

Statistical Features of Short-Period and Long-Period Near-Source Ground Motions

Masumi Yamada, Anna H. Olsen, and Thomas H. Heaton

Abstract:

This study collects recorded ground motions from the near-source region of large earthquakes and considers to what extent this historic record can inform expectations of future ground motions at similar sites. The distribution of observed peak ground acceleration (PGA) is well approximated by the lognormal distribution, and we expect the observed distribution to remain unchanged with the addition of data from future earthquakes. However, the distribution of peak ground displacements (PGD) will likely change after a well recorded large earthquake. Specifically we expect future observations of PGD greater than those previously recorded. We use seismic scaling relations to motivate the expected distribution of PGD as uniform on the logarithmic scale, or at least fat-tailed. Since PGA does not scale with fault rupture area or slip on the fault, there are no such scaling relations to predict the observed distribution of PGA. The observed records show that there is essentially no correlation between PGD and PGA for near-source ground motions from large events. The large uncertainty in a future value of PGD in the near-source region of a large earthquake exists despite the ability of earth scientists to accurately model long-period ground motions. In contrast, the relative certainty in a future value of PGA exists despite the inability to model short-period ground motions reliably. The stability of the observed distribution of PGA with respect to new ground motion records enables us to predict the distribution of future PGA and to calculate the probability of exceeding the largest recorded PGA.

1 Introduction

The current practice of seismic structural design assumes an expectation of future ground motions. An engineer first defines the likely and most severe ground motions that may excite a structure at a

particular site in its anticipated lifetime. Then the engineer specifies a lateral force resisting system with a capacity sufficient to withstand these expected demands. This approach to design problems works well if the definition of future ground motions is reliable. Unfortunately, predicting the future is a notoriously difficult endeavor.

Past experience primarily informs the seismic design of structures. ASCE 7-05, the standard adopted by the 2006 International Building Code for seismic design, requires the use of at least three recorded ground motion time histories for structural designs based on seismic response histories (American Society of Civil Engineers, 2006). If there are fewer than three appropriate recorded ground motions, simulated motions may be used to meet the requirement. Also, structural engineers develop their judgments through their own experiences and the experiences of their predecessors. Experience and the historic record may be necessary, but not sufficient, evidence for anticipating the future ground motions at a given site. This study considers the extent to which recorded ground motions can inform the prediction of future ground motions.

Throughout this study we distinguish ground motions dominated by short-period versus long-period energy content. All seismic ground motions can be decomposed according to the relative amounts of energy in harmonic waves of every period. Generally speaking, the energy in ground motions from smaller, more frequent earthquakes is mostly from the short-period content, whereas the energy in ground motions from larger, less frequent earthquakes is primarily in the long-period content. This distinction between short- and long-period energy also categorizes a structure as short- or long-period according to the structure's fundamental period. The lateral force-resisting system of a structure with a short fundamental period tends to be susceptible to short-period ground motions. Similarly, a structural system with a long fundamental period is particularly sensitive to long-period ground motions.

This study collects recorded ground motions from the near-source region of large earthquakes and considers to what extent this historic record can inform expectations of future ground motions at similar sites. We find that future peak ground accelerations (a good intensity measure for short-period ground motions) can be readily predicted from existing records, but the experience of past large earthquakes does not adequately predict future peak ground displacements (a good intensity measure for long-period ground motions). These observations directly affect the design of structures, since building codes rely explicitly or implicitly on recorded ground motions: existing measures of peak ground acceleration from records are a reliable predictor of future peak ground accelerations, but future peak ground displacements may differ significantly from measures of peak ground displacement from existing records.

2 Data

We collect ground motions from ten historic earthquakes to generate distributions of observed peak ground acceleration and displacement. Specifically, we use ground motions only from earthquakes with magnitudes greater than 6.0 and from sites in the near-source. We define the near-source region to be within 10 km of the surface projection of the fault rupture. We choose this number because it is small enough to demonstrate the source scaling, but large enough to include a sufficient number of records to characterize the statistics. We disregard the soil class at the instrumented sites.

Table 1 lists important features of the data set. The earthquakes include the major events in the years 1979 through 2004. The data set covers a range in moment magnitude from 6.0 to 7.8, with wide gaps between 6.0 and 6.5 and between 6.9 and 7.3. Nonetheless, the data well represent the ground motions of interest. The Chi-Chi and Parkfield earthquakes are well recorded, each with over forty near-source records. (In fact, the Chi-Chi strong ground motion data set is the largest for a major earthquake (Shin and Teng, 2001).) However, the Landers, Hyogoken-nanbu, Izmit, and Denali ruptures are poorly recorded, since each has fewer than five near-source records. Table 1 also lists the focal depth of each hypocenter, ranging from 12.0 km to 21.2 km, as well as the fault rupture model used to define the near-source area.

We obtained ground acceleration time histories from several databases around the world. The COSMOS Virtual Data Center, which includes the California Strong Motion Instrumentation Program and the United States Geological Survey seismic networks, was the source for ground motions from the Imperial Valley, Loma Prieta, Landers, Northridge, Denali, and Parkfield earthquakes. The ground motions for the Hyogoken-nanbu earthquake were from: the Japan Meteorological Agency; the Committee of Earthquake Observation and Research in the Kansai Area in Japan (Toki et al., 1995); and the Japan Railway Institute, whose records were scanned and digitized by Wald (1996). (Seismometers in the CEORKA network record velocity, but CEORKA processes the records to provide acceleration time histories.) The national strong-motion network in Turkey recorded ground motions in the Izmit earthquake (Akkar and Gülkan, 2002), and we collected them from two sources: the Earthquake Research Department of the General Directorate of Disaster Affairs in Turkey (Earthquake Research Department, 2004); and the COSMOS Virtual Data Center, which archived records from the National Earthquake Monitoring Center, Kandilli Observatory and from the Earthquake Research Institute, Bögaziçi University. The Central Weather Bureau Seismic Network in Taiwan recorded the ground motions for the Chi-Chi earthquake, and Lee et al. (2001) disseminated them. The K-NET (Kyoshin Net, 1996) and KiK-net (KiK-net, 2000) seismic networks, operated by the National Research Institute for Earth Science and Disaster Prevention in Japan, were our sources for the strong ground motions of the

Niigataken-chuetsu earthquake.

The geographic distributions of instrumented sites in the near-source regions are not uniform. The density of sites per near-source area is not constant for all events, and the sites do not evenly represent all near-source regions. For nine of the ten earthquakes used in this study, Yamada et al. (2007) showed the sites in comparison to the fault rupture surface projection and the near-source region. As examples of particularly poor distributions of instrumented sites (for our purposes), many of the ground motions from the 1979 Imperial Valley earthquake are from sites forming a line perpendicular to the fault, and most ground motions from the 1999 Chi-Chi earthquake are from sites on the foot wall. Despite the non-uniform geographic distribution of the recorded ground motions, we treat each record as a sample of the ground movement in the near-source region.

To determine the peak ground acceleration and displacement of a recorded ground motion, we process each time history. We first remove the bias from the acceleration record by subtracting the mean of the pre-event interval. Then we integrate twice to generate the displacement time history. We remove the longest period content of the displacement time history with a fourth-order Butterworth filter (13.3 s corner period) to avoid problems caused by a shift in the baseline. Filtering in this way can produce a peak ground displacement smaller than what would have been calculated from the unfiltered record. Nonetheless, these filtered records still contain rich long-period content. We then find the square root of the sum of the squares of the north-south and east-west components at each time step of the record. The peak ground measure is the largest such value over all time steps. This measure of peak ground motion is also known as the vector amplitude (Kanno et al., 2006). Yamada (2007) compared PGAs and PGDs calculated from the horizontal components or from the vertical component and found that they have a similar trend with respect to moment magnitude. We use only the horizontal components to calculate the peak values since most engineered structures are more vulnerable to lateral than vertical excitation.

In addition to the recorded ground motions, we also consider a recurrence of the 1906 San Francisco earthquake. We use the simulated ground motions generated by Aagaard et al. (2008) closest to sites in the California Integrated Seismic Network as of December 2007. To be consistent with the choice of recorded ground motions, we only include sites within 10 km of the historic rupture on the San Andreas fault. These selection criteria result in 87 “records” of a recurrence of the 1906 San Francisco earthquake. These ground motions are long period (periods greater than 2 s), so we use a fourth-order Butterworth filter (25.0 s corner period) and calculate only the peak ground displacement of each record.

3 Observed Distributions of PGA and PGD

For our collected set of recorded ground motions, Figure 1 plots the peak ground accelerations and displacements as functions of the moment magnitude of the generating earthquake. Each data point is denoted with a symbol, and the least squares regression lines for the peak ground measures as functions of magnitude are shown. At each magnitude, the peak ground acceleration (PGA) and peak ground displacement (PGD) data are approximately lognormally distributed about their respective regression lines. Note that PGA saturates at magnitudes greater than 6.0, as evidenced by the approximately zero slope of the regression line. The saturation of short-period ground motions in the near-source has been observed previously by Rogers and Perkins (1996) and Somerville (2003). In contrast, the logarithm of PGD increases linearly with respect to magnitude, with a slope of 0.6. Previous studies have discussed this linear trend (e.g., Wells and Coppersmith, 1994; Campbell and Bozorgnia, 2008). We expect this slope to be in the range of 0.5 to 0.75 based on the following argument.

Given the definition of seismic moment, M_0 , ($\log M_0 = \log \mu + \log S\bar{D}$) and the relationship between seismic moment and moment magnitude, M_w , ($\log M_0 = \frac{3}{2}(M_w + 10.73)$), we find:

$$\log S\bar{D} = \frac{3}{2}(M_w + 10.73) - \log \mu, \quad (1)$$

where S is the rupture area, \bar{D} is the average slip, and μ is the crustal rigidity. If we assume the rupture area is approximately the square of the rupture length (reasonable for small to moderate earthquakes (Aki, 1972; Kanamori and Anderson, 1975)) and PGD is proportional to the average slip, Equation 1 implies that:

$$\log \text{PGD} = \frac{1}{2}M_w + \text{a constant.}$$

If we assume the rupture area is proportional to the rupture length (reasonable for large earthquakes, (Scholz, 1982)) and PGD is proportional to the average slip again, Equation 1 then implies that:

$$\log \text{PGD} = \frac{3}{4}M_w + \text{another constant.}$$

Thus we expect the slope of the relationship between the logarithm of PGD and moment magnitude to be between 0.5 and 0.75 for the combination of moderate and large earthquakes.

As mentioned previously, PGA is known to saturate at magnitudes greater than 6, and the logarithm of PGD is known to increase linearly with respect to magnitude. Thus, there is essentially no correlation between these two intensity measures for near-source ground motions from large events. Figure 2 plots PGA versus PGD for the near-source records used in this study and for additional far-source records. The far-source PGDs and PGAs are from ground motions recorded

in magnitude greater than 6 earthquakes at distances greater than 10 km from the fault. Unlike the far-source records, the near-source PGD and PGA are very weakly correlated.

Figure 3 compares the marginal distributions of PGA and PGD to a base-10 lognormal distribution. The histogram of the PGA data is consistent with a lognormal distribution: on a logarithmic scale, the PGA data are clearly peaked, symmetric about the mean, and have thin tails. This observation is consistent with Figure 1 since the mean PGA is approximately constant for magnitudes greater than 6.0, and the PGA data appear to be lognormally distributed about the mean with constant variance. The near-source PGA of earthquakes with magnitude greater than 6.0 can be approximated by the lognormal distribution:

$$p(PGA) \propto \frac{1}{0.56PGA\sqrt{2\pi}} e^{-(\ln PGA - \ln(4.64))^2/0.64}, \quad (2)$$

where \ln represents the natural logarithm. The histogram of the PGD data, however, is inconsistent with a lognormal distribution: on a logarithmic scale, the PGD data are roughly uniform over the range of observed PGD. The PGD data are not peaked because the mean PGD is correlated with magnitude. At a given magnitude, the PGD data appear lognormally distributed, but since the mean increases with magnitude, the projection of the data as a marginal distribution becomes more uniform than peaked.

Figures 4 and 5 are histograms of the PGA and PGD data, respectively. For the PGA data, we show a subset of the observed data with the Chi-Chi data removed as well as the full data set. The distribution of the subset without the Chi-Chi data appears lognormal. The addition of the Chi-Chi data does not change the shape of the distribution; including the Chi-Chi PGA data simply increases the height of the histogram. This exercise suggests that a lognormal distribution is a good approximation for existing, recorded PGA at near-source sites in earthquakes with magnitudes greater than 6.0, and this distribution will remain approximately lognormal with the addition of future data. The saturation of PGA at large magnitudes supports our claim: since the mean and standard deviation of PGA are approximately constant for each large earthquake, data from future earthquakes will also have the same mean and standard deviation and thus cannot change the shape of the PGA distribution.

Now consider the observed distribution of PGD in Figure 5. Similar to the analysis of the PGA data, we show the distribution of PGD for a subset of our recorded data set that does not include ground motions from the Chi-Chi earthquake. This distribution is roughly uniform on a logarithmic scale over the range of this subset. Adding the PGDs from the Chi-Chi earthquake extends the range of PGD at large values; there are several PGDs from the Chi-Chi ground motions that are larger than observations from other large earthquakes. The spread in the distribution of observed near-source PGD widens notably with the inclusion of this single earthquake.

We use a simulation of ground motions in the 1906 San Francisco earthquake to imagine the effect of the next large, well-recorded earthquake on this distribution of PGD. If the 1906 earthquake recurred and was recorded with existing instrumentation, Figure 5 shows the effect on the “observed” distribution. The simulated data add to the histogram at the largest observed PGDs. In other words, many of the simulated PGDs are in the range of the largest recorded PGDs. Certainly the next large, well-recorded earthquake will not be a repeat of the 1906 San Francisco earthquake, but a future rupture has the potential to produce numerous recorded ground motions with PGDs at the high end of, or greater than, those recorded at present.

Earthquakes with magnitudes greater than 6.0 change the observed distribution of PGD, either by extending the distribution or by adding data at previously recorded values. The increase of PGD with magnitude (Figure 1) helps to explain this point. Although PGD is lognormally distributed at a given magnitude, the mean of the logarithm of PGD increases linearly with magnitude. Thus, the magnitude of each event determines where its observed PGDs add to the distribution of PGD from all events, either beyond the observed range or within it. This finding is in contrast to our statements about the distribution of PGA: records from future, well-recorded earthquakes will not change the observed distribution of PGA, whereas they may change—and many change substantially—the observed distribution of PGD.

4 Theoretical Distributions of PGA and PGD

In this section we first use empirical scaling relationships to derive the expected marginal distribution of PGD in the near-source region of large earthquakes. First, we assume that the number of earthquakes N with magnitudes greater than M is given by the Gutenberg-Richter relationship:

$$N = 10^{a-bM} = \int_M^{\infty} N_M dM, \quad (3)$$

where N_M is the number of earthquakes with magnitude between M and $M + \Delta M$. In a different form,

$$N_M = -\frac{dN}{dM} = b \ln(10) 10^{a-bM} = 10^{A-bM}, \quad (4)$$

where $10^A \triangleq b \ln(10) 10^a$, a constant. The moment magnitude can be rewritten in terms of the seismic moment, M_0 , which is equal to the product of the crustal rigidity, μ , average slip on the fault, \bar{D} , and the area of fault rupture, S :

$$M_w = \frac{2}{3} \log_{10} M_0 - 10.73 = \frac{2}{3} \log_{10} (\mu \bar{D} S) - 10.73.$$

Thus,

$$\begin{aligned} N_M &= 10^{A+10.73b} 10^{-\frac{2}{3}b \log_{10}(\mu \bar{D} S)} \\ &= C (\bar{D} S)^{-\frac{2}{3}b}, \end{aligned}$$

where $C \triangleq 10^{A+10.73b} \mu^{-\frac{2}{3}b}$, a constant.

One of two assumptions can be made at this point, depending on the size of the earthquake. For moderate earthquakes, the fault rupture length, L , is approximately equal to the fault rupture width, which implies that the fault rupture area is approximately L^2 . For large earthquakes, the rupture length is much larger than the approximately constant rupture width, which implies that the rupture area is proportional to the rupture length. We also assume in both cases that the average slip on the fault is proportional to the length of the fault rupture. Thus,

$$N_M \propto \begin{cases} (\overline{DL}^2)^{-\frac{2}{3}b} & \text{if } S \approx L^2 \text{ (moderate earthquakes)} \\ (\overline{DL})^{-\frac{2}{3}b} & \text{if } S \propto L \text{ (large earthquakes)} \end{cases}$$

$$\propto \begin{cases} \overline{D}^{-2b} \\ \overline{D}^{-\frac{4}{3}b}. \end{cases}$$

The total rupture area for all earthquakes of magnitude M is:

$$N_M S \begin{cases} \approx N_M L^2 \propto \overline{D}^{-2b} \overline{D}^2 \stackrel{b=1}{=} \text{constant} \\ \propto N_M L \propto \overline{D}^{-\frac{4}{3}b} \overline{D} \stackrel{b=1}{=} \overline{D}^{-\frac{1}{3}}. \end{cases} \quad (5)$$

Assuming a uniform geographic distribution of seismic instrumentation, Equation 5 implies that the total number of near-source records at a particular magnitude is either constant or weakly dependent on magnitude, depending on the assumed geometry of the fault rupture. Furthermore, the peak ground displacement in the near-source is proportional to the average slip on the fault (Aagaard et al., 2001). These results can be combined with the observations from Figure 1 to generate a rough approximation of the marginal probability density function for the logarithm of PGD.

If we assume the rupture area is approximately square (that is, consider moderate earthquakes), then the total number of near-source records is constant for all magnitudes. Given enough time, all magnitude 6 earthquakes will produce the same number of near-source records as all magnitude 7 earthquakes or all magnitude 8 earthquakes. Figure 6 (top) combines this result with the observations that the mean PGD increases with magnitude and PGD is lognormally distributed at a given magnitude. The resulting marginal distribution of PGD is approximately uniform because the total number of recorded ground motions is constant for all magnitudes.

Alternatively, if we assume the rupture length is much larger than the rupture width (that is, consider large earthquakes), then the total number of near-source records decreases with magnitude, since magnitude depends on average slip. Figure 6 (bottom) shows the relative number of near-source ground motions at each magnitude given that the total number is proportional to the mean PGD raised to the negative one-third power. The resulting marginal distribution of PGD is not uniform, but it is also not lognormal; it is something in-between the two.

We cannot develop a similar argument for the marginal distribution of PGA because the relationship between PGA and faulting is not clear. At this time, seismologists are developing finite element models to simulate broadband ground motions and thus predict PGA in the near-source region. However, these models require sophisticated fault and velocity models to propagate short-period seismic waves or require a stochastic description of PGA. Unlike PGD, which can be related to slip on the fault, it is difficult to use a fundamental, physical understanding of faulting and wave propagation to predict PGA. Yamada et al. (2007) shows that the distribution for near-source PGA seems to be compatible with the hypothesis that the total radiation of high-frequency energy scales with the rupture area. This hypothesis is consistent with the observation that the PGA weakly depends on the slip on the fault.

5 Expectation of the Largest PGA and PGD

An important question for the design of an engineered structure is: what is the largest expected ground motion at the site of interest? This question can be reformulated as: what are the largest expected PGA and PGD at a site within 10 km of a fault capable of producing at least a magnitude 6 earthquake? The analyses of observed PGAs and PGDs presented in this paper should inform the answer to this question. We have shown that the empirical distribution of PGA is approximately lognormal and stable with respect to new data, but at present, there is no theoretical basis to explain this observation. The empirical marginal distribution of PGD changes after well-recorded earthquakes, and seismic scaling relations suggest that the marginal distribution of PGD is roughly uniform on a logarithmic scale. Thus, we can readily address the question of the largest expected PGA, but determining the largest expected PGD is not clear.

We consider the historic deployment of seismic instrumentation to place recorded peak ground measures in context. Figure 7 shows the history of the largest recorded PGA as well as the estimated number of seismic stations in the United States and Japan. (In this section, we calculate PGA from three orthogonal components of ground motion instead of two horizontal components, as calculated in all other sections.) We collect this information from several sources. Çelebi et al. (1999) report the largest recorded PGA until 1999. Trifunac and Todorovska (2001) tally the number of seismic stations before 1982. We estimate the current number of seismic stations in the United States from the United States Geological Survey and California Strong Motion Instrumentation Program websites. The current number of seismic stations in Japan is the sum of K-NET (Kyoshin Net, 1996), KiK-net (KiK-net, 2000), and Japan Meteorological Agency instrumentation. After an interval of 23 years, the largest recorded PGA increased by about a factor of two in 2008 (Aoi et al., 2008). A larger number of seismic instrumentation makes a record of PGA in excess of one previously recorded more likely. However, the empirical marginal distribution of

PGA suggests that it will become increasingly more difficult to record ever larger PGAs, as such extreme PGAs are vanishingly unlikely.

To quantify the probability of observing a PGA in excess of the current largest value, we account for an increase of instrumentation and the observed distribution of PGAs. Following Section 3, we assume that the lognormal distribution with mean m and standard deviation σ approximates the future distribution of PGA. Thus, the probability that one measure of PGA is less than x is given by the cumulative distribution function:

$$F(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln(x) - m}{\sigma\sqrt{2}} \right]. \quad (6)$$

We assume that records of PGAs in the near-source of major earthquakes are independent with respect to space and time. Therefore, if there are n_y independent measures of PGA in year y , then the probability that none exceeds x is $F(x)^{n_y}$, and the probability that at least one measure in the next Y years exceeds x is:

$$P(\text{PGA} > x) = 1 - \prod_{y=1}^Y F(x)^{n_y}. \quad (7)$$

The number of recorded PGAs is the product of the station density, the near-source area, and the number of earthquakes in a given year. We assume the station density, $\rho(y)$, is a function of the year but uniform in space. We also assume that the fault rupture area approximates the near-source area, and we use the empirical relationship from Irikura and Miyake (2001):

$$S = 10^{0.75M - 2.29}, \quad (8)$$

where S has a unit of km^2 . The Gutenberg-Richter frequency-magnitude relationship defines the expected number of earthquakes in any time period, taken here as one year. Therefore:

$$n_y = \rho(y) \int_{M_L}^{M_U} SN_M dM \quad (9)$$

$$= \rho(y) 10^{A-2.29} \int_{M_L}^{M_U} 10^{(0.75-b)M} dM \quad (10)$$

$$= \frac{\rho(y) 10^{A-2.29}}{(0.75-b) \ln(10)} \left[10^{(0.75-b)M} \right]_{M_L}^{M_U}, \quad (11)$$

where M_U and M_L are the upper and lower bounds of the considered range of magnitudes. The probability that a recorded PGA will exceed a value x in the next Y years can be calculated from Equation 7, with substitutions of Equations 8 and 9.

We calculate the probabilities of recording a PGA greater than 20 m/s^2 (approximately the pre-2008 largest PGA) and greater than 40 m/s^2 (approximately the current largest recorded PGA). We

let $m = 1.53$ and $\sigma = 0.57$, based on the empirical distribution of PGA (Figure 3). We assume the station density increases by 5% of ρ_0 per year, so $\rho(y) = \rho_0(1.05 * y)$. Since the current station density in Japan is roughly spaced at 15 km, we estimate the current station density as $\rho_0 = 0.004\text{stations}/\text{km}^2$. We let $A = 7.0$, as determined from the seismic catalog of Japanese crustal earthquakes in 2008, and $b = 1.0$, since b is typically close to this value. We consider earthquakes on the interval $M_L = 6$ to $M_U = 8.5$.

Figure 8 shows the probabilities that a recorded PGA exceeds 20 m/s^2 or 40 m/s^2 in the next 0 to 100 years. Given that another recorded PGA in excess of 20 m/s^2 is more than likely in the next 30 years, it is not surprising that there are currently two such records from the past seventy-five years of ground motion observations. However, the probability of a future PGA greater than 40 m/s^2 is only 10% in the next fifty years, assuming a steady increase in station density of 5% of ρ_0 per year. Aoi et al. (2008) and Yamada et al. (2009) proposed a physical mechanism for extremely large vertical accelerations. This may explain future records with PGAs much larger than the current, largest recorded PGA. Since we have shown in the previous section that the empirical distribution of PGA is stable with respect to new data and well characterized by a lognormal distribution, we can use this distribution to quantify the probabilities that future records of PGA will exceed the previous and current largest recorded values. This type of analysis could be used to predict reliably future large PGAs in the near-source of active faults capable of producing at least a magnitude 6 earthquake.

A similar analysis cannot be developed at present for PGD. We have shown that the empirical marginal distribution of PGD changes with well-recorded earthquakes. Therefore, this distribution alone is inadequate to predict future PGDs. Furthermore, seismic scaling relations suggest that very large PGDs are roughly as numerous as large PGDs; the distribution of PGD is fat-tailed. We expect that the probability of recording large PGDs—for example, in excess of 5 m or 10 m—in the next fifty years is not negligible. Nonetheless, there must be an upper-bound to the marginal distribution of PGD based on a physical limit to PGD. We do not propose a value for this physical limit.

6 A Paradox

This discussion of predicting PGA and PGD from the fundamentals of faulting and wave propagation leads directly to a paradox: the PGA in a future earthquake can be easily predicted from past earthquakes even though short-period ground motions cannot be easily simulated, whereas the PGD in a future earthquake cannot be predicted even though long-period ground motions are routinely simulated. Given only the information that a site is in the near-source of an earthquake with

magnitude greater than 6.0, the recorded ground motions used in this study show that PGA is well approximated by a lognormal distribution, but the distribution of the logarithm of PGD is roughly uniform with an unknown point of truncation. Thus, using past observations of ground motions can produce a relatively certain prediction of PGA and a quite uncertain prediction of PGD. This is in contrast to the present capabilities of simulating ground motions. Long-period ground motions can be readily generated from models of faults and seismic velocity structures. Currently, however, generating short-period ground motions is not routinely done as it requires much more detailed models or stochastic descriptions.

The resolution of this paradox considers the assumed information. Simulating ground motions assumes a magnitude and either generates a slip distribution (dynamic source modeling) or assumes a slip distribution (kinematic source modeling). Our study of recorded ground motions assumes only that the earthquake has a magnitude greater than 6.0. Knowing the magnitude does not help predict the PGA since PGA saturates with magnitude, but knowing the magnitude helps to predict PGD since the logarithm of PGD is proportional to the magnitude. However, we need to predict more than the magnitude of future events; the near-source PGD is best predicted by slip on a nearby segment of a fault. Thus, the ability of seismologists to generate reliable long-period ground motions (and thus reliable predictions of PGD) in an earthquake of known magnitude does not imply that they can predict the PGD in the next large-magnitude earthquake, unless they can also reliably predict slip.

7 Implications for Seismic Risk

Engineers often use ground motion intensity measures in place of full time histories to predict seismic structural responses. Although the interaction of the complete ground motion time history with the structure defines the response, a characteristic of the ground motion—as measured by the intensity measure—is often sufficient to predict the structural response for many types of structures. The most common intensity measure is spectral acceleration at the fundamental period of the structure (abbreviated $S_a(T_1)$). Other scalar and vector intensity measures based on spectral quantities have been proposed as well (for example, Baker and Cornell (2005) and Luco and Cornell (2007)). For six- or twenty-story, steel, moment-resisting frames, Olsen (2008) showed that, among the considered intensity measures, the peak inter-story drift ratio was best predicted by $S_a(T_1)$, assuming that the lateral force-resisting system was not compromised. However, large permanent drifts and collapse due to P- Δ are predicted better by a combination of large peak ground velocity (PGV) and PGD. Near-source ground motions with large PGDs are almost always accompanied by large PGVs. Thus, in the near-source region, PGDs are often associated with simulated collapse. Thus, although $S_a(T_1)$ is widely used as an intensity measure, there is evidence that an

alternate intensity measure may predict structural response better, depending on the structure type and response of interest.

The maximum value of the appropriate intensity measure likely to be experienced at a site is often of interest for seismic risk assessment. This maximum value depends on an expectation of future ground motions, which is informed by existing recorded ground motions as well as the current understanding of the physical mechanisms. This study shows that the maximum likely PGA within 10 km of a fault that produces a magnitude greater than 6.0 earthquake can be reliably derived from a lognormal distribution, since the empirical marginal distribution of such PGAs is stable with respect to new records and consistent with a lognormal distribution. Determining the maximum likely PGD in the same geographic region is more problematic given the available recorded ground motions with large PGDs. This study shows that the current empirical marginal distribution of PGD is fat-tailed, if not uniform on a logarithmic scale to a physical limit. This physical limit has not yet been determined, and simulation of ground motions from plausible large earthquakes (for example, the repeat of the 1906 San Francisco earthquake) suggest that the next large, well-recorded earthquake will contribute several measured PGDs between the current largest recorded PGDs and the suspected physical limit. Given this evidence, we find the calculation of the maximum likely PGD difficult—if not impossible—given the current knowledge.

These results have a direct bearing on seismic risk assessment. If the appropriate intensity measure for a structural response of interest is PGA, then recorded ground motions with the largest observed PGAs can be used as representative of future ground motions with the largest PGAs. The predicted structural responses based on existing large PGAs would be reliable even after the next large earthquake. However, if the appropriate intensity measure is PGD, the largest observed PGDs may not be representative of the future largest PGDs. Consequently, structural responses based on the largest recorded PGDs may not be representative of structural responses in the future largest PGDs. Consider tall buildings as an example. Their response to ground motions with PGDs of 0.2 m or 2 m or 10 m range from elastic to catastrophic failure. If this type of building is located within 10 km of an earthquake with magnitude greater than 6, each response is roughly equally likely if each level of ground motion is approximately equally likely. This study provides evidence to caution against assuming that recorded ground motions with the largest PGDs are representative of unrecorded or future ground motions with the largest PGDs.

This study supports the argument made in Heaton (2007). Among other questions, Heaton (2007) asks whether the current structural design philosophy of performance-based earthquake engineering is appropriate for the design of long-period structures. PBEE assumes that the reliability of a structure's performance—be it, seismic response or lifetime cost—can be predicted during the design phase. Heaton (2007) points out that such a reliability analysis is credible only if the

hazard is well-described by a knowable and finite uncertainty. The present study suggests that a future PGA (and thus the peak amplitude of short-period ground motions) is readily predictable from past records with a known, finite uncertainty. In contrast, a future PGD (or peak amplitude of long-period ground motions) cannot be anticipated with certainty because first, each large earthquake changes the observed marginal distribution of PGD, and second, the uncertainty of this distribution at present is quite large.

8 Conclusions

For ground motions recorded within 10 km of a fault producing an earthquake with magnitude greater than 6.0, the observed distribution of peak ground acceleration is lognormal, but the observed distribution of peak ground displacement is approximately uniform on a logarithmic scale, or at least fat-tailed. We expect the distribution of PGA to remain approximately lognormal with the addition of data from future earthquakes. However, the distribution of PGD will likely change after a well-recorded large earthquake, and specifically we expect the range of recorded PGD to expand at the largest values. The observed records show that there is essentially no correlation between PGD and PGA for near-source ground motions from large events.

Seismic scaling relations motivate the expected distribution of PGD as uniform on the logarithmic scale, or at least fat-tailed. This result implies that every peak ground displacement is approximately equally likely at a given site—to a physical limit—even though an earthquake large enough to produce large displacements is infrequent. There are no such scaling relations to predict the lognormal distribution of PGA because PGA does not scale with fault rupture area or slip on the fault. The large uncertainty in a future value of observed PGD exists despite the ability of earth scientists to accurately model long-period ground motions. In contrast, the certainty in a future value of PGA exists despite the inability to model short-period ground motions reliably.

The design of short-period structures with modern building codes is reliable in part because earthquakes with magnitude greater than 6 produce a stable marginal distribution of PGA with respect to future earthquakes. The mean and standard deviation of PGA in the next large earthquake can be predicted reliably from the historic distribution of PGA. The design of long-period structures, however, should not rely on the same argument: the observed marginal distribution of PGD may change after the next well-recorded large earthquake. The emphasis on recorded ground motions for structural design using seismic response histories assumes that future ground motions will be sufficiently similar to recorded ground motions. This study demonstrates that this assumption is not valid. Thus, there must be a different approach to the design of long-period structures that does not rely heavily on past observations of PGD to predict future amplitudes of long-period

ground motions. This approach should acknowledge that, for any site in the near-source region of a major fault, every PGD is similarly likely in the logarithmic sense. That is, 1 m of future peak ground displacement is approximately as likely as 10 m of future peak ground displacement. The optimal design of long-period structures should be the most economical design that is robust in large ground motions.

9 Data and Resources

All data used in this study are publicly available. Please see the Data section for more detail.

10 Acknowledgments

We are deeply appreciative of comments provided by the associate editor and reviewers. We acknowledge all institutions providing strong motion records, especially the COSMOS data center, SCEC data center, NIED, and JMA. Dr. David Wald of the United States Geological Survey and Dr. Hiroo Kanamori of the California Institute of Technology helped obtain data for the Hyogoken-nanbu earthquake. Dr. Egill Hauksson provided the CISN station list. This research was supported by the United States Geological Survey and Promoting Science and Technology (SCF) commissioned by the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- Aagaard, B., Hall, J., and Heaton, T. (2001). Characterization of near-source ground motions with earthquake simulations. *Earthquake Spectra*, 17(2):177–207.
- Aagaard, B. T., Brocher, T. M., Dolenc, D., Dreger, D., Graves, R. W., Harmsen, S., Hartzell, S., Larsen, S., McCandless, K., Nilsson, S., Petersson, N. A., Rodgers, A., Sjögren, B., and Zoback, M. L. (2008). Ground-motion modeling of the 1906 San Francisco earthquake, part II: Ground-motion estimates for the 1906 earthquake and scenario events. *Bulletin of the Seismological Society of America*, 98(2):1012–1046.
- Aki, K. (1972). Earthquake mechanism. *Tectonophysics*, 13(1-4):423–446.
- Akkar, S. and Gülkan, P. (2002). A critical examination of near-field accelerograms from the Sea of Marmara region earthquakes. *Bulletin of the Seismological Society of America*, 92(1):428–447.
- American Society of Civil Engineers (2006). *Minimum Design Loads for Buildings and Other Structures*. ASCE/SEI 7-05, Reston, VA.

- Aoi, S., Kunugi, T., and Fujiwara, H. (2008). Trampoline effect in extreme ground motion. *Science*, 322:727–729.
- Baker, J. W. and Cornell, C. A. (2005). A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon. *Earthquake Engineering and Structural Dynamics*, 34:1193–1217.
- Campbell, K. and Bozorgnia, Y. (2008). Nga ground motion model for the geometric mean horizontal component of pga, pgv, pgd and 5ranging from 0.01 to 10 s. *Earthquake Spectra*, 24:139.
- Çelebi, M., Toprak, S., and Holzer, T. (1999). Strong-motion, site-effects and hazard issues in rebuilding turkey: In light of the 17 august, 1999 earthquake and its aftershocks. *Proc ITU-IAHS Int Conf on the Kocaeli Earthquake, 17 Aug*, page 1999.
- Earthquake Research Department (2004). <ftp://angora.deprem.gov.tr/>. General Directorate of Disaster Affairs, Republic of Turkey Ministry of Public Works and Settlement.
- Hartzell, S. H. and Heaton, T. H. (1983). Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America*, 73(6A):1553–1583.
- Heaton, T. H. (2007). Will performance-based earthquake engineering break the power law? *Seismological Research Letters*, 78(2):183–185.
- Honda, R., Aoi, S., Morikawa, N., Sekiguchi, H., Kunugi, T., and Fujiwara, H. (2005). Ground motion and rupture process of the 2004 Mid Niigata Prefecture earthquake obtained from strong motion data of K-NET and KiK-net. *Earth, Planets and Space*, 57:527–532.
- Irikura and Miyake (2001). <http://sms.dpri.kyoto-u.ac.jp/irikura/M8.pdf>. Fault Parameters of the Earhtquake of M8 Class.
- Ji, C., Choi, K. K., King, N., Larson, K. M., and Hudnut, K. W. (2004). Co-seismic slip history and early afterslip of the 2004 Parkfield earthquake. *Eos, Transactions, American Geophysical Union*, 85(47). Fall Meeting Supplement, Abstract S53D-04.
- Ji, C., Helmberger, D. V., Wald, D. J., and Ma, K.-F. (2003). Slip history and dynamic implications of the 1999 Chi-Chi, Taiwan, earthquake. *Journal of Geophysical Research*, 108(B9):2412.
- Kanamori, H. and Anderson, D. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, 65(5):1073–1095.

- Kanno, T., Narita, A., Morikawa, N., Fujiwara, H., and Fukushima, Y. (2006). A new attenuation relation for strong ground motion in Japan based on recorded data. *Bulletin of the Seismological Society of America*, 96(3):879–897.
- KiK-net (2000). <http://www.kik.bosai.go.jp/>. National Research Institute for Earth Science and Disaster Prevention, Japan.
- Kyoshin Net (1996). <http://www.k-net.bosai.go.jp/>. National Research Institute for Earth Science and Disaster Prevention, Japan.
- Lee, W. H. K., Shin, T. C., Kuo, K. W., Chen, K. C., and Wu, C. F. (2001). CWB free-field strong-motion data from the 21 September Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America*, 91(5):1370–1376.
- Luco, N. and Cornell, C. A. (2007). Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra*, 23(2):357–392.
- Olsen, A. H. (2008). *Steel Moment-Resisting Frame Responses in Simulated Strong Ground Motions: or How I Learned to Stop Worrying and Love the Big One*. PhD thesis, California Institute of Technology, Pasadena, California.
- Rogers, A. and Perkins, D. (1996). Monte carlo simulation of peak-acceleration attenuation using a finite-fault uniform-patch model including isochrone and extremal characteristics. *Bulletin of the Seismological Society of America*, 86(1A):79–92.
- Scholz, C. (1982). Scaling laws for large earthquakes: Consequences for physical models. *Bulletin of the Seismological Society of America*, 72(1):1–14.
- Sekiguchi, H. and Iwata, T. (2002). Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms. *Bulletin of the Seismological Society of America*, 92(1):300–311.
- Shin, T.-C. and Teng, T.-I. (2001). An overview of the 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America*, 91(5):895–913.
- Somerville, P. (2003). Magnitude scaling of the near fault rupture directivity pulse. *Physics of the Earth and Planetary Interiors*, 137(1):201–212.
- Toki, K., Irikura, K., and Kagawa, T. (1995). Strong motion records in the source area of the Hyogoken-Nambu earthquake, January 17, 1995, Japan. *Journal of Natural Disaster Science*, 16(2):23–30.

- Trifunac, M. and Todorovska, M. (2001). Evolution of accelerographs, data processing, strong motion arrays and amplitude and spatial resolution in recording strong earthquake motion. *Soil Dynamics and Earthquake Engineering*, 21(6):537–555.
- Tsuboi, S., Komatitsch, D., Ji, C., and Tromp, J. (2003). Broadband modeling of the 2002 Denali fault earthquake on the Earth Simulator. *Physics of The Earth and Planetary Interiors*, 139(3–4):305–313.
- Wald, D. J. (1996). Slip history of the 1995 Kobe, Japan, earthquake determined from strong motion, teleseismic, and geodetic data. *Journal of Physics of the Earth*, 44:489–503.
- Wald, D. J. and Heaton, T. H. (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America*, 84(3):668–691.
- Wald, D. J., Heaton, T. H., and Hudnut, K. W. (1996). The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data. *Bulletin of the Seismological Society of America*, 86(1B):S49–S70.
- Wald, D. J., Helmberger, D. V., and Heaton, T. H. (1991). Rupture model of the 1989 Loma Prieta earthquake from the inversion of strong-motion and broadband teleseismic data. *Bulletin of the Seismological Society of America*, 81(5):1540–1572.
- Wells, D. and Coppersmith, K. (1994). New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America*, 84(4):974–1002.
- Yamada, M. (2007). *Early warning for earthquakes with large rupture dimension*. PhD thesis, California Institute of Technology.
- Yamada, M., Heaton, T., and Beck, J. (2007). Real-time estimation of fault rupture extent using near-source versus far-source classification. *Bulletin of the Seismological Society of America*, 97(6):1890–1910.
- Yamada, M., Mori, J., and Heaton, T. (2009). The slapdown phase in high acceleration records of large earthquakes. *Seismological Research Letters*, 80(4):559–564.

<i>Earthquake</i>	M_w	N	<i>Focal Depth [km]</i>	<i>Fault Model</i>
1979 Imperial Valley	6.5	14	12.0	Hartzell and Heaton (1983)
1989 Loma Prieta	6.9	8	19.0	Wald et al. (1991)
1992 Landers	7.3	1	15.0	Wald and Heaton (1994)
1994 Northridge	6.6	17	16.8	Wald et al. (1996)
1995 Hyogoken-Nanbu	6.9	4	20.3	Wald (1996)
1999 Izmit	7.6	4	17.0	Sekiguchi and Iwata (2002)
1999 Chi-Chi	7.6	42	21.2	Ji et al. (2003)
2002 Denali	7.8	1	15.0	Tsuboi et al. (2003)
2004 Parkfield	6.0	47	12.0	Ji et al. (2004)
2004 Niigataken-Chuetsu	6.6	9	13.0	Honda et al. (2005)

Table 1: Earthquake dataset used for the near-source ground motion analysis. Moment magnitude (M_w) and focal depth are cited from the Harvard Centroid Moment Tensor solution. The preliminary determination of epicenter is used for the focal depth. The number of near-source records for each earthquake (N) is also tabulated. The fault models are used to define near-source stations.

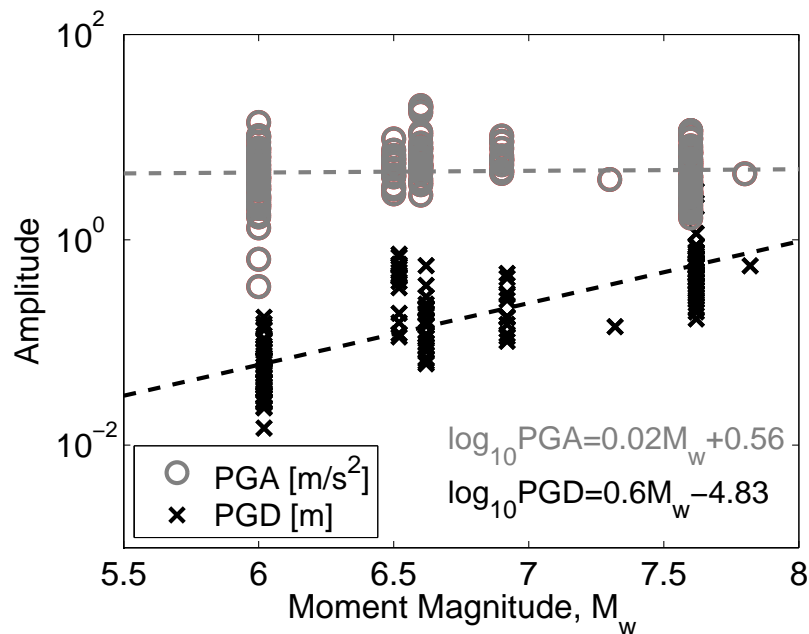


Figure 1: Near-source PGA and PGD as functions of moment magnitude. The dashed lines are least-squares regression lines for the two intensity measures, and the equations for these lines are also given.

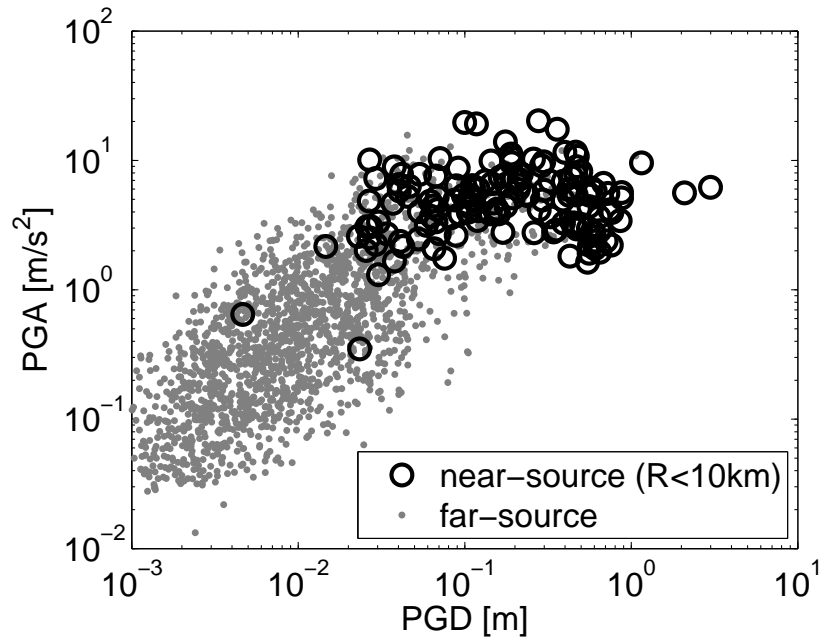


Figure 2: Relationship between PGD and PGA. The near-source records are the same data used in this study. The far-source records are the records from earthquakes with magnitude greater than 6.0.

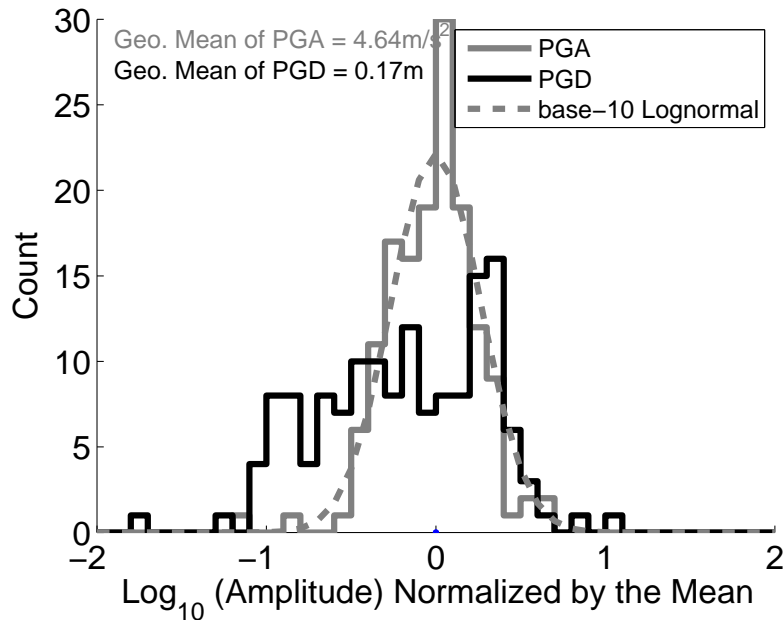


Figure 3: Comparison of observed near-source PGA and PGD histograms to a base-10 lognormal distribution. The lognormal distribution shown has a geometric mean (4.64 m/s^2) and geometric standard deviation (1.77 m/s^2) calculated from the PGA data.

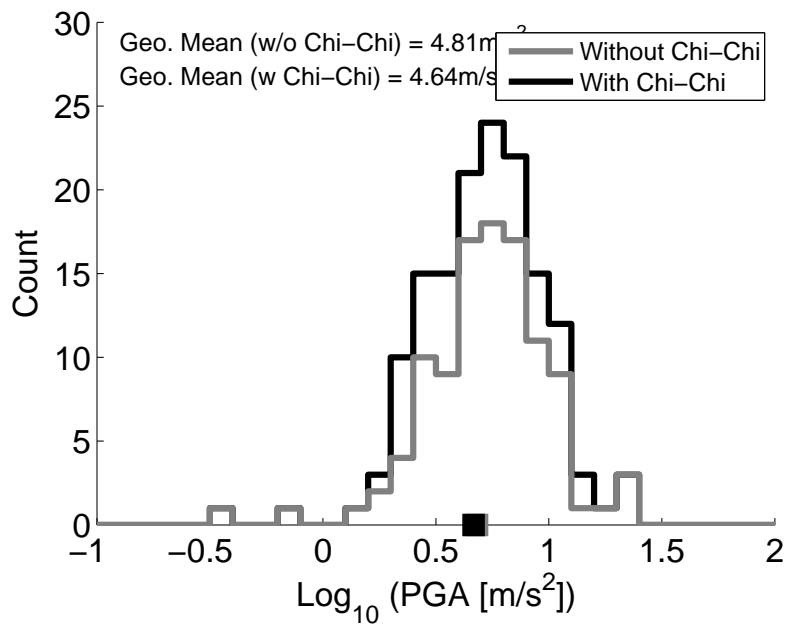


Figure 4: Histogram of the near-source PGA for earthquakes with magnitude greater than or equal to 6.0. The black and grey lines are histograms of the full dataset with and without the Chi-Chi earthquake records, respectively. The squares on the horizontal axis indicate the geometric mean of each component and their values are shown on the plot.

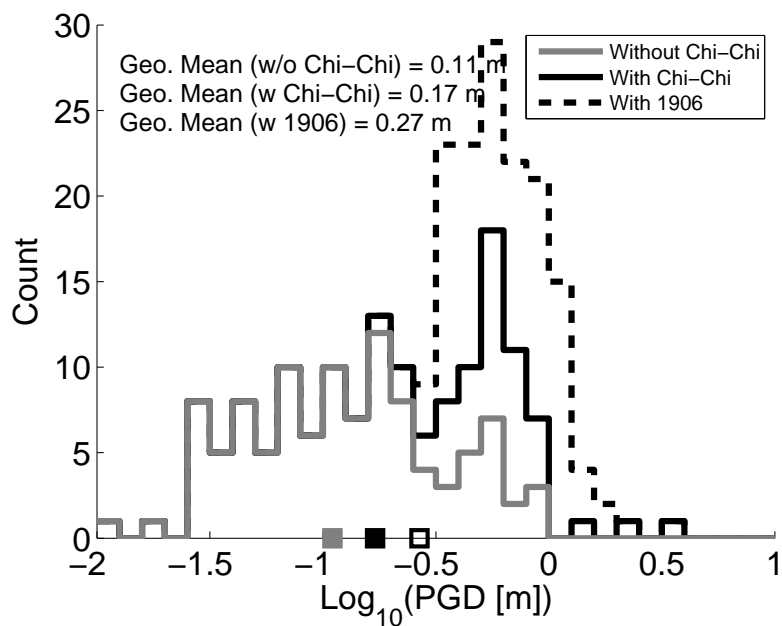


Figure 5: Histogram of the near-source PGD for earthquakes with magnitude greater than or equal to 6.0. The black and grey lines show the distribution of recorded PGD in our data set with and without the Chi-Chi observations, respectively. The dashed lines augment the recorded data set with an imagined repeat of the 1906 San Francisco earthquake. The squares on the horizontal axis indicate the geometric mean of each set of data, and the values are listed in the plot.

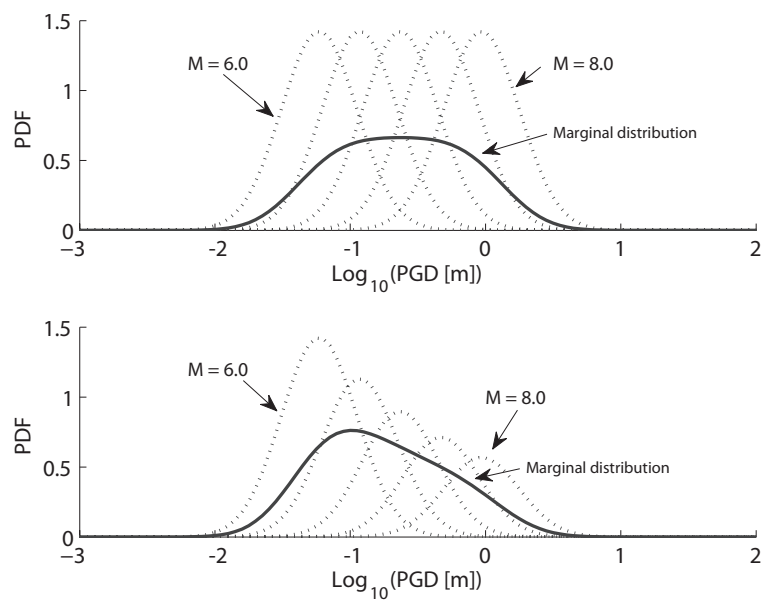


Figure 6: Theoretical marginal distribution of PGD. Seismic scaling relations predict that the number of recorded PGDs at a given magnitude in the near-source region is constant (top, moderate earthquakes) or weakly dependent on magnitude (bottom, large earthquakes). Given enough observations, the resulting, expected marginal distribution of PGD more closely resembles a uniform—not normal—distribution on the logarithmic scale.

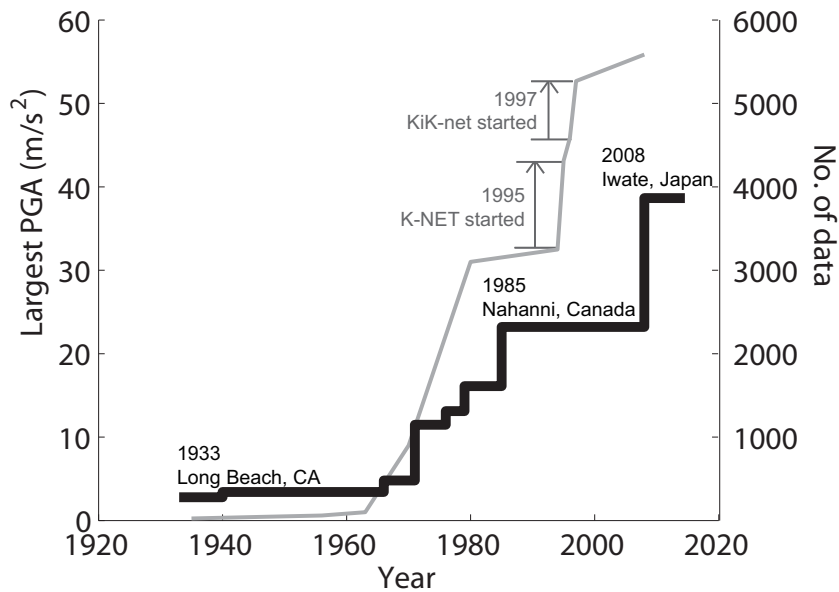


Figure 7: Approximate number of seismic stations in the United States and Japan and the largest recorded PGA. As the number of seismic stations recording ground motions increases, the largest recorded PGA increases as well. With more instrumentation deployed, it is more likely to record larger PGAs. (Adapted and extended from Çelebi et al. (1999))

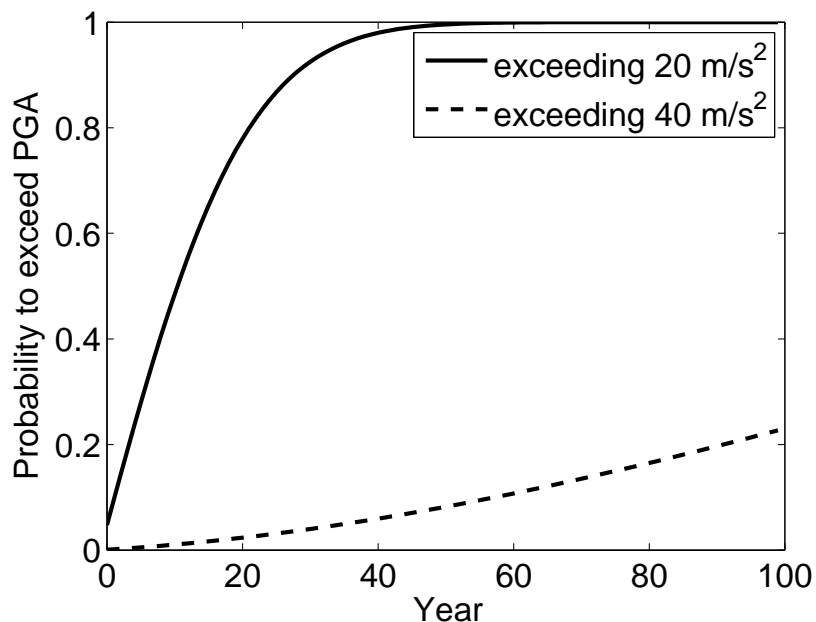


Figure 8: Probability that the PGA exceeds 20 and 40 m/s^2 as a function of time. The parameters used for this simulation are: $\rho=0.004$ station/ km^2 , $A=6.0$, $b=1.0$ $m=1.54$, $\sigma=0.57$, $M_0=6.0$, and $M_U=8.5$. The station density is assumed to increase 5% every year.