UNIVERSITY OF SOUTHERN CALIFORNIA

Department of Civil Engineering

ON UNIFORM RISK FUNCTIONALS WHICH DESCRIBE STRONG EARTHQUAKE GROUND MOTION: DEFINITION, NUMERICAL ESTIMATION, AND AN APPLICATION TO THE FOURIER AMPLITUDE OF ACCELERATION

by

John G. Anderson and M. D. Trifunac

Report No. CE 77-02

A report on research conducted under a contract from the U.S. Nuclear Regulatory Commission

Los Angeles, California

March, 1977

TABLE OF CONTENTS

ABSTRACT

A uniform risk functional (e.g. Fourier spectrum, response spectrum, duration etc.) has been defined so that the probability that it is exceeded during any earthquake is independent of the frequency of the seismic waves. Such a functional is derived by an independent calculation, for each frequency, of the probability that the quantity being considered will be exceeded at the frequency. Different aspects of the seismicity can control the amplitude of a uniform risk functional in different frequency ranges, and a uniform risk functional does not necessarily describe the strong shaking from any single earthquake. Two methods for calculating uniform risk functionals which represent a synthesis and generalization of many of the existing procedures are presented. Thus, much of the existing work on seismic risk could be regarded as a special case of the method presented here.

The new method for scaling of Fourier amplitude spectra of acceleration given by Trifunac (1976) has been applied in the analysis of uniform risk Fourier amplitude spectra. This scaling relationship leads to improved risk analysis by describing the scatter of amplitudes about the mean trend, as well as by allowing the amplitudes to be estimated in several frequency bands to derive uniform risk spectra.

It has been found that the small local events can significantly affect the level of a uniform risk spectrum, especially at high frequencies and that aftershocks of a large event also affect the level of a uniform risk spectrum (URS) slightly. The shape of a URS changes with location of the site with respect to seismic zone, thus emphasizing

the problems involved in scaling a fixed shape spectrum to an appropriate level.

This paper discusses in some detail the problem of whether to use an extended rupture zone in the model of the earthquake. For earthquakes where the epicenter is constrained to occur at a single point, we contour the difference of the spectral amplitudes associated with an extended rupture and a point process. This indicates that when there is no constraint on the direction of rupture, there is up to a factor of two which remains to be resolved by future improvements in risk models and scaling relationships.

In an application to finding the uniform risk Fourier amplitude spectrum at a realistic site, we find that this method is highly sensitive to the description of seismicity, and that distinct models of seismicity (all consistent with our current level of knowledge of an area) can give significantly different risk estimates.

CHAPTER I

DEFINITION AND NUMERICAL ESTIMATION OF UNIFORM RISK FUNCTIONALS

INTRODUCTION

The vulnerability of modern society to earthquake hazards calls for the development of rational methods in earthquake risk evaluation. With continued population growth and with increasing technological complexity of engineering structures in seismically active zones, it now appears useful to reexamine some fundamental principles which form the basis for seismic risk models.

The object of a seismic risk analysis is usually to describe, for the purposes of earthquake resistant design, the nature of possible future shaking. With the development of plate tectonics a significant advancement toward this goal has been achieved. Through plate tectonics, we now have a basic understanding of where, on a global scale, most earthquakes are likely to occur. Plate tectonics also provides a sound basis for identifying, for example, some regions with a large chance of a major earthquake in the near future (e.g. Kelleher et al., 1973). However, much work remains to be done before seismologists will be able to predict, on a time scale of tens of years, the largest earthquakes which will cause strong shaking at any one site, and provide a description of what that shaking will be. On a local scale, considerable judgment is involved in attempts to identify even the locations of future large events for seismic risk estimates.

Recognition of these uncertainties in seismic risk analyses on a local scale, plus the increasing dependence of modern society on complex installations requiring a high degree of seismic safety, has resulted in frequent upgrading of standards which govern earthquake resistant design during the past several decades. For that reason it appears desirable to formulate seismic standards and a corresponding seismic risk model whose validity may extend over a longer time span.

One purpose of this paper is therefore to construct a risk model which describes both the nature of strong shaking and the degree of confidence in the result. To accomplish this the model considered must incorporate a description of the uncertainties in the assumptions and empirical results used as inputs. The result then implies only the appropriate degree of confidence. With this approach only the probability functions which are input need to be upgraded with increased understanding of the phenomena involved; the methodology itself, if sufficiently general may remain in its original form.

DEFINITION OF A UNIFORM RISK FUNCTIONAL

The probability functions which are supplied to the model must describe numerous random processes which affect the amplitudes and duration of the strong earthquake ground motion at a site. These random processes, in addition to partially understood regional and physical processes, affect the number and size of earthquakes, their locations, and the relationship of source parameters and distance to the nature of shaking. Since these random processes affect the ground motion, they also affect functionals which describe various aspects of this motion (e. g. peak response of certain instruments, spectral amplitudes,

duration, etc.). Thus, for any site, there exists a relationship between each value in the range of such a functional of shaking and the probability that this value will occur. The form that such a probability function, $p[S(\omega)]$, say, must take is shown in Fig. 1.1. This function gives the probability that the functional $S(\omega)$, for example the spectral amplitude at frequency ω, will be exceeded at least once in Y years. At very low spectral amplitudes, $S(\omega)$ is certain to be exceeded by the background seismic noise. On the other hand, there is no chance that sufficiently large values of $S(\omega)$ will be exceeded, because the finite material strength of the earth sets an upper limit to the amount of energy which can be released in a single earthquake. Clearly, the probability of exceeding $S(\omega)$ must decrease monotonically for the intermediate values from certainty at small values of $S(\omega)$ to zero at large values of $S(\omega)$. Note that $p[S(\omega)]$ is a physically defined function for any point on the earth; the object of a risk analysis is to estimate this function. The derivative $\frac{-d\,p[S(\omega)]}{d\,S(\omega)}$ gives the density function of Y year extremes of $S(\omega)$.

In this paper, $S(\omega)$ is used in a very general manner to represent the amplitude of nearly any functional of strong ground motion at frequency ω. Thus, it can represent Fourier amplitude, response spectral amplitude, peak response of any particular instrument, or even the duration of strong shaking.

Ideally, one would like to estimate the probability function $p[S(\omega)]$ by explicit consideration of the physical processes (e.g. seismicity, the seismic source, and attenuation) which affect it. These physical processes all have an inherent physical variability which can only be

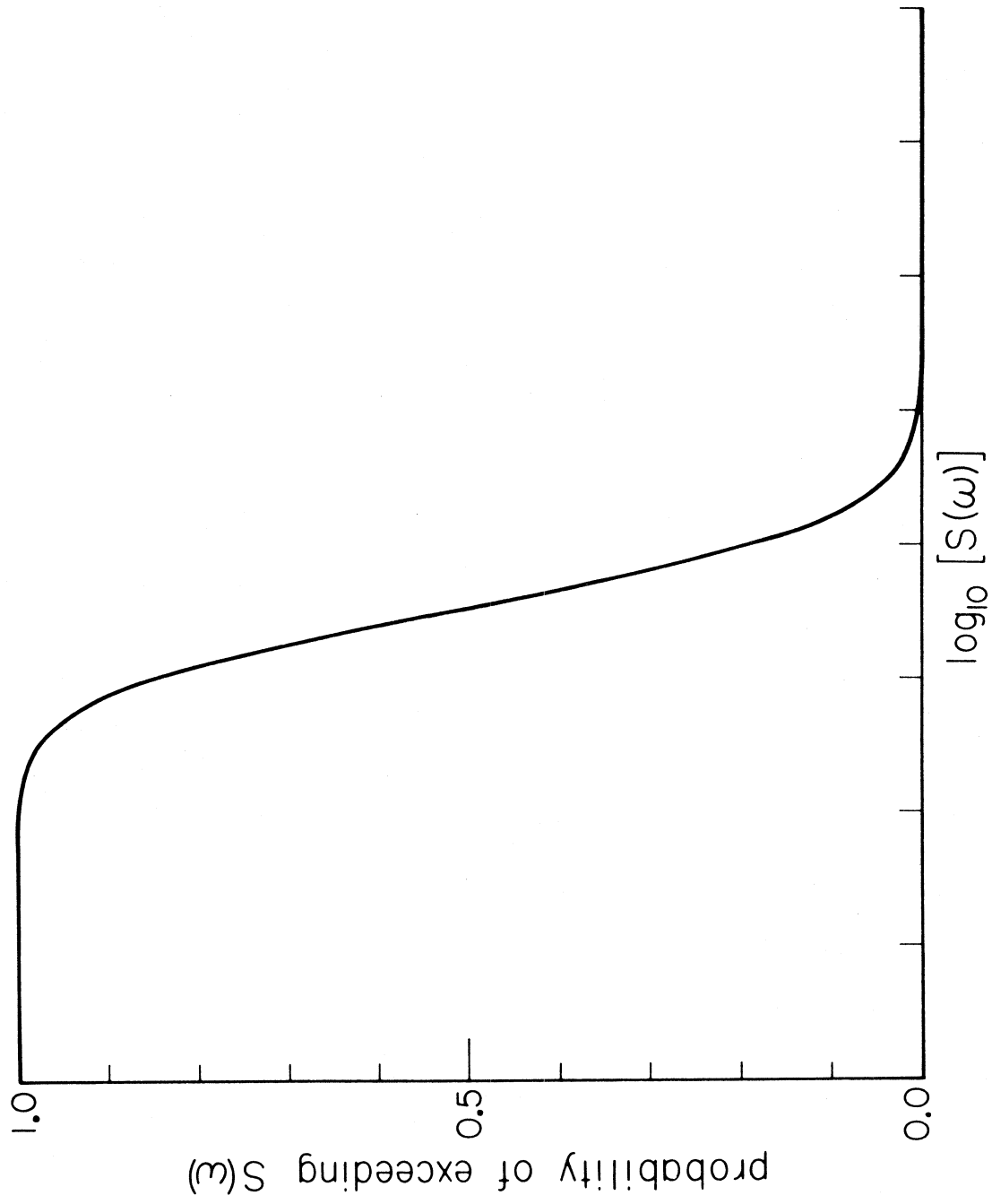

Figure 1.1. Example of the function p[S(ω)], the probability that S(ω) will be exceeded at least once in Y years.

described by probabilistic methods. Furthermore, we do not yet have a complete model of any of these physical processes. Thus, present estimates of probability functions such as that in Fig. 1.1 necessarily incorporate both the randomness in the physical processes and the lack of knowledge of the process. However, even though the functions used to estimate $p[S(\omega)]$ may change, the manner in which $p[S(\omega)]$ enters into a risk model will likely remain unchanged. If $p[S(\omega)]$ can be reliably estimated at each of several frequencies, then it is possible to construct a uniform risk functional using the method shown in Figs. 1.2 and 1.3. Figure 1.2 shows the probability of exceeding $S(\omega)$ at three frequencies: ω_1, ω_2, and ω_3. To construct a spectrum (or any other functional) which at each frequency has a probability of 0.1 that it will be exceeded, one simply reads from these probability functions the spectral amplitudes S_1, S_2, and S_3 which have the probability 0.1. From this, it is easy to construct a uniform risk spectrum, as shown in Fig. 1.3. Of course, one would use more than three frequencies to construct a complete uniform risk spectrum. When defined in this way, it is clear that a uniform risk functional has the property that the probability that it is exceeded during any event within the specified time period is independent of frequency.

Most studies of seismic risk can be regarded as a special case of the method just described above. Often the magnitude, some peak amplitude of ground shaking, the Modified Mercalli Intensity, or some other single parameter are used as a measure of some characteristic of shaking [$S(\omega)$ in our notation], and the return period is estimated as a function of that parameter (e.g. Cornell, 1968; Milne and

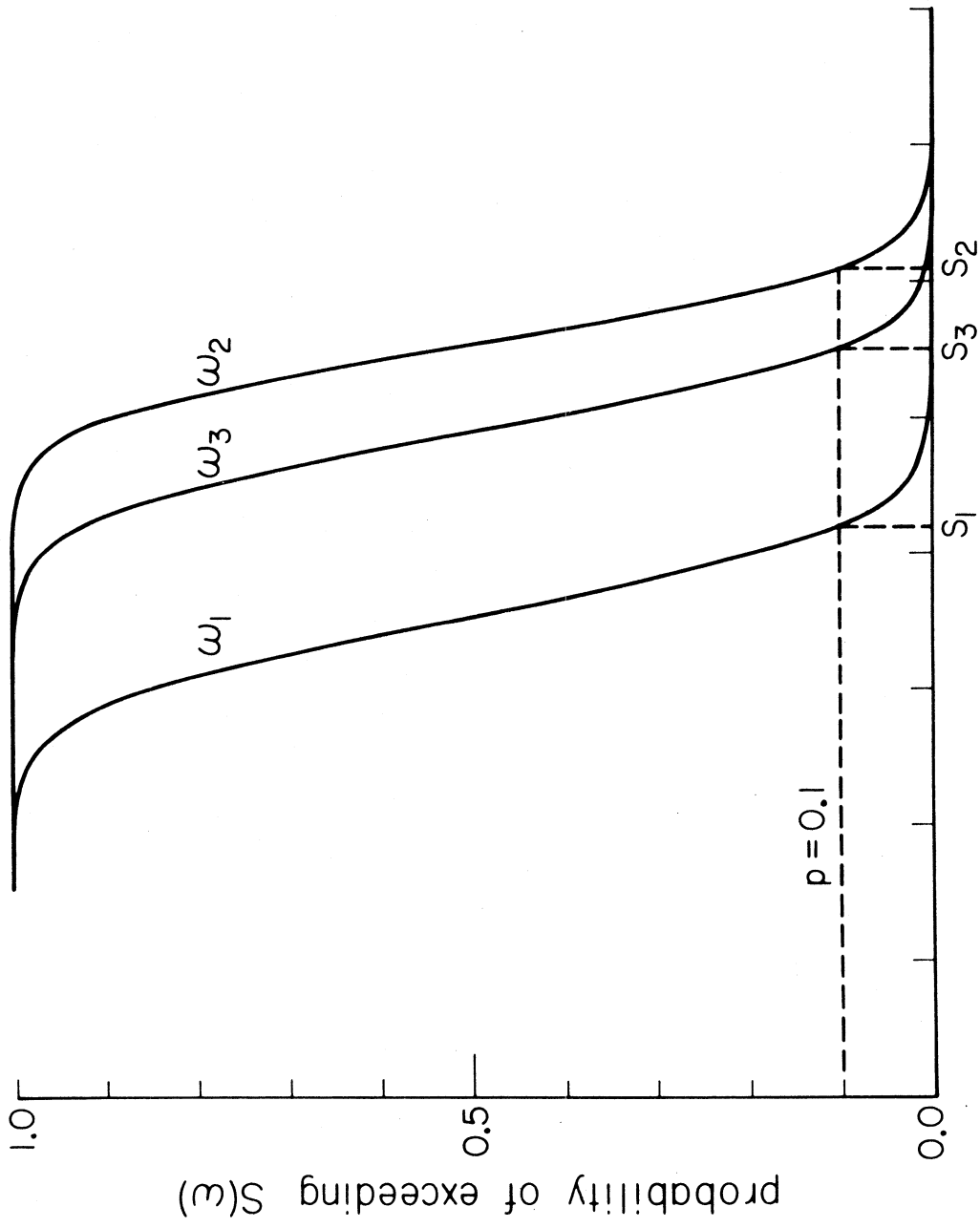

Figure 1.2. Example of the function p[S(ω)] at three different frequencies: ω_1, ω_2 & ω_3.

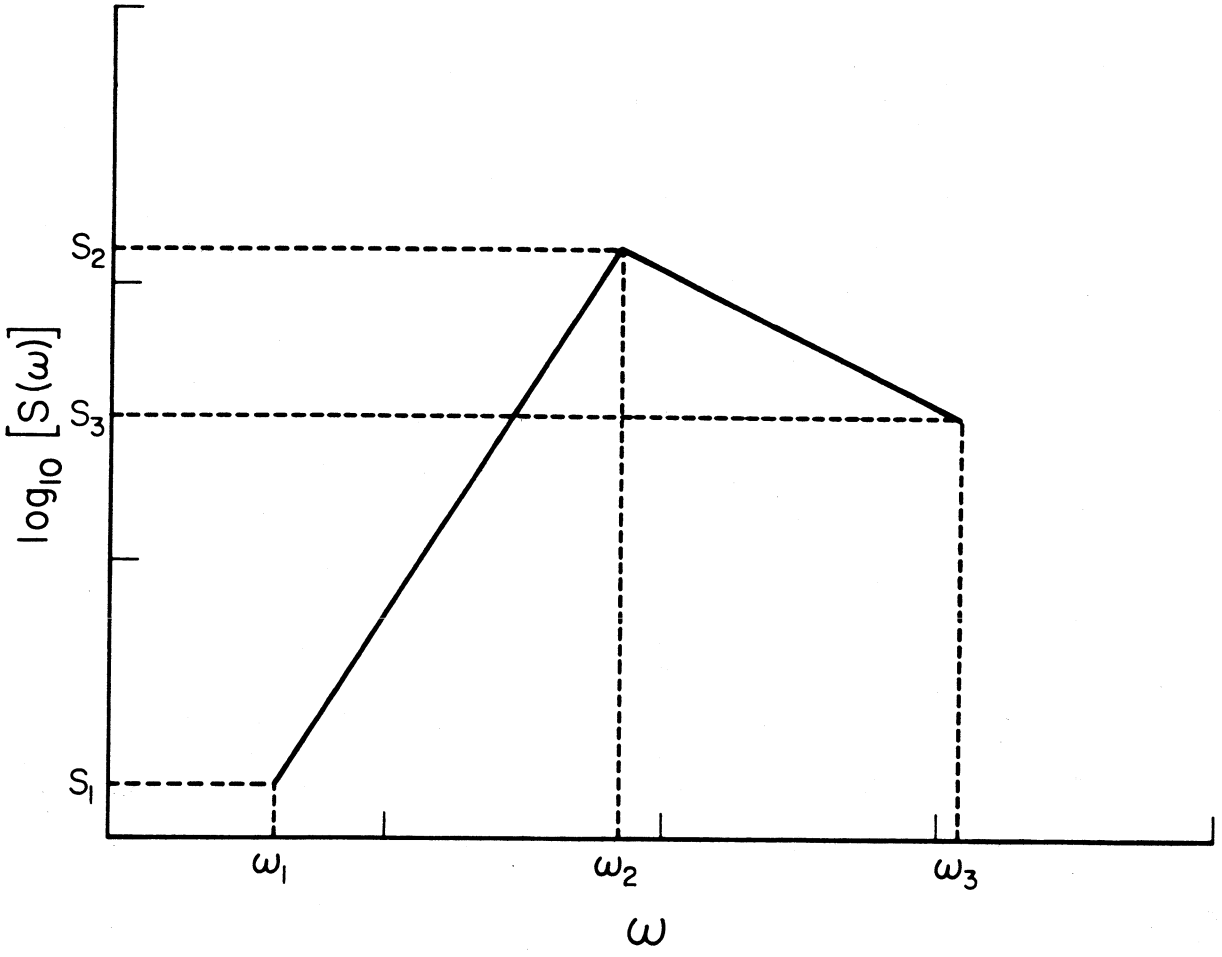

Fig. 1.3. A uniform risk functional derived from the functions $p[S(\omega)]$ shown in Fig. 1.2. At each frequency, the probability is 0.1 that $S(\omega)$ will be exceeded at least once in Y years. This figure uses only three frequencies for schematic purposes; in an actual calculation, one would use many frequencies and the resulting functional would have a smoother appearance.

Davenport, 1969; Liu and Fagel, 1972; Dalal, 1973; DeCapua and Liu, 1974; Douglas and Ryall, 1975; Algermissen and Perkins, 1976). The return period is, of course, closely related to the annual probability of excedance (e.g. Cornell, 1968). In such studies, the detailed spectral nature of ground motion is often overlooked, or as mentioned, a standard spectrum shape (i.e. shape which is independent of earthquake magnitude, epicentral distance, site conditions, etc.) may be scaled to the parameters considered. The disadvantage of this approach is that the probability that such a spectrum amplitude will be exceeded during future earthquakes may depend on the wave frequency (Trifunac, 1977). Clearly, if a uniform risk response spectrum is derived using the procedure outlined above, this problem will be eliminated.

McGuire (1974) was first to treat each frequency separately to obtain uniform risk response spectra in a manner similar to that described above. Der-Kiureghian and Ang (1977) determine the level of risk independently for peak acceleration, peak velocity, and peak displacement, and discuss how a response spectrum derived from the three peak values changes shape with changes in the seismicity distribution and level of risk. The present paper generalizes these methods to any functional of shaking, incorporates a more realistic model of the seismicity, and proposes a second, independent method to obtain uniform risk functionals. Subsequent papers will apply this method to evaluate several uniform risk functionals.

Although the function $p[S(\omega)]$ must take the form shown in Fig. 1.1, one must also ask what the conditions are for $p[S(\omega)]$ to be meaningful. As mentioned, $p[S(\omega)]$ is defined as the probability that

$S(\omega)$ will be exceeded in a time interval of Y years. This implies that if a sufficiently large number of intervals of this length could be observed, then $S(\omega)$ would be exceeded in about the fraction $p[S(\omega)]$ of these intervals. Clearly, using historical data, any estimate for $p[S(\omega)]$ cannot, so far, be adequately verified.

Two factors could significantly diminish the usefulness of any estimate for $p[S(\omega)]$, if they do not render it meaningless. The first is a long term change in the seismicity of a region. Although the geological record proves such changes occur, the reasonable and practical working assumption might be that such changes are so slow that they can be neglected.

The second factor, which cannot be neglected, is that earthquakes occur as part of a cycle of stress buildup and release. Even though these cycles are complex and certainly do not seem to have a regular period, their presence needs to be considered because, for nearly all engineering structures, the lifetime interval of Y years might be of the same order as the time required for stresses to accumulate and be released in a major earthquake sequence. Therefore, in many cases, a modified definition of $p[S(\omega)]$ may be the most useful, for example, that $p[S(\omega)]$ is the probability that $S(\omega)$ will be exceeded considering both the best estimates of the rates of tectonic processes which occur in a region and the recent historical seismicity. Theoretically, one could verify such an estimate of $p[S(\omega)]$ by considering a large number of previous time intervals of length Y for which the preceeding major earthquake activity was similar to that observed in the recent past. Any attempt to make such an estimate

is, at present, clearly dependent on scientific judgment. This judgment may strongly affect the resulting estimate of the distribution $p[S(\omega)]$.

Much effort in risk studies has been devoted to finding ways to describe the probability function such as in Fig. 1.1, or closely related functions. Cornell (1968) showed how the seismic risk at a site can be described by a function relating the intensity of shaking with the return period for that intensity. The method used in this paper to derive the probability functions $p[S(\omega)]$ is analogous to that used by Cornell, but it incorporates several improvements suggested by more recent research and also introduces some modifications of our own.

In the following sections, we shall show how this probability function can be derived having first formulated the seismicity model and its physical characteristics.

MODEL OF THE SEISMICITY

In this paper, seismicity refers to the description of where earthquakes occur, the frequency of occurrence of earthquakes of various sizes (e.g. magnitudes), and the determination of the largest events expected to occur in any particular region or on any particular fault.

In this section, we suggest that for the purposes of risk analysis the spatial distribution of earthquakes in any region can be described as a superposition of five geometries of source zones. Our object is to realistically model the seismicity with as little complexity as

necessary. There are, undoubtedly, numerous variations of the zones suggested here and completely different approaches are possible. If subsequent applications find that some different descriptions of source zones are necessary, they could be handled in the risk analysis of the following section much as the zones we use are handled.

The geometries we use are a point source, a line source (where the line is, in general, not straight), an areal source zone with arbitrary boundary, an arbitrary surface, and a volume of arbitrary shape. Without loss of generality, the boundaries of these zones are chosen so that the probability of the epicenter of an earthquake occurring at each point within one source zone is uniform. Cornell (1968) has described the seismicity with the first three of the source geometries mentioned above, although his analysis was developed only for a straight line source.

A point source may describe concentrated seismicity, such a geothermal area of a volcanic source, when this source is far from the site. A line source can describe events concentrated near a shallow, vertical fault. A diffused zone can describe areas where the faults are too numerous and too small to describe individually, areas where the faults cannot be mapped because of alluvial or other cover, or areas where even major earthquakes have not been correlated with faults. These three sources all assume the relative depths of the earthquakes are not a factor in determining the risk.

In cases where the relative depths of the earthquakes are important, one can use events on a fault and in a volume source. These sources might, so far, be less useful in engineering risk

analysis, because to date effects of hypocentral depth on strong shaking have not been adequately quantified. Thus, although we have developed the general formalism for this type of a source, its importance will not be known until strong motion data for deep earthquakes become available.

Figure 1.4 presents a map view of the three source types which do not consider the depth of the events. It shows an irregular shaped diffused region of seismicity, two line sources, and two point sources. One of the line sources extends beyond the boundary of the diffused region. In this case, the line source is not straight, in contrast to most previous risk studies which build on the model of Cornell (1968).

For any single source zone, let $N(\underset{\sim}{e}_j)$ be the number of earthquakes with "size" ranging from $\underset{\sim}{e}_j$ to $\underset{\sim}{e}_j + \Delta\underset{\sim}{e}_j$ that are expected to occur in the time period for which the risk model is to be developed. This description of the seismicity avoids building into the risk model any explicit assumptions about the nature of earthquake recurrence relations, and allows judgment to be more easily included in the analysis.

For generality, the size of the earthquake is described by the vector $\underset{\sim}{e}$. Often, $\underset{\sim}{e}$ will be a scalar quantity such as magnitude or maximum intensity. There may be cases where two or more parameters will have to be included in $\underset{\sim}{e}$, such as magnitude and rupture length or moment and stress drop, for example. The analysis assumes that a reasonable estimate of the number of events with size $\underset{\sim}{e}_j$ can be made for each source zone; often this estimate will be a description of a probability distribution.

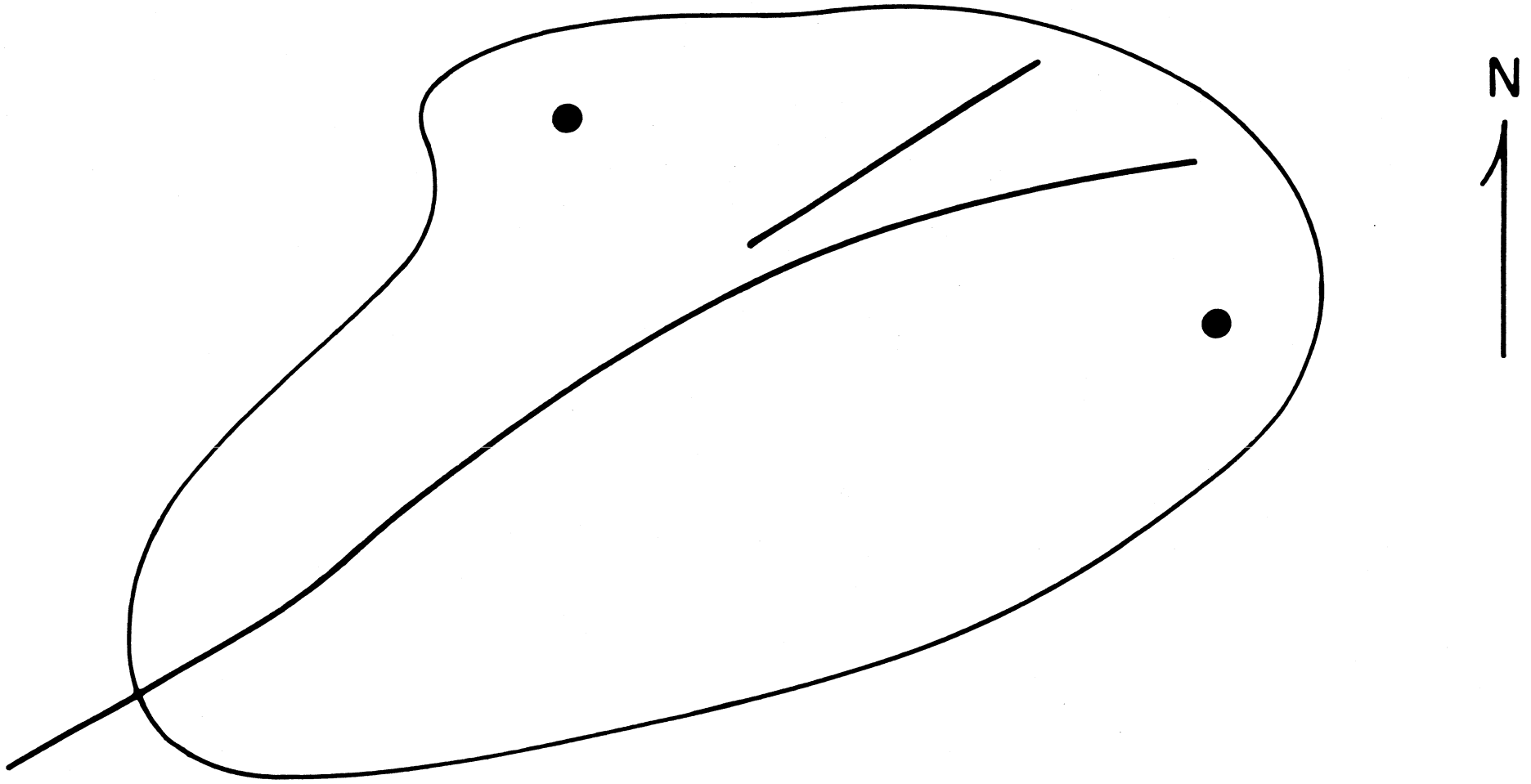

Figure 1.4. Map with the three kinds of source zones which do not explicitly consider the depth of earthquakes. Two point sources and two line sources are superimposed on an irregular shaped areal source of diffuse seismicity.

There are several ways in which $N(\underset{\sim}{e}_j)$ may be estimated. In one extreme, there could, in the future, be cases where one or more earthquakes are predicted for a region. In the other extreme, teleseismic seismicity data may be the only information available. In intermediate cases, seismicity data may be supplemented by geological and geophysical information, insights obtained from plate tectonics, and intuition.

Because previous seismicity of a region will often heavily influence estimates of $N(\underset{\sim}{e}_j)$, it is important to consider briefly how this could cause a biased result. We previously mentioned long term changes in the seismicity and cycles of stress buildup and release in major earthquake sequences. Even in cases where neither of these factors are operating, estimates of $N(\underset{\sim}{e}_j)$ could be influenced by the statistical nature of earthquake occurrence and the short duration of the historical record. If we assume earthquake occurrences are Poissonian (Gardner and Knopoff, 1974), if we treat the historic seismicity as one time interval for a region, and if there were n events (of size $\underset{\sim}{e}_j$) in that time interval, then (n+1) is the best estimate of the mean rate (of size $\underset{\sim}{e}_j$ events) for the region (Appendix 1.1). On the other hand, if the historic record is long enough to divide into several time intervals (as it may be for smaller events in some regions), then it becomes possible to consider explicitly the distribution of the number of events (of size $\underset{\sim}{e}_j$) occurring in one time interval. Such a test could verify the Poissonian assumption or provide grounds for replacing it with a better distribution.

Allen (1976) points out that in places where the historical record of earthquakes is longest, the dangers of extrapolating from a short seismicity history are most clearly illustrated. This is further supported by his observations that major earthquakes in several parts of the world have occurred on faults where previous Quaternary, and particularly Holocene, activity could have been recognized, and his identification of several active faults which have not ruptured historically in regions with long records of earthquakes.

MODEL FOR THE RISK

Consider a functional $S(\omega)$. We shall determine two functions of $S(\omega)$ and Y.

1. The expected number of times, $N_E[S(\omega)]$, that $S(\omega)$ will be exceeded at the site in Y years, and
2. $p[S(\omega)]$, the probability that $S(\omega)$ will be exceeded at least once in Y years.

From these results, the uniform risk functionals are derived as described in the introduction.

To derive $N_E[S(\omega)]$ and $p[S(\omega)]$, one needs a description of the seismicity, and a function which relates the source parameters of the earthquake to the shaking at the site. In the previous section, the seismicity was modeled as a superposition of source zones. To evaluate the risk, each of these zones is divided into small source elements, and to each of these source elements, we assign $n_i(\underset{\sim}{e}_j)$, the expected number of earthquakes with source parameters $\underset{\sim}{e}_j$ which occur in the i-th element in Y years.

In Appendix 1.2, $n_i(\underset{\sim}{e}_j)$ is derived from $N(\underset{\sim}{e}_j)$ for different geometries of source zones. The detailed manner in which the source parameters are related to the shaking at the site (as discussed in the next paragraphs) can affect the way $n_i(\underset{\sim}{e}_j)$ is evaluated.

Let the function relating source parameters to shaking at a site be $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$. Physically, q is the probability that $S(\omega)$ will be exceeded in an event with source parameters $\underset{\sim}{e}$, where the path from the source to the receiver is described by $\underset{\sim}{c}$ and $\underset{\sim}{b}$ stands for factors specific to the site. Parameters grouped under $\underset{\sim}{c}$ might include the distance from source to site and attenuation along the path. The type of soil at the site or the depth of the alluvium might be included under the parameter $\underset{\sim}{b}$.

Most risk studies to date have assumed that the amplitude of shaking is completely determined by the magnitude and distance of the earthquake so that q jumps discontinuously from the value of 1.0 which it takes at small amplitudes to the value of zero which it takes at large amplitudes. This approach was followed by Cornell (1968), Milne & Davenport (1969), Dalal (1973), Douglas and Ryall (1973), and for the most part, Kiureghian and Ang (1975). Milne & Davenport (1969) recognized that the uncertainty in the attenuation equation ought to be considered, and Kiureghian & Ang (1975) did consider its effect. McGuire (1974) explicitly considered the scatter in the way response spectrum attenuates at sixteen frequencies.

Until recently, except for McGuire (1974), there have been no systematic studies which can be applied to find the function q. However, Trifunac (1976) has recently proposed an empirical model for

the Fourier amplitude spectrum of strong ground acceleration which can be inverted to find the function we need. In Chapter II, we show how this is done; we now assume q is known and examine how it is applied.

Consider one small source element, the i-th region, say, and assume that $n_i(\underset{\sim}{e}_j)$ is the expected number of earthquakes of size $\underset{\sim}{e}_j$ in the i-th element. For this i-th element, describe the path to the site by c_i, and let $q_{ij} = q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}_j, \underset{\sim}{c}_i]$.

If the i-th region is small, then $n_i(\underset{\sim}{e}_j)$ will also be small, and is not likely to be an integer. However, clearly, in the time interval of concern either an event of size $\underset{\sim}{e}_j$ occurs in the i-th element or it does not. Thus, following Cornell (1968) and others, we first assume that $n_i(\underset{\sim}{e}_j)$ is the mean of a Poissonian distribution.

Then the probability of exactly k events is $\dfrac{e^{-n_i(\underset{\sim}{e}_j)}[n_i(\underset{\sim}{e}_j)]^k}{k!}$.

If there are k events in the i-th element, the expected number of times that $S(\omega)$ will be exceeded is kq_{ij}. Thus the expected number of times that $S(\omega)$ will be exceeded from an event of size e_j in the i-th element is

$$E_{ij} = \sum_{k=0}^{\infty} kq_{ij} \frac{e^{-n_i(\underset{\sim}{e}_j)}[n_i(\underset{\sim}{e}_j)]^k}{k!}$$

$$= q_{ij} n_i(\underset{\sim}{e}_j)$$

Then, the expected number of times $S(\omega)$ will be exceeded for all source elements and all source sizes is the sum of E_{ij} over all

such regions ($i = 1, 2, \ldots, I$) and source sizes ($j = 1, 2, \ldots, J$). Thus,

$$N_E[S(\omega)] = \sum_{i=1}^{I} \sum_{j=1}^{J} q_{ij} n_i(\underset{\sim}{e}_j) \quad . \tag{1.1}$$

To find the probability $p[S(\omega)]$ that $S(\omega)$ will be exceeded in the Y years time interval, we first find the probability p_{ij} that at least one event with size $\underset{\sim}{e}_j$ in element i will cause $S(\omega)$ to be exceeded. Again using the Poisson assumption,

$1 - q_{ij}$ is the probability of not exceeding for one event

$(1 - q_{ij})^k$ is the probability of not exceeding for k events

$1 - (1 - q_{ij})^k$ is the probability of exceeding at least once for k events, and thus

$$p_{ij} = \sum_{k=0}^{\infty} [1 - (1 - q_{ij})^k] \frac{e^{-n_i(\underset{\sim}{e}_j)} [n_i(\underset{\sim}{e}_j)]^k}{k!}$$

$$= 1 - e^{-n_i(\underset{\sim}{e}_j) q_{ij}}$$

To derive $p[S(\omega)]$, proceed as follows:

$(1 - p_{ij})$ = probability that no source of size $\underset{\sim}{e}_j$ at point i will cause $S(\omega)$ to be exceeded

$\prod_{j=1}^{J} (1 - p_{ij})$ = probability that no source of any size at point i will cause $S(\omega)$ to be exceeded.

$$1 - p[S(\omega)] = \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - p_{ij}) = \text{probability that no source will exceed } S(\omega)$$

Thus

$$p[S(\omega)] = 1 - \prod_{i=1}^{I} \prod_{j=1}^{J} (1 - p_{ij})$$

Using (1.1), this simplifies to

$$p[S(\omega)] = 1 - \exp\{-N_E[S(\omega)]\} \qquad (1.2)$$

This shows the close relationship between the expected number of exceedances and the probability of exceedance when a Poisson assumption is used to describe the seismicity rate. The complete function $p[S(\omega)]$, as illustrated in Fig. 1.1, can be easily calculated using eq. (1.1) and (1.2). As described earlier, when $p[S(\omega)]$ is known for several frequencies, a uniform risk functional can be defined.

The recurrence time of a given amplitude $S(\omega)$ is just $T[S(\omega)] = N_E[S(\omega)]^{-1}$ where, from the way the problem is set up, the time unit is Y years. Thus a functional with a uniform expected recurrence time can be found from the functions $E[S(\omega)]$. For small $E[S(\omega)]$, note that $p[S(\omega)] \simeq N_E[S(\omega)]$.

Note also, as is pointed out by Gumbel (1958, p. 23) and by Algermissen & Perkins (1976), that the probability that $S(\omega)$ is exceeded in one return period $\{N_E[S(\omega)] = 1\}$ is $p[S(\omega)] = 1 - 1/e \cong 0.63$. To achieve a small probability that $S(\omega)$ will be

exceeded in a structure which is designed to last for a time T, Gumbel (1958) points out that $S(\omega)$ should be chosen to correspond to a return period of approximately T/ϵ.

In some cases, for example an earthquake prediction, one may be told that an earthquake and its aftershock sequence will occur some place within a diffused region or along a specific fault. In this event, one may want to know the risk at various sites in order to guide decisions such as if preventive actions should be recommended. For this, one would want to know the quantities $p^*[S(\omega)]$ and $N_E^*[S(\omega)]$ derived on the assumption that one knows how many earthquakes will occur within the source zone, rather than on the assumption that the input seismicity is an estimate for the mean of a Poisson process. We refer to this as the "literal" model because the seismicity is interpreted literally.

Suppose we have a source region, and we know that $N(\underset{\sim}{e}_j)$ events of size $\underset{\sim}{e}_j$ will occur some place in that region but we do not know where. As discussed in Appendix 1.2, we divide the source zone into I smaller regions. We shall let γ_i be the probability that an event occurs in the i-th region. From Appendix 1.2, it is clear that $n_i(\underset{\sim}{e}_j) = \gamma_i N(\underset{\sim}{e}_j)$ and that $\sum_{i=1}^{I} \gamma_i = 1$.

Suppose $M_i(\underset{\sim}{e}_j)$ events occur in the i-th element, where $\sum_{i=1}^{I} M_i(\underset{\sim}{e}_j) = N(\underset{\sim}{e}_j)$. There are many different ways in which the $N(\underset{\sim}{e}_j)$ events can be distributed among the I small regions. The probability of any one particular choice is, by the multinomial theorem,

$$\frac{N!}{M_1!\,M_2!\ldots M_I!}\ \gamma_1^{M_1}\gamma_2^{M_2}\ldots\gamma_I^{M_I}\ ,$$

where we have abbreviated $M_i(\underset{\sim}{e}_j) = M_i$ and $N(\underset{\sim}{e}_j) = N$. Note

$$\underbrace{\sum_{M_1}\sum_{M_2}\cdots\sum_{M_I}}_{\forall\text{ combinations such that }\Sigma M_i = N}\ \frac{N!}{M_1!\,M_2!\ldots M_I!}\ \gamma_1^{M_1}\ \gamma_2^{M_2}\ \ldots\ \gamma_I^{M_I}$$

$$= (\gamma_1 + \gamma_2 + \ldots + \gamma_I)^N = 1^N = 1\ .$$

To determine $N_E^*[S(\omega)]$, we proceed as before to find

$$E_{ij}^* = \sum_{M_i=0}^{N} q_{ij}M_i \times \text{prob}\left\{\begin{array}{l} M_i \text{ events of size } \underset{\sim}{e}_j \\ \text{occur in element } i \end{array}\right\}\ .$$

Since the locations of the events which occur outside element i do not matter for this analysis,

$$E_{ij}^* = \sum_{M_i=0}^{N} q_{ij}M_i\ \frac{N!}{M_i!\ (N-M_i)!}\ \gamma_i^{M_i}\ (1-\gamma_i)^{N-M_i}$$

$$= q_{ij}N\gamma_i\ .$$

Because $N\gamma_i$ is the quantity identified as $n_i(\underset{\sim}{e}_j)$ earlier, it is clear that,

$$N_E^*[S(\omega)] = N_E[S(\omega)]\ . \qquad (1.3)$$

To find $p^*[S(\omega)]$, consider:

$(1 - q_{ij})$ = probability that $S(\omega)$ is not exceeded from 1 event in element i.

$(1 - q_{ij})^{M_i}$ = probability that $S(\omega)$ is not exceeded from M_i events in element i.

$\prod_{i=1}^{I} (1 - q_{ij})^{M_i}$ = probability that $S(\omega)$ is not exceeded for a particular combination of the M_i.

Since the different combinations are mutually exclusive and exactly one of them will occur, the probability that $S(\omega)$ is not exceeded by any one combination of $S(\omega)$ is

$$\underbrace{\sum_{M_1} \sum_{M_2} \cdots \sum_{M_I}}_{\substack{\forall \text{ combinations} \\ \text{such that } \Sigma M_i = N}} \left\{ \begin{array}{l} \text{probability that } S(\omega) \text{ not} \\ \text{exceeded for a parti-} \\ \text{cular combination} \end{array} \right\} \left\{ \begin{array}{l} \text{probability of that} \\ \text{particular combi-} \\ \text{nation} \end{array} \right\}$$

$$= \sum \cdots \sum \left\{ \prod_{i=1}^{I} (1 - q_{ij})^{M_i} \right\} \left\{ \frac{N!}{M_i!\, M_2! \cdots M_I!} \gamma_1^{M_1} \gamma_2^{M_2} \cdots \gamma_I^{M_I} \right\}$$

$$= \sum \cdots \sum \frac{N!}{M_1!\, M_2! \cdots M_I!} [(1 - q_{ij})\gamma_i]^{M_1} \cdots [(1 - q_{ij})\gamma_I]^{M_I}$$

$$= \left\{ \sum_{i=1}^{I} [(1 - q_{ij})\gamma_i] \right\}^N = \left\{ \sum_{i=1}^{I} \gamma_i - \sum_{i=1}^{I} \gamma_i q_{ij} \right\}^N$$

$$= \left(1 - \sum_{i=1}^{I} \gamma_i q_{ij} \right)^N$$

Since the different magnitudes are independent:

$$p\{S(\omega) \text{ is not exceeded by any magnitude}\}$$

$$= \prod_{j=1}^{J} p\{S(\omega) \text{ is not exceeded by magnitude } j\}$$

$$= 1 - p^*[S(\omega)]$$

$$1 - p^*[S(\omega)] = \prod_{j=1}^{J} \left(1 - \sum_{i=1}^{I} \gamma_i q_{ij}\right)^{N(\underset{\sim}{e}_j)},$$

and thus,

$$p^*[S(\omega)] = 1 - \exp\left\{\sum_{j=1}^{J} N(\underset{\sim}{e}_j) \ln \left(1 - \sum_{i=1}^{I} \gamma_i q_{ij}\right)\right\} \qquad (1.4)$$

There is another way to derive the probabilities under the assumption that the number of earthquakes $N(\underset{\sim}{e}_j)$ is to be taken literally. This is done by considering a very long time interval of duration MY in which $MN(\underset{\sim}{e}_j)$ events occur in the source regions. Then we assume that in the i-th region $Mn_i(\underset{\sim}{e}_j)$ events will occur, where $Mn_i(\underset{\sim}{e}_j)$ is now an integer. For sufficiently large number of events, this assumption does not introduce any significant error. Following the same logic used earlier to find $N_E[S(\omega)]$ and $N_E^*[S(\omega)]$ and $p[S(\omega)]$ and $p^*[S(\omega)]$, we obtain:

$$N_E^{\dagger M} = M \sum_{i=1}^{I} \sum_{j=1}^{J} n_i(\underset{\sim}{e}_j) q_{ij}$$

$$\left(1 - p^{\dagger M}\right) = \exp\left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} M n_i(\underset{\sim}{e}_j) \ln (1 - q_{ij}) \right\}$$

If the sources in all of the M time intervals are independent, then:

$$N_E^{\dagger}[S(\omega)] = \frac{1}{M} N_E^{\dagger M} = N_E[S(\omega)] \qquad (1.5)$$

and

$$1 - p^{\dagger}[S(\omega)] = \{1 - p^{\dagger M}[S(\omega)]\}^{1/M} ,$$

which implies

$$p^{\dagger}[S(\omega)] = 1 - \exp\left\{ \sum_{j=1}^{J} \sum_{i=1}^{I} N(\underset{\sim}{e}_j) \gamma_i \ln (1 - q_{ij}) \right\} . \qquad (1.6)$$

It can be shown that

$$p^{\dagger}[S(\omega)] \geq p^{*}[S(\omega)] \geq p[S(\omega)] . \qquad (1.7)$$

Equality occurs between $p^{\dagger}$ and p^{*} only when there is just one source element (or when q_{ij} is the same for all source elements). Equality between p^{*} and p occurs only in the case where $q_{ij} = 0$ for all source elements. For large $N(\underset{\sim}{e}_j)$, the differences between $p^{\dagger}$ and p approach zero. Also, as $S(\omega)$ increases, and the probabilities of exceeding $S(\omega)$ thus decrease, the differences between $p^{\dagger}[S(\omega)]$ and $p[S(\omega)]$ decreases toward zero. The differences may still be quite significant, however, for a probability of exceedance of around 0.1.

These probabilities have been expressed as sums over elements of source regions. Clearly, for the limit of the size of these elements approaching zero, each of these formulae could be expressed as an integral. Because they are generally evaluated numerically, however, this transformation is not necessary, and we will omit it.

EXTREME VALUE STATISTICS

The previously described method to estimate risk at a site is referred to by Lomnitz (1974, p. 119-124) as a "composition method" because shaking at the site is estimated using a description of the seismicity and the attenuation. As pointed out by Lomnitz (1974), for the unusual cases where a long and complete record of past shaking at the site exists, it is possible to apply extreme value statistics.

The first step in this application would be to examine the functionals $S(\omega)$ of all the events recorded at the site, and for each frequency ω, find the largest value of $S(\omega)$ that was recorded in each year of the records. Then, using the methods of extreme value statistics described by Gumbel (1958), one could derive a function giving the return period for each value of $S(\omega)$. If a set of these functions were derived, then the method described in the introduction could be applied to obtain a uniform risk functional.

Because the necessary data is not likely to exist, this method as described above has only very limited usefulness. However, Milne and Davenport (1969) have shown that one may apply extreme value

statistics where there are no instrumental records of shaking by (1) using standard attenuation formulae to estimate what the shaking at a site may have been in all historical events, and then (2) applying extreme value statistics to these derived amplitudes of shaking. Probability could be introduced by calculating for each event the amplitude which, with probability p, was not exceeded.

At present, there is no theoretical reason to prefer either the extreme value method or the method we develop in this paper; ideally both would give the same result. On a regional scale, Yegulalp and Kuo (1974) have found the maximum annual magnitudes of earthquakes are consistent with extreme value statistics. For practical purposes, however, the "composition" method more easily allows judgment by the seismologist or engineer to be included, for example, in recognizing the existence of a seismic gap, and is not necessarily sensitive to individual large events in short historic records.

DISCUSSION

Uniform risk functionals of strong earthquake ground motion have the same probability of being exceeded at each frequency, when all of the seismicity of a region is considered.

The steps in the derivation of a uniform risk functional are as follows:

(1) Describe the geometry of earthquake zones, by point, line, and areal sources (after Cornell, 1968), and for completeness a dipping surface and a diffused volume source. For each of

these source zones, the expected number of events of each size $[N(\underset{\sim}{e})]$ is defined by studying previous seismicity, by insights obtained from geological studies and plate tectonics, and by scientific judgment.

(2) Find a frequency-dependent description of the way strong motion attenuates in the region, including a description of how amplitudes of $S(\omega)$ scatter about mean trends. From this, define the function $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$ which gives the probability that $S(\omega)$ will be exceeded for site conditions $\underset{\sim}{b}$, source $\underset{\sim}{e}$, and path $\underset{\sim}{c}$.

(3) Divide each source zone into small source elements, and, assuming the epicenter of each event of size $\underset{\sim}{e}$ is equally likely to occur any place in the source zone, distribute the seismicity in a manner consistent with the definition of the function $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$.

(4) Using the formulae (1.1)-(1.6), [usually (1.1) and (1.2)] find the functions $p[S(\omega)]$ for several frequency bands.

(5) Using the method described in the introduction, derive the uniform risk functional from the functions $p[S(\omega)]$.

Steps (1) through (4) can in some cases be replaced by the methods of extreme value statistics to derive a function equivalent to $p[S(\omega)]$ in each frequency band.

The method described synthesizes the method presented by Cornell (1968) with improvements suggested by Der-Kiureghian and Ang (1975), Dalal (1973), and McGuire (1974). The result is that it

uses a realistic description of the seismicity to derive uniform risk functionals. In deriving these functionals, probability is taken into account at every step in a uniform way, so that at the end one has a result with a known degree of confidence. All the assumptions can be clearly identified, so that the effect of each assumption can easily be tested.

The method is applied independently in each of several frequency bands. Thus the factors which are important in each frequency band control the risk there. This is in contrast, for example, to common methods which determine one factor, such as a peak of motion, and extrapolate from there to an entire spectral shape.

APPENDIX 1.1

Suppose that we have a Poisson process operating, and the mean rate is not known. Furthermore, assume that we have only one observation of this process, and in this observation, n events occurred. We want to know what is the best estimate of the mean of the Poisson process given this information. The mean need not be integer. Assume

$$p\{\text{mean is between } \mu \text{ and } \mu + d\mu \mid n \text{ events occurred}\}$$

$$= p\{n \text{ events occur} \mid \text{mean is between } \mu \text{ and } \mu + d\mu\}$$

$$= \frac{e^{-\mu}\mu^n}{n!}$$

The expected value of the mean becomes

$$E(\text{mean}) = \int_0^\infty \mu p\{\mu \mid n\}\, d\mu$$

$$= \int_0^\infty \mu \frac{e^{-\mu}\mu^n}{n!}\, d\mu$$

$$= \frac{1}{n!}\int_0^\infty e^{-\mu}\mu^{n+1}\, d\mu = \frac{(n+1)!}{n!} = n + 1$$

APPENDIX 1.2

Derivation of $n_i(\underset{\sim}{e}_j)$

The way $n_i(\underset{\sim}{e}_j)$, the number of events in a small element, i, of a source zone is derived from the total number of events $N(\underset{\sim}{e}_j)$ in that source zone will depend upon the selected parameters in the analysis. For example, if the path parameters $\underset{\sim}{c}_i$ represent only the epicentral distance r, then q_{ij} is the probability that an earthquake located in element i at epicentral distance r with source parameters $\underset{\sim}{e}_j$ causes shaking at the site to exceed $S(\omega)$, and $n_i(\underset{\sim}{e}_j)$ must be the expected number of earthquakes with epicenters in the i-th element of the source region. However, if $\underset{\sim}{c}_i$ represents the distance to the closest point on the fault (r_c, say), then q_{ij} must be the probability that an earthquake with closest point to the fault at distance r_c and with source parameters $\underset{\sim}{e}_j$ will cause $S(\omega)$ to be exceeded. In this case, $n_i(\underset{\sim}{e}_j)$ must represent the expected number of earthquakes which will have their closest point to the site in the i-th element.

In reality, the amplitudes of ground shaking are determined by contributions from the entire rupture surface, as in the representation theorem of Burridge and Knopoff (1964). Thus, for an earthquake with known fault breakage, the best method to determine q would be to use a scheme which considers the probable amplitudes of radiation from each part of the fault. At present, rather than applying such a complicated deterministic analysis, we may consider two simplifications which should bracket the results of such an analysis.

We shall derive $n_i(\underset{\sim}{e}_j)$ for the two forms of q described above. We will set up the problem for the closest point using the extreme assumptions that the rupture is unilateral and the direction of rupture is random. For such rupture in a diffuse zone, we allow the end of the fault to be outside of the source zone.

This is an extreme assumption which deserves some discussion. Considering several attenuation curves which have been used so far (Trifunac & Brady, 1975), there appears to be no systematic difference between those using epicentral distance and those using the closest distance to the fault. This, of course, does not mean that these different descriptions of the distance between the source and the station should not be considered in a future improved and more detailed model, but merely indicates that numerous other uncertainties and scatter of the existing data, as well as the availability of data only for epicentral distances larger than about 10 km, do not, at this time, suggest large and significant differences.

The consequence of the overall similarity of existing attenuation functions is that derived functions $q[S(\omega),\ldots]$ would also have no significant differences between those using epicentral distance and those using hypocentral distance. Therefore, in Chapter II, for lack of any real alternative, we will use the same functions (formulated in terms of epicentral distance) for both cases. As shown by Der-Kiureghian and Ang (1975), this results in larger amplitudes of spectra when the distance in our analysis is measured to the closest point. By using the extreme case described above, we suggest that the result will give an upper bound for the actual spectrum. On the

other hand, analysis using the epicentral distance, which is logically consistent for our choice of attenuation function, may lead to a less accurate estimate of the uniform risk functional for all we know, but we expect that it will, if anything, tend to underestimate a spectral amplitude. Ideally, the use of the proper functions should compensate for the different ways of defining the distance, and thus both methods should ultimately give similar estimates of the function $p[S(\omega)]$.

We now derive $n_i(\underset{\sim}{e}_j)$ defined above for two definitions of source to station distance.

Method (1). The function q is assumed to depend only on the distance between the epicenter and the site. Cornell (1968), Dalal (1973) and McGuire (1974) used this form.

a. Point source. If there are $N(\underset{\sim}{e}_j)$ events of size $\underset{\sim}{e}_j$, and the point source is designated as the i-th element, then $n_i(\underset{\sim}{e}_j) = N(\underset{\sim}{e}_j)$.

b. Line source. Designate a short element of length L_i as the i-th element. Then $n_i(\underset{\sim}{e}_j) = \dfrac{N(\underset{\sim}{e}_j)L_i}{L}$, where the line source has length L.

c. Areal source. The areal source (source depth is assumed to be constant) covers a region with area A. Designate a small element of area A_i as the i-th element. Then $n_i(\underset{\sim}{e}_j) = \dfrac{N(\underset{\sim}{e}_j)A_i}{A}$.

d. Dipping surface. Designate a small element of the fault area Σ_i as the i-th element. Then $n_i(\underset{\sim}{e}_j) = \dfrac{N(\underset{\sim}{e}_j)\Sigma_i}{\Sigma}$ where the surface has a total area Σ.

e. Volume source. Divide the volume into small elements, and let the i-th element have volume V_i. Then $n_i(\underset{\sim}{e}_j) = \dfrac{N(e_j)V_i}{V}$, where the source has a total volume V.

Method (2). The function q depends only on the distance of the closest point of rupture to the site. Der-Kiureghian and Ang (1975) derived a method similar to this treating the sources as straight line segments; Douglas and Ryall (1975) have applied a similar method.

In this case, the vector $\underset{\sim}{e}_j$ has at least two components, which we shall call S_j, which represents magnitude or moment, and ℓ_k, which represents rupture length in most cases. We have given the rupture length a different subscript from the source "size" to emphasize that for any particular size S_j, several possible values of rupture length are possible. For a rupturing fault surface, we replace ℓ_k by a_k, the area of the surface which ruptures. Now $N(\underset{\sim}{e}_j)$ becomes $N(S_j, \ell_k)$, and a number of assumptions would be required about the nature of an earthquake source to describe $N(S_j, \ell_k)$ explicitly. However, this can be handled approximately by assuming that ℓ_k is determined by S_j, or that for each S_j, there is a distribution of possible values of ℓ_k. For the following, we assume that $N(S_j, \ell_k)$ describes the number of sources with epicenter in the source region, and we wish to find, for each nearby element of area (designated by i), $n_i(S_j, \ell_k)$, the expected number of events with source (S_j, ℓ_k) which have their closest point of approach to the site in that element.

a. Short straight surface source. This is the equivalent of the point source of the case which depends on the epicentral distance. There are N events with size S_j and dimension ℓ_k, as described above, which occur with epicenters at the point source. A source

zone where all the epicenters occur at a point but where the ruptures are radial away from the point and a significant fraction of the distance to the site is not physically realistic. However, we develop this case in detail because a superposition of such sources is directly applicable to the areal source and because this case is useful for studying characteristics of the model.

Consider one event. The geometry of this situation is illustrated in Fig. A1.2-1. The epicenter is shown at point E, and for rupture of length ℓ, the entire fault is confined to within a circle of radius ℓ centered at E. The site in question is shown as point O on the figure, and the distance from O to E is r. If the rupture is straight, then the nearest point of the rupture to point O must be within the heavily shaded region. Allowing a curved fault would slightly increase the allowed area by adding the lightly shaded region.

It is possible to build into the risk model a detailed description of the probability distribution of the closest distance, as a function of the direction of rupture, the distance of the epicenter from the end of the rupture and even the deviation of the rupture from a straight line. Such detail, however, seems to be unjustified, because the assumption that the shaking is determined entirely by the closest point on the fault is wrong in general. We will make a simplifying assumption for this case that the rupture is unilateral and in a straight line. In addition to the considerations described earlier in this section, we have chosen this assumption because, from considering the rupture of large earthquakes, we get a general impression that the rupture is often predominantly unilateral. With this assumption,

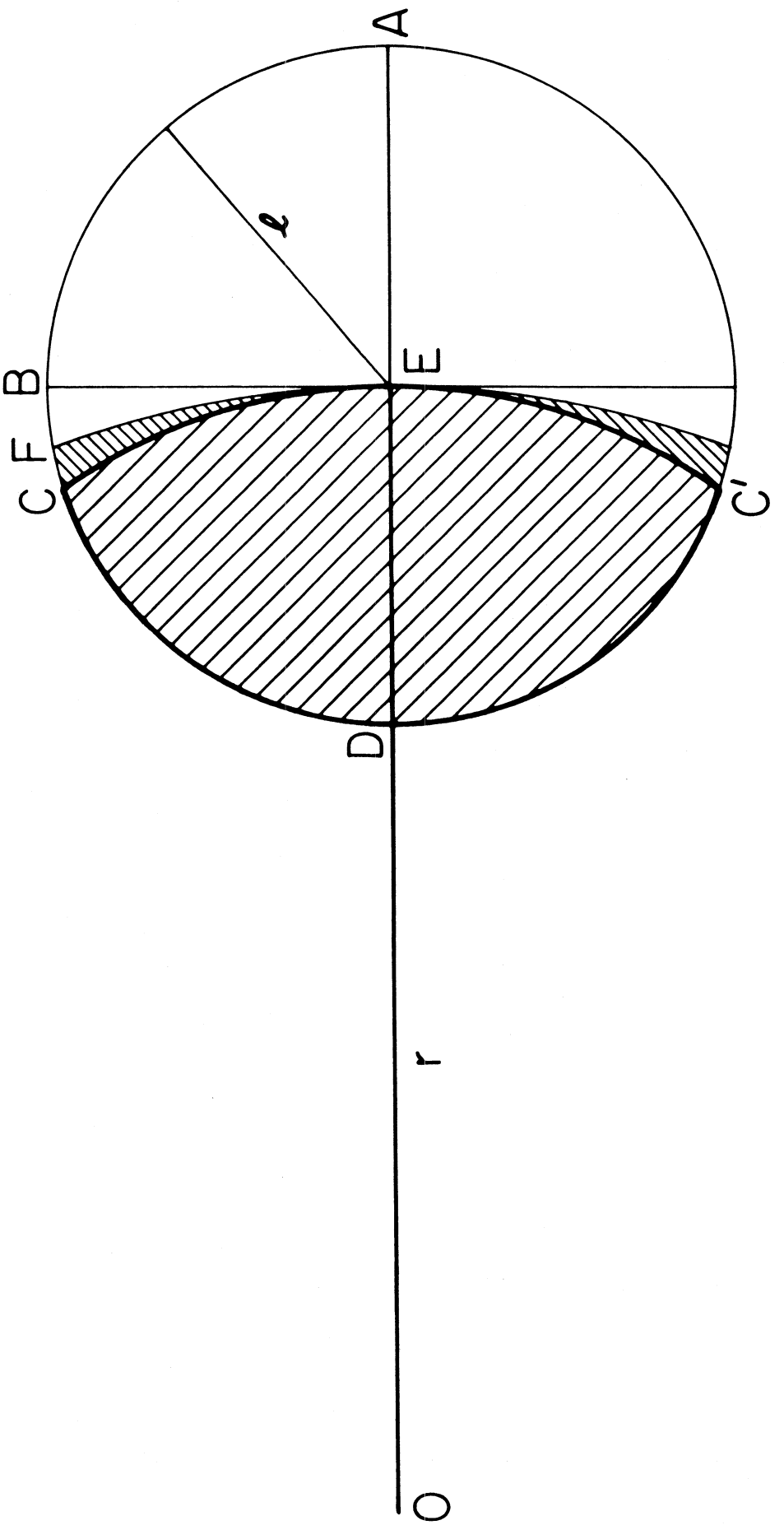

Figure A1.2-1. Possible locations (shaded) of the closest point to the site O of a fault with epicenter E and length of rupture l.

the closest point to the site always lies on the curve E C D C' E which bounds the heavily shaded region.

By assuming that the fault has equal probability of rupturing in any direction, it is a simple matter to find the probability that the closest point is along a given portion of the curve. The closest distance from the fault to the site, d, is a function of the angle ϕ, as shown in Fig. A1.2-2. The distance is

$$d(\phi) = \begin{cases} r & 0 \leq |\phi| < \pi/2 \\ r \sin |\phi| & \pi/2 \leq |\phi| < \pi/2 + \theta \\ (r^2 + \ell^2 + 2r\ell \cos \phi)^{\frac{1}{2}} & \pi/2 + \theta \leq |\phi| \leq \pi \end{cases}$$

where $\theta = \tan^{-1}\left(\frac{\ell}{\sqrt{r^2 - \ell^2}}\right)$, ℓ is the length of the fault, and r is the distance from the site to the epicenter. Next, assume q is independent of the direction of wave approach and any angle ϕ is equally probable. Then $d(\phi)$ can be used to find the probability that d is in a distance range $d_1 \leq d \leq d_2$. This is done by first finding the inverse function $\phi(d)$. If $\ell < r$, then:

$$\phi(d) = \begin{cases} 0 & d \geq r \\ \sin^{-1}\left(\frac{d}{r}\right) & r > d > (r^2 - \ell^2)^{\frac{1}{2}} \\ \cos^{-1}\left(\frac{d^2 - r^2 - \ell^2}{2r\ell}\right) & (r^2 - \ell^2) > d \geq r - \ell \\ \pi & r - \ell \geq d \end{cases}$$

If $\ell \geq r$, then

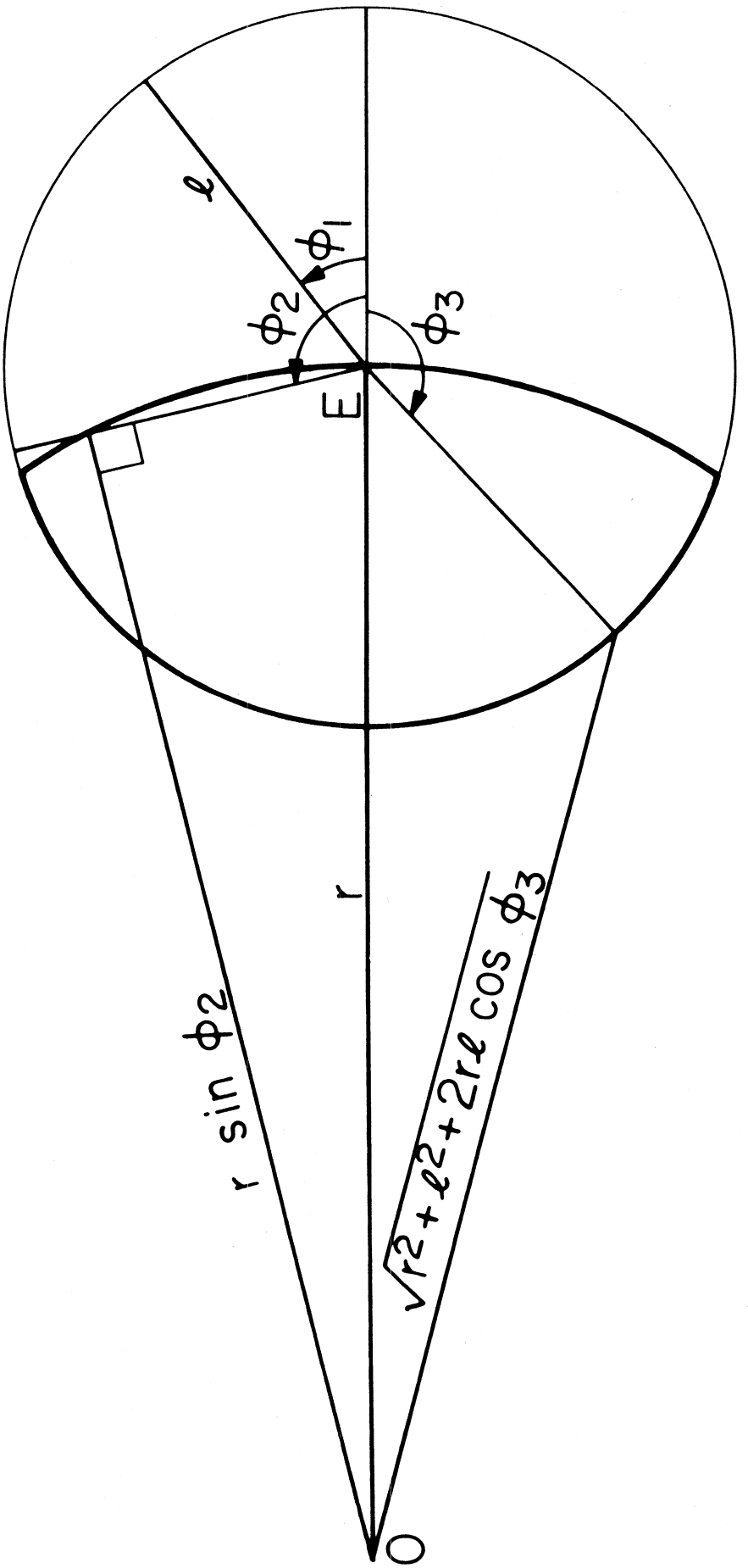

Figure A1.2-2. Locus of closest point to the site O (heavy line) of a straight fault with epicenter E and unilateral rupture of length l.

$$\phi(d) = \begin{cases} 0 & d \geq r \\ \sin^{-1}\left(\frac{d}{r}\right) & r \geq d \geq 0 \end{cases}$$

Here the inverse sine and inverse cosine functions are always taken in the second quadrant and ϕ is in radians. The probability that $d_1 \leq d \leq d_2$ is

$$\frac{\phi(d_1) - \phi(d_2)}{\pi} .$$

b. Line source. Let a line source be represented by M equally spaced points with spacing $\frac{L}{M-1}$, where L is the length of the fault. This does not require the fault to be a straight line. Consider an earthquake which ruptures the fault along a length ℓ_k. Then the center, say, of the rupture zone can occur any place along a section of the fault with length $L - \ell_k$, and we assume that each point in this length occurs with equal probability. Thus for the m-th allowed epicenter on the fault there is a number $d_m(S_j, \ell_k) = \frac{N(S_j, \ell_k)}{L - \ell_k}$ of events which are allowed to occur with center at that point. The center of the rupture zone, as used here, is strictly a geometrical description, and has nothing to do with the epicenter, which could occur any place on the length of the fault. Thus for this case, there is no assumption of "unilateral" or "bilateral" rupture.

For this m-th allowed location, let the i-th element of the fault be the point on the rupture zone closest to the point where the risk is to be calculated. We attribute to that i-th point the number

$d_m(S_j, \ell_k)$ of events with center at the m-th point. Numerically, this is easily done for an arbitrary shape of the line source. By considering in order each of the m points on the fault, the function $n_i(S_j, \ell_k)$ is built up. When ℓ_k is larger than L, the fault length, we assume that the entire fault ruptures, but do not extend the fault or eliminate events with such a long rupture. This is because any estimated fault length is approximate at best.

c. Areal source. Having derived a relationship for the short straight line source, it is a simple matter to proceed to the areal source. First, one finds the number of events with epicenters in each small element of area in the source region, and treats that small element as a short straight line source.

d. Dipping surface. We can in this case account fully for the way that shallow earthquakes rupture a portion of a fault surface. As mentioned earlier, in this case we assume that $\underset{\sim}{e}_j$ has two components, S_j, which may be magnitude or moment, and a_k, the area of the fault surface which is ruptured. To find the distribution of closest elements of the fault, we proceed in a manner similar to that for the line source.

First, divide the fault surface into J equal sized elements with area $\Sigma_i = \frac{\Sigma}{J}$. Then an event with area a_k must rupture a_k/Σ_i adjacent elements. A reasonable approximation is that the rupture area is approximately square for small events and rectangular with a width equal to the width of the fault surface for large events. In a manner analogous to that for the line source, one can successively consider each location on the fault where the event can occur (say

there are K such locations), find the closest element for each of these locations, and then add $N(S_j, a_k)/K$ events to that element.

e. Volume source. At present, there are insufficient empirical results to motivate selection among possible methods of modifying the distribution of epicenters to account for the finite dimensions of rupture for this type of source zone.

CHAPTER II

UNIFORM RISK FOURIER AMPLITUDE SPECTRA OF STRONG GROUND MOTION USING EARTHQUAKE MAGNITUDE, SOURCE TO STATION DISTANCE, AND RECORDING SITE CONDITIONS

INTRODUCTION

This Chapter is a continuation of the work presented in the first part of this report. Chapter I discusses the meaning of uniform risk functionals, and attempts to extend and generalize the methods and results of previous investigators.

This Chapter applies the methods discussed in Chapter I to calculation of uniform risk Fourier amplitude spectra of acceleration and uses these spectra to examine the characteristics of the method of Chapter I. As defined there, a uniform risk Fourier spectrum has the property that the probability that it is exceeded by the Fourier spectra of any event within a specified time period (any size, at any distance) is independent of frequency.

This uniform risk Fourier spectrum applies to a specific site and to a time interval of Y years. Suppose this site were observed for a large number of intervals with duration Y. In some circumstances, as mentioned in Chapter I, it may be desirable to choose all these time intervals in the same portion of the cycles of accumulation and release of seismic stress (to the extent that this is possible). Then the largest Fourier amplitude in a frequency band centered at frequency ω from the earthquakes that occurred could be used to define a distribution function of Fourier amplitudes; there would be one data point for each Y year interval. As data

accumulated, this empirical distribution function should converge toward a smooth curve which correctly characterizes the site and which was called $p[S(\omega)]$ in Chapter I. There we described how functions $p[S(\omega)]$ can be estimated from present knowledge of seismicity and attenuation. Once known they can easily be used to obtain uniform risk spectra.

A problem arises in attempts to numerically estimate the function $p[S(\omega)]$, for at present, the seismic history of most regions of the world is not known adequately. Chapter I summarized and generalized two methods which have been used to estimate these, or similar, functions. Chapter I developed, in the terminology of Lomnitz (1974), a "composition method" which generalizes the work of Cornell (1968). Chapter I also discussed how the extreme value approach of Milne and Davenport (1969) can be generalized to find uniform risk functionals. Here, as in Chapter I, the "composition method" is emphasized. This "method" requires the knowledge of two functions. The first is a description of the seismicity. In this paper, seismicity refers to the description of where earthquakes occur, the frequency of occurrence of earthquakes of various sizes (e.g., magnitudes), and the determination of the largest events expected to occur in any particular region or on any particular fault. The second is a function {called $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$ in Chapter I} which gives the probability that an earthquake with a given source and location will cause a given spectral level to be exceeded.

The computations in this paper are made possible because of recent work by Trifunac (1976) which characterizes the dependence

of Fourier amplitude spectra of acceleration in several frequency bands on the parameters of earthquake magnitude, epicentral distance, site condition, and confidence level. His empirical functions can be inverted to obtain the function $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$ which gives the probability that $S(\omega)$ is exceeded by a single earthquake. Thus, we do not assume that the amplitudes of shaking from a single event are uniquely determined, but rather that the amplitudes at each frequency fall on a distribution function which is determined by the earthquake magnitude, epicentral distance, and site condition.

Because we are using the Fourier amplitude spectrum to describe the risk, we cannot easily compare our results with those of previous studies which have used a single peak value such as peak acceleration to scale response spectra. The method we use is substantially different, however, because we recompute the Fourier spectral amplitude independently at each frequency. We can, and do, contrast the results of the method with those resulting from scaling a standard spectral shape to an appropriately chosen amplitude.

The organization of this Chapter is as follows: First, we discuss the details of how the results of Trifunac can be inverted to find the function $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$. This section may be skipped without loss of continuity. Then, after briefly discussing the computer code, we apply the method to some artificial examples of seismicity to show several characteristics of the model. Finally, we apply the method to find the risk for an actual example, and discuss the effects of differing descriptions of the seismicity.

DERIVATION OF THE FUNCTION $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$ FOR NUMERICAL APPLICATION USING FOURIER SPECTRUM

Recent work by Trifunac (1976) makes it possible to apply the concepts of the previous sections to Fourier spectra. His work characterizes the dependence of Fourier amplitude spectra of acceleration in several frequency bands on the parameters of earthquake magnitude, epicentral distance, site condition, and confidence level. His empirical functions can be inverted to obtain the function $q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}]$ which gives the probability that $S(\omega)$ is exceeded by a single earthquake. Thus, we do not assume that the amplitudes of shaking from a single event are uniquely determined, but rather that the amplitudes at each frequency fall on a distribution function which is determined by the earthquake magnitude, epicentral distance, and site condition.

Because we are using the Fourier amplitude spectrum to describe the risk, we cannot easily compare our results with those of previous studies which used a single peak value such as peak acceleration to scale response spectra. The method we use is substantially different, however, because we recompute the Fourier spectral amplitude independently in several frequency bands. We can, and do, contrast the results of the method with those resulting from scaling a standard spectral shape to an appropriately chosen amplitude.

The function $q_{ij} = q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}_j, \underset{\sim}{c}_i]$ gives the probability that $S(\omega)$ will be exceeded given the site, source, and path conditions. For the Fourier amplitude spectrum, it can be derived from the

results of (Trifunac, 1976):

$$\log_{10} FS(T) = M + \log_{10} A_0(R) - a(T)P_\ell - b(T)M - c(T) - d(T)s - e(T)v - f(T)M^2 - g(T)R \quad . \qquad (2.1)$$

In this relationship, M is the magnitude, R is the distance, v = 0 for horizontal components and v = 1 for vertical components, and s = 0 for the alluvium sites, s = 2 for a site on sound basement rock with s = 1 for "intermediate" geologic conditions. The terms a(T), b(T),...,g(T) are regression coefficients found independently at each period (T) considered. Following Trifunac (1976), for magnitudes $M \leq M_{min} = \frac{-b(T)}{2f(T)}$, the terms $-b(T)M - f(T)M^2$ are replaced by the constant $-b(T)M_{min} - f(T)M_{min}^2$, and for magnitudes $M \geq M_{max} = \frac{1-b(T)}{2f(T)}$, the terms $M - b(T)M - f(T)M^2$ are replaced by the constant $M_{max} - b(T)M_{max} - f(T)M_{max}^2$.

The term $\log_{10} A_0(R)$ is the attenuation function given by Richter (1958) for the local magnitude scale in Southern California. To correctly apply this method to another region, the coefficients in equation (2.1), or a comparable equation, should be derived entirely from data of that region. However, in the absence of such data, a reasonable approximation is to replace the term $\log_{10} A_0(R)$ derived for Southern California with one applicable to the region considered.

The term $a(T)P_\ell$ describes the way the data is distributed about the mean curve, where P_ℓ is approximately the probability that FS(T) will not be exceeded for a given M, R, s, and v. To illuminate its meaning, suppose that we designate each of the 546 data points used by Trifunac by an index i, and find the quantity:

$$\epsilon_i = -\frac{1}{a(T)}\left[\log_{10} FS_i(T) - M_i - \log_{10} A_0(R_i) + b(T)M_i + c(T) + d(T)S_i + e(T)V_i + f(T)M_i^2 + g(T)R_i\right]$$

The number of residuals $N(\epsilon)d\epsilon$ for which $\epsilon - \frac{d\epsilon}{2} \leqslant \epsilon_i < \epsilon + \frac{d\epsilon}{2}$ can easily be found. Let us represent the sum of $N(\epsilon)$ over a range $a \leqslant \epsilon \leqslant b$ by the integral $\int_a^b N(\epsilon)d\epsilon$. Clearly $\int_{-\infty}^{\infty} N(\epsilon)d\epsilon = N$, and $N = 546$ in our case.

Now consider the relationship:

$$P_a = \frac{1}{N}\int_{-\infty}^{P_\ell} N(\epsilon)d\epsilon \ . \qquad (2.2)$$

Here $P_a(P_\ell)$ is a distribution function of the residual ϵ; it gives the probability that any one residual ϵ_i will have a value less than P_ℓ. The meaning of P_ℓ here is the same as in equation (2.1). The reason for this is that a(T) and c(T) were chosen so that $P_a \approx P_\ell$ over the range $0.1 < P_a < 0.9$.

Using equation (2.2), Trifunac (1976) found values of P_a as a function of T for $P_\ell = 0.1, 0.2, \ldots, 0.8$, and 0.9. These are plotted in the top of Figure 2.1.

For the risk analysis in this paper, we shall identify:

$$q[S(\omega), \underset{\sim}{b}, \underset{\sim}{e}, \underset{\sim}{c}] = q[FS(T), M, R, s, v] = 1 - P_a$$

Thus the site parameters $\underset{\sim}{b}$ are represented by s and v, the source parameters $\underset{\sim}{e}$ are represented only by the magnitude M, and the path parameters $\underset{\sim}{c}$ are represented only by the epicentral distance R.

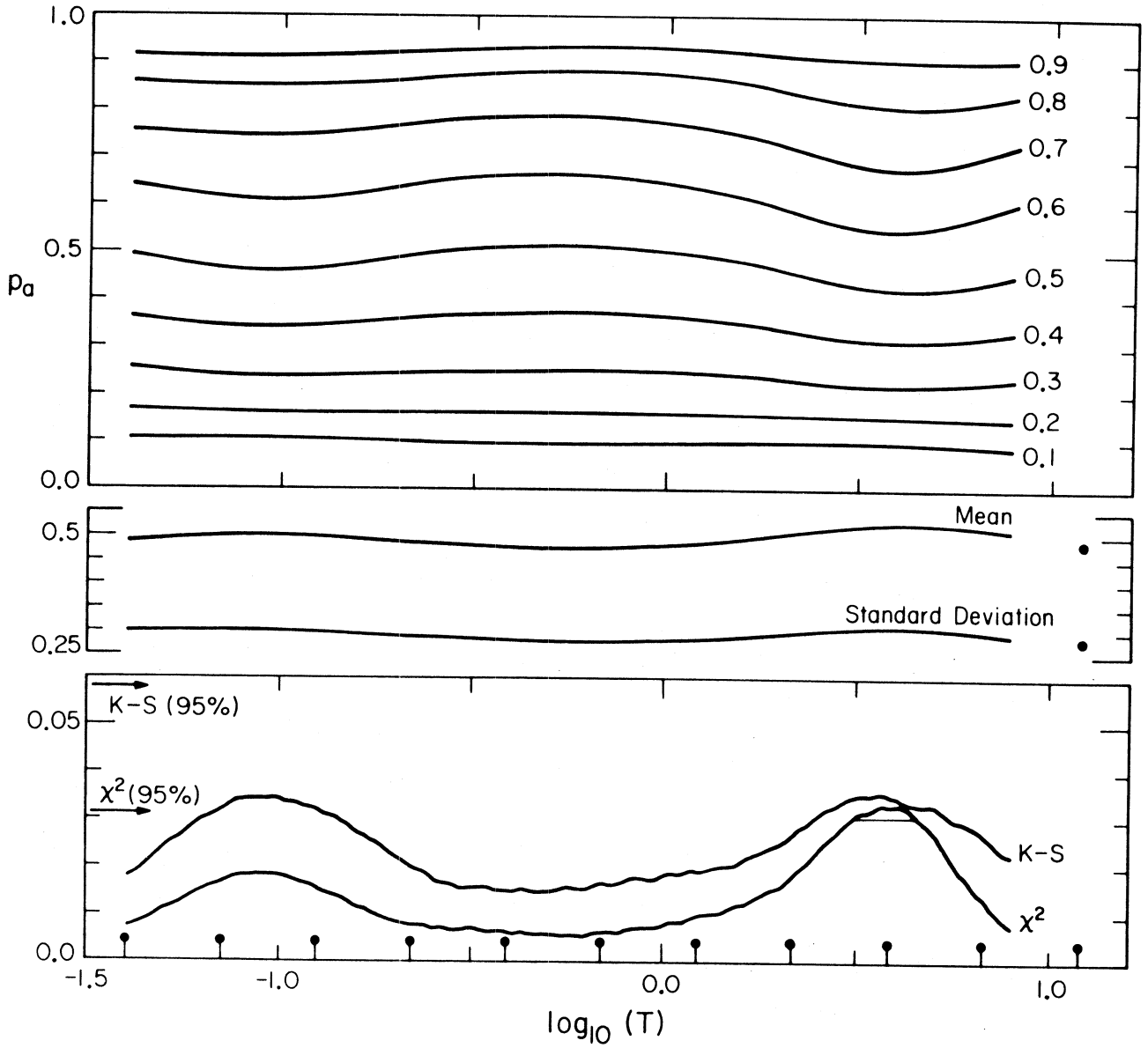

Figure 2.1. Derivation of P_a from P_ℓ. The data in the upper portion show P_a as a function of the period of seismic waves for the values of P_ℓ indicated at the right. The central portion gives the mean and standard deviation of the best fitting Gaussian distribution relating P_a to P_ℓ. The lower portion measures the quality of fit of these parameters according to the Kilmogorov-Smirnov and the χ^2 statistics. The eleven periods used for further calculations are shown on the base.

One means by which $q(\omega)$ could be found for the later application is to invert equation (2.1) to find P_ℓ for each event in the seismicity model, and use this with the data in Figure (2.1) to interpolate to the appropriate P_a. However, this approach would have two problems. The interpolation would be too slow for a procedure which must be repeated thousands of times in any risk analysis. More important, when P_ℓ is outside the range of 0.1 to 0.9, it is necessary to extrapolate. To handle these problems, we have fit a probability distribution to the set of points $P_a(P_\ell)$, shown that it is consistent with these points, and then used it in the later calculations.

Trial calculations showed that for the Fourier spectrum the relationship of P_a and P_ℓ could be approximated by a Gaussian distribution function at all periods. Specifically, we substitute

$$\frac{N(\epsilon)}{N} = \frac{1}{\sigma(T)\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}\left(\frac{\epsilon-\mu(T)}{\sigma(T)}\right)^2\right]$$ in equation (2.2), find the best values of the mean $\mu(T)$ and the standard deviation $\sigma(T)$ (in a least square sense) from the data in Figure (2.1) as described in Appendix 2.1, and use these values in the later calculations. In Figures 2.2, we show this approximation for two frequency bands. Figure 2.2a shows one of the "best" correlations, and Figure 2.2b shows one of the "worst" correlations. Figure 2.1 (center) shows the mean $[\mu(T)]$ and the standard deviation $[\sigma(T)]$ of these Gaussian curves determined at each of the frequency bands used by Trifunac. As one would expect from the way P_ℓ is determined, the mean is approximately 0.5; the standard deviation is also approximately constant and equal to about 0.3.

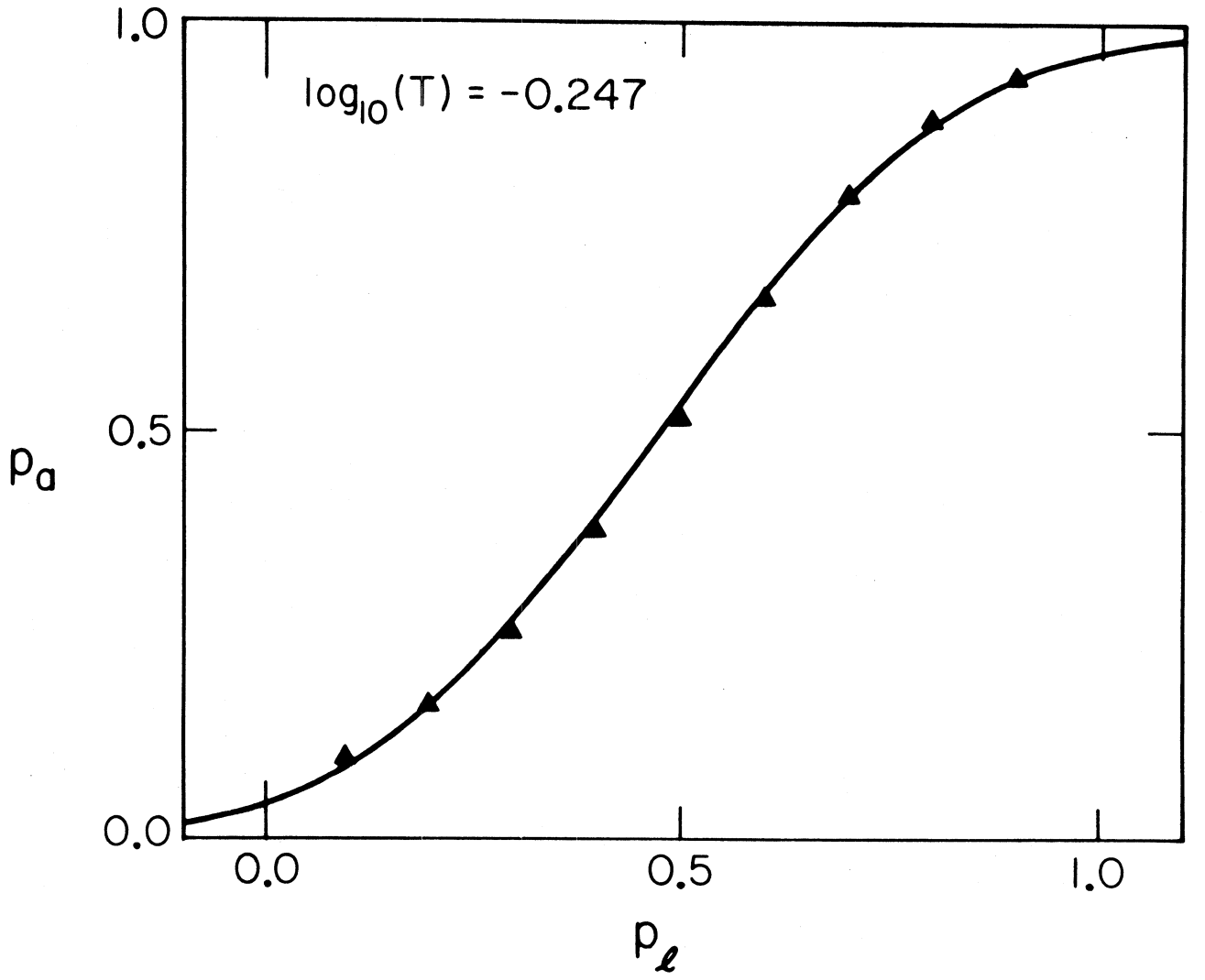

Figure 2.2a. Relationship of P_a and P_ℓ at a period where the Gaussian distribution is an excellent fit.

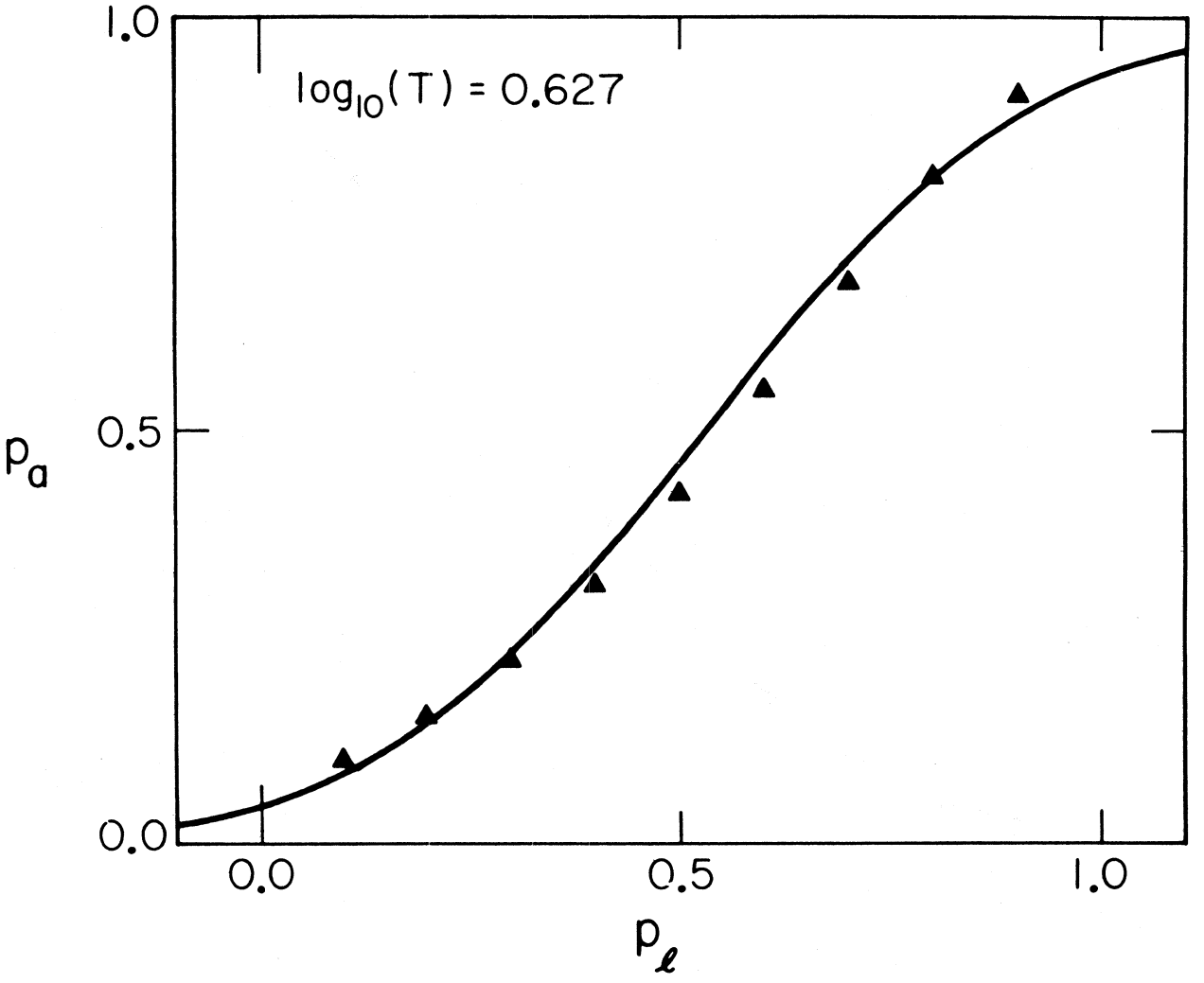

Figure 2.2b. Relationship of P_a and P_ℓ at a period where the Gaussian distribution is a poor fit.

Two approximate statistical tests demonstrate that the approximation by a Gaussian distribution is not contradicted by the values of P_a found by Trifunac. The first is Kolmogorov-Smirnov statistic (Hoel, 1971, p. 324), which applies to a probability distribution function directly. This statistic gives, for a given number of data points, the largest deviation that one may expect (at a given confidence level) between the distribution function derived from the data and the theoretical distribution function. We were forced to apply this in an approximate way because we have only 9 points chosen from a distribution function of 546 data points. Figure 2.1 (bottom) thus shows the maximum deviation of the data points from the theoretical curves at each of the periods studied by Trifunac. If the deviation at any period exceeded the 95% confidence level for the Kolomogorov-Smirnov test, as indicated on Figure 2.1, then we would have to reject the theoretical distribution at that period (at the 95% confidence level). It is clear that by this approximate test the distribution curves we have derived are not contradicted by the data.

We also applied a Chi-square test to see if it would agree with the results of the Kolmogorov-Smirnov test. We found the value of Chi-square at each period, shown in Figure 2.1, by numerically differentiating the theoretical and the observed distribution functions. The 95% confidence level for χ^2 is also shown in Figure 2.1. At log (periods) between 0.51 and 0.65, we might reject the Gaussian distribution functions on the basis of this test. We decided not to do this because (1) these distribution functions could not be rejected at the 99% confidence level, (2) because at other periods the Gaussian

functions are not rejected according to the Chi-square test, (3) because the Gaussian functions passed the approximate K-S test at all periods, and (4) because the preliminary correlations of Trifunac involve approximations which probably do not justify finding a more precise distribution function.

Note that even though use of a Gaussian distribution function makes it possible to extrapolate to values of P_a much outside the range of 0.1 to 0.9, such extrapolation is not supported by the data. We use a Gaussian distribution function for later calculations only because it approximates the distribution of the available data for Fourier spectrum. Clearly, there is no reason to assume that a Gaussian distribution would be applicable to other types of spectra, or even to a larger set of Fourier spectra.

The use of a Gaussian distribution carries with it the assumption that extremely large amplitudes are possible from any event, although they are very unlikely. It is possible that this assumption may introduce a bias into the amplitude of the uniform risk spectrum. However, we expect any such bias to be small, except possibly for very small probabilities of exceedance (e.g., less than 10^{-2}).

As noted by Trifunac (1976), the Fourier amplitude spectra in terms of his original parameters a(T),..., and g(T) still contain some processing noise at long periods, and thus as period increases they do not fall off as rapidly as may be expected. He suggests a fall off of T^{-1} in the near field; other considerations (e.g., Brune, 1970) suggest that T^{-2} could be more reasonable for periods of 3 to 15 sec. In any case, a T^{-1} decay appears to be a reasonable upper limit for this frequency range. Therefore, to minimize possible

distortions of spectral amplitudes which would be caused by the processing noise, we have modified the original parameters of Trifunac, to obtain a T^{-1} decay of amplitudes at periods longer than where the spectral shapes in Trifunac suggest such a decay should begin. The modified parameters are given in Table 1.

In summary, we have applied the scaling relationship derived by Trifunac (1976) for Fourier spectra to find the function $q[FS(T), M, R, s, v]$ needed in risk calculations. Although Trifunac's relationship is the best one available at present, we modified it somewhat in order to obtain greater computational convenience and in order to force the results at long periods to coincide more nearly with theoretical concepts. The result is acceptable; however, an improved scaling relationship would be welcome.

Incorporation of the scatter of amplitudes about the mean trend in the attenuation equation, as we have done, is one of the most significant improvements of our risk analysis over much of the previous work. Der-Kiureghian (1977) suggests that the uncertainty associated with the attenuation equation far exceeds the uncertainties associated with other aspects of the modeling, particularly the seismicity. Thus, by using the scaling relationship based on Trifunac (1976), we have eliminated a major source of potential errors in the seismic risk analysis.

TABLE I

Model parameters used to compute uniform risk Fourier amplitude spectra

$\log_{10}$ (T)	-1.398	-1.150	-0.903	-0.655	-0.407	-0.159	0.088	0.336	0.584	0.831	1.079
a(T)	-1.688	-1.620	-1.517	-1.445	-1.460	-1.514	-1.549	-1.570	-1.601	-1.630	-1.633
b(T)	-1.086	-1.380	-1.418	-1.216	-1.053	-1.129	-1.499	-2.592	-4.042	-4.699	-4.872
c(T)	7.615	7.892	7.344	6.249	5.587	5.913	7.328	11.230	16.381	18.875	19.715
d(T)	-0.018	-0.080	-0.068	0.011	0.102	0.163	0.189	0.197	0.200	0.204	0.203
e(T)	-0.098	-0.026	0.094	0.229	0.304	0.319	0.309	0.288	0.281	0.292	0.297
f(T) × 10	1.320	1.527	1.542	1.364	1.206	1.227	1.469	2.250	3.300	3.775	3.900
g(T) × 1000	-0.441	-0.869	-1.052	-0.940	-0.709	-0.610	-0.753	-1.033	-1.258	-1.352	-1.375
μ (T)	0.492	0.502	0.500	0.488	0.479	0.479	0.488	0.511	0.532	0.522	0.492
σ (T)	0.301	0.300	0.299	0.289	0.281	0.280	0.287	0.301	0.312	0.302	0.289

COMPUTER PROGRAM

We have written a Fortran computer program which evaluates the uniform risk spectra using the formulae in Chapter I and using the function q[FS(T), M, R, s, v] as just described. The program includes point, line and areal sources and a dipping planar source. This last source is included in an arbitrary manner by replacing the epicentral distance with the hypocentral distance in Trifunac's relationships. This decision cannot be justified by the data which is now available or by any simple physical arguments, and we have done it only because we lack a more realistic method. For point and areal sources, options allow either the assumption that the epicentral distance should be used or the assumption that the closest point to the fault should be used to calculate the risk. For the line source, the program automatically uses the closest point. We discuss this decision in detail in the next section.

The number of events of each magnitude is input in the form of a table; there is no requirement that this number is an integer, even though we have used integers for most of the calculations shown here.

The computer code computes the probabilities using both the Poisson and the "literal" assumptions described in Chapter I. For these cases, either the seismicity is treated entirely as the mean of a Poisson distribution or all the seismicity is treated literally. The code does not handle the case where some of the seismicity is treated as Poissonian and some of it is treated literally, but this refinement could be useful for future addition. For the "literal"

case, the program calculates $p^{\dagger}[S(\omega)]$ rather than $p^{*}[S(\omega)]$ because it is more convenient computationally and because $p^{\dagger}[S(\omega)]$ and $p[S(\omega)]$, from the Poisson case, provide an upper and lower bound to $p^{*}[S(\omega)]$.

APPLICATION TO IDEALIZED CASES

To illustrate the properties of the proposed uniform risk spectra, we calculated the spectral amplitudes at the six sites shown in Figure 2.3, and for the earthquakes located on the line source zone. We used two descriptions of the seismicity for this source: one is a single M = 7.5 earthquake which ruptures the entire length of the fault; the other has one M = 7.5 earthquake, and also smaller events on the fault. These smaller events can represent either aftershocks of the M = 7.5 earthquake or minor seismicity on the fault over a longer period of time. The numbers and magnitudes of the small events on the line source are given in Table II. Uniform risk spectra for this case are shown in Figure 2.4.

The spectra in Figure 2.4 show that at all six sites, the smaller events cause a small increase in the level of the uniform risk spectrum. The shapes of the spectra are somewhat affected by the small events, as the level is increased more at the high frequencies (about 40%) than at the long periods (around 10%). This is to be expected, because the spectra of small events derived by Trifunac show a considerably greater proportion of high frequency energy relative to long periods than the spectra of large events.

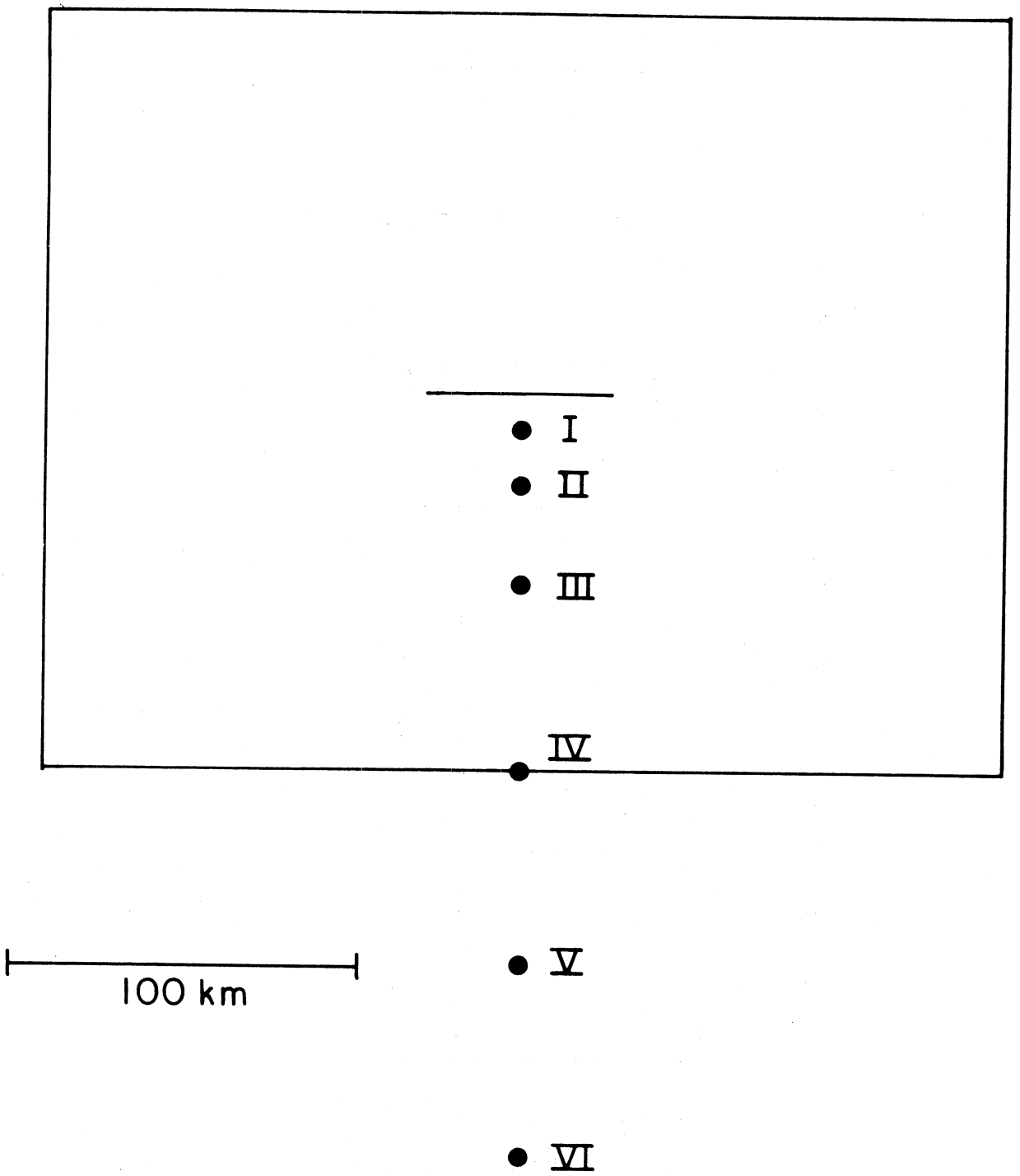

Figure 2.3. Geometry used for several calculations. The line in the center represents a short line source, and the box is the boundary of a diffuse region of seismicity. The six points labeled I, II,...,VI are sites where uniform risk spectra for this geometry have been calculated. Differing seismicities assigned to the source regions shown here are given in Table II.

TABLE II

Number of earthquakes in 50 years for the idealized source zones in Figure 2.3.

Source	Size of Source	Magnitude											
		3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
Line source	50 km	1109	407	150	55	20	7	3	1	0	1	0	0
Diffuse zone with low activity	6.08×10^4 km^2	3833	1212	383	121	38	12	4	1	0	0	0	0
Diffuse zone with high activity	6.08×10^4 km^2	11220	4217	1585	596	224	84	32	12	4	2	0	0

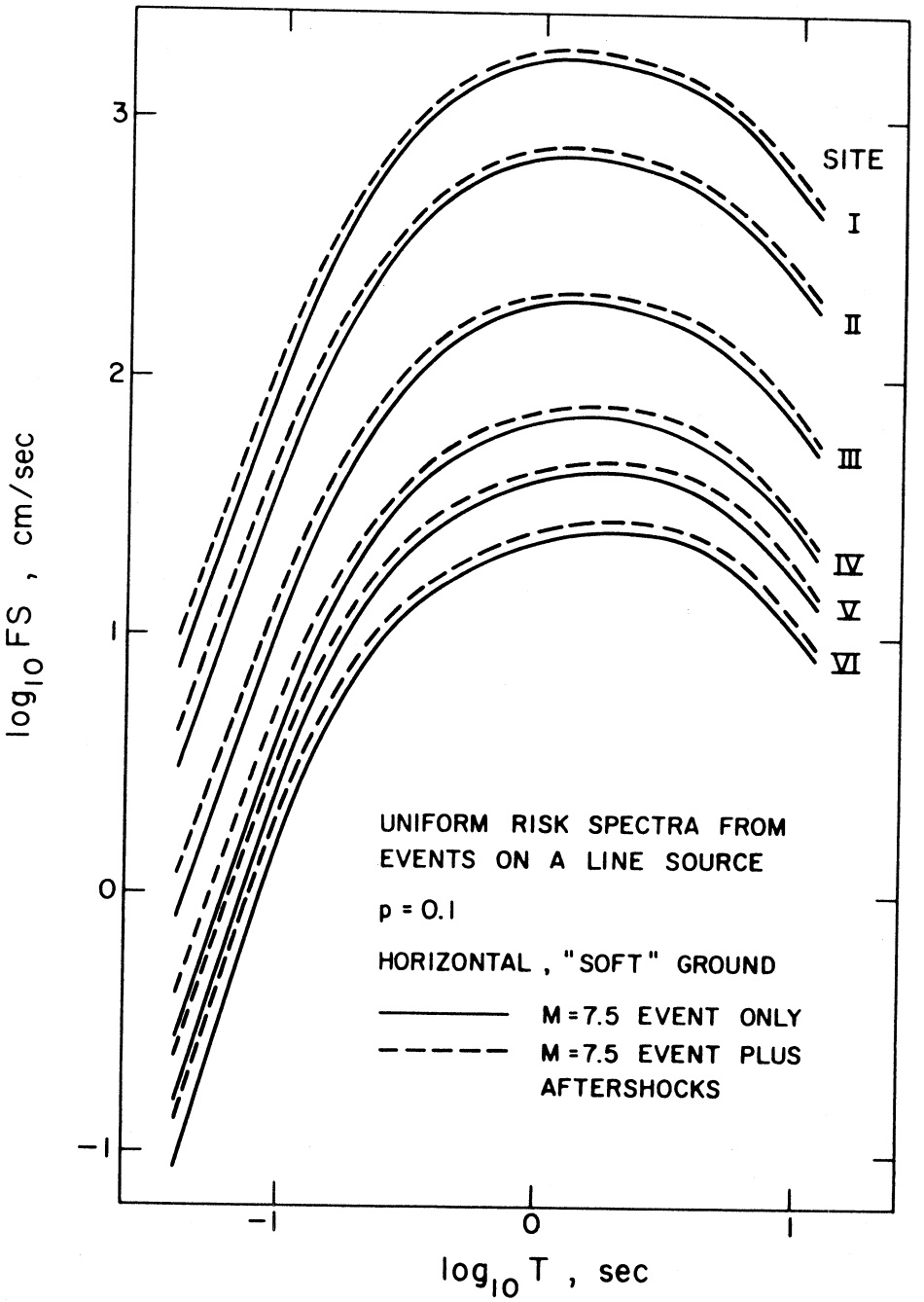

Figure 2.4. Uniform risk spectra for the line source shown in Fig. 2.3 with two examples of seismicity: first, one M = 7.5 event only, and second a M = 7.5 event plus minor seismicity which could represent either aftershocks or smaller unrelated events on the fault. The total seismicity for the model including minor seismicity is that listed in Table II for the "line source." These seismicities are treated as the mean of a Poisson distribution.

To study this effect further, we again employed the artificial geometry of Figure 2.3, and considered several additional seismicity distributions. The first (A) is the occurrence of the magnitude 7.5 event plus minor seismicity on the fault, as used in Figure 2.4 above. The second seismicity distribution (B) is the occurrence of events only in the diffuse zone outlined in the figure, with no events constrained to occur on the fault. The third model (C) is a superposition of both the previous sources (A) and (B). The fourth model (D) has the same total seismicity as model (C), but it is all uniformly distributed in the diffuse zone. The computation used the epicentral distance for events in the diffuse zone.

Computed uniform risk spectra with probability of exceedance of 0.1 for these models are shown in Figures 2.5 and 2.6 for two distinct rates of background seismicity in the diffuse zone (B). Figure 2.5 has a rather low rate, while Figure 2.6 has a rate about three times higher, and also has larger events. The number of events per unit area in Figure 2.6 is comparable, for example, to that which could occur in the Imperial Valley of California in a time span of 50 years (Hileman et al., 1973). The number of events per unit area in the diffuse zone for Figure 2.5 is more nearly comparable to the number which could occur in the Santa Barbara Channel region of Hileman et al. (1973) in a period of 50 years. The assumed seismicity rates are given in Table II.

The results for these cases should be illustrative of what might be expected in a more complex situation. We first consider Figure 2.5 with the low seismicity rates. Close to the fault, at stations I

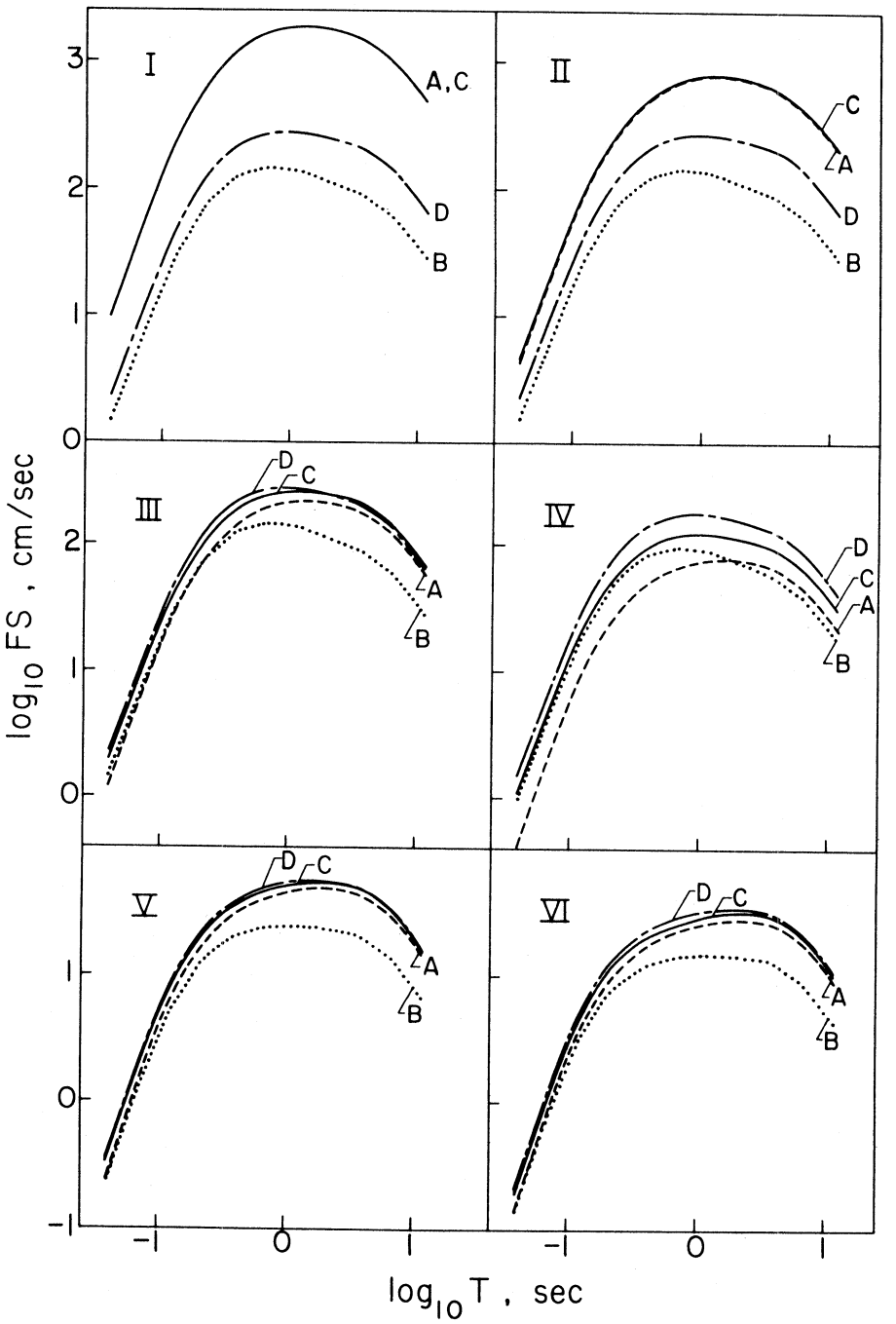

Fig. 2.5. Uniform risk spectra for a probability p = 0.1 of exceedance and for horizontal motion on alluvium. The Roman numeral (upper left) identifies the site for each set of spectra, and corresponds to the sites shown in Fig. 2.3. The letters A, B, C, D refer to four combinations of seismicity which are described in the text and which use the geometry of Fig. 2.3.

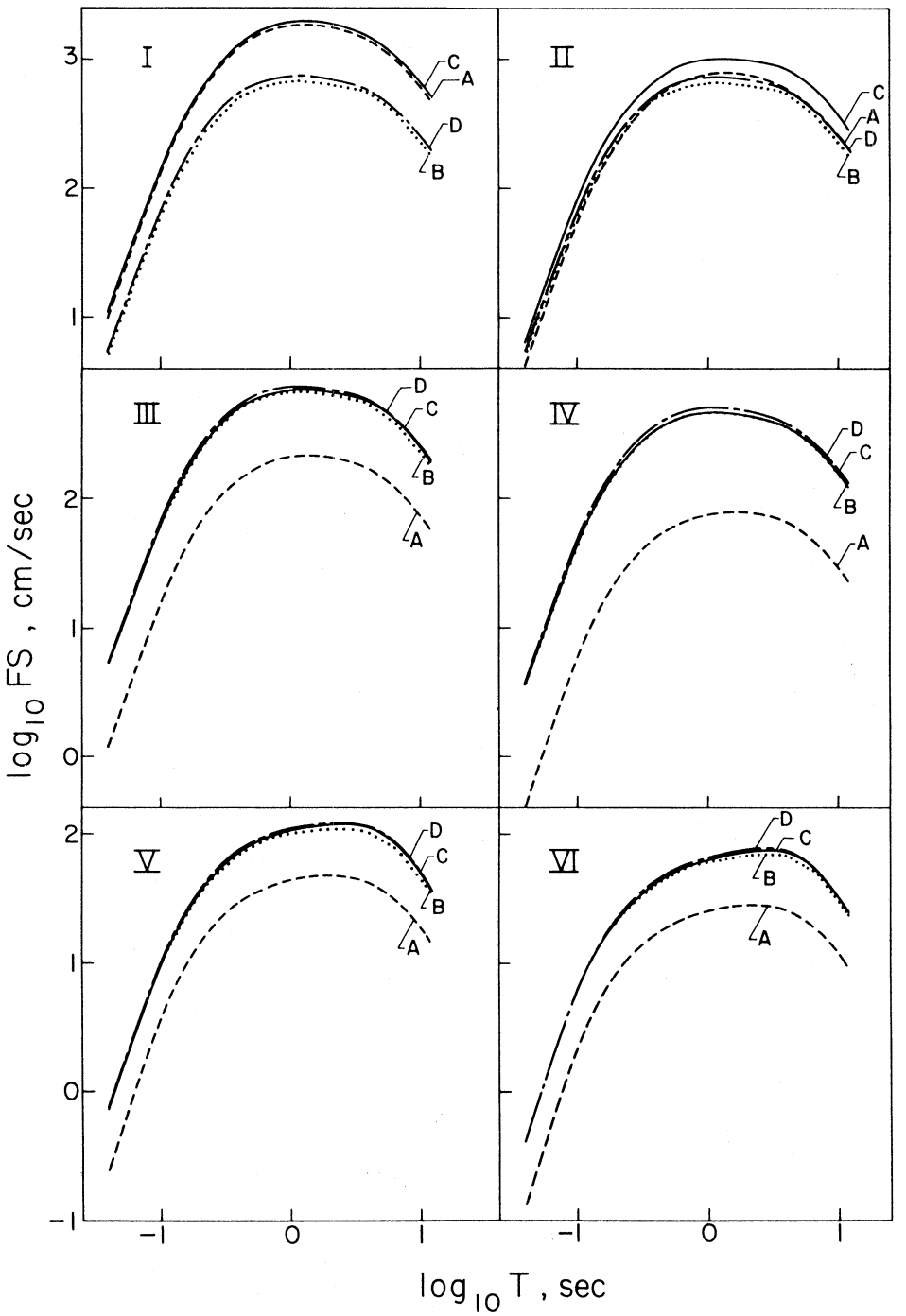

Figure 2.6. Same as Fig. 2.5 except that the low seismicity in the diffuse region for seismicity models B, C, and D is replaced by the high rate given in Table II.

and II (about 10 and 25 km away), the addition of the local seismicity does not affect the result. At station III (about 55 km away), the local seismicity has a small effect on the spectra at long period and a larger effect at short periods. At station IV, about 110 km away, this trend continues. At this site, it is interesting to compare the uniform risk spectra resulting from model C with the spectra resulting from models A and B. The combination of seismicity in model C causes larger spectral amplitudes at all frequencies than either model A or model B alone. At long periods, model C is controlled by the large events on the relatively distant ($\sim$ 110 km) line source of model A. At high frequencies, however, the more numerous small events in the diffuse zone (model B) control the amplitudes. At sites V and VI, the amplitudes of the uniform risk spectra begin to decrease because the sites are outside the seismic zone. Cases A and C tend to converge again as the effect of the local seismicity is absent. Case D, which has the same total seismicity as C but lacks the information about where the magnitude 7.5 event may occur, gives a risk level similar to case C at site III, and leads to the largest amplitudes at greater distances. Close to the fault, it gives much smaller amplitudes. Cases B and D give a constant spectrum every place interior to the region, so the amplitudes they predict do not change between sites I and III.

In Figure 2.6, the effects of the higher seismicity and the larger events in the diffuse zone nearly mask the effects seen in Figure 2.5. Already at site II, 25 km from the fault, the activity on the fault does not have much effect on the level of spectra, and at site III, the fault has essentially no effect.

A traditional seismic risk analysis of the cases just described would lead to a fixed shape spectrum being scaled up or down. The changing shapes of the spectra, although somewhat subtle in this case, emphasize the advantage of estimating the spectral amplitudes independently for each frequency. For example, the uniform risk spectrum for case C shifts between sites I and IV to have higher levels of high frequency waves relative to the long period amplitudes. On leaving the seismic zone, however, the high frequency levels decrease more rapidly than the long period levels.

Although it may not be apparent from these figures, the spectrum in cases (B) and (D) is the same at stations I, II, and III. Its amplitudes begin to decrease as the station approaches the boundary of the diffuse zone. Additional calculations demonstrated that the spectral levels of cases (B) and (D) begin to decrease between 30 and 50 km from the boundary of the diffuse zone.

ON THE INFLUENCE OF CERTAINTY IN THE DESCRIPTION OF SEISMICITY

Chapter I described a variation of the composition method which might be applied to cases where there is some reason to believe that the number of events in the source region is known. Such a situation could result from an accurate earthquake prediction, where precursory phenomena indicate an approximate magnitude of the earthquake and a region or, for example, a segment of a fault within which the earthquake will occur. In this case, for specific sites, we may want to know the probable range of the amplitudes of seismic

shaking to determine, for example, if preliminary safety precautions are advisable. This method is referred to as the "literal" assumption because the specified seismicity is treated literally, in contrast to the "Poisson" assumption, where the specified seismicity is treated as the mean of a Poissonian sequence in time.

These assumptions have been studied for four cases:

1. A single magnitude 7.5 earthquake on the line source in Figure 2.3. The rupture length was assumed approximately equal to the fault length. This case is shown in Figure 2.7.
2. A single magnitude 7.5 earthquake on a line source like that in Figure 2.3, but extended to twice the length of the source shown by adding an equal length onto each end. The spectra for this case are not shown.
3. A single magnitude 7.5 earthquake on a line source like that in Figure 2.3, but extended to five times the length of the source shown by adding an equal length onto each end. The spectra for this case are shown in Figure 2.8.
4. A single magnitude 7.5 earthquake in the diffuse region outlined in Figure 2.3. Again, the epicentral distance was used in the attenuation relationships. The spectra for this case are shown in Figure 2.9.

In Chapter I we noted that using the "literal" assumption always leads to larger amplitude than the "Poisson" assumption. This can be seen in Figures 2.7-2.9. In Figure 2.7, where the event is

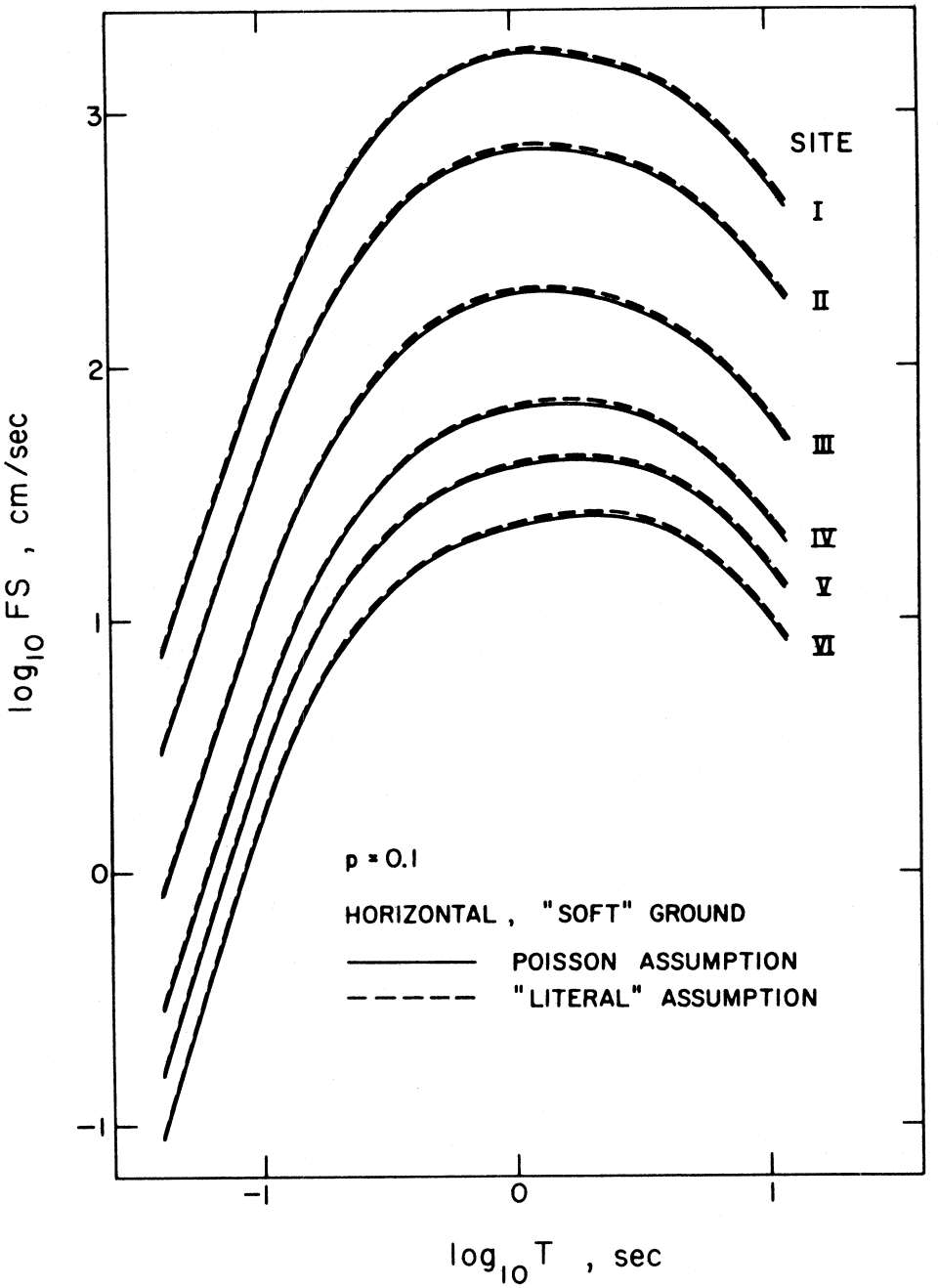

Fig. 2.7. Comparison of uniform risk spectra derived using the Poisson assumption and the literal assumption for a single earthquake on a short fault, of length comparable to the length of rupture. The fault and the sites are as shown in Fig. 2.3.

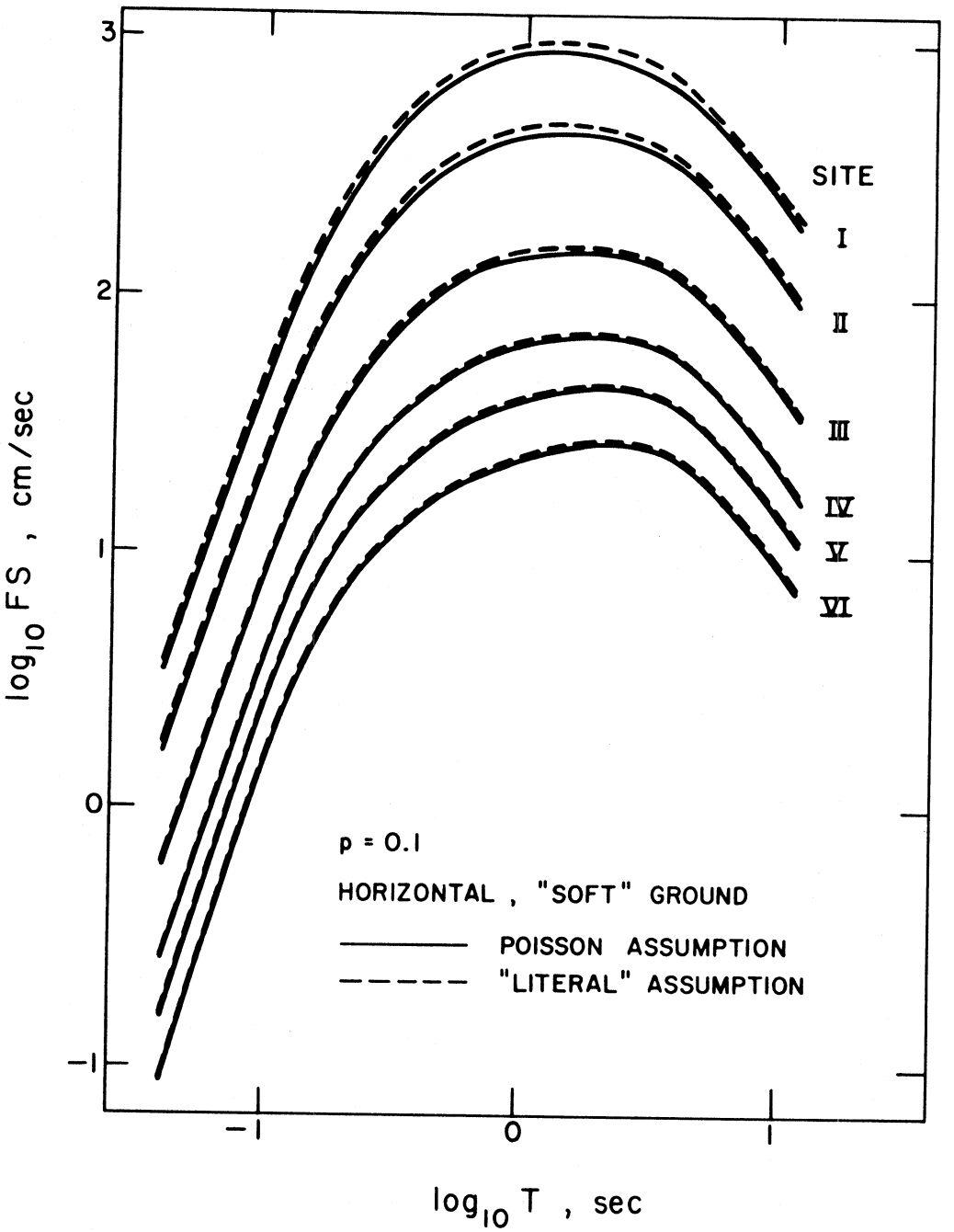

Figure 2.8. Similar to Fig. 2.7, but the fault is extended symmetrically so that its length is five times the length of rupture for the M = 7.5 earthquake.

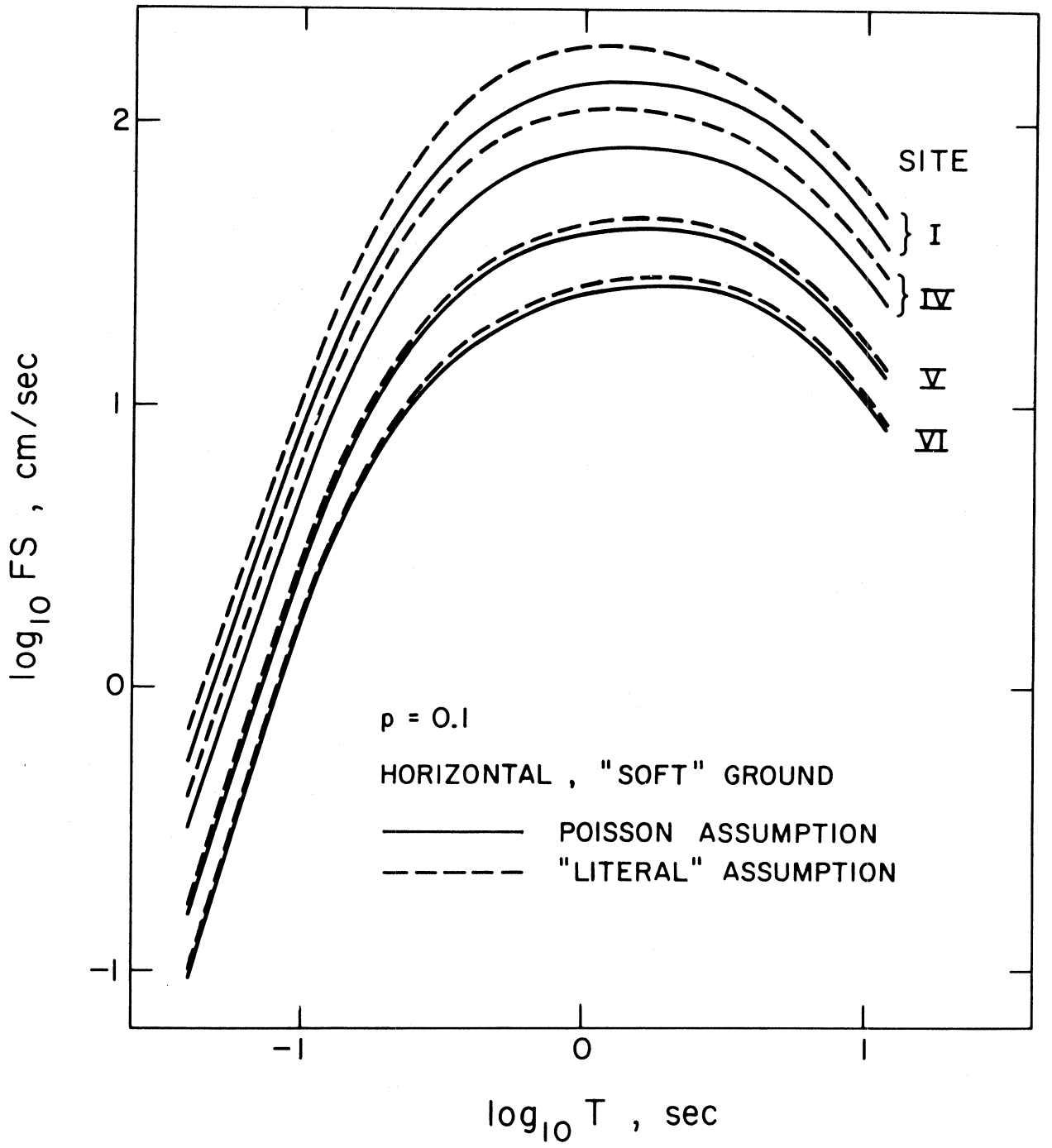

Figure 2.9. Uniform risk spectra derived using the Poisson and the literal assumptions for a single M = 7.5 earthquake in the diffuse zone shown in Fig. 2.3.

constrained to a relatively small region, the differences are minor and almost insignificant. In Figure 2.9, however, the stations inside the diffuse zone display large differences between the "literal" and the "Poisson" assumptions of about 25 to 40 percent (0.1 to 0.15 on the logarithmic scale). These differences decrease to under 10 percent outside the source region. Thus it appears that when an earthquake is predicted to occur within a diffuse region, a treatment of the risk similar to the "literal" treatment here is important.

The second and third cases are intermediate between case (1) and (4). Case (2) gives the same results as case (1), thus the spectra are not shown. This result occurs because the closest point of rupture is at the closest point of the fault to the site. If the rupture had been extended asymmetrically while maintaining a 2:1 ratio of fault length to rupture length, then the spectral levels would have been reduced somewhat, as they are in case 3 (Figure 2.8).

Figures 2.7 and 2.8 show how the relationship of fault length and rupture length affects amplitudes of uniform risk spectra. Because the fault length is greater for the third model (Figure 2.8) rupture is not forced to occur as close to site I. As a result, the spectral amplitudes at site I for the third assumption of seismicity (Figure 2.8) is considerably smaller than the spectral amplitude for the first seismicity model (Figure 2.7). At site VI, however, the differences between the two models do not significantly affect spectral amplitudes.

Comparison of cases 1 and 3 with case 4 (Figure 2.9) shows the large increase in amplitudes at sites I, II, and III caused by

constraining the event to occur on the fault. In contrast, at the distant sites V and VI, the amplitudes resulting from events in the diffuse zone (case 4) are slightly larger than those resulting from events constrained to occur on the fault (cases 1 and 3).

EFFECT OF AN EXTENDED SOURCE

Field observations have consistently shown that large, shallow earthquakes offset the surface of the ground along a distance of at least several kilometers (Allen, 1976). Intuitively, therefore, it appears necessary to account for such an extended source zone in a risk analysis. Der-Kiureghian and Ang (1975) and Douglas and Ryall (1975) have shown methods by which this can be done. However, it is not obvious that including an extended rupture length is always necessary.

To demonstrate this, consider the description of the attenuation of the amplitude of strong ground shaking with distance. In deriving these relationships, different authors have not agreed on how to define the distance between the earthquake and the site. For example, Figure 1 of Trifunac and Brady (1976) shows some curves that are related to the epicentral distance, some that are related to the hypocentral distance, and some that use the distance to the causative fault. It follows from elementary considerations that ideally, these relations should not have the same form. For example, the distance to the causative fault is always less than or equal to the distance to the hypocenter. Therefore, when a datum is plotted to define these two curves, the same value of acceleration, say, will occur at

a closer distance to the origin when the distance to the causative fault is used. As a result, since these attenuation curves decrease with increasing distance, then average acceleration for a given hypocentral distance will have a larger amplitude than the average acceleration for the same fault distance.

When these attenuation functions are applied to find the risk, it is helpful to remember that a seismic risk calculation attempts to estimate the numerical value of a physically well-defined function, such as $p[S(\omega)]$. To do this, the method used to determine $p[S(\omega)]$ must be internally consistent. However, there is no reason to assume that there is only one correct method to obtain the correct result.

As an example, a model which is formulated using the epicentral distance, such as Cornell (1968), needs an attenuation equation based on epicentral distance. A model which is formulated using the closest distance, such as Der-Kiuerghian and Ang (1975), needs an attenuation equation based on the closest distance. Inside a region of diffuse seismicity, the smaller fault to site distances in the latter method should be compensated for by somewhat smaller accelerations predicted for each distance by the attenuation equations formulated using the closest distance. Thus, once the statistical properties of the attenuation function and of rupture dimensions are known in a region, the two methods may be expected to give the same results.

The situation described above applies to the interior of a region of diffuse seismicity. In the presence of a line fault, a different situation holds. An attenuation equation based on epicentral distance incorporates, to a first approximation, random directions

of rupture relative to the site. However, the line fault constrains the rupture to a certain direction. In order to maintain logical consistency for this case and still avoid introducing unreasonable complexity when the direction of rupture is prescribed, it appears necessary to use an attenuation equation based on the shortest distance.

By doing this, new uncertainties are introduced into the risk analysis, particularly in describing the seismicity. Rather than N(M), we now need $N(M,\ell)$, where ℓ is the rupture length. For large magnitudes, the data necessary to estimate such a distribution are not available. Thus, we are nearly forced to assume that ℓ is determined by M. Several persons have proposed empirical relationships of the form $\log \ell = a + BM$, but it is hard to evaluate whether these are widely applicable. Conceptually, of course, describing an earthquake by a rupture length is still an approximation to the physical situation, in which motion occurs in part of a fault zone.

At present, the above considerations are almost entirely academic, because among the presently existing relationships for the attenuation of peak acceleration with distance, there is no systematic difference between those using the closest point on the fault and those using epicentral distance (Figure 1, Trifunac and Brady, 1975). This apparently results from the scatter in the data and the lack of sufficient data for events with rupture dimensions large enough to have an effect. Therefore, for all we know, the attenuation equation of Trifunac (1976) may be nearly equal to the correct equation based on the closest distance, even though it was derived

using the epicentral distance. Because of this similarity and the consideration described earlier, we have used an extended rupture zone for all events on the line source.

For a diffuse source region it is not clear whether we should use the epicentral distance or the closest distance with the attenuation equations of Trifunac, and either method can be used as an option in our computer program. We made a number of numerical calculations to determine the discrepancy in amplitudes which is involved. By using the same attenuation equation for both methods of measuring the distance from the fault to the site, we first determine when the two assumptions lead to different results, and then we consider an idealized but realistic case to learn how great the difference could be. The calculations treat rupture as unilateral, with a random direction of propagation from the epicenter.

The first calculation consisted of studying the uniform risk spectra from single events. Two computations were done for each of many combinations of fault lengths and epicentral distances, one using the epicentral distance (Method 1 of Appendix 1.2) and the other considering different possible closest distances (Method 2). Differences in the spectra for these two methods were calculated, and the results are summarized in Figure 2.10. In that figure, the axes are rupture length and epicentral distance. The contours indicate approximately the percentage increase of the level of the uniform risk spectrum caused by using the extended rupture over the level which results from a point source. Using Trifunac's (1976) model of attenuation, this increase for any one combination of

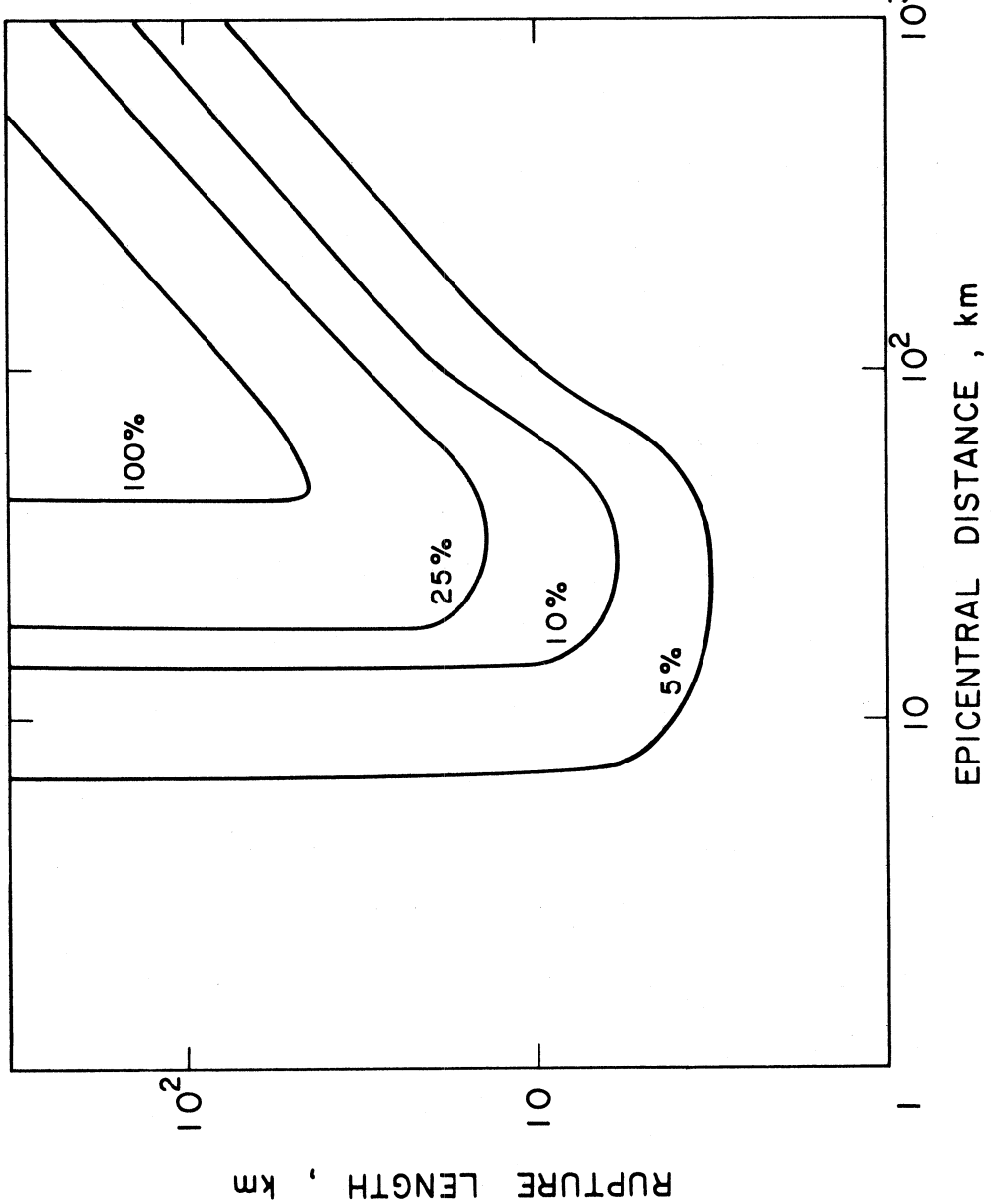

Fig. 2.10. Approximate contours of the percent increase in the level of the uniform risk spectrum (Poisson, p = 0.1) resulting from a single earthquake when it is treated as a unilateral rupture in a random direction rather than as a point source.

epicentral distance and rupture length is not a strong function of wave frequency, and thus the values contoured are averaged over the entire frequency band for which spectra are studied here.

These results are independent of both the magnitude of the earthquake and the probability of exceedance. The increase in the spectral level results entirely from the extent to which the average distance between the fault and the site is reduced from the epicentral distance, and the way the attenuation function behaves over that distance range. Thus, for large epicentral distances, the percentage of increase diminishes as distance increases because the slope of the attenuation curve generally decreases with distance. For short distances, the amplitude of the effect reaches an upper limit when the fault length surpasses the epicentral distance.

As described in Appendix 1.2 of Chapter I, our method of including an extended rupture in an areal source zone assumes that the rupture is unilateral. Other possibilities can be included in the risk model, such as bilateral or non-linear rupture. The effect of any of these assumptions on the contours in Figure 2.10 would be to decrease the effect of the extended rupture relative to the calculations based on the epicentral distance. Thus, these contours give an upper bound on the effect.

The results of Figure 2.10 give a way to evaluate an upper bound of the effect of the extended rupture on the set of accelerograms used by Trifunac (1976) to derive the attenuation relationship. To this end, in Figure 2.11 we have superimposed a set of data points to indicate where each of the accelerograms used by Trifunac were recorded with respect to these contours.

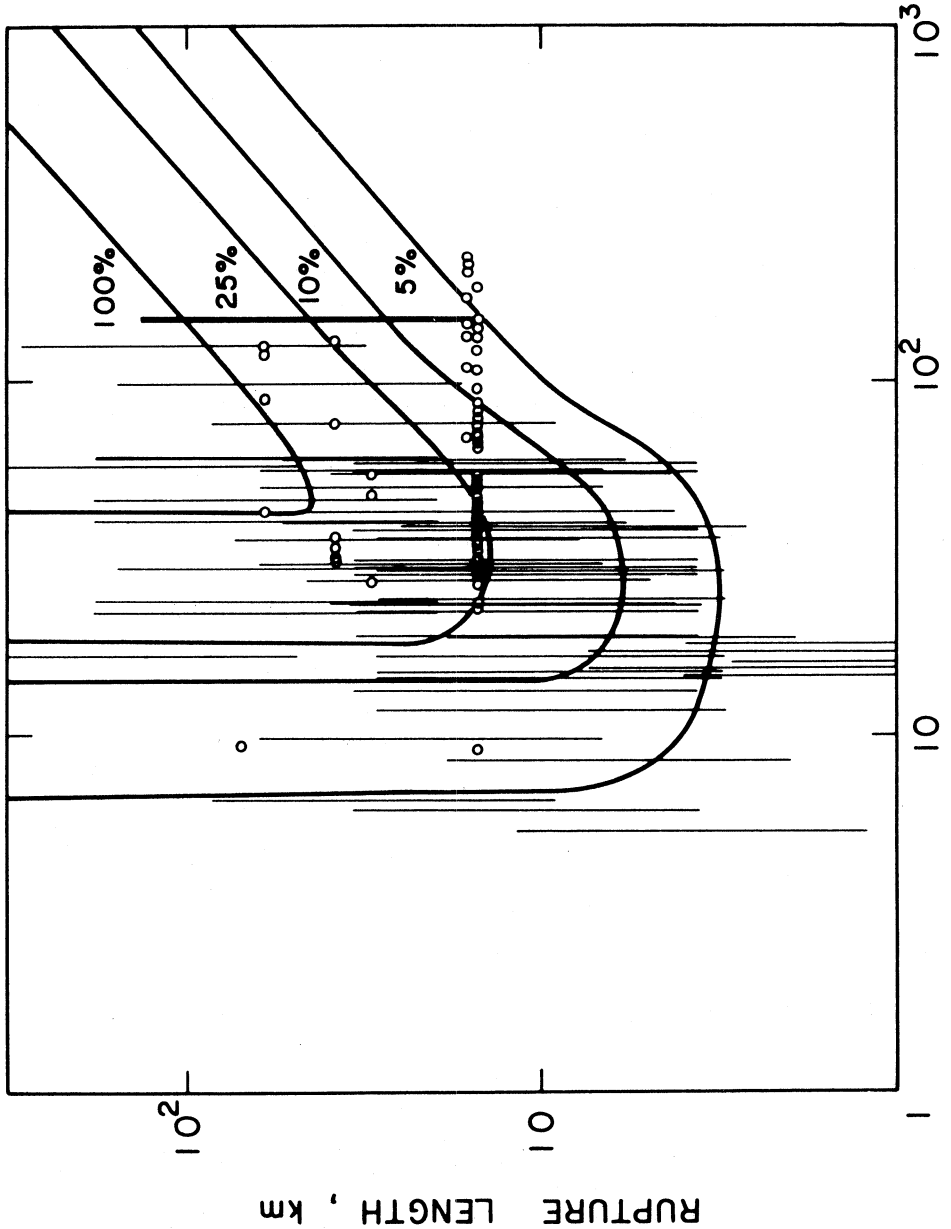

Fig. 2.11. Data used by Trifunac in studying scaling relations of strong ground motion superimposed on Fig. 2.10. Each circle represents one accelerogram from one of the six earthquakes listed in Table III; each vertical line represents one accelerogram from one of the remaining 51 events. The lines span the range of rupture lengths predicted by the equations of Thatcher & Hanks (1973) for 1 bar and 100 bars stress drop.

One or several estimates of the rupture lengths are available for six earthquakes which lead to most (126) of the accelerograms used by Trifunac (1976). These earthquakes, with the rupture length we used, are listed in Table III. The estimate of rupture length for the San Fernando earthquake is an average of the two dimensions of the fault given by Trifunac (1974). For the remaining 60 records (corresponding to 51 earthquakes) we plotted a vertical line in Figure 2.11 which represents the range of the rupture length to be expected from the relationships of Thatcher and Hanks (1973). The larger rupture length for each event represents a low stress drop event (1 bar), while the lower rupture length represents a high stress drop (100 bars).

From the way the data fall on Figure 2.11, it appears that the extent of the rupture might be expected to have some effect on the relationships derived by Trifunac (1976). Since his relationship is based almost entirely on earthquakes with magnitude less than about $6\frac{1}{2}$, it may be reasonable to suggest that in a diffuse source zone with no events greater than $M \approx 6\frac{1}{2}$, it is reasonable to apply his relationship using the epicentral distance. The next example investigates the consequences of this suggestion.

For the next numerical example, we compare the uniform risk spectra in a region of diffuse seismicity derived using the two forms: first using the epicentral distance, and second using the shortest distance to the fault. The geometry for this problem again is shown in Figure 2.3. We use the diffuse region of activity of the previous sections, and examine the uniform risk spectra at the same six sites and for the low activity rate given in Table II.

TABLE III

Rupture lengths and sources of information for six earthquakes which lead to 126 of the 186 accelerograms used by Trifunac (1976). These rupture lengths and the epicentral distances for the 126 accelerograms are plotted in Fig. 2.11.

Event	Date	Rupture length (km)	Reference
Long Beach Calif.	March 10, 1933	30	Geller (1976)
Imperial Valley Calif.	May 18, 1940	70	Geller (1976)
Kern County Calif.	July 21, 1952	60	Geller (1976)
Parkfield Calif.	June 27, 1966	38	Brown & Vedder (1967)
Borrego Mtn. Calif.	April 8, 1968	16	Burdick & Mellman (1976)
San Fernando Calif.	Feb. 9, 1971	15	Trifunac (1974)

Two of the many proposed relationships between magnitude and fault length were used in this calculation. They are:

(1) $\log_{10} \ell = \frac{2}{3}M - 3.41$ — Thatcher and Hanks (1973) $\Delta\sigma = 100$ bars

(2) $\log_{10} \ell = 0.53M - 1.47$ — Wyss & Brune (1968)

A fixed relationship such as this is not necessary; with little additional difficulty we could include a range of rupture lengths for each magnitude.

The uniform risk spectra for these relationships and also for calculations based on the epicentral distance are shown in Figure 2.12. Figure 2.12 only shows the uniform risk spectra at four of the six sites in Figure 2.3; at the two sites left out, sites II and III, the spectra are nearly identical to those at station I.

As expected, when the shortest distance to the fault is used, the spectral amplitudes are increased from the level which results from using the epicentral distance. The increase is greater at the 0.1 probability of exceedance level than at the 0.9 probability of exceedance level. The amount of the increase cannot be predicted from Figure 2.10 because in these examples the sources are in a diffuse zone rather than at a point and the effects in Figure 2.12 are an average from all the possible locations in the zone. However, the general trends are what would be expected. For example, the Wyss & Brune (1968) relationship gives the largest event in the zone ($M = 6\frac{1}{2}$) about a 100 km rupture length, while the Thatcher & Hanks (1973) relationship gives the same event about a 10 km

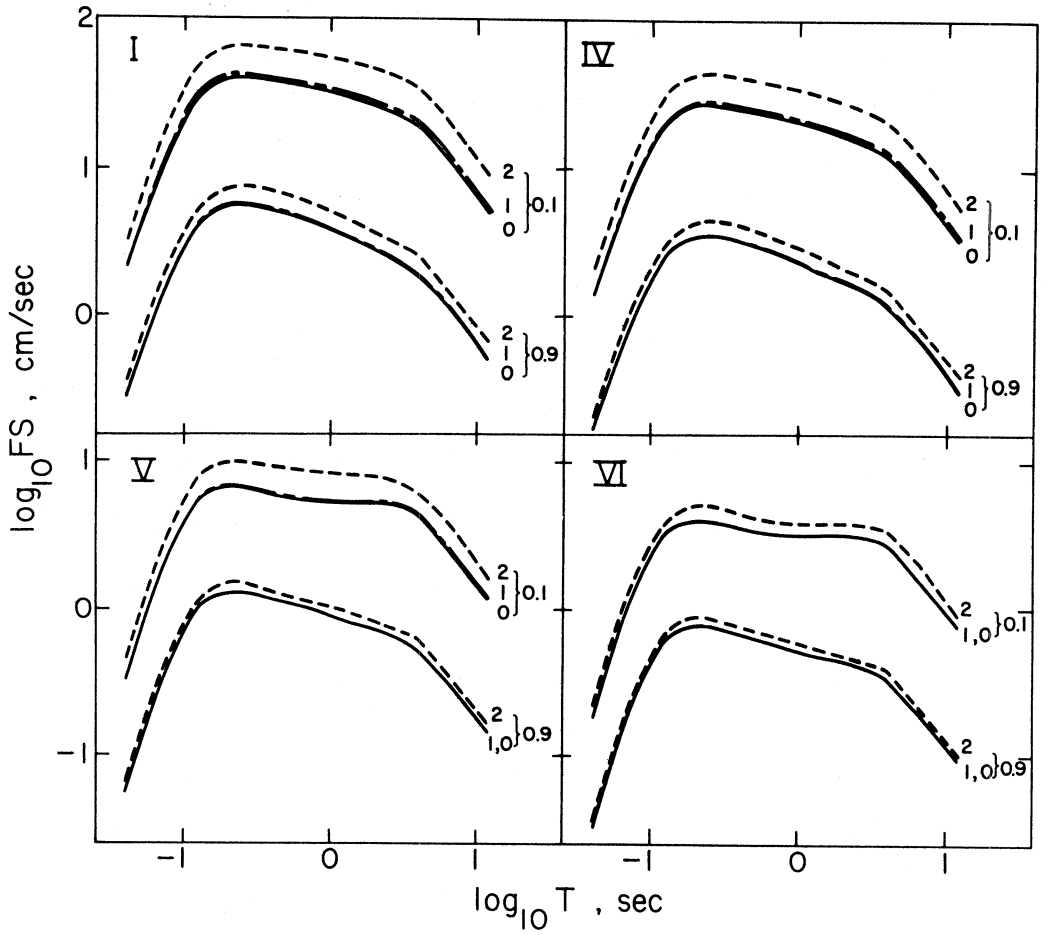

Fig. 2.12. Uniform risk spectra at four of the sites of Fig. 2.3 for low seismicity in the diffuse region and three assumptions: spectra labeled "0" use the epicentral distance in calculating the risk; those labeled "1" use the closest point and the magnitude-fault length equation of Thatcher & Hanks (1973) for a 100 bar stress drop; those labeled "2" use the closest point and the equation of Wyss & Brune (1968). The two sets of spectra at each site are for probabilities of exceedance of 0.1 and 0.9. The spectra are for vertical acceleration on rock sites. The seismicity in the diffuse zone is the low rate given in Table II.

rupture length. For a 10 km rupture, Figure 2.10 shows that the amplitudes of uniform risk spectra are never increased by more than 10 to 15 percent, and at all the sites in Figure 2.12, the effect of this rupture is about 5 to 10 percent, or less. The increase for a 100 km rupture length in Figure 2.10, on the other hand, is several tens of percent and can reach a factor of over two, and a large effect (about 75% at site I) is observed for the diffuse source (Figure 2.12). Also, as expected, the effect of the extended rupture diminishes as the sites move away from the source zone.

Thus a considerable discrepancy remains to be resolved. As suggested above, it is possible that the scaling relationships of Trifunac (1976) may be adequate to use the epicentral distance in this case; unfortunately, we cannot be certain. It is also possible that there was not enough data with extended ruptures for the scaling relationship to be used this way; if that is right, then the epicentral distance may lead to too low an estimate of the risk even in this case, and the extended rupture may give a more realistic estimate of the risk.

Some of the spectra from Figure 2.12 are shown again in Figure 2.13, this time to compare the rates of attenuation. Here it becomes apparent that the spectra for the extended rupture zones are more strongly affected at the long periods than at the higher frequencies. It is also clear that the spectra within the diffuse zone are considerably less than the spectra for an M = 6.5 event, which is the largest event in the zone. This results from the large average distances from the site to the earthquake.

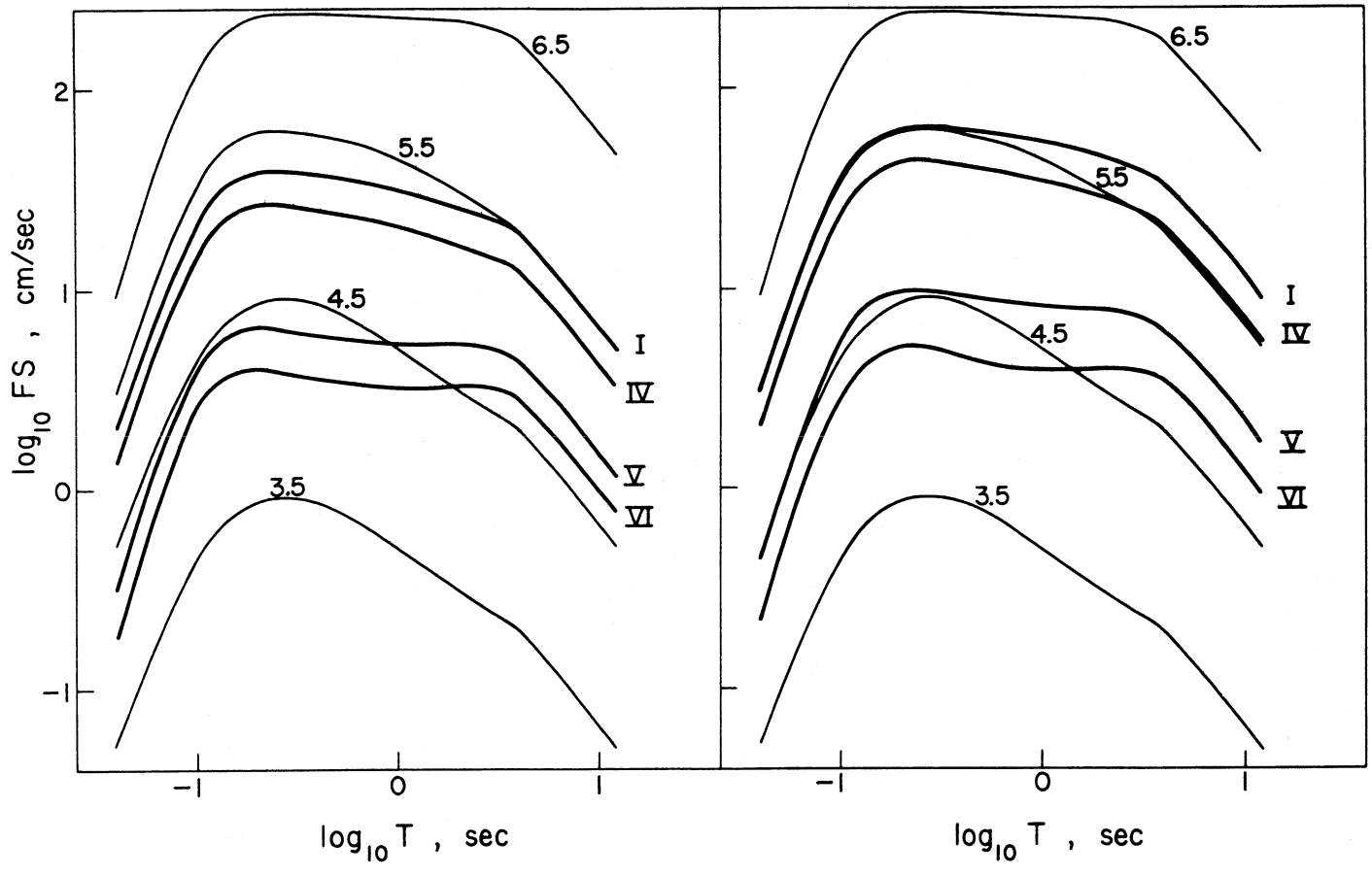

Fig. 2.13. Uniform risk spectra with probability of exceedance of 0.1 for the sites I, IV, V, and VI in Fig. 2.3 and a low rate of seismicity in the diffuse zone (Table II). The spectra on the left are calculated with the epicentral distance, and those on the right are calculated with the closest point on the fault and the magnitude-rupture length relationship of Wyss & Brune (1968). The light lines are the spectra of a single earthquake of the indicated magnitudes at the site. All spectra are for the vertical component on rock (s = 2).

The effect of using an extended rupture depends on the seismicity model; for example a larger maximum magnitude would have a larger rupture length, and thus the spectra for the extended source model would be increased by a greater percentage over the spectra of a model using the epicenters. Also, for a larger rupture length, the effect of the extended rupture would not diminish as rapidly as it does here for sites outside of the diffuse zone.

AN APPLICATION

To study the properties of the proposed model in a realistic setting, we apply it to finding the seismic risk for a site on the north coast of Puerto Rico. This particular site was studied by Der-Kiureghian and Ang (1975). They used response spectrum rather than Fourier amplitude spectrum, so that our current results are not directly comparable. A future study dealing with response spectra will permit a direct comparison. Here, we concentrate on the effects of differing descriptions of the same seismicity on the uniform risk spectra. To ease comparison in the anticipated work, we described the seismicity using the data available to Der-Kiureghian & Ang (1975), rather than obtaining a more current listing of epicenters. That data was, essentially, a list of all events known to have occurred in the vicinity between 1915 and 1971. The coverage of small earthquakes improved dramatically in this time period.

Two aspects of the seismicity are shown in Figure 2.14. The top, left seismicity map shows the locations of all known, major earthquakes since 1900. The locations shown are those listed by Der-Kiureghian and Ang (1975), rather than the improved locations given by Kelleher et al. (1973), but the discussion of Kelleher et al. was considered in the subsequent models. The top, right seismicity map shows the locations of smaller earthquakes during the period 1964 through 1971.

On the basis of these maps, the complete listing of events, and some additional information in Kelleher et al., four models of the seismicity of this region were found by four separate individuals, who were asked to describe the seismicity they expected in the region in the next 50 years. These four models are shown in Figure 2.14 and are labeled A., B., C. and D.; the uniform risk spectra derived from these models are shown in Figure 2.15. Although some of these models are more realistic than others, they demonstrate four distinct philosophies in describing the seismicity. For each model, the map is divided into zones and the number of events in each of the zones is given in Table IV.

Model A assumes that the seismicity in the future will essentially be the same as the seismicity that is known to have occurred in the past. The map is divided into five zones, and the seismicity of each zone is scaled slightly to account for the difference in time periods between the period of data and the period for which risk is to be estimated. The model does not consider the incomplete coverage of small earthquakes for much of the historical record.

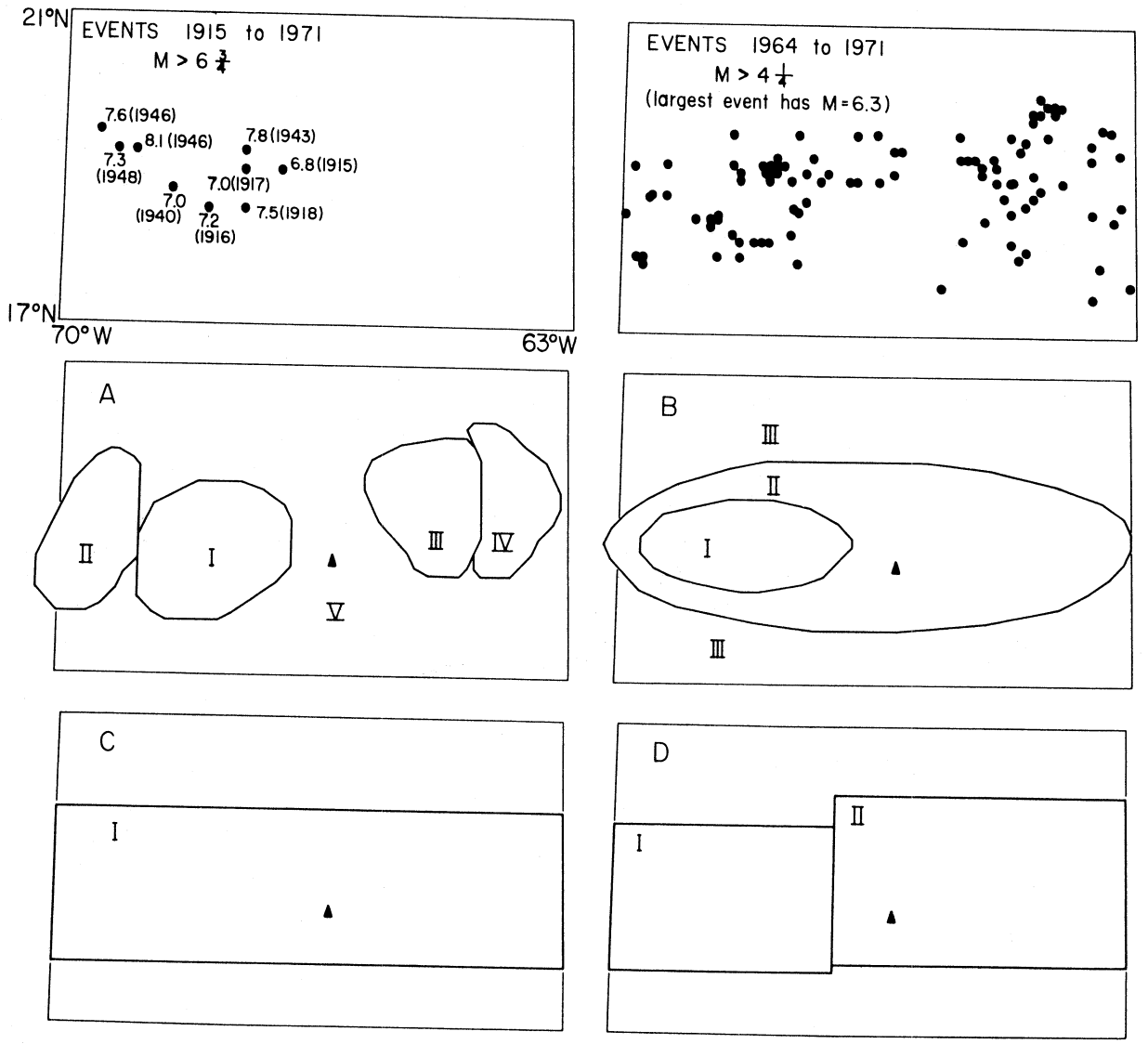

Fig. 2.14. Seismicity in the vicinity of Puerto Rico, as given by Der-Kiureghian & Ang (1975), and 4 models to describe this seismicity. The boundary of all 6 maps is the same as shown in the upper left. The map at the upper left shows the epicenters of major earthquakes since 1915; the map at the upper right shows events with magnitude greater than $4\frac{1}{4}$ which occurred between 1964 and 1971. The estimated seismicity rates in each zone of the 4 models (A, B, C, D) are in Table IV. These zones are the projections of the dipping planar source region for the risk estimates in Fig. 2.15 (right) and are diffuse shallow zones for the risk estimates in Fig. 2.15 (left). Models are for the site shown by a solid triangle.

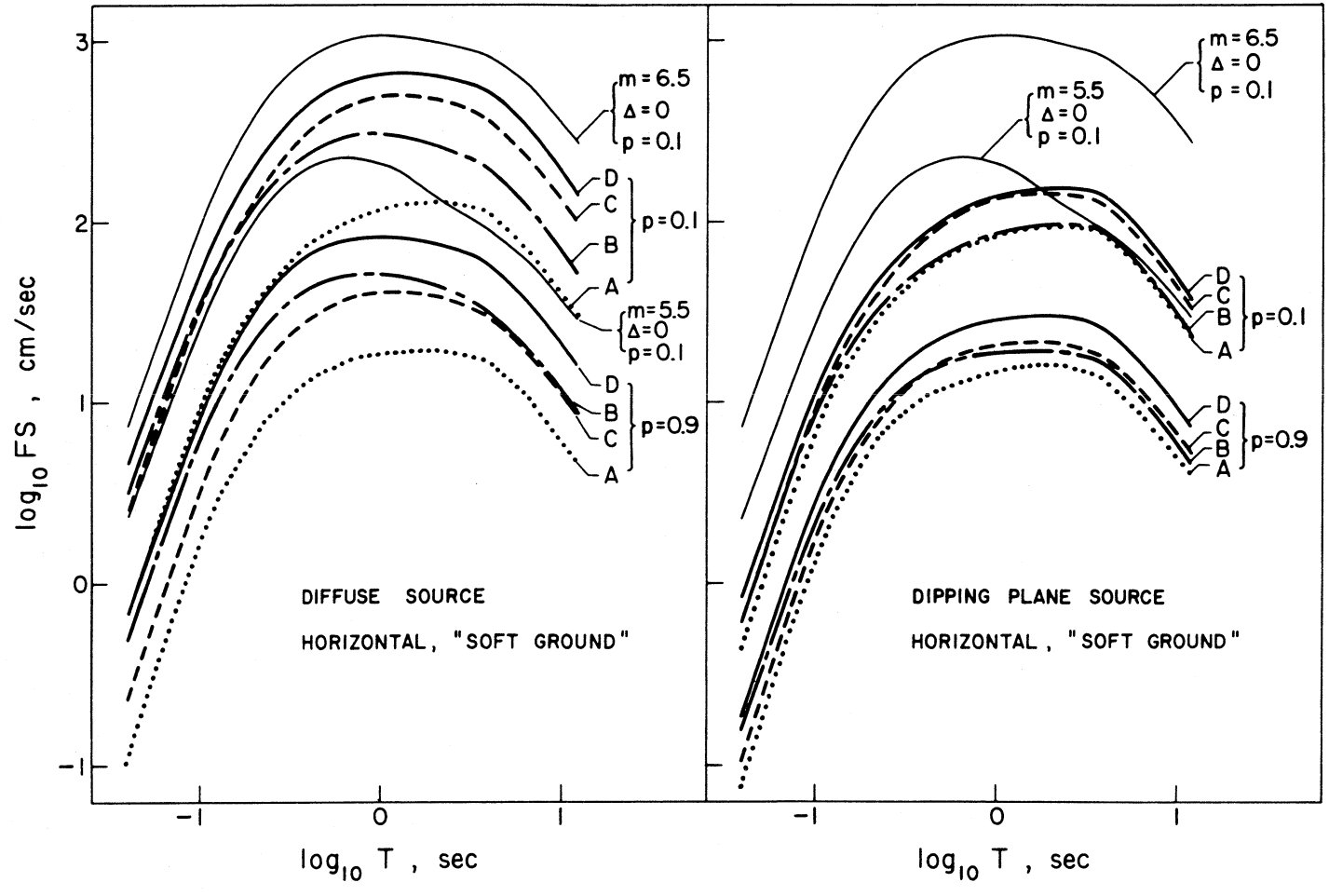

Fig. 2.15. Uniform risk spectra for the seismicity models of the Puerto Rico vicinity described in Table IV and illustrated in Fig. 2.14. These are shown for the probabilities of exceedance of both 0.1 and 0.9. For scale, we have also superimposed the spectra of an M = 6.5 and M = 5.5 earthquake at $\Delta = 0$. On the left, all the earthquakes in the source zones are assumed to have a shallow focus; on the right, they occur on a plane dipping to the south from the Puerto Rico trench.

TABLE IV

Number of earthquakes assigned to the source zones of Figure 2.14 for the Puerto Rico applications

Case	Zone	Area 10^4 km^2	Magnitude											
			3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
A	I	3.64	91	59	38	25	16	10	7	4	3	2	1	0
	II	2.48	23	17	12	9	6	5	3.3	2.4	1.7	1.3	0.9	0
	III	2.60	58	35	22	13	8	5	3	2	0	0	0	0
	IV	1.97	83	44	23	12	6	3	2	1	0	0	0	0
	V	22.24	107	54	28	14	7	4	2	1	0	0	0	0
B	I	3.09	3583	1095	319	146	44	29	13	7	5	3	0.5	0
	II	11.69	13555	4142	1209	554	166	111	50	0	0	0	0	0
	III	17.92	20779	6350	1852	0	0	0	0	0	0	0	0	0
C	I	16.33				66	40	23	13	6	5	3	2	1
D	I	6.65				501	178	63	22	8	3	1	0	0
	II	10.25				889	315	112	40	14	5	3	1.5	0

Model B assumes that the major seismic activity in the future will continue to occur where the major activity of the past was recorded. The numbers of smaller events are scaled from the rates recorded in most recent times.

Model C observes that the entire band of seismic activity crossing the seismicity map on the top right in Figure 2.14 is associated with an active plate margin; it makes no differentiation between the western portion which has a history of large events and the eastern portion which does not have that history.

Model D, finally, considers the possibility that because, as Kelleher et al. (1973) point out, the eastern region has not had any serious historic earthquakes, that this portion of the plate margin must be considered a more likely candidate for a large earthquake in the future. The western region was given a description of the seismicity which is lower than the historic rates on the assumption that some of the strain there was relieved by past events, and would take some time to accumulate to a dangerous level again.

We have calculated the uniform risk spectra for two variations of each of these models. The first variation (the left side of Figure 2.15) assigns each of the zones to be a diffuse region on the surface. The extent of rupture is not included; clearly, from Figure 2.10, including this could have a significant effect on the amplitudes of the uniform risk spectra.

The second variation (the right side of Figure 2.15) assigns each of these zones to be a diffuse region on a plane dipping south from the northern edge of the region of high activity at an angle of about 35 degrees. For this variation, which probably is more

realistic in describing the geometry of the seismic zone, we had to make an arbitrary extension of the model for attenuation. We did this by replacing the epicentral distance with the hypocentral distance. We emphasize that this is an arbitrary decision, and the results we obtain using it are perhaps not as reliable as one would desire. Thus, the results of these calculations represent a more realistic source geometry but a less reliable attenuation function.

Several observations can be made on the spectra in Figure 2.15. First, the differences between the spectral amplitudes resulting from cases A-D for a given probability level is considerably greater for the diffuse source than for the dipping plane source. This arises because the large events, which contribute most to the spectral amplitudes, are in a narrower distance range when they are placed on the dipping plane. This smaller spread would thus occur for nearly any model of the attenuation, and is not an artifact of our particular choice. Because it so badly underestimates the numbers of small events, for the diffuse source, model A is strongly depleted in high frequencies relative to the other three models. The differences of spectral amplitudes between models B, C, and D for the diffuse source are nearly a factor of three at long periods, but they diminish at higher frequencies. These large differences emphasize that the estimate of uniform risk spectra are, in general, strongly dependent on the description of seismicity. Although the surface distribution in this case is not necessarily realistic, differences of this type could occur in other portions of the world where the activity is all shallow. Perhaps in situations like this, where the philosophy of

the seismicity descriptions are not flawed so far as current knowledge can show, caution would advise the selection of the model which gives the larger amplitudes of the uniform risk spectra.

Finally, we note that the levels in Fig. 2.15 are considerably lower than the spectra which correspond to the largest earthquakes allowed in the regions. This is because of the large area in which these events might occur, leading to generally large distances between the source and the site, and thus to low levels of uniform risk spectra. As mentioned previously, use of the closest distance to extended rupture zones would have increased the spectral amplitudes somewhat, but the spectral amplitudes would still have been lower than those resulting from the largest earthquakes allowed in the region. Notice also that the uniform risk spectra have shapes which differ considerably from that for a single $M = 5\frac{1}{2}$ earthquake at the site, even though the amplitudes are somewhat comparable.

CONCLUSIONS

We have investigated the empirical model for scaling Fourier amplitude spectra of strong ground motion given by Trifunac (1976) to find the probability that an arbitrary earthquake will cause the spectral amplitude $S(\omega)$ to be exceeded at the site. This relationship, together with a description of seismicity, is sufficient to evaluate the uniform risk spectra at any point in California. By replacing the attenuation function for California with one applying

to another region, we might also use this method to evaluate the risk in other regions.

To investigate some characteristics of the model, we have numerically evaluated the risk at several sites using some idealized models of the seismicity. The conceptual difference between the risk model used here and methods commonly applied elsewhere is that the spectral amplitude for a given risk of exceedance is determined independently in each of several frequency bands; thus the spectra are referred to as "uniform risk" spectra. This relaxes the constraint of a constant shape to the spectrum, and our results show that the shape and level adjust according to changes in the seismicity patterns. For example, the possibility of a large event close to the station leads to a uniform risk spectrum identical to the spectrum predicted by the Trifunac model for an event of that size and distance. However, when that large event is at a greater distance from the site, and when low level activity is possible any place near the site, the uniform risk spectrum reflects this low level activity by increasing, sometimes strongly, the amplitudes at high frequencies compared to the spectrum of the single large event, while increasing spectral amplitudes at long periods to a lesser extent. We found that even in the case of aftershocks on a line source, the uniform risk spectral amplitudes are increased slightly.

Normally, the seismicity which is input to our model is treated as the mean of a Poisson sequence. However, in the important case of an earthquake prediction, it may be necessary to find the risk for a site when the number and magnitude of events is taken

literally. We considered the differences between this "literal" assumption and the Poisson assumption for the two cases of a single event occurring on a known shallow fault and in a large diffuse zone. The differences between the two assumptions for the event on the fault are small but significant, for the event in the diffuse zone the differences are quite large.

Some of the confusion centered around the problem of how to measure the distance from the earthquake to the site can be alleviated by remembering that the object of a risk analysis is to estimate a physically well defined function. However, at present the decision to use either the epicentral distance or the nearest point on an extended rupture must still depend on mathematical consistency, a critical assessment of the attenuation function used in the risk calculations, and judgement. Our attenuation function, derived from Trifunac (1976) is based almost entirely on magnitudes less than $6\frac{1}{2}$, and uses the epicentral distance. It is not systematically different from relationships based on the nearest point of rupture, although elementary considerations suggest the two forms should differ. Therefore, we have used it with the nearest fault distance for line sources. For a diffuse source zone with a maximum magnitude of $6\frac{1}{2}$, it may be mathematically and logically consistent to use the epicentral distance with the Trifunac scaling relationships rather than nearest distance; the difference between the two methods is as much as a factor of 2, however. Use of the epicentral distance in this case may underestimate the amplitude of a uniform risk spectrum; use of the extended fault, on the other hand, may

overestimate the amplitudes. For magnitudes 7 and greater in a diffuse zone, the assessment of the attenuation function and judgement indicate the extended fault must be used.

We have also examined a number of cases based on the actual distribution of seismicity found in the vicinity of Puerto Rico. These cases suggest that the greatest source of uncertainty in any description of uniform risk spectra is related to the characterization of seismicity. Four descriptions of the seismicity which are all more or less consistent with the historic activity lead to differences of up to a factor of 3 in the amplitudes of uniform risk spectra. However, as pointed out by Allen (1976), the regions of the world where the longest seismic history is known give the greatest reason to doubt that extrapolation from past seismicity is reliable. Thus, the uncertainty associated with the characterization of seismicity is probably even greater than what our example suggests. Unfortunately, this uncertainty is present in any method of arriving at descriptions of possible ground motion to be used for earthquake resistant design; the calculations in this report only illustrate the uncertainty clearly. Thus significant improvements in seismic risk estimates will have to come from improved seismicity estimates.

APPENDIX 2.1

We want to find the mean (μ) and the standard deviation (σ) of a Gaussian distribution such that

$$P_a = \int_{-\infty}^{P_\ell} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad (1)$$

We treat P_ℓ as the independent variable, and P_a as the dependent variable (with possible noise) in the following procedure. Substituting $Z = \left(\frac{x-\mu}{\sigma}\right)$, results in

$$P_a = \int_{-\infty}^{\left(\frac{P_\ell - \mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} dZ \quad . \qquad (2)$$

For an integral of the form

$$P_a = \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} dZ \quad , \qquad (3)$$

Abramowitz and Stegun (1964, p. 932) give a polynomial approximation to s given P_a. Given the value of s for each P_a, we find by least squares the line $s = aP_\ell + b$. By comparing (2) and (3), $\sigma = \frac{1}{a}$ and $\mu = -\frac{b}{a}$.

ACKNOWLEDGMENTS

We are indebted to J. E. Luco, A. Der-Kiureghian and H. L. Wong for numerous suggestions and critical review of the manuscript. Their constructive criticism lead to numerous additions and improvements in the final manuscript.

This research was initiated under a United States Geological Survey Contract and with the support from the Earthquake Research Affiliates program, while the authors were at the California Institute of Technology. The work was completed under a Nuclear Regulatory Commission Contract and with support from the National Science Foundation at the University of Southern California.

REFERENCES

Abramowitz, M. and I. A. Stegun (1965). A Handbook of Mathematical Functions, Dover Publications, Inc., New York, 1046 pp.

Algermissen, S. T. and D. M. Perkins (1976). A probabilistic estimate of maximum acceleration in rock in the contiguous United States, United States Department of the Interior Geological Survey, Open File Report 76-416.

Allen, C. R. (1976). Geological criteria for evaluating seismicity, in Seismic Risk and Engineering Decisions, C. Lomnitz and E. Rosenblueth, Editors, Elsevier Scientific Publishing Company, Amsterdam.

Brown, R. D. and J. G. Vedder (1967). Surface tectonic fractures along the San Andreas Fault, in the Parkfield-Cholame, California Earthquakes of June-August 1966, U. S. Geol. Surv. Profess. Paper 579, 2-22.

Brune, J. (1970). "Tectonic stress and the spectra of seismic shear waves from earthquakes," J. Geophys. Res. 75, 4997-5009.

Burdick, L. J. and G. R. Mellman (1976). Inversion of the body waves from the Borrego Mountain earthquake to the source mechanism, Bull. Seism. Soc. Am. 66, 1485-1499.

Burridge, R. and L. Knopoff (1964). Body force equivalents for seismic dislocations, Bull. Seism. Soc. Am. 54, 1875-1888.

Cornell, C. A. (1968). Engineering seismic risk analysis, Bull. Seism. Soc. Amer. 58, 1583-1606.

Dalal, J. G. (1973). Probabilistic seismic exposure and structure risk evaluation, Ph. D. Thesis, Stanford University.

DeCapua, N. J. and S. C. Liu (1974). Statistical analysis of seismic environment in New York State, Fifth Symposium on Earthquake Engineering, Roorkee, India, 389-396.

Der-Kiureghian, A. (1977). Analysis of uncertainties in seismic risk evaluation (in preparation).

Der-Kiureghian, A. and A. H-S. Ang (1975). A line source model for seismic risk analysis, University of Illinois, Urbana.

Der-Kiureghian, A. and A. H-S. Ang (1977). Risk consistent earthquake response spectra, Proc. Sixth World Conference on Earthquake Engineering Research, in press.

Douglas, B.M. and A. Ryall (1975). Return periods for rock acceleration in western Nevada, Bull. Seism. Soc. Am. 65, 1599-1611.

Gardner, J.K. and L. Knopoff (1974). Is the sequence of earthquakes in Southern California Poissonian?, Bull. Seism. Soc. Am. 64, 1363-1367.

Geller, R.J. (1976). Scaling relations for earthquake source parameters and magnitudes, Bull. Seism. Soc. Am. 66, 1501-1523.

Gumbel, E.J. (1958). Statistics of extremes, Columbia University Press, New York, 375 pp.

Hileman, J.A., C.R. Allen and J.M. Nordquist (1973). Seismicity of the Southern California region, 1 January 1932 to 31 December 1972, California Institute of Technology, Pasadena, California.

Hoel, P.G. (1971). Introduction to mathematical statistics, Fourth edition, John Wiley & Sons, Inc., New York, 409 pp.

Kelleher, J., L. Sykes and J. Oliver (1973). Possible criteria for predicting earthquake locations and their application to major plate boundaries of the Pacific and the Caribbean, J. Geophys. Res. 78, 2547-2585.

Liu, S.C. and L.W. Fagel (1972). Earthquake environment for physical design: a statistical analysis, The Bell System Technical Journal 51, 1957-1982.

Lomnitz, C. (1974). Global tectonics and earthquake risk, Elsevier Scientific Publishing Company, Amsterdam, 320 pp.

McGuire, R.K. (1974). Seismic structural response risk analysis, incorporating peak response regressions on earthquake magnitude and distance, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge.

Milne, W.G. and A.G. Davenport (1969). Distribution of earthquake risk in Canada, Bull. Seism. Soc. Am. 59, 729-754.

Thatcher, W. and T.C. Hanks (1973). Source parameters of Southern California earthquakes, J. Geophys. Res. 78, 8547-8576.

Trifunac, M.D. (1974). A three dimensional dislocation model for the San Fernando, California, earthquake of Feb. 9, 1971, Bull. Seism. Soc. Am. 64, 149-172.

Trifunac, M.D. and A.G. Brady (1976). Correlations of peak acceleration, velocity and displacement with earthquake magnitude, distance and site conditions, Earthquake Engineering and Structural Dynamics 4, 455-471.

Trifunac, M.D. (1976). Preliminary empirical model for scaling Fourier amplitude spectra of strong ground motion in terms of earthquake magnitude, source to station distance, and recording site conditions, Bull. Seism. Soc. Am. 66, 1343-1373.

Trifunac, M.D. (1977). Forecasting the spectral amplitudes of strong earthquake ground motion, Sixth World Conference on Earthquake Engineering, New Delhi, India.

Wyss, M. and J.N. Brune (1968). Seismic moment, stress drop, and source dimensions for earthquakes in the California-Nevada region, J. Geophys. Res. 73, 4681-4694.

Yegulalp, T.M. and J.T. Kuo (1974). Statistical prediction of the occurrence of maximum magnitude earthquakes, Bull. Seism. Soc. Am. 64, 393-414.