Comment on “The Gutenberg-Richter or Characteristic Earthquake Distribution, Which Is It?” by Steven G. Wesnousky

by Yan Y. Kagan

In a recent article, Wesnousky (1994) (henceforth referred to as Wesnousky) analyzes seismicity and geologic deformation rates for five faults in southern California and arrives at the conclusion that the characteristic model yields a better description of the earthquake size statistical distribution than the Gutenberg-Richter (G-R) magnitude-frequency law. To analyze earthquake occurrences in the fault zones, Wesnousky compares seismicity levels during the past 50 to 60 yr with the return rate of characteristic earthquakes. The characteristic recurrence time is calculated on the assumption that all of the seismic moment rate evaluated on the basis of geologic slip-rate measurements is released by the characteristic events (equations 2 and 3 in Wesnousky). Wesnousky finds that for four out of five zones, the characteristic earthquake rate thus estimated is larger by a factor of more than 10 compared to the rate obtained by the extrapolation of instrumental seismic records. Although the results of paleoseismic studies are sometimes used to corroborate the recurrence periods, the major proof for the characteristic rate discrepancy is based on a comparison of magnitude-frequency plots and the return periods calculated using equations (2) and (3) of Wesnousky.

Thus, Wesnousky’s results depend crucially on whether the seismic activity (productivity) of small and intermediate earthquakes in fault zones can be effectively correlated with the geologic deformation rate, as well as whether the above equations correctly evaluate the rate of occurrence for large earthquakes. Even if we accept the estimates of geologic slip-rate deformation rate proposed by Wesnousky, two factors can significantly change his conclusions: (1) parameters of the earthquake size distribution measured for each of the faults and (2) calculations of the characteristic recurrence rate. The latter depends on the contribution from earthquakes that are smaller or larger than the characteristic limit. Wesnousky briefly discusses smaller earthquakes and finds that their contribution is negligible; the possibility of earthquakes larger than the characteristic limit is not considered.

The guiding principle of this commentary is the idea that the simple null hypothesis (model) needs to be first fully and critically investigated, and only if it can be shown that the null hypothesis should be rejected, we should formulate an alternative model. The null hypothesis for the earthquake size statistical distribution is that it follows the G-R law. The difference between the G-R law that approximates the earthquake size distribution for large regions and the size distribution for geologically defined faults is explained as the consequence of the distribution of fault sizes and slip rates. Wesnousky et al. (1983) argue that whereas earthquake size distribution for “individual” faults follows the characteristic model, as the faults themselves are distributed according to a power law, the resulting magnitude-frequency relation for a large region is the G-R law. However, for southern California, application of the characteristic model yields cumulative curves that are significantly different from the historical/instrumental magnitude-frequency relation (Working Group, 1995, see their Fig. 14 and discussion on pp. 399 and 417). The “characteristic” curves also are less resembling the linear G-R relation than the experimental curve. Only by invoking the “cascade” procedure in which an earthquake ruptures through several characteristic fault segments, the discrepancy can be reduced to a manageable factor of 2.

The cascade hypothesis is not a part of the original characteristic model (Schwartz and Coppersmith, 1984); in the strict interpretation of the model, no earthquake larger than the characteristic one is allowed.

As a measure of earthquake size, I use the scalar seismic moment of an earthquake, which I denote by symbol $M$. Occasionally, I also use an empirical magnitude of earthquakes denoted by $m$. To transform the seismic moment into magnitude, I use the standard relation

$$ m = \frac{2}{3} (\log_{10} M - 16), \tag{1} $$

where $M$ is measured in dyne cm.

In this commentary, I first remark on the statistical problems of verifying the characteristic hypothesis using relatively small seismic zones, then I analyze how earthquakes that are smaller and larger than the characteristic size contribute to the seismic moment rate. To quantify the latter comparison, I introduce two null hypotheses that can be tested against the characteristic model as shown in Figure 1b of Wesnousky. The first null hypothesis, $H_{\text{null}}$, assumes that the size of tectonic earthquakes is distributed according to the gamma distribution (see equation 5 below) with the universal values for its parameters over the world. The hypothesis tests whether very large events ($m \geq 8$) may occur everywhere in the world; these earthquakes carry most of the tectonic deformation (Kagan, 1993, 1994, 1995b). The second null hypothesis, $H_{\text{null}}$, assumes that earthquakes follow the cumulative G-R relation up to a value of $m_{\text{max}}$ (maximum magnitude) and no earthquake can occur with $m > m_{\text{max}}$. The maximum magnitude is specific to each fault and is allowed
to vary from region to region. In Kagan (1993), this relation is called the "maximum moment (magnitude) distribution."

Seismicity Statistics

Statistical Properties of Seismicity of Small Regions

In this section, I discuss a procedure for determining the recurrence rate for characteristic earthquakes using catalog data. A spatial window has to be first defined for the data. Wesnousky selects this box window as an approximately rectangular polygon centered on the mapped fault trace. The window, therefore, has a width \( W \) and a length \( L \). The selection of \( W \) and \( L \) needs to be justified. But before we turn to the \( W \) and \( L \) selection, we should first consider more general statistical problems of working with earthquake catalogs.

Although the Caltech earthquake catalog is indisputably one of the best well-documented, long-term compilations of the seismic data, it has several drawbacks that are common to other catalogs. The major problem is the inhomogeneity of the catalog in time, space, and magnitude. This problem is caused by the evolution of a seismographic network, methods of seismogram interpretation, and the combination of several magnitudes in one catalog (Hutton and Jones, 1993). Habermann (1987) emphasizes the influence of these factors on the statistical properties of seismicity and provides many examples when a change in the earthquake occurrence pattern can be traced to man-made causes. The catalog inhomogeneity makes the extrapolation of the magnitude-frequency curves to large magnitudes an uncertain and error-prone procedure. Although a detailed analysis of network procedures and resulting biases can alleviate the problem, many systematic errors and deficiencies of technical and human origin cannot be definitely discovered and corrected. We have to treat many such biases as random errors and try to limit their influence by applying statistical methods. However, the proper use of statistical procedures requires that data have a property of statistical stability: if the data are contaminated by a few large systematic errors, the statistical result is meaningless. This is usually the case for the data from small seismic regions: for example, an opening, closure, or malfunction of one seismographic station can significantly change the data in ways that are difficult to predict. The seismic data from larger regions or from global earthquake catalogs have many similar irregularities; nonetheless, they can be treated as random noise.

Another source of statistical bias is the data-based selection: by this I mean the direct or implicit use of catalog data to select boundaries, magnitude range, time interval, etc., for the region that is analyzed statistically. In such a case, it is often very difficult to establish the extent to which statistical properties of the data have been modified by the selection, and the reliability of the results is always questionable. This may contribute to a general inconsistency of statistical studies of local seismicity: these results may depend strongly on the choice of initial conditions. Again, it is much easier to perform such biased selection for a relatively small region; for large provinces, the selection effects can be treated essentially as a random influence, thus the application of statistics to large regions is more likely to yield robust results.

Seismic Zone Selection

How can one define the width \( W \) of a zone? The selection of a fault zone is problematic when seismic data are available during the selection—limiting width of a fault zone, one can obtain a "characteristic" distribution with the characteristic earthquake and a few of the noncharacteristic events left in a sample. A zone selection procedure that produces such potentially implausible results has little persuasive power.

Ideally, a seismic sample should match the geologic determination of deformation in a fault zone. However, the zone width is not easy to define unambiguously. First, the geologic measurements are easier to make in places of localized high strain, whereas zones of widely distributed deformation may not be as uniformly covered in a survey. Second, not all strands of a seismic fault may be active during the available catalog time span. Moreover, geometric complexities of fault systems such as the Big Bend of the San Andreas fault necessitate the occurrence of earthquakes which are inconsistent with the general trend of the major fault, and, on the other hand, are clearly part of the fault system. Should the fault zones of such earthquakes be included in the San Andreas fault zone? Earthquakes with diverse mechanisms on faults close to the "main" fault affect the stress environment within the San Andreas fault and contribute to the plate motion. All these questions make the definition of the zone width a difficult and subjective operation.

Wesnousky discusses the selection of the fault zone length \( L \) in many places of the article. Most of the zones are analyzed using several scenarios, in which the length of the characteristic rupture is allowed to vary, sometimes by an order of magnitude. However, no objective, testable criteria are offered to justify the selection and show, for example, that rupture lengths that are larger than those listed by Wesnousky in Table 1 cannot exist. A nonuniqueness of the fault segmentation is demonstrated by the Working Group (1995) report where the San Andreas fault zone is subdivided into seven segments. Freeman et al. (1986) as well as Wheeler and Krystinik (1988) also discuss the ambiguity and uncertainties of fault segmentation.

Wesnousky also argues that the San Jacinto fault zone, which is the only zone that shows little or no difference between the observed magnitude-frequency law and his calculations for the characteristic event rate, can be subdivided into smaller individual fault segments (see his Fig. 12). For these smaller segments, the discrepancy between the characteristic rates and extrapolation of observed earthquake distribution increases.
Let us consider a general case of a fault segmentation: I assume that the geologic deformation rate $M_g$ and the characteristic rate $a_{ch}$ based on the extrapolation of the magnitude-frequency relation, are statistically uniform over the length of the fault. The seismic moment for a characteristic earthquake is calculated as

$$M_{ch} = CL^d,$$  \hspace{1cm} (2)

where $L$ is the fault length, and $d$ and $C$ are empirical constants (Wesnousky, 1994, Fig. 4). The ratio of the geologic rate to the observed characteristic rate

$$r_o = \frac{M_g}{M_{ch}a_{ch}}$$ \hspace{1cm} (3)

demonstrates the degree of discrepancy between the seismic and geologic data, according to the characteristic model. Suppose we calculate a similar ratio $r_p$ for the part (segment) $l$ of the fault, $l = pL$, ($p < 1$); taking into account change in the moment of the characteristic earthquake and its rate $a_{ch}$ due to the fault subdivision, we obtain

$$r_p = r_o p^{-d(1-\beta)}.$$ \hspace{1cm} (4)

The rate $a_{ch}$ in equation (4) is influenced both by the size of the segment and by the reduction of the characteristic magnitude. For example, if $p = 0.2$, $d = 1.3$, and $\beta = 2/3$, using equation (4) we obtain $r_p \approx 2.0 r_o$, i.e., the discrepancy increases by a factor of 2.

From equation (4), $r_p = r_o$ only if $d = 0$. The above does not mean that the earthquake size does not depend on the fault or the rupture length. According to the null hypothesis, the magnitude-frequency relation can be measured for any spatial window; it counts the number of epicenters in the window. The null hypothesis also assumes that the earthquake size distribution should have the same form for any window. An epicenter of a large earthquake can be located in any part of its rupture zone, thus the ratio (3) is on average the same for any segment of a fault. Of course, the null hypothesis assumes no characteristic earthquakes, and the ratio $r_o$ should be reinterpreted accordingly (see Distribution of Earthquake Numbers below).

According to equation (4), any subdivision of a fault zone into smaller segments on average “improves” arguments in favor of the characteristic model: the smaller the segment, the larger the discrepancy between the recorded seismic activity and a calculated characteristic rate. Since the fault geometry distribution is fractal (Kagan, 1994), such segmentation can continue up to very small fault strands and segments, and $r_p$ could be made as large as desired. Again, an argument that can produce such results cannot be regarded as strong evidence.

Seismic Moment Rate Distribution

In this section, I discuss two problems of the seismic moment rate calculation using the tectonic deformation rate. The tectonic deformation can be evaluated either using geologic data (Wesnousky) or geodetic measurements. We need to evaluate the contribution to the total deformation rate coming from earthquakes that are smaller than the assumed characteristic event or the possible contribution of larger earthquakes.

Wesnousky (pp. 1954–1955) argues that the contribution of small earthquakes is negligible for most of the southern California fault zones and calculates the sum of seismic moments for small events as proof. To analyze this argument, we need to know the statistical distribution for the sum. I discuss the distribution of the seismic moment sum in further detail below, since similar problems are often encountered in the interpretation and analysis of earthquake data.

Regarding the second problem, Wesnousky does not discuss why earthquakes larger than the assumed characteristic limit should be ignored when calculating the balance of a seismic moment release. Such a possibility is not implausible. Two types of calculations with characteristic events of different size are proposed for the San Jacinto zone (see Figs. 7 and 12 in Wesnousky, 1994). Furthermore, most of the analysis for other seismic zones employs several scenarios that also envision that characteristic earthquakes of significantly different sizes may occur in the same zone. The use of different scenarios undermines one of the fundamental assumptions of the characteristic model: the lack of earthquakes slightly smaller than the characteristic limit (Schwartz and Coppersmith, 1984). The Landers 1992 earthquake demonstrated, for example, that four faults that were described as independent geologic entities, each with a separate estimate of the maximum magnitude (Wesnousky, 1986), can be ruptured in one earthquake (Johnson et al., 1994).

If we assume that very large earthquakes ($m \geq 8.0$) can originate on all fault zones in California, their recurrence rate can be calculated by extrapolating the magnitude-frequency relation. Using the plots in Wesnousky (1994), we obtain a recurrence time varying from 0.6 to 30 Kyr (depending on the catalog time limits and assumed $b$ value). Surface traces of such earthquakes can be obliterated, at least partially, especially if they rupture through multiple fault zones as in the Landers earthquake, thus geologic investigations cannot rule out the possibility of such rare events. Moreover, the geologic studies indicate that the major faults, like those considered by Wesnousky, have an age of many million years. With the slip rate of several millimeters per year several kilometers or tens of kilometers of displacement should accumulate along the fault trace. It is mechanically implausible that all this strain is distributed along the length of faults that according to Wesnousky (his Table 1) measure 50 to 250 km. One possible solution of this paradox is to
assume the occurrence of very large earthquakes transmitting a displacement outside of the fault trace now known.

Kagan (1993, equations 20 through 24) considers the contribution of large events and shows that if earthquakes actually follow the G-R relation, but the characteristic model is assumed, the estimated rate of characteristic events could exceed 5 to 10 times the actual rate of a similar size earthquake. In this commentary, I want to pursue a different argument: I would like to show that if the tectonic deformation rate and seismic activity levels are known, the estimates of the maximum magnitude for each fault zone are consistent with the worldwide value, which is based only on earthquake statistics (Kagan, 1994, 1995b). This would mean that the null hypothesis $H_{00}$ does not contradict the southern California data. If, for example, $m_{\text{max}}$ for the faults considered by Wesnousky is shown to be significantly less or more than the worldwide value, it would suggest that $H_{00}$ is incorrect. To carry out such analysis, we need to know possible variations of both geologic and seismic results. Wesnousky supplies the upper and lower limits for the geologic slip rate, $M_e$. To evaluate possible fluctuation of the seismic activity rate, we should study the statistical distribution of earthquake numbers.

Distribution of Seismic Moment Sum

To explore the sum behavior, we need to review the distribution of the earthquake scalar seismic moment. The moment distribution can be approximated by several relations that are transformations of the G-R law with a modification at the large moment end (Anderson and Luco, 1983; Kagan, 1993, 1994). In effect, the moment distribution has a Pareto form for small and intermediate earthquakes and an exponential or a step-function decay (truncation) at the maximum moment, $M_{\text{max}}$. Below I assume that the distribution density is a modified gamma law.

\begin{equation}
\phi(M) = C^{-1} M^{-1-\beta} \exp(-M/M_{\text{max}}),
\end{equation}

where $M \geq M_{\text{cut}}$; $M_{\text{cut}}$ corresponds to a minimum size (cutoff) earthquake covered uniformly in time and space by a seismographic network; $C$ is a normalizing coefficient; and $\beta = 2/3b$, where $b$ is a traditional parameter in the G-R relation. The value of $\beta$ for shallow earthquakes is close to 2/3 (Kagan, 1994, 1995b). For purposes of this commentary, the approximation by other statistical distributions tabulated in Anderson et al. (1993) or in Kagan (1993) yields similar results. Kagan (1995b) discusses the cases when the application of truncated distributions leads to significantly different results compared to the gamma law.

The distribution of the sum of a random variable (equation 5), i.e., the sum of earthquake scalar seismic moments, is complicated and cannot be studied in a closed form. However, there is a class of so-called stable statistical distributions for which the exact distribution of the sum is easily described (Feller, 1966). Tails of stable distributions are power laws, and one of these distributions can be applied to approximate the moment size distribution. The stable distributions started recently to be widely employed in many scientific fields, but since these distributions are not familiar to most geophysicists, below I discuss it in more detail.

The stable distribution density can be represented as (Zolotarev, 1986)

\begin{equation}
f(x, \beta, \delta, \eta),
\end{equation}

where $\beta$ is the main parameter (exponent) of a stable distribution that characterizes its form; $\delta$ is a scale parameter; $\eta$ is the degree of the distribution asymmetry; and $\eta = 0$ corresponds to a symmetric distribution, whereas for $\beta < 1, \eta = 1$ represents a distribution defined only at $x \geq 0$. Since the scalar seismic moment is positive, $\eta = 1$.

The stable distribution density cannot usually be expressed in analytic or even special functions. Fortunately, for $\beta = 2/3$ and $\eta = 1$, the representation exists and can be demonstrated to show its relevance for our purposes. The expression for density is given by Zolotarev (1986, p. 157, equation 2.8.33):

\begin{equation}
f(x, 2/3, 1, 1) = \frac{\sqrt{3}}{2x^{4/3}} \exp \left( -\frac{2}{27} x^2 \right) W_{1/2, 1/6} \left( \frac{4}{27} x^2 \right),
\end{equation}

where $W_{1/2, 1/6}$ is a Whittaker function (Gradshteyn and Ryzhik, 1980, p. 1059) and $x \geq 0$.

Distribution (7a) has been evaluated using the Mathematica package (Wolfram, 1991). In Figure 1, I display the distribution in density and cumulative forms. The cumulative curve is obtained by numerical integration (cf. Zolotarev, 1986, equation 2.2.27).

\begin{equation}
F(x, 2/3, 1, 1) = \frac{1}{\pi} \int_0^x \exp \left[ -\frac{2}{27} t^2 \right] \sin(t/3) \sin(2t/3) \sin(t)^{-3} \, dt.
\end{equation}

The distribution tail decays according to a power law:

\begin{equation}
f(x, \beta) \propto x^{-\beta-1}, \quad \text{for } x \to \infty,
\end{equation}

for the distribution density (see Fig. 1a), and

\begin{equation}
1 - F(x, \beta) \propto x^{-\beta}, \quad \text{for } x \to \infty,
\end{equation}

for the function $1 - F$ (see Fig. 1b). The decay (8) signifies that distribution (7) does not have the first statistical moment (a mean) for $\beta = 1.0$.

The curves in Figure 1 have a scale parameter $\delta$ equal to 1.0. In order to approximate the seismic moment distribution, $\delta$ should be modified; for example, $X = 1$ could be identified with $M_{\text{cut}}$. A change in $\delta$ causes only a horizontal shift of curves in Figure 1, or as another option, we rescale the abscissa. Figure 1 suggests that stable distribution (7)
Figure 1. Stable probability distribution \( f(x, 2/3, 1, 1) \) (see equation 7): (a) distribution density; (b) cumulative function.

differs from the moment distribution (5) in two regions: for small and for very large earthquakes. In the former case, distribution (5) is controlled by a magnitude cutoff level, whereas the decay for very large events is a consequence of the conservation of energy argument. Nonetheless, unless a few very large events are present in a sample, stable distribution (7) is a reasonable approximation for the moment sum.

Feller (1966, vol. 2, p. 169) discusses the stable distribution behavior with exponent \( \beta < 1 \); the average of stable variables

\[
(X_1 + X_2 + \cdots + X_k + \cdots + X_n)/n = S_n/n, \quad 1 \leq k \leq n, \quad (10)
\]

has the same distribution as \( X_n^{n^{-1+\nu}} \), i.e., in our case (\( \beta = 2/3 \)), the average seismic moment \( S_n/n \) is distributed as \( X_n^{1/3} \). Thus, for example, the distribution for a sum of seismic moments for 100 earthquakes is the same as shown in Figure 1, but the abscissa should be scaled by a factor of 1000. This means that the random fluctuation of the mean of the variable \( X \) increases as \( \sqrt{n} \), the summation of the stable variables with \( \beta < 1 \) accentuates randomness instead of suppressing it. This is contrasted with the behavior of the average of Gaussian variables, which is more familiar in geophysics: \( S_n/n \) is distributed as \( X_n/\sqrt{n} \); i.e., the standard deviation of the mean decreases as \( 1/\sqrt{n} \). The reason for such a surprising result is that the maximum term in the sum (10), \( M_n = \max(X_1 + \cdots + X_n) \), is comparable in magnitude to the sum \( S_n \); the ratio \( M_n/S_n \) tends to \( (1 - \beta)^{-1} \) (Feller, 1966, vol. 2, p. 169); i.e., for \( \beta = 2/3 \), the seismic moment of the largest earthquake contributes on average \( 1/3 \) of the seismic moment total.

The above result has been described by Kagan (1993). For the distribution of earthquake size called the “maximum magnitude (moment)” with the value of parameter \( \beta = 2/3 \), \( 1/3 \) of the total seismic moment rate is contributed by the earthquakes with \( M_{\text{max}} \), and the rest of the moment rate is supplied by smaller events. Thus, we can interpret the maximum moment distribution so that the largest earthquake in the observed sequence of earthquakes is declared as the maximal possible event, and the remaining smaller earthquakes are expected to satisfy the G-R relation; i.e., the maximum moment distribution is a statistical artifact.

What are the conclusions one can draw from the seismic moment sum distribution described above? Since all statistical moments, including the average and variance, are infinite for the stable distribution with \( \beta \leq 1.0 \), no regular statistical error estimate is possible for the sum. The sum, in general, has rather poor statistical properties; it depends strongly on the largest one or two earthquakes in a catalog, making a selection bias more probable. Such a bias may involve declaring any earthquake that is close in magnitude to the assumed characteristic limit as belonging to the characteristic event population. This strategy “works” with available data, but it should fail in a real-time prediction: whatever limit is established for a characteristic earthquake, according to the null hypothesis, there will always be events with a magnitude close to this limit, and these earthquakes should supply a significant part of the total moment rate.

The instability of the moment sum estimates can be demonstrated by the calculation of the total seismic moment for the Newport-Inglewood fault zone: depending on which time period we use, 1932 to 1994 or 1944 to 1994, the seismic rate changes by two orders of magnitude. This property of the total moment makes it less appropriate for statistical conclusions. A similar calculation of the seismic moment rate, using the magnitude-frequency relation, is much more stable. Unfortunately, such evaluations are model dependent.

There is no indication in the magnitude-frequency plots of observational seismic data (Figs. 6 through 10 in Wesnousky, 1994) for the significant upward “bend” in the cumulative curves. Such a bend would correspond to the lack of earthquakes close to the characteristic limit. Upward fluc-
tations of the curves should be tested for statistical significance before they are proposed as a proof for characteristic earthquakes (see more in Kagan, 1993a). Thus the presence of a large bend in these plots is a consequence of combining the instrumental and geologic data.

The comparison of geologic and seismic rates is dependent on the appropriateness of the characteristic hypothesis and on the segmentation scenario for each fault. If several scenarios with widely different segment lengths are proposed, all of the earthquakes occurring on such segments should be introduced into the moment budget. Since these earthquakes are infrequent, the historic and instrumental data are not usually sufficient to draw valid conclusions. If, as Wesnousky suggests, the larger segments could be subdivided into smaller ones (see his Figs. 6 and 12), the range of presumed characteristic earthquakes overlaps with smaller events tabulated in a catalog.

Wesnousky plots the rates corresponding to each scenario as if they were mutually exclusive. However, the calculation should take into account the sum of the seismic moment of all of the possible characteristic earthquakes, each scenario weighted with the appropriate probability. Suppose three fault rupture scenarios are proposed. Assuming equal probabilities for each “script,” the discrepancy between the seismic data and characteristic rates would be reduced by a factor of 3. For cumulative curves, the result of such calculation would change their shape in a critical region between the instrumental data and the geologically based results and possibly suppress the horizontal part of the curve in Wesnousky’s Figures 2b and 6b through 10b. Thatcher (1990) documents several cases in which the same fault segment has been ruptured by earthquakes covering different subsegments.

If we assume that the null hypothesis $H_0$ is correct, the calculations of the characteristic rate should yield the value that is at least 3 times lower. If the actual distribution of earthquake sizes is different from $H_0$, assuming the same value of $M_{max}$, we may obtain an even larger bias in the estimates of a geologic recurrence period (Kagan, 1993). A possible bias in estimating the characteristic earthquake rate due to events larger than the characteristic limit is even higher (see below).

Distribution of Earthquake Numbers

Several approaches can be used in the study of earthquake number distribution. The distribution for catalogs of large earthquakes that have relatively few aftershocks is considered to follow the Poisson distribution (Gardner and Knopoff, 1974). However, for catalogs of small and intermediate events, such approximation is clearly not permissible. One can decluster the catalogs that include small earthquakes, as Wesnousky suggests. Unfortunately, there are no universally accepted procedures for aftershock removal. Moreover, since aftershocks are smaller in general than mainshocks, such declustering should lower the $b$ value (Frohlich and Davis, 1993). The $b$ parameter plays an important role in the extrapolation of the magnitude-frequency relation to large magnitudes.

My attempts to decluster the Caltech catalog using the Reasenberg (1985) method do not yield unambiguous results. Although declustered catalogs are closer to the Poisson process, as measured by parameters of the negative binomial distribution (see below), the $b$ values for modified catalogs do not exhibit an consistent pattern. This may be due to the magnitude inhomogeneity of the catalog. Hutton and Jones (1993, p. 319), for example, suggest that the local magnitude be replaced by the moment magnitude for earthquakes with $m \geq 6$. This makes the use of the magnitude-frequency curves for prediction of large earthquake rates unreliable and subject to additional uncertainties caused by the parameter selection for declustering. Therefore, I use the original catalogs to compare seismic and geologic data: at least their statistical properties can be studied and fewer systematic errors are introduced into our procedures.

The Poisson distribution has been extensively used to approximate the number of earthquakes in various space-time domains. The distribution approximates reasonably well the numbers of large earthquakes in seismic regions or the numbers of events in declustered earthquake catalogs. The probability of $k$ events is

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

where $k = 0, 1, 2, \ldots$; the rate of earthquakes $\lambda \geq 0$. The average and variance of $k$ is

$$E(k) = D(k) = \lambda.$$

The negative binomial distribution (Feller, 1966, vol. 1, p. 166; Kotz and Johnson, 1985, vol. 6, pp. 169–177) has been suggested by Shlien and Toksoz (1970) and Kagan (1973) to approximate the number of earthquakes at large space-time domains, especially in those cases when an earthquake catalog contains extensive aftershock sequences. The probability of $k$ events is

$$p(k) = \binom{\tau + k - 1}{k} \theta^k (1 - \theta)^\tau,$$

where $k = 0, 1, 2, \ldots$; $\tau \geq 0$; and $0 \leq \theta \leq 1$. The average of $k$ is

$$E(k) = \tau \frac{1 - \theta}{\theta},$$

and its variance

$$D(k) = \tau \frac{1 - \theta}{\theta^2}.$$
We can see that the negative binomial distribution generally has a larger standard deviation than the Poisson law. For $\theta \to 1$ and $\tau(1 - \theta) \to \lambda$, expression (13) tends to (11) (Feller, 1966, vol. 1, p. 281); the negative binomial distribution becomes the Poisson distribution; i.e., the latter distribution is a special case of the former one.

Figure 2 displays the distribution of the yearly numbers of events $m \geq 3.0$ in the Caltech catalog 1932 to 1994 for the southern California region inside the box [32.5° to 36.5° N, 114.0° to 122.0° W]. The largest number of events—1560—for 1992 is due to the Landers sequence. A large stochastic variation of yearly event numbers can be seen in Figure 3 by Wesnousky. Using the statistical moments of the distributions (12), (14), and (15), we can determine the values of parameters for distributions (11) and (13): $\lambda = 219.6$, $\theta = 0.006$, and $\tau = 1.29$. If the magnitude cutoff is increased, the value of $\theta$ also increases, for instance, for $m_{cut} = 6$, $\theta = 0.54$; i.e., the catalog is closer to the Poisson model (see the previous paragraph).

The same plot shows several theoretical distributions: the negative binomial and Poisson laws with parameters listed above, as well as the negative binomial distribution with $\theta = 0.05$. The Poisson cumulative distribution is calculated using the following formula:

$$F(k) = P(N < k) = \frac{1}{k!} \int_0^\infty y e^{-\lambda} dy = 1 - \gamma(k + 1, \lambda),$$  \hspace{1cm} (16)

where $\gamma(k + 1, \lambda)$ is an incomplete gamma function. For the negative binomial distribution,

$$F(k) = P(N < k) = \frac{1}{B(\tau, k + 1)} \int_0^\tau \gamma^{k-1} (1 - \gamma)^\tau d\gamma,$$  \hspace{1cm} (17)

where $B(\tau, k + 1)$ is a beta function (Kotz and Johnson, 1985, v. 6, pp. 169–177) and the right-hand part of the equation corresponds to an incomplete beta function, $\beta(\tau, k + 1, x)$ (Gradshteyn and Ryzhik, 1980).

The comparison of the curves in Figure 2 suggests that the Poisson distribution is not a good approximation for earthquake numbers, but the negative binomial distribution is not fully appropriate for this purpose either. The negative binomial distribution with $\theta = 0.05$ approximates the cumulative curve for most of the years but fails in the years when large aftershock sequences occur. However, to calculate the standard deviation $\sigma$ for earthquake numbers, the negative binomial distribution is a more appropriate choice than the Poisson distribution. Whereas for the latter, the deviation is equal to $\sqrt{k}$, and for the former distribution, it is $\sqrt{k/\theta}$ (see equations 12 and 15); i.e., it is larger by a factor of more than 10. In the calculations below, I use $\sigma(a_{cut}) = 10\sqrt{k}$.

Maximum Magnitude

For the gamma distribution, I calculate the maximum magnitude as (Kagan, 1993, equation 14)

$$M_{max} = \log_10 \left( \frac{1}{\Delta} \right) + \log_10 \left( \frac{1}{2} \right) + \log_10 \left( \frac{a_{max}}{1000} \right)$$

Figure 2. Distribution of the yearly numbers of earthquakes in southern California in 1932 to 1994 (solid curve). Three approximations of the observed numbers are plotted: dashed curve is a negative binomial distribution with $\theta = 0.006$, dashdot curve is a negative binomial distribution with $\theta = 0.05$, and a dotted curve is a Poisson distribution.

Figure 3. Hypothetical magnitude distributions giving the same seismic moment rate. If the actual distribution is gamma (5) with $m_{max} = 7.5$ (case A), and the moment is ascribed to characteristic earthquakes with magnitude 7, then the rate of magnitude 7 and above earthquakes is overestimated by a factor of 6. If the actual distribution is gamma with a magnitude limit of 8.25 (case B), the overestimation is almost an order of magnitude.
where $\Gamma$ is a gamma function, $a_{\text{cut}}$ is the yearly rate of events above the cutoff level, $M_g$ is a geologic rate of deformation, and $\lambda_g = \log_{10}$. For the $\beta$ range of 0.5 to 0.8, $\Gamma(2 - \beta)$ is almost constant, changing from $-0.053$ to $-0.037$, respectively.

Table 1 lists the calculated maximum magnitudes $m_{\text{max}}$ for two cases: (a) $\beta$ value is universal over the world and equal to $2/3$ (Kagan, 1994, 1995b) and (b) is determined by the maximum likelihood method for each of the fault zones. For the geologic slip rates on major faults, $M_g$, I use Wesson's values. For southern California in the box $[32.5^\circ$ to $36.5^\circ$ N, $114.0^\circ$ to $122.0^\circ$ W], I assume the shear modulus $3 \times 10^{11}$ dyne/cm², slip rate 4.5 cm/yr, fault length 500 to 700 km, and thickness of seismogenic zone 10 to 15 km. The numbers of events ($k$) are counted for magnitude $m \geq 3$ ($M_{\text{cut}} = 10^{2.0}$ dyne cm); to obtain $a$, the numbers $k$ should be divided either by 62- or 50-yr span of the catalog.

For the standard error in $m_{\text{max}}$ estimates due to earthquake number random fluctuations, I take

$$\sigma_m(a_{\text{cut}}) \approx \frac{20 \log e}{3(1 - \beta)\sqrt{k}}$$

(see above). For $\beta = 2/3$, $\sigma_m \approx 8.7/\sqrt{k}$. For variable $\beta$, we need also to consider the $m_{\text{max}}$ errors due to $\beta$ variation,

$$\sigma_m(\beta) \approx \frac{2\sigma_g}{3(1 - \beta)} \left[ 1.5m_{\text{max}} + 16 - \log M_{\text{cut}} - \frac{\log e}{\beta(1 - \beta)} \right]$$

and the total standard error (see more in Kagan, 1995b),

$$\sigma_m = \sqrt{\sigma_m^2(\beta) + \sigma_m^2(a_{\text{cut}})}$$

(see Table 1).

Many of the $m_{\text{max}}$ confidence limits in Table 1 are very broad, due to large variations of the geologic rate and an insufficient number of earthquakes. The application of a similar procedure to global seismicity (Kagan, 1995b) yields the maximum magnitude estimates with significantly smaller confidence limits: the numbers of large earthquakes in the Flinn-Engdahl seismic regions are sufficiently high to render more informative results. Very large, mechanically implausible values of the maximum magnitude in Table 1 can be ruled out on the strength of the prior information—the worldwide estimate of $m_{\text{max}} = 8.0 - 9.0$ (Kagan, 1994, 1995b). Moreover, because of the exponential decay of the density function (5), the maximum magnitude earthquakes for the gamma distribution occur infrequently. In almost all of the cases, the range of the $m_{\text{max}}$ variation due to the geologic slip-rate uncertainties and errors connected with the seismic activity and $\beta$ fluctuations includes the worldwide estimate.

For $\beta = 2/3$, the case I consider to be more relevant (Kagan, 1995b), the only fault zone that shows a significantly different $m_{\text{max}}$ is the San Andreas fault. In the Seismicity Statistics section, I argue that the width of the San Andreas zone may be understated: the whole southern California seismicity is a part of the San Andreas fault system. If we use similar calculations for southern California, the resulting $m_{\text{max}}$ value limits are consistent with the worldwide result (see Table 1). Anderson and Luco (1983) obtain similar values for $m_{\text{max}}$ in California using analogous arguments. Thus the null hypothesis $H_0$—that the size of an earthquake is distributed according to the gamma distribution (5) with universal values for its parameters—cannot be rejected.

Site-Specific Calculations

The discussion above is based on the point model of an earthquake: we assume that the moment release is concen-

Table 1
Calculation of Maximum Magnitude $m_{\text{max}}$ for California Faults

<table>
<thead>
<tr>
<th>Fault</th>
<th>Years</th>
<th>Event Number</th>
<th>$k$</th>
<th>$\beta$</th>
<th>Geologic Rate $M_g$</th>
<th>Maximum Magnitude</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elsinore</td>
<td>1932-94</td>
<td>682</td>
<td>2/3</td>
<td>1.6-9.7</td>
<td>7.8-9.4</td>
<td>±0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1944-94</td>
<td>438</td>
<td>0.72±0.04</td>
<td>8.7-10.6</td>
<td>±0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garlock</td>
<td>1932-94</td>
<td>198</td>
<td>0.52±0.05</td>
<td>9.8-10.5</td>
<td>±0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1944-94</td>
<td>169</td>
<td>0.81±0.07</td>
<td>8.6-9.2</td>
<td>±0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newport</td>
<td>1932-94</td>
<td>519</td>
<td>2/3</td>
<td>0.03-1.7</td>
<td>4.6-8.1</td>
<td>±0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1944-94</td>
<td>138</td>
<td>0.81±0.07</td>
<td>6.3-12.6</td>
<td>±2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>San Jacinto</td>
<td>1932-94</td>
<td>1,533</td>
<td>2/3</td>
<td>7.0-11.4</td>
<td>8.4-8.8</td>
<td>±0.2</td>
<td></td>
</tr>
<tr>
<td>San</td>
<td>1932-94</td>
<td>1,215</td>
<td>0.60±0.02</td>
<td>7.7-8.1</td>
<td>±0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andreas</td>
<td>1944-94</td>
<td>2,546</td>
<td>2/3</td>
<td>47-137</td>
<td>9.6-10.6</td>
<td>±0.2</td>
<td></td>
</tr>
<tr>
<td>Southern</td>
<td>1932-94</td>
<td>1,231</td>
<td>0.56±0.01</td>
<td>8.3-9.0</td>
<td>±0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>1944-94</td>
<td>11,335</td>
<td>0.60±0.006</td>
<td>8.5-9.1</td>
<td>±0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For $\beta = 2/3$, the case I consider to be more relevant (Kagan, 1995b), the only fault zone that shows a significantly different $m_{\text{max}}$ is the San Andreas fault. In the Seismicity Statistics section, I argue that the width of the San Andreas zone may be understated: the whole southern California seismicity is a part of the San Andreas fault system. If we use similar calculations for southern California, the resulting $m_{\text{max}}$ value limits are consistent with the worldwide result (see Table 1). Anderson and Luco (1983) obtain similar values for $m_{\text{max}}$ in California using analogous arguments. Thus the null hypothesis $H_0$—that the size of an earthquake is distributed according to the gamma distribution (5) with universal values for its parameters—cannot be rejected.
treated at the epicenter, thus we often consider probabilities of very large earthquakes in a small spatial window that is clearly insufficient to contain fully the earthquake source region. The reason that the results of such an analysis yield reasonable values for the maximum magnitude is that the geologic rate and the average rate a scale similarly in equation (18), thus $M_{\text{max}}$ can be defined even for an infinitesimal box. For a large earthquake with the epicenter in the window, only a small part of the moment would be released in the box, but it is compensated by other earthquakes that originate in different sections of a fault system and rupture through the window of interest.

I use the earthquake point model because this is the only model with the magnitude-frequency relation documented by reliable and extensive data. Whereas the seismic moment exhibits a power-law scaling over a broad range, the distribution of a rupture length or a coseismic slip apparently does not follow such a simple scaling (see below).

However, I attempt to use an extended model of an earthquake source to see whether any significant changes in the conclusions are warranted. The first problem that needs to be addressed is whether an earthquake that starts on a relatively small fault segment can develop into a very large event possibly intersecting major tectonic discontinuities. As I mentioned above, the Landers 1992 earthquake is an example of such an event, and simple mechanical and paleoseismic arguments suggest that very large earthquakes may be necessary to explain displacements accumulated during the long geological history of relatively small faults. Sieh et al. (1989) indicate that large earthquakes on the San Andreas fault do not always stop at the Big Bend. Dolan et al. (1995, see items 48 and 50 in their reference section) discuss the possibility of an $m = 8$ event in the Los Angeles area or of a large earthquake combining several modes of deformation. Two of the largest earthquakes of the twentieth century in southern California ($m = 7.5$ Kern County and $m = 7.3$ Landers) occurred on the fault systems that were not appropriately recognized before the earthquakes. Thus, unless it can be convincingly shown that a very large earthquake cannot rupture on two or more faults, such a possibility needs to be analyzed.

Ideally, it would be interesting to determine a slip distribution due to earthquakes at a specific site (possibly around a fault trace—cf. Anderson and Luco, 1983). Unfortunately, such a program encounters serious difficulties. First, earthquakes do not occur exactly in the same spot as previous large events, and they often rupture through a broad deformation zone (Johnson et al., 1994); thus, the site specification is ambiguous. Second, there is the absence of reliable data on geometry of earthquake source; the relation between the length of the rupture and the seismic moment is a subject of controversy (Scholz, 1994a, 1994b; Romanowicz, 1994, and references therein).

For illustration, I assume $d = 2$ (see equation 2) in the magnitude range from 6 to 8.5 (Scholz, 1994a), the length of the fault zone $L = 700$ km, the thickness of seismic zone $w = 15$ km, and the coseismic slip $u$ proportional to the length of the rupture: $u = 5 \times 10^{-5} l$ (Scholz, 1994b). The distribution of the coseismic slip at a particular site can then be deduced from the distribution of the moment. To simplify the calculations, I use the Pareto distribution of seismic moment truncated at both ends (equations 6 through 8 in Kagan, 1993; or the case $i = 2$ in Anderson and Luco, 1983, p. 474). The distribution density for the seismic moment $M$, $l$, and $u$ is

$$\phi(X) \propto X^{-1-\kappa},$$

(22)

where $\kappa = \beta$ for the moment distribution, $\kappa = \nu = d\beta$ for $l$, and $\kappa = \zeta = d\beta - 1$ for the slip. The slip distribution is site specific; we take into account the fact that strong earthquakes have a higher probability of rupturing through a site (Anderson and Luco, 1983), because of their larger length. For $\beta = 2/3$, the values of exponents are $\nu = 4/3$ and $\zeta = 1/3$; for the magnitude distribution at a site, the exponent is $1/4$ (Anderson and Luco, 1983). Thus for this choice of parameters, large earthquakes rupture a fault trace at a rate only slightly lower than that of smaller earthquakes. If we take $d = 3/2$ in equation (2), i.e., assume that log $(l)$ is proportional to the first power of magnitude, all large earthquakes rupture through a site at an equal rate.

The average of each quantity can be calculated as follows:

$$\bar{X} = \frac{\kappa X_1}{1 - \kappa} \frac{Y^\kappa}{Y^\kappa - 1} [Y^{-\nu} - 1].$$

(23)

where $Y = X_2/X_1$ and $X_1$ and $X_2$ are the variable’s $X$ lower and upper limits, respectively. The calculated values for earthquake geometrical parameters are listed in Table 2. I calculate $\lambda$ and $T$, the rate and the recurrence time for earthquakes at a site, by

$$\lambda = \lambda \bar{u} L.$$

The seismic moment rate $\dot{M}$ is $\mu \lambda \bar{u} l w$. It is possible that the coseismic slip values for the largest events in Table 2 are excessive; Scholz (1994b) suggests that the slip saturates for earthquakes with the rupture length over 200 km. If we cap the slip at 12.5 m for the $m \geq 8.0$ events (cf. Sieh, 1978), it would lead to the decrease of $T$ by a factor of about 2.

The obtained $T$ values agree with the available data on the earthquake occurrence on the San Andreas fault at Pallette Creek (Anderson and Luco, 1983; Sieh et al., 1989). Similar calculations for the Newport fault yield the return rate for $m \geq 8$ earthquakes 6 to 300 Kyr, the values analogous to those obtained through the extrapolation of the magnitude-frequency relation.
Table 2

<table>
<thead>
<tr>
<th>n</th>
<th>$\log M$ (dyne cm)</th>
<th>Length $l$ (km)</th>
<th>Slip $w$ (m)</th>
<th>Rate $\lambda$ (eq/yr)</th>
<th>$\ell$ (km)</th>
<th>$a$ (m)</th>
<th>$M_{10}^{11}$ (dyne cm/yr)</th>
<th>$\ell = 1/\lambda$ (yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>25</td>
<td>6.67</td>
<td>0.33</td>
<td>0.199</td>
<td>13.0</td>
<td>0.82</td>
<td>9.5</td>
<td>272</td>
</tr>
<tr>
<td>7</td>
<td>26.5</td>
<td>37.5</td>
<td>1.9</td>
<td>0.0199</td>
<td>72.9</td>
<td>4.69</td>
<td>30.2</td>
<td>482</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>210.8</td>
<td>10.5</td>
<td>0.0015</td>
<td>308.4</td>
<td>16.3</td>
<td>34.3</td>
<td>1513</td>
</tr>
<tr>
<td>8.5</td>
<td>28.75</td>
<td>500</td>
<td>25.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

What is the bias in the estimates for the rate of expected characteristic earthquakes? The method used by Wesnousky (1994) ascribes the entire slip rate on a fault to the effects of characteristic earthquakes alone. If creep and earthquakes of other sizes contribute to the slip, the rate of characteristic earthquakes will be significantly overestimated. Kagan (1993) shows that if earthquakes actually follow the G-R relation, the estimated rate of characteristic events could exceed their actual rate 5 to 10 times. This relationship is illustrated in Figure 3, which shows three hypothetical magnitude-frequency relations all having the same moment release rates. The curve with a sharp "knee" is a characteristic earthquake distribution, which models almost all the moment release as due to magnitude 7 earthquakes with a mean recurrence rate of 10 yr. The other two sets of curves show the G-R and gamma distributions, the latter with limiting magnitudes of 7.5 (A) and 8.25 (B). The characteristic assumption would overestimate the rate of magnitude 7 earthquakes by almost an order of magnitude for the higher cutoff and by a factor of 6 for the lower threshold. (Even for the characteristic distribution shown in Fig. 3, 16.7\% of the seismic moment rate is released by noncharacteristic earthquakes.)

From (18),

$$a_{out} \propto 10^{-3(1-\beta)\mu_{max}/2};$$

i.e., for $\beta = 2/3$, the seismic activity should increase by $\sqrt{10}$ for each unit decrease in the maximum magnitude (see Fig. 3). Figure 5 by Anderson et al. (1993) also illustrates this behavior.

The only comprehensive, large-scale application of the characteristic model is by Nishenko (1991), who gave specific probabilities for characteristic earthquakes in each of about 100 zones of the circum-Pacific rim during the periods of 5, 10, and 20 yr beginning on 1 January 1989. Only the first prediction can be tested so far. Kagan and Jackson (1995) have tested it and found that (a) the Nishenko model predicts too many characteristic earthquakes—9.2 predicted events versus two earthquakes, which have actually occurred during the last 5 yr and strictly satisfy the prediction criteria, and (b) the observed size distribution of earthquakes is inconsistent with the characteristic hypothesis: instead of a deficit of earthquakes below and above the characteristic limit, earthquake numbers during the last 5 yr are distributed according to the standard G-R relation. The prediction failed most of the statistical tests, and it is apparent that the characteristic hypothesis significantly exaggerates the seismic hazard. The statistical test failure of the 5-yr prediction (Nishenko, 1991) contradicts Wesnousky's (p. 1958) suggestion that the characteristic model has a predictive potential at least for the next few decades. Contrary to that conjecture, the null hypothesis that includes the Poisson temporal occurrence of large earthquakes and the G-R relation is shown to be superior in predicting future seismicity (Kagan and Jackson, 1995) for the first 5 yr after Nishenko's prediction.

Nishenko's predictions are based to a large degree on historic and geologic (paleoseismic) earthquake occurrence dates. Only for 17 out of 98 zones, a "direct" method, similar to the Wesnousky model has been applied. To rigorously test the characteristic model, one should compare two quantitative data bases available now: the seismicity level and the tectonic deformation rate. Contrary to historical earthquake data and paleoseismic results, possible numerical errors for instrumental catalogs and geodetic deformation rates can be quantitatively evaluated.

As mentioned in the Seismicity Statistics section, the possibility of a selection bias and other systematic errors cannot be ruled out when dealing with the existing data; however, the forward predictions based on the characteristic model can be rigorously tested within a reasonable time, if applied to a large area such as the circum-Pacific earthquake belt. Since many of the calculations of earthquake hazard are based presently on the characteristic hypothesis, and their results have very serious societal and financial consequences (cf. Working Group, 1995), it is important that such tests be carried out. If, for example, the comparison of the seismic activity and tectonic deformation rate yields results similar to those of Table 1, there will be strong evidence that the maximum moment earthquakes have a uniform size over the earth, i.e., one of the basic assumptions of the characteristic hypothesis would not be validated by this test.

To illustrate what inconsistency and difficulties the ap-
plication of the characteristic model leads to, let us consider
the recent evaluation of a seismic risk for the Los Angeles
metropolitan area by Dolan et al. (1995) and by Hough
(1995). Using the geodetic and geologic estimates for the
deflection rate in the Los Angeles area, Dolan et al. (1995)
argue that to accommodate the strain, 15 to 20 earthquakes
similar to the San Fernando (1971) or the Northridge (1994)
are needed during the last 200 yr, whereas only two earth-
quakes listed above are known to have occurred. Even if one
assumes that stronger earthquakes with magnitude 7.2 to 7.5
are possible in the Los Angeles area, the return time for these
and smaller events is still too high (Dolan et al., 1995;
Hough, 1995) to explain satisfactorily the seismic history of
the area.

The Caltech catalog contains 819 earthquakes $m \geq 3.5$
during 1932 to 1995 in the box with corners $34.71^\circ$ N,
118.89$^\circ$ W; 34.18$^\circ$ N, 117.45$^\circ$ W; 33.40$^\circ$ N, 117.87$^\circ$ W; and
33.92$^\circ$ N, 119.32$^\circ$ W. Using the deformation rate suggested
by Hough (1995, p. 212), from $7.6 \times 10^{-4}$ to $10.1 \times 10^{-4}$
dyne cm/yr and $\beta = 2/3$, I obtain (see equation 18) $m_{\text{max}} =
8.0$ to $8.3 \pm 0.3$; i.e., the values are consistent with the
worldwide estimate (Hough, 1994, 1995b; see also Table 1).

As I explained in earlier sections, this value of $m_{\text{max}}$ does not
mean that the fault rupture of such an earthquake should be
fully contained inside the box, the moment-frequency rela-
tion for the box measures the occurrence rate of earthquakes
with epicenters inside the box, thus the spatial density rate
can be defined for any infinitesimal region.

An extrapolation of the moment-frequency relation sug-
jects that the expected number for $m \geq 6.5$ earthquakes in
200 years is 2.2 to 3.5 (depending on assumed value of $\beta$),
which is in a rough agreement with the observed seismicity.
Assuming $m_{\text{max}} = 8.5$, the return time for very large earth-
quakes ($m \geq 8.0$) that release most of the accumulated strain
is about 3.5 to 6 Kyr (the last computation takes into account
the tail decay for large earthquakes, as shown in Fig. 3).

In conclusion, I emphasize that the null and character-
istic hypotheses do not compete on an equal basis. The G-R
relation has been extensively used to approximate the earth-
quake size distribution ranging over many orders of
magnitude. It satisfactorily describes the seismicity of the
Earth’s tectonic plates and smaller tectonic regions, as well
as microearthquakes in mines and acoustic emission events
during rock specimen testing. Therefore, we need only to
prove that the model is not strongly rejected by the available
evidence (cf. Anderson, 1992). The characteristic model, on
the other hand, still needs to be validated. The characteristic
hypothesis has undergone many modifications since it was
proposed a decade or so ago, and its present content is un-
clear. In particular,

• It seems that the fault segmentation, which is a fundamen-
ental feature of the characteristic model, cannot be defined
in any objective, quantitative manner. If the segmentation
cannot be unambiguously carried out, various rupture sce-
narios can always be fitted to available historical and in-
strumental data; however, such constructs have little if any
predictive power (Kagan and Jackson, 1995).
• If several scenarios for a fault segmentation are feasible,
then the absence of earthquakes slightly smaller than the
characteristic limit (corresponding to a horizontal interval
in the characteristic curve in Fig. 3) is to be questioned. It
implies that there is no fundamental difference between
smaller earthquakes and characteristic events.
• The cascade model (Working Group, 1995) undermines
another basic aspect of the characteristic hypothesis—absence
of earthquakes larger than the characteristic limit.

Thus, I propose to answer the question “The Gutenberg-
Richter or characteristic earthquake distribution, which is it?” by
Wesnousky (1994) the following way: The statistical
analysis supporting the characteristic model can be shown
to be significantly deficient; the null hypothesis has not been
sufficiently analyzed and shown to be untenable. Accumu-
lating evidence calls into question the whole basis of the
characteristic hypothesis, at least in the form it has been
formulated and practiced.

Acknowledgments

This research was supported in part by the National Science Foundation
through Cooperative Agreement EAR-8920136 and by USGS through Co-
operative Agreement 14-08-0001-A0899 to the Southern California Earth-
quake Center (SCEC) and in part by USGS through Grant 1434-94-G-2425.
I thank D. D. Jackson of UCLA; D. Vere-Jones of Victoria University,
Wellingon; S. E. Hough of USGS, Pasadena; and Y. Ben-Zion of Harvard
University for useful comments. Publication 4302, Institute of Geophysics
and Planetary Physics, University of California, Los Angeles. Publication
219, SCEC.

References

on earthquake occurrence relation, Bull. Seism. Soc. Am. 73, 471–496.
Anderson, J. G., M. Ellis, and C. DePol0 (1993). Earthquake rate analysis,
Tectonophysics 218, 1–21.
for larger or more frequent earthquakes in the Los Angeles metropoli-
тан region, Science 267, 199–205.
Frohlich, C. and S. D. Davis (1993). Teleseismic b-values; or, much ado
about 1.0, J. Geophys. Res. 98, 631–644.
southern California, with aftershocks removed, Poissonian?, Bull.
Seism. Soc. Am. 64, 1363–1367.
Gradshteyn, I. S. and I. M. Ryzhik (1980). Table of Integrals, Series, and
Seism. Soc. Am. 77, 141–159.
a possible fractal distribution of rupture size, Science 267, 211–213.


Institute of Geophysics and Planetary Physics
University of California
Los Angeles, California 90095-1567

Manuscript received 4 May 1995.