

An Introduction to Probabilistic Seismic Hazard Analysis (PSHA)

Jack W. Baker

Version 1.3

October 1st, 2008

Acknowledgements

This document is based in part on an excerpt from a report originally written for the US Nuclear Regulatory Commission. Financial support for the writing of that material is greatly appreciated.

Thanks to Marcello Bianchini, Tom Hanks, Ting Ling, Nirmal Jayaram and Seok Goo Song for pointing out errors and making suggestions to improve this document. Any remaining errors are the sole responsibility of the author.

Contents

Acknowledgements	2
Contents	3
Section 1 An overview of PSHA	5
1.1 Introduction.....	5
1.2 Deterministic versus probabilistic approaches.....	6
1.2.1 Variability in the design event.....	6
1.2.2 Variability of ground motion intensity for a given earthquake event.....	8
1.2.3 Can we use a deterministic approach, given these uncertainties?	9
1.3 Probabilistic seismic hazard analysis calculations.....	10
1.3.1 Identify earthquake sources	13
1.3.2 Identify earthquake magnitudes.....	13
1.3.3 Identify earthquake distances	17
1.3.4 Ground motion intensity	21
1.3.5 Combine all information.....	25
1.4 Example PSHA calculations	27
Section 2 Extensions of PSHA	37
2.1 Deaggregation.....	37
2.2 Impact of bounds on considered magnitudes and distances	43
2.3 Probabilities, rates, and return periods.....	44
2.4 Summarizing PSHA output: the uniform hazard spectrum	46
2.5 Joint distributions of two intensity measures.....	47
Section 3 Conclusions	52
Section 4 A review of probability	53
4.1 Random events.....	53
4.2 Conditional probabilities.....	55
4.3 Random variables	58
4.4 Expectations and moments	66
Section 5 Further study	68
5.1 Origins and development of PSHA.....	68
5.2 Books and summary papers	68
References	71

Section 1 An overview of PSHA

“The language of probability allows us to speak quantitatively about some situation which may be highly variable, but which does have some consistent average behavior. Our most precise description of nature must be in terms of probabilities.”

–Richard Feynman

1.1 Introduction

The goal of many earthquake engineering analyses is to ensure that a structure can withstand a given level of ground shaking while maintaining a desired level of performance. But what level of ground shaking should be used to perform this analysis? There is a great deal of uncertainty about the location, size, and resulting shaking intensity of future earthquakes. Probabilistic Seismic Hazard Analysis (PSHA) aims to quantify these uncertainties, and combine them to produce an explicit description of the distribution of future shaking that may occur at a site.

In order to assess risk to a structure from earthquake shaking, we must first determine the annual probability (or rate) of exceeding some level of earthquake ground shaking at a site, for a range of intensity levels. Information of this type could be summarized as shown in Figure 1.1, which shows that low levels of intensity are exceeded relatively often, while high intensities are rare. If one was willing to observe earthquake shaking at a site for thousands of years, it would be possible to obtain this entire curve experimentally. That is the approach often used for assessing flood risk, but for seismic risk this is not possible because we do not have enough observations to extrapolate to the low rates of interest. In addition, we have to consider uncertainties in the size, location, and resulting shaking intensity caused by an earthquake, unlike the case of floods where we typically only worry about the size of the flood event. Because of these challenges, our seismic hazard data must be obtained by mathematically combining models for the location and size of potential future earthquakes with predictions of the potential shaking intensity caused by these future earthquakes. The mathematical approach for performing this calculation is known as Probabilistic Seismic Hazard Analysis, or PSHA.

The purpose of this document is to discuss the calculations involved in PSHA, and the motivation for using this approach. Because many models and data sources are combined to create results like

those shown in Figure 1.1, the PSHA approach can seem opaque. But when examined more carefully, the approach is actually rather intuitive. Once understood and properly implemented, PSHA is flexible enough to accommodate a variety of users' needs, and quantitative so that it can incorporate all knowledge about seismic activity and resulting ground shaking at a site.

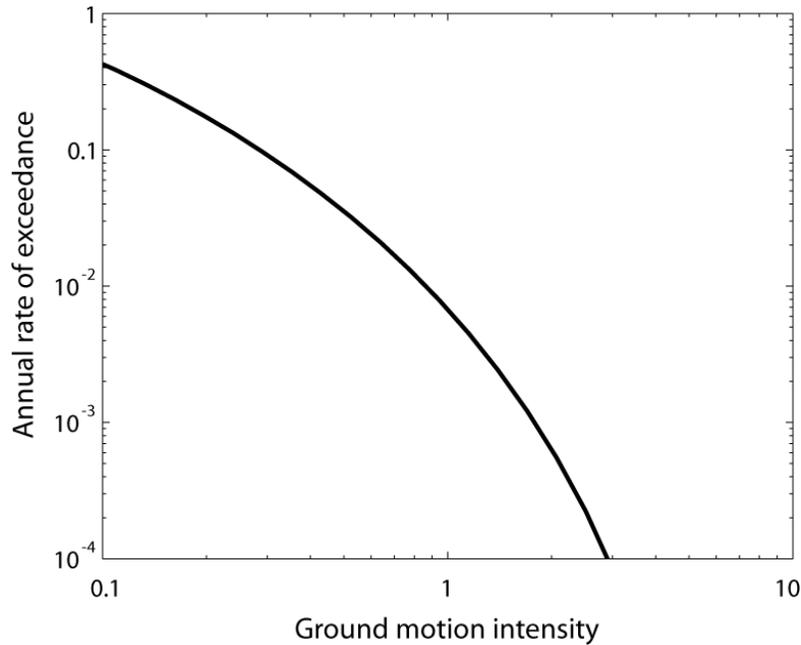


Figure 1.1: Quantification of the possibility of intense ground shaking at a site.

Probability calculations are a critical part of the procedures described here, so a basic knowledge of probability and its associated notation is required to study this topic. A review of the concepts and notation used in this document is provided for interested readers in Section 4.

1.2 Deterministic versus probabilistic approaches

The somewhat complicated probabilistic evaluation could be avoided if it was possible to identify a “worst-case” ground motion and evaluate the facility of interest under that ground motion. This line of thinking motivates an approach known as deterministic hazard analysis, but we will see that conceptual problems arise quickly and are difficult to overcome.

1.2.1 Variability in the design event

A designer looking to choose a worst-case ground motion would first want to look for the maximum magnitude event that could occur on the closest possible fault. This is simple to state in theory, but several difficulties arise in practice. Consider first the hypothetical site shown in Figure 1.2a, which is located 10 km from a fault capable of producing an earthquake with a maximum magnitude of 6.5. It

is also located 30 km from a fault capable of producing a magnitude 7.5 earthquake. The median predicted response spectra from those two events are shown in Figure 1.2b. As seen in that figure, the small-magnitude nearby event produces larger spectral acceleration amplitudes at short periods, but the larger-magnitude event produces larger amplitudes at long periods. So, while one could take the envelope of the two spectra, there is not a single “worst-case” event that produces the maximum spectral acceleration amplitudes at all periods.

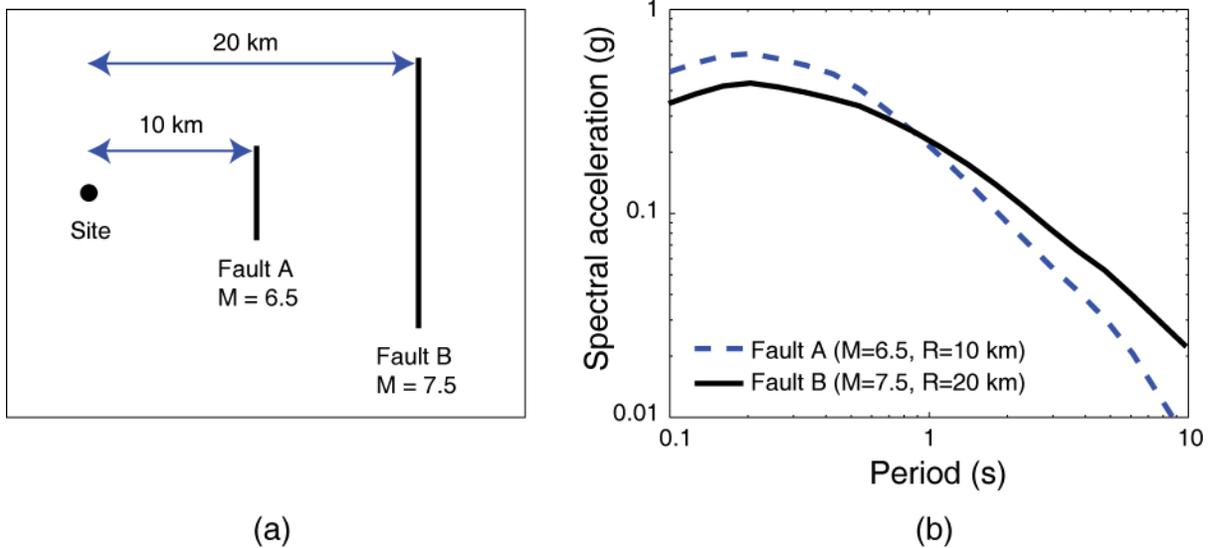


Figure 1.2: (a) Map view of an illustrative site, with two nearby sources capable of producing earthquakes. (b) Predicted median response spectra from the two earthquake events, illustrating that the event producing the maximum response spectra may vary depending upon the period of interest (prediction obtained from the model of Campbell and Bozorgnia 2008).

While the site shown in Figure 1.2a produces some challenges in terms of identifying a worst-case event, an even greater challenges arise when faults near a site are not obvious and so the seismic source is quantified as an areal source capable of producing earthquakes at any location, as shown in Figure 1.3. In this case, the worst-case event has to be the one with the maximum conceivable magnitude, at a location directly below the site of interest (i.e., with a distance of 0 km). This is clearly the maximum event, no matter how unlikely its occurrence may be. For example, in parts of the Eastern United States, especially near the past Charleston or New Madrid earthquakes, one can quite feasibly hypothesize the occurrence of magnitude 7.5 or 8 earthquakes immediately below a site, although that event may occur very rarely.

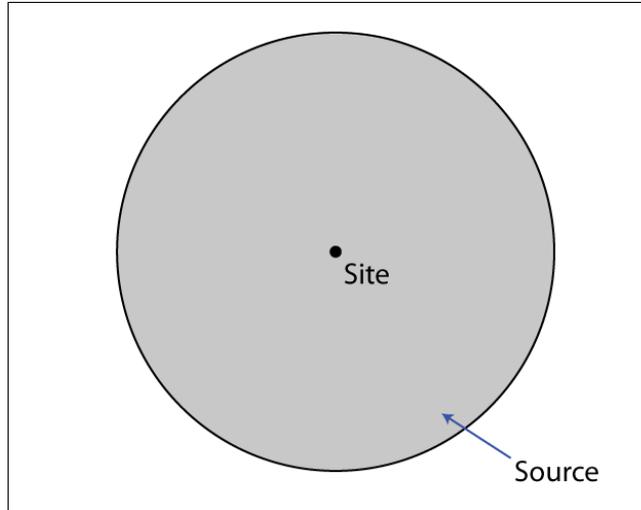


Figure 1.3: Example site at the center of an area source, with potential earthquakes at zero distance from the site.

1.2.2 Variability of ground motion intensity for a given earthquake event

While the choice of a “worst-case” earthquake can be difficult and subjective, as discussed in the previous section, an even greater problem with deterministic hazard analysis is the choice of worst-case ground motion intensity associated with that earthquake. The response spectra plotted in Figure 1.2 are the median¹ spectra predicted by empirical models calibrated to recorded ground motions. But recorded ground motions show a very large amount of scatter around those median predictions. By definition, the median predictions shown in Figure 1.2b are exceeded in 50% of observed ground motions having the given magnitudes and distances.

An example of the large scatter around those ground motion prediction models is seen in Figure 1.4, which shows spectral acceleration values at 1 second that were observed in a past earthquake (1999 Chi-Chi, Taiwan), plotted versus the closest distance from the earthquake rupture to the recording site. Note that observations at distances between 1 and 3 km vary between 0.15g and more than 1g—nearly an order of magnitude. Also plotted are the mean predicted $\ln SA$ values, along with bounds illustrating one standard deviation above and below that mean. The scatter of the log of

¹ There is considerable opportunity for confusion when referring to means and medians of predicted ground motion intensity. Ground motion prediction models, such as the one used to make Figure 1.4, provide the mean and standard deviation of the *natural logarithm* of spectral acceleration ($\ln SA$) or peak ground acceleration ($\ln PGA$). These $\ln SA$ values are normally distributed, which means that the non-logarithmic SA values are lognormally distributed. The exponential of the *mean* $\ln SA$ value can be shown to equal the median SA value. It is easiest to work with $\ln SA$ values in the calculations that follow, so this text will often refer to mean $\ln SA$ values rather than median SA values. Plots such as Figure 1.4 will show non-logarithmic SA , because the units are more intuitive, but the axis will always be in logarithmic scale so that the visual effect is identical to if one was viewing a plot of $\ln SA$.

spectral accelerations around the mean prediction is well-represented by a normal distribution (leading to symmetric scatter in Figure 1.4, which is plotted in logarithmic scale).

The one-standard-deviation bounds should enclose about 2/3 of the observed values if the variations are normally distributed, and that is the case here. To account for this scatter, deterministic hazard analyses sometimes specify a “mean plus one standard deviation” response spectra, but even that will be exceeded 16% of the time². Because the scatter is normally distributed, there is no theoretical upper bound on the amplitude of ground motion that might be produced at a given magnitude and distance³.

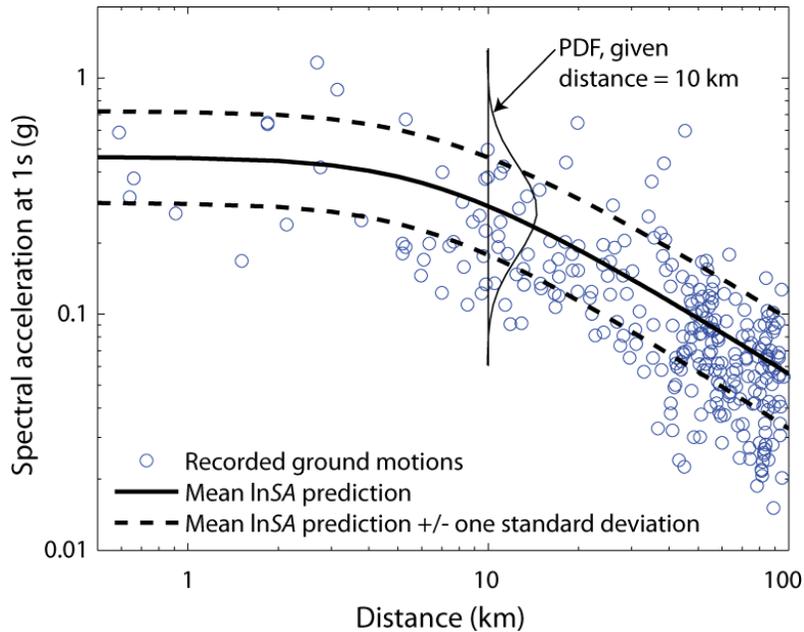


Figure 1.4: Observed spectral acceleration values from the 1999 Chi-Chi, Taiwan earthquake, illustrating variability in ground motion intensity. The predicted distribution comes from the model of Campbell and Bozorgnia (2008).

1.2.3 Can we use a deterministic approach, given these uncertainties?

Given these challenges, it is clear that whatever deterministic design earthquake and ground motion intensity is eventually selected, it is not a true “worst-case” event, as a larger earthquake or ground motion can always plausibly be proposed. Without a true worst-case event to consider, we are left to

² This number comes from considering normally-distributed residuals. As seen in Table 4.1, the probability of a normal random variable being more than one standard deviation greater than its mean (i.e., $1 - \Phi(1)$) is 0.16.

³ There is almost certainly some true physical upper bound on ground motion intensity caused by an inability of the earth to carry more intense seismic waves without shattering or otherwise failing. Current research suggests that this limit may be important to structures designed for extremely intense ground motions, such as nuclear waste repositories, but it almost certainly has no practical impact on more common structures such as buildings or bridges, which are analyzed for ground motion intensities that are exceeded once every few thousand years. Thus, the assumption of not theoretical upper bound is reasonable and appropriate in most cases.

identify a “reasonably large” event. That is often done by choosing a nearby large-magnitude event, and then identifying some level of reasonable intensity associated with this event. While it is possible to proceed using this type of approach, two issues should be made clear. 1) The resulting ground motion is not a “worst-case” ground motion. 2) The result may be very sensitive to decisions made with regard to the chosen scenario magnitude and ground motion intensity. An event chosen in this manner was historically described as a “Maximum Credible Earthquake,” or MCE. More recently, however, the acronym has been retained but taken to mean “Maximum Considered Earthquake,” in recognition of the fact that larger earthquakes (and larger ground motion intensities) are likely to be credible as well. This “worst-case” thinking will be abandoned for the remainder of the document, although the problems identified here will serve as a useful motivation for thinking about probability-based alternatives.

1.3 Probabilistic seismic hazard analysis calculations

In this section, we will describe a probability-based framework capable of addressing the concerns identified above. Rather than ignoring the uncertainties present in the problem, this approach incorporates them into calculations of potential ground motion intensity. While incorporation of uncertainties adds some complexity to the procedure, the resulting calculations are much more defensible for use in engineering decision-making for reducing risks.

With PSHA, we are no longer searching for an elusive worst-case ground motion intensity. Rather, we will consider all possible earthquake events and resulting ground motions, along with their associated probabilities of occurrence, in order to find the level of ground motion intensity exceeded with some tolerably low rate. At its most basic level, PSHA is composed of five steps.

1. Identify all earthquake sources capable of producing damaging ground motions.
2. Characterize the distribution of earthquake magnitudes (the rates at which earthquakes of various magnitudes are expected to occur).
3. Characterize the distribution of source-to-site distances associated with potential earthquakes.
4. Predict the resulting distribution of ground motion intensity as a function of earthquake magnitude, distance, etc.
5. Combine uncertainties in earthquake size, location and ground motion intensity, using a calculation known as the total probability theorem.

These steps will be explained in more detail below.

The end result of these calculations will be a full distribution of levels of ground shaking intensity, and their associated rates of exceedance. The illusion of a worst-case ground motion will be removed, and replaced by identification of occurrence frequencies for the full range of ground motion intensities of potential interest. These results can then be used to identify a ground motion intensity having an acceptably small probability of being exceeded.

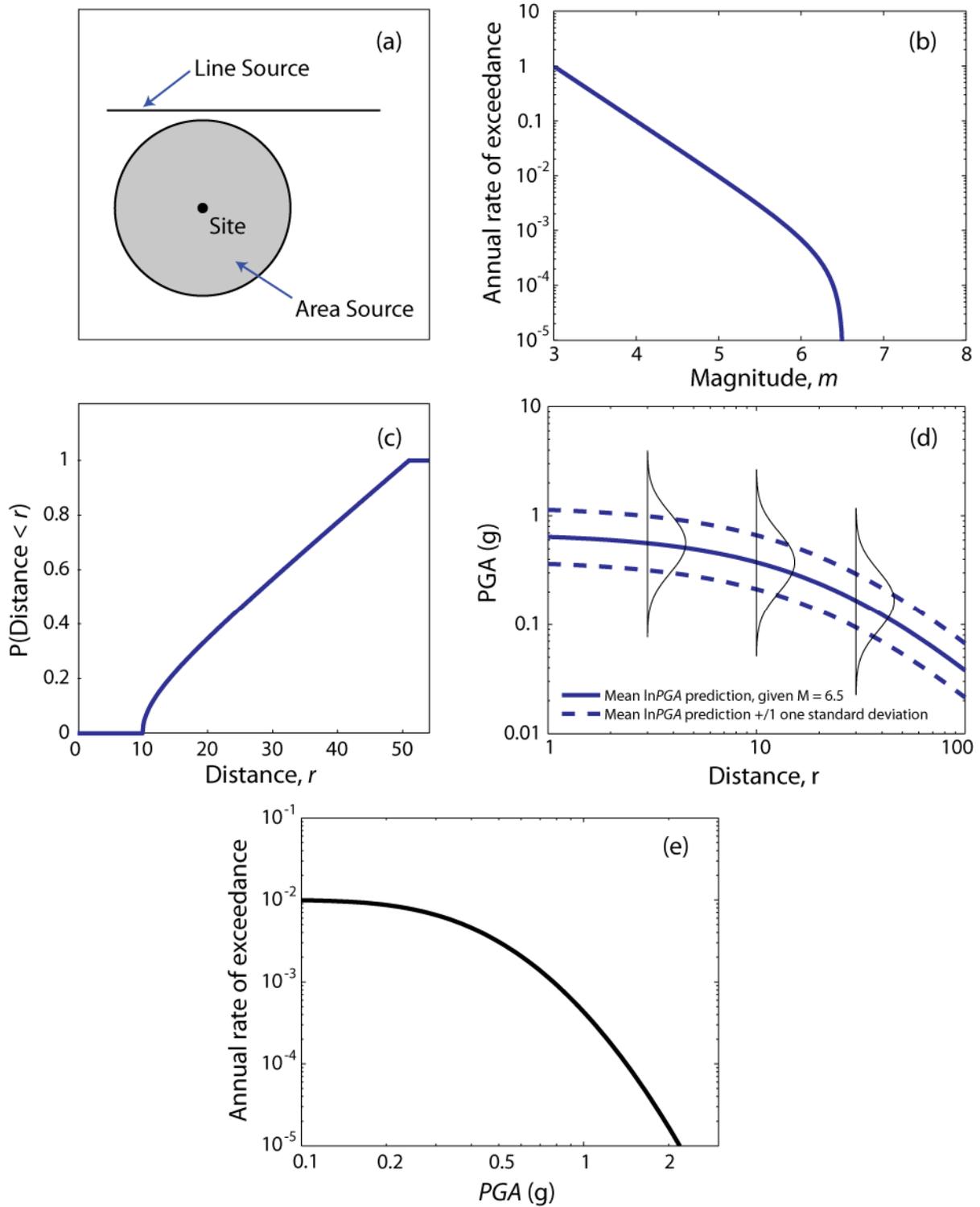


Figure 1.5: Schematic illustration of the basic five steps in probabilistic seismic hazard analysis. (a) Identify earthquake sources. (b) Characterize the distribution of earthquake magnitudes from each source. (c) Characterize the distribution of source-to-site distances from each source. (d) Predict the resulting distribution of ground motion intensity. (e) Combine information from parts a-d to compute the annual rate of exceeding a given ground motion intensity.

1.3.1 Identify earthquake sources

In contrast to the deterministic thinking above, which focused only on the largest possible earthquake event, here we are interested in all earthquake sources capable of producing damaging ground motions at a site. These sources could be faults, which are typically planar surfaces identified through various means such as observations of past earthquake locations and geological evidence. If individual faults are not identifiable (as in the less seismically active regions of the eastern United States), then earthquake sources may be described by areal regions in which earthquakes may occur anywhere. Once all possible sources are identified, we can identify the distribution of magnitudes and source-to-site distances associated with earthquakes from each source.

1.3.2 Identify earthquake magnitudes

Tectonic faults are capable of producing earthquakes of various sizes (i.e., magnitudes). Gutenberg and Richter (1944) first studied observations of earthquake magnitudes, and noted that the distribution of these earthquake sizes in a region generally follows a particular distribution, given as follows

$$\log \lambda_m = a - bm \quad (1.1)$$

where λ_m is the rate of earthquakes with magnitudes greater than m , and a and b are constants. This equation is called the *Gutenberg-Richter recurrence law*. Figure 1.6 illustrates typical observations from a fault or region, along with the Gutenberg-Richter recurrence law given by equation 1.1.

The a and b constants from equation 1.1 are estimated using statistical analysis of historical observations, with additional constraining data provided by other types of geological evidence⁴. The a value indicates the overall rate of earthquakes in a region, and the b value indicates the relative ratio of small and large magnitudes (typical b values are approximately equal to 1).

Equation 1.1 can also be used to compute a cumulative distribution function⁵ (CDF) for the magnitudes of earthquakes that are larger than some minimum magnitude m_{min} (this conditioning is used because earthquakes smaller than m_{min} will be ignored in later calculations due to their lack of engineering importance).

⁴ Note that some care is needed during this process to ensure that no problems are caused by using historical data that underestimates the rate of small earthquakes due to the use of less sensitive instruments in the past. Methods have been developed to address this issue (e.g., Weichert 1980), but are not considered further in this document.

⁵ Probability tools such as cumulative distribution functions and probability density functions are necessary for much of the analysis that follows. See Section 4 for a review of this material.

$$\begin{aligned}
F_M(m) &= P(M \leq m | M > m_{\min}) \\
&= \frac{\text{Rate of earthquakes with } m_{\min} < M \leq m}{\text{Rate of earthquakes with } m_{\min} < M} \\
&= \frac{\lambda_{m_{\min}} - \lambda_m}{\lambda_{m_{\min}}} \\
&= \frac{10^{a-bm_{\min}} - 10^{a-bm}}{10^{a-bm_{\min}}} \\
&= 1 - 10^{-b(m-m_{\min})}, \quad m > m_{\min}
\end{aligned} \tag{1.2}$$

where $F_M(m)$ denotes the cumulative distribution function for M . One can compute the probability density function (PDF) for M by taking the derivative of the CDF

$$\begin{aligned}
f_M(m) &= \frac{d}{dm} F_M(m) \\
&= \frac{d}{dm} [1 - 10^{-b(m-m_{\min})}] \\
&= b \ln(10) 10^{-b(m-m_{\min})}, \quad m > m_{\min}
\end{aligned} \tag{1.3}$$

where $f_M(m)$ denotes the probability density function for M .

Note that the PDF given in equation 1.3 relies on the Gutenberg-Richter law of equation 1.1, which theoretically predicts magnitudes with no upper limit, although physical constraints make this unrealistic. There is generally some limit on the upper bound of earthquake magnitudes in a region, due to the finite size of the source faults (earthquake magnitude is related to the area of the seismic rupture). If a maximum magnitude can be determined, then equation 1.2 becomes

$$F_M(m) = \frac{1 - 10^{-b(m-m_{\min})}}{1 - 10^{-b(m_{\max}-m_{\min})}}, \quad m_{\min} < m < m_{\max} \tag{1.4}$$

and equation 1.3 becomes

$$f_M(m) = \frac{b \ln(10) 10^{-b(m-m_{\min})}}{1 - 10^{-b(m_{\max}-m_{\min})}}, \quad m_{\min} < m < m_{\max} \tag{1.5}$$

where m_{\max} is the maximum earthquake that a given source can produce. This limited magnitude distribution is termed a *bounded Gutenberg-Richter recurrence law*. Example observations of earthquake magnitudes are shown in Figure 1.6, along with Gutenberg-Richter and bounded Gutenberg-Richter recurrence laws fit to the data.

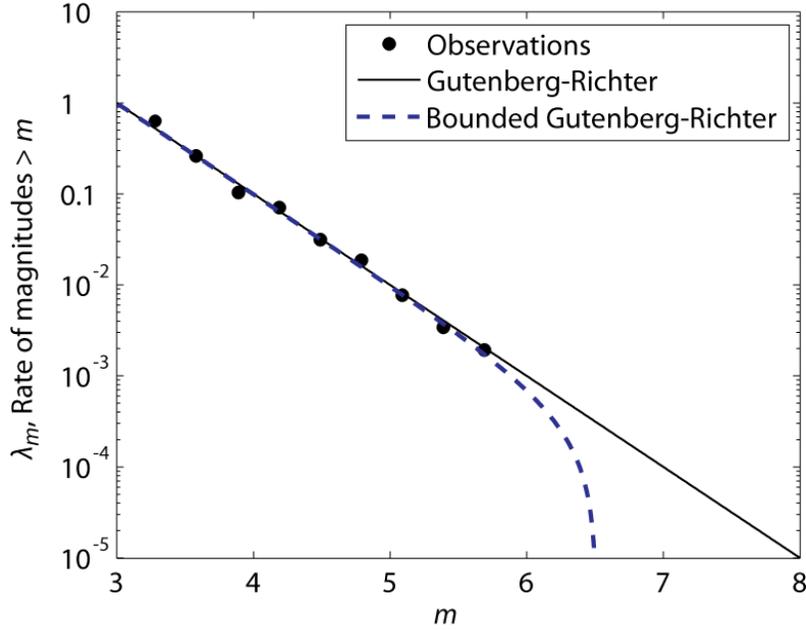


Figure 1.6: Typical distribution of observed earthquake magnitudes, along with Gutenberg-Richter and bounded Gutenberg-Richter recurrence laws fit to the observations.

For our later PSHA equations, we will convert the continuous distribution of magnitudes into a discrete set of magnitudes. For example, consider a source with a minimum considered magnitude of 5, a maximum magnitude of 8, and a b parameter equal to 1. Table 1.1 lists probabilities of interest for this example source. The first column lists 13 magnitude values between 5 and 8. The second column lists the cumulative distribution function, as computed using equation 1.4. The third column lists probabilities of occurrence of these discrete set of magnitudes, assuming that they are the only possible magnitudes; they are computed as follows

$$P(M = m_j) = F_M(m_{j+1}) - F_M(m_j) \quad (1.6)$$

where m_j are the discrete set of magnitudes, ordered so that $m_j < m_{j+1}$. This calculation assigns the probabilities associated with all magnitudes between m_j and m_{j+1} to the discrete value m_j . As long as the discrete magnitudes are closely spaced, the approximation will not affect numerical results. Magnitudes are spaced at intervals of 0.25 for illustration in Table 1.1 so that the table is not too lengthy, but a practical PSHA analysis might use a magnitude spacing of 0.1 or less.

Table 1.1: Magnitude probabilities for a source with a truncated Gutenberg-Richter distribution, a minimum considered magnitude of 5, a maximum magnitude of 8, and a b parameter of 1. The numbers in this table were computed using equations 1.4 and 1.6.

m_j	$F_M(m_j)$	$P(M = m_j)$
5.00	0.0000	0.4381
5.25	0.4381	0.2464
5.50	0.6845	0.1385
5.75	0.8230	0.0779
6.00	0.9009	0.0438
6.25	0.9447	0.0246
6.50	0.9693	0.0139
6.75	0.9832	0.0078
7.00	0.9910	0.0044
7.25	0.9954	0.0024
7.50	0.9978	0.0014
7.75	0.9992	0.0008
8.00	1.0000	0.0000

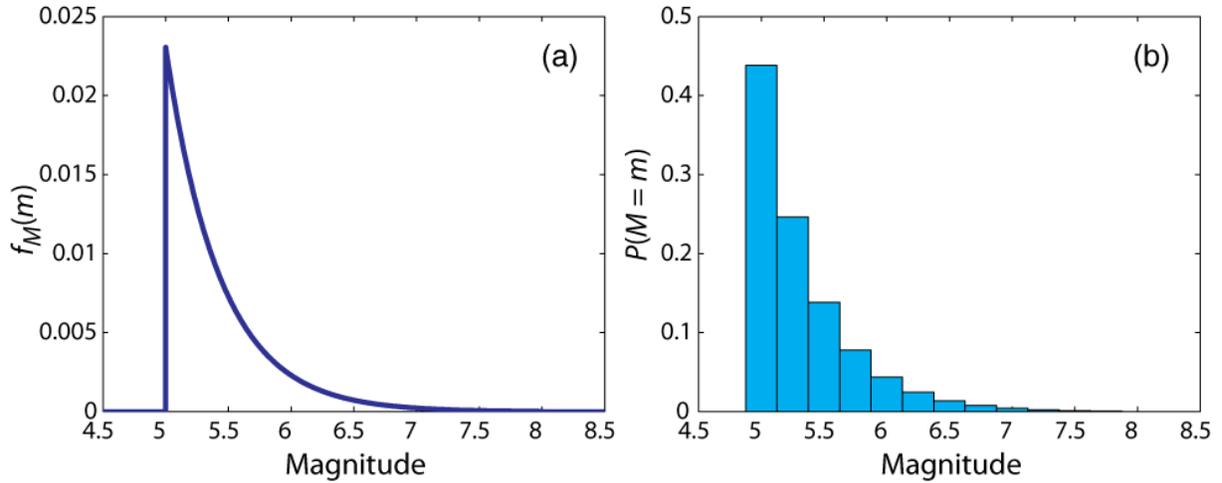


Figure 1.7: Illustration of discretization of a continuous magnitude distribution for a source with a truncated Gutenberg-Richter distribution, a minimum considered magnitude of 5, a maximum magnitude of 8, and a b parameter of 1. (a) Continuous probability density function from equation 1.5. (b) Discrete probabilities from equation 1.6.

An aside: The Gutenberg-Richter models above are not the only models proposed for describing the distribution of earthquake magnitudes. One common alternative is the Characteristic Earthquake model, which proposes that some faults have repeated occurrences of a characteristic earthquake with a reasonably consistent magnitude (Schwartz and Coppersmith 1984). This characteristic magnitude occurs more often than predicted by the Gutenberg-Richter models proposed above. All that is required to adopt an alternative recurrence model is to replace equation 1.5 with a suitably modified

probability density function (e.g., Lomnitz-Adler and Lomnitz 1979; Youngs and Coppersmith 1985). All other PSHA equations remain identical.

1.3.3 Identify earthquake distances

To predict ground shaking at a site, it is also necessary to model the distribution of distances from earthquakes to the site of interest. For a given earthquake source, it is generally assumed that earthquakes will occur with equal probability at any location on the fault⁶. Given that locations are uniformly distributed, it is generally simple to identify the distribution of source-to-site distances using only the geometry of the source. Example calculations are shown in this section for an area source and a line source.

An aside: While “distance” sounds like a well-defined term, there are in fact several definitions commonly used in PSHA. One can use distance to the epicenter or hypocenter, distance to the closest point on the rupture surface, or distance to the closest point on the surface projection of the rupture. Note that some distance definitions account for the depth of the rupture, while others consider only distance from the surface projection of the rupture. Note also that epicenter- and hypocenter-based definitions need only consider the location of rupture initiation; some other definitions need to explicitly account for the fact that ruptures occur over a plane rather than at a single point in space. The analyst’s choice of distance definition will depend upon the required input to the ground motion prediction model. Here we will consider only distance to the epicenter, for simplicity.

1.3.3.1 Example: Area source

Consider a site located in an area source. The source produces earthquakes randomly and with equal likelihood anywhere within 100 km of the site. (In actuality, the source may be larger, but is typically truncated at some distance beyond which earthquakes are not expected to cause damage at the site.) Area sources are often used in practice to account for “background” seismicity, or for earthquakes that are not associated with any specific fault. The example source is illustrated in Figure 1.8.

⁶ In a few special cases, analysts use non-uniform distributions for future earthquake locations based on models for stress transfer and time-dependent earthquake occurrence. Those situations will not be considered here.

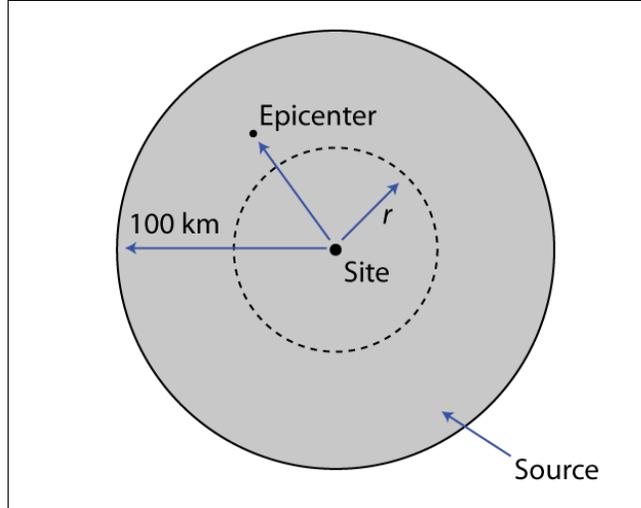


Figure 1.8: Illustration of an example area source.

We can easily compute a probabilistic description for the distances to earthquakes in this case by noting that if the earthquakes are equally likely to occur anywhere, then the probability of an epicenter being located at a distance of less than r is equal to the area of a circle of radius r , divided by the area of the circle of radius 100

$$\begin{aligned}
 F_R(r) = P(R \leq r) &= \frac{\text{area of circle with radius } r}{\text{area of circle with radius 100}} \\
 &= \frac{\pi r^2}{\pi(100)^2} \\
 &= \frac{r^2}{10,000}
 \end{aligned} \tag{1.7}$$

Equation 1.7 is only valid, however, for distance (r) values between 0 and 100 km. Accounting for other ranges gives the more complete description

$$F_R(r) = \begin{cases} 0 & \text{if } r < 0 \\ \frac{r^2}{10,000} & \text{if } 0 \leq r < 100 \\ 1 & \text{if } r \geq 100 \end{cases} \tag{1.8}$$

We can also find the PDF for the distance by taking a derivative of the CDF

$$f_R(r) = \frac{d}{dr} F_R(r) = \begin{cases} = \frac{r}{5000} & \text{if } 0 \leq r < 100 \\ = 0 & \text{otherwise} \end{cases} \tag{1.9}$$

Plots of this PDF and CDF are shown in Figure 1.9. We see that distances close to 0 km are possible but unlikely, because there are few locations in the source that are associated with such small distances. Unlike the deterministic “worst-case” distance of 0 km, the PSHA calculations below will use these distributions to account for the differing probabilities of observing earthquakes at various distances.

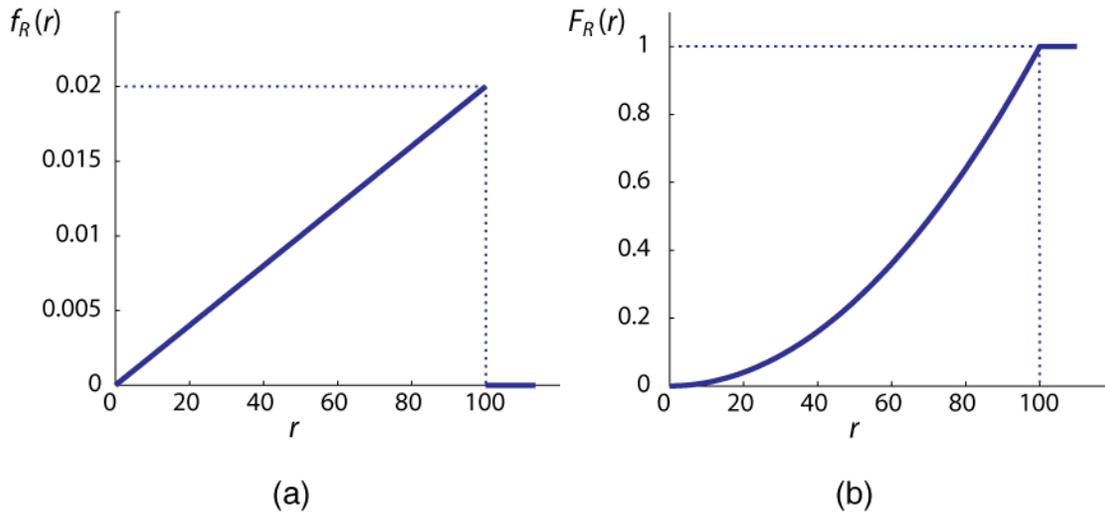


Figure 1.9: PDF and CDF of the source-to-site distance for future earthquakes from the example area source.

1.3.3.2 Example: Line source

Earthquake sources are also sometimes quantified as line sources⁷. This is particularly appropriate for modeling identified faults that exist on the boundary of two tectonic plates (as is the case in much of coastal California).

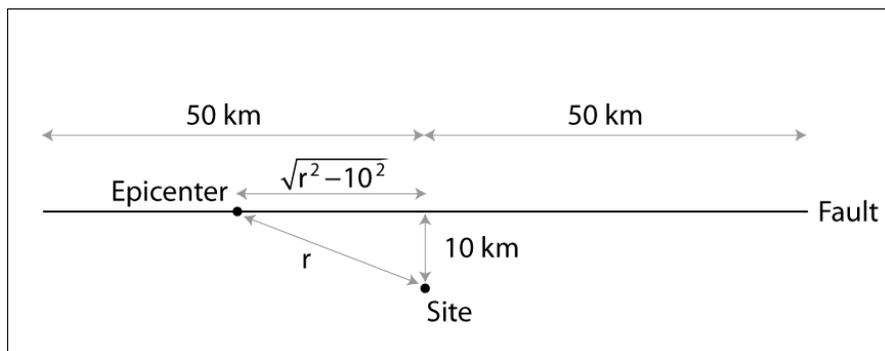


Figure 1.10: Illustration of an example line source.

⁷ It is also common to treat the earth’s structure in three dimensions, meaning that faults will be represented as planes rather than lines. The examples in this document will all be two-dimensional for simplicity, but the extension to three dimensions is straightforward.

Consider a 100 km fault, modeled as a line source in Figure 1.10, with a site located 10 km from the center of the fault. We again assume that earthquake epicenters are equally likely at all locations. In this case, the probability of observing a distance of less than r is equal to the fraction of the fault located within a radius of r . Using the Pythagorean theorem, we can see that the distance from the center of the fault to a point a distance r from the site is $\sqrt{r^2 - 10^2}$.

Using this information, we can then compute the CDF of R

$$F_R(r) = P(R \leq r) = \frac{\text{length of fault within distance } r}{\text{total length of the fault}} \quad (1.10)$$

$$= \frac{2\sqrt{r^2 - 10^2}}{100}$$

but that equation is only true for distances of less than 10 km or greater than 51 km. Distances outside that range are not possible on this fault, so the complete CDF is

$$F_R(r) = \begin{cases} 0 & \text{if } r < 10 \\ \frac{2\sqrt{r^2 - 10^2}}{100} & \text{if } 10 \leq r < 51 \\ 1 & \text{if } r \geq 51 \end{cases} \quad (1.11)$$

The PDF can be obtained from the derivative of the CDF

$$f_R(r) = \frac{d}{dr} F_R(r) = \begin{cases} = \frac{r}{50\sqrt{r^2 - 100}} & \text{if } 10 \leq r < 51 \\ = 0 & \text{otherwise} \end{cases} \quad (1.12)$$

The PDF and CDF are plotted in Figure 1.11.

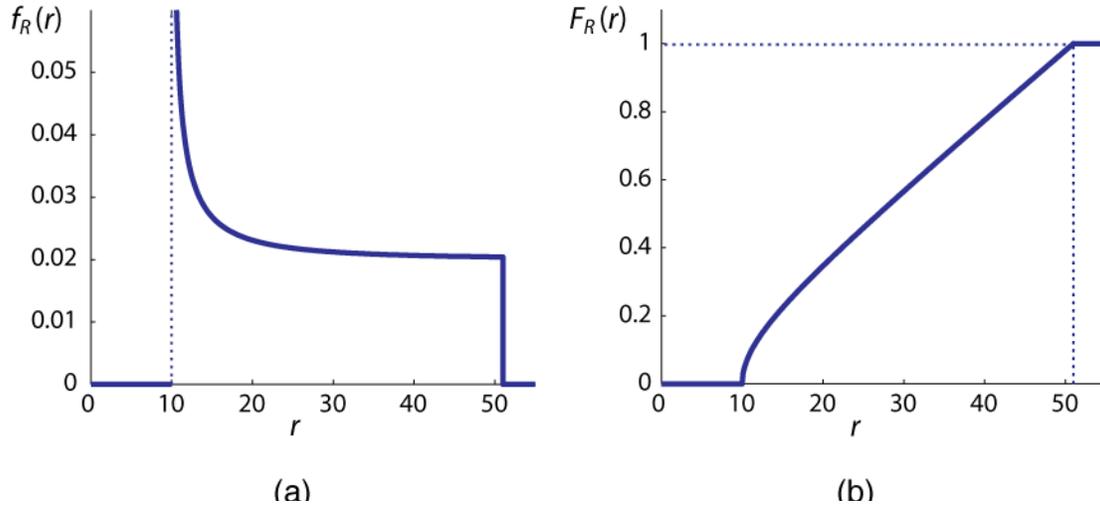


Figure 1.11: PDF and CDF of the source-to-site distance for future earthquakes from the example line source.

These two example sources are shown to provide simple examples of distance distributions. Distributions for more irregular sources can also be computed in the same manner. These distributions are important inputs for the calculations that follow.

1.3.4 Ground motion intensity

We have now quantified the distribution of potential earthquake magnitudes and locations, but we are interested in analyzing ground motions—not earthquakes. The next step is therefore a ground motion prediction model⁸. These models predict the probability distribution of ground motion intensity, as a function of many predictor variables such as the earthquake’s magnitude, distance, faulting mechanism, the near-surface site conditions, the potential presence of directivity effects, etc. Because the number of predictor variables is large, we often write that the model predicts ground motion intensity given “magnitude, distance, etc.”

Ground motion prediction models are generally developed using statistical regression on observations from large libraries of observed ground motion intensities. For example, spectral acceleration (SA) values at 1 second observed in the 1999 Chi-Chi, Taiwan, earthquake were shown previously in Figure 1.4, along with lines showing the predicted mean (and mean \pm one standard deviation) of the $\ln SA$ values from an example ground motion prediction model (Campbell and Bozorgnia 2008). That prediction model, like other modern models, was fit to thousands of observed intensities from dozens of past earthquakes.

⁸ These models were called “attenuation models” or “attenuation relations” until recently. Those names have fallen out of favor, however, because the prediction model accounts for a great number of effects, of which attenuation is only one.

It is apparent from Figure 1.4 that there is significant scatter in observed ground motion intensities, even after accounting for the effect of magnitude, distance, etc. Thus, these predictive models must provide a probability distribution for intensities, rather than just a single intensity. This is important, because our later PSHA calculations need to account for the possibility of unlikely outcomes such as extreme intensities much larger than the predicted mean (Bommer and Abrahamson 2006).

To describe this probability distribution, prediction models take the following general form:

$$\ln IM = \overline{\ln IM}(M, R, \theta) + \sigma(M, R, \theta) \cdot \varepsilon \quad (1.13)$$

where $\ln IM$ is the natural log of the ground motion *intensity measure* of interest (such as spectral acceleration at a given period); this $\ln IM$ is modeled as a random variable, and has been seen to be well-represented by a normal distribution. The terms $\overline{\ln IM}(M, R, \theta)$ and $\sigma(M, R, \theta)$ are the outputs of the ground motion prediction model; they are the predicted mean and standard deviation, respectively, of $\ln IM$. These terms are both functions of the earthquake's magnitude (M), distance (R) and other parameters (generically referred to as θ). Finally, ε is a standard normal random variable that represents the observed variability in $\ln IM$. Positive values of ε produce larger-than-average values of $\ln IM$, while negative values of ε produce smaller-than-average values of $\ln IM$.

Over decades of development and refinement, the prediction models for $\overline{\ln IM}(M, R, \theta)$ and $\sigma(M, R, \theta)$ have become complex, consisting of many terms and tables containing dozens of coefficients. These modern models are no longer simple to calculate using pencil and paper, so here we will use an older and much simpler (but obsolete) model to illustrate the example calculations. The approach is identical when using modern prediction models, but this simple model keeps us from being distracted by tedious arithmetic.

Cornell et al. (1979) proposed the following predictive model for the mean of log peak ground acceleration (in units of g)

$$\overline{\ln PGA} = -0.152 + 0.859M - 1.803 \ln(R + 25) \quad (1.14)$$

The standard deviation of $\ln PGA$ was 0.57 in this model, and was constant for all magnitudes and distances. The natural logarithm of PGA was seen to be normally distributed, so we can compute the probability of exceeding any PGA level using knowledge of this mean and standard deviation

$$P(PGA > x | m, r) = 1 - \Phi\left(\frac{\ln x - \overline{\ln PGA}}{\sigma_{\ln PGA}}\right) \quad (1.15)$$

where $\Phi(\)$ is the standard normal cumulative distribution function, as shown in Table 4.1 on page 33. Modern prediction models also provide a mean and standard deviation to be used in equation 1.15,

so the general procedure is identical when using newer models; the equations for predicting the mean and standard deviation are just more complicated.

Equation 1.15 used the cumulative distribution function to compute $P(PGA > x | m, r)$, but sometimes it may be useful to use an alternate formulation incorporating the probability density function for PGA . Noting that the cumulative distribution function is equivalent to an integral of the probability density function (equation 4.25), we can also write

$$P(PGA > x | m, r) = \int_x^{\infty} f_{PGA}(u) du \quad (1.16)$$

where $f_{PGA}(u)$ is the probability density function of PGA , given m and r . Unlike the cumulative distribution function $\Phi(\cdot)$, $f_{PGA}(u)$ can actually be written out analytically. Substituting in this PDF gives

$$P(PGA > x | m, r) = \int_x^{\infty} \frac{1}{\sigma_{\ln PGA} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln u - \overline{\ln PGA}}{\sigma_{\ln PGA}}\right)^2\right) du \quad (1.17)$$

This integral can then be evaluated numerically within the PSHA software.

To connect these equations to a visual display of ground motion predictions, consider Figure 1.12, which shows PGA predictions for a magnitude 6.5 earthquake, as a function of distance. The mean and the mean +/- one standard deviation of the Cornell et al. prediction is plotted for distances between 1 and 100 km. At distances of 3, 10 and 30 km, the entire PDF of the predicted normal distribution is also superimposed. Suppose we were interested in the probability of $PGA > 1g$. At those three distances, equation 1.14 gives predicted means of -0.5765, -0.9788 and -1.7937, respectively⁹. At all three distances, the standard deviation of $\ln PGA$ is 0.57. So we can use equation 1.15 to compute the probabilities of exceedance as

$$\begin{aligned} P(PGA > 1 | 6.5, 3) &= 1 - \Phi\left(\frac{\ln 1 - (-0.5765)}{0.57}\right) = 1 - \Phi(1.01) = 0.16 \\ P(PGA > 1 | 6.5, 10) &= 1 - \Phi\left(\frac{\ln 1 - (-0.9788)}{0.57}\right) = 1 - \Phi(1.72) = 0.043 \\ P(PGA > 1 | 6.5, 30) &= 1 - \Phi\left(\frac{\ln 1 - (-1.7937)}{0.57}\right) = 1 - \Phi(3.15) = 0.0008 \end{aligned} \quad (1.18)$$

⁹ All of the example calculations will provide answers with more significant figures than should reasonably be used or reported. This is done to allow readers to reproduce the example calculations exactly, and because many answers are intermediate results for later calculations.

These probabilities correspond to the fraction of the corresponding PDFs in Figure 1.12 that are shaded. This visual interpretation may provide some intuition regarding the previous equations, which will likely be rather unfamiliar to readers who do not regularly perform probability calculations.

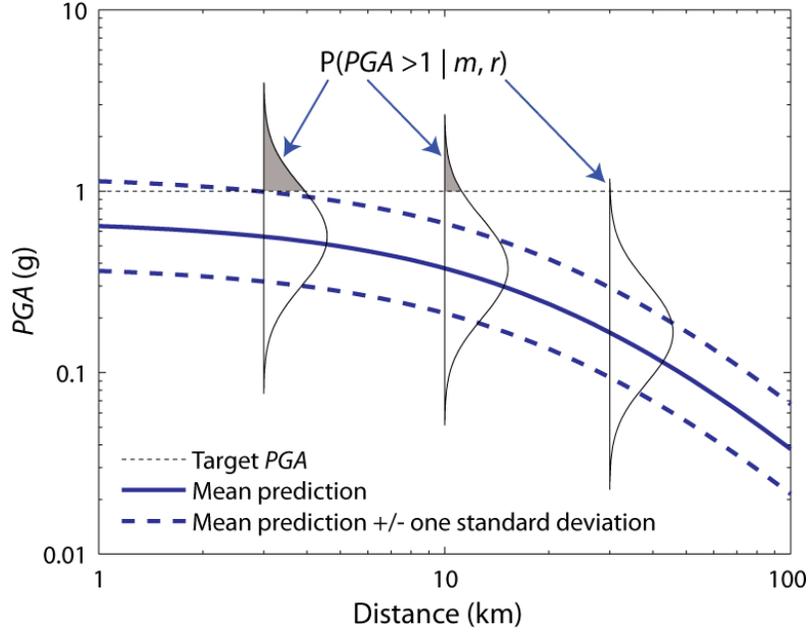


Figure 1.12: Graphical depiction of the example ground motion prediction model for a magnitude 6.5 earthquake, and the probability of $PGA > 1g$ at several source-to-site distances.

Let us consider a second example, which will provide intermediate results for some calculations that follow. Assume magnitude 5 earthquake has occurred at a distance of 10 km. The Cornell et al. ground motion prediction model provides a mean $\ln PGA$ of -2.2673, and a standard deviation of 0.57. The first column of Table 1.2 lists a series of PGA values of potential interest. The second column lists the probabilities of exceeding those PGA values, using equation 1.15. The third column lists the probability of equaling those PGA values, using the same discretization approach we used previously for the continuous magnitude distribution

$$P(PGA = x_j) = P(PGA > x_j) - P(PGA > x_{j+1}) \quad (1.19)$$

Table 1.2: PGA probabilities associated with a magnitude 5 earthquake at 10 km.

x_j	$P(PGA > x_j)$	$P(PGA = x_j)$
0.20	0.12418	0.06312
0.25	0.06106	0.03004
0.30	0.03102	0.01471
0.35	0.01631	0.00745
0.40	0.00886	0.00390
0.45	0.00496	0.00211
0.50	0.00285	0.00117

0.55	0.00168	0.00067
0.60	0.00101	0.00039
0.65	0.00062	0.00024
0.70	0.00038	0.00014
0.75	0.00024	0.00009
0.80	0.00015	0.00006
0.85	0.00009	0.00004
0.90	0.00005	0.00002
0.95	0.00003	0.00002
1.00	0.00001	0.00001

An aside: At first glance, one might wonder whether the large variability must necessarily be a part of prediction models, or whether it is a fundamental error caused by over-simplifications or inappropriate mixing of observational data. The uncertainty arises because we are trying to predict a highly complex phenomenon (ground shaking intensity at a site) using very simplified predictive parameters (magnitude, distance, and perhaps a few other parameters). Earthquake rupture is a complex spatial-temporal process, and here we represent it simply by “magnitude,” which measures the total seismic energy released in the complex rupture. Similarly, nonlinear wave scattering and propagation through a complex structure such as the earth is represented by simply the distance between the source and the site. It is true that if we had more detailed knowledge of the rupture and propagation process, we might predict ground shaking intensity with less uncertainty. But we don’t always have detailed models for rupture and wave propagation from past earthquakes to use in calibrating predictive models, and even if we were able to develop complex predictive models, then our predictions of future earthquake events would have to be much more detailed than simply predicting their distributions of magnitudes and distances. Ground motion prediction equations of the type used to produce Figure 1.4 have evolved over a period of 40 years, and are now developed using thousands of observed ground motions and are constrained using many theoretical and physical insights. This level of refinement suggests that there is little hope of this scatter being reduced significantly in the near future. The scatter is best thought of as an inherent variability in the earthquake hazard environment that must be accounted for when identifying a design ground motion intensity.

1.3.5 Combine all information

With the above information in place, we can now combine it using the PSHA equations. We will first consider two intermediate calculations as we build towards a PSHA equation that considers multiple sources.

First, let us compute the probability of exceeding an IM intensity level x , *given* occurrence of a future earthquake from a single source. The ground motion prediction model of Section 1.3.4 allows us to compute the probability of exceeding that IM level for a given magnitude and distance. The magnitude and distance of the future earthquake are not yet known, but we can find their probability distributions using Sections 1.3.2 and 1.3.3. We then combine this information using the total probability theorem

$$P(IM > x) = \int_{m_{\min}}^{m_{\max}} \int_0^{r_{\max}} P(IM > x | m, r) f_M(m) f_R(r) dr dm \quad (1.20)$$

where $P(IM > x | m, r)$ comes from the ground motion model, $f_M(m)$ and $f_R(r)$ are our PDFs for magnitude and distance, and we integrate over all considered magnitudes and distances¹⁰. The integration operation adds up the conditional probabilities of exceedance associated with all possible magnitudes and distances (with the PDFs weighting each conditional exceedance probability by the probability of occurrence of the associated magnitude and distance).

Equation 1.20 is a probability of exceedance given an earthquake, and does not include any information about how often earthquakes occur on the source of interest. We can make a simple modification to that equation, to compute the rate of $IM > x$, rather than the probability of $IM > x$ given occurrence of an earthquake.

$$\lambda(IM > x) = \lambda(M > m_{\min}) \int_{m_{\min}}^{m_{\max}} \int_0^{r_{\max}} P(IM > x | m, r) f_M(m) f_R(r) dr dm \quad (1.21)$$

where $\lambda(M > m_{\min})$ is the rate of occurrence of earthquakes greater than m_{\min} from the source, and $\lambda(IM > x)$ is the rate of $IM > x$.

To generalize the analysis further, we would like to consider cases with more than one source. Recognizing that the rate of $IM > x$ when considering all sources is simply the sum of the rates of $IM > x$ from each individual source, we can write

$$\lambda(IM > x) = \sum_{i=1}^{n_{\text{sources}}} \lambda(M_i > m_{\min}) \int_{m_{\min}}^{m_{\max}} \int_0^{r_{\max}} P(IM > x | m, r) f_{M_i}(m) f_{R_i}(r) dr dm \quad (1.22)$$

where n_{sources} is the number of sources considered, and M_i/R_i denote the magnitude/distance distributions for source i . Since we will nearly always be performing this calculation on a computer, it

¹⁰ More generally, we should use a joint distribution for magnitude and distance, $f_{M,R}(m,r)$, rather than the product of their marginal distances $f_M(m)f_R(r)$. The above formulation is correct only if the magnitudes and distances of events are independent. The above formulation is helpful, however, for easily incorporating the magnitude and distance distributions computed earlier.

is practical to discretize our continuous distributions for M and R , and convert the integrals into discrete summations, as follows

$$\lambda(IM > x) = \sum_{i=1}^{n_{\text{sources}}} \lambda(M_i > m_{\min}) \sum_{j=1}^{n_M} \sum_{k=1}^{n_R} P(IM > x | m_j, r_k) P(M_i = m_j) P(R_i = r_k) \quad (1.23)$$

where the range of possible M_i and R_i have been discretized into n_M and n_R intervals, respectively, using the discretization technique discussed earlier.

Equation 1.22 (or, equivalently, 1.23) is the equation most commonly associated with PSHA. It has integrated our knowledge about rates of occurrence of earthquakes, the possible magnitudes and distances of those earthquakes, and the distribution of ground shaking intensity due to those earthquakes. Each of those inputs can be determined through scientific studies of past earthquakes and processing of observed data. The end result—the rate of exceeding IM levels of varying intensity—is very useful for engineering decision making, and can be determined even for rare (low exceedance-rate) intensities that are not possible to determine through direct observation. In the next section, we will perform some example calculations to demonstrate how this equation is used in practice.

1.4 Example PSHA calculations

To illustrate the procedure described in the previous section, several numerical examples will be performed below, starting from basic calculations and building to more complex cases. These examples will compute rates of exceeding varying levels of Peak Ground Acceleration, using the procedures described above.

1.4.1.1 Example: a source with a single magnitude and distance

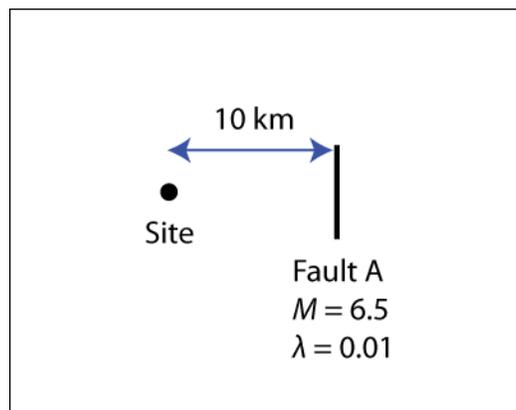


Figure 1.13: Map view of the example site, with one earthquake source.

We first start with a simple site shown in Figure 1.13. There is a single fault (Fault A) that produces only magnitude 6.5 earthquakes. We assume that this earthquake will rupture the entire fault,

so that the source-to-site distance is exactly 10 km in every case (that is, we will not consider random locations). Assume also that earthquakes with this magnitude and distance occur at a rate of $\lambda = 0.01$ times per year. Although the magnitudes and distances of all future earthquakes are fixed at 6.5 and 10 km, respectively, we still expect variations in observed peak ground accelerations at the site, due to differences from event to event that are not captured by our simple measures of magnitude and distance.

Using the Cornell et al. model presented in equation 1.14, we predict a median PGA of 0.3758g (i.e., a mean $\ln PGA$ of -0.979), and log standard deviation of 0.57. We can easily compute that the probability of exceeding 0.3758g given an earthquake is 0.5, since this is the median predicted value. The annual rate of exceeding 0.3758g is thus $0.5 * 0.01 = 0.005$ per year. This quick calculation is done to develop intuition regarding the calculations, but we can also use the more formal equations presented above. When considering equation 1.23, we see that $P(M = 6.5) = 1$, $P(R = 10) = 1$ and $\lambda(M_i > m_{\min}) = 0.01$. There is only one magnitude, distance and source to consider, so all of the summations are replaced by a single term. We thus have

$$\begin{aligned}\lambda(PGA > x) &= \lambda(M > m_{\min})P(PGA > x | 6.5, 10)P(M = 6.5)P(R = 10) \\ &= 0.01P(PGA > x | 6.5, 10)\end{aligned}\tag{1.24}$$

Next, since we know the mean and standard deviation of $\ln PGA$, we can compute the probability of exceeding a given PGA value using equation 1.15

$$P(PGA > x | 6.5, 10) = 1 - \Phi\left(\frac{\ln x - \overline{\ln PGA}}{\sigma_{\ln PGA}}\right) = 1 - \Phi\left(\frac{\ln x - \ln(0.3758)}{0.57}\right)\tag{1.25}$$

We can use Table 4.1 to evaluate the standard normal cumulative distribution function $\Phi(\)$. For example, the probability of $PGA > 1g$ is

$$P(PGA > 1g | 6.5, 10) = 1 - \Phi(1.72) = 0.044\tag{1.26}$$

Substituting this into equation 1.24, we can find the annual rate of exceeding 1g

$$\lambda(PGA > 1g) = 0.01P(PGA > 1g | 6.5, 10) = 0.00044\tag{1.27}$$

By repeating these calculations for many PGA levels, one can construct the curve shown in Figure 1.14. This “ground motion hazard curve” for PGA summarizes the rates of exceeding a variety of PGA levels. The two calculations performed explicitly above (for $PGA > 0.3758g$ and $PGA > 1g$) are labeled on this figure as well. Note that because both axes often cover several orders of magnitude, and the y axis contains very small values, one or both axes of ground motion hazard curves are often plotted in log scale.

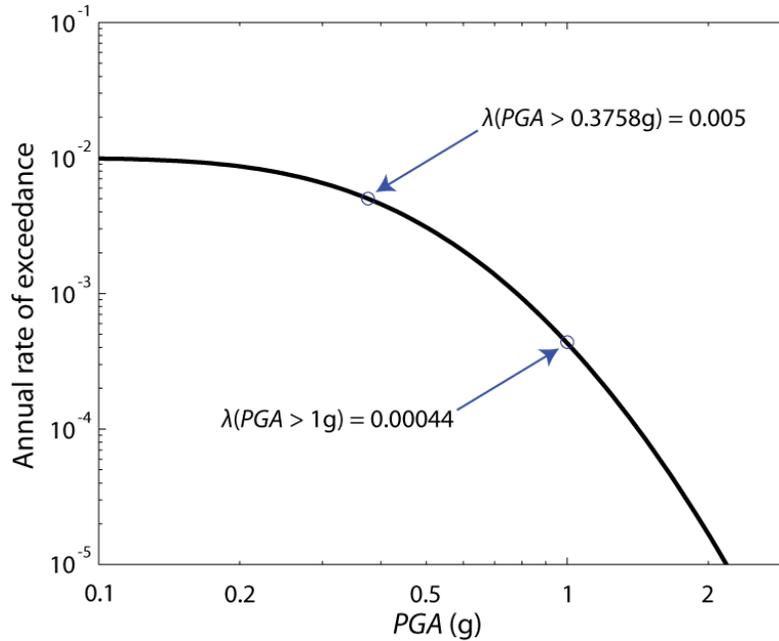


Figure 1.14: *PGA* hazard curve for the single-source site.

This example demonstrates the essence of a PSHA calculation. All that remains to perform a more realistic calculation is to consider a range of feasible magnitudes and distances, rather than just the single event in this hypothetical example.

1.4.1.2 Example: two magnitudes and distances

Before moving to an example with a continuous range of magnitudes, let us first try another hypothetical example with only two possible magnitude/distance combinations. The first source, “Fault A,” is identical to the source in the immediately preceding example. The second source, “Fault B,” produces only magnitude 7.5 earthquakes at a distance of 20 km. The earthquake on Fault B occurs at a rate of $\lambda = 0.002$ times per year. The map of this example site is shown in Figure 1.15. We will continue using the Cornell et al. model presented in equation 1.14 to predict *PGA* at the site.

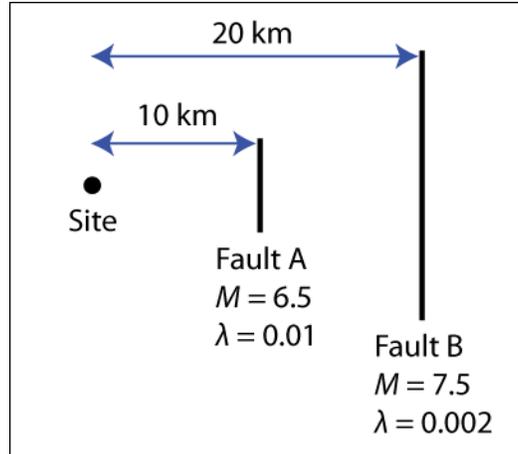


Figure 1.15: Map view of an example site with two earthquake sources.

The previous example to quantified the hazard from Fault A, so let us focus on calculating the hazard from Fault B. Using the Cornell et al. ground motion model, we predict a median PGA of $0.5639g$ if the earthquake on Fault B occurs, and a log standard deviation of 0.57 . Now consider the two PGA values considered in the previous example. The probability of $PGA > 0.3758g$, given an earthquake on Fault B, is

$$P(PGA > 0.3758g | 7.5, 20) = 1 - \Phi\left(\frac{\ln(0.3758) - \ln(0.5639)}{0.57}\right) = 1 - \Phi(-0.71) = 0.761 \quad (1.28)$$

We can then multiply this probability by the rate of occurrence of earthquakes on Fault B (0.02), to get the rate of $PGA > 0.3758g$ due to earthquakes on Fault B. But the PSHA formula of equation 1.23 includes a summation over all sources, so we add these probabilities to the corresponding probabilities for Fault A to find the overall rate of $PGA > 0.3758g$

$$\begin{aligned} \lambda(PGA > 0.3758g) &= \overbrace{0.01 P(PGA > 0.3758 | 6.5, 10)}^{(\text{Fault A})} + \overbrace{0.002 P(PGA > 0.3758 | 7.5, 20)}^{(\text{Fault B})} \\ &= 0.01(0.5) + 0.002(0.761) \\ &= 0.00500 + 0.00152 \\ &= 0.00652 \end{aligned} \quad (1.29)$$

Similarly, we can compute the probability of $PGA > 1g$, given an earthquake on Fault B

$$P(PGA > 1g | 7.5, 20) = 1 - \Phi\left(\frac{\ln(1) - \ln(0.5639)}{0.57}\right) = 1 - \Phi(1.01) = 0.158 \quad (1.30)$$

and then combine this with our other known information to compute the overall rate of $PGA > 1g$

$$\begin{aligned}
\lambda(PGA > 1g) &= \overbrace{0.01P(PGA > 1 | 6.5, 10)}^{(\text{Fault A})} + \overbrace{0.002P(PGA > 1 | 7.5, 20)}^{(\text{Fault B})} \\
&= 0.01(0.043) && + 0.002(0.158) \\
&= 0.000430 && + 0.000316 \\
&= 0.000746
\end{aligned} \tag{1.31}$$

The two rates computed above are plotted in Figure 1.16, along with rates for all other *PGA* levels. Also shown in Figure 1.16 are curves showing the rates of exceedance for the two individual faults. The intermediate rate calculations shown above (0.005 and 0.00152 for $PGA > 0.3758g$, and 0.0043 and 0.00316 for $PGA > 1g$) are also noted with circles on Figure 1.16 for reference. A few observations can be made regarding this figure and its associated calculations. First, note that the hazard curve for Fault A in the figure is identical to the curve in Figure 1.14. We have simply added the additional rates of exceedance due to Fault B in order to get the total hazard curve shown in Figure 1.16. Second, we can note that the relative contributions of the two faults to the ground motion hazard vary depending upon the *PGA* level of interest. At relatively lower *PGA* values such as in the calculation of equation 1.29, Fault A contributes much more to the overall rate of exceedance. This is because it has a higher rate of earthquakes. At larger *PGA* levels such as in the calculation of equation 1.31, the relative contributions of the two faults are close to equal; this is because the smaller-magnitude earthquakes from Fault A have a low probability of causing high *PGA* values, even though they are more frequent than the larger-magnitude earthquakes from Fault B. We see in Figure 1.16 that for *PGA* values greater than about 1.5g, Fault B actually makes a greater contribution to the hazard than Fault A, even though its rate of producing earthquakes is only 1/5 of Fault A's. This is a typical situation for real-life PSHA calculations as well: low-intensity shaking is generally dominated by frequent small earthquakes, while high-intensity shaking is caused primarily by large-magnitude rare earthquakes.

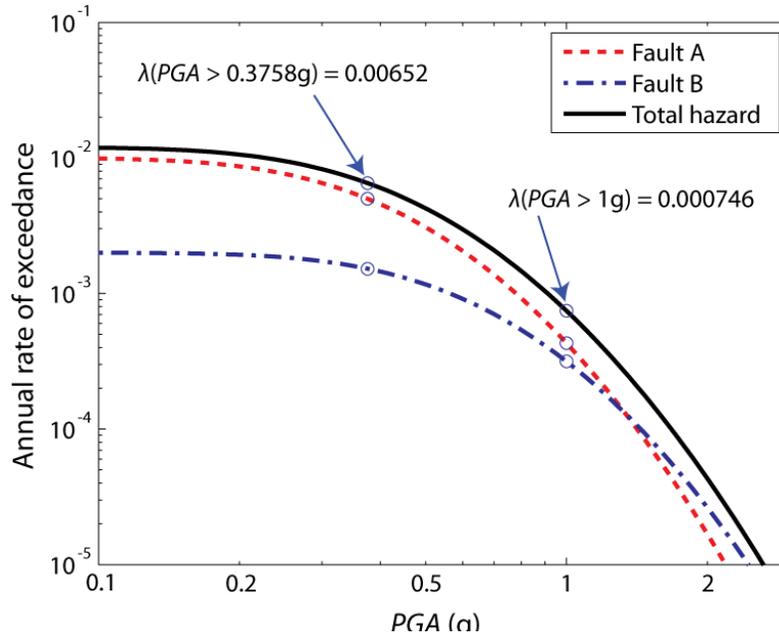


Figure 1.16: *PGA* hazard curve for the example two-source site.

1.4.1.3 Example: point source with Gutenberg-Richter magnitudes

In this example we will now consider a source that is capable of producing earthquakes with a variety of magnitudes. The source produces events with $M \geq 5$ at a rate of 0.02 events per year. The distribution of those earthquakes follows the bounded Gutenberg-Richter model, with a maximum magnitude of 8 and “ b ” value of 1 (the b parameter in equation 1.5). We will thus use equation 1.5 to describe the PDF of earthquake magnitudes, with $m_{min} = 5$ and $m_{max} = 8$. We again assume that all earthquakes occur at the same distance of 10 km, so that we can simplify the PSHA summations.

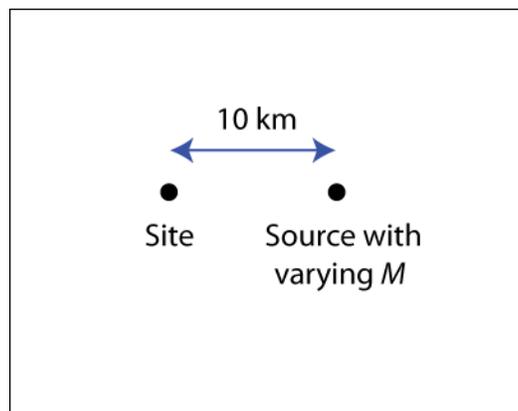


Figure 1.17: Map view of an example site with one source producing earthquakes having a variety of magnitudes.

We will use equation 1.23 to perform the PSHA calculation for peak ground acceleration, using the Cornell et al. ground motion model from the previous examples. Noting that there is only one source and one distance, and substituting $P(R=10)=1$ and $\lambda(M_i > m_{\min})=0.02$, we get

$$\lambda(PGA > x) = 0.02 \sum_{j=1}^{n_M} P(PGA > x | m_j, 10) P(M = m_j) \quad (1.32)$$

To compute this rate of exceeding some PGA level, we simply need to compute the probabilities of observing various earthquake magnitudes, compute the probabilities of exceeding our PGA level given those magnitudes, and then sum the product of those two terms evaluated for the range of feasible magnitudes. Table 1.3 shows those probabilities for calculations of $\lambda(PGA > 0.2g)$. The first column lists the discrete set of magnitudes considered (a magnitude increment of 0.25 has been used). The second column lists the probabilities of observing those magnitudes, as computed using equation 1.6 (note these probabilities are identical to those in Table 1.1, because the assumed magnitude distribution is the same). The third column lists the probability of $PGA > 0.2g$, given occurrence of an earthquake having the specified magnitude. This is computed, as was done in the earlier examples, by evaluating the Cornell et al. ground motion model using each magnitude value. The fourth column lists the products of the second and third columns. We see that equation 1.32 is simply a summation of the terms in the fourth column, multiplied by the rate of occurrence of earthquakes. The sum of the fourth column is 0.269, so the rate of $PGA > 0.2g$ at the site of interest is $\lambda(PGA > 0.2) = 0.02(0.269) = 0.0054$.

Table 1.3: Probabilities used to compute $\lambda(PGA > 0.2g)$.

m_j	$P(M = m_j)$	$P(PGA > 0.2 m_j, 10)$	$P(PGA > 0.2 m_j, 10) \cdot P(M = m_j)$
5.00	0.4381	0.1242	0.0544
5.25	0.2464	0.2185	0.0538
5.50	0.1385	0.3443	0.0477
5.75	0.0779	0.4905	0.0382
6.00	0.0438	0.6379	0.0279
6.25	0.0246	0.7672	0.0189
6.50	0.0139	0.8657	0.0120
6.75	0.0078	0.9310	0.0073
7.00	0.0044	0.9686	0.0042
7.25	0.0025	0.9873	0.0024
7.50	0.0014	0.9955	0.0014
7.75	0.0008	0.9986	0.0008
			Sum = 0.269

We continue the hazard analysis by repeating the calculations of Table 1.3 for more *PGA* values. In Table 1.4, the same calculation is repeated for $PGA > 1g$. Here we see that the summation of the right-hand column is 0.0048, so the rate of $PGA > 1g$ at the site of interest is $\lambda(PGA > 1) = 0.02(0.0048) = 9.6 \cdot 10^{-5}$.

Table 1.4: Probabilities used to compute $\lambda(PGA > 1g)$.

m_j	$P(M = m_j)$	$P(PGA > 1 m_j, 10)$	$P(PGA > 1 m_j, 10) \cdot P(M = m_j)$
5.00	0.4381	0.0000	0.0000
5.25	0.2464	0.0002	0.0000
5.50	0.1385	0.0006	0.0001
5.75	0.0779	0.0022	0.0002
6.00	0.0438	0.0067	0.0003
6.25	0.0246	0.0181	0.0004
6.50	0.0139	0.0430	0.0006
6.75	0.0078	0.0901	0.0007
7.00	0.0044	0.1676	0.0007
7.25	0.0025	0.2786	0.0007
7.50	0.0014	0.4168	0.0006
7.75	0.0008	0.5662	0.0004
			Sum = 0.0048

By repeating this calculation for many more *PGA* values, we can create the *ground motion hazard curve* shown in Figure 1.18. The two individual rates of exceedance calculated above are labeled on this curve.

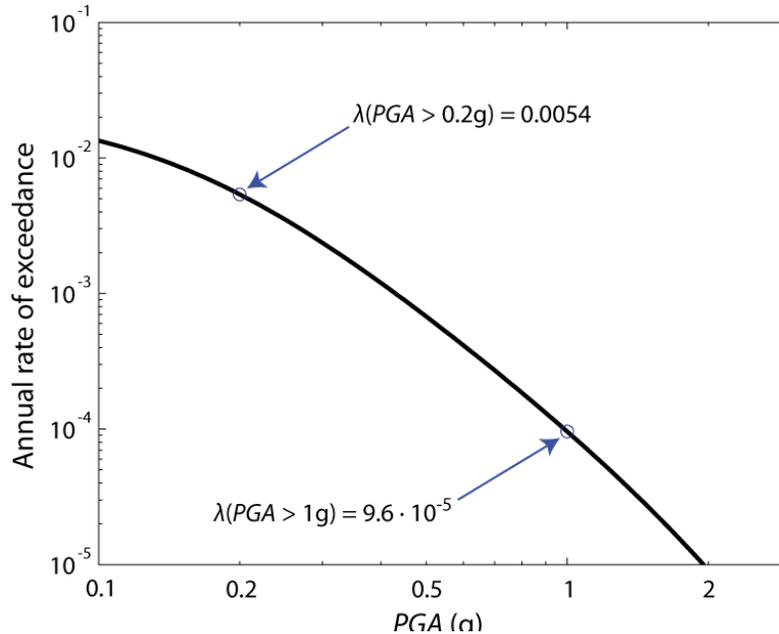


Figure 1.18: *PGA* hazard curve for the example site with one source and a Gutenberg-Richter magnitude distribution.

Comparing Table 1.3 to Table 1.4, we can make several observations. The first two columns of both tables are identical, as they are only describing the earthquake magnitudes and so are not affected by changes in the *PGA* level of interest. We see that all probabilities in the third column are much larger in Table 1.3 than in Table 1.4: the *PGA* threshold was lower in Table 1.3, so the probability of exceeding the threshold is therefore higher.

In Table 1.4, the probability of $PGA > 1g$ is equal to zero for the smallest magnitude considered. This means that considering even smaller magnitudes would have no impact on our final answer, because smaller magnitude earthquakes have effectively zero probability of causing a *PGA* greater than $1g$. In Table 1.3, however, we see that even magnitude 5 earthquakes have a non-zero probability of causing $PGA > 0.1g$; this is somewhat worrisome because lower magnitude earthquakes could also cause $PGA > 0.1g$; this is somewhat worrisome because lower magnitude earthquakes could also cause $PGA > 0.1g$, so including them would have changed our answer. This suggests that the choice of the minimum considered earthquake can be important in some cases. We will return to this issue in Section 2.2.

By looking at the right-hand column of these tables, we can also identify the magnitudes that make the greatest contribution to the probability of exceeding the *PGA* level of interest. Each number in this column is a product of the probability of occurrence of some magnitude and the probability of exceedance of the *PGA* given that magnitude, which is equal to the probability of both events occurring (given that an earthquake has occurred). In Table 1.3, we see that the probabilities are largest for the small magnitudes; this is because small-magnitude earthquakes are much more likely to

occur than large-magnitude earthquakes (as seen in column 2), *and* because these small-magnitude earthquakes have a reasonable probability of causing $PGA > 0.1g$. In Table 1.4, on the other hand, these small magnitude earthquakes have very small probabilities in the fourth column because they have a very small probability of causing $PGA > 1g$. In Table 1.4, the moderate- to large-magnitude earthquakes have the highest probabilities in the fourth column, because even though they are relatively rare, they are the only ones with significant probabilities of causing $PGA > 1g$. Intuitively, we can imagine that this information would be useful for identifying the earthquake scenarios most likely to damage a structure at the site of interest. We will revisit this information in a more quantitative manner in Section 2.1 below.

For more complex sites than the simple cases shown in the above examples, the PSHA summations can quickly get lengthy. For this reason, PSHA is performed using computer software in all practical analysis cases. The software's purpose is to perform the calculations shown here, but for more complicated cases involving many earthquake sources, while also using modern ground motion prediction models that are much more complex than the one used here. Note that in the example above we used a relatively wide magnitude spacing of 0.25 units in our discretization in order to keep the length of Table 1.3 reasonably short. When performing these calculations in a computer program, it is also easy to use a fine discretization of the magnitudes and distances of interest.