SEISMOLOGICAL
RESEARCH
LETTERS

Volume 61, Number 2, April - June 1990

Estimation theory for peak ground motion ..............................................

G.-B. Ou and R. B. Herrmann 99

Observational test for wave propagation effects in local earthquake
seismograms .............................................................................................

D. E. Williams and C. E. Langston 109

Characteristics of seismically induced liquefaction sites and features
located in the vicinity of the 1886 Charleston, South Carolina earthquake
........................................................................................................

D. Amick, G. Maurath and R. Gelinas 117

Research Note
The probability of a major earthquake in the eastern United States ...........

M. S. Sibol, J. A. Snoke, G. A. Bollinger, M. C. Chapman, J. B. Birch,
and A. C. Johnston ............................................................................

131

Comments on "The probability of a major earthquake in the eastern
United States" by M. S. Sibol, J. A. Snoke, G. A. Bollinger, M. C.
Chapman, J. B. Birch, and A. C. Johnston .............................................

J. E. Beavers and V. R. Uppuluri 135

Monthly Listings - PDE ........................................................................

USGS/NEIS
(Inside Back Cover)
ESTIMATION THEORY FOR PEAK GROUND MOTION

Gwo-Bin Ou and Robert B. Herrmann

Department of Earth and Atmospheric Sciences
Saint Louis University
3507 Laclede Avenue
Saint Louis, MO 63103

ABSTRACT

The application of estimation theory for predicting peak ground motion is critically examined in order to be more precise in its application. Estimation theory relates peak ground motion to the duration and spectrum of the signal. Using vertical component data from the Eastern Canada Telemetered Network, at distance range of 100-1000 km, we find that a duration must be defined by the interval where the cumulative energy of the main signal increases linearly, here between 5% and 75% of the cumulative power. This duration, when used with the spectra within this window, adequately replicates observed peak motions. This duration used differs significantly from that used by Herrmann (1985) and Toro and McGuire (1987) beyond 500 km. The estimation theory is extended to estimate confidence limits on the peak motion. Finally, the relation between various spectral level estimators, linear, logarithmic, and RMS, is considered to point out the need for consistency in spectral level estimation using smooth models.

INTRODUCTION

Theoretical estimation of seismic motion as a function of source strength and anelastic attenuation is often expressed in terms of a source spectral scaling (Hanks and Thatcher, 1972; Hanks, 1979) and frequency dependent absorption (Lomnitz, 1957; Futterman, 1962). The observations of interest to ground motion seismology, however, are in the time domain. A simple technique, random process theory (Boore and Atkinson, 1987) or random vibration theory (Boore, 1983), relates the two domains for applications requiring only the peak amplitude (Cartwright and Longuet-Higgins, 1956; Udwadia and Trifunac, 1974). Hanks and McGuire (1981), Boore (1983), and Atkinson (1984) used such an approach to predict the peak ground motion. Boore and Joyner (1984) extended this theory for use in estimating response spectra. They pointed out the importance of oscillator duration when trying to model response spectra. Herrmann (1985) showed, using synthetic seismogram techniques, how this theory can be extended to large distances for waves propagating in a crustal wave guide. Toro and McGuire (1987) used appropriate equations by Hanks and McGuire (1981), Boore (1983), Herrmann and Kijko (1983), and Herrmann (1985) to investigate ground motion characteristics. To use this technique, both the ground amplitude spectrum and the signal duration have to be specified. In general, the peak motion estimation assumes a random and stationary signal. The seismic signal is, however, far from stationary. Finding a duration within which the signal approximates a stationary process is needed. The theoretical level which can match the theoretical use also has to be considered. Since a random process is assumed, the predicted peak amplitude is required to give a confidence interval associated with its probability from estimation theory.

ASYMPTOTIC DISTRIBUTION OF THE LARGEST VALUE

The most general way for predicting the largest value is that based upon a known distribution for an observation and upon the times of observations. If we assume that all observations are mutually independent and have the same probability distribution, the cumulative distribution function of the largest value can be expressed as

\[ F_n(y) = F^n(y) \]  \hspace{1cm} (1)

and the probability density function of the largest value is

\[ f_n(y) = nF^{n-1}(y)f(y) \]  \hspace{1cm} (2)

for \( n \) observations, where \( F(y) \) and \( f(y) \) are the cumulative distribution function and the probability density function of the largest value for a single observation. In the case of large \( n \), Gumbel (1954) simplified the calculation by an asymptotic approach. The mode, \( y_n \), of the largest value for \( n \) observations can be determined
approximately by
\[ F(\hat{y}_n) = 1 - \frac{1}{n}. \] (3)

A concentration of the largest value for \( n \) observations is defined by
\[ \alpha_n = n \cdot f(\hat{y}_n). \] (4)

The mode and the concentration are the two parameters of the asymptotic distribution. The asymptotic cumulative distribution function of the largest value for \( n \) observations is then written as
\[ F_n(y) = \exp\left(-\alpha_n(y - \hat{y}_n)\right) \] (5)

and its derivative is the asymptotic probability density function
\[ f_n(y) = \alpha_n e^{-\alpha_n(y - \hat{y}_n)} F_n(y) \] (6)

\( f_n(y) \) is asymmetrical about the mode \( \hat{y}_n \) and its values drop more slowly to zero at \( y = \infty \) than at \( y = -\infty \). For smaller \( \alpha_n \), \( f_n(y) \) is lower and spreads farther. An estimate of the mean and standard deviation for the asymptotic distribution can be derived by means of the moment generation function (Fisher and Tippett, 1928; Gumbel, 1944). The mean, \( \mu_n \), and the standard deviation, \( \sigma_n \), of the largest value for \( n \) observations are expressed as
\[ \bar{y}_n = \hat{y}_n + 0.57722 \frac{0.57722}{\alpha_n} \] (7)

and
\[ \sigma_n = \sqrt{\frac{\pi}{6}} \frac{1}{\alpha_n}. \] (8)

The mean \( \bar{y}_n \) is always larger than the mode \( \hat{y}_n \) by \( 0.57722/\alpha_n \). A confidence interval is used to estimate the largest value for \( n \) observations from this asymptotic distribution. From the asymptotic cumulative distribution function (5), a value \( y_c \) such that the probability of the largest value \( y \) less than \( y_c \) is \( c \) can be expressed as
\[ y_c = \hat{y}_n - \frac{1}{\alpha_n} \ln(-\ln c). \] (9)

Thus, a \((1-\beta)100\%\) confidence interval of the largest value \( y \) is given by
\[ y_{\beta/2} < y < y_{1-\beta/2}, \] (10)

where \( y_{\beta/2} \) and \( y_{1-\beta/2} \) are determined by (9). For example, to obtain a 95\% confidence interval we choose \( \beta = 0.05 \). Applying (9), we find
\[ y_{0.025} = \hat{y}_n - 1.305\alpha_n \quad \text{and} \quad y_{0.975} = \hat{y}_n + 3.676\alpha_n. \]

Thus, the 95\% confidence interval of the largest value \( y \) is
\[ \hat{y}_n - \frac{1.305}{\alpha_n} < y < \hat{y}_n + \frac{3.676}{\alpha_n}. \] (11)

**PEAK AMPLITUDE IN A RANDOM SIGNAL**

Consider a random signal, \( s(t) \), such that it is both the sum of an infinite number of sinusoidal waves with random phase and a stationary process. The probability density function of a local maximum \( y \) can be obtained from a ratio
\[ f(y) = \frac{M(y)}{N}, \] (12)

where \( M(y) \) is the density frequency of a local maximum \( y \) and \( N \) is the total density frequency of local maxima. The density frequency is an average number of the particular signal level in a unit time. From Cartwright and Longuet-Higgins (1956),
\[ M(y) = \frac{1}{(2\pi)^{3/2}} \frac{\Delta^{1/2}}{m_0 m_2^{1/2}} e^{-\eta^2/2} \]
\[ \cdot \left[ e^{-\eta^2/(2\beta^2)} + (\eta \beta) \int_{-\infty}^{\infty} e^{-x^2/2} dx \right] \]

and
\[ N = \frac{1}{2\pi} \left( \frac{m_4}{m_2} \right)^{1/2}, \] (14)

where \( m_k \) is the \( k \)'th moment of the power density spectrum of \( s(t) \), \( \Delta = m_0 m_4 - m_2^2 \), \( \eta = y/m_0^{1/2} \), and \( \delta = \Delta^{1/2}/m_2 \). The \( k \)'th moment of the power density spectrum is defined by
\[ m_k = \frac{1}{\pi} \int_0^\infty \omega^{k-1} S(\omega) \frac{1}{T} d\omega, \] (15)

where \( S(\omega) \) is the Fourier transform of \( s(t) \) and \( T \) is duration. It is noted that the root mean square amplitude of a signal is equivalent to \( m_0^{1/2} \) (Parseval's theorem). We now rewrite the probability density function of a local maximum \( y \) as
\[ f(y) = \frac{\eta(1-\epsilon^2)^{1/2} e^{-\eta^2/2}}{m_0^{1/2}} \]
\[ + \frac{1}{(2\pi m_0)^{1/2}} \left[ e^{-\eta^2/(2\beta^2)} - \eta(1-\epsilon^2)^{1/2} e^{-\eta^2/2} \int_{-\infty}^{\infty} e^{-x^2/2} dx \right], \] (16)
where
\[ \epsilon^2 = 1 - \frac{m_2^2}{m_0 m_4}. \] (17)

\( \epsilon \) is a measure of the width of the power density spectrum and its value ranges from 0 to 1. The cumulative distribution function of a local maximum \( y \) is then written as
\[ F(y) = 1 - \int_y^\infty f(\xi) \, d\xi. \] (18)

The effect of the width of the power density spectrum on \( \epsilon \) was studied by Udawadia and Trifunac (1974). If the spectrum is a narrow band centered at some frequency, \( \epsilon \) will approach 0. On the other hand, \( \epsilon \) approaches 1 when the spectrum infinitely broadens. According to (16), \( f(y) \) is a Rayleigh distribution when \( \epsilon = 0 \) and \( f(y) \) is a Gaussian distribution when \( \epsilon = 1 \). For \( \epsilon \) between 0 and 1, \( f(y) \) lies between the Rayleigh and Gaussian distributions. To estimate the peak amplitude in a random signal, all extrema, including local maxima and minima, have to be considered. Because a random signal is symmetrical about the mean level, each of extrema are regarded as an observation and have the same probability distribution if local minima are multiplied by a reverse sign. From the total frequency of local maxima (14), the number of extrema can be calculated by
\[ n = \frac{1}{\pi} \left( \frac{m_4}{m_2} \right)^{1/2} T, \] (19)

where \( T \) is duration of the signal. Referring to the previous section, the asymptotic distribution of the largest value for \( n \) extrema is formed by two parameters, the mode determined by (3) and the concentration (4), which are related to the number of extrema and the probability distribution for a single extremum. For a convenience of evaluating \( f(y) \) and \( F(y) \) for a single extremum, two cases, \( 0 \leq \epsilon < 1 \) and \( \epsilon = 1 \), are discussed.

For \( 0 \leq \epsilon < 1 \). An asymptotic approach is used to evaluate the integral in \( f(y) \) in (16). Consider a function
\[ g(u) = \int_u^\infty e^{-x^2/2} \, dx \] (20)
of the exponential type which converges toward zero at least as quickly as an exponential function with increasing \( u \). According to L'Hôpital's rule, we have an asymptotic relation of the form
\[ \frac{g'(u)}{g(u)} \approx \frac{g''(u)}{g'(u)} \] (21)
Thus, for large \( u \)
\[ \int_u^\infty e^{-x^2/2} \, dx \approx \frac{1}{u} e^{-u^2/2}. \] (22)
The probability density function for an extremum then becomes
\[ f(y) = \frac{y}{m_0} (1 - e^{-y^2/2m_0})^{1/2}. \] (23)

By (18), the cumulative distribution function for an extremum is
\[ F(y) = 1 - (1 - e^{-y^2/2m_0})^{1/2}. \] (24)

Note that (23) and (24) are applicable only for small \( \epsilon \) or at large \( y \). According (3) and (4), the distribution only at large \( y \) is concerned as the number of extrema increases. From (3) and (24), the mode for \( n \) extrema is
\[ y_n^* = m_0^{1/2} \sqrt{2 \ln[n (1 - \epsilon^2)^{1/2}]}. \] (25)

Relating the probability density function (23) for an extremum and the mode (25) for \( n \) extrema to (4), the concentration for \( n \) extrema is
\[ \alpha_n = m_0^{-1/2} \sqrt{2 \ln[n (1 - \epsilon^2)^{1/2}]]. \] (26)

The mean \( \bar{y}_n \) (7) and the standard deviation \( \sigma_n \) (8) of the peak amplitude in a random signal having \( n \) extrema are now given as
\[ \bar{y}_n = m_0^{1/2} \left( \sqrt{2 \ln[n (1 - \epsilon^2)^{1/2}]} + \frac{0.57722}{\sqrt{2 \ln[n (1 - \epsilon^2)^{1/2}]}} \right) \] (27)
and
\[ \sigma_n = \frac{m_0^{1/2}}{\sqrt{6 \ln[n (1 - \epsilon^2)^{1/2}]}}. \] (28)

It is obvious that the mode \( y_n^* \) and the mean \( \bar{y}_n \) are increased but the standard deviation \( \sigma_n \) is decreased as a function of the number of extrema \( n \), which is proportional to duration of the signal. The term \([n (1 - \epsilon^2)^{1/2}]\) is equivalent to \((m_2/m_0)^{1/2} T / \pi\), interpreted as the number of zero crossings in a random signal (Rice, 1944; 1945) and seen in formulas (e.g., Boore, 1983; Toro and McGuire, 1987). Thus, \( m_4 \) which is highly dependent on the upper tail of the spectrum is not used directly. In the past decade, seismologists used random process theory to study peak ground motions. Some referred the predicted value to the mode (e.g., Hanks and McGuire, 1981; Atkinson, 1984; Boore and Atkinson, 1987), and other preferred the mean (e.g., Boore, 1983; Herrmann, 1985; Toro and McGuire; 1987). Since this
theory is based on a random process, we would like to establish a confidence interval for the peak amplitude prediction. A (1-β)100% confidence interval for the peak amplitude can be easily established by (9) and (10). The 95% confidence interval for the peak amplitude is then expressed from (11) as

\[ m_0^{1/2} \left( \sqrt{2 \ln[n(1-e^2)^{1/2}]} - 1.305 \right) < y < m_0^{1/2} \left( \sqrt{2 \ln[n(1-e^2)^{1/2}]} + 3.676 \right) \]  

(29)

The uncertainty of the predicted peak amplitude is decreased as a function of the number of zero crossings, controlled by the duration and the width of the power density spectrum of the random signal.

\[ \epsilon = 1 \]. The above formulas for 0 ≤ ε < 1 are no longer valid in this case. ε = 1 is substituted directly into (16), the probability density function is reduced to

\[ f(y) = \frac{1}{\sqrt{2\pi}} m_0^{1/2} e^{-y^2/(2m_0)} \]  

(30)

and the cumulative distribution function, by the aid of (18) and (22), becomes

\[ F(y) = 1 - \frac{1}{\sqrt{2\pi}} m_0^{1/2} e^{-y^2/(2m_0)} \]  

(31)

for an extremum. Following a similar procedure as before, we obtain the mode

\[ \hat{y}_n = m_0^{1/2} \left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2} \]  

(32)

the concentration, using (3.9) of Ou (1990),

\[ \alpha_n = m_0^{-1/2} \left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2} \]  

(33)

the mean

\[ \bar{y}_n = m_0^{1/2} \left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2} + \frac{0.57722}{\left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2}} \]  

(34)

the standard deviation

\[ \sigma_n = \frac{\pi}{\sqrt{6}} \frac{m_0^{1/2}}{\left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2}} \]  

(35)

and the 95% confidence interval

\[ m_0^{1/2} \left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2} - \frac{1.305}{\left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2}} \]  

< \bar{y}_n < \left( \frac{2 \ln(n/\sqrt{2\pi})}{2 \ln(n/\sqrt{2\pi})} \right)^{1/2} + \frac{3.676}{\left( 2 \ln(n/\sqrt{2\pi}) \right)^{1/2}} \]  

(36)

for the peak amplitude in a random signal having \( n \) extrema. A similar form of (32) is seen in (6.20) of Cartwright and Longuet-Higgins (1956).

We have discussed the asymptotic distributions of the peak amplitude in a random signal for 0 ≤ ε < 1 and ε = 1. The formulas in both cases are equivalent at ε = 0.917, then a choice between them is decided. We also assume that \( \sqrt{2m_0^{1/2}} \) is a low limit of the mode for a small number of extrema. Under this assumption, we must have \( (1-\epsilon^2)^{1/2} \geq 2.7183 \) to use (25) and \( n \geq 6.8137 \) to use (32). Therefore, the asymptotic distribution of the peak amplitude is made as follows: (1) If ε=-0.917, we will use the formulas (25-29) in the first case, 0 ≤ ε < 1, and (1-ε^2)^{1/2} will be either 2.7183 or \( (1-\epsilon^2)^{1/2} \), whichever is larger. (2) If ε>0.917, we will use the formulas (32-36) in the second case, ε=1, and \( n \) will be the larger of 6.8137 or \( n \). Thus, the controlling factors are functions of the duration \( T \) and the spectral moments, \( m_k \).

VALIDITY IN SEISMIC SIGNALS

The process of making estimates of the peak amplitude from the spectrum has been used since Udawadia and Trifunac (1974). Hanks and McGuire (1981), Boore (1983), Atkinson (1984), Boore and Joyner (1984), Herrmann (1985), Boore and Atkinson (1987), and Toro and McGuire (1987) used this statistical technique in the prediction of ground motions for \( S \) waves, \( Lg \) waves, and response spectra. From their studies, we realize that there are two questions to be solved before using the statistical technique. First, the window within which ground motion can be treated as a random and stationary signal must be defined. Second, the theoretical spectrum model which can represent a real spectrum is needed. There are several definitions for the window in ground motion seismology. One common definition is based on the cumulative power. Trifunac and Brady (1975) defined the duration of strong ground motion as a time interval between 5% and 95% of the cumulative square amplitude, whereas 5% and 75% are adopted by Kennedy et al. (1985). Vanmarcke and Lai (1980) related the total ground motion intensity to the square of the peak ground motion to obtain the duration. McGuire and Hanks (1980), Hanks and McGuire (1981), and Boore (1983) defined the duration of \( S \) wave as the inverse of the source corner frequency. Herrmann (1985) and Toro and McGuire (1987) used a distance-dependent duration when they tried to extend the application from short distances to large distances. Usually, an overestimate of duration violates the assumption of a stationary process for ground motion and then underestimates the peak amplitude. In contrast, if we use an underestimate of duration, the source strength within the window used has to be deliberated. A theoretical
amplitude spectrum is always formed as a smooth function, but the actual signal amplitude spectrum is quite variable. Thus, the spectral level which can match the theoretical use has to be chosen.

We use the Eastern Canada Telemetered Network data of seven 1982 Miramichi earthquakes in New Brunswick, Canada, to examine the validity of estimation theory on the peak amplitude. The earthquakes and stations were described by Shin and Herrmann (1987). Epicentral distances range from 135 to 994 km. The main phases in the seismograms are \( L_g \) waves which travel with a velocity about 3.6 km/sec and have a noticeable duration in this range. In general, the main phase lasts longer at larger distance. It rises sharply and dies gradually. We analyze the main phase which begins

GSG1  
(218 km)

![Graph of Duration vs Distance](image)

Fig. 2. Duration of the ergodic window versus epicentral distance from our whole set of records. The dashed line shows the distance-dependent duration of Toro and McGuire (1987), where the source duration is neglected.

at a small level before it rises and ends at the same level after it diminishes. The cumulative power percentage curve of the main phase is constructed to determine the power change at any instant. We found that the cumulative power of the main phase increases linearly within a window from 5% to 75% of its total. Therefore, the signal can be treated as a stationary process within this window. Outside this window, the power increase is less and more gradual. Such a duration, defined from 5% to 75% of the cumulative power of the main phase, will be called the ergodic window if a random process is assumed. Figure 1 shows an example at station GSG for EQ1, where the epicentral distance is 218 km. The analyzed main phase is shown in the upper left. The ergodic window is the boxed region where the signal is represented as a stationary process. The upper right shows the cumulative power percentage curve. Two flat lines cut the cumulative curve at two instants when the cumulative power is 5% and 75% of its total. Between these two instants, the power increase is very consistent. The jagged curve in the lower portion of the figure is the amplitude spectrum of the signal within the ergodic window. The observed peak amplitude is shown in the top box at the center right, and the mode, mean, and 95% confident interval of the peak amplitude predicted from the signal amplitude spectrum are displayed in the middle box. Comparing the predicted values in the middle box to the observed value, the defined ergodic window seems to provide us a good choice for use.

Duration of the ergodic window versus epicentral distance for the whole set of records is shown in Figure 2. The distance-dependent duration used by Toro and McGuire (1987) is also plotted as a dashed line where the source duration is neglected. It is seen that both

---

103
Fig. 3. Validity of estimation theory for the signal within the ergodic window. Symbols for the mode, mean, and 95% confidence interval of the peak amplitude predicted from the signal amplitude spectrum versus the observed peak amplitude are shown at top left corner. Including some records of poor quality, 7 of 74 observed peak amplitudes lie outside the predicted 95% confidence interval.

Definitions of duration rapidly increase with distance from 100 to 200 km, but ours increases less rapidly than the Toro and McGuire (1987) relation at large distances. This results from our narrower window which may include a more stationary signal.

To examine the validity of our defined duration, the mode, mean, and 95% confidence interval of the peak amplitude predicted from the amplitude spectrum of the signal within the ergodic window are plotted against the observed peak amplitude for each of the 74 traces in the entire data set (Figure 3). The diagonal line denotes equivalent values between the predicted and observed peak amplitudes. Only 7 of 74 observed peak amplitudes exceed their predicted 95% confident limits. Two of these outliers are illustrated in Figures 4 and 5. Figure 4 shows the record, observations, and estimates at station WBO for EQ1. Because most of amplitudes are clipped, the predicted peak amplitudes exceed the clipped values as expected. Figure 5 shows that the peak amplitude of the record at station WBO for EQ3 is dominated by a single isolated peak which results in an underestimation of the peak amplitude. From evidence above, it appears that our defined ergodic window is a good choice for the use in estimation theory. At this point, the ability of estimation theory to predict the peak amplitude from the signal amplitude spectrum is based on the defined ergodic window.

For making a ground motion prediction, we are forced to ask which measure of the spectral level should be used to fit a theoretical spectrum to the signal spectrum to obtain appropriate results. Three estimates of the spectral level can be made using the signal logarithmic amplitude, amplitude, and power spectrum fits by a least-squares approach. Obviously, since random process theory uses Parseval's theorem, a fit to the power spectrum is appropriate but often a manual spectral fit is made using log-log plots of the spectra.

The first case fits the logarithmic amplitude spectrum. A smooth amplitude spectrum model similar to that of Boore (1983) is used. Its shape is controlled by distance, corner frequency, Q, f_max, and an instrument transfer function. We adjust the unknown parameters in the theoretical spectrum model to find a spectral level using the Marquardt method (1963). An estimated root mean square amplitude is then calculated from the spectral level by Parseval's theorem. Plotting the estimatedversus the actual root mean square amplitudes for the entire data set as shown in Figure 6, we find that the spectral level of the logarithmic amplitude fit underestimates the actual root mean square amplitude. The second case fits the amplitude spectrum. Figure 7 shows that the spectral level of the amplitude fit works better than that of the logarithmic amplitude fit but it is still underestimated the root mean square amplitude. The

Fig. 4. A test of estimation theory using station WBO for EQ1 at an epicentral distance of 702 km. This is a record of poor quality. Because most of amplitudes are clipped, the predicted peak amplitudes exceed the clipped values.
Fig. 5. A test of estimation theory using station WBO for EQ3 at an epicentral distance of 702 km. This is a record of poor quality. Because the peak of the record is isolated, the predicted peak amplitudes are underestimated.

Fig. 6. A comparison of the estimated to the actual root mean square amplitudes. The estimated root mean square amplitude is calculated from a smooth amplitude spectrum having a least-squares fit of the logarithmic amplitude spectrum.

Fig. 7. A comparison of the estimated to the actual root mean square amplitudes. The estimated root mean square amplitude is calculated from a smooth amplitude spectrum having a least-squares fit of the amplitude spectrum.

Fig. 8. A comparison of the estimated to the actual root mean square amplitudes. The estimated root mean square amplitude is calculated from a smooth amplitude spectrum having a least-squares fit of the power spectrum. The third case fits the power spectrum. Figure 8 shows that the root mean square amplitude estimated from the spectral level of the power fit is very close to the actual root mean square amplitude. Besides the comparison of the expected values in these three cases, the variance in the logarithmic amplitude fit is further larger than those in the other two fits. Thus, the spectral level of the
From 5% to 75% of the cumulative power of the main phase is required to use this theory. One of the most important factors in predicting the peak amplitude is to correctly estimate the root mean square amplitude from the amplitude spectrum. Although the real signal amplitude spectrum should be used, a smooth theoretical amplitude spectrum is usually assumed in the prediction. From Parseval's theorem, a direct relation exists between the power spectrum and the root mean square amplitude. We have shown that the smooth amplitude spectrum which is defined such that its power spectrum is a least-squares fit to the signal power spectrum works very well in the estimation of the root mean square amplitude and hence in the prediction of the peak amplitude. If the logarithmic amplitude fit is used, the spectral level should be multiplied by a factor of 1.76. The accuracy for the amplitude fit, the multiplication factor is 1.16. The functional forms for predicting duration and spectral shape from local to regional distances can be theoretically derived (Ou and Herrmann, 1990) for a basic layered earth model.

In ground motion seismology, the predicted peak amplitude is sometimes represented by the mode (Hanks and McGuire, 1981; Atkinson, 1984; Boore and Atkinson, 1987) or by the mean (Boore, 1983; Herrmann, 1985; Toro and McGuire, 1987). Our study shows that there is no significant difference in using one or the other to estimate the observed peak amplitude. The mode and the mean will become close to each other if a concentrated distribution is present in estimation theory. Two factors affect the shape of the distribution. For a large number of extrema and a narrow-band signal, the predicted peak amplitudes tend to cluster at the mode such that the mean closes to the mode. Therefore, comparing to S waves, Lg waves have a more precise prediction due to a longer duration and a narrower shape of the amplitude spectrum. Besides a single predicted value, the mode or the mean, the accuracy of the prediction is also useful. This can be indicated by the interval estimate. From the probability distribution of the peak amplitude, we have constructed a 95% confidence interval for the prediction. Except for the records of poor quality, the 95% confidence interval formula has provided us a good prediction for the peak amplitude. The concept of what the peak amplitude might be is required to formulate engineering design criteria for which allowance must be made for the uncertainties. Here, we provide the uncertainty of the peak amplitude for a fixed amplitude spectrum and a given duration. The variations of these two parameters can be used to analyze the uncertainties from site-to-site and event-to-event (Boore and Atkinson, 1987; Toro and McGuire, 1987).

ACKNOWLEDGMENTS

We thank B. J. Mitchell, W. V. Stauder, and K. D. Hutchenson for reviewing the manuscript. The comments from the anonymous reviewers were helpful. This research was supported by the U. S. Nuclear Regulatory Commission under Contract NRC-04-86-121 and the National Science Foundation under Grant CEE-8406577.
REFERENCES


Received November 2, 1989
Revised March 6, 1990
Accepted March 18, 1990