hanson, and M. Hopper (1996). National seismic-hazard maps: documentation, June 1996, U.S. Geol. Surv. Open-File Rept. 96–532, 110 pp.
- Hall, J. F., T. H. Heaton, M. W. Halling, and D. J. Wald, (1995). Nearsource ground motion and its effects on flexible buildings, *Earthquake Spectra* 11, 569–605.
- Harmsen, S., and A. Frankel (2001). Geographic deaggregation of seismic hazard in the United States. *Bull. Seism. Soc. Am.* **91**, 13–26.
- Harmsen, S., D. Perkins, and A. Frankel (1999). Deaggregation of probabilistic ground motions in the central and eastern United States, *Bull. Seism. Soc. Am.* 89, 1–13.
- International Code Council, Inc. (ICC) (2000). International Building Code, Building Officials and Code Administrators International, Inc., International Conference of Building Officials, and Southern Building Code Congress International, Inc., Birmingham, Alabama.
- International Conference of Building Officials (ICBO) (1997). Uniform Building Code, Whittier, California.
- Leyendecker, E. V., R. Hunt, A. Frankel, and K. Rukstales (2000). Development of maximum considered earthquake ground motion maps, *Earthquake Spectra* **16**, 21–40.
- McGuire, R. K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop, *Bull. Seism. Soc. Am.* 85, 1275–1284.
- Naeim, F. (1998). Research overview: Seismic response of structures, in Proc. of the NEHRP Conf. and Workshop on Research on the Northridge, California Earthquake of January 17, 1994, Vol. III-A (Engineering), 1–23.
- Petersen, M. D., W. Bryant, C. Cramer, T. Cao, M. Reichle, A. Frankel, J. Lienkaemper, P. McCrory, and D. Schwartz (1996). Probabilistic seismic hazard assessment for the state of California, *Calif. Dept. Conserv., Div. Mines Geol. OFR 96–08* and U.S.Geol. Surv. OFR 96– 706, Appendix A.
- Petersen, M. D., T. R. Toppozada, T. Cao, C. H. Cramer, M. S. Reichle, and W. A. Bryant (2000). Active fault near-source zones within and bordering the state of California for the 1997 Uniform Building Code, *Earthquake Spectra* 16, 69–83.
- Rogers, A. M., S. Harmsen, E. Corbett, K. Priestley, and D. dePolo (1991). The seismicity of Nevada and some adjacent parts of the Great Basin. Decade Map, in *Neotectonics of North America*, Slemmens, Engdahl, Zoback, and Blackwell (Editors), *Geological Society of America*, Boulder, Colorado, 153–184.
- Scientists from the United States Geological Survey (USGS), Southern California Earthquake Center (SCEC), and California Division of Mines and Geology (CDMG) (2000). Preliminary report on the 16 October 1999 M 7.1 Hector Mine, California, earthquake, Seism. Res. Lett. 71, 11–23.
- Senior Seismic Hazard Analysis Committee (SSHAC) (1997). Probabilistic Seismic Hazard Analysis: a Consensus Methodology, Senior Seismic Hazard Analysis Committee (R. Budnitz, G. Apostolakis, D. Boore, L. Cluff, K. Coppersmith, A. Cornell, P. Morris), U.S. Dept. Energy, U.S. Nuclear Reg. Comm., Elec. Power Res. Inst. NUREG/CR-6372, Vols. I and II.
- Somerville, P. G., N. F. Smith, R. W. Graves, and N. A. Abrahamson (1997). Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity, *Seism. Res. Lett.* 68, 199–222.
- Toro, G., N. Abrahamson, and J. Schneider (1997). Model of strong ground motions from earthquakes in central and eastern North America: Best estimates and uncertainties, *Seism. Res. Lett.* 68, 41–57.
- Wesson, R. L., A. D. Frankel, C. S. Mueller, and S. C. Harmsen (1999). Probabilistic seismic hazard maps of Alaska, U.S. Geol. Surv. Open-File Rept. 99–36, 43 pp.
- Whitman, R. V. (1995). Toward a strategic plan for the national earthquake risk reduction program, *Earthquake Spectra* **11**, 299–317.
- Youngs, R. R., S.-J. Chiou, W. J. Silva, and J. R. Humphrey (1997). Strong ground motion attenuation relationships for subduction zone earthquakes, *Seism. Res. Lett.* 68, 58–73.

#### Appendix: Caveats

- 1. Analysts of strong-motion data agree that the empirical SA distribution conditional on (M,R) is approximately lognormal; however, there is no published consensus about the maximum SA that is possible for a given (M,R). Establishing a realistic upper bound on ground motion becomes important when performing PSHA for very long return times such as 10<sup>5</sup> yr, and can affect the analysis for shorter return times as well. For sites near the Brawley seismic zone of southern California,  $\hat{\varepsilon}_0$  for the 2% in 50 yr PE can exceed 2. This means that the probability mass corresponding to motion more than  $2\sigma$  above the median is the only part of the probability mass function that the PSHA considers for those sources. Errors in its assumed shape (such as infinite upper tail) affect estimates of SA<sub>0</sub> at such sites. Frequently, seismologists recommend truncating the ground-motion uncertainty distribution at  $\mu_{\rm A} + 3\sigma_{\rm A}$ .
- 2. Most empirical-data-based SA models are published with specific domains where the analysts believe the models are valid. Boore et al. (1997, p. 146) state, "The equations are to be used for M 5.5–7.5 and d (R) no greater than 80 km." What then is done in PSHA for M,R pairs outside this domain of applicability? For example, USGS deaggregations indicate that an M 7.8 or greater earthquake on the SAF is the modal event at many sites in western California, some more than 100 km from the SAF. In practice, USGS PSHA simply extrapolates the attenuation function outside the recommended domain using the same carrier variable coefficients and same estimate of  $\sigma_A$ . Some attenuation models use large shallow subduction-zone earthquake records to define the large (M 7.8 and greater) event response (Campbell, 1997). In the 1996 USGS hazard model, the only M 7.8 and greater crustal earthquakes are strike-slip. Thus Campbell's implied extrapolation is not only between South American and North American subduction events, but between thrust and strike slip for these magnitudes as well. Estimates of the location of and uncertainty on both  $\mu$  and  $\sigma$ for the largest SAF events are a problem when considering the current strong-motion database.
- 3. There is no strong-motion data to confirm validity of  $\mathbf{M}$  9 attenuation models.  $\mathbf{M}$  9 is an important hazard for coastal sites in the Pacific Northwest (Harmsen and Frankel, 2000) and megathrust events with  $\mathbf{M} > 9$  are important for sites in southern Alaska (Wesson *et al.*, 1999). Youngs *et al.* (1997), whose model is used exclusively for  $\mathbf{M} \ge 9$  sources in USGS PSHA, do not indicate any

upper bound on  $\mathbf{M}$  when they specify the domain of applicability of their attenuation function.

- 4. In general, the list of probabilistic  $M, R, \varepsilon$  triples having no or very limited strong-motion data to support the attenuation model(s) used to compute SA<sub>0</sub>,  $\varepsilon_0$ , modal  $M, R, \varepsilon$ , and so on, is much larger than can be indicated here. Even where recorded strong motion data are used, the number of events used to estimate the median and spread parameter is often quite small. A fundamental statistical fact is "there is no law of small numbers." To apply published attenuation functions in PSHA or in deterministic analysis, we are frequently required to rely on regression-determined  $\mu$ ,  $\sigma$  values, which at M,R are based on data from one earthquake or none (extrapolation). On the other hand, to the extent that the attenuation model is a valid representation of the physics of earthquakes and crustal propagation, extrapolations may be justifiable. Much of the heated debate about the appropriateness of different attenuation models results from the need for answers where the event and similar-site strongmotion data sample is small. Applying a weighted sum of attenuation models in PSHA (e.g., equations 2 and 4) does not really change the nature of the debate, as the dominant model (least attenuation) largely determines the estimates engineers need (e.g., modal  $M, R, \varepsilon$ ).
- 5. The near-source forward directivity pulse, which has been observed to produce major damage, receives limited attention in attenuation models used in 1996 USGS PSHA calculations. For spectral response corresponding to long and intermediate oscillator periods (t > 1 sec), this pulse can be several times stronger than the median predicted ground motion. For some quantitative work on potential building damage from the directivity pulse of an M 7 thrust event in an urban area, see Hall et al. (1995). These pulses are known to accompany many major earthquakes, well-known examples being Landers, Northridge, and Kobe (Somerville et al., 1997). Somerville et al. (1997) propose an attenuation model that specifically accounts for the directivity pulse. Naeim (1998) states that further research is needed before a consensus can be reached among seismologists and structural engineers about the ability of modern design to successfully resist the nearsource pulse(s).

U.S. Geological Survey Denver, Colorado

Manuscript received 5 December 2000.