NUREG/CR-4145

Earthquake Recurrence Intervals at Nuclear Power Plants

Analysis and Ranking

Prepared by J. A. Hileman, L. Knopoff, N. R. Mann, R. K. McGuire

Earth Technology Corporation

Prepared for
U.S. Nuclear Regulatory Commission

NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 1717 H Street, N.W.
   Washington, DC 20555

2. The NRC/GPO Sales Program, U.S. Nuclear Regulatory Commission,
   Washington, DC 20555

3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the NRC/GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Division of Technical Information and Document Control, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

GPO Printed copy price: $5.50

NUREG/CR-4145
RA

Earthquake Recurrence Intervals at Nuclear Power Plants

Analysis and Ranking

Manuscript Completed: September 1984
Date Published: March 1985

Prepared by
J. A. Hileman, L. Knopoff, N. R. Mann, R. K. McGuire

Earth Technology Corporation
3777 Long Beach Boulevard
Long Beach, CA 90807

Prepared for
Division of Radiation Programs and Earth Sciences
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN B7226
Under Contract No. NRC-04-81-167

ABSTRACT

Five methods for estimating earthquake recurrence were ranked. The methods represent those used, or proposed, in nuclear power plant studies through 1982 and include Log Linear Poisson, Extreme Value, Semi-Markov, Bayesian, and Uniform Hazard Method. Ranking focused on recurrence estimates for earthquake sources, excluding attenuation and site response. Scores were assigned to each method for 12 criteria such as accuracy, use of geologic data, and subjective judgment. Criteria scores were weighted by their importance and summed. Different scoring and weighting schemes were used to identify any sensitivities. To aid in scoring statistical criteria, methods were tested on synthetic earthquake catalogs with known statistics, and natural catalogs were tested against theoretical magnitude distributions.

The Uniform Hazard Method scored high because, in principal, expert judgment draws upon all seismologic knowledge. The Bayesian Method scored low because data requirements are severe for practical cases. The other methods were intermediate. These observations seem insensitive to scorer, scoring approach, or weighting scheme. The semi-Markov Method scores were sensitive to the weighting scheme. The ranking is traceable, so other scoring and weighting can be used.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

SUMMARY

This report describes Task 3, Analysis and Ranking, for the project Earthquake Recurrence Intervals at Nuclear Power Plants. The project objective is to provide the most appropriate method for estimating earthquake recurrence intervals at nuclear power plant sites. The earlier tasks, (1) literature survey of recurrence methods and (2) developing the ranking methodology, are reviewed here to provide a context for the ranking results. The ranking steps are shown in detail to provide traceability through the methodology and to enable alternative choices for ranking criteria, criteria scoring, or criteria weighting.

Project goals identify concern for rare earthquakes with return periods on the order of 10^3, 10^4, or 10^5 years. Annual probabilities may be a preferable concept for such events. For return periods longer than about 10 years, the annual probability for a Poisson process is very nearly the reciprocal of the return period in years. The probability that a particular site will experience strong ground motion depends on probability distributions for earthquake occurrence at seismic sources, for strong motion generation from a given magnitude shock, for path attenuation, and for site-specific response. If the upper tails of each distribution are combined, events with very low annual probability, say 10^{-5}, can be defined. This study focuses on earthquake recurrence methodologies at the seismic sources, so methodology testing was limited to earthquake magnitudes that equal or approach typical maximum earthquake estimates regardless of return periods.

Task 1 studies led to the selection of five methodologies to be ranked. Throughout the ranking studies, these methodologies are identified using the following terms:

- Log Linear Poisson Model. Earthquake occurrence in time follows a Poisson distribution; magnitudes follow the usual Gutenberg-Richter log-linear distribution. Parameters for the magnitude distribution can be estimated using the cumulative distribution, the differential distribution, or maximum likelihood techniques.

- Extreme Value Statistics. Only the largest earthquake in each year (or other time period) is required. The formulation is based on the Poisson and log linear distributions.

- Semi-Markov Model. The approach by Patwardhan et al (1980) is used. The probability for an earthquake with a particular magnitude depends on the magnitudes of the most recent earthquakes and the times since they occurred.

- Bayesian Method. The approach by Mortgat and Shah (1978) is used. A prior distribution for earthquake occurrence (based on site-analogous data or expert opinion) is multiplied by a sample distribution (based on site-related data) to produce a posterior distribution.

- Uniform Hazard Method. From the methodology described by Bernreuter (1980), the steps for earthquake recurrence at seismic sources are used. Seismicity parameters, such as those for the Poisson Log Linear Model, are solicited from a panel of experts.

The ranking methodology uses twelve criteria that consider technical performance, theoretical bases, data requirements, convenience, and acceptability. Each methodology is scored on a 1-to-7 scale for each criterion. The criteria have been assigned weighting factors according to their relative importance. Accuracy is given a greater weight than convenience, for example. The criteria scores are weighted and summed to obtain the final scores which are relative only; the absolute values have no particular significance. Several ranking criteria relate to statistical capabilities. To support the ranking, some methodologies were tested against synthetic earthquake catalogs having known statistical properties. Real seismicity catalogs were also tested for goodness-of-fit to statistical distributions assumed by the methodologies.

The results to date are the following:

- The Uniform Hazard Method consistently shows higher scores than the other methods.

- The Bayesian Method consistently shows lower scores than the other methods.

- The Poisson Log-Linear methods and the Semi-Markov Method are grouped mid-way in the ranking.

- The above three observations appear to be insensitive to the choice of scorer, scoring approach, or weighting scheme.

- Within the Poisson Log-Linear methods, the Cumulative Least Squares and Differential Least Squares generally seem to come out a little better than Maximum Likelihood and Extreme Value, though perhaps not significantly so.

- The Semi-Markov Method shows variable results, and appears sensitive to the weighting scheme applied.

Scores for the Uniform Hazard Method (earthquake recurrence steps) are strongly related to the criteria chosen in the ranking methodology. The criteria for generality, completeness accommodation, use of geologic data, use of geologic theory, agreement with best knowledge, robustness, and use of subjective judgment all can be expected to rate high scores for any method that can use a great deal of subjective, expert opinion. The potential for expert panels probably exceeds the performance of a specific panel, but scoring here has not been adjusted for any imagined performance factor.

In the Bayesian Method, the need for sample data at the large magnitudes for rare events is a critical requirement. Lack of such data invalidates the method. The equations of Mortgat and Shah (1978) do provide an output if the sample data are zero, but the output is just the unaltered prior distribution and not really a Bayesian result.

The Poisson Log-Linear methods cluster in mid-range. The strongest criteria scores were for convenience, generality, and public acceptance. The weakest criteria scores were for use of geologic data/theory. Cumulative and differential least squares tend to have slightly better scores than Maximum Likelihood and Extreme Value. Maximum Likelihood seemed to give the best accuracy and uncertainty results in the synthetic catalog studies by Knopoff. The Extreme Value method also seems to invoke extreme scores, very high for convenience, completeness sensitivity and robustness; and very low for use of geologic data/theory and use of subjective judgment.

The Semi-Markov Method also placed mid-range in the ranking, but shows more variable scores than the other methods. Different weighting schemes for the criteria affected the ranking position. Many criteria for which the Semi-Markov Method scored well are related to correspondence between the methodology's model and the presumed actual earthquake processes. The Semi-Markov Method is the only method ranked that has a model that may reflect some causal aspects of the earthquake process.

For one study testing recurrence methods on synthetic earthquake catalogs, three types of catalogs were generated. One catalog used the log-linear relation for magnitudes and Poisson statistics for occurrence times. A second catalog was derived from the first by adding aftershocks. the third catalog was based on a Semi-Markov process in which strain is added to a model at a constant rate, and released by random earthquakes. Both Eastern and Western US earthquake statistics were modeled. These catalogs with known statistics were used to test recurrence methods based on differential least squares, cumulative least squares, maximum likelihood, and extreme values. The results from the maximum likelihood method were generally preferred over those from other methods. The extreme value results were the least reliable for these catalogs.

Natural earthquake processes are more complex than the simple statistical models, and seem to include some memory and spatial/temporal causality. In a second model study, the recurrence methods were tested against synthetic catalogs whose generating process is at least one level more complex than the processes implied by the recurrence estimation methods. Synthetic catalog generation used a model with five separate, but interdependent, seismic sources. Strain is added to the system, and a source's cumulative energy release depends on both the current state and states of neighboring sources. None of the recurrence methods gave consistently good estimates for the long-term catalog properties. The methods did give good statistical fits to the short-term, or local, statistics. This difference between short-term and long-term statistics may be representative of some areas having only low to moderate seismicity.

Three natural earthquake catalogs having reasonably good seismicity data were analyzed to determine if there was a best statistical distribution type describing earthquake recurrence. Extreme value techniques were used to avoid aftershock problems. The Gumbel Type III distribution was the most appropriate model. Clearly, a maximum earthquake should be used in fitting seismicity data. Distribution parameters can be estimated and then used to predict magnitudes for various return periods. The results were consistent with the catalogs for moderate magnitudes. For long return periods, the estimated magnitudes tended to be too large.

1. INTRODUCTION

The project "Earthquake Recurrence Intervals at Nuclear Power Plants" has comprised three tasks to date. Work on these tasks was conducted from May, 1981 through June, 1983. Task 1 included a literature review surveying the various recurrence methods, either in use or proposed, and included developing a methodology for ranking. In Task 2, the methods to be ranked were chosen. Both the ranking methodology and the earthquake recurrence methods to be ranked were developed with concurrence of the Nuclear Regulatory Commission staff. The overall project objective is to provide the most appropriate methodology for estimating earthquake recurrence intervals at nuclear power plant sites. The methodology finally adopted could include a system of methods, each applied under various site conditions.

The Task 1 and Task 2 results were described in task reports to the Nuclear Regulatory Commission (Earth Technology, 1981; 1982). This report describes the Task 3 results on ranking the selected methods. Some portions of the prior task reports are incorporated herein so this document can better stand alone. The ranking studies were carried out by The Earth Technology Corporation and employed Leon Knopoff (University of California at Los Angeles), Robin McGuire (Dames and Moore Consultants, and formerly with The Earth Technology Corporation), and Nancy Mann (University of California at Los Angeles) as consultants. Sections 5.1 and 5.3 are essentially verbatum reports from Leon Knopoff and Nancy Mann, respectively. Section 5.2 was prepared at The Earth Technology Corporation and describes the studies directed by Robin McGuire. He also participated in the criteria scoring.

The studies described in this report were conducted to provide a relative ranking for five selected earthquake recurrence methodologies. The methodologies are:

1) Log Linear Poisson Model,
2) Extreme Value Statistics,
3) Semi-Markov Model,
4) Bayesian Method, and
5) Uniform Hazard Method

The above terms are used to identify the various earthquake recurrence methodologies, but certain conventions and limitations should be kept in mind. The Log Linear Poisson Model includes three closely related techniques: cumulative recurrence curve statistics, differential recurrence curve statistics, and maximum likelihood statistics. The Semi-Markov Model is the technique described by Patwardhan et al (1980), although other similar semi-Markov approaches have been described. The Bayesian Method is the technique described by Mortgat and Shah (1978), although Bayes rule can be used in many ways. The Uniform Hazard Method in its entirety develops a site-specific uniform hazard spectrum. However, the ranking studies here consider only the earthquake recurrence steps in the methodology, and "Expert Opinion Method" might be a more accurate term.

As this report was in preparation, a sixth methodology, the Full Historic Method (Veneziano et al, 1984), was nominated for ranking along with the five methodologies listed above. Though not reported herein, the Full Historic Method should be considered for future extensions of this work.

Traceability and full disclosure of the ranking steps is an important aspect of the ranking method. The bulk of this report is intended to show just how the steps were executed and the results obtained. Alternative choices for certain criteria, criteria scoring, or criteria weighting can be propagated through the steps with minimal effort.

Each earthquake recurrence methodology to be ranked is described in Section 3 so the methods, as we have considered them, will be clearly defined. Some methods have been used widely and are defined by conventional practice more than in some particular document. Other methods did not have a concise description appropriate for this report.

Section 4 describes the ranking per se. The ranking approach was developed in the Task 1 (Earth Technology Corporation, 1981). Twelve criteria were chosen such as convenience, accuracy, use of expert judgment, etc. For each criterion, each earthquake recurrence methodology was compared to the others and assigned scores on a scale of 1-to-7. In some cases, separate scores were assigned for a method's application to Eastern U.S. seismicity data and to Western U.S. seismicity data. The scoring is given in Section 4.2 so complete traceability of the ranking choices is present. Scoring for several statistics-dependent criteria is partially based on extensive studies using synthetic earthquake catalogs having known statistical properties.

The synthetic catalog studies to support ranking are reported in Section 5 and include the following:

1) Leon Knopoff, University of California at Los Angeles. Synthetic catalogs based on Poisson-log linear statistics and catalogs based on stochastic release of strain accumulation were used to test the Poisson-Log Linear Methods and Extreme Value Statistics.

2) Robin McGuire, Earth Technology, and Dames and Moore. Synthetic catalogs whose generating process is meant to be more complex than any of the methodology's statistical models were used to test the Poisson-Log Linear Methods, Extreme Value Statistics, and the Bayesian Model.

3) Nancy Mann, University of California at Los Angeles. Observed seismicity catalogs were tested against several statistical distribution models to determine best fits.

The ranking is meant to be unbiased by using the following sequence: select the criteria and their weighting, select the recurrence methodologies to be ranked, score all methodologies for each criterion at the same time, and finally compute the weighted sums. Two independent scorers, two ways of using the 1-to-7 scale, and two criteria weighting schemes were used to check for any strong scoring bias or numerical sensitivities in the scheme. The scorers were in agreement, more or less.

Probably other scorers could be found to disagree on some point or other. Traceability in the procedure permits the effects of any such disagreement to be determined. No particular sensitivities were noted except as discussed in Section 4 for the Semi-Markov Model and weighting, and the Uniform Hazard Method and criteria selection.

2. RARE EVENT CONSIDERATIONS

This section discusses the role of rare earthquakes, those having return periods up to 500,000 years, in the ranking studies. Justification is given for limiting the recurrence methodology testing to earthquake magnitudes that either equal or approach typical maximum earthquake estimates irrespective of the return periods. Some of the ranking criteria are not closely related to the return periods.

The project goals identify particular concern for rare earthquakes with return periods on the order of 10^3, 10^4, or 10^5 years. Geologists assign about the last 1.1×10^4 years to Holocene time, and the last 2×10^6 years to Quaternary time. Both Holocene and Quaternary faults are usually given careful scrutiny during geotechnical investigations for nuclear power plant sites. NRC guidelines (10 CFR Part 100, Appendix A) indicate consideration must be given to faults that have moved more than once in the last 5×10^5 years. For earthquakes that might recur with intervals of 10^4 or 10^5 years, the geologic data are usually inadequate to show that such earthquakes do in fact have characteristic return periods.

Conceptually, annual probabilities provide the best way to describe rare earthquakes. For return periods longer than about 10 years, the annual probability for a Poisson process is very nearly the reciprocal of the return period in years. The probability of at least one earthquake occurrence is $p(1,2,3,...) = 1 - p(0)$. For a Poisson process with rate λ per year,

$$p(0) = (\lambda)^0 e^{-\lambda}/0! = e^{-\lambda}.$$

then

$$1-e^{-\lambda} \approx \lambda \text{ for } \lambda << 1.$$

Using annual probabilities avoids the implication of event recurrence and permits consideratio of rare events that may, or may not, occur within the current tectonic environment for a site region.

An important distinction must be made between probabilities estimated from observed earthquake histories, probabilities estimated by fitting some statistical model, and "real" probabilities (though unknown) that would best describe the actual physics of an earthquake source. Good statistical fits have been obtained for the earthquake magnitude distributions where the data are abundant at low and moderate magnitudes (Earth Technology, 1981). However, the goodness-of-fit at large magnitudes (long return periods or low annual probabilities) beyond the observed data is not demonstrated. There must be a departure between data and the fitted statistical model if a maximum earthquake is not considered. Even using a maximum earthquake, the fit is uncertain because the maximum earthquake itself has some uncertainty. Continued development of statistical analyses to determine the best possible distribution and estimate its fitting parameters is needed for rare earthquake statistics.

The earthquake recurrence methodologies selected for ranking, except the Semi-Markov Model, use some analytic statistical distribution function(s) to fit the observed data. The usual log-linear magnitude distribution underlies the Extreme Value Statistics and the Uniform Hazard Method as well as the Poisson-Log Linear Methods. For the Extreme Value Statistics, Bayesian Model, and Poisson-Log Linear Methods, synthetic catalogs were used to test the methodologies. In these tests, the assumption is made that the relative performance of the methods at moderate to large magnitudes will be indicative of their relative performance for rare events with very low annual probabilities. If a method produces a statistical fit that does poorly at moderate to large magnitudes, the performance is not expected to improve for magnitudes with extremely low annual probabilities. By chance, the apparent performance could improve at very low annual probabilities if the true magnitude distribution differed significantly from the statistical fit. However, many data sets from the synthetic seismicity catalogs were analyzed to ascertain the average performance of the methods tested (see Sections 5.1 and 5.2).

If a maximum earthquake is included in the analysis, considering very low annual probabilities may not significantly affect the maximum ground motion to be expected from a seismic source. For example, a hypothetical site in California affected by the San Andreas fault might be characterized as being exposed to a magnitude 8-1/4 or greater earthquake with an annual probability of 4×10^{-3} (250-year return period), and a maximum earthquake with magnitude 8.5. Considering annual probabilities of 10^{-4} or 10^{-5} will not greatly increase the magnitude and strong ground motion. For many eastern sites, the appropriate maximum magnitudes are subject to more speculation. The probabilities implied by the experts as they chose maximum earthquakes for eastern sites in the Uniform Hazard Method were on the order of 10^{-3} according to D. Bernreuter (ACRS Subcommittee Meeting, Santa Monica, Oct. 1982). He stated that the experts made little distinction between maximum earthquakes with 10^{-3} probability and any others with smaller probability. L. Reiter, in the same meeting, concluded that the experts were implying annual probabilities on the order of 2×10^{-4} to 10^{-4}. In any case, going to annual probabilities lower than those associated with the maximum earthquake was not intended by the experts to imply larger magnitudes.

The current tectonic environment for a site region, its neotectonics, is disclosed by Quaternary tectonic and geomorphic features, historical seismicity, seismological data from instrumental observations, and various geophysical measurements. These data can contribute to estimating some maximum earthquake as a characteristic parameter for the current tectonic regime. The maximum earthquake may have an uncertainty or a distribution of values. Often in seismic design studies, a conservative value (higher than the best estimate) is required. Certain earthquakes still could be postulated having magnitudes greater than the chosen maximum earthquake and very low annual probabilities. The clear implication is that such earthquakes would be caused by tectonic forces that differ from those characterizing the current regime. All the earthquake recurrence methodologies to be ranked in this study are based on geological data, seismicity data, and maximum earthquakes that describe the current tectonic environment in a site region. Therefore, testing the methods for magnitudes at or approaching the maximum earthquakes is considered appropriate for ranking purposes, even if the associated annual probabilities are greater than 10^{-4} to 10^{-5}.

This ranking study focuses on methodologies for estimating earthquake recurrence at the seismic sources. Specifically excluded are the effects of attenuation as seismic energy is propagated from its source to a nuclear power plant site, and local effects from the site geology and soil column. Attenuation and site effects are also probabilistic, but they can be incorporated equally with each recurrence methodology tested. (Note: The full Uniform Hazard Method does include attenuation and site effects, but only the earthquake source portion was compared to the other methods.) Earthquake recurrence as experienced at some particular site from earthquakes in the surrounding region is the recurrence needed for seismic design and risk decisions. This site-recurrence distribution combines the following:

1) The probabilistic description for magnitude occurrence at the various sources ("recurrence" as used throughout the rest of this report),

2) A probabilistic description of strong ground motion generation as a function of magnitude and possibly other source parameters,

3) Path attenuation effects which may have uncertainty, and

4) Site effects which may have uncertainty.

Because the site-recurrence distribution is the product of several probability distributions, it can take on very low annual probabilities, 10^{-4} or 10^{-5}, for maximum ground motion experienced at the site. Such low values follow when the tails of each distribution are used. The ranking in this study is to identify the best methodology for earthquake recurrence at the seismic sources, the first step in the site-recurrence problem. For this step, the annual probabilities for maximal events do not necessarily extend to the low values of 10^{-4} or 10^{-5}. Further studies are needed to address the best methodologies for estimating the distributions for strong motion generation, attenuation, and site effects.

3. METHODOLOGY DESCRIPTIONS

Descriptions in this section define the methodologies as we have considered them for ranking. This is done because there can be variations in the methodologies, and questions could arise concerning our "standardization." Where differences are deemed to cause possible variations in the scores assigned to criteria, the variations and their effect on scoring are discussed in the scoring text. These variations are also considered again in the section on sensitivity studies of the ranking process. The descriptions are given in more detail in the Task 1 report, Literature Review (Earth Technology, 1981).

3.1 Selection of Methods to be Ranked

In the literature on estimating recurrence intervals, very few formal methodologies are defined. Rather, geological and statistical methods have been applied on a more or less ad hoc basis. One notable exception is the Uniform Hazard Method developed in the Site Specific Response Spectra program.

The analysis and ranking in this project focus on the underlying methods that are most likely to find application in a recurrence interval methodology. Because maximum earthquakes can be expected generally to have return periods much longer than the time period represented by most data sets, there is an emphasis on evaluating statistical methods. As pointed out in the iterature review, there are often variations in methods are they are used or presented by different investigators. Representative methods are ranked, and consideration is given to the significance of any variations. Literature review showed the recurrence methods could be grouped broadly as Poisson Models, Memory Models, Bayesian Statistics, and Geologic Methods.

The Poisson Models category includes Log-Linear models, Quadratic models, and Extreme Value Theory. The Log-Linear models have been widely used in recurrence estimates, and have a high priority for ranking. Even if other methods or models are found superior, the relative ranking of Log-Linear models is important for comparison. The Quadratic models seem to be of much less importance, and they have not been used to any great extent. If the Log-Linear approach were judged best of the methods ranked, then the Quadratic models could be further considered as a possibility for improved estimates. Extreme Value theory has a moderately high priority for ranking because the seismicity data for many site areas is so amenable to this type of analysis.

The Memory Models category includes Markov models, Semi-Markov models, Weibull models, and Clustering models. The Markov process (having a one-step memory) can be thought of as a special case of the Semi-Markov process (having a finite-length memory). The Semi-Markov process is given a high priority for ranking, because it can reflect some aspects of physical models proposed for earthquake occurrence. The Semi-Markov process is also important because it permits consideratle variation in both temporal and spatial relationships as expressed in the transition matrices. The Wiebull models are given a low priority because they are probably best considered as a variant of the Poisson models. Clustering models seem to have appealing physical bases and perhaps they should be ranked, however the Semi-Markov models were judged more important.

In the Bayesian Statistics category, the Uniform Hazard Method was indicated by the NRC for ranking. The Uniforn Hazard Method is a Bayesian method in the sense that experts are in effect acting as Bayesian processors by combining data with their expert opinion. A fully Bayesian approach as described by Mortgat and Shaw (1978) also has high priority for ranking because this method allows formally weighted input of both expert opinion and observational data.

In reviewing the Geological Methods, a clear picture emerges that there are a large number of geological techniques, each having unique requirements and applicability. In practice, geological methods are almost always used to supplement or corroborate other estimates of earthquake recurrence intervals at nuclear power plant sites. Definitive recurrence estimates using geological methods have been done for individual seismic source structures only. We are not aware of any power plant site for which a geological method, or methods, was used exclusively. It seems best to consider the Geological Methods as a necessary part of any methodology that is to be employed, and not to attempt independent ranking.

In light of the above considerations, the following methods were selected for ranking and analysis.

1. Log-Linear Poisson Model with consideration of the presence or absence of an upper limit to the magnitude.

2. Extreme Value Statistics for both the Type I distribution (unbounded magnitudes) and Type III distribution (magnitudes bounded by an upper limit).

3. Semi-Markov Model, which can include a Markov Model as a limiting case.

4. Bayesian Method as described by Mortgat and Shaw (1978).

5. Uniform Hazard Method.

3.2 Log-Linear Model

3.2.1 Theoretical Basis

The Log-Linear Model presumes an exponential distribution of earthquake magnitudes according to an empirical relationship proposed by Gutenberg and Richter and having the form

$$\text{Log } N_C = a - bM$$

where N_C is the number of earthquakes having magnitude M or greater in some time interval, a and b are constants that characterize a particular relation. A differential form of the equation

$$\text{Log } N_I = a' - bM$$

describes the number of earthquakes N_I in a magnitude range ΔM that contains the magnitude M. The slope b is the same for both forms.

Because a time interval is specified, the equations are often called recurrence relations for magnitudes. The actual times of occurrence for the earthquakes are assumed in this methodology to follow a Poisson process. In a Poisson process, the probability of n earthquakes in an interval of time dt is given by

$$P\ (n,dt) = \frac{(\lambda)^n\ e^{-\lambda dt}}{n!}$$

and λ is the average rate of occurrence per unit time dt. The probability that the interoccurrence time lies in the range from t to t+dt follows the exponential distribution (Lomnitz, 1966).

$$p(t) = \lambda\ e^{-\lambda t}\ dt$$

Both of these distributions provide only emperical fits to observed seismicity catalogs. The earthquake generation process is not yet well enough understood to derive magnitude and occurrence distributions from first principals of earthquake processes.

Theoretical arguments and observational experience suggest that the magnitude distribution must have an upper bound. The methodology is applied both with and without an upper bound. Both cases are considered here.

3.2.2 Data Required

Because the magnitude and interoccurrence time distributions cannot now be derived from physical principals of the earthquake generation process, the distribution constants must be estimated from seismicity data. The data must represent the geographical area under consideration, so earthquake times, locations and magnitudes are required.

In the Log-Linear Model methodology, the constants a and b could be found, at least conceptually, using data from only a limited magnitude range and a limited time period. However, the magnitude data would have to be exceptionally precise, and the time period would have to be truly representative of long term observations. In practice, the methodology is used with confidence only when:

1) The seismicity catalogs are considered to report all earthquakes above some threshold magnitude M_o that have occurred during the reporting period.

2) The range of magnitudes from M_o to the maximum magnitude observed is at least several magnitude units so that the slope b of the recurrence curves is reasonably well constrained.

3) The magnitude data do in fact fit the log-linear distribution reasonably.

4) The observed rates of seismicity seem consistent with the local geology and tectonics. The earthquake occurrences during the period of observation should be representative of the long term process; i.e., the earthquake process is stationary in a statistical sense.

In some applications, not all requirements may be fully met. The methodology can still be applied with caution, but care must be exercised to avoid undue confidence in the resulting recurrence estimates.

3.2.3 Methodology Steps

Three generalized steps are necessary: 1) data collection and evaluation, 2) parameter estimation, and 3) recurrence period estimation.

3.2.3.1 Data Collection and Evaluation

This task is included here as the first task because the user must be aware of the quality and characteristics of the data. Seismicity catalogs have been derived from many different seismographic networks operating in various tectonic environments with various degrees of geographic coverage, and for various purposes. The data often are not uniform over extended time periods. Critical evaluation is needed to estimate the threshold magnitude M_o, above which a catalog reports all earthquakes that have occurred during the observation period. Different M_o values may apply to different time periods such as $M_o = 3.0$ for the past 10 years and $M_o = 5.0$ for the past 80 years.

The earthquake process is presumed to be Poisson, and the earthquake occurrences in time as independent of each other. Therefore, aftershocks should be identified and removed from the data. The term aftershock implies a causal relation between the aftershock and its mainshock without identifying the relation. A clear definition of aftershocks, serving to identify them in the record, has not been made. Identifying and removing aftershocks is somewhat arbitrary at best (Knopoff et al, 1983). Often, an arbitrary temporal and spatial window relative to the mainshock is assumed; all shocks within the window are discarded as aftershocks. A better system is needed.

Consideration should be given also to the magnitude scale because some catalogs may not be homogeneous in their magnitude scale use. Small shocks may be M_L or M_b, while larger shocks may be M_L or M_S.

For nuclear power plant sites, special studies may be done to extend the data base as far back in time as possible and to evaluate the data quality.

3.2.3.2 Parameter Evaluation

The parameters to be estimated are the constants $\underline{a}$ and $\underline{b}$ in the recurrence relation, and sometimes the maximum magnitude. Usually, maximum magnitude is estimated separately using geologic and tectonic data.

The parameters $\underline{a}$ and $\underline{b}$ may be estimated by graphically fitting a straight line, $\log N = a - bM$, to data points for log N versus M. In a graphical fit, subjective allowance can be made for the threshold magnitude M_o and for individual data points that may appear anamalously high or low. Sometimes the data may be poor and suggest that analytical fitting techniques may not be warranted except to give a formal estimate of the parameter uncertainties.

In dire circumstances, a reasonable value for the slope $\underline{b}$ may be assumed and a recurrence curve passed through some particular data point that evokes the best confidence. For example, in a region with poorly recorded seismic history, the small earthquakes may be judged to be under-reported and large earthquakes too infrequent to be properly represented statistically. But the record for moderate shocks, say magnitude 4 to 5, may be complete because of reported felt effects. Clearly, this approach leads to large uncertainty in the eventual estimates of return periods for rare events.

The straight line may be determined analytically with well known regression formulas derived from a least squares criterion (Benjamin and Cornell, 1970). For this recurrence period methodology, care must be taken that log N is regressed on M, rather than the opposite, so that a recurrence rate for the desired earthquake magnitudes can be correctly obtained (Bolt, 1978). Only data for M_o and larger magnitudes are used. This method also provides formal estimates of the standard deviations of the parameter estimates. The analytical fit is not quite the same conceptually as the graphical fit, because the graphical fit is attempting to fit data points rather than regress one variable on the other.

The straight line parameters a and b for the cumulative form may be obtained analytically also using maximum likelihood expressions (Weichert, 1980) as follows:

$$b = 1/(\ln 10(\overline{M}-M_o))$$

where M_o is the threshold magnitude and $\overline{M}$ is the average magnitude,

$$\overline{M} = 1/n \sum_{i=1}^{n} M_i.$$

a is found from b and the total number of events above the threshold M_o.

3.2.3.3 Return Periods

The final step is to extrapolate the straight line Log N = a-bM beyond the largest observed magnitude to the desired rare event magnitudes. Return periods are taken as simply the inverse of the annual recurrence rate and uncertainty in the return period is estimated from uncertainties in a and b.

3.2.4 Uncertainties

Uncertainties in the estimated return periods arise from several sources. The methodology depends solely on statistical data describing earthquake recurrence. An implied assumption is that the data are stationary and truly portray the long term behavior. Uncertainty arises here for nearly all geographic regions except ones for which earthquakes reasonably near the maximum earthquake have been observed repeatedly. The amount of this uncertainty is not determined analytically and can only be estimated subjectively, perhaps by comparison with long term geologic data. All the methodologies are subject to this uncertainty.

An estimate of uncertainty can be obtained in graphically fitting a straight line to the log N-versus-M data points by trying to bound the acceptable range of fits.

The analytic methods, least squares and maximum likelihood, give formal estimates of uncertainty. It must be remembered that these formal estimates are only a measure of how well the data points fit the assumed log-linear relationhip. If the earthquake process differs from the assumed distributions, the estimated uncertainties may not apply.

If the data truly fit a log-linear recurrence relation, the least squares and maximum likelihood results would be identical. A criticism of the maximum likelihood estimate is that undue weight is given to the small-magnitude earthquakes as they strongly control the value of $\bar{M}$, the average magnitude. The least-squares method is similarly criticized for giving equal weight to all data points. The large-magnitude data points are not well determined because they may come from only one or two earthquakes, whereas smaller magnitude data points may be much better determined.

3.2.5 Regional or Tectonic Limitations

The Log-Linear Method has no regional limitations other than that the region, or fault structure, to which the recurrence statistics apply must provide reasonably good statistics. If the region chosen is too small, the statistics may be sparse or non-stationary for the time interval used. In general, the region of concern around a nuclear power plant site is divided into zones of similar geology and tectonics. A separate recurrence curve is constructed for each zone. Separate maximum earthquakes are used, and their return periods estimated.

3.2.6 Similarities/Affinities to Other Methodologies

The Extreme Value Statistics Method is also based on the exponential distribution of magnitudes and the Poisson process for interoccurrence times. However, the statistics are handled differently.

The Semi-Markov Method can assume an exponential distribution for earthquake magnitudes.

The Uniform Hazard Method can use earthquake recurrence curves, identical to those of the Log-Linear Model, as a way of describing the seismicity.

The Bayesian Method uses a Poisson distribution for earthquake occurrence.

3.2.7 Prior Applications

Earthquake recurrence curves are used in many safety analysis reports to describe seismicity. However, a formal methodology for return periods of rare events has not been applied. There are numerous applications in the seismological and geological literature that estimate return periods for large earthquakes using recurrence curves.

3.3 Extreme Value Statistics

3.3.1 Theoretical Basis

Extreme value methods generally assume earthquake magnitudes are distributed according to the Gutenberg-Richter exponential relationship, and times of occurrence follow a Poisson process. Then, for earthquake recurrence applications, the statistic used is the largest earthquake annually, or some other convenient time period. Other distributions could be used. This method has the same basic assumptions as the Log-Linear Model, but the statistics are handled differently.

Epstein and Lomnitz (1966) combined the magnitude and interoccurrence time distributions to show the largest annual earthquake M is distributed with a cumulative distribution G(M) where

$$G(m) = P(M<m) = \exp(-\alpha e^{-\beta m})$$

where m is the the distribution variable, α and β are related to a and b of the log-linear distribution by $\ln\alpha = a \ln 10$ and $\beta = b \ln 10$. Taking natural logarithms of the equation gives

$$-\ln(-\ln G(M)) = \ln \alpha - \beta M$$

in which the ln-ln term, left hand side, has a linear relation to magnitude M. The parameters α and β can be found after fitting a straight line to probability-versus-maximum magnitude points plotted according to the above equation.

The derivation can be extended to give the mean return period T between earthquakes with magnitude M or greater as

$$T = \exp(\beta M)/\alpha.$$

Return periods can also be found directly from the annual probabilities G(M) extrapolated to the desired magnitudes along the fitted straight line. The parameters α and β can be converted to a and b in the log-linear equation and that equation applied as in the Log-Linear Model methodology. Finally, Knopoff and Kagan (1977) use the form

$$T = 10^{(M-M_o)b}$$

where M_o is the magnitude for which $-\ln(-\ln G(M)) = 0$ (the magnitude with annual probability 1/e, or the average annual magnitude) after finding b from β.

3.3.2 Data Required

The main attraction of this method is its very modest data requirements. Only the largest earthquake in each successive time period is required. Maximum annual earthquakes are generally used, but the theory is equally amenable to shorter or longer time periods.

An important requirement is that the seismicity data actually fit the assumed exponential distribution for magnitudes and the Poisson distribution for interoccurrence times. If the distribution assumptions are met and enough time periods are sampled, the method results will converge with those from the Log-Linear Method. Because only maximum annual earthquakes are used, the total time span with reliable statistics may be considerably longer than that for many other recurrence methodologies. The longer time span may provide a more stationary representation of seismicity.

3.3.3 Methodology Steps

Three generalized steps are necessary: 1) data collection and evaluation, 2) parameter estimation, and 3) recurrence period estimation.

3.3.3.1 Data Collection and Evaluation

Collecting and evaluating the seismicity data is included as the first step because the methodology user must be aware of the quality and characteristics of the data. The considerations for extreme value statistics are similar to those for the Log-Linear Model, but less demanding because only maximum earthquake magnitudes are used. However, some knowledge of the distribution characteristics for smaller shocks is needed to judge if the methodology assumptions are reasonable.

Aftershocks are not generally considered. However, an aftershock of a large earthquake may also be the largest earthquake in a succeeding time interval. Then, the aftershock should be removed and the largest independent earthquake used. Identification of aftershocks is no different here than for the Log-Linear Method, except only the large aftershocks generally need to be considered.

3.3.3.2 Parameter Estimation

The parameters to be estimated are α, β and perhaps M_o. If b is required, it is derived from β. These parameters are obtained from the straight line

$$-\ln(-\ln G(M)) = \ln \alpha - \beta M$$

fitted to a probability plot of the maximum annual magnitudes. First the maximum annual magnitudes for N years are ranked monotonically according to magnitude with the largest magnitude first, R = 1, and the smallest last, R = N. Then a cumulative probability is assigned to each rank R according to a plotting rule such as

$$G(M) = \frac{R - 1/2}{N}$$

Other plotting rules can be used (Knopoff and Kagan, 1977). The points M and $-\ln(-\ln G(M))$ are fit with a straight line, either analytically or graphically, which fixes the parameters $\ln \alpha$, β and M_o. If $\ln \alpha$ and β are to be converted to a and b in the log-linear equation, $\log N = a - bM$, then the straight line should regress M onto $-\ln(-\ln G(M))$. If the annual probabilities G(M) are to be used, the straight line should regress $-\ln(-\ln G(M))$ onto M.

Formal uncertainties in the straight line fitting regression can be converted to uncertainties in the parameters α, β, M_o, a and b as needed. However, the plotted data points should also be examined to judge their linearity, and thereby the distribution assumptions.

Maximum magnitude is usually assumed to be unbounded, or at least larger than any of the observed annual extrema. The analysis can be modified to fit a line incorporating the effects of an assumed maximum magnitude (see Section 5.1). If the time span covered by the data includes several occurrences of near maximum earthquakes, a maximum earthquake may also be estimated.

3.3.3.3 Return Periods

Return periods for large earthquakes can be estimated by several methods using the straight line parameters. As noted in Section 2.2.3.1, the fitted line gives the annual probabilities G(M) very simply from the -ln(-ln G(M)) value for the desired magnitude. A direct estimate is also obtained from the equation

$$T = \exp(\beta M)/\alpha.$$

Alternatively, the estimated parameters ln α and β may be converted to a and b, and the equation

$$\text{Log } N = a - bM$$

used as in the Log-Linear Method. The form

$$T = 10^{(M-M_o)b}$$

is equivalent, and can be applied directly. The formal uncertainties in ln α and β can be transformed to uncertainties in the estimated return periods.

3.3.4 Uncertainties

This methodology is subject to the same sources of uncertainty as the Log-Linear Model: the data stationarity and the validity of the assumed statistical distributions. Analytical methods for curve fitting provide formal estimates of the variance in the estimated parameters ln α and β. However, these uncertainty estimates describe only the linearity of the data points and do not necessarily reflect the validity of the methodology assumptions.

3.3.5 Regional or Tectonic Limitations

Limitations for this methodology are the same as those described for the Log-Linear Method.

3.3.6 Similarities/Affinities to Other Methodologies

This methodology is closely related to the Log-Linear Method in that both methods are based on an exponential distribution for earthquake magnitude and a Poisson distribution for earthquake interoccurrence times.

3.3.7 Prior Applications

No formal use of the Extreme Value Statistics methodology has been incorporated into safety analysis reports for nuclear power plant sites. However, there are several applications in the seismological literature (Earth Technology, 1981).

3.4 Semi-Markov Model

3.4.1 Theoretical Basis

Markov and Semi-Markov models are a broad class of statistical models with the general property of having some finite memory of their past history. The true Markov process has a one-step (one unit of time) memory; the Semi-Markov process can have a multi-step, but finite, memory. For a Semi-Markov earthquake process, the probabilities that earthquakes with various magnitudes will occur in the next time step depend on the earthquakes occurring during the current time step, or perhaps back to the most recent large earthquake. This general process can be modeled with various choices for process variables and distributions.

The process form modeled by Patwardhan et al (1980), is adopted here as the Semi-Markov methodology to be ranked. In the current literature, their model seems to be the best application of the Markov process to the occurrence of damaging earthquakes. They estimate earthquake probability rather than return periods, and there is an important distinction. Return periods are a property of an earthquake generation process. They can be expressed as a constant annual probability, and do not depend on the current state of the process. However, return periods can be combined with information on the time since the last earthquake to estimate earthquake probabilities. Patwardhan et al (1980), estimate earthquake probabilities directly using the most recent great earthquake time and magnitude among the process variables.

Basic elements of the model are the state description, the probabilities for changing from one state to another, and the probabilities that the state will hold unchanged for various times. The model has been developed for great earthquakes, $M \geq 7.8$, around the Circum-Pacific Belt. Earthquake magnitudes are descritized into N intervals and the system state is defined by the most recent great earthquake magnitude i, i = 1 to N. When another great earthquake occurs, the state changes to j according to the new magnitude. The probability that the system will change from state i to state j is P_{ij}, the transition probability. There is a matrix of P_{ij} for all combinations of the indices i and j. The time that the system will hold in state i before changing to state j is an integer random number of time steps with a distribution h_{ij}.

Because the system can change to any of the possible states (each earthquake magnitude is possible), the waiting time between state changes (earthquakes) is a distribution W_i obtained from the h_{ij}, $j = 1$ to N.

Past seismicity is used to estimate the various distributions. These distributions are modified for the initial step in the process because some waiting time since the last great earthquake has already elapsed. The model is allowed to run to predict the number and magnitudes of earthquakes for some desired period of years.

3.4.2 Data Required

The methodology requires probability distributions for the occurrence of magnitude j earthquakes given that the previous earthquake was magnitude i. Probability distributions are also needed for the interoccurrence times for earthquakes with magnitudes i and j. When $i = j$, this is just the return period for earthquakes with magnitude i, and the methodology could become a circular argument for estimating return periods. However, geological data and subjective judgment (prior distribution) are combined with the seismicity data (likelihood function) using a Bayesian procedure to obtain the needed distributions (posterior distribution). Therefore, this methodology is very closely akin to the Bayesian methodologies.

The number of variables to be estimated increases with the square of the number of possible system states, $n = N^2 + N$. Even for only 3 system states as in the Patwardhan et al (1980) example, there are 12 variables required for the analysis. The required variables can quickly overwhelm the available data base in detailed applications.

The basic data required are seismicity data for the system states that are assigned to the model. The data quality and completeness for these states should be known with the same care as for the Log-Linear Model. However, the Bayesian use of subjective judgment and geological data can ease somewhat the seismicity data requirements. A wide range of data can be used through the Bayesian approach, but the method presents no specific requirements on this data.

3.4.3 Methodology Steps

The methodology comprises 6 steps: 1) data collection and evaluation, 2) defining states and time intervals for the Markov process, 3) defining initial conditions, 4) estimating model parameters, 5) calculating model predictions, and 6) converting results to return periods.

3.4.3.1 Data Collection and Evaluation

This task is indicated because the analyst should be familiar with the data quality and its limitations. General considerations are discussed in Section 2.1.3.1 for the Log-Linear Model. For the Semi-Markov model, only the data relating to the discrete system states are required. There may be less total data available, and they assume relatively greater importance because many parameters are to be estimated. Assessment of data quality becomes extremely important.

Subjective judgment and geological data can be incorporated into some parameter estimates using a Bayesian approach. Care must be taken here if the model is to be kept close to the physical system under consideration. The method requires reasonably good data if the results are to reflect anything other than the influence of subjective judgment.

3.4.3.2 Defining System States and Time Intervals

The states of the system are a set of discrete magnitude intervals. The magnitude range and intervals will depend on the tectonic environment under consideration. The time interval should be short enough that there is negligible probability of two earthquakes, state transitions, occurring within a single time interval. Lesser shocks could occur but they are ignored. The interval should be as long as the above criterion will permit so that computations are minimized.

3.4.3.3 Defining Initial Conditions

The initial conditions are the magnitude and elapsed time since the last earthquake in the magnitude range of the system states. These are self-evident if there has been such an earthquake historically. In other cases, estimates must be made using geologic data. If the problem is divided into several geologic zones, each with its own Markov process, initial conditions and process parameters are needed for each zone.

3.4.3.4 Estimating Model Parameters

The required model parameters are the transition probabilities P_{ij} and the holding times h_{ij}. These are determined from the seismicity data base if the data are adequate. In the Patwardhan et al (1980) example, and probably for most potential applications, the data do not provide good statistics for rare earthquakes. Therefore, a Bayesian approach is used to estimate the required parameters (see Section 3.4). The observed seismicity is used for the likelihood functions and subjective judgment, augmented by geologic and tectonic data, are used to compose the conjugate prior distributions.

3.4.3.5 Calculating Model Predictions

The Semi-Markov model used here describes how an earthquake generating system can transition from state i to another state j. The time holding in a state is a random variable depending on the states i and j. A time period is selected, such as the 50-year design life of a project. During the 50-year period, there may be 0, 1, or several earthquakes in the Semi-Markov process, and each earthquake may be in any of the allowed magnitude intervals. If there are 3 system states, allowed magnitude ranges, and we use the notation (a, b, c) to describe the number of earthquake occurrences in each magnitude range during the 50-year period, typical histories could be (0,0,0), (1,0,0), (0,1,0), (2,1,0), etc., independent of the exact magnitude sequence. The probability that a particular history will occur is the joint probability for a transitions to state 1, b transitions to state 2, and c transitions to state 3. Patwardhan et al (1980) give equations to compute the joint probabilities recursively.

These results apply only to the specified time period rather than to general time periods with the same length. The initial conditions imposed, time and magnitude for the most recent great earthquake, cause the probabilities to be specific to the time period used.

Combining the joint probabilities can give the probability that an earthquake with magnitude M or greater will occur during the time period considered, or the probability of other earthquake senarios.

3.4.3.6 Converting Results into Return Periods

The method applied by Patwardhan et al (1980) does not estimate return periods for the earthquake process. The results are probabilities specific for the particular time period considered. Therefore, it would be incorrect to convert the time-period probabilities into return periods. If the initial conditions were circumvented by ignoring them, and the time period were taken as 1 year to obtain annual probabilities, the method would just degenerate into a Bayesian approach: see Sections 3.3.2.

3.4.4 Uncertainties

The method described by Patwardhan et al (1980) does not give a formal estimate of uncertainties in the estimated probabilities. They conducted a parametric analysis showing that altering holding times by a 1.5 factor changed the computed probabilities by 1.3 to 4.6 factors. The factors for probability change were strongly dependent on the time period length and the earthquake magnitude considered. The absolute uncertainties depend greatly on the data quality as in the Log-Linear Model, and also on the quality of the subjective input.

3.4.5 Regional or Tectonic Limitations

The method requires observational data on each allowed system state to form the needed likelihood functions. Therefore, the method can be applied only for areas in which earthquakes with magnitudes up to the maximum magnitude of concern have been observed. The Patwardhan et al (1980) example uses great earthquakes in the Circum-Pacific Belt. Except for highly seismic regions with historic occurrences of near maximum earthquakes, most regions do not have adequate statistics for the Semi-Markov method. The method is still quite useful though to evaluate the consequences of various assumed seismicity models.

3.4.6 Similarities/Affinities to Other Methods

As noted in Section 2.3.2, the methodology described by Patwardhan et al (1980) is strongly dependent on Bayesian statistics. Conceptually, Bayes Rule is not required for a Semi-Markov process, but limited data can lead to a Bayes Rule application.

3.4.7 Prior Applications

There have been no formal applications to nuclear power plant sites. The paper by Patwardhan et al (1980) is the most prominent application.

3.5 Bayesian Model (Mortgat and Shah, 1978)

3.5.1 Theoretical Basis

Bayesian models or Bayesian statistics are founded on Bayes Rule, which is broadly applicable to many statistical problems. Perhaps the major attraction of the Bayesian models is that they allow data from different, even disparate, sources to be rationally combined. Thus, long term geologic data or historical records of earthquakes can be combined with modern instrumental observations. "Expert opinion" can be combined with observational data, and each weighted according to its perceived importance.

Bayes Rule can be expressed as

$$P[B_i \mid A] = \frac{P[A \mid B_i] \; . \; P[B_i]}{P[A]}$$

where

$$P[A] = \sum_{i=1}^{n} P[A \mid B_i] \cdot P[B_i].$$

$P[B_i \mid A]$ is the conditional probability for event B_i given that event A has occurred. The variable n is the number of different states B_i can assume. Benjamin and Cornell (1970) show that Bayes Rule can be equivalently written for a variable λ as

$$f''(\lambda) = N_1 \; f'(\lambda) \; L(\lambda)$$

where

$f''(\lambda)$ is the posterior distribution of λ

$f'(\lambda)$ is the prior distribution of λ

$L(\lambda)$ is the likelihood function of λ

N_1 is a normalizing constant.

The prior distribution represents a starting point that may be "expert opinion" or a particular class of data. The likelihood function represents obsevations (a sample function), or perhaps a second class of data. For convenience, the distribution types for the prior and likelihood functions are generally chosen from well known distributions that are mathematically orthogonal, so the product distribution will also be a well known distribution. Distribution types that can fit a large variety of shapes are used so no strong limitations are imposed on modeling the data.

Mortgat and Shah (1978) use one application of Bayes rule to develop a posterior distribution for earthquake occurrences independent of magnitude. They use a second application for a posterior distribution on the probability p_i that r_i earthquakes will occur in a given magnitude range M_i, such as $6.5 \leq M < 7.0$, given that n earthquakes have occurred. These two distributions are combined to give posterior distributions for r earthquakes in the range M_i (their equations 2.21 and 2.23), or for zero earthquakes in the range M_i (their equations 2.22 and 2.24).

Earthquake Occurrence. Earthquake occurrences, independent of magnitude are assumed to follow a Poisson process. A Poisson distribution for the occurrence rate λ is taken as the likelihood function

$$L(\lambda \mid N,T) = \frac{(\lambda T)^N e^{-\lambda T}}{N!} \quad .$$

A gamma function form is chosen to describe the prior distribution for the rate λ, because the gamma distribution is orthogonal to the Poisson distribution and can model a variety of shapes.

$$f'(\lambda) = \frac{e^{-l\lambda}(l\lambda)^{k-1}\, l}{\Gamma(k)}$$

where the parameters l and k control the distribution shape. The product $f''(\lambda)$ of these prior and likelihood functions is also a gamma function because of their orthogonality. Finally, the posterior distribution for n earthquake occurrences is obtained by combining the Poisson occurrence model for n-given-λ with the posterior distribution on λ and removing the conditional aspect by itegrating over all possible λ.

$$p(n) = \int_0^\infty p(n \mid \lambda) \quad f''(\lambda)\ d\lambda$$

which results in the Mortgat and Shah (1978) equations 2.9 and 2.10.

Magnitude Distribution. Earthquake magnitudes are descretized into a number of magnitude bins M_i, where each bin may contain magnitudes ranging over perhaps 1/2 magnitude unit. The development treats each magnitude bin M_i separately. A Bernoulli process is assumed, i.e. independent trials with only two possible outcomes. When an earthquake occurs, its magnitude is either in bin M_i with probability p_i, or not in bin M_i with probability $1-p_i$. This generating process is a binomial process and the likelihood function for observing the probability p_i is given by

$$L(p_i \mid N_1, r_i) = p_i^{r_i} (1-p_i)^{N-r_i}$$

where N is the total number of earthquakes and r_i is the number with magnitude in bin M_i. A beta function is chosen as the form for the prior distribution because it is orthogonal to the binomial distribution and is defined on the interval 0 to 1 as probabilities are. The beta function chosen by Mortgat and Shah (1978) has the form

$$f'(p_i) = \frac{p_i^{(a_i-1)} \quad (1-p_i)^{(b_i-a_i-1)}}{B(a_i,\ b_i)}$$

where

$$B(a_i,\ b_i) = \frac{\Gamma(a_i)\ \Gamma(b_i)}{\Gamma(a_i + b_i)} .$$

The parameters a_i and b_i control the distribution shape. The product $f''(p_i)$ of these prior and likelihood functions is also a beta function because of their orthogonality. Finally, the posterior distribution for r_i, the number of earthquakes in magnitude bin M_i given that n earthquakes have occurred, is obtained by combining the binomial law for occurrences r_i with the posterior distribution on the probability p_i. The condition on p_i is removed by integrating over all possible probabilities.

$$p(r_i \mid n) = \int_0^1 p(r_i \mid n,\ p_i)\ f''(p_i)\ dp_i$$

which results in the Mortgat and Shaw (1978) equation 2.20.

Magnitude Occurrence. The final step is to combine the results of the two previous paragraphs to remove the condition on n earthquakes being given. The condition is removed by summing over all possible n

$$p(r_i) = \sum_{n=o}^{\infty} p(r_i \mid n)\ p(n)$$

which results in the Mortgat and Shah (1978) equations 2.21 through 2.24 which include cases for $r_i = 0$ and when equation parameters take on integer values. Thus, using prior data (or expert opinion) and sample data, the probability for earthquakes with a particular magnitude range can be estimated. Earthquake return periods can then be taken as the reciprocals of the annual probabilities when the probabilities are small.

3.5.2 Data Required

One of the outstanding features of this method is the extremely broad range of permissible data. Mortgat and Shaw (1978) formulated their equations for the prior distributions in terms such as either number of years for the prior data or "the equivalent time period over which the analyst bases his subjective input." Thus, many kinds of data are amenable if they can be translated into an equivalent number of earthquakes for an equivalent time period.

In the final equations, seven input parameters are required for each magnitude bin M_i, see Figure 5.2-B.

- For the sample, or likelihood, functions representing the observed data.

  N = The total number of earthquakes observed above some threshold M_o.

  R_{M_i} = The number of earthquakes in the magnitude bin M_i.

  T = The sample time, duration of the observed data.

- For the prior distribution on λ, the earthquake occurrence rate.

  λ' = The equivalent time used by the expert on which to base his estimate of λ. λ' enters the equations through the relation $\lambda'' = \lambda' + T$.

  ν' = The equivalent number of earthquake occurrences (all shocks above the threshold M_o) during the period λ'. ν' enters the equations through the relation $\nu'' = \nu' + N$.

- For the prior functions on p_i, the probability of an earthquake in magnitude bin M_i given an earthquake has occurred.

  η' = The equivalent number of trials used by the expert; the total number of earthquakes above the threshold magnitude M_o. η' enters the equations through the relation $\eta'' = \eta' + N$.

  ξ' = The equivalent number of successes (r_i, the earthquakes in magnitude bin M_i) during the η' trials. ξ' enters the equations through the relation $\xi'' = \xi' + R_{M_i}$.

If the same data are used to estimate the prior distributions for both the rate λ and the probabilities ρ_i, then $\eta' = \xi'$ and $\eta'' = \xi''$. Input parameters must be provided for each magnitude bin M_i for which statistical estimates of return periods are desired. The method does not extrapolate to high magnitudes for which there are no observed data.

Relative weighting between the prior distribution on the rate λ and the sample distribution on λ is effected by the ratio between the "equivalent time period" λ' used by the expert and the sample time T. Relative weighting for the prior distribution on the probability p_i and the sample distribution on p_i is effected by the ratio between the "equivalent number of events" n' used by the expert and the observed events N. If only the prior data or the sample data are present, the equations place full weighting on the available data.

3.5.3 Methodology Steps

The methodology includes four steps: 1) data collection and evaluation of the sample data, 2) either data collection and evaluation of the prior distributions, or solicitation and evaluation of expert opinion, 3) parameter estimation, and 4) converting results into return periods.

3.5.3.1 Sample Data Collection and Evaluation

This task is indicated because the analyst should be familiar with the data quality and its limitations. General considerations are discussed in Section 3.1.3.1 for the Log Linear Model. Data collection for the Bayesian approach incudes here only the sample data, which are presumed to be reliable catalogs that are complete for all earthquakes above the chosen threshold magnitude M_o. Different time periods may be used for different magnitude ranges, so the data will include large earthquakes for which statistical estimates are desired.

Alternatively, the Bayesian approach can be used to combine and weight two classes of data. The sample data may simply be one data class.

3.5.3.2 Prior Data Collection and Evaluation, or Expert Opinion

If an actual data set is to be used to represent the prior data, the considerations are the same as those for the sample data. In addition, the choice must be made for the relative weights to be assigned to the prior data and the sample data. One simple choice weights each data set by the time period corresponding to the data set. The Mortgat and Shah (1978) equations accommodate this weighting automatically. If some other weighting is desired, perhaps according to the analyst's confidence in the data sets, one or the other data set should be scaled (numbers of events and time periods). The time periods used for the prior and sample data sets should be in the ratio of the desired weighting.

If expert opinion is to be used to construct the prior data set, the expert's estimates of rates or probabilities must be translated into the equation parameters "equivalent earthquakes" for "equivalent time" periods. Careful solicitation of expert opinion is itself an extensive topic beyond the scope of this discussion.

3.5.3.3 Parameter Estimation

The parameters estimated by the Mortgat and Shah (1978) equations are $p(r_i)$ or $p(r_i=o)$ where r_i is the number of earthquakes in a magnitude bin M_i. Often the calculation will be for the probability of zero earthquakes, $p(r_i=o)$. The probability of at least one earthquake with magnitude M_i is $1-p(r_i=o)$. This calculation must be done for each desired magnitude bin M_i. The time period chosen for the probability calculations is taken as 1 year, and annual probabilities are obtained. Probabilities could also be obtained for a facility's expected design lifetime.

If the seismic hazard at a particular site is modeled as arising from more than one source area, the annual probabilities are obtained for each contributing source area and summed.

3.5.3.4 Return Periods

The return periods are taken as just the reciprocals of the annual probabilities.

3.5.4 Uncertainties

The method as described by Mortgat and Shah (1978) indicates formal variances for the posterior distribution $f''(\lambda)$, the earthquake occurrence rate, and the posterior distribution $f''(p_{M_i})$, magnitude probability. The undertainty estimates are not carried forward into the final posterior distributions $p(n)$ in Step 1, $p(r_i \mid n)$ in Step 2 or $p(r_i)$ in Step 3.

Because the method is so broadly applicable in the way the prior and sample distributions may be derived and combined, there is no generalized formulation for uncertainty estimates. In some cases, relative uncertainties in the prior data (or expert opinion) and/or the sample data will determine how the analyst chooses to weight the prior and sample distributions. In other cases, the weighting may be simply in proportion to the time periods represented by the two distributions. The method doesn't include a means to relate these uncertainties to that of the final estimates.

Other sources of uncertainty include the data quality as in the Log Linear Model (Section 3.1.4), and the quality of the subjective judgment used as input. The method assumes that earthquake occurrence follows a Poisson process so shocks are time independent. No assumptions are made about the magnitude distribution; each magnitude bin M_i is treated separately.

3.5.5 Regional or Tectonic Limitations

The method requires data for prior and sample distributions for each seismic source to be considered. This requirement may be met in regions with high seismicity where earthquakes as large as design earthquakes have occurred to provide reasonable statistics. For regions with moderate or low seismicity in which the design earthquakes have not occurred, the method does not really extrapolate to the rare events. If there is no sample data for a given large-magnitude bin, the method simply gives an estimate based on the prior distribution. Such estimates are not Bayesian because no sample distribution is used. Data requirements place a strong limitation on the method when probability estimates for rare earthquakes are desired.

3.5.6 Similarities/Affinities to Other Methods

The Semi-Markov Model described by Patwardhan et al (1980) uses a Bayesian procedure to estimate conditional probability distributions on earthquake magnitude (see Section 3.3.2). The Uniform Hazard Method (Bernreuter, 1980) includes a mathematical model for earthquake occurrence that follows the Mortgat and Shah (1978) formulation exactly. However, the Uniform Hazard Method solicitation of expert opinion requested the experts to incorporate sample data in their judgments, so the uniform hazard spectrum is considered a subjective probability estimate.

The Bayesian model assumes a Poisson distribution for earthquake occurrence as do the other methods.

3.5.7 Prior Applications

The Bayesian model has not been applied formally to nuclear power plant sites. Mortgat and Shah (1978) illustrate their methodology with an application to Nicaragua.

3.6 Uniform Hazard Method

3.6.1 Theoretical Basis

The Uniform Hazard Method (UHM) described by Smith et al (1981) and Bernreuter (1980) is founded philosophically on the Bayesian approach of Mortgat and Shah (1978). The Mortgat and Shah mathematical development is reproduced in early draft reports on the methodology. However, in its final form, the Uniform Hazard Method is described as a probabilistic model which systematically incorporates subjective judgment into the evaluation of seismic hazard (Bernreuter, 1980). Bayesian analysis is not used formally, but experts are asked to act as "Bayesian processors" in formulating their subjective judgment. The experts' answers are considered in Bayesian terminology as posterior estimates.

The method does not estimate earthquake return periods. The method's objective is to develop a uniform hazard spectrum, a spectrum of peak ground motion (acceleration, velocity, or displacement) versus frequency, whose ordinates each have the same probability of being exceeded in a given number of years. Such a spectrum is not limited to the predicted effects of any one design earthquake. High frequencies in the uniform hazard spectrum may be controlled by the potential occurrence of moderate-sized nearby earthquakes, and low-frequencies by larger more distant shocks. Each frequency has the same probability. To achieve this objective, the Uniform Hazard Method has four steps:

- Identify and specify discrete seismic source zones, either as area sources or line sources. Each zone has roughly uniform seismic activity.

- Develop an earthquake occurrence model for each seismic source zone comprising a probability distribution for magnitudes and another for occurrence times.

- Develop a ground motion model that describes ground motion at the seismic source, attenuation and modification as energy propagates to the site, and modification by site effects.

- Combine the data from the first three steps to estimate the uniform hazard spectrum and uncertainties associated with its ordinates.

The Uniform Hazard Method extends beyond the other methodologies ranked in this study by including a ground motion description, attenuation, and site effects. The full methodology estimates ground motion recurrence experienced at a particular site, rather than explicitly describing occurrences at the sources. For purposes of this ranking analysis we will consider only the first two methodology steps which deal with earthquake occurrence per se.

In the first step, experts are asked to delineate seismic sources using seismicity data and their subjective judgment. Thus a broad range of geologic, tectonic, and seismological theory can be used, or ignored. The seismic sources are descritized into manageable source elements and the experts' uncertainties are quantified.

In the second step, earthquake occurrence is presumed to follow a Poisson time sequence, as in the Log Linear Model; and the rate λ is supplied by the experts. No particular magnitude distribution is implied by the methodology, but in practice the experts supplied their magnitude information using the usual log-linear relation $\log N_C = a-b_M$. In addition, a maximum magnitude, or range of magnitudes, was specified so the log-linear relation is bounded for large shocks.

The theoretical basis of Steps 3 and 4 are not included here because the ranking study does not extent to them.

3.6.2 Data Required

The Uniform Hazard Method is quite flexible in the amount and type of data required. In princple, the method could be applied using just expert opinion and no overt incorporation of geologic, tectonic, or seismological data. The experts are required only to specify parameters such as seismic source geometry, recurrence curve constants, and maximum earthquake magnitude. Because various experts may not be equally familiar with all potential source regions, a standard, comprehensive package of seismicity and tectonic data can be supplied to each expert (Bernreuter, 1980). In this way, the experts are asked to act somewhat as Bayesian processors using the data set as sample data and their opinions as prior data. Supplying the best possible data set should lead to less variation between expert opinions and less uncertainty in the final uniform hazard spectrum. For regions such as the East Coast where there is only a brief seismic history, little strong ground motion data, and uncertainty in earthquake mechanisms, the method leans heavily on expert opinion.

3.6.3 Methodology Steps

As described above, the Uniform Hazard Method comprises four steps. However for purposes of this ranking study, and in an attempt to "equalize" our perspective of each methodology, the ground motion model and the spectrum calculations are not included here. The success or failure of the Uniform Hazard Method is certainly strongly dependent on the last two steps. However, each of the other methodologies being ranked could be used to provide input for the ground motion and spectrum calculations. For this study, the Uniform Hazard Method is considered to have the following tasks: 1) collection of geologic, tectonic and seismicity data, 2) solicitation of expert opinion to delineate seismic sources and earthquake occurrence models, and 3) compilation of expert opinion and formatting for input to later methodology stages. Viewed in this manner, the portion of the methodology to be ranked is essentially using experts to act as Bayesian processors operating on sample data and subjective opinion to produce an earthquake recurrence model.

3.6.3.1 Data Collection

The first task is somewhat different here than all the other methodologies. The seismicity data should be compiled as described in Section 2.1.2.1 under the Log Linear Model. However, much of the data evaluation is the responsibility of the experts in the second task. Information about data quality and completeness is developed to be provided to the experts; the final evaluation and selection, or rejection, of various data are left to expert opinion. In addition, the experts are also provided with geologic and tectonic information that may supplement their knowledge.

3.6.3.2 Solicitation of Expert Opinion

The second task begins with determining what parameters and descriptions will be required. Questions are formulated to quantify answers and uncertainties, avoid prejudicing answers, and evaluate the internal consistency of answer sets. There may be need to iterate the question-reply cycle to address issues precisely. This same solicitation can be conducted by interview if the interviewers are trained in the procedure. Accurate solicitation and quantification of expert opinion, whether by questionnaire or interview, can have many subtleties because few persons recognize many cues that influence opinion (Spetzler and Stael von Holstein, 1975). In the TERA application (Bernreuter, 1980), experts were asked to delineate seismic source zones (perhaps with alternate interpretations and associated confidences), characterize the seismicity of each source in terms of the recurrence relation parameters $\underline{a}$ and $\underline{b}$ with uncertainties, and to estimate a maximum magnitude or range of maximum magnitudes.

3.6.3.3 Compilation

The third task is to convert each experts' responses into the parametric forms needed as input to subsequent methodology steps.

3.6.4 Uncertainties

Considerable attention is given throughout the methodology to uncertainties. The experts are asked to indicate the uncertainties in each of their estimates. They also express their degree of self confidence for each study area. Because the methodology objective is the uniform hazard spectrum, uncertainties are developed for that spectrum and not for the aggregate expert opinion on the earthquake occurrence model. Each experts' uncertainties are carried individually through the computational steps to a resultant uniform hazard spectrum estimate, one for each expert. Finally, a weighted average spectrum is calculated along with its uncertainties.

The method recognizes uncertainties in the source definitions, earthquake occurrence models, the ground motion model, and local site effects. The seismic hazard estimate results from a sequence of probability calculations, and the formal uncertainty estimate is large. Even with only the tasks defined for this ranking, the uncertainty can be large because of differing expert opinion, or artificially small if the experts are in close agreement.

Because the method as applied to East Coast sites is too dependent on assumed magnitude distributions, the TERA report (Bernreuter, 1980, p. 3) says "one should be cautioned against using seismic hazard values for rare events (e.g. events with return periods in excess of 5,000 years)."

Uncertainties are treated as random errors to simplify analysis. A clear distinction is not available between systematic errors and random errors. There may be interdependence among some uncertain quantities.

3.6.5 Regional or Tectonic Limitations

The method assumes only that earthquake occurrence can be modeled by seismic source zones, each having relatively uniform seismicity. Any lack of data can always be obviated if the expert is willing to give an opinion. As a practical matter, the experts must have a reasonable data base to have confidence in their estimates, and to inspire confidence in the methodology results. Still, the method does provide a rational way of using various expert opinions in cases where the data are inadequate.

3.6.6 Similarities/Affinites to Other Methods

The method is closely related to the Bayesian Method through its mathematical approach. However, application of Bayes Rule becomes subjective with the experts as they combine the prior data (their prior opinions) with the sample data provided to arrive at a posterior opinion.

The method, as it has been applied, has a basis similar to the Log Linear Model because expert opinion is used to specify parameters in a Poisson distribution for earthquake occurrence and a bounded log-linear magnitude distribution.

3.6.7 Prior Applications

This method was developed under NRC sponsorship and has been applied so far exclusively to nuclear power plant sites. The sites include Zion, Palisades, Big Rock Point, Dresden, La Crosse, Yankee Rowe, Oyster Creek, Connecticut Yankee, Millstone and Gina.

4. RANKING

Ranking the five selected recurrence methods was conducted according to the ranking methodology developed in Task 1. For this ranking, the Log Linear Poisson Model was subdivided into Cumulative Least Squares, Differential Least Squares, and Maximum Likelihood methods. Thus, seven separate methods were carried through the ranking process.

The ranking procedure is intended to be objective and highly visible. Obtaining the objectivity is not a trivial task because the mere choice of ranking criteria and relative weights assigned to them can reflect personal biases. But, if each of the steps and judgments can be clearly seen, then any subjectivity can more easily be identified by users or independent reviewers. Ranking criteria arise from several broad aspects of the earthquake recurrence interval estimation problem such as definition of sources, assumptions about the earthquake process, data limitations, and the perception of the method.

For this ranking study, we presume seismic sources have been delineated as they may be required for input to a recurrence method.

There are obvious criteria addressing the accuracy of recurrence estimates and their associated uncertainty. Several criteria relate to the earthquake process itself. How well does a method use current knowledge of the earthquake process? Earthquake occurrence is tabulated in seismicity catalogs, and various interpretations of those catalogs are made using proposed underlying statistical distributions. There are some insights, although not enough, into the physical process of earthquakes.

Seismological data limitations lead to another group of criteria. How well does a method deal with data that may be incomplete, inaccurate, or even non-stationary? Can all of the required data be measured, or are assumptions required?

Another group of criteria includes less technical and non-technical considerations such as acceptance by the scientific community, public credibility, cost, convenience of application, and generality. These criteria clearly are less important than those that relate to accuracy and uncertainty. However, they may serve as useful discriminators if some methods are found closely comparable on purely technical criteria.

The ranking criteria chosen and their proposed relative importances are as follows:

Criterion		Importance	Values	100% Basis
1.	Convenience	Low	1	2
2.	Generality	Average	3	5
3.	Uncertainty of Estimate	Very High	10	17
4.	Accuracy	Very High	10	17
5.	Completeness Sensitivity	Average	3	5
6.	Completeness Accommodation	High	5	8
7.	Use of Geological Data	Average	3	5
8.	Use of Geological Theory	High	5	8
9.	Agreement with Best Knowledge	Very High	10	17
10.	Robustness	Average	3	5
11.	Subjective Judgment	High	5	8
12.	Public Acceptance	Low	1	2

These criteria are described further in Section 4.2, Criteria Scoring. Not all of the criteria are completely independent. Trying to achieve truly independent criteria would tend to produce very narrow criteria that could be unwieldy to apply. Some interdependence is reasonable if it is clearly recognized. Assigning criteria scores is intended to help discriminate between recurrence methods, so the perception of average performance can be taken as a score of 4. Relative ranking of methods is the objective rather than some arbitrary scoring threshold.

Each criterion is also assigned a weighting factor relative to the other criteria. Care is taken so that the combined importance of several moderate or average criteria should not be allowed to overwhelm some critical criterion. Development of one set of weights on a numerical scale is given above, but these weights should be considered one of the primary subjects for review and possible modification.

Several criteria are related to statistical aspects of determining earthquake recurrence intervals. Statistical analysis and testing was conducted to aid in scoring the criteria that relate to the statistical properties of the methodologies. Items such as accuracy, specification of uncertainty, and sensitivity to data quality, completeness, and length were tested using synthetic catalogs having known statistical properties. In addition, an interactive computer routine was used to test how well proposed distributions fit real data catalogs.

The following sections summarize the ranking results, present the tabulated criteria scores, and discuss the individual criteria scoring in detail.

4.1 Summary of Ranking Results

The individual criteria scores and resulting ranking scores (weighted sums) are given in tabular form in Tables 4-1, 4-2, and 4-3 for three

Table 4-1

CRITERIA SCORING BY SCORER A, SPREAD SCORING

Criteria	Weights		Methodologies[1]						
	W-1	W-2	A CLS	B DLS	C ML	D EV	E SM	F B	G UHM
1. Convenience	1	1	6	6	6	7	4	4	1
2. Generality	3	3	5	5	5	4	1	3	7
3. Est. Uncertainty	10	10	4	4	4	4	1	1	3
4. Accuracy	10	10	4	4	4	4	4	1	4
5. Comp. Sens.	1	3	4	4	3	5	7	1	4
6. Comp. Accom.	1	5	5	5	4	3	7	1	4
7. Use Geol. Data	3	3	1	1	1	1	5	3	7
8. Use Geol. Theory	3	5	1	1	1	1	5	3	7
9. Best Knowledge	3	10	3(1)	4(2)	4(2)	3(1)	7	1	4
10. Robustness	3	3	6(3)	6(3)	7(3)	7(3)	1	4	4
11. Sub. Judgment	3	5	2	2	1	1	4	4	7
12. Pub. Accept.	1	1	7	6	6	3	1	1	4
Unweighted Sums			48(43)	48(43)	46(40)	43(37)	47	27	56
W-1 Weighted Sums			156(141)	158(143)	156(138)	149(131)	138	81	191
W-2 Weighted Sums			211(182)	220(191)	210(178)	196(164)	247	108	271

[1] The letters A,B,C, etc. key the methodologies to Figures 4-1, 4-2, etc. The methodologies are:
CLS, Cumulative Least Squares
DLS, Differential Least Squares
ML, Maximum Likelihood
EV, Extreme Value
SM, Semi-Markov
B, Bayesian
UHM, Uniform Hazard Method
The notation X(Y) when used indicates scores for Eastern (Western) data.

Table 4-2

CRITERIA SCORING BY SCORER B, STANDARD SCORING

Criteria	Weights		Methodologies[1]						
	W-1	W-2	A CLS	B DLS	C ML	D EV	E SM	F B	G UHM
1. Convenience	1	1	4(5)	4(5)	4(5)	6	2	3	1
2. Generality	3	3	4	3	3	5	3	2(3)	7
3. Est. Uncertainty	10	10	2(4)	2(4)	2(4)	2(4)	1	1	2
4. Accuracy	10	10	3(4)	2(4)	3(4)	3(4)	1	2(3)	3(4)
5. Comp. Sens.	1	3	3(4)	2(4)	2(4)	6	4	2	4
6. Comp. Accom.	1	5	3(5)	3(5)	3(5)	2	6	2	7
7. Use Geol. Data	3	3	2(4)	2(4)	2	2	2(5)	2(4)	4
8. Use Geol. Theory	3	5	3	3	2	2	6	4	7
9. Best Knowledge	3	10	4(3)	4(3)	3	3	4	2	4
10. Robustness	3	3	4	4	4	5	2	1	4
11. Sub. Judgment	3	5	5(4)	5(4)	4(3)	2	6	6(5)	7
12. Pub. Accept.	1	1	6	6	4	3	2	3	6
Unweighted Sums			43(50)	40(49)	36(43)	41(44)	39(42)	30(33)	56(57)
W-1 Weighted Sums			132(166)	118(163)	117(149)	124(154)	103(112)	91(107)	167(177)
W-2 Weighted Sums			194(229)	178(226)	166(208)	173(203)	187(196)	137(151)	259(269)

[1] The letters A,B,C, etc. key the methodologies to Figures 4-1, 4-2, etc. The methodologies are:
CLS, Cumulative Least Squares
DLS, Differential Least Squares
ML, Maximum Likelihood
EV, Extreme Value
SM, Semi-Markov
B, Bayesian
UHM, Uniform Hazard Method
The notation X(Y) when used indicates scores for Eastern (Western) data.

Table 4-3

CRITERIA SCORING BY SCORER B, SPREAD SCORING

Criteria	Weights		Methodologies[1]						
	W-1	W-2	A CLS	B DLS	C ML	D EV	E SM	F B	G UHM
1. Convenience	1	1	4(5)	4(5)	4(5)	7	2	3	1
2. Generality	3	3	4	4	4	5	2	1	7
3. Est. Uncertainty	10	10	3(5)	3(5)	3(5)	3(5)	1	1	3
4. Accuracy	10	10	2(4)	2(4)	2(4)	2(4)	2	1	3
5. Comp. Sens.	1	3	4	3	3	7	4	1	4
6. Comp. Accom.	1	5	3(4)	3(4)	3(4)	2	6	1	7
7. Use Geol. Data	3	3	2(3)	2(3)	2(3)	1	3(4)	4	7
8. Use Geol. Theory	3	5	3	3	2	1	4	4	7
9. Best Knowledge	3	10	4(3)	4(3)	3(2)	2	7	1	5
10. Robustness	3	3	6	6	6	7	2	1	4
11. Sub. Judgment	3	5	4(3)	4(3)	4(3)	1	6	5	7
12. Pub. Accept.	1	1	6(7)	5(6)	4(5)	3	1	2	5
Unweighted Sums			45(51)	43(49)	40(46)	41(45)	40(41)	25	60
W-1 Weighted Sums			136(176)	134(174)	127(167)	120(160)	115(118)	75	188
W-2 Weighted Sums			198(233)	194(229)	178(213)	160(200)	216(219)	106	287

[1] The letters A,B,C, etc. key the methodologies to Figures 4-1, 4-2, etc. The methodologies are:
CLS, Cumulative Least Squares
DLS, Differential Least Squares
ML, Maximum Likelihood
EV, Extreme Value
SM, Semi-Markov
B, Bayesian
UHM, Uniform Hazard Method
The notation X(Y) when used indicates scores for Eastern (Western) data.

variations in the criteria scoring and two variations in weighting. The variations are discussed under Criteria Scoring, Section 4.2, but they are mentioned briefly here. Scoring was done first on the 1-to-7 scale with pre-chosen performance standards defining low, medium and high scores. Another scoring variation was obtained by giving the best method a score of 7, the least method a score of 1, and distributing the remaining methods in the interval. The two criteria weighting variations are: the weighting developed along with the ranking methodology, and a slightly different weighting suggested by Robin McGuire. The tabulated data permit alternative weighting schemes to be implemented easily, or the effects of varying certain criteria scores to be checked.

The summed scores for the methodologies are given in Figures 4-1 and 4-2. These figures do not show any functional relationships, but merely help display the results in a convenient manner. Each methodology is coded A, B, C, etc. as shown in the tables. The ranking scores were normalized so a perfect score would be 100. For each data subset (scorer/scoring-approach/East-or-West data), three sets of weighting results are shown: the unweighted sums, sums weighted according to the proposed ranking scheme, and sums weighted according to McGuire's proposed weights.

Figures 4-1 and 4-2 quickly show some of the major results of the ranking.

- o The Uniform Hazard Method (G) consistently shows higher scores than the other methods.
- o The Bayesian Method (F) consistently shows lower scores than the other methods.
- o The Poisson Log-Linear methods (A,B,C,D) and the Semi-Markov Method (E) are grouped mid-way in the ranking.
- o The above three observations appear to be insensitive to the choice of scorer, scoring approach, or weighting scheme.
- o Within the Poisson Log-Linear methods (A,B,C,D), the Cumulative Least Squares (A) and Differential Least Squares (B) generally seem to come out a little better than Maximum Likelihood (C) and Extreme Value (D), though perhaps not significantly so.
- o The Semi-Markov Method (E) shows variable results, and appears sensitive to the weighting scheme applied.

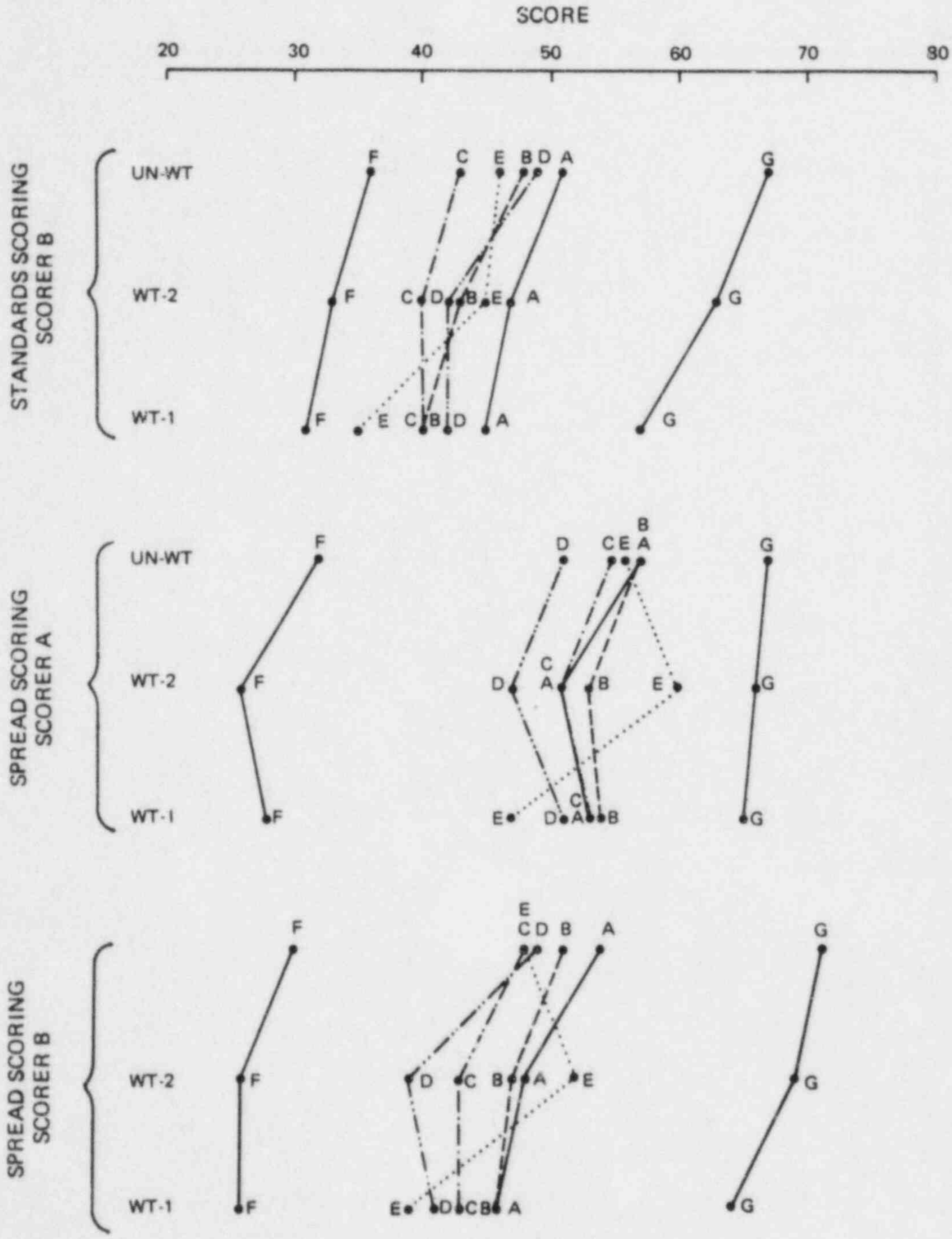


FIGURE 4-1 SCORES FOR EASTERN APPLICATIONS, NORMALIZED TO 100 FOR A PERFECT SCORE. NOTATION FOLLOWS THE FOOTNOTE TO TABLE 4-1

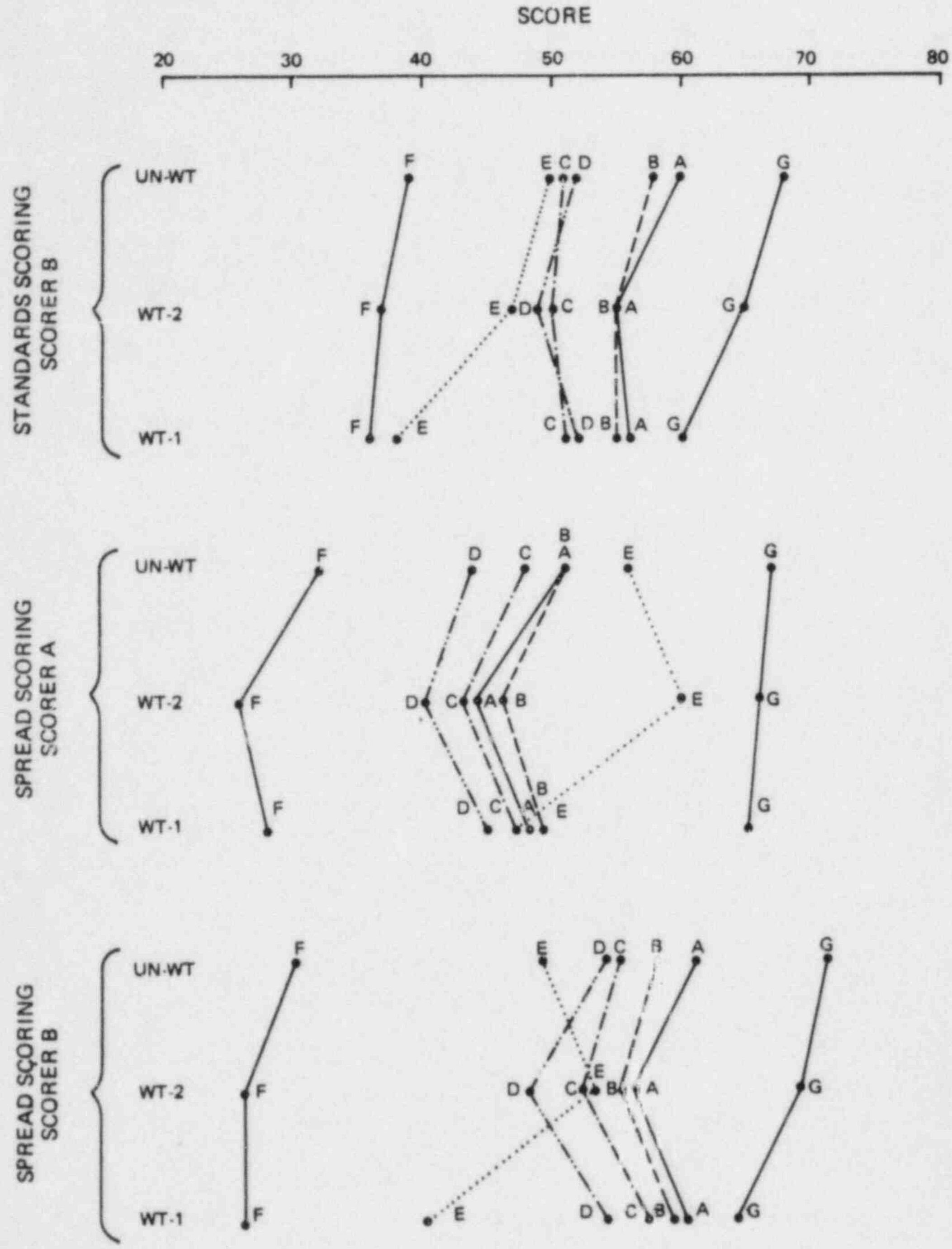


FIGURE 4-2 SCORES FOR WESTERN APPLICATIONS, NORMALIZED TO 100 FOR A PERFECT SCORE. NOTATION FOLLOWS THE FOOTNOTE TO TABLE 4-1

4.1.1 Uniform Hazard Method

The high scores for the Uniform Hazard Method were not anticipated beforehand, but they are strongly related to the criteria chosen in the ranking methodology. The criteria for generality, completeness accommodation, use of geologic data, use of geologic theory, agreement with best knowledge, robustness, and use of subjective judgement all can be expected to rate high scores for any method that can use a great deal of subjective, expert opinion. In effect, the experts are able to draw on all other methodologies, if they wish, to formulate their opinions. Clearly, the weighting scheme has permitted these criteria to overwhelm the effects of "convenience", in which UHM ranked least, and key criteria of "accuracy" and "uncertainty", in which UHM was ranked as average or below. Heavier weighting on convenience, accuracy, uncertainty estimate, and public acceptance would affect the method's score. Further consideration of the weighting scheme, with input from several sources, may be desirable before adopting the ranking results.

In choosing criteria scores, we have usually presumed the experts would make the best possible use of geologic data/theory, or the best accommodation to incomplete data, etc. The potential for expert panels probably exceeds the performance of a specific panel, but scoring here has not been adjusted for any imagined performance factor. In the TERA application, the experts were limited to choosing seismicity zones and Poisson Log-Linear model seismicity parameters for those zones.

Finally, the full Uniform Hazard Method has not been rated in this ranking, only those aspects relating to earthquake recurrence within the seismic zones. This limitation was necessary because none of the other methodologies to be ranked consider attenuation, site response, and response spectra calculations.

4.1.2 Bayesian Method

One important attribute of the Bayesian Method as proposed by Mortgat and Shah (1978) is the critical need for a sample distribution that includes data at the magnitudes that correspond to large, rare earthquakes. The Bayesian approach is quite simple: a prior statistical distribution (perhaps based on expert opinion) is multiplied by a sample statistical distribution (observed data) to produce the desired estimate, a posterior statistical distribution. The method is formulated for earthquake magnitudes in discrete bins. If no sample data are observed for a particular magnitude bin, the Bayesian output is zero. The equations of Mortgat and Shah (1978) do provide an output if the sample data are zero, but the output is just the unaltered prior distribution and not really a Bayesian result. This need for observed data at large magnitudes is a critical flaw in the method as described. The method also scores low because uncertainty is not treated, and the equations are rather obscure.

4.1.3 Poisson Log-Linear Methods

The Poisson Log-Linear methods tend to cluster in mid-range in the ranking results. For the most part, these methods usually clustered in the mid-range for the individual criteria scores also, most scores were near average. The strongest criteria scores were for convenience, generality, and public acceptance. The weakest criteria scores were for use of geologic data/theory. The least squares methods, Cumulative and Differential, tend to have slightly better scores than Maximum Likelihood and Extreme Value. Among these methods, Maximum Likelihood seemed to give the best accuracy and uncertainty results in the synthetic catalog studies by Knopoff.

The Extreme Value method also seems to invoke extreme scores, very high for convenience, completeness sensitivity and robustness; and very low for use of geologic data/theory and use of subjective judgment.

4.1.4 Semi-Markov Method

The Semi-Markov Method also placed mid-range in the scoring, but shows more variable scores than the other methods. For western data (Figure 4-2), high seismicity areas, the standards-scoring approach with unweighted sums placed the Semi-Markov (E) among the Poisson Log Linear methods (A,B,C,D); the spread-scoring placed E above A,B,C,D for one independent scorer and E below A,B,C,D for the other independent scorer. For eastern data (Figure 4-1), low seismicity, the analogous results are E within A,E,C,D; E near top of A,B,C,D; and E near bottom of A,B,C,D. The weighting scheme proposed in the ranking methodology consistently enhances the relative position of E. The weighting scheme proposed by McGuire consistently lowers the relative position of E. The main cause of these effects is that the McGuire weighting scheme places lower weights on the criteria for completeness sensitivity, completeness accommodation, use of geologic theory, agreement with best knowledge, and use of subjective judgment. On most of these criteria, the Semi-Markov Method is scored higher than the Poisson Log Linear Methods, but the other criteria usually scored lower. This distribution of scores causes the Semi-Markov Method to be sensitive to the weighting schemes used.

Many criteria for which the Semi-Markov Method scored well are related to correspondence between the methodology's model and the presumed actual earthquake processes. The Semi-Markov Method is the only method ranked that has a model that may reflect some causal aspects of the earthquake process. All the other methods directly, or ultimately, use a Poisson random process for earthquake occurrence. Thus with a spread-scoring approach, the Semi-Markov scores very high. The standards-scoring approach can give good, but not necessarily excessive, scores for model-related and expert-related criteria.

4.2 Criteria Scoring

Twelve criteria were selected for ranking the methodologies. In each following section, all methodologies are considered simultaneously relative to a given criterion to achieve the best relative scoring for the criterion. Each criterion is scored on a 1-to-7 scale. A seven-point scale has precedence in psychological testing and probably represents a practical limit to the distinctions that can be judged. Two approaches to scoring were used, and these are referred to as "standards-scoring" and "spread-scoring."

In the standards-scoring approach, standards were defined for low, average, and high scores. The methodologies were assigned scores relative to the standards and to each other. Scores for a particular criterion could be spread out or clustered according to the scorer's perception of the methods. Most of the discussion below is for the standards-scoring approach, and similar considerations are not restated for the spread-scoring. Often, the Log Linear Method was arbitrarily chosen as average because it has considerable prior usage. An average performance is scored as 4, and scores are completely relative for the methodologies considered here.

In the spread-scoring approach, the method perceived as best was usually assigned a score of 7. The method perceived as least was assigned a score of 1. The remaining methods were then distributed in the interval. Sometimes they were grouped toward one extreme or the other.

Separate scores for applications to eastern sites (low seismicity regions) or to western sites (high seismicity regions) seemed appropriate for some criteria. When they differ, scores are given for eastern sites and western sites in the format X(Y) where X is the eastern score and Y is the western score.

Most of the discussion below is for the standards-scoring approach done initially by one author, and later updated after ranking discussions. Two independent scorings were done using the spread scoring approach. These results were compared with the standards scoring results to verify there were no significant differences that should be resolved in the sequence of relative scores, such as A > B, B = C, C > D, etc. No attempt was made to adjust the independent scores, nor are there any differences that represent strong, significant disagreement.

4.2.1 Convenience

Convenience is clearly a general term and tends to encompass many other criteria as well as, perhaps, a subjective attitude toward a method. Ideally, a method would be quick, relatively easy to apply, and require only readily available data. An average method is taken as using readily available data and having no excessive computation requirement. Significantly greater effort rates a lower score. A simple method with minimal need for expert intervention rates a high score.

The Log Linear Method rates an average score, almost by tradition. If adequate data are available, the method can be applied quickly and conveniently. Any difficulties are usually related to cases where data are sparse for any statistical method. There seem to be no differences in convenience between the cumulative least squares and the differential least squares. Maximum likelihood has slightly simpler computations. Convenience is down graded some for Eastern seismicity because much more attention must be given to data completeness and representativeness. The scores are: Cumulative Least Squares 4(5), Differential Least Squares 4(5), and Maximum Likelihood 4(5).

The Extreme Value method is easier than the Log Linear method because only maximum earthquakes in successive time periods are considered. In many cases, less expert input is needed, and the method could be applied practically by rote. However, if the method were developed further, statisticians may wish to include sophisticated tests for distribution type and goodness of fit. Then, the method would lose some of its present convenience. The score is 6.

The Semi-Markov method requires considerable data interpretation to estimate probabilities for the transition matrix and distributions on the various i,j interoccurrence times. Subjective judgment and expert input are required for data on rare events. The number of variables to be evaluated increases as the square of the number of systems states considered. The method has been illustrated with circum-pacific plate boundary earthquakes rather than eastern or western seismicity. The score is 2.

The Bayesian Method can be applied in a straightforward manner if the data are adequate. Subjective expert input is needed to define the data sets and possibly to form the prior distributions. Reliance on expert input increases quickly as the data are less available. The score is 3.

Because it requires extensive interactions with a panel (or panels) of experts, the Uniform Hazard Method requires considerable time and effort in most practical cases. The method analysts also have to be more intimately involved with the many steps than in other procedures, excep perhaps the Semi-Markov method. The score is 1.

For the spread-scoring, the clear choice as the least convenient method is the Uniform Hazard Method, score of 1, because of the cumbersome cycle to elicit review and revise expert opinion. The best score, 7, is assigned to the Extreme Value Method because of its simplicity. The various Poisson methods tend to rank high because they are nearly as convenient as the Extreme Value Method.

4.2.2 Generality

A desirable characteristic is that a method should be applicable to all regions and site types. All methods considered here have broad generality in their theoretical bases, so no tectonic restrictions are evident. However, data availability in various regions does strongly influence the application of some methods. An average score for this criterion is given if the method has at least some applicability to both eastern and western seismicity.

The Log Linear Method can be applied universally, at least in some form and with expert input. A rote application would not be dependable though for an area with sparse seismicity. The method is very much empirical using whatever data is available. Although widely applicable, the method is scored average because difficulties can arise as less data are available. The maximum likelihood method could be downgraded a little because it seems a little less flexible in the presence of data problems. The cumulative least squares is more widely used and applicable because it smoothes the recurrence data points: it also fits some mathematical developments more easily because it uses cumulative distribution. The scores are: Cumulative Least Squares 4, Differential Least Squares 3, Maximum Likelihood 3.

The Extreme Value Method applications are essentially similar to the Log Linear Model, or even a little more broad because only largest earthquakes are needed. The useful data in an eastern seismicity situation may span a longer time period than that for complete magnitude distributions. The score is 5.

The Semi-Markov Method requires either statistical data for all earthquake magnitudes considered, or expert input for those magnitudes. The score is less than average because of the need for many input parameters that may be based on sparse data. In a sense, the inclusion of expert opinion could overcome any local problems, but this approach tends to circumvent the basic method. The method is intended to estimate probabilities for facility lifetimes rather than annual probabilities or return periods. The score is 3.

The Bayesian Method can combine both expert opinion and observed data. Expert opinion can be used to overcome data deficiencies. However if there are no sample data to represent the site, then the method degrades into just expert opinion and perhaps use of data from analogous sites. The Mortgat and Shaw (1978) equations 2.21 through 2.24 provide just the prior distribution as output if the sample distribution is zero, but this result is not really Bayesian. The sensitivity of the Bayesian Method to sample data (site related) at the actual magnitudes of rare earthquakes is a serious problem that affects many of the criteria scores. The scores for generality are 2(3).

Because the Uniform Hazard Method depends heavily on expert opinion, it is the most universally applicable method. The method is limited (if at all) only in being able to find experts to render their judgments. The score is 7.

For the spread-scoring approach, the Uniform Hazard Method easily ranks as the most general method, score of 7, because the experts can call on any techniques they wish to estimate the seismicity parameters. The Bayesian Method was chosen as least general, score of 1, in one independent ranking because of the high potential for zero sample data at many sites and magnitudes of interest. The Semi-Markov Method was chosen as least general, score of 1, in the other independent ranking because data may often be too sparse to estimate the many input parameters. Both rankings place the Bayes and Semi-Markov methods low.

4.2.3 Uncertainty of Estimate

The important attribute here is the ability to estimate an uncertainty such that the true annual probability is included in the interval of the return period estimate plus-and-minus the uncertainty. Magnitude of the uncertainty is a useful criterion only if the above requirement is met. A problem can arise from actual non-stationarity in the earthquake generation process (comparing sample duration to return periods for rare earthquakes). All methods that aren't strongly constrained in some way by the geologic rates may sometimes have unrealistically small uncertainties. Arguments can be made that long-term rates are not necessarily most relevant to a facility with limited design life; current statistics may be more representative than long-term statistics if the two differ. For this ranking, we have modeled earthquake catalogs and used long-term statistics as a standard against which to rate the short-term based estimates (see Section 5.1 and 5.2). For a high score, a method's estimate plus-or-minus one σ should include the long term rate about 70% of the time, and the σ's should be comparable to the actual long-term σ'. For average score, the σ's are larger than the long term σ, perhaps a factor of 2.

Tests using synthetic catalogs based on the most common statistical assumptions for the earthquake generation process (Poisson process with log-linear magnitude distribution) compared several Log Linear Methods and the Extreme Value Method (see Section 5.1). Maximum Likelihood was generally best with Cumulative Least Squares not far behind. Means and sigmas of the estimates were compared to the long-term mean. Other tests were conducted using an earthquake catalog from a generating process meant to be one step more complex than any of the methodology assumptions (see Section 5.2). Visually, the catalog sequence appears clustered and irregular on the time scales appropriate to modeling seismicity estimates; the long term statistics closely model the Poisson process with a log-linear magnitude distribution. The methods tested (the Log Linear variations, Extreme Value, and Bayesian) showed estimates plus-and-minus one sigma that included the long term mean about 20% to 30% of the time. Using plus-and-minus two sigma raised the performance much closer to 70%. To the short term estimates, again particularly the eastern seismicity, the catalogs appeared non-stationary. None of these tests were able to model the reasonable and quite valuable influence of an analyst judging the input data and perhaps modifying or rejecting some data, or using geologic information and theory.

The Log Linear Model generates uncertainty estimates that measure how well the data points fit a linear relation. A fortuitous data set can give an extremely good fit. The method clearly suffers when applied to eastern data sets, uncertainties can vary widely because of the sparse data. However, using expert judgment can keep the input data set reasonable, and the results are at least reasonable. Maximum likelihood approximates a weighted least squares emphasizing data points for small magnitude earthquakes. This weighting is particularly important if the rates for the two lowest magnitudes are not co-linear (log-linear relationship) with rates observed for higher magnitudes, and can lead to much variation in estimates when using eastern data. The synthetic catalog studies results have been tempered somewhat for the least squares methods because expert input would clearly be used in most applications to carefully evaluate the input data. The scores are Cumulative Least Squares 2(4), Differential Least Squares 2(4), and Maximum Likelihood 2(4).

The Extreme Value Method provides uncertainties quite similar to the Log Linear Method when the data are generous because the underlying theory is the same. However, for sparse data as eastern seismicity, the various short term estimates vary considerably. Expert judgment of the input data cannot identify possibly spurious data points so easily here as in the Log Linear Method. In practical cases, there may be extreme value data available for a longer time period, and the statistics could improve somewhat. The scores are 2(4).

In applying the Semi-Markov Method, considerable unquantified variation can be introduced through using subjective opinion to develop the transition matrices. In most cases, the data for rare events are probably few and uncertainties on the transition probabilities are large. The way these uncertainties propagate through the Semi-Markov process that is used to model earthquake generation is not practical to calculate. The method's authors (Patwardhan et al 1980) used a parametric analysis and explored only a few combinations of uncertainties. Because the uncertainties are not defined, the score is 1.

For the Bayesian Method, uncertainties in the final estimate are not defined by Mortgat and Shah (1978). Each posterior distribution, rate and magnitude, is analytic with defined uncertainties so a combined uncertainty after removing the condition on n earthquakes may be feasible. However, such uncertainty would be rather artificial because of the choices for prior functions. The prior function on the rate λ is a gamma function for the mathematical convenience, and it is used without regard for the expert's subjective uncertainty in the rate. If data are used for the rate λ, there is no incorporation of a σ_{λ}. Furthermore the posterior distribution is strongly influenced by the relative weighting given to the prior distribution and the sample function. Similar arguments apply to the magnitude distributions. Uncertainties are not defined. The score is 1.

The Uniform Hazard Method is based upon the Bayesian approach, but experts are asked to perform as the Bayesian processor. They are asked to give estimates of uncertainty in their results. Thereafter in the method, detailed attention is given to uncertainties and how they propagate through the procedure. As a result, the method gives a good analytical result of the cumulative effects of various sources of uncertainty. However, in practice, these uncertainties are strongly subjective rather than based on statistics. There is an unknown interdependence among various uncertain quantities, and the very basis of the method is such that the estimated uncertainties tend to measure the expert's relative confidence rather than any physical quantity. The values are sometimes quite large, they serve best for comparison with other sites analyzed in the same manner. The Uniform Hazard Method does not provide uncertainties at the earthquake recurrence rate stage, because the expert's input is combined only at the end of the full procedure. The score is 2 because the uncertainty values are best used for comparison with other sites analyzed similarly.

For the spread-scoring approach, both the Semi-Markov Method and the Bayesian Method are ranked as least, score of 1, because they do not provide uncertainty estimates for a specific site calculation. The Uniform Hazard Method is also ranked low because its uncertainty estimates are not intended to be used except for comparisons. None of the methodologies was perceived as deserving a 7 score, the remaining methods were given average, or essentially neutral, scores.

4.2.4 Accuracy

The standard proposed for accuracy testing with synthetic catalogs was that an estimate for return period should be within some percentage of the true long-term return period: 10% for a high score, 30% for average, and 50% for a low score. This standard was not so straightforward in practice. Testing using some methods on the Knopoff catalogs (see Section 5.1)can be argued to be somewhat circular because the catalog synthesis was based on the same statistical assumptions as the methodologies. Thus we probably obtained predicably acceptable accuracies for some tests on the Knopoff catalogs. Testing with the McGuire catalogs (see Section 5.2) has the problem that the catalogs are quite variable in their statistical properties for time periods comparable to actual data catalogs. This variability may, in fact, be an excellent model of real catalogs, especially for eastern seismicity. Some results for the McGuire catalog differ from the long-term average, but are very good local estimates. Some methodologies such as Semi-Markov and Uniform Hazard Methodology were not tested against the synthetic catalogs. Scoring this criteria does use the test results, but has to be more subjective than originally envisioned.

The Log Linear Model showed predictably good results on the Poisson catalogs when the method assumptions and the catalog distribution were similar. Some synthetic catalogs, particularly eastern seismicity, were erratic at the upper magnitudes. The automatic procedure for least squares produced some fits that an analyst would probably modify by rejecting some data points. Differential least squares was most sensitive to "questionable" data points. With the McGuire catalogs, some recurrence curves and extreme value plots were very linear indicating a good data fit. At the same time, these portions of the catalog were clearly different statistically from the long term catalog properties. These results have pointed out the need to analyze real data catalogs for their most appropriate distribution form and the goodness-of-fit. Scores assigned are: Cumulative Least Squares 3(4), Differential Least Squares 2(4), and Maximum Likelihood 3(4).

The Extreme Value Method was tested along with the Log Linear methods, and the above discussion applies here as well. Extreme Value seemed to be at least comparable to Cumulative Least Squares. The scores are 3(4).

The Semi-Markov method is essentially a modeling procedure rather than a way of analyzing seismicity data. The results are heavily dependent on the initial conditions and the process assumptions. When enough statistics are known that the transition matrix elements and the interoccurrence time distributions are well constrained, the method's results may be comparable to other methods. But in most cases, the statistics of rare events will not be known adequately. The method seems most appropriate for testing the implications of various seismicity hypotheses. The accuracy of the results are generally unknown. The score is 1.

The accuracy of the Bayesian Method results can be greatly affected by the subjective input for the prior distributions and the weighting applied to these distributions. Thus, the accuracy is poorly controlled if expert opinion is given much weight. The very act of trying to combine disparate data such as geologic opinion and seismic history can cause the accuracy to be ill-defined. Comments in Section 3.2.3 about method uncertainties are equally applicable to the method accuracy. If two classes of data are used to define the prior distribution and the sample distribution, accuracy is a function of the representativeness of these distributions. However, even with pure data input the accuracy could be variable. The Bayesian Method is scored less than average for this criterion primarily because the accuracy is not well defined. Furthermore, if there is no sample data for the desired magnitudes, the Bayesian approach indicates zero output. The scores are 2(3).

In the Uniform Hazard Method, accuracy is not readily controlled. There is nothing in the method that distorts the inherent accuracy of the expert opinions, but their inherent accuracies are totally unknown. In general, we would like to believe that the expert opinions are at least reasonable, and perhaps as accurate as any formal statistics could be for the same site. The scores are 3(4).

For the spread-scoring approach, the Semi-Markov Method was judged least, score of 1, because the accuracy is not defined, and the procedure is just an attempt to model the seismicity process. The Bayesian Method was also given a score of 1 here because a lack of sample data at the rare-event magnitudes leads to a nul answer. As in the uncertainty estimate criterion, no method was perceived as deserving a 7 score. The Poisson-based methods were given average, or neutral, scores.

4.2.5 Completeness Sensitivity

Data for a particular site region may be incomplete because some small or moderate shocks that have occurred may be omitted from the seismicity record. Incompleteness to some degree is practically unavoidable as the coverage for seismicity is extended to earlier dates and/or smaller magnitudes. An average score for this criterion is given if a method recognizes the sensitivity, and the sensitivity is not excessive. Low scores are given if the sensitivity is not well defined. High scores if the method is insensitive. The ability to overcome any sensitivity is considered in the subsequent criterion.

In the Log Linear Method, sensitivity effects are recognized in the distribution of the recurrence data points, particularly the differential recurrence data points. The least squares method are not considered excessively sensitive for Western data assuming the analyst recognizes incomplete rates and either avoids or accommodates them. For the more sparse Eastern data, there is less opportunity to recognize or accommodate incomplete rates. The Maximum Likelihood method places less emphasis on recurrence curves, so incompleteness is less likely to be observed unless the curves are plotted. The scores are: Cumulative Least Squares 3(4), Differential Least Squares 2(4), and Maximum Likelihood 2(4).

The Extreme Value Method is among the least sensitive to data completeness because incomplete data affect mostly the small-magnitude shocks. The main virtue of the method is overcoming data incompleteness. However, if the time periods used are extended back too far, then some earthquakes could be missing even from the extreme value record. The score is 6.

The Semi-Markov Method uses statistics for only the larger-magnitude earthquakes, those likely to be near or at the design levels. Completeness for smaller magnitude shocks is irrelevant. However, completeness for the larger magnitude earthquakes is quite important. The method recognizes the sensitivity and deals with it using expert opinion to fill in missing data. The score is 4.

In the Bayesian Method, the equations presume that data are complete for the magnitude ranges used. The requirements here are the same as those for the Log Linear methods, except the Bayesian method doesn't itself give indications of the data completeness. The Bayesian Method permits expert opinion to be used to reduce or overcome completeness sensitivity in the prior data. The sample data are presumed complete and must be non-zero, and in some cases may be chosen with that criterion in mind. The rate λ will be sensitive to data completeness because the majority of shocks are at the low magnitudes. The score is 4.

The Uniform Hazard Method is judged to have average sensitivity to data completeness, but expert opinion is used to accommodate all deficiencies. The score for sensitivity is 4.

For the spread-scoring approach, the Bayesian Method was ranked as least because it has fatal sensitivity to the need for sample data at the desired magnitudes. Although Mortgat and Shaw's (1978) equations just pass through the prior distributions in the presence of a zero sample distribution, this is judged to be a degeneration of the Bayesian approach. One independent scorer ranked the Semi-Markov Method as having the best score for completeness sensitivity because expert opinion can be used to great extent. The other scorer considered Extreme Value as best because it obviates the need for much of the lesser magnitude data.

4.2.6 Accommodation of Incomplete Data Sets

This criterion is not completely independent of that for Completeness Sensitivity. However, the ability to reduce adverse effects of incomplete data or to modify the estimated uncertainty appropriately is the quantity measured by this criterion. None of the methods ranked attempt to modify the earthquake recurrence uncertainty using the data completeness. The methods that are quite insensitive to data completeness are given an average, or neutral, score so this criterion will neither help nor hurt their total scores. High score is given to methods that have, or easily could have, a formalism for accommodating incompleteness. Average scores are given to methods that consider incompleteness only in the input data evaluation and selection.

In the Log Linear Method, the least squares approaches can overcome incompleteness in many data sets using the technique described by Stepp (1972). This technique tries to develop stable earthquake recurrence rates for magnitudes that will define the recurrence curve. The analyst can also evaluate the seismicity data and choose time periods for which the data are complete for each magnitude range. For the Maximum Likelihood approach, Weichert (1980) described a technique to use different time periods for different magnitude ranges. These techniques, Stepp (1972) and Weichert (1980), are not widely used in the Log Linear Methodology, but they could be incorporated. For sparse, Eastern data, accommodating incompleteness becomes more difficult. The scores are: Cumulative Least Squares 3(5), Differential Least Squares 3(5), and Maximum Likelihood 3(5).

The Extreme Value Method is given a low score here because very little is done to accommodate any data incompleteness other than to adopt a threshold magnitude above which the data are presumed complete. The duration of useful data depends on the threshold magnitude chosen. The score is 2.

In the Semi-Markov Method, expert opinion can be used to overcome any lack of data. However, there is nothing in the method, other than expert input, by which to judge the data completeness. The score is 6.

Expert opinion can also be injected into the Bayesian Method to overcome data completeness for the prior distributions. However, no treatment for completeness (either testing or correction) is specified for the sample distributions. Although the analyst may try to select the sample distribution to avoid incompleteness, a lack of sample data for large-magnitude earthquakes causes the method to fail. The score is 2.

The Uniform Hazard Method is given the maximum score because the experts can be used to overcome any data completeness. The score is 7.

For the spread-scoring aproach, the Uniform Hazard Method is ranked best, score of 7, because the experts can provide reasonable seismicity parameters in all cases using data or theory. The Baysian Method is ranked least, score of 1, because there is no provision to overcome a lack of data at the desired rare-event magnitudes, and still retain the Bayesian approach.

4.2.7 Incorporation of Geologic Data

This criterion considers how geologic and tectonic information for the site region are incorporated into the recurrence rate estimates. The data usually relate to specific faults and tectonic structures, or the characteristics of observed large earthquakes. Data may be used directly such as in supplying the average rate for large earthquakes during Holocene time from fault studies. Data may be used in a confirmatory sense in judging the reasonableness of methodology estimates. General geologic/tectonic theories or hypotheses are considered in the subsequent criterion. High scores are given if the geologic data are used directly as input to the methodology estimates. Average scores are given if the data are used to verify or modify input data. Low scores are given if the geologic data is used only for comparison with the methodology estimates.

In the Log Linear Method, geologic data are sometimes used to supply or at least substantiate the rates input for larger magnitude earthquakes. These rates can be used for the Cumulative Least Squares and the Differential Least Squares. Geologic data are not used as input for the Maximum Likelihood estimates. Geologic data for recurrence rates are essentially non-existent for most eastern regions, so Eastern scores are downgraded. The scores are: Cumulative Least Squares 2(4), Differential Least Squares 2(4), Maximum Likelihood 2(2).

The Extreme Value Method is nearly a pure statistical approach. Geologic data are not used to modify the input data in any way, but only for comparison with the recurrence estimates. The score is 2.

The Semi-Markov Method generally requires geologic data to construct the transition matrix probabilities for large earthquakes. Seismological data are likely to be inadequate, or even missing for the rare events. Patwardhan et al (1980) demonstrated the method using regional data for the circum-Pacific. When applied to a siting region, geologic data for large earthquake probabilities would be essential. The Eastern score is downgraded because there are so little applicable geologic data. The scores are 2(5).

In the Bayesian approach, little geologic data are used directly. Experts should incorporate geologic data in specifying their input; but expert input is not required. Here too the Eastern score is down from lack of useful data. The scores are 2(4).

The Uniform Hazard Method relies on expert opinion input, and the experts should temper their judgements with geologic data. Geologic data are not used as direct input. The Eastern score is down from lack of useful data. The Eastern experts probably use more geologic theory and hypothesis than data. The score is 4.

For the spread-scoring approach, the Uniform Hazard Method is rated best, score of 7, because of its flexibility in how data can be used and encouragement for experts to use geologic data to the fullest extent. The Extreme Value Method is rated least, score of 1, because geologic data are essentially ignored. One independent scorer also placed the Poisson-Log Linear methods equally low because they can be used with very little geologic data input. The other independent scorer placed these methods a little higher because geologic data is sometimes used to derive or corroborate rates for moderate to large shocks.

4.2.8 Use of Geologic and Tectonic Theory/Hypothesis

Geologic and Tectonic theory, whether supported by recurrence data or other geologic data, may be used to help estimate recurrence rates of rare events. Theories and hypotheses tend to describe the regional setting and compare local earthquake sources with similar sources in other regions by analogy. Methodology input data or constraints may be deduced from some regional tectonic theories/hypotheses. High scores are given to methods that strongly depend on applying geologic theory. Average scores are given if a method can use geologic theory to provide or modify some input data. Low scores are given if a method uses geologic theory only as a comparison check on the methodology estimates.

The Log Linear Methods with least squares can use geologic theory to place reasonable constraints on rates for large earthquakes. Some geologic theory may be translated into occurrence rates for large shocks to use either as input data or for checking results. Geologic theory does not provide input data for the Maximum Likelihood approach. The scores are: Cumulative Least Squares 3, Differential Least Squares 3, Maximum Likelihood 2.

The Extreme Value Method is very nearly a purely statistical approach. The input values for maximum earthquakes during sequential time periods are not altered by geologic theory. Theory can be compared with the method results. The score is 2.

The Semi-Markov Method is flexible in using geologic theory. However, in most practical cases, both seismological data and geological data are inadequate to provide the required input data. The transition matrix probabilities for rare events and the interoccurrence time distributions must sometimes be supplied from theoretical arguments. The score is 6.

In the Bayesian approach, expert opinion may be used to develop the prior distributions. The experts in turn can be expected to incorporate geologic theory into their judgments. The sample distributions will not usually be modified by geologic theory. The score is 4.

The Uniform Hazard Method uses geologic theory very heavily because input data is based on expert opinion. The actual degree to which geologic theory is incorporated into the method depends completely on the individual preferences of each expert. The reports by Bernreuter (1980) indicate that the experts do use theoretical considerations extensively. The score is 7.

For the spread-scoring approach, the Uniform Hazard Method is scored best, score of 7, and the Extreme Value Method is scored least, score of 1, for the same reasons given in scoring use of geologic data.

4.2.9 Correspondence With Best Knowledge of the Earthquake Process

Our best knowledge of the earthquake process represents two broad areas: 1) the geologic, tectonic, and physical processes for generating earthquakes, and 2) the statistical distribution, or distributions, that describe earthquake occurrence either theoretically or empirically. The earthquake process is not well understood, and attempting to describe the state of knowledge itself is far outside this discussion. For this criterion, high scores are given if the method's theoretical basis (relative to the earthquake process or the statistical distributions) are likely to be viewed favorably. Average scores are given if the bases are likely to be only acceptable. Low scores are given if the bases are questionable or not present. Some of the statistical testing described in Section 5 was aimed at determining how well various statistical distributions fit observed earthquake occurrences.

The Log Linear Method presumes a Poisson process for earthquake occurrence and the usual log-linear distributions for earthquake magnitude, $\log N = a - bM$. The magnitude distribution may be unbounded or bounded at some maximum magnitude. The bounded, log-linear distribution is taken as "average" for this criterion. This distribution is widely accepted as evidenced by common practice. There are some theoretical arguments for the mathematical form of the recurrence relation (Kanamori and Anderson, 1975), but the relation is for the most part empirical. The Poisson assumption is a convenient, useful approximation, but it too is clearly incorrect for many discrete sources from individual faults in the western U.S. The scores are downgraded a little for the western seismicity where deterministic seismicity models may sometimes be more appropriate. Methods using unbounded forms for the magnitude distribution and extrapolating large magnitudes for rare earthquakes would be scored lower. Maximum magnitudes need some geological or theoretical constraint if not evidenced by the data. Unbounded methods are not scored separately. The Maximum Likelihood Method is downgraded slightly because it has been criticized for placing too much emphasis on the smaller magnitude earthquakes, very much like a weighted least squares fit to the recurrence data. The criticism is moot if the rates do follow the distribution exactly, but real data are always subject to variations. The scores are: Cumulative Least Squares 4(3), DifferentialLeast Squares 4(3), and Maximum Likelihood 3(3).

The Extreme Value Method has the same theoretical bases as the Log Linear Method. However, Extreme Value is downgraded slightly because it has been criticized for not using much valuable data at the lower magnitudes. Here too, as is the case for Maximum Likelihood, the effect criticized depends on how well the data follow the presumed statistical distribution. The score is 3.

The Semi-Markov Model attempts to model the earthquake process by including some process memory. Conceptually, the occurrence time for earthquakes and their magnitudes should be related to past history. The Semi-Markov process provides a statistical relation, which may or may not be founded on knowledge of the earthquake process. If expert opinion is used to develop some input parameters, knowledge of the earthquake process will influence the parameter values selected. In principle, this method could score much higher than the Log Linear models because the method involves at least some memory process which is clearly more physical than the Poisson process. An average score is given here because in practice, as described by Patwardhan et al (1980), the transition probabilities are mainly empirical. The score is 4.

The Bayesian Method uses the same earthquake occurrence distribution (Poisson) as the Log Linear Model. The probability distributions on magnitudes, P_{M_i}, are mostly empirical, unless supplied by expert opinion. The forms for prior distributions on rate λ and probabilities P_{M_i} are completely arbitrary, being chosen for mathematical convenience. The score is 2.

In the Uniform Hazard Method, the experts are free to incorporate their best knowledge of the earthquake process, but they are not required to do so. No particular theories are specified or excluded. The score is 4.

For the spread-scoring, the Semi-Markov Model was given the highest score, 7, because the statistical model in the method is somewhat similar to current physical models of the earthquake generation process.

All of the Poisson-related methods were scored lower because the actual earthquake process is not likely to be Poisson-random for large-magnitude earthquakes. Various fault studies show a much more deterministic process. The Bayesian Model was scored least, 1, because quite arbitrary distributions, gamma and beta, are used to fit the prior data, or to describe the expert opinion.

4.2.10 Robustness

Each methodology has statistical assumptions such as distribution types for magnitudes or interoccurrence times, data variability, or independence of various input parameters. A method is considered robust if its results are useful even if the assumptions are not strictly true. The more the assumptions can be relaxed, the more robust the method. The average performance for this criterion is taken as the Log Linear Method (Least Squares) because estimates seem reasonable in so many instances, at least for moderate magnitudes, even though the independent earthquake assumption is not strictly true. Low scores are given if all statistical assumptions must be true for the method to work.

The Log Linear Method is taken as the norm as mentioned above. The magnitude distribution is also a critical statistical assumption, particularly the maximum magnitude aspect. Both geological and geophysical arguments require some maximum magnitude, and the Cumulative magnitude distribution cannot be linear near the maximum. The Differential distribution probably isn't linear there either. The scores are: Cumulative Least Squares 4, Differential Least Squares 4, and Maximum Likelihood 4.

The Extreme Value Method has the same statistical bases as the Log Linear Method. But the Extreme Value Method is considered much more robust because only the largest earthquake in each sequential time period need be considered rather than the whole distribution. Within the method, plots of maximal earthquakes versus the assigned probabilities are theoretically linear. The plots can provide indications about the validity of the statistical assumptions, just as the recurrence curves can be used in the Log Linear Method. The score is 5.

In the Semi-Markov Method, a Semi-Markov process is used to model the gross properties of the earthquake occurrence. The input parameters are empirical or derived from expert input. The main statistical assumptions are that the current system state affects only the succeeding system state, and that the transitions and interoccurrence times can be modeled by probability distributions. The model parameters, once chosen, define a very specific earthquake process. As the model parameters depart from their best physical representation, the results may be degraded rapidly. The score is 2.

In the Bayesian Method, the sample distribution on the rate λ is Poisson as in the Log Linear Model. The magnitude sample distributions are strictly empirical and independent of any statistical assumptions. The prior distributions on both rate and magnitude are arbitrary and quite flexible. If the sample data are non-zero, the method is very robust because few assumptions are made. However, the possible lack of sample data is a critical condition for the method. The score is 2.

In the Uniform Hazard Method, there are no statistical assumptions about the earthquake process. The experts act as Bayesian processors and absorb all robustness problems within their judgments. The score is 4.

For the spread-scoring approach, the Extreme Value Method was scored highest, 7, because it has minimal dependence on data completeness. The Poisson Log-Linear forms are widely applicable. The Poisson Log-Linear methods were scored high, but downgraded by one scorer for Western data because some faults may be more predictable seismic sources. The Semi-Markov Method was scored lowest, 1, because it has critical dependence upon the presence of sample data at the desired earthquake magnitudes.

4.2.11 Incorporation of Subjective Judgment

Subjective judgment can enter a methodology in several forms. Expert opinion itself may be used as input data. Geological and seismological data must be judged for its quality and completeness. Some methods may require expertise in their application. And finally, the recurrence estimates must be judged for their reasonableness. For this criterion an average score is given if the method rationally uses subjective judgment concerning the data, application, and results. Low scores are given if the method bypasses judgment, or could be applied by rote. High scores are given for incorporation of expert opinion as input data and formal review of results.

The Log Linear Method is taken as average because judgment should be used in selecting and using the seismicity data. Rate data for large magnitudes are sometimes obtained from geologic data through expert opinion. The method can lead to questionable results if subjective judgment is not used in its application. The Maximum Likelihood approach is downgraded somewhat because it tends to bypass evaluating the distribution of the recurrence data. Higher scores are given for Eastern application because the data are fewer and more judgment is required. The scores are: Cumulative Least Squares 5(4), Differential Least Squares 5(4), and Maximum Likelihood 4(3).

The Extreme Value Method can be applied with little subjective judgment. Considerable potential seems available for expert analysis of the distribution selection and goodness-of-fit, but these are not properly part of the method as currently applied. The score is 2.

The Semi-Markov Model tends to require considerable expert judgment in formulating the process input parameters, because the data are few for the desired large earthquakes. The score is 6.

The Bayesian Method normally uses expert opinion as input data for its prior distributions on rate and magnitude. Observed data are used as the sample distributions. The degree to which subjective judgment is used, and the weight assigned to that judgment can be varied widely. The Eastern applications are scored higher because less data are available. The scores are 6(5).

The Uniform Hazard Method uses subjective judgment directly for all its input parameters. The experts may use hard data, but that usage does not directly affect the methodology calculations. The score is 7.

For the spread-scoring, the Uniform Hazard Method clearly rates the highest use of subjective judgement, score of 7. The Semi-Markov Method and Bayesian Method are somewhat less, but still higher than Poisson-based methods. The least use of subject judgment is in the Extreme Value Method as it is currently defined, score of 1. One independent scorer placed Maximum Likelihood equally low, with the least squares methods only slightly better because all of these methods can be applied with very little input from expert opinon. The other independent scorer considered Maximum Likelihood and the least squares methods closer to an average score because expert opinion is usually incorporated into the data evaluation.

4.2.12 Public Credibility

The public and scientists from other disciplines will be asked to accept earthquake risk estimates that depend on estimated earthquake recurrence. Correctness is the ultimate criterion, but the public's perception of that correctness is also important. Esoteric methods are more difficult to understand, and probably less likely to instill confidence in their results. Scores for this criterion are high, average or low according to whether the methods are likely to be perceived with favor, acceptance, or skepticism.

In the Log Linear Model, the Cumulative and Differential Least Squares methods are reasonably straightforward to non-specialists. The recurrence curves are easy to understand. However, the Maximum Likelihood approach is more obscure and subject to misunderstanding. The scores are: Cumulative Least Squares 6, Differential Least Squares 6, and Maximum Likelihood 4.

The Extreme Value Method relies on a mathematical derivation that might seem obscure to many persons. Logrithms of logrithms are not necessarily intuitive quantities. Some might argue that the method is questionable because it requires so little input data. However, the method deals with a simple concept, the largest earthquake in each successive year, and probably could be perceived as acceptable. The score is 3.

The Semi-Markov Model presents itself as a way to construct something like the earthquake occurrence process. The required input parameters, such as the transition matrix probabilities, are not easily available for rare events from observed data, nor are they readily derived from theory. The results seem likely to be perceived as interesting, but not necessarily a firm basis for risk decisions. The scores are 2.

The Bayesian approach uses sophisticated statistics to develop a formalism for application. The equation to be applied is itself complicated and the action of the various terms is not easily discerned. Acceptance of the method is helped by direct use of expert opinion, and the concept of blending both opinion and data. The scores are 3.

The Uniform Hazard Method has sophisticated theoretical bases, but the actual implementation of the method can be presented in a straightforward way. This ranking, of course, is considering only the UHM steps that treat earthquake occurrence at the seismic sources. In addition, the method can claim to make rational use of all the best expert opinion. The score is 6.

For the spread-scoring approach, the method likely to be most acceptable to the public is judged to be the Cumulative Least Squares, score of 7, particularly for Western data. The method has the simplest concepts and good emperical support. Differential Least Squares and Maximum Likelihood are nearly as good. The Semi-Markov Process is scored lowest, 1, because it is a complex modeling process. The Bayesian Method has complicated mathematics, and scores low.

5. ANALYSES TO SUPPORT STATISTICAL CRITERIA

The earthquake recurrence methodologies to be ranked have various statistical bases such as the way the data are manipulated, or the distributions used to describe magnitudes and occurrence times. Accordingly, several criteria were chosen to evaluate the methodologies for accuracy, uncertainty of estimates, sensitivity to incomplete data, accommodation of known incomplete data, and robustness. Studies conducted by consultants to this program provide support for scoring the statistics-related criteria. Three approaches were used: 1) testing methodologies on synthetic catalogs based on the usual Poisson, log-linear models for earthquake occurrence, 2) testing methodologies on synthetic catalogs generated from physically based models more complex than the usual statistical models, and 3) fitting real seismicity catalogs to statistical models. Sections 5.1 and 5.3 are essentially verbatim reports provided by Leon Knopoff and Nancy Mann, respectively. Section 5.2 was prepared at Earth Technology Corporation describing the studies Robin McGuire directed.

5.1 Synthetic Catalog Testing, Leon Knopoff

This is a report on investigations into the statistical properties of synthetic earthquake catalogs. We have two prototype expressions which are the starting points for our calculation. These are the linear log frequency-magnitude formula for a system with unbounded upper magnitudes,

$$\log\dot{N} = a - bM \qquad (1)$$

where $\dot{N}$ is the number of earthquakes per unit time with magnitudes greater than M, and the linear log energy-magnitude formula

$$\log E = \alpha + \beta M \qquad (2)$$

In what follows, we shall refer to the cases

I. $a = 4.77$, $b = 0.85$

II. $a = 3.77$, $b = 0.85$

as the Western and Eastern models respectively. The Western model was indicated by Richter (1958, page 360) to be appropriate for Southern California. We have arbitrarily reduced the number of Western earthquakes tenfold for the Eastern model in an effort to model the occurrence of a sparse population of earthquakes. We take $\beta = 1.5$ in (2), when it becomes necessary to invoke this equation (Gutenberg and Richter, 1956).

We shall also have recourse to a differential form of (1). By direct differentiation, the number of earthquakes per unit time $\Delta\dot{N}$ in a magnitude band of width ΔM is

$$\log\Delta\dot{N} = a' - bM$$

where

$$a' = \log b + \log M + \log(\ln 10) + a \qquad (3)$$

If we assume (3) applies over a finite range of magnitudes from M_o to M_{max}, integration of (3) yields the version of (1) with bounded upper magnitudes:

$$\log\dot{N} = a - bM + \log(1 - 10^{-b(M_{max}-M)}) \qquad (4)$$

The fractional number of earthquakes with magnitudes greater than M, if there are upper and lower cutoff magnitudes, is

$$f = \frac{10^{a-bM} - 10^{a-bM_{max}}}{10^{a-bM_o} - 10^{a-bM_{max}}} \qquad (5)$$

5.1.1 Synthetic Catalogs

We have generated three catalogs for analysis for each of the two values of a. In the first of these, the catalog is a Poisson independent model of (5). The sequence of earthquake magnitudes M_i is generated from (5) by the formula

$$f = Ra_i$$

and solving the resultant for M_i where Ra_i is the ith random number; Ra_i has uniform probability of occurrence over the interval 0 to 1. We have generated three subcases of this catalog corresponding to

$$M_{max} = 6,\ 8,\ 12$$

Practically the condition $M_{max} = 12$ should be equivalent to setting $M_{max} = \infty$. We choose $M_o = 4.0$. (Useful constants for our later discussion are the numbers $\lambda = 23.44 = 10^{4.77-0.85(4.0)}$ and $\lambda = 2.344 = 10^{3.77-0.85(4.0)}$. This means that there are 23.44 or 2.344 earthquakes per year with $M \geq 4.0$ in the two Poisson catalogs). To adjust the time of each event we use the probability $(1 - e^{-\lambda t})$ that the next earthquake will occur within the time t after the last one. Thus the time t_k of the kth event is

$$1 - e^{-\lambda(t_k - t_{k-1})} = g_k$$

where g_k is a random number between 0 and 1; magnitudes are assigned as above.

In the second set of catalogs, we equip catalogs of the first type with aftershocks. To do this we first assign a location to each event in a catalog of the first type. These events are distributed randomly over a line whose length is taken to be 400 km, which is roughly the distance along the San Andreas Fault from the Los Angeles-Kern County border to the Mexican border. We assign aftershocks to each main shock of magnitude M_i (catalog Type I) according to the prescription given by Kagan and Knopoff (1978). We divide the magnitude interval $M_i - 4.0$ into 0.1 units. Let there be r_i such intervals: $r_i = 10^{(M_i-4.0)} + 1$. We calculate

$$n_k = -R\{0.01979(1.188)^{k-1}\ln h_k\} \tag{6}$$

where R is a round-off operator, h_k is a random number between 0 and 1, and $1 \leq k \leq r_i$. The quantity n_i is the number of aftershocks of magnitude M_k: $k = 10(M_i - M_k) + 1$. Having identified the magnitude of all aftershocks, we indicate when and where they are located with respect to the main shock.

The time after each identified main shock is calculated according to the Kagan and Knopoff formula:

$$t_m = .0935 \times 3^{M_i - 7.0} \times 10^{4 j_m} \tag{7}$$

where M_i is the magnitude of the parent earthquake, m is the identification number of an aftershock and j_m is a random number between 0 and 1. Aftershocks are chosen to occur either bilaterally or unilaterally distributed about a main shock; these two distributions can occur with equal probability; if the distribution is unilateral, the sequence has equal probability of lying to the north or south of the mainshock; if the distribution is bilateral, successive aftershocks are alternately north and south of the main shock. The distance of the mth aftershock from the main shock is also given by Kagan and Knopoff to be

$$x_m(\text{km}) = \frac{37.5(3.0)^{M_i - 7.0}}{(1 - 0.646446\ell_m)^2}$$

where ℓ_m is a new random number between 0 and 1. If an aftershock is located outside the 400 km band, it is not removed from the catalog.

The third catalog is a modification of the semi-Markov model of Patwardhan et al (1980), which is in reality a revised version of the Poisson stochastic model of Knopoff (1971). We model the stochastic distribution of Benioff strain accumulation and release, instead of energies as originally suggested by Knopoff. The reason for this is computational. We imagine that there is a uniform rate of increase of some function (strain or energy) in the interval between earthquakes. If we use energy as the function, the range of the energies spanning earthquakes from $4.0 \leq M \leq 8.0$ (for example) is 10^6 while the ratio of energies for a 0.1 increment in magnitude is $10^{0.15} = 1.4$ if $\beta = 1.5$. This means that we have to explore a range of 10^6 in maximum energy steps of 0.4 units. The amount of computer time required is prohibitive. On the other hand if we use strain = energy$^{1/2}$, with an effective $\beta = .75$, the range is only 10^3 in maximum steps of 0.19 units, which is about 5000 steps.

The stochastic model of Knopoff (1971) requires the specification of two probability densities. One is $\lambda(S)$ such that $\lambda(S)dt$ is the Poissonian probability that an earthquake will occur at a given time, where dt is the time step of the calculation; the probability is allowed to vary with the state of deformation of the earth. In our calculations we have arbitrarily chosen

$$\lambda(S) = CS \qquad 0 < S < S_{max} \qquad (9)$$

$$= \infty \qquad S = S_{max}$$

We choose the value of λ at the midpoint of the range $(1/2)S_{max}$ to be the mean rate of occurrence of earthquakes, which is 23.44 year^{-1} for the Western catalog. Thus $C = \frac{2}{23.44}$ years^{-1}.

The other probability function concerns the likelihood that if an earthquake occurs while the system is in the state X, the final state of strain will be $S < X$, i.e., that the strain release will be $(X - S)$. We use the cumulative probability function

$$T(S,X) = \frac{S^{-n}-X^{-n}}{d^{-n}-X^{-n}}$$

where d is the minimum shock size ($d = 10^3$ for $M_O = 4.0$); the power law S^{-n} is chosen so that we simulate the power law (2) as much as possible.

The exponent $n = (\frac{b}{\beta} - 1) = \frac{0.85}{0.75} -1 = .13.$

The function T(S,X) is properly normalized so that T = 0 if S = X and T = 1 if S = d. We assume an input of strain at a rate α that is constant. We fix α by computing the total average strain release per unit time by integrating (2) from M_O to M_{max}. We use values of 73535, 119932, and 167678, strain/day for M_{max} = 6, 8, 12 respectively. The program uses time increments of 1 day. The strain is raised by α units from the preceding value. Via a random number selector and (9), we decide whether an earthquake will occur. If no earthquake occurs, we increment the time (and the strain) once again. If an earthquake occurs, we determine the final state of strain from (10) by another random number selection.

To summarize, we generate three types of catalogs: Poissonian, Poissonian plus aftershocks, and the Strain/Probabilistic stochastic model. Three values of M_{max} were used: 6.0, 8.0, and 12.0. Two values of the coefficient a were used: 4.77 and 3.77. This creates 18 catalogs in all. Each of the catalogs was simulated 15 times for a total span of 50 years per subcatalog. We also the chained the 15 realizations to form 18 single catalogs, each 750 years in duration.

5.1.2 Processing

Four models of the processing of the above catalogs were considered: differential least-squares, maximum likelihood, cumulative least squares, and extreme value statistics. An example of the plotted distributions from one of the synthetic catalogs is given in Figure 5-1.

5.1.2.1 Differential Least-Squares

We construct a histogram of the logarithm of the number of events in each 0.1 magnitude interval and fit according to (3) by least squares. We tabulate the average of the quantities a, b, and their rms deviations

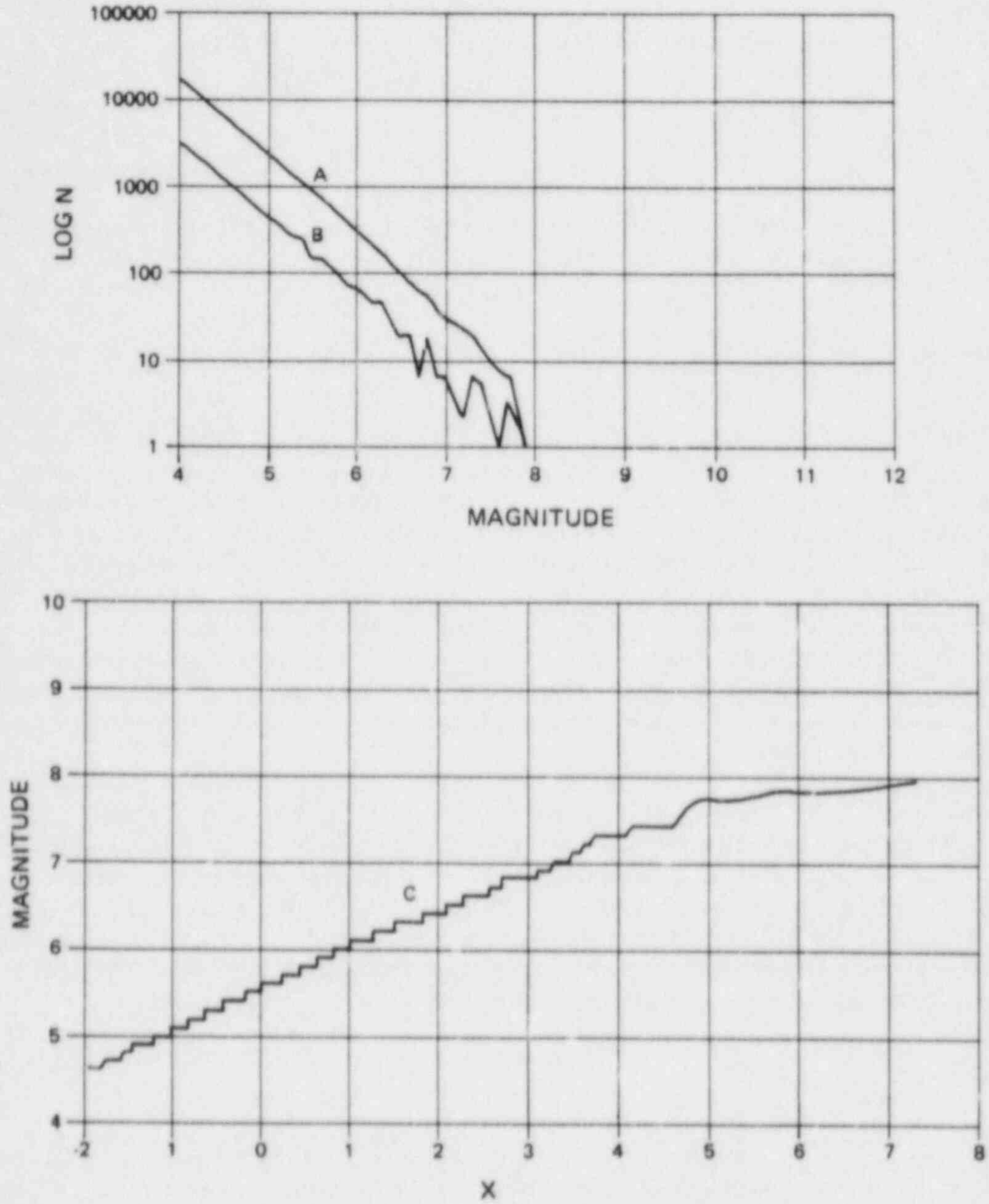


FIGURE 5-1 TYPICAL PLOT OF SYNTHETIC CATALOG DISTRIBUTIONS. THESE CURVES ARE FOR POISSON GENERATED SHOCKS, WESTERN SEISMICITY RATE, NO AFTERSHOCKS ADDED, AND MAXIMUM MAGNITUDE 8.
(A) CUMULATIVE LEAST SQUARES
(B) DIFFERENTIAL LEAST SQUARES
(C) EXTREME VALUE PROBABILITIES

over the 15 realizations and compare with that for the 750 year chained catalog. We also compute the recurrence time T7 for an earthquake of magnitude 7.0 or greater from (1) using the values of (a,b) as determined from the fit (even though we may have privileged information that the catalog has $M_{max} < 7.0$). We report the average value of the recurrence time T7 and its standard deviation. These latter standard deviations are computed in two ways: from (1), using the standard deviations of a and b already determined (these are reported as a percentage of the estimate of T7 and as a standard deviation in years), and from the standard deviation of the 15 values of T7.

5.1.2.2 Maximum Likelihood

We use the maximum likelihood method to fit the histogram of part A. The maximum likelihood estimate for a Poisson distribution involves the maximization of the function

$$(a+\log_{10}(b\Delta M))\sum_k N_k - b\sum_k N_k M_k - \sum_k (b\Delta M)10^{a-bM_k}$$

with respect to (a,b). In the above formula N is the number of earthquakes with magnitude M_k in the kth magnitude band of width M = 0.1. We report quantities a, b, T7 and error estimates as above.

5.1.2.3 Cumulative Least Squares

We construct the logarithm of the cumulative distribution from the distribution of part A. We fit this distribution by the usual least squares according to (1). We report quantities a, b, T7 and error estimates as above.

5.1.2.4 Extreme-Value Statistics

For each 50 year catalog (or 750 years in the chained cases) we identify the largest annual earthquakes and rank order them by magnitude. We fit the function

$$\frac{-\ln(-\ln R_n)}{\ln 10} = a - bM_n$$

by least squares where

$$R_n = \frac{n-1/2}{50}$$

and n is the nth rank order integer, with n = 50 corresponding to the largest values of M_{max} and n - 1 corresponding to the least value of M_{max}. We tabulate a, b, T7 and as before.

In many cases of the Eastern synthetic catalog, no earthquake occurred during a given year. There were two choices: 1) take an interval longer than 1 year in order to identify at least one earthquake per interval, or 2) assume that there were unobservable earthquakes with $M < M_{co} = 4.0$ in the years with no event reported, rank order the annual events observed, normalized to 50 years as above, and then fit by least squares to the part of the curve that could be observed. We have chosen the second option since the first choice requires an a posteriori decision regarding sampling interval, and it can be shown that for a Poisson distribution, there will always be some a priori chosen interval for which no earthquake with $M > M_{co}$ will be observed for a sufficiently long catalog, thus requiring an a posteriori readjustment of interval.

5.1.3 Aftershocks

In one view, natural aftershocks represent a population of earthquakes that is superimposed on "main shocks". Furthermore they represent an image of the past rather than the future, by virtue of their reference to earlier seismic activity. Thus if estimates of future seismic risk are to be derived from earthquake catalogs, it is more appropriate to consider the statistical analysis of seismic catalogs with aftershocks removed. It is the view of Knopoff that this position is not a valid one, and that aftershocks are an organic part of the complete process of seismicity and gives information about the future as well as the past. But, in the interest of evaluating a popular proposal, we construct a suitable test.

We separate aftershocks from a main shock catalog by a simple windowing procedure described by Knopoff and Gardner (1972). We apply the same window to the Eastern catalogs. The window we have used here is reproduced in Table 5-1. The procedure is as follows: For a given shock in the catalog with magnitude M, a search is made to identify subsequent shocks which occurred within the time and distance intervals T(M), L(M) of the given shock. Shocks occurring within this window are labelled aftershocks. This filter is applied to all of the shocks in the catalog in turn.

The aftershock filter window was applied to all three types of catalogs, including the first and the third which are free of calculated aftershocks, since they are Poissonian in character. (A spatial assignment to all shocks in all catalogs is made, based on a length 400 kilometers, as indicated above). The number of events identified as aftershocks in the 18 750-year chained catalogs is given in Table 5-2.

Table 5-1

PARAMETERS DEFINING AFTERSHOCK WINDOWS

Magnitude	Distance (Km)	Time (years)
3.5	26	0.0603
4.0	30	0.1151
4.5	35	0.2274
5.0	40	0.4247
5.5	47	0.7945
6.0	54	1.3973
6.5	61	2.1644
7.0	70	2.5068
7.5	81	2.6301
8.0	94	2.6986
8.5	109	2.7671
9.0	124	2.8356

Table 5-2

MAIN EVENT AND AFTERSHOCK IDENTIFICATION

Region	Catalog	M_{max}	Main Events Created	After-shocks Created	Main Events Identi-fied, I	Aftershocks Identi-fied, I	Main Events Identi-fied, II	Aftershocks Identified, II
Western	Poisson Independent	6	17642	0	6780	10862	15353	2289
		8	17526	0	5531	11995	14546	1980
		12	17558	0	5375	12271	14561	3085
	Poisson + Aftershocks	6	17494	398	6874	11018	15514	2378
		8	17693	3473	5391	15775	15454	5712
		12	17442	3922	5430	15934	15181	6183
	Strain Probabilistic	6	17558	0	6824	10734	15501	2057
		8	17633	0	5215	12418	14657	2976
		12	9410	0	4887	4523	8472	938
Eastern	Poisson Independent	6	1770	0	1593	177	1748	22
		8	1748	0	1544	204	1716	32
		12	1853	0	1646	207	1820	33
	Poisson + Aftershocks	6	1772	39	1615	196	1782	29
		8	1800	526	1568	758	1900	426
		12	1824	374	1608	590	1909	289
	Strain Probabilistic	6	1774	0	1659	115	1764	10
		8	1639	0	1459	180	1609	30
		12	828	0	783	45	822	6

Case I: Main shocks distributed randomly on a line
Case II: Main shocks distributed randomly on a rectangle

The average number of main events for 750-year catalogs of the first two types should be about 17,500 for the Western catalogs and about 1,750 for the Eastern catalogs, since an average of 23.44 (or 2.344) earthquakes per year is the expectation; the computer-generated catalogs have yielded close to the expected values. What is more remarkable is the large number of events identified as aftershocks in the Western Catalogs and the correspondingly smaller number of aftershocks identified in the Eastern catalogs. We interpret this to mean that our effort to project the equivalent of the Southern California catalog onto a line produces such a high density of events on the line that a very large number of main shocks are misidentified as aftershocks, mainly because they fall close to others in space and time. In the case of the Eastern catalog, the density is still low and main shocks are far enough from neighbors to avoid being labelled as aftershocks. We have tried a second window that identifies aftershocks that has a spatial span of 1/2 of that of the first window which operates on the same catalog as before. The window effectively reduces the number of aftershocks that are identified, but surprisingly the number of aftershocks in the Eastern catalog is more distinctly reduced than in the number in the Western catalog.

For the above reasons we have generated a second version of the spatial distribution of main shocks. In this second version, main shocks are no longer distributed along a line but are instead randomly distributed within a rectangle of dimensions 360 km x 480 km. Aftershocks were distributed as before on a north-south line through the epicenter. The aftershock location routine was then applied to the new catalogs as a search within an area surrounding the epicenter, including aftershocks, rather than on a line, as in the first case. Both the distribution on a line and on a rectangle are unsatisfactory approximations to a real and therefore non-uniformly distributed set of earthquake faults in a region, but the number of shocks misidentified is considerably smaller in the second case than in the first.

In one respect these investigations testify to the incompatibility of the aftershock generating and aftershock identification programs. But that is inevitable since the aftershock generating program of Nature is also inconsistent with the best endeavors of seismologists to construct identification programs. Despite this comment, we remark that, by purely random processes, some independent events are likely to be identified as aftershocks simply because they have been located near one another. In Table 5-2 we can obtain rough measures of the relative importance of these two competing ingredients in the analysis. For example in the case of the Western catalog with $M_{max} = 8$, a catalog with no aftershocks generated originally, 2980/17526 or 17% of the "main events" were identified as aftershocks, if the events were distributed in a rectangle. If we added 3473 aftershocks to the main shock catalog, the aftershock identification program identified 5712 aftershocks, which is roughly 2700 events more than the 2980 mainshocks originally identified as aftershocks. Hence, in this case at least, the aftershock distribution routine does a pretty good job of identifying aftershocks; in addition, it identifies a sizeable number of main shocks as aftershocks, due to the random process of determining locations.

We have created an additional set of catalogs that is designed to test some features of the statistical analysis packages. In this new set, the first two catalogs of main shocks (Poissonian without and with aftershocks) for the West are all generated with the same set of random numbers. The earthquake magnitude in the catalogs for different values of M_{max} are simply one-dimensional maps of one another (a magnitude 5.9 shock in the case M_{max} = 6 might be mapped into a shock of magnitude near 7.8 or 7.9 for the mase M_{max} = 8), but the time and space relationships are preserved. This has significance mainly in regard to the manner of selection of the aftershock window. We do not find significant differences between the two sets of catalogs, but the results are summarized for comparison in the Appendix.

From Table 5-2 we also note that few aftershocks are created by the generation routine with M_{max} = 6. This is because most aftershocks are more than two magnitudes less than the parent shock in our formula (consistent with Bath's observation). With M_o = 4.0, few aftershocks occur in the window $4.0 < M < M_i$ (M_{max} = 6.0) where M_i is the magnitude of the parent. We also observe from Table 5-2 that the number of aftershocks created for the case M_{max} = 12 does not differ significantly from the number for the case M_{max} = 8. (These numbers are 3922 and 3473 respectively for the Western catalogs). We can explain this seeming inconsistency with the above "law" (that the number of aftershocks above the magnitude cutoff threshold of 4.0 should increase as the magnitudes of the largest shocks increase), by the remark that even with M_{max} = 12, the likelihood of generating a magnitude 10 earthquake, for example, is extremely small in a 750 year interval. Indeed the numbers and magnitudes of events in the case M_{max} = 12 are only slightly different from those with M_{max} = 8.

5.1.4 Three-Parameter Fits

The method of generation of the catalog yields numbers of earthquakes in each 0.1 magnitude interval band that are independent of M_{max}. We have assumed that the main shocks are generated according to the log frequency-magnitude relation (1) at magnitudes less than M_{max}, and are uninfluenced by the value of M_{max}. Thus any procedure for data analysis that describes differential least-squares or the maximum likelihood procedures cannot yield information regarding M_{max}, except we can note that there are no earthquakes greater than a certain magnitude in the catalog.

On the other hand, cumulative distributions are strongly biased, as are the methods of the theory of extremes, which depend on rank ordering procedures. Both these procedures generate monotonic distributions, and in these cases there is zero probability that an ordinate will be less than that of its nearest neighbor in the direction of decreasing monotonicity. Purists among statisticians would prefer to eschew consideration of biased distributions. Nevertheless, the magnetism of trying to get "something for nothing" by studying the curvature of cumulative distributions to determine M_{max} is popular enough that we should make an effort to assess these procedures. It is obvious that we display our intuitive indication of disdain for these procedures, by these preforatory remarks: the curvature of the two distributions is generated by the method of data processing and is not an intrinsic property to be unearthed by data analysis. Our preference is to determine M_{max} by looking at the catalogs and identifying the largest earthquakes.

In the statistical analysis thus far of the two procedures that concern us here, we have assumed that $M_{max} = \infty$. (Whether or not we have made the assumption in the unbiased cases is irrelevant, since this assumption does not enter into the data analysis). Hence we have, up to this point, a two parameter fit involving the parameters a, b, having fixed the third parameter M_{max} at infinity. In an attempt to determine M_{max} we must consider the three parameter case. In general, the larger the number of parameters, the better the fit and in these cases, as is to be expected, the quality of the fit is improved. But we pay a penalty for this improvement, as we discuss below in our assessment of the results of the data analysis.

The expressions we have used in the cases of fitting with three parameters are:

$$\log\dot{N} = a - bM + \log(1 - 10^{-b(M_{max}-M)})$$

and

$$M = M_{max} - \left\{ \frac{bM_{max}-a}{a} \right\} \exp\left\{ - \frac{x}{\ln 10\,(bM_{max}-a)} \right\}$$

in the cumulative and extreme-value cases respectively. In the first case we minimize

$$\sum_i \left\{ \log \dot{N}_{th}(M_i) - \log \dot{N}_{obs}(M_i) \right\}^2$$

with respect to (a, b, M_{max}) with N_{th} given by the above formula, and in the second case we minimize

$$\sum \left\{ M_{th}(x_i) - M_{obs}(x_i) \right\}^2$$

with M_{th} given by the above formula. The quantity x is of course the double logarithm of the rank ordering parameter.

5.1.5 Appendices

This section describes the computer plots and listings that comprise the appendices to the original report prepared by Leon Knopoff. The computer outputs are not reproduced as part of this report.

In the appendices we present the summaries of all the results we have obtained. In Appendix I, we provide the differential, cumulative, and extreme-value distributions for all eighteen 750-year catalogs. In the cases of the differential distributions, some raggedness is evident at the large-magnitude end of the distributions, as is to be expected when dealing with small numbers. The cumulative curves provide appropriately smoothed versions of these distributions.

In Appendix II we provide complete sets of computer output for five cases of analysis as follows:

Appendix IIa. The main shocks have been distributed over the 360 km x 480 km rectangle. Twelve different analysis blocks are displayed.

1.	Western Poisson independent	no aftershocks removed
2.	Western Poisson plus aftershocks	no aftershocks removed
3.	Western Strain probabilistic	no aftershocks removed
4.	Eastern Poisson independent	no aftershocks removed
5.	Eastern Poisson plus aftershocks	no aftershocks removed
6.	Eastern Strain probabilistic	no aftershocks removed
7.	Western Poisson independent	aftershocks removed
8.	Western Poisson plus aftershocks	aftershocks removed
9.	Western Strain probabilistic	aftershocks removed
10.	Eastern Poisson independent	aftershocks removed
11.	Eastern Poisson plus aftershocks	aftershocks removed
12.	Eastern Strain probabilistic	aftershocks removed

Analysis by four different methods for each of the catalogs is performed for three different values of M_{max}. Each method of analysis uses only a two parameter fit, i.e. in the analysis we assume M_{max} is infinite, even though it is not infinite in the catalog.

Near the top of each block we list the average values of a and b and their standard deviations for the 15 50-year catalogs. This is followed by the average return time for a magnitude 7 earthquake (in years) for the 15 catalogs as determined from the a and b values above it, the percentage standard deviation and the absolute standard deviation (in years) of the return time for a magnitude 7 earthquake as determined from the standard deviations of the a and b values and finally the standard deviation of this return time as determined from the 15 values of the return time.

Finally, at the bottom of each block, we list the a and b values of the chained catalog and the return time for a magnitude 7 earthquake as determined from these a and b values. The "correct" values from equation (1) are listed at the top of each block of output.

Appendix IIb. We complete the analysis of Appendix IIa for the two cases of fit to the catalogs by three parameter systems. The parameters are a, b, and M_{max}. These are applicable as discussed above to the cases of the cumulative distributions and the theory of extremes. In many cases of the three-parameter analysis, both the cumulative and the extreme value curves have the "wrong" curvature, i.e. they were concave upward. In these cases the least squares searches "blew up", i.e. division by zero occurred. We have listed as an additional entry in these tables the number of cases for which successful solutions (out of 15 possible) that were obtained, i.e. for which no blow up occurred. In some cases of analysis, the distributions for the chained 750 year catalogs also had reversed curvature. These cases are identified by blank entries.

Appendix IIc. This is a repetition of Appendix IIa, except that the main shocks are distributed along a straight line as described above. For the first six blocks of each analysis, the results are identical with those of Appendix IIa, since no aftershocks removal procedures were used in these cases.

Appendix IId. Same as Appendix IIc, except the same random number initializes all catalog sequences, while in all the other appendices, different random numbers are used throughout. This case is included to separate the intrinsic effects of differences among the catalogs and the data processing techniques from those of the Monte Carlo procedures we have used. (The Monte Carlo effects are very small).

Appendix IIe. Same as Appendix IIc, except the aftershock identification routines use windows that are 1/2 of those in case IIc.

Appendices c, d, and e are included for completeness. However of the two alternatives for locating aftershocks, we prefer that which distributes them on a rectangle to that which distributes them on a line. In the discussion we refer to Appendices IIa and b.

5.1.6 Discussion

Before offering more speculative interpretations of the results of Appendices IIa and b, we make the following more concrete observations.

For the more abundantly endowed Western catalogs, the least squares fits to the cumulative distributions and the extreme-value distributions, not unexpectedly, improve monotonically (approach the theoretical value of a=4.77, b=0.85) as M_{max} increases for all three catalog types. For the Eastern catalogs, there are discrepancies in the results of these two methods for M_{max}=6, but trends for the cases M_{max}=8 and 12 are not identifiable. The methods of least squares fit to the cumulative distributions and the extreme-value distributions are strongly biased for small values of M_{max} because of the curvature of these distributions: a linear fit to these curved distributions give high values of both a and b. As the curvature decreases with increasing M_{max}, the results of the fit become more appropriate.

As we expected, the quality of the fit in these two cases improves as we go to three parameter fits. The number of instances of "blowup" due to reversed curvature increases as M_{max} increases. Generally, the estimate of M_{max} from the three parameters fits are best for catalogs with M_{max}=6; it is in these cases that the curvature of the distribution is most pronounced and hence most easily identified in curve-fitting. For the catalogs with larger values of M_{max}, the distributions are often almost linear (see Appendix I) (or have reversed curvature) and it becomes extremely difficult to identify M_{max} from the curve-fitting. Because of the large number of "blowups", we refrain from further discussion of the cases of three parameter fit represented by Appendix II-b, except to note that estimates of M_{max} from the output are better done in all cases by directly noting the largest magnitude in the catalog, rather than by curve fitting.

The mere coincidence of an average of 15 values of a, b, or T7 with the "theoretical" values is not of itself an identification of merit since, in reality, a processing of a single 50-year catalog could give values of these three quantities widely divergent from expectation if the standard deviations are large. Conversely, a small standard deviation in a, b, T7 does not recommend a method for use if the mean of the analysis is far from the theoretical values. To evaluate the relative "performance" in each of the analyses, we construct the following ad hoc criterion:

We list as "acceptable" those procedures whose mean estimates of both a and b lie within two standard deviations of the theoretical values. We then identify, from among these acceptable analyses, the method which has the least of the standard deviations. Table 5-3 gives this listing. The end result of this analysis is that

1) If we do not remove aftershocks from the catalog, maximum likelihood analysis is the preferred method in all cases, with only one clear-cut exception (catalog W, P+A, M_{max}=8). This exception seems to be due to the statistical quirk (of random

numbers) of having given very small values of standard deviations for a and b, i.e. all catalogs generated were similar to one another. The standard deviation for this case is less than for either M_{max}=6 or 12 for the same catalogs and is therefore slightly suspect.

2) If we remove aftershocks, the conclusion is the same, namely, maximum likelihood methods are preferred over all others for the Eastern catalogs. However for the Western catalogs there is some ambiguity in choice between maximum likelihood and cumulative least square methods.

   This last conclusion implies that there is perhaps some intermediate procedure that might be generated that falls between the two sets of weights: maximum likelihood methods give large weight to the data which are most abundant, namely to the smallest earthquakes, while cumulative distributions give greater weight to the larger earthquakes (but not as much as in the case of the theory of extremes).

3) The theory of extremes gives consistently the least reliable results of statistical analysis for these catalogs of finite length.

Table 5-3

COMPARISON OF METHOD PERFORMANCES

Region	Catalog	After-shocks Removed	M_{max}	Acceptable[1]	Least σ	Remarks
W	PI	No	6	DM	M	
			8	DMCE	M	
			12	MCE	M	
	P+A	No	6	DM	M	
			8	CE	C	
			12	MCE	M	
	SP	No	6	DM	M	
			8	DMCE	M	
			12	MCE	M	
	PI	Yes	6	DM	M	{M barely better than D}
			8	CE	C	
			12	CE	C	
	P+A	Yes	6	DM	M	
			8	CE	C	
			12	MCE	M	
	SP	Yes	6	DM	M	
			8	CE	C	
			12	CE	C	
E	PI	No	6	MC	M	{M barely better than C}
			8	MCE	M	
			12	MCE	MC	About equal
	P+A	No	6	DMC	M	
			8	MCE	M	
			12	MCE	M	
	SP	No	6	DMC	D	{D barely better than M}
			8	MCE	M	
			12	MCE	M	

Table 5-3 (Continued)

Region	Catalog	After-shocks Removed	M_{max}	Acceptable[1]		Least σ	Remarks
	PI	Yes	6	MC		MC	About equal
			8	MCE		M	
			12	MCE		MC	About equal
	P+A	Yes	6	DMC		M	
			8	MCE		M	
			12	MCE		M	
	SP	Yes	6	DMC		D	{D barely better than M}
			8	MCE		M	
			12	MCE		M	
W	PI	No	6	DM	C_x	M	
			8	DMCE,	C_x, E_x	M	
			12	MCE,	E_x	M	
	P+A	No	6	DM,	C_x	M	
			8	CE,	C_x, E_x	C_x	
			12	MCE,	C_x, E_x	M	
	SP	No	6	DM,	C_x	M	
			8	DMCE,	C_x, E_x	M	
			12	MCE,	C_x, E_x	M	
	PI	Yes	6	DM,	C_x	M	{M barely better than D or C_x}
			8	CE,	C_x, E_x	C_x	
			12	CE,		C	
	P+A	Yes	6	DM,	C_x	M	
			8	CE,		C	
			12	MCE,	C_x, E_x	M	
	SP	Yes	6	DM,	C_x	M	
			8	CE,	C_x, E_x	C	
			12	CE,		C	
E	PI	No	6	MC,	C_x, E_x	M	{M barely better than C}
			8	MCE,	E_x	M	
			12	MCE,	E_x	MC	About equal

Table 5-3 (Continued)

Region	Catalog	After-shocks Removed	M_{max}	Acceptable[1]		Least σ	Remarks
	P+A	No	6	DMC,	C_x, E_x	M	
			8	MCE,	E_x	M	
			12	MCE,	C_x, E_x	M	
	SP	No	6	DMC,	C_x, E_x	D	{D barely better than M}
			8	MCE,	C_x, E_x	M	
			12	MCE,	E_x	M	
	PI	Yes	6	MC,	C_x, E_x	MC	About equal
			8	MCE,	E_x	M	
			12	MCE,	E_x	MC	About equal
	P+A	Yes	6	DMC,	C_x, E_x	M	
			8	MCE,	E_x	M	
			12	MCE,	C_x, E_x	M	
	SP	Yes	6	DMC,	C_x, E_x	D	{D barely better than M}
			8	MCE,	C_x, E_x	M	
			12	MCE		M	

[1] D = Diff. L.S.
M = Max. Like.
C = Cum. L.S.
E = Extreme

5.2 Synthetic Catalog Testing, Robin McGuire

The Poisson distribution for earthquake occurrence times and the log-linear distribution for magnitudes have been shown empirically (Earth Technology, 1981) to provide reasonably good fits to observed seismicity data. However, the natural processes generating earthquakes clearly include some spatial/temporal causality and memory. Earthquake recurrence methodologies that are based on the Poisson and log-linear distributions might be expected to perform reasonably well when tested against synthetic catalogs generated using Poisson and log-linear distributions. The recurrence methodologies should also be tested against synthetic catalogs whose generating proces is at least one level more complex than the processes implied by the methodologies. A more complex process may be a better model for natural processes.

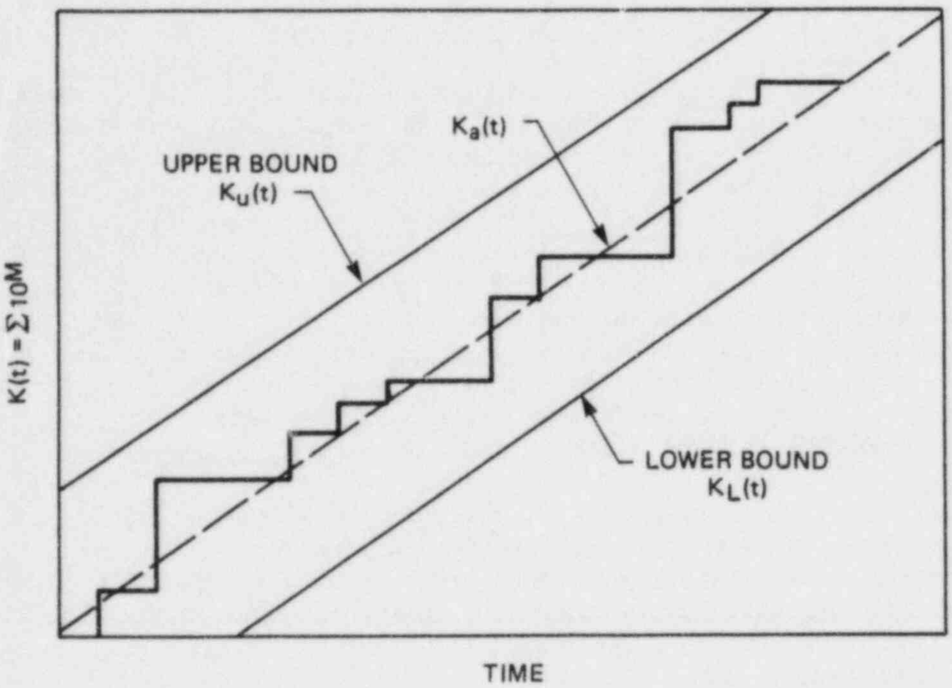
This section describes ranking support studies designed by Robin McGuire. The efforts include catalog generation, an analysis program for several Poisson-Log Linear methods (POISS), an analysis program for Extreme Value Statistics (EXTRM), an analysis program for the Bayesian Method (BAYES), and evaluation of the results for both eastern and western seismicity rates.

When the Poisson-Log Linear, Extreme Value, and Bayesian Statistics methodologies were tested against the McGuire catalogs, none gave consistently good estimates for the long-term catalog properties. The estimated values for large-magnitude earthquake return periods and the recurrence curve slope varied considerably. The catalog value plus-or-minus one sigma contained the estimates only in about 30 percent of the cases. When the local statistics in the catalog were examined, the statistical fits were often found to be quite good. In effect, the short-term statistics, equivalent to 200 years of eastern seismicity data, were often different from the catalog's long-term statistics. This situation may be quite representative of real earthquake processes in many areas having low to moderate seismicity. Several examples have been reported for which observed seismicity has shown significant temporal changes (Earth Technology, 1981).

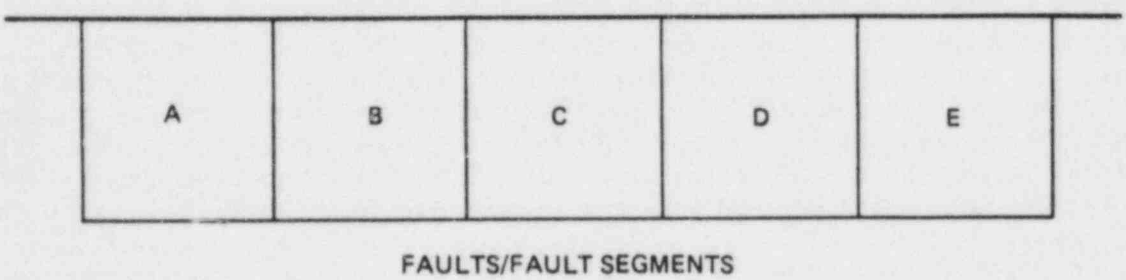
5.2.1 Synthetic Catalogs

The earthquake generation process is modeled using five separate source zones, or these may be five adjacent segments along a major fault as shown in Figure 5-2b. K(t) is a process parameter for each source zone and is related to cumulative energy release as a function of time.

$$K(t) = \sum_{i=1}^{j} 10^{M_i}$$



(a)



(b)

FIGURE 5-2 SCHEMATIC EARTHQUAKE GENERATION PROCESS
a. CONCEPTUAL RELATION BETWEEN PARAMETER K AND TIME.
b. ADJACENT FAULT SEGMENTS.

where M_i is the magnitude for the ith earthquake and j is the most recent earthquake before time t. The earthquake process for a source can be shown schematically by the stairstep line in Figure 5-2a. Vertical bars are proportional to earthquake size and horizontal bars are quiescent times between earthquakes. The average rate of energy release, and the average rate of energy input to the source, is indicated by the slope of the line $K_a(t)$.

As time passes after the previous earthquake, the horizontal lines move toward the lower bound $K_l(t)$ and the probability for an earthquake increases. This lower bound can be analogous to a limit set by the rupture strength of a fault. If an earthquake occurs its magnitude is shown by the vertical bar. The maximum magnitude for each shock is limited by an upper bound $K_u(t)$ and can be analogous to a limit set by the fault area available for rupture. The maximum earthquake for the source is the difference between the lower and upper bounds.

The interoccurrence times are generated from the exponential distribution function

$$f(t) = A \cdot Re^{-Rt}$$

where A is a constant, R is the average occurrence rate, and t is the interoccurrence time. The corresponding cumulative distribution function is

$$F(t) = A - Ae^{-Rt}$$

from which

$$A = \frac{1}{1 - e^{-Rt_{max}}}$$

because $F(t) \to 1$ as $t \to \infty$. Also $A \to 1$ as $t_{max} \to \infty$. The complementary cumulative distribution function is $G(t) = 1 - F(t)$. Taking logarithms of G(t), solving for t, and letting $A \to 1$ gives

$$t = - \frac{\ln G(t)}{R} \quad .$$

Because $0 < G(t) < 1$, an exponential distribution of t-values for interoccurrence times can be found using a random number generator to supply a uniform distribution from 0 to 1. The interoccurrence times are dependent on the occurrence rate R.

The magnitude distribution is generated similarly starting with the exponential distribution for magnitude M.

$$f(m) = A'\beta e^{-\beta(M-M_o)}$$

where A' is a constant, β is the recurrence slope (β = b ln 10 and log N = a - bm), M_o is the lower magnitude cutoff. Proceeding as above,

$$A' = \frac{1}{1 - e^{-\beta(M-M_o)}}$$

$$G(M) = 1 - F(M)$$

and $$M = \frac{\ln A' - \ln(G(M) - 1 + A)}{\beta}$$

Because $0 < G(M) < 1$, the magnitude distribution can be generated using another random number generator to give a uniform distribution between 0 and 1. The magnitudes will depend on M_{max}, M_o, and the slope β.

Dependence between the five seismic sources and temporal memory are introduced into the model by adjusting the rate R and the slope β. A parameter K(t) is kept for each seismic source; $K_A(t)$, $K_B(t)$, etc. A ΔK is defined for each source and expresses whether its cumulative energy release is ahead or behind that of the neighboring sources. For example,

$$\Delta K_C = K_C - \frac{K_B + K_D}{2} \quad .$$

ΔK_C implies an energy difference equal to some magnitude M by

$$M = \log | \Delta k_c | \ .$$

The rate R is modified by a factor F, $R_{new} = R_{old} \cdot F$, where

$$F = 1 - \text{sign}(\Delta K_C) \cdot \frac{| M-M_o |}{| M_{max}-M_o |} \quad .$$

Similarly, slope β is modified by F, $\beta_{new} = \beta_{old} \cdot F$. Thus, if seismic source C is behind the average of its neighbors, the source is encouraged to catch up. Increasing the rate R raises the probability for an earthquake, and increasing the slope β raises the probability for a larger shock. If a source is behind its neighbors, the converse occurs. Using the neighbors inserts spatial dependence; using the K parameter inserts temporal dependence. The end sources have only one neighbor considered. The factor F is arbitrarily bounded so changes can not happen too fast and lead to instability in the process.

Catalog generation follows the sequence: 1) predict occurrence times for each seismic source, 2) choose the source with the earliest time, 3) calculate a magnitude for the chosen source, 4) update the source's K(t), R, and β, and 5) return to step 1.

Synthetic catalogs for the five seismic sources were generated for a 10,000-year time duration. The process parameters were chosen so the catalogs produce recurrence curves with slope about b = 0.85 (β = 1.96) and rate about 0.75 earthquakes with M $\geq$4 per year. Statistics for these master catalogs are given in Table 5-4, and recurrence curves are shown in Figure 5-4. A typical segment of the time sequence, Figure 5-3, shows clustering and lower activity levels after large shocks, mucn like histograms of natural catalogs. The synthetic catalogs were time scaled to represent the Eastern data; 0.5 catalog years equals 1 data year for a rate of about 0.37 earthquakes M $\geq$4.0 per year. For the Western data, 30 catalog years equals 1 data year for a rate of about 22 earthquakes M $\geq$4.0 per year. Earthquake rates and return periods for these two data sets are given in Tables 5-5 and 5-6 respectively.

TABLE 5-4

MASTER CATALOG STATISTICS, 10,000 YEARS

Interval Data

Magnitude	4-4.5	4.5-5	5-5.5	5.5-6	6-6.5	6.5-7	7-7.5	7.5-8	8-8.5	>8.5
Source No. 1	4718	1411	517	194	56	39	17	8	0	1
Source No. 2	5248	1729	628	284	91	63	11	5	2	0
Source No. 3	5191	1635	559	248	96	48	18	6	2	0
Source No. 4	5114	1616	619	231	114	40	13	3	3	0
Source No. 5	5058	1532	518	217	62	35	20	4	3	0
Sum	25329	7923	2841	1174	419	225	79	26	10	1

Cumulative Data

Magnitude	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5
Source No. 1	6961	2243	832	315	121	65	26	9	1	1
Source No. 2	8061	2813	1084	456	172	81	18	7	2	0
Source No. 3	7803	2612	977	418	170	74	26	8	2	0
Source No. 4	7753	2639	1023	404	173	59	19	6	3	0
Source No. 5	7449	2391	859	341	124	62	27	7	3	0
Sum	38027	12698	4775	1934	760	341	116	37	11	1

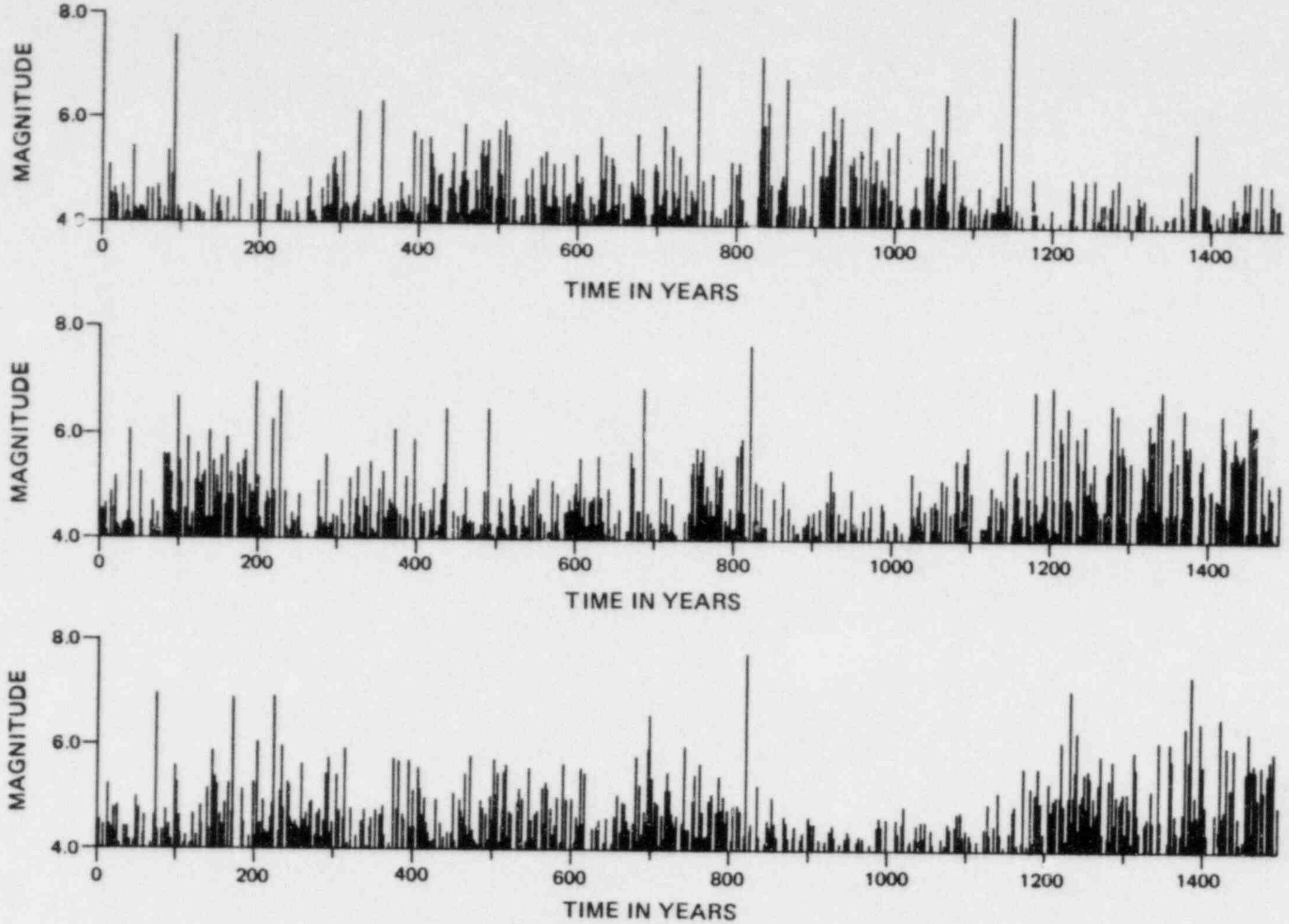


FIGURE 5-3 EARTHQUAKE HISTORY. EARTHQUAKE OCCURRENCES FOR THREE FAULT SEGMENTS. RATE OF ACTIVITY IS APPROXIMATELY 0.75 SHOCKS GREATER THAN MAGNITUDE 4 PER YEAR OVER THE LONG TERM FOR EACH SOURCE.

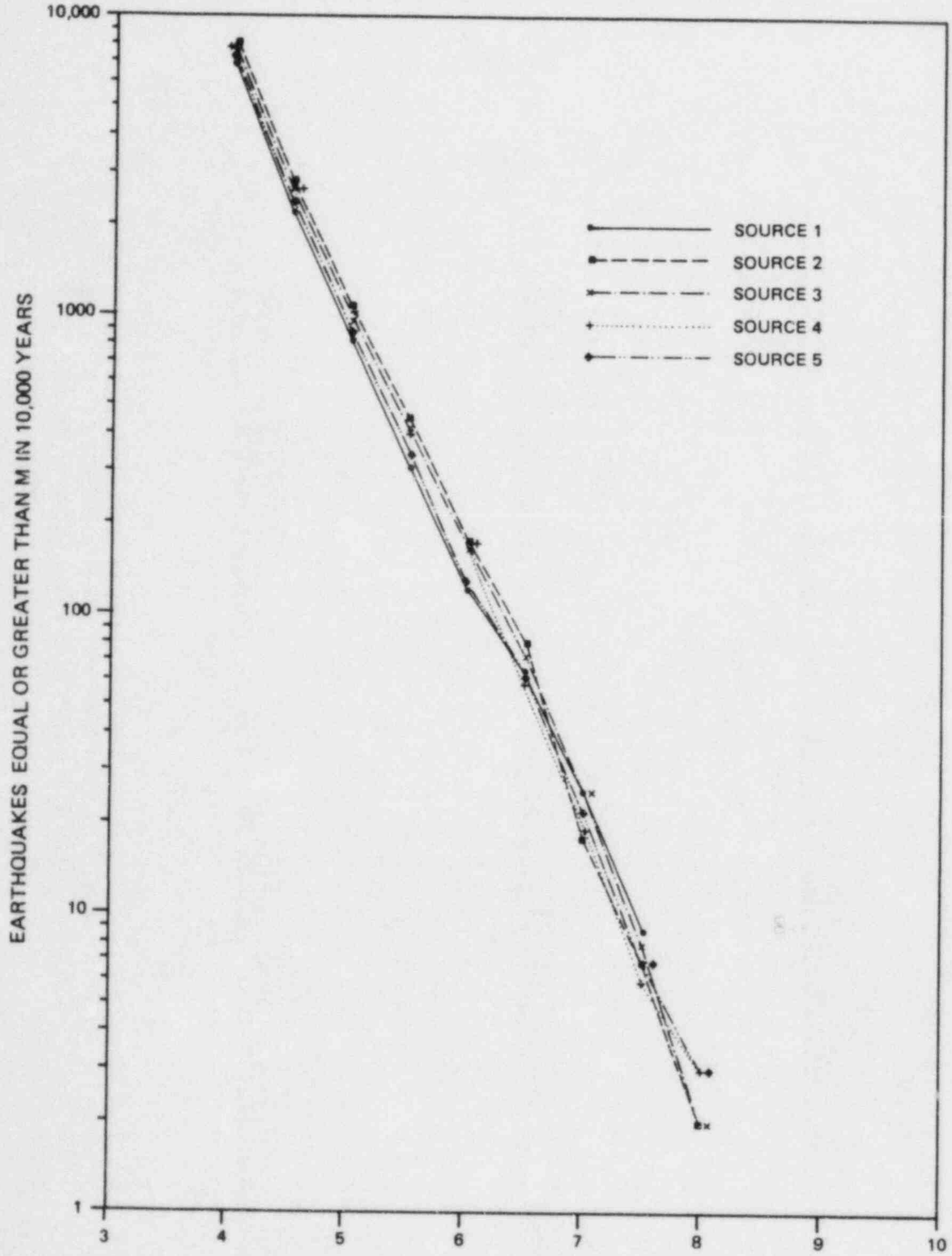


FIGURE 5-4 SYNTHETIC CATALOG RECURRENCE CURVES

Table 5-5

CATALOG RATES, "EASTERN"

0.38 EQ M≥4/Year

MAG	NO. EARTHQUAKES	AVE	RETURN PERIODS	AVE
≥4.0	3.48 E-1 - 4.03 E-1	3.80 E-1	2.48 - 2.87 Yrs.	2.6 Yrs.
≥5.0	4.15 E-2 - 5.42 E-2	4.78 E-2	18.5 - 24.1	20.9
≥6.0	6.05 E-3 - 8.65 E-3	7.60 E-3	116 - 165	132
≥7.0	9.00 E-4 - 1.35 E-3	1.16 E-3	741 - 1,111	862
≥8.0	5.00 E-5 - 1.50 E-4	1.10 E-4	6,667 - 20,000	9,091

Table 5-6

CATALOG RATES, "WESTERN"

20 EQ M≥4/Year

MAG	NO. EARTHQUAKES	AVE	RETURN PERIODS	AVE
≥6.5	59 - 81	68.2	4.6 - 6.3 Yrs	5.4 Yrs
≥7.0	18 - 27	23.2	13.7 - 20.6	16.0
≥7.5	6 - 9	7.4	41.2 - 61.7	50.1
≥8.0	1 - 3	2.2	123.5 - 370.4	168.4

5.2.2 Program POISS

The program POISS was written to analyze synthetic catalog data according to the Cumulative Least Squares, Differential Least Squares, and Maximum Likelihood methods. The analyses are done both with and without considering an upper bound M_{max} for the magnitude distributions. A generalized program flowchart is shown in Figure 5-5.

In the least squares fitting for both cumulative and interval statistics, the parameters a and b in the usual earthquake magnitude distribution, log N = a-bM, were found by standard linear regression. The maximum likelihood estimate for β was made using a technique that permits input data from unequal observation periods for different magnitudes (Weichart, 1980).

The scheme for introducing a probabilistic upper bound M_{max} uses the following equation for annual probability G.

$$G = p(M \mid M_x) = R \left[1 - \frac{1 - e^{-\beta(\bar{M}-M_o)}}{1 - e^{-\beta(M_x-M_o)}} \right] \left[1 - \frac{1 - e^{-\beta(M-\bar{M})}}{1 - e^{-\beta(M_x-\bar{M})}} \right]$$

where

> $p(M \mid M_x)$ is the probability for magnitude M given that the maximum magnitude is M_x,
>
> R is the rate for magnitudes $M \geq M_o$,
> M_o is the threshold magnitude,
> M is the average magnitude, and
> β is the slope parameter.

The first and second bracketed terms are complementary cumulative distribution functions normalized so probability goes to zero at M_{max}. R times the first bracketed term is $\bar{R}$ the rate for the average magnitude $\bar{M}$. $\bar{R}$ times the second bracketed term gives the rate, or annual probability, for magnitude M.

To introduce a probabilistic M_x, let $M_x = M_i$ where i = 1,N and each M_i has probability 1/N, a uniform distribution for the maximum magnitude. Then $G_i = p(M \mid M_i)$ and

$$G = \sum_i G_i \cdot \rho(M_i) .$$

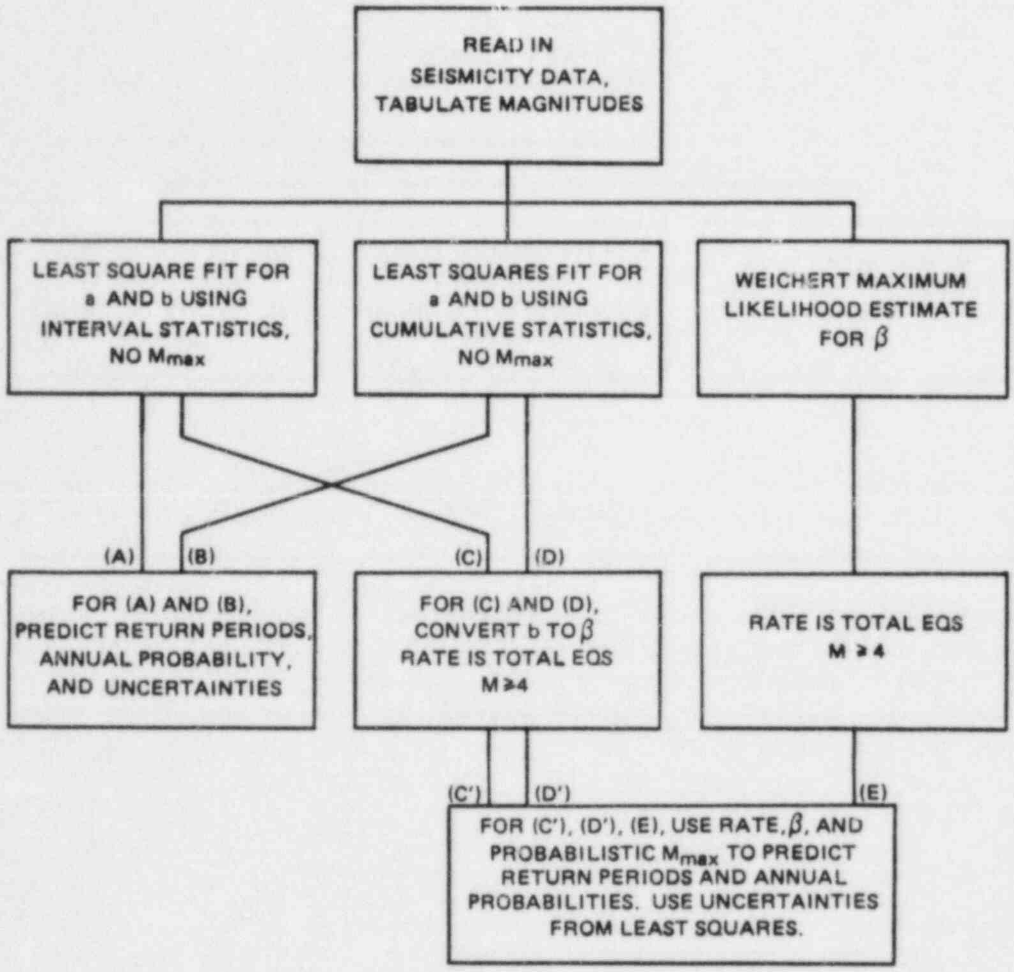


FIGURE 5-5 PROGRAM POISS, GENERALIZED FLOW

If N = 5 and the M_i are 6.0, 6.5, 7.0, 7.5, 8.0, each with probability 0.2, the return period RT computed for M ≥ 6 earthquakes using the probabilistic upper bound is

$$RT_{6.0} = \frac{1}{G} = \frac{1}{(G_1 + G_2 + G_3 + G_4 + G_5) \cdot (0.2)} .$$

Similarly for M ≥ 7.5,

$$RT_{7.5} = \frac{1}{G} = \frac{1}{(G_4 + G_5) \cdot (0.2)} .$$

The probabilistic upper bound has the effect of increasing the return period length for larger magnitudes. Other distributions for M_x could be easily implemented.

Uncertainties for the return period estimates are estimated using the linear regression analysis given by Benjamin and Cornell (1980, Section 4.3.2). The parameters a and b are found by linear regression on the form log N = a-bM. If $(\log N)_k$ is a predicted value for log N using M_k, a and b, the varience σ^2 of $(\log N)_k$ is

$$\sigma^2_{(\log N)_k} = \frac{\sigma^2}{n} \left(1 + \frac{(M_k-\bar{M})^2)}{S^2_M}\right),$$

where

$$\sigma^2 = \frac{1}{n-2} \sum_i (\log N_i - (a + bM_i))^2 ,$$

N is the number of data points,

$\bar{M}$ is the mean value for the M_i's, and

$$S^2_M = \frac{1}{N} \sum_i (M_i-\bar{M})^2.$$

The varience in $(\log N)_k$ depends on M which can be changed arbitrarily by selecting the coordinate origin for M. A least squares linear fit to data points always passes through the point ($\overline{\log N}$, $\bar{M}$). Heuristically, the uncertainty in log N can be considered to have two components: 1) the uncertainty in b causes the slope of the fitted line to increase or decrease, and the uncertainty becomes larger at points farther away from the pivot at the point ($\overline{\log N}$, $\bar{M}$), and 2) the uncertainty in a causes the fitted line to shift up and down. For the form log N = a-bM, a is being estimated at M=0. The fitted line must pass through ($\overline{\log N}$, $\bar{M}$), so uncertainty in b causes some uncertainty in a. The uncertainty in a then is coupled to uncertainty in b. However, if M' = M-$\bar{M}$ and log N = a-bM' is used for regression, the estimate for a is made at M'=0, or M-M, and there is no longer any coupling between the uncertainties in a and b. Program POISS uses the form log N = a-bM' a=bM' so the uncertainty estimates are not biased.

Separate uncertainties were not estimated for the cases using probabilistic upper bounds. Instead, the uncertainties for the cumulative and differential least squares fits were just passed through to the new return period estimates.

The uncertainty estimates are taken as one standard deviation, the square root of the varience. The term log N then becomes $\log N \pm \sigma$. When antilogarithms are taken to get the rate, or annual probability, the uncertainty becomes a multiply/divide factor.

5.2.3 Program EXTRM

The program EXTRM analyzes synthetic catalog data using extreme value statistics. As shown in Section 3.2, analysis of extreme values leads to an equation:

$$-\ln(-\ln G(M)) = \ln\alpha - \beta M$$

in which the ln-ln term, the left hand side, has a linear relation to magnitude M. M is the largest magnitude annually (or some other convenient time period). G(M) is a cumulative probability assigned to each magnitude M. These probabilities are assigned by monotonically ranking the largest earthquakes from each of N successive years. The largest earthquake is first, or rank R=1; the smallest earthquake is last or R=N. Cumulative probabilities are assigned using a plotting rule

$$G(M) = \frac{R - 1/2}{N} \quad .$$

The program compiles a set of data points $x = -\ln(-\ln G(M))$ and $y = M$, which can be fitted with a straight line by least squares linear regression to find the parameters $\ln\alpha$ and β.

For Case A, the G(M) term is regressed on M. The resulting estimates for $\ln\alpha$ and β are converted to a and b in the usual log-linear equation, $\log N = a + bM$, and return periods are estimated for various M. For case B, M is regressed on the G(M) term, and annual probabilities G(M) are used directly to estimate return periods. A generalized program flowchart for EXTRM is shown in Figure 5-6.

Uncertainties in $\ln\alpha$ and β are given by the linear regression. Corresponding uncertainties in a and b result from simple conversion. Uncertainties in the predicted return periods are not simple conversions, but follow a general formulation such as that by Meyer (1975, Chapter 10). If $z = f(x,y)$ and the σ_x and σ_y are uncorrelated,

$$\sigma_z^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 .$$

For the linear regression in EXTRM, the σ's for $\ln\alpha$ and β are not strictly uncorrelated, because the coordinate origin is not at the mean values for the variables. However, the lack of correlation is judged to be approximately true. Uncertainties were also estimated using $\beta + \sigma_\beta$ and $\beta - \sigma_\beta$ in all computations following the linear regression. The two methods for estimating the uncertainties give comparable results for the data analyzed.

A typical plot with the data points and the linear regression lines for a Western data set is shown in Figure 5-7. Usually, there is very little difference in the straight lines fitted by regressing M on the G(M) term, or the G(M) term on M. A similar plot for Eastern data is shown in Figure 5-8. For Eastern data, many years did not have a maximum earthquake equal or above the 4.0 magnitude threshold. Data points for such years were suppressed, and analysis was carried out on the remaining subset of points under the assumption the smaller magnitude years would fit the same trend.

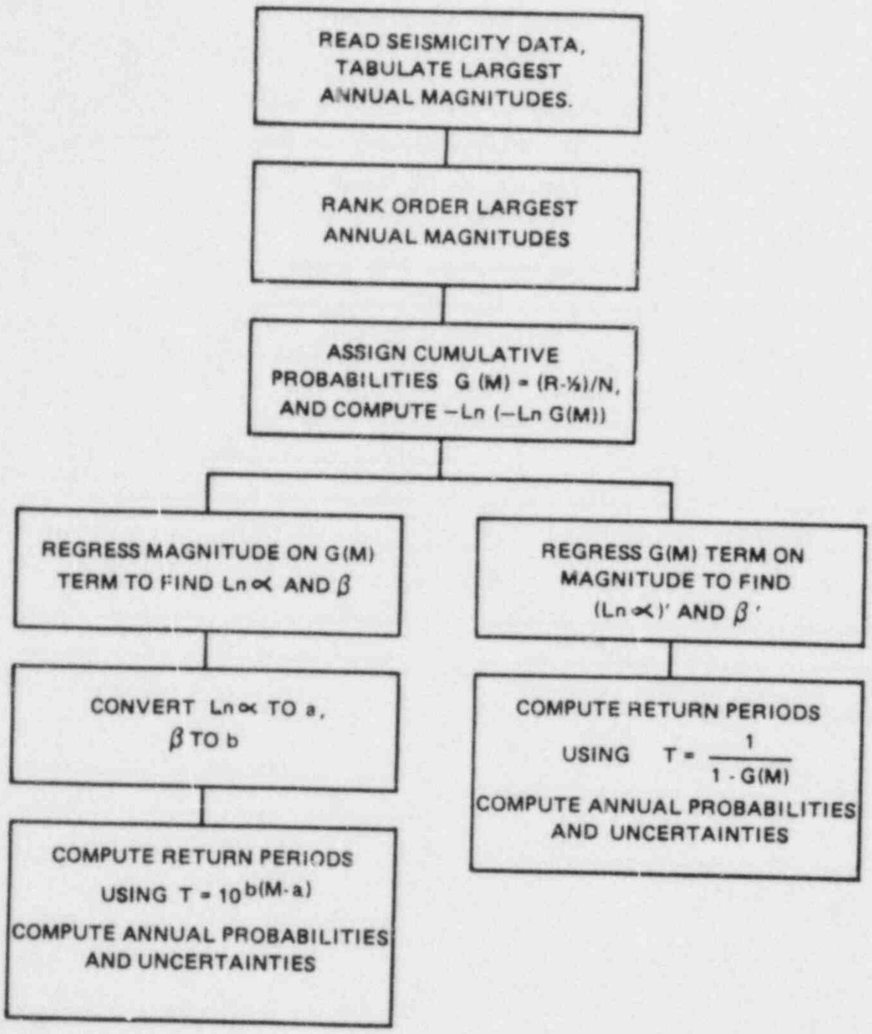


FIGURE 5-6 PROGRAM EXTRM, GENERALIZED FLOW

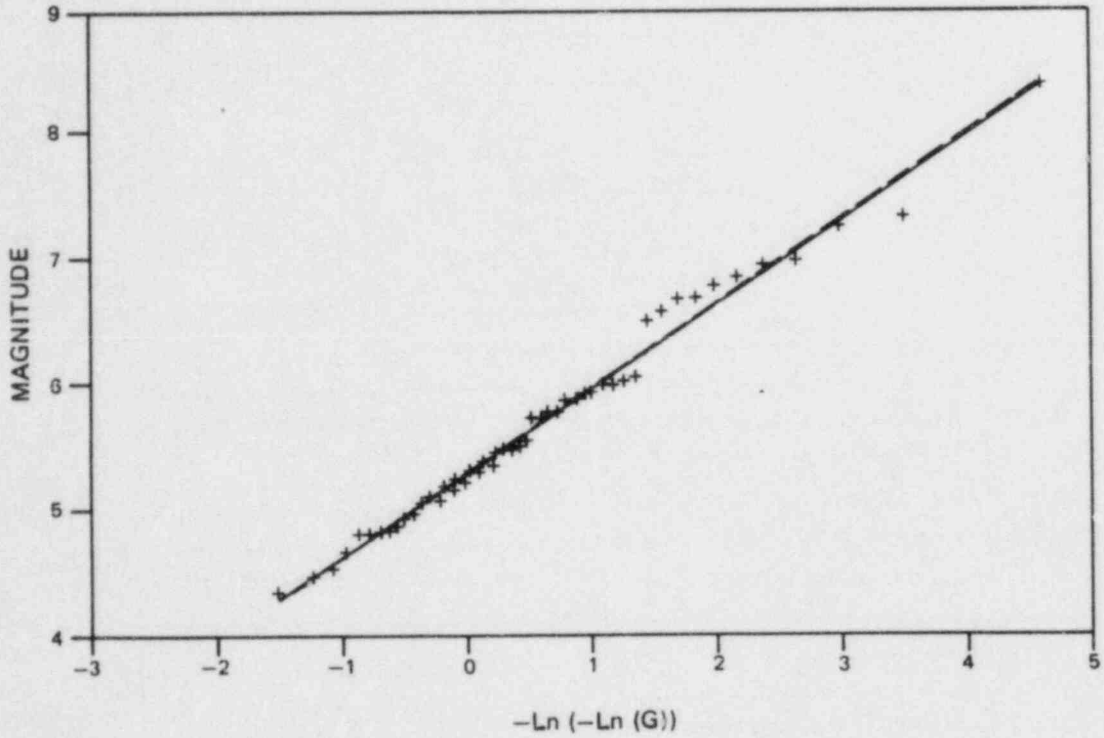


FIGURE 5-7 EXTREME VALUE PLOT FOR WESTERN SEISMICITY RATE. G IS PROBABILITY FOR EACH EXTREME VALUE. SOLID LINE IS LEAST SQUARES REGRESSION FOR MAGNITUDE ON G. DASHED LINE IS LEAST SQUARES REGRESSION FOR G ON MAGNITUDE.

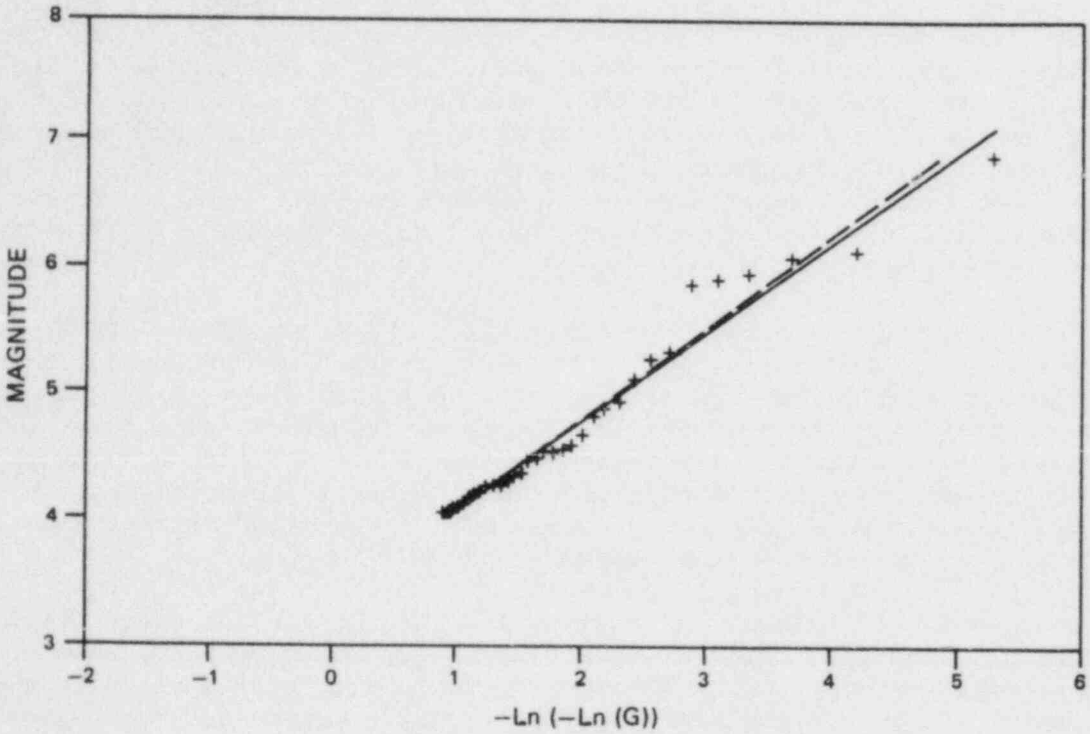


FIGURE 5-8 EXTREME VALUE PLOT FOR EASTERN SEISMICITY RATE. CURVES AS IN FIGURE 5.2-6

5.2.4 Program BAYES

Program BAYES computes Bayesian posterior probability for earthquake occurrence using the formulation proposed by Mortgat and Shah (1978). The Mortgat and Shah equation 2.22 gives the probability $p(r_{Mi}=0)$ that no earthquakes in magnitude bin M_i occur. If a one-year time period is used, annual probability for earthquakes with magnitudes in bin M_i follows from $1-p(r_{Mi}=o)$. Figure 5-9 shows equation 2.22 and defines the variables. An equivalent form given in equation 2.24 with factorials rather than gamma functions could be used in BAYES because the variables are all integers in this application. Evaluating either gamma functions or factorials for large arguments can be a problem. The gamma function form was chosen and the gamma functions evaluated using a double precision polynomial approximation, equation 6.1.41 from Abramowitz and Stegun (1968).

As the synthetic catalogs were generated, five separate earthquake histories were maintained, one for each seismic source. In this Bayesian application, one seismicity catalog represents a region around a power plant site and produces the sample distribution. The remaining four seismicity catalogs represent regions that could be selected by experts as analogous tectonic regimes and the basis for prior distributions. Catalogs A, B, D, and F were used for the prior distributions. Catalog C was used for the sample distribution.

The basic concept in Bayesian statistics is that the prior distribution is multiplied by the sample distribution to produce a posterior distribution (see Section 3.1). A weakness in the method, when applied to the statistics of rare earthquakes, is the critial need for non-zero values in the sample distribution. Philosophically, a zero value for the number of earthquakes in some magnitude bin M_i should lead to a zero estimate for the posterior probability. Equation 2.22 combines the prior and sample data by addition (see definitions for r'', ξ, λ'' and ν'' in Figure 5-9). Only the term ξ'' is affected by a zero-value R_{M_i} and the equation yields an estimate for the posterior probability. In accord with the basic Bayesian concept, program BAYES arbitrarily sets the posterior probability equal to zero whenever R_{M_i} is zero.

No uncertainties on the predicted posterior probabilities are defined in the method.

$$p(r_{mi}=o) = \sum_{n=o}^{\infty}\left[\frac{1}{n!}\cdot\frac{\Gamma(\eta'')}{\Gamma(\eta''-\xi'')}\cdot\frac{\Gamma(n+\eta''-\xi'')}{\Gamma(n+\eta'')}\cdot\frac{\Gamma(n+\nu'')}{\Gamma(\nu'')}\cdot\frac{t^n \lambda''^{\nu''}}{(t+\lambda'')^{n+\nu''}}\right]$$

$$\eta'' = \eta' + N \qquad \lambda'' = \lambda' + T$$

$$\xi'' = \xi' + R_{mi} \qquad \nu'' = \nu' + N$$

SAMPLE DATA (LIKELIHOOD FN)

N = NUMBER OF EQS

R_{mi} = NUMBER OF EQS IN MAGNITUDE BIN M_i

T = TIME BASE DATA

PRIOR DATA

RATE p(n):

λ' = TIME BASE (EQUIVALENT TIME)

ν' = NUMBER OF OCCURENCES (EQUIVALENT NO.)

MAG p(ri/n):

η' = NUMBER OF OCCURRENCES (EQUIVALENT NO.)

ξ' = NUMBER OF EQS IN MAGNITUDE BIN M_i (EQUIVALENT NO.)

FIGURE 5-9 MORTGAT AND SHAH EQUATION 2.22

5.3 Goodness-Of-Fit Tests, Nancy Mann

The model most often applied for analysis of earthquake magnitudes is the exponential distribution resulting from an assumption of Poisson distributed events. Such an assumption, in conjunction with the parameters estimated for the exponential model, leads to the result that there must be a finite upper bound for the magnitude of earthquakes. (See Newmark and Rosenblueth, 1971).

Knopoff and Kagan (1977) demonstrate this same result empirically by contradiction, with California data, showing that an assumption of an exponential model leads to California earthquakes of magnitudes 8.0 or greater in a 40 year period with probability 0.992 and earthquakes of 9.0 or greater with probability about 0.5, if there is no upper bound.

One can also use methods described in Mann, Schafer, Singpurwalla (1974, Section 5.2.4(b)) with California data to demonstrate analytically a result similar to that demonstrated empirically by Knopoff and Kagan. For example, if, as they assume, the exponential scale parameter Θ is 0.85 for quakes having magnitude λ = 4.0 or greater, then the expected magnitude of the largest of N = 800 earthquakes (assuming an average number of 20 per year with magnitude 4.0 or greater) is

$$X_{(N)} = \lambda + \Theta \sum_{i=1}^{N} (1/(N-i+1))$$

$$= 4.0 + 0.85 \sum_{i=1}^{800} (1/(800-i+1))$$

$$\cong 4.0 + 0.85\,(\ln(800) + 0.577215\epsilon) \cong 10.17$$

where ln(x) indcates the natural logarithm, and 0.5772156 approximates Euler's constant. For 10,000 years (i.e. 200,000 quakes), the expected largest magnitude for this exponential model is 14.86.

Also, one can use an F approximation in conjunction with these assumptions to obtain an approximate 99 percent lower confidence bound of 7.85 for $X_{(800)}$, the largest magnitude in 800 successive quakes. This is similar to the value of 8.00 obtained by Knopoff and Kagan. The derivation of the F approximation that gives this prediction is given in Section 5.3.2 under Parameter Estimation and Prediction.

Thus, if the exonential model is appropriate, it appears that there is an upper bound, or threshold. Alternative models are: 1) a different distribution for earthquake magnitudes, one with a right tail that is shorter than the exponential tail, a distribution with at least one mode, or 2) a model suggested by Knopoff and Kagan (1977) that is exponential with the scale parameter decreasing with magnitude, "so that the energy integral remains finite."

This suggests that it is difficult to approach the problem of analyzing earthquake magnitudes directly for the following reasons.

a) The problem of lack of sensitivity of instrumentation for lower magnitudes is not important if a classical exponential model applies, because of the lack-of-memory property for exponential data. Thus, one can simply include in the analysis earthquakes of magnitude larger than 4.0, say, transform the data by subtracting 4.0 from each magnitude, and treat the resulting data like ordinary one-parameter exponentially distributed observations ranging from zero to infinity. If, however, the data are truncated somewhere in the upper ranges and the point of truncation is not known or if the scale parameter is changing as a function of magnitude, the analysis becomes much more complicated. It is not at all clear what might be happening in the upper tail of the distribution, even though the lower part of the distribution has been fitted by some investigators with the standard exponential model.

b) If the exponential model is completely inapplicable and another distribution, or mixture of distributions, applies, then the lower tail also becomes a problem. This results because the exponential distribution alone exhibits the lack-of-memory property. The instrumentation problem dictates that the population is truncated on the left at different magnitudes, depending upon the time and place of collection of the data. If one chooses to measure earthquakes of magnitude greater than 4.0, the fact that one does not know the number of earthquakes ignored by doing so is the source of this problem for nonexponential models. Clearly, not ignoring known earthquakes of smaller magnitude does not solve the problem since there still remain many, perhaps hundreds of quakes for a single catalog, that are not detected.

c) In addition to the upper- and lower-tail problems, there is the problem of aftershocks. Does one include them or not, and if not, how does not define an aftershock?

Extreme-value analysis can circumvent all three of these problems. If the time interval of interest is long enough, then the smallest of the largest earthquakes occurring in that time interval will be no less than about 4.0. Furthermore, the problem of aftershocks is circumvented if the interval is long enough. The longer the length of the interval of interest, the more the largest earthquakes for the various intervals will tend to be independent. On the other hand, the smaller the interval, the more data one has to analyze. Thus, there exists a need in choosing an interval length to optimize so as to get sufficient data without including magnitudes that are too small or introducing dependency. Once a suitable interval length has been chosen, the data themselves can tell us whether or not there is an upper bound. In order to make any sort of predictions for the future, we assume a steady-state condition. Without such an assumption we cannot be justified in making any inferences concerning future seismic events.

One point to consider is whether or not extreme-value theory is applicable since the aftershocks, particularly, indicate a lack of independence of events. A theorem proved by Barlow and Singpurwalla (1974) applies here. If the events within each interval are either independent, or only "associated," (i.e. positively correlated), then for a very large number of events, extreme-value theory is applicable. Since the number of events within each interval is extremely large, the majority being unmeasurable, use of extreme-value theory appears to be justified. Graphical and Fourier analysis lend corroboration, as will be shown in the sequel.

It should be pointed out that in using the extreme-value analysis we are concerned with predicting largest earthquakes in a specified period of time, not simply with estimating parameters. Thus, an exponential mean Θ can be estimated more efficiently by using all individual measurements of magnitude, rather than in using yearly extremes alone; that is, the variance in estimating the mean is smaller when all the data are used. In estimating the largest magnitude $X_{(N)}$ for this exponential model in a large number N of time intervals, however, the variance of the estimate is $\Theta^2\Pi^2/6$ whether one uses all the data or only the extreme magnitudes applying to the various intervals. The smaller sample size for the extreme-magnitude data is offset by the fact that the largest observation in a future sample from this population is not so far out on the distribution tail as is the same observation with respect to the distribution corresponding to the population of all observable earthquake magnitudes.

5.3.1 Analysis Method

Analysis of a California earthquake catalog representing the last forty-eight years demonstrates a consistency of patterns for one-month, three-month, six-month and yearly extreme values. The samples based on three- and six-month extremes are devoid of the small values appearing in the samples of one-month extremes, are informative in terms of amount of data and are well behaved (stable).

The preliminary analysis was done graphically using both an interactive computer program called Grafstat (see Tarter, 1978) and a noninteractive computer program for plotting largest magnitudes on extreme-value probability paper by means of a Versatec plotter.

The Grafstat program uses a Fourier analysis and gives a graphical representation of an underlying distribution. For magnitudes of both largest monthly and largest semi-annual earthquakes, Grafstat shows for each a bi-modal frequency function representing a mixture of two distributions, one ranging from magnitudes of about 4.0 to perhaps 6.5 and another applying to larger magnitudes. See Figure 5-10.

The behavior of these distributions is such that one might conceive of yet another distribution, or series of distributions, associated with increasing magnitudes, with a last extremely small distribution degenerate at an upper threshold.

Versatic plots of $x_{(i)}$, the ith smallest of the n largest periodic magnitudes versus $-\ln(-\ln \{\frac{i-1/2}{n}\})$ yield straight lines if the magnitudes have type-I distributions of largest extremes. Figure 5-11 shows, in agreement with the Grafstat analysis, what appears to be two straight lines indicating a mixture of at least two distributions for largest earthquakes in six-month intervals.

A type-I extreme-value distribution is consistent with a classical exponential model and, more generally, any single distribution that does not have an upper threshold. The mixture model exhibited by the type-I extreme-value analysis suggets, as one possibility, an asymptotic result somewhat similar to the model suggested by Knopoff and Kagan (1977) -- a distribution scale parameter decreasing (but not continously decreasing) with magnitude, and perhaps the existence of an upper threshold.

Other investigators who have used the extreme-value type-I distribution to analyze largest periodic earthquakes are Epstein and Lomnitz (1966), Rikitake (1974), Schenkova' and Schenk (1975) and Schenkova' and Karnik (1970). In each case the interval has been one year, and the model has been assumed with or without corrections.

The other of three possible extreme-value distributions suggested for analysis of largest earthquakes is the type-III distribution of largest extremes. This distribution applies when there is an upper threshold and has been used for analysis of earthquake magnitude data by Makjanic (1972) and Yegulalp and Kuo (1974). In both cases the interval length used was a year and the number of years considered ranged from 37 to 88. The methods used for estimation of the parameters were the method of

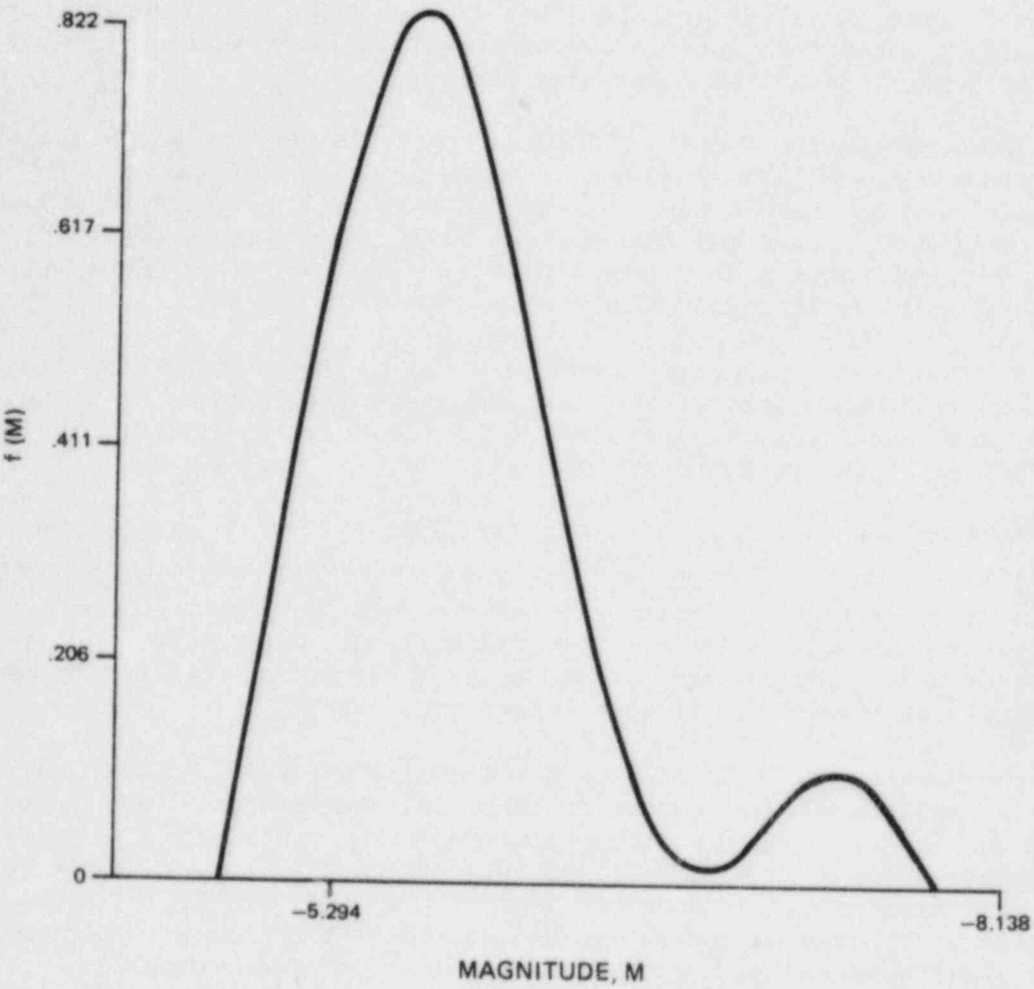


FIGURE 5-10 GRAFSTAT ANALYSIS

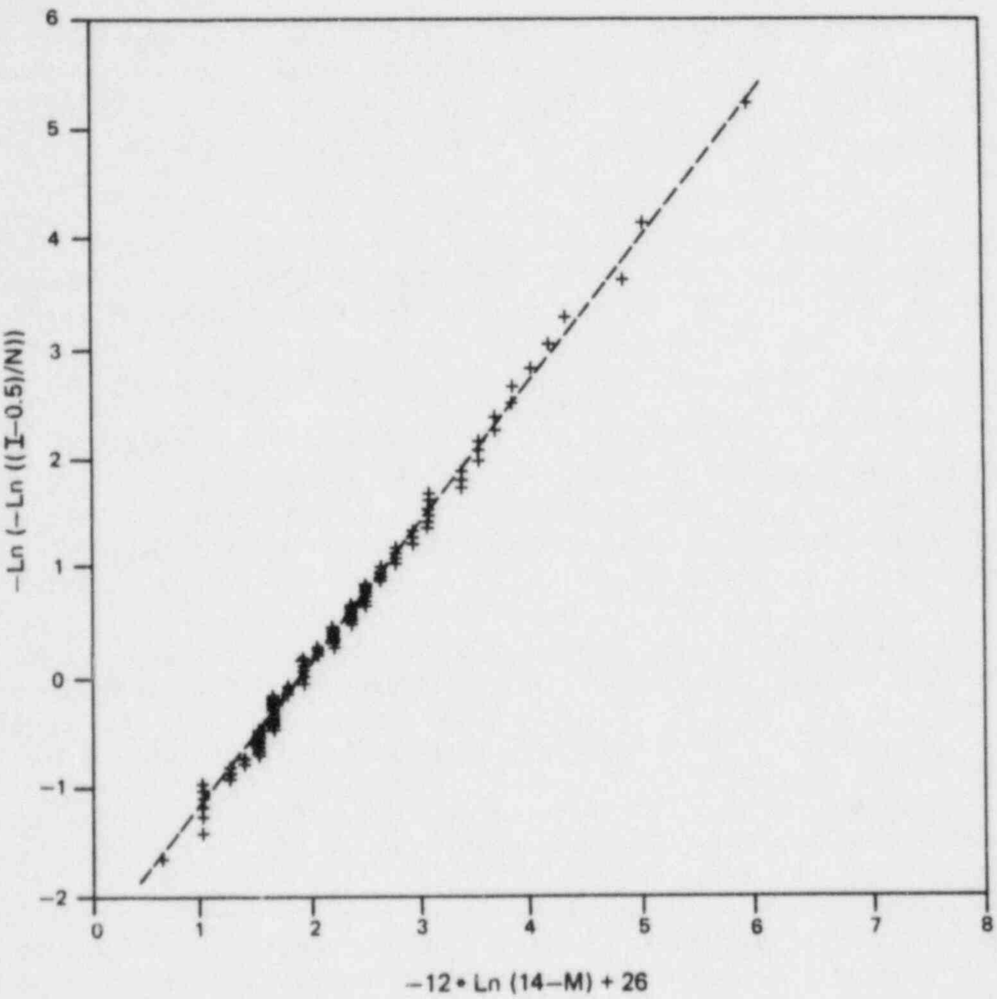


FIGURE 5-11 WEIBULL PROBABILITY PLOT OF MAGNITUDE. CALIFORNIA DATA, SIX-MONTH INTERVALS, TYPE III DISTRIBUTION WITH UPPER LIMIT OF 14.

moments and least squares. Yegulalp and Kuo analyzed 46 regional catalogs and found that only 3 of these could be represented as well by a type-I extreme-value distribution as by a type-III. This provides more evidence for an upper threshold.

If largest earthquake magnitude X has a type-III distribution of largest extremes with upper threshold m_o then $Z = m_o - X$ has a two-parameter Weibull distribution with shape parameter β and scale parameter δ. A straight-line probability plot for this model can be produced by plotting $\ln(m_o - X_{(i)})$ versus $\ln(-\ln\{1\ \frac{i-1/2}{n}\})$, as for a two-parameter Weibull, or by plotting $-\ln(m_o - X_{(i)})$ versus $-\ln(-\ln\{\frac{1-1/2}{n}\})$.

The latter gives the original values from smallest to largest, while in the former the original values are plotted along the abscissa from largest to smallest. The latter plot is the mirror image of the former.

In Figure 5-12, the value 14.0 has been used for m_o, and $-\ln(m_o - x_{(i)})$ is plotted versus $-\ln(-\ln\{\frac{i-1/2}{n}\}$, with $x_{(i)}$ the ith smallest of the six-month magnitudes. (Actually the variable plotted on the abscissa is $28 - 12 \ln(14 - X_{(i)})$. The factor 12 expands the plot so that details of the shape can be observed, and the added term makes the plotted values positive.)

The value 14.0 (actually about 12.0 to 14.0) was determined after several iterations. For values less than about 12, the plotted curve is concave downward, and for values larger than about 14, it is concave upward. Figure 5-12 shows a plot that describes a straight line.

Using $-\ln(14.0 - x)$ for the variable of interest in the Grafstat analysis, we find also a single distribution, rather than the mixture seen in analyzing the untransformed x.

The California Catalog is the one of highest quality available for this investigation. A series of measurements of magnitudes of South American earthquakes was also analyzed, but had the disadvantage that different measures of magnitude were used at different times. Still, the magnitude distribution made up of the various types of measures appeared to exhibit behavior similar to that of California data (see Figures 5-13 and 5-14). The untransformed data plot like a mixture of 3 distributions, and assuming an upper threshold equal to a value (8.8), estimated by analytical procedures, yields a reasonable straight-line plot with a few stray observations at both the lower and upper ends.

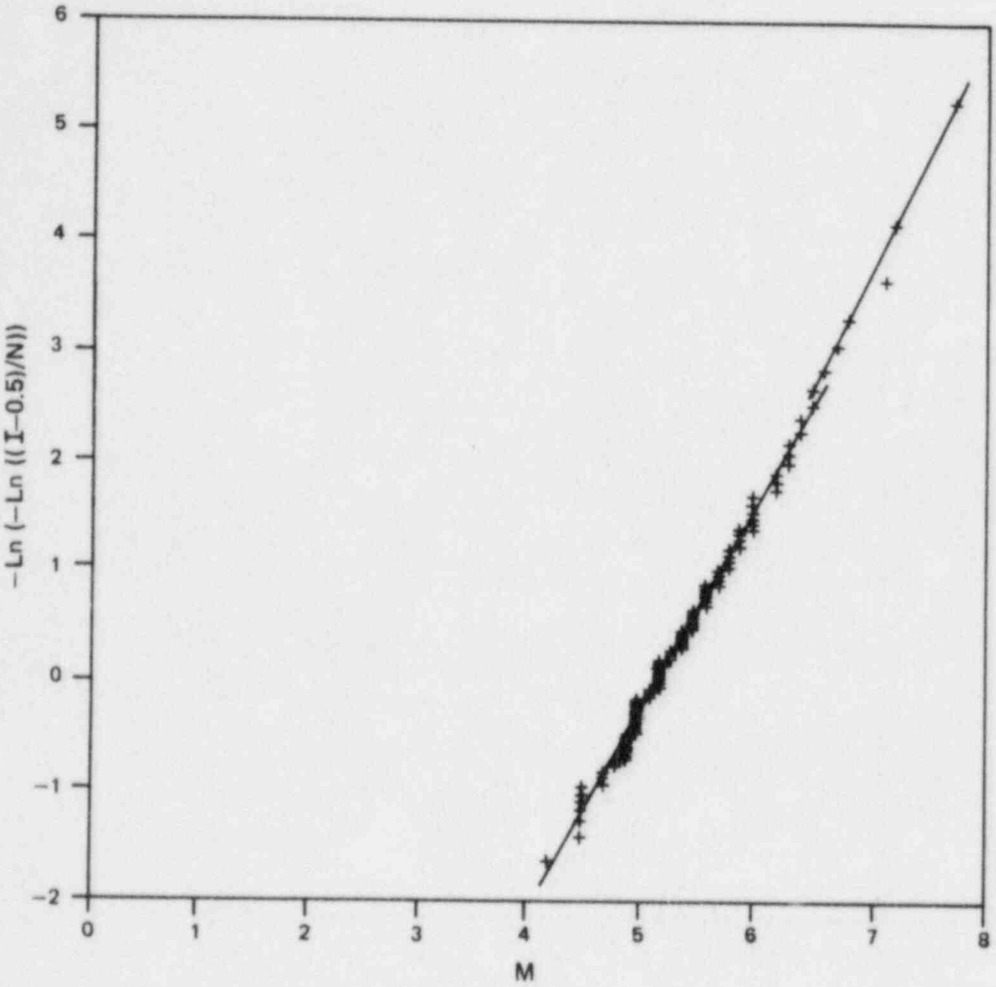


FIGURE 5-12 WEIBULL PROBABILITY PLOT OF MAGNITUDE. CALIFORNIA DATA, SIX-MONTH INTERVALS, TYPE I DISTRIBUTION WITH NO UPPER LIMIT.

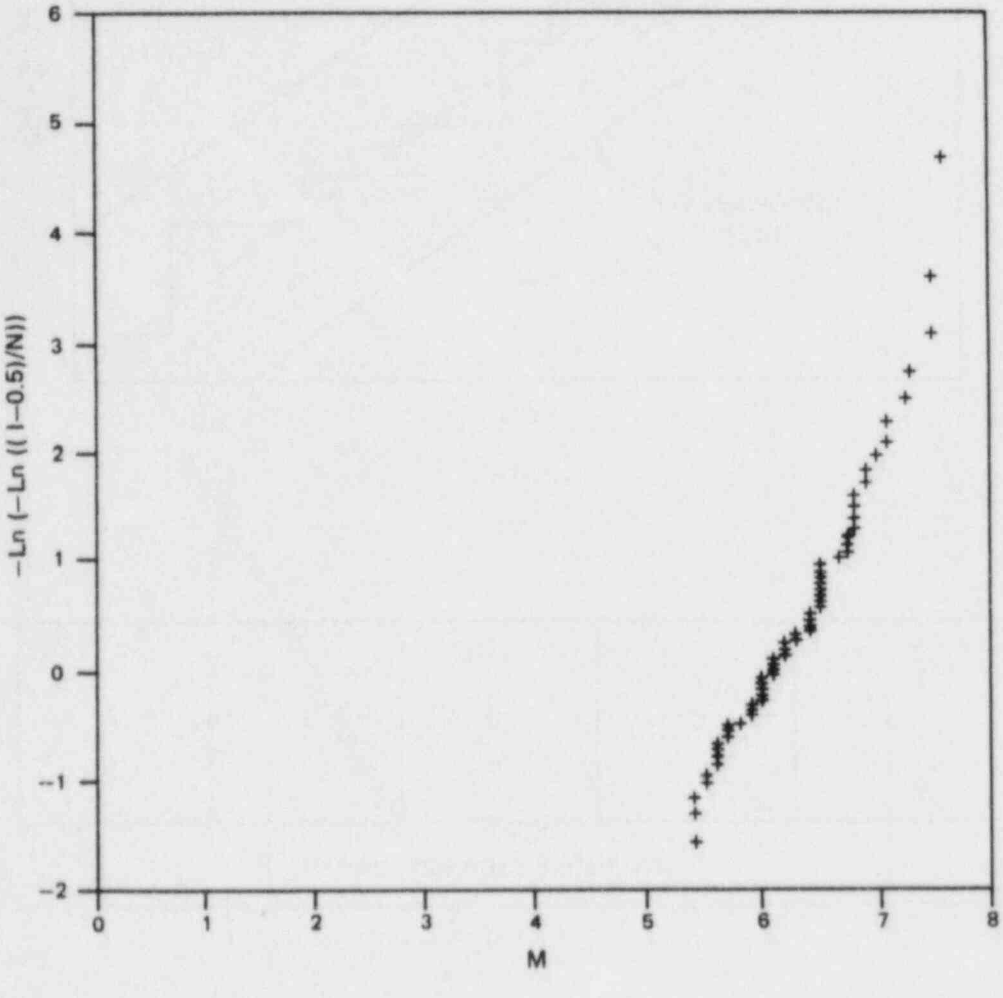


FIGURE 5-13 WEIBULL PROBABILITY PLOT OF MAGNITUDE. SOUTH AMERICA DATA, SIX-MONTH INTERVALS, TYPE I DISTRIBUTION WITH NO UPPER LIMIT.

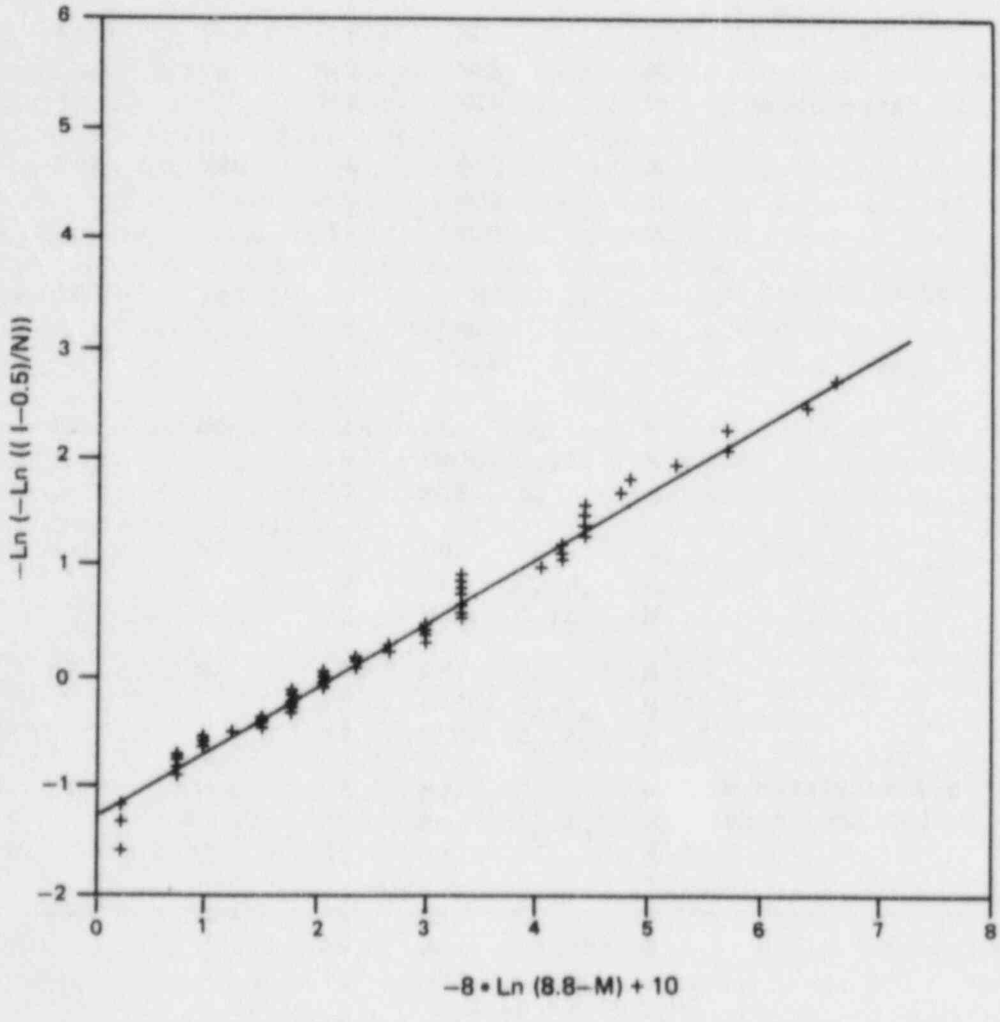


FIGURE 5-14 WEIBULL PROBABILITY PLOT OF MAGNITUDE. SOUTH AMERICA DATA, SIX-MONTH INTERVALS, TYPE III DISTRIBUTION WITH UPPER LIMIT OF 8.8.

5.3.2 Parameter Estimation and Prediction

The three-parameter extreme-value distribution that appears to characterize the samples of extreme magnitudes that have been observed is more difficult with which to deal, in terms of prediction than the obviously inappropriate exonential model often applied to earthquake magnitude data sets. It is, however, not so difficult as a mixture model, which has 5 or more parameters.

One sees in the 3-parameter model involving an upper threshold a very tight distribution around a characteristic magnitude. The tight distribution indicates a very large value for the shape parameter, β, which creates difficulty in the estimation of the upper threshold and hence in the estimation of β as well. This difficulty arises from the fact that an estimate of the threshold is effectively made up of two components; one which is the largest observeable magnitude with variance $\Pi^2/(6\beta^2)$ and another whose variance is divided by $n^{2/\beta}$. For values of β greater than 10 or 50, the variance of estimates of m_o decreases very slowly with increases in sample size.

For the 48-year California catalog, a graphical analysis indicates an upper threshold of about 12.0 to 14.0 for earthquake magnitude and an estimate of β for six-month data (based on the slope of the line plotted with $m_o = 14$) equal to approximately 14.1. A percentile estimate of the characteristic magnitude δ is 5.2.

The "median-unbiased" estimate of Mann and Fertig (1975) for m_o is 13.6 based on the same data set. The procedure from which this value results is based on a ratio of linear combinations of successive differences of logarithms of observed magnitudes ordered from smallest to largest. Each magnitude has an estimated value of m_o subtracted from it. The ratio is set equal to a tabulated value of the fiftieth percentile of the distribution of the random variate, and the value of the estimate of m_o is iteratively modified until the value of the ratio satisfies the equation. For more details, see the section on goodness-of-fit.

Use of the M-F median-unbiased estimate of m_o to obtain estimates of $1/\beta$ and $\ln\delta$ that are linear in the logarithms of the observed magnitudes yields 14.2 for β (for six-month extremes) and 5.3 for δ.

The corresponding moment estimates of m_o, β and δ (see Cohen and Whitten, 1982) are 12.3, 12.5 and 5.2, respectively.

For the California Catalog, there appears to be some discrepancy between the moment estimates of m_o and β and those obtained from the median unbiased estimation procedure for m_o (13.6) and the plot obtained from $m_o = 14.0$. Estimates of largest earthquake magnitudes based on these different estimation procedures do not, however, vary significantly when the number of time intervals is large.

Let us assume that the observable largest magnitude X for six-month intervals has a type-III distribution of largest extremes. Then $Y = -X$ has a 3-parameter Weibull distribution (see Mann and Singpurwalla, 1982) with lower threshold equal to $\lambda = -m_o < 0$, shape parameter $\beta > 0$ and scale parameter $-\delta-(-m_o)>0$. Then, for N intervals,

$$E(Y_{(1)}) = m_o + \frac{-\delta-(-m_o)}{N^{1/\beta}} \Gamma(1 + 1/\beta), \text{ and}$$

$$E(X_N) = -E(Y_{(1)}).$$

If we use the moment estimates with N=96 to predict the magnitude of the largest earthquake in N/2 years, we obtain

$$E\ (Y_{(1)}) = -12.3 + \frac{7.1}{96^{(1/12.5)}} \Gamma(1 + 1/12.5)$$

For $N/2 = 48$ years, $E(X_N) = -E(Y_{(1)}) = 7.6$ for $N/2 = 1000$ and 10,000 years, respectively, then the expected largest magnitudes are 8.6 and 9.3, respectively.

Using the median unbiased estimate of m_o and the associated linear estimates for $1/\beta$ and $\ln \delta$. we get for the expected magnitude of the largest earthquake in 48, 1,000 and 10,000 years, 7.5, 8.9 and 9.6, respectively.

5.3.3 Confidence Bounds

Obtaining a confidence bound for $X_{(N)}$ for N larger than an available sample is usually possible for a location-scale-parameter family. Consider the lower confidence bound on a largest 40-year earthquake under the assumption of the exponential model discussed in the Introduction. We assume a lower imposed threshold of 4.0 and an average of 20 quakes larger than 4.0 per year.

We derive an F approximation similar to those used extensively in Mann, Schafer, and Singpurwalla (1974). It consists of the ratio of two unbiased estimates of the exponential scale parameter θ. In the numerator is an expression involving the unknown $X_{(N)}$, namely

$$(X_{(N)} - \lambda) / \sum_{i=1}^{N} (1/_{N} {}^{-i\ +1}).$$

This variate has expectation $m = \Theta$ and variance

$$n = \Theta^2 \sum_{i-1}^{n} (1/n-i+1)^2 \Big/ \sum_{i=1}^{n} [(1/n-i+1)]^2$$

since the lower threshold is known. For large N, $u \cong \Theta^2\Pi^2/6(\ln N + 0.5772156)^2$, where $\ln(\cdot)$ indicates natural logarithm and 0.5772156 approximates Euler's constant. In the denominator we use the best available estimate of Θ, namely $\hat{\Theta} = 0.85$ with variance Θ^2/n where n is the sample size upon which $\hat{\Theta}$ is based. The two estimates of Θ have zero covariance; and, each, being an unbiased estimate of a scale parameter, has (when multiplied by 2) approximately a chi-square distribution with $2\ m^2/\nu$ degrees of freedom. In circumstances such as these, an F approximation works well when the numbers of degrees of freedom, ν_1, and ν_2 are large. In this case they are

$$\nu_1 = 2\ (\ln N + 0.5772156)^2 / \ (\Pi^2/6)$$

and $\qquad \nu_2 = 2n.$

Thus

$$F = (X_{(N)} - \lambda) \ / \ [\Theta\ (\ln N + 0.5772156)]$$

for N = 800 quakes (20 per year for 40 years) and n = 40 has degrees of freedom 64 and 80. If we assume, as did Knopoff and Kagan, that $\hat{\Theta}$ is given, as if n were infinitely large, then the degrees of freedom are 64 and ∞. This gives $F \cong 1.5$ for a 99 percent lower confidence bound on $X_{(800)}$, and hence a value of 7.85, close to the value of 8.0 obtained by Knopoff and Kagan in their simulation.

It would be nice to be able to derive similar sorts of confidence bounds for $X_{(N)}$, pertaining to the three-parameter extreme-value model, but the shape parameter complicates this.

One heuristic approach that we can use to obtain an approximate confidence bound is to find an expression for a non-parametric confidence bound pertaining to the largest earthquake magnitude in N/2 years in terms of the unknown parameters and then substitute estimated values of the parameters as if they were known. Consider $Y_1 = -X_N$, the smallest Weibull order statistic. The distribution of Y_1 is given by

$$1-\exp \ [-N(\frac{m_o+z}{m_o-\delta})^{\beta}] = G_{Y_1}(z)$$

$$= 1 - [1-F(y)]^N ,$$

where $F(y)$ is the distribution of $Y=-X$. Hence, $1-G_{Y_1}(z) \geq 1-\alpha$ with probability $1-\alpha$, and $Y_1 \geq -m_o + (m_o-\delta)\ [-\ln(1-\alpha)^{1/N}]^{1/\beta}$ with probability $1-\alpha$. Substituting $m_o = 12.3$, $\beta = 12.5$, and $\delta = 5.2$, the moment estimates for the California Catalog, and $N = 2(48) = 96$, we obtain a 95 percent lower confidence bound for Y_1, of -8.4, hence an upper confidence bound for $X_{(96)}$ (48 years) of 8.4 at confidence level .95. Using the parameter estimates associated with the Mann-Fertig median unbiased estimation procedure gives 8.8 as a 95 percent upper confidence bound for $X_{(96)}$.

Use of this same technique for obtaining an approximate 95 percent upper confidence bound for the largest 1000 year earthquake gives 9.3 and 9.7 by substituting the moment estimates and median unbiased procedures, respectively. These bounds do not take into account the uncertainty in the parameter estimates.

5.3.4 Goodness-of-Fit

Fourier analyses, such as Figure 5-10, indicate that the untransformed data constitute a mixture of at least 2 distributions. To obtain a goodness of fit test using more mundane statistical techniques, we consider results of Mann (1983).

In that article it is shown that classical outlier tets that use ratios of dispersion estimates as test statistics are essentially testing the differences of slopes of two lines that would be plotted on probability paper, one line using all observations and one using all but the suspected outliers. Such a test is most efficient when the set of n-k suspected outliers(applying to the smaller sized distribution in the assumed mixture) plot as a line with slope greater (or possibly smaller) than the slope of the line formed in the plot by the other k observations.

In Mann (1983), critical values are published for testing for outliers in various situations, but none applying to the situation at hand.

To find applicable tables for an appropriate test statistic, we use a result shownin Mann and Fertis (1975) indicating that in tests of goodness-of-fit it is the choice of spacings used rather than the way the spacings are combined that determine the power of the test.

Thus, as in Mann, Scheuer and Fertig (1973) and Mann and Fertig (1975) we use linear combinations of $(X_{(i+1)} - X_{(i)})$, $i=1, \ldots, n-1$ to estimate the slope of the line formed by the $k = 7$ largest observations and the slope of the line formed by all the observations.

We let $L_i = (X_{(i+1)} - X_{(i)})/E(Z_{(i+1)} - Z_{(i)})$, $i = 1, \ldots, n-1$

with $Z_k = (X_{(k)} - M)/\xi$, $k = 1, \ldots, n$.

Here n is the mode of a type-I extreme value distribution of largest values and ξ is a scale parameter, with $\Pi\xi/\sqrt{6}$ the distribution standard deviation. The negatives of the values of $E(Z_k)$ are found in White (1967). These values were reordered from smallest to largest to take into consideration the change in sign (since White's tabulated values apply to type-I extreme-value distributions of smallest values).

The test statistic, approximately Beta with parameters k-1 and n-k, is

$$\sum_{i=n-6}^{n-1} L_i \Bigg/ \sum_{i=1}^{n-1} L_i$$

and effectively tests whether the ratio of the slopes of the line formed by all the observations and the line formed by the suspected outliers is significantly different from unity.

For the parameters of interest, 6 and 89, the fiftieth percentile of a Beta distribution is 0.0652 and the tenth is 0.041. The calculated value from the California data set is 0.041. Thus we would reject an hypothesis of a single extreme-value type-I distribution (with no upper bound) at the 0.10 significance level.

Testing the goodness of fit of the three parameter extreme-value threshold model other than by the Fourier analysis approach of Grafstat presents problems because of the third parameter assumed under the null hypothesis.

To devise some sort of test, let us consider the median-unbiased estimator of m_o discussed earlier. To obtain a median-unbiased estimate for m_o (one that would be too large with probability 0.5 and too small with probability 0.5) by the procedure of Mann and Fertig, one would find a good first guess m_o^* for m_o, subtract each observation x_i from m_o^* and form $\ln(m_o^* - x_i)$, $i=1,\ldots,n$. The transformed observations are then ordered from smallest to largest to form

$$w^*(i) = \ln(m_o^* - x_{(n-i+1)})$$

estimates for $w_{(i)} = \ln(m_o - x_{(n-i+1)})$, $i=1,\ldots,n$.

The variates $W_{(i)} = Z_{(i)}/\beta + \ln\delta$, $i=1,\ldots,n$ have expectation $E(Z_{(i)})/\beta + \ln\delta$, with $E(Z_{(i)})$ the expectation of i^{th} reduced, parameter-free extreme-value order statistic. The initial statistic used for estimating m_o iteratively is

$$S^* = \sum_{i=k+1}^{n-1} L_i^* \Big/ \sum_{i=1}^{n-1} L_i^* ,$$

where

$$L_i^* = (W^*_{(i+1)} - W^*_{(i)}) \Big/ (E(Z_{(i+1)}) - E(Z_{(i)})).$$

Somerville (1977) shows that optimal results are obtained when k in (2.1) is equal to $(n-1)/5$. Since S^* is monotonic in m^*, as shown by Mann and Fertig, an iterative procedure like a bisection technique can be used to find an approximately median-unbiased estimate for m_o corresponding to a value of S^* equal to the fiftieth percentile of a Beta distribution with parameters $4(n-1)/5$ and $(n-1)/5$ 1(when n-1 is divisible by 5).

The fact that the analytical estimate of m_o found by this approach and the associated linear estimates of β and δ so closely match the graphical estimates of these parameters and demonstrate relatively small discrepancies with the moment estimates which involve the first 3 moments of the data indicates that completely different manipulations of the data under the given model give very close to the same results. This is encouraging and is reinforced by the results of the Grafstat analysis.

Using 14 for m_o^* yields 0.803 for S^*, very close to the 50th percentile 0.8023 of S^*.

In general, the 3-parameter extreme value model appears to predict well, and the predicted quakes ranging over past years are at least as large in magnitude as actual observed magnitudes. In other words, this model is more conservative than the exponential model, but possibly not conservative enough. Thus we are erring in the right direction. Further corroboration of the model is given by the fact that linear plots are exhibited for various subsets of the data, i.e. largest 3-month, 6-month, 12-month earthquakes and by the results of Yegulalp and Kuo (1974).

6. REFERENCES

Abramowitz, M. and I. A. Stegun (1972). Handbook of Mathematical Functions: Dover, New York, 1046 p.

Barlow, R. E. and N. D. Singpurwalla (1974). In Proceedings of the Symposium on Statistical Aspects of Air Quality Data, Research Triangle Institute, N.C., Ed. L. D. Kornreich, U.S. E.P.A. Pub. No. EPA 650/4-74-038.

Benjamin, J. R. and C. A. Cornell (1970). Probability Statistics, and Decision for Civil Engineers: McGraw-Hill Book Co., New York, 683 p.

Bernreuter, D. L. (1980). Seismic hazard analysis, a methodology for the eastern United States: U.S. Nuclear Regulatory Commission NUREG/CR-1582, Vol. 2.

Bolt, B. A. (1978). Incomplete formulations of the regression of earthquake magnitude with surface fault rupture length: Geology, v. 6, p. 233-235.

Cohen, A. C. and B. Whitten (1982). Modified maximum likelihood and modified moment estimators for the three parameter weibull distribution. Communications in Statist., Theory and Methods v. 11, p. 2631-2656.

Epstein, B. and C. Lomnitz (1966). A model for the occurrence of large earthquakes: Nature, v. 211, 954-956.

Earth Technology (1981). Earthquake recurrence intervals at nuclear power plants; Task 1A, Literature review; Task 1B, Proposed ranking of methodologies: prepared for U.S. Nuclear Regulatory Commission under Contract NRC-04-81-167, December, 1981.

Earth Technology (1982). Earthquake recurrence intervals at nuclear power plants; Task 2, Selection of methods to be ranked: prepared for U.S. Nuclear Regulatory Commission under Contract NRC-04-81-167, March, 1982.

Gutenberg, B and C. F. Richter (1956). Magnitude and energy of earthquakes: Ann. Geofisica, v. 9, p. 1-15.

Kagan, Yan, and L. Knopoff (1978). Statistical study of the occurrence of shallow earthquakes: Geophys. J. R. Astr. Soc., v. 55, p. 67-86.

Knopoff, L. (1971). A stochastic model for the occurrence of main sequence earthquakes: Revs. Geophys. and Space Physics, v. 9, p. 1975-188.

Knopoff, L., and J. K. Gardner (1972). Higher seismic activity during local night on the raw worldwide earthquake catalog: Geophys. J. R. Astr. Soc., v. 28, p. 311-313.

Knopoff, L., and Y. Y. Kagan (1977). Analysis of the theory of extremes as applied to earthquake problems: Jour. Geophys. Res., v. 82, p. 5647-5657.

Knopoff, L., Y. Y. Kagan and R. Knopoff (1983). b-values for foreshocks and aftershocks in real and simulated earthquake secuences: in press.

Lieblein, J. (1953). On the exact evaluation of the variances and covariances of order statistics in samples from the extreme value distribution: Ann. Math. Statist., v. 24, p. 282-287.

Lomnitz, C. (1966). Statistical prediction of earthquakes: Reviews of Geophysics, v. 4, p. 377-393.

Makjanic, B. (1972). A contribution to the statistical analysis of Zagreb earthquakes in the period 1869-1968. Pure Appl. Geophys. v. 95, p. 80-88.

Mann, N. R. (1983). Optimal outlier tests for a Weibull model - to identify process changes or to predict failure times: in Optimization in Statistics - Eds. S. Zanakis and J. Rustagi, The Institute of Management Science, Providence.

Mann, N. R. and K. W. Fertig (1975). A goodness-of-fit test for the two parameter vs. three-parameter Weibull; confidence bounds for threshold. Technometrics, v. 17, p. 237-245.

Mann, N. R., R. R. Schafer and N. D. Singpurwalla (1974). Methods for Statistical Analysis of Reliability and Life Data. New York, John Wiley.

Mann, N. R., E. M. Scheuer, and K. W. Fertig (1973). A new goodness of fit test for the two-parameter Weibull or extreme-value distribution with unknown parameters. Commun. in Statist., v. 2, p. 383-400.

Mann, N. R. and N. D. Singpurwalla (1982). Extreme-value distributions: in Encyclopedia of Statistical Sciences, ed. N. L. Johnson and S. Kotz, John Wiley.

Mortgat, C. P. and H. C. Shah (1978). A Bayesian approach to seismic hazard mapping; development of stable design parameters: John A. Blume Earthquake Engineering Center, Dept. of Civil Engr., Report No. 28, Stanford University, 233 p.

Newmark, N. M., and E. Rosenmuth (1971). Fundamentals of Earthquake Engineering: Prentice Hall, New Jersey.

Parwardhan, A. S., R. B. Kulkarni and D. Tocher (1980). A semi-Markov model for characterizing occurrence of great earthquakes: Bull. Seism. Soc. Am., v. 70, p. 323-347.

Richter, C. F. (1958). Elementary Seismology: W. H. Freeman and Company, San Francisco.

Rikitake, T. (1976). Earthquake Prediction: Elsevier, Amsterdam.

Schenkova, Z. and V. Karnik (1970). The probability of occurrence of largest earthquakes in the European area, II: Pure Appl. Geophys., v. 80, p. 152-161.

Schenkova, Z. and V. Schenkova (1975). Return periods of earthquakes and trends of seismic activity. Pure Appl. Geophys., v. 113, p. 683-693.

Smith, P. D., R. G. Dong, D. L. Bernreuter, M. P. Bohn, T. Y. Chuang, G. E. Cummings, J. J. Johnson, R. W. Mensing and J. E. Wells (1981). Seismic Safety Margins Research Program - Phase 1 Final Report Overview: U.S. Nuclear Regulatory Commission NUREG/CR-2015 Vol. 1.

Somerville, P. N. (1977). Some aspects of the use of the Mann-Fertig statistic to obtain confidence interval estimates for the threshold parameter of the Weibull: in Theory and Applications of Reliability, Vol. 1. Eds. C. P. Tsokos and I. N. Shimi. Academic Press, New York.

Spetzler, C. S. and C. S. Stael von Holstein (1975). Probability encoding in decision analysis: Management Science, v. 22, p. 340-357.

Stepp, J. C. (1972). Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard: in Proceedings of the International Conference on Microzonation, Seattle, Washington, p. 897-909.

Tarter, M. E. (1978). Implementation of harmonic data analysis procedures: in Proceedings of Computer Science and Statistics: Eleventh Annual Symposium on the Interface. Eds. R. Gallant and T. Gerig. North Carolina State University, Raleigh.

Veneziano, D., C. A. Cornell, and T. O'Hara (1984). Historic method of seismic hazard analysis: Electric Power Research Institute Report NP-3438, prepared by Yankee Atomic Electric Company.

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes: Bull. Seism. Soc. Am., v. 70, p. 1337-1346.

White, J. S. (1967). The moments of log-Weibull order statistics: General Motors Research Publication GMR-717, General Motors Corporation, Warren, MI.

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

NRC FORM 335
(11-81)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned ...)
NUREG/CR-4145

4. TITLE AND SUBTITLE (Add Volume No., if appropriate)
Earthquake Recurrence Intervals at Nuclear Power Plants: Analysis and Ranking

2. (Leave blank)

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)
J.A. Hileman, L. Knopoff, N.R. Mann, R. K. McGuire

5. DATE REPORT COMPLETED

MONTH	YEAR
December	1984

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Earth Technology Corporation
3777 Long Beach Boulevard
Long Beach, California 90807

DATE REPORT ISSUED

MONTH	YEAR
March	1985

6. (Leave blank)

8. (Leave blank)

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)
Division of Radiation Programs & Earth Sciences
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

10. PROJECT/TASK/WORK UNIT NO.

11. FIN NO.
B7226

13. TYPE OF REPORT
Final Technical

PERIOD COVERED (Inclusive dates)
November 1981--December 12, 1984

15. SUPPLEMENTARY NOTES

14. (Leave blank)

16. ABSTRACT (200 words or less)

Five methods for estimating earthquake recurrence were ranked. The methods represent those used, or proposed, in nuclear power plant studies through 1982 and include Log Linear Poisson, Extreme Value,, Semi-Markov, Bayesian, and Uniform Hazard Method. Ranking focused on recurrence estimates for earthquake sources, excluding attenuation and site response. Scores were assigned to each method for 12 criteria such as accuracy, use of geologic data, and subjective judgment. Criteria scores were weighted by their importance and summed. Different scoring and weighting schemes were used to identify any sensitivities. To aid in scoring statistical criteria, methods were tested on synthetic earthquake catalogs with known statistics, and natural catalogs were tested against theoretical magnitude distributions.
The uniform Hazard Method scored high because, in principal, expert judgement draws upon all seismologic knowledge. The Bayesian Method scored low because data requirements are severe for practical cases. The other methods were intermediate. These observations seem insensitive to scorer, scoring approach, or weighting scheme. The semi-Markov Method scores were sensitive to the weighting scheme.

17. KEY WORDS AND DOCUMENT ANALYSIS

Earthquake recurrence
Probabilities
Nuclear Power Plants
b-Value
Bayesian method

17a DESCRIPTORS

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlimited

19. SECURITY CLASS (This report)
Unclassified

21. NO. OF PA...

20. SECURITY CLASS (This page)

NRC FORM 335 (11-81)

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D.C. 20555

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

FOURTH CLASS MAIL
POSTAGE & FEES PAID
USNRC
WASH. D.C.
PERMIT No. G-67

120555078877 1 1ANIRA
US NRC
ADM-DIV OF TIDC
POLICY & PUB MGT BR-PDR NUREG
W-501
WASHINGTON DC 20555