UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

PROCEEDINGS OF CONFERENCE XXXIV

A WORKSHOP ON "PROBABILISTIC EARTHQUAKE-HAZARDS ASSESSMENTS"

November 25-27, 1985
San Francisco, California

CONVENED UNDER THE AUSPICES OF THE
NATIONAL EARTHQUAKE HAZARDS REDUCTION PROGRAM

SPONSORED BY
U.S. Geological Survey
U.S. Nuclear Regulatory Commission

EDITOR
Walter W. Hays
U.S. Geological Survey
Reston, Virginia 22092

Open-File Report 86-185

Compiled by
Carla Kitzmiller

Reston, Virginia
1986

CONTENTS

158 40008

SUMMARY AND HIGHLIGHTS OF THE WORKSHOP ON PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENTS

By
Walter W. Hays
U.S. Geological Survey
Reston, Virginia 22092

INTRODUCTION

The U.S. Geological Survey (USGS) and the U.S. Nuclear Regulatory Commission (NRC) sponsored a 3-day workshop in San Francisco, California, on November 25-27, 1985. The objectives were: 1) to review the methodologies currently used to assess earthquake hazards probabilistically, especially in the Eastern United States, and 2) to identify practical and innovative ways to improve the overall state-of-knowledge and to foster the implementation of this knowledge in the siting of nuclear power reactors and other applications.

This workshop, the thirty-fourth in a series of conferences and workshops sponsored since 1977 by USGS in cooperation with other Federal Agencies since 1977, was attended by 45 scientists and engineers representing industry, academia, architectural and engineering firms, national laboratories, and the Federal Government. Representatives of agencies of three foreign governments, Spain, Italy, and Chile, were included. The participants, listed below and in Appendix A, represent a large percentage of those who are actively involved in research on probabilistic earthquake hazards assessments, and the application of the research results throughout the United States.

Keitti Aki	University of Southern California
Shelton Alexander	Pennsylvania State University
Ted Algermissen	U.S. Geological Survey
Walter Arabasz	University of Utah
Rodrigo Araya	Chile
Bernice Bender	U.S. Geological Survey
Don Bernreuter	Lawrence Livermore National Laboratory
Rafael Blazquez	Spain
Steve Brocoum	U.S. Nuclear Regulatory Commission
Kenneth Campbell	U.S. Geological Survey
Kevin Coppersmith	Geomatrix Consultants
Allin Cornell	Stanford University
John Dwyer	Law Engineering Testing Company
Gus Giese-Koch	U.S. Nuclear Regulatory Commission
Walter Hays	U.S. Geological Survey
Abou Bakr Ibrahim	U.S. Nuclear Regulatory Commission

John Jacobson	Yankee Atomic Electric Company
William Joyner	U.S. Geological Survey
Jerry King	Electric Power Research Institute
Armen Der Kiureghian	University of California, Berkeley
Ram Kulkarni	Woodward Clyde Consultants
Cole McClure	Bechtel
Robin McGuire	Risk Engineering Inc.
Richard Mensing	Lawrence Livermore National Laboratory
Giuliano Milana	Italy
Christian Mortgat	TERA Corporation
Andrew Murphy	U.S. Nuclear Regulatory Commisson
David Perkins	U.S. Geological Survey
Paul Pomeroy	Rondout Associates, Inc.
Maurice Power	Geomatrix Consultants
Leon Reiter	U.S. Nuclear Regulatory Commission
Albert Rogers	U.S. Geological Survey
Robert Rothman	U.S. Nuclear Regulatory Commisson
Jean Savy	Lawrence Livermore National Laboratory
Leonello Serva	Italy
Haresh Shah	Stanford University
Jogeshwar Singh	Harding-Lawson Associates
Burton Slemmons	University of Nevada
Phyllis Sobel	U.S. Nuclear Regulatory Commission
Tom Stratton	Woodward Clyde Consultants
Carl Stepp	Electric Power Research Institute
Paul Thenhaus	U.S. Geological Survey
Nafi Toksoz	Massachusetts Institute of Technology
Daniele Veneziano	Massachusetts Institute of Technology
Ian Wall	Electric Power Research Institute

This workshop is the fifth conference jointly sponsored by the USGS and NRC. The prior workshops and the references for their proceedings are listed below:

1) Conference XVI, The Dynamic Characteristics of Faulting Inferred from Recordings of Strong Ground Motion, October 21-23, 1981. (Reference: Boatwright, 1982, U.S. Geological Survey Open-File Report 82-591.)

2) Conference XX, The 1886 Charleston, South Carolina Earthquake and Its Implications for Today, May 23-26, 1983. (Reference: Hays and Gori, 1983, U.S. Geological Survey Open-File Report 83-843.)

3) Conference XXII, Site-Specific Effects of Soil and Rock on Ground Motion and the Implications for Earthquake-Resistant Design, July 25-27, 1983. (Reference: Hays, 1983, U.S. Geological Survey Open-File Report 83-845.)

4) Conference XXXII, Earthquake Hazards in the Puget Sound, Washington, Area, October 29-31, 1985. (Reference: Hays and Gori, 1986, U.S. Geological Survey Open-File Report 86-93.)

BACKGROUND

The ongoing research and technical assistance programs of the USGS and the NRC have dedicated significant resources to emphazing of both assessment of earthquake hazards, deterministically and probabilistically. The objective of the USGS is to conduct fundamental research and to prepare maps and other documents showing the broad variation of seismic hazards and risk throughout the Nation (Executive Office of the President, 1978). The broad objective of the NRC research program is to focus on fundamental issues in earthquake hazards created by the need to develop techniques to deal in a regulatory environment with the uncertainties associated with: 1) seismic source zones, 2) propagation of seismic energy, and 3) site-specific ground-motion response, including soil amplification (U.S. Nuclear Regulatory Commission, 1985). NRC's technical assistance program focuses on specific regulatory issues related to individual or groups of nuclear power plants.

Research related to earthquake hazards is also being performed by other organizations within the United States. For example, the National Science Foundation (NSF) and the Department of Energy (DOI) through its national laboratories (such as the Lawrence Livermore National Laboratory (LLNL) have significant programs. A private organization, the Electric Power Research Institute (EPRI) also has a strong program. Many individuals in academia and in industry are involved in the research on earthquake hazards under the USGS, NRC, NSF, DOE, LLNL, and EPRI programs.

THE FIELD OF PROBABILISTIC EARTHQUAKE ASSESSMENTS

Since 1968 when Cornell published one of the classic papers in the emerging scientific field of probabilistic earthquake hazards assessments, rapid advances have occurred in: the technical methodologies, and computer and analytical modeling capabilities. The number of applications and researchers has grown substantially since 1968 and many important and insightful contributions have been made to the technical literature. A few references are cited below to illustrate the broad diversity of these contributions (note: a comprehensive review of the literature is beyond the scope of this report):

1) Methodology (Cornell, 1968; Algermissen and Perkins, 1976; Woodward Clyde Consultants, 1982; Bender, 1984; Electric Power Research Institute, 1985a).

2) Technical Issues (Chung and Bernreuter, 1981; McGuire and Barnhard, 1981; McGuire and Shedlock, 1981; Hays, 1984; Electric Power Research Institute, 1985b; Campbell, 1985).

3) Applications (TERA, 1980; Algermissen and others, 1982; Bernreuter and others, 1984; Budnitz and others, 1985; Thenhaus and others, 1985; Shieh and others, 1985; Boissonnade and Shah, 1985; Cummings, 1985).

Figure 1 illustrates schematically the types of hazards that can occur in an earthquake. Almost all of them are now being modeled probabilistically. Damage losses are also being modeled probabilistically.

As the field of probabilistic earthquake hazards assessments has evolved, to facilitate communication, a glossary of technical terms is included as Appendix B. Standard usage of terms is very important.

WORKSHOP PROCEDURES

The procedures followed in the workshop were designed to increase the interaction and communication between participants. Each participant was encouraged to join in the articulation and debate of the technical issues, maintaining a broad collegiate point of view on both the definition of the technical issues and their proposed solutions (Figures 2 and 3). The emphasis was placed on innovation, not confrontation. The goal was to seek ways to eliminate or reduce the controversy associated with various topical subjects.

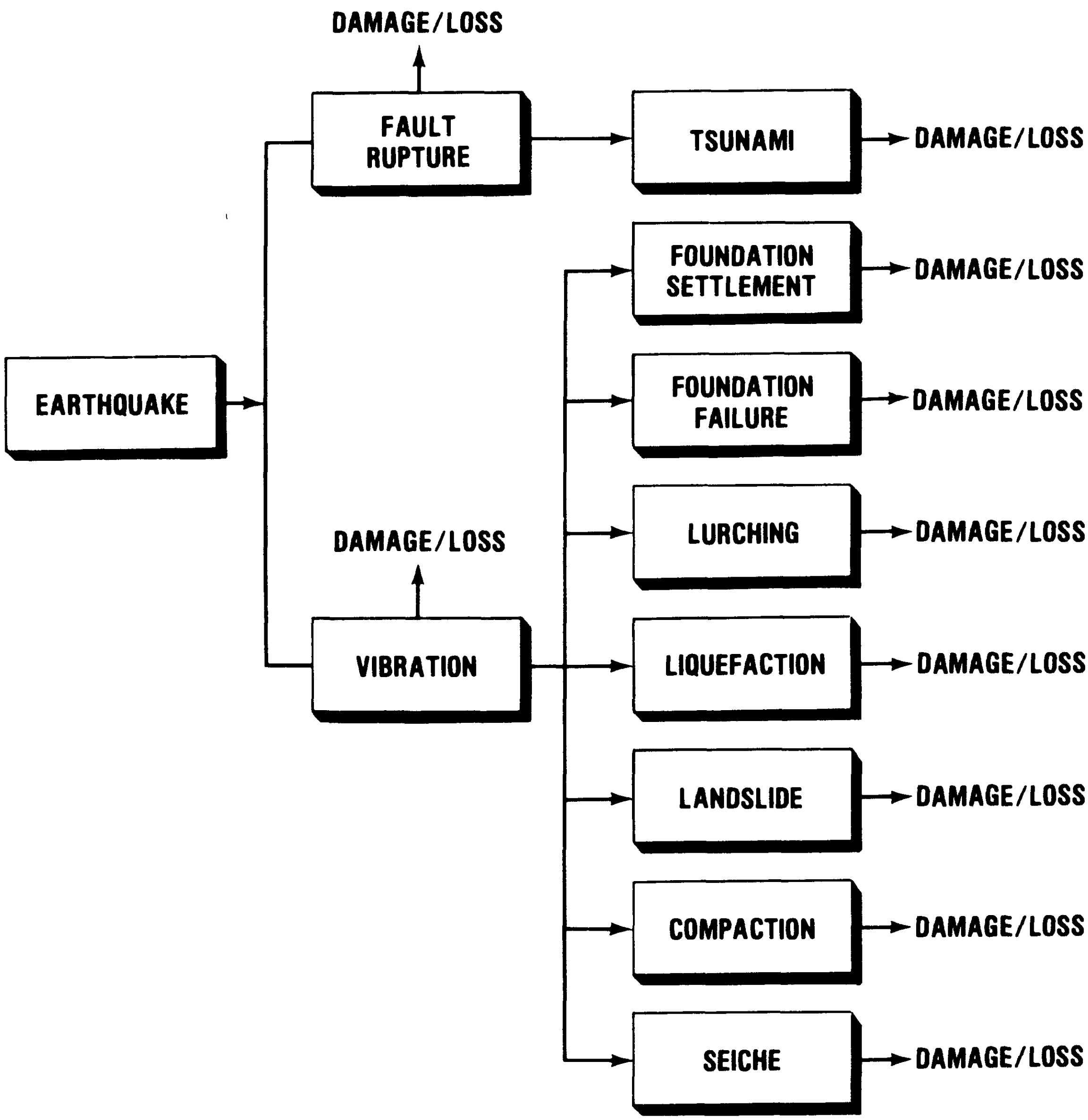


Figure 1.--Schematic illustration of the types of physical phenomena (hazards) that can occur in an earthquake and cause damage and losses. Almost all of these hazards as well as damage and losses are now being assessed probabistically.

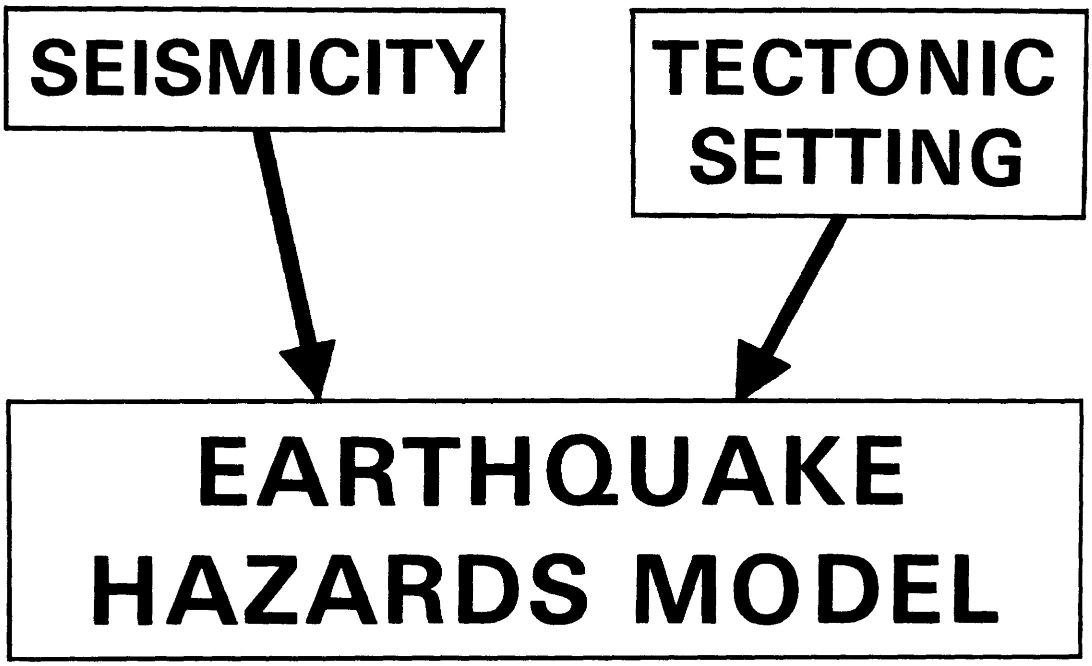


Figure 2.--Schematic illustration of the earthquake hazards model and the two main sources of data that are used to define the key parameters. The process of defining the parameters of the model forces researchers to deal with technical issues associated with seismotectonics, seismogenic zones, earthquake recurrence, magnitudes, and complexities of the earthquake rupture process. Questions such as the following must be addressed: 1) Where have earthquake occurred in the past. 2) Where are they occurring now? 3) Why are they occurring? 4) How often do they recur? 5) How big have they been? 6) How big can they be? 7) How severe have the physical effects been? and 8) How severe can the physical effects be in the future? Inability to produce explicit answers can lead to controversy and hinder applications.

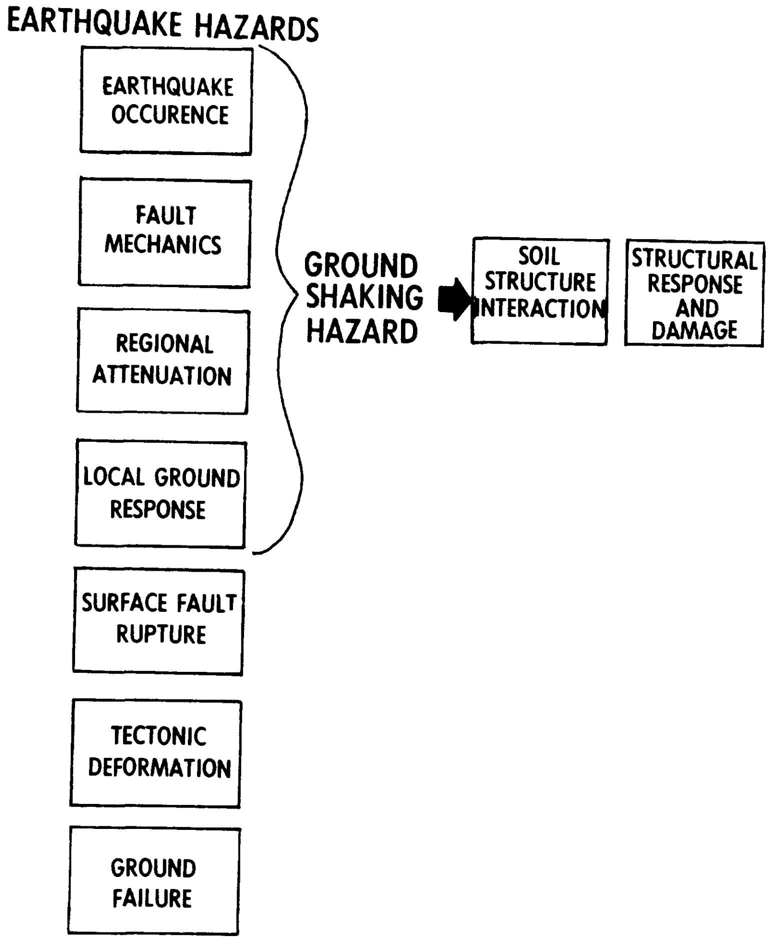


Figure 3.--Schematic illustration of the key elements in probabilistical assessments of the ground-shaking hazard. The hazard is typically expressed in terms of three parameters: 1) ground motion (for example, peak bedrock acceleration), an exposure time (for example, 50 years), and a probability of nonexceedance (for example, 90 percent). The application usually drives the selection of these three parameters. Each box represents a broad category of typical research studies, each of which has unresolved technical issues. The extent to which the technical issues are resolved determines in large measures the degree of controversy affects believability and applicability of the ground shaking hazards products.

153 40008

Seven procedures were followed throughout the workshop. They are summarized below for completeness:

Procedure 1: The participants were provided with a broad overview of the range of technical issues in probabilistic earthquake hazards assessments. This activity provided a broad framework for discussion throughout the meeting.

Procedure 2: Twelve researchers gave 20-30 minute presentations on selected topical subjects, focusing on the technical issues and ways to resolve them.

Procedure 3: Two discussion stimulators, along with the moderator, initiated the group discussion after each topical presentation. The goal was to focus on the technical issues to seek to draw out opinions from the participants about the extent to which each technical issue had (or had not) been resolved. The time allocated for free wheeling discussion was approximately equal to the time allocated for presentation.

Procedure 4: Two teams of researchers representing Lawrence Livermore National Laboratory and Electric Power Research Institute described the procedures each organization used to make probabilistic earthquake hazards assessments.

Procedure 5: Two discussion stimulators, along with the moderator, initiated a group discussion of each organization's methodology and results, focusing on the extent to which technical issues had been resolved and suggesting ways improvements might be made.

Procedure 6: Two researchers gave their perspectives on fruitful directions that the field of probabilistic earthquake hazards assessments might take in the near future.

Procedure 7: A panel of six individuals representing EPRI, LLNL, NRC, and USGS presented ideas and suggestions for the most appropriate next steps to foster the continued development and application of probabilistic earthquake hazards assessments methodologies.

Following the workshop, another procedure was used to create the proceedings; a permanent record of the workshop.

Procedure 8: All speakers were requested to provide a manuscript for the conference proceedings. Each speaker had a period of sixty days after the workshop to finalize the manuscripts for the proceedings.

WORKSHOP THEMES AND OBJECTIVES

The themes, objectives, speakers, and panelists for each session of the workshop are described below. The papers contained in this report contain detailed information on each session

SESSION 1: OVERVIEW OF TECHNICAL ISSUES

Objective: To provide a framework for discussion throughout the workshop.

Speaker: Ted Algermissen, U.S. Geological Survey

SESSION 2: DISCUSSION OF IMPORTANT TOPICS IN PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENTS

Objective: By means of a series of topical presentations provide a broad range of perspectives about critical technical issues. Through interactive discussion, determine the extent to which each technical issue has (or has not) been resolved and the related controversy, if any.

Topics and Speakers:

Seismotectonics (Relations between earthquake and earth structure, seismogenic zones)
--Paul Thenhaus, U.S. Geological Survey

Characterization and Representation of Future Seismicity (Balance between geologic and historical indicators of seismic activity, representation of historical seismicity and allocation of regional seismicity to constituent seismogenic zones).
--Daniele Veneziano, Massachusetts Institute of Technology
--Kenneth Campbell, U.S. Geological Survey
--David Perkins, U.S. Geological Survey

Earthquake Recurrence Models (Time-independent (Poissonian) versus time-dependent (Weibull, others) models for time occurrence; Gutenberg-Richter relation, segmented Gutenberg-Richter relation, and characteristic earthquake models for magnitude distribution).
--Allin Cornell, Stanford University
--Kevin Coppersmith, Geomatrix Consultants

Earthquake Magnitudes (Maximum and minimum magnitudes)
--Burton Slemmons, University of Nevada

Modeling Earthquake Sources (Point, line, and plane ruptures; complex ground motions; directivity; uncertainty)
--Armen Der Kiureghian, University of California, Berkeley

Seismic Wave Attenuation (Functional models, near- and far-field characteristics, directivity, stress-drop variability, path effects variability).
--Robin McGuire Risk Engineering Inc.
--Kenneth Campbell, U.S. Geological Survey

Local Site Effects (Site geologic models, frequency- and strain-dependent effects, empirical data, variability).

Applications (Considerations of regional versus site-specific applications, short versus long return periods, short versus long exposure times).

Discussion:
--Christian Mortgat, TERA Corp;
--Paul Pomeroy, Rondout Stimulators Associates
--Haresh Shah, Stanford University

SESSION III: REVIEW OF METHODOLOGIES FOR PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENTS

Objective: A general review and discussion of methodologies

Topics and Speakers: Overview and results of recent studies conducted by Lawrence Livermore National Laboratory
--Don Burneuter, Lawrence Livermore National Laboratory
--Richard Mensing, Lawrence Livermore National Laboratory
--Jean Savy, Lawrence Livermore National Laboratory

Overview and results of recent studies sponsored by Electric Power Research Institute
--Carl Stepp, Electric Power Research Institute
--Kevin Coppersmith, Geomatrix Consultants
--Robert Youngs, Geomatrix Consultants (Note: although unable to attend the workshop, his paper is included in the proceedings.
--Daniele Veneziano, Massachusetts Institute of Technology
--Robin McGuire, Risk Engineering Inc.

Discussion Stimulators:
Maurice Power, Geomatrix Consultants
Kenneth Campbell, U.S. Geological Survey

SESSION IV: DISCUSSION OF FUTURE STEPS

Objective: To identify strenghts and weaknesses in the present state-of-knowledge and state-of-practice and to suggest innovative ways to new odd pg

PREFACE

The field of probabilistic earthquake hazards assessments has evolved rapidly in the past two decades. Many organizations and many individuals are now using probabilistic methods routinely for a wide variety of applications that include:

1) Construction of national and regional ground-shaking hazard maps.

2) Construction of regional ground-failure hazard maps.

3) Construction of regional surface-faulting hazard maps.

4) Construction of regional tsunami hazard maps.

5) Construction of macro- and microzoning maps.

6) Making decisions about the siting of critical facilities (dams, hospitals, nuclear power reactors, etc.).

7) Establishing criteria for earthquake-resistant design.

8) Preparation of PRA's.

9) Construction of fragility curves for various types of structures, facilities, and lifelines.

In spite of the large increase in applications, much work remains to be done. Technical issues need to be resolved through focused research. Common usage of technical terminology needs to be encouraged. Basic publications need to be developed and widely disseminated to geologists, geophysicists, seismologists, and engineers.

This workshop was organized to bring together many of the people who are actively working on the leading edge of the field of probabilistic earthquake hazards assessments. The goal of this publication, the third in a new series on knowledge utilization, is to foster improved communication and utilization of fundamental knowledge in the field. The challenge is for the technical community to make innovative advances in theory, models, methodology, and applications, eliminating, or at least minimizing, controversy associated with technical elements of the problem. This publication will be a success if it stimulates improvements and progress in the field of probabilistic earthquake hazards assessments.

1B8 40008

Walter W. Hays
U.S. Geological Survey

advance them through improvements in theory, models, methodology, and applications.

Speakers: --Allin Cornell, Stanford University
--Haresh Shah, Stanford University

Panelists: --Keitti Aki, University of Southern California
--Robin McGuire, Risk Engineering Inc.
--Carl Stepp, Electric Power Research Institute
--Don Bernreuter, Lawrence Livermore National Laboratory
--Leon Reiter, U.S. Nuclear Regulatory Commission
--David Perkins, U.S. Geological Survey

SUMMARY OF TECHNICAL ISSUES

A number of technical issues have been identified in probabilistic earthquake hazards assessments. Unless these issues are resolved by improved or new data, increased understanding of the physics and analysis procedures, controversy occurs. Controversy affects the credibility and acceptability of various applications of the methodology. Some of the most important technical issues discussed at the workshop are summarized below; the papers contained in the report identify additional technical issues:

1) Lack of knowledge about seismogenic sources

   a. What are the preferred tectonic models? Why?
   b. How can critical seismotectonic data be acquired either to eliminate or to validate specific tectonic models?
   c. How is historical seismicity best used in the selection of the final tectonic model that will be used in ground-shaking hazard assessments?
   d. How are tectonic data best used in the selection of the final tectonic model that will be used in ground-shaking hazard assessments.

2) Utilization of existing seismicity data

   a. How should historical seismicity data be used in modeling seismogenic sources in both space and time?

2514000 8

b. How should historical seismicity data be distributed in a seismogenic source?

c. What weight should be given to historical seismicity data? Geologic data?

d. How should they be combined, if at all?

e. How should the maximum and minimum magnitudes of a seismogenic source be determined?

f. How should the maximum and minimum magnitudes be allowed to vary within and between seismogenic sources?

3) <u>Limited data on seismic wave attenuation</u>

a. Are intensity data useful or useless for modeling attenuation? How can the usefulness of these data be extended?

b. What is the best form for an attenuation law?

c. How can realistic frequency-dependent attenuation laws be derived for the East in light of limited data?

4) <u>Limited data or site geology and its effects</u>

a. What is the best way to incorporate the frequency-dependent effects of local site geology in probabilistic earthquake hazards assessments?

5) <u>Applicability of analytical modeling techniques</u>

a. Do earthquakes correspond best with the memoryless Poisson occurrence model having memory? What are the strengths and weaknesses of each analytical model?

b. How significant are the ways that boundaries of seismogenic sources are modeled?

c. How do they affect the level of the ground shaking hazard

d. How significant are the assumptions made in modeling fault rupture lengths?

e. How do they affect the level of the ground-shaking hazard?

 2BS 4000 8

6) <u>Utilization of expert opinion</u>

a. What are the strengths and weaknesses, biases, and pitfalls in the use of expert opinion? How can strengths of the procedures be improved and weaknesses be eliminated?

7) <u>Quantification of uncertainty and parameter sensitivity</u>

a. How precisely do we know the median values of important parameters? How precisely do we need to know them?
b. Which results of the ground-shaking hazard most sensitive to small changes in values of the physical parameters?
c. What types of surprises have occurred in past earthquakes? (For example, the 1976 Tangshan, China earthquake and the 1985 Mexico earthquake).

8) <u>Surprises</u>

a. What types of surprises have occurred in past earthquakes? Why?
b. Will these and other surprises occur in the future? How do we minimize the probability of surprises occurring that will have a negative impact on the current efforts in probabilistic earthquake hazards assessments?

CONCLUSIONS AND RECOMMENDATIONS

The participants concluded that although significant progress has been made in the past decade, great care must be taken at this stage to ensure that certain goals are met. These goals include:

1) A concentrated effort to eliminate or resolve critically important technical issues that now contribute substantially to the present controversy associated with the ground-shaking hazard.

2) Wise use of probabilistic models.

288 4000 8

3) Cross-education of geologists, geophysicists, seismologists, and engineers.

4) Increased understanding of the physics of the earthquake generation process.

5) Incorporation of more physics and mechanics into the analytical models.

If these goals are not attempted and achieved the field of probabilistic earthquake hazards assessments could suffer a setback. At the present time, the basic criticisms and perceptions include (rightly or wrongly):

1) The capability to create statistical and analytical models is far ahead of the data required to validate the models.

2) Basic understanding of the physics of the earthquake process in the Eastern United States is weak in spite of large expenditures for research and data acquisition.

3) Applications to many engineering problems are still controversial.

4) The benefit/cost ratio is not always clear.

These criticisms and perceptions must be dealt with through accelerated and focused research. Research results must be carefully documented and widely communicated.

ACKNOWLEDGMENTS

A special note of appreciation is extended to each of the following individuals for their contributions:

1) The Steering Committee of Leon Reiter (NRC), Ted Algermissen (USGS), Carl Stepp (EPRI), Don Bernreuter (LLUN), and Walter Hays (USGS).

 254 40008

2) The speakers and discussion stimulators who provided the information and stimulii for vigorous discussion.

3) The participants who joined in the discussion of the technical issues. Their vigorous and healthy exchange of ideas make the workshop interesting and stimulating

4) Carla Kitzmiller, Paula Gori, Joyce Costello, Lynne Downer, Wanda Fuller, and Shirley Carrico, USGS, who provided strong and capable administrative support.

REFERENCES

Algermissen, S. T., and Perkins, D. M., 1976, A Probabilistic Estimate of Maximum Acceleration in Rock in the Contiguous United States: U.S. Geological Survey Open-File Report 76-416, 45 p.

Algermissen, S. T., Perkins, D. M., Thenhaus, P. C., Hanson, S. L., and Bender, B. L., 1982, Probabilistic Estimates of Maximum Acceleration and Velocity in Rock in the Contiguous United States: U.S. Geological Survey Open-File Report 82-1033, 99 p.

Bender, B. L., 1984, Seismic Hazard Estimation Using a Finite Fault Rupture Model: Bulletin of the Seismological Society of America, v. 74, p. 1899-1923.

Bernreuter, D. L., Savy, J. B., Mensing, R. W., and Chung, D. H., 1984, Seismic Hazard Characterization of the Eastern United States: Methodology and Interim Results for Ten Sites: Lawrence Livermore National Laboratory, NUREG/CR--3756, UCRL-53527, 435 p.

Boatwright, John (Editor), 1982, The Dynamic Characteristics of Faulting Inferred from Recordings of Strong Ground Motion: Proceedings of Conference XVI, U.S. Geological Survey Open-File Report 82-591, 2 volumes, 976 p.

Budnitz, R. J., Amico, P. J., Cornell, C. A., Hall, W. J., Kennedy, R. P., Reed, J. W., and Shiumozuka, M., 1985, An Approach to the Quantification of Seismic Margins in Nuclear Power Plants: Lawrence Livermore National Laboratory, NUREG/CR-4334, UCID-20444, 398 p.

Buissonnade, A. C., and Shah, H. C., 1985, Earthquake Insurance Risk: John A. Blume Earthquake Engineering Center, Stanford University, Report no. 68, 55 p.

Campbell, K. W., 1985, Strong Ground Motion Attenuation Relations: A Ten Year Perspective: Earthquake Spectra, v. 1, pp. 759-804.

Chung, D. H., and Bernreuter, D. L., 1981, The Effect of Regional Variation of Seismic Wave Attenuation on the Strong Ground Motion From Earthquakes: Lawrence Livermore National Laboratory, NUREG/CR-1655, UCRL-53004, 173 p.

Cornell, C. A., 1968, Engineering Seismic Risk Analysis: Bulletin of the Seismological Society of America, v. 58, p. 1583-1606.

Cummings, G. E., 1985, Summary Report on the Seismic Safety Margins Research Program: Lawrence Livermore National Laboratory, NUREG/CR-4431, UCID-20549, 58 p.

Electric Power Research Institute, 1985a, Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States: EPRI/SOG Draft Report 85-1, 3 vols., 444 p.

Electric Power Research Institute, 1985b, Tectonic Framework and Seismic Source Zones of the Eastern United States: EPRI/SOG, Draft Reports 85-1 to 85-7, 1111 p.

Executive Office of the President, 1978, The National Earthquake Hazards Reduction Program: Government Publishing Office, Washington, D.C., p. 10.

Hays, W. W. (Editor), 1983, Site-Specific Effects of Soil and Rock on Ground Motion and the Implications for Earthquake-Resistant Design: Proceedings of Conference XXII, U.S. Geological Survey Open-File Report 83-845, 501 p.

Hays, W. W., and Gori, P. L., (Editors), 1983, The 1886 Charleston, South Carolina Earthquake and Its Implications for Today: Proceedings of Conference XX, U.S. Geological Survey Open-File Report 83-843, 502 p.

Hays, W. W., 1984, Technical Problems in the Construction of a Map to Zone the Earthquake Ground-Shaking Hazard in the United States: Engineering Geology, v. 20, p. 13-23.

Hays, W. W. and Gori, P. L., (Editors), 1986, Evaluation of Regional and Urban Earthquake Hazards and Risk in the Puget Sound, Washington Area: Proceedings of Conference XXXII, U.S. Geological Survey Open-File Report 86-93, 301 p.

McGuire, R. K., and Barnhard, T. R., 1981, Effects of Temporal Variations in Seismicity on Seismic Hazard: Bulletin of the Seismological Society of America, v. 71, p. 321-334.

McGuire, R. K., and Shedlock, K. M., 1981, Statistical Uncertainties in Seismic Hazard Evaluations in the United States: Bulletin of the Seismological Society of America, v. 71, p. 1287-1308.

Shieh, L. C., Johnson, J. J., Wells, J. E., Chen, J. C., and Smith, P. D., 1985, Simplified Seismic Probabilistic Risk Assessment--Procedures and Limitations: Lawrence Livermore National Laboratory, NUREG/CR-4331, UCID-20468, 235 p.

Tera Corporation, 1980, Seismic Hazard Analysis--Solicitation of Expert Opinion: Lawrence Livermore National Laboratory, NUREG/CR-1582, v. 3, 146 p.

 2B1 40008

Thenhaus, P. C., Ziony, J. I., Diment, W. H., Hopper, M. G., Perkins, D. M., Hanson, S. L., and Algermissen, S. T., 1985, Probabilistic Estimates of Maximum Seismic Horizontal Ground Acceleration on Rock in Alaska and the Adjacent Continental Shelf: Earthquake Spectra, v. 1, p. 285-306.

U.S. Nuclear Regulatory Commission, 1985, Seismic Safety Research Program Plan: NUREG-1147, p. 2-1.

Woodward Clyde Consultants, 1982, Development of Initial Application of Software for Seismic Exposure Evaluation, v. 2, 22 p.

255 40008

WHERE DO WE GO FROM HERE TO INCREASE THE STATE OF KNOWLEDGE ?

by

C. Allin Cornell
Stanford University
Stanford, California 94305

I don't pretend to have any unexpected answers to the questions in the title that was assigned to me. Like the rest of you, I am here to find out where we are today as a first step in deciding what to do tomorrow. The subject of seismic hazard analysis is advancing very rapidly. My own experience with the EPRI project, for example, has been that it was absorbing, developing and using ideas as fast as those involved in it could come with them. I simply take advantage of the space offered me to make a few comments hopefully relevant to the title.

A CAUTION

Although we may speak of increasing the state of knowledge in seismic hazard analysis, we should all recognize that probabilistic models of nature do not in themselves increase our information. At best they help us organize our knowledge in a useful way. That use may be to help us apply the information, for example, to an engineering decision or to help stimulate the search for more knowledge. Unfortunately, at their worst, probabilistic models may tend to force our knowledge into an ill-fitting mold that distorts or obscures our state of knowledge.

Let me give some positive and negative examples. When it was developed nearly 20 years ago, the "classical" model of seismic hazard analysis was created to meet engineering needs.

Relevant knowledge could at that time be expressed in the form of physical features such as faults, frequency data in the form of Gutenberg-Richter diagrams, and ground motion prediction as crude attenuation laws. Using a "minimalist principle", these elements were fashioned into the simplest possible model. The model was apparently successful in meeting engineers demand for probabilistic statements about the earthquake hazard. It permitted them to compare these threats with other hazards, both natural phenomena and accidents and to fit the information in a uniform way into structural codes and engineering risk assessments. Further, the model probably stimulated us to look for deviations from the simple exponential distribution of magnitudes and deviations from the Poisson law. More broadly, generalized versions of seismic hazard analysis have provided a framework into which to place more elegant models of space-time behavior of earthquakes and their effects; examples include clustering models, spatial migration, new physically-based ground motion prediction, etc. On the other hand, it is easy to point to examples of situations where the elements of the simple model are still applied without critical examination. In many cases, there is no questioning of the exponential magnitude assumption or the stationary, memoryless Poisson model. In contrast, Bob Wallace has argued for jumping of the seismicity from region to region. Others have argued that at the feature-specific level seismicity may occur in an on-and-off pattern. Admitting such representations of seismicity may change not only our probabilistic predictions but the entire way we formulate our professional assessments and inferences.

I do not believe that the question at any such professional scientific-engineering meeting should be: are probabilistic methods acceptable? Rather, it should be: how do we make best use of all of the tools available to the scientific and engineering profession, where those tools include mechanics, photography, probability, Indian folk tales, animal behavior or whatever else may be on the agenda. In short, I would like to encourage the widest and wisest possible use of probabilistic models and uncertainty treatment to insure that they enhance rather than inhibit the state of knowledge.

CROSS EDUCATION

I certainly need not tell the geoscientists in the audience how ignorant we engineers are of their field. I would not recognize a mafic pluton if it were not crawling across my sleeve. At the risk of being patronizing, however, I would like to express my concern with what I perceive to be a rather widespread lack of education in the seismological community at large in the fundamentals of statistics and probability. Elementary statistics are used frequently in the science and application of seismology. Much of the science has, particularly in the past, been empirical. To illustrate my point, I only need cite the recent experience in the profession with the unfortunate misunderstandings about which way to conduct a regression analysis when studying magnitude and rupture lengths. Seismic hazard analysis requires, in addition to statistics, an acquaintance with probability theory and the interested scientist's education must be broaden still further. A clear example of geoscientists' difficulty with probability can be found in the experience the first Lawrence Livermore Study Team had in trying to communicate to their geoscience experts the meaning of the maximum magnitude as a truncation point on a distribution function. My recent experience is that many seismologists are learning these tools very rapidly. They already represent a much larger percentage of their profession than do the engineers knowledgeable in probability. But this is as it should be for few engineers are asked to use statistics and probability on a day-to-day basis in their profession.

My concern becomes deeper when we observe that the methodology groups in the large projects on seismic hazard analysis that we have reviewed these past days are predominantly populated by engineers not geoscientists. It concerns me, because as modelers we may be much too naive. We may not know enough to create other than simple models, unknowingly excluding options that these models should include. We may be guilty of producing the bad mold discussed above. We badly need informed, experienced, controlled guidance and critique by geoscientists trained and experienced in applied probability theory. Clearly then, a major need in order that the state of knowledge of seismic hazard analysis be increased is that the geoscientists and engineers (or more generally the probabilistic modelers) educate one another more thoroughly in their respective fields. Let me turn next to a few individual specific topics in seismic hazard analysis.

UNCERTAINTY TREATMENT

Certainly the major change in the last few years in the way seismic hazard analysis is typically treated is the explicit uncertainty assessment and propagation. It aids in the reporting, in the tracking, and in the assignment of responsibility in the analysis process. It has made it easier for a scientist to assign numbers to parameters or hypotheses for which he may have been reluctant to make simple unqualified point estimates. The important question here is whether our current techniques of dealing with uncertainty treatment, and more particularly the use of multiple expert opinions, are properly capturing the community's level of knowledge about seismic hazard estimation. The LLNL projects were the first to address the multiple-expert question head on. The later EPRI project has tried to structure these assessment and opinion aggregation issues even more carefully. One of the project reviewers, Dr. Peter Morris, has stated that this may be the most elaborate and successful attempt in the scientific-technology-public-policy arena to encode and express multiple expert opinion. Nonetheless, there is clearly room for improvement in the procedures and the wide airing of the issues involved in the opinion aggregation.

GROUND MOTION ATTENUATION

This subject remains the single largest concern especially in the eastern U.S. assessments. If I had some good ideas about how to solve this problem, I would not be standing here speaking I would be out exercising them. Certainly the derivatives seem to be correct, that is, away from Modified Mercalli Intensity and toward theory and data. In the west, the major changes we have seen recently include the increased use of theoretical and numerical modeling coupled with empirical data. The need here is to develop techniques for using the modeling predictively rather than descriptively.

EARTHQUAKE RECURRENCE MODELS

Non-trivial (i.e., non-Poisson-exponential) models have been available for many years, pioneered by scientists such as Aki, Vere-Jones, Toksoz, Knopoff, and others. Few of these models, however, have had a major impact on practice. I believe seismic hazard analysis provides the formulation into which those models can be inserted with effective results. This development is moving fast in the western U.S. where feature-specific applications have been conducted. Nonetheless, the models are probably already outstripping the data available. The pressing need is for more physical, mechanistic models to permit some kind of generalization that might be extended to the eastern U.S. as well. Given the complexity of the physical process, however, it may be much too naive to believe that the same or similar models apply to all or even most physical situations. Some short-term gain appears to be possible but confirmed, specific models for any specific feature appear to be a long way off.

STATISTICAL METHODS OF CATALOGUE TREATMENT

Again the use of advanced statistical methods is not new to seismology. I think, for example, of the pattern-recognition work of Keilis-Borok and others. Nonetheless, it appears that Dr. Veneziano's recent contributions to this area will be particularly effective because they have been inserted within the larger framework of seismic hazard analysis. The EPRI project has, after seriously re-investigating many alternatives, reconfirmed the important role of seismicity in the eastern U.S. hazard assessment. I agree with Dr. Reiter that we are, therefore, obliged to dig into this historical seismicity information and get as much out of it as is possible. Veneziano's new methods should facilitate the process. Again, I would suggest that the burden is upon the seismologists to prepare themselves to critique and understand these methods well. The door is open to exploit spatial-temporal statistical techniques and to implement them immediately in seismic hazard assessment.

SOURCE IDENTIFICATION

The recent EPRI project has made a much needed step, going backward from the rather loosely defined areal source zones in customary eastern U.S. seismic hazard analysis models to a feature-by-feature interpretation. They have been successful in bringing a much larger fraction of the geoscience community face-to-face with seismic hazard analysis. The project has, I believe, given everyone the satisfaction of having gone back to first principles and to the causative mechanisms of eastern U.S. earthquakes before the application of seismic hazard analysis. This process should induce a much greater stability over time in the interpretation of the seismic threat in the eastern U.S. by reducing the impact of new theories and local events. The advances in knowledge in this area will come from the geoscientist community, for example, through paleo-seismicity or from new interpretations of the physics of the failure process. These advances will not come quickly but seismic hazard analysis per se will probably not have an impact upon the rate at which this knowledge is gained.

THE APPLICATION INTERFACE

From past experience we can anticipate that many of the demands for new knowledge in seismic hazard analysis will come from the users. For example, the engineers have demanded that hazard assessments be in terms of spectral velocity rather than simply PGA, and other users, policy makers, have asked that our results communicate the degree of uncertainty in our hazard assessments.

In the future we can anticipate that the engineers will ask for perhaps still better scala measures of ground motion for their use in conventional design, e.g., Dr. Kennedy's "effectiveness" measure. Similarly, they may ask for more sophisticated multi-parameter measures for use in design margin assessments or in probabilistic risk assessments. Fortunately, it is the engineering user who will settle such current arguments as to what the lower bound magnitude value should be used in seismic hazard computations. Policy makers who utilize the results of seismic hazard analysis also will create questions that need to be addressed in future seismic hazard analysis development. These users find themselves making comparative evaluations, acceptable risk decisions, and expected cost computations. It is from such applications that resolution of issues such as whether the hazard analyst should compute means or medians, how he should display the uncertainty and how he should aggregate the opinion of multiple experts will arise.

PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENT--WHERE DO WE GO FROM HERE?

by

H. C. Shah and
W. M. Dong
Stanford University
Stanford, California 94305

During the last 15 years, methods for evaluating seismic hazards have reached a certain level of maturity. There has been an evolution of methodology from deterministic to probabilistic procedures. Researchers have realized that there is a great deal of uncertainty involved in each step of hazard analysis. Faced with such uncertainties, deterministic methods adopt conservative values for all factors involved in hazard analysis, such as maximum magnitude, shortest distance etc. to cover uncertain adverse situations, resulting in extremely large design requirements. For most structures, even critical facilities, these highly conservative design values cannot be justified economically for use. It was thus reasonable to develop a probabilistic approach which takes into account the uncertainties in the size, location, and attenuation of seismic events. Use of probabilistic seismic hazard and risk analysis has been widely accepted by engineers, planners, and regulatory bodies. With this acceptance has come the wider use of such models in developing load and design criteria for facilities in seismic regions. This is indeed a positive development which will and should continue in this direction to refine models and methodologies.

With the increasing use of probabilistic methods have come the problems of misuse, overuse, and, sometimes, over-reliance and faith in the results. We have to be aware of the weaknesses in our current technology. The authors of this review paper have just come back from China where our Chinese colleagues shared some interesting statistics. Out of the 11 major earthquakes in the last ten years or so, seven occurred in regions which were estimated to be low seismic zones. These seven events resulted in 81% of the overall economic and human losses.

Of course, these numbers do not suggest that our hazard analysis was wrong because there is not sufficient data to refute our original hypothesis.

However, we should remember that many of our hypotheses are based upon insufficient data. Thus, it is also not strong evidence to support our hypotheses.

Uncertainty is caused not only by randomness but also by ignorance. In the probabilistic setting, we need a distribution to work with. If we know the distribution, then we can say that the uncertainty is only related with randomness. However, we know neither the exact distribution nor a reliable frequency distribution due to scarcity of data. Lack of knowledge about distribution has a limited place in the formal probability theory, because the result we get is a single probability value. In this regard, Dampster and Shafer's evidence theory (Ref.1) would be appropriate to reflect randomness and ignorance by providing the bound estimation (called plausibility and credibility). The essence of evidence theory is to distinguish between disbelief and lack of belief which seems more appropriate regarding our state of knowledge.

On the other hand, because of lack of data we can not get sufficient statistical information. We might use our experience and meta-knowledge for reasoning from evidence to reach a hypothesis. This is a kind of knowledge engineering. Experience is a loosely structured knowledge which is usually summarized in the "if-then" form with natural language statements, e.g. "if the epicenter is far away, then the predominant frequency at the site would be low". Note that in these inference rules, we seldom use numbers but words, so-called linguistic values. Often, these are fuzzy, imprecise statements without crisp boundaries. Some natural phenomenon are so complex that perhaps only vague assertion might be justified simply because these assertions are compatible with the broad range of observed facts. Also the association of if-then might not be uncertain; sometimes the associations exist, sometimes not. This is a new frontier in our research efforts, a real challenge which convolve the following four disciplines (See Fig. 1).

- o Earthquake Engineering
- o Artificial Intelligence
- o Probability Theory
- o Fuzzy Theory

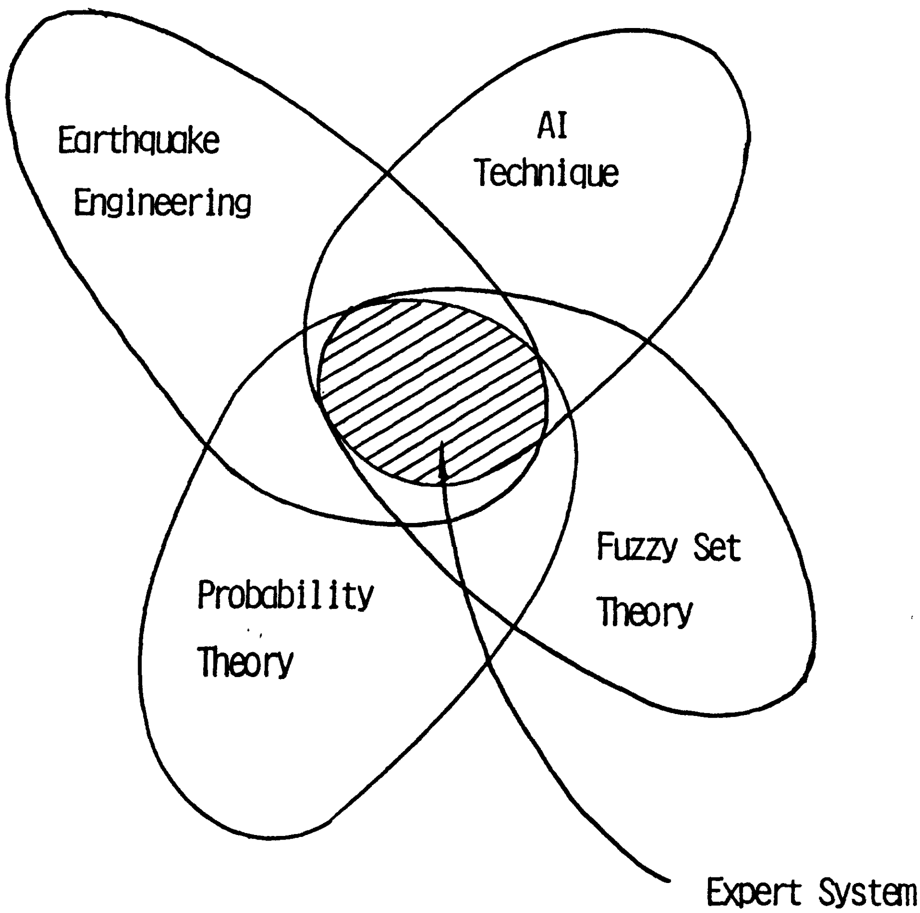


Fig. 1 Expert system for earthquake engineering (Ref. 2).

From our point of view, until such time as we have sufficient data,the best solution is to combine the statistical data with our experience, subjective judgment, and knowledge. This is our general feeling. More specifically, we will mention about our suggestions for future research at each step of a hazard analysis.

SOURCE MODELING

In the regions where seismic sources are relatively well defined along the plate boundaries or faults, the active fault approach is used for seismic hazard evaluation. Seismic sources could be modeled as point, line, or area source. In this case, the seismicity is usually relatively high and the source-to-site distance is reasonably well defined. However, for many intraplate regions, such as the eastern United States, the seismicity is diffused over a large area without any identifiable faults. In these regions, the tectonic province approach is used to delineate different provinces as area sources. Each province is assumed to have a homogeneous seismicity, resulting in large uncertainty. Due to lack of sufficient historical and geological evidence, there is no unique and consistent way to delineate the provinces. Considering the successful development of pattern recognition in many fields, the authors presume that the clustering analysis combining with fuzzy information from the experts can be used in order to reduce the arbitrariness and uncertainty (Ref. 3). For example, in delineating basically "round" provinces, the clustering analysis can be conducted by minimizing the objective functional

$$J_w(U,v) = \sum_{i=1}^{c} \left(\sum_{\underline{x}_k \in \underline{u}_i} w_k \, ||\underline{x}_k - \underline{v}_i||^2 \right) \qquad (1)$$

where U is the ensemble of the provinces,

c is the number of provinces,

$\underline{u}_i$ is the i^{th} province,

$\underline{v}_i$ is the center vector of province i,

$\underline{x}_k$ is the vector of historical epicenter,

w_k is the weight of each epicenter, considering the magnitude and the confidence of the data.

3B5 40008

For other shapes of provinces, the appropriate clustering criterion has to be used to get a rational delineation (such as single-linkage criterion for strip-type provinces).

Recent research of association seismicity with tectonic features from expert's subjective evaluation is in the right direction (Ref. 4). Usually the expert opinion is very vague, mostly using language descriptions. For example, when determining the activity of a feature, one has to investigate the ground surface features which are described by visual inspection, imagery technique, drilling, and trenching, etc. The information obtained from these methods is usually very imprecise and not well defined. Hence fuzzy logic is needed to infer the activity from that information. Also such a logic can help to combine opinions of different experts in a consistent manner.

SOURCE SEISMICITY

The assessment of source seismicity has mostly depended on the recorded events, but the database on a given source is often incomplete and nonhomogeneous in time. In the past, due to low population density and lack of interest in earthquake activity, only large events were recorded. With increasing instrumental coverage, intermediate and small earthquakes have been recorded with more frequency, producing an apparent increase in seismicity with time which biases the statistics from the catalog of data. A data adjustment is usually necessary to get a reliable hazard analysis.

For the occurrence of earthquakes, there are two kinds of models: time-independent and time-dependent models. Traditionally, the homogeneous Poisson model has been used to describe the occurrence of earthquakes. The common Gutenberg & Richter (GR) relationship $\ln N(m) = a + b\ m$ implies the exponential distribution of magnitude with unlimited maximum magnitude. The fact that the magnitude in a specific region has an upper bound leads to a modified GR relationship by truncating the magnitudes at the maximum possible magnitude (Ref. 5). This truncated model has good agreement with world-wide data as well as with data for smaller regions. However, for some regions where the historical data are usually scarce, one would not expect a reliable prediction based only on the data, especially for large events. As an example, estimation of seismic hazard in the San Francisco Bay Area from data obtained between 1907 and 1983 will underestimate the hazard. On the other hand, data from 1900 to 1983 may overestimate the hazard if the mean return period for large events is more than 200 years. In such cases, it would be prudent to use geological and/or geophysical information to anchor the rate of occurrence of large events.

There are three methods by which geological and geophysical information about the occurrence of large earthquakes can be used. One way is to use the average seismic moment to "anchor" the values of M_{max}, a , and b (Ref. 6). The geophysical data are assumed to have no uncertainty. This method does not consistently combine the short-range historical data with long-range geological and geophysical data. The second method uses the concept of modified maximum entropy. Based on the information on M_{max} and its average return period, the probability distribution function for magnitudes is obtained. This distribution has the property that it is minimally biased and is consistent with the type and level of information (Ref. 7). The third method considers the uncertainty in the geological/geophysical information. Combination of such data with historical data is done by incorporating relative "weights" through Bayes' model (Ref. 8). Any one of these methods achieves one objective and that is to use all the available information. This gets us away from relying too much on short historical data.

All the models discussed above assume that earthquakes (small, medium, or large) are independent events. The elastic rebound theory and the available evidence suggest that the assumption of independence may be practical but not realistic. To get around this problem, researchers have suggested time-dependent models (Refs. 9, 10, 11, 12).

All time-dependent models under investigation seem to be more realistic than the time-independent models currently used. However, we have to be careful in our unconditional adoption of these models. They require estimation of many constants and parameters that are currently either not known or the database is so small that they cannot be estimated with any reasonable reliability or confidence. Thus estimation of hazards from such "realistic" models may still be full of assumptions and uncertainty. Perhaps, the best conclusion we can get a the qualitative statement such as: the longer the holding time, the bigger the event due to a constant rate of increase of stress. We need tools to handle such fuzzy information.

ATTENUATION OF GROUND MOTION

The type and amount of attenuation of seismic ground motion depends on many factors such as the size of the event, the type of fault mechanism, transmission path, distance and local soil condition of the site. The commonly used empirical attenuation relationship incorporates some of these parameters but generally leaves out important variables such as the azimuth between the source and the site, and the parameters that identify the fault rupture mechanism. Equation 2 shows the commonly used empirical attenuation function

$$PGA = f(M, R, b_1, b_2, b_3, c) \qquad (2)$$

where PGA = Peak Ground Acceleration

M = magnitude,

R = distance from the source to the site,

b_1, b_2, b_3 = regression constants, which depend on the type of data, site condition, transmission path, etc.,

C = saturation effect, depending on magnitude

The uncertainty in explaining the "load effect" at a site due to an earthquake using such crude empirical equations is considerable. Large error terms are common, indicating that attempts to quantify the site severity parameter are at best crude. The problems in using Eq. 2 are two-fold. First, the ground motion severity cannot be truly represented by a single parameter such as the PGA. Second, the equation leaves out important contributing factors such as the azimuth, stress drop, velocity of rupture, etc. In recent years, attempts have been made to rectify this but without much success due to lack of reliable data. Generating ground motion severity parameters by using geophysical models such as the normal mode analysis (Ref. 13) may provide one bit of additional input to our database. Such analytically generated values, which would be functions of distance, magnitude, azimuth and geophysical properties of fault rupture may be helpful in refining and improving our attenuation relationships. Some recent work on the use of pattern recognition may also provide a better tool in processing the data from various diverse sources.

CONCLUSION

It is often of great concern when we see users of the current probabilistic risk analysis models put unreasonable confidence in the numbers they get. Often, we see engineers, planners, regulators, and public officials argue about the level of peak ground acceleration one would get for, say, a 100-year return period. A 10-20% variation from the estimated value is sometimes argued between the various parties as if the analyst had the ability to "fine-tune" the numbers. This "overreliance" or "faith" in the results seems to be inversely proportional to the level of understanding one has about the uncertainty in each step of the hazard analysis.

In practice, researchers and engineers have used models which did not necessarily fit the phenomenon but which were available at the time of analysis. Thus, deterministic models were used to describe highly probabilistic and uncertain events. Similarly, some researchers insist on using statistical and probabilistic models for problems which are based on qualitative and fuzzy information. This is synonymous with the case of a person looking for his key near a lamp post even though he has dropped his key somewhere else where there was darkness. Just because deterministic and probabilistic models are highly developed and are available, does not mean that we have to use them, no matter what the information base and the physical problems are.

There is also a trend among users and developers of probabilistic hazard analysis procedures to "get more out" of the data than the data can provide. This is especially true in source modeling and in attenuation studies. There seem to be as many attenuation relationships as there are researchers in seismic hazard analysis. They all use the same data and they all come out with conceptually similar empirical models with minute differences in constants or the numerical procedures. We seem to be all trying to squeeze water out of stones. In our opinion, the time has come when we should look at the available information from an entirely fresh perspective instead of redoing everything with an "episilon" type of variation, just to get a method or an equation which is given the name of the researcher. Use of pattern recognition in sorting past intensity data and in combining recurrence information from historical and geologic databases is one such fresh approach. Use of fuzzy set theory may not provide the ultimate in combining expert opinion, but it will provide a new and fresh look at the way we do things. Only through such innovative and imaginative tools will we be able to improve our ability to reduce uncertainty in seismic hazard estimation.

REFERENCES

1. Shafer, C.T. (1976), A Mathematical Theory of Evidence, Princeton University Press, Princeton, NJ.

2. Dong, W.M., Boissonnade, A.C. and Shah, H.C. (1985), "Treatment of Vague Information in the Development of a Risk Evaluation System Application to Seismic Risk Analysis," to be presented at the 1st International Conference on Applications of Artificial Intelligence to Engineering Problems, Southampton, U.K.

3. Bezdek, J.C. (1982). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York.

4. Woodward-Clyde Consultants (1984), "Approaches to Developing Tectonic Frameworks and Seismic Sources," Proc. for Workshop No. 4.

5. Cornell, C.A. and Vanmarcke, E.H. (1969), "The Major Influences on Seismic Risk," Proc. 4th World Conference Earthquake Engineering, Santiago, Chile, Vol. 1.

6. Anderson, J.G. and Luco, J.E. (1983), "Consequences of Slip Rate Constraints on Earthquake Occurrence Relations," Bull. Seism. Soc. Am., Vol. 73.

7. Dong, W.M., Shah, H.C. and Bao, A.B. (1983), "Use of Maximum Entropy Principle in Earthquake Recurrence Relationships," Proc. of U.S.-Japan Cooperative Research on Generalized Seismic Risk Analysis and Development of a Model Seismic Format.

8. Dong, W.M., et al. (1983), "Utilization of Energy Flux, Seismic Moment and Geological Information in Bayesian Seismic Hazard Models," Proc. of U.S.-Japan Cooperative Research on Generalized Seismic Risk Analysis and Development of a Model Seismic Format.

9. Savy, J.B., Shah, H.C. and Boore, D. (1980), "Nonstationary Risk Model with Geophysical Input," J. Struc. Div., ASCE, Vol. 106.

10. Nishioka, T. and Shah, H.C. (1980), "Application of the Markov Chain on Probability of Earthquake Occurrence, Proc. of J.S.C.E., No. 298.

11. Patwardhan, A.S., Kulkarni, R.B. and Tocher, D. (1980), "A Semi--Markov Model for Characterizing Recurrence of Great Earthquakes, Bull. Seism. Soc. Am., Vol. 70.

12. Kiremidjian, A.S. and Anagnos, T. (1983), "Stochastic Models for Earthquake Occurrence," Technical Report, The John A. Blume Earthquake Engineering Center, Stanford University.

13. McCann, M.W. (1981), "A Bayesian Geophysical Model for Seismic Hazard," Technical Report, The John A. Blume Earthquake Engineering Center, Stanford University.

SOME PROBLEMS IN SEISMIC HAZARD ASSESSMENT

by

S. T. Algermiseen
U.S. Geological Survey
Golden, Colorado

ABSTRACT

A preliminary examination of Modified Mercalli intensity data in the central and southeastern United States suggests ground motion levels in some areas that are quite different than those given by conventional, regional probabilistic ground motion assessments. In these areas, variations in ground motion resulting from regional or subregional variations in attenuation or site response may be as important as the ground motion generated by the seismicity associated with the seismic source zones.

Variations in techniques used to model seismic source zone boundaries in probabilistic ground motion assessment may result in significant differences in expected ground motions near source zone boundaries. Variations in techniques used to model fault rupture lengths may also result in significant differences in expected ground motion at selected locations.

Modeling of viable tectonic hypotheses with limited observational data is recognized as a major problem in several parts of the country. For example, assessment of probabilistic ground motion in the Puget Sound area is highly uncertain because two of the three suggested sources of large earthquakes (large or great offshore plate boundary shocks and large onshore shallow shocks) cannot be adequately modeled with the available data. Earthquakes with magnitudes in the range $6.5 \leq M_S \leq 7.5$ at depths of 40-70 km can be adequately modeled but the spatial distribution of these shocks is uncertain.

INTRODUCTION

Many users are unaware of the substantial effects alternative modeling techniques can have on estimated ground motions in probabilistic seismic hazard analysis. Viable alternative models arise from uncertainties in both data and knowledge. This paper discusses several modeling aspects of earthquake hazard assessment that may significantly affect ground motion estimates.

The following problems in hazard assessment are discussed: (1) the significance of persistent geographical areas of anomalous Modified Mercalli (MM) intensity in the central and southeastern United States; (2) modeling of seismic source zone boundaries; (3) modeling of fault ruptures; and (4) difficulties in modeling viable seismotectonic hypotheses with minimal observational data. Only a qualitative discussion of these problems is attempted here, but it is believed that all of these aspects of seismic hazard assessment deserve additional, extensive research and quantification.

ATTENUATION AND SITE RESPONSE--CENTRAL AND EASTERN UNITED STATES

Probable ground shaking assessed on a regional basis is routinely referenced to a standard bedrock (Algermissen and Perkins, 1976; Applied Technology council, 1978; Algermissen and others, 1982). For example, "rock" as used by Algermissen and Perkins (1976) means: "....material having a shear wave velocity of between 0.75 and 0.90 km/sec...." However, even a cursory examination of Modified Mercalli (MM) intensity data shows that regional ground motion estimates based on "rock" do not account for a variety of observed intensities. In the eastern and central United States the problem is difficult to resolve since not many recordings of strong ground motion are available and the geotechnical properties of sites are often not well known. Modified Mercalli intensity observations from reasonably well observed shocks provide some clues as to the extent and amplitude of attenuation variation and site response, at least on a subregional basis (over areas on the order of 10 to 1000 km^2). A qualitative examination of the relevance of intensity data is attempted here using intensity observations and maps for the eight earthquakes listed in table 1.

Table 1.--Earthquakes in the central and eastern United States used and related data in this study[1]

Date	Location	Maximum M.M. intensity (I_o)	Magnitude (m_b)	Source of intensity data (see footnotes)	Constants A	Constants B	Constants C	Standard deviation $\sigma(I_C)$
Jan. 5, 1843	Near Memphis, Tenn.	VIII	(6.0)	4	10.0571	-.0054	-1.162	.147
Aug. 31, 1886	Near Charleston, S.C.	X	(6.8)	5	12.6605	-.0036	-1.930	.200
Oct. 31, 1895	Near Charleston, Mo.	VIII-IX	(6.2)	6	10.9615	-.0040	1.503	.144
May 31, 1897	Giles County, Va.	VIII	5.8	7	9.9300	-.0066	-1.366	.325
Sept. 20, 1931	Near Anna, Ohio	VII	5.3[2]	8	9.2535	-.0020	-2.481	.118
Nov. 9, 1968	South Central Ill.	VII	5.5	9	9.9530	-.0032	-1.670	.081
July 27, 1980	Near Sharpsburg, Ky.	VII	5.1	10	7.8496	-.0053	-.929	.269
Jan. 31, 1986	Near Chardon, Ohio	VI	5.0[3]	11	A, B and C for 1931 earthquake used; VI < I_C < IX			

[1]Data (except for source of intensity map and as noted) compiled by Algermissen (1983) from various sources. Brackets indicate magnitudes estimated from intensity.
[2]Nuttli and others (1978)
[3]U.S. Geological Survey, Preliminary Determination of Earthquakes
[4]Hopper and Algermissen (1983)
[5]Bollinger (1977)
[6]Hopper and Algermissen (1980)
[7]Map by Margaret Hopper (1980, unpublished) based primarily on data from Law Engineering Testing Company (1975)
a. The Monthly Weather Review
b. Campbell (1898)
c. Hopper and Bollinger (1971)
d. Bollinger and Hopper (1971)
e. Various newspapers
[8]Map by Margaret Hopper (1979, unpublished) based primarily on data from Neumann (1932).
[9]Unpublished map by Hopper (1980) based on data from:
a. Coffman and Cloud (1970)
b. Gordon and others (1970)
c. Unpublished USGS intensity questionnaire cards
[10]Map by Stover and Reagor (1980, preliminary unpublished map). Except for a few minor details, this map is identical to the published map in Stover and von Hake (1982).
[11]Unpublished preliminary field data by Hopper (February 1986).
[12]Constant s in $I_C = A + B\Delta + C\log\Delta$; $I_C \leq I_o$

For each earthquake listed in table 1, the area shaken at each intensity level was computed. A least squares regression equation of the form

$$I_C = A + B\Delta + C\log\Delta \qquad (1)$$

was used to compute the intensity I_C given the average outer radius (Δ) for each intensity level. In (1), Δ is the radius of a circular isoseismal with an area equivalent to the average area shaken at intensity I_C and greater; A, B and C are constants. The difference between the observed intensity I_{OB} and I_C, called here the residual intensity I_R was computed using

$$I_R = I_{OB} - I_C \qquad (2)$$

I_{OB} was obtained by estimating intensities at points at intervals of 0.5° of latitude and 0.65° of longitude. The points at which actual intensity data was observed and the isoseismal maps for each earthquake were used to estimate I_{OB}. Intensities were estimated on the same grid of points for each earthquake to facilitate averaging the residual intensities I_R for several earthquakes.

Figures 1, 2, and 3 show some of the results of these analyses. In figure 1, contours of residual (or anomalous) intensity were obtained by averaging the residual intensities among the three largest earthquakes in the central and southeastern United States since 1812 at each grid point. Figure 2 shows the results of a similar analysis for all eight earthquakes and figure 3 compares the results shown in figure 1 with those on figure 2. Figure 3 indicates that there is quite reasonable correlation between the residual intensities obtained for the three largest earthquakes (1843, 1886, and 1895) and the total sample of eight earthquakes considered. The source of the residual intensities is not known. Interestingly, however, some anomaly patterns (fig. 1) lie astride or closely parallel the southern margin of Pleistocene glacial drift (P. C.Thenhaus, oral commun., 1986).

The basic intensity observations (or isoseismals) from which the residual (intensity anomaly) maps have been constructed have at least two sources of uncertainty. First, there is the uncertainty resulting from the

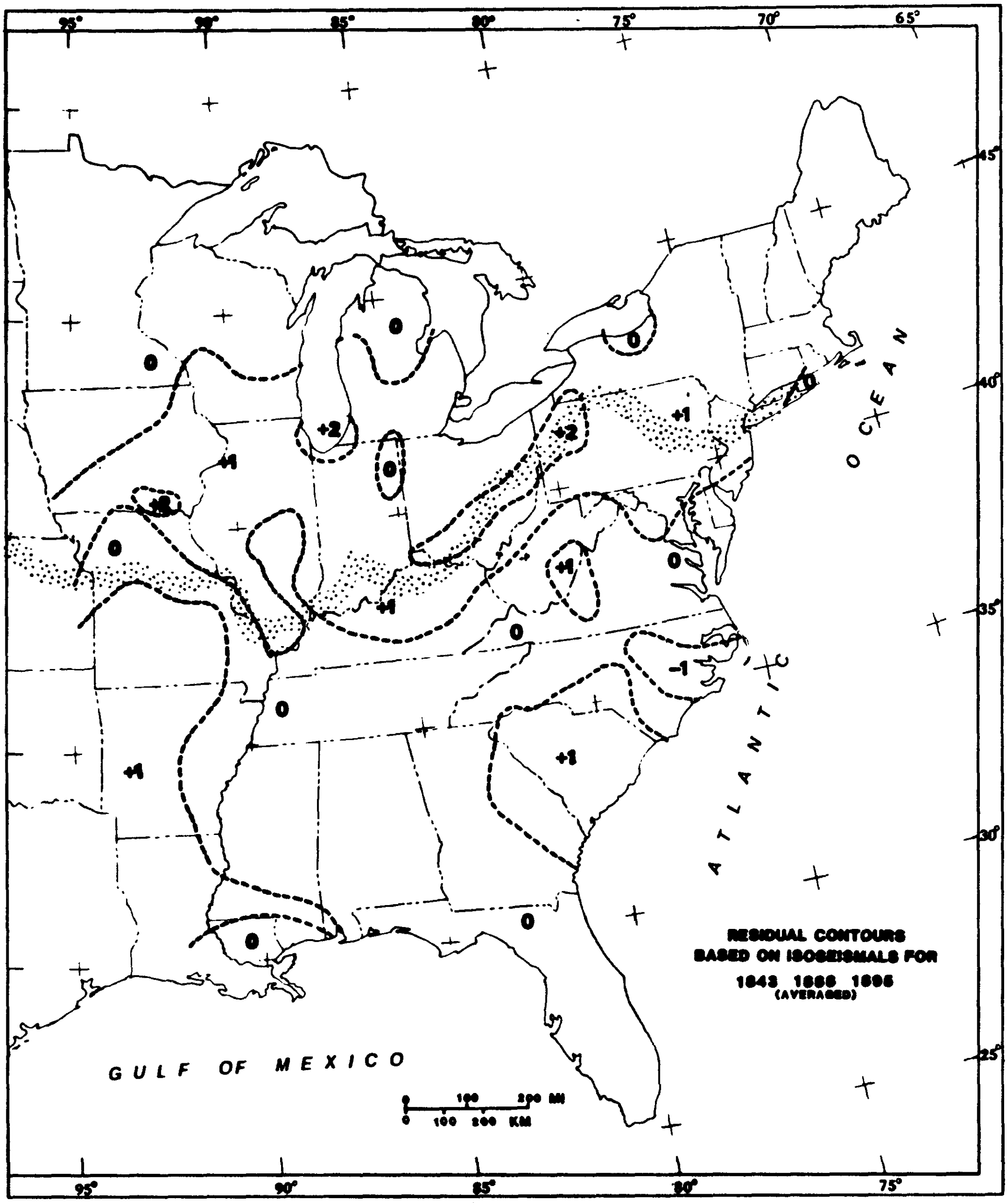


Figure 1.--Anomalous intensity in the central and southeastern United States based on the analysis of the three largest earthquakes in these areas since 1812. Contours separate areas having differences in Modified Mercalli intensities of one or more unit of intensity. Stippling shows the southern extent of Pleistocene glacial drift.

353 40008

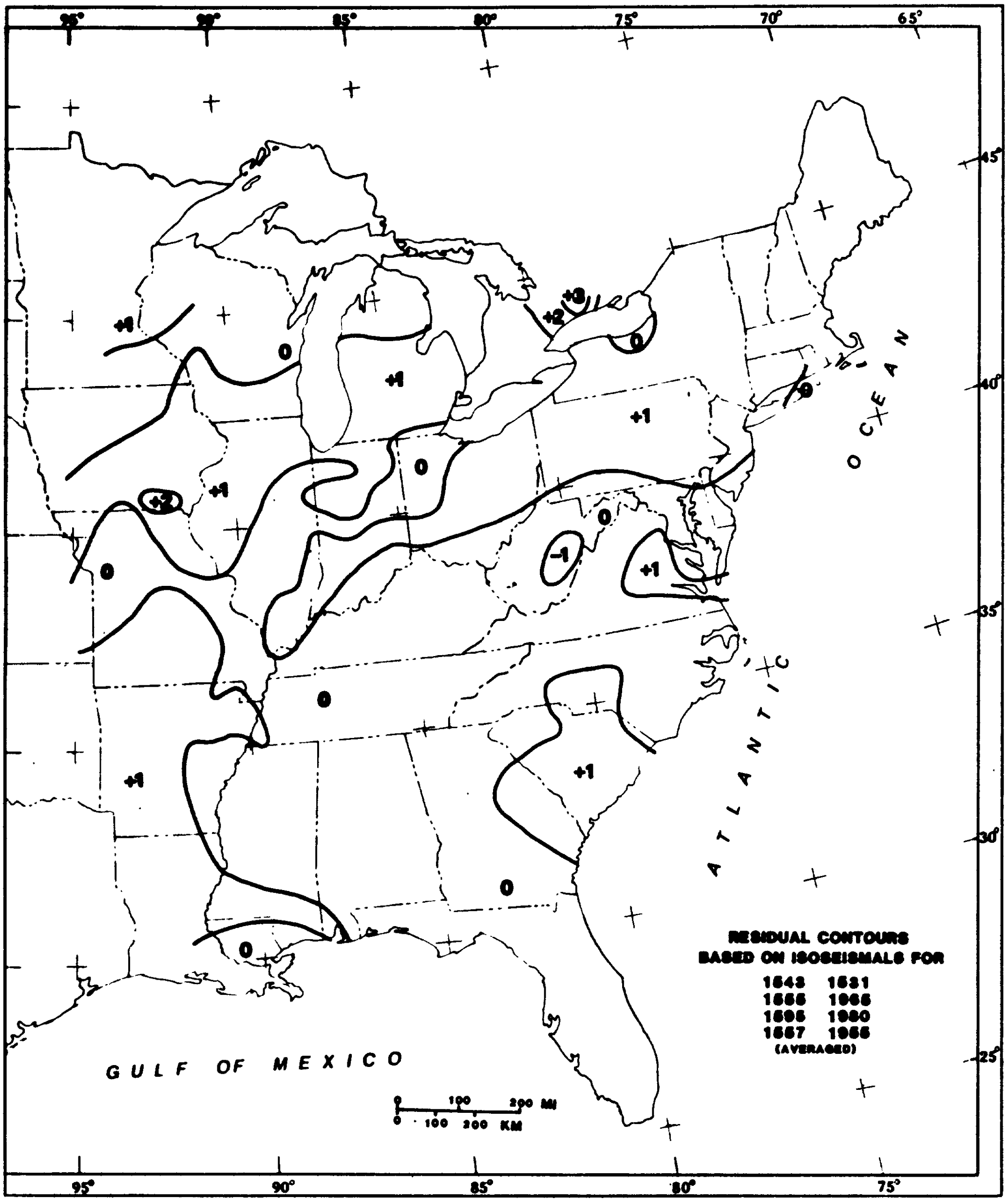


Figure 2.--Anomalous intensity in the central and southeastern United States based on the analysis of eight selected earthquakes in these areas since 1812. Contours separate areas having differences in Modified Mercalli intensities of one or more unit of intensity.

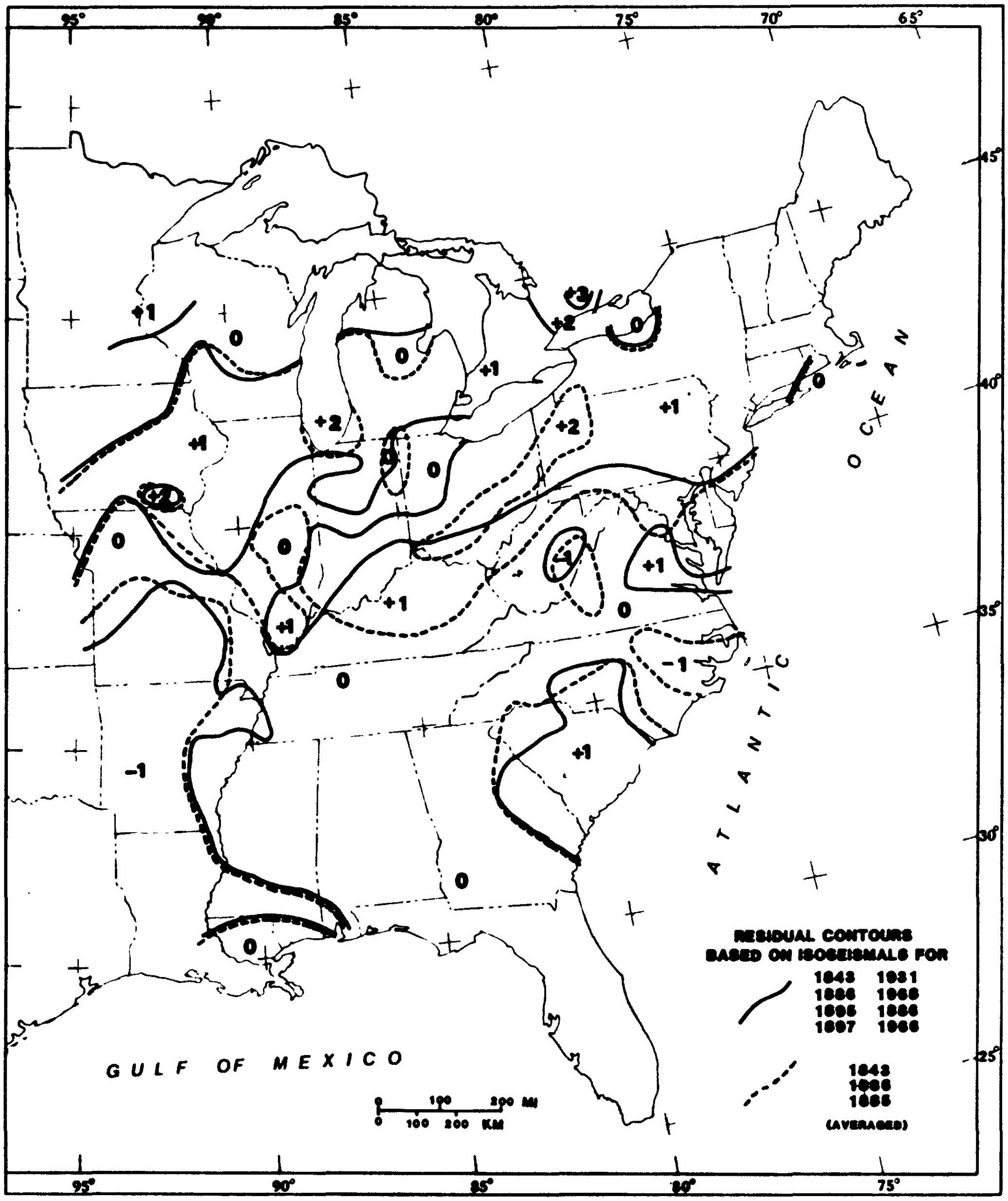


Figure 3.--Comparison of the areas of anomalous intensity in the central and southeastern United States based on studies of eight selected earthquakes and the three largest earthquakes in these areas since 1812. Contours separate areas having differences in Modified Mercalli intensities of one or more unit of intensity.

interpretation of the available data by various investigators and, second, there is uncertainty with regard to the area represented by the intensity assignment. Is the intensity assignment based on observed damage to a single or small group of buildings located in a small area or has the intensity assignment been averaged over a considerable area? These uncertainties are difficult to resolve and in many cases, it may not be possible to resolve unambiguously the areal extent of a particular intensity assignment in the literature because of lack of precision in the historical data. Despite the difficulties with the historical intensity data, an interesting result is that there appears to be correlation of the various anomalous intensity areas when they are averaged over several earthquakes.

What is the significance of these anomalous areas of intensity to seismic hazard analysis? Figure 4 shows the expected ground velocity in a 50-year period with a 90-percent probability of not being exceeded taken from Algermissen and Perkins (1976) and Algermissen and others (1982). This map does not contain parameter uncertainty although the effects of parameter uncertainty are discussed by Algermissen and others (1982). Inclusion of parameter uncertainty would increase the velocities shown in figure 4 by a factor in the range of about 1.5 to 2.0. Thus, as examples, the velocities (with uncertainty) in South Carolina would be approximately 10-14 cm/sec, in St. Louis, Missouri to about 9-12 cm/sec, and in northern Kentucky and central Ohio, about 3-8 cm/sec with parameter uncertainty included.

If the residual intensities shown in figure 1 for the three example areas are combined with the velocities (with parameter uncertainty) estimated from figure 4, velocities in the three example areas increase to approximately 20 cm/sec, 16 cm/sec and 6-24 cm/sec, respectively. The approximation

$$I = \frac{\log 14v}{\log 2} \qquad (3)$$

of Rosenbeuth (1964), where I is intensity and v is velocity in cm/sec has been used to convert the residual intensities to velocities.

352 40008

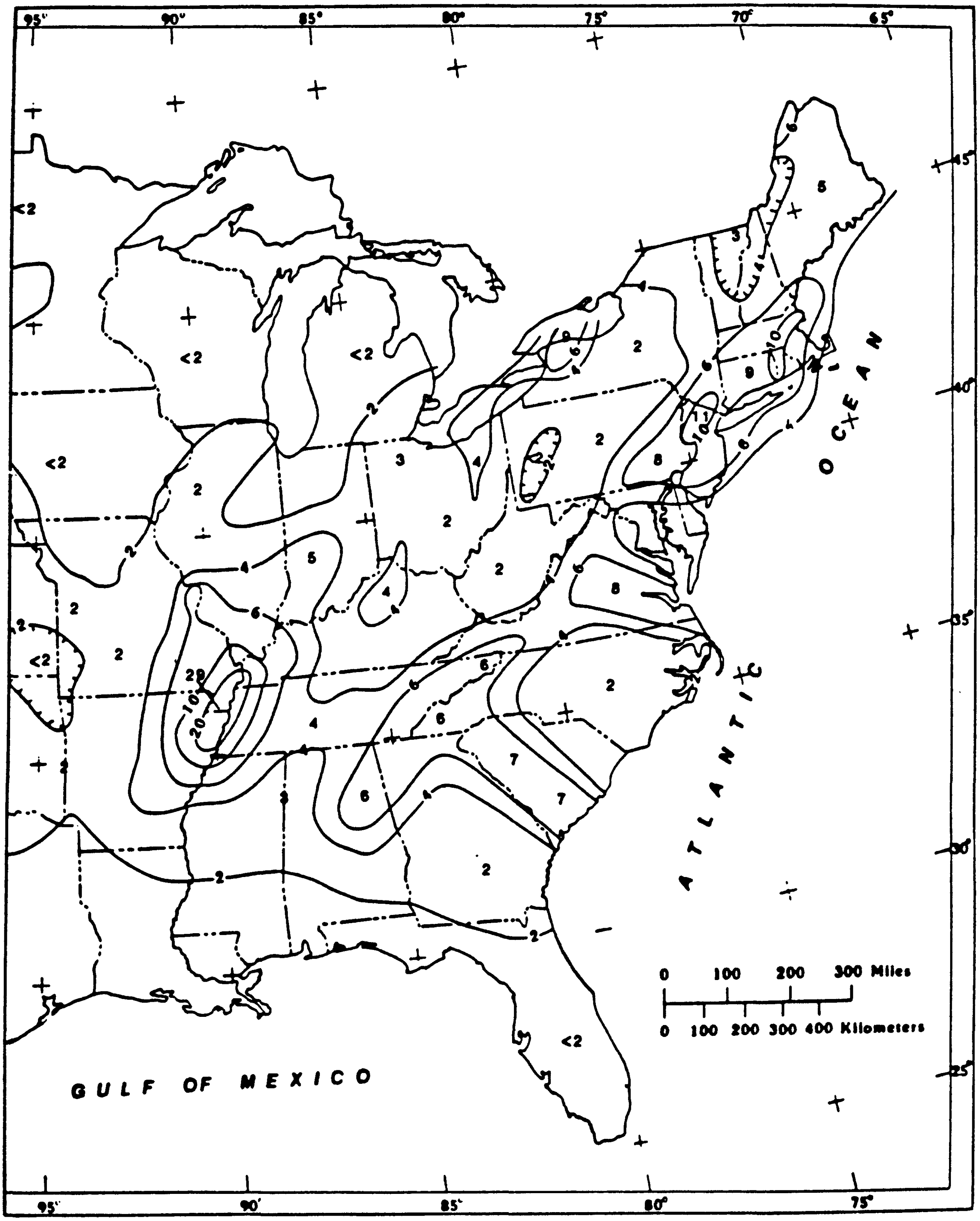


Figure 4.--Probabilistic ground velocity (cm/sec) in rock in the central and eastern United States for a 50-year exposure time and a 90-percent probability of not being exceeded (from Algermissen and others, 1982).

451 40008

Some regions not within important earthquake source zones (such as, St. Louis, Missouri, and northern Kentucky - central Ohio) would have appreciably increased expected velocities. Velocities within the important South Carolina source zone (fig. 5) are approximately doubled. This analysis of intensity data seems to indicate that the ground motion associated with subregional or local site response in the central and southeastern United States is about the same order of magnitude as the ground motion obtained in a conventional probabilistic ground motion assessment in which ground motions are reduced to a reference bedrock material. The analysis of the intensity data undertaken here was aimed at ascertaining whether or not intensity anomalies in the central and southeastern United States might affect expected ground motion assessments. A more quantitative review and analysis including a larger ensemble of intensity data will be necessary to more accurately detail both the areas of abnormal ground response and that amplitude of the response.

An important issue is whether the intensity anomalies represent reasonably large areas or essentially point sources. As previously noted, the areal extent of the intensity anomaly is difficult to resolve because of lack of precision of many historical intensity observations. A further generalization is introduced since intensity anomalies were obtained by estimating the observed intensities from isoseismals rather than at discrete observation points. However, estimation of residual intensities at points where intensity observations exist indicate that at a minimum some locations show intensity residuals of at least two intensity degrees.

SEISMIC SOURCE ZONE BOUNDARIES

Conventional regional probabilistic hazard analysis (Algermissen and Perkins, 1976; Algermissen and others, 1982) makes use of seismic source zones (see fig. 5) that (1) are assigned discrete boundaries; and, (2) have a uniform rate of earthquake activity throughout the source zone. In many cases adjoining source zones are assigned very different rates of seismic activity. When probabilistic ground motion maps are prepared using zones of this type, quite often there are sharp changes in ground motion from one zone to another along the conterminous zone boundary. These sharp changes in ground motion associated with seismic source zone boundaries are, in many

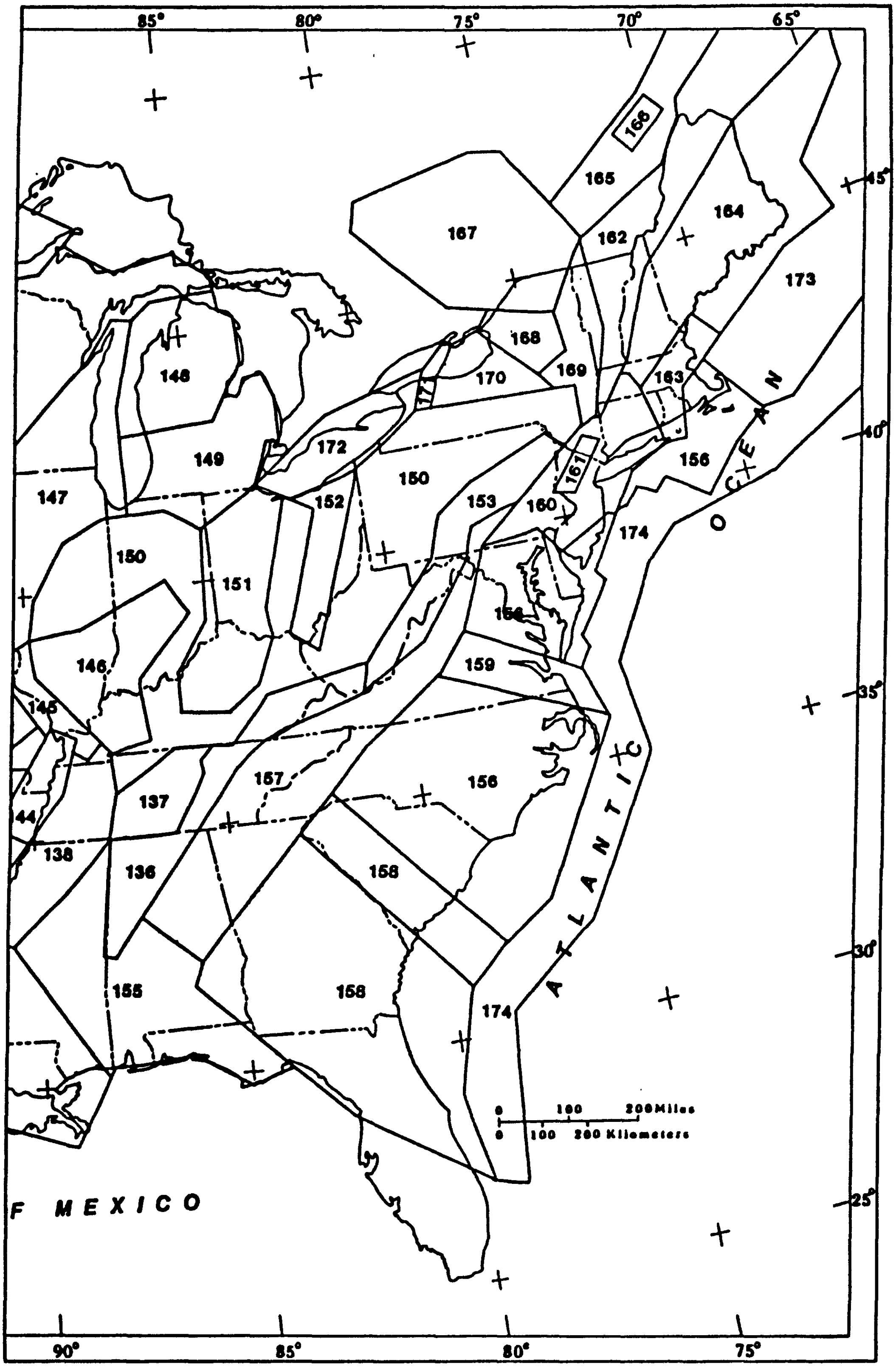


Figure 5.--Seismic source zones used by Algermissen and others (1982) in the eastern United States.

488 40008

instances, artificial, and may strongly affect sites located near or on the seismic source zone boundaries. In reality, most (but not all) seismic source zone boundaries are transition zones between seismic areas. This transitive nature of the boundary is not generally accurately depicted in conventional regional probabilistic hazard maps. Bender (1986) has recently discussed a technique for dealing with the problem of seismic source boundaries by assuming that the seismicity originally associated with each point in a source zone is normally distributed with standard deviation of σ about each point.

The technique suggested by Bender (1986) modifies source zone boundaries in an effective and generally realistic manner. It serves to remove the areas of large artificial changes in ground motion that frequently occur at zone boundaries. It will probably be necessary to develop additional techniques for modifying seismic source zone boundaries since earthquake location uncertainty is not the only kind of uncertainty contributing to the arbitrariness of zone boundaries. Difference in seismotectonic hypotheses also contribute to uncertainty. At any rate, it will probably be necessary to develop additional techniques for modeling seismicity consistent with advances in our understanding of seismotectonic models.

Modeling of Faults

For regional probabilistic ground motion assessment, earthquakes with magnitudes greater than about M_S=6.5 should be modeled as two-dimensional seismic sources rather than point sources in order to properly account for the spatial distribution of ground shaking. Important differences in expected ground motion occur depending upon the modeling technique used and the seismotectonic assumptions adopted. Modeling techniques differ substantially but have not been much discussed in papers on hazard analysis. For example, the fault modeling technique used by Algermissen and others (1982) assumes that no earthquake within a seismic source zone will rupture outside of the source zone. The source zones must be designed to accommodate this assumption. Some other techniques allow modeled faults to rupture to length specified only by a relationship of the form log L=f(M) where L is the length of the fault and M is the magnitude. Other modeling techniques use sequences of point sources to approximate two-dimensional rupture. Fault modeling techniques must be used very carefully so that they correctly represent the seismotectonic model preferred. Conversely, different seismotectonic assumptions regarding the spatial distribution of faults may also substantially affect the resulting ground motion. Some examples of these kinds of effects are given in Algermissen and others (1982).

Seismotectonic Hypothese--Minimal Data: Puget Sound Area

Recently, Heaton and Kanamori (1984) have suggested the possibility of very large, shallow subduction zone earthquakes at the Juan de Fuca-America plate boundary. No historical large plate boundary earthquakes are known in this region. Historically, all of the recent damaging earthquakes (1939, 1946, 1949, 1965) are believed to have occurred at depths of 40-70 km either within the region of bending of the subducted Juan de Fuca plate or near the plate interface.

In the Pacific Northwest, very little attention has been given to the possibility of a large $M_S \approx 7.0$, shallow earthquake, even though one is known to have occurred and there is other evidence of recent significant shallow activity. Evidence of the occurrence of an earthquake in 1872 with a magnitude of approximately 7.0 M_S has been extensively reviewed by a number of investigators, most recently by Hopper and others (1982) who believe that the earthquake was located near Lake Chelan, Washington and had a shallow focus (fig. 6). Other recent significant shallow activity has occurred in the Elk Lake (Grant and others, 1984) and Goat Rocks (Zollweg and Crosson, 1981) areas of Washington, and there is evidence of Holocene faulting west of the Hood Canal (Gower, 1978). A more conservative modeling of earthquake occurrence in the Puget Sound area with regard to shallow earthquakes was taken by Algermissen and others (1982) than by Algermissen and Perkins (1976). For the national ground motion maps developed in 1982, 25 percent of the earthquakes with M_S magnitudes greater than 6.5 were assumed to occur at shallow depth. The choice of 25 percent was, however, very arbitrary. All large shocks were assumed to occur at depths of 60 km in the development of the 1976 national map. Neither the 1982 maps or the 1976 map considered the possibility of a large plate boundary earthquake west of Puget Sound. Thus, there are large uncertainties in probabilistic ground motion assessment in the Puget Sound

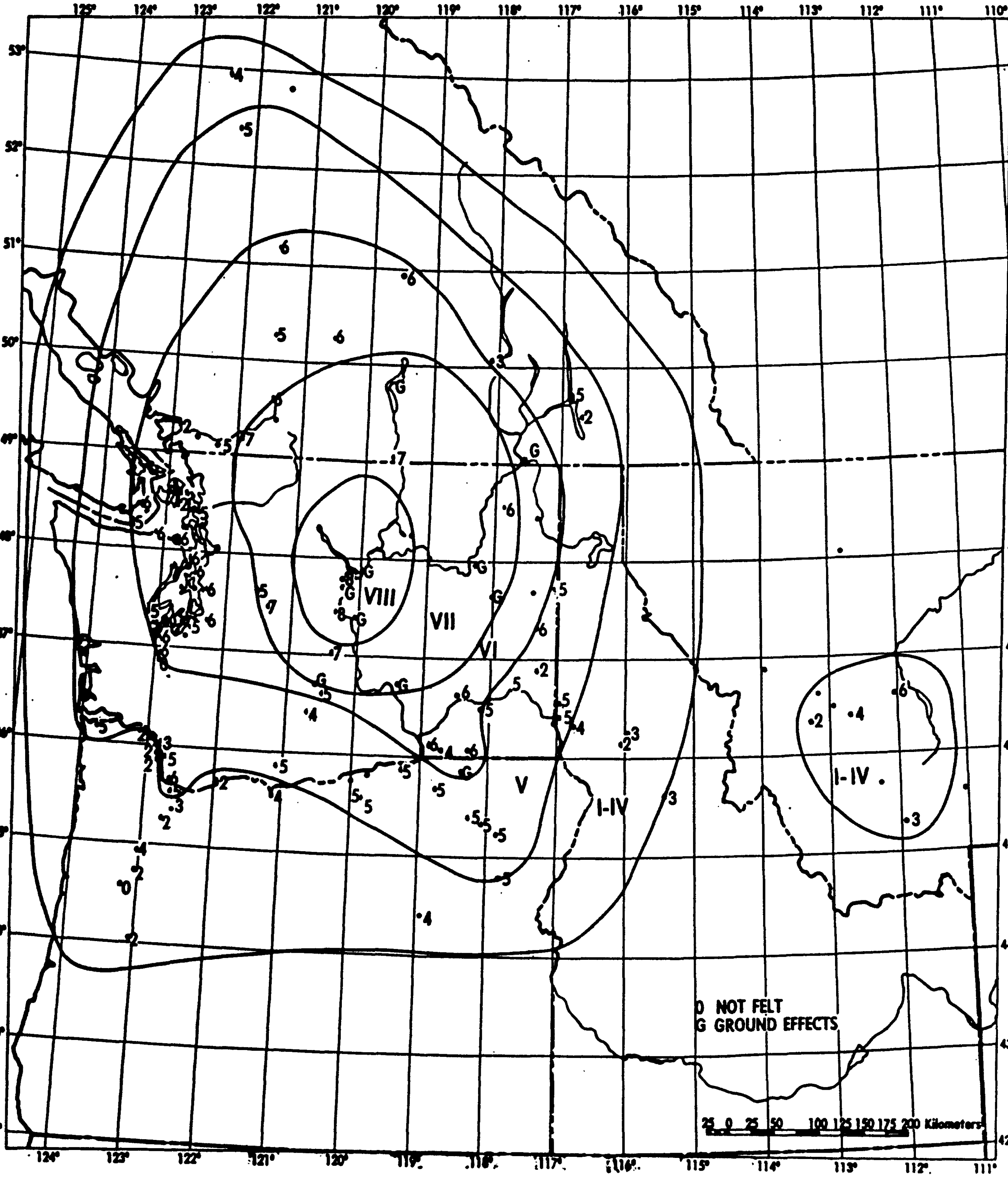


Figure 6.--Isoseismal map of the central Washington earthquake of December 14, 1872 (after Hopper and others, 1982).

484 40008

area, because little progress has been made in the development of a viable Examination of table 2 shows that only large ($M_S \approx 7.0$) earthquakes at depths of 40-70 km can be reasonably well modeled with the presently available data although it is difficult to restrict their spatial distribution. The two other postulated sources of potentially large ground motion in the Puget Sound area can only be modeled with very great uncertainty due to the lack of observational data in the region, both seismological and geological, that would serve to constrain the loci of potential large-earthquake sources.

Table 2.--Uncertainties in ground motion hazard assessment in the Puget Sound area

Hypothesis	Evidence
Very large plate boundary earthquakes $M_W \approx 9.0$ might occur.	No known historical or paleoseismic evidence. Conflicting views regarding the rate of subduction of the Juan de Fuca plate and the accumulation of strain. No adequate ground motion attenuation relationships for such an earthquake.
Large, shallow ($M_S \approx 7.0$) earthquakes might occur onshore.	Evidence of an $M_S \approx 7.0$ shock near Lake Chelan in 1872 but location and magnitude very uncertain (Hopper and others, 1982). Evidence of Holocene faulting west of the Hood Canal (Gower, 1978). Very limited available seismotectonic or seismological data to identify possible source areas of large shallow shocks.
Large ($M_S \approx 7.0$) earthquakes occur at depths of 40-70 km.	Well-documented historical shocks, but the possible spatial distribution is uncertain.

DISCUSSION

All of the sources of uncertainty discussed can produce order of magnitude changes in regional and national ground motion maps over areas of a few square kilometers up to areas as large as western Washington state. They are discussed here because they are considered important problems in earthquake hazard analysis. Considering the magnitude of the changes in estimated ground motion associated with the sources of uncertainty discussed here, they have received relatively little attention in the geophysical and engineering literature.

REFERENCES

Algermissen, S. T., 1983, An introduction to the seismicity of the United States: Earthquake Engineering Research Institute, 148 p.

Algermissen, S. T., Perkins, D. M., Thenhaus, P. C., Hanson, S. L., and Bender, B. L., 1982, Probabilistic estimates of maximum acceleration and velocity in rock in the contiguous United States: U.S. Geological Survey Open-File Report 82-1033, 99 p., 6 pl.

Algermissen, S. T., and Perkins, D. M., 1976, A probabilistic estimate of maximum acceleration in rock in the contiguous United States: U.S. Geological Survey Open-File Report 76-416, 45 p.

Applied Technology Council, 1978, Tentative provisions for the development of seismic regulations for buildings: National Science Foundation Publication 78-8, 505 p.

Bollinger, G. A., 1977, Reinterpretation of the intensity data for the 1886 Charleston, South Carolina, earthquake, in Rankin, D. R., ed., 1977, Studies related to the Charleston, South Carolina, earthquake of 1886--A preliminary report: U.S. Geological Survey Professional Paper 1028, p. 17-32.

Bollinger, G. A., and Hopper, M. G., 1971, Virginia's two largest earthquakes--December 22, 1875, and May 31, 1897: Seismological Society of America Bulletin, v. 61, no. 4, p. 1033-1039.

Campbell, M. R., 1898, Earthquake shocks in Giles County, Virginia: Science, v. 7, p. 233-235.

Coffman, J. L., and Cloud, W. K., 1970, United States earthquakes, 1968: U.S. Coast and Geodetic Survey, 111 p.

Gordon, D. W., Bennett, T. J., Herrmann, R. B., and Rogers, A. M., 1970, The south-central Illinois earthquake of November 9, 1968: Macroseismic studies: Seismological Society of America Bulletin, v. 60, p. 955-971.

Gower, H. D., 1978, Tectonic map of the Puget Sound region, Washington: U.S. Geological Survey Open-File Report 78-426, 22 p.

Grant, W. C., Weaver, C. S., and Zollweg, J. E., 1984, The 14 February 1981 Elk Lake, Washington, earthquake sequence: Seismological Society of America Bulletin, v. 74, p. 1289-1309.

Heaton, T. H., and Kanamori, H., 1984, Seismic potential associated with subduction in the northwestern United States: Seismological Society of America Bulletin, v. 74, p. 933-941.

Hopper, M. G., 1985, Historical seismicity of the Mississippi Valley, in Hopper, M. G., ed., 1985, Estimation of earthquake effects in the New Madrid seismic zone: U.S. Geological Survey Open-File Report 85-457, p. 7-30.

Hopper, M. G., and Algermissen, S. T., 1983, The earthquakes of January 5, 1843, and October 31, 1895, in the central Mississippi Valley [abs.]: Earthquake Notes, v. 54, no. 1, p. 7

Hopper, M. G., Algermissen, S. T., Perkins, D. M., Brockman, S. R., Arnold, E. P., 1982, The earthquake of December 14, 1872, in the Pacific Northwest: Program of the Annual Seismological Society of America Meeting.

Hopper, M. G., and Bollinger G. A., 1971, The earthquake history of Virginia 1774-1900: Virginia Polytechnic Institute and State University, 87 p.

Law Engineering Testing Company, 1975, Report on evaluation of intensity of Giles County, Virginia, earthquake of May 31, 1897: Marietta, Georgia, 94 p.

Neumann, Frank, 1932, United States earthquakes 1931: U.S. Coast and Geodetic Survey Serial No. 553, 26 p.

Nuttli, O., and Herrmann, R. B., 1978, State-of-the-art for assessing earthquake hazards in the United States--credible earthquakes for the central United States: U.S. Corps of Engineers, Papers 5-13-1, report 12, 99 p.

Rosenblueth, E., 1964, Probabilistic design to resist earthquakes: Proceedings, American Society of Civil Engineers, v. 90 (EMS), p. 189-219.

Schnabel, P., and Seed, H. B., 1973, Acceleration in rocks for earthquakes in the western United States: Seismological Society of America Bulletin, v. 63, p. 501-516.

Stover, C., and Von Hake, C., 1982, United States earthquakes, 1980: U.S. Geological Survey and National Oceanic and Atmospheric Administration, 182 p.

Weaver, C. S., and Smith, S. W., 1983, Regional tectonic and earthquake hazard implications of a crustal fault zone in southwestern Washington: Journal of Geophysics Research, v. 88, p. 10371-10383.

Zollweg, J. E., and Crosson, R. S., 1981, The Goat Rocks wilderness, Washington, earthquake of 28 May 1981: EOS, v. 62, no. 45, p. 966.

SEISMIC SOURCE ZONES IN PROBABILISTIC ESTIMATION OF THE EARTHQUAKE GROUND-MOTION HAZARD: A CLASSIFICATION WITH KEY ISSUES

by

Paul C. Thenhaus
U.S. Geological Survey
Denver, Colorado

INTRODUCTION

A fundamental initial step in probabilistic earthquake hazard analysis is delineation of seismic source zones and identification of seismically active faults. Seismic source zones define areas that share common seismologic, tectonic, and (or) geologic attributes under the assumption that these areas also share similar seismotectonic origins of seismicity that can be described by a unique magnitude-frequency relation. The map thus defines (1) the historical earthquake data base to be used in statistical analyses of earthquake recurrences, (2) the particular geographical distribution of expected future earthquakes and, in so doing, (3) the probable future distribution of earthquake ground motions.

Recent efforts to define seismic source zones for regional hazard assessment reflect an increasingly heavy reliance on available tectonic and paleoseismic data to establish the location of source zone boundaries (Thenhaus, 1983; Algermissen and others, 1982); the equivocal association of historical earthquakes with geologic structure and lack of insight into recently active structures and neotectonic processes throughout much of the United States preclude reliance on verifiable seismotectonic models. Empirical approaches based primarily on qualitative spatial association predominate and, hence, the definition of particular seismic source zones admits to a wide range of possible interpretations.

PREMISE OF SOURCE ZONE TYPES

The purpose of defining seismic source zones for hazard analysis is to model likely contrasts in the future distribution of seismicity. The model may be regional or local in scope and may or may not include distinctions in maximum magnitude earthquakes among zones although rates of activity may vary significantly. The degree to which available seismologic, geologic, and geophysical information can be effectively applied to this task is extremely variable due in large part to differences in both the long-term seismotectonic research effort expended and the relative ease with which a region yields seismotectonic insights. The integrated results of these two factors among a variety of tectonic settings is a broad range of certainty with which earthquakes can be associated with causal faults or geologic structures. Figure 1 schematically illustrates four types of seismic source zones in terms of developing seismotectonic knowledge for various regions of the United States. The body of knowledge relating to any of the types of source zones is divided into a seismic history and a structural geologic history. The degree to which these two histories are completely known not only constitutes the certainty with which causal geologic structures can be associated with seismicity, but also determines the primary methodologies by which earthquake recurrences and estimated maximum magnitudes are determined. Critical gaps in these histories impose assumptions on any hazard evaluation.

IDENTIFYING ISSUES AMONG ZONE TYPES

Key issues among source zone types in probabilistic ground-motion hazard analysis relate in a complex manner to (1) the map scale of the hazard investigation (that can range from large map-scale, site specific to small map-scale, regional studies), (2) the desired probability level of the ground-motion estimate (annual exceedance probability), and (3) the rate of earthquake activity in the region of concern. Study of the consequences on estimated ground motion of statistical variability in estimated seismicity parameters and ground-motion attenuation models is the purpose of hazard sensitivity studies that explicitly investigate methodological procedures and parameter assumptions (for example, Bender, 1983; 1984a,b). These studies reveal that (1) while the consequences due to issue-related decisions may vary

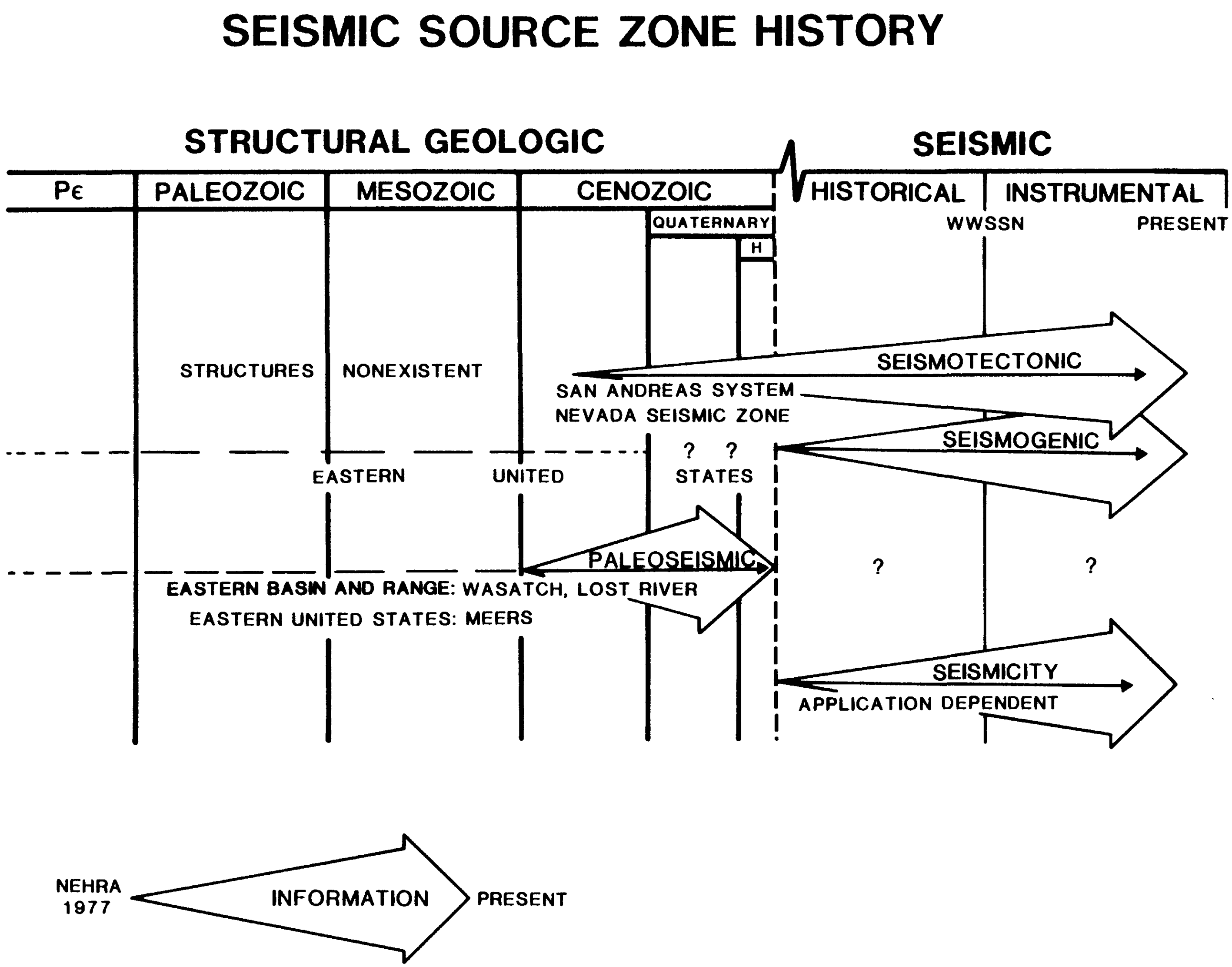


Figure 1.--Schematic of the relative preponderance of available geological and seismological information developed under the National Earthquake Hazards Reduction Act (NEHRA) for four types of seismic source zones.

453 40008

among methodological procedures, these issues remain important regardless of methodology, and (2) not all geological issues impact ground-motion hazard estimates significantly in all hazard applications. For example, concerning maximum estimated magnitudes, Bender (1984a) states, "Because high accelerations can result from high magnitude earthquakes, the selection of m_{max} has been of considerable interest in seismic hazard analysis. When a single acceleration is associated with earthquakes of a given magnitude and distance, a maximum acceleration a_{max} from earthquakes along the fault will be calculated at a site. This maximum acceleration is produced only by ruptures of magnitude m_{max} earthquakes that include the point on the fault nearest the site. If the maximum magnitude is increased to $m_{max\ (new)}$, occurrences of accelerations near a_{max} may be greatly increased, and higher accelerations will be possible. However, for accelerations considerably below a_{max}, m_{max} will have a much smaller effect. Earthquakes at lower magnitudes will produce a high fraction of the lower accelerations, and changes in m_{max} will have a smaller effect as the acceleration decreases." The size of the effect is dependent upon the attenuation function used. Nonetheless, for source zones having a relatively low rate of seismic activity, high exceedance probability ground-motion estimates (for example, 1 in 500) are not critically sensitive to choice of maximum magnitude. Conversely, maximum magnitude can become a critical factor in the hazard estimate at low exceedance probabilities (for example, 1 in 5,000) depending, of course, on the rate of activity in the source zone.

In the following section that describes the source zone types, issues impacting ground-motion hazard estimates order themselves around the current state of knowledge concerning seismotectonic processes among regions. The issues among the source zone types range from unknown earthquake causal faults (a virtual lack of knowledge) to choices among predictive hazard models (the best available knowledge). The classification implicitly includes a qualitative measure of rate-of-seismic-activity and, therefore, aids the identification of issues relative to changing exceedance probabilities. It further defines primary methodologies that are currently available for dealing with seismic parameter estimation. As in any classification, examples can be (and are) cited that fall into "gray areas," not fitting well into any defined slot. This is not troublesome as it merely indicates our growing knowledge regarding the presently operating seismotectonic processes in the area of concern. Considerable statistical illustration could have been added showing the change of ground-motion values with changing seismicity parameters, different source zones, different attenuation functions, different hazard models and methodologies, etc., as in the study area of sensitivity analysis. However, that would obscure a primary objective of the paper which is to provide a general geological context precisely for those types of studies.

In the general outline of methods and key issues that follows, "reliable" means that a particular method with its inherent uncertainties (whether measurable or not) is dependable in predicting activity rates and maximum magnitudes. "Important" means that the issue can influence the hazard results to a moderate degree relative to the stated conditions of rate-of-activity and exceedance probability level. "Critical" means that the issue can influence the hazard results to a great degree relative to the stated conditions of the rate-of-activity and exceedance probability level.

CHARACTERISTICS AND ISSUES OF ZONE TYPES

Seismotectonic Zones

A seismotectonic zone is a seismic source zone in which a causal relationship has been established between a geologic structure (usually a fault) and earthquakes. Processes of earthquake mechanism and generation can be studied from both a structural geologic aspect and a seismological aspect. There is continuity between the seismic and structural histories of the zone and knowledge of these histories is developing simultaneously (fig. 1). The main task in the hazard analysis is to characterize the future temporal and spatial occurrence of earthquakes on the known structures(s). Recently, Lindh (1983) and Sykes and Nishenko (1984) calculated probabilities of the near-future occurrence of large earthquakes on the San Andreas fault and selected other faults of the San Andreas fault system (conditional probabilities for large, fault-rupturing earthquakes ranging in size from M_W 6.0 to M_W 7.9 for future time periods of interest). Their procedures incorporated a fault segmentation model based primarily upon rupture extents of historic earthquakes but also

4B6 40008

incorporated (1) distinctions in geologically and geodetically determined slip rates along the faults, (2) estimates of repeat times of the fault-rupturing events described by a Gaussian distribution with standard deviation 0.33 about the mean, and (3) the date of the last large earthquake. Estimates of the ground motion that would accompany large, fault-rupturing events have the same probability of occurrence as the events themselves. However, because such estimates are discreetly associated with a single earthquake event on a single fault segment, they obviously do not describe the full amplitude-frequency distribution of possible ground motions at a site and, hence, are not probabilistic ground-motion estimates. Needed for a standard probabilistic representation of ground-motion hazard are: (1) a magnitude-frequency relation for earthquakes smaller than segment-rupturing events, if such earthquakes cannot be ruled out; (2) an upperbound magnitude that is possible on each fault segment (this would be derived from a scenario of multiple segments rupturing in a single event); (3) a magnitude-frequency relation (or probability distribution) relating the upperbound events to the segment-breaking events; (4) a fault rupture length-magnitude relation (using moment magnitudes); (5) representation and magnitude-frequency characterization of all geologic structures within some radius of the site that could contribute damaging ground motions (however defined) at the site.

Bender (1984a) has shown that ground motion at sites near the terminous of modeled faults are highly sensitive to fault-model assumptions. With respect to a fault consisting of a single segment located on the x-axis and extending from $0 \leq x \leq L$, she states, "As the site location (X,P) is moved parallel to the fault from the center of the fault to the end, the acceleration with a fixed return period may decrease by 50 percent. Much of the decrease occurs as the site moves to within 10 or 20 km of the end of the fault. Moving the site past the end of the fault another 10 km parallel to the extended fault line may result in another 25-percent decrease in acceleration level. As P, the perpendicular distance from the site to the fault, increases, acceleration values become less sensitive to the x coordinate of the site." The effects Bender describes result from a fault-contained rupture model but are analogous to the situation where an inferred fault segment boundary juxtaposes two segments having highly contrasting rates (or probabilities) of earthquake occurrence. Inspection of the conditional probability maps of Sykes and

Nishenko (1984) and Lindh (1983) indicate such a boundary located on the San Francisco peninsula (boundary between segments 3-4 of Sykes and Nishenko, and the San Francisco peninsula-San Juan Bautista boundary of Lindh). Significant to ground-motion hazard estimates in the western San Francisco Bay region, the location of that segment boundary, as well as the length of the high-potential segment, differs between the two interpretations. The point is, that locations of segment boundaries are interpretive and not unique but are potentially dominant influences on the distribution of ground motions for areas located near boundaries that juxtapose segments of highly contrasting rates (or probabilities).

Typically, seismotectonic zones are highly active and would be characterized by high regional hazard even without geological investigations. However, refined large map-scale hazard studies and predictive hazard methodologies have little basis without them. Benefits of geological investigations are:

(1) accurate determination of active fault locations,
(2) compilation of prehistoric fault-rupturing events,
(3) determination of age of last faulting event,
(4) determination of fault-slip rates,
(5) determination of changes in fault attitude or strike, which holds potential for accurate determination of fault-segment ends (King and Yielding, 1984; King and Nabelek, 1985).

Seismotectonic zones need not be confined to single, through-going faults as the San Andreas. Indeed, much work in the Nevada Seismic Zone (Wallace, 1984; Van Wormer and Ryall, 1980) illustrates the seismotectonic zone type. With the seismotectonic zone unequivocally identified, primary methods for estimating maximum magnitude with currently available data are:

Method	Comments
1. Historical record	Reliable in high-rate zones.
2. Magnitude-on-fault length regressions.	Generally reliable as data are from seismotectonic zones.
3. Analogous tectonic settings.	Reliability unknown. What constitutes "analogous"?

551 40008

4. Models from geodynamic and mechanical principles. | Reliability variable: 1. Models waiting for verification in future earthquake occurrence. 2. Models virtually verified (for example, circum-Pacific seismic gaps).

Primary methods for establishing recurrence with currently available data are:

Method	Comments
1. Historical record.	Reliable regional averages; inadequate for fault-specific rates.
2. Paleoseismic faulting data.	Reliable for large-earthquake fault-specific rates.
3. Seismic moment.	Reliability variable. How well are fault dimensions, slip rate, b-value and maximum moment known?

Key hazard-model issues.--The following issues are considered critical because these zones are typically located in high-rate areas where ground-motion values near active faults are highly sensitive to fault-model assumptions. Modeling a linear fault source is assumed.

* Fault segmentation definitions and possible rupture between adjacent segments.
* Maximum magnitude assessment for individual segment ruptures as well as for potential multisegment ruptures.
* Stochastic recurrence model versus time-predictable versus slip-predictable models for particular sizes of earthquakes and incorporation into magnitude-frequency distributions of moderate-to-large earthquakes (say, M_W 5 to M_W 8).
* Fault- and segment-end location uncertainty.

Paleoseismic Zones

Paleoseismic zones are those zones having an important Quaternary-Holocene structural history that indicates they constitute a seismic threat in the future. However, these zones lack a seismic history (fig. 1). Faults having Holocene displacements in the eastern Basin and Range, including most of the Wasatch fault, as well as faults in other regions such as the Meers fault of southern Oklahoma (Gilbert, 1985) fall into this category. The Lost River fault, having Holocene displacement but no recorded seismic activity prior to 1983, just recently ruptured in the 1983, M_S-7.3 Borah Peak, Idaho, earthquake (Crone and Machette, 1984; Scott and others, 1985). The zone thus could be classed marginally as a seismotectonic zone. The marginal classification is due to the meager seismic history of the fault.

To motivate a key issue concerning treatment of paleoseismic data, consider a simple model of the Wasatch fault given the fault segments and segment lengths of Schwartz and Coppersmith (1984) and their preferred average recurrence interval of 444 years for large earthquakes along the fault. Assuming that the six segments break their entire lengths independently and randomly in earthquakes 6.75-7.75 (M_S) with uniformly distributed rates in that magnitude range, calculated acceleration values (using the same attenuation function as in Algermissen and others, 1982) at sites near the fault are consistently lower than values in Algermissen and others (1982). At a 10-percent exceedance probability for exposure times of 10 and 50 years, acceleration values are a factor of 10 less than values in Algermissen and others (1982). At a 250-year exposure time, differences in values are less than 10 percent. The point is, that although the geologic recurrence estimates of large earthquakes along the Wasatch fault are an order of magnitude higher than recurrences estimated from the historical catalog (Bucknam and Algermissen, 1984; Schwartz and Coppersmith, 1984), the seismic hazard for short-exposure times is not necessarily increased above existing estimates. Depending on assumptions applied, substantially lower hazard estimates could result. Critical in this regard is the treatment of low-to-moderate magnitude earthquakes in the recurrence relationship. The above illustration excludes earthquakes $4.0<M_S<6.75$; a literal interpretation of the contemporary earthquake history along both the Wasatch and Lost River faults. It is characterizing the recurrence of, or perhaps exclusion of, earthquakes in this range of magnitudes that will most significantly influence high-exceedance probability hazard estimates. Unfortunately, recurrences of earthquakes of this size along these faults will be the most difficult to resolve satisfactorily as there is no reason to expect unequivocal identification of such events in fault-trenching studies, and historical and instrumental records indicate a dearth of such activity, except along the southernmost segment of the Wasatch (Arabasz and others, 1980; Zoback, 1983).

Primary hazard-related issues for paleoseismic zones are those of characterizing future recurrences of earthquakes based virtually upon geologic history. The characteristic earthquake model (Shwartz and Coppersmith, 1984) was derived from a need to reconcile the geologic history with the seismic history of individual faults. However, for relatively high-exceedance probabilities (say, 1 in 500), ground-motion exceedance contributions are dominated by the more frequent low-to-moderate magnitude earthquakes. In paleoseismic zones having no seismic history, the problem of characterizing the recurrence of events smaller than fault-rupturing earthquakes persists and the paradox arises that low-probability exceedance ground motions can be estimated more accurately than high-probability exceedance ground motions.

At least some geological data needs to be available to even identify a paleoseismic zone. Primary methods for estimating maximum magnitudes with the available data are:

Method	Comments
1. Magnitude-on-fault length or displacement regressions.	Reliable if regression data is from the same structural province, otherwise reliability unknown.
2. Analogous tectonic settings.	Reliability unknown. What constitutes "analogous"?

Primary methods for establishing recurrences with currently available data are:

1. Paleoseismic faulting data.	Reliable for large-earthquake, fault-specific rates.
2. Historical record.	Reliable for average regional rates, inadequate for fault-specific rates.

Key hazard-model issues.--Modeling a linear fault source is assumed.

* Maximum magnitude assessment. Important in all hazard assessments. May be critical in low-exceedance probability estimates depending on rate of activity and exceedance probability.

554 40008

* Appropriate use of geologic recurrence estimates for the hazard level being estimated. Recurrence models that fit geologic data on large earthquakes do not necessarily result in "conservative" ground-motion values at high-probability exceedance estimates.

* Fault segmentation based on paleoseismic information alone: Is the entire fault active if only one segment has Holocene displacement? If so, how active is it? (Example: Meers fault). Important for all hazard estimates. Critical for low-probability exceedance estimates.

* Fault domains: Should all faults of a fault domain (however defined) be considered active if one has Holocene displacement? If so, how active are they? (Example: Wichita frontal fault zone.) Important for all hazard estimates. Critical for low-probability exceedance estimates.

* Fault- or segment-end location uncertainty. Important in all hazard estimates at sites located near a fault or segment terminous. May be critical for these sites in low-probability exceedance estimates depending on rate of activity and exceedance probability.

Seismogenic Zones

A seismogenic zone lacks development of a clear history relating contemporary seismic activity to geologic structure (fig. 1). Critical gaps in the Quaternary geologic history preclude direct evidence of active faulting and may be due to a number of reasons:

(1) lack of geologic investigations aimed at identifying young fault movement,

(2) unfavorable geologic conditions for preservation of evidence of geologically young fault movement,

(3) lack of surface displacement from fault rupture at depth, etc.

Seismogenic zones are, by far, the most common type of source zone employed in probabilistic hazard analyses. Commonly, seismogenic zones are area sources, but the zone type applies also to inferred associations of seismicity with individual faults. Probabilistic methodologies can subdivide the seismogenic-zone type into a variety of zone classifications to more explicitly describe probabilistic treatments of possible causal structures. Seismogenic zones are obviously nonunique. Their ubiquitous use stems from the oft-cited fact that seismotectonic processes and deep-crustal structure of large intraplate regions are so poorly known that mere identification of earthquake-causal structures is, at best, highly uncertain. Inferences relating earthquakes with structure are, therefore, based on judgement which, in turn, is most often based either implicitly or explicitly on analogy of geologic or tectonic setting; the only guideline being not to grossly disrupt the historical regional seismicity pattern. For example, in the west-central United States a poor understanding of seismotectonics dictates that the spatial pattern of seismicity serve as the primary guide for the definition of seismic source zones. However, a spatial correlation appears to exist between much of the activity and high-basement features. This association has been noted by others; in the Great Plains (Becker and Zeltinger, 1983; Brill and Nuttli, 1983) and the Cincinnati arch of the east-central interior (Barstow and others, 1981). Microseismicity has been noted along the Chadron-Cambridge arch of southern Nebraska, the central Kansas arch and Nemaha ridge (Steeples, 1978; Rothe and others, 1981; Steeples and others, 1979). Although some of the activity in the west-central region can be related to oil field pumping, deeper events are thought to be tectonic. Boundaries of source zones in Algermissen and others (1982) were extended along these basement features where the basement structures could be associated with at least a number of low-intensity earthquakes. Intervening areas between these zones show a markedly lower frequency of earthquake occurrence. The area-normalized rate of seismic activity of the high-basement zones is an order of magnitude higher than the rate for intervening areas (excluding zones in the southern Illinois basin and Mississippi Embayment). While the rates cannot be construed as proof of the association, the association is compelling because of its regional persistence and therefore serves as a useful independent guide to source zone boundaries. As there is no generally accepted seismotectonic model accounting for the association, the issue remains whether all areas of high-basement features should be included in the higher rate zones (i.e., even those high-basement areas appearing aseismic historically).

Seismogenic zones are typical of most seismic source zones developed for the Central and Eastern United States. However, the New Madrid seismic zone is distinctive in that a sketchy Quaternary-Holocene structural history has been developed for the area (Russ, 1979; 1982), and a structurally disturbed zone has recently been recognized along the length of the well-defined seismicity trend (Crone and others, 1985). Such aspects approach those of a seismotectonic zone; however, the fact remains that the causal fault (or faults) of the 1811-12 earthquake sequence has yet to be identified and the overlap in the seismic and structural histories is still vague enough to preclude a meaningful analysis of temporal and spatial seismicity characteristics.

Primary methods of estimating maximum magnitudes with currently available data are:

Method	Comments
1. Historical record.	Poor reliability due to short-time period relative to recurrence time. Subjective "conservative" estimates predominate.
2. Analogous tectonic settings.	Reliability unknown. What constitutes "analogous"?

Primary methods of establishing recurrences with currently available data are:

1. Historical record.	Reliable average regional rates.

Key hazard-model issues.--Seismogenic zones assume geologic associations and typify areas of low-to-moderate earthquake activity. The following issues are considered important for all hazard investigations. Issues may become critical for low-exceedance probability estimates depending on rate of activity and exceedance probability. Issues are critical for such estimates if local zones of high seismic activity are defined.

* Maximum magnitude assessment.

* Appropriate balance between circumstantial geologic arguments and spatial distributions in the seismic history. To what extent should speculative hypotheses be allowed to perturb the historic spatial distribution of seismicity on (1) a regional scale? (2) On a local scale?

* Lacking knowledge of definite earthquake-causal structures, regional consistency in approach to seismogenic zone delineation arises as one measure of the reasonableness of the seismic source zone map. Does spatial association with one

distinct structural or tectonic feature imply seismic potential of other similar features that have no spatially associated seismicity? Does this judgement change when several or more distinct features have spatially associated seismicity (examples: high-basement features of the central interior, mafic plutons along the eastern seaboard)?

* In light of (1) the limited seismological data, (2) the limited methods available for establishing maximum magnitudes and recurrences, and (3) the likelihood that earthquake behavior of individual seismic faults does not conform to a simple exponential magnitude-frequency distribution, are seismogenic fault-specific hazard estimates meaningful?

* Boundary location uncertainty assuming geologic associations. Discrete zone boundaries are only approximations. For example, considering the broad basement arches of the central interior, the question arises as to where the arch ends and the basin begins with respect to the seismogenic zone boundaries. Considering, for example, mafic plutons as seismogenic zones, it should not be the pluton itself that defines the seismogenic zone boundaries, but rather, some inferred area of stress amplification around the pluton in which stresses decrease with distance from the pluton.

Seismicity Zones

Seismicity zones are those seismic source zones that do not assume any relations with geologic structure. They are defined solely on the spatial distributions of the seismic history and their use and reasonableness can only be judged relative to the intended use of the final hazard estimate. They serve a legitimate purpose by providing useful hazard guidelines when available seismotectonic information is irreconcilable with accepted seismologic and tectonic theory. In that respect, seismicity zones are no better nor worse than seismogenic zones, the only distincting being that seismogenic zones use guides independent of seismicity. Fundamental issues involve the integrity of, and indiscriminate use of the historical earthquake catalog. The identification of such zones could easily be biased by local, temporary or long-term earthquake monitoring and historical variations in population and settlement.

The "Charleston earthquake problem" of the eastern seaboard has been discussed at great length (Hays and Gori, 1983; Dewey, 1985) and is the leading example of difficulties involved with delineating seismicity zones. In conclusions drawn from a worldwide search for intraplate earthquakes that might provide tectonic and seismologic insights to the Charleston earthquake, Dewey (1983) states, "Data from the other midplate source regions suggest that the Charleston region is more likely to experience a strong earthquake in future decades than a random midplate site, but that strong Eastern United States earthquakes will also occur in the future at sites that have not previously experienced strong earthquakes. Data from the other regions do not provide conclusive seismological or geological guidelines for identifying sources of future strong earthquakes in the absence of a historical record of strong earthquakes." The difficult and presently intractable question posed in the delineation of seismic source zones is: What should be the relative balance of hazard between the Charleston area and the large Mesozoic extensional province in which it is located acknowledging the fact that large earthquake occurrences cannot be ruled out provincewide, but available data does not provide guidelines sufficient for identification of future large-earthquake sources? The answer is judgemental.

A strict use of seismicity zones (i.e., no generalization of the seismic history) results in localized high hazard for sites that have historically experienced a large earthquake and does not attempt to identify other areas that may be susceptible to similar-sized events. Broad seismogenic zones encompassing seismicity zones with all zones having the same maximum magnitude, such as the eastern seaboard zones of Algermissen and others (1982), is one attempt at addressing this problem. Nonetheless, area-normalized rates of activity are the dominant influence on the ground-motion values and it is determined by the judgemental seismicity zones.

Primary methods of estimating maximum magnitudes and recurrences are the same as for seismogenic zones.

Key hazard-model issues.--The following issues are considered critical. The arbitrary nature of seismicity zone delineation can lead to extreme contrasts in area-normalized activity rates that can greatly influence the hazard estimate.

* Maximum magnitude assessment in high-rate zones.

* Appropriate use of seismicity zones relative to the intended application of the hazard estimate.

SUMMARY

Issues concerning the delineation of seismic source zones can be subsumed under a classification of four types of seismic zones. These types are: (1) seismotectonic zones, (2) paleoseismic zones, (3) seismogenic zones, and (4) seismicity zones. Each is defined by a different level of understanding concerning seismic faulting and seismotectonic processes. Issues concerning each source zone type gain or lose importance relative to the map scale of the investigation, the desired probability level of the ground-motion estimates, and the rate of earthquake activity.

The regions and faults discussed in the text are intended to be examples of type localities of the different seismic source zones but are not intended to be a comprehensive inventory. It should not be assumed that only one type of source zone persistently characterizes each of the regions. The nature of successful, multidisciplinary geologic investigations of earthquake hazards is one of concentrating efforts on a subregional, even local level. Hence, it is not unusual for optimum use of available information to result in a regional mix of source zone types.

REFERENCES CITED

Algermissen, S. T., Perkins, D. M., Thenhaus, P. C., Hanson, S. L., and Bender, B. B., 1982, Probabilistic estimates of acceleration and velocity in the contiguous United States: U.S. Geological Survey Open-File Report 82-1033, 99 p. 6 pl., scale 1:7,500,000.

Arabasz, W. J., Smith, R. B., and Richins, W. D., 1980, Earthquake studies along the Wasatch Front, Utah; network monitoring, seismicity, and seismic hazards: Seismological Society of America Bulletin, v. 7, p. 1479-1499.

Barstow, N. L., Brill, K. G., Jr., Nuttli, O. W., and Pomeroy, P. W., 1981, An approach to seismic zonation for siting nuclear electric generation facilities in the Eastern United States: U.S. Nuclear Regulatory Commission Report Nureg/CR-1577, 143 p.

Becker, D. J., and Zeltinger, J. M., 1983, Comment on Seismicity of the Colorado lineament, by K. G. Brill, Jr., and O. W. Nuttli, 1983 (in Geology, v. 11, p. 20-24: Geology, v. 11, p. 559-560.

Bender, Bernice, 1983, Maximum likelihood estimation of b-values for magnitude grouped data: Seismological Society of America Bulletin, v. 73, p. 831-851.

______1984a, Seismic hazard estimation using a finite fault rupture model: Seismological Society of America Bulletin, v. 74, p. 1899-1923.

______1984b, Incorporating acceleration variability into seismic hazard analysis: Seismological Society of America Bulletin, v. 74, p. 1451-1462.

Brill, K. G., Jr., and Nuttli, O. W., 1983, Reply to Comment on Seismicity of the Colorado lineament by D. J. Becker and J. M. Zeltinger, 1983, (in Geology, v. 11, p. 559-560): Geology, v. 11, p. 559-560.

Bucknam, R. C., and Algermissen, S. T., 1984, A comparison of geologically determined rates of seismic activity and historic seismicity data in the Great Basin, western United States, in A collection of papers of the International Symposium on continental seismicity and earthquake prediction, September 8-14, 1982: Seismological Press, Bejing, China, p. 169-176.

Crone, A. J., McKeown, F. A., Harding, S. T., Hamilton, R. M., Russ, D. P., and Zoback, M. D., 1985, Structure of the New Madrid seismic source zone in southeastern Missouri and northeastern Arkansas: Geology, v. 13, p. 547-550.

Crone, A. J., and Machette, M. N., 1984, Surface faulting accompanying the Borah Peak earthquake, central Idaho: Geology, v. 12, p. 664-667.

Dewey, J. W., 1983, A global search of continental midplate seismic regions for scientific characteristics bearing on the 1886 Charleston, South Carolina, earthquake, in Hays, W. W., and Gori, P. L., eds., Proceedings of Conference XX, A workshop on The 1886 Charleston, South Carolina, earthquake and its implications for today, May 23-26, 1983, Charleston, South Carolina: U.S. Geological Survey Open-File Report 83-843, p. 391-426.

______1985, A review of recent research on the seismotectonics of the southeastern seaboard and an evaluation of hypotheses on the source of the 1886 Charleston, South Carolina, earthquake: U.S. Nuclear Regulatory Commission Report NUREG/CR-4339, 44 p.

Gilbert, M. C., 1985, The Meers fault--Its nomenclature and structural setting [abs.]: Earthquake Notes, v. 55, p. 1.

Hays, W. W., and Gori, P. L., eds., 1983, Proceedings of Conference XX, A workshop on The 1886 Charleston, South Carolina, earthquake and its implications for today, May 23-26, 1983, Charleston, South Carolina: U.S. Geological Survey Open-File Report 83-843, 502 p.

King, Geoffrey, and Nabelek, John, 1985, Role of fault bends in the initiation and termination of earthquake rupture: Science, v. 228, p. 984-987.

King, Geoffrey, and Yielding, Graham, 1984, The evolution of a thrust fault system--Processes of rupture initiation, propagation and termination in the 1980 El Asnam (Algeria) earthquake: Geophysical Journal of the Royal Astronomical Society, v. 77, p. 915-933.

Lindh, A. G., 1983, Preliminary assessment of long-term probabilities for large earthqukes along selected fault segments of the San Andreas fault system in California: U.S. Geological Survey Open-File Report 83-63, 15 p.

Rothe, G. H., Lui, C. Y., and Steeples, D. W., 1981, Recent seismicity on the Chadron-Cambridge arch, south-central Nebraska [abs.]: Earthquake Notes, v. 52, p. 61..

Russ, D. P., 1979, Late Holocene faulting and earthquake recurrence in the Reelfoot Lake area, northwestern Tennessee: Geological Society of America Bulletin, v. 90, pt. 1, p. 1013-1018.

______1982, Style and significance of surface deformation in the vicinity of New Madrid, Missouri, in McKeown, F. A., and Pakiser, L. C., eds., Investigations of the New Madrid, Missouri, earthquake region: U.S. Geological Survey Professional Paper 1236, p. 95-114.

Schwartz, D. P., and Coppersmith, K. J., 1984, Fault behavior and characteristic earthquakes--Examples from the Wasatch and San Andreas fault zones: Journal of Geophysical Research, v. 89, p. 5681-5698.

Scott, W. E., Pierce, K. L., and Hait, M. H., Jr., 1985, Quaternary tectonic setting of the 1983 Borah Peak earthquake, central Idaho: Seismological Society of America Bulletin, v. 75, p. 1053-1066.

Steeples, D. W., 1978, Microearthquake network in Kansas, in Andreas Vogel, ed., Terrestrial and space techniques in earthquake prediction: Braunschweig/Weisbaden Publishers, p. 115-130.

557 40008

Steeples, d. W., DuBois, S. M., and Wilson, F. W., 1979, Seismicity faulting, and geophysical anomalies in Nemaha County, Kansas--Relationship to regional structures: Geology, v. 7, p. 134-138.

Sykes, L. R., and Nishenko, S. P., 1984, Probabilities of occurrence of large plate rupturing earthquakes for the San Andreas, San Jacinto, and Imperial faults, California: Journal of Geophysical Research, v. 89, p. 5905-5927.

Thenhaus, P. C., ed., 1983, Summary of workshops concerning regional seismic source zones of parts of the conterminous United States, convened by the U.S. Geological Survey 1979-1980, Golden, Colorado: U.S. Geological Survey Circular 898, 36 p.

Van Wormer, J. D., and Ryall, A. S., 1980, Sierra Nevada-Great Basin Boundary Zone--Earthquake hazard related to structure, active tectonic processes, and anomalous patterns of earthquake occurrence: Seismological Society of America Bulletin, v. 70, p. 1557-1572.

Wallace, R. E., 1984, Patterns and timing of Late Quaternary faulting in the Great Basin province and relation to some regional tectonic features: Journal of Geophysical Research, v. 89, p. 5763-5769.

Zoback, M. L., 1983, Structure and Cenozoic tectonism along the Wasatch Fault zone, Utah, in Tectonic and stratigraphic studies in the eastern Great Basin: Geological Society of America Memoir 157, p. 3-27.

THE INTEGRATION OF GEOLOGICAL AND HISTORICAL DATA IN THE PROBABILISTIC ESTIMATION OF EXTREME EARTHQUAKE OCCURRENCES

by

Kenneth W. Campbell
U.S. Geological Survey
Golden, Colorado

ABSTRACT

Bayesian probability theory in conjunction with the model of extremes is used to develop a Bayesian distribution of extreme earthquake occurrences by assuming that earthquakes represent a Poisson process with exponential distribution of magnitudes. The Bayesian distribution represents the probability that M_{max}, the largest earthquake expected to occur within a period of t years, will exceed some specified magnitude m, and may be computed from the relationship,

$$\tilde{P}(M_{max} > m \mid t) = 1 - \left(\frac{t''}{t'' + t[1 - \tilde{F}(m)]}\right)^{n''}$$

where n" and t" represent updated (posterior) Bayesian estimates of the number of earthquakes and the time period of observation, respectively, and $\tilde{F}(m)$ is the Bayesian distribution of magnitudes, each updated from prior estimates of seismicity using historical observations of earthquake occurrences. The Bayesian extreme-value distribution of earthquake occurrences is tested and applied to the estimation of seismic hazards for the San Jacinto fault zone of southern California. Prior estimates of seismicity are developed from seismotectonic data based on standard seismological relationships among seismic moment, slip rate, earthquake recurrence rate, and magnitude. These estimates are then updated using Bayes' theorem and historical estimates of seismicity associated with the San Jacinto fault zone.

This paper is an integration of two papers previously published in the BULLETIN OF THE SEISMOLOGICAL SOCIETY OF AMERICA, vol. 72, p. 1689-1705 and vol. 73, p. 1099-1115.

INTRODUCTION

Bayesian probability theory contains two features that make it extremely valuable in the estimation of seismic hazards. The first feature provides a rigorous means of combining prior information on seismicity, whether it be judgmental, geological, or statistical, with historical observations of earthquake occurrences. Such prior information may be used to supplement seismicity data when they are incomplete, inaccurate, or cover too short a period of time. In this way, emphasis may be placed on the seismic potential of faults and small seismogenic zones that contribute most significantly to the seismic hazards at a site, rather than on the relatively large source regions that are required to obtain statistically significant samples of earthquakes. This feature also imparts an important dynamic quality to the analysis of seismic hazards by allowing one to incorporate new information on earthquake occurrences directly as they become available.

The second feature provides a means of incorporating the statistical uncertainty associated with the estimation of the parameters used to quantify seismicity in addition to the probabilistic uncertainty associated with the inherent randomness of earthquake occurrences. Both features provide for more reliable estimates of seismic hazard than do conventional methods, when applied by experienced professionals.

The basis of the first feature is Bayes' theorem. As applied to the estimation of seismicity, this theorem states that the posterior probability, $f''(\theta|z)$ or simply $f''(\theta)$, that a specified value θ of a seismicity parameter is the true value, given an observed set of earthquake occurrences, Z, is proportional to the prior probability, $f'(\theta)$, that θ is the true value, times the probability or likelihood of observing Z given θ, $f(z|\theta)$. Mathematically, this theorem is represented by the relationship,

$$f''(\theta) = \frac{f(z|\theta)f'(\theta)}{\int f(z|\theta)f'(\theta)\, d\theta}.$$

The integral function in the denominator is required to normalize $f''(\theta)$ to represent a proper probability density function and arises from application of the total probability theorem. Having applied Bayes' theorem, we may say that f"(θ) has been "updated" from the prior distribution f'(θ) using observational data on earthquake occurrences.

The second feature of Bayesian probability theory has as its basis a compound distribution, which arises when a parameter of a distribution of a random variable is itself treated as a random variable. In applying this to the estimation of seismicity parameters, let the random variable X represent either the number of earthquake occurrences or earthquake magnitude, and let the random variable Θ again represent the parameter of either distribution, such that the probability that X is equal to some specified value x is given by the density function $f(x|\theta)$. Then a new compound distribution, $\tilde{f}(x)$ may be defined that includes both inherent (model) uncertainty and statistical (parameter) uncertainty through application of the theorem of total probabilities, such that

$$\tilde{f}(x) = \int f(x \mid \theta) f''(\theta)\, d\theta.$$

The distribution $\tilde{f}(x)$ is referred to as a "Bayesian distribution," prompted by the treatment of unknown parameters as random variables, a central concept in Bayesian statistics. More specifically, $\tilde{f}(x)$ is a posterior Bayesian distribution, since the posterior distribution of Θ, f"(θ), was used to determine $\tilde{f}(x)$ in order to incorporate the updating feature of Bayes' theorem. The new distribution $\tilde{f}(x)$ can be interpreted as a weighted average of all possible density functions $f(x/\theta)$ which are associated with different values of Θ .

The first application of Bayesian probability theory to earthquake engineering was made by Benjamin (1968). He used a Poisson distribution to develop a Bayesian distribution of earthquake occurrences that accounted for the uncertainty in the estimated mean rate of occurrence. A similar application based on other distributions was later presented by Chou et al. (1971). Esteva (1969) applied Bayes' theorem to the estimation of seismicity within limited geographic regions in Mexico. He based his prior estimate of seismicity on the large statistical sample of earthquakes occurring within the Circumpacific Belt. Lomnitz (1969) suggested a similar approach in Chile, where he used the long historical record of large earthquakes as a prior estimate of the mean rate of occurrence. Other investigators (Esteva and Villaverde, 1973; Cornell and Merz, 1974; McGuire, 1977; Nair and Cluff, 1977) have used discrete Bayesian procedures to include subjective information and uncertainty on maximum magnitude, surface faulting, strong-motion attenuation, and the geometry of source regions in the analysis of seismic hazards. The direct application of Bayesian probability theory to the probabilistic assessment of seismic hazards was first proposed by Cornell (1972), and has only recently found limited application in earthquake engineering (Esteva and Bazan, 1978; Campbell, 1979, 1977; Eguchi and Hasselman, 1979; Mortgat and Shah, 1979).

The methodology presented in this paper was developed from an earlier procedure proposed by the author (Campbell, 1977) for estimating the seismic hazard potential of a fault. Extreme-value theory is used to develop a Bayesian distribution of earthquake hazard from conventional models of earthquake occurrences. For this purpose, earthquake hazard is defined as the probability that the largest earthquake expected to occur within a given period of time will exceed a specified magnitude. This definition is consistent with that currently used in engineering practice for quantifying the seismic hazards associated with strong-motion parameters and earthquake magnitude and is well-established in the literature (e.g., Cornell, 1968; Lomnitz, 1974; Algermissen and Perkins, 1976).

The Bayesian extreme-value distribution of earthquake occurrences developed in this study is applied to the San Jacinto fault zone of southern California. This fault was selected because there is sufficient information available in the literature with which to establish prior seismotectonic estimates of seismicity and sufficient historical activity with which to demonstrate the updating features of the model.

BAYESIAN POISSON-GAMMA DISTRIBUTION OF EARTHQUAKE OCCURRENCES

The temporal occurrence of earthquakes may, for all practical purposes, be represented by a Poisson process if we can assume that earthquakes are independent random events and that no two events can occur at the same instant in time. Such an assumption, although commonly accepted, is inconsistent with periodic strain release mechanisms or earthquake clustering. To be consistent with the development of the conventional extreme-value distribution, we will accept common practice and use the Poisson model of occurrence for this study. The distribution is given by the expression,

$$P(N = n \mid \nu, t) = \frac{(\nu t)^n e^{-\nu t}}{n!} \quad (1)$$

where $P(N = n|\nu, t)$ is the probability that the number of earthquakes occurring within a specified period of time t will be equal to n, given that the mean rate of earthquake occurrences is ν.

To account for the statistical uncertainty in the estimation of ν, equation (1) is more accurately represented by a Bayesian (compound) distribution. Following the discussion in the previous section and applications by Benjamin and Cornell (1970) and Benjamin (1968), a Bayesian distribution representing equation (1) may be obtained by evaluating the integral equation,

$$\tilde{P}(N = n \mid t) = \int_0^\infty P(N = n \mid \nu, t) f''(\nu)\, d\nu \quad (2)$$

in which $f''(\nu)$ represents the posterior probability density function of ν, updated from the prior distribution of ν by incorporating, through Bayes' theorem, observations on the occurrence of earthquakes (i.e., the number of occurrences within a specified period of time).

By assuming that earthquake occurrences are a Poisson process and that the uncertainty in ν may be represented by a gamma distribution, Cornell (1972), Campbell (1977), and Mortgat and Shah (1979) have shown by application of Bayes' theorem that $f''(\nu)$ may be represented by another gamma distribution*,

$$f''(\nu) = K_1 \nu^{n''-1} e^{-\nu t''} \quad (3)$$

where the normalizing constant $K_1 = t''^{n''}/\Gamma(n'')$, and where $\Gamma(n'')$ represents the gamma function with parameter n". If n" is an integer, then the gamma function reduces to the factorial (n" - 1)!. The parameters n" and t" represent updated values of the number of earthquake occurrences and the time period of observation, respectively, and may be computed from the relationships,

$$n'' = n_0 + \left(\frac{\bar{\nu}'}{\sigma_\nu'}\right)^2 \tag{4a}$$

$$t'' = t_0 + \frac{\bar{\nu}'}{(\sigma_\nu')^2} \tag{4b}$$

where n_o is the number of earthquakes observed within a time period of t_o years, and $\bar{\nu}'$ and σ_ν' represent the prior "best estimates" of the mean and standard deviation of the mean rate of occurrence parameter ν.

Equation (2) may now be evaluated by substituting for $P(N = n|\nu, t)$ and $f''(\nu)$ their equivalent expressions given in equations (1) and (3), respectively, and integrating to obtain,

*This results from the choice of a gamma distribution to represent the prior distribution of ν. Being a "conjugate" of the Poisson distribution used to represent the likelihood of observing the number of historical earthquakes which were known to occur, one obtains through the application of Bayes' theorem the mathematically convenient result that the posterior distribution of ν is the same type as its prior and that the parameters of the posterior distribution are simply related to the parameters of the prior distribution and to simple statistics of the sample. The use of a gamma distribution poses no limitation to the specification of uncertainty, since its two parameters allow one to independently specify both a mean and variance for ν. Since the gamma distribution is also a conjugate of the exponential distribution used to represent the likelihood of observing the earthquake magnitudes which were known to occur historically, similar logic was used to select a gamma distribution for the prior distribution of β, leading to the development of equation (9).

$$\tilde{P}(N = n \mid n'', t'', t) = \frac{\Gamma(n + n'')}{n!\Gamma(n'')} \left(\frac{t''}{t + t''}\right)^{n''} \left(\frac{t}{t + t''}\right)^{n}. \tag{5}$$

Because the derivation of equation (5) was based on a Poisson distribution of earthquake occurrences and a gamma distribution for ν, $\tilde{P}(N = n/n'',t'',t)$ is referred to as a Bayesian Poisson-gamma distribution.

BAYESIAN EXPONENTIAL-GAMMA DISTRIBUTION OF EARTHQUAKE MAGNITUDES

The distribution of earthquakes with respect to their size, usually represented by their magnitude, has been found empirically to obey the Gutenberg-Richter relationship (Richter, 1958) given by,

$$\log_{10} N = a - b(m - m_l) \tag{6}$$

where N is the number of earthquakes of $m \geqq m_1$ occurring within a specified period of time, m is earthquake magnitude, and a and b are empirical constants.

Epstein and Lomnitz (1966) found that equation (6) was consistent with a singly truncated exponential distribution of earthquake magnitudes of the form,

$$\begin{aligned} F(m \mid \beta, m_l) &= P(M \leqq m \mid \beta, m_l) \\ &= 1 - \exp[-\beta(m - m_l)] \end{aligned} \tag{7}$$

where $P(M \leqq m|\beta, m_1)$ is the probability that an earthquake has a magnitude less than or equal to m, given a specified value of the frequency parameter β and a threshold magnitude m_1 below which earthquakes may be neglected. The magnitude frequency parameter is related to b in equation (6) through the relationship $\beta = b \ln 10$.

654 40008

Consistent with the treatment of ν, we may account for the uncertainty in β through the evaluation of the Bayesian distribution,

$$\tilde{F}(m \mid m_l) = \int_0^\infty F(m \mid \beta, m_l) f''(\beta)\, d\beta \qquad (8)$$

where f"(β) represents the posterior probability density function of β, updated from its prior distribution by incorporating observations on the number and magnitude of earthquakes through Bayes' theorem. By assuming that earthquakes are independent, exponentially distributed events and that the variation in β may be represented by a gamma distribution, Campbell (1977) and Cornell (1972) have shown that f"(β) may be represented by another gamma distribution,

$$f''(\beta) = K_2 \beta^{\eta''-1} e^{-\beta m''} \qquad (9)$$

where the normalizing constant $K_2 = m''^{\eta''}/\Gamma(\eta'')$.

The parameters η" and m" represent updated Bayesian estimates of the number of earthquake occurrences† greater than the minimum value m_1 and the sum of the differences between their magnitudes and m_1, respectively, and are given by the expressions,

$$\eta'' = n_0 + \left(\frac{\bar{\beta}'}{\sigma_\beta'}\right)^2 \qquad (10a)$$

$$m'' = n_0(\bar{m} - m_l) + \frac{\bar{\beta}'}{(\sigma_\beta')^2} \qquad (10b)$$

†This estimate of the updated number of earthquake occurrences η" is based on $\bar{\beta}'$, σ_β' and n_0, and is independent of the updated number of occurrences n" based on $\bar{\nu}'$, σ_ν' and n_0. If the coefficients of variation of the prior estimates of β and ν are identical, then η" = n" . The assumptions of independence of the two seismicity parameters will be discussed in a later section.

where $\bar{m}$ is the mean magnitude of the historically observed earthquakes. $\bar{\beta}'$ and σ_β' represent the prior "best estimates" of the mean and standard deviation of the frequency parameter β.

Equation (8) may now be evaluated by substituting for $F(m|\beta, m_1)$ and $f''(\beta)$ their equivalent expressions given by equations (7) and (9), respectively, and integrating to obtain,

$$\tilde{F}(m \mid m_l) = \begin{cases} 1 - \left(\dfrac{m''}{m'' + m - m_l}\right)^{\eta''} & m_l \leqq m \leqq \infty \\ 0 & m < m_l. \end{cases} \qquad (11)$$

Because the derivation of the above expression was based on an exponential distribution of earthquake magnitudes and a gamma distribution for β, $\tilde{F}(m|m_1)$ is referred to as a Bayesian exponential-gamma distribution.

When dealing with small probabilities of occurrence, a physical upper limit to earthquake magnitude is required to realistically characterize earthquake occurrences. To account for this finite limit, equation (11) must be normalized such that $\tilde{F}(m|m_1)$ is equal to unity at the specified upper limit m_u rather than at infinity. The normalizing constant K required to do this may be computed from the equality,

$$K[\tilde{F}(m_u \mid m_l) - \tilde{F}(m_l \mid m_l)] = 1.$$

Recognizing that $F(m_1|m_1) = 0$ and substituting for $F(m_u|m_1)$ its expression given by equation (11), we obtain,

$$K = \left[1 - \left(\frac{m''}{m'' + m_u - m_l}\right)^{\eta''}\right]^{-1}. \qquad (12)$$

6B4 40008

Thus, the doubly truncated Bayesian exponential-gamma distribution of earthquake magnitude becomes,

$$\tilde{F}(m \mid m_l, m_u) = \begin{cases} 1 & m_u < m \\ K\left[1 - \left(\dfrac{m''}{m'' + m - m_l}\right)^{\eta''}\right] & m_l \leqq m \leqq m_u \\ 0 & m < m_l. \end{cases} \tag{13}$$

MODEL OF EXTREMES

In many earthquake engineering applications, the largest load to which a structure will be subjected is cause for concern. The ability of many structures and systems to function under the maximum demand, not simply expected values, will, in many situations, determine their success or failure. This is why earthquake engineers traditionally have been interested in knowing the probability that the largest earthquake expected to occur within some specified period of time, usually the economic lifetime of the structure, will be exceeded.

The widespread acceptance of the extreme-value approach to estimating earthquake occurrence probabilities can also be attributed to several advantages associated with an extreme-value distribution (Lomnitz, 1974). These advantages are: (1) a detailed knowledge of the probability distribution of the process is not required, only the behavior of the distribution in its upper tail; (2) the extreme values of a statistical variable are better known, more homogeneous, and more accurately determined than the mean event in a time sequence of data; and (3) it is simple to use and understand, involving few assumptions.

To develop this concept mathematically, let us assume that the random variable M_{max} is the largest magnitude earthquake in a sequence of n earthquakes specified by magnitudes (random variables) M_1, M_2,..., M_n. Then the probability that M_{max} will be less than some specified magnitude m within a period of t years may be represented by the expression,

$$F_{max}(m \mid t) = P(M_{max} \leqq m \mid t)$$
$$= P(\text{no } M_i > m \mid t).$$

To develop this further, we may invoke a simple characteristic of a Poisson process relating to random selection. Quoting Benjamin and Cornell (1970), "...if a random variable Z is Poisson-distributed, then so too is the random variable X, which is derived by (independently) selecting only with probability p each of the incidents counted by Z..." Therefore, using the notation of Benjamin and Cornell, if Z is Poisson-distributed with mean rate of occurrence ν, then a randomly selected subset X is also Poisson-distributed with mean rate of occurrence $p\nu$.

In terms of the earthquake occurrence models used in this paper, the above characteristic implies that earthquakes of magnitudes greater than m (where $m \geqq m_1$) may be represented by a Poisson process with mean rate of occurrence $\nu_m = \nu[1 - F(m)]$, where F(m) is used generically to refer to either the singly or doubly truncated exponential distribution of magnitudes. Thus, the probability $P(\text{no } M_i > m|t)$ appearing in the above equation may be replaced by a Poisson distribution with mean rate of occurrence ν_m, resulting in the expression,

$$F_{max}(m \mid t) = P(N = O \mid \nu_m, t). \qquad (14a)$$

An alternate distribution for $F_{max}(m|t)$ was derived by Campbell (1977) and Algermissen and Perkins (1976)†† based on the well-known result that, for independent and identically distributed random variables, the maximum of a sequence of fixed size n has a distribution equal to the distribution of the variable raised to the power n, or in terms of our hypothetical earthquake sequence, $P(M_{max} \leqq m|t) = F(m)^n$. Since the number of earthquakes in time t is a Poisson-distributed random variable, then by the total probability theorem,

††The distribution of Algermissen and Perkins (1976) was derived for peak acceleration, but the logic is identical to that described here for magnitude.

$$F_{max}(m \mid t) = P(N = O \mid \nu, t) + \sum_{n=1}^{\infty} P(N = n \mid \nu, t) F(m)^n. \quad (14b)$$

By substituting equations (1) and (7) into either equation (14a), consistent with the approach taken by Cornell (1968), or equation (14b), consistent with the approach taken by Algermissen and Perkins (1976), the conventional extreme-value distribution is obtained,

$$F_{max}(m \mid t) = \exp\{-\nu t \exp[-\beta(m - m_l)]\}. \quad (15)$$

The double exponential form of the above distribution is seen to be very similar to the type 1 asymptotic extreme-value distribution of largest values first proposed by Gumbel (1945) as a plotting equation and later developed as a distribution (Gumbel, 1958). This explains the early success of Nordquist (1945), Dick (1965), and Milne and Davenport (1965) in applying the type 1 extreme-value distribution to earthquake data. The common acceptance of equation (15) in earthquake engineering practice began when Epstein and Lomnitz (1966) found that the type 1 extreme-value distribution was consistent with the common assumption of the Poisson occurrence of earthquakes and the empirically established exponential distribution of magnitudes. Since then, independent derivations, such as those resulting from equations (14a) and (14b), have firmly established equation (15) and its varietal forms as the convention in earthquake engineering practice in the United States (e.g., Cornell, 1971; Donovan, 1973; Lomnitz, 1974; Algermissen and Perkins, 1976; Whitman, et al., 1977) as well as in other countries throughout the world (e.g., IASPEI, 1981).

Since the greatest interest lies in the probability that the largest magnitude will exceed some specified value m, probabilities are usually specified in terms of an exceedance probability $P(M_{max} > m|t)$ given by,

$$P(M_{max} > m \mid t) = 1 - F_{max}(m \mid t). \quad (16)$$

In the following section, we will use the model of extremes, given by equation (14a), to derive a Bayesian extreme-value distribution of earthquakes consistent with the conventional distribution of equation (15). An alternate and more complex derivation based on equation (14b) is given in the Appendix.

BAYESIAN EXTREME-VALUE DISTRIBUTION

The model of extremes represented by equation (14a) may be used to derive a Bayesian extreme-value distribution of earthquakes by substituting for $P(N = n|\nu_m, t)$ and ν_m Bayesian distributions of earthquake occurrences and magnitudes. In order to proceed, we must first derive a Bayesian representation of ν_m in terms of updated Bayesian parameters n" and t" to use in conjunction with equation (5).

From the known properties of a gamma distribution, the updated mean rate of occurrence ν'' may be represented in terms of n" and t" by the relationship,

$$\bar{\nu}'' = \frac{n''}{t''}. \tag{17}$$

Recognizing as before that earthquakes of M > m form a Poisson process with mean rate of occurrence $\nu_m = \nu[1 - F(m)]$, then the "updated" mean rate of occurrence of such events becomes,

$$\bar{\nu}_m'' = \bar{\nu}''[1 - \tilde{F}(m)] \tag{18}$$

where $\tilde{F}(m)$ is a generic representation of both the singly and doubly truncated Bayesian distributions of magnitudes. Substituting this expression into equation (17) results in an expression for $\bar{\nu}_m''$ in terms of n" and t",

$$\bar{\nu}_m'' = \frac{n''}{t''}[1 - \tilde{F}(m)]. \tag{19}$$

If we characterize the above expression as representing equation (17) with parameters,

$$n_m'' = n'' \tag{20a}$$

$$t_m'' = \frac{t''}{1 - \tilde{F}(m)} \tag{20b}$$

then the Bayesian representation of equation (14a) becomes,

$$\tilde{F}_{\max}(m \mid t) = \tilde{P}(N = O \mid n_m'', t_m'', t). \tag{21}$$

Upon substituting equations (5), (20a), and (20b) in the above expression, we may then obtain the following expression for $\tilde{F}_{max}(m|t)$ in terms of physically meaningful and easily computed parameters,

$$\tilde{F}_{\max}(m \mid t) = \left(\frac{t_m''}{t + t_m''}\right)^{n_m''}$$

$$= \left(\frac{t''}{t'' + t[1 - \tilde{F}(m)]}\right)^{n''}. \tag{22}$$

This expression for the Bayesian extreme-value distribution of earthquakes may be evaluated by replacing $\tilde{F}(m)$ by either Bayesian exponential-gamma distributions given by equations (11) or (13), recognizing that equation (22) is only valid over the limits of magnitude appropriate for the distribution used.

Note that the distribution $\tilde{F}_{max}(m|t)$ has a finite probability at the lower magnitude threshold m_1 given by the expression,

$$\tilde{F}_{max}(m_l \mid t) = \left(\frac{t''}{t'' + t}\right)^{n''}. \quad (23)$$

At first glance, this property seems inconsistent with our assumptions regarding earthquake occurrences (we might expect it to be equal to unity). However, if we recognize that the extreme-value distribution represents the probability that no earthquake above a specified magnitude occurs, then the probability that no earthquake above the lower limit occurs is simply the Poisson probability $\tilde{P}(N = 0 \mid n'', t'', t)$, exactly equal to the value of $\tilde{F}_{max}(m \mid t)$ at $m = m_1$.

Two quantities of interest to earthquake engineers are the annual probability of exceedance of the largest-magnitude earthquake and the return period, or mean time between earthquakes of magnitudes exceeding m. From equation (22), the annual probability of exceedance is simply,

$$\tilde{P}(M_{max} > m \mid t = 1) = 1 - \tilde{F}_{max}(m \mid t = 1)$$

$$= 1 - \left(\frac{t''}{t'' + 1 - \tilde{F}(m)}\right)^{n''}. \quad (24)$$

Since the return period (T_m), as defined in the preceding paragraph, is the reciprocal of the mean rate of occurrence of earthquakes having M > m, then from equation (19),

$$T_m = \frac{1}{\bar{\nu}_m''}$$

$$= \frac{t''}{n''[1 - \tilde{F}(m)]}. \quad (25)$$

By comparing equations (24) and (25), we find that the return period is not equal to the reciprocal of the annual probability of exceedance as is commonly assumed. However, by expanding equation (24) as a binomial series, we find

GS7 40008

that equation (25) does represent the reciprocal of the first element of this series. Therefore, we may consider return period to adequately represent the reciprocal of the annual probability of exceedance for large values of t", representing probabilities less than approximately 0.05.

A limited number of Bayesian extreme-value distributions have been presented in the literature. Besides the model developed in this study, equations (22) and (24), only three other Bayesian extreme-value distributions have been proposed (Cornell, 1972; Campbell, 1977). The more recent Bayesian hazard models proposed by Esteva and Bazan (1978), Mortgat and Shah (1979), and Eguchi and Hasselman (1979) represent a probabilistic estimate of ground motion without incorporating the model of extremes, and as such represent a marked departure from the extreme-value models being considered here.

A comparison of equation (22) with the three Bayesian extreme-value distributions proposed in the literature reveals some interesting results. The model proposed by Cornell (1972) was found to give exceedance probabilities substantially larger than the model presented in this paper for the entire range of magnitudes considered. This results from simplifications used in the development of his model. Of the two models proposed by Campbell (1977), his distribution based on a numerical integration of the conventional extreme-value distribution of earthquake occurrences was found to give results very similar to equation (22).

A second model proposed by Campbell (1977) was derived by substituting updated Bayesian estimates of the seismicity parameters into the conventional extreme-value distribution of earthquake occurrences (e.g., Epstein and Lomnitz, 1966; Lomnitz, 1974). This model is appealing because of the extensive use of the conventional extreme-value distribution in seismic hazard analyses. It differs from the Bayesian distribution only in the use of conventional rather than Bayesian Poisson and exponential distributions to model the occurrence of earthquakes. Therefore, it lacks the first feature inherent in the currently proposed Bayesian distribution--the incorporation of statistical uncertainty in the seismicity parameters. Because of the similarity between the two models, the distribution proposed by Campbell (1977) is presented here and compared to the currently proposed distribution in the application presented later in this paper. It is represented by the expression,

$$\tilde{P}(M_{\max} > m \mid t) = 1 - \exp\{-\bar{\nu}'' t[1 - F(m)]\} \qquad (26)$$

where the magnitude distribution F(m) is represented by the doubly truncated exponential distribution (Cornell and Vanmarcke, 1969) with parameter $\bar{\beta}''$, where for $m_l \leqq m \leqq m_u$,

$$F(m) = K\{1 - \exp[-\bar{\beta}''(m - m_l)]\} \qquad (27a)$$

$$K = \{1 - \exp[-\bar{\beta}''(m_u - m_l)]\}^{-1}. \qquad (27b)$$

EFFECT OF HISTORICAL EARTHQUAKE DATA

One of the significant features of the Bayesian analysis described in this paper is the ability to combine prior information with historical observations of earthquake occurrences to form a "posterior" or "updated" estimate of the seismicity parameters ν and β. Since the posterior distribution for both seismicity parameters was found to be of the gamma type, it can easily be shown that the updated estimates of the mean and coefficient of variation (i.e., the standard deviation divided by the mean) of ν and β are given by the relationships,

$$\bar{\nu}'' = \frac{n''}{t''}; \qquad V_\nu'' = \frac{1}{\sqrt{n''}} \qquad (28a)$$

$$\bar{\beta}'' = \frac{\eta''}{m''}; \qquad V_\beta'' = \frac{1}{\sqrt{\eta''}} \qquad (28b)$$

where ν'' and β'' represent the means, and V_ν'' and V_β'' the coefficients of variation of the mean rate of occurrence and magnitude frequency parameter,

6B7 40008

respectively, and n", t", η", and m" are parameters of the posterior gamma distributions of ν and β.

By replacing the updated gamma parameters in equation (28a) with their equivalent expressions given in equations (4a) and (4b), we may then obtain expressions for ν" and V_ν" in terms of their prior estimates and information on the historical occurrence of earthquakes, thus

$$\bar{\nu}'' = \frac{(\bar{\nu}')^2 + n_0(\sigma_\nu')^2}{\bar{\nu}' + t_0(\sigma_\nu')^2} \qquad (29a)$$

$$V_\nu'' = \frac{V_\nu'}{\sqrt{1 + n_0(V_\nu')^2}}. \qquad (29b)$$

The above relationships clearly show the effect of historical earthquake occurrences on the updating process. For instance, as the number and observation period of earthquakes become relatively large, or as the uncertainty in the prior estimate becomes relatively large, the updated estimates are essentially controlled by the historical data, such that $\bar{\nu}'' \to n_0/t_0$ and $V_\nu'' \to \sqrt{1/n_0}$. In contrast, as the number and observation period of earthquakes become relatively small, or as the uncertainty in the prior estimate becomes relatively small, the updated estimates are essentially controlled by the prior estimates $\bar{\nu}'$ and V_ν' . These effects are graphically demonstrated in Figure 1, where the effect of the number of historical earthquakes is shown for a fixed observation period of 44 yr and prior values of the coefficient of variation ranging from 0.25 to 2.0, and in Figure 2, where the effect of the time period of observation is shown for a fixed prior coefficient of variation of 1.0 and historically observed mean rate of occurrences ranging from 0 to 0.4 events/yr.

Substituting for the updated gamma parameters in equation (28b) their equivalent expressions in equations (10a) and (10b), relationships between updated estimates for the mean and coefficient of variation of β as a function of the prior estimates and historical data are obtained, thus

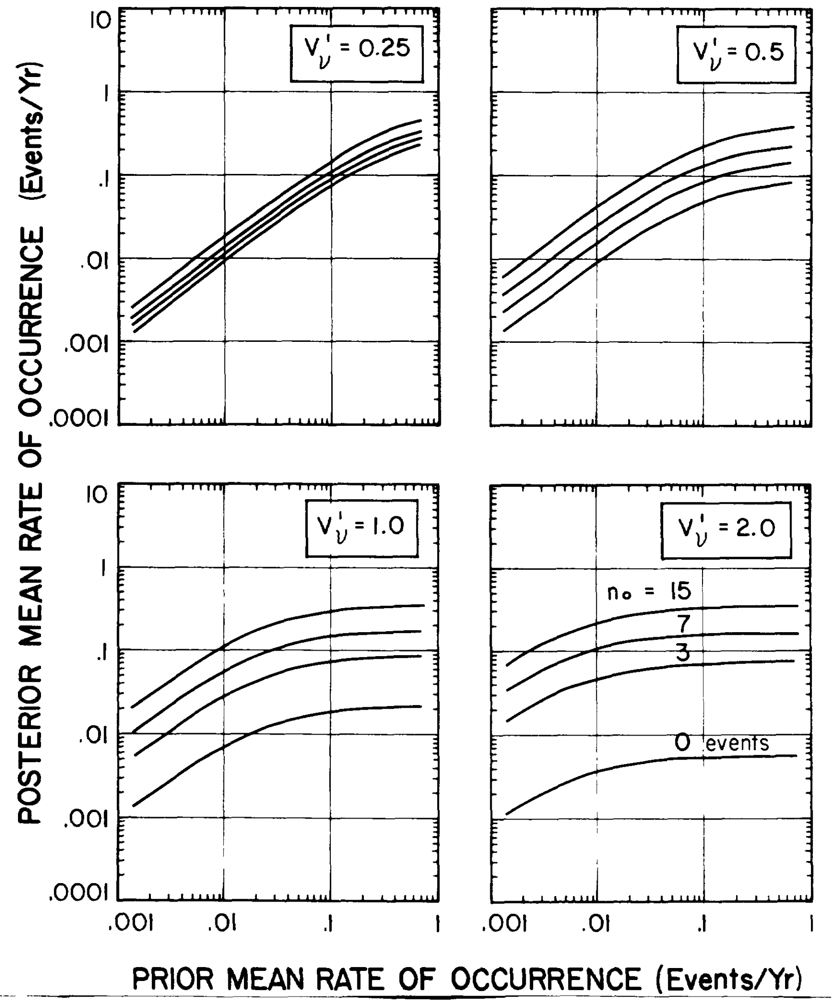


FIG. 1. The effect of the observed number of earthquake occurrences (n_0) on updating the prior value of the mean rate of occurrence ($\bar{\nu}'$). Prior values of the coefficient of variation of $\nu(V_\nu')$ range from 0.25 to 2.0, and the time period of observation is 44 yr.

652 40008

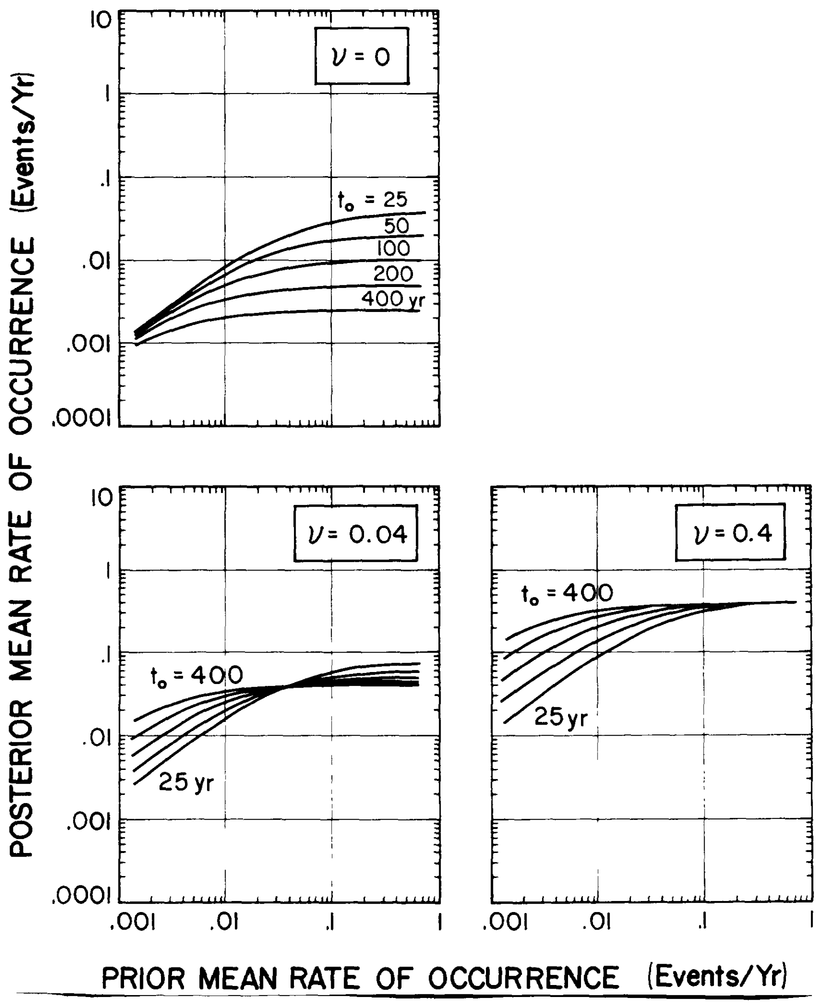


FIG. 2. The effect of the time period of observation (t_0) on updating the prior value of the mean rate of occurrence ($\tilde{\nu}'$). Observed rates of earthquake occurrence (ν) range from 0 to 0.4 events/yr, and the prior value of the coefficient of variation (V_ν') is equal to 1.0.

751 40008

$$\bar{\beta}'' = \frac{(\bar{\beta}')^2 + n_0(\sigma_\beta')^2}{\bar{\beta}' + n_0(\bar{m} - m_l)(\sigma_\beta')^2} \quad (30a)$$

$$V''_\beta = \frac{V'_\beta}{\sqrt{1 + n_0(V'_\beta)^2}}. \quad (30b)$$

From the above expressions we find that, as the number of historical earthquakes becomes relatively large or as the uncertainty in the prior estimate of β becomes relatively large, the updated estimates are controlled by the historical data, such that $\bar{\beta}'' \rightarrow 1/(\bar{m} - m_1)$ and $V_\beta'' \rightarrow \sqrt{1/n_o}$. When the converse is true, then the updated esimates approach $\bar{\beta}'$ and V_β, the prior estimates of the parameters. These effects are graphically demonstrated in Figure 3, where the effects of the number of historical earthquakes are shown for fixed prior estimates of 2.0 and 0.25 for the mean and coefficient of variation of β, respectively, and for values of $m - m_1$ ranging from 0 to 3.0. The prior estimates were fixed at values suggested by Campbell (1977) as being representative of seismically active regions.

From the above analyses, we find that the effect of the historical record on updating the prior estimates of the seismicity parameters is a complex interaction between their prior estimates and associated uncertainties, and the number and observation period of the historical events. In general, the more complete the historical record and the smaller the prior estimates, or the greater the uncertainty in the prior estimates, the greater is the modification of the prior estimates during updating. Although the updating feature allows for great flexibility in the use of prior information, it should be noted that results can be easily biased by poor judgment or inaccurate data. Therefore, we strongly recommend that if such analyses are to be carried out that both the prior estimates of the parameters and the historical data be carefully evaluated by experienced professionals.

SEISMOTECTONIC ESTIMATE OF SEISMICITY

An important feature of the Bayesian distribution proposed in this study is its ability to combine prior estimates of seismicity with historical

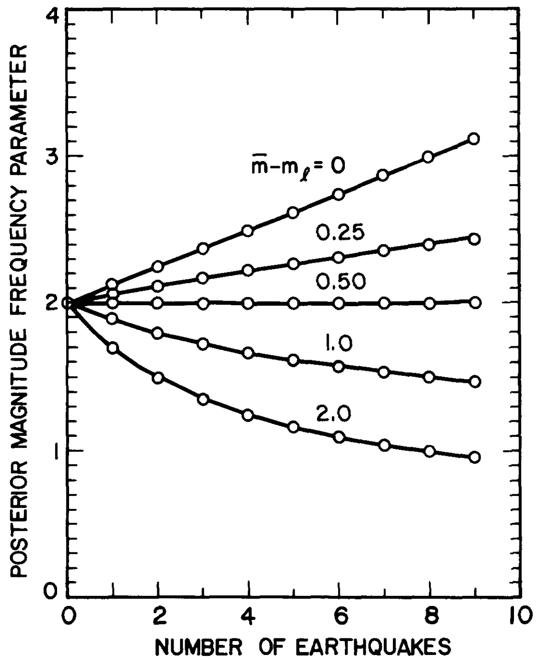


FIG. 3. The effect of the observed number of earthquake occurrences (n_0) on updating the prior value of the magnitude frequency parameter ($\bar{\beta}'$). Observed values of $\bar{m} - m_l$ range from 0 to 2.0, and prior values of the magnitude frequency parameter and its coefficient of variation (V_{β}') are fixed at 2.0 and 0.25, respectively.

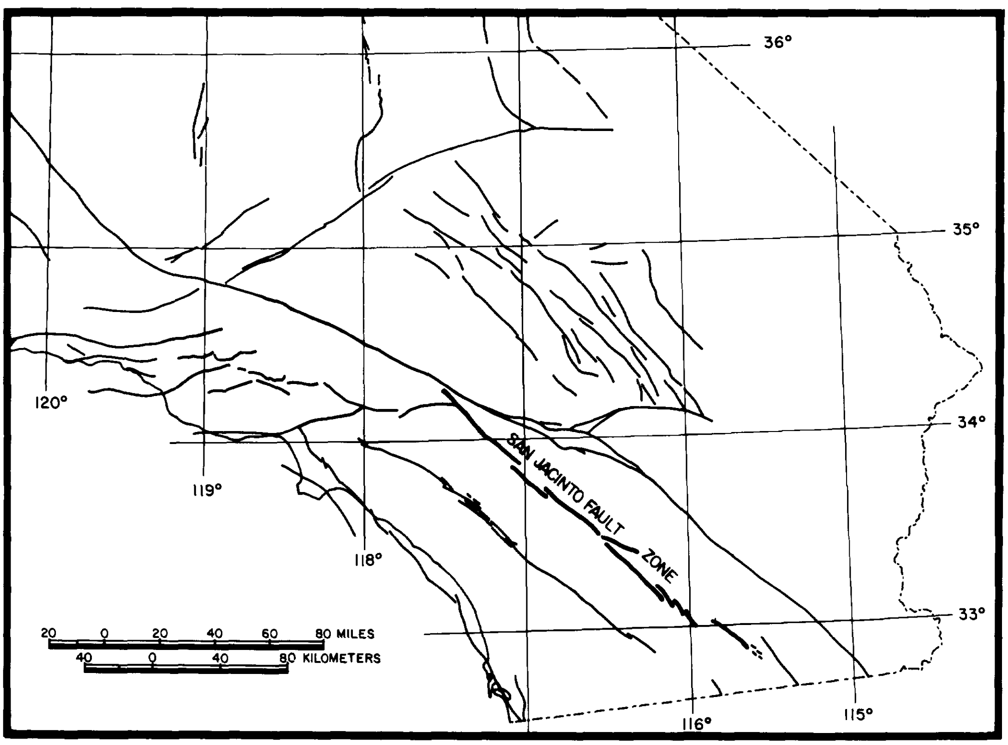


FIG. 4. Fault map of southern California showing the location of the San Jacinto fault zone.

788 40008

earthquake occurrences in computing seismic hazards. Statistical estimates of seismicity become unreliable when there are very few historical data on which to base them. Therefore, it is important that a consistent and physical basis for estimating these parameters be used if one is to reliably estimate the hazard associated with individual faults or small seismogenic zones.

Esteva (1966, 1976) has suggested that statistical estimates of seismicity from large, geotectonically similar regions be used as prior estimates of ν and β for the region of interest. For instance, for the Pacific Coast he has recommended that the seismicity of the Circumpacific Belt, normalized to the area of interest, be used to establish $\bar{\nu}'$ and $\bar{\beta}'$ and their uncertainties. Lomnitz (1969) used the long historical record of great earthquakes in Chile to establish a prior estimate for the mean rate of occurrence for these rare events.

Where there exists a clear relationship between geotectonics and earthquake occurrences, the inferred history of deformation would seem to be the best means of establishing the potential for future activity (Allen, 1975; Esteva, 1976). For faults, this deformation is usually given in terms of slip rate, defined as mean relative displacement per year. Regional deformation is usually given in terms of strain rate. These rates of deformation may be established from geologic, geodetic, seismicity, and plate tectonic information (Brune, 1968; Wallace, 1970; Davies and Brune, 1971; Matsuda, 1975; Savage and Prescott, 1976; Smith, 1976; Campbell, 1977; Anderson, 1979; Molnar, 1979; Shedlock et al., 1980). This information is collectively referred to as seismotectonic data in the discussion that follows.

Prior estimate of ν. The first attempt at estimating the mean rate of occurrence of earthquakes from slip rate was made by Wallace (1970). His approach assumed that all of the displacement on a fault was relieved by the occurrence of earthquakes of similar magnitudes. However, if earthquakes of various sizes are expected to occur, then Wallace's estimates are only valid for magnitudes very near the upper bound for the fault (Campbell, 1978). This limitation may be overcome if all events are included in the estimation of ν.

Using standard seismological relationships among seismic moment, slip rate, earthquake occurrence rate, and magnitude, Smith (1976), Campbell (1977, 1978), Anderson (1979), Molnar (1979), Eguchi and Hasselman (1979), and Shedlock et al. (1980) have all developed expressions relating seismicity to seismotectonic data. Each has assumed earthquake magnitudes to be exponentially distributed, thereby generalizing the procedure proposed by Wallace (1970).

The relationship developed by Shedlock et al. (1980) is used in this study to estimate a prior value for the mean rate of occurrence of earthquakes because their use of a doubly truncated exponential distribution to model the relative frequency of earthquake magnitudes is consistent with that used in the development of the Bayesian extreme-value distribution. In terms of the notation used in this paper, their expression for the mean rate of occurrence of earthquakes having magnitudes greater than m_o is given by the relationship,

$$\bar{\nu}_0' = \frac{\mu A_0 S(c_2 - \bar{b}')}{K_0 \bar{b}'} [M_0(m_u) 10^{-\bar{b}'(m_u - m_0)} - M_0(m_0)]^{-1} \tag{31}$$

where μ is shear modulus, A_o is the total area of the fault plane, S is slip rate (tectonic rate minus creep rate), $M_o(m_u)$ is the seismic moment of the upper bound magnitude, and $M_o(m_o)$ is the seismic moment of m_o. The parameter $\bar{b}'$ represents the prior estimate of b and is related to the prior estimate of β by the expression $\bar{b}' = \bar{\beta}' \log_{10}e$. The truncation factor K_o is given by equation (27b), substituting $\bar{\beta}'$ in place of $\bar{\beta}''$ and m_o in place of m_1. The magnitude m_o represents a physical lower limit below which earthquakes either are not expected to occur or do not contribute to the observed slip on the fault, and is not to be confused with m_1, the arbitrary lower threshold of magnitude used to quantify seismicity (note that $m_o \leqq m_1$). Seismic moment and the parameter c_2 are defined from the expression,

$$\log_{10} M_0(m) = c_1 + c_2 m. \tag{32}$$

758 4008

In order to estimate $\bar{\nu}'$, the mean rate of occurrence of earthquakes greater than m_l, we must evaluate the expression,

$$\bar{\nu}' = \bar{\nu}_0'[1 - F(m)] \tag{33}$$

which, based on the doubly truncated distribution of magnitudes used to establish $\bar{\nu}_0'$, becomes,

$$\bar{\nu}' = \bar{\nu}_0'(1 - K_0\{1 - \exp[-\bar{\beta}'(m_l - m_0)]\}). \tag{34}$$

If there exists no physical basis to support a lower bound for magnitude on a fault (i.e., $m_o \ll o$) or if $m_u \gg m_o$, then equation (34) can be simplified considerably, resulting in the relationship,

$$\bar{\nu}' = \frac{\mu A_0 S}{M_0(m_u)} \frac{c_2 - \bar{b}'}{\bar{b}'} 10^{\bar{b}'(m_u - m_l)}. \tag{35}$$

This expression for $\bar{\nu}'$ is equivalent to that originally developed by the author (Campbell, 1977, 1978) and consistent with the expressions for seismic moment release rate developed by Anderson (1979) and Molnar (1979).

Prior estimate of β. The magnitude distribution parameter β (or b) is proportional to the inverse of the mean magnitude of a sequence of events. This makes its estimation from geological data alone very difficult. Microfracturing studies of rock in the laboratory and observations of earthquake sequences have led several investigators to suggest a possible relationship between b and seismotectonic data (e.g., Scholz, 1968; Wyss, 1973).

Until a more thorough understanding of the relationship between β and seismotectonic data becomes available, it will be necessary to estimate this parameter from actual earthquake sequences (Esteva, 1969, 1976; Newmark and Rosenblueth, 1971). Microearthquakes, mainshocks, aftershocks, and earthquake swarms occurring within the region of interest or within geotectonically similar regions may give statistically significant estimates of $\bar{\beta}'$.

APPLICATION TO THE SAN JACINTO FAULT ZONE

The Bayesian procedures presented earlier in this paper have been used to estimate the seismic hazard associated with the San Jacinto fault zone of southern California (Figure 4). Equations (24) and (26) were used to compute the seismic hazard, where for purposes of comparison, the hazard is given in terms of the annual probability of exceedance (t = 1 yr). Before proceeding, it is appropriate to have a general discussion on the development of the seismotectonic data required to estimate $\bar{\nu}'$ and $\bar{\beta}'$ for faults in southern California.

Estimation of $\bar{\nu}'$ and $\bar{\beta}'$. Equations (31) and (34) indicate that the prior estimate of the mean rate of occurrence requires estimates of several seismotectonic parameters: slip rate S, total area of the fault A_o, shear modulus μ, seismic moments $M_o(m_u)$ and $M_o(m_o)$, $\bar{b}'$, and the magnitude limits m_o, m_1, and m_u. The following discussion indicates how these parameters may be established in general for southern California faults.

The average shear modulus for southern California basement rock is commonly reported to be equal to 3×10^{11} dyne/cm^2. However, if direct measurements are available for a specific fault or region of interest, they should be used in lieu of this value. The total fault area is simply taken as the length times the width of the fault plane, where the length is estimated from published fault maps of the region (e.g., Jennings, 1975) or inferred from seismicity trends. The width, measured along the dip of the fault plane, may be computed from the relationship,

$$W_0 = \frac{H}{\sin \alpha} \tag{36}$$

7B1 40008

where α is the average dip of the fault plane as measured from the horizontal plane and H is the thickness of the seismogenic zone. For southern California, H is estimated to be about 15 km (Hileman et al., 1973; Eguchi et al., 1979).

The lower bound magnitude m_l is normally selected to include only those earthquakes of specific interest. If only those earthquakes that pose a serious threat to structures are of interest, this magnitude is approximately 5.0 M_L (Richter, 1958). This limit may have to be adjusted to accommodate the available record of earthquake occurrences for the fault; but in no case may it be less than m_o. The upper bound magnitudes for faults of the region may be taken from Greensfelder (1974), from Anderson (1979), or from available fault length-magnitude relationships (e.g., Bonilla and Buchanan, 1970; Slemmons, 1977, 1982). The physical lower bound magnitude m_o must be evaluated on an individual basis for each fault by evaluating its seismogenic characteristics. Faults that tend to lock will release most of their accumulated slip as a series of large earthquakes, requiring a relatively large value for m_o. Fault zones that release their slip as it accumulates in a continuous series of events will require relatively small values of m_o (which for most cases may be neglected).

Slip-rate data are generally available for major faults in the region (Anderson, 1979). These rates are primarily established from offsets of geological formations or geomorphic features along the faults. However, other information such as scarp erosional characteristics, tree ring analysis, dated organic material, and fossils can be used to establish these rates (Wallace, 1977). Slip rates may also be determined from reliable geodetic measurements when available. Such measurements are valuable since they may indicate that the strain accumulation rate on the fault is substantially different at present than indicated by long-term geologic offsets. For minor faults of the region, it is currently necessary to infer a slip rate from faults having similar tectonic and deformational characteristics, keeping in mind that the total slip rate along a section of crust perpendicular to the plate boundary is limited to about 5.5 cm/yr from plate tectonic data (Minster et al., 1974).

A relationship between seismic moment and magnitude is required to establish $M_o(m_u)$, $M_o(m_o)$ and the coefficient c_2 of equation (32). Although it has been suggested that such a relationship for southern California may be regionally dependent, a single relationship for all of southern California is considered appropriate for the purposes of this study. Data taken from Wyss and Brune (1968), Hanks and Wyss (1972), Thatcher (1972), Wyss and Hanks (1972), Thatcher and Hanks (1973), and Hanks et al. (1975) were used to establish this relationship for earthquakes of M_L = 2.0 to 6.8 and M_s = 6.7 to 8.3. A least-squares fit to the 176 data points resulted in the expression,

$$\log_{10} M_0(m) = 16.2 + 1.43m \qquad (37)$$

where $M_o(m)$ is seismic moment in dyne-cm and m represents either M_L (for $m \lesssim 6.8$) or M_s (for $m \gtrsim 6.8$). The standard error of estimate for $\log_{10} M_o$ was found to be 0.41. From this analysis, the coefficient c_2 of equation (32) was found to be 1.43 with a standard error of estimate of 0.03. The use of M_s to characterize earthquakes of magnitudes greater than 6.8 avoids the problems associated with the saturation of the M_L scale at $M_L \approx 7$, increasing the validity of equation (37) to include the largest magnitudes expected to occur in southern California.

Esteva (1969, 1976) has suggested that earthquakes occurring within the Circumpacific Belt may be used to establish prior estimates of β and its uncertainty for regions such as southern California. His suggested values for $\bar{\beta}'$ and V_β' are 2.16 and 0.32, respectively. Campbell (1977) suggested values of $\bar{\beta}' = 2.05$, $V_\beta' = 0.12$ for subregions within southern California and $\bar{\beta}' = 1.90$, $V_\beta' = 0.21$ for regions within the Western United States, based on published recurrence relationships. Ideally in the future, it may become feasible to estimate $\bar{\beta}'$ for individual faults based on geologic information.

A prior seismotectonic estimate of the mean rate of earthquake occurrence ($\bar{\nu}'$) for faults in southern California may be taken from equations (34) or (35) of this study or from similar expressions offered by Anderson (1979), Molnar (1979), Eguchi and Hasselman (1979), and Shedlock et al. (1980). The coefficient of variation V_ν' should be chosen with great care to avoid arbitrarily biasing the results in favor of either the prior estimates or the historical data. In general, this uncertainty can be quite large. For example, Campbell (1977) conservatively estimated V_ν' to be 2.7 for faults in southern California, where V_ν' was estimated from an expression similar to equation (35). Esteva (1969, 1976) found similarly large values from a statistical analysis of earthquakes within the Circumpacific Belt. For the purposes of this paper, several values of V_ν' were used in the analysis of seismic hazard. In this way, some insight into the sensitivity of the analysis to this parameter may be gained.

San Jacinto Fault Zone. The San Jacinto fault zone represents a system of faults extending some 300 km from Cajon Pass to El Centro, California (Figure 4). The various branches included in this zone are the Casa Loma, Hot Springs, Buck Ridge, Clark, Coyote Creek, and San Jacinto faults. The Imperial fault, source of the 1940 El Centro and 1979 Imperial Valley earthquakes, lies just to the south of this zone and is not included as part of the system for the purposes of this study. The San Jacinto fault zone has been one of the most active fault systems in southern California. The record of main shocks from 1932 through 1971 listed in Table 1 reveals 13 events of magnitudes 5.0 M_L or greater that can be directly attributed to the zone (Lamar et al., 1973; Hileman et al., 1973). Aftershocks are excluded from this list, since their relatively large numbers would give substantial weight to the historical data during updating, when such a sequence actually reflects only a single event in time and contributes little to the total amount of slip on the fault.

The seismotectonic data used to compute a prior estimate of the mean rate of occurrence for this zone were established in accordance with the procedures outlined in the previous section and appear in Table 2. Several entries of a fault-specific nature require further explanation. The slip rate of 2.0 cm/yr was taken from Anderson (1979) and is consistent with both recent geodetic measurements (Savage and Prescott, 1976) and the upper limit established for slip rate from the offset of Quaternary deposits (Sharp, 1967). The value of 0.3 cm/yr determined by Lamar et al. (1973) for this fault zone was based on

TABLE 1

HISTORICAL SEISMICITY ASSOCIATED WITH THE SAN JACINTO FAULT ZONE, CAJON PASS TO EL CENTRO, CALIFORNIA (MAIN SHOCKS OF $M_L \geq 5.0$ FROM 1932 TO 1972)

Event	Date (UTC)	Latitude (North)	Longitude (West)	Magnitude
1	25 Mar. 1937	33° 24.5′	116° 15.7′	6.0
2	4 June 1940	33° 0.0′	116° 26.0′	5.1
3	21 Oct. 1942	32° 58.0′	116° 0.0′	6.5
4	15 Aug. 1945	33° 13.0′	116° 8.0′	5.7
5	8 Jan. 1946	33° 0.0′	115° 50.0′	5.4
6	24 Jan. 1951	32° 59.0′	115° 44.0′	5.6
7	14 June 1953	32° 57.0′	115° 43.0′	5.5
8	19 Mar. 1954	33° 17.0′	116° 11.0′	6.2
9	26 May 1957	33° 13.9′	116° 0.3′	5.0
10	23 Sept. 1963	33° 42.6′	116° 55.5′	5.0
11	2 Sept. 1968	33° 11.4′	116° 7.7′	6.4
12	28 Apr. 1969	33° 20.6′	116° 20.8′	5.8
13	12 Sept. 1970	34° 16.2′	117° 32.4′	5.4

TABLE 2

SEISMOTECTONIC AND HISTORICAL DATA USED IN THE BAYESIAN ESTIMATE OF SEISMICITY FOR THE SAN JACINTO FAULT ZONE

Parameter Description	Symbol	Units	Value
Slip rate	S	cm/yr	2.0
Shear modulus	μ	dyne/cm^2	3×10^{11}
Fault dimensions	L_0	km	300
	W_0	km	15
Magnitude limits	m_0	M_L	$\ll 0$
	m_l	M_L	5.0
	m_u	M_S	7.5
Historical seismicity	n_0	—	13
	t_0	yr	40
	m_l	M_L	5.0
	$\bar{m}$	M_L	5.66

TABLE 3

BAYESIAN ESTIMATES OF SEISMICITY FOR THE SAN JACINTO FAULT ZONE

Prior (seismotectonic) Estimates			Historical Estimates		Posterior (updated) Estimates		
$\bar{\nu}'$	$\bar{\beta}'$	V'_ν, V'_β	$\bar{\nu}$	$\bar{\beta}$	$\bar{\nu}''$	$\bar{\beta}''$	V''_ν, V''_β
3.087	2.00	0	0.325	1.51	3.087	2.00	0
3.087	2.00	0.10	0.325	1.51	1.561	1.93	0.094
3.087	2.00	0.25	0.325	1.51	0.642	1.75	0.186
3.087	2.00	1.00	0.325	1.51	0.347	1.54	0.267

the maximum estimate of the age associated with the Quaternary deposits described by Sharp (1967) and probably represents a lower limit for slip rate.

The upper bound magnitude of 7.5 M_S for the fault zone was taken from Greensfelder (1974) and Anderson (1979). Consideration of discussions by Esteva (1969, 1976) and Campbell (1977) led us to select a value of 2.0 for $\bar{\beta}'$, while values of V_β' were selected to be equal to those chosen for the mean rate of occurrence. The selected value for $\bar{\beta}'$ is similar to the value of 1.8 determined empirically by Eguchi and Campbell (1977) for earthquakes of $M_L \geqq 4.0$ which occurred from 1932 through 1975 within a narrow zone centered on the fault.

The results of the updating process on the estimation of the seismicity parameters are summarized in Table 3. The prior estimate of ν was computed using equation (35). The historical estimates were computed from events in Table 1 using the maximum likelihood expressions,

$$\bar{\nu} = \frac{n_0}{t_0} \tag{38a}$$

$$\bar{\beta} = \frac{1}{\bar{m} - m_l}. \tag{38b}$$

Posterior (updated) estimates of the seismicity parameters were established from equations (29a) and (30a) for prior coefficients of variation of 0, 0.1, 0.25, and 1.0. Posterior estimates of the coefficient of variation were computed from equations (29b) and (30b).

Table 3 indicates that the prior estimates of seismicity have been modified substantially during updating, reflecting the relatively large number of historical earthquakes. In fact, for $V_\nu' = V_\beta' = 1.0$, the posterior estimates of the seismicity parameters are very nearly equal to the historical estimates, and their coefficients of variation have been reduced by 73 percent to a value of 0.267.

757 40008

Annual probabilities of exceedance based on the Bayesian distribution, equation (24), are plotted in Figure 5 for values of V_ν' and V_β' equal to 0.1, 0.25, and 1.0. Also plotted as the line labeled $V_\nu' = V_\beta' = 0$ in this figure is the annual probability of exceedance based on equation (26). The two equations were found to give identical values of probability when V_ν' and V_β' were set to equal 0.01 in equations (4a) to (4b) and (10a) to (10b). If smaller coefficients of variation are used, unreliable estimates of probability may be obtained, due to numerical complexities of equation (24). Figure 5 indicates that relatively small amounts of uncertainty can result in relatively large variations in computed probability or magnitude. Differences in computed magnitude greater than 0.1 and differences in computed probability greater than 20 percent are found to result from prior coefficients of variation exceeding values as little as 0.05.

To further explore these variations, recall that uncertainty enters the computation of seismic hazard twice, once as uncertainty in the prior estimates of ν and β during the updating process of seismicity, and once as uncertainty in the posterior (updated) estimates of ν and β in the development of the Bayesian Poisson-gamma and Bayesian exponential-gamma distributions of earthquake occurrences. What we would like to know is how each type of uncertainty contributes to the variation in hazard we see in Figure 5.

To study this, we have computed the hazard separately using the seismotectonic, historical, and updated estimates of seismicity given in Table 3. For this purpose, we found it convenient to represent the hazard in terms of magnitude rather than probability. Differences in magnitude are more easily interpreted in terms of their impact on engineering design criteria. The expression used to estimate magnitude is given by

$$\tilde{m}_p = m_l - m''\left[1 - \left(1 - \frac{1}{K''}\{1 - t''[(1-p)^{-(1/n'')} - 1]\}\right)^{-(1/\eta'')}\right] \qquad (39)$$

where $\tilde{m}_p$ is the magnitude associated with p, a specified value of the annual probability of exceedance $\tilde{P}(M_{max} > m|t = 1)$. The expression is derived directly from equations (4), (10), (12), (13), (24), and (28).

753 40008

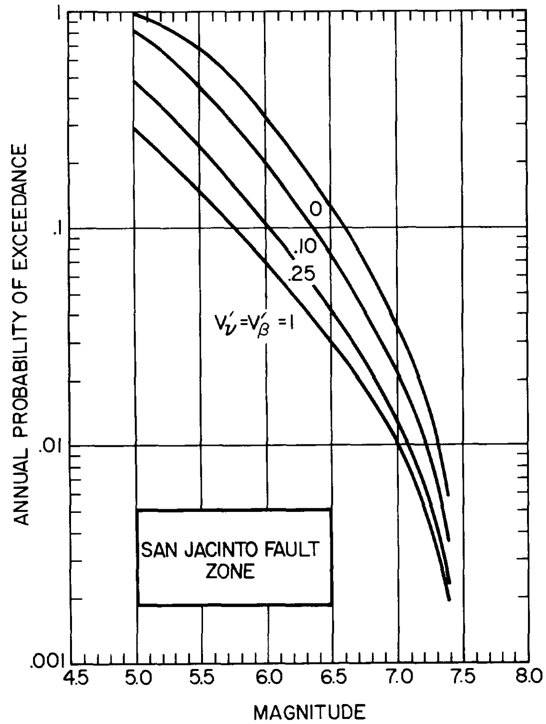


FIG. 5. Bayesian estimates for the annual probability of exceedance of magnitude for the San Jacinto fault zone. V'_ν and V'_β represent prior seismotectonic estimates of the coefficients of variation of ν, the mean rate of occurrence of earthquakes, and β, the magnitude frequency parameter.

TABLE 4
MAGNITUDE ESTIMATES FOR SELECTED RETURN-PERIOD EARTHQUAKES ASSOCIATED WITH THE SAN JACINTO FAULT ZONE

Return Period (yr)	Annual Probability of Exceedance	V'_ν, V'_β	Seismicity Estimate		
			Seismotectonic	Historical	Updated
10	0.095	0	6.62	5.76	6.62
		0.1	6.64	5.76	6.38
		0.25	6.71	5.76	6.04
		1.0	7.06	5.76	5.79
50	0.020	0	7.16	6.68	7.16
		0.1	7.18	6.68	7.02
		0.25	7.24	6.68	6.83
		1.0	7.40	6.68	6.71
100	0.010	0	7.30	6.98	7.31
		0.1	7.31	6.98	7.21
		0.25	7.35	6.98	7:08
		1.0	7.45	6.98	7.01
500	0.002	0	7.45	7.36	7.45
		0.1	7.46	7.36	7.42
		0.25	7.47	7.36	7.38
		1.0	7.49	7.36	7.37
1000	0.001	0	7.48	7.43	7.48
		0.1	7.48	7.43	7.46
		0.25	7.48	7.43	7.44
		1.0	7.50	7.43	7.43

787 40008

It is also convenient to relate the annual probability of exceedance to return period, the average time between exceedances. To do this, we combine equation (24) with equation (25) to derive the following expression between return period, T_m in years, and probability of exceedance, p

$$T_m = \{n''[(1-p)^{-(1/n'')} - 1]\}^{-1}. \tag{40}$$

For small probabilities, where $p \leq 0.05$, return period is closely approximated by the reciprocal of p, consistent with previous observations based on conventional distributions of earthquake occurrences.

The results of the study appear in Table 4, where magnitudes associated with return periods ranging from 10 to 1000 yr (annual exceedance probabilities ranging from 0.001 to 0.095) are given for seismotectonic, historical, and updated estimates of seismicity. Magnitudes computed from the seismotectonic estimate of seismicity assumed $n_o = t_o = 0$ in equations (4a) to (4b) and (10a) to (10b), so that n" and t" would reflect only the prior estimates of the seismicity parameters and their uncertainties. On the other hand, those computed from the historical estimate of seismicity used $\bar{\nu}' = \bar{\beta}' = 0$ in these same equations so that $n'' = n_o$ and $t'' = t_o$. Note that in this latter case, $V_\nu'' = V_\beta'' = 1/\sqrt{n_o}$ making uncertainty dependent only on the number of historical earthquake occurrences. Also note that the estimates of magnitude appearing in Table 4 are given to two decimal places only as a means of comparison and are not intended to imply this level of accuracy in the estimates. They represent M_L for magnitudes less than about 6.8 and M_S for events of larger magnitude.

The seismotectonic estimates of magnitude in Table 4 may be used to isolate the effects of the second type of uncertainty discussed above. Since no updating of the seismicity parameters was allowed in these computations, the observed variation in $\tilde{m}_p$ reflects only that uncertainty inherent in the Bayesian distributions of earthquake occurrences. Further inspection indicates that the variation in $\tilde{m}_p$ is dependent on return period, increasing with decreasing T_m. If we use variations in magnitude greater than about 0.1

to indicate significant differences in $\tilde{m}_p$, then this second type of uncertainty is found to be negligible for coefficients of variation of about 0.25 or less.

The updated estimates of $\tilde{m}_p$ are found to range between their seismotectonic estimates for $V_\nu' = V_\beta' = 0$ and their historical estimates for $V_\nu' = V_\beta' = 1.0$. Although the updated estimates contain both types of uncertainty, the effects of these uncertainties are easily isolated by recognizing from Table 3 that the largest coefficient of variation assocated with the updated estimates of the seismicity parameters is only 0.267. In light of the above discussion, we may then conclude that the variation in the Bayesian estimate of $\tilde{m}_p$ results almost exclusively from the variation in the updated estimate of seismicity observed in Table 3. We may also conclude from this and observations concerning the seismotectonic estimates of $\tilde{m}_p$ that equations (24) and (26) may be used interchangeably provided $V_\nu'' = V_\beta'' \lesssim 0.25$.

Because of the significance of V_ν' and V_β' in the Bayesian estimation of seismic hazard, the selection of these parameters should be done with as much care as the selection of $\bar{\nu}'$ and $\bar{\beta}'$. It is not immediately apparent, although, what considerations should go into such a selection. Although a detailed discussion is beyond the scope of this study, some insight is provided by Algermissen et al. (1982). They found from Monte Carlo simulations of earthquake occurrences that at least 40 earthquakes were required before one obtained statistically stable estimates of seismicity. This would suggest that values of V_ν' and V_β' substantially less than 1.0 (for which the historical earthquake occurrences control the hazard) are appropriate for the San Jacinto fault zone. Campbell (1977) adds further support for values less than unity by observing that variations in earthquake recurrence relations throughout southern California and the Western United States are consistent with coefficients of variation for β ranging from 0.12 to 0.21.

Based on these simple arguments, we suggest that values of $V_\nu' = V_\beta' = 0.1$ to 0.25 be used as a basis for selecting appropriate values of $\tilde{m}_p$ or annual probabilities of exceedance for the San Jacinto fault zone. Values of $\tilde{m}_p$ so selected are found to lie approximately halfway between seismotectonic and historical estimates of $\tilde{m}_p$, and are represented by magnitudes of 6.2, 6.9, 7.1, 7.4, and 7.5 for return periods of 10, 50, 100, 500, and 1000 yr, respectively. By comparison, the seismotectonic and historical estimates of magnitude for the 100-yr earthquake are 7.0 and 7.3, respectively.

The results presented in the preceding paragraph point out the significance of upper bound magnitude on the computation of seismic hazard for active fault zones such as the San Jacinto. If a value of m_u of 7.0 M_s, a value closer to the historical upper bound for the fault zone, was to be used in the analysis, the estimates of $\tilde{m}_p$ for return periods of 50 to 1000 yr would be considerably smaller than those presented in Table 4. This is confirmed in Table 5, where values of $\tilde{m}_p$, based on an upper bound magnitude of 7.0, are compared to similar estimates of $\tilde{m}_p$ based on $m_u = 7.5$.

While no difference in $\tilde{m}_p$ is observed for a return period of 10 yr, the 100-yr magnitude is found to decrease to a value of 6.8. Since the 1000-yr events reflect the upper bound for the fault zone, their differences are directly related to the assumed differences in m_u.

DISCUSSION

Throughout this paper, we have assumed the two seismicity parameters ν and β to be independent of one another for mathematical convenience. However, if we assume that the long-term rate of energy release for a specific fault or region is relatively constant, then these two parameters would be expected to be positively correlated. Small values of β would require relatively small numbers of earthquakes to maintain the same rate of energy release as would large values of β and relatively large numbers of earthquakes. The independent updating of these parameters proposed in this study contradicts this mechanism, which requires these parameters to be jointly distributed in a rigorous mathematical analysis. On the other hand, the historical earthquake data may be indicative of a rate of energy release that is different than that assumed "a priori," not requiring updated values of ν and β to maintain the prior rate of energy release. Considering the possible errors involved in assuming earthquakes to be independent random events, we feel the assumption of independence of the seismicity parameters is not a significant limitation

TABLE 5

THE EFFECT OF UPPER BOUND MAGNITUDE ON ESTIMATES OF SEISMIC HAZARD FOR THE SAN JACINTO FAULT ZONE ($V'_\nu = V'_\beta = 0.1 - 0.25$)

Return Period (Yr)	Updated Estimate of Magnitude	
	$m_u = 7.0$	$m_u = 7.5$
10	6.2	6.2
50	6.7	6.9
100	6.8	7.1
500	7.0	7.4
1,000	7.0	7.5

in the present analysis. The effect of this assumption may be minimized by taking $V_{\nu}' = V_{\beta}'$ in the analysis if possible.

Although consistent with common practice, the assumption that earthquakes may be represented as a Poisson process with an exponential distribution of magnitudes represents a possible weakness in both the conventional and Bayesian extreme-value distributions. These models do not account for earthquake clustering or cyclic strain build-up and release mechanisms that have been observed or hypothesized for specific regions and faults. However, the problem is currently a topic of research, and its incorporation in the assessment of earthquake hazards will represent the next major step in the evolution of the extreme-value distribution of earthquake occurrences.

The Bayesian extreme-value distribution of earthquakes presented in this paper represents a major step toward the ability to use information from geologists, seismologists, and geophysicists, relating to the seismic potential of various faults or regions, together with historical seismicity, in assessing the hazards associated with earthquake occurrences. Based on the familiar Poisson and exponential models of earthquake occurrence and size, the Bayesian extreme-value distribution has the same strengths as the widely accepted conventional extreme-value distribution. However, through the application of Bayesian probability theory, the new distribution gains two features that make it even more powerful than the conventional model.

The first feature is its ability to incorporate uncertainty in the estimate of seismicity in addition to the inherent or physical uncertainty associated with the random nature of earthquake occurrences modeled by the Poisson and exponential distributions. Unlike the probabilistic uncertainty associated with the model, the statistical uncertainty in seismicity estimates becomes smaller as more observations of the process become available to provide more reliable estimates of the parameters. The second feature of the Bayesian model, which is closely related to the first, is its ability to rigorously combine prior information on earthquake occurrence with historical observations. The prior information can be nonstatistical in nature, such as that derived from expert opinion (e.g., Eguchi et al., 1979, TERA Corporation, 1980), from geological investigations (e.g., Wallace 1970; Lamar et al., 1973;

 888 4008

Sieh, 1977), or from seismotectonic data (e.g., Campbell, 1977, 1978; Anderson, 1979; Molnar, 1979), or it can be derived statistically from a similar seismotectonic region or fault (e.g., Newmark and Rosenblueth, 1971; Esteva, 1979) or from an ancient catalog of felt earthquakes (e.g., Lomnitz, 1969). This updating feature makes the model dynamic, since as new observations become available, they may be used to further update the estimates of seismicity, thereby updating the extreme-value distribution.

The analyses presented in this paper demonstrate the importance of uncertainty in the Bayesian estimate of seismicity and subsequently in the estimate of seismic hazard. Modest amounts of uncertainty in the prior estimate of seismicity can produce substantial variations in computed probabilities and magnitudes. For example, the Bayesian analysis of seismic hazard for the San Jacinto fault zone of southern California indicated that differences in computed magnitude exceeding 0.1 and variations in computed probabilities exceeding 20 percent result from coefficients of variation associated with prior estimates of seismicity as small as 0.05.

Although these large differences admittedly result from the rather large differences in seismicity associated with the prior seismotectonic and historical estimates of ν and β, it does point out the need for reliable data on which to base these estimates. In particular, careful selection of the prior value of uncertainty is necessary in order that proper weight may be given to the prior (seismotectonic) estimate of seismicity.

It would appear from the present study that prior estimates of the coefficients of variation should be restricted to relatively small values (say, values less than about 0.25), if one is to control the amount of updating that can result from limited historical data. Actual values should be chosen on a case-by-case basis taking into consideration the reliability of the seismotectonic data used to estimate seismicity and the quality of the historical data.

From sensitivity analyses used to compare the Bayesian distribution developed in this study with others that have been proposed, we have found: (1) the Bayesian distribution proposed by Cornell (1972) tends to overestimate probabilities of exceedance and associated magnitudes; (2) Campbell's (1977) generalization of the distribution proposed by Cornell, which requires numerical integration, gives results consistent with the currently proposed distribution; and (3) the conventional extreme-value distribution proposed by Epstein and Lomnitz (1966) and Cornell (1968), incorporating updated Bayesian estimates of the seismicity parameters, is identical to the currently proposed distribution for updated coefficients of variation of seismicity equal to 0.01, with negligible differences between the two distributions for updated coefficients of variation of 0.25 or less.

ACKNOWLEDGMENTS

This study is based in part on research sponsored by the National Science Foundation at the University of California, Los Angeles. I am greatly indebted to C. Martin Duke for his guidance during the course of the project. I also wish to thank Ron Eguchi, Dave Boore, and an anonymous reviewer for their critical review of the manuscript and their many helpful suggestions.

REFERENCES

Algermissen, S. T., D. M. Perkins, P. C. Thenhaus, S. L. Hanson, and B. L. Bender (1982). Probabilistic estimates of maximum acceleration and velocity in rock in the contiguous United States, *U.S. Geol. Surv. Open-File Rept. 82-1033.*

Allen, C. R. (1975). Geological criteria for evaluating seismicity, *Bull. Geol. Soc. Am.* **86,** 1041–1057.

Anderson, J. G. (1979). Estimating the seismicity from geological structure for seismic risk studies, *Bull. Seism. Soc. Am.* **69,** 135–158.

Benjamin, J. R. (1968). Probabilistic models for seismic force design, *J. Struct. Div., ASCE,* **94(ST5),** 1175–1196.

Benjamin, J. R. and C. A. Cornell (1970). *Probability, Statistics and Decision for Civil Engineers,* McGraw-Hill, New York.

Bonilla, M. G. and J. M. Buchanan (1970). Interim report on worldwide historic surface faulting, *U.S. Geol. Surv. Open-File Rept.*

Brune, J. N. (1968). Seismic moment, seismicity, and rate of slip along major fault zones, *J. Geophys. Res.* **73,** 777–784.

Campbell, K. W. (1977). The use of seismotectonics in the Bayesian estimation of seismic risk, Report UCLA-ENG-7744, School of Engineering and Applied Science, University of California, Los Angeles, California.

Campbell, K. W. (1978). An estimate of earthquake recurrence times from seismotectonic data on a fault, in: *Geologic Guide and Engineering Geology Case Histories, Los Angeles Metropolitan Area,* First Annual California Section Conference, Los Angeles, May 12–14, 1978, Association of Engineering Geologists, Los Angeles, California, 95–101.

Campbell, K. W. (1979). A Bayesian procedure for incorporating seismotectonics in the estimation of seismic risk on a fault, *Proceedings of the Specialty Conference on Probabilistic Mechanics and Structural Reliability,* A. H-S. Ang and M. Shinozuka, Editors, Tucson, Arizona, January 10–12, 1979, American Society of Civil Engineers, New York, 290–294.

Chou, I. H., W. J. Zimmer, and J. T. P. Yao (1971). Likelihood of strong motion earthquakes, Bureau of Engineering Research, University of New Mexico, Technical Report CE 27(71).

Cornell, C. A. (1968). Engineering seismic risk analysis, *Bull. Seism. Soc. Am.* **58,** 1583–1606.

Cornell, C. A. (1971). Probabilistic analysis of damage to structures under seismic loads, in *Dynamic Waves in Civil Engineering,* D. A. Howells *et al.*, Editors, John Wiley and Sons, London, England.

Cornell, C. A. (1972). Bayesian statistical decision theory and reliability-based design, *Proceedings of the International Conference on Structural Safety and Reliability,* A. M. Freudenthal, Editor, April 9–11, 1969, Washington, D.C., Smithsonian Institute, 47–66.

Cornell, C. A. and H. A. Merz (1974). Seismic risk analysis of Boston, ASCE National Structural Engineering Meeting, Cincinnati, Ohio, Preprint 2260.

Cornell, C. A. and E. H. Vanmarcke (1969). The major influences on seismic risk, *Proceedings of the Fourth World Conference on Earthquake Engineering,* Santiago, Chile, vol. 1, 69–83.

Davies, G. F. and J. N. Brune (1971). Regional and global fault slip rates from seismicity, *Nature* **229,** 101–107.

Dick, I. D. (1965). Extreme-value theory and earthquakes, *Proceedings of the Third World Conference on Earthquake Engineering,* Auckland, New Zealand, vol. 1, sect. III, 45–53.

Donovan, N. C. (1973). Earthquake hazards for buildings, in *Building Practices for Disaster Mitigation,* R. Wright *et al.*, Editors, National Bureau of Standards Building Science Series 46, U.S. Government Printing Office, Washington, D.C., 82–111.

Eguchi, R. T. and K. W. Campbell (1977). Seismicity and site effects on earthquake risk, *Proceedings of the Sixth World Conference on Earthquake Engineering,* New Delhi, India, vol. II, 399–404.

Eguchi, R. T. and T. K. Hasselman (1979). A Bayesian seismic risk study of California, *Proceedings of the Second U.S. National Conference on Earthquake Engineering,* Stanford University, August 22–24, 1979, Earthquake Engineering Research Institute, Berkeley, California, 52–61.

Eguchi, R. T., K. W. Campbell, and J. H. Wiggins (1979). A survey of expert opinion on active and potentially active faults in California, Nevada, Arizona and Northern Baja California, J. H. Wiggins Company Technical Report 79-1328-2, Redondo Beach, California.

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

Esteva, L. (1969). Seismicity prediction: a Bayesian approach, *Proceedings of the Fourth World Conference on Earthquake Engineering,* Santiago, Chile, vol. I, 172–184.

Esteva, L. (1976). Seismicity, chapter 6 in *Seismic Risk and Engineering Decisions,* C. Lomnitz and E. Rosenblueth, Editors, Developments in Geotechnical Engineering 15, Elsevier Scientific Publishing Company, New York, 179–224.

Esteva, L. and E. Bazan (1978). Seismicity and seismic risk related to subduction zones, *Proceedings of the Second International Conference on Microzonation for Safer Construction-Research and Application,* San Francisco, California, November 26–December 1, 1978, National Science Foundation, Washington, D.C., vol. II, 657–668.

Esteva, L. and R. Villaverde (1973). Seismic risk, design spectra and structural reliability, *Proceedings of the Fifth World Conference on Earthquake Engineering*, Rome, Italy, vol. II, 2586–2596.

Greensfelder, R. W. (1974). Maximum credible rock acceleration from earthquakes in California, map sheet 23, California Division of Mines and Geology.

Gumbel, E. J. (1945). Simplified plotting of statistical observations, *Trans. Am. Geophys. Union* **26,** 69–82.

Gumbel, E. J. (1958). *Statistics of Extremes,* Columbia University Press, New York.

Hanks, T. C. and M. Wyss (1972). The use of body wave spectra in the determination of seismic source parameters, *Bull. Seism. Soc. Am.* **62,** 561–589.

Hanks, T. C., J. A. Hileman, and W. Thatcher (1975). Seismic moments of the larger earthquakes of the southern California region, *Bull. Geol. Soc. Am.* **86,** 1131–1139.

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

IASPEI (1981). Symposium on earthquake prediction and risk, Twenty-first General Assembly of the International Association of Seismology and Physics of the Earth's Interior, University of Western Ontario, London, Ontario, Canada, Book of Abstracts, A2.1–A.2.47.

Jennings, C. W. (1975). Fault map of California, Geologic Data Map 1, California Division of Mines and Geology.

Lamar, D. L., P. M. Merifield, and R. L. Proctor (1973). Earthquake recurrence intervals on major faults in southern California, in *Geology, Seismicity and Environmental Impact,* D. E. Moran *et al.*, Editors, Association of Engineering Geologists Special Publication, 265–276.

Lomnitz, C. (1969). An earthquake risk map of Chile, *Proceedings of the Fourth World Conference on Earthquake Engineering,* Santiago, Chile, vol. I, 161–171.

Lomnitz, C. (1974). *Global Tectonics and Earthquake Risk,* Developments in Geotectonics 5, Elsevier Scientific Publishing Company, New York.

Matsuda, T. (1975). Magnitude and recurrence interval of earthquakes from a fault, *J. Seism. Soc. Japan* **28,** 269–283.

McGuire, R. K. (1977). Effects of uncertainty in seismicity on estimates of seismic hazard for the east coast of the United States, *Bull. Seism. Soc. Am.* **67,** 827–848.

Milne, W. G. and A. G. Davenport (1965). Statistical parameters applied to seismic regionalization, *Proceedings of the Third World Conference on Earthquake Engineering*, Auckland, New Zealand, vol. 1, Sect. III, 181–193.

Minster, J. B., *et al.* (1974). Numerical modeling of instantaneous plate tectonics, *Geophys. J.* **36,** 541–576.

Molnar, P. (1979). Earthquake recurrence intervals and plate tectonics, *Bull. Seism. Soc. Am.* **69,** 115–133.

Mortgat, C. P. and H. C. Shah (1979). A Bayesian model for seismic hazard mapping, *Bull. Seism. Soc. Am.* **69,** 1237–1251.

Nair, K. and L. S. Cluff (1977). An approach to establishing design surface displacements for active faults, *Proceedings Sixth World Conference on Earthquake Engineering*, New Delhi, India, preprint vol. 2, 2-449–2-454.

Newmark, N. M. and E. Rosenblueth (1971). *Fundamentals of Earthquake Engineering,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Nordquist, J. N. (1945). Theory of largest values applied to earthquake magnitudes, *Trans. Am. Geophys. Union* **26,** 29–31.

Richter, C. F. (1958). *Elementary Seismology,* W. H. Freeman and Company, San Francisco, 768 pp.

Savage, J. C. and W. H. Prescott (1976). Strain accumulation on the San Jacinto fault near Riverside, California, *Bull. Seism. Soc. Am.* **66,** 1749–1754.

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes, *Bull. Seism. Soc. Am.* **58,** 399–415.

Sharp, R. V. (1967). San Jacinto fault zone in the peninsular ranges of southern California, *Bull. Geol. Soc. Am.* **78,** 705–730.

Shedlock, K. M., R. K. McGuire, and D. G. Herd (1980). Earthquake recurrence in the San Francisco Bay region, California, from fault slip and seismic moment, *U.S. Geol. Surv. Open-File Rept. 80-999.*

Sieh, K. E. (1977). A study of Holocene displacement history along the south-central reach of the San Andreas Fault, *Ph.D. Dissertation,* Stanford University, Stanford, California.

Slemmons, D. B. (1977). Faults and earthquake magnitude, in *State-of-the-Art for Assessing Earthquake Hazards in the United States,* Misc. Paper S-31-1, Report 6, U.S. Army Waterways Experiment Station, Vicksburg, Mississippi.

Slemmons, D. B. (1982). Determination of design earthquake magnitudes for microzonation, *Proceedings of the Third International Earthquake Microzonation Conference,* Seattle, Washington, June 28–July 1, 1982, vol. 1, 119–130.

Smith, S. W. (1976). Determination of maximum earthquake magnitude, *Geophys. Res. Letters* **3,** 351–354.

TERA Corporation (1980). Seismic hazard analysis: solicitation of expert opinion, U.S. Nuclear Regulatory Commission, Washington, D.C., NUREG/CR-1582, vol. 3.

Thatcher, W. (1972). Regional variation of seismic source parameters in the northern Baja California area, *J. Geophys. Res.* **77,** 1549–1565.

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

Wallace, R. E. (1970). Earthquake recurrence intervals on the San Andreas fault, *Bull. Geol. Soc. Am.* **81,** 2875–2890.

Wallace, R. E. (1977). Time-history analysis of fault scarps and fault traces—A longer view of seismicity, *Proceedings of the Sixth World Conference on Earthquake Engineering,* New Delhi, India, vol. II, 409–411.

Whitman, R. V., N. C. Donovan, B. Bolt, S. T. Algermissen, and R. Sharpe (1977). Seismic design regionalization maps for the United States, *Proceedings of the Sixth World Conference on Earthquake Engineering,* New Delhi, India, vol. 2, 387–392.

Wyss, M. (1973). Towards a physical understanding of the earthquake frequency distribution, *Geophys. J. R. Astron. Soc.* **31,** 341–359.

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

Wyss, M. and J. N. Brune (1971). Regional variations of source properties in southern California

APPENDIX

An alternate derivation of the Bayesian extreme-value distribution of earthquakes [equation (22)] may be developed based on the model of extremes represented by equation (14b) by substituting for $P(N = n \mid \nu, t)$ and $F(m)$ their Bayesian distribution counterparts giving,

$$\tilde{F}_{\max}(m \mid t) = \tilde{P}(N = O \mid n'', t'', t) + \sum_{n=1}^{\infty} \tilde{P}(N = n \mid n'', t'', t)\tilde{F}(m)^n. \quad \text{(A1)}$$

The above expression may be evaluated by replacing $\tilde{P}(N = O \mid n'', t'', t)$ and $\tilde{P}(N = n \mid n'', t'', t)$ with their respective Bayesian Poisson-gamma relationships given in equation (5) and simplifying, such that

$$\begin{aligned}\tilde{F}_{\max}(m \mid t) &= \left(\frac{t''}{t + t''}\right)^{n''} + \sum_{n=1}^{\infty} \frac{\Gamma(n + n'')}{n!\Gamma(n'')}\left(\frac{t''}{t + t''}\right)^{n''}\left(\frac{t}{t + t''}\right)^{n}\tilde{F}(m)^n \\ &= \left(\frac{t''}{t + t''}\right)^{n''}\left[1 + \frac{1}{\Gamma(n'')}\sum_{n=1}^{\infty}\frac{\Gamma(n + n'')}{n!}\left(\frac{t\tilde{F}(m)}{t + t''}\right)^{n}\right]. \quad \text{(A2)}\end{aligned}$$

Recognizing that a property of the gamma function is that $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, the gamma function appearing within the summation may be expressed by the series,

$$\begin{aligned}\Gamma(n + n'') &= (n'' + n - 1)(n'' + n - 2)\ldots n''\Gamma(n'') \\ &= \Gamma(n'')\prod_{i=0}^{n-1}(n'' + i).\end{aligned}$$

Replacing $\Gamma(n + n'')$ in equation (A2) by the above expression and simplifying,

$$\tilde{F}_{\max}(m \mid t) = \left(\frac{t''}{t + t''}\right)^{n''}\left[1 + \sum_{n=1}^{\infty}\frac{\prod_{i=0}^{n-1}(n'' + i)}{n!}\left(\frac{t\tilde{F}(m)}{t + t''}\right)^{n}\right].$$

Further, recognizing that the term in brackets represents a binomial expansion of the form $(1 - x)^{-n''}$, then

$$\begin{aligned}\tilde{F}_{\max}(m \mid t) &= \left(\frac{t''}{t + t''}\right)^{n''}\left(1 - \frac{t\tilde{F}(m)}{t + t''}\right)^{-n''} \\ &= \left(\frac{t''}{t + t''}\right)^{n''}\left(\frac{t + t''}{t'' + t[1 - \tilde{F}(m)]}\right)^{n''} \\ &= \left(\frac{t''}{t'' + t[1 - \tilde{F}(m)]}\right)^{n''}. \quad \text{(A3)}\end{aligned}$$

Although this derivation is more complicated than that based on equation (14a), we find that the resulting distribution is identical, thereby serving as an independent verification of equation (22).

MODELS OF SEISMICITY AND USE OF HISTORICAL DATA IN

EARTHQUAKE HAZARD ANALYSIS

by

Daniele Veneziano and Jozef Van Dyck
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

INTRODUCTION

A frequent use of historical earthquake data in seismic hazard analysis is to estimate parameters of recurrence models; for example, the parameters a_i and b_i of

$$\Lambda_i(m) = 10^{a_i - b_i m} \quad (1)$$

where $\Lambda_i(m)$ is the rate of events with magnitude larger than m and epicenter in the i^{th} "seismogenic province". The same data, appropriately "attenuated," can be used to estimate the parameters of a recurrence model at a site $\underline{X}$, for example the parameters $c_{\underline{X}}$ and $d_{\underline{X}}$ in

$$\Lambda_{\underline{X}}(y) = 10^{c_{\underline{X}} - d_{\underline{X}} y} \quad (2)$$

where Y is a measure of site intensity, for example Y = ℓn(PGA), and $\Lambda_{\underline{X}}(y)$ is the rate of events at $\underline{X}$ with Y > y.

Representations of regional seismicity exemplified by Eq. 1 are referred to here as regional models, whereas representations of site activity like that of Eq. 2 are called site models.

This paper has a threefold objective: 1. examine the genesis of some popular site models, 2. show that conventional regional and site models are special cases of a broader class of seismicity models, and 3. indicate the role of historical data in parameter estimation and model selection. Topic 1 concerns mainly properties of existing methods, whereas Topics 2 and 3 identify modeling possibilities and data analysis techniques that have just begun to be explored.

After making the distinction between parametric and nonparametric earthquake occurrence models (both types are used in practice), we bring unity to the subject by considering the continuum of seismicity representations generated by a nonparametric recurrence model constrained by prior knowledge. What gives rise to different representations of seismicity in this case is the type and amount of prior knowledge, which comprises information not processed through formal statistical data analysis. We also examine several model-fitting procedures, including among others ordinary maximum likelihood (ML), maximum cross-validated likelihood (MCVL), and maximum penalized likelihood (MPL). In the case of no prior information, ML produces site models otherwise known as nonparametric historic models. At the end of the paper, models and data analysis techniques are exemplified through reference to the EPRI (1985) project and to other recent studies. Because of space limitations, treatment is at a conceptual level, with a minimum of analytical details and numerical examples.

PARAMETRIC AND NONPARAMETRIC MODELS AND STATES OF UNCERTAINTY

Both regional and site models come in a variety of forms. Naturally, the recurrence relationships need not be exponential and probabilistic dependencies of many types may exist among the locations $\underset{\sim}{X}_{o_i}$, times t_i, and sizes m_i of different earthquakes. Another aspect in which models may differ is that they may represent seismicity either parametrically or nonparametrically. In the former case, the seismicity model is specified by parameters that are few compared with the number of historical earthquakes; examples are the recurrence

8B3 40008

relationships in Eqs. 1 and 2. By contrast, nonparametric models require estimation of a large number of variables or are entirely data based. For instance, models with many variables result from using small seismogenic provinces so that the index i in Eq. 1 ranges from 1 to a large number N, whereas an example of data-based estimation is when the parameters a and b are allowed to vary continuously in space.

Except for the case when the parametric form has a theoretical or a very strong empirical justification, nonparametric models are conceptually superior to parametric alternatives and include the latter as special cases. A common misconception is that nonparametric models necessarily require more data to be estimated than parametric models. In reality, estimation accuracy depends on the amount of empirical and prior information, not on the parametric or nonparametric form of the model. Prior information is expressed through constraints and penalties in estimation procedures such as least squares and maximum likelihood, through a prior distribution in Bayesian analysis. The real limitations of nonparametric models are the need to express prior information in high-dimensional spaces and the complexity of fitting models and quantifying estimation uncertainty.

The next two sections consider nonparametric models and indicate statistical estimation procedures that are appropriate under different states of prior information. Some of the models and estimation methods have been long in use, while others have been proposed only recently or have not been tried yet.

SEISMICITY MODELS AND ESTIMATION PROCEDURES: NO PRIOR INFORMATION

A widely used nonparametric estimator of $\Lambda_{\underline{X}}$ in Eq. 2 is the empirical exceedance rate

$$\hat{\Lambda}_{\underline{X}}(y) = \frac{1}{T} \sum_{i} \frac{1}{P_{D_i}} P[Y_i > y] \tag{3}$$

where T is the time period of the catalog, P_{D_i} is the probability of detection for events with the characteristics of the i^{th} earthquake (see Eq. 4 below), and the summation extends over all historical earthquakes. The exceedance probability in the right-hand side of Eq. 3 accounts for uncertainty on ground motion attenuation to the site. Different forms of the detection probability P_{D_i} can be used, producing different estimators. Two possibilities, which will be shown to have a precise statistical meaning, are

$$P_{D_i} = \begin{cases} P_D(m_i, \underline{X}_{o_i}, t_i) & \text{(a)} \\ \overline{P}_D(m_i, \underline{X}_{o_i}) = \frac{1}{T}\int_0^T P_D(m_i, \underline{X}_{o_i}, t)dt & \text{(b)} \end{cases} \tag{4}$$

where $\underline{X}_o$ is epicentral location and $\overline{P}_D$ is the average over time of P_D in Eq. 4a. The actual exceedance rate $\Lambda_{\underline{X}}(y)$ is given by

$$\Lambda_{\underline{X}}(y) = \int_{m,\underline{X}_o} \lambda(m,\underline{X}_o)\; P[Y(m,\underline{X}_o)>y]\; dm\; d\underline{X}_o \tag{5}$$

in which $\lambda(m,\underline{X}_o)$ is the rate density of events with epicentral location $\underline{X}_o$ and magnitude m.

It can be shown that $\hat{\Lambda}_{\underline{X}}(y)$ in Eq. 3 with P_{D_i} in Eq. 4a or 4b is the maximum likelihood estimator of $\Lambda_{\underline{X}}(y)$ under the following conditions:

For P_{D_i} in Eq. 4a,

Earthquakes occur as a Poisson process, with rate density $\lambda(m,\underline{X}_0,t)$. This rate density is an unknown function of magnitude, space, and time; in particular, λ is not necessarily exponential in m, constant in t, or piecewise-constant in $\underline{X}_0$. Eqs. 3 and 4a give the maximum likelihood estimator of the <u>average</u> exceedance rate during the period of the catalog.

For P_{D_i} in Eq. 4b,

Earthquakes occur as a Poisson process, with rate density $\lambda(m,\underline{X}_0,t) \equiv \lambda(m,\underline{X}_0)$; i.e., the dependence of λ on m and $\underline{X}_0$ is unknown, but λ is considered to be constant in time (stationarity assumption).

Again omitting derivations, the estimator in Eq. 3 has the property that, for P_{D_i} in either Eq. 4a or 4b, $E[\hat{\Lambda}_{\underline{X}}(y)|\Lambda_{\underline{X}}(y)] = \Lambda_{\underline{X}}(y)$ for any given $\Lambda_{\underline{X}}(y)$. In the case of Eq. 4a, this unbiasedness condition holds for $\Lambda_{\underline{X}}(y)$ any average exceedance rate during the period of the catalog. Under the assumption of stationary Poisson seismicity, one can further show that the variance of $\hat{\Lambda}_{\underline{X}}(y)$ is given by

$$\mathrm{Var}[\hat{\Lambda}_{\underline{X}}(y)] = \begin{cases} \frac{1}{T} \int_{m,\underline{X}_o} \lambda(m,\underline{X}_o)\, \underset{t}{E}\left[\frac{1}{P_D(m,\underline{X}_o,t)}\right] P^2[Y(m,\underline{X}_o)>y]\,dm\, d\underline{X}_o, & \text{(a) for } P_{D_i} \text{ in Eq. 4a} \\ \frac{1}{T} \int_{m,\underline{X}_o} \frac{\lambda(m,\underline{X}_o)}{\bar{P}_D(m,\underline{X}_o)} P^2[Y(m,\underline{X}_o)>y]\, dm\, d\underline{X}_o, & \text{for } P_{D_i} \text{ in Eq. 4b (b)} \end{cases} \quad (6)$$

In the last equation, $\underset{t}{E}$ denotes time averaging over the period of the catalog, say from 0 to T. Notice that $\underset{t}{E}[1/P_D(m,\underline{X}_o,t)] \geq 1/\bar{P}_D(m,\underline{X}_o)$, meaning that the variance of the first estimator is never smaller than the variance of the second estimator. In the special case of complete reporting, $P_D(m,\underline{X}_o,t) \equiv 1$, both Eqs. 6a and 6b reduce to

8B7 40008

$$\mathrm{Var}[\hat{\Lambda}_{\underline{X}}(y)] = \frac{1}{T}\,\Lambda_{\underline{X}}(y) \tag{7}$$

In fact, in this case, the summation in Eq. 3 has Poisson distribution with mean and variance equal to $\Lambda_{\underline{X}}(y)T$.

The expressions in Eq. 6 are for a catalog that extends in time from 0 to T, but they can be easily generalized to the case when only a portion of the catalog is used, say since $t_o > 0$: all one must do is replace 1/T with $1/(T-t_o)$ and extend the averages E and $\overline{P}_D$ over the interval from t_o to T. Whereas the variance in Eq. 6b is a nonincreasing function of t_o, the variance in Eq. 6a may have a minimum for $t_o > 0$, indicating that earlier data should not be used. This happens if, for small t, $P_D(m,\underline{X}_o,t)$ is close to zero. Another consideration for limiting the range of catalog data used in the estimation of $\Lambda_{\underline{X}}$ is that, for early events, the estimates of m, $\underline{X}_o$ and P_D may be unreliable.

The property of unbiasedness noticed earlier does not by itself qualify the estimator in Eq. 3 as a good estimator: for values of y that are of interest in practice, 1. the variances in Eq. 6 may be unacceptably large and 2. the distribution of $\hat{\Lambda}_{\underline{X}}(y)$ is skewed with a long upper tail, so that the median of $\hat{\Lambda}_{\underline{X}}(y)$ is much lower than the true rate $\Lambda_{\underline{X}}(y)$.

One can of course do better if one has prior information on the earthquake process. Even if prior information is not available, one can improve on the ML estimators of Eqs. 3 and 4, e.g. by using <u>cross validation</u>. To exemplify this technique, consider the simpler problem of estimating the probability density function f_X of a variable X from a random sample $X_1,\ldots,X_n$. No a priori information is supposed available on f_X, so that the unconstrained maximum-likelihood (ML) estimator is

8B6 40008

$$\hat{f}_X(x) = \frac{1}{n} \sum_{i=1}^{n} \delta(x-X_i) \qquad (8)$$

in which δ is the Dirac delta function. Consider now a family of nonparametric estimators $\hat{f}_{X,\Theta}(x)$ that depend on a variable Θ and include $\hat{f}_X$ in Eq. 8 as a special case. For example, the estimator

$$\hat{f}_{X,\Theta}(x) = \frac{1}{n}\frac{1}{\Theta} \sum_{i=1}^{n} \phi\left(\frac{x-X_i}{\Theta}\right), \quad \Theta \geqslant 0 \qquad (9)$$

where ϕ is the standard normal density, approaches $\hat{f}_X$ in Eq. 8 as $\Theta \to 0$. The idea of maximum cross-validated likelihood (MCVL) is to split the data in several ways into an estimation and a prediction sample. For each Θ, the cross-validated likelihood is the average over the sample splits of the likelihood of the prediction sample given the model fitted to the estimation sample using that Θ. For example, one might split the sample $X_1,\ldots,X_n$ in n different ways, each time putting one value X_j in the prediction sample and all the others in the estimation sample. Then one calculates the probability densities

$$\hat{f}^{(j)}_{X,\Theta}(X_j) = \frac{1}{n-1}\frac{1}{\Theta} \sum_{i \neq j} \phi\left(\frac{X_j-X_i}{\Theta}\right), \quad j = 1,\ldots,n \qquad (10)$$

and maximizes with respect to Θ the cross-validated likelihood

$$\ell^*(\Theta) = \prod_{j=1}^{n} \hat{f}^{(j)}_{X,\Theta}(X_j) \qquad (11)$$

The maximum of $\ell^*(\Theta)$ is typically found for $\Theta^* > 0$, meaning that the MCVL estimator $\hat{f}_{X,\Theta^*}$ is smoother and has smaller variance than the ML estimator in Eq. 8, although it is unbiased only asymptotically as $n \to \infty$. Model fitting methods of this type have not been developed for seismic analysis, although they have the potential of substantially improving earthquake hazard estimators for the case of no prior information.

The cross-validation principle is quite general and can be applied to quantities other than the likelihood; for example one might minimize a cross-validated measure of the squared deviation between actual and predicted earthquake counts for a suitable discretization of space, time, and magnitude. Also, Θ may be a vector of parameters not just a scalar.

The main limitation of MCVL is the large amount of computation needed to evaluate and maximize the cross-validated likelihood. A method that produces smooth estimates of the recurrence rate through a much reduced effort is <u>maximum penalized likelihood</u> (MPL). In this case, one estimates the regional rate density $\lambda(m,\underline{X}_0)$ [or the rate density at the site, $\lambda_{\underline{X}}(y)$] so that a penalized version of the likelihood function is maximum. If no prior information is available, the penalty must be viewed as an expedient to counteract the extreme "spikiness" of ordinary ML estimates. On the other hand, if prior information exists and is expressed through a multiplicative penalty, then MPL corresponds to maximum-a-posteriori (MAP) Bayesian estimation. It is important that one distinguishes between these two interpretations of MPL and realizes that MPL is an appropriate estimation method also under noninformative priors. In the case of no prior, a conceptual disadvantage of MPL with respect to MCVL is that the former method imposes smoothness externally, whereas the latter method uses the data to find the optimal degree of smoothing.

9514000 8

USE OF PRIOR INFORMATION

In many cases, prior information exists which reflects physical knowledge of the seismogenic process in the region of study or knowledge of the seismicity of "similar" regions. For example, physical characteristics of a fault and the current stress regime might constrain the size of possible earthquakes, while data from different regions might give some credibility to a parametric form of the magnitude-recurrence relationship.

One may have prior information on the rate density function $\lambda_{\underline{X}}(y)$ at site $\underline{X}$ or, more frequently, on the regional rate density function $\lambda(m,\underline{X}_0)$. Information may further correspond to a nonparametric model (meaning that, from prior considerations alone, it is not possible to constrain the rate functions to a parametric form) or to a parametric model. Although the latter case is rare, it is frequent in practice to use parametric earthquake rate models because they are commonly accepted and easier to analyze. The development of nonparametric alternatives should be a priority of future research.

As prior information increases without restricting the function $\lambda(m,\underline{X}_0)$ to a parametric family (case of informative nonparametric state of prior knowledge), ordinary ML estimation becomes less attractive. One can however use other methods mentioned previously: MPL, in which the penalty reflects now prior knowledge and possibly the desire to reduce the estimator variance of $\lambda(m,\underline{X}_0)$, and MPA, which coincides with MPL if the penalty in the latter method is chosen as a multiplicative prior density. A penalized version of cross-validated likelihood (MPCVL) is also possible which, as the name says, maximizes a penalized form of the cross-validated likelihood. The penalty can be taken to be the same as in MPL or MAP.

A difficult step in informative nonparametric analysis is the formulation of constraints, penalties, or priors on $\lambda(m,\underline{X}_o)$ that reasonably reflect the initial state of knowledge. Penalties that pull the MPL estimate of $\lambda(m,\underline{X}_o)$ towards simple parametric functions have been proposed and implemented by the authors in the EPRI (1985) study and will be mentioned briefly in the next section.

Another limitation of the previous nonparametric techniques (for both noninformative or informative prior states of knowledge) is that ML, MPL, MAP, MCVL, and MPCVL generate just one estimate $\hat{\lambda}(m,\underline{X}_o)$ and therefore do not characterize uncertainty on the function $\lambda(m,\underline{X}_o)$. In practice, using just one estimate is not acceptable if uncertainty on the seismic hazard function $\Lambda_{\underline{X}}(y)$ comes mainly from uncertainty on $\lambda(m,\underline{X}_o)$. For the Bayesian analyst, uncertainty on a nonparametric function such as $\lambda(m,\underline{X}_o)$ is described through a random process. In the related area of probability density estimation, work has been done to characterize such processes but results to date are limited and have little potential for earthquake rate modeling. Two simpler approaches, which are also used in nonparametric probability density estimation, have been implemented in the EPRI study: One is to directly assess the posterior probabilities of a discrete set of estimates $\hat{\lambda}_i(m,\underline{X}_o)$, $i = 1,\ldots,r$. The estimates are obtained through MPL using penalties specified by the seismologist, who also judgementally assigns the posterior probabilities. As an aid to judgement, goodness-of-fit statistics are provided for each estimate $\hat{\lambda}_i(m,\underline{X}_o)$, which describe the degree to which the fitted model explains the historical data.

The other procedure used in the EPRI study to quantify uncertainty on $\lambda(m,\underline{X}_o)$ is bootstrapping. For any given penalty specified in MPL, this technique generates not just one estimate $\hat{\lambda}(m,\underline{X}_o)$, but a number of estimates, which can be viewed as equiprobable rate density functions given the historical

9B8 40008

data and the prior state of knowledge (the latter is expressed through the chosen penalty).

Prior knowledge may be such that parametric constraints can be placed on the function $\lambda(m,\underline{X}_o)$ [or on the function $\lambda_{\underline{X}}(y)$]; for example, one might find it reasonable to assume exponentiality of λ as a function of m, or to assume that λ is piecewise constant on the geographical plane, or that the function $\lambda_{\underline{X}}(y)$ has a certain analytical form.

The transition between nonparametric and parametric representations of seismicity is a gradual one: not only can one envision a model that is partly parametric, partly nonparametric, but one can use penalties on the function $\lambda(m,\underline{X}_o)$ such that certain parametric forms are included as special or limiting cases. An example in which λ depends parametrically on m but depends parametrically on $\underline{X}_o$ only in the limit of very strong penalty is the model of the EPRI study; see next section.

If the available information points clearly at a parametric form of $\lambda(m,\underline{X}_o)$, then a fully parametric model, with informative or noninformative prior on the parameters, becomes appropriate. A state of prior knowledge of this rather special type is implied by the classical exponential seismogenic-province model of regional seismicity.

As the previous review of alternative representations of seismicity indicates, there is a continuum of modeling possibilities, which range from noninformative nonparametric models and ML estimation (so-called "nonparametric historic methods") to completely parametric models (e.g., of the exponential seismogenic-province type). The full range of possibilities is spanned by nonparametric models with varying type and amount of prior knowledge.

There are also many possible estimation procedures. Those reviewed here are variants of maximum likelihood. Some of them (maximum penalized

954 40008

likelihood, maximum cross-validated likelihood, maximum a-posteriori density, and maximum penalized cross-validated likelihood) are flexible enough to express a wide range of prior states of knowledge. Their limitation is that they produce single estimates of $\lambda(m,\underline{X}_o)$. Two methods have also been suggested to characterize uncertainty on this function, for the case when λ is not constrained parametrically.

EXAMPLES OF NONPARAMETRIC MODELING AND DATA ANALYSIS TECHNIQUES

We now turn to example applications of the previous modeling ideas and to data analysis techniques that the authors have recently proposed. Some of the models and techniques have been developed and extensively tested in the course of the EPRI study, while others have been originated independently of that project.

Seismicity Model and Statistical Data Analysis in the EPRI Project

A partly parametric, partly nonparametric regional model is used in the EPRI (1985) project. The rate density $\lambda(m,\underline{X}_o)$ is written as

$$\lambda(m,\underline{X}_o) = 10^{a(\underline{X}_o)-b(\underline{X}_o)m}, \quad m\varepsilon[m_o,m_1(\underline{X}_o)] \qquad (12)$$

where a and b are spatially varying parameters and $m_1(\underline{X}_o)$ is a given upper-bound magnitude function. The model is parametric (truncated exponential) in m and nonparametric in $\underline{X}_o$.

Estimation of the functions $a(\underline{X}_o)$ and $b(\underline{X}_o)$ is through maximum penalized likelihood (MPL), with penalty imposed on departures of a and b from locally interpolated values. Therefore, the penalty is minimal if the functions a and b display regional trends but are locally smooth. This penalty may be regarded as either a prior on the roughness of a and b or a means of counteracting the

958 40008

erraticity of the ordinary ML solution, or a combination of the two. Additional prior information is included as follows:

1. Penalization for erratic variation in space of a and b can be restricted inside given "smoothing regions". There regions replace and generalize the classical notion of seismogenic provinces. Uncertainty on the configuration of the smoothing regions, e.g. reflecting uncertainty on the seismogenic process, can be expressed by providing alternative configurations and the probability that each configuration is the correct one.
2. The upper bound magnitude $m_1(\underline{X}_o)$ is assumed constant inside each smoothing region. The upper bound for Region i is treated as a random variable, with distribution provided by the user.
3. It is commonly believed that the parameter b is spatially more stable than the parameter a. This means that, for a location $\underline{X}_o$ inside Region i, $b(\underline{X}_o)$ should be estimated using not only local earthquake data, but also knowledge of b from other parts of the world. This external information on $b(\underline{X}_o)$ is expressed through a prior distribution.

The model of Eq. 12 is parametric in m and nonparametric in $\underline{X}_o$. By varying the amount of penalty on the likelihood, one can represent a variety of prior states of information on the functions $a(\underline{X}_o)$ and $b(\underline{X}_o)$. In particular, the use of very high roughness penalties produces constant estimates of a and b inside each smoothing region. This corresponds to the traditional model with homogeneous seismogenic provinces and exponential magnitude distribution. At the other extreme, one may impose no penalty on $a(\underline{X}_o)$ and $b(\underline{X}_o)$. In this case one obtains a solution that constrains future events to occur in space at the location of the historical earthquakes.

Uncertainty on the functions $a(\underline{X}_o)$ and $b(\underline{X}_o)$ is accounted for in a variety of ways:

1. Through uncertainty on the configuration of the smoothing regions.
2. For any given configuration of the smoothing regions, several pairs of functions $(a(\underline{X}_o), b(\underline{X}_o))$ can be generated by selecting different roughness penalties and prior distributions for b. Because the penalty does not strictly reflect prior beliefs but is used also to reduce the erraticity of ML solutions, the analyst is not asked to assign probabilities to the penalties for a and b and to the prior on b; rather, probabilities are assigned directly to the estimates $(\hat{a}(\underline{X}_o), \hat{b}(\underline{X}_o))$ obtained under various input conditions. (For comments on this procedure, see previous section.)
3. For any given configuration of the smoothing regions, any given penalty on $a(\underline{X}_o)$ and $b(\underline{X}_o)$, and any given prior on $b(\underline{X}_o)$, bootstrapping can be used to quantify uncertainty on the seismicity parameters; see previous section.

The assignment of probabilities at Point 2 is based primarily on judgement. One might aid the user in this task by ranking the various solutions in terms of their cross-validated likelihood. This procedure is however computationally very demanding. As an alternative, we provide several goodness-of-fit tests, which are simpler to make and are in some respects more informative than the cross-validated likelihood. The idea is to discretize space into cells (i_1,i_2), time into intervals i_t, and magnitude into ranges i_m and then compare the actual earthquake counts $N(i_1,i_2,i_t,i_m)$ in the various categories with the corresponding expected counts $E[N(i_1,i_2,i_t,i_m)]$. The expected counts depend of course on the functions $a(\underline{X}_o)$ and $b(\underline{X}_o)$ and on the detection probability

$P_D(m,X_o,t)$. One way to make the comparison is to test the assumption that the counts in each space-time-magnitude cell are Poisson, with the mean value given by the model. However, the power of this test is low due to the small counts and the results are difficult to display because of the 4-way classification of seismicity. More useful and statistically more meaningful results are obtained by testing the previous hypothesis on aggregated counts. For example, one might test the validity of the model on the geographical plane by using

$$\overline{N}(i_1,i_2) = \sum_{i_t,i_m} N(i_1,i_2,i_t,i_m)$$

and (13)

$$E[\overline{N}(i_1,i_2)] = \sum_{i_t,i_m} E[N(i_1,i_2,i_t,i_m)]$$

i.e. by testing for each (i_1,i_2) the hypothesis that $\overline{N}(i_1,i_2)$ is a sample from the Poisson distribution with mean $E[\overline{N}(i_1,i_2)]$. If the test fails for a fraction of geographical cells larger than the significance level at which the test is performed, then one should conclude that too much smoothing has been applied to obtain $\hat{a}(\underline{X}_o)$ and $\hat{b}(\underline{X}_o)$, or that the smoothing regions are not sufficiently honogeneous and should be redefined.

Similarly, one may compare

$$\overline{N}(i_m) = \sum_{i_1,i_2,i_t} N(i_1,i_2,i_t,i_m)$$

with (14)

$$E[\overline{N}(i_m)] = \sum_{i_1,i_2,i_t} E[N(i_1,i_2,i_t,i_m)]$$

to detect statistically significant departures from the assumed exponential distribution of magnitude. Tests performed on counts aggregated in space and magnitude but not in time further indicate whether the assumption of stationarity is plausible.

955 40008

As an illustration of the method, results of two analysis cases are presented for the "Boston-Ottawa region" in Fig. 1. The earthquake data used in this analysis is taken from the catalog compiled for the EPRI study. In the same study, this catalog has been preprocessed to convert all size measures to body wave magnitude m_b, to remove clustered events and to quantify incompleteness of the data as a function of time, location, and magnitude. For the results shown here, only events with $m_b \geqslant 3.3$ have been used and incompleteness of the catalog has been explicitly accounted for in the estimation of the recurrence parameters.

Table 1 shows estimates of the parameters a and b as functions of geographical location. Whereas b has the same meaning as in Eq. 1, the parameter a given in the table is the logarithm (base 10) of the annual rate of events with magnitude between 3.3 and 3.9, for a unit area of $(111.11\ \text{Km})^2$.

The estimates of Table 1 have been obtained by discretization of the region into half-degree cells and by imposing a certain amount of smoothing on the spatial variation of a and b. Smoothing is controlled independently for a and b and is imposed through a penalty on the likelihood function. The location of the half-degree cells is indicated by the longitude and latitude of the southeast corner of each cell.

The estimate of a is clearly a bimodal function on the geographical plane, with peaks in the northwestern and southestern parts of the region. Between these two modes, the region is relatively quiescent. The spatial variation of b is less pronounced.

Also indicated in Table 1 is the result of local statistical tests of observed and predicted counts in each half-degree cell, performed at two levels of significance (0.10 and 0.02). The counts on which the tests are performed are obtained through summation over the entire time period analyzed for each

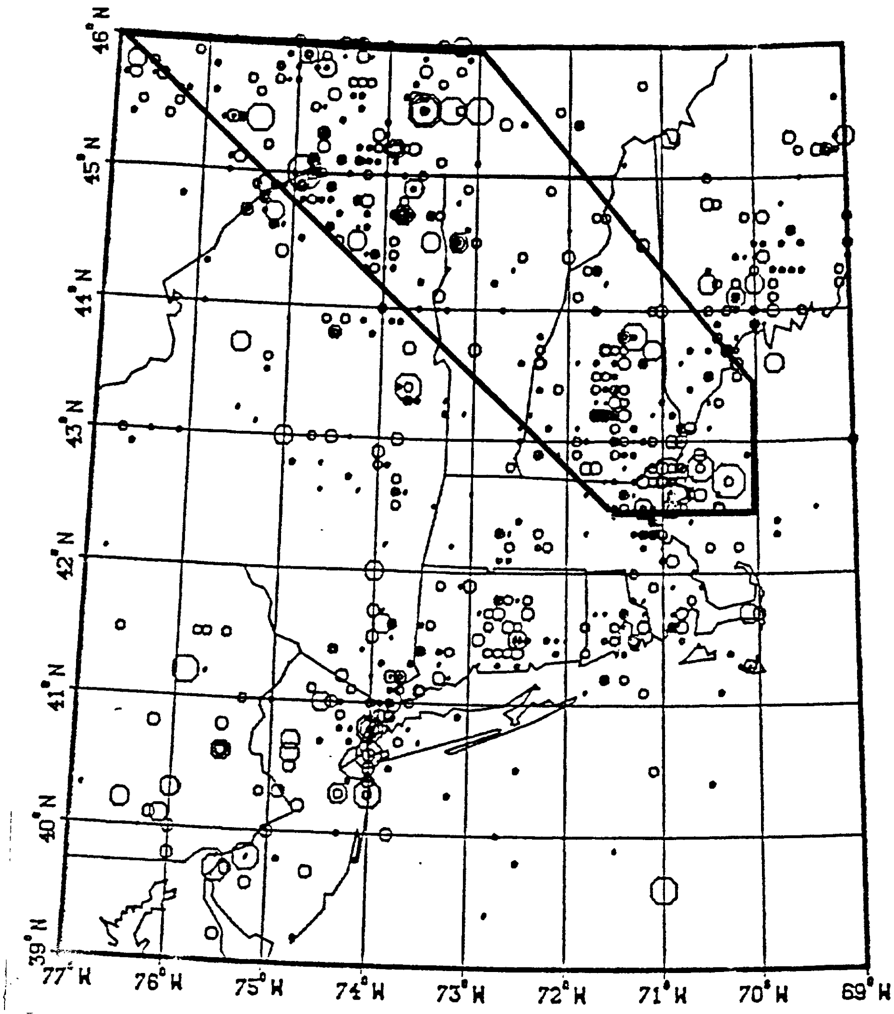


Fig. 1 - Boston-Ottawa region used to exemplify the seismicity model and diagnostic testing of the EPRI study.

9B2 40008

A-estimates

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5	-0.81	-0.80	-0.84	-0.66	-0.66	-0.67	-0.74	-0.95	-1.28	-1.39				
45.0		-0.82	-0.74	-0.66	-0.60	-0.68	-0.60	-1.00	-1.29	-1.42	-1.36			
44.5				-0.59	-0.46	-0.62	-0.69	-0.98	-1.30	-1.34	-1.26	-1.16		
44.0					-0.52	-0.57	-0.94	-1.08	-1.23	-1.24	-1.05	-1.05	-0.92	
43.5							-0.99	-1.23	-1.25	-1.10	-0.90	-0.73	-0.86	-0.68
43.0								-1.20	-1.20	-1.02	-0.64	-0.58	-0.71	-0.70
42.5										-0.90	-0.68	-0.47	-0.47	-0.75

B-estimates

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5	1.16	1.15	1.15	1.11	1.05	1.02	0.96	0.91	0.95	0.99				
45.0		1.16	1.14	1.10	1.08	1.05	0.89	0.92	0.95	1.02	1.05			
44.5				1.11	1.08	1.07	1.00	0.99	1.03	1.06	1.08	1.06		
44.0					1.10	1.09	1.09	1.07	1.09	1.11	1.07	1.04	1.02	
43.5							1.10	1.14	1.15	1.14	1.07	0.99	1.02	0.99
43.0								1.15	1.17	1.16	1.10	1.05	1.00	1.00
42.5										1.15	1.12	1.05	0.93	0.95

Significance test of observed versus expected counts

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5			-											
45.0						-	+							
44.5									-					
44.0							-							
43.5													-	+
43.0													-	
42.5													+	-

mb	Actual Count	Model Prediction	
3.30-3.90	132.00	130.52	
3.90-4.50	42.00	40.63	
4.50-5.10	9.00	14.69	-
5.10-5.70	7.00	5.67	
5.70-6.30	3.00	1.50	

	Test Flag	α
quiescence	<	0.02
	-	0.10
excessive	+	0.10
activity	>	0.02

Table 1 - Estimates of parameters a and b and goodness-of-fit tests for the "Boston-Ottawa region" in Fig. 1. Semi-parametric model of seismicity used in the EPRI study. Case of moderate smoothing.

magnitude and over all magnitudes, i.e. in a way similar to Eq. 13. The amount of flagging is about what one would expect at a significance level of 0.10. A second test has been made to compare observed and predicted counts summed over all locations and time periods but differentiated by magnitude. Again, the fit of the model is reasonably good.

Similar results have been otained using very high spatial smoothing of a and b; see Table 2. In this case, the solution corresponds to that obtained under the conventional assumption of homogeneity of the earthquake process inside the region. If one considers the amount of flagging, it is quite clear that such an assumption is very strongly contradicted by the historical data. The flagging also clearly indicates the presence of two modes, separated by a quiescent "valley". This analysis exemplifies the way in which diagnostic testing can be used to screen out unlikely models and to select appropriate source configurations and levels of smoothing of a and b within each source.

Automatic Identification of Earthquake Sources

In regions where the seismogenic process is not well understood or earthquakes cannot be associated with specific tectonic or geologic features, it is common to define "seismogenic provinces" or "sources" as regions inside which the earthquake process displays some statistical regularity. For example, sources may be identified as regions with homogeneous Poisson activity or, as in the EPRI study, with slowly varying parameters a and b.

In current practice, the identification of seismic sources is made judgementally on the basis of historical seismicity and geological and geophysical data. Because of the many qualitative elements that must be considered, judgement plays an important role in the definition of sources. However, the exclusive reliance on judgement, unaided by analysis, may become itself a source of error and of unnecessary differences among seismologists.

approximate. Also, one may use several variants of the splitting and merging phases and various postprocessing options to refine the boundary of the sources or "clean up" the solution; see Veneziano and Pais (1986).

Here, we just show an application example to the northeastern U.S. and adjacent Canada. The region of analysis extends in longitude between 64 and 80°W and in latitude between 40 and 48°N. The catalog used for analysis is that of Chiburis (1981), limited to main events. The region is progressively divided into smaller rectangular cells, which however are not allowed to have side lengths smaller than one longitude and one-half latitude degree. After splitting and merging, the solution is improved by optimizing the classification of boundary cells. Two source configurations are shown in Fig. 2, one obtained using the entire catalog, the other using only data since 1900. The dashed lines in the figure give source boundaries estimated by the algorithm and the square at the center of each cell has a side length proportional to the activity rate inside the source to which the cell belongs. Notice for example the high activity rates in Fig. 2a for the La Malbaie region and for Southern New Hampshire. Diagonal links indicate continuation of a source. It is interesting to note that, although the solutions in Figs. 2a and 2b are in qualitative agreement (e.g., they both display activity highs along the coast, in the La Malbaie region, and along the St. Lawrence valley near Ottawa), they also indicate that the level of seismicity of some of the sources has changed significantly during the past three centuries. In relation to the "Boston-Ottawa source" considered in the last section, analyses of the type in Fig. 2 could be used, together with geological and tectonic information, to define earthquake source configurations that are more compatible with historical seismicity.

A-estimates

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5	-0.81	-0.81	-0.81	-0.80	-0.80	-0.80	-0.80	-0.81	-0.81	-0.81				
45.0		-0.81	-0.81	-0.80	-0.80	-0.80	-0.80	-0.81	-0.81	-0.81	-0.81			
44.5				-0.80	-0.80	-0.80	-0.80	-0.81	-0.81	-0.81	-0.81	-0.81		
44.0					-0.80	-0.80	-0.81	-0.81	-0.81	-0.81	-0.81	-0.81	-0.81	
43.5							-0.81	-0.81	-0.81	-0.81	-0.81	-0.81	-0.81	-0.80
43.0								-0.81	-0.81	-0.81	-0.80	-0.80	-0.80	-0.80
42.5										-0.81	-0.80	-0.80	-0.80	-0.80

B-estimates

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03				
45.0		1.03	1.03	1.03	1.03	1.03	1.02	1.03	1.03	1.04	1.04			
44.5				1.03	1.03	1.03	1.03	1.03	1.04	1.04	1.04	1.04		
44.0					1.03	1.03	1.03	1.03	1.04	1.04	1.04	1.03	1.03	
43.5							1.03	1.04	1.04	1.04	1.03	1.03	1.03	1.03
43.0								1.04	1.04	1.04	1.03	1.03	1.03	1.03
42.5										1.04	1.03	1.03	1.02	1.02

Significance test of observed versus expected counts

	76.5	76.0	75.5	75.0	74.5	74.0	73.5	73.0	72.5	72.0	71.5	71.0	70.5	70.0
45.5			-	+					-					
45.0							>	-	-	<				
44.5					>		+		<	-				
44.0						>	-		-	-		-		
43.5								<	-	-		+	-	>
43.0									-	-	+	+		
42.5												>	>	-

mb	Actual Count	Model Prediction	
3.30-3.90	132.00	130.34	
3.90-4.50	42.00	40.86	
4.50-5.10	9.00	14.72	-
5.10-5.70	7.00	5.60	
5.70-6.30	3.00	1.47	

Test flag		α
quiescence	<	0.02
	-	0.10
excessive activity	+	0.10
	>	0.02

Table 2 - Estimates of parameters a and b and the goodness-of-fit tests for the "Boston-Ottawa region" in Fig. 1. Case of very high smoothing.

For example, it often happens that earthquake sources are identified on the basis of physical homogeneity but are treated as if they were also statistically homogeneous, without checking compatibility of the latter assumption with the historical earthquake record. In this section, we briefly describe and illustrate a procedure for the automatic identification of earthquake sources from historical catalog data (Veneziano and Pais, 1986). The procedure can account for uneven incompleteness of the catalog over the region of study, but does not include errors in the location of historical events, geologic or historic information, and nonstationarity or non-Poisson characteristics of the earthquake process. Because of these limitations, the method in its current formulation should not be used as a substitute but only as an aid to judgement.

The method is an adaptation of the algorithm of visual image analysis known as "split/merge". The latter is often used to "segment" digitized images, i.e. to partition two-dimensional pictures into regions with uniform characteristics such as intensity of gray, color, or texture; see Pavlidis (1977). The main idea is to progressively partition the region until each subregion can be considered internally homogeneous. This is the split phase of the method. One then aggregates neighboring subregions that pass certain homogeneity criteria (merging phase). For example, a splitting operation might consist of partitioning a rectangular cell into four subcells, whereas merging might replace two neighboring sources with a single source.

In the case of earthquake source identification and Poisson seismicity, criteria for region splitting and merging can be related to the statistical acceptance or rejection of the hypothesis of Poisson homogeneity. There are several statistical tests of such hypothesis; some are exact but apply only to special situations, others have a wider range of applications but are

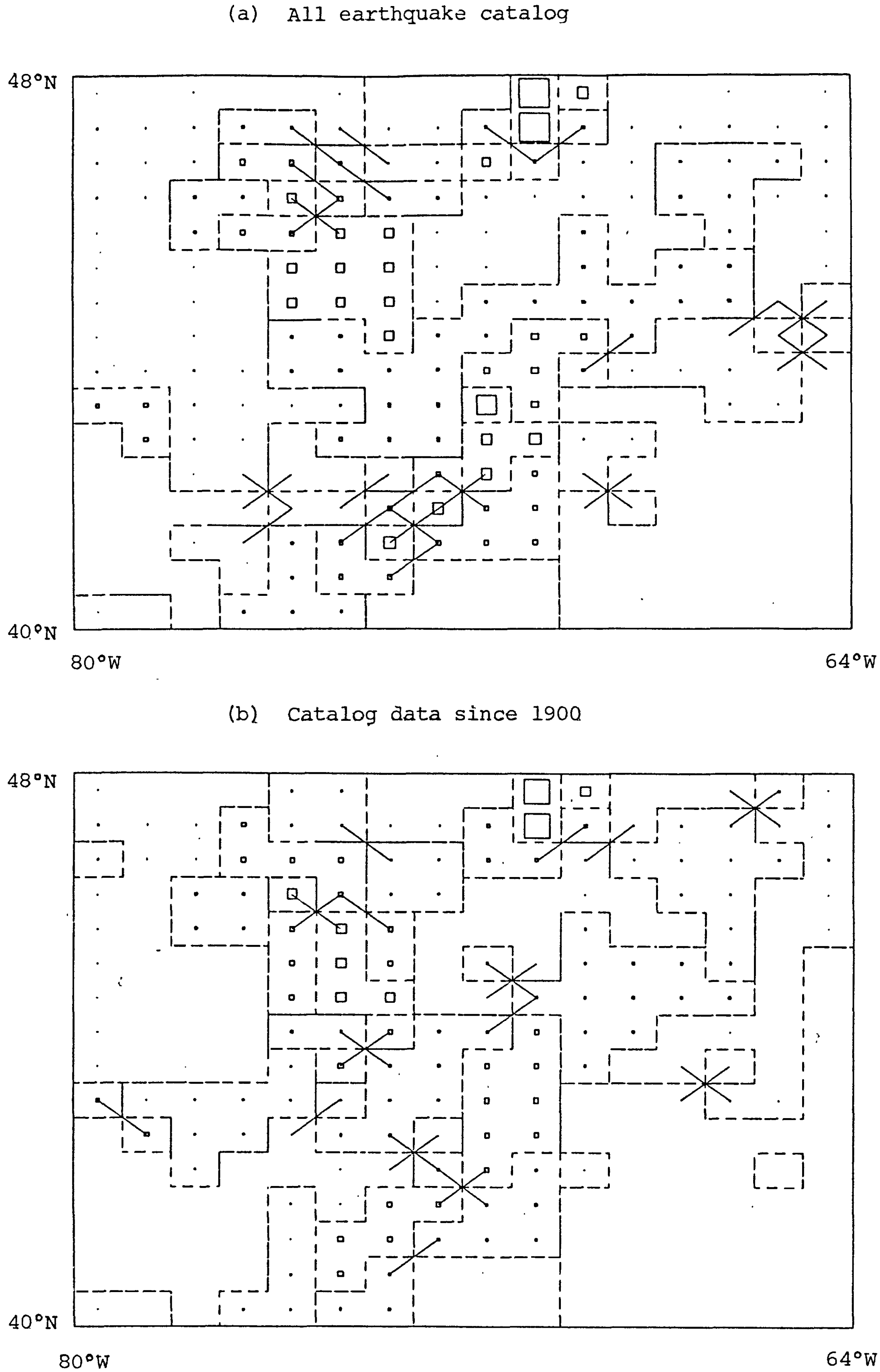


Fig. 2 - Automatic identification of earthquake sources. Example application to the Northeastern United States and adjacent Canada.

952 40008

CONCLUSIONS

Seismic risk analysts, regulatory agencies, and public officials have become increasingly concerned with the objective validity of earthquake hazard estimates. The main cause for concern is the variability in calculated hazard due to differences in judgement among experts and to the fact that different models currently in use (e.g., historic and source models) sometimes produce inconsistent results.

The main objective of this paper is to introduce a broad class of seismicity models, which includes those presently in use, and to provide a unified framework for the interpretation of such models. It is demonstrated that, at a conceptual level, differences in the models express differences in prior knowledge. Such a conclusion may appear obvious, but it has important consequences; for instance, it implies that comparison of different models is rather meaningless, unless one first establishes the credibility of the prior knowledge implied by them. In particular, models that postulate the existence of homogeneous seismogenic provinces incorporate strong prior assumptions, which need be verified before results are considered plausible. According to the unified approach presented here, one needs just one model of seismicity, with the following characteristics:

1. The model should be sufficiently general to include as special cases all current representations of earthquake activity. The model should therefore be of the nonparametric type.
2. When fitting the model to data, it should be possible to account for a variety of prior states of information.

These conditions are necessary for the comparison of results obtained by different users and for the accurate representation of prior information. An example of semi-parametric model which comes close to satisfying both requirements is that used in the EPRI study.

There are many ways in which models can be fitted to data, including the standard procedures of maximum likelihood and least squares. Better methods can be developed, based on maximization of a penalized and/or cross-validated likelihood. The main advantage of cross validation is that decisions about smoothing of the fitted model are made internally based on data, rather than externally based on judgement.

Finally, the paper suggests innovative ways in which historical earthquake data can be used. The most important one is the quantification of the goodness-of-fit of proposed models. Analyses of this type, exemplified in the previous section, have been routinely performed in the EPRI study. There is, however, an obvious need for further development of statistical procedures to validate or reject commonly made assumptions, for example exponentiality of the magnitude distribution, Poisson occurrences, stationarity in time, local homogeneity in space, and independence of earthquake magnitudes.

The scope of the present paper is limited to the modeling of earthquake occurrences and related data analysis techniques. Problems of errors in earthquake data, magnitude conversion, clustering, and incompleteness of the historical catalog are not considered. These problems have been addressed by the authors in the EPRI study and are discussed in detail in Van Dyck (1986).

ACKNOWLEDGMENT

The authors are grateful to Artur Luna Pais for his cooperation in the part of the paper that deals with the automatic identification of earthquake sources.

References

Chiburis, E.F. (1981), "Seismicity, Recurrence Rates and Regionalization of the Northeastern U.S. and Adjacent Southeastern Canada," NUREG/CR-2309.

EPRI (1985), "Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States," Research Project P101-29, Palo Alto, California.

Pavlidis, T. (1977), Structural Pattern Recognition, Springer-Verlag, New York.

Van Dyck, J. (1986), "Statistical Analysis of Earthquake Catalogs," Ph.D. Thesis., Dept. of Civil Engineering, MIT, Cambridge, Massachusetts.

Veneziano, D. and Pais, A.L. (1976), "Automatic Source Identification Based on Historical Seismicity," 8th European Conference on Earthquake Engineering, Lisbon, Portugal.

SOURCE MODELS AND ANALYSIS OF UNCERTAINTIES IN EARTHQUAKE HAZARD ASSESSMENT

By

Armen Der Kiureghian and Rodrigo Araya

Department of Civil Engineering, University of California

Berkeley, California 94720

Introduction

If, for each site in the United States, we were lucky enough to have recordings of ground motion intensities (e.g., peak ground accelerations) for all past earthquakes, the assessment of seismic hazard (i.e., the probability of the intensity exceeding a given threshold over a specified interval of time) would be a simple matter of statistics. The inexistence of such a data for any given site has necessitated the development of indirect probabilistic methods whereby the hazard is computed based on stochastic modeling of: (a) the occurrences of earthquakes in time and space, (b) the earthquake source, (c) the attenuation of seismic waves through the ground medium, and (d) the local site effects. The basic idea is that one can evaluate non-site-specific such models with relatively limited data and then proceed with computing the seismic hazard using standard probabilistic techniques.

The level of sophistication in selecting a stochastic model depends on three items: (a) the level of our understanding of the physical processes involved, (b) the amount of data available for predicting the parameters of the model, and (c) our ability to carry out the computation of the hazard. One objective of the present paper is to describe a general, efficient methodology for computing the hazard, which essentially removes the last limitation in our selection of the aforementioned stochastic models.

Irrespective of what models are selected, it is essential to properly account for the uncertainties inherent in such models in seismic hazard assessment. These uncertainties might be directly incorporated in the estimated hazard (Bayesian approach) or expressed in terms of a confidence interval or distribution on the estimated hazard. Clearly, one would expect that a more sophisticated model (i.e., one which is a more realistic representation of the reality) contains less uncertainty and, hence, the corresponding estimated hazard is less dispersed.

As an example, consider the models used for describing the earthquake source and the attenuation law. (These two have to go together.) At the most crude level, one may choose to represent the source in terms of a single point, e.g., the epicenter or the focus, and the attenuation law in terms of a corresponding distance, i.e., the epicentral or focal distance. In that case, the effects of source dimension (i.e., rupture length, width, depth, and dip angle), directivity, rupture velocity, etc., are all lumped into the model uncertainty of the attenuation law. Whereas this crude model may have a large uncertainty, it has the advantage of a large data base, since the epicentral or focal locations of many earthquakes are known. At the next level, one may choose to represent the source in terms of a rupture plane with dimensions dependent on the magnitude of the earthquake. In that case, the most appropriate choice for the distance term in the attenuation law is the shortest distance from the recording site to the rupture plane (Der Kiureghian, 1977). Only the effects of directivity, rupture velocity, etc., are then lumped into the model uncertainty of the attenuation law. One may expect that this model will have less uncertainty; however, it will have the disadvantage of a smaller data base, as there may not exist many earthquakes for which the rupture geometry is known. At an even more refined level, the directivity effect may also be included in terms of the relative azimuth of the site with respect to the direction of the rupture. In that case, one will have to work with an even smaller data base, as the characteristics of the directivity phenomenon are seldom known for recorded earthquakes.

Lack of a large statistical data base in constructing refined attenuation laws may be compensated by mathematical modeling of the source mechanism and the propagation of seismic waves. Source models such as Haskell's (1966) can be used to derive semi-empirical attenuation laws which can be evaluated with relatively small data base. Attempts in this direction have already been made (Faccioli, 1983) and there is no doubt that continued progress in this area is forthcoming. With increasing worldwide instrumentation, it is only a matter of time before refined models of attenuation laws can be verified by statistical analysis.

It is the intention here to point out that model uncertainties are an integral part of seismic hazard assessment and that they should be properly accounted for in such assessments. Also, refined stochastic models (of earthquake occurrences, sources, attenuation laws, and local site characteristics) are needed in seismic hazard analysis, even if existing data cannot support such refined models. With increasing improvement in our data base and understanding of the physical processes involved, such refined models can be verified in time, thus, resulting in reduction of the dispersion in our assessed seismic hazard. In the meantime, refined models can be used to investigate the importance of various factors on seismic hazard, which will help us understand the areas where further refinements or additional data are needed.

In the following sections, a new formulation for seismic hazard assessment including the effect of directivity is presented. It is shown that the effect of directivity on the hazard is significant and its neglect may lead to serious underestimation of the hazard. A general methodology for incorporating model uncertainty in seismic hazard assessment is presented. This formulation and methodology are based on modern concepts of structural reliability theory and provide a practical tool for computation of the hazard when a large number of variables, representing refined models or model uncertainties, must be included in the analysis. Both topics are currently under development by the authors.

Source Model Including Directivity

Recent experience with the 1979 Coyote Creek, the 1979 Imperial Valley, the 1980 Livermore Valley, and the 1984 Morgan Hill earthquakes has shown that the azimuthal change in the level of ground shaking around the seismic source due to the direction of the moving rupture can be significant, particularly for near-field regions. This effect, known as the *directivity* effect, only recently has gained attention in the context of seismic hazard analysis. To include the effect in such analysis, models for the earthquake source and the attenuation law that account for the direction of rupture propagation are needed. Whereas source models that account for this effect have been around for many years (e.g., Haskell (1966)), it is only recently that attempts at formulating attenuation laws including the directivity effect have been made (e.g., Monzón (1980), Schoof et al. (1984), Faccioli (1983)). The models by Monzón and Schoof are rather crude and are not supported by either recorded data or theoretical grounds. The attenuation law by Faccioli (1983) is based on the far-field displacement spectrum of the source model by Haskell (1966) and has a theoretical foundation. This law is expressed in the form

$$A = \frac{M_0}{4\pi\rho V_s^3} \frac{\exp(-q R)}{R} \frac{K_t^3}{\sqrt{K_t T_r}} D(\alpha, m) \tag{1}$$

in which A is the root-mean-square acceleration, M_0 is the seismic moment, V_s is the shear-wave velocity, R is the distance, ρ is the density of the medium of

10B1 40008

propagation of waves, K_t^{-1} is the correlation time of the stochastic rupture model, q is an attenuation parameter, T_r is the duration of rupture, and $D(\alpha, m)$ is the directivity factor given by

$$D(\alpha, m) = \frac{1}{2 - m \cos\alpha} \left\{ \frac{3 - m \cos\alpha}{1 - m \cos\alpha} \right\}^{1/2} \tag{2}$$

in which $m = V_r / V_s$, the ratio of the rupture velocity V_r to the shear wave velocity V_s, is the seismic Mach number and α is the angle between the direction of the rupture propagation and the radius vector from the epicenter to the site. Using empirical relations between K_t, T_r and the seismic moment M_0, and between the seismic moment M_0 and the moment magnitude M, the above expression may be written in the familiar form

$$A = \frac{c_1 \exp(c_2 M - c_3 R)}{R} D(\alpha, m) \tag{3}$$

where c_1, c_2 and c_3 are regional constants that must be obtained through regression analysis. These constants may also include the effects of the local conditions at the recording site.

The above model is not applicable to the near-field region since it diverges for R approaching zero. Furthermore, it does not account for the rupture dimension which is important in the near field. To account for these effects, R is interpreted as the nearest distance from the recording site to the rupture plane and Eq. 3 is modified by adding a constant term in the denominator to read

$$A = \frac{c_1 \exp(c_2 M - c_3 R)}{R + c_4} D(\alpha, m) \tag{4}$$

At the present time, this equation should be regarded merely as a plausible model for the attenuation law. Figure 1 shows contours of equal acceleration obtained with this model for various sets of the parameters, demonstrating the effects of variation in the distance, the magnitude, the seismic Mach number, and the rupture length. Regression tests with recorded data are planned in

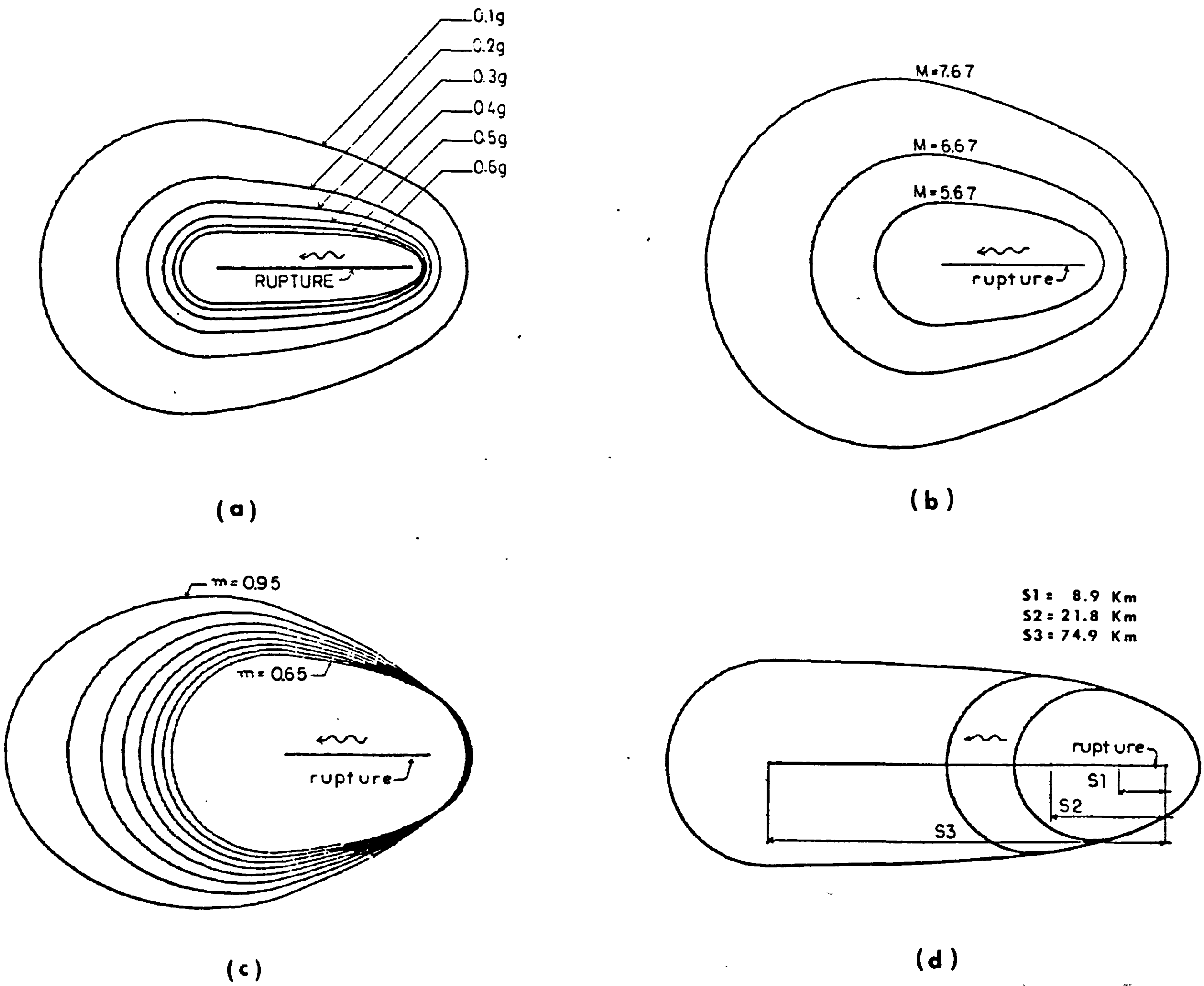


Figure 1. Contours of Equal Acceleration Showing Variations with (a) Distance, (b) Magnitude, (c) Seismic Mach Number, and (d) Rupture Length

1056 40008

order to ascertain the validity of the model and to evaluate the parameters c_1-c_4.

Consistent with the above attenuation law, the earthquake source is described in terms of the magnitude, M, the rupture length and width, S and W, the depth below the fault trace, H, the dip angle, δ, the directional azimuth of the rupture propagation, φ, and the seismic Mach number, m. The location of the source with respect to a specific site may be described by the Cartesian (or polar) coordinates X and Y of the epicenter. For future earthquakes, these variables in general are random. In addition, one has to account for the uncertainty in the attenuation law, which arises from the imperfection of the model in Eq. 1, the uncertainty in the relationships used for K_t, T_r, M_0, etc., as well as the uncertainty arising from the stochastic nature of the propagation path and the local site conditions. This may be accomplished by considering the constants c_1-c_4 as random variables. In practice, this entire uncertainty is often lumped onto c_1.

Now, let the vector $\mathbf{Z} = (M, S, W, H, \delta, \varphi, m, X, Y, c_1, c_2, c_3, c_4)$ represent the set of all random variables considered in the hazard analysis. Then, for a given earthquake, the probability of exceeding the intensity level a is given by

$$P(A > a) = \int_{A(\mathbf{z}) > a} f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \tag{5}$$

where $f_{\mathbf{Z}}(\mathbf{z})$ is the joint probability density function of $\mathbf{Z}$ and the integration is over all values of $\mathbf{Z} = \mathbf{z}$ for which the computed intensity from Eq. 3 exceeds a. In order to compute the seismic hazard for a specified interval of time, T, the above probability has to be used in conjunction with a stochastic model of earthquake occurrences in time. For example, if the Poisson model is used, then

$$P(A > a \text{ in } T) = 1 - \exp[-\nu T P(A > a)] \tag{6}$$

where ν is the mean rate of earthquake occurrences in time.

Offhand, evaluation of the integral in Eq. 5 appears to be a formidable task. One apparent difficulty is the formulation of the joint distribution of **Z**. The other is that, even with today's high-speed computers, the integration over so many variables may not be feasible. These difficulties, however, are easily resolved by the use of fast integration techniques developed for structural reliability, as described in the following paragraphs.

The joint distribution of **Z** can be formulated by conditioning. We first condition on the magnitude, M. The conditional distributions of S and W may then be obtained from regression analysis of recorded (or inferred) data. In general, these two variables for any M are correlated and their correlation coefficient can be estimated from data. Having this information, a joint distribution model for S and W can be formulated using a Nataf-type model (Der Kiureghian et al., 1985). The variables H and δ may be regarded as independent of the other variables and their distributions may be described for each potential earthquake source in a region of interest. The seismic Mach number, m, may also be regarded as independent of other variables and its distribution may be assigned on regional basis. The distribution of φ depends on the geometry of the seismic source. For a known fault, φ may assume one of two values indicating the direction of the rupture propagation along the fault. For area sources, φ may be assigned a distribution based on the perceived orientation of the faults. In particular, if the orientation of the faults is entirely unknown, a uniform distribution over the interval 0-2π is appropriate. The joint distribution of X and Y depends on the locations of earthquake sources in the region and the distribution of the expected seismic activity within each source. For example, for a known fault, X and Y may assume values satisfying the governing equation of the faultline in the assumed coordinate system and the distribution may be obtained by assigning a distribution to the expected seismic activity along the fault. For an area source, X and Y may assume values over the entire area of the source and the distribution may be obtained by assigning a distribution to the relative activity over the area. In most cases, the distribution of activity within an area may be assumed to be uniform. The coordinates X and Y of the epicenter in general will depend on the variables S and φ. (This is because certain locations of the epicenter might be impossible for a given S and φ.) For idealized source geometries, the conditional distribution of X and Y for given S and φ can be easily formulated using standard probability techniques. Finally, the joint distribution of c_1-c_4 may be obtained from the residuals of the regression law.

As stated earlier, a straight-forward numerical evaluation of the integral in Eq. 5 is difficult because of the large dimension. However, integrals of this kind have been of interest in structural reliability for a long time and efficient techniques for their evaluation are now available (Ang et al. 1984, Madsen et al. 1986). Briefly stated, these techniques consist of two steps: (a) Transformation of the variables into the standard normal space, i.e., $\mathbf{U} = \mathbf{T}(\mathbf{Z})$, where $\mathbf{U}$ is an uncorrelated normal vector with zero means and unit standard deviations. The integration domain in the transformed space is expressed by $A(\mathbf{T}^{-1}(\mathbf{u}))>\alpha$, where $\mathbf{T}^{-1}$ is the inverse transformation; (b) Replacement of the boundary of the integration domain in the standard space with an approximating boundary for which the exact or approximate solution of the integral exists. In particular, in the first-order approach the boundary is approximated by the tangent hyperplane at the nearest point to the origin and the approximation to Eq. 5 is

$$P(A>\alpha) \approx \Phi(-\beta) \tag{7}$$

where β is the distance from the origin to the hyperplane, and in the second-order approach the boundary is approximated by a hyperparaboloid at the nearest point to the origin and the approximation to Eq. 5 is

$$P(A>\alpha) \approx \Phi(-\beta) \prod_{i=1}^{n-1} (1-\beta\kappa_i)^{-1/2} \tag{8}$$

where n is the number of variables and κ_i are the principal curvatures of the boundary at the nearest point. These approximations work well because most of the contribution to the probability of interest comes from the neighborhood of the nearest point, where the integration boundary is well approximated.

One important byproduct of the aforementioned approach is a set of sensitivity factors which are obtained in the process of finding the nearest point. These factors represent partial derivatives of the hazard with respect to each variable or parameter involved in the formulation. Clearly, such information would be of vital interest in seismic hazard analysis as well as other areas of earthquake engineering.

The Importance of the Directivity Effect

As an illustration of the significance of the directivity effect, a comparison is made in Fig. 2 between computed hazards with and without the directivity effect for a hypothetical source. Shown in this figure are ratios of computed hazards including the directivity effect to hazards excluding the directivity effect for four sites, plotted against the root-mean-square acceleration. The results without the directivity effect were obtained by using the expected value of $D(\alpha, m)$ in Eq. 4, where α was assumed to be uniformly distributed over the interval 0-2π and m was assumed to be uniformly distributed over the interval 0.60-0.95. For the results with the directivity effect, it was assumed that the rupture is equally likely to propagate in one or the other direction along the fault. In order to isolate the directivity effect, uncertainties in the attenuation law and the relation between S and M were ignored in this analysis.

As can be seen in Fig. 2, the effect of directivity is rather significant, particularly at high acceleration levels (low hazard levels), which are of engineering interest. At such values, the effect of directivity is an increase in the hazard by a factor of around 10 to 20 for the example considered. As one would expect,

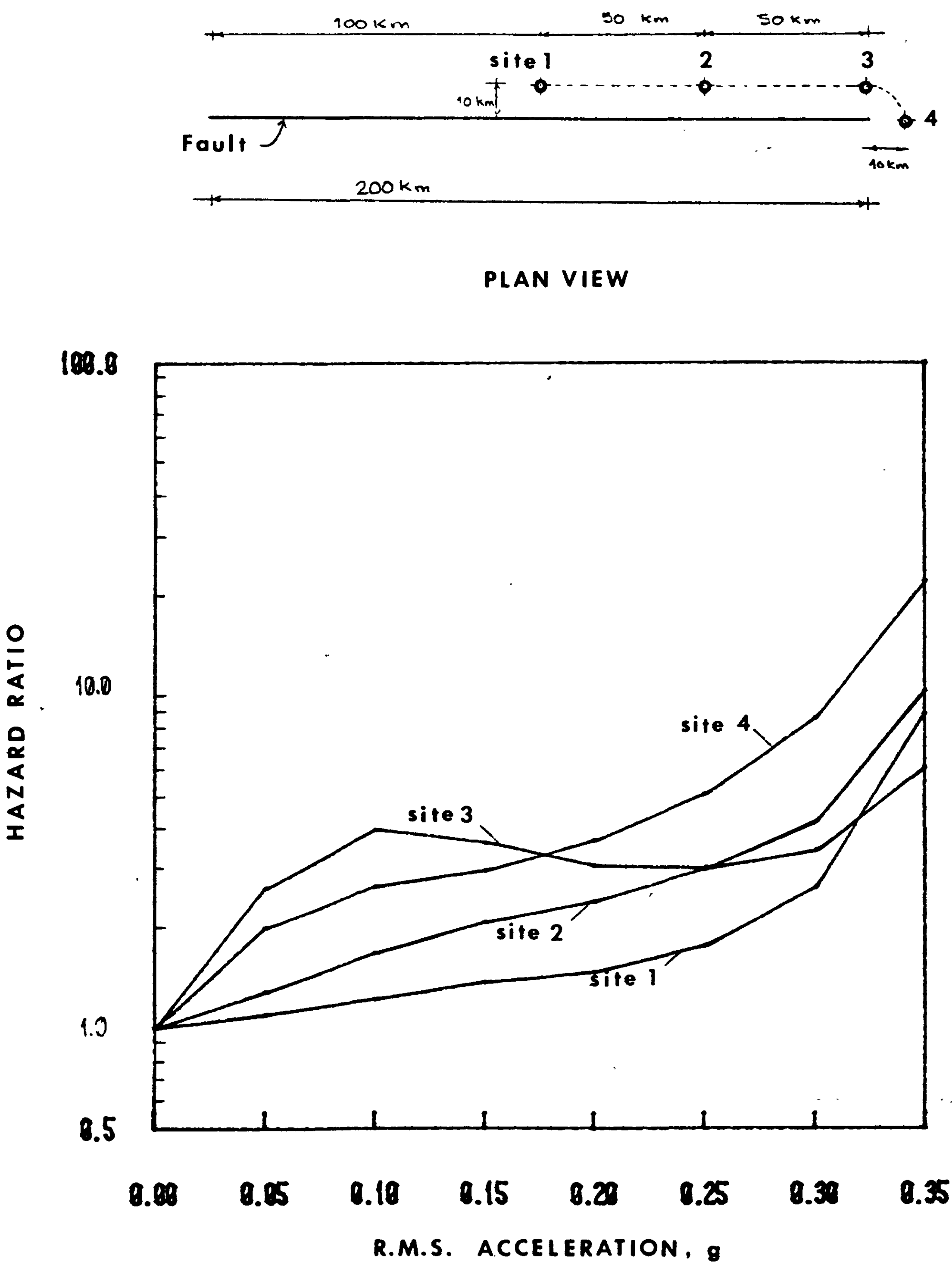


Figure 2. Ratio of Estimated Hazards With and Without the Directivity Effect

10B7 40008

this effect appears to be more significant for sites located at the extremity of the fault. These results clearly demonstrate the need for serious consideration of the directivity effect in the assessment of seismic hazard.

Analysis of Uncertainties

In the formulation presented in the previous section some of the elements of model uncertainty were incorporated. These included, for example, the uncertainty in the attenuation model and the uncertainties in the relations of S and W to M. Whether these are prediction or model uncertainties or are inherent variabilities of the underlying processes, is a matter of judgement. In any case, there are other uncertainties that need to be considered in the analysis. These include, for example, the uncertainties in parameters of the various distributions and the uncertainties in the parameters of the stochastic model describing the occurrences of earthquakes in time. Good examples of these are the upper-bound magnitude (a parameter of the distribution of magnitudes) and the mean occurrence rate, ν. Methods for incorporating such uncertainties in seismic hazard assessment are described in this section.

Let the vector $\mathbf{Z}$ denote the set of random variables representing inherent variabilities and the vector $\boldsymbol{\Theta}$ denote the set of variables representing model uncertainties. As stated earlier, this separation might be quite arbitrary and judgemental. In the Bayesian approach, one does not make a distinction between $\mathbf{Z}$ and $\boldsymbol{\Theta}$ and directly incorporates the uncertainty in $\boldsymbol{\Theta}$ in the estimated hazard. This makes sense, since the distinction is judgemental in the first place. In terms of computation of hazard, this is particularly attractive as one only needs to include $\boldsymbol{\Theta}$ as a subset of $\mathbf{Z}$ in the previous formulation.

In recent years, in PRA studies of nuclear power plants, there has been a tendency to separate the contribution of modeling uncertainties to seismic hazard from the contribution of variables perceived to represent inherent variabilities. Although the writers do not see a compelling reason for this separation, it is the intention here to show that the methodology described in the previous section can be effective in carrying out a seismic hazard assessment where the contributions of the two sources of uncertainty are separately analyzed.

Let $h(a_0, \vartheta)$ denote the estimated hazard for acceleration level a_0 for a given set of model parameters, $\Theta = \vartheta$. For each such set ϑ, $h(a_0, \vartheta)$ may be regarded as a point estimate of the hazard for level a_0. In particular, for ϑ equal to the mean of Θ, the estimated hazard may be regarded as a "best" estimate. To represent the effect of model uncertainty, one may compute a distribution for $H(a_0, \Theta)$ for random Θ. Using the aforementioned methodology,

$$P(H(a_0, \Theta) < h) = \int_{H(a_0, \vartheta) < h} f_\Theta(\vartheta)\, d\vartheta \tag{9}$$

where $f_\Theta(\vartheta)$ is the joint PDF of Θ and the computation is repeated for each h. This distribution can be numerically differentiated to compute a probability density function for the hazard at acceleration level a_0, from which the mean, standard deviation, or various confidence intervals of the hazard may be computed. This idea is illustrated in Fig. 3, where the Bayesian hazard curve as well as the "best" estimate curve are also shown.

Alternatively, for each hazard h_0, the corresponding acceleration level may be expressed as a function of Θ, i.e., $A(h_0, \Theta)$. The distribution of the acceleration for the fixed hazard, resulting from the uncertainty in Θ, is then obtained from

$$P(A(h_0, \Theta) < a) = \int_{A(h_0, \vartheta) < a} f_\Theta(\vartheta)\, d\vartheta \tag{10}$$

Again, this distribution may be numerically differentiated to obtain the probability density function or the mean, and standard deviation, or confidence intervals of the acceleration at the selected hazard level. This idea is also shown in Fig. 3.

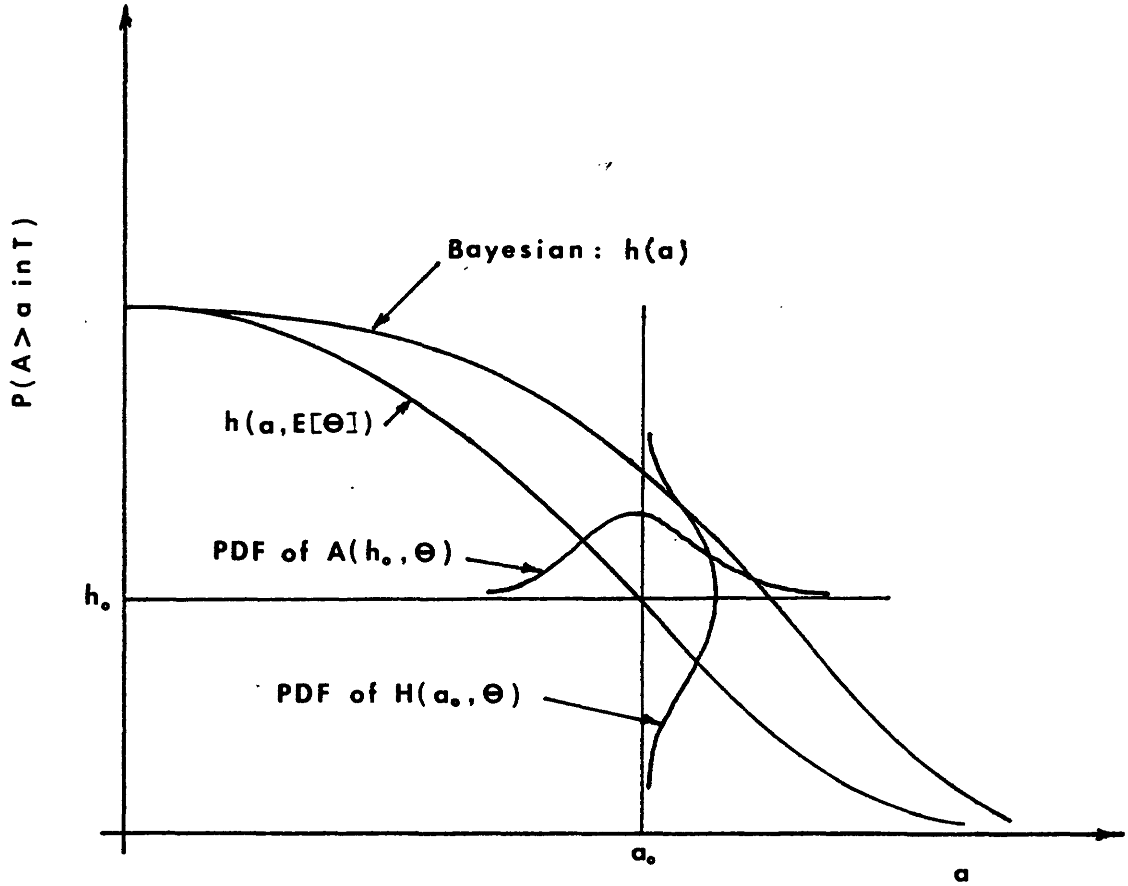


Figure 3. Earthquake Hazard Showing Contributions of Uncertainties

The above methodology is currently under development by the authors and further details and results will be forthcoming.

References

1. Ang, A-H. S., and W. H. Tang, *Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability*, John Wiley & Sons, New York, N.Y., 1984.

2. Der Kiureghian, A. and A. H-S. Ang, "A Fault-Rupture Model for Seismic Risk Analysis," *Bulletin of the Seismological Society of America*, Vol. 67, No.4, August 1977, pp. 1173-1194.

3. Der Kiureghian, A., and P-L. Liu, "Structural Reliability Under Incomplete Probability Information," *Report UCB/SESM-85/01*, Division of Structural Engineering and Structural Mechanics, Department of Civil Engineering, University of California, Berkeley, CA., January 1985.

4. Faccioli, E., "Measures of Strong Ground Motion derived from a Stochastic Source Model," *Soil Dynamics and Earthquake Engineering*, Vol. 2, No.3, 1983, pp. 135-149.

5. Haskell, N. A., "Total Energy and Energy Spectral Density of Elastic Wave Radiation from Propagating Faults. Part 2: A Statistical Source Model," *Bulletin of the Seismological Society of America*, Vol. 56, No. 1, February 1966, pp. 125-140.

6. Madsen, H. O., S. Krenk, and N. C. Lind, *Methods of Structural Safety*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1986.

7. Monzón-Despang, H., "Seismic Performance Analysis of Spatially Distributed Systems," *Report No.43*, The John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA., June 1980.

8. Schoof, C. C., R. N. Haar, and R. J. Geller, "Normal Mode Synthetics of Strong Ground Motion," *Proceedings*, 8th World Conference on Earthquake Engineering, San Francisco, CA., July 1984, Vol. II, pp. 449-456.

APPLICATION OF EARTH SCIENCE INFORMATION IN SEISMIC HAZARD STUDIES IN THE EASTERN UNITED STATES

BY

Kevin J. Coppersmith
Geomatrix Consultants
One Market Plaza, Spear Tower
San Francisco, California 94105

INTRODUCTION

Key words in the title of this paper are "earth science information". Classical seismic hazard assessments (SHA) have been conducted for some time in the eastern United States (EUS), but only recently has there been a strong move to incorporate tectonic information into the analyses. This paper will focus on the motivation, the conceptual basis, and the methods that have recently been employed to identify and characterize seismic sources for SHA in the East. I will not be considering other important aspects of the EUS hazard assessment problem such as ground motion attenuation or hazard modelling.

A principle motivation of the recent hazard studies in the EUS has been the recent geologic studies at the locations of past large-magnitude earthquakes, particularly that at Charleston, South Carolina in 1886. Despite rather extensive geologic investigations in the meisoseismal region of the 1886 earthquake carried out over the past decade, the causitive structure for this event has not yet been unequivocally identified. Instead, several potential source structures have been found, none of which are unique to the Charleston region or even to the southeastern United States. This raises the possibility that other locations in the EUS containing geologic structures similar to those at Charleston might be capable of generating Charleston-type earthquakes (m_b 6.8). This possibility has become known as the "Charleston issue".

Most seismic ground motion studies for the design and verification of design bases for nuclear power plants in the EUS have relied fairly heavily on the historical seismicity record to define the locations of future earthquake occurrences that may affect these sites. Coupled with the historical data has been a reliance on the assumption that the locations of higher (and lower) levels of seismic activity in the historical past will respectively remain so in the next few decades, i.e., the seismicity is essentially stationary in space, magnitude and rate. This assumption has led to the definition of seismic source zones around the Charleston and New Madrid regions that would be expected to contain future large events and, outside of which, the likelihood of large earthquake occurence is assumed to be much less. Paleoseismic investigations at both New Madrid and Charleston have suggested the recurrence of large events at these locations over the past few thousand years, consistent with a stationarity assumption. However, the absence of large events at locations away from these zones during the same time period has not yet been verified.

In order to address the Charleston issue as well as the associated issue of stationarity, a seismic hazard assessment should contain the following:

- Incorporation of uncertainty in the seismic potential of known tectonic features
- Assessments of the causal association of structure with large historical earthquakes
- Consideration of the scientific criteria that may be used to evaluate which structures are seismogenic

In considering the most appropriate methods to accomplish the above, it is important to bear in mind the rather severe restrictions caused by the limited data available in the eastern United States to make seismotectonic evalulations, as summarized below:

Data Not Available in EUS

- High-resolution instrumental seismicity data, except at selected locations
- Instrumental recordings of large earthquakes ($m_b \geq 6$)

- Calibrated source property data at large magnitudes (stress drop, seismic moment, etc.)
- Historical surface rupture
- Identified active faults
- Comprehensive mapping of late Cenozoic deformation

Several types of geologic and geophysical data are available to some degree and should provide a basis for any SHA methodology that is designed to include tectonic considerations. These are summarized below.

Data Available in the EUS

- Historical seismicity (~200 yr) and local instrumental seismicity (~10 yr)
- Potential field data (gravity, aeromag, etc.)
- Deep reflection locally (e.g. COCORP)
- Basic geologic mapping/borehole data
- In-situ stress locally
- Post-Cretaceous fault mapping locally
- Paleoseismicity studies (Charleston, New Madrid, Meers fault)

The remainder of this paper will summarize the recent attempts that have been made to include tectonic information into SHA, followed by a personal opinion of the likely future directions of SHA in the East.

Seismic Source Identification

Whereas the seismic sources defined for SHA in the western United States are typically defined by a knowledge of known active faults, sources in the EUS are typically defined primarily from the spatial distribution of seismicity, tempered by qualitative geologic intuition. The boundaries of seismic source zones are usually influenced by interpretations of the location of "seismotectonic provinces", although no explicit identification of the causal structures is usually given.

The reason for the reliance of seismicity data to identify seismic sources in our poor understanding of the causal mechanism(s) of earthquakes in the EUS; this includes both the geologic structures that are potential seismic sources as well as the nature and causes of crustal stresses. As a result, no attempt has been made to define seismic sources for SHA from the standpoint of first principles (e.g., earthquake process, strength properties, strain accumulation, etc.)

A variety of hypotheses have been proposed regarding earthquake occurrences in parts of the EUS that have direct implications to source identification. Such hypotheses include reactivation of regional decollement underlying the Appalachians, stress concentration near plutons, and reactivation of Appalachian structures or Mesozoic rift-related structures as reverse faults. These hypotheses are generally not well-developed in terms of the physical processes leading to reactivation and earthquake occurrence. Therefore, their use in predicting which features may be seismogenic is limited. For example, at some locations such as Charleston, several candidate structures are present and, as a result, several competing hypotheses have been presented for what generates the earthquakes in this region. At other locations, structures exist but they have not been associated with historical earthquakes and geologic evidence of recent deformation has not been documented. At still other locations, seismicity has occurred but no causitive structures have been identified. The proposed hypotheses specify the type of structures that might be active but they do not specify the criteria for evaluating the activity of any particular site. Before a meaningful hazard analysis can be conducted, these criteria must be defined and individual tectonic features must be evaluated. The following section discusses one method for making this evaluation.

EPRI Methodology

As part of a study sponsored by the Electric Power Research Institute (EPRI) to assess the probabilistic seismic hazard at nuclear power plant sites east of the Rocky Mountains, a methodology was developed to incorporate tectonic information into the identification of seismic sources for the SHA.

The primary goals of the source identification part of the EPRI methodology were to establish a framework to explicitly elicit expert opinion from six tectonic evaluation teams to describe the state of knowledge regarding two critical issues:

1. The causes of crustal stress in the eastern United States and the present state of the stress regime.

2. The identification and characterization of the tectonic features in the East that may localize or generate moderate-to-large earthquakes ($m_b \geq 5.0$). The result of this analysis is called a Tectonic Framework.

Rather than attempt to resolve the uncertainties associated with these issues, the methodology was designed to capture the present state-of-knowledge and uncertainty as a "snap-shot" of the understanding of the key issues, with the assumption that the interpretations will possibly change with further data collection and analysis. To document the consideration of these issues, the methodology involves a series of interim results or porducts that stand alone, as well as become integral parts of a process that ultimately leads to estimates of ground motions at sites in the eastern United States. In this aspect, the EPRI program is unlike other seismic hazard methodologies that only implicitly rather than explicitly consider the scientific knowledge and uncertainties that are the basis for assessments regarding seismic sources and seismic hazards.

It is generally agreed that earthquakes are the result of a combination of stress and failure mechanisms. A logical division of the methodology, then, falls along these lines (Figure 1). First, the present tectonic stress regime is evaluated in the eastern United States. Second, the tectonic framework of the eastern United States is evaluated. This second step involves a consideration of the tectonic features that may be seismogenic, criteria for evaluating activity, and characterization of each tectonic feature in terms of its likelihood of being seismogenic.

Tectonic Stress Regime

A number of stress-generating mechanisms have been proposed to account for lithospheric stresses within intraplate tectonic environments. These processes range in scale from local influences affecting relatively small regions (hundreds of kilometers in dimension) to lithospheric scale processes affecting entire plates. In addition, the contribution to the magnitude of stress due to any single process can range from a few bars to kilobars.

An important aspect of the evaluation of tectonic processes for the EPRI program is the manifestation of each tectonic process in terms of observable characteristics. That is, in order to determine which tectonic processes are operative in a region, one must have a knowledge of the particular set of observable characteristics that serve to verify that a particular process is operative and to distinguish this process from all others. Once these characteristics have been established, they can be compared with the appropriate data in the eastern United States to allow interpretations of which tectonic processes are operative in the East, where they operate, and their magnitude.

The second component of the tectonic stress regime evaluation is the actual interpretation of the orientation and magnitude of stresses throughout the eastern United States. The first sources of data for this interpretation are in-situ stress indicators such as hydraulic fracture measurements, stress relief, well-bore breakouts, focal mechanisms, etc. Each indicator is subject to a variety of uncertainties in measuring the stress field. In addition, the relatively sparse and uneven data set leaves significantly large regions of the East without any direct observations. One approach to addressing this problem is through a consideration of tectonic processes. The scale over which an operative tectonic process is expected to occur may provide extrapolations of the stress regime into areas where few direct stress indicators are present. Some of the tectonic processes considered by the expert teams include "ridge-push" tectonic mechanisms, sediment loading and unloading, density changes at the continental margin, and glacial effects. Ultimately, the product of this component of the methodology is an interpretative map of stress orientations and a discussion of constraints on stress magnitudes.

Tectonic Framework

The methodology for assessing Tectonic Frameworks (Figure 1) consists of three principal steps: 1) identification of the tectonic features that are judged to be potentially active, 2) definition of physical characteristics or criteria that are most useful and diagnostic in assessing activity, and 3) assessment of the probability that each tectonic feature is active. Each of these steps involves a number of assessments.

The identification of tectonic features that may be potential sources of moderate-to-large earthquakes first involves compilation of several data sets including geologic structure maps, potential field maps (aeromagnetics, gravity, etc.), interpretive maps of crustal structure, etc. The goal in this effort is to identify significant regional tectonic features that may generate earthquakes of $m_b \geq 5$.

The next step in the methodology is the development and definition of criteria for assessing the probability of activity of tectonic features. These criteria must have two essential attributes: 1) they must have credibility as being scientifically diagnostic of seismic activity, and 2) they must be based on observable data that are generally available in the East. If a criterion is weak on either or both of these two grounds, then its utility in the assessment of activity is limited. An exhaustive list of the criteria given some level of consideration by the experts is given below:

Criteria for Assessing Activity

- Spatial association with seismicity
- Geometry and most recent sense of slip relative to present stress regime

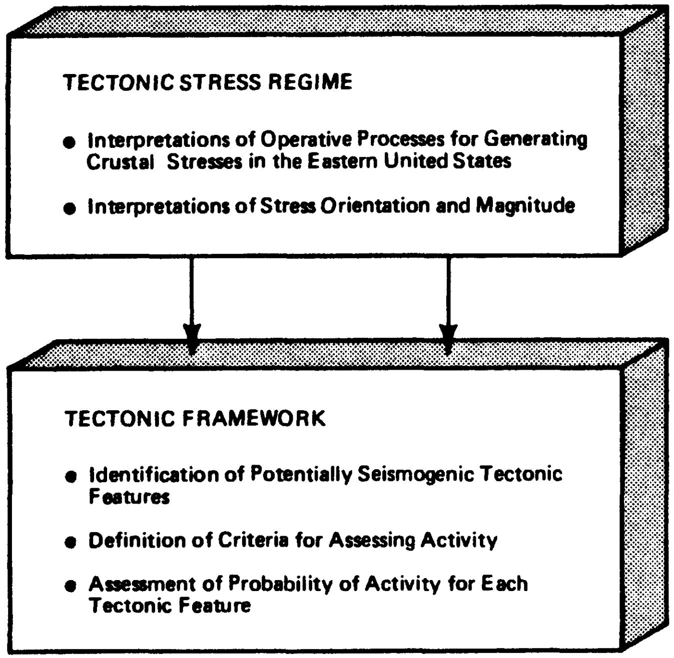


Figure 1. Key Components and Products of Tectonics Stress Regime and Tectonic Framework Assessments

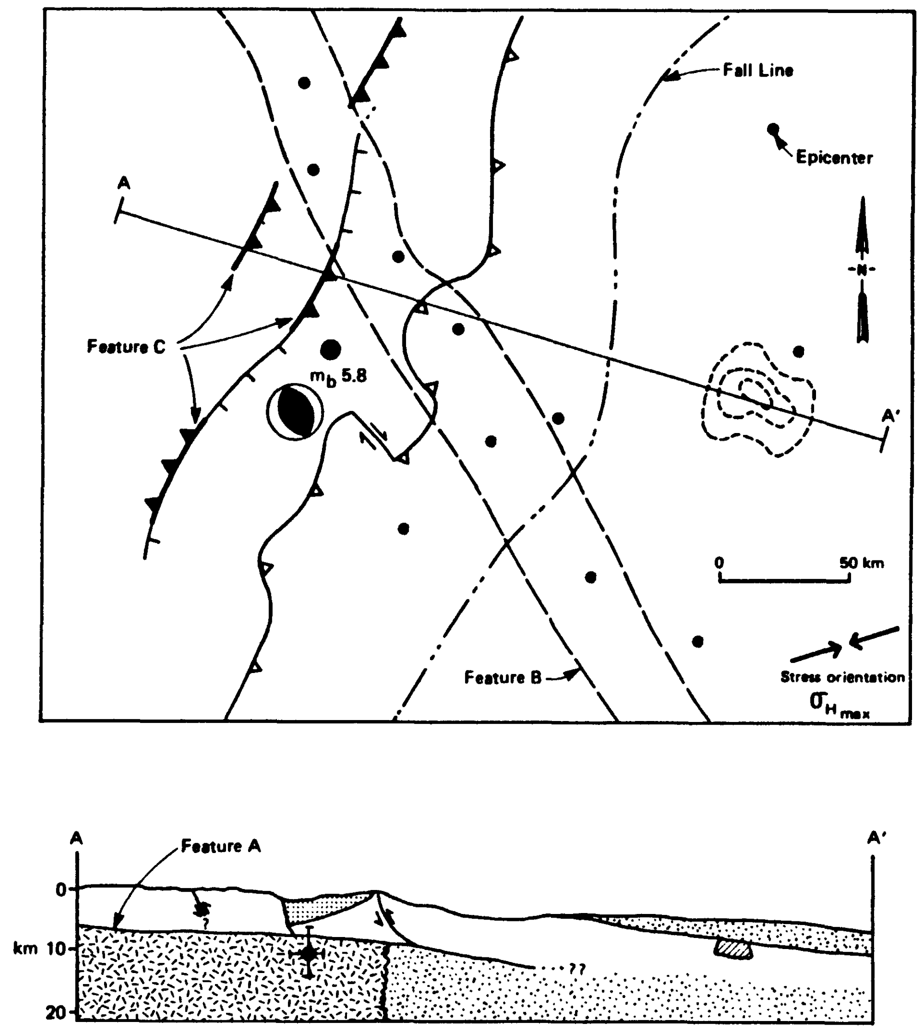


Figure 2. Hypothetical Map of Significant Tectonic Features and Earthquake Epicenters

- Crustal scale expression (expressed in the deep crust)
- Proximity to structural barriers or intersections
- Persistence of deformation through a significant period of geologic time
- Recent regional strain, including that recognized from geodetic measurements or from geomorphic indicators
- Feature-specific brittle slip in a mode consistent with the present stress regime
- Feature lies within a region of high seismic flux
- Evidence of local stress amplification
- Paleoseismicity (prehistoric earthquake evidence)

In practice, only about 3-4 criteria were judged by each expert team to be generally applicable to the full set of tectonic features in the EUS.

The next key step in this part of the methodology is an evaluation by each expert team of the discriminating power of each of their selected criteria. That is, an assessment is made of how diagnostic each criterion is in estimating the activity of a tectonic feature, in an absolute sense and relative to the other criteria. This assessment is made using the matrix of physical characteristics, discussed in the example below.

Once the particular criteria for assessing activity have been defined and evaluated, then an evaluation is made of the probability that each tectonic feature is active. Here, the available data and uncertainties pertaining to each particular tectonic feature are analyzed and documented in the feature characteristics assessments. The result of this analysis is a probability of activity for each tectonic feature in the eastern United States.

The final product of the Tectonic Framework assessment is a map of potentially active tectonic features, an evaluation of observational criteria for assessing activity, and an estimate of the probability of activity for each tectonic feature in the eastern United States. In modelling for SHA, the tectonic features evaluated as part of the tectonic framework serve as a basis for constructing seismic source zones. Background source zones are used to model the possible occurrence of seismicity away from known tectonic features. The probabilities of activity assigned to the features provide the basis for apportioning the seismicity among the sources for hazard analysis.

Example Tectonic Framework Assessment

To illustrate the procedures for developing a Tectonic Framework, an illustrative example is presented here.

Within a hypothetical region of interest, three tectonic features have been identified that are judged to have some potential to generate moderate-to-large earthquakes (Figure 2). Feature A is a prominent subhorizontal reflector in the deep seismic reflection profiles and is inferred to be a detachment related to Paleozoic compression. Feature B is a northwest-striking gravity gradient that extends for at least 400 km. Feature C is a group of high-angle reverse faults. Epicenters from historical and instrumental seismicity data are shown in Figure 2 including the location and focal mechanism of the m_b 5.8 earthquake. The maximum horizontal stress direction, inferred from focal mechanisms and in-situ stress measurements is also shown.

The criteria that were used to assess the activity of the tectonic features are shown in Figure 3. The matrix of physical characteristics was completed based on subjective judgments of the probability of a hypothetical feature being active given the conditions of each cell of the matrix. Note that to assess features such as Feature B, which is buried and known only from geophysical data, a second matrix was constructed that does not include observations of brittle slip. The matrix assessment reflects the scientific uncertainty regarding which criteria are most diagnostic of activity. Next, each feature was assessed relative to the criteria that have been identified (Figure 4). This assessment reflects the informational uncertainty regarding each feature.

Spatial Association with Seismicity; Focal Mechanism Consistent / Geometry Relative to Stress/Sense of Slip / Brittle Slip on a Feature (Most recent age)	Moderate-to-Large Earthquakes		Small Earthquakes Only		No Seismicity	
	Favorable	Unfavorable	Favorable	Unfavorable	Favorable	Unfavorable
Pleistocene Holocene Slip	1.0	0.9	0.8	0.7	0.5	0.4
Cretaceous Tertiary Slip	0.7	0.5	0.5	0.3	0.2	0.1
Pre Cretaceous Slip or No Brittle Slip	0.4	0.3	0.1	0.05	0	0

Spatial Association with Seismicity; Focal Mechanism Consistent / Geometry Relative to Stress/Sense of Slip	Moderate-to-Large Earthquakes	Small Earthquakes Only	No Seismicity
Favorable	1.0	0.5	0.1
Unfavorable	0.8	0.3	0

Figure 3 . Matrix of Physical Characteristics for Example Evaluation

Physical Characteristic	Probability that the Given Feature Exhibits a Given Level of Each Characteristic: Feature #A	Feature #B	Feature #C
1. **Spatial Association with Seismicity; Focal Mechanisms Consistent with Feature Geometry**			
1. Moderate-to-Large Earthquakes	0.3	0.1	0.2
2. Small Earthquakes Only	0.2	0.5	0.1
3. No Seismicity	0.5	0.4	0.7
	Σ=1.0	1.0	1.0
2. **Geometry of Feature Relative to Stress Orientation and/or Sense of Slip**			
1. Favorable Geometry/Sense of Slip	0.3	0.3	0.1
2. Unfavorable Geometry/Sense of Slip	0.7	0.7	0.9
	Σ=1.0	1.0	1.0
3. **Brittle Slip on a Feature**			
1. Pleistocene-Holocene Slip	0		0.1
2. Cretaceous-Tertiary Slip	0.1		0.9
3. Pre-Cretaceous Slip or No Brittle Slip	0.9		0
	Σ=1.0		1.0

Figure 4. Example Feature Characteristics Assessment

The combination of the matrix assessment and the feature assessment results in a calculation of the probability of activity of each feature, which for Features A, B, and C is 0.13, 0.28, and 0.25, respectively.

The details of the EPRI methodology for developing seismic sources and specifying their seismicity parameters is not given here.

Future Directions for SHA in the EUS

On the basis of the recent increasing attempts to incorporate tectonic data into EUS hazard assessments, I here offer my opinion of the future directions that I expect SHA to follow in the next few years.

1. Further testing and verification of tectonic processes responsible for crustal stress: By establishing the operative tectonic processes in the EUS, the stress regime can be further characterized by its orientation and magnitude. Better knowledge of crustal stresses can be important to evaluating the activity of tectonic features due to reactivation processes and to understanding the constancy of the crustal loading process.

2. Increasing geologic/tectonic bases for identifying seismic sources: Continued seismological and geologic/geophysical studies will help to clarify the causal association of seismicity with structure, at least locally in well-studied areas. Local geologic studies will likely identify paleoseismic evidence of Quaternary deformation, such as that along the Meers fault in Oklahoma. In order to generally identify the tectonic structures that are seismic sources, diagnostic criteria for assessing whether or not particular structures are seismogenic will continue to be refined.

3. Compilation and evaluation of data from historical earthquakes within intraplate regions: Global studies of regions analogous to the EUS will be made to validate or contradict recent comparisons between interplate and intraplate regions in terms of source scaling

relations, maximum earthquakes, source properties, stationarity, etc. These studies may identify unique physical properties related to the intraplate earthquake generation process or they may show similarities with earthquake processes in more active interplate regions. An increased, but still limited, use of physical properties of known structures will be used to constrain models of future earthquake occurrences (e.g., rupture location, maximum magnitude, etc.).

4. Evaluation/resolution of EUS paradox: Paleoseismic studies and historical seismicity data suggest relatively high rates of recurrence of large earthquakes (few thousand years) at locations such as Charleston and New Madrid. However, the geologic data regarding Quaternary deformation effectively preclude maintaining these rates for long periods of time (few million years). Further studies will provide insights into this problem, which has importance to understanding the temporal nature of intraplate seismicity. Periodic models proposed for plate boundary environments, whereby earthquakes repeat regularly according to an average recurrence interval, may not be appropriate for intraplate environments. Instead, more episodic behavior may be typical whereby a seismic source goes into an active state for several seismic cycles and then "turns off" for long periods of time. This type of behavior would explain the EUS paradox. Any model of this type has important implications to SHA forecasts, which are usually made for a short time period of interest (say 50 years).

5. Quasi-predictive tools for hazard assessments without detailed source characterization: In the future, geologic and geodetic data will continue to identify locations where Quaternary or more recent deformation is taking place. Examples are the localized strain anomalies in the lower crust observed in the northern New Jersey - southeastern New York area from triangulation data, and the rapid subsidence occurring in the Passamaguoddy Bay region of Maine observed from both geodetic and geologic data. Further definition of deforming areas in terms of the extent and rates of deformation and associations with

seismicity will likely lead to characterizing them as potential seismic sources for SHA, regardless of whether the exact seismogenic structure can be determined.

6. Further development of probabilistic tools to incorporate uncertainties in intraplate tectonics: It is unlikely that in the next few years the key problems related to EUS seismotectonics will be resolved. Therefore, acknowledged probabilistic tools will be adapted to allow uncertainties to be explicitly incorporated into hazard analyses. As these tools become more widely used and accepted by earth scientists, increased interest will develop to further specify the intraplate earthquake process including basic uncertainties such as stress-generating mechanisms and likely locations of future seismicity.

ADVANCES IN SEISMIC SOURCE DEFINITION FOR SEISMIC HAZARD ANALYSES IN THE EASTERN UNITED STATES

by

Robert R. Youngs
Geomatrix Consultants
One Market Plaza, Spear St. Tower, Suite 717
San Francisco, California 94105

INTRODUCTION

In terms of probabilistic seismic hazard analyses, a seismic source represents a region of the Earth's crust in which future seismicity is assumed to follow specified probability distributions for occurrence in time, earthquake size, and location in space. In seismically active areas, such as the western United States, seismic sources are defined primarily on the basis of identified faults or fault zones. However, in the eastern United States there is a general lack of knowledge of which geologic structures have generated past earthquakes and a lack of understanding of the causal mechanisms for earthquakes. Consequently, seismic sources have been defined primarily on the basis of the spatial distribution of historical and instrumental seismicity. The inability to relate observed seismicity to specific geologic structures leads naturally to uncertainty in specifying the boundaries of seismic sources drawn around spatial clusters of seismicity and several seismic hazard analyses conducted for sites in the eastern United States have explicitly incorporated this uncertainty in the analysis generally through the use of alternate source zonation maps.

The Electric Power Research Institute (EPRI) is currently sponsoring a study to assess probabilistically the seismic hazard at nuclear power plant sites in the eastern United States. In that study a major emphasis has been placed on incorporating tectonic information into the process of defining seismic sources for seismic hazard analysis. In a companion paper, Coppersmith (this volume) describes the methodology developed for the EPRI study to evaluate the available tectonic information with an aim toward identifying potentially seismogenic tectonic features. The purpose of this paper is to describe the methods used to define seismic sources for the eastern United States on the basis of tectonic features and the methods used to specify the interdependencies between sources in developing a complete representation of potential future seismicity for seismic hazard analysis.

DEFINITION OF INDIVIDUAL SEISMIC SOURCES

The product of the tectonic evaluations consists of a set of tectonic features that may be capable of generating future moderate-to-large earthquakes (defined as $m_b \geq 5$) and for each feature an assessment of the probability that it is active. Activity is considered to be a binary attribute, a feature either can generate moderate-to-large earthquakes or it cannot. The assessed probability that a feature is in an active state is a marginal probability as it is evaluated without consideration of what other features might be active in a region.

The identified tectonic features form the primary basis for defining seismic sources. Definition of each seismic source consists of two parts, specification of the geometry of the source and specification of the probability that the source is active. The types of seismic sources developed in the EPRI study and the guidelines developed for specifying their probability of activity are described below.

Feature-Specific Seismic Sources

The majority of seismic sources are drawn to directly represent individual tectonic features and hence are termed feature-specific sources. An individual feature-specific source can represent:

(1) A single tectonic feature The source geometry is represented by either a line source or an area source encompassing the feature. The probability that the source is active is equal to the probability that the feature represented by the source is active;

(2) A class of features A class of features represents a set of similar geologic structures whose seismogenic potential is assessed as a class using all of the characteristics that could be attributed to the individual members of the class. The source geometry is represented by an area source encompassing all of the features or by individual area or line sources for each feature. The probability that the source is active is equal to the probability that the class of features is active; and

(3) A group of features Due to the regional nature of the EPRI study, it is possible to treat several small features located in close proximity or overlapping as a single source when they are located at a significant distance from the sites of interest. The source geometry is represented by an area source encompassing the features. In this case, the probability that the source is active is equal to the probability that at least one of the individual features is active. The procedures used to specify the joint probability of activity of several features are presented below.

Default Seismic Sources

Regions in which a moderate-to-large earthquake has occurred historically are usually interpreted to contain at least one active feature with certainty (i.e. there is zero probability that there are no active features in the region). Usually in such regions one or more potentially active features will have been identified, but none of the features will have been judged to be active with certainty. Thus there may exist other, unknown features in the region. A default area source is used to represent the geographical limits within which the unknown feature is judged to lie. The probability that the default source is active is equal to the probability that none of the identified features are active.

Background Sources

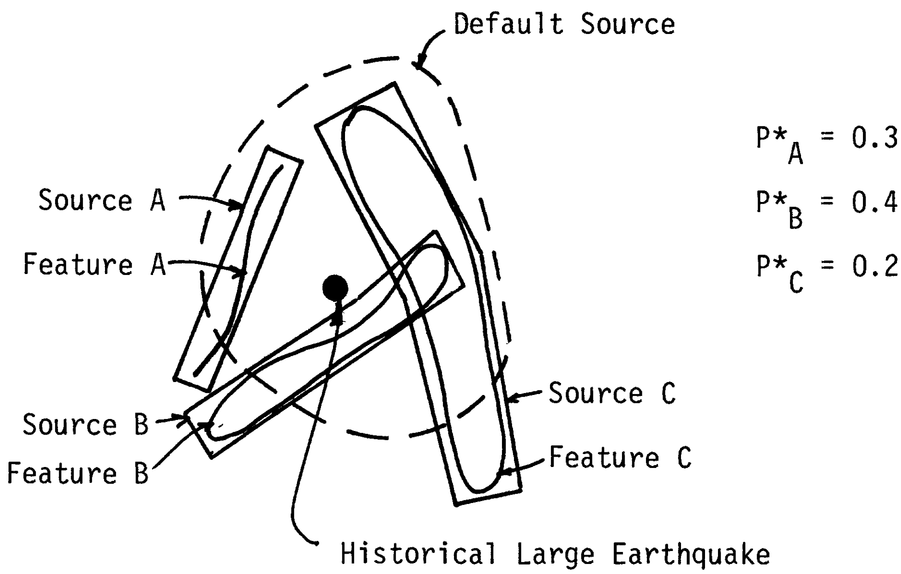
A background zone represents a region in which no distinguishable tectonic features or patterns of seismicity have been identified, but it is judged that there is some potential for moderate-to-large earthquakes to occur. The probability that the background source is active is evaluated in terms of the probability that, at any randomly chosen point in the background, a moderate-to-large earthquake could occur. This interpretation involves a "spatial average" of areas that are active and areas that are inactive within the background.

Default sources and background sources are used to address unknown sources of seismicity. With default sources, the focus is on unknown causes of known seismic activity in a relatively small area; with background sources the focus is on unknown but possible seismic activity at unknown locations in a broad geographical region. Background sources can also serve as default sources when there is no basis for restricting the location on an unknown source of known activity to a specific region surrounding a the location of a past moderate-to-large earthquake.

SPECIFICATION OF INTERDEPENDENCIES OF SOURCE ACTIVITY

The probability of activity of the individual feature-specific sources described above represent marginal probabilities without consideration of the possible dependencies between sources in a region. However, probabilistic seismic hazard computations utilize the joint probability of occurrence of earthquakes on all sources that could affect a site. As the probability that a source is active is an integral part of the probability of the source producing ground motions at a site, the joint probabilities of source activity are required.

The approaches used for the specification of the joint probability of source activity are illustrated by the hypothetical example in Figure 1. Shown are three tectonic features located in an area in which a large earthquake has occurred. The marginal probabilities that the features



Possible States of Joint Source Activity

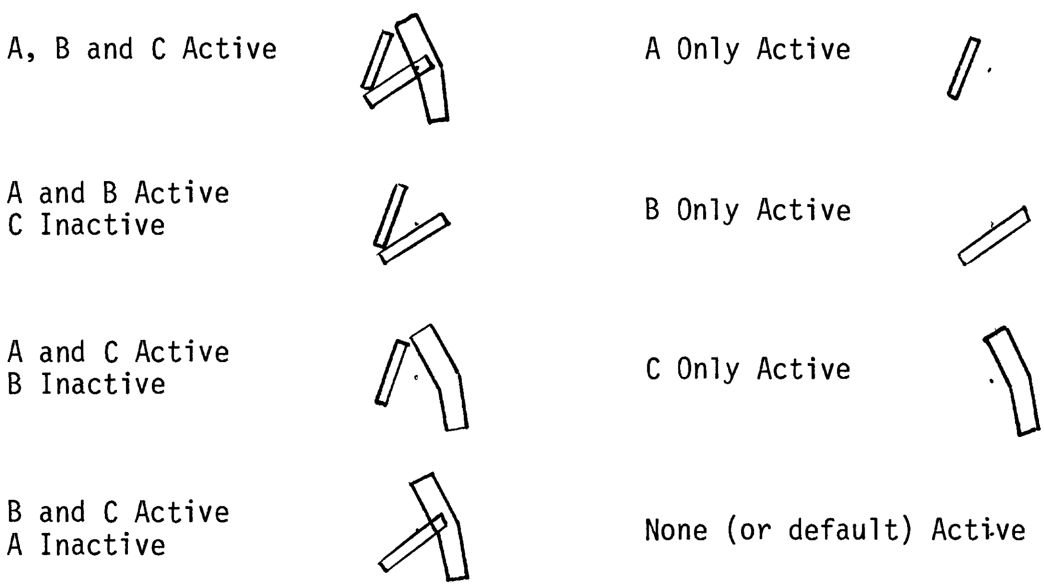


Figure 1 Specification of Joint Source Activities

1284 40008

are active, P^*, are assessed using the methodology described in Coppersmith (this volume). These tectonic features are the only ones identified in the vicinity of the large earthquake. For seismic hazard computations, we represent each of the features by a feature-specific source, as shown, with P^*_A=0.3, P^*_B=0.4, and P^*_C=0.2.

As indicated in Figure 1, there are eight possible states of joint activity for the three sources, including the state of none of the three being active. The fact that the probabilities of activity for the identified sources are less than 1.0 may lead to a finite probability that there are no active sources. This result may be an unacceptable outcome in this example as at least one large earthquake is known to have occurred in the region.

One possible interpretation is that some other, yet undiscovered feature in the region may exist and may have been the source of the past large event. The location of this feature is unknown but it presumably lies in the vicinity of the epicenter of the large earthquake. A default source encompassing possible locations for the unknown feature could be drawn, as shown in Figure 1. The default source would be considered active only if the three feature-specific sources are all not active, that is P(default active) = P(A and B and C not active).

The marginal probabilities of activity of the individual sources in a region do not provide enough information to specify the probabilities of the joint states of activity of all of the sources, requiring additional evaluations by the analyst. The possible interpretations fall into three general classes, discussed below.

Independent Sources

One possible interpretation is that the activity state of each identified source is independent of the assessment for other sources. Given independence, the probabilities of the various states of joint source activity are obtained by multiplying the marginal probabilities of activity for the individual sources or their complements. For the example in Figure, 1 these are:

$$P(\text{A and B and C active}) = P^*_A \cdot P^*_B \cdot P^*_C = 0.024$$

$$P(\text{A active only}) = P^*_A \cdot (1-P^*_B) \cdot (1-P^*_C) = 0.056$$

$$P(\text{B active only}) = (1-P^*_A) \cdot P^*_B \cdot (1-P^*_C) = 0.036$$

$$P(\text{C active only}) = (1-P^*_A) \cdot (1-P^*_B) \cdot P^*_C = 0.096$$

$$P(\text{A and B active only}) = P^*_A \cdot P^*_B \cdot (1-P^*_C) = 0.084$$

$$P(\text{A and C active only}) = P^*_A \cdot (1-P^*_B) \cdot P^*_C = 0.144$$

$$P(\text{B and C active only}) = (1-P^*_A) \cdot P^*_B \cdot P^*_C = 0.224$$

$$P(\text{none active}) = (1-P^*_A) \cdot (1-P^*_B) \cdot (1-P^*_C) = 0.336$$

If a default source is considered to be active when neither A or B is active, then:

$$P(\text{default active}) = (1-P^*_A) \cdot (1-P^*_B) \cdot (1-P^*_C) = 0.336$$

<u>Mutually Exclusive Sources</u>

Another possible interpretation is that one, and only one, source in a region is active. In this case, if one of the identified feature-specific sources is active, then there are no other active sources. Under the assumption that the states of activity of the sources are mutually exclusive the probabilities of the states of joint activity for the example in Figure 1 become:

$$P(\text{A active only}) = P^*_A = 0.3$$

$$P(\text{B active only}) = P^*_B = 0.4$$

$$P(\text{C active only}) = P^*_C = 0.2$$

$$P(\text{none active}) = 1-P^*_A-P^*_B-P^*_C = 0.1$$

If a default source is considered to be active when neither A nor B is active, then:

$$P(\text{default active}) = 1-P^*_A-P^*_B-P^*_C = 0.1$$

1255 40008

It is important to note that if the sum of the marginal assessments of source activity is greater than one ($\Sigma P^{*} > 1$), then the interpretation of mutually exclusive sources is inconsistent with the individual marginal assessments as it leads to a negative probability that none of the identified sources are active.

Dependent Sources

The third possible interpretation is that the activity state of the individual source is dependent in some manner that must be specified by the analyst. Given N sources in a region, there are 2^N possible states of joint activity. The marginal assessments of the probability of activity for the individual sources provide N constrains and the assumption that the 2^N states are a mutually exclusive and collectively exhaustive set provides one additional constraint. This leaves 2^N-N-1 probabilities that must be specified by the analyst. For the example in Figure 1, four additional constraints must be provided.

One possibility is that Source C is mutually exclusive with the other sources and that given C is inactive, A and B are independent. This leads to the following probabilities of joint source activity:

P(C active only)	$= P^{*}_{C}$	= 0.2
P(A and B active only)	$= P^{*'}_{A} \cdot P^{*'}_{B} \cdot (1-P^{*}_{C})$	= 0.15
P(A active only)	$= P^{*'}_{A} \cdot (1-P^{*'}_{B}) \cdot (1-P^{*}_{C})$	= 0.15
P(B active only)	$= (1-P^{*'}_{A}) \cdot P^{*'}_{B} \cdot (1-P^{*}_{C})$	= 0.25
P(none active)	$= (1-P^{*'}_{A}) \cdot (1-P^{*'}_{B}) \cdot (1-P^{*}_{C})$	= 0.25

where $P^{*'}_{A}$ and $P^{*'}_{B}$ are conditional probabilities of activity for the two sources [$P^{*'}_{A}$ = P(A active| C inactive) = $P^{*}_{A}/(1-P^{*}_{C})$].

If a default source is considered to be active when none of the identified sources are active, then:

1256 40008

$$P(\text{default active}) = (1-P^{*'}_A)\cdot(1-P^{*'}_B)\cdot(1-P^{*}_C) = 0.25$$

Again it should be noted that the analyst is not completely free to select any value for the various probabilities of joint activity. For example, an assumption that the probability of all three sources being active is equal to 0.4 would be inconsistent with the marginal assessment that the probability of activity for Sources A and C as they each have marginal probabilities of activity less than 0.4.

The general rules for specifying joint probabilities of source activity applied in the study can be summarized as follows:

Option 1 - Assume Sources Independent

- Default source required if area considered active and no source has $P^* = 1$ with $P^*_{default} = \pi(1-P^*_i)$

Option 2 - Assume Sources Mutually Exclusive

- Inconsistent if $\Sigma P^*_i > 1.0$
- Default source required if area considered active and $\Sigma P^*_i < 1.0$ with $P^*_{default} = 1 - \Sigma P^*_i$

Option 3 - Assume Sources Dependent with Specified Joint Probabilities

- Default source required if area considered active and $\Sigma P^*_i < 1.0$
- For N sources requires 2^N-N-1 additional assessments beyond marginal assessments for individual sources

The procedures used for specifying the seismicity parameters for the sources and for conducting the seismic hazard analyses are described in companion papers in this volume.

DETERMINATION OF EARTHQUAKE SIZE

by

D. Burton Slemmons and Craig M. dePolo
Mackay School of Mines
University of Nevada
Reno, Nevada 89557

INTRODUCTION

One of the most crucial elements of a seismic hazard study is the determination of earthquake size that could provide strong ground motion at a site. Geologists, seismologists, and engineers have strived to understand the size, distribution and character of seismic sources. One of the most commonly used scaling parameters for earthquakes is the magnitude scale (Coppersmith, 1982; Slemmons, 1982; Schwartz and others, 1984; Smith, 1976). Although other source parameters have been considered to be significant to seismic studies, most studies and relationships generally include an estimation of magnitude.

The methods that are commonly used for determining earthquake magnitude are discussed for active faulting, with emphasis on uncertainties.

EARTHQUAKE SIZE (MAGNITUDE AND INTENSITY)

The quantification or size scaling of earthquakes is a complex relation. Many scales are used for both "magnitude", a quantity determined from seismographic records to indicate the strength of earthquake kinetic energy release, and "intensity", a measure of the effects of an earthquake on man, on structures built by him, and on the earth's surface.

Common magnitude scales include the well-known and widely used Richter or local magnitude (ML), surface wave magnitude (Ms), body wave magnitude (Mb), and moment-magnitude (Mw), as noted by Kanamori in 1983 (Figure 1). Surface wave magnitudes are utilized in this report for design or maximum earthquake

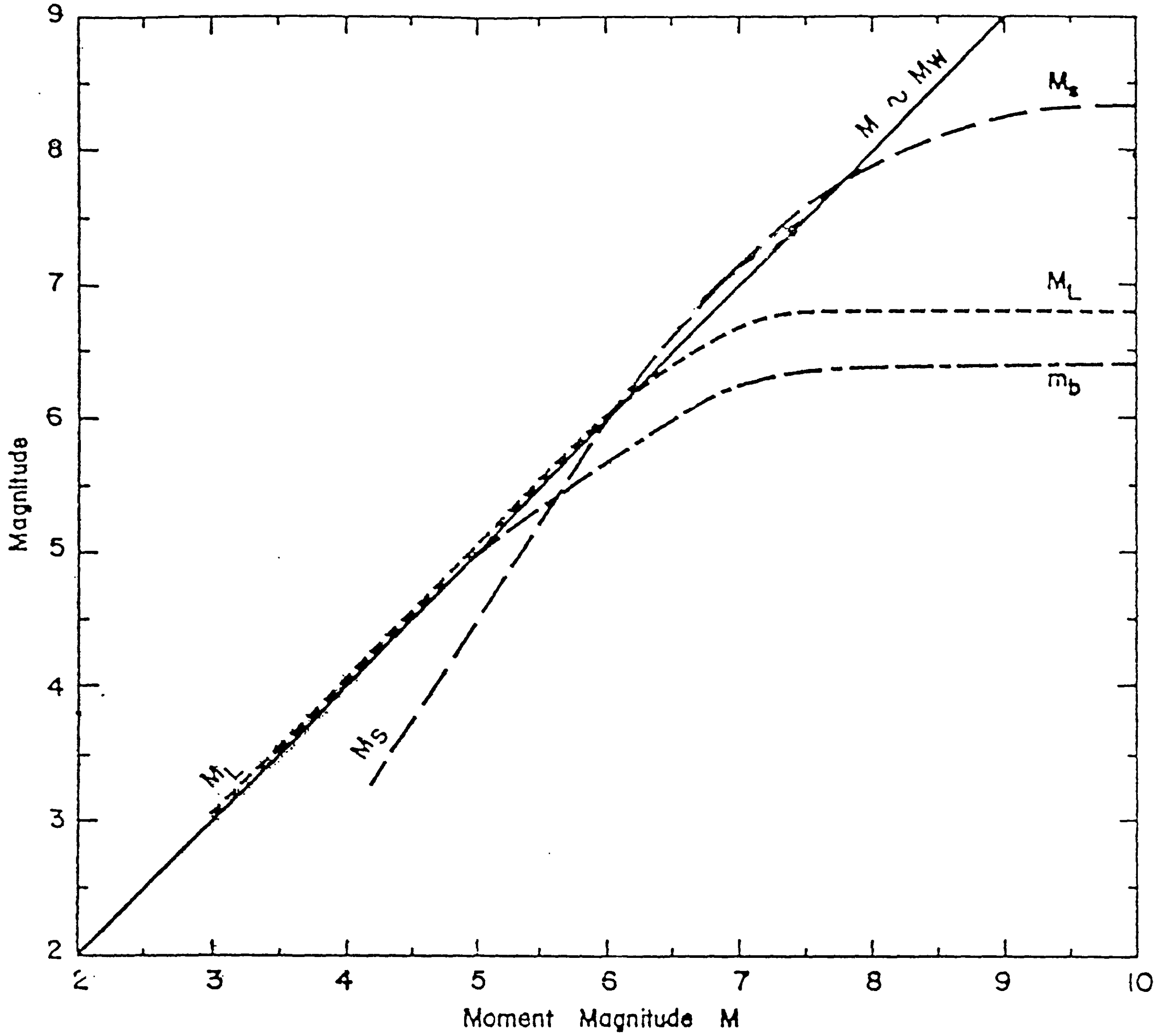


Figure 1. Lines showing relationship between magnitude and moment magnitude (modified from Kanamori, 1983). Lines show saturation for mb, body wave magnitude, ML, local or Richter magnitude, and Ms. surface wave magnitude.

1257 40008

determinations of non-subduction zones. In subduction zones, moment magnitude (Hanks and Kanamori, 1979) is more appropriate for great earthquakes.

Intensity determinations and mapping of isoseismal lines for areas of the same earthquake intensity are especially used for early historical earthquakes, activity estimates and compilations for interplate regions, and for detailed effects within a local area or region. With good macroseismic data, the intensity values show a good correlation with magnitude, but for sparse data, the magnitudes determined are generally too low, and the time frequency graphs provide a b-value that also is too low. The use of historical earthquake data may lead to results that are in disagreement with geological evidence (Schwartz and Coppersmith, 1984; Youngs and Coppersmith, 1985). This discrepancy between seismological and geological data leads to questions about the applicability of the b-value technique. An example for an intraplate region is the Meers fault zone in Oklahoma, which is historically aseismic but has strong geologic and geomorphic evidence of recurrent Holocene and late Quaternary activity. This fault has the potential for a magnitude 6-1/2 to 7-1/2 earthquake (Slemmons, Ramelli and Brocoum, 1985).

ESTIMATING EARTHQUAKE MAGNITUDE

In order to examine the estimation procedure, and consider ways for a scientist to estimate the uncertainties and understand the data they use in the estimation process, we have tried to categorize the process into three categories: approaches, methods, and techniques (Table 1). The table serves as a reminder so that the most magnitude estimates of a fault can be generated. This is known as the "multi" approach in general and is often a powerful aid in decisions.

The approaches listed in this table are actually second order approaches, the first order approach decided earlier on is whether the analysis is going to be a deterministic approach or a probabilistic approach. Parts of both of these first order approaches then trickle into the terms listed in the table. The second order approaches involve looking at different data or aspects of a fault. The different approaches have relative qualities which can be considered when rationalizing the different magnitude estimations. For

Table 1 APPROACHES, METHODS, AND TECHNIQUES THAT CAN BE USED AS GUIDES IN ESTIMATING EARTHQUAKE MAGNITUDES

APPROACHES	METHODS	TECHNIQUES
Historical Seismicity Approach - involves the analysis of historical earthquake data	Historical Seismicity Method	historical earthquake data
	fault seismicity	magnitudes
	regional seismicity	temporal relationships
		spacial relationships
	Fault Length Method	magnitude/freq. eqns.
	segmentation	moment tensor
Paleoseismic Approach - involves using parameters from delineated pre-historic surface ruptures	fractional fault length	seismic slip
	half length	
	total fault length	fault length correlations
	total fault system length	fract. fault length correl.
	Fault Displacement Method	total fault length correl.
Geological Approach - involves using geological information about fault	maximum displacement	total fault system correl.
	average displacement	
	modal displacement	fault displacement correl.
	Slip Rate Method	length X displ. correl.
Geophysical Approach - involves geophysical data and theory (source modeling)	long term (geologic)	slip rate correlations
	recent (paleoseismic)	
	active (historical & geodetic)	
	Fault Area Method	fault area correlations
	Moment Magnitude Method	moment magnitude calculations
Relative Approach - compare other similar faults which are well studied or have historical events	Relative Comparison Method	geophysical data & theory
	to another fault	heat flow models
	to another earthquake	geodetic measurements
	to another region	
	to another tectonic regime	tectonic models

1287 4008

example, the Historical Seismicity Approach and the Paleoseismicity Approach both examine direct evidence of earthquakes, which is data which can often be somewhat confidently used as evidence to constrain estimates.

"Methods" are methods a scientist uses to analyze a particular fault or area, and reduce geologic and geophysical information into parameters which can be used comparatively as a guide to magnitude estimation. These are questions such as "how long is the fault?", or "is the fault segmented?", or "what is the slip rate of the fault?".

"Techniques" are the actual correlations, calculations, models, and data used to convert the fault or areal parameters into a value or number which can be used as a guide in estimating magnitude. A scientist reviewing the values arrived at using the various methods and techniques should know the quality of the data, quality of the technique, and applicability of the method and techniques. Although there are many of these guides, only a few are usually applicable to a specific fault due to a lack of data or understanding of the active nature of the fault.

A decision then needs to be made as to a final value. This decision varies of course with the different types of earthquake magnitudes being estimated.

When a final value has been arrived at for a fault or area, the value can be used in the overall analysis process for the design earthquake (Figure 2).

APPROACHES AND METHODS FOR INTRAPLATE REGIONS

Intraplate earthquakes are poorly understood with respect to interplate earthquakes. Recently, earthquakes of these regions have been recognized to be very important because of the high population and industry concentration, low attenuation of ground motion, recognition of active faults, and occurrence of high magnitude paleoseismic or early historical earthquakes. This suggests that evaluation of intraplate areas should combine traditional "province or intraplate type", and typical interplate active fault assessments.

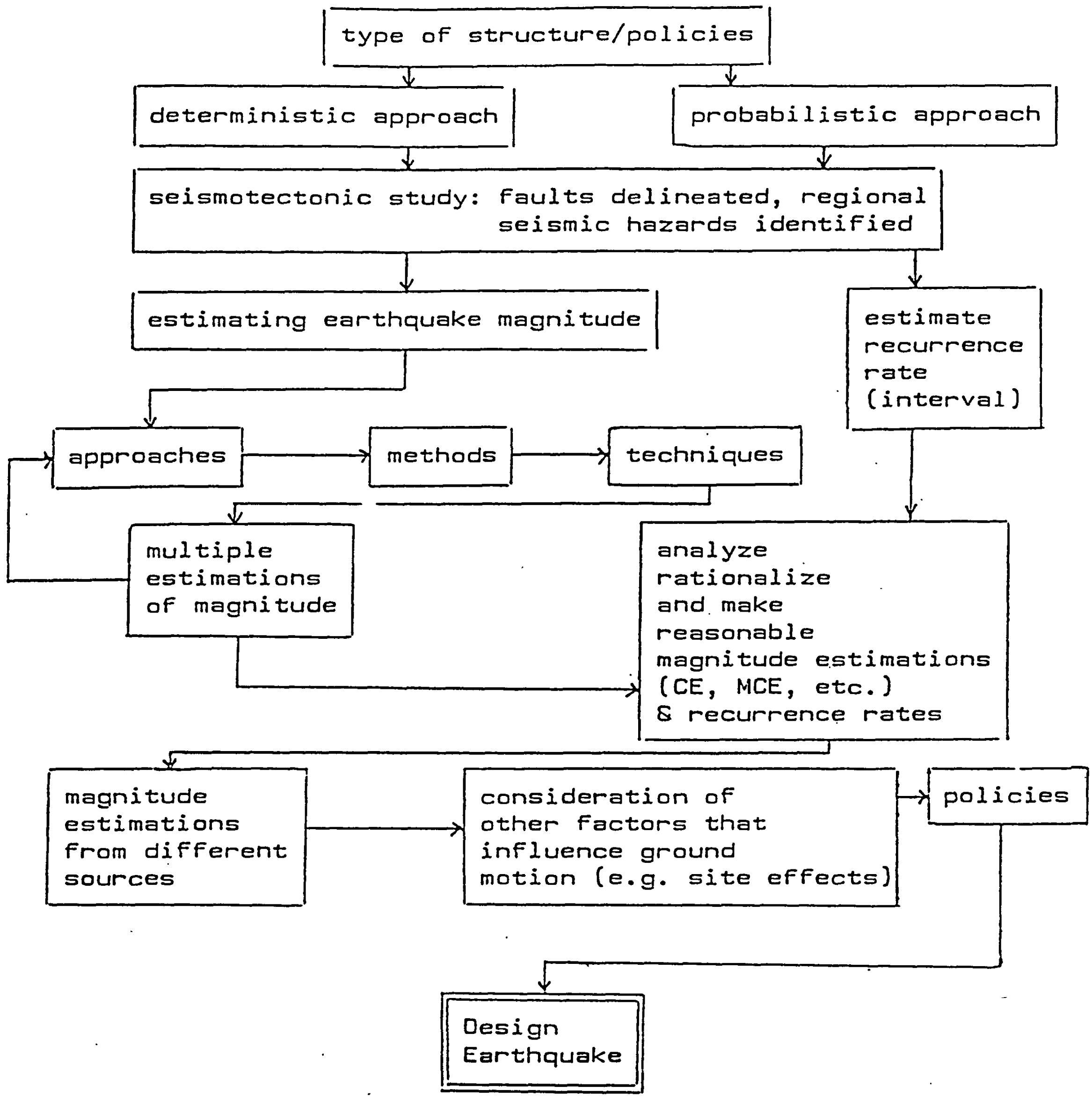


Figure 2. Flow chart of a progression of steps toward assignment the design earthquake.

1252 40008

Typical evaluations have used historical seismicity for provinces or regions, with seismological determination of earthquake recurrence or frequency to magnitude relations (Gutenberg and Richter, 1954). The determination of provinces may be difficult, since the active tectonics concept is different than traditional geologic structural provinces and the rates of activity are very low. For provinces or faults where the activity may be related to a seismic cycle, the brief window of historical activity may not give a reasonable indication of future activity. Maximum earthquakes appear to be between 5.2 to 6.0 magnitude, although longer intervals of time for recording, or higher historical magnitude values of up to about 6.25 magnitude for earthquakes in the oceanic crustal materials may indicate the possibility of over 6 magnitude earthquakes.

The evaluation of these regions commonly is made by statistical analyses that predict from b-values the likelihood of an earthquake within a homogeneous region or province. This approach may assume a random location or "floating" earthquake source.

Within recent years a growing number of evaluations of historical or paleoseismic activity in eastern and central United States. Three areas may be local, or perhaps partly interconnected zones of active faulting and higher area of 1811 and 1812 (Russ, 1979; Zoback, 1979), the Charleston, South Carolina earthquake area of 1886 (Obermeir and others, 1985; Talwani, 1985), the Meers fault zone in Oklahoma (Slemmons and others, 1985 and abstracts presented at the Meers fault symposium, Seismological Society of America in 1985), and the Kentucky Valley fault in Kentucky (VanArsdale, 1985). This growing number of exceptions to a simple province approach indicates a need for more complex evaluations of several regions in central and eastern United States. It also suggests a possibility that there may be more widespread break-up of into several other isolated or possibly interconnected faults that extend into this interplate region. Some of these zones are listed in Slemmons and dePolo (in press).

APPROACHES AND METHODS FOR INTERPLATE REGIONS

Historical Seismicity Approach

The maximum historical earthquake method assumes that the largest historical earthquake that has occurred along a fault or in a region is the maximum or maximum credible earthquake. This method provides reasonable values for about ten percent of the examples where fault slip rates are very high, or recurrence intervals are very short, for very long fault zones with similar maximum earthquakes along their length, for regions with long historical records, or for fortuitous cases where the observational record happens to include a maximum event. Figure 3 provides an indication of whether the maximum historical earthquake has occurred if the fault slip rate and average recurrence interval are known. In California, for example, the Ms = 8.3 that has occurred twice along the San Andreas fault zone may be near the maximum for those segments of the fault zone.

Paleoseismic Approach

Paleoseismic methods are used for all types of faults and active tectonic regions of the world to infer the size of prehistorical earthquakes by using evidence of past events that are recorded along a fault from geomorphic, soil-stratigraphic, and stratigraphic evidence. Many active faults have recurrent activity (Wallace, 1970; Slemmons, 1977) and develop fresh or youthful geomorphic features (Slemmons, 1977 and 1981; various papers in Morisawa and Hack (eds.) in 1985; Wallace, 1977; Wallace and Whitney, 1984). The first important tool for inferring the size of prehistorical earthquakes was from the regression analysis of Tocher (1958) to estimate earthquake magnitude from fault rupture length, maximum displacement, or both. More recent analyses of magnitude determinations (Lienkaemper, 1984) and fault regression analysies by Bonilla and others (1984), Slemmons (1982), and Slemmons and Chung (1982) permit a quantitative assessment by use of measured fault parameters. The commonly used regressions include correlations between Ms, fault surface rupture length, and maximum displacement.

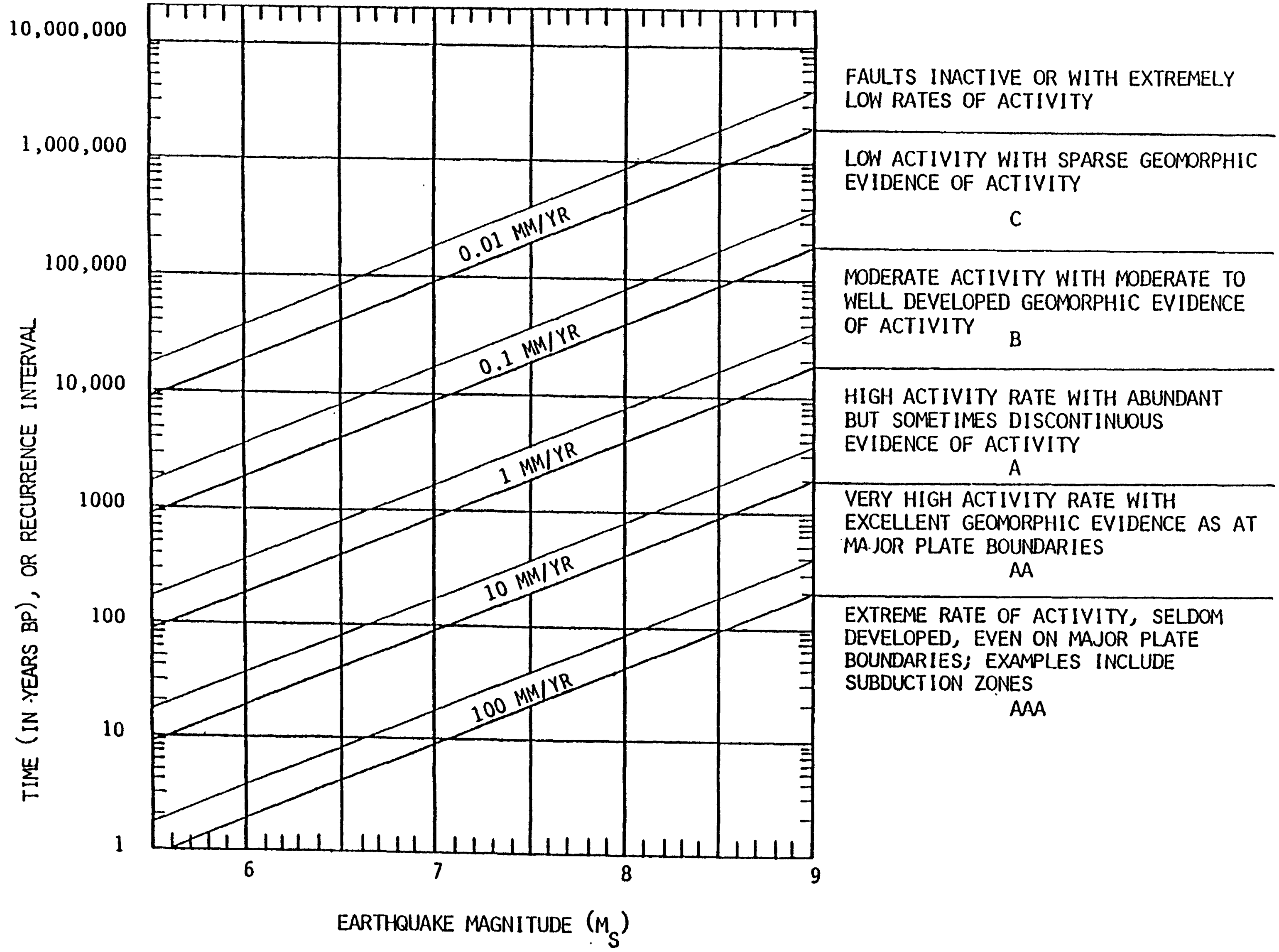


Figure 3. Relation between time or recurrence interval between earthquakes, earthquake magnitude, and slip rate across the fault zone. This chart assumes that most of the energy is released by seismogenic rather than aseismic activity and that the average displacement is one half the maximum.

1388 4008

In addition to evaluating the size of the maximum earthquake, it is important to also evaluate the recurrence or freqency of earthquakes. Some of these methods are outlined in Coppersmith (1980), Schwartz and Coppersmith (1984 and in press).

Segmentation Method

The segmentation method involves the identification of individual fault segments that have a continuity, character, and orientation that suggests the segment will rupture as a unit (Slemmons, 1982). Individual fault segments have different characteristics relative to adjacent segments or are separated from adjacent segments by identifiable discontinuities.

Use of the segmentation approach was suggested by Swan et al (1980) for the six to ten segments for the 370 km long Wasatch fault zone in Utah. These segments were approximately 40 to 60+ km in length and were believed to be capable of generating earthquakes in the 6-3/4 to 7-1/2 Ms magnitude range. More recently, Schwartz and Coppersmith (in press) and Slemmons and dePolo (in press) noted that scarp morphology and fault geometry, location of transverse structural trends, cross-fault and gravity data suggested that there are six segments along this zone. The segments are about 30 to 70 km in length and are separated by transition zones of from a few to more than ten km in length. Slemmons (1980) noted that patterns of offshore faults and seismic reflection profiles along the offshore zone of deformation in the San Diego-Long Beach area also suggests a similar segmentation of fault geometry and transition zone development. These and other studies suggest a possible full realistic values for magnitude.

In other cases the segmentation of fault zones is more difficult and a reasonable case for segmentation may not be able to be made. For example, the multiple section failure accompanying the 1915 Pleasant Valley earthquake, would be difficult to predict without conducting extensive and sophisticated geological and geophysical studies.

Although multiple earthquakes appear to occasionally be capable of activating more than one segment, the use of 100 percent rupture length for fault segments appears to be reasonable for most cases and lead to realistic values for magnitude. The use of one-half rupture lengths for segments, or of very short fault lengths of a few km to ten or so km length is not conservative and does not appear to be valid for many cases.

Slip Rate Method

The use of fault slip rate was proposed by Woodward-Clyde Consultants in 1979 and in subsequent responses of Southern California Edison in response to questions for the San Onofre Nuclear Generating Station (Figure 4). This method was restricted to strike-slip faults, particularly of the branching type found in California. Faults with high magnitudes, for example of Ms magnitudes of above 8, all had high slip rates that were measured in cm/yr. Progressively lower slip rates were associated with faults that had maximum historical earthquakes with low slip rates. Faults with slip rates that were about 1/2 mm/yr were accompanied by maximum earthquake magnitudes that were about 6 or 6 1/2 magnitude. Application of this method could be based on a best fit (6.3 magnitude), extreme limits of error of measurement boxes (6.8 magnitude), or of somewhat higher and more conservative values as was used by the U. S. Nuclear Regulatory Commission NUREG-0712.

This method does not appear to provide similar progressive relations and moderate dipping normal-slip and reverse-slip faults with very low slip rates may large historical earthquakes of 7-1/2 to 8+ magnitude. These include examples from Pleasant Valley (Ms 7.6), China in 1932 (Ms 7.7), China in 1951 (Ms 7.4), Arvin-Tehachapi in 1952 (Ms 7.7-), Mongolia in 1957 (Ms 7.9), and Hegben Lake in 1959 (Ms 7.6), 1977 (Ms 7.4). These areas appear to have low slip rates.

ACKNOWLEDGEMENTS

Although the material presented in this paper represents our own efforts and conclusions, it was influenced by work for and with the staff of the U. S. Nuclear Regulatory Commission and earlier work for the Corps of Engineers, Waterways Experiment Station in Vicksburg. We are grateful for examples and discussions associated with this earlier work. The authors are grateful for

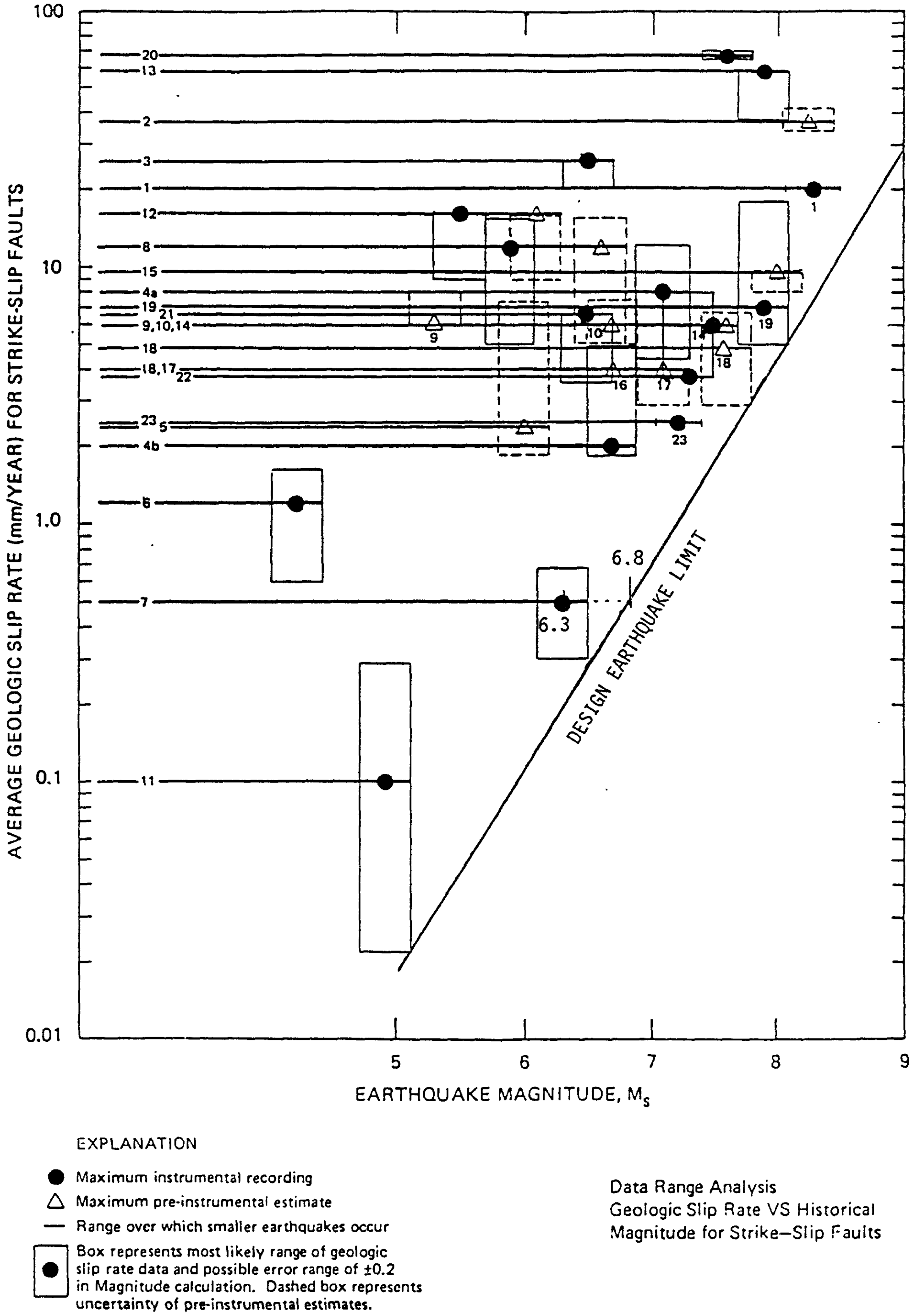


Figure 4. Geologic slip rate vs. earthquake magnitude for strike-slip faults from various parts of the world with the maximum earthquake magnitude for each fault indicated by the dots and triangles. The approximate error in determination of the magnitude is indicated by the sides of the error box and the geologic slip rate by the top and bottom of each box. The approximate bounding values for the range in natural values is shown by the diagonal line (modified from Woodward-Clyde Consultants, 1979).

editorial and word processing assistance from Barbara Matz of the Mackay School of Mines, University of Nevada Reno. We also are grateful for the helpful suggestions of Ellis Krinitzsky of Corps of Engineers, Waterways Experiment Station in Vicksburg.

REFERENCES

Albee, A. L. and Smith, J. L., 1966, Earthquake characteristics and fault activity in southern California: in Engineering Geology in Southern California, Los Angeles Section of Engineering Geologists, p. 9-34.

Bakun, William H. and McEvilly, Thomas V., 1984, Recurrence Models and Parkfield, California, Earthquakes: Journal of Geophysical Research, v. 89, no. B5, p. 3051.

Bonilla, M. G., Mark, R. K. and Lienkaemper, J. J., 1984, Statistical relations among earthquake magnitude, surface rupture length, and surface fault displacement: Seismological Society of America Bulletin, v. 74, p. 2379-2411.

Coppersmith, K. J., 1982, Probabilistic evaluations of earthquake hazards; California Division of Mines and Geology, Special Paper 62, p. 125-134.

Gutenberg, B., and Richter, C. F., 1954, Seismicity of the Earth and Associated Phenomena (2d ed.): Princeton University Press, Princeton, New Jersey, 310 p.

Hanks, T. C. and Kanamori, H., 1979, A moment magnitude scale: Journal of Geophysical Research, v. 79, p. 1215-1229.

Kanamori, H., 1983, Strong motion seismology [seminar reference material]: Evaluation of seismic hazards and decisionmaking in earthquake-resistant design, EERI Annual Meeting, Feb. 10, Reno, Nevada.

Lienkaempter, J. J., 1984, Comparison of two surface-wave magnitude scales: M of Gutenberg and Richter (1954) and Ms of "preliminary determination of epicenters": Seismological Society of America Bull., v. 74, p. 2357-2378.

Morisawa, M., and Hack, J. T. (eds.), 1985, Tectonic geomorphology: Allen & Unwin, Boston, 390 p.

Russ, D. P., 1979, Late Holocene faulting and earthquake recurrence in the Reelfoot Lake area, northwestern Tennessee: Geological Society of America Bulletin, v. 90, p. 1013-1018.

Schwartz, D. P., Coppersmith, K. J., Swan, F. H. III, Somerville, P. and Savage, W. U., 1981, "Characteristic" Earthquakes on Intraplate Normal Faults: Seismological Society of America, Earthquake Notes, v. 52, p. 71.

Schwartz, D. P. and Coppersmith, K. J., 1984, Fault Behavior and Characteristic Earthquakes: Examples from the Wasatch and San Andreas Fault Zones: Journal of Geophysical Research, in press.

Schwartz, D. P., and Coppersmith, K. J., in press.

Schwartz, D. P., Coppersmith, K. J. and Swan, F.H., 1984, Methods for estimating maximum earthquake magnitudes: 8th World Conference on Earthquake Engineering Proceedings, v. 1, p. 279-285.

Slemmons, D. B., 1977, Faults and earthquake magnitude: U. S. Army Engineers Waterways Experiment Station, Vicksburg, MS, Misc. Paper S-73-1, 166 p.

Slemmons, D. B., 1981, Geologic Considerations for Earthquake Microzonation: U.S. - China Workshop on Microzonation for Earthquakes, January 1981.

Slemmons, D. B., 1982, Determination of Design Earthquake Magnitudes for Microzonation: in Proc. 3rd Intl. Earthquake Microzonation Conference, v. 1, p. 119.

Slemmons, D. B., 1981, A procedure for analyzing fault-controlled lineaments and the activity of faults: in O,Leary, D. W., and Earle, J. L., Proceedings of the Third International Conference on Basement Tectonics, no. 3, p. 33-49.

Slemmons, D. B., and Chung, D. H., 1982, Maximum credible earthquake magnitudes for the Calaveras and Hayward fault zones, California: California Division of Mines and Geology, Special Publication 62, p. 115-124.

Slemmons, D. B. and dePolo, C. M., in press, 1985, Evaluation of active faulting and associated hazards.

Smith, S. W., 1976, Determination of maximum earthquake magnitude: Geophysical Research Letters, v. 3, p. 351.

Slemmons, D. B., Ramelli, A. R., and Brocoum, S., 1985, Earthquake potential of the Meers Fault, Oklahoma, Seismological Society of America Abstracts, 80th Annual Meeting.

Tocher, D., 1958, Earthquake energy and ground breakage; Seismological Society of America Bulletin, v. 48, p. 141-153.

U.S. Nuclear Regulatory Commission,

U.S. Nuclear Regulatory Commission, 1984, Part 100-Reactor Site Criteria: Code of Regulations, v. 10, Office of the Federal Register, National Archives and Records Service, p. 789 [10 CFR-100].

VanArsdale, R. B., 1985, Post Pliocene displacement on faults within the Kentucky River fault zone of east-central Kentucky: Geological Society of America Abstract with Programs, v. 17, p. 739.

Wallace, R. E., 1970, Earthquake recurrence intervals on the San Andreas Fault, California; Geological Society of America Bulletin, v. 81, p. 2875-2890.

Wallace, R.E., and Whitney, R. A., 1984, Long-term history of the Stillwater seismic gap, Nevada, as indicated by prehistoric fault scarps: Bulletin Seismological Society of America, v. 74, p. 301-314.

Woodward-Clyde Consultants, 1979, Report of the evaluation of maximum earthquake and site ground motion parameters associated with the offshore zone of deformation, San Onofre Nuclear Generating Station; Woodward-Clyde Consultants, 30 p. and Tables, Figures and Appendices.

Youngs, R. R. and Coppersmith, K. J., 1985, Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates: Bulletin Seismological Society of America, v. 75, p. 939.

Zoback, M. D., 1979, Recurrent faulting in the vicinity of Reelfoot Lake, northwestern Tennessee: Geological Society of America Bulletin, v. 90, p. 1019-1024.

STRONG-MOTION ATTENUATION RELATIONSHIPS*

by

Kenneth W. Campbell
U.S. Geological Survey
Golden, Colorado 80401

ABSTRACT

Research on strong ground-motion characteristics conducted in the United States within the last 10 years (1974-1984) forms the basis for a detailed discussion of important factors to be considered when selecting or developing strong-motion attenuation relations for use in earthquake engineering and seismic hazard studies. While emphasis is placed on the empirical prediction of ground-motion amplitudes, a brief discussion of procedures is presented that can be used when insufficient strong-motion data are available to perform an adequate statistical analysis. The discussion is followed by a tabulated summary of selected strong-motion attenuation relations proposed and developed in the last 10 years (1974-1984) to acquaint the reader with the types of relationships currently available.

INTRODUCTION

Studies concerned with evaluating seismic hazards related to ground shaking require the prediction of strong ground motion from earthquakes that pose a potential threat to the public, either by injury or damage to property. To make such a prediction, one must know certain fundamental characteristics of these earthquakes, as they relate to the source of the seismic waves, the medium through which the waves propagate, the local geology of the site, and the structures located at the site. If a sufficient number of strong-motion recordings from earthquakes and sites having the same or similar characteristics as those being evaluated are available, then it is straightforward to select an ensemble of these recordings for evaluating or

---

* Reprinted from Earthquake Spectra: The Professional Journal of the Earthquake Engineering Research Institute, v. 1, No. 4, p. 759-804, August 1985.

designing structures located at these sites for seismic loads (Fallgren et al., 1974; Guzman and Jennings, 1976; Werner, 1976b; Bernreuter, 1981a; Kimball, 1983; Campbell, 1984a; Heaton et al. 1984). Estimates of strong ground motion using this approach are currently referred to as site-specific.

For most applications, site-specific procedures are not feasible because a sufficient number of recordings with appropriate characteristics are generally unavailable. This is especially true for probabilistic analyses, where a wide range of earthquake sizes and locations are hypothesized, or for analyses where near-source estimates of ground motion are required. In such cases, a predictive model is needed. Such a model, commonly referred to as an attenuation relation, is expressed as a mathematical function relating a strong-motion parameter (e.g. peak acceleration) to parameters characterizing the earthquake, propagation medium, local site geology, and structure (Figure 1).

The remainder of this paper will be concerned with a discussion of factors to be considered in selecting or developing a strong-motion attenuation relation for use in deterministic or probabilistic seismic hazard studies. The discussion is divided into five elements: (1) the selection of parameters, (2) the selection of a data base, (3) the selection of a model or functional equation, (4) the selection of an analysis procedure, and (5) the evaluation of the relationship. While emphasis is placed on relationships derived from ground-motion recordings, there is a brief discussion of procedures that may be used when sufficient strong-motion data are unavailable. Following this is a summary of selected attenuation relations for peak acceleration, peak velocity, and other simple indices related to strong-motion amplitudes that have been developed in the last 10 years. Other general discussions on this subject may be found in Hofman (1974), Werner (1976a), Idriss (1978), Hays (1980), Young (1980a,b), and Boore and Joyner (1982). Discussions related to strong-motion recordings and their parameterization, including specific engineering applications, are presented by Hudson (1979), Housner and Jennings (1982), and Campbell (1984d).

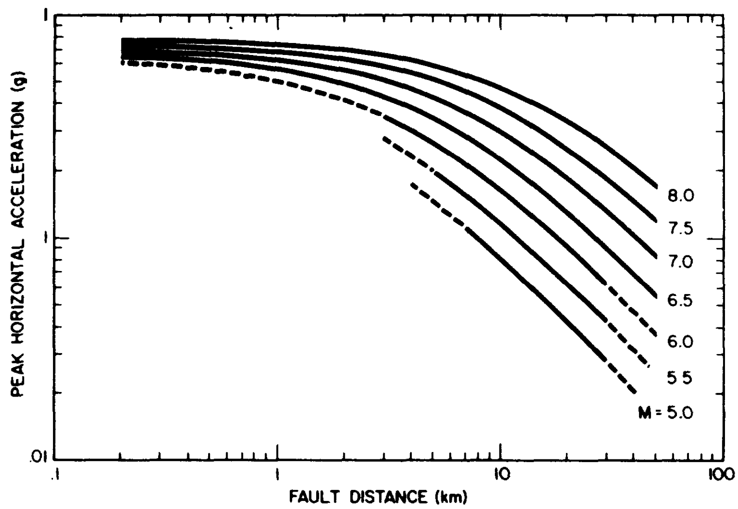


Figure 1. Near-source attenuation relation for peak horizontal acceleration. The dashed lines indicate an extrapolation of the relation based on little or no data. M is earthquake magnitude. Figure taken from Campbell (1981a).

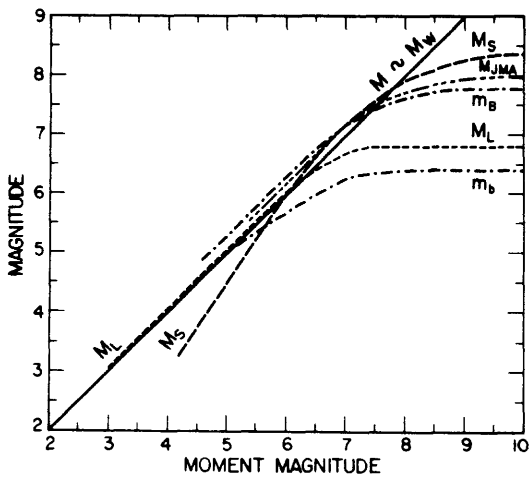


Figure 2. Relationship between moment magnitude and various magnitude scales: M (moment magnitude); M_s (surface-wave magnitude); M_L (local magnitude); M_{JMA} (Japan Meteorological Agency magnitude); m_b (short-period body-wave magnitude); m_B (long-period body-wave magnitude). Redrawn from Heaton et al. (1984).

PARAMETER SELECTION

In statistical terminology, the parameter to be predicted--in this case a strong-motion parameter--is referred to as a dependent variable. The parameters used to predict this variable are referred to as independent variables. There are two important considerations when selecting a parameter as an independent variable. First, the parameter should be reliable--that is, it should be characteristic of the earthquake, propagation medium (path), site, or structure it is meant to represent, and its estimation from existing data should be reasonably accurate and precise. Second, since the attenuation relation will be used to predict strong ground motion for future hypothesized events, the parameter should be predictable--that is, it should be easily estimated from known seismotectonic characteristics of the region.

Concerning the selection of a dependent variable, one should choose a strong-motion parameter (or parameters) that best relates to the purpose of the prediction, whether it be for zoning, planning, or design. While all would agree that the parameter selected should be representative of the seismic performance or damageability of the structure under consideration, there remains considerable controversy as to what parameters best relate to these effects. This stems from a poor understanding of what characteristics of ground motion cause damage in specific structures, a topic the subject of two recent workshops (Applied Technology Council, 1984; Earthquake Engineering Research Institute, 1984). It has become increasingly clear to the earthquake engineering community that peak acceleration alone is not adequate to characterize the seismic performance of structures (Sharpe, 1982; Campbell and Murphy, 1983; Kennedy, et al. 1984), although this has been known for sometime by experienced structural dynamicists (e.g. Housner, 1971). The remainder of this section will present a discussion of the various parameters that may be used to represent dependent and independent variables and factors that should be considered in selecting specific parameters to be used in predicting strong ground motion.

Strong-Motion Parameters. One must first decide what strong-motion parameter is to be predicted. Peak ground acceleration is most commonly used; however, as discussed above, it has come under much criticism for its lack of correlation with observed structural performance during past earthquakes. This has led several investigators to study a number of other peak parameters, including peak velocity and response spectra, as well as several energy related parameters such as Arias intensity, r.m.s. (root-mean-square) acceleration, Fourier spectra, power spectral density, and spectrum intensity. A list of references to publications related to these strong-motion parameters may be found in the Bibliography.

The acceleration time history is the most comprehensive description of ground motion one could use in earthquake engineering applications. It has the potential for incorporating all salient features of ground motion--in both time and frequency--and can be used in elastic and inelastic analyses of any type of structure. However, such time-domain analyses are extremely expensive and time consuming to perform and are not feasible for most engineering applications.

The response spectrum is probably the most comprehensive description of ground motion that is easily used by design engineers (Applied Technology Council, 1974; Sharpe, 1982), but these data are not as readily available nor as completely reported as peak acceleration. As an added disadvantage, the prediction of response spectra requires the development of several attenuation relations, one for each structural period and damping of interest. To simplify the development of response spectra, engineers have adopted the use of standard response spectra shapes whose ordinates are proportional to peak ground-motion parameters (e.g. Newmark et al., 1973; Mohraz, 1976; Seed et al., 1976a; Werner, 1977; Applied Technology Council, 1978; Wong and Yun, 1979; Newmark and Hall, 1982). Although simple to use, standardized spectra generally lack the sensitivity to earthquake size and source-to-site distance exhibited by recorded spectra. However, this disadvantage can be overcome to a large extent by the use of spectral shapes that scale to peak acceleration at short periods, to peak velocity at intermediate periods, and to peak displacement at long periods (e.g. Newmark and Hall, 1982). In the future, standardized spectra will no doubt be replaced by relationships for individual response spectral ordinates as they become more readily available.

Because strong gound motions are usually recorded on three orthogonal components, one must decide which component(s)--horizontal or vertical--are to be predicted. In addition, one must decide how the horizontal components are to be treated. In the past, treatment of strong-motion parameters from the two horizontal components have included the use of (1) the largest of the two components, (2) both components, (3) the mean of both components, (4) the vectoral combination of both components, and (5) the random selection of components (e.g. see Table 2). The use of both horizontal components results in a prediction representing a random selection of component orientations and is found to give median predictions identical to, but standard deviations larger than, those using the mean of the two horizontal components (Campbell, 1982a). Because of the strong correlation between the two horizontal components, the use of both components as independent data points will artificially increase the statistical significance of the resulting analyses unless their dependence is appropriately accounted for in the analyses. To insure consistency, the choice of a method for combining components should take into consideration the intended use of the predictions by the engineer responsible for the seismic analyses (Donovan, 1982a).

For most engineering applications, the amplitude of the vertical component of ground motion is simply taken to be two-thirds of the amplitude of the horizontal components when estimating peak parameters and response spectra (Young, 1980a; Newmark and Hall, 1982). However, recent experience suggests that this rule of thumb may not be appropriate near the source where peak vertical accelerations have been observed to equal or exceed peak horizontal values for moderate-to-large earthquakes (Bureau, 1981; Campbell, 1982). This same experience indicates that the ratio of vertical to horizontal components tends to decrease with distance, dropping below two-thirds at large distances. Therefore, one should, if at all possible, avoid the pragmatic use of a constant vertical to horizontal ratio when statistically estimating vertical amplitudes of strong ground motion.

Earthquake Parameters. The parameter most commonly used to characterize earthquake size in strong-motion attenuation relations is earthquake magnitude. Magnitude is the only source parameter routinely reported by seismographic networks. Other source parameters used in the past have included source dimensions (Ts'ao, 1980; Bernreuter, 1981b), seismic moment or moment magnitude (Hanks, 1979; McGuire and Hanks, 1980; Hanks and McGuire, 1981; Joyner and Boore, 1981, 1982; McGuire et al., 1984), and stress drop (Hanks and Johnson, 1976; Hanks, 1979; Ts'ao, 1980; Bernreuter, 1981b; McGuire and Hanks, 1980; Hanks and McGuire, 1981; McGuire et al., 1984).

While stress drop is an important source parameter from a theoretical point of view, in practice its estimation is associated with a large degree of uncertainty. This, coupled with the results of several studies suggesting that localized stress drop may be relatively independent of other measures of earthquake size (Hanks and McGuire, 1981; Aki, 1982; Papageorgiou and Aki, 1983)--and that static stress drop does not correlate with r.m.s. acceleration (Hanks and McGuire, 1981), peak ground-motion parameters (Boore, 1983a; Atkinson, 1984), or response spectra (McGuire, et al., 1984)--would indicate that stress-drop parameters are not very reliable. Seismic moment, or its equivalent moment magnitude (Hanks and Kanamori, 1979), is preferred by some investigators (e.g. Boore and Joyner, 1982) because it corresponds to a well-defined physical property of the source. Its use is currently hindered, however, because routine calculations of seismic moment have only recently become available (Bolt and Herraz, 1983). For many past earthquakes, as well as for most smaller events, seismic moment is unavailable or only crudely estimated. For example, Joyner and Boore (1981) found it necessary to use local magnitude (M_L) in place of moment magnitude for several earthquakes in their data set.

Earthquake magnitude, although routinely reported and universally used as a measure of earthquake size, is not without its limitations. The variety of magnitude scales that exist can lead to confusion in comparing various predictions. There is also a clear tendency for all scales, except moment magnitude, to reach a limiting value (saturate) as the size of the earthquake increases (Figure 2). Because most magnitude scales are based on the peak amplitude of an instrumental recording, one might expect a good correlation between magnitude and a ground-motion parameter of similar frequency. For example, Boore (1980) found a strong correlation between peak velocity of strong-motion recordings and peak amplitude of Wood-Anderson seismographs, suggesting a direct relationship between peak velocity and M_L. Extending this logic, short-period estimates of ground motion might be expected to correlate best with m_b (body-wave magnitude) or M_L (local magnitude), and long-period estimates of ground motion might be expected to correlate best with M_S (surface-wave magnitude) or **M** (moment magnitude). However, this simple concept has not been verified empirically.

A critical element in the choice of a magnitude scale involves the specification of the magnitude of a future hypothetical earthquake. Because of limitations of most magnitude scales, magnitudes are usually specified in terms of one or more different scales. For instance, surface-wave magnitudes are, in general, not reliably determined for magnitudes less than about 6, and because of saturation, M_L and m_b become relatively independent of earthquake size for magnitudes near 7. Therefore, magnitudes are generally specified in terms of m_b or M_L for smaller earthquakes and M_S or **M** for larger earthquakes. This dual use of magnitude scales is consistent with the interpretation of the Richter magnitude scale by Nuttli (1979), who suggests that the widely used Richter scale represents M_L for magnitudes less than about 6 and M_S for larger earthquakes. A similar generic scale has been used in relationships among earthquake source dimensions and magnitude (Slemmons, 1977), which form the basis for estimating maximum magnitudes for many faults. Therefore, if the attenuation relation is to represent a wide range of magnitudes, it may be desirable to use a dual magnitude scale to be consistent with the application of the relationship (e.g. Campbell, 1981a). For regions outside the Western United States, it is probably more appropriate to replace M_L by m_b, a standard measure of magnitude worldwide, or some similar regional magnitude scale (e.g. m_{bLg} in the Eastern United States). Whatever scale is used, it is important to clearly state the choice and be consistent in its use.

Chung and Bernreuter (1981), Herrmann and Nuttli (1982), and Nuttli and Herrmann (1982) have observed regional differences in magnitude determinations for m_b that should also be considered in the development and application of attenuation relations. They found that the determination of m_b is strongly affected by regional variations in the Q structure (attenuation characteristics), composition, and physical state within the earth. For example, because of differences in attenuation properties between the Western and Eastern United States, a regional m_b magnitude bias exists, which, depending on where the earthquake occurs and where the ground motion is recorded, can lead to magnitudes as much as one-third unit larger in the Eastern United States. Chung and Bernreuter (1981) also point out that when using regional catalogs to obtain magnitudes, it is often necessary to determine how the reported magnitudes were determined. This may also be true for more universal scales such as m_b and M_s. For example, a significant change in the m_b scale occurred in the early 1960's when the World-Wide Standard Seismograph Network (WWSSN) was established. This change in instrumentation had a significant effect on estimated magnitudes and the saturation level of the m_b scale (e.g. compare m_b and m_B in Figure 2). The older, longer-period instruments resulted in larger magnitudes than are currently determined from the WWSSN instruments.

Another earthquake source parameter found to be related to strong ground motion is focal mechanism. Campbell (1983, 1984c), in his empirical analysis of near-source ground motion, found that reverse and reverse-oblique mechanisms are associated with ground motions approximately 30-40 percent larger than strike-slip mechanisms. Young (1980a) attributes such differences to regional variations in stress drop. Young's interpretation is consistent with theoretical analyses of McGarr (1982, 1984) and Anderson and Luco (1983), which predict differences in strong motion based on differences in the amplitude and orientation of tectonic stress. Other source effects found to influence strong ground motion are source radiation pattern, source directivity, and geometry of the fault plane (Berrill, 1975; Arnold and Vanmarcke, 1977; Bureau, 1978; Boatwright and Boore, 1982; Anderson and Luco, 1983; Singh, 1983), focal depth (McGarr, 1984), and near-field pulses (Bolt, 1981; Luco and Anderson, 1984). The latter effect is especially significant for sites located near the fault.

Propagation Parameters. Propagation parameters characterize the effects of wave scattering, geometrical attenuation, and anelastic attenuation of ground motion as it travels from the source to the site. The independent variable universally used to characterize these parameters is distance. The attenuation parameters themselves are usually determined from strong-motion data or seismological network data. Exceptions to this will be discussed later. Because earthquake rupture can extend over tens to hundreds of kilometers, a number of distance measures have come into use (Figure 3). The measure used should depend on the specific application. For sites located several source dimensions from the earthquake, there is little difference between distance measures. However, for shorter distances, the difference between measures becomes significant. In the near-source region, where predictions are of greatest concern, the use of epicentral or hypocentral distance (M1 and M2 in Figure 3) leads to considerably greater scatter in estimates of strong ground motion than the use of distance measures representing closest distance to the fault (M4 and M5) (Campbell, 1982a; Huang et al., 1982). Schnabel and Seed (1973) first recognized the importance of using a fault-distance measure for sites near the rupture, and many recent studies have adopted such a measure (see Table 2). Notable exceptions are those relationships in Table 2 (e.g. Trifunac, 1976b; McGuire, 1978b) which utilize epicentral or hypocentral distance. Most of these investigators, however, acknowledge that their relationships should not be used at near-fault distances (i.e. within several source dimensions of the rupture zone).

Some investigators have argued that using a fault-distance measure can lead to biased predictions, especially if the strong-motion stations used in the analysis have a nonrandom distribution around the fault or if the strong motions come from a localized source (or sources) on the fault (Shakal and Bernreuter, 1981; Toro, 1981). This latter concept of a localized source is represented by distance measure M3 in Figure 3. While the distribution of stations is rarely random for a particular earthquake, the randomization introduced by considering recordings from an ensemble of earthquakes should help reduce this possible bias. In any case, Boore and Joyner (1982) point

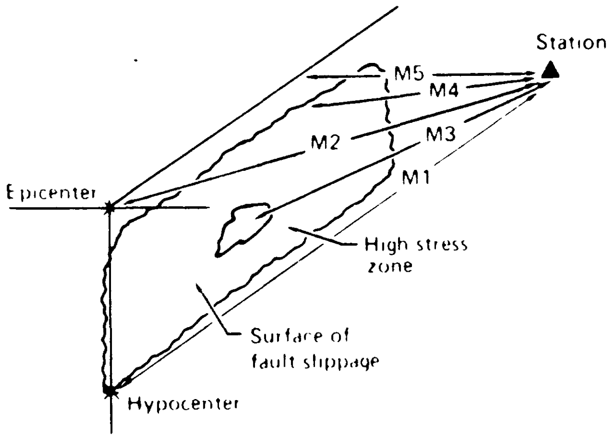


Figure 3. Schematic showing distance measures used in strong-motion attenuation relations: M1 (hypocentral distance); M2 (epicentral distance); M3 (distance to energetic zone); M4 (closest distance to rupture zone); M5 (closest distance to surface projection of rupture zone). Figure taken from Shakal and Bernreuter (1981).

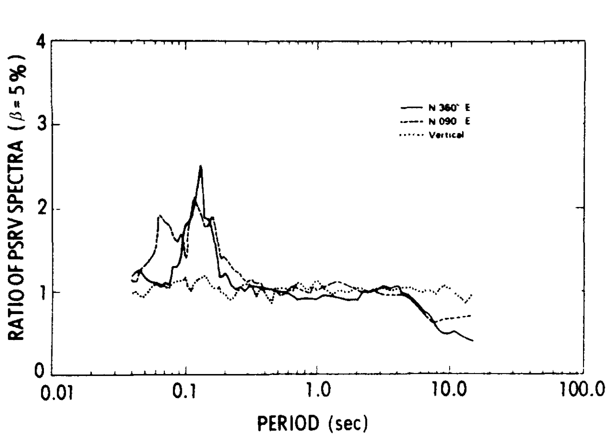


Figure 4. Ratio of 5%-damped pseudo-relative velocity (PSRV) response spectra, recorded in an instrument shed, to the spectra obtained at an adjacent free-field instrument at the Differential Array site in El Centro, Calif. Recordings were obtained during the 1979 Imperial Valley earthquake. Figure taken from Campbell (1983).

40008
1458 ~~40007~~

out that the placement of recording instruments from which the data are obtained and the placement of structures (or sites) for which predictions are to be made are comparable sampling processes from a statistical point of view.

If the strong motions are radiated from small areas of the fault rupture surface (referred to as asperities), then a fault-distance measure would tend to underestimate the actual distance to these localized sources. This is not, however, a serious limitation in practice. While it may be possible to identify these asperities for some past earthquakes, it is virtually impossible to anticipate their locations during future events. Thus, such a distance measure is unpredictable. Because of this, most earthquake scenarios, whether for probabilistic or deterministic applications, use the closest approach of the fault, tectonic structure, or earthquake rupture as the representative distance from the hypothesized earthquake. This is completely consistent with the definition of the closest distance measures M4 and M5 in Figure 3 and justifies their use in attenuation relations used to predict strong ground motions from such events. If, however, an analysis hypothesizes earthquake sources to be equally distributed along a fault or within an area, with no accommodation of source rupture, then epicentral distance, hypocentral distance, or distance to the energy center would be the more appropriate measure to use. In this case, attenuation relations in terms of fault distance will indeed underestimate the true ground motions.

Site Parameters. Traditionally, site parameters have been related to simple geologic descriptions of the recording stations. For example, a summary of selected site classification schemes used in the last 10 years (1974-1984) is presented in Table 1. Entries in Table 1 are keyed to the attenuation relations appearing in Table 2 and were selected using the same criteria. These criteria are presented later in the text. Exceptions are the classification schemes of Borcherdt and Gibbs (1976) and Mohraz (1976), which are included by virture of meeting all criteria except those specifically related to the selection of specific ground motion parameters included in Table 2. For completeness, the original reference to the classification scheme is included if known. More sophisticated classifications of the site have been based on wave-propagation velocity (Blume, 1977; Joyner et al.,

TABLE 1

SUMMARY OF SELECTED SITE CLASSIFICATION SCHEMES (1974-1984)

Classification	Description	Reference
Soil	Alluvium and soft deposits >10 m deep	McGuire (1978b)
Rock	Basement rock, sedimentary rock, and soil ≤10 m deep	
Soil	Alluvium, soil, mud, fill, and glacial deposits >4-5 m deep	Boore *et al.* (1980), Joyner and Boore (1981), Bolt and Abrahamson (1982)
Rock	Basement rock, sedimentary rock, and soil ≤4-5 m deep	
Soft	Soft alluvium	Trifunac and Brady (1975a), Trifunac (1976b)
Intermediate	Sedimentary rock	
Hard	Basement or crystalline rock	
Shallow alluvium	Alluvium 7-20 m deep	Duke *et al.* (1972), Campbell and Duke (1974a,b)
Deep alluvium	Alluvium >20 m deep	
Sedimentary rock	Sedimentary rock with alluvium <7 m deep	
Crystalline rock	Igneous and metamorphic rock with alluvium <7 m deep	
Rock	Basement and sedimentary rock	Mohraz (1976)
Alluvium	Alluvium of unspecified thickness	
Alluvium (<9 m)	Alluvium <9 m deep	
Alluvium (9-60 m)	Alluvium 9-60 m deep	
Granite	Pre-Tertiary basement rock	Borcherdt and Gibbs (1976)
Franciscan formation	Sedimentary and volcanic rock of the Franciscan melange	
Great valley sequence	Sedimentary rock of pre-Tertiary age	
Santa Clara formation	Sedimentary rock of Pliocene and early Pleistocene age	
Alluvium	Soils of late Pleistocene and Holocene age with less than 40 weight percent water	
Bay mud	Recently deposited clays and silts with more than 50 weight percent water	
Rock	Shale-like and harder rock with V_s >760 m/s	Schnabel and Seed (1973), Seed *et al.* (1976b), Donovan and Bornstein (1978)
Stiff soils	Stiff clay, sand, or gravel <45 m deep	
Deep cohesionless soils	Generally cohesionless soils >75 m deep	
Soft soils	Soft to medium-stiff clays with sand and gravel	
Hard rock	Crystalline, hard metasedimentary, and hard volcanic rock	Campbell (1981a)
Soft rock	Sedimentary, soft metasedimentary, and soft volcanic rock	
Pleistocene deposits	Pleistocene age soils >10 m deep	
Recent alluvium	Holocene age soils >10 m deep	
Shallow soils	Soils <10 m deep	
Soft soils	Extremely soft or loose soils	

1983; Joyner and Fumal, 1984) and depth of deposits (Trifunac and Lee, 1978a, 1979; Rogers et al., 1983, 1984; Campbell, 1984c). The diversity of site classifications used in the past attests to the complex and poorly understood relationship between strong ground motion and site characteristics.

While the classifications in Table 1 may be used as a guide in establishing site parameters for strong-motion attenuation relations, they should not be adopted without careful consideration of several factors. One such factor is the complex relationship between site and structure effects. Crouse (1978) suggests that effects attributed to the free-field response of the recording site in the past may actually reflect a modification of the ground motion by the structure housing the instrument. This was confirmed by Campbell (1983, 1984b), who found that factors such as fault mechanism, site topography, soil depth, instrument embedment, and structure size, if not properly accounted for in the development of strong-motion attenuation relations, can significantly influence the quantification of site effects.

Campbell (1981a, 1983), Chiaruttini and Siro (1981), and Faccioli (1981) have recently observed a large amplification (as much as a factor of two) in accelerations associated with shallow soil deposits for sites located near the source of small to moderate earthquakes. Even larger amplifications have been observed for short-period spectral parameters (Mueller et al., 1982; Rogers et al., 1983, 1984). The classification of these shallow sites as rock--a common practice in the past--can significantly increase estimates of short-period components of strong ground motion for rock if enough of these sites are included in the analysis. One should also be aware of the possible effects of site topography. The significant influence of topography was first documented for the 1971 San Fernando earthquake by Boore (1973), Davis and West (1973), and Mickey et al. (1973). More recent empirical evidence has been presented by Chang (1980), Campbell (1983), and Tucker et al. (1984). Campbell (1983) finds that the majority of rock recording sites in the United States are situated in areas of steep topography, suggesting that this may have influenced the relationship between rock and soil response found in past analyses.

The site classification scheme selected should be compatible with the strong-motion parameter being predicted. Different site characteristics will influence each strong-motion parameter differently, and their effects will vary depending on the distance of the recording from the source and the size of the earthquake. These differences relate to differences in the frequency content of the ground motion (Rogers et al., 1983, 1984). For example, shallow soils have been observed to amplify accelerations at sites located relatively near the source, while peak velocities are found to be virtually unaffected by the presence of such shallow deposits (Campbell, 1983). The depth of sediments (i.e. the depth to basement rock) is found to correlate only with moderate- to long-period components of strong ground motion (Hanks, 1975; Trifunac and Lee, 1978a, 1979; Rogers et al., 1983, 1984; Campbell, 1984c; King and Tucker, 1984; Tucker and King, 1984).

Structure Parameters. If free-field predictions of strong ground motion are desired, then parameters characterizing the effect of the structure in which the recording was obtained may be required. These effects have usually been neglected or confused with the effects of site response in the past. However, recent empirical studies have indicated that ground motions can be significantly affected by the size and embedment of a building (Crouse, 1978; Boore et al., 1980; Seed and Lysmer, 1980; McCann and Boore, 1982; Campbell, 1979, 1983, 1984b, 1984c), confirming the results of theoretical soil-structure interaction analyses. Boore et al. (1980) classified structures into large buildings (greater than two stories in height) and small buildings or shelters, and they found significant differences in peak accelerations recorded during the 1971 San Fernando earthquake. This formed the basis for excluding large buildings in the subsequent analyses of Joyner and Boore (1981, 1982) and Campbell (1982a, 1983). Campbell (1979, 1983, 1984b) used parameters to account for differences in peak accelerations between embedded and ground-level buildings and found reductions as large as 50 percent due to instrument embedment.

The analysis of strong-motion recordings obtained primarily in basements of buildings during the 1971 San Fernando earthquake (Crouse, 1976; Lee et al, 1982) suggests that building embedment rather than foundation size might be the factor controlling the reduction of short-period ground motions observed for ordinary buildings in the past. In contrast, the kinematic scattering effects of extremely large, rigid foundations, such as those associated with nuclear powerplants, can cause significant reductions in amplitudes whether embedded or not when the wavelengths of ground motion are smaller than the characteristic size of the foundation (Scanlan, 1976; Bycroft, 1977; Loh et al., 1982; Smith et al., 1982). As with site parameters, these effects will vary depending on the strong-motion parameter investigated, the distance to the source, and the size of the earthquake.

Mickey et al. (1973) and Riemer et al. (1973) document the effects of the response of Pacoima Dam on the abutment instrument recording during the 1971 San Fernando earthquake. Their studies suggest that the response of dams may also have to be considered in the development of strong-motion attenuation relations. Joyner and Boore (1981, 1982) removed recordings on the abutments and toes of dams for this reason. Studies of Bycroft (1978), Crouse (1983), McNeill (1983), Campbell (1983, 1984b), and Crouse et al. (1984) indicate that so-called free-field recordings can be amplified substantially by small instrument shelters, especially if they are founded on very soft soils. Recordings obtained at the Differential Array in El Centro, Calif., during the 1979 Imperial Valley earthquake are evidence of this potentially important effect (Figure 4).

DATA SELECTION

Once the dependent and independent parameters have been selected, a data base must be chosen. Selection criteria should be established to insure that minimum standards of quality and consistency are met. If this is not done, biases will be introduced into the analyses, resulting in increased scatter in the predictions. Significant bias and scatter can be largely avoided if records are selected to represent (1) tectonic provinces of similar attenuation and source characteristics, (2) recording instruments of similar response characteristics, (3) consistent and accurate record-processing techniques, and (4) consistent definitions of strong-motion, earthquake, path, site, and structure parameters. Data should be selected to represent the range of parameters for which predictions are to be made. Inclusion of data outside this range can also result in increased bias and scatter in the predictions. Another potential source of bias arises when independent variables are statistically correlated. This can result in biased estimates of coefficients during regression analyses. Scatter plots (Figure 5) or correlation analyses may be used to identify any significant correlations that might exist. A modification of the selection criteria or the use of special analysis techniques should be considered if significant biases are found.

Consistency among the data can be obtained by either excluding those records that do not meet the recording characteristics to be predicted or by including parameters that adequately account for these characteristics. The first technique is used when undesirable recordings make up a relatively small percentage of the total data set or when there is sufficient data having the appropriate characteristics for a statistically stable analysis. The most common application of this technique has been the selection of data based on site characteristics (Schnabel and Seed, 1973; Seed et al., 1976a, 1976b; Donovan and Bornstein, 1978; Sadigh et al., 1978; Blume, 1980; Boore et al. 1980; Seed and Idriss, 1982; Idriss, 1983; Sadigh, 1983); however, others have used this procedure to segregate data by magnitude (Seed et al., 1976b; Sadigh et al., 1978; Boore et al., 1980; Bolt and Abrahamson, 1982) and structure size (Boore et al., 1980; Campbell, 1982a, 1983; Joyner and Boore, 1981, 1982). The second technique is used when a parameter represents an independent variable required for the prediction, such as magnitude or distance, or when excluding the undesirable data would leave too few data for a stable statistical analysis.

Data should not be removed from the data base when they represent a random characteristic of the earthquake, path, site, or structure. A random characteristic is one that cannot be reliably predicted in the future. For example, the azimuthal variations in ground motion due to source radiation patterns and directivity (directional focusing) require a knowledge of the location and direction of rupture, characteristics generally not known in advance. In this case, the scatter represented by these data would reflect a true random uncertainty in the prediction of a strong-motion parameter. Such random uncertainty can be appropriately accounted for in both probabilistic

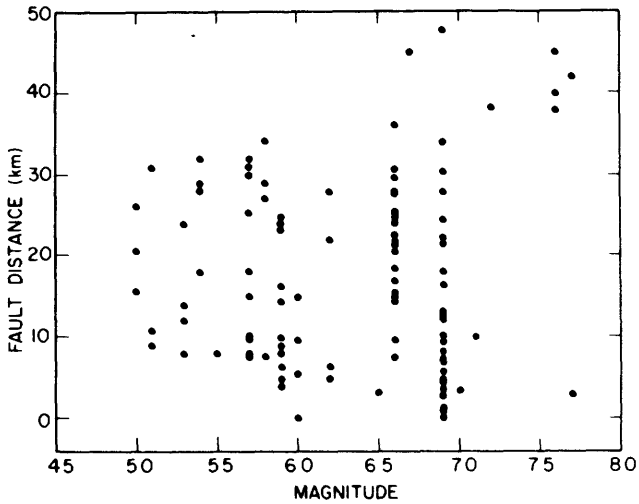


Figure 5. Scatter plot of strong-motion data used to develop the attenuation relation shown in Figure 1. Figure taken from Campbell (1981a).

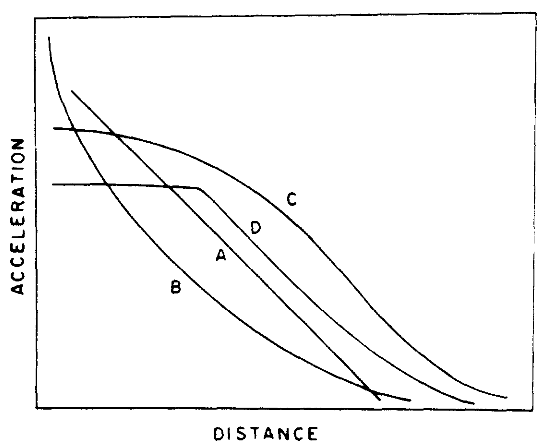


Figure 6. Schematic showing attenuation characteristics of selected functional models used to develop strong-motion attenuation relations: Curve A (linear relation); Curve B (Equation 3a with b_5=0); Curve C (Equation 3a or 3b); Curve D (relation proposed by Bolt and Abrahamson, 1982). Both axes are linear with respect to the variables being plotted. Figure taken from Bolt and Abrahamson (1982).

1457 40008

and deterministic analyses through the specification of confidence limits (discussed later in the text). Inclusion of data that represent a systematic characteristic of the earthquake, path, site, or structure will lead to greater uncertainty in the predicted strong-motion parameter. This bias is extremely critical when uncertainty is treated as random scatter in probabilistic analyses for which predictions are made for small probability levels (long return periods).

MODEL SELECTION

The next step in the development of a strong-motion attenuation relation is the selection of a mathematical function or model relating the independent and dependent variables. The functional form of the model will depend, in general, on the use of the relationship and the data base selected. If the data base represents a relatively uniform subset of data, then a function having only a few parameters would be appropriate. If predictions are to be restricted to a range of parameters well represented by the data (e.g. near the centroid of the magnitude-distance space defined in Figure 5), then a relatively simple empirical model would be justified. However, if the attenuation relation is to be extrapolated much beyond the centroid of the data, then it is important that the model have a physical basis for such an extrapolation to be meaningful.

The physical basis of attenuation relations used in the past have been restricted to the most fundamental principles of seismology and geophysics. However, this has given little information on what form the function should take at distances close to the fault, where details of the rupture process become important. Modeling this process has helped to define the form of the function in this critical region (e.g. McGarr et al., 1981; Hadley et al., 1982; Scholz, 1982; Anderson and Luco, 1983; Gusev, 1983).

The general form chosen by most investigators can be characterized by the expression

$$Y = b_1 \, f_1(M) \, f_2(R) \, f_3(M,R) \, f_4(P_i) \, \varepsilon \qquad (1)$$

where Y is the strong-motion parameter to be predicted (dependent variable); $f_1(M)$ is a function of the magnitude scale M; $f_2(R)$ is a function of the distance measure R; $f_3(M,R)$ is a joint function of M and R; $f_4(P_i)$ is a function representing parameters of the earthquake, path, site, or structure; and ε is a random variable representing the uncertainty in Y.

In its most common form, the function $f_1(M)$ is an exponential function of magnitude,

$$f_1(M) = e^{b_2 M} \qquad (2)$$

which comes from the basic definition of magnitude as a logarithmic measure of ground motion amplitude (Richter, 1958). However, others have used the exponential of a quadratic of magnitude and the reciprocal of magnitude to represent this function (e.g. Werner et al., 1979; Joyner and Boore, 1982).

The most common form for $f_2(R)$ is

$$f_2(R) = e^{b_4 R} \, [R + b_5]^{-b_3} \qquad (3a)$$

where the term in brackets accounts for attenuation due to geometrical spreading (b_3 representing the geometrical attenuation rate) and the exponential of R accounts for anelastic attenuation--that is, material damping and scattering (b_4 representing the coefficient of anelastic attenuation). Both of these functions come from basic principles of wave propagation in elastic media. The coefficient b_5 is used by some investigators to limit the value of Y at zero distance, a property referred to as saturation (in this case saturation with distance). This is especially necessary when a distance measure, such as epicentral distance or distance to the fault trace, is used, which can take on values of zero. An alternate expression commonly used in place of equation (3a) is

$$f_2(R) = e^{b_4 R} \, \left[\sqrt{R^2 + b_5^2}\,\right]^{-b_3} \qquad (3b)$$

where the term in brackets is analogous to the definition of hypocentral distance. Some investigators have replaced b_3 in equations (3a) and (3b) by a

1487 40008

logarithmic function of R (Donovan and Bornstein, 1978; Campbell, 1979, 1982a; Espinosa, 1979, 1980), while others (e.g. Bolt and Abrahamson, 1982; Brillinger and Preisler, 1984) have used more complicated expressions for $f_2(R)$ to account for the distance-saturation properties of strong ground motion. Variations in distance scaling characteristics due to differences in functional models is demonstrated in Figure 6.

The function $f_3(M,R)$ is used to account for differences in magnitude scaling with distance. In its most common form, $f_3(M,R) = 1$; however, in studies that incorporate such a function, it is generally characterized by the expression

$$f_3(M,R) = [R + b_6 e^{b_7 M}]^{-b_3} \quad (4)$$

which simply replaces b_5 in equation (3a) with an exponential function of magnitude (Esteva, 1970; Idriss, 1978; Campbell, 1981a; Sadigh, 1983). A similar function has been used in place of b_5 in equation (3b) by some investigators. For negative values of b_7, as is generally the case, this function reduces the amount of magnitude scaling at short distances, another form of saturation (in this case saturation with magnitude).

Magnitude saturation of peak acceleration near the fault has been proposed on both empirical and physical grounds (e.g. Campbell, 1981a; Chung and Bernreuter, 1981; McGarr, 1982, 1984; Campbell and Niazi, 1982; Hadley et al., 1982; Gusev, 1983; Munguia and Brune, 1984; see also Campbell, 1981a for a list of earlier references), but there is still considerable controversy regarding this characteristic. Joyner (1984) presents an earthquake dislocation model that requires strong-motion parameters to saturate with magnitude at all distances, not just near the fault. This latter characteristic was adopted by Trifunac (1976a, 1976b, 1978) and Joyner and Boore (1982), all of whom used a quadratic function of magnitude in place of $b_2 M$ in equation (2) to incorporate magnitude saturation in their attenuation relations. Another expression for $f_3(M,R)$ used in the past involves replacing b_3 in equation (3a) by a linear function of magnitude, $b_3 M$ (Donovan and Bornstein, 1978; Campbell, 1979; Idriss, 1983).

The function $f_4(P_i)$ is usually represented by an expression of the form

$$f_4(P_i) = \sum e^{b_i P_i} \tag{5}$$

While somewhat arbitrary, this expression agrees with empirical evidence suggesting that most source and site effects are multiplicative. The most common parameter included in this expression is that related to geologic classifications of the site; however, parameters related to characteristics of the earthquake and structure have also been included in this way (Table 2; Campbell, 1983, 1984c). Although not commonly done, one could add functions of magnitude and distance to equation (5) if P_i is found to correlate with these parameters (e.g. Campbell, 1984c).

The random variable ε is usually assumed to be lognormally distributed, though this is not a requirement in most regression procedures. Some justification comes from the exponential form of the functions used in equation (1), recognizing that the product of lognormally distributed variables is itself lognormally distributed. An a posteriori empirical justification in support of a lognormal distribution for ε comes from statistical tests on the observed scatter about the predicted values of Y (Esteva, 1970; Donovan, 1973; Donovan and Bornstein, 1978; McGuire, 1978a, 1979; Campbell, 1981a), but these results might be biased by the assumed functional form of the relationship.

Brillinger and Preisler (1984) present statistical procedures for determining the optimal functions (transformations) for Y, M, R, and P_i in equation (1) from strong-motion data. This is generally referred to as Exploratory Data Analysis or EDA. This powerful technique would eliminate the need for a priori assumptions regarding the form of the model. For example, using the data base of Joyner and Boore (1981), Brillinger and Preisler found that the optimal transformation of Y was $Y^{1/3}$, not ln Y. However, one must be extremely cautious in adopting such statistically determined functions, since they may not comply with known physical characteristics of ground motion and could thus lead to unrealistic estimates when extrapolated. Some of the

1452 40008

recent controversy surrounding the choice of an appropriate functional form may be found in Donovan (1982a, b), Bolt and Abrahamson (1983), and Joyner and Boore (1983).

SELECTION OF AN ANALYSIS PROCEDURE

Having selected a model, one must choose a procedure for determining the unknown coefficients (the b_i's) in equations (1) through (5). Such a procedure is referred to as regression analysis. Because of the apparent lognormal (or near lognormal) distribution for Y--the strong-motion parameter to be predicted--regressions are usually performed on the logarithm of Y, based on the model

$$y = \ln b_1 + \ln[f_1(M)] + \ln[f_2(R)] + \ln[f_3(M,R)] + \ln[f_4(P_i)] + \varepsilon' \quad (6)$$

where y=ln Y, ε'=ln ε, and ε' is a random variable with a mean of zero and a standard deviation of σ. The term σ is referred to as the standard error of estimate of y. If ε is lognormally distributed, then ε' will have a Normal or Gaussian distribution. This, however, is not a necessary requirement for regression analyses. It is required in order to make certain statistical statements about the results, as will be discussed in the next section. While the discussions that appear in the remainder of this paper are based on the transformation y=ln Y, they are equally applicable to any other transformation of Y. One need only replace y by the appropriate transformation in the equations that follow.

Regressions in the past have been performed almost exclusively using a least-squares procedure. The least-squares procedure minimizes the sum square error

$$\sum_{i=1}^{n} w_i \, (\hat{y} - y_i)^2 \quad (7)$$

where $\hat{y}$ is the predicted value of y, y_i is the $i^{\underline{th}}$ observed value of y, and w_i is the weight assigned to y_i (in the case of weighted regressions). Other procedures, such as the least absolute sum criterion, may be used to minimize the influence of outlying observations on the results. If equation (6) (or some alternate model used in place of this expression) is linear with respect to the coefficients to be determined, then standard linear least-squares procedures can be used. I not, then nonlinear procedures (e.g. Gallant, 1975; More et al., 1980; SAS Institute, 1980) must be used. If the model is linear and the coefficents are Normally distributed, then a t-test may be used to establish the statistical significance of the coefficients. Any coefficient not meeting the required significance level (for example, a 90 percent probability of not being zero) should then be removed from the model. Stepwise regression procedures are useful for this purpose, especially if it is not known in advance which parameters are important. If the model is nonlinear, the distributions for the nonlinear coefficients must be developed empirically using Monte Carlo simulation techniques (Gallant, 1975; Campbell, 1981a; Boore and Joyner, 1982).

Biased estimates of the coefficients will be obtained if the data are not distributed evenly among the parameters, for example, if magnitude and distance are statistically correlated, or if the data are dominated by many recordings from a few earthquakes or recording sites. Attempts at reducing this bias in the past have included: restricting the data sample to no more than a certain number of recordings from a given earthquake and a given site (McGuire, 1978a, 1978b; Cornell et al., 1979); the use of weighted regression procedures to equalize the impact of recordings from individual earthquakes or from specified ranges of magnitude and distance (Campbell, 1979, 1981a, 1981b, 1982a, 1982b, 1983, 1984c; Askins and Cornell, 1979; McCann, 1983); the use of a two-step regression procedure to separate the estimation of the distance and magnitude scaling coefficients (Joyner and Boore, 1981, 1982); a regression on Y rather than on ln Y to increase the impact of the larger values of Y (Bolt and Abrahamson, 1982); incorporating parameter uncertainty in the regression analysis (Bolt, 1978); segregating the data by distance or magnitude (Boore et al., 1980; Blume, 1980; Bolt and Abrahamson, 1982); and the use of a random effects model to separate the uncertainties associated with between-earthquake and within-earthquake variations (Brillinger and Preisler, 1984).

Of all the proposed procedures for handling data bias, the most traditional procedure statistically is weighted regression, where weights are assigned on the basis of data quality. Although the quality of an observation is usually based on some estimate of the accuracy or precision of the measurement, judgment is often used when many factors are known to contribute to the bias. For example, Mosteller and Tukey (1977) offer a broad interpretation of weighted regression in which weights are used to ensure the "high performance" in estimation one desires from regression models. It was in the spirit of this latter interpretation that Askins and Cornell (1979), Campbell (1981a), and McCann (1983) applied weighted regression to the analysis of strong-motion data as a means of reducing the bias associated with the uneven distribution of recordings with respect to individual earthquakes, magnitude, and distance.

At the present time (1984), there is considerable controversy concerning the choice of an appropriate procedure for reducing the bias associated with the distribution of strong-motion recordings. It is not possible to discuss at length all the strengths and weaknesses of the various procedures that have been proposed to date; suffice it to say that they have been elaborated at length at the various conferences and workshops held over the last few years. The notoriety of some of the more recent procedures (e.g. Campbell, 1981a; Joyner and Boore, 1981; Bolt and Abrahamson (1982); Brillinger and Preisler, 1984) has brought about an unusual amount of discussion regarding their use. The greatest controversy has involved the use of weighted regression. W. B. Joyner (personal communication, 1984) suggests that the weighting scheme proposed by Campbell (1981a) gives undue influence to singly recorded earthquakes. On the other hand, the writer believes that the two-step regression procedure employed by Joyner and Boore (1981) precludes an optimum fit of the data and seems to diminish the effect of magnitude saturation of strong-motion parameters at near-source distances. The regression on Y rather than on ln Y suggested by Bolt and Abrahamson (1982) has been observed to bias predictions of Y at small values of Y (Joyner and Boore, 1983), due primarily to a lack of homoscedasticity (K. W. Campbell, written communication, 1983). Brillinger and Preisler (1984) prefer the use of a random effects model, which they believe "handles the problem of weighting" by including individual random error terms for within-earthquake and between-earthquake variability. Research during the next few years should help to resolve these issues. For the time being, the selection of a procedure must be based on one's own judgment.

$$\hat{y} + z_{\alpha/2}\sigma \qquad (10)$$

where $z_{\alpha/2}$ is the standard normal variable associated with a cumulative probability of $1-\alpha/2$. This involves two assumptions: first, $t_{\alpha/2,\nu}$ is assumed to be equal to $z_{\alpha/2}$, valid only for a large number of degrees of freedom (say $\nu \geq 30$); second, $\sigma_{\hat{y}}$ is assumed to be zero, thus neglecting any uncertainty in the mean prediction of $\hat{y}$. This second assumption is approximately true only for predictions near the centroid of the data for which $\sigma_{\hat{y}} \approx \sigma/n^{0.5}$. For extrapolation of the model as is common practice in design applications, the uncertainty associated with the mean of $\hat{y}$ can be significant, making equation (10) an inappropriate representation of equation (8).

In addition to the uncertainties associated with the dispersion of the data (characterized by σ) and the uncertainties in $\hat{y}$ associated with the estimation of the b_i's (characterized by $\sigma_{\hat{y}}$), a third source of uncertainty arises from the assumed form of the regression model. This later uncertainty is not reflected in the confidence limits computed from equation (8). The better the functional form models the true state of nature (not necessarily how well it fits the data), the smaller this type of uncertainty will be. It can be largely avoided by using a nonparametric procedure together with EDA to establish optimum functions of the variables (Brillinger and Preisler, 1984), but this procedure precludes predictions which require extrapolation of the data base. If a model is required, then model uncertainty is best evaluated through the use of multiple functional forms, each having approximately the same goodness of fit (represented by similar multiple correlation coefficients and standard errors) with respect to the data.

ADEQUACY OF THE MODEL

The usefulness of the regression model may be evaluated by an F-test associated with an analysis of variance; however, the overall adequacy of the model is best assessed from an analysis of residuals (Draper and Smith, 1981). A residual is simply the difference between the observed and predicted values of y. Before analysis, it may be convenient to normalize the residuals to have a mean of zero and a standard deviation of unity. By the very nature of the regression analysis, the residuals will have a mean approximately equal to zero. If a weighted regression is used, then it will also be necessary to weight the residuals. Letting n equal the total number of observations used in the regression, the normalized weighted residual NWR_i for the i^{th} observation may be computed from the expression

$$NWR_i = \frac{[w_i(y_i - \hat{y})] - MWR}{\sigma} \tag{11}$$

where

$$MWR = \frac{1}{n} \sum_{i=1}^{n} w_i(y_i - \hat{y})$$

and

$$\sum_{i=1}^{n} w_i = n$$

In these expressions, w_i is the weight of the observation and MWR is the mean weighted residual. For unweighted analyses, one simply substitutes w_i=1 in the above expressions.

The first step in the analysis is to plot the NWR_i's (hereafter referred to simply as the residuals) versus the predicted value of y and the independent variables. Such a plot is shown in Figure 7. If no apparent trend in the residuals is observed in these plots, then the model can be considered adequate. A trend would indicate an inadequacy in the model to predict the data and would require modifying the functional form. Figure 8 gives an example of residuals that exhibit such trends. If trends appear to exist, then correlation analysis can be used to test the statistical significance of the trends, and EDA or some equivalent procedure can be used to modify the model if necessary [see Draper and Smith (1981) for a discussion of correlation analysis and Tukey (1977) for a description of Exploratory Data Analysis (EDA)].

The second step is to plot the residuals against parameters that were not formally included in the model but are suspected of having some effect on the independent variable. This is particularly important when iterative procedures, such as backward elimination or stepwise regression, are not used to statistically accept or reject parameters during the model building

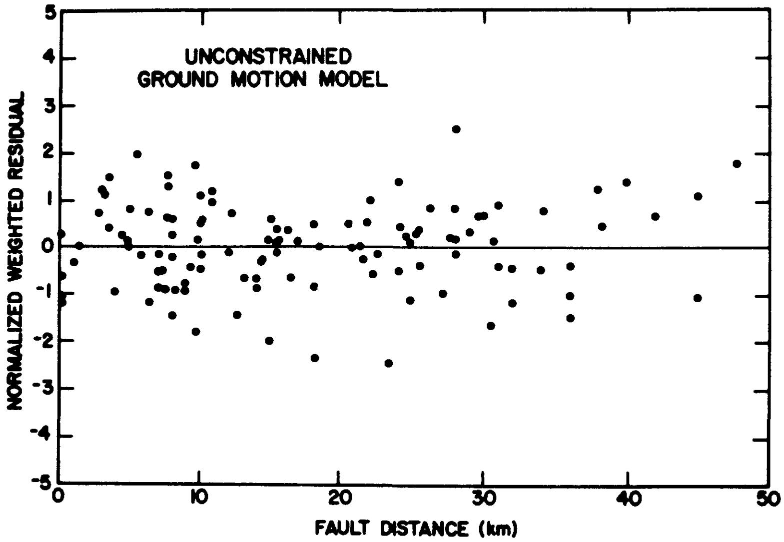


Figure 7. Normalized weighted residuals plotted as a function of distance for the least-squares regression analysis used to establish the attenuation relation in Figure 1. Figure taken from Campbell (1981a).

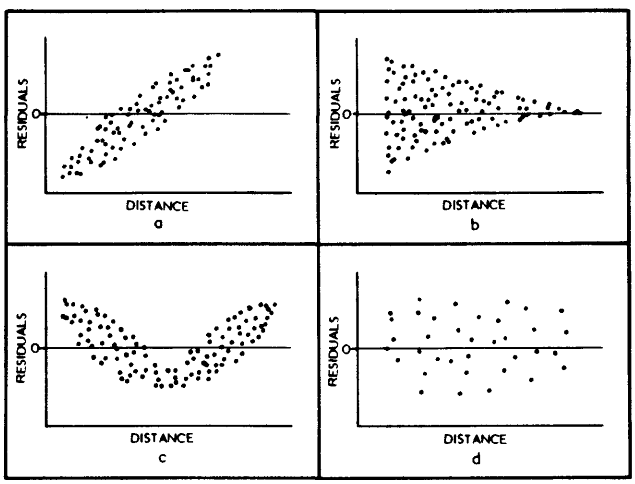


Figure 8. Examples of residual plots showing trends with distance: (a) linear trend suggesting $y = a + bR$; (b) lack of homoscedacity suggesting $y^{-b} = f(R)^{-b}$ or $\ln y = \ln [f(R)]$; (c) quadratic trend suggesting $y = a - bR + cR^2$; (d) no trend.

15B4 40008

process. As with independent variables, correlation analysis and EDA can be used to identify significant trends and establish the appropriate transformation of the parameter to include in the model. For parameters, such as site geology, that are not easily represented by continuous variables, a different approach is required. In this case subsets of residuals, selected on the basis of the class variable being investigated, can be plotted against the dependent and independent variables, and the plots inspected for two types of variation: (1) trends in the residual plots and (2) differences in trends between plots. The statistical significance of observed differences between plots can be established through hypothesis-testing techniques (e.g. Bowker and Liebermann, 1972; Freund and Walpole, 1980).

The application of hypothesis-testing techniques to the examination of residuals associated with regression analysis of strong-motion data is briefly described by Campbell and Davis (1981) and Campbell (1983). Rigorous application of the procedures requires several tests, depending on the observed trend of residuals. These include testing residuals or subsets of residuals for (1) differences in statistical correlation, (2) differences in mean values, (3) differences in variances, and (4) differences in medians. The first three are parametric tests, requiring that the residuals have a Normal distribution (tests for normality are described below). This requirement can be relaxed, however, if both samples are large enough for the central limit theorem to be invoked. The last test is nonparametric and does not require an assumption regarding the distribution of the residuals.

Application of the above procedures requires the formulation and testing of two alternative hypotheses. These procedures are best described using an example. Let us assume that we are interested in testing the significance in observed differences between the mean values of two subsets of residuals, representing, for example, soil and rock. The mean of each subset can be computed from the expression for MWR in equation (11), where n is replaced by the number of recordings in each subset, n_1 and n_2. The procedure involves testing the null hypothesis--that the two means are equal--against the alternative hypothesis--that they are different--by means of a t-test. If $|t| \geq t_{\alpha/2,\nu}$, where $|t|$ is the absolute value of the computed t-statistic, α is the acceptable level of significance, and $\nu = n_1 + n_2 - 2$ is the number of degrees

of freedom, then the null hypothesis can be rejected. However, if $|t|<t_{\alpha/2,\nu}$, then the null hypothesis cannot be rejected and differences in the two means cannot be considered statistically significant. The level of significance α is referre
of rejecting the null hypothesis when it is true. Acceptance of the null hypothesis when it is false is called a type II error; its probability is denoted by β. The tests are designed to minimize β; however, if β is found to be too large, then one must either increase α or increase the sample size in order to lower β. Campbell and Davis (1981) and Campbell (1983) adopted α=0.1 for their analyses, but the selected level of significance should be chosen to reflect the level of conservatism required in the results.

Trends or differences in residuals determined to be significant based on correlation analyses or hypothesis tests can be accommodated in several ways. Subsets found to be different can simply be removed from the data base and the regression repeated. This is appropriate when a subset represents a relatively small number of recordings. Alternatively, the data base can be segregated and a regression analysis performed on individual subsets. This approach is feasible when each subset consists of a significant number of recordings. Lastly, parameters can be added to the model to accommodate the observed trends in residuals and the analysis repeated to establish the new coefficients. In each case, the significance of all coefficients should be tested and removed from the model if they are not found to be significantly different from zero.

The analytical computation of confidence intervals for $\hat{y}$ requires that the coefficients and residuals have Normal distributions, although Monte Carlo simulation could be used to establish confidence intervals for any type of distribution. A qualitative assessment of normality may be obtained by inspecting a histogram of the residuals, like the one appearing in the inset of Figure 9. It should resemble the standard bell-shaped curve of the Normal distribution. A Kolmogorov-Smirnoff or Chi-square test may be used to statistically test the hypothesis that the distribution is Normal. Alternatively, a graphical procedure closely related to the Kolmogorov-Smirnoff test can be used. This procedure involves making a Normal Probability Plot--a plot of the normal score or estimate of the standard normal variable versus the normalized residual. If this plot (Figure 9) represents a straight line, then the residuals can be considered Normally distributed. Although this latter technique requires judgment on the part of the investigator, it does allow a more rigorous assessment than is possible from inspection of a histogram alone.

SEMI-EMPIRICAL METHODS

In many regions of the world, strong-motion recordings may be unavailable or extremely limited. For these regions, the development of strong-motion attenuation relations cannot rely on empirical procedures alone. One of the most common practices in such regions has been the prediction of ground motion from intensity, a qualitative measure of the severity of ground motion [e.g. see the description of the Modified Mercalli intensity scale in Richter (1958)]. This approach requires relationships between strong-motion parameters and intensity, such as those offered by Trifunac and Brady (1975a, b), Murphy and O'Brien (1977), Trifunac (1976c, 1978, 1979), Werner (1978), Krinitzsky and Marcuson (1983), and McGuire (1984). These can either be used in conjunction with an intensity attenuation relation (e.g. Howell and Schulz, 1975; Gupta and Nuttli, 1976; Anderson, 1979; Chandra, 1979; Atkinson, 1984; McGuire, 1984; Nuttli et al., 1984) or site-specific estimates of intensity to establish estimates of strong ground motion. McGuire (1977b), Cornell et al. (1979), and Bernreuter et al. (1984) describe the procedures and assumptions required for such an approach, and specific applications may be found in Nuttli and Herrmann (1978), Battis (1981), Bernreuter (1981a), Hasegawa et al. (1981), Nuttli et al. (1984), Atkinson (1984), and McGuire (1984).

Theoretical earthquake models can also be used to predict ground motion in regions where strong-motion recordings are limited. However, at present, such models are not commonly used for engineering applications due to their relative complexity and unknown reliability. These models fall into three basic categories. The first type uses kinematic and dynamic models of the fault-rupture process to generate deterministic predictions of ground motion. Swanger et al. (1980, 1981) and Aki (1982)

1556 40008

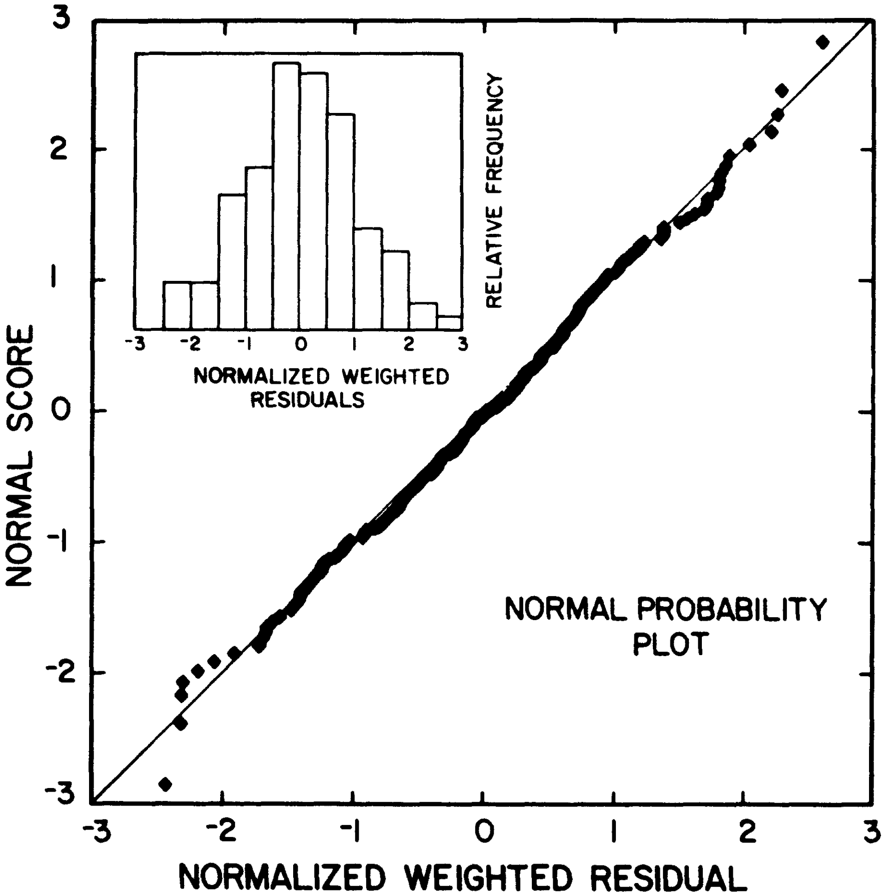


Figure 9. Normal probability plot for the normalized weighted residuals shown in Figure 7. The inset displays a histogram of the residuals. Figure taken from Campbell (1981a).

describe the characteristics of this type of model. The second category of theoretical models uses stochastic simulation of ground motions based on simple seismological source models (sometimes in conjunction with random vibration theory) to produce random predictions of strong ground motion. Most recent examples of this type of model are found in Joyner and Boore (1980), Hadley et al. (1982), Hanks and McGuire (1981), Boore (1983a), Gusev (1983), Atkinson (1984), McGuire et al. (1984), and Joyner (1984). The third type of model uses simple seismological source models (sometimes calibrated empirically) to deterministically predict strong ground motions. Because of its simplicity, this type of model has been most widely used to generate attenuation relations (e.g. Campbell and Duke, 1974a; Hanks and Johnson, 1976; Bureau, 1978; Hanks, 1979; Ang and Mohammadi, 1981; Bernreuter, 1981b; McGuire and Hanks, 1980; McGarr, 1981; Scholz, 1982; Nuttli and Herrmann, 1984). The reader is referred to Boatwright (1982), Boore (1983b), Luco and Anderson (1984), and Earthquake Engineering Research Institute (1984) for a comprehensive review and compilation of recent work on theoretical earthquake modelling.

Even if sufficient data are available with which to develop a strong-motion attenuation relation, it may still be desirable to constrain some of the coefficients of the model. This is particularly useful when specific coefficients are highly correlated with one another and neither can be determined accurately, or when the data are not distributed well enough to give robust estimates of some coefficients. Some investigators have simply preferred to use constraints in the development of their relationships to be consistent with well-established seismological principles.

The most common seismological constraints used in past regression studies are those related to geometrical and anelastic attenuation (e.g. Schnabel and Seed, 1973; Trifunac, 1976a, 1976b, 1978; Campbell, 1981b, 1982b; Joyner and Boore, 1981, 1982; Nuttli and Herrmann, 1984). Others have used constraints on magnitude scaling (e.g. Trifunac, 1976a, 1976b, 1978; Espinosa, 1979, 1980; Boore, 1980), while others have constrained coeffcients based on strong-motion recordings of nuclear explosions (Blume, 1977; Orphal and Lahoud, 1974) or have based constraints on the results of other empirical studies (Eguchi, 1980; Battis, 1981; Campbell, 1981a; Hasegawa et al., 1981). The most common and least-supported empirical constraint used to develop strong-motion attenuation relations has involved the coefficient b_5 in equations (3a) and (3b). Typically, values of 20 to 25 km have been assumed for this coefficient in order to control the amplitudes of strong-motion parameters at small distances (e.g. see Table 2). Only recently have values of $b_5 > 0$ been justified statistically (Campbell, 1981a, 1982a, 1984c; Boore and Joyner, 1982).

REVIEW OF EXISTING RELATIONSHIPS

There have been a vast number of strong-motion attenuation relations that have been proposed throughout the years. Table 2 contains a summary of some of the more significant relationships proposed within the last decade (i.e. from 1974 to 1984). The 10-year criterion is used to limit the number of relationships tabulated to those commonly used in practice. Attenuation relationships published before 1974 have generally been revised or have become obsolete due to the rapid advancement in the field of engineering seismology.

To further limit the number of attenuation relations summarized in Table 2, the compilation has been restricted to North American relationships that (1) predict peak acceleration, peak velocity, or some other single index of ground motion such as Arias intensity, r.m.s. acceleration, etc. (peak displacement is excluded for reasons specified below); (2) are available in the open literature--that is, in professional journals; and (3) are based at least in part on strong-motion data. The restriction to single indices is required to eliminate spectral values from the listing due to their large number of parameters (one for each period and damping). Peak displacements are not included because of their generally poor accuracy, resulting from errors in the record-processing procedures used to integrate and filter the accelerograms and from long-period noise inherent in the records themselves (Trifunac and Lee, 1978b; Sunder and Connor, 1982). The restriction to relationships published in the open literature limits the compilation to those relations generally available to engineers and seismologists and, thus, have had the

TABLE 2

SUMMARY OF SELECTED STRONG-MOTION ATTENUATION RELATIONS (1974-1984)
[Note: Definition of symbols appears in Appendix]

Reference	Parameters	Applicability	Y	Attenuation Relation		$\sigma_{\ln Y}$
Campbell and Duke	$M = M_S$	California;	AI_v	$Y = 3.31\times10^3\ e^{M(0.33M-1.47)}\ R^{-3}\ S$		
(1974a)	$R = R_c$	$M = 4.5\text{-}8.5$;		$S = 1.80$ (basement rock)		
		$R = 15\text{-}160$		3.63 (sedimentary rock)		
				3.74 (alluvium $\leq$60 ft)		
				5.12 (alluvium >60 ft)		
Campbell and Duke	$M = M_S$	California;	AI_v	$Y = 3.13\times10^4\ e^{M(0.33M-1.47)}\ R^{-3.79}\ S$		
(1974b)	$R = R_c$	$M = 4.5\text{-}8.5$;		$S = 0.57\ R^{0.46}$ (basement rock)		
		$R = 15\text{-}160$		$1.02\ R^{0.51}$ (sedimentary rock)		
				$0.37\ R^{0.81}$ (alluvium $\leq$60 ft)		
				$0.65\ R^{0.74}$ (alluvium >60 ft)		
Orphal and Lahoud	$M = M_R$	California;	PHA_L	$Y = 0.066\ e^{0.92M}\ R^{-1.39}$		0.69
(1974)	$R = R_h$	$M = 4.1\text{-}7.0$;				
		$R = 30\text{-}350$	PHV_L	$Y = 0.726\ e^{1.20M}\ R^{-1.34}$		0.64
Seed *et al.* (1976b)	$M = M_L$	western U.S.;	PHA_b	$Y = aR^{-c}$	(graphical)	
	$R = R_z$	$M = 6.5$;				
		$R = 20\text{-}350$	PHV_b	$Y = aR^{-c}$	(graphical)	
Trifunac (1976b)	$M = M_R$	Western U.S.;	PHA_b, PVA	$Y = Y_oA_oe^{2.3M}$		0.71
	$R = R_e$	$M = 3.0\text{-}7.7$;		$Y_o = 1.741\times10^{-9}\ e^{-0.138S}\ e^{-0.762V}\ F(M)$		
		$R = 20\text{-}200$		$F(M) = e^{4.12M}\ e^{-0.428[M^2-(M-7.5)^2]}$	$(M\geq7.5)$	
				$e^{4.12M}\ e^{-0.428M^2}$	$(7.5>M\geq4.8)$	
				2.023×10^4	$(M\leq4.8)$	
			PHV_b, PVV	$Y = Y_oA_oe^{2.3M}$		0.85
				$Y_o = 1.536\times10^{-8}\ e^{-0.309S}\ e^{-0.792V}\ F(M)$		
				$F(M) = e^{4.74M}\ e^{-0.463[M^2-(M-7.61)^2]}$	$(M\geq7.61)$	
				$e^{4.74M}\ e^{-0.463M^2}$	$(7.61\geq M\geq5.12)$	
				1.856×10^5	$(M\leq5.12)$	
				$S = 0$ (alluvium)		
				1 (intermediate rock)		
				2 (basement rock)		
				$V = 0$ (horizontal components)		
				1 (vertical components)		
McGuire (1977a)	$M = M_R$	Western U.S.;	PHA_b	$Y = 0.482\ e^{0.64M}\ (R+25)^{-1.30}$		0.51
	$R = R_h$	$M = 5.3\text{-}7.6$;				
		$R = 15\text{-}125$;	PHV_b	$Y = 5.640\ e^{0.92M}\ (R+25)^{-1.20}$		0.63
Donovan and Bornstein	$M = M_R$	California;	PHA_b	$Y = A\ e^{BM}\ (R+25)^C$		Graphical
(1978)	$R = R_c$	$M = 5.0\text{-}7.7$;		$A = 2.198\times10^3\ R^{-2.1}$		
		$R = 5\text{-}320$		$B = 0.046 + 0.193 \ln R$		
		Rock and stiff soil		$C = -2.515 + 0.211 \ln R$		
McGuire (1978b)	$M = M_R$	Western U.S.;	PHA_b	$Y = 0.0306\ e^{0.89M}\ R^{-1.17}\ e^{-0.20S}$		0.62
	$R = R_h$	$M = 4.5\text{-}7.7$;	PHV_b	$Y = 0.3680\ e^{1.07M}\ R^{-0.96}\ e^{0.07S}$		0.64
		$R = 10\text{-}200$		$S = 0$ (rock)		
				1 (soil)		
Cornell *et al.* (1979)	$M = M_R$	Western U.S.;	PHA_r	$Y = 0.863\ e^{0.86M}\ (R+25)^{-1.80}$		0.57
	$R = R_h$	$M = 3.0\text{-}7.7$;				
		$R = 20\text{-}200$;	PHV_r	$Y = 8.248\ e^{1.07M}\ (R+25)^{-1.55}$		0.64
Espinosa (1979)	$M = M_L$	Western U.S.;	PHV_b	$Y = 6.17\times10^{-4}\ e^{2.3M}\ R^{-1.35}$		
	$R = R_e$	$M = 4.0\text{-}7.2$;		$Y = 1.17\times10^{-4}\ e^{2.3M}\ R^C$		
		$R = 5\text{-}300$		$C = -0.28 - 0.16 \ln R$		
Boore *et al.* (1980)	$M = M_L$	Western N.A.;	PHA_L	$Y = 1.58\ R^{-0.9}$	$(M_L=5.0\text{-}5.7,\ R=5\text{-}30)$	Graphical
	$R = R_f$	Soil and rock		$Y = 10.0\ R^{-1.2}$	$(M_L=6.0\text{-}6.4,\ R=15\text{-}55)$	
				$Y = 398\ R^{-2.0}$	$(M_L, M_S=7.1\text{-}7.6,\ R=40\text{-}150)$	
			PHV_L	$Y = 251\ R^{-1.2}$	$(M_L=5.3\text{-}5.7,\ R=5\text{-}30)$	
				$Y = 79.4\ R^{-0.6}$	$(M_L=6.4,\ R=15\text{-}55)$	

1553 40008

TABLE 2

SUMMARY OF SELECTED STRONG-MOTION ATTENUATION RELATIONS (1974-1984)--continued

Reference	Parameters	Applicability	Y	Attenuation Relation		$\sigma_{\ln Y}$
Espinosa (1980)	$M = M_L$ $R = R_e$	Western U.S.; $M = 4.0-7.5$; $R = 5-300$	PHA_b	$Y = 5.235\times10^{-7}\, e^{2.3M}\, R^{-0.06}$ $Y = 1.776\times10^{-5}\, e^{2.3M}\, R^{-1.59}$ $Y = 4.153\times10^{-3}\, e^{2.3M}\, R^{-2.93}$ $Y = 1.119\times10^{-6}\, e^{2.3M}\, R^{C}$ $C = -0.11 - 0.22 \ln R$	$(R \le 10)$ $(10 \le R \le 60)$ $(60 \le R \le 300)$ $(5 \le R \le 300)$	
Battis (1981)	$M = m_b$ $R = R_e$	$M = 5.0-6.5$; $R = 10-350$	PHA_b	$Y = 0.3480\, e^{1.21M}\, (R+25)^{-2.08}$ $Y = 0.0239\, e^{1.24M}\, (R+25)^{-1.24}$	(California) (Central U.S.)	0.71 0.71
Campbell (1981a)	$M = M_L$ $(M<6)$ $M = M_S$ $(M \ge 6)$ $R = R_f$	Worldwide; $M = 5.0-7.7$; $R \le 50$; Rock; Soil >10 m deep	PHA_m	$Y = 0.0159\, e^{0.868M}\, [R + C(M)]^{-1.09}$ $C(M) = 0.0606\, e^{0.7M}$		0.37
Hanks and McGuire (1981)	$M = M_L$ $R = R_h$	California; $M = 4.0-6.5$; $R = 10-100$	$RMSA_b$ PHA_b	$Y = 0.119\, R^{-1} \left[\frac{f_{max}}{f_o}\right]^{0.5}$ $Y = 0.119\, R^{-1} \left[\frac{2f_{max}}{f_o} \ln\left(\frac{2f_{max}}{f_o}\right)\right]^{0.5}$		
Hasegawa *et al.* (1981)	$M = M_L$ (W) $M = m_b$ (E) $R = R_h$	Canada; $M = 4.0-7.0$; $R = 10-200$	PHA_b PHV_b	$Y = 1.02\times10^{-2}\, e^{1.3M}\, R^{-1.5}$ $Y = 3.47\times10^{-3}\, e^{1.3M}\, R^{-1.1}$ $Y = 4.00\times10^{-4}\, e^{2.3M}\, R^{-1.3}$ $Y = 1.80\times10^{-4}\, e^{2.3M}\, R^{-1.0}$	(W. Canada) (E. Canada) (W. Canada) (E. Canada)	
Joyner and Boore (1981)	$M = M$ $R = R_s$	Western N.A.; $M = 5.0-7.7$; $R \le 370$; Small structures; Soil and rock	PHA_L PHV_L	$Y = 0.0955\, e^{0.573M}\, D^{-1}\, e^{-0.00587D}$ $D = (R^2 + 7.3^2)^{0.5}$ $Y = 0.214\, e^{1.13M}\, D^{-1}\, e^{-0.0059D}\, e^{0.39S}$ $D = (R^2 + 4.0^2)^{0.5}$ $S = 0$ (rock) 1 (soil)		0.60 0.51
Bolt and Abrahamson (1982)	$M = M$ $R = R_s$	Western N.A.; $R \le 370$; Small structures; Soil and rock	PHA_L	$Y = 1.20\, [(R+23)^2 + 1]^{0.033}\, e^{-0.066(R+23)}$ $Y = 1.20\, [(R+25)^2 + 1]^{0.042}\, e^{-0.044(R+25)}$ $Y = 0.24\, [(R+15)^2 + 1]^{0.100}\, e^{-0.022(R+15)}$	$(M=5.0-5.9)$ $(M=6.0-6.9)$ $(M=7.0-7.7)$	0.06g (σ_y) 0.10g (σ_y) 0.05g (σ_y)
Nuttli and Herrmann (1984)	$M = m_b$ $R = R_e$	Mississippi Valley; $M = 4.5-7.5$; $R = 3-200$	PHA_m PHV_m	$Y = 3.79\times10^{-3}\, e^{1.15M}\, D^{-0.83}\, e^{-0.00159R}$ $Y = 2.51\times10^{-4}\, e^{2.30M}\, D^{-0.83}\, e^{-0.00076R}$ $D = (R^2 + H_{min}^2)^{0.5}$ $H_{min} = 0.0186\, e^{1.05M}$		0.55 0.55

1587 40008

opportunity of being subjected to peer review and acceptance. That is not to say that those relationships omitted from tabulation are not of equal or even greater value, only that such relationships are generally not widely known and have not had as much opportunity for peer review. The restriction to relationships based on strong-motion data merely requires that the relationships have at least some empirical basis, eliminating the large number of theoretical models that have been proposed recently. A more complete listing of strong-motion attenuation relations for the last 10 years, including those available in reports, appears in the Bibliography. Also included in the Bibliography are references to some attenuation relations proposed by foreign investigators within the last few years (1983-1984). Additional compilations and bibliographies are available in Donovan (1973), Hofman (1974), Trifunac and Brady (1975c), Idriss (1978), Eguchi and Wiggins (1979), Hays (1980), Young (1980a, b), Boore and Joyner (1982), Boore (1983b), and Campbell (1984e).

The summary provided in Table 2 is organized into six categories: (1) a reference, (2) a definition of parameters, (3) a statement of applicability, (4) a description of the strong-motion parameter being predicted, (5) the attenuation relation, and (6) the standard error. The table is provided only as a summary of those relationships that have been developed in the past 10 years. Specific relationships should not be used without careful consideration of the application for which they are intended. The guidelines for developing attenuation relations presented earlier in this paper can serve as a framework to be used in evaluating these existing relationships for specific applications.

ACKNOWLEDGMENTS

This paper was originally developed as part of a U.S. Geological Survey course on Geologic and Hydrologic Hazards held in March 1984 in Denver, Colorado. The original concept was conceived by Dave Boore and Bill Joyner, to whom I am greatly indebted. I also wish to thank Bill Joyner and Paul Thenhaus for their helpful suggestions.

APPENDIX--GLOSSARY OF TERMS

Term	Definition
A_o	Standard distance attenuation factor for computing M_L (Richter, 1958)
AI	Arias intensity (cm/sec)
b	Subscript denoting the use of both horizontal components
b_i	Regression coefficients
C	Covariance matrix of regression coefficients
$e^{(*)}$	Exponential of (*)
f(*)	Functional of parameter (*)
f_o	Spectral corner frequency (Hz)
f_{max}	Maximum spectral frequency (Hz)
H_{min}	Minimum focal depth (km)
ln(*)	Natural logarithm of (*)
L	Subscript denoting use of maximum horizontal component
m	Subscript denoting use of mean horizontal component
m_b	Short-period body-wave magnitude
m_B	Long-period body-wave magnitude
M	Earthquake magnitude (generic)
M	Moment magnitude
M_{JMA}	Japan Meteorological Agency magnitude
M_L	Local magnitude
M_R	Richter magnitude
M_s	Surface-wave magnitude
MWR	Mean weighted residual
n	Number of recordings used in regression analysis
n_o	Number of observations for which confidence limits on $\hat{y}$ are desired
n_1,n_2	Number of recordings in subsets of residuals
p	Number of independent variables in regression model
P_i	Parameter representing earthquake, site or structure effects
PHA	Peak horizontal acceleration (g)
PHV	Peak horizontal velocity (cm/sec)
PVA	Peak vertical acceleration (g)
PVV	Peak vertical velocity (cm/sec)
r	Subscript denoting use of random horizontal component
R	Distance (generic, km)
R_c	Distance to center of energy release (km)
R_e	Epicentral distance (km)
R_f	Closest distance to fault rupture (km)
R_h	Hypocentral distance (km)
R_s	Closest distance to surface projection of fault rupture (km)
R_z	Closest distance to zone of energy release (km)
RMSA	Root-mean-square acceleration (g)
S	Site classification variable
$t_{\alpha/2,\nu}$	t-statistic associated with a level of significance α and ν degrees of freedom
$\|t\|$	Computed absolute value of t-statistic

1552 40008

v	Subscript denoting use of vectoral horizontal component
V	Strong-motion component variable
V_s	Shear-wave velocity
w_i	Weight associated with y_i
$\mathbf{X}_o$	Vector containing specific values of independent variables
$\mathbf{X}'_o$	Transpose of $\mathbf{X}_o$
y	Transformed value of Y (e.g. y=ln Y)
$\hat{y}$	Predicted value of y (i.e. from a regression model)
y_i	Observed value of y
Y	Strong-motion parameter being predicted (i.e. dependent variable)
$z_{\alpha/2}$	Standard normal variable associated with cumulative probability $1-\alpha/2$
α	Level of significance or type I error ($1-\alpha$ is level of confidence)
β	Type II error
ε	Random error term in regression model
ε'	Natural logarithm of ε ($\varepsilon'=\ln \varepsilon$)
ν	Number of degrees of freedom
Π	Symbol denoting multiplication operation
σ	Standard error of estimate of regression
$\sigma_{\ln Y}$	Standard deviation of ln Y
$\sigma_{\hat{y}}$	Standard deviation of $\hat{y}$
Σ	Symbol denoting summation operation

BIBLIOGRAPHY

Aki, K. (1982). Strong motion prediction using mathematical modeling techniques, Bull. Seism. Soc. Am., **72**, S29-S41.

Anderson, J. G. (1979). On the attenuation of Modified Mercalli intensity with distance in the United States, Bull. Seism. Soc. Am., **68**, 1147-1179.

Anderson, J. G., and J. E. Luco (1983). Parametric study of near-field ground motions for oblique-slip and dip-slip dislocation models, Bull. Seism. Soc. Am., **73**, 45-57.

Ang, H-S., and J. Mohammadi (1981). Development of an attenuation equation for near-source regions, in Proceedings of Review Meeting of U.S.-Japan Cooperative Research on Seismic Risk Analysis and Its Application to Reliability-Based Design of Lifeline Systems, Honolulu, Hawaii, 1981, 1-22.

Applied Technology Council (1974). An evaluation of a response spectrum approach to seismic design of buildings, Applied Technology Council Rept. ATC-2, Berkeley, California.

Applied Technology Council (1978). Tentative provisions for the development of seismic regulations for buildings, Applied Technology Council Rept. ATC-3-06, Berkeley, California.

Applied Technology Council (1984). Proceedings of seminar and workshop on Earthquake Ground Motion and Building Damage Potential, San Francisco, California, 1984, Applied Technology Council Rept. ATC-10-1, Palo Alto, California.

Arnold, P., and E. H. Vanmarcke (1977). Ground motion spectral content: the influence of local soil conditions and site azimuth, in Proceedings of Sixth World Conference on Earthquake Engineering, New Delhi, India, 1977, **I**, 445-451.

Askins, R. C., and C. A. Cornell (1979). SHA-based attenuation model parameter estimation, in Proceedings of Second U.S. National Conference on Earthquake Engineering--1975, Stanford, California, 1979, 703-712.

Atkinson, G. M. (1984). Attenuation of strong ground motion in Canada from a random vibrations approach, Bull. Seism. Soc. Am., **74**, 2629-2653.

Battis, J. (1981). Regional modification of acceleration functions, Bull. Seism. Soc. Am., **71**, 1309-1321.

Bernreuter, D. L. (1981a). Seismic hazard analysis--Application of methodology, results, and sensitivity studies, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-1582, **4**.

Bernreuter, D. L. (1981b). Scaling and estimation of earthquake ground motion as a function of the earthquake source parameters and distance, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-2103.

Bernreuter, D. L., J. B. Savy, R. W. Mensing, and D. H. Chung (1984). Development of Eastern United States ground motion models, Appendix C in Seismic Hazard Characterization of the Eastern United States: Methodology and Interim Results for Ten Sites, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-3756, C1-C114.

Berrill, J. B. (1975). A study of high-frequency ground motion from the San Fernando earthquake, Ph. D. thesis, California Institute of Technology, Pasadena, California.

Blume, J. A. (1977). The SAM procedure for site-acceleration-magnitude relationships, in Proceedings of the Sixth World Conference on Earthquake Engineering, New Delhi, India, 1977, **I**, 416-422.

Blume, J. A. (1980). Distance partitioning in attenuation studies, in Proceedings of Seventh World Conference on Earthquake Engineering, Istanbul, Turkey, 1980, **2**, 403-410.

Boatwright, J. (1982). Proceedings of workshop XVI--The dynamic characteristics of faulting inferred from recordings of strong ground motion, J. Boatwright, Editor, Incline Village, California, 1981, U.S. Geol. Surv. Open-File Rept. 82-591.

Boatwright, J. A., and D. M. Boore (1982). Analysis of the ground accelerations radiated by the 1980 Livermore Valley earthquakes for directivity and dynamic source characteristics, Bull. Seism. Soc. Am., **72**, 1843-1865.

Bolt, B. A. (1978). Fallacies in current ground motion prediction, in Proceedings of Second International Earthquake Microzonation Conference, San Francisco, California, 1979, **II**, 617-634.

Bolt, B. A. (1981). Interpretation of strong ground motion records, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 17, Vicksburg, Mississippi.

Bolt, B. A., and N. A. Abrahamson (1982). New attenuation relations for peak and expected acclerations of strong ground motion, Bull. Seism. Soc. Am., **72**, 2307-2321.

Bolt, B. A., and N. A. Abrahamson (1983). Reply to W. B. Joyner and D. M. Boore's "Comments on 'New attenuation relations for peak and expected accelerations of strong ground motion,'" Bull. Seism. Soc. Am., **73**, 1481-1483.
Bolt, B. A., and M. Herraz (1983). Simplified estimation of seismic moment from seismograms, Bull. Seism. Soc. Am., **73**, 735-748.
Boore, D. M. (1973). The effect of simple topography on seismic waves--Implications for the accelerations recorded at Pacoima Dam, San Fernando, California, Bull. Seism. Soc. Am., **63**, 1603-1609.
Boore, D. M. (1980). On the attenuation of peak velocity, in Proceedings of the Seventh World Conference on Earthquake Engineering, Istanbul, Turkey, 1980, **II**, 577-584.
Boore, D. M. (1983a). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, Bull. Seism. Soc. Am., **73**, 1865-1894.
Boore, D. M. (1983b). Strong-motion seismology, in U.S. National Report to International Union of Geodesy and Geophysics, 1979-1982, Rev. Geophys. and Space Phys., **21**, 1308-1318.
Boore, D. M., and W. B. Joyner (1982). The empirical prediction of ground motion, Bull. Seism. Soc. Am., **72**, S43-S60.
Boore, D. M., W. B. Joyner, A. A. Oliver, III, and R. A. Page (1980). Peak acceleration, velocity, and displacement from strong-motion records, Bull. Seism. Soc. Am., **70**, 305-321.
Borcherdt, R. D., and J. F. Gibbs (1976). Effects of local geological conditions in the San Francisco Bay region on ground motions and the intensities of the 1906 earthquake, Bull. Seism. Soc. Am., **66**, 467-500.
Bowker, A. H., and G. J. Liebermann (1972). Engineering statistics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
Brillinger, D. R., and H. K. Preisler (1984). An exploratory analysis of the Joyner-Boore attenuation data, Bull. Seism. Soc. Am., **74**, 1441-1450.
Bureau, G. J. (1978). Influence of faulting on earthquake attenuation, in Proceedings of ASCE Specialty Conference on Earthquake Engineering and Soil Dynamics, Pasadena, California, 1978, **I**, 290-307.
Bureau, G. J. (1981). Near-source peak ground acceleration [abstract], Earthquake Notes, **52**, 81.
Bycroft, G. N. (1977). Soil-structure interaction at higher frequency factors, J. Earthquake Eng. Struct. Dyn., **5**, 235-248.
Bycroft, G. N. (1978). The effect of soil-structure interaction on seismometer readings, Bull. Seism. Soc. Am., **68**, 828-843.
Campbell, K. W. (1979). Preliminary evaluation of near-source attenuation of peak acceleration in the United States [abstract], Earthquake Notes, **50**, 16-17.
Campbell, K. W. (1981a). Near-source attenuation of peak horizontal acceleration, Bull. Seism. Soc. Am., **71**, 2039-2070.
Campbell, K. W. (1981b). A ground motion model for the Central United States based on near-source acceleration data, in Proceedings of Conference on Earthquakes and Earthquake Engineering: The Eastern U.S., Knoxville, Tennessee, 1981, **1**, 213-232.

1688 40008

Campbell, K. W. (1982a). Near-source scaling characteristics of peak horizontal acceleration for moderate-to-large earthquakes, in Proceedings of Workshop XVI--The Dynamic Characteristics of Faulting Inferred From Recordings of Strong Ground Motion, Incline Village, California, 1981, U.S. Geol. Surv. Open-File Rept. 82-591, **1**, 120-184.

Campbell, K. W. (1982b). A preliminary methodology for the regional zonation of peak acceleration, in Proceedings of Third International Conference on Earthquake Microzonation, Seattle, Washington, 1982, I, 365-376.

Campbell, K. W. (1982c). A study of the near-source behavior of peak vertical acceleration (abstract), EOS, **63**, 1037.

Campbell, K. W. (1983). The effects of site characteristics on near-source recordings of strong ground motion, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 280-309.

Campbell, K. W. (1984a). An empirical assessment of near-source strong ground motion for a 6.6 m_b (7.5 M_s) earthquake in the eastern United States, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-3839.

Campbell, K. W. (1984b). Observed structural modification of recorded strong ground motion, in Proceedings of Seminar and Workshop on Earthquake Ground Motion and Building Damage Potential, San Francisco, California, 1984, Applied Technology Council Rept. ATC-10-1, Palo Alto, California.

Campbell, K. W. (1984c). Near-source attenuation of strong ground motion for moderate to large earthquakes--an update and suggested application to the Wasatch fault zone of north-central Utah, in Proceedings of Workshop on Evaluation of Regional and Urban Earthquake Hazards and Risk in Utah, Salt Lake City, Utah, 1984, U.S. Geol. Surv. Open-File Rept. 84-763, 483-499.

Campbell, K. W. (1984d). Strong ground motion data and its use in earthquake engineering, in Proceedings of Conference on Pathways and Future Directions for Environmental Data and Information Users, Denver, Colorado, 1984, 240-261.

Campbell, K. W. (1984e). Selection and development of strong motion scaling relationships for seismic hazard studies, in Proceedings of Course on Geologic and Hydrologic Hazards, Denver, Colorado, 1984, U.S. Geol. Survey Open-File Rept. 84-760, 28-82.

Campbell, K. W., and B. J. Davis (1981). Statistical analysis of earthquake ground motion characteristics [abstract], Earthquake Notes, **52**, 19.

Campbell, K. W., and C. M. Duke (1974a). Bedrock intensity attenuation and site factors from San Fernando earthquake records, Bull. Seism. Soc. Am., **64**, 173-185.

Campbell, K. W., and C. M. Duke (1974b). A reply to a discussion on bedrock intensity attenuation and site factors from San Fernando earthquake records by P. C. Jennings, Bull. Seism. Soc. Am., **64**, 2009-2010.

Campbell, K. W., and J. R. Murphy (1983). Report and recommendations of special study group three, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 19-22.

Campbell, K. W., and M. Niazi (1982). Evidence for saturation of peak acceleration in the near field [abstract], Earthquake Notes, **53**, 91.

Chandra, V. (1979). Attenuation of intensitites in the United States, Bull. Seism. Soc. Am., **69**, 2003-2024.

Chang, F. K. (1980). The effects of elevation and site conditions on ground motion of the San Fernando, California, earthquake, 9 February 1971, in Proceedings of Seventh World Conference on Earthquake Engineering, Istanbul, Turkey, 1980, **2**, 207-214.

Chang, F. K., and E. L. Krinitzsky (1977). Duration, spectral content, and predominant period of strong motion earthquake, records from western United States, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 8, Vicksburg, Mississippi.

Chiaruttini, C., and L. Siro (1981). The correlation of peak ground horizontal acceleration with magnitude, distance, and seismic intensity for Friuli and Ancona, Italy, and the Alpide Belt, Bull. Seism. Soc. Am., **71**, 1993-2009.

Chung, D. H., and D. L. Bernreuter (1981). Regional relationships among earthquake magnitude scales, Rev. Geophys. and Space Phys., **19**, 649-663.

Cornell, C. A., H. Banon, and A. F. Shakal (1979). Seismic motion and response prediction alternatives, J. Earthquake Eng. Struct. Dyn., **7**, 295-315.

Crouse, C. B. (1976). Horizontal ground motion in Los Angeles during the San Fernando earthquake, J. Earthquake Eng. Struct. Dyn., **4**, 333-347.

Crouse, C. B. (1978). Prediction of free-field earthquake ground motions, in Proceedings of ASCE Specialty Conference on Earthquake Engineering and Soil Dynamics, Pasadena, California, 1978, I, 359-379.

Crouse, C. B. (1983). Soil-structure interaction effects of accelerograph stations, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 357-360.

Crouse, C. B., and B. E. Turner (1978). Analysis of ground motion spectra, in Proceedings of Second International Earthquake Microzonation Conference, San Francisco, California, 1978, **II**, 1117-1132.

Crouse, C. B., G. C. Liang, and G. R. Martin (1984). Experimental study of soil-structure interaction at an accelerograph station, Bull. Seism. Soc. Am., **74**, 1995-2013.

Davis, L. L., and L. R. West (1973). Observed effects of topography on ground motion, Bull. Seism. Soc. Am., **63**, 283-298.

1658 4000 8

Donovan, N. C. (1974). A statistical evaluation of strong motion data, including the February 9, 1971, San Fernando earthquake, in Proceedings of the Fifth World Conference on Earthquake Engineering, Rome, Italy, 1973, **1**, 1252-1261.

Donovan, N. C. (1982a). Strong motion attenuation equations--a critique, in Proceedings of Third International Earthquake Microzonation Conference, Seattle, Washington, 1982, **I**, 377-388.

Donovan, N. C. (1982b). Attenuation of vertical acceleration and a review of attenuation equation processes, in Proceedings of Workshop XVI--The Dynamic Characteristics of Faulting Inferred from Recordings of Strong Ground Motion, Incline Village, California, 1981, U.S. Geol. Surv. Open-File Rept. 82-591, 185-202.

Donovan, N. C., and A. E. Bornstein (1978). Uncertainties in seismic risk procedures, J. Geotech. Engrng. Div., ASCE, **104**, 869-887.

Draper, N. R., and H. Smith (1981). Applied regression analysis, second edition, John Wiley and Sons, New York.

Duke, C. M., K. E. Johnsen, L. E. Larson, and D. C. Engman (1972). Effects of site classification and distance on instrumental indices in the San Fernando earthquake, University of California School of Engineering and Applied Science Rept. UCLA-ENG-7247, Los Angeles, California.

Earthquake Engineering Research Institute (1984). Proceedings of Workshop on Strong Ground Motion Simulation and Earthquake Engineering Applications, Los Altos, California, 1984, Earthquake Engineering Research Institute, Berkeley, California.

Eguchi, R. T. (1980). An alternative approach to modeling earthquake ground motion attenuation in the Western United States, in Proceedings of the Seventh World Conference on Earthquake Engineering, Istanbul, Turkey, 1980, **2**, 113-120.

Eguchi, R. T., and J. H. Wiggins (1979). A Bayesian seismic risk study of California, with loss estimates, J. H. Wiggins Company Technical Rept. 79-1328-1, Redondo Beach, California.

Espinosa, A. F. (1979). Horizontal partical velocity and its relation to magnitude in the western United States, Bull. Seism. Soc. Am., **69**, 2037-2061.

Espinosa, A. F. (1980). Attenuation of strong horizontal ground accelerations in the western United States and their relation to M_L, Bull. Seism. Soc. Am., **70**, 583-616.

Esteva, L. (1970). Seismic risk and seismic design decisions, in Seismic Design for Nuclear Power Plants, R. J. Hanson, Editor, M.I.T. Press, Cambridge, Masachusetts, 142-182.

Fallgren, R. B., P. C. Jennings, J. L. Smith, and D. K. Ostrom (1974). Aseismic design criteria for electrical facilities, J. Power Div., ASCE, **100**, 1-14.

Faccioli, E. (1981). The use of RMS acceleration as a measure of soil amplification in strong earthquakes, in State-of-the-Art in Earthquake Engineering, 1981, Erguray, O., and M. Erdik, Editors, Turkish National Committee on Earthquake Engineering, Ankara, Turkey.

Freund, J. E., and R. E. Walpole (1981). Mathematical statistics, third edition, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Gallant, A. R. (1975). Nonlinear regression, The American Statistician, **29**, 73-81.

Gupta, I. N., and O. W. Nuttli (1976). Spatial attenuation of intensities for central U.S. earthquakes, Bull. Seism. Soc. Am., **66**, 743-751.

Gusev, A. A. (1983). Descriptive statistical model of earthquake source radiation and its application to an estimation of short-period strong ground motion, Geophys. J. Roy. Astr. Soc., **74**, 787-808.

Guzman, R. A., and P. C. Jennings (1976). Design spectra for nuclear power plants, J. Power Div., ASCE, **102**, 165-178.

Hadley, D. M., D. V. Helmberger, and J. A. Orcutt (1982). Peak acceleration scaling studies, Bull. Seism. Soc. Am., **72**, 959-979.

Hanks, T. C. (1975). Strong ground motion of the San Fernando, California, earthquake: Ground displacements, Bull. Seism. Soc. Am., **65**, 193-226.

Hanks, T. C. (1979). b values and $\omega^{-\gamma}$ seismic source models: implications for tectonic stress variations along active crustal fault zones and estimation of high frequency strong ground motion, J. Geophys. Res., **84**, 2235-2241.

Hanks, T. C., and D. A. Johnson (1976). Geophysical assessment of peak accelerations, Bull. Seism. Soc. Am., **66**, 959-968.

Hanks, T. C., and H. Kanamori (1979). A moment magnitude scale, J. Geophys. Res., **84**, 2348-2350.

Hanks, T. C., and R. K. McGuire (1981). The character of high-frequency strong ground motion, Bull. Seism. Soc. Am., **71**, 2071-2095.

Haroun, M. A., and G. W. Housner (1982). Characteristics of ground motion of the San Fernando earthquake, in Proceedings of Third International Earthquake Microzonation Conference, Seattle, Washington, **I**, 389-400.

Hasegawa, H. S., P. W. Basham, and M. J. Berry (1981). Attenuation relations for strong seismic ground motion, Bull. Seism. Soc. Am., **71**, 2071-2095.

Hays, W. W. (1980). Procedures for estimating earthquake ground motions, U.S. Geol. Surv. Profess. Paper 1114.

Heaton, T. H., F. Tajima, and A. W. Mori (1984). Estimating ground motions using recorded accelerograms, unpublished manuscript.

Herrmann, R. B., and O. W. Nuttli (1982). Magnitude: The relation of M_L to m_{bLg}, Bull. Seism. Soc. Am., **72**, 389-397.

Hofman, R. B. (1974). Factors in the specification of ground motions for design earthquakes in California, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 3, Vicksburg, Mississippi.

Housner, G. W. (1971). Strong ground motion, in Earthquake Engineering, R. L. Wiegel, Editor, Prentice-Hall, Englewood Cliffs, New Jersey, 93-106.

Housner, G. W., and P. C. Jennings (1982). Earthquake design criteria, Earthquake Engineering Research Institute Monograph, Berkeley, California.

Howell, B. F., and T. R. Schulz (1975). Attenuation of modified Mercalli intensity with distance from the epicenter, Bull. Seism. Soc. Am., **65**, 651-665.

Huang, M., M. A. Haroun, and G. W. Housner (1982). A study on strong motion data of the 1979 Imperial Valley earthquake, in Proceedings of Third International Earthquake Microzonation Conference, Seattle, Washington, **I**, 401-412.

Hudson, D. E. (1979). Reading and interpreting strong motion accelerograms, Earthquake Engineering Research Institute Monograph, Berkeley, California.

Idriss, I. M. (1978). Characteristics of earthquake ground motions, in Proceedings of ASCE Specialty Conference on Earthquake Engineering and Soil Dynamics, Pasadena, California, **III**, 1151-1265.

Idriss, I. M. (1983). Site response and design ground motions, in Proceedings of EERI Seminar on Evaluation of Seismic Hazards, and Decision Making in Earthquake Resistant Design, Reno, Nevada, 1983, Earthquake Engineering Research Institute, Berkeley, California.

Joyner, W. B. (1984). A scaling law for the spectra of large earthquakes, Bull. Seism. Soc. Am., **74**, 1167-1188.

Joyner, W. B., and D. M. Boore (1980). A stochastic source model for synthetic strong-motion seismograms, in Proceedings of Seventh World Conference on Earthquake Engineering, Istanbul, Turkey, 1980, **1**, 1-8.

Joyner, W. B., and D. M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, Bull. Seism. Soc. Am., **71**, 2011-2038.

Joyner, W. B., and D. M. Boore (1982). Prediction of earthquake response spectra, U.S. Geol. Surv. Open-File Rept. 82-977.

Joyner, W. B., and D. M. Boore (1983). Comments on "New attenuation relations for peak and expected accelerations of strong ground motion," by B. A. Bolt and N. A. Abrahamson, Bull. Seism. Soc. Am., **73**, 1479-1480.

Joyner, W. B., and T. E. Fumal (1984). Use of measured shear-wave velocity for predicting geologic site effects on strong ground motion, in Proceedings of the Eighth World Conference on earthquake Engineering, San Francisco, California, 1984, II, 777-784.

Joyner, W. B., T. E. Fumal, and J. C. Tinsley (1983). Estimating site effects of strong ground motion using local shear-wave velocity, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 251-260.

Katayama, T. (1982). An engineering prediction model of acceleration response spectra and its application to seismic hazard mapping, J. Earthquake Eng. Struct. Dyn., **10**, 149-163.

Kennedy, R. P., S. A. Short, T. R. Kipp, H. Banon, F. J. Tokarz, and K. L. Merz (1984). Engineering characterization of ground motion--Task I: Effects of characteristics of free-field motion on structural response, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-3805.

Kimball, J. K. (1983). The use of site dependent spectra, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 401-422.

King, J. L., and B. E. Tucker (1984). Observed variations of earthquake motion across a sediment-filled valley, Bull. Seism. Soc. Am., **74**, 137-151.

Krinitzsky, E. L., and W. F. Marcuson III (1983). Principles for selecting earthquake motions in engineering design, Bull. Assoc. Eng. Geol., **XX**, 253-265.

Lee, V. W., M. D. Trifunac, and C. C. Feng (1982). Effects of foundation size on Fourier spectrum amplitudes of earthquake accelerations recorded in buildings, Int. J. Soil Dyn. Earthq. Eng., **1**, 52-58.

Loh, C. H., J. Penzien, and Y. B. Tsai (1982). Engineering analysis of SMART 1 array accelerograms, J. Earthquake Eng. Struct. Dyn., **10**, 575-591.

Luco, J. E., and J. G. Anderson (1984). Near source ground motion from kinematic fault models, in Proceedings of Workshop on Strong Ground Motion Simulation and Earthquake Engineering Applications, Los Altos, California, 1984, Earthquake Engineering Research Institute, Berkeley, California.

McCann, M. W. (1983). Uncertainty in ground motion predictions, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 331-355.

McCann, M. W., and D. M. Boore (1983). Variability in ground motions, Root mean square acceleration and peak acceleration for the 1971 San Fernando, California, earthquake, Bull. Seism. Soc. Am., **73**, 615-632.

McGarr, A. (1981). Analysis of peak ground motion in terms of a model of inhomogeneous faulting, J. Geophys. Res., **86**, 3901-3912.

McGarr, A. (1982). Upper bounds on near-source peak ground motion based on a model of inhomogeneous faulting, Bull. Seism. Soc. Am., **72**, 1825-1841.

McGarr, A. (1984). Scaling of ground motion parameters, state of stress, and focal depth, J. Geophys. Res. **89**, 6969-6979.

McGarr, A., R. W. E. Green, and S. M. Spottiswoode (1981). Strong ground motion of mine tremors--some implications for near-source ground motion parameters, Bull. Seism. Soc. Am., **71**, 295-319.

McGuire, R. K. (1974). Seismic structural response risk analysis, incorporating peak response regression on earthquake magnitude and distance, M.I.T. Dept. of Civil Engineering Research Rept. R74-51, Cambridge, Massachusetts.

McGuire, R. K. (1977a). Seismic design spectra and mapping procedures using hazard analysis based directly on oscillator response, J. Earthquake Eng. Struct. Dyn., **5**, 211-234.

McGuire, R. K. (1977b). The use of intensity data in seismic hazard analysis, in Proceedings of the Sixth World Conference on Earthquake Engineering, New Delhi, India, 1977, **1**, 709-714.

McGuire, R. K. (1978a). A simple model for estimating Fourier amplitude of horizontal ground acceleration, Bull. Seism. Soc. Am., **68**, 803-822.

McGuire, R. K. (1978b). Seismic ground motion parameter relations, J. Geotech. Engrng. Div., ASCE, **104**, 481-490.

1656 40008

McGuire, R. K. (1984). Ground motion estimation in regions with few data, in Proceedings of Eighth World Conference on Earthquake Engineering, San Francisco, California, 1984, **II**, 327-334.
McGuire, R. K., and T. P. Barnhard (1979). Four definitions of strong motion duration: Their predictability and utility for seismic hazard analysis, U.S. Geol. Surv. Open-File Rept. 79-1515.
McGuire, R. K., and T. C. Hanks (1980). RMS accelerations and spectral amplitudes of strong motion during the San Fernando, California, earthquake, Bull. Seism. Soc. Am., **70**, 1907-1919.
McGuire, R. K., A. M. Becker, and N. C. Donovan (1984). Spectral estimates of seismic shear waves, Bull. Seism. Soc. Am., **74**, 1427-1440.
McNeill, R. L. (1983). Some possible errors in recorded free-field ground motions, in Proceedings of Conference XXII--A workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 361-400.
Mickey, W. V., V. Perez, and W. K. Cloud (1973). Response of Pacoima Dam to aftershocks of San Fernando earthquake, in The San Fernando, California, earthquake of February 9, 1971, EERI/NOAA, **II**, 403-415.
Milne, W. G. (1977). Seismic risk maps for Canada, in Proceedings of Sixth World Conference on Earthquake Engineering, New Delhi, India, 1977, **1**, 930.
Mohraz, B. (1976). A study of earthquake response spectra for different geological conditions, Bull. Seism. Soc. Am., **66**, 915-935.
More, J. J., B. S. Garbow, and K. E. Hillstrom (1980). User guide for MINPACK-1, Argonne National Laboratory Rept. ANL-80-74, Argonne, Illinois.
Mosteller, F., and J. W. Tukey (1977). Data analysis and regression, Addison-Wesley Publishing Company, Reading, Massachusetts.
Mueller, C. S., D. M. Boore, and R. L. Porcella (1982). Detailed study of site amplification at El Central strong-motion array station no. 6, in Proceedings of Third International Earthquake Microzonation Conference, Seattle, Washington, 1982, **I**, 413-424.
Munguia, L., and J. N. Brune (1984). Local magnitude and sediment amplification observations from earthquakes in the Northern Baja-Southern California Region, Bull. Seism. Soc. Am., **74**, 107-119.
Murphy, J. R., and L. J. O'Brien (1977). The correlation of peak ground acceleration amplitude with seismic intensity and other physical parameters, Bull. Seism. Soc. Am., **67**, 877-915.
Newmark, N. M., and W. J. Hall (1982). Earthquake spectra and design, Earthquake Engineering Research Institute Monograph, Berkeley, California.
Newmark, N. M., J. A. Blume, and K. K. Kapur (1973). Seismic design spectra for nuclear power plants, J. Power Div., ASCE, **99**, 287-303.
Nuttli, O. W. (1979). The relation of sustained maximum ground acceleration and velocity to earthquake intensity and magnitude, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 16, Vicksburg, Mississippi.

Nuttli, O. W., (1981). Similarities and differences between western and eastern United States earthquakes, and their consequences for earthquake engineering, in Proceedings of Conference on Earthquakes and Earthquake Engineering: The Eastern U.S., Knoxville, Tennessee, 1981, **1**, 24-51.

Nuttli, O. W., (1982). Advances in seismicity and tectonics, in Proceedings of Third International Earthquake Microzonation Conference, Seattle, Washington, 1982, **I**, 3-38.

Nuttli, O. W., and R. B. Herrmann (1978). Credible earthquakes for the central United States, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 12, Vicksburg, Mississippi.

Nuttli, O. W., and R. B. Herrmann (1982). Earthquake magnitude scales, J. Geotech. Engrng. Div., ASCE, **108**, 783-786.

Nuttli, O. W., and R. B. Herrmann (1984). Ground motion of Mississippi Valley earthquakes, J. Tech. Topics in Civil Engrng., ASCE, **110**, 54-69.

Nuttli, O. W., R. Rodriquez, and R. B. Herrmann (1984). Strong ground motion studies for South Carolina earthquakes, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-3755.

Okubo, T., T. Arakawa, and K. Kawashima (1983). Attenuation of peak motions and absolute acceleration response spectra with use of Japanese strong motion data, in Proceedings of International Symposium on Lifeline Earthquake Engineering and Fourth U.S. National Conference on Pressure Vessels and Piping Technology, Portland, Oregon, 1983, American Society of Mechanical Engineers.

Orphal, D. L., and J. A. Lahoud (1974). Prediction of peak ground motion from earthquakes, Bull. Seism. Soc. Am., **64**, 1563-1574.

Papageorgiou, A. S., and K. Aki (1983). A specific barrier model for the quantitative decription of inhomogeneous faulting and the prediction of strong ground motion; Part II, Applications of the model, Bull. Seism. Soc. Am., **73**, 953-978.

Reimer, R. B., R. W. Clough, and J. M. Raphael (1973). Evaluation of the Pacoima Dam accelerogram, in Proceedings of the Fifth World Conference on Earthquake Engineering, Rome, Italy, 1973, **2**, 2328-2337.

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

Rogers, A. M., J. C. Tinsley, and W. W. Hays (1983). The issues surrounding the effects of geologic conditions on the intensity of ground shaking, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 32-67.

Rogers, A. M., J. C. Tinsley, and R. D. Borcherdt (1984). Geographic variation in ground shaking as a function of changes in near-surface properties and geologic structure near Los Angeles, California, in Proceedings of the Eighth World Conference on Earthquake Engineering, San Francisco, California, 1984, **II**, 737-744.

1683 40008

Sadigh, K. (1983). Considerations in the Development of Site-Specific Spectra, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 423-458.

Sadigh, K., M. S. Power, and R. Youngs (1978). Peak horizontal and vertical accelerations, velocities and displacements on deep soil sites for moderately strong earthquakes, in Proceedings of Second International Earthquake Microzonation Conference, San Francisco, California, 1978, **II**, 801-811.

SAS Institute (1980). Statistical analysis system, SAS Institute, Raleigh, North Carolina.

Scanlan, R. H. (1976). Seismic wave effects on soil-structure interaction, J. Earthquake Eng. Struct. Dyn., **4**, 379-388.

Schenk, V., (1984). Relations between ground motions and earthquake magnitude, focal distance and epicentral intensity, Engineering Geology, **20**, 143-151.

Schnabel, P. B., and H. B. Seed (1973). Accelerations in rock for earthquakes in the western United States, Bull. Seism. Soc. Am., **63**, 501-516.

Scholz, C. H. (1982). Scaling relations for strong ground motion in large earthquakes, Bull. Seism. Soc. Am., **72**, 1903-1909.

Seed, H. B., and I. M. Idriss (1982). Ground motions and soil liquefaction during earthquakes, Earthquake Engineering Research Institute Monograph, Berkeley, California.

Seed, H. B., and J. Lysmer (1980). The seismic soil-structure interaction problem for nuclear facilities, University of California, Report to Lawrence Livermore National Laboratory, Berkeley, California.

Seed, H. B., C. Ugas, and J. Lysmer (1976a). Site-dependent spectra for earthquake-resistant design, Bull. Seism. Soc. Am., **66**, 221-243.

Seed, H. B., R. Murarka, J. Lysmer, and I. M. Idriss (1976b). Relationships of maximum acceleration, maximum velocity, distance from source and local site conditions for moderately strong earthquakes, Bull. Seism. Soc. Am., **66**, 1323-1342.

Shakal, A. F., and D. L. Bernreuter (1981). Empirical analysis of near-source ground motion, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-2095.

Sharpe, R. L. (1982). An investigation of the correlation between earthquake ground motion and building performance, Applied Technology Council Rept. ATC-10, Palo Alto, California.

Singh, J. P. (1983). Causes of variation in strong ground motion in the near-field for a given soil condition, in Proceedings of Conference XXII--A Workshop on Site Specific Effects of Soil and Rock on Ground Motion and Their Implications for Earthquake-Resistant Design, Santa Fe, New Mexico, 1983, U.S. Geol. Surv. Open-File Rept. 83-845, 89-114.

Slemmons, D. B. (1977). Faults and earthquake magnitude, in State-of-the-Art for Assessing Earthquake Hazards in the United States, U.S. Army Engineer Waterways Experiment Station Misc. Paper S-73-1, Rept. 6, Vicksburg, Mississippi.

Smith, S. W., J. E. Ehrenberg, and E. N. Hernandez (1982). Analysis of the El Centro Differential array for the 1979 Imperial Valley earthquake, Bull. Seism. Soc. Am., **72**, 237-258.

Sunder, S. S., and J. J. Connor (1982). A new procedure for processing strong-motion earthquake signals, Bull. Seism. Soc. Am., **72**, 643-661.

Swanger, H. J., S. M. Day, J. R. Murphy, and R. Guzman (1981). State-of-the-art study concerning near field earthquake ground motion, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-1978.

Swanger, H. J., J. R. Murphy, T. J. Bennett, and R. Guzman (1980). State-of-the-art concerning near field earthquake ground motion--annual report for period Sept. 1978 to Sept. 1979, U.S. Nuclear Regulatory Commision Rept. NUREG/CR-1340.

Toro, G. R. (1981). Biases in seismic ground motion prediction, M.I.T. Dept. of Civil Engineering Research Rept. R81-22, Cambridge, Massachusetts.

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

Trifunac, M. D. (1976b). Preliminary analysis of the peaks of strong earthquake ground motion--Dependence of peaks on earthquake magnitude, epicentral distance, and recording site conditions, Bull. Seism. Soc. Am., **66**, 189-219.

Trifunac, M. D. (1976c). A note on the range of peak amplitudes of recorded accelerations, velocities and displacements with respect to modified Mercalli intensity scale, Earthquake Notes, **47**, 9-24.

Trifunac, M. D., (1978). Response spectra of earthquake ground motion, J. Engrng. Mech. Div., ASCE, **104**, 1081-1097.

Trifunac, M. D. (1979). Preliminary empirical model for scaling Fourier amplitude spectra of strong ground acceleration in terms of modified Mercalli intensity and recording site conditions, J. Earthquake Eng. Struct. Dyn., **7**, 63-74.

Trifunac, M. D., and A. G. Brady (1975a). On the correlation of seismic intensity scales with the peaks of recorded strong ground motion, Bull. Seism. Soc. Am., **65**, 139-162.

Trifunac, M. D., and A. G. Brady (1975b). On the correlation of seismoscope response with earthquake magnitude and modified Mercalli intensity, Bull. Seism. Soc. Am., **65**, 307-321.

Trifunac, M. D., and A. G. Brady (1975c). On the correlation of peak acceleration of strong motion with earthquake magnitude, epicentral distance and site conditions,in Proceedings of U.S. National Conference on Earthquake Engineering--1975, Ann Arbor, Michigan, 1975, 43-52.

Trifunac, M. D., and V. W. Lee (1978a). Dependence of the Fourier amplitude spectra of strong motion acceleration on the depth of sedimentary deposits, University of Southern California Dept. of Civil Engineering Rept. CE78-14, Los Angeles, California.

Trifunac, M. D., and V. W. Lee (1978b). Uniformly processed strong earthquake ground accelerations in the western United States of America for the period from 1933 to 1971: Corrected acceleration, velocity and displacement curves, University of Southern California Dept. of Civil Engineering Rept. CE78-01, Los Angeles, California.

Trifunac, M. D., and V. W. Lee (1979). Dependence of pseudo-relative velocity spectra of strong motion acceleration on the depth of sedimentary deposits, University of Southern California Dept. of Civil Engineering Rept. CE79-02, Los Angeles, California.

Tucker, B. E., and J. L. King (1984). Dependence of sediment-filled valley response on input amplitude and valley properties, Bull. Seism. Soc. Am., **74**, 153-165.

Tucker, B. E., J. L. King, D. Hatzfeld, and I. L. Nersesov (1984). Observations of hard rock site effects, Bull. Seism. Soc. Am., **74**, 121-136.

Tukey, J. W. (1977). Exploratory data analysis, Addison-Wesley Publishing Company, Reading, Massachusetts.

Ts'ao, H-S. (1980). Correlations of peak earthquake ground acceleration in the very near field, Agbabian Associates Rept. SAN/1011-125, El Segundo, California.

Ts'ao, H-S. (1981). Statistical analysis of response spectral amplitudes including near-field data, Agbabian Associates Rept. SAN/1011-129, El Segundo, California.

Wang, P. C., and C. B. Yun (1979). Site-dependent critical design spectra, J. Earthquake Eng. Struct. Dyn., **7**, 569-578.

Werner, S. D. (1976a). Engineering characteristics of earthquake ground motion, Nucl. Eng. Des., **36**, 367-395.

Werner, S. D. (1976b). Procedures for developing vibratory ground motion criteria at nuclear power plant sites, Nucl. Eng. Des., **36**, 411-441.

Werner, S. D. (1977). A study of vertical ground response spectrum shapes, Agbabian Associates Rept. SAN/1011-114, El Segundo, California.

Werner, S. D. (1978). Correlation of ground response spectra with modified Mercalli site intensity, Agbabian Associates Rept. SAN/1011-113R, El Segundo, California.

Werner, S. D., H-S. Ts'ao, and D. Rothman (1979). Statistical analysis of earthquake ground motion parameters, U.S. Nuclear Regulatory Commission Rept. NUREG/CR-1175.

Young, G. A. (1980a). Definition of design earthquake vibratory ground motion--recommended guidelines, Agbabian Associates Rept. SAN/1011-124, El Segundo, California.

Young, G. A. (1980b). Earthquake vibratory ground-motion intensity attenuation, Nuclear Safety, **21**, 205-214.

U.S. Geological Survey
Box 25046, MS 966
Denver Federal Center
Denver, Colorado 80225

SITE EFFECTS:
A Generic Method for Modeling Site Effects in Seismic Hazard Analyses*

By Jean Savy, Don Bernreuter and J.C. Chen
Lawrence Livermore Laboratory
P.O. Box 808, Livermore, CA

1. DEFINITION OF SITE EFFECTS

We define site effects as the departure of the strong motion observed at a site from the strong motion which would have been observed, had the local site characteristics been those of a generic case. Thus, the consideration of site effects is necessarily a relative concept where the objects of the comparison have to be defined. Figure 1 shows some possible travel paths of the seismic energy from the seismic source to the site. The seismic waves travel from the source directly to the location of the site, others are refracted and reflected at greater depths and yet others propagate at the surface of the earth before reaching the location of the site. The characteristics of the motion at the site are therefore a function of the following parameters:

- The characteristics of the source geometry of the source, orientation, and kinematic properties, and all other properties of the source which are affecting the spatial description of the energy content of the source.

- The properties of the travel paths. The physical properties of the medium (i.e. stress-strain and attenuation properties) and regional characteristics , such as discontinuities in the upper crust (plate boundaries, subduction zones, etc.).

- The local topography and geology.

*"This work was supported by the United States Nuclear Regulatory Commission under a Memorandum of Understanding with the United States Department of Energy."

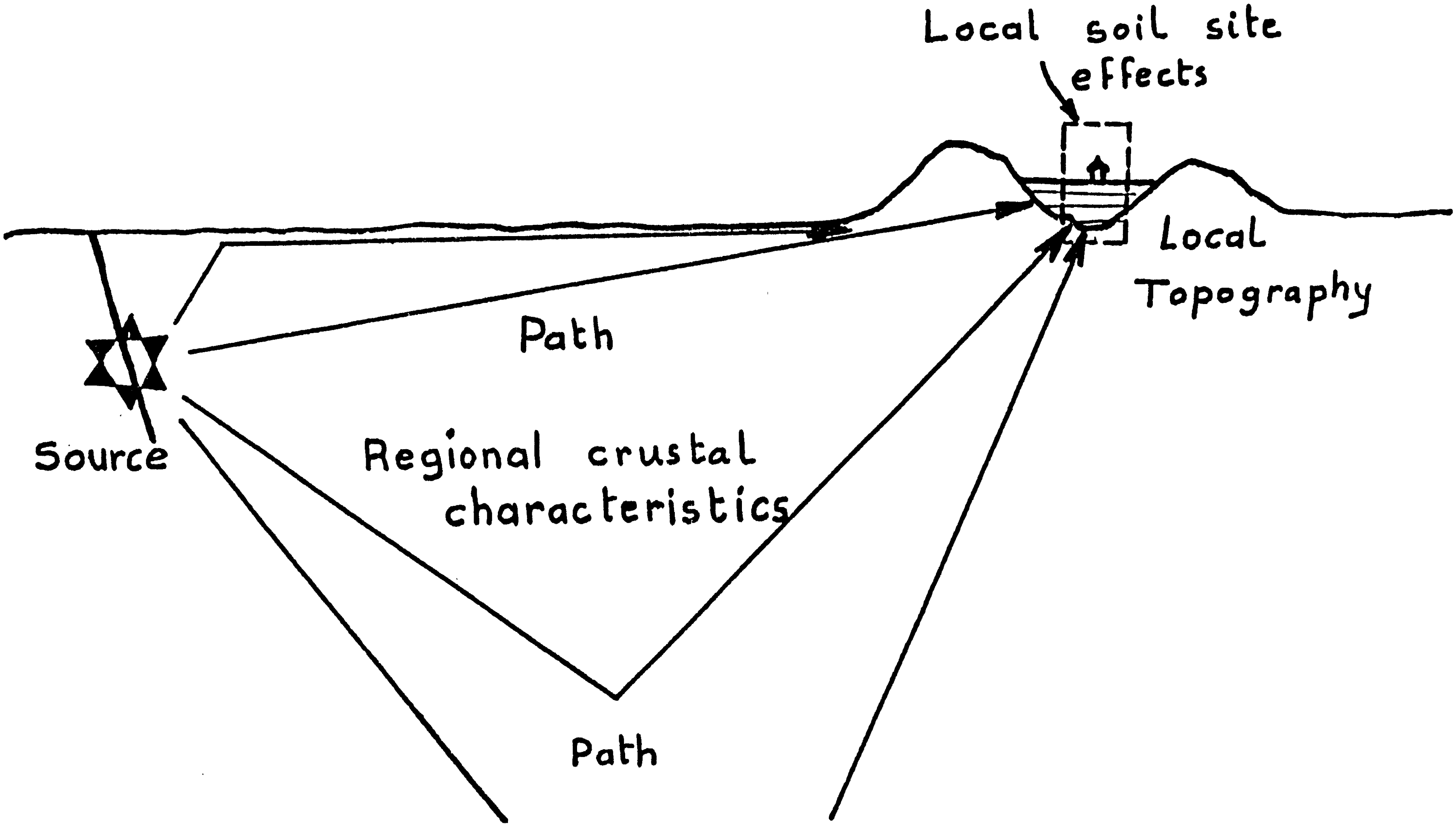


Fig. 1: Parameters influencing the motion at a Site. Schematic representation of the Seismic Source, various wave travel paths and the local site characteristics.

1652 40008

The above definition of site effects assumes that it is possible to estimate the motion at a site from a generic description of the source and of the travel path, provided that the characteristics of the site are consistent with the assumed characteristics of the travel path in the vicinity of the site.

In practice, the generic site properties are either a hard rock or deep stiff soil and the topography is a flat surface. In seismic hazard studies, the choice of the generic soil properties is dependent on the type of ground motion attenuation models used for the analyses.

The ground motion models used in the seismic hazard analyses of the eastern United States (EUS), (Bernreuter el al. 1985) were in large part based on western United States (WUS) data which was mostly recorded on medium to stiff deep soil sites. Thus, the generic site type of the EUS study consists of a deep soil site. Section 3 of this paper describes how the sites with characteristics considered different from the generic case are treated as a departure form the generic case by applying a simple correction to the ground motion prediction.

Section 2 of this paper presents a case study where the existence of site effects is demonstrated by comparing strong ground motions observed at several stations during the 1976 Friuli earthquake in Italy. Section 3 presents the method adopted to generically account for the soil site conditions at any site of a nuclear power plant located in the EUS.

2. CASE STUDIES

2.1 The Friuli Earthquake (Italy), 1976

The 1976 Friuli (Italy) earthquake provides a good opportunity to identify some effects of soil site conditions on strong ground motion because the ground motion has been recorded at several sites of different soil conditions and very close to one another. Thus, the source characteristics and travel paths can be assumed to be the same at these sites and the only differences observed between two closely located sites can be attributed to the differences in local soil conditions at those two sites. (Bohn et al, 1984)

The two sites under consideration here are:

- o Rocco, located on rock
- o Cornino Forgaria, located on soil

Figure 2 shows the local soil profile at the two sites (Fig. 2a), as well as the strong motion recorded (Fig. 2b) and the corresponding 5% damping response spectra (Fig. 2c). Although an examination of Figs. 2b and 2c shows that the motions at the two sites are different, a better tool for their comparison is the plot of the spectral ratios. The spectral ratios (i.e, Spectral values calculated for the Forgaria site divided by the spectral values for the Rocco site) are calculated for ten pairs of horizontal components recorded for several earthquakes in the close vicinity of the sites. Figure 3 shows the mean value and an estimate of the uncertainty in these spectral ratios. Using the mean curve as an estimate of the difference in effects at the two sites, it is seen that both the rock and soil sites have the same amplification for frequencies below 1-1.5 Hz. At higher frequencies (4 Hz and above) the soil site amplification is approximately twice as much as the amplification on the rock. In the intermediate range of frequencies (1.5 to 4 Hz) the characteristics of the soil site (Forgaria) are such that the motion is amplified as much as four times more than for the rock.

2.2 The 1975 Oroville and 1977 Briones and Richmond Earthquakes

Spectral ratios were also calculated for records obtained at the D. Johnson Ranch (DJR) from the August and September 1975 Oroville earthquakes and the Richmond Field Station (RFS) from the 1977 Briones and Richmond earthquakes. These spectral ratios are shown in Figs. 4 and 5 respectively. The effects observed at these two sites (DJR and RFS) are similar to those observed at the

Fig. 2b: Sample accelerograms collected at the two sites during the Friuli 2/12/76 earthquake.

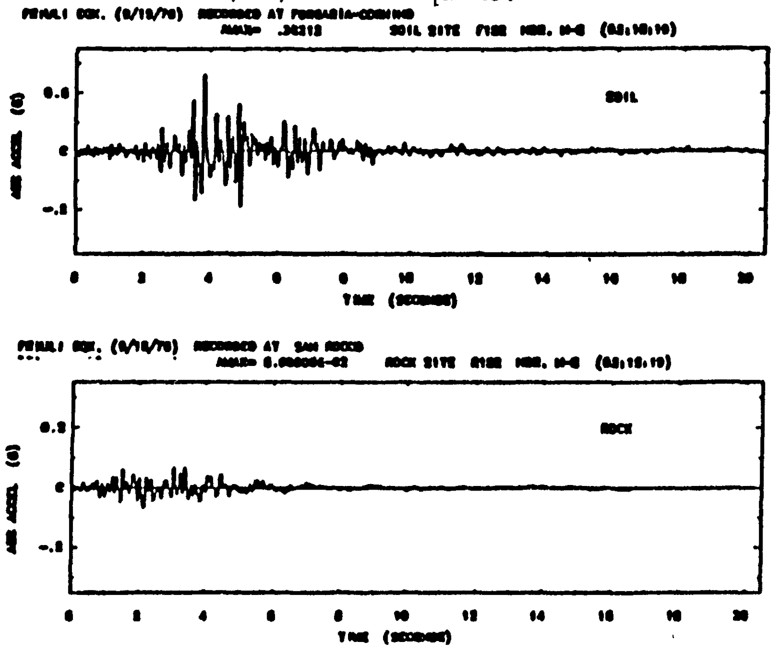


Fig. 2c: Response spectra of the motions described in Fig. 2b.

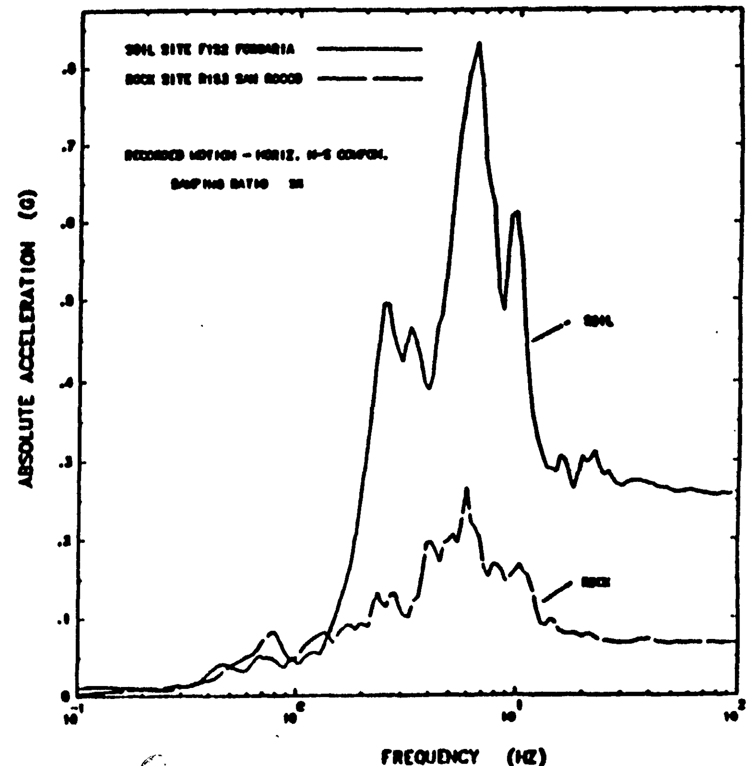
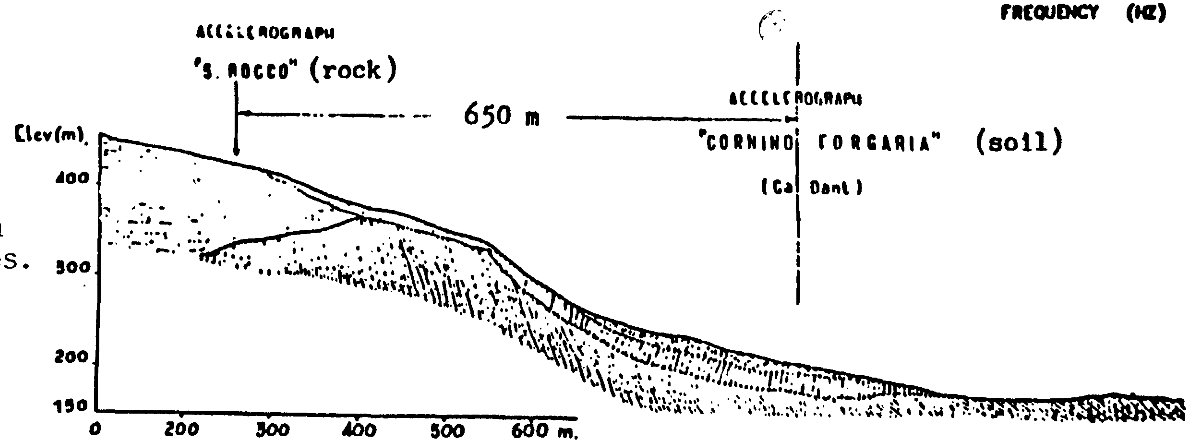


Fig. 2a: Description of the sites.



SITE PROFILE ALONG NORTH-SOUTH DIRECTION

Fig. 2: Example of the [redacted]

1788 4008

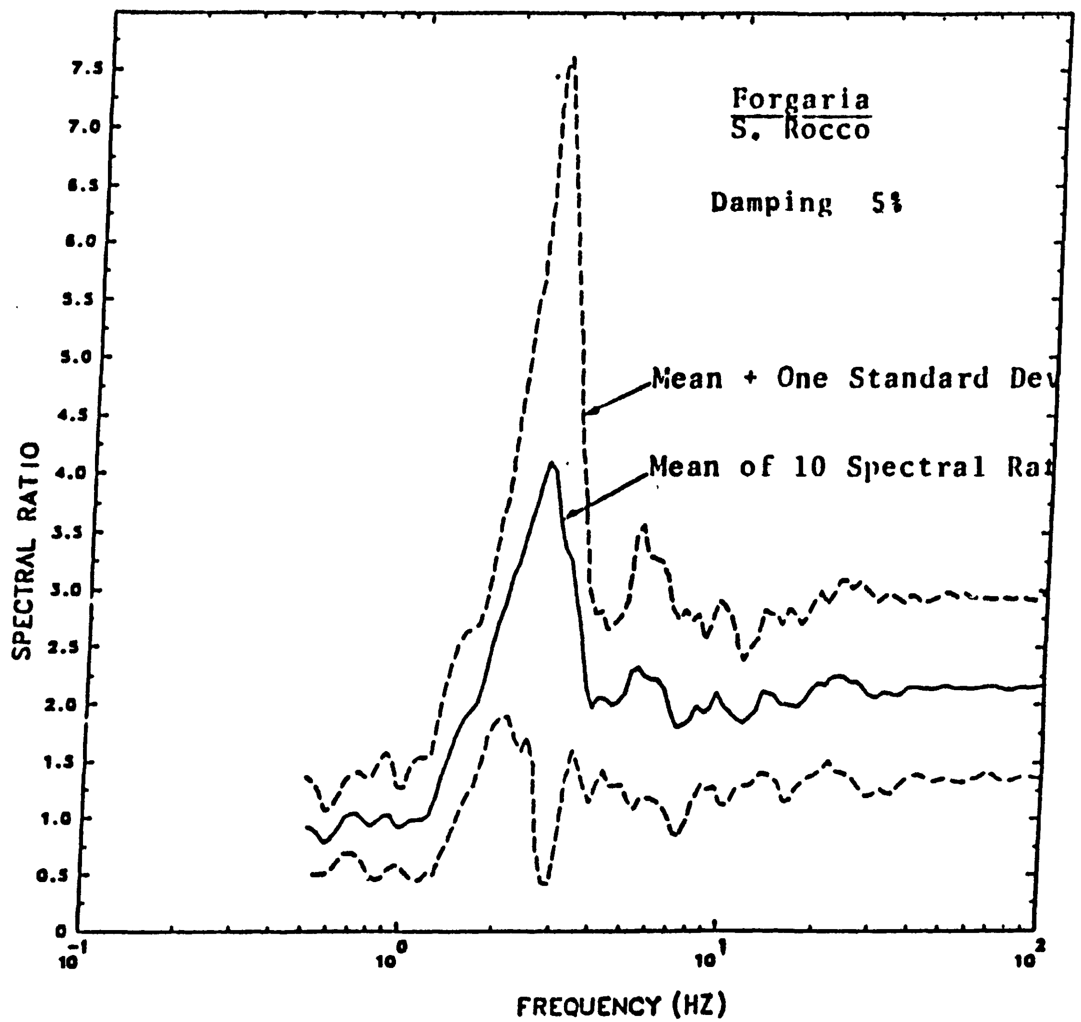


Fig. 3: Average spectral ratio at the Rocco-Cornino Forgaria sites based on 10 horizontal components of groung motion acceleration.

1754 40008

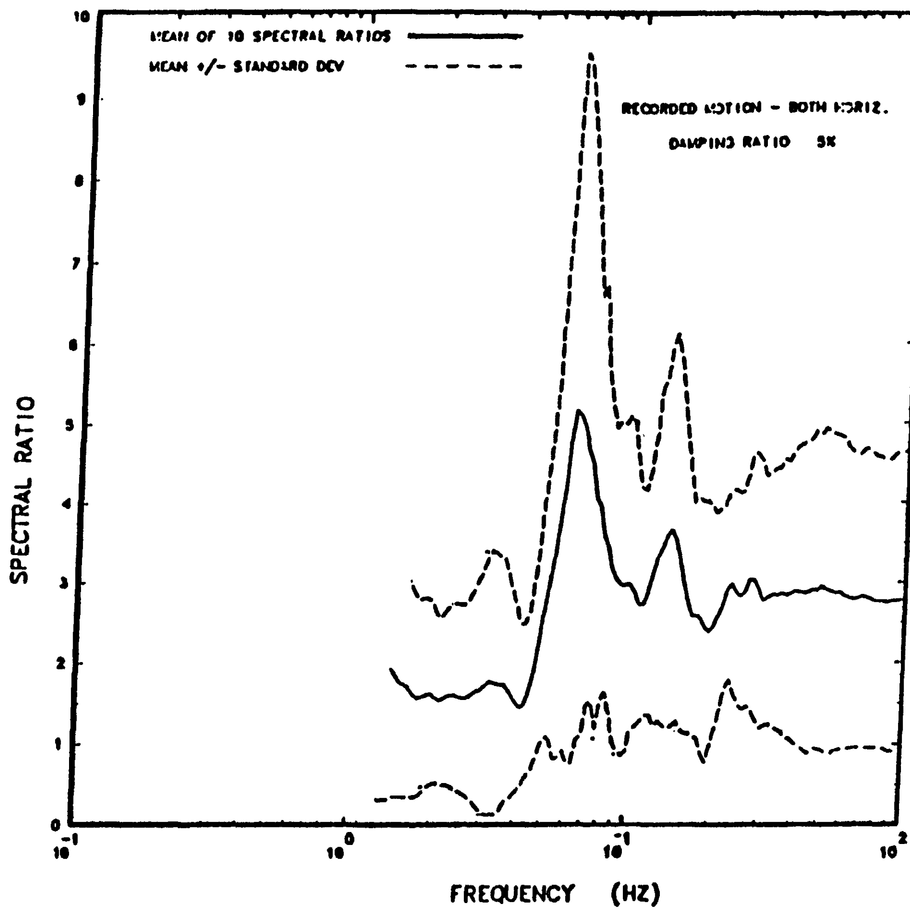


Fig. 4: Average acceleration spectral ratio estimated for D. Johnson Ranch from the August and September 1975 Oroville Earthquakes.

1758 40008

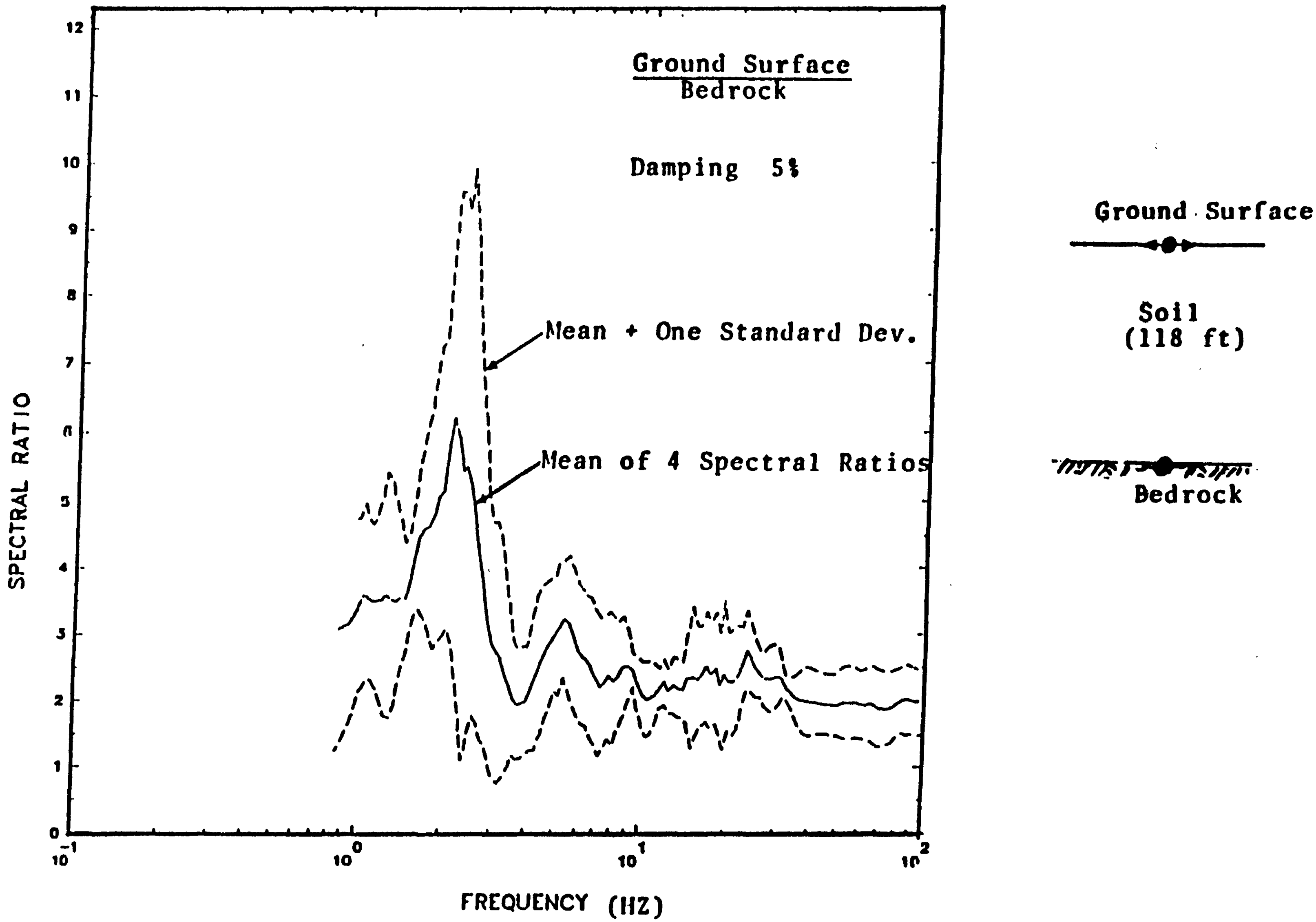


Fig. 5: Average acceleration spectral ratio at the Richmond field station site from the 1977 Briones and Richmond earthquakes.

Rocco-Forgaria site. However, the details of the spectral ratio curves are different in all three cases and depend on the actual sites' characteristics. For instance, the dominant frequency is not the same in all cases, and the peak spectral ratios are also different. Table 1 summarizes the various values obtained in each case. Of particular interest to us is the peak value of the spectral ratio (called PEAK, in Table 1) and the frequency at which this peak is reached (called FREQ. (Hz)).

In analyzing these results, it is worth noting that the characteristics of the sites at DJR are not as well known as those at the Rocco-Forgaria site, furthermore the comparisons at RFS are for records at the same location for different elevations, as shown in Fig. 5.

3. A GENERIC METHOD OF ACCOUNTING FOR SITE SOIL CONDITIONS FOR THE SEISMIC HAZARD ANALYSIS OF THE EUS

3.1 Overview

One of the main objectives of the EUS project was to assess the uncertainty in the estimate of the seismic hazard at selected nuclear power plant sites in the EUS. In keeping with this objective, we wanted to include the uncertainty introduced by the local site conditions at various power plant sites. This uncertainty has both random and systematic components. The systematic component can be accounted for by using several different approaches to obtain the correction factors varying from "no correction" to that obtained by a linear 1-D analysis such as performed in the SHAKE computer program. The random aspect which arises from our uncertainties in the soil column and energy and frequency content of the potential seismic ground motion at the site can be accounted for by including uncertainty in the correction factors for each systematically different method used to develop the correction factors.

TABLE 1

SITES	AVERAGE IN TWO HORIZONTAL DIRECTIONS					AVERAGE IN VERTICAL DIRECTION				
	PGA	COV	PEAK	COV	FREQ.(Hz)	PGA	COV	PEAK	COV	FREQ.(HZ)
FORGARIA	2.3	(.32)	4.2	(.80)	2.8	2.2	(.23)	4.1	(.52)	5.5
DJR	2.8	(.64)	5.2	(.86)	6.1	2.4	(.75)	2.9	(.59)	12.0
RFS	2.0	(.25)	6.2	(.56)	2.1	4.2	(-)	7.5	(-)	13.0

The general methodology for collecting input data was to elicit experts' opinions in each of the relevant fields. A panel of experts in ground motion modeling was formed. The role of the Panel Members was first to select one approach/correction. They were then asked to assign to their choice a value representing their degree of belief that it was the true approach/correction.

There were a few important limitations to what was possible, e.g. schedule and budget requirements precluded the development of some new approaches. It restricted what could be done, i.e., acceptable approaches must fit into our analysis scheme.

For the short term we proposed the following approaches/corrections factors:

1) No correction.

2) Use only a simple soil or rock classification if available -- otherwise, no correction.

3) Develop correction factors for each ground motion model based on several generic site classifications, 1-D analysis data, and judgment.

4) Do a site specific analysis.

5) Other - as proposed by panel members.

Each of these are discussed in detail in the following sections:

3.2 No Correction

Here, the argument might be that both our knowledge of EUS ground motion is so poor and the methods we have to assess local site effects so uncertain, that it would be better to do nothing. All site types would be treated the same.

Here, it should be noted that some of the ground motion models adopted by the experts fell into this category. For example, Nuttli's model makes no reference to site type. One could argue that it is for "generic" soil sites. Certainly, Campbell's models fall into the "generic soil" category as they were developed using only soil data. For other models, e.g. what we have labeled the Trifunac-Anderson model, have a simple site correction term included. If this case is selected and such a ground motion is included, then the value for the site correction term should be specified.

3.3 Simple Rock/Soil Correction

For this case, the site types are put into two (rock or soil) or (stiff or soft) or three categories (soft, stiff, basement rock) and a simple constant (for each category) correction factor is applied.

Figure 6 shows typical correction factors going from soil to rock as a function of period found from WUS data. The curve labeled 1 is based on Joyner-Boore (1982) regression analysis, the curves labeled 2 and 3 are based on Trifunac and Anderson (1977) and the curve labeled 4 is based on the SEP results Bernreuter (1981). For Trifunac's model, curve 2 is between soft alluvium and hard sedimentary rock and curve 3 is between soft alluvium and basement or crystalline rock. Both Joyner-Boore and the SEP use only rock/soil categories.

It can be seen from Table 2 that there is considerable variation between different studies. These differences arise from several causes, including the applicability of the form of the model, i.e., the influence of site type might be a function of magnitude and distance. This is very hard to verify because there are too few sites which are truly rock sites.
In addition to possible deficiencies in the form of the mathematical model used for the regression analyses, all of the regression analyses were performed using less than perfect data sets. All the data sets suffer from the use of poor criteria for the identification of rock sites. Recently, more

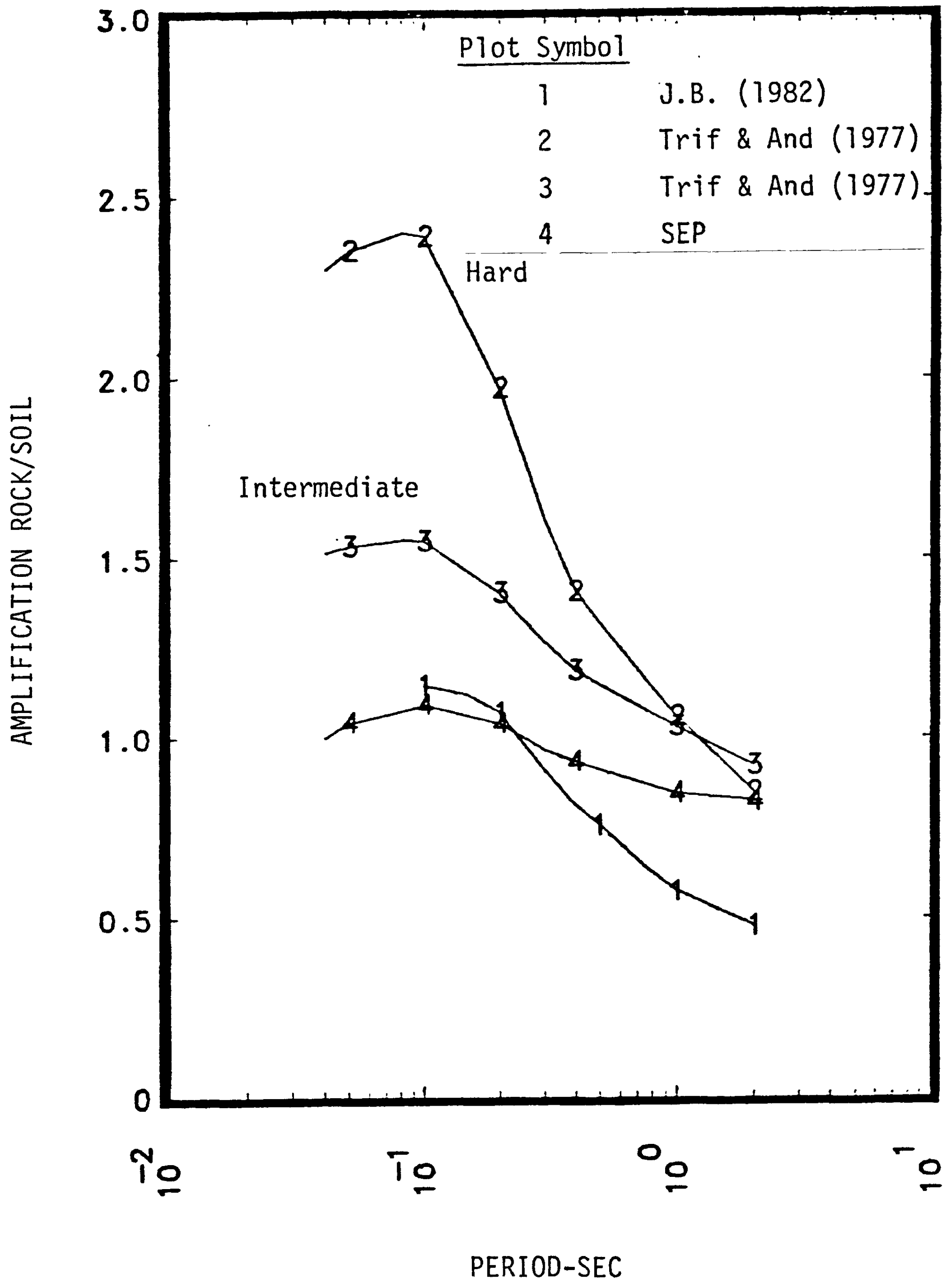


Fig. 6 : Simple correction factors obtained by regression analysis of WUS data.

1783 4000 8

TABLE 2

Model	Ratio PGA	PGV
Trifunac (1976a) Basement Rock	1.93	1.07
(Intensity Based) Sedimentary Rock	1.40	1.03
Joyner-Boore (1982)	1.00	0.68
SEP	1.00	0.87
McGuire (1978)	1.22	0.93
Trifunac (1976b) Crystalline Rock	0.76	0.55
Sedimentary Rock	0.87	0.74
Campbell (1981)	1.00	----

site boring data has become available to assist in properly sorting the data into categories. The Joyner-Boore (1981) data set is the best in this regard, but it contains a number of questionable sites identified as rock sites. In contrast with the Joyner-Boore data set, most data sets contain data recorded in large buildings and/or in basements. These data sets were used to obtain the results plotted on Fig. 6 and given in Table 2. Joyner and Boore (1981), Campbell (1981) and (1983) and others have shown that building type and location of the recorder in a sub-basement can have a significant effect on both the PGA and PGV.

Thus, one possible approach to account for the rock would be to introduce a correction factor for these models based on WUS data, i.e., use one of the factors plotted on Fig. 6 to convert from "soil" to rock. This is a somewhat arbitrary approach but would include some correction for the systematic difference that exists between sites. PGA could be corrected using one of the factors from Table 2. In a somewhat more complex model, correction factors could be given along with their uncertainty.

3.4 Generic Correction Factors

3.4.1 Overview of Approach

The simple model proposed in Section 3.3 might be adequate if enough categories are used; however, the data base is too sparse to define many categories. The approach developed in this section, consists in supplement the empirical data set with analysis. Our proposed procedure is as follows:

1) Use available soil/rock pairs (soil and rock stations in close proximity to each other that record the same earthquakes) to compute observed amplification factors. This provides a measure of the range of realistic amplification and the uncertainty introduced by source, travel path and rheological effects. These results are used to calibrate analytic results.

2) Ideally, one would like to have a sufficient number of soil/rock pairs to put them into a "reasonable" number of categories based on soil type, depth, bedrock shear wave velocity, etc. and develop "generic" median amplification factors for each category and the uncertainty associated with each category. Due to insufficient data we resorted to analytic modeling. We defined eight categories based on three soil depth categories and two soil type categories plus a rock category and a deep soil category. Two basic soil types were chosen: (1) primarily a sandy type soil column and (2) primarily a "till like" column. Granted, most soil columns are mixed, but defining too many categories becomes pointless and a site specific approach such as discussed in 3.5 should be used. Each category contains several different soil columns which are based on actual soil columns at nuclear power plant sites. We selected a set of time histories recorded at rock sites with a range of magnitudes and distances to incorporate the uncertainty from the source and travel path effects in the analysis. These time histories were used as input to the SHAKE computer program and the PGA and amplification factors were computed for each category and for each time history. Then the median and standard deviation of the correction factor were estimated for each category.

One major question arises -- what should these correction factors be applied to? Ideally, we should predict the PGA hazard curve and Uniform Hazard Spectrum (UHS) at each site at a hypothetical unweathered basement rock outcrop and then apply the correction for the appropriate site category to obtain the PGA hazard curve and UHS corrected for local site effects. Unfortunately, we think that none of the ground motion models available can be considered as predicting the ground motion at an unweathered hard rock outcrop. In fact, the ground motion models used in our analysis can at best be considered as applicable to only a "generic" soil site. A few of the models chosen can be considered applicable to weathered rock sites. But in our opinion, these models have the serious problem that the different field data sets used to develop them have shallow soil sites listed as rock sites and they have data from large buildings/basements intermixed with free-field data.

Given the lack of data recorded at <u>true</u> rock sites and the possible complexity of the systematic differences between rock and soil sites, it is not clear that even with added analysis we could develop an acceptable ground motion model for even weathered rock sites. For this reason we included a "generic" soil category in our analyses. Amplification factors were computed relative to the generic soil category as well as rock.

Several avenues are possible:

1) Only use the soil version of the various ground motion models and correct using computed correction factors for the site's category.

2) Use a mixed set; i.e., use the rock version, if available, and use the rock to site category correction factor; otherwise use the soil model and soil to site category correction factor.

3) Convert all models to a rock version using the approach suggested in Section 3.3 and then use the rock to site category correction factor.

3.4.2 Selection of Time Histories

Ideally, we would like to have a set of time histories recorded on unweathered hard rock from earthquakes with m_b magnitudes ranging from 4.0 to 7.0 and distances ranging from 2-1/2 km to several hundred kilometers. With such a set, we could examine the dependence on magnitude and distance. Unfortunately, the available set of time histories does not match the ideal set. Most records are recorded on weathered rock -- and in fact, the shear wave velocity of the "rocks" at many sites is closer to a "soil" than a rock. Because only ratios are involved, this may not be a major problem except at hard unweathered rock sites where we do not have a good measure of what the difference might be between the ground motion recorded at hard rock sites as compared to soft weathered rock sites.

Table 3 gives a list of the records selected. We restricted our choice to recordings made either in the free-field or in small buildings. It can be seen from Table 3 that we have a reasonable distribution of magnitudes but not a very good distribution on distance.

3.4.3 Definition of Site Categories

In order to define site categories, site data for more than 60 nuclear power plant sites throughout the United States has been reviewed. Such site data includes geologic profile (layering and depth to rock), soil parameters (soil type, shear wave velocity, compressional wave velocity, density, shear modulus and damping ratio at high strain levels) and bedrock properties (shear wave velocity, compression wave velocity and density). Like most site classification systems, soil depth to the bedrock and soil type are the primary site parameters used to define site categories. Based on our review of the available data from FSAR and PSAR of U.S. nuclear plants sites we have defined the following categories based on the range of the thickness of the soils above bedrock and primary soil type:

Site Class I: Rock Sites. This category includes sites with exposed bedrock including plutonic, igneous, metamorphic, crystalline and sedimentary rock. Sites where the thickness of the soil is less than 25 feet are also assumed to fall in this category and the surface material is neglected as it is generally removed. The mean shear wave velocity from 60 sites is about 6200 fps with a coefficient of variation (COV) of 40%.

Site Class II: Intermediate thickness soil sites. This category includes sites having soil layering thickness ranges from 25 to 300 ft over bedrock. Based on the samples distribution of available sites, we further classify this site class into three subclasses IIa, IIb and IIc.

Table 3
Rock Records Used in the Analysis

Station	Earthquake	Mag. M_L	Dist. R	Accel. g's
Helena Fed. Bld.	Helena, Mont. 10/31/35	6.	8	0.15
Golden Gate Park	Daly City 3/22/57	5.3	8	0.13
Temblor	Parkfield 6/21/67	5.5	11	0.41
Pacoima Dam	San Fernando 2/9/72	6.4	3	1.20
Pacoima Dam	After Shock	5.4	12	0.11
Cal. Tech. Seism. Lab	San Fernando 2/9/72	6.4	18	0.19
Griffith Park Obs.	San Fernando 2/9/72	6.4	17	0.18
Cape Mendocino	Cape Mendocino 6/7/75	5.3	25	0.20
Oroville Seism. Sta.	Oroville 8/1/75	5.7	8	0.11
Gilroy Array No. 1	Coyote Lake 8/6/79	5.9	9	0.13
Gilroy Array No. 6	Coyote Lake 8/6/79	5.9	4	0.42
Superstition Mt.	Imperial Valley 10/15/79	6.6	25	0.21
Cerro Prieto	Imperial Valley 10/15/79	6.6	24	0.17
Superstition Mt.	Westmoreland 4/26/81	5.7	13	0.11
Rocca	Ancona, Italy 6/14/72	4.7	6	0.55
Rocca	Ancona, Italy 6/14/72	4.2	6	0.45
San Rocco	Friuli, Italy 9/15/76	6.1	9	0.12
San Rocco	Friuli, Italy 9/15/76	6.0	19	0.23
Bagnoli	Campunia Lucania 11/23/80	6.7	12	0.18
Sturno	Campania Lucania 11/23/80	6.7	18	0.23

(IIa): Soil deposit of 25 to 80 ft over rock. The mean shear wave velocity for each site is calculated by weighted sublayer sites, the mean shear wave velocity is 1500 fps with a COV of 40%. The mean thickness is 48 ft with a COV of 30%. The mean and COV of bedrock shear wave velocity are 6000 fps and 30% respectively.

(IIb): Soil thickness of 80 to 180 feet over rock The mean and the COV of sites are 1550 fps and 40% respectively. The mean soil thickness is about 120 feet with a COV of bedrock is 6400 fps with a COV of 40%.

(IIc): For soil depth of 180 to 300 feet over rock. Only four sites fell into this category. The mean shear wave velocities of soil and rock are 2000 fps and 9350 fps respectively. The mean soil thickness is 250 ft. No COVs were computed due to insufficient site data available to us this time.

Site Class III: Thick soil sites. This is our generic soil category site and includes those sites having soil deposit more than 300 feet over the bedrock. The mean and COVs of the shear wave velocity among a set of 14 sites are 2115 fps and 26% respectively. The median value of soil thickness was found to be 650 feet with a COV of 40%. The mean shear wave velocity of the bedrock is 5700 fps with a COV of 45%.

As these site profiles show a great deal of variability in site parameters, a generic site class can only be defined in a loose manner. For each site class, we only consider two extreme soil types, namely till-like soil and cohesionless soil (sand-like). Since the variation of shear wave velocity with depth for a till-like soil is not significant, a mean constant shear wave velocity was assumed in each site class for the generalized till-like site model for response calculation. However, because the shear modulus for cohesionless soils is more sensitive to the soil depth, it is assumed that the shear modulus varies with the square root of the effective over burden pressure for the sand-like site models.

For the deep soil site Class III, only a sand-like soil was used, i.e., it was assumed that the shear modulus varied as the square root of the effective over burden stress.

3.4.4 Analysis Procedure

The site response was calculated by assuming one-dimensional vertically propagating SH waves. Sites were modeled as a system of horizontal layers of infinite extent. Viscoelastic material model for each layer were assumed -- shear modulus, density, Poisson's ratio, and material damping.

Site response calculation should account for the uncertainty contributed by the variation of depth of the soil model, dynamic soil properties, and the impedance ratio between soil and bedrock. All of these factors contribute significant uncertainty to calculated response. In addition, the seismic input motions are also an important contributor to the uncertainty. In our analysis to account for the above sources of uncertainty, we perform repeated deterministic analysis, each analysis simulating an earthquake occurrence. By performing many such analyses and by varying the values of the above input parameters, a mean response and its coefficient of variation can be obtained. Variability in the seismic input is included by sampling one of the twenty time histories listed in Table 2 to obtain a different earthquake time history for each simulation. Variability in the dynamic modeling was introduced by sets of input parameters (mainly shear wave velocities of soil and rock, damping ratio of soil and the depth of soil deposit) from assumed probability distribution for each simulation. A log normal distribution was assumed for this study.

Table 4 shows the mean and the COV of four of the input parameters used in the simulation for site response analysis.
The influence of non-linear soil behavior on site amplification is still an open research area. Currently, very little field data has been obtained to address this question. Tucker and King (1984) show that the observed site amplifications are not much different between groups of strong and weak motions.

Practical non-linear soil constitutive models are not yet available. The non-linear behavior of soil materials cannot be fully described by constant elastic moduli and damping coefficients. However, a good approximation of the effects of soil non-linearities on the response can be obtained by the use of constant strain compatible moduli and damping ratios in a sequence of linear analyses. This method is known as the equivalent linear method (Seed and Idriss, 1969).

Both linear and equivalent linear analyses were performed for site classes IIa, IIb and IIc. The results of our analysis are discussed in the following section.

3.4.5 Computed Correction Factors

Our analysis of both actual and simulated data indicates that there is considerable variability in the correction factors from earthquake to earthquake (see the case studies in Section 2). The variability can easily be accounted for by including it in the simulation process. The main problem, in our opinion, is in defining the ground motion levels and frequency content for the generic (base) case. The difficulties in defining the base case are threefold:

1854-10008

TABLE 4

The Means and COVs of the Input Parameters Used for Numerical Simulation for Site Response Analysis

	Class IIa		Class IIb		Class III	
	Mean	COV	Mean	COV	Mean	COV
H (ft; soil	48.	0.25	120.	0.25	650.	0.40
Vs (fps; soil)	1500.	0.40	1550.	0.40	2115.	0.26
D (%; soil)	7.	0.60	7.	0.60	7.	0.60
Vs (fps; rock)	6000.	0.40	6400.	0.40	6200.	0.40

H = Layer Thickness
V_s = Shear Wave Velocity
D = Damping Ratio

1858 40008

1) Few (possibly none) of the ground motion models are applicable for rock sites.

2) What soil column should represent the generic soil models.

3) The set of rock records contain many records obtained at soft weathered sites.

As a starting point we have taken the "generic soil site" to be represented by a deep sand-like (shear-wave velocity function of the depth) site with linear viscoelastic properties, i.e. site Class III defined in Section 3.4.3. We have examined two sets of correction factors -- one set relative to the rock outcrop records and one set relative to the Class III generic soil sites. The spectra at the surface of the Class III sites were computed using the set of rock time histories given in Table 3 and for twenty sets of soil column properties obtained by simulation using the approach discussed in Section 3.4.4.

Because it is conceptually the simplest, we first summarize the results relative to "rock"; i.e., relative to the set of time histories/spectra given in Table 3. Figure 7 shows a comparison of the median amplification factors relative to the rock site category computed for sand-like sites for categories IIa (25'-80'), IIb (80'-180'), IIc (80'-300') and III (deep). Also shown are the amplification factors computed by Joyner-Boore (1982). The match between Category III and Joyner and Boore's results is good. Note that our modeling results give a peak acceleration amplification factors of unity, in agreement with the results obtained by Joyner and Boore and Campbell. There is considerable departure at longer periods (greater than 1.5 seconds). This might be due, in part, to the fact that some of the rock records were not all base line corrected so that they contain some longer period noise. It also might be due to the fact that the damping is not a function of frequency; hence, the long period motion has the same damping as the short period high frequency motion.

A sensitivity analysis indicates that the damping of the soil is one of the most important parameters. We considered three cases for damping of the soil:

1) Median damping value of 2% with COV of 40%.

2) Median damping value of 7% with a COV of 40%.

3) Equivalent linear case with a best estimate curve based on available data.

Figure 8 illustrates the importance of damping for IIb sand like soils by comparing the three damping cases.

As noted earlier, because many of the ground motion models are generic soil models, we also give the amplification factors relative to Category III on Fig. 9 for sand-like sites and Fig. 10 for till-like sites. Also shown is the amplification factors relative to Category III for the rock set. Figure 11 shows the envelope, median and 1-Sigma amplification factors computed for the IIb sand-like category relative to Category III.

One example of the differences in amplification factors that can result between using a category approach and a site specific approach is illustrated in Fig. 12. For the site specific case we selected sand-like Category IIb and greatly reduced the COV used to simulate the site models. We used a COV of 5% on depth and rock shear wave velocity and a COV of 15% for the shear wave velocity of the soil column. We selected these values to be consistent with the range of uncertainty that one would have relative to these parameters at any particular site. Also shown on Fig. 12, for comparison, is the curve for sand-like category IIb from Fig. 7. These amplification factors are relative to rock. As expected, the site specific case has higher amplification factors and less dispersion about the peak. Figure 13 shows similar comparison for the till-like IIb case. For these site types the difference between site specific and the Category approach is similar.

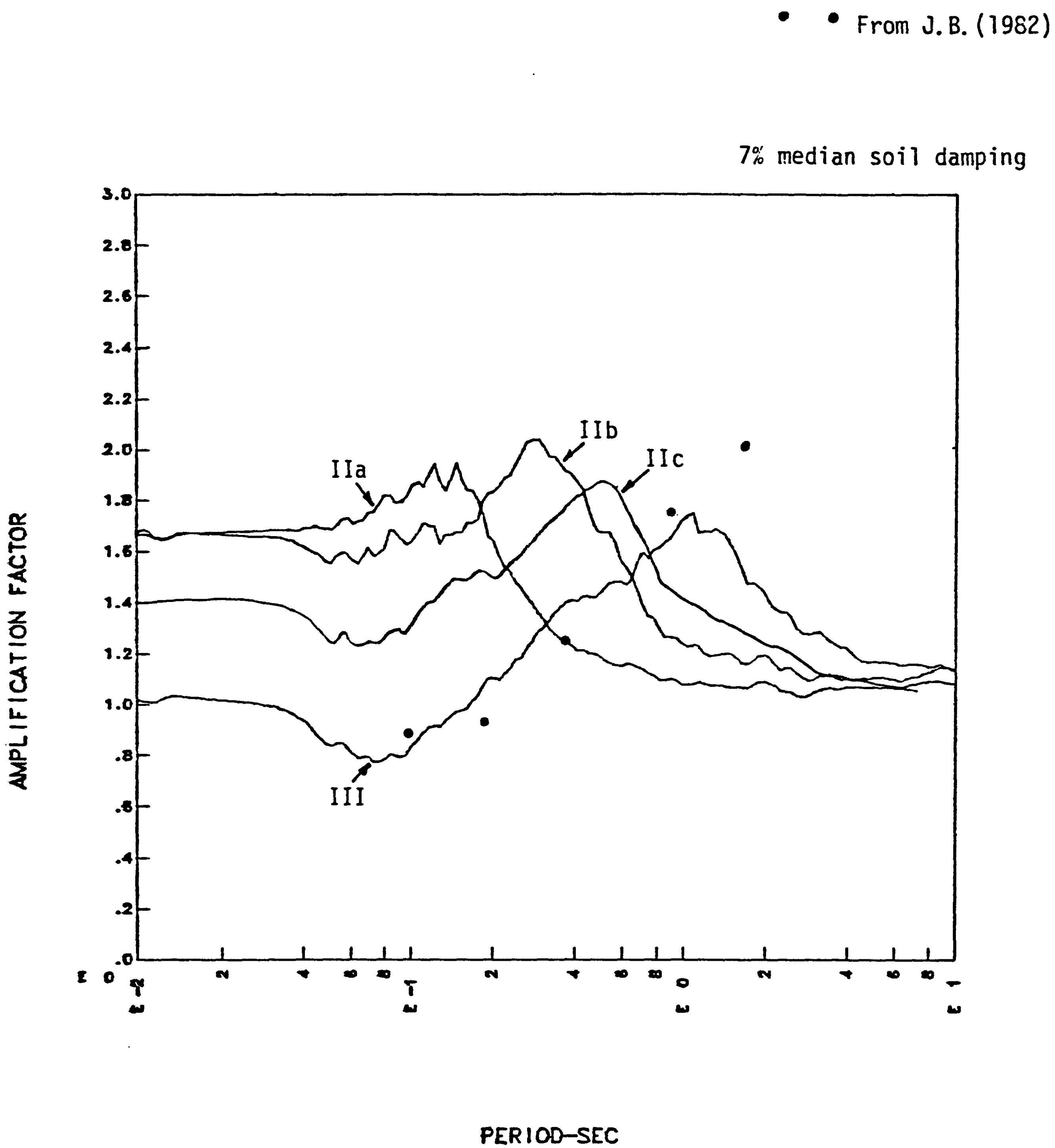


Fig. 7: Comparison of amplification factors for sand-like sites for Categories IIa, IIb and III. Amplification is relative to rock. Also shown is the results from J.B. (1982) (Fig. 6) , Regression Analysis.

1855 ~~1855~~ 40008

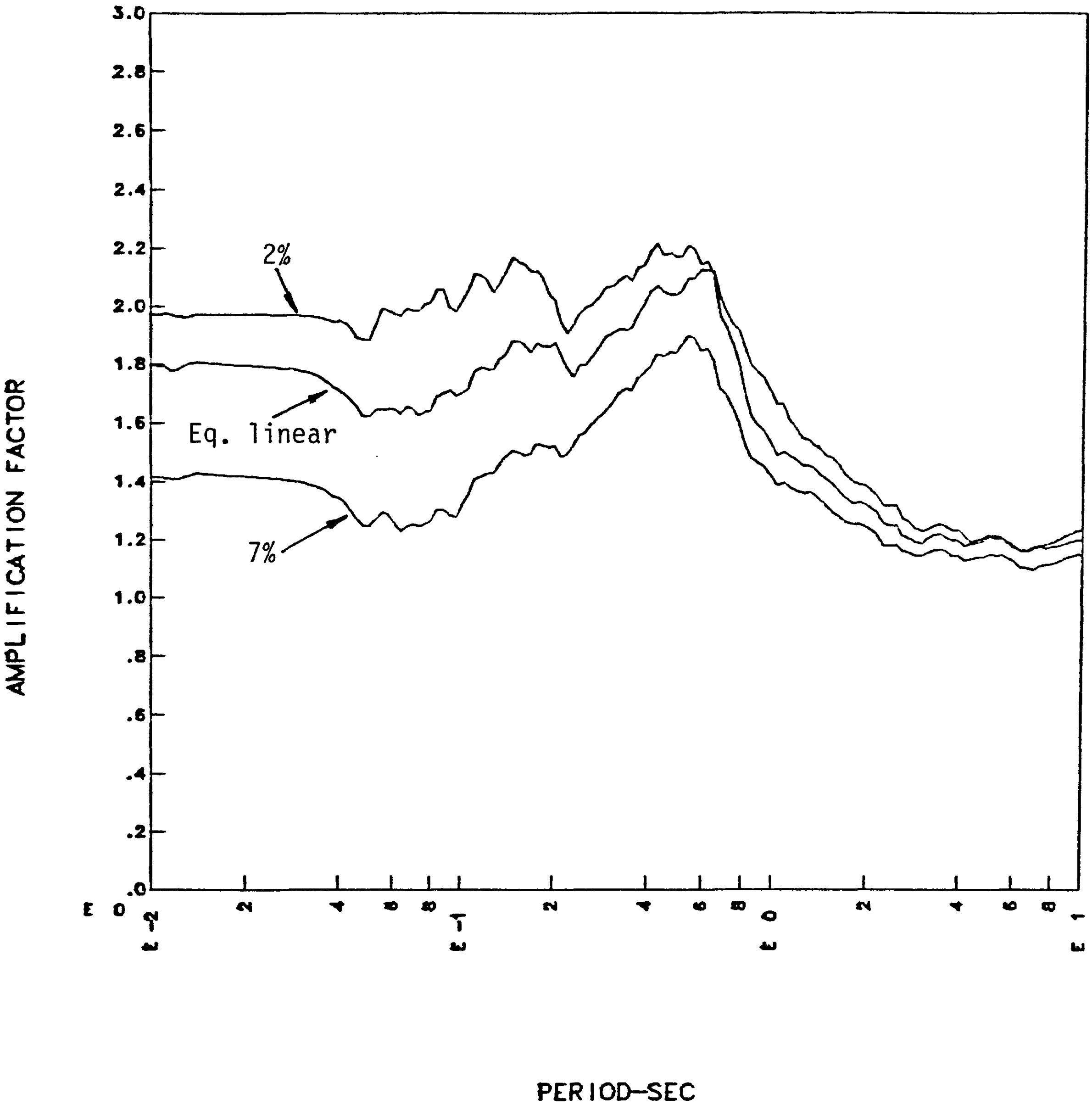


Fig. 8: Importance of damping illustrated for sand-like IIc category.

1856 40008

7% Median Soil Damping

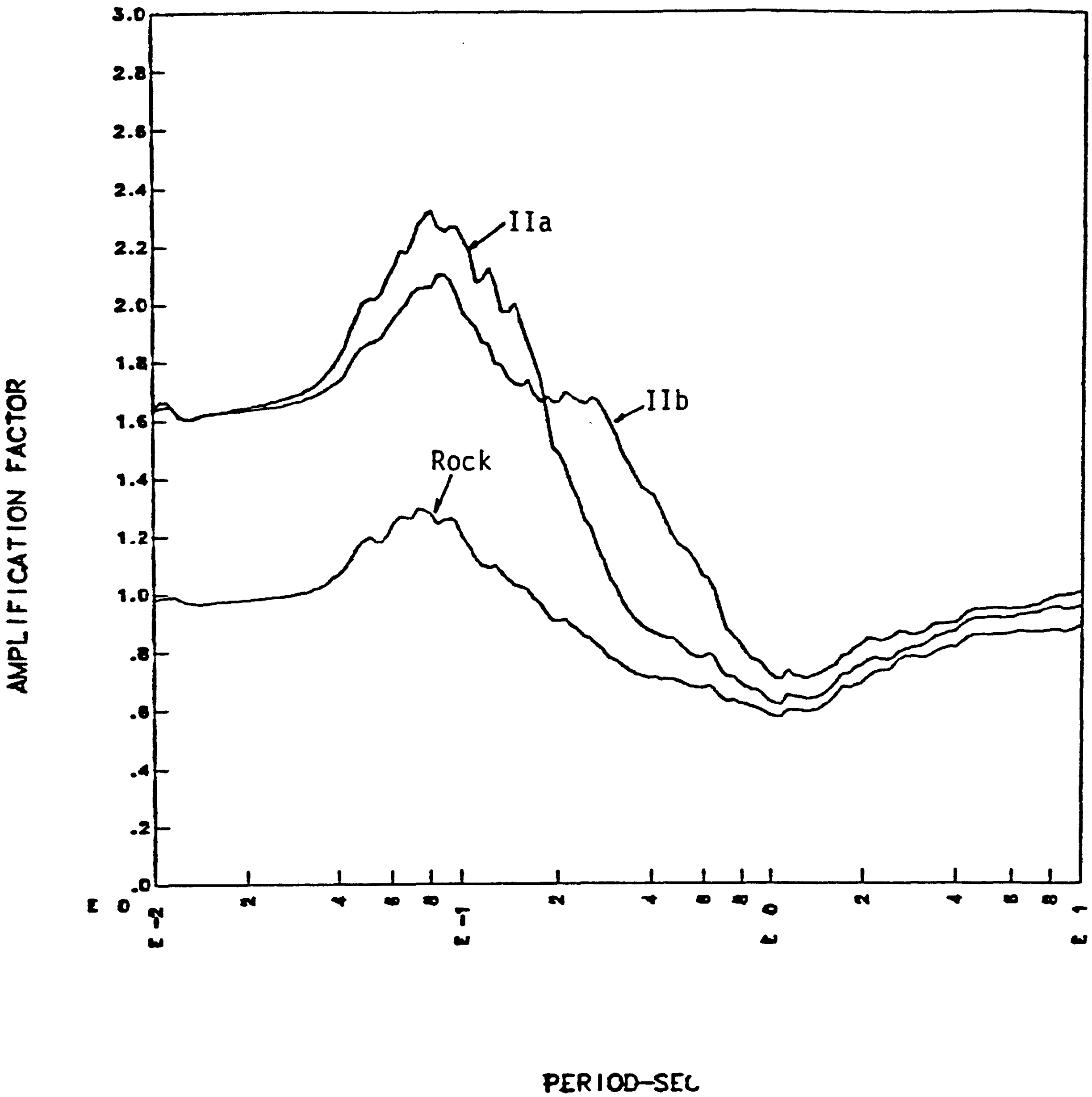


Fig. 9: Amplification factors for sand-like soils relative to Category III. Also shown is the amplification of the rock category relative to Category III.

18B2 40008

7% Median soil damping

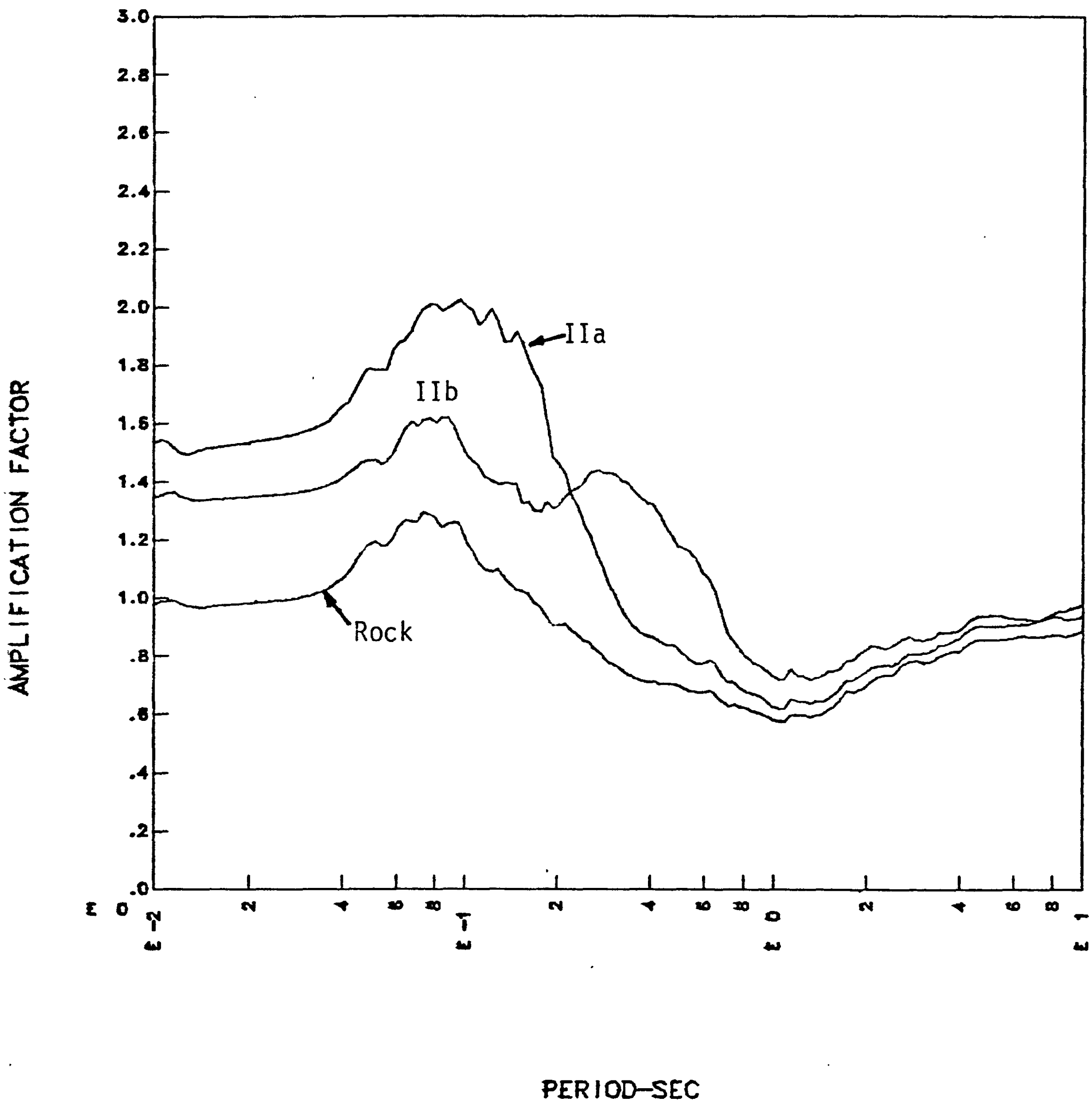


Fig. 10: Amplification factors for till-like soils relative to Category III. The rock category relative to Category III is also shown for reference.

18B3 40008

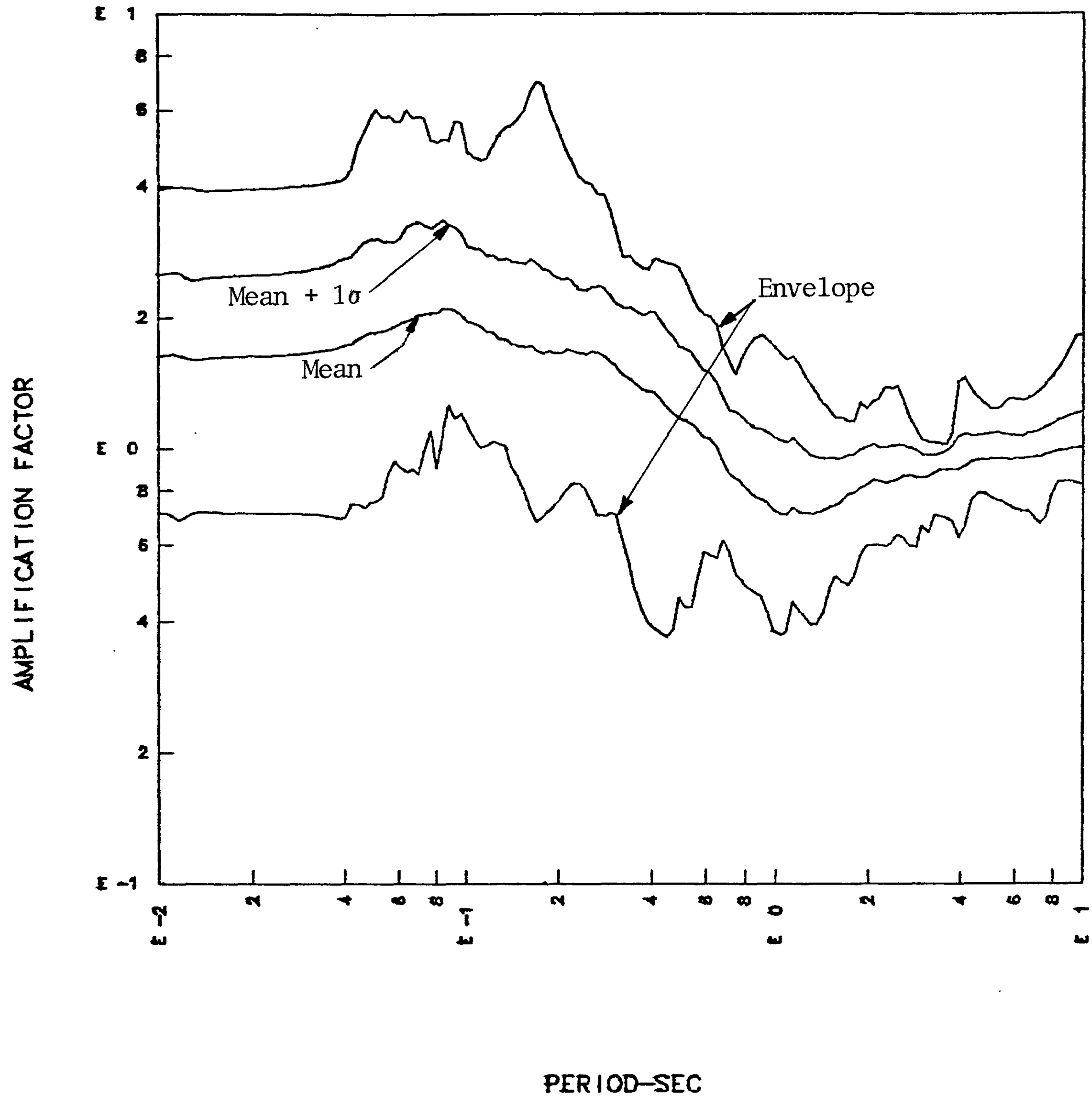


Fig. 11: Increases in uncertainty in the computed amplification factors when Category III is used as reference. Envelope, median and +1σ curves shown for Category IIb sand-like relative to Category III.

1857 40008

"Site Specific" site COV on depth and rock shear wave velocity was reduced to 5%. COV of soil shear wave velocity was reduced to 15% relative to larger values used in the values used for simulation.

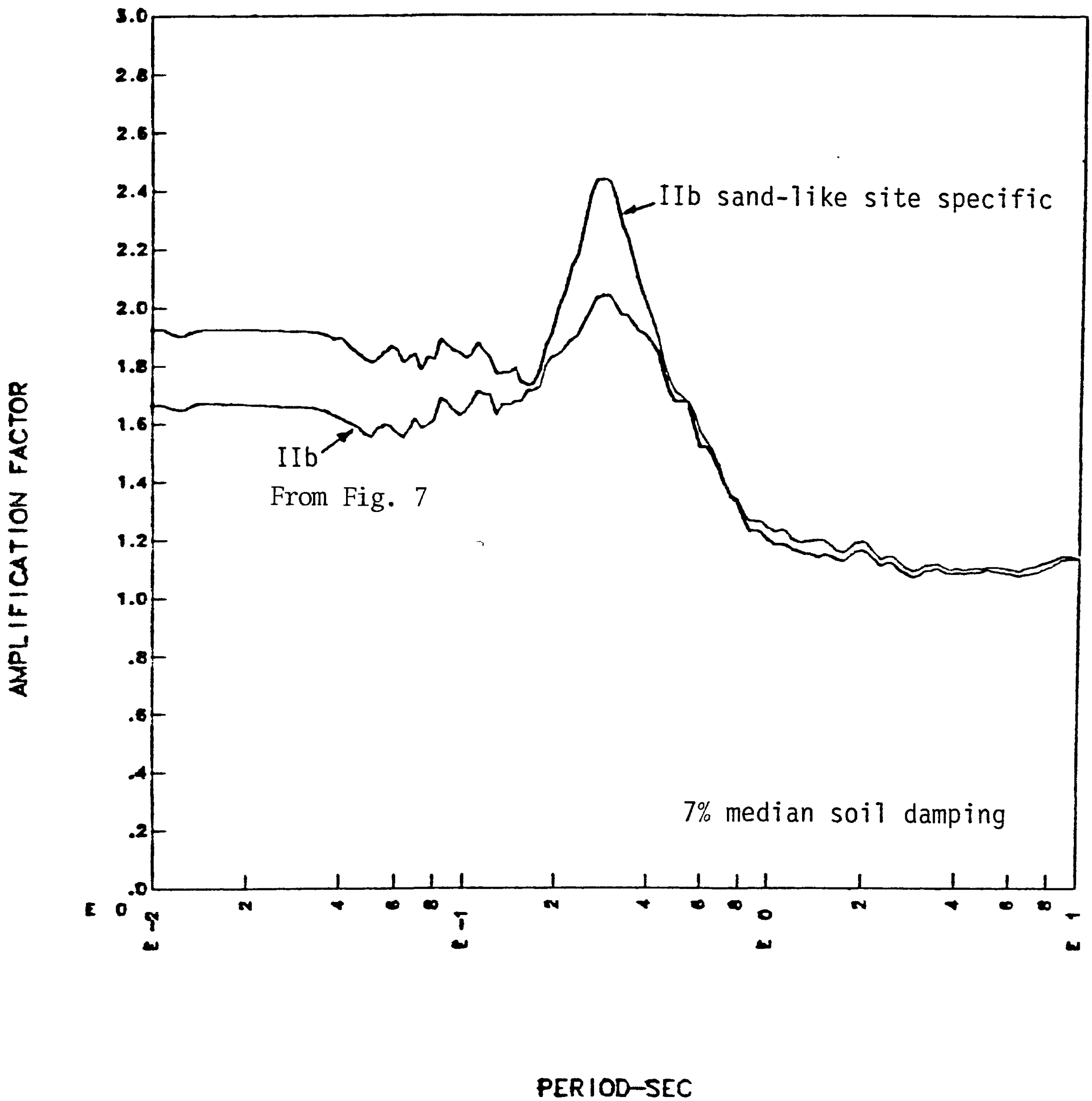


Fig. 12: Comparison of computed amplification factors for a "Site Specific" case category IIb sand-like to the median curve for Category IIb shown on Fig. 7.

1853 40008

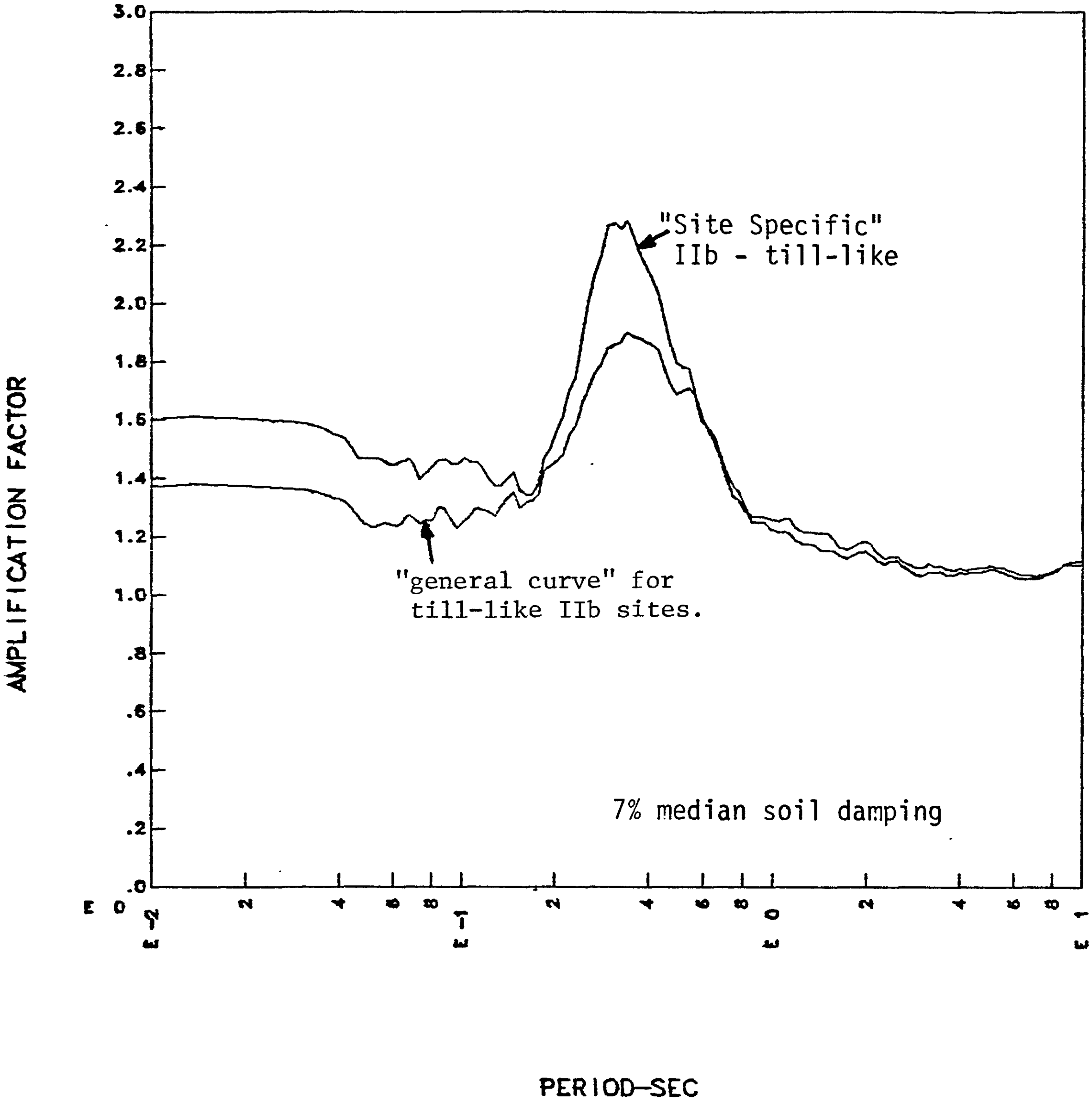


Fig. 13: Comparison of "Site Specific" case for till-like IIb site to general curve for till-like IIb sites

18B7 40008

3.5 DISCUSSION

Interaction with the panel of experts contributing to the project identified several areas needing discussion. In particular it was suggested that surface waves and other non-vertically incident waves could be important. Also focusing and defocusing of rays are not considered.

To include the above considerations would require a very detailed site specific analysis. For a Western U.S. site where the configuration of major nearby active faults is known, it would be possible to perform such complex studies and examine them. It should be noted that such studies are almost beyond the current state-of-the-art and few even limited studies have been performed to address them. For the EUS these questions are even more difficult to assess because it is assumed that the earthquakes occur randomly around the site.

Our proposed approach evolved from the following observations. First, it is very difficult to separate out the wave type in the strong motion accelerograms. In part because strong motion accelerograms are generally recorded within 100km of the source. Our analysis shows that much of the hazard is contributed by earthquakes located within 100km of the site. It must also be kept in mind that we are primarily interested in the high frequency end of the ground motion spectrum--i.e. for frequencies greater than 1 or 2Hz.

Even for distant sources--in particular site amplification observed at sites from underground nuclear explosions (Hays, 1980; Murphy et al, 1971) it has been found that the simple linear theory similar to our proposed approach, has been adequate to explain the important feature of the observed ground motion.

Also to address this issue as part of NRC funded SSMRP, we funded an analysis using earthquake source modeling to (in part) characterize the type and direction of incoming seismic waves at a typical EUS site as compared to WUS site (Apsel, et al, 1980). This analysis found for a soil site (falling into our IIb category with a soil depth of about 100' over a bedrock with a shear wave velocity of 3.3km/Sec) that all of the energy emerged almost vertically at all frequencies. For a deeper soil WUS site the results were much different with waves generally emerging at angles of 20° or more relative to the vertical. It is assumed that he results for a deep soil EUS site would be similar.

The issue of ray focusing is very difficult to deal with given that relatively random nature of seismic activity around any particular site. Except for a few very special potential earthquake locations around a few sites, it is not evident how to even approach this question.

REFERENCES

1. Apsel, R.J., Frazier G.A., Jurkevis, A., and Fried, J.C. Application of Earthquake Source Modeling to Assess the Relative Differences Between Seismic Ground Motion in the Eastern and Western Regions of the United States, and to Characterize the Type and Direction of Incoming Seismic Waves: Final Report. Del Mar Technical Associates, Del Mar, CA, DELTA-R-043, (September, 1980).

2. Bernreuter, D.L., Seismic Hazard Analysis: Application of Methodology, Results and Sensitivity Studies, NUREG/CR 1582, UCRL-53030 (1981).

3. Bernreuter, D.L., Savy, J.B., Mensing, R.W., Chen, J.C. and Davis, B.C. Seismic Hazard Characterization of the Eastern United States, UCID 20421, 2 Volumes, (April, 1985).

4. Bohn, M.P., Shieh L.L., Wells, J.E., Cover, L.C., Bernreuter, D.L., Chen, J.C., Johnson, J.J., Bumpus, S.E., Mensing, R.W., O'Connell, J., and Lappa, D.A., Application of the SSMRP Methodology to the Seismic Risk at the Zion Nuclear Power Plant, NUREG/CR-3428, UCRL-53483, (1984).

5. Campbell, K.W., "Near-Source Attenuation of Peak Horizontal Acceleration", Bull. Seism. Soc. Am., 71, No. 6 pp. 2039-2070, (December 1981)

6. Campbell, K.W. "The effects of Site Characteristics on Near-Source Recordings of Strong-Ground Motion", Proceedings of Conference XXII on Site Specific Effects of Soil and Rock on Ground Motion and the Implications for Earthquake-Resistant Design. USGS open-file report, 83-845, Walt Hays ed., Santa Fe, New Mexico, (July 25-27, 1983).

7. Joyner, W.B. and Boore, D.M. "Peak Horizontal Accelerations and Velocity from Strong Motion Records Including Records from the 1979 Imperial Valley, California Earthquake", Bull. Seism. Soc. Am., 71, No.6, pp 2011-2038 (December 1981).

8. Joyner, W.B. and Boore, D.M. Prediction of Earthquake Response Spectra, USGS open-file report 82-977, (1982).

9. Hays, W.W, Procedure for Estimating Earthquake Ground Motions, USGS Professional Paper 1114, (1980).

10. Murphy, J.R., Davis, A.H. and Weaver, N.L., "Amplification of Seismic Body Waves by Low-Velocity Surface Layers", Bull. Seism. Soc. Am., 61, No.1, pp 109-145, (February 1971).

11. Seed, H.B. and Idriss, I.M., "Influence of Soil Conditions on Ground Motions During Earthquakes", Journal of the Soil Mechanics and Foundation Division, Proceedings of the A.S.C.E., pp 99-137, (January 1969).

12. Trifunac, M.D., (1976a),"Preliminary Analysis of the Peaks of Strong Earthquake Ground Motion-Dependence of Peaks of Earthquake Magnitude, Epicentral Distance, and Recording Site Conditions", Bull, Seism. Soc. Am., 66, No. 1, pp 189-219, (February 1976).

13. Trifunac, M.D., (1976b),"Preliminary Empirical Model for Scaling Fourier Amplitude Spectra of Strong Ground Acceleration in Terms of Earthquake Magnitude, Source-to-Station Distance, and Recording Site Conditions", Bull. Seism. Soc. Am., 66, No. 4, pp 1343-1374, (August 1976).

14. Trifunac, M.D. and Anderson, J.G., "Preliminary Empirical Models for Scaling Absolute Acceleration Spectra", University of Southern California Dept. of Civil Engineering Report, No. CE77-03, (1977).

15. Tucker, B.E. and King, J.L., "Dependence of Sediment-filled Valley Response on Input Amplitude and Valley Properties", Bull. Seism. Soc. Am., 74, No.1, pp 153-166, (February 1984).

TECHNICAL ISSUES ASSOCIATED WITH THE PHENOMENON OF LOCAL SITE AMPLIFICATION OF GROUND MOTION

by
Walter W. Hays
U.S. Geological Survey
Reston, Virginia 22092

ABSTRACT

The patterns and distribution of damage in an earthquake correlate directly with the physical parameters of the total system consisting of: the earthquake source, wave propagation path, local soil-rock column, and the structure. In past earthquakes, the primary cause of damage has often been correlated with the frequency-dependent effects of the local soil-rock column. When the dominant period of the ground motion developed in the rock column is the same as the period of the dominant response of the soil column and the structure, soil-structure interaction occurs. This phenomenon was identified as the cause of severe damage in earthquakes such as the 1967 Caracas, 1970 Gediz, and 1985 Mexico earthquakes. Site amplification, the frequency-dependent response of the soil column to ground motion, is controversial because the strain-dependent properties of the soil control how the soil column filters the input body and surface seismic waves, modifying their peak amplitude, spectral composition, and duration. Consequently, site amplification causes considerable variability in the ground motion recorded at the surface and is an important consideration in hazard and risk assessments. A number of urban areas in the United States (for example, San Francisco, Los Angeles, Seattle, Las Vegas, Salt Lake City, Memphis, St. Louis, Charleston, Boston, and San Juan) have been shown to have soil-rock columns that amplify ground motion under conditions of low-to-intermediate levels of dynamic shear strain. Using the available data, sites in these urban areas can be categorized in terms of the period band of the dominant site response: 1) short period (0.05 - 0.5 seconds), 2) intermediate period (0.5-3 seconds), and 3) long period (3-10 seconds). The overall assessment of the ground-shaking hazard and the risk in each urban area requires careful evaluation of the physical properties of the local soil-rock columns and the structure.

WORLDWIDE DATA ON SITE AMPLIFICATION

Scientists and engineers throughout the world have recognized and documented the occurrence of site amplification (the frequency-and strain-dependent response of a soil-rock colunm to seismic waves) since the 1800's (Macmurdo, 1824; Idriss and Seed, 1968; Seed and Idriss, 1969; Seed and others, 1972; Tezcan and others, 1977; Rosenblueth, 1986; Savy and others, 1986). The classic examples are:

- The 1967 Caracas, Venezuela earthquake,
- The 1970 Gediz, Turkey, earthquake
- The 1976 Friuli, Italy, earthquake
- The 1985 Mexico earthquake.

These earthquakes, and others, have reminded scientists and earthquake engineers that two frequency-dependent phenomena, site response and structural response, are very important considerations in earthquake-resistant design. The most important lessons and facts derived from past earthquakes include:

1. In any city in any earthquake, the characteristics of the earthquake ground motions can vary widely depending on the local soil-rock columns.

2. The damage to a structure at a site in an earthquake is complexly related to the dynamic frequency-dependent properties of the earthquake source, the low-pass filtering characteristics of the wave propagation path, and the band-pass filtering characteristics of both the soil-rock column underlying the structure and the structure (Figure 1). The physical parameters that cause the soil-rock column and the structure to vibrate with the same period contribute most to the potential for damage (Yamahara, 1970).

3. The ground motion recorded in an earthquake at a free-field location is the best dynamic representation of how the ground moved--its time histories of acceleration, velocity, and displacement, spectral composition, level of dynamic strain, and duration of shaking. Physical parameters of the source, propagation path, and soil-rock column contribute distinctive frequency-dependent signatures to these ground motion parameters. For example: a) source - increasing the magnitude increases the peak amplitudes of all

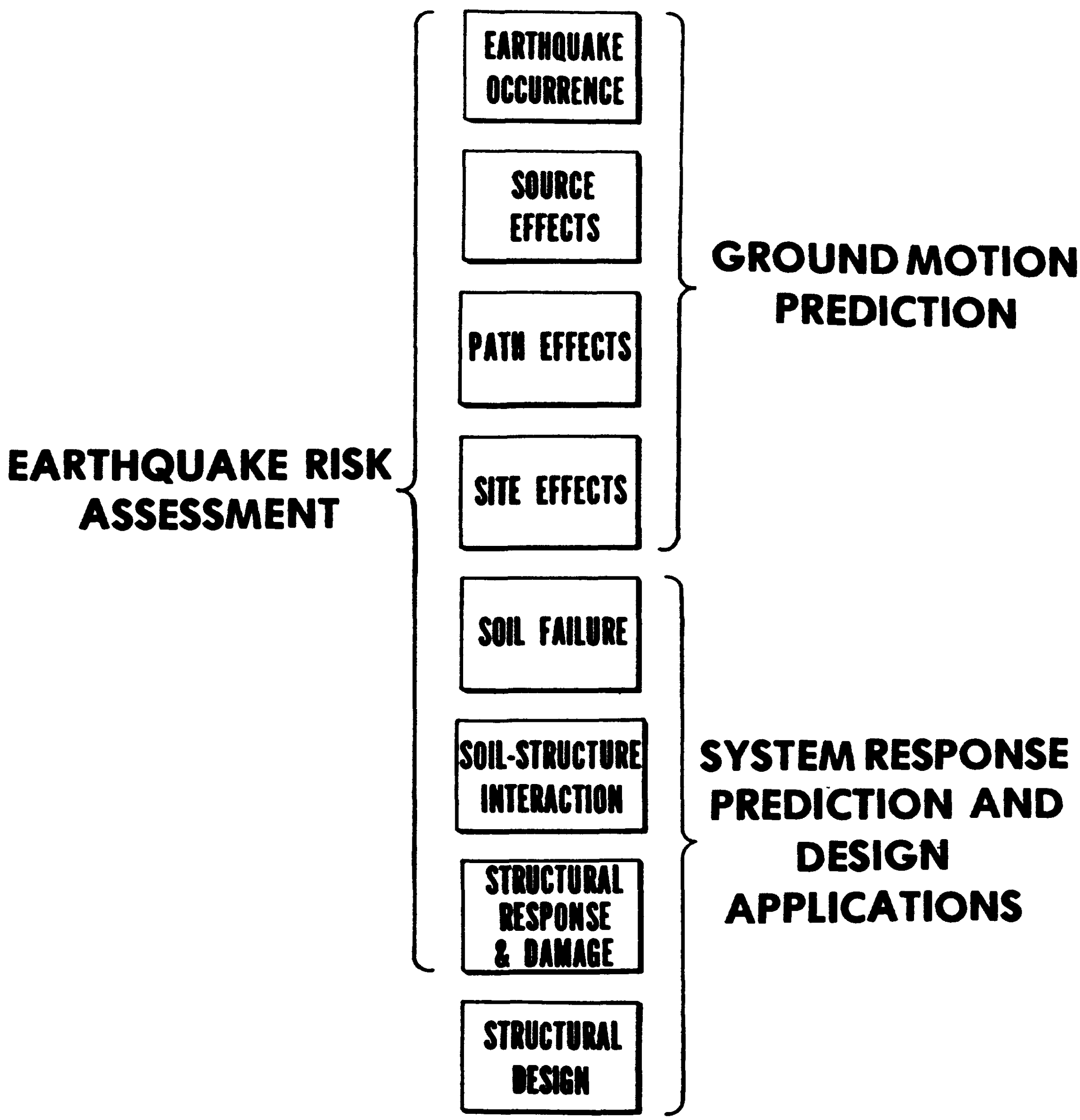


Figure 1.--Schematic illustration of the elements, comprising the earthquake-site-structure system. Physical parameters of the source, path, and the soil-rock column underlying the structure affect the amplitude, spectral compostion, and duration of surface ground shaking.

1958 40008

periods, enhancing the long periods most, b) propagation path - the path acts like a low-pass filter, attenuating the peak amplitude of the short periods more rapidly than the peak amplitudes of the long periods, and c) site - the soil-rock column acts like a band-pass filter, increasing the peak amplitudes of the surface ground motion in a narrow-period band and diminishing it in other period bands (Hays, 1980).

4. The level of dynamic shear strain and its effects on soil properties are the most controversial aspects of site amplification. The level of strain induced in the soil column by the ground motion increases as the magnitude increases and decreases as the distance from the center of energy release increases.

5. The response of the soil-rock column strongly depends on the strain-dependent properties of the soil. Based on the level of dynamic shear strain and the contrast in physical properties of the soil and rock, the soil acts either as an energy transmitter or an energy dissipator. As an energy transmitter, the soil column acts like a band-pass filter, modifying the amplitude and phase spectra of the incident body and surface seismic waves (Murphy and others, 1971) and increasing the duration of shaking (Hays, 1975). As an energy dissipator, the soil column damps the earthquake ground motion, transmitting part of the vibrational energy of both the soil column and the structure back into the Earth and permitting: vertical movement, rocking, and side-to-side movement of the structure on its base (Wolf, 1985).

6. Site amplification, the frequency- and strain-dependent response of the soil-rock column to body and surface seismic waves, increases the surface ground motion in a narrow-period band that is related to the thickness, shear wave velocity, bulk density, properties, and geometry of the soil column. The site transfer function (Figure 2) is a way to categorize the dominant spectral response in terms of the period band where it occurs. Three period bands are typically used to categorize the effect: a) short period (0.05 - 0.5 second, b) intermediate period (0.5 - 3 seconds), and c) long period (3 - 10 seconds). Each period band correlates directly with buildings of various heights. The dominant spectral response for a site underlain by soil has been as much as 1,000 percent greater (factor of 10) than the response for a

1984 40008

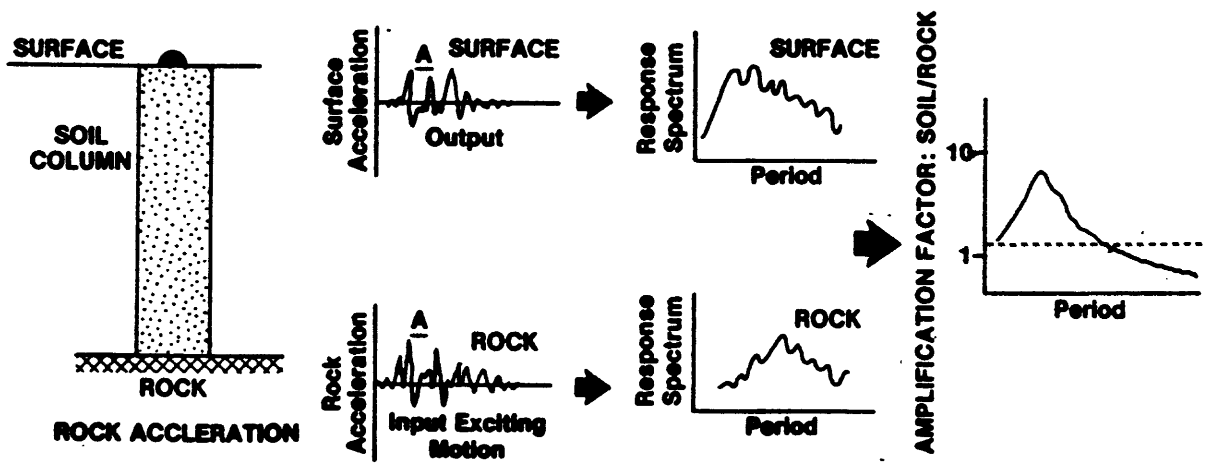


Figure 2.--Schematic illustration of the procedure used to define a site transfer function. The transfer function can also be defined in terms of two surface recordings.

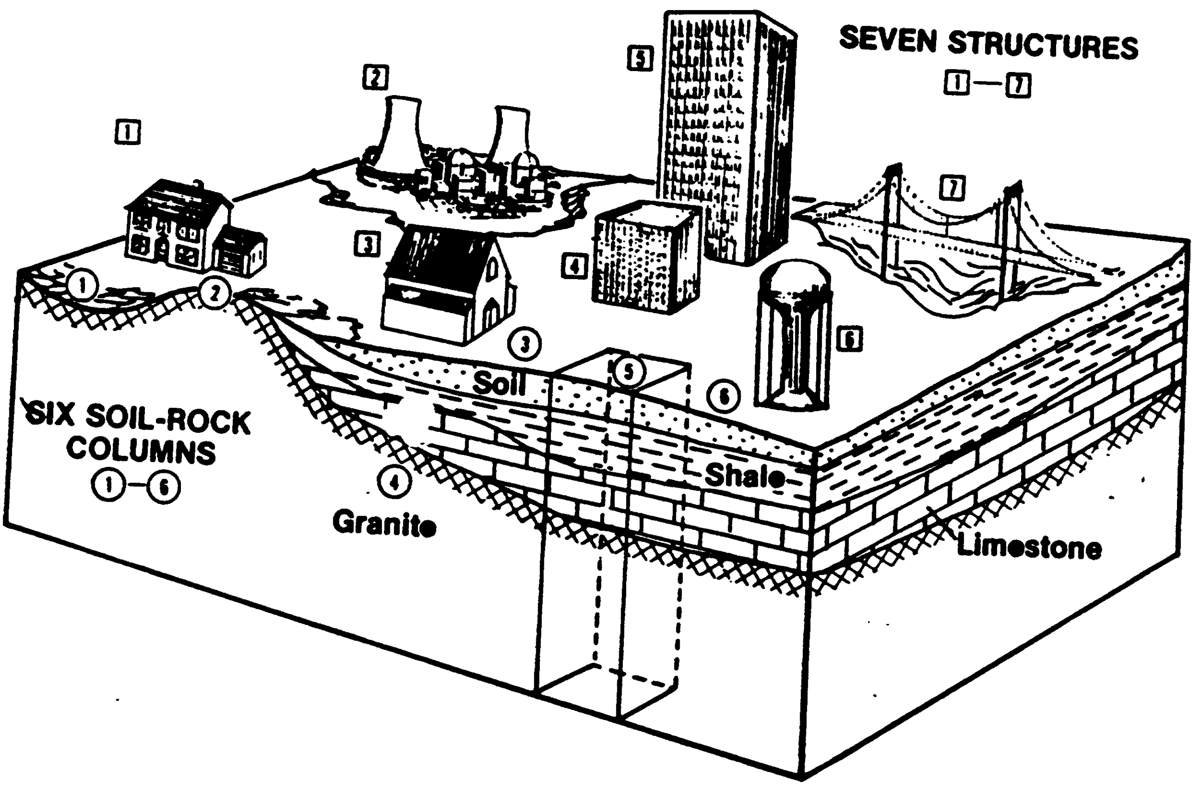


Figure 3.--Schematic illustration of seven typical structures (top) and six typical soil-rock columns (bottom). Each structure and soil-rock column has a natural period of vibration. In earthquake-resistant design, care should be taken to ensure that the natural periods of the response of the soil and structure do not coincide.

1981 40008

site underlain by rock; whereas, the level of peak acceleration (generally caused by body waves and restricted to the short-period band) has been only as much as 250 percent greater (factor of 2.5) (Seed, and others, 1976).

7. The site transfer function depends on many physical parameters, including level of dynamic shear strain, shear wave velocity, density, material damping, thickness, water content, surface and subsurface geometry of the soil-rock column, and the types of seismic waves that excite the soil-rock column--their wavelengths and directions of vibration.

8. The structure also acts like a band-pass filter as it responds to ground motion. The response of the structure can be increased or decreased, depending on the type of structure, the construction materials, the lateral and vertical dimensions, the physical properties of the soil-rock column, and the wavelengths and strengths of the incident seismic waves. The worst case is when the dominant period of the rock motion, the fundamental natural period of vibration of the structure, and the natural period of the soil column are the same, creating a condition of resonance (Figures 3 and 4).

SITE AND BUILDING PERIODS

Earthquake-resistant design must take into account the conditions that cause site amplification of ground motion and damaging soil-structure interaction. Careful evaluation is required to identify the wide range of soil columns and their physical properties, the various types of buildings, and the physical conditions that cause the soil and building responses to occur at or close to the same period.

A soil column, like a building or structure (see Figure 3 and 4), has a natural period of vibration. The characteristic period of vibration T_s of a soil column is given by the relation

$$T_s = \frac{4H}{V_s} \qquad (1)$$

where H is the thickness of the soil column and V_s is the shear wave velocity measured at low levels of strain. Soils, depending on their physical properties,

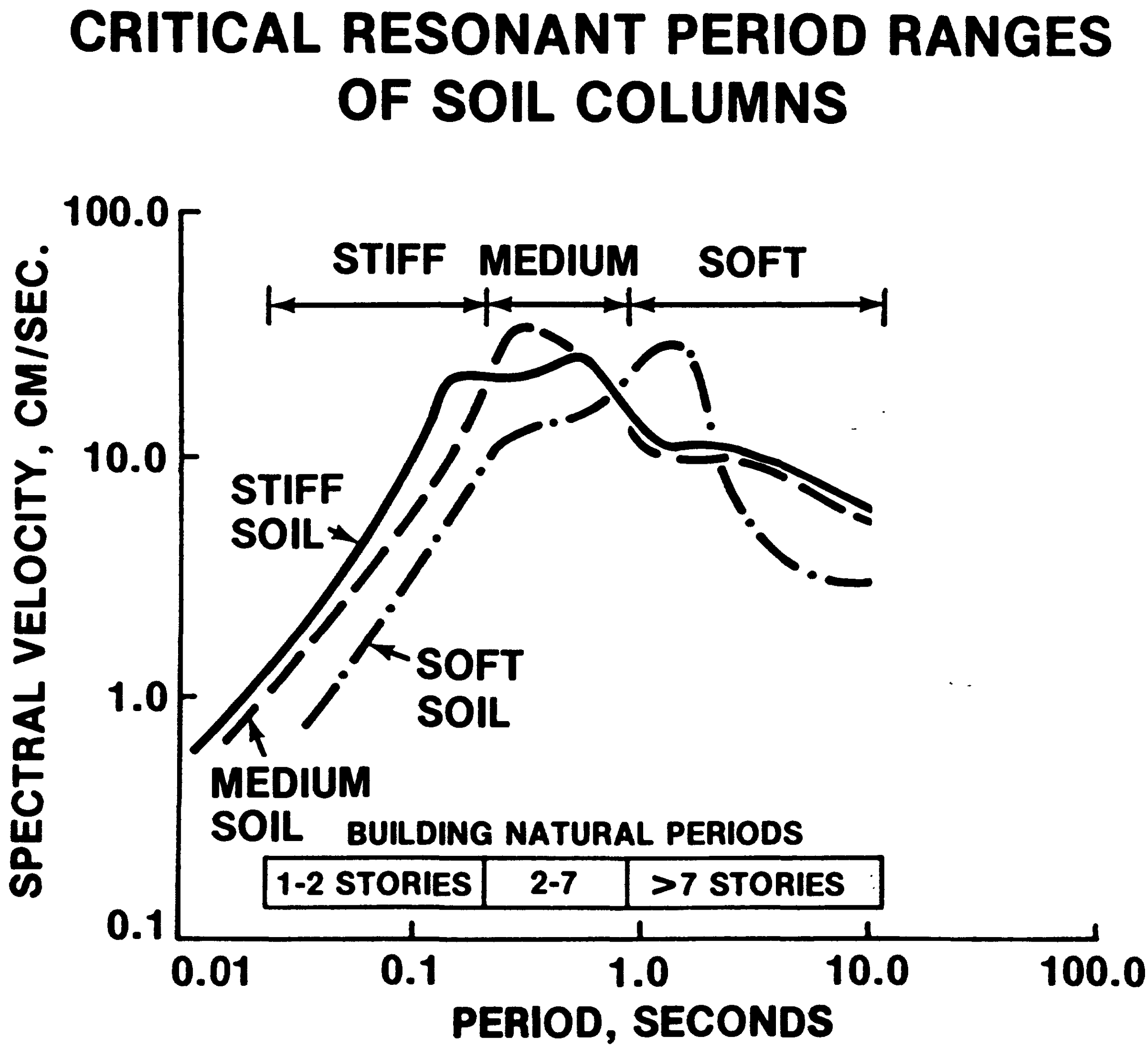


Figure 4.--Schematic illustration showing the period bands where the dominant response of three different types of soil columns occur. The approximate correlation with the response of various types of structures is also shown.

1956 40008

typically have shear-wave velocities ranging from 50 m/sec to 600 m/sec; whereas, rock-like material and rock have shear wave velocities of 765 m/sec or greater.

Soil columns exhibit properties that are strain-dependent. Laboratory tests (Seed and Idriss, 1969) have shown that as the level of dynamic shear strain increases the material damping increases and the modulus of shear decreases. The result is that T_s increases as the level of shear strain increases. The basic relation becomes:

$$T_s = \frac{4H}{RV_s} \qquad (2)$$

where R is an empirical factor (Seed, 1975) having the following values:

-- 0.9 for a magnitude 6 earthquake producing a peak effective acceleration of 0.1 g.

-- 0.8 for a magnitude 6 earthquake producing a peak effective acceleration of 0.2 g.

-- 0.67 for a magnitude 7 earthquake producing a peak effective acceleration of 0.3 - 0.4 g.

The fundamental natural period of vibration T_b of a building is given approximately by the relation

$$T_b = \frac{N}{10} \qquad (3)$$

where N is the number of stories. However, the actual fundamental natural period of a building can be shorter or longer, depending on the engineering design to make the building stiffer or more flexible. Observations from postearthquake investigations have shown that T_b lengthens as the thresholds of various states of damage are reached. In an earthquake the "worst" case for damage is when the value of T_s coincides with T_b. This situation causes resonance of the building and can result in severe damage or collapse unless the building has been designed to withstand the forces generated by this phenomenon.

TECHNICAL CONSIDERATIONS IN THE EVALUATION OF SITE AMPLIFICATION

Evaluation of the potential for site amplification of ground motion requires careful consideration of the following factors:

1. Types of seismic waves - Understanding the physics of local site amplification requires consideration of the ground-motion time histories. Typical horizontal acceleration, velocity, and displacement time histories observed at sites located a distance of about 20 kilometers from the 1971 San Fernando earthquake are shown in Figure 5. These time histories represent the superposition in time of elastic waves which have traveled a wide variety of paths between the earthquake source and the recording site. It is impossible to delineate all of the travel paths involved because one would need to know the details of the geology between the source and the receiver to a depth of perhaps the Mohorovicic discontinuity (i.e., in the order of 30 km (18 mi)). Although this detailed information is not typically available, both theoretical considerations and experience indicate that the seismogram is composed of body and surface waves. The body waves are the familiar compressional (P) and shear (SV and SH) waves which travel from the source to the recording site along paths which extend deep into the Earth's crust. Because of the nature of these travel paths, the energy associated with these wave types is vertically incident on the site geology from below. These waves mainly cause short-period (high-frequency--frequencies greater than 1 Hertz) vibrations which are most efficient in causing low-rise buildings to vibrate. The surface waves (Love and Rayleigh), on the other hand, propagate through channels or wave guides which are bounded above by the surface of the Earth. Thus, they traverse the site geology laterally rather than being incident from below. They mainly cause long-period (low-frequency--frequencies less than 1 Hertz, vibrations which are most efficient in causing high-rise buildings to vibrate. Because the body and surface waves travel at different velocities, they tend to be separated in time on seismograms recorded some distance from the epicenter. In general, both types of elastic waves must be examined in order to evaluate local site amplification effects in a comprehensive manner. However, the type of structure being sited can reduce the scope of the evaluation. For example, surface wave amplification is not typically considered in siting nuclear

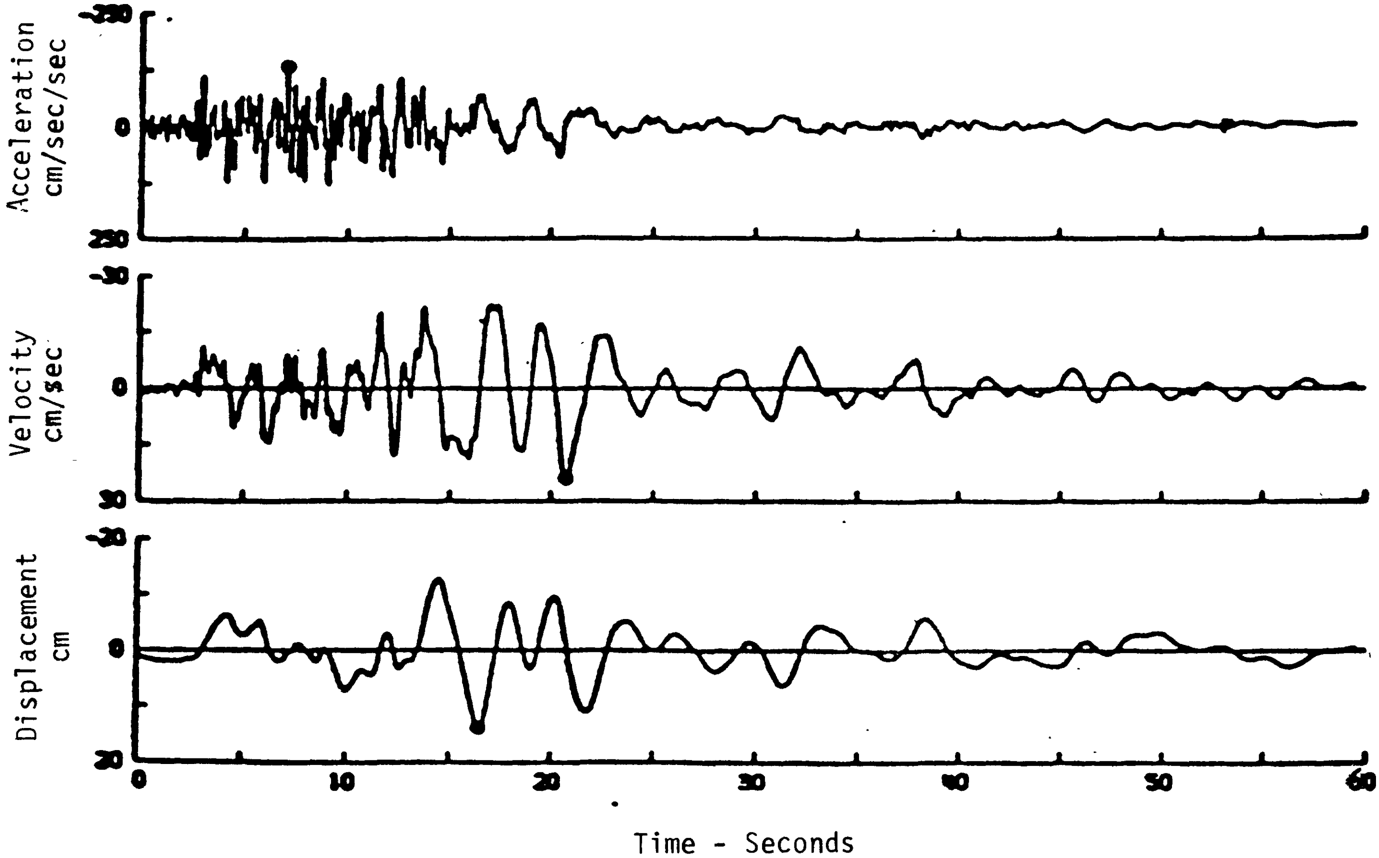


Figure 5.--Horizontal accelerogram recorded at Holiday Inn and the velocity and displacement seismograms derived from it--1971 San Fernando, California earthquake. The peak amplitudes of velocity and displacement are caused by the Rayleigh wave; the peak amplitude of acceleration by body waves.

1957 40008

power plants which are more sensitive to short-period vibrations than long-period vibrations.

2. Level of dynamic shear strain and the dynamic physical properties of the soil column - Careful judgment must be used when assessing the level of dynamic shear strain and its effects on the physical properties of the soil column. One of the sources of controversy comes from the fact that laboratory measurements have demonstrated that soils have shear moduli and damping characteristics that depend on the level of strain. These facts suggest that, under certain conditions, nonlinearities and inelasticities in the soil will attenuate rather than amplify ground motoin at sites underlain by soil. Unfortunately, the high levels of strain produced in the laboratory have not been duplicated by actual strong motion records of past earthquakes. For example, the greatest value of peak ground velocity ever recorded (in the 1971 San Fernando, California and 1979 Imperial Valley, California earthquakes) is 110 cm/sec. Using the empirical rule that

$$\text{Strain} = \frac{\text{peak velocity recorded at the site}}{\text{shear wave velocity of the soil column at the site}} \quad (4)$$

one can conclude that the greatest level of strain induced in soil columns in past earthquakes reached only about 0.5 percent.

Some researchers (for example, Hays and others, 1979; Hays and King, 1982) have shown that site response is essentially linear up to strain levels of about 0.5 percent for some soil-rock columns and that the epicentral distance to the strain level of 0.5 percent is only a few km (about 1 mi) when the shear wave velocity of the soil column is 200 m/sec; less when the shear wave velocity is higher.

Selection of the dynamic properties of the soil is especially complicated below depths of 30 m (100 ft). For the deeper zone, the average shear wave velocity (V_s) can be estimated fairly accurately from values of the compressional wave velocity (V_p) determined from seismic reflection or refraction surveys or from measurements in boreholes, using a value of 0.4 to 0.45 for Poisson's ratio.

3. Thickness of the soil column - Two different points of view have been used to select the critical thickness of the soil column that affects the dominant period of response. One view (Seed, 1975) considers that the soil column can be terminated without appreciable error when rock-like material having a shear wave velocity of about 765 m/sec (2,600 ft/sec) is reached. The other view (Kobayashi and Nagahashi, 1982) considers that the soil column can be terminated without appreciable error only when bedrock having a compressional wave velocity of at least 3,600 m/sec (12,000 ft/sec) is reached. In the first case, surface motions are assumed to be affected mainly by a short soil column, frequently about 30 m (100 ft) thick; whereas, in the second case, surface motions are assumed to be affected by a much thicker soil column.

4. Near field - The near-field is the most complex part of the problem. Analyses of strong ground motion data recorded in the near field (that is, locations within 15 km (9 mi) of the source) have been made by a number of investigators (for example, Idriss, 1978; Hays, 1980; Singh, 1985). For the near field, these analyses indicate that:

   -- Separation of the frequency-dependent effects of the source from the effects of the soil-rock column is very difficult, because the source effects tend to dominate the path and site effects. The directivity of the source appears to cause most of the large variability in the values of peak ground accelerations, peak ground velocity, peak ground displacement, and spectral velocity (Singh, 1985).

   -- A "killer pulse", a pulse of approximately 1 second duration that typically does not have the greatest amplitude but which has the greatest kinetic energy, is generated in some cases in the near field as a consequence of the "fling" of the fault (Bertero and others, 1978). Breakout and stopping phases related to the fault rupture can also occur.

5. Rock Motions - Specification of the ground motions developed in rock by the earthquake source is one of the most difficult tasks in the analysis of site amplification. The frequency-dependent characteristics of the ground motion input to the soil column depend on the details of the

geology of the propagation path, which are usually imprecise. Therefore, analytical calculations must be augmented with a suite of strong motion records acquired in past earthquakes at sites underlain by rock. The ideal data are those from sites underlain by rock located at about the same distance from the zone of energy release and having the same geology for the propagation path as the site being evaluated.

Damage in past earthquakes is worst when the natural periods of both the soil column and the structure are similar to the predominant period of the ground motion developed in the rock.

6. Aftershock ground motion data - Broadband records of the aftershock sequence of past earthquakes can be used, but the strengths and weaknesses of the analysis procedure must be carefully considered. The strength is that aftershock records have the signature of the same travel path and soil-rock columns traversed by the main shock, only the source parameters differ. The weakness is that the lower levels of dynamic shear strain developed in an aftershock may cause overestimation of the amplification factor and underestimation of the dominant period of site respnse.

7. Angle of incidence - Analysts typically assume vertical incidence of the body waves at the base of the soil column. Violation of this assumption, does not introduce significant error (Murphy and others, 1971).

8. Variability in the mean site transfer function - Several investigators (for exmaple, Murphy and others, 1971; Hays, 1980) have shown that the site transfer function in the intermediate and far fields is fairly repeatable. The degree of repeatability of the site transfer function is unknown for both the near field and conditions of strain exceeding 0.5 percent.

EMPIRICAL DATA ON SITE AMPLIFICATION

Worldwide - Scientists and engineers throughout the world have recognized and documented site amplification phenomena since the 1800's (Macmurdo, 1924; Idriss and Seed, 1968; Seed and Idriss, 1969; Seed and others, 1972; Tezcan and others, 1977; Rosenblueth, 1986; Savy and others, 1986). Four classic examples are described below in terms of the spectral response relative to rock and the period band of the dominant site response.

1. The 1967 Caracas, Venezuela, earthquake - Soil-structure interaction occurred in Caracas, 56 km (35 mi) from the epicenter of this moderate (magnitude 6.4) earthquake. Tall buildings (14 stories and greater) sited on soil columns of at least 160 m (520 ft) thickness were damaged severely. The dominant site response occurred in the intermediate period band, centered around 1.2 - 1.6 seconds (Seed and others, 1972).

2. The 1970 Gediz, Turkey, earthquake - Soil-structure interaction caused the collapse of a one-story garage and paint workshop (a part of the Tofias automobile factory) located 225 km (135 mi) from the epicenter of this large (magnitude 7.0) earthquake. The cause was the similarity of the predominant periods of: a) the bedrock motions, b) the respone of the 120-135 m (390-440 ft) column of alluvium, and c) the response of the building, all of which occurred in their intermediate period and centered around 1.2 seconds (Tezcan and others, 1977).

3. The 1976 Friuli, Italy, earthquake - Site amplification of a factor of 4 occurred in the short- to intermediate-period band (0.2 - 0.7 seconds) for a site underlain by 15 m (50 ft) column of alluvium located 25 km (15 mi) from the epicenter. The input rock accelerations ranged from 0.1 g to 0.53 g (Savy and others, 1986).

4. The 1985 Mexico earthquake - This great (magnitude 8.1) earthquake produced two surprises: a) the low value of peak acceleration (0.18 g) in the epicentral region, and b) the high (0.18 g) value of peak acceleration in certain parts of Mexico City located 400 km (250 mi) from the epicenter (Figure 6). Extensive damage occurred to 5- to 20-story buildings sited in the lake bed zone of Mexico City (Rosenblueth, 1986). The largest ground motions in Mexico City occurred at sites underlain by 35 to 50 meter-thick columns of soft lake bed deposits having a shear wave velocity of about 100 m/sec. The dominant site response in the lake bed zone occurred at 2 seconds, an amplification by about a factor of 5 relative to the level of ground motion observed at nearby sites underlain by stiffer,

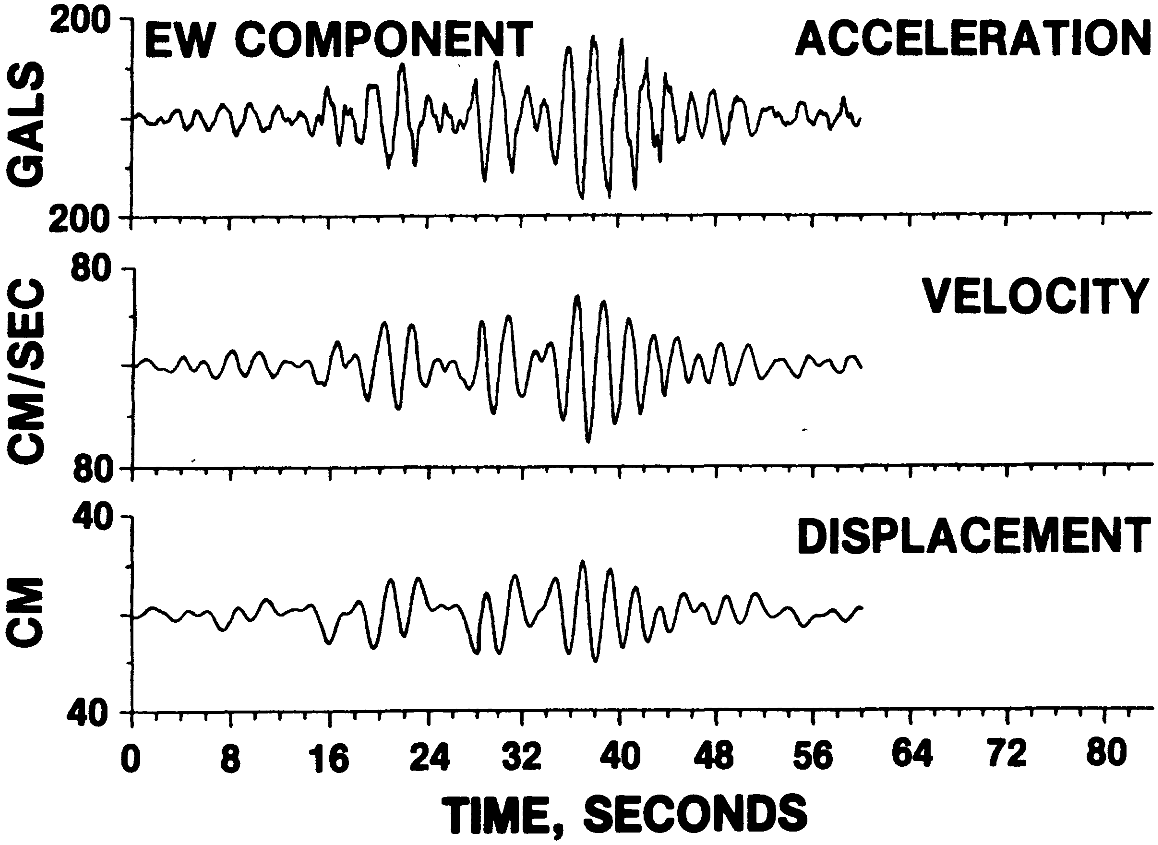


Figure 6.--Accelerogram (top) recorded at a free field location on the surface of the 50-meter thick lake beds forming the foundation in parts of Mexico City. The epicenter of the September 19, 1985, Mexico earthquake was located some 400 km to the west. The strong 2 second period energy in the accelerogram and the velocity (middle) and displacement (bottom) time histories derived from it are a consequence of the filtering effect of the lake beds which amplified the ground motion, (relative to adjacent sites underlain by firmer rock-like materials) about a factor of 5. The coincidence of the dominant period of ground shaking (2 seconds) with the fundamental period of vibration of tall buildings contributed to their collapse. These records were provided by the Universidad nacional Autonoma de Mexico.

2051 40008

rock-like material. Soil-structure interaction occurred at many locations in the lake bed zone, resulting in severe damage and collapse of buildings having a fundamental natural period near 2 seconds.

United States - Since the 1960's, many ivnestigators have studied site amplification phenomena in various parts of the United States. Results obtained in each area are summarized below in terms of spectral response and the period band of the dominant site response:

1. San Francisco Bay Region - The most significant contributors to knowledge on site amplification were: a) the 1906 San Francisco earthquake, b) the 1957 Daly City earthquake, and c) the extensive program of geologic and engineering seismology data acquisition conducted by the U.S. Geological Survey in the 1970's. The most significant results include:

   -- Inferrences from the 1906 earthquake that the soil-rock column underlying a structure can have a significant effect on the surface ground motions and the damage patterns (Wood, 1908).

   -- Strong ground motion data from the 1957 Daly City earthquake that provided a basis for concluding that the amplitude and spectral composition of the ground motions varied as a direct function of the propagation path and the physical properties of the soil-rock column (Idriss and Seed, 1968).

   -- Empirical data showing that each geologic unit in the San Francisco Bay region has a characteristic and predictable response to low-strain seismic excitation (Borcherdt, 1975; Borcherdt and others, 1975; 1978; Joyner and others, 1981).

   -- Empirical data showing that the San Francisco Bay mud exhibits the most spectacular response, amplifying the short-period energy by a factor of 10 or more under conditions of low-strain ground shaking. Other soil-rock columns also cause amplification, mostly in the short- and intermediate-period bands.

2. Los Angeles Region - The most significant contributors to knowledge on site amplification were: a) the 1971 San Fernando earthquake which produced 241 3-component strong motion accelerograms for buildings and for of free-field locations within 75 km (45 mi) of the epicenter of a magnitude 6.4 earthquake, b) the extensive program to monitor the aftershocks of the San Fernando earthquake at more than 100 locations, and c) the comprehensive program of data acquisition conducted by the U.S. Geological Survey in the 1970's and 1980's. Important results included:

-- Site transfer functions derived from ground motion data recorded from the mainshock, selected aftershocks, and nuclear explosions which were similar, even though the levels of rock motions and strain varied markedly (Hays and others, 1979; Rogers and others, 1982, 1985).

-- Amplification of short-period seismic energy amplified along the boundary of the San Fernando Valley, a zone of damage (Hays, 1977) and in Glendale (Murphy an others, 1971b).

-- Amplification of the long-period surface waves by the thick alluvium in the Los Angeles basin (Hanks, 1976).

-- Amplification of the ground motion by some topographic highs (Boore, 1973; Davis and West, 1973).

-- Amplification occurred at soil sites in the Long Beach (Rogers and others, 1982) and Los Angeles areas (Rogers and others, 1985). The short-, intermediate-, and long-period bands were enhanced by factors ranging from 2 to 5, relative to rock.

3. Nevada - The main contributors to knowledge on site amplification was the Ground Motion and Structural Response Program of the U.S. Atomic Energy Commission, conducted in the 1960's and 1970's. More than 3,000 strong motion records were obtained at locations such as Tonopah, Las Vegas, and Beatty where the regional geology and the soil-rock columns were fairly well known. The most significant results included:

2088 40008

-- Documentation of the similarities of the strong ground motion records of earthquake and nuclear explosions within a few hundred miles of the source (Hays, 1975; 1980).

-- Acquisition of site amplification data at locations having a wide range of soil-rock columns (Murphy and others, 1971) and representing levels of strain as great as 0.5 percent (Hays and others, 1979).

-- Demonstration of classic short-period body-wave amplification in Tonopah (Figure 7) where the soil amplification factor was 7 (Murphy and others, 1971).

-- Demonstration of classic intermediate-to-long-period surface wave amplification (Figure 8) in Las Vegas where the soil amplification factor was 10 (Murphy and Hewlett, 1975).

-- Demonstration of site amplification as a function of depth at Beatty where the rock motion was reduced by a factor of 4 at the characteristic site period, T_s (Murphy and West, 1975).

4. Seattle, Washington - Ihnen and Hadley (1984) modeled the strong ground motion of the 1965 Seattle earthquake using a ray tracing technique. Their results indicated that the thick, soft soil deposits of the Duwamish River caused short- to intermediate-period site amplification of a factor of about 5 in western Seattle, the area experiencing the greatest damage in 1965.

5. Wasatch Front, Utah - Salt Lake City, Ogden, and Provo are adjacent to the 370-km-long (222 mi) Wasatch fault zone. These cities are founded on several soil deposits, ranging from coarse gravels and sands close to the Wasatch front to fine grained silts and clays in the valley center. The soils were deposited as lakes filled the Great Salt Lake basin in the Pleistocene epoch. The silts and clays have an average shear-wave velocity of about 200 m/sec; the gravels and sand have a higher velocity.

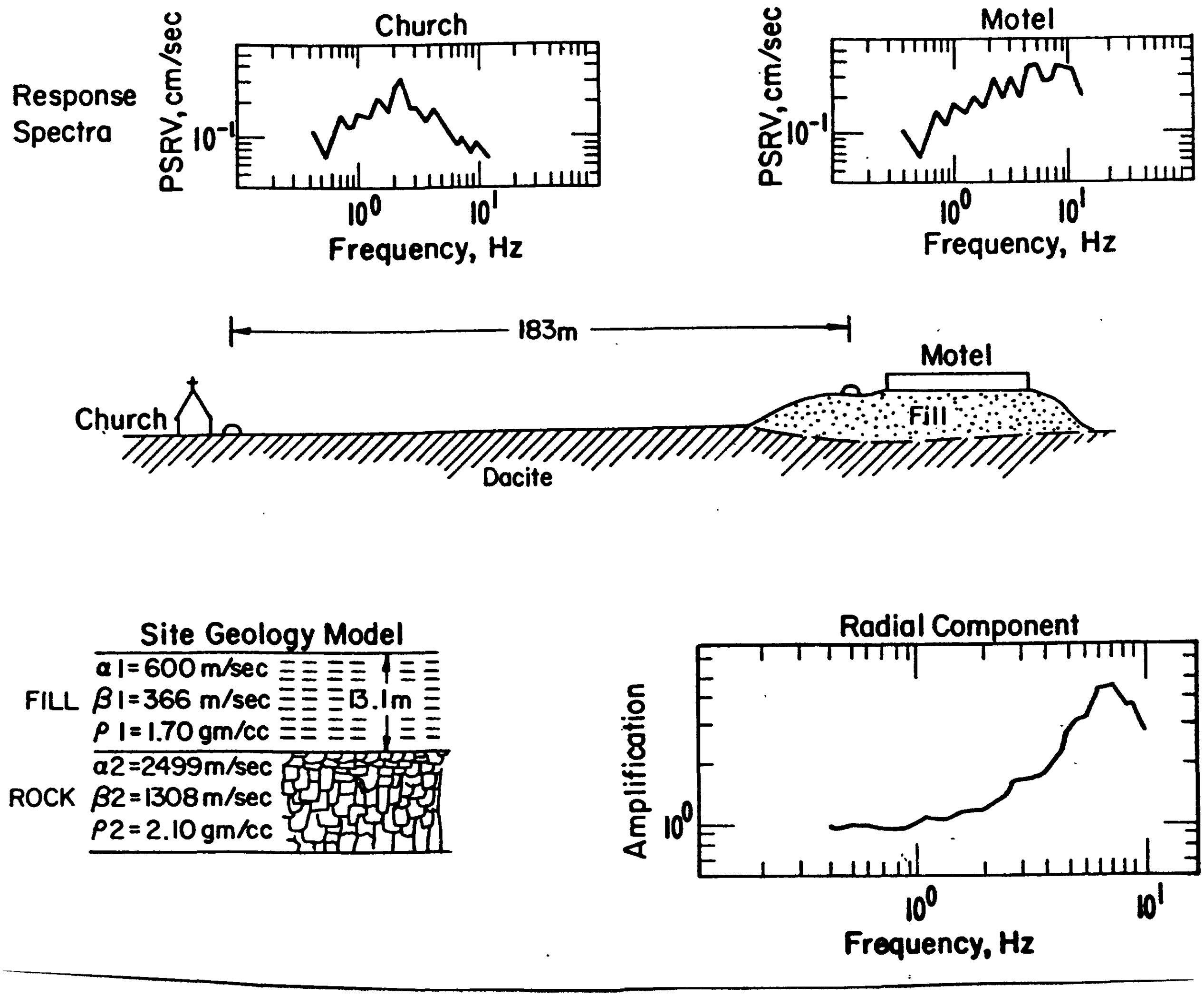


Figure 7.--Example of the frequency-dependent effect of local site geology on nuclear-explosion ground motion, Tonopah, Nevada. The horizontal transfer function, which characterizes the ground response of the site, is defined by dividing the horizontal response spectrum for the soil (motel) site by the corresponding spectrum for the rock (Church) site. The ground response in this example is caused by amplification of SH waves.

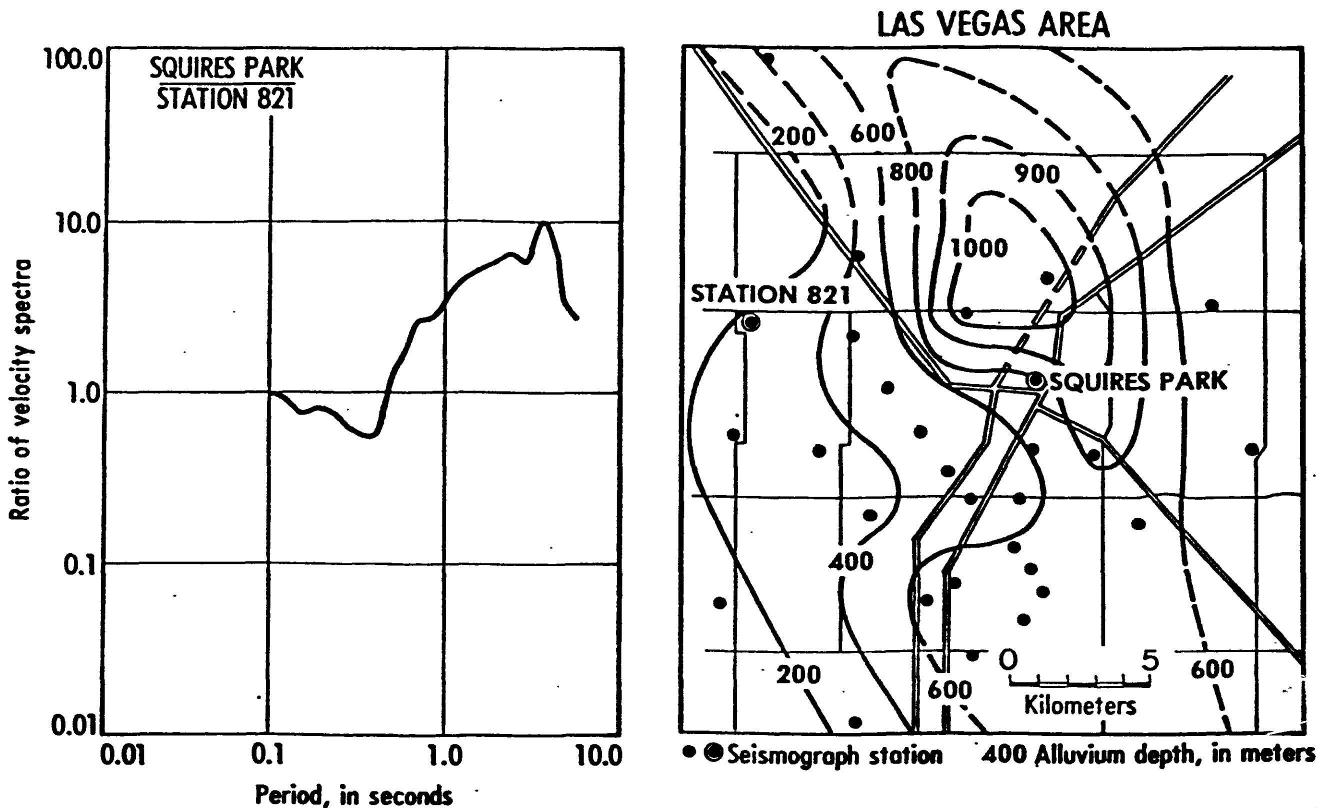


Figure 8.--Example of the frequency-dependent effect of local site geology on nuclear explosion ground motion, Las Vegas Valley, Nevada. The ground response in Las Vegas is controlled by amplification of Rayleigh waves by the varying thickness of alluvium.

2084 40008

Only one strong ground motion record from past earthquakes exists in Utah; it was recorded in Logan in 1965 from the magnitude 5.7 Cache Valley earthquake. In order to define the potential for site amplification from earthquake ground motion at various locations in the Salt Lake City-Odgen-Provo urban corridor, measurements of nuclear-explosion ground motions were made at selected sites using broad band strong motion instruments. These sites were characterized by fairly well defined soil columns ranging in thickness from about 100 m to 1,000 m (328 to 3,280 ft). The underlying rock included limestone, quartz monzonite, shale, and sandstone. Recording about a dozen events over a decade, ground-motion measurements suitable for defining site amplification were made at 40 locations in Salt Lake City, 13 locations in Ogden, 11 locations in Provo, 5 locations in Logan, and 5 locations in Cedar City. The recording sites were located 400-500 km (240-300 mi) from the energy source, so Raleigh waves were dominant on the time histories.

Soil transfer functions were derived in each city from the nuclear explosion ground-motion data (Hays and King, 1982, 1984). Maps of ground response for portions of the short- and intermediate-period bands (0.05-3 seconds) were prepared to show the spatial variation of ground motion. These data showed that:

-- The level of site amplification (relative to a site underlain by rock on the Wasatch front) increases with distance from the Wasatch fault zone and offsets the normal decay of peak amplitude with distance. This effect dominated out to an epicentral distance of about 30 km (18 mi).

-- The dominant site response occurs around 1 second. The low-strain response in the short- and intermediate-period bands, the dominant periods of response of low to mid-rise buildings, is enhanced at all sites underlain by soil.

-- The level of site amplification in the short-period band (0.2-0.7 seconds) is as much as a factor of 10 greater at sites underlain by the thick columns of clay and silt in the center of the valleys. The level is less--about a factor of 2--when the site is underlain by the thinner columns of coarse sands and gravels near the Wasatch front (Figure 9). The potential for soil-structure interaction exists in parts of Salt Lake City, Odgen, and Provo.

6. Other locations--Although the data are sparse, a number of investigators have shown in a preliminary way that Boston (Whitman, 1983), Memphis (Sharma and Kovacs 1980), Charleston (Elton and Martin, 1986), and San Juan (Molinelli, 1985) have soil-rock columns that will cause site amplification under conditions of low-to-intermediate-strain ground shaking. The effects are restricted to the short- and intermediate-period bands.

SUMMARY OF SITE AMPLIFICATION EFFECTS

The effects of site amplification in various urban areas of the United States can be categorized in terms of the period band where the dominant response occurs. The table below gives this informtion.

	Period Band of Dominant Response		
Location	Short	Intermediate	Long
San Francisco	x	x	
Los Angeles	x	x	x
Seattle	x	x	
Las Vegas	x	x	x
Salt Lake City	x	x	
Memphis	x	x	
St. Louis	x	x	
Charleston	x	x	
Boston	x	x	
San Juan	x	x	

RECOMMENDATIONS FOR FUTURE RESEARCH

The research on site amplification must be continued to reduce the variability and to eliminate uncertainty and controversy that affects implementation. The following technical issues are not completely answered and require research:

1. To what degree do site amplification phenomena (peak amplitudes of ground motion, spectral composition, duration) derived from small earthquakes or

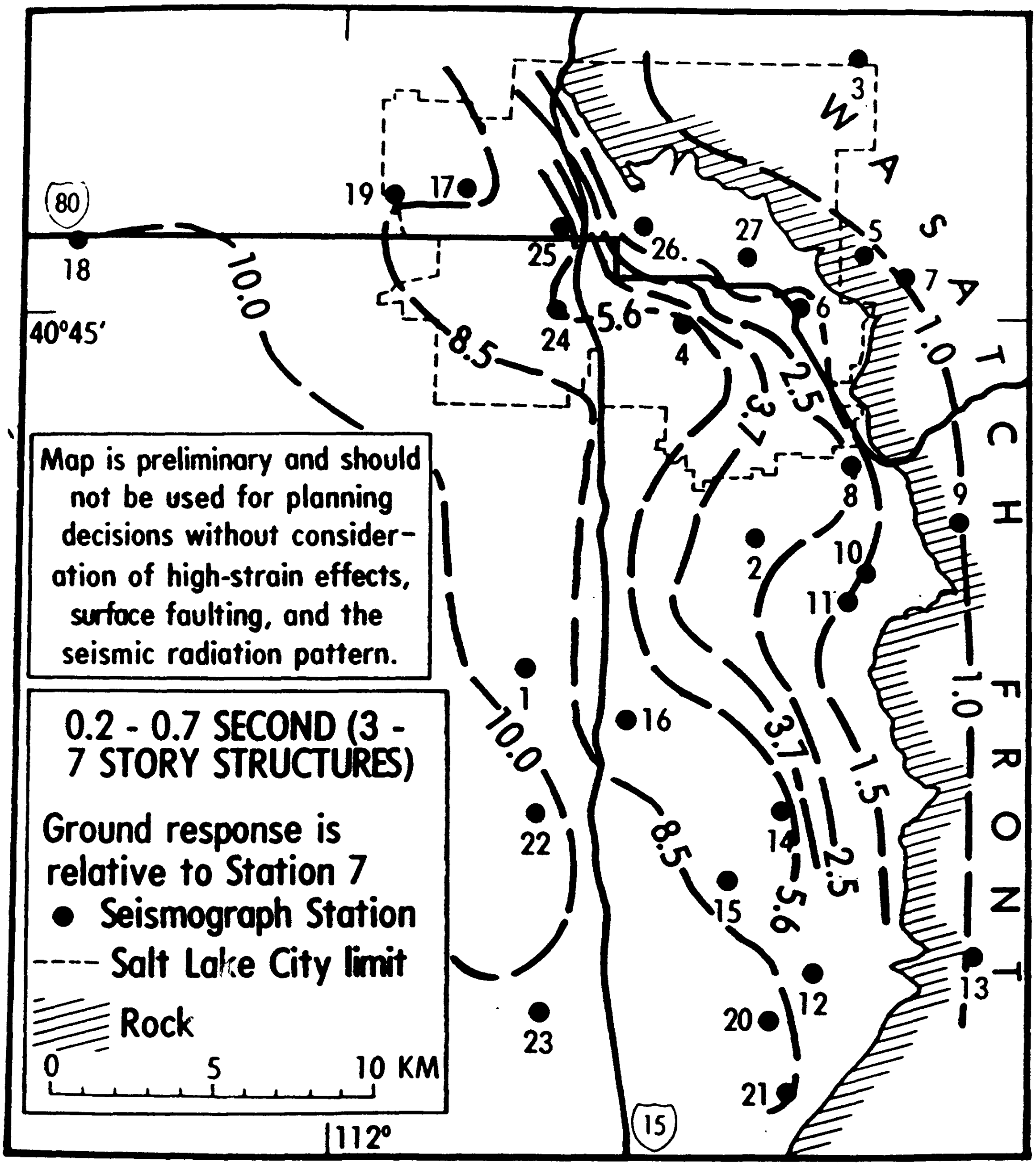


Figure 9.--Example of soil amplification factors for the period band 0.2-0.7 seconds, Salt Lake City, Utah. These factors were derived from low-strain nuclear-explosion ground-motion data.

2056 40008

distant nuclear-explosion ground-motion data represent the phenomena expected to occur in future large earthquakes?

2. Will the characteristics of peak amplitude, spectral composition, and duration of ground shaking that are controlled by physical parameters of the local soil/rock columns vary significantly from those derived from either small earthquakes or nuclear-explosion ground motion data in the case of either a distant or a nearby earthquake having magnitudes of 6 to 7.5?

3. How sensitive are site amplification phenomena to the level of dynamic shear strain? To what extent should strain-dependent phenomena be incorporated in building code zoning maps and design criteria for other structures?

4. To what degree can site amplification effects be represented accurately in hazard maps in building codes?

The research should integrate the empirical ground motion data base with analytical models to define the ground-shaking hazard more completely. Additional ground moiton data from earthquakes in the area or in areas having an analogous seismotectonic setting should be acquired as soon as possible to quantify the variability of site effects in the overall assessemnt of the ground shaking hazard.

REFERENCES

Bertero, V. V., Mahin, S. A., and Herrera, R. A., 1978, Aseismic design implications of near-field San Fernando records: Journal of Earthquake Engineering and Structural Dynamics, v. 6, pp. 31-42.

Boore, D. M., 1973, The effect of simple topography on seismic waves: implications for the accelrations recorded at Pacoima Dam, San Fernanco Valley, California: Seismological Society of America Bulletin, v. 63, pp. 1603-1610.

Borcherdt, R. D., 1975, Studies of seismic zonation of the San Francisco Bay region: U.S. Geological Survey Professional Paper 941-A, 102 p.

Borcherdt, R. D., Joyner, W. D., Warrick, R. E., and Gibbs, J. F., 1975, Response of local geologic units to ground shaking, in Borcherdt, R. D., (Editor), Studies for Seismic Zonation of the San Francisco Bay Region: U.S. Geological Survey Professional Paper 941-A, pp. 52-67.

Elton, D. J., and Martin, J. R., 1986, Site Period Study for Charleston, S.C.: 3rd U.S. Conference on Earthquake Engineering, Proceedings, Charleston, S.C. (in press).

Hanks, T. C., 1976, Observations and estimation of long period strong ground motion in the Los Angeles Basin: International Journal of Earthquake Engineering and Structural Dynamics, v. 4, pp. 473-488.

Hays, W. W., 1975, A note on the duration of earthquake and nuclear explosion ground motions: Seismological Society of America Bulletin, v. 65, pp. 875-844.

Hays, W. W., 1977, Evaluation fo the seismic response in the Sylmar-San Fernando area, California, from the 1971 San Fernando earthquake: American Society of Civil Engineers, Mechanics Division Specialty Conference on Dynamic Response of Structures, Los Angeles, Proceedings, pp. 502-511.

Hays, W. W., 1980, Procedures for estimating earthquake ground motions: U.S. Geological Survey Professional Paper 1114, 77 p.

Hays, W. W., and King, K. W., 1984, Seismic microzoning along the Wasatch fault zone, Utah: World Conference on Earthquake Engineering, 8th, Proceedings, v. 1, pp. 1-12.

Hays, W. W., and King, K. W., 1982, Zoning of the earthquake ground-shaking hazard along the Wasatch fault zone, Utah: International Earthquake Microzonation Conference, 3rd Proceedings, v. 3, pp. 1307-1317.

Hays, W. W., Rogers, A. M., and King, K. W., 1979, Empirical data about local ground response: Earthquake Engineering Research Institute, National Conference on Earthquake Engineering, 2nd, Proceedings, pp. 223-232.

Idriss, I. M., 1978, Characteristics of earthquake ground motions: Earthquake Engineering and Soil Dynamics, Proceedings of Specialty Conference, American Society of Civil Engineers, v. 3, pp. 1151-1267.

Idriss, I. M., and Seed, H. B., 1968, Analysis of ground motions during the 1957 San Francisco earthquake: Seismological Society of America Bulletin, v. 58, pp. 2013-2032.

Ihnen, S., and Hadley, D. M., 1984, Prediction of strong ground motion in the Puget Sound region: The 1965 Seattle earthquake: Sierra Geophysics Report, SGI-R-84-113 to U.S. Geological Survey, 38 p.

Joyner, W. B., Warrick, R. E., and Fumal, T. E., 1981, The Effects of Quaternary Alluvium on Strong Ground Motion in the Coyote Lake, California Earthquake of 1979: Seismological Society of American Bull., v., 71, p. 1333-1349.

Kobayaskhi, Hiroyoshi, and Nagahashi, Sumio, 1982, Response spectra on seismic bedrock during earthquakes: Engineering Seismology, Tokoyo Institute of Technology, pp. 22-27.

Macurdo, J., 1824, Papers relating to the earthquake which occurred in India in 1819, Philadelphia Magazine, v. 63, pp. 105-177.

Molinelli, Jose, 1985, Earthquake Vulnerability Study for the Metropolitan Area of San Juan, Puerto Rico: in Hays, W. W., and Gori, P. L., (Editors), A Workshop on Reducing Potential Losses from Earthquakes in Puerto Rico, Proceedings of Conference XXX, U.S. Geological Survey, Open-File Report 85-731, p. 211-277.

Murphy, J. R., and Hewlett, R. A., 1975, Analysis of seismic response in the city of Las Vegas, Nevada: A preliminary microzonation: Seismological Society of America Bulletin, v. 65, pp. 1575-1598.

Murphy, J. R., Lynch, R. D., and O'Brien, L. J., 1971b, Predicted San Fernando earthquake spectra: Environmental Research Corporation Report NVO-1163-TM30, to U.S. Atomic Energy Commission, 38 p.

Murphy, J. R., and West, L. R., 1974, An analysis of surface and subsurface seismic measurements demonstrating the amplification effect of near-surface geology: Environmental Research Corporation Report NVO-1163-TM41 to U.S. Atomic Energy Commission, 21 p.

Murphy, J. R., Weaver, N. L., and Davis, A. H., 1971a, Amplification of seismic body waves by low-velocity layers: Seismological Society of America Bulletin, v. 61, pp. 109-146.

Rogers, A. M., Tinsley, J. C., and Borcherdt, R. D., 1985, Predicting relative ground response, in Ziony, J. F., (Editor), Evaluating earthquake hazards in the Los Angeles region--An earthquake science perspective: U.S. Geological Survey Professional Paper 1360, pp. 221-248.

Rosenblueth, Emilio, 1986, The Mexican earthquake: A first-hand report: Civil Enginering, American Society of Civil Engineers, New York, January, pp. 38-40.

Savy, Jean, Bernreuter, Don, and Chen, J. C., 1986, Site effects: A generic method for modeling site effects in seismic hazard analyses, in Hays, W. W., (Editor), Proceedings of Conference XXXIV, U.S. Geological Survey Open File Report 86- , 35 p.

Seed, H. B., 1975, Design provisions for assessing the effects of local geology and soil conditions on ground and building response during earthquakes, in new earthquake design provisions: Proceedings of seminar sponsored by Professional DEvelopment Committee of Structural Enginers Association of Northern California and San Francisco Section of America Society of Civil Engineers, pp. 38-63.

2053 40008

Seed, H. B., and Idriss, I. M., 1969, Influence of soil conditions on ground motions during earthquakes: Journal on Soil Mechanics Foundations Division of America Society of Civil Engineers, v. 95, pp. 1199-1218.

Seed, H. B., Murarka, R., Lysmer, John, and Idriss, I. M., 1976, Relationships of maximum acceleration, maximum velocity, distance from source, and local site conditions for moderately strong earthquakes: Seismological Society of America Bulletin, v. 66, pp. 1323-1342.

Seed, H. B., Whitman, R. V., Dezfulian, H., Dobry, R., and Idriss, I. M., 1972, Soil Conditions and building damage in the 1967 Caracas earthquake, Journal of the Soil Mechanics Foundations Division, America Society of Civil Engineers, v. 98, pp. 787-806.

Singh, J. P., 1985, Earthquake ground motions: Implications for designing structures and reconciling structural damage: Earthquake Spectra, v. 1, pp. 239-270.

Sharma, S., and Kovacs, W. D., 1980, The Microzonation of the Memphis, Tennesseee Area; U.S. Geological Survey Open-File Report 80-914, 129 p.

Tezcan, S. S., Seed, H. B., Whitman, R. V., Serff, N., Christian, J. T., Durgunoglu, H. T., and Yegian, M., 1977, Resonant period effects in the Gediz, Turkey earthquake of 1970: Earthquake Engineering and Structural Dynamics, v. 5, pp. 157-179.

Whitman, R. V., 1983, An Engineer's Perspective on the Cape Anne, Massachusetts Earthquake of 1755: Earthquake Notes, v. 54, no. 1, p. 8 (Abstract).

Wolf, J. P., 1985, Dynamic Soil-Structure Interaction, Prentice-Hall Publishing Company, Englewood Cliffs, New Jersey, 466 p.

2087 40008

Wood, H. O., 1908, Distribution of Apparent Intensity in San Francisco, in The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission, Carnegie Institution of Washington, Washington, D.C., pp. 220-245.

Yamahara, H., 1970, The interrelation between frequency characteristics of ground motion and earthquake damage to structure: Soils and Foundations, v. 10, pp. 57-74.

THE LLNL APPROACH TO SEISMIC HAZARD ESTIMATION IN AN ENVIRONMENT OF UNCERTAINTY*

BY

D. L. Bernreuter, J. B. Savy, and R. W. Mensing
Lawrence Livermore National Laboratory
P.O. Box 808, Livermore, CA

1.0 INTRODUCTION

Because the seismic hazard at a site depends on the seismicity in the vicinity of the site and the attenuation of the ground motion between the source of the earthquake and the site, an estimate of the seismic hazard at the site depends on having:

- o a description of the seismicity of the region affecting the site.
- o a model for the attenuation of the ground motion between the source and the site.

The main sources of information about seismicity have been based on:

- o physical information such as
  - catalogs of past events
  - maps/trends of past events
  - tectonic structures/states of stress information
  - geophysical data
- o mathematical models such
  - models of tectonic/stress processes
  - distributions of magnitudes
  - recurrence models relating frequency to magnitude

---

*"This work was supported by the United States Nuclear Regulatory Commission under a Memorandum of Understanding with the United States Department of Energy".

2052 40008

Similarly, ground motion models have been developed using:

- o ground motion data from past earthquakes in the region of interest or throughout the world combined with

- o appropriate earthquake parameter data.

And models describing the effects of local site conditions have been based on

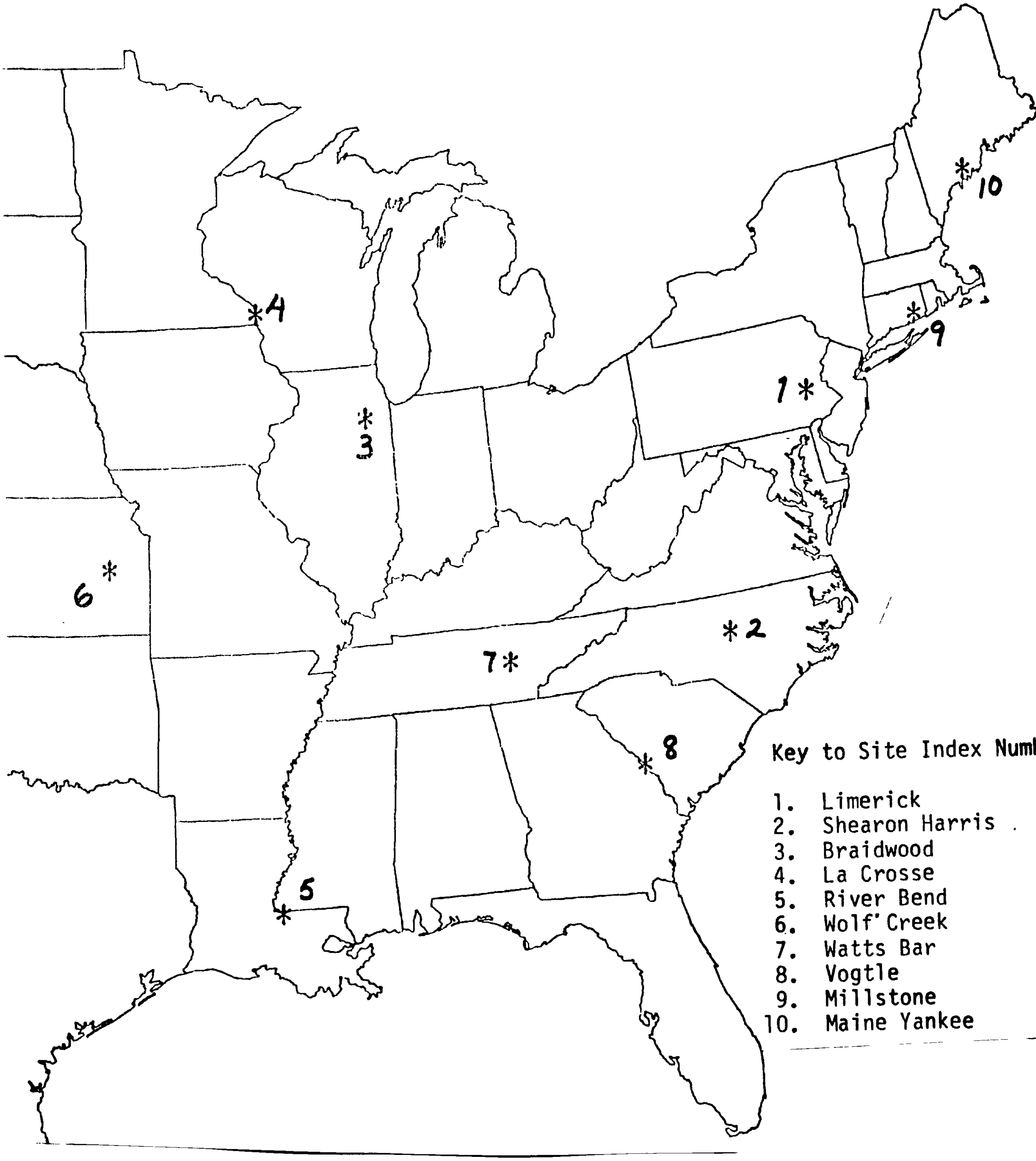
- o local site soil data combined with

- o local ground motion recordings showing local site amplification

Both seismicity descriptions and ground motion modeling rely on having an extensive data base of past events. Because of the short historical record, low rate of earthquake occurrence and a general lack of agreement as to the causes of earthquakes in the eastern United States (EUS) both the physical data alone and/or mathematical models are inadequate for describing the seismic hazard throughout that region. Therefore, it is a common practice to supplement the data with professional judgement and opinions when attempting to estimate the future seismic hazard in the EUS. Because of the limited historical record and the use of subjective judgements it can be expected that diverse opinions and large uncertainties will surround seismicity and ground motion descriptions. Therefore, any estimation of future seismic hazard in the EUS must deal with this uncertainty and diversity of opinions.

Recognizing these facts, the U.S. Nuclear Regulatory Commission (NRC) funded the Lawrence Livermore National Laboratory (LLNL) to develop a seismic hazard assessment methodology which deals with the diverse opinions and uncertainties and to implement the methodology at ten test sites in the EUS, shown in Fig. 1.1.

When using professional opinions as a source of information, e.g. for seismicity descriptions and ground motion models, a number of basic questions must be addressed. For example,



Key to Site Index Numbers

1. Limerick
2. Shearon Harris
3. Braidwood
4. La Crosse
5. River Bend
6. Wolf Creek
7. Watts Bar
8. Vogtle
9. Millstone
10. Maine Yankee

Fig. 1.1 Location of the Sample Sites

2185 40008

- Recognizing that individuals may have a diversity of opinions, should individual opinions, hence diversity, be retained or should a consensus be established?

- How should information be elicited (e.g. using questionnaires, face-to-face meetings, or in group sessions)?

- How much data should the experts be provided? Should they be expected to develop their own data sources?

- How much group interaction and/or feedback should there be?

- To what extent and degree should the quality and consistency of responses be monitored?

- Should the individuals' or groups' opinions be aggregated, and if so, how?

- How should the uncertainties of an individual (group) and diversities between individuals (groups) be assessed and included in the estimation of hazard?

These are some of the fundamental issues considered in developing the LLNL methodology for estimating future seismic hazard in an environment of uncertainty.

The LLNL approach, which is based on:

- the use of individual professional judgements (in conjunction with data) including limited group interaction, feedback and monitoring

- the development of a flexible computational framework that incorporates a Monte Carlo simulation technique to describe the uncertainty in estimating seismic hazard due to the diversity of opinions and uncertainties surrounding the estimation process, is described, briefly, in Section 3. The LLNL method is related to other seismic hazard methodologies in Section 2. The fundamental issues, as they relate to the LLNL approach, are also discussed in that section. Section 4 describes a few of the important results from the analysis at the ten sites (Bernreuter et al. 1985). Finally, a comparison of the LLNL method with other recent studies is discussed in Section 5.

2.0 SEISMIC HAZARD METHODOLOGIES INCLUDING EXPERTS' OPINIONS

2.1 Hazard Model and Calculation

The LLNL methodology is similar in many ways to the well established methods developed by Cornell 1968 and 1971, McGuire 1976, DerKuireghian and Ang 1977, Mortgat and Shah 1979, and Algermissen et al. 1982. These studies all are based on the four elements, described in Fig. 2.1,

- o Identification of source zones affecting the site.
- o Description of the seismicity of a source zone using a recurrence model.
- o Identification of an appropriate ground motion model.
- o Estimation of the hazard by a hazard curve.

It is assumed that the region affecting the ground motion at a site can be divided into discrete areas, referred to as source zones, of uniform seismicity characteristics. The seismicity of each source zone is described by the recurrence model which expresses the expected number of earthquakes, per year, exceeding a magnitude as a function of magnitude. For the LLNL methodology, the zonation and estimation of the parameters (a,b) of the recurrence model are performed by seismicity experts using geophysical data (such as tectonic stresses, plate motions, geology) and observed seismicity (analysis of earthquake catalogs). Numerous ground motion models exist for

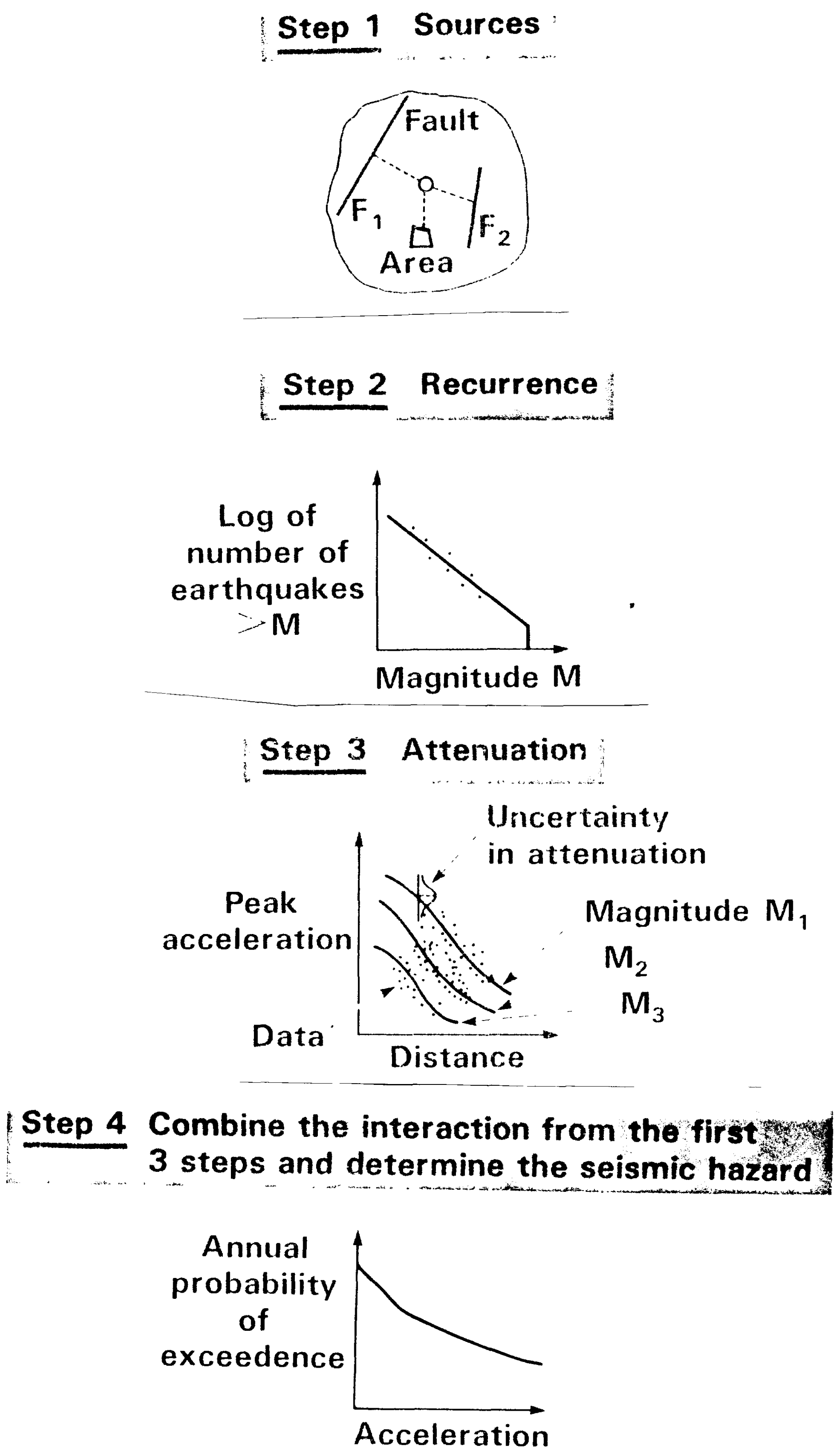


Fig. 2.1 Typical Steps in a Seismic Hazard Analysis

describing the attenuation of ground motion between the source and the site. The models used in the LLNL methodology are based on the assessment of these numerous models by ground motion experts. Finally, these inputs are combined to estimate the seismic hazard at a site in terms of a hazard curve which describes the probability that the maximum value of some ground motion parameter exceeds a certain prescribed level for different values of that level.

The main difference between the LLNL methodology and the other studies is the way in which the uncertainties related to the hazard calculations are handled. The LLNL methodology makes a clear distinction between:

1) the random variation inherent in the occurrence of earthquakes affecting a site and the propagation of the related ground motion at the source to ground motion at the site.

and

2) the uncertainties, due to diverse opinions between experts and the limited data base of past events, associated with the estimation process,

The uncertainties in the estimation process, referred to as modeling uncertainties, form the bases for the experts to describe their state of knowledge and level of confidence in the information used in formulating their opinions. Modeling uncertainties were introduced into the hazard analysis by having the experts provide alternative zonations and/or models as well as ranges of values for the seismicity parameters, e.g., (a,b) in the recurrence model. These uncertainties were themselves modeled in terms of probability distributions which were sampled, using Monte Carlo simulation, to describe the resulting uncertainty in the estimation of the seismic hazard.

Other recent studies, e.g., Yankee Atomic 1981 and EPRI 1985, have also attempted to deal with modeling uncertainties. The approach in these studies, based on logic tree methods, limits the description of uncertainty in the inputs to specifying a small number, usually 2 or 3, of alternative values for the inputs. We believe this approach is too limiting to properly reflect the uncertainty surrounding the description of seismicity and ground motion, particularly in the EUS.

2.2 Role of Experts' Opinions in Developing Hazard Analysis Inputs

The limited historical record of earthquakes and overall understanding of the tectonic processes occurring in the EUS made it necessary to rely on the judgements of knowledgeable individuals to adequately describe the seismicity and ground motion attenuation throughout that region. The LLNL hazard methodology is based on using individual experts' opinions as the sole source of seismicity and ground motion descriptions. However, every effort was made to assure that the experts had the available historical data at their disposal in formulating their opinions.

As part of the LLNL approach two panels of experts were formed. As indicated in Table 2.1, one panel, the S-Panel, included 11 experts knowledgeable about seismicity and zonation. The second panel, the G-Panel, consisted of 5 experts knowledgeable about ground motion prediction. The individuality of the opinions of the experts was emphasized by having them formulate their descriptions of zonation and seismicity without prior interaction and by encouraging them to use their own tectonic and seismicity information and data bases. The intent of this approach was to avoid the screening of non-classical interpretations which would likely occur if consensus descriptions of seismicity were required or even if agreement with historical data or models of tectonic processes was necessary.

Initially, information was elicited from the panel members through a series of written questionnaires. This was followed up by joint feedback meetings in which each expert was provided with calibration hazard results based on their own input and during which the panel members were encouraged to interact and share their views. The experts' final opinions were elicited through feedback written questionnaires.

TABLE 2.1

List of Panels and Panel Members

EUS ZONATION AND SEISMICITY PANEL (S-Panel)

Professor Gilber A. Bollinger

Mr. Richard J. Holt

Professor Arch C. Johnston

Dr. Alan L. Kafka

Professor James E. Lawson

Professor L. Tim Long

Professor Otto W. Nuttli

Dr. Paul W. Pomeroy

Dr. J. Carl Stepp

Professor Ronald L. Street

Professor M. Nafi Toksoz

EUS GROUND MOTION MODEL PANEL (G-Panel)

David M. Boore

Kenneth Campbell

Professor Otto W. Nuttli

Professor Nafi Toksoz

Professor Mihailo Trifunac

PEER REVIEW PANEL

Professor G.B. Baecher

Professor J.E. Ebel

Professor L.T. Long

Professor D. Veneziano

The responses provided by the experts were extensively analyzed to check for consistency and gross errors. Any discrepancies or gross deviations from known data were brought to the attention of the expert who was given every opportunity to change his responses. However, the results, based on any individual's inputs, were not required to match the historical data. This was felt to be appropriate because of the incompleteness of the historical data and because it is not necessary to expect everyone to believe the past is a perfect model of the future. However, the soundness of the overall methodology, as well as the quality of the data, was subject to critique by our Peer Review Panel (see Table 2.1). This review contributed to assuring the quality of the inputs into the hazard analysis.

Retention of the diversity of opinions between experts is an important consideration in the LLNL methodology. A hazard estimate, i.e., calculation of a hazard curve, requires input from a member of each of the two panels. Thus, the hazard is estimated based on the inputs for every pair, i.e., an S-expert and G-expert, of experts. The variation in the hazard estimates between the 55 (11 x 5) pairs is representative of the diversity of opinions between experts.

2.3 Aggregation of Expert's Opinions

Frequently it is appropriate to have a single estimate of hazard, thus it is necessary to aggregate the results over the experts. Since the LLNL method is based on aggregating only the hazard estimates, a pooled estimated hazard curve is based on a weighted average of the 55 individual hazard curves. The weights for the G-experts are normalized values of self weights provided by the experts. The weights for the S-experts are themselves a weighted average of four regional self weights provided by the experts.

The uncertainties each expert associates with their inputs are modeled in the LLNL methodology by associating a probability distribution, either discrete or continuous, with each input, e.g., zonation, seismicity parameters, ground motion models. Monte Carlo simulation is used to translate the uncertainties in the inputs to uncertainty in the hazard estimate. The uncertainty in the hazard for each pair of experts is pooled in an analogous way to pooling the individual hazard curves to develop the overall uncertainty in the hazard estimate.

3.0 LLNL HAZARD ANALYSIS METHODOLOGY

3.1 Introduction

As discussed earlier, the LLNL methodology is based on using the professional judgements of experts in seismology and ground motion predictions as the source of inputs for a hazard analysis. Elicitation of experts' opinions was primarily through written questionnaires. The questionnaires were designed to obtain the complete opinion of the experts about any given parameter, including their uncertainty.

The questionnaires elicited information about a model or parameter in two forms:

- a most likely value or model (referred to as the 'best estimate' (BE)); this value or model represented, in the expert's opinion, the most realistic estimate, based on all the information available to the expert, of the state of nature.

- a set of alternative models (or range of values) which span, with 'high' confidence in the expert's opinion, the potential set of models (or values) describing the state of nature.

The above inputs were used to develop probability distributions (either discrete or continuous) for each of the inputs. These distributions provide the inputs for the Monte Carlo simulation technique used in the uncertainty analysis portion of the LLNL method.

3.2 Hazard Analysis Inputs

The inputs for a seismic hazard analysis, as designed in the LLNL methodology, can be divided into three groups, a description of the seismic source zones, specification of the seismicity of each source zone and identification of appropriate ground motion models.

3.2.1 Seismic Zonation Maps

It is difficult, in general, to associate historic events to specific geologic or tectonic features in the EUS. Thus, in the LLNL approach, the seismicity of the EUS is modeled by partitioning the region into source zones, assumed to be areas of diffuse seismicity in which all potential earthquakes occurring within a zone have the same expected characteristics, such as spacial and temporal occurrence and magnitude distribution including potential maximum magnitude. In this approach, the historical seismicity is not necessarily associated with specific known features. This approach, we believe, introduces a degree of flexibility into the elicitation process in the sense that it does not require the experts to identify a unique tectonic feature and/or process to each of the historic events.

Elicitation of the seismic zonation maps for the EUS was based on developing two forms of zonation information. First the S-experts were asked to provide a map (called their BE map) of the source zones which they believed, based on their knowledge and all available information, most realistically modeled the zonation of the EUS. Secondly, the S-experts were asked to suggest alternative zonations. These alternatives were intended to reflect the individuals uncertainty in how well the source zones in the EUS could be identified. Alternative zonations were generated by having the experts

- assign a degree of belief (confidence) that a specific zone should be identified versus being a part of the surrounding zone(s)

- suggest alternative boundaries (shapes) for a zone or cluster of zones with an associated degree of belief.

This information was used to generate a collection of maps for each S-expert. The degrees of belief were used as probabilities to associate a probability with each map. This probability represented, for each expert, his degree of belief that the map represented the zonation of the EUS. The collection of maps along with their associated probabilities, are inputs for the hazard estimation and uncertainty analysis. To assure the tractability of the procedure the collection was limited to the 30 maps with the highest degrees of belief.

3.2.2 Seismicity

The seismicity of a source zone is determined by the frequency of earthquake occurrences and the distribution of earthquake magnitudes. This distribution is bounded above by the largest possible magnitude (referred to as the upper magnitude cutoff, M_u) which the source zone tectonic conditions are capable of producing. In most hazard analyses, the seismicity is described in terms of a recurrence model expressing the expected number of earthquakes greater than a given magnitude (or intensity) as a function of magnitude. Usually the model is described by the Gutenberg-Richter equation

$$\text{Log}_{10} N = a - bm \text{ (or I)} \qquad (3.1)$$

where (a,b) are constants which vary between zones and m denotes magnitude (I is epicentral intensity). Equation 3.1 must be modified to ensure that N is zero at $m = M_u$.

In the LLNL methodology, the S-experts were asked to describe the seismicity of the EUS by estimating the values of a,b and M_u for each zone identified in their zonation maps. Each expert provided a BE value, i.e., the value which they believe best represents the true state of nature, and a range of values which they believe, based on their state of knowledge, with high confidence represents the potential seismicity within each zone. In addition, the experts were given the opportunity to state whether the uncertainties they have regarding the appropriate values of (a,b) should be considered as independent or correlated.

Estimation of the seismicity by each of the experts was based on their own sources of information and review and analysis of a catalog of earthquakes of their choosing. Analysis of past events depends on making corrections for catalog incompleteness as well as adjusting, if appropriate, for aftershocks. The experts were asked to consider these latter problems as they saw fit.

Similar to the treatment of the alternative zonations, the BE values and ranges for the seismicity parameters were used to develop probability distributions for each of the uncertain parameters. These distributions provided the inputs to the uncertainty analysis portion of the LLNL methodology.

3.2.3 Ground Motion Models

The purpose of a ground motion model is to predict the ground motion at a site caused by an earthquake of known location and magnitude. The approach used in the LLNL methodology to obtain ground motion models for use in the hazard analysis was to present the G-panel members with descriptions of the available models and to ask them to select a set of ground motion models applicable in each of four regions (northeast, southeast, northcentral and southcentral) of the EUS. For each region, the experts were asked to select the model which they believed provided the best predictions of ground motion at sites within that region. They were also asked to select as many as six alternative models and to associate degrees of belief to each of the models selected. The selection of several models was intended to reflect the experts' uncertainty in the models predicting the true ground motion within a region. Again, the models and the associated normalized degrees of belief were used to develop discrete probability distributions for use in the uncertainty analysis. The experts were also asked to estimate the random variation in ground motions caused by earthquakes of known magnitude and distance from the site. Point estimates (BE) and a range of values were elicited from the experts.

A second model used in the LLNL methodology was a model for considering the effects of local site soil conditions on the ground motion at the site. In

2153 40008

most hazard analyses this effect is modeled by a simple multiplicative factor depending on whether the site is considered to be either a soil or a rock site. This approach was generalized in the LLNL methodology (Savy 1986) to allow for other models. Three alternatives were considered:

(1) make no change in the predictions from the ground motion model

(2) use the simple multiplicative factor to adjust for either rock or soil conditions

(3) use a set of adjustment factors applicable to eight generic classes of sites including thick soil sites (base case), hard rock, and three depths of soil deposits, either sand-like or till-like.

The experts were asked to assess the potential of each of these three methods to properly predict, in conjunction with the ground motion model, the ground motion at a site. The weights assigned by the experts for each of these alternatives were used to develop the appropriate probability distribution for the uncertainty analysis.

3.3 Hazard Analysis Calculations

The hazard analysis portion of the LLNL methodology is generally consistent with other studies and is based on assuming that

- earthquakes occur within a source zone 'at random' over time and uniformly in space throughout the zone

- the distribution of magnitudes, given an earthquake, is approximately a truncated (at M_u) exponential distribution

- for all but one acceleration ground motion model (there were a total of 33 models of acceleration) the variation in the ground motion parameter at the site, caused by an earthquake of known magnitude and distance from the site, can be approximated by a lognormal distribution.

Following standard analyses, the seismic hazard at a site is described by the hazard curve defined as the probability the maximum value of the ground motion parameter (usually peak acceleration or velocity) per unit time (e.g., per year) is greater than the value a, written as a function of a. Although it is not often clearly stated but it is important to recognize, particularly when comparing results from different studies, that the hazard is only based on earthquakes of magnitudes above a minimum level. In the LLNL study this minimum magnitude was $m_0 = 3.75$ on the M_b scale.

Using the assumptions stated above, it can be shown (Bernreuter et al. 1985) that the hazard $P(A \geq a | m_0 = 3.75)$ can be expressed as

$$P(A \geq a | m_0 = 3.75) = 1 - \prod_{zone,z} e^{-\lambda_0 \int_{m_b} \int_{d} f_M(m_b|z)\, f_0(d|z)\, P(A \geq a|m_b,d)\, dm\, dd} \tag{3.2}$$

where $f_M(\cdot|z)$ and $f_d(\cdot|z)$ represent the truncated exponential distribution of magnitudes and geometric distribution of distances (from the site) within a zone respectively. $P(A \geq a|m_b,d)$ represents the lognormal distribution of the ground motion parameter. The actual hazard calculation, in the LLNL methodology, is based on approximating the probability integral by summations. The hazard, i.e., evaluation of Eq. 3.2, can be estimated for each set of inputs, source zonation, seismicity parameters (a,b,M_u) values, choice of ground motion model and value of random variation, and choice of local site correction.

3.4 Uncertainty Analysis

An important part of the LLNL methodology is the treatment of the uncertainties associated with describing seismicity and ground motion predictions. As described earlier these modeling uncertainties were handled by associating a probability (uncertainty) distribution to each of the inputs (models and parameters). To reflect these uncertainties in the hazard estimates a Monte Carlo simulation was developed. For each pair of S-G experts, a typical simulation is as follows:

- Draw a map from the distribution of maps for this S-expert.

- For each one of the seismic sources in a sample map, draw a set of seismicity parameters from their respective distributions, i.e., draw

  - a value for the a parameter of the recurrence law
  - a value for the b parameter of the recurrence law. b is allowed to have three levels of correlation with a. This correlation level is chosen by the individual experts.
  - the value of the upper magnitude (or intensity) cutoff.

- Draw a ground motion model from the distribution of models, for the appropriate region (NE, SE, NC or SC).

- Draw a value for the random variation parameter which is associated with the selected ground motion model, for the appropriate region (NE, SE, NC or SC).

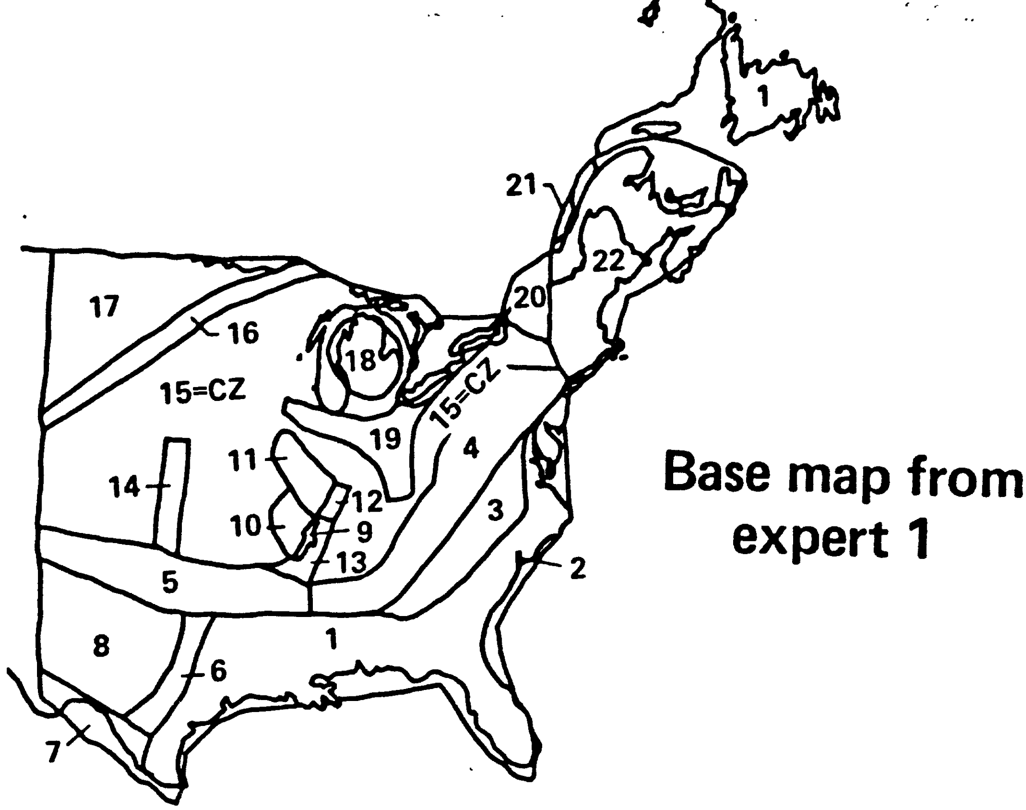
  - Draw a site correction method.

The hazard is calculated for each of the seismic sources and combined over sources, as given in Eq. 3.2.

Each simulation results in a possible hazard curve. For the NRC study applying this methodology to ten sites, 2750 such curves (50 simulations per G-expert x 5 G-experts x 11 S-experts) were generated for each site.

3.5 Hazard Analysis Outputs

Generally, the hazard at a site has been described in terms of a hazard curve, i.e., a graph of the probability that within a period of one year the maximum value of a ground motion parameter, e.g., peak ground acceleration or velocity, will exceed a given level, say A, as a function of a. A set of such hazard curves, using peak ground acceleration as the parameter, is shown in Fig. 4.2.



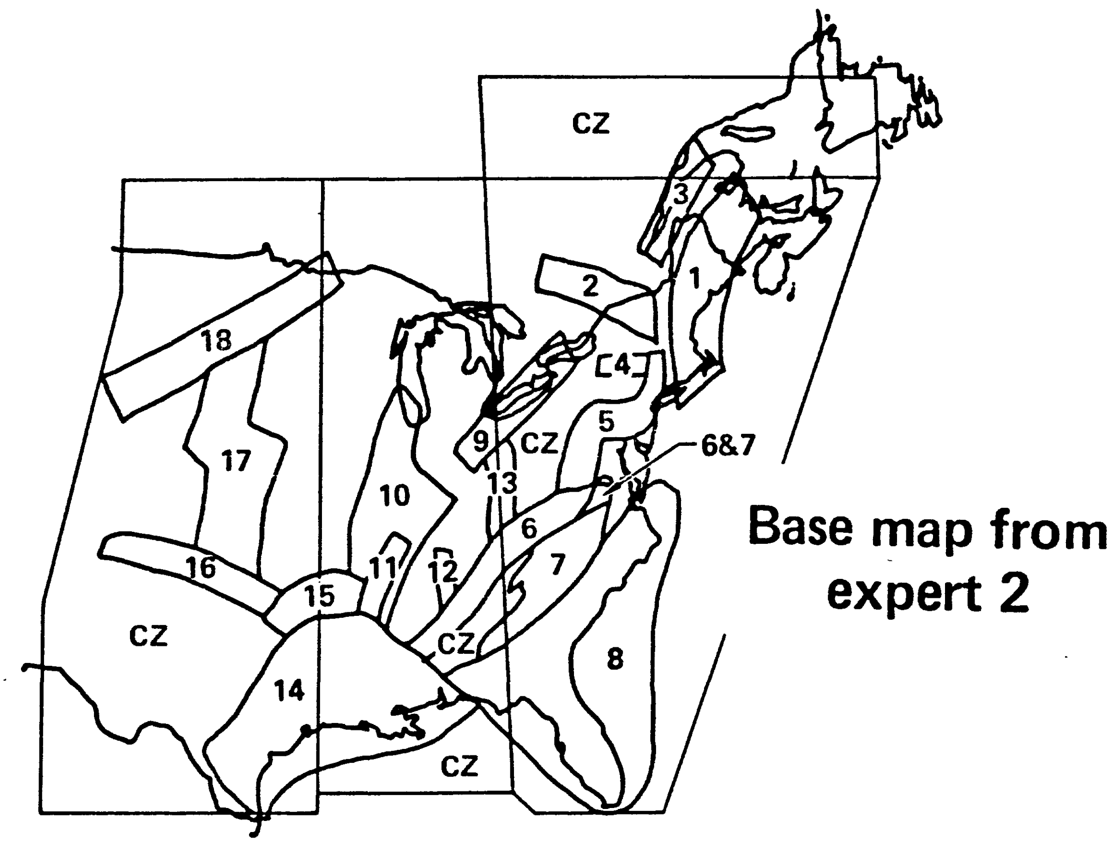


Fig. 4.1 Typical Variation Between LLNL's S-Experts' Best Estimate Seismic Zonation Maps

2251 40008

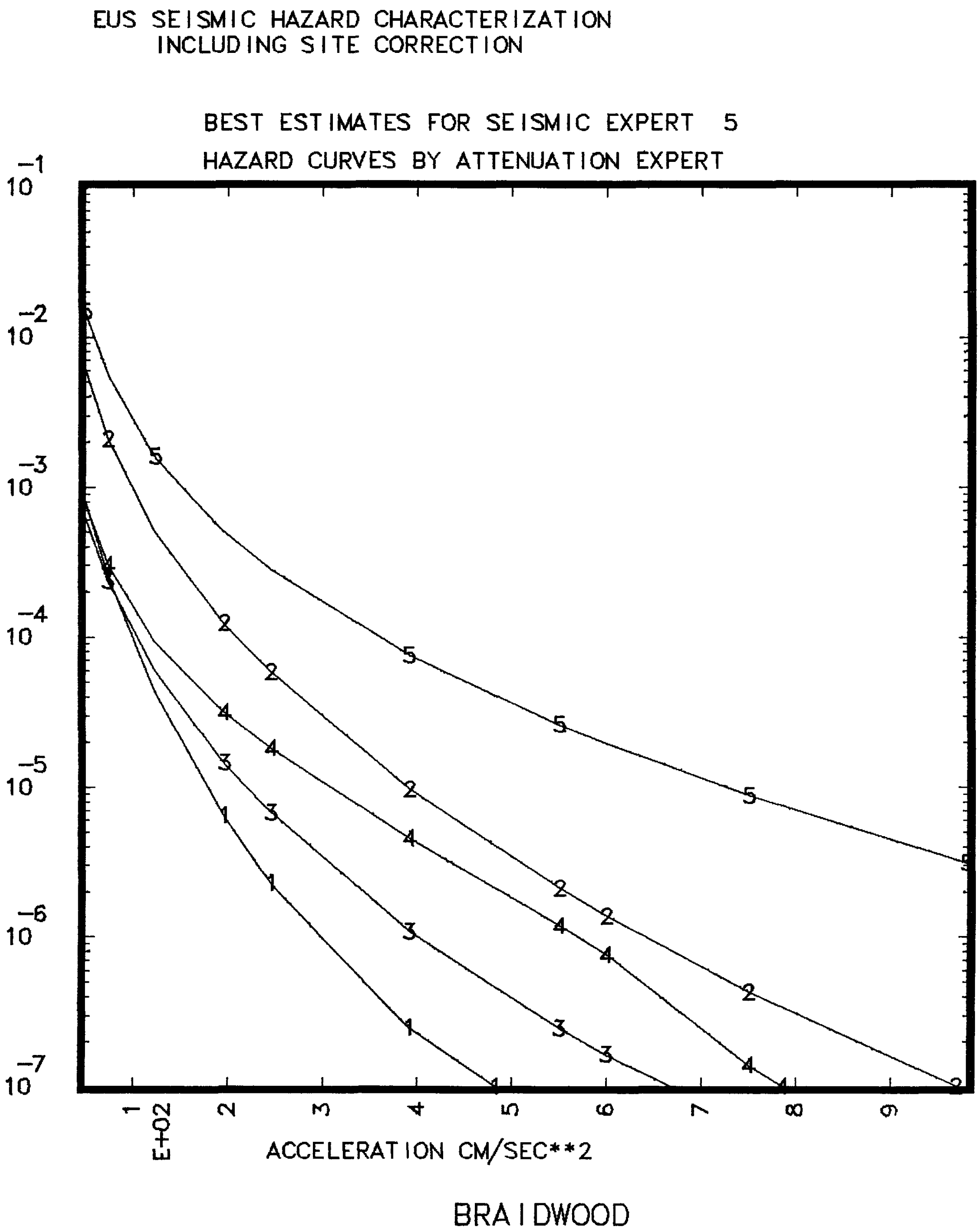


Fig. 4.2 Best Estimate Hazard Curves For S-Expert 5 for Each of the G-Experts For the Braidwood Site

22B5 40008

Introducing modeling uncertainty into the LLNL hazard analyses expands considerably the number of estimators of hazard that can be used. One estimator produced by the LLNL methodology is referred to as the best estimate hazard curve (BEHC). This is the hazard curve, for a particular pair of seismicity and ground motion experts, based on using the BE models and parameter values given by the experts. This corresponds to the hazard curve that would be produced if only a single source of seismicity and ground motion information is available and no uncertainty information is elicited. The BEHC is not necessarily the 'best estimator', but is simply one possible estimator of the seismic hazard at a site.

A second type of estimator produced by the LLNL methodology, referred to as constant percentile hazard curve (CPHC), is based on using the uncertainty information provided by the experts. By treating all the input models and parameters as uncertain variables and using simulation (see Section 3.4) a probability (uncertainty) distribution for the hazard at each value, a, of the level, is developed. Combining the percentiles of the hazard over all levels (over the range of a) gives a CPHC. The 15th, 50th and 85th CPHC's were most often used in the LLNL studies. Just as the BEHC, the CPHC's can be produced for each S-expert and G-expert pair. Such curves describe the uncertainty expressed by a particular pair of experts. However, CPHC's were most often produced when all of the experts were considered, hence producing an uncertainty distribution for the hazard which describes both experts' uncertainties as well as diversity of opinions between experts. Such CPHC's are shown in Fig. 4.5.

In addition to generating BEHC for each pair of S- and G-experts, the methodology includes aggregations of curves. Such combinations of hazard curves were based on using the self-weights provided by the experts. One level of aggregation consists in combining the BEHC over ground motion experts for a given seismicity expert. These aggregated curves can also be combined over seismicity experts to form a second level of aggregation. Two types of aggregated estimators are considered in the LLNL methodology. One is the arithmetic weighted average of the hazards (AMHC) and the second is the geometric weighted average of the hazards (GMHC).

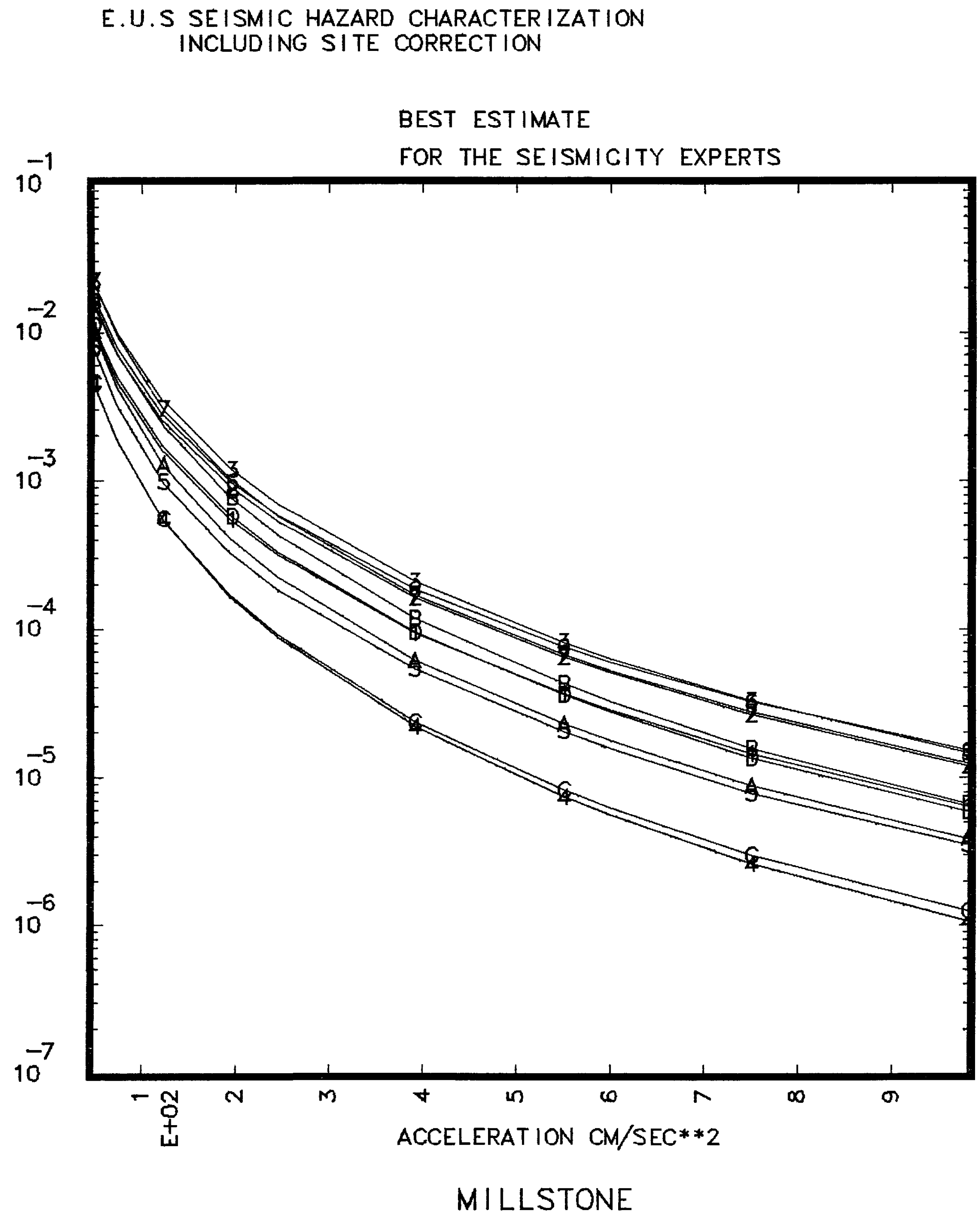


Fig. 4.4 Best Estimate Hazard Curves For Each of the S-Experts Aggregated Over All G-Experts For the Millstone Site

2258 40008

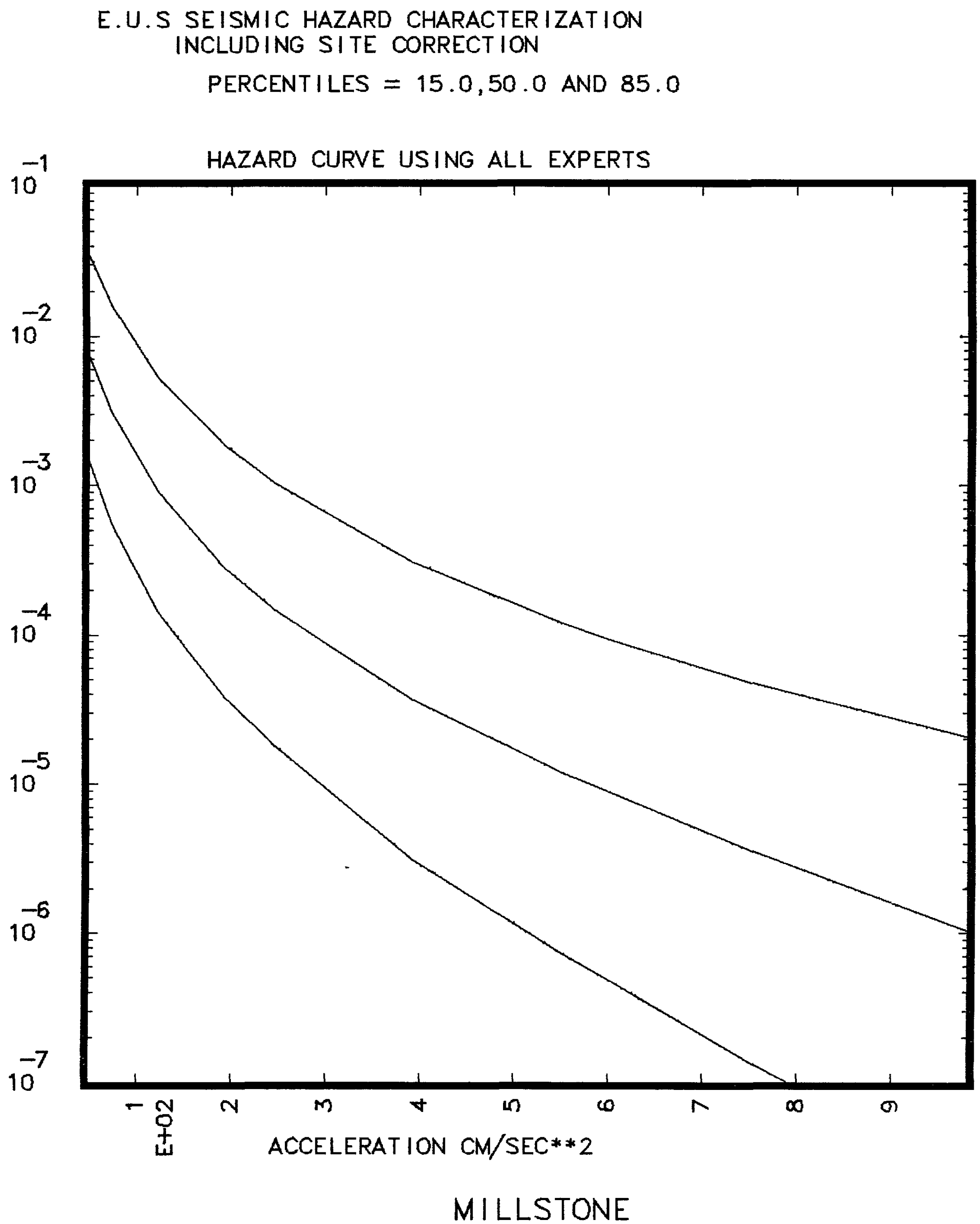


Fig. 4.5 15th, 50th and 85th Constant Percentile Hazard Curves Obtained Using All S- and G-Experts for the Millstone Site.

2284 40008

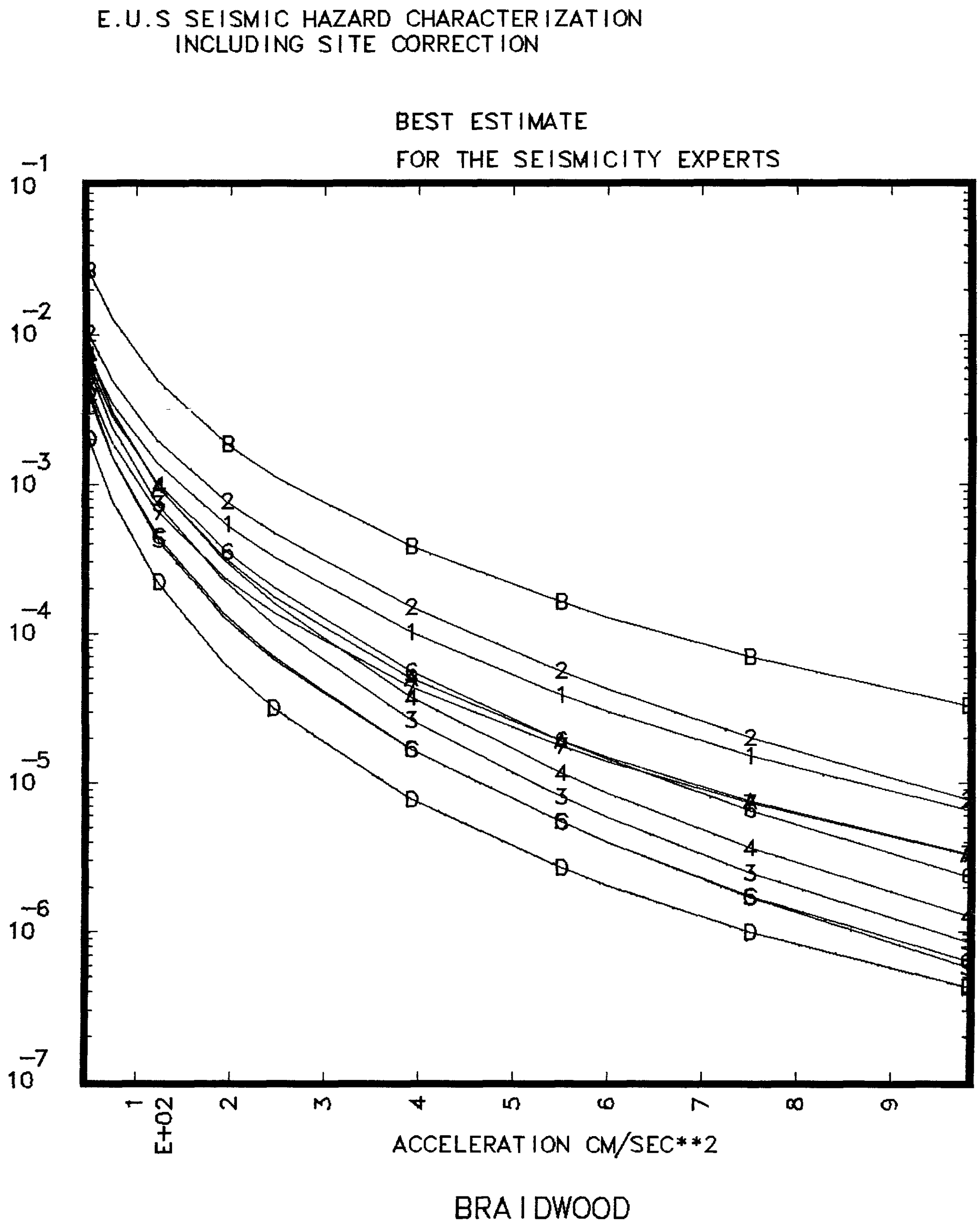


Fig. 4.3 Best Estimate Hazard Curves For Each of the S-Experts Aggregated Over All of the G-Experts for the Braidwood site

A plot of the different estimators is shown in Fig. 4.6 for the Braidwood site. The AMHC is most significantly affected by extreme inputs. On the other hand, the GMHC and the 50th (median) CPHC are less influenced by extreme values, and are in more general agreement.

The LLNL methodology also produces 'best estimate' uniform hazard spectra (BEUHS) and constant percentile uniform hazard spectra (CPUHS). By definition, the uniform hazard spectrum is a spectrum in which each spectral amplitude has the same probability of being exceeded. In the development of the spectrum each frequency is considered independently thus the correlation between the spectral amplitudes is not taken into account.

4.0 TYPICAL RESULTS

There are significant differences between the experts at every step of the analysis process. For example, Fig. 4.1 shows the BE maps of two experts and illustrates some of the diversity in zonation as seen by two different experts. The same variation carried over to other parameters, as well as to the choice of ground motion models. All of these judgemental differences lead to significant differences in the BEHC for each S-G expert pair. For example, Fig. 4.2 shows the BEHC for S-expert 5 for each of the five G-experts for the Braidwood site and Fig 4.3 shows the BEHC for each S-expert aggregated over the G-experts for the Braidwood site. These figures show that there is a significant difference between the BEHC of various experts. The spread between experts is also site dependent as can be seen by comparing Fig 4.3 with Fig. 4.4.

The inter-expert differences in opinion are illustrated by the spread between the 15th and 85th CPHC shown in Fig. 4.5. These large inter-expert differences lead to significant differences between estimators. For example Fig. 4.6 compares the median CPHC to the GMHC and the AMHC for the Braidwood site. Also shown on Fig. 4.6, for reference, are the 15th and 85th percentile CPHC.

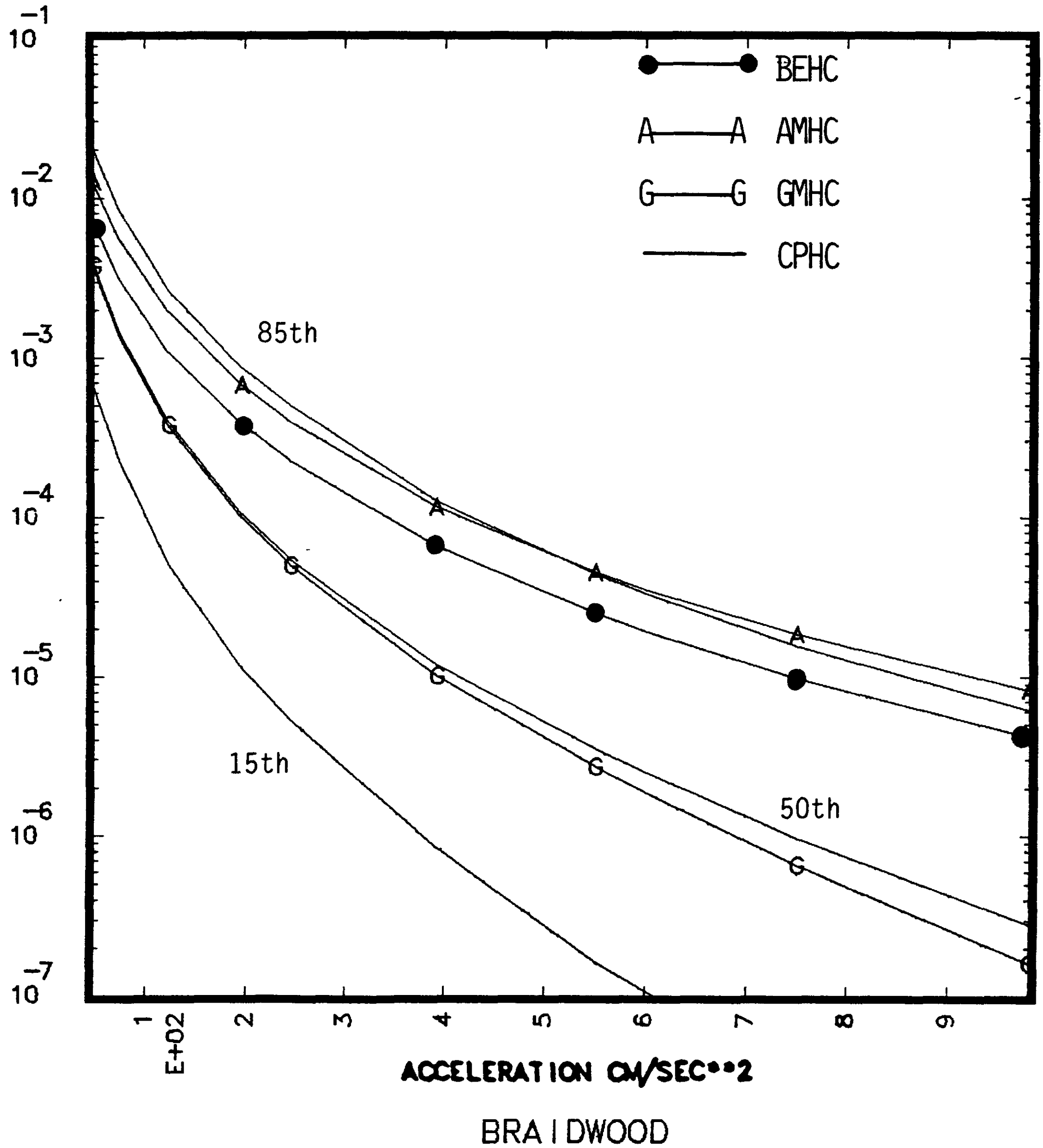


Fig. 4.6 Comparison of Various Extimators of the Seismic Hazard for the Braidwood Site.

2255 40008

Figure 4.7 shows the 1000 year return period CPUHS for the Braidwood site. Complete results for each site shown on Fig. 1.1 are given in Bernreuter et al. (1985).

5.0 COMPARISON WITH OTHER TECHNIQUES

5.1 Historical Analyses

The characteristic feature of historical analyses is that they do not require the identification of sources and their seismicity. In place of this information, they use catalogs of past earthquakes (historic catalogs). The concept of a historical analysis is based on assessing the strong motion at the site, due to past earthquakes, by using ground motion models to "attenuate" each event of the catalog to the site, from which a prediction is made. In the "non-parametric" method (Bernreuter 1981; Veneziano et al. 1984), no a priori assumption is made on the form of the expected number of events (λ) that produce site accelerations greater than a certain value. In the "parametric" method (Veneziano et al. 1984), a functional form is assumed for λ. Figures 5.1 and 5.2 show typical comparisons of hazard calculations using the LLNL methodology and the non-parametric historical method. The curves without labels refer to the 11 seismicity experts of the LLNL study and the curve labeled H refers to the historical analysis. Although the non-parametric method is known to underestimate the hazard by comparison with parametric studies the agreement between the historical and LLNL results is generally reasonable.

5.2 Comparison with the Algermissen et al. Zonation

In the study made by Algermissen et al. 1982 the data for zonation were developed in regional workshops. The aggregation of the experts opinions was made by the analyst at the input level, thus only a single BE zonation map was developed. The final BE seismicity parameters were developed by the analyst. The hazard calculated with the zonation and seismicity data using the G-experts set of ground motion models is shown by the curves labeled (X) on Figs. 5.1 and 5.2. In both cases the hazard is within the range of values

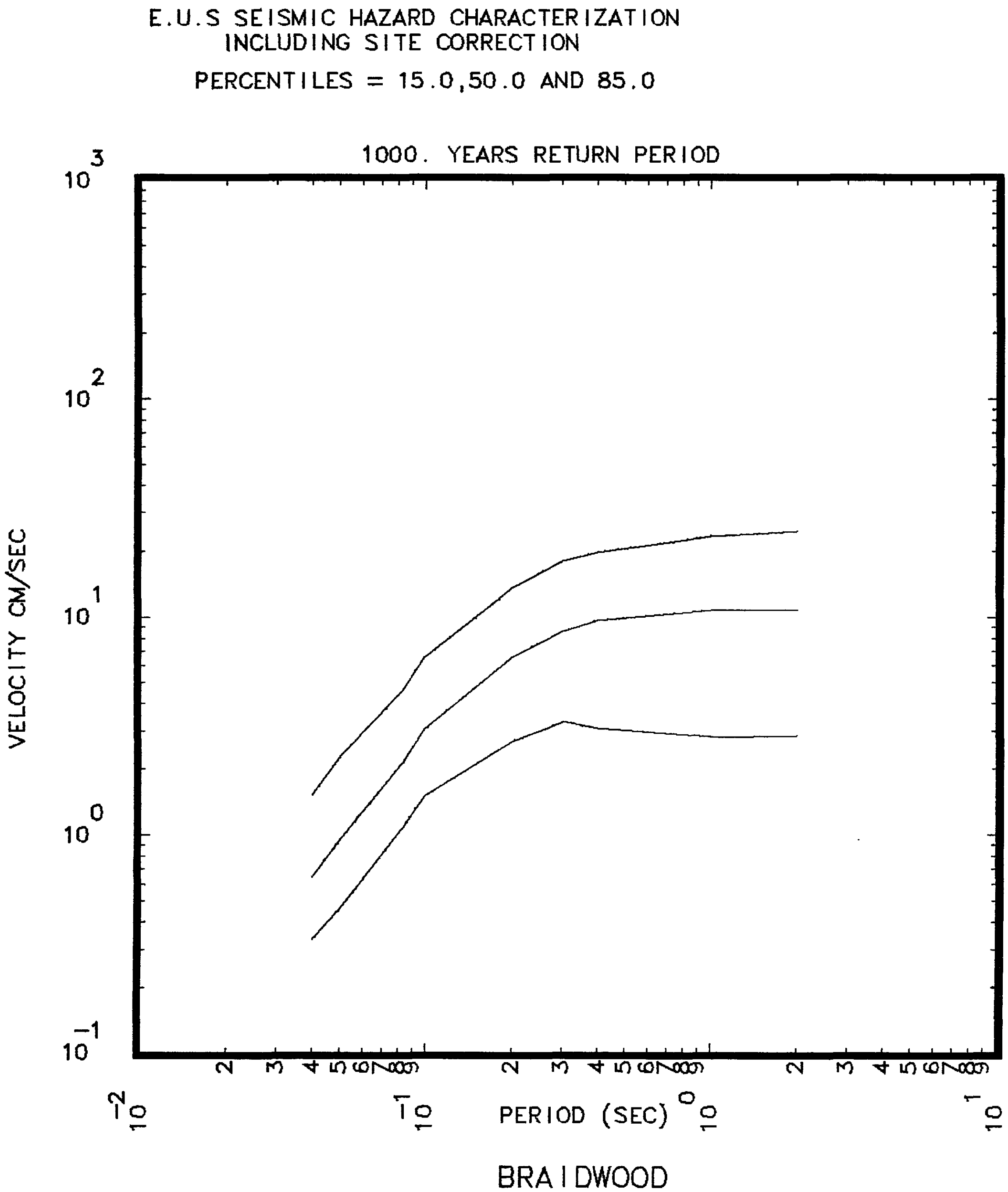


Fig. 4.7 1000 Year Return Period 15th, 50th and 85th Constant Percentile Uniform Hazard Spectra for the Braidwood Site Based on all S- and G- Experts' Inputs.

2282 40008

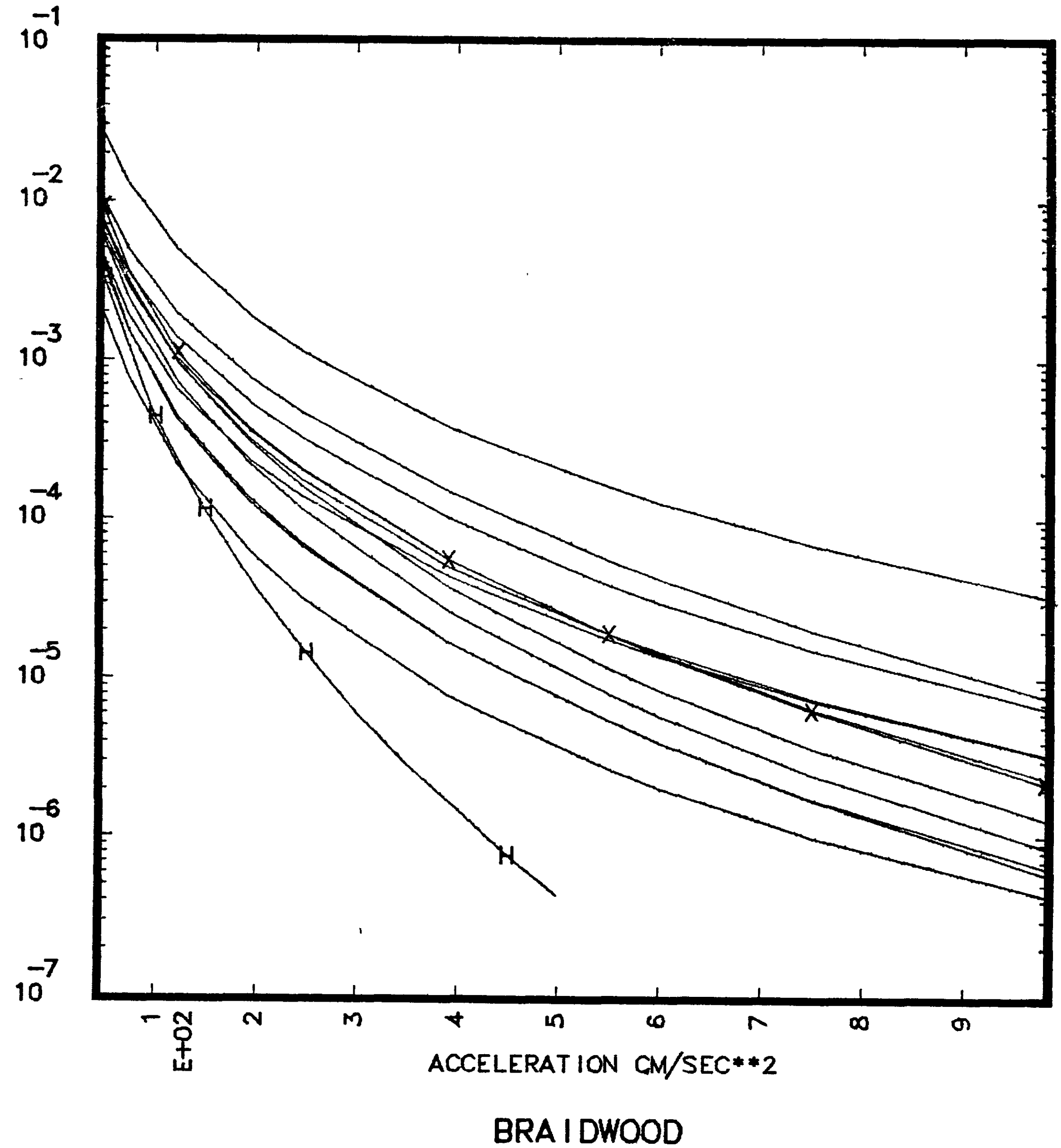


Fig. 5.1 Comparison Between the LLNL S-Experts' BEHC (unlabeled curves) the Historical Hazard Curve (Curves Labeled H) and the BEHC (curves labeled X) Obtained Using the Algermissen et al. (1982) Seismicity Model.

2283 40008

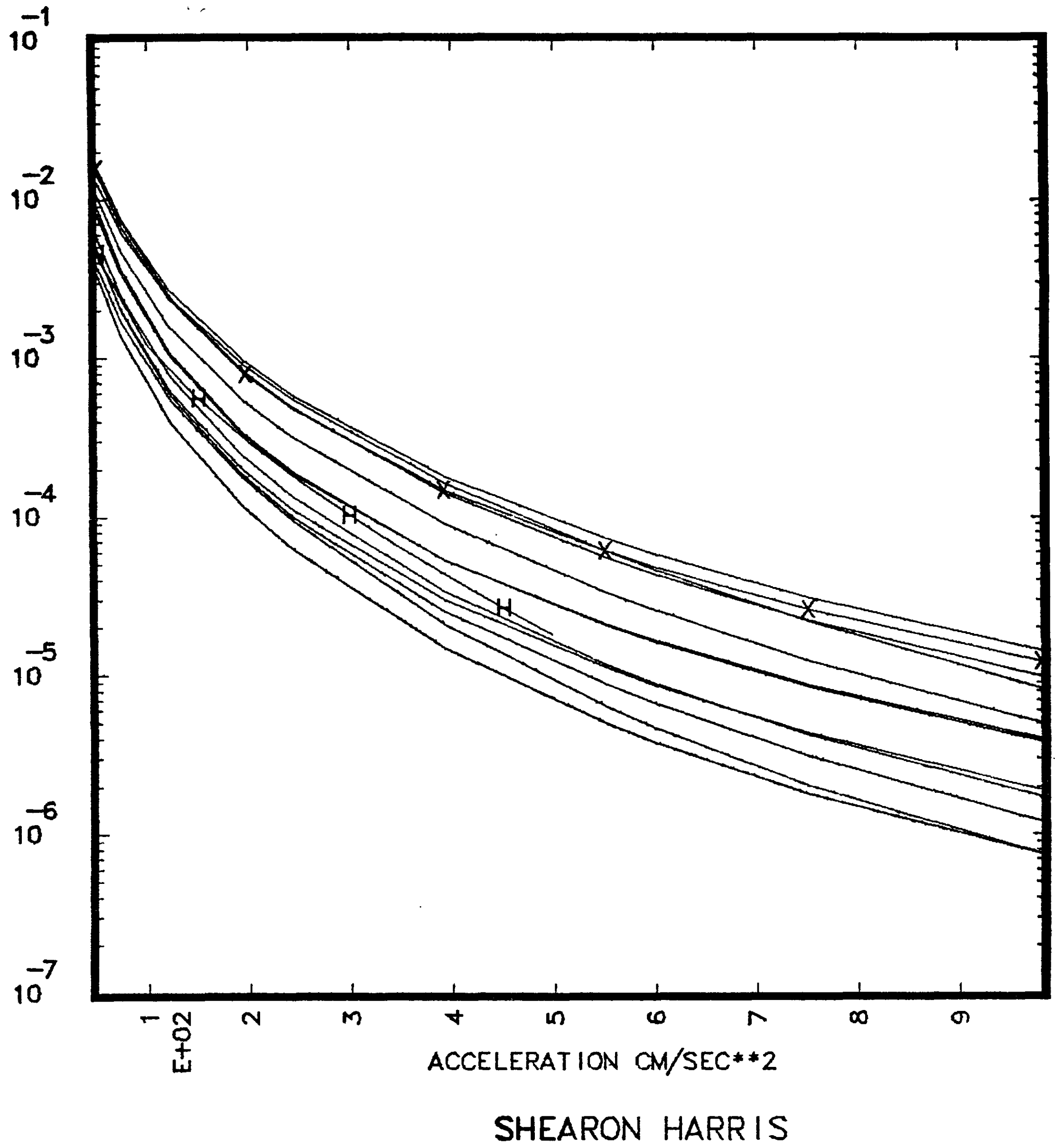


Fig. 5.2 Comparison Between the LLNL S-Experts' BEHC (unlabeled curves) the Historical Hazard Curve (curves labeled H) and the BEHC (curves labeled X) obtained Using the Algermissen et al. (1982) Seismicity Model.

2257 40008

obtained by using the models developed by the LLNL experts. Other comparisons are given in Section 6 of Bernreuter et al. 1985.

5.3 Comparisons To The EPRI Study

There are significant differences between the results of the LLNL study and the results presented in EPRI 1985b. However, much of the difference between the two studies are because EPRI used m_b = 5.0 as the lower bound of integration whereas in the LLNL study m_b = 3.75 was the lower bound of integration. In addition, significantly different ground motion models were used between the two studies. To simplify the comparisons between the two studies the hazard at the ten test sites using the LLNL method was recomputed using m_b = 5.0 as the lower bound of integration and the EPRI 1985b version of Nuttli's ground motion model.

When the same lower bound of integration and same ground motion model is used there is:

1) Excellent agreement of the median CPHC at four sites as is illustrated by Figs. 5.3 and 5.4.

2) Poorer agreement of the medians at five sites; e.g. Figs 5.5 and 5.6 (worst case).

3) EPRI's uncertainty bounds are smaller than LLNL's at five sites, e.g. Fig. 5.3, and larger at four sites, e.g. Fig. 5.6.

4) Generally, LLNL's uncertainty bounds are more symmetric with respect to the median than EPRI's, e.g. Fig. 5.4 and 5.6.

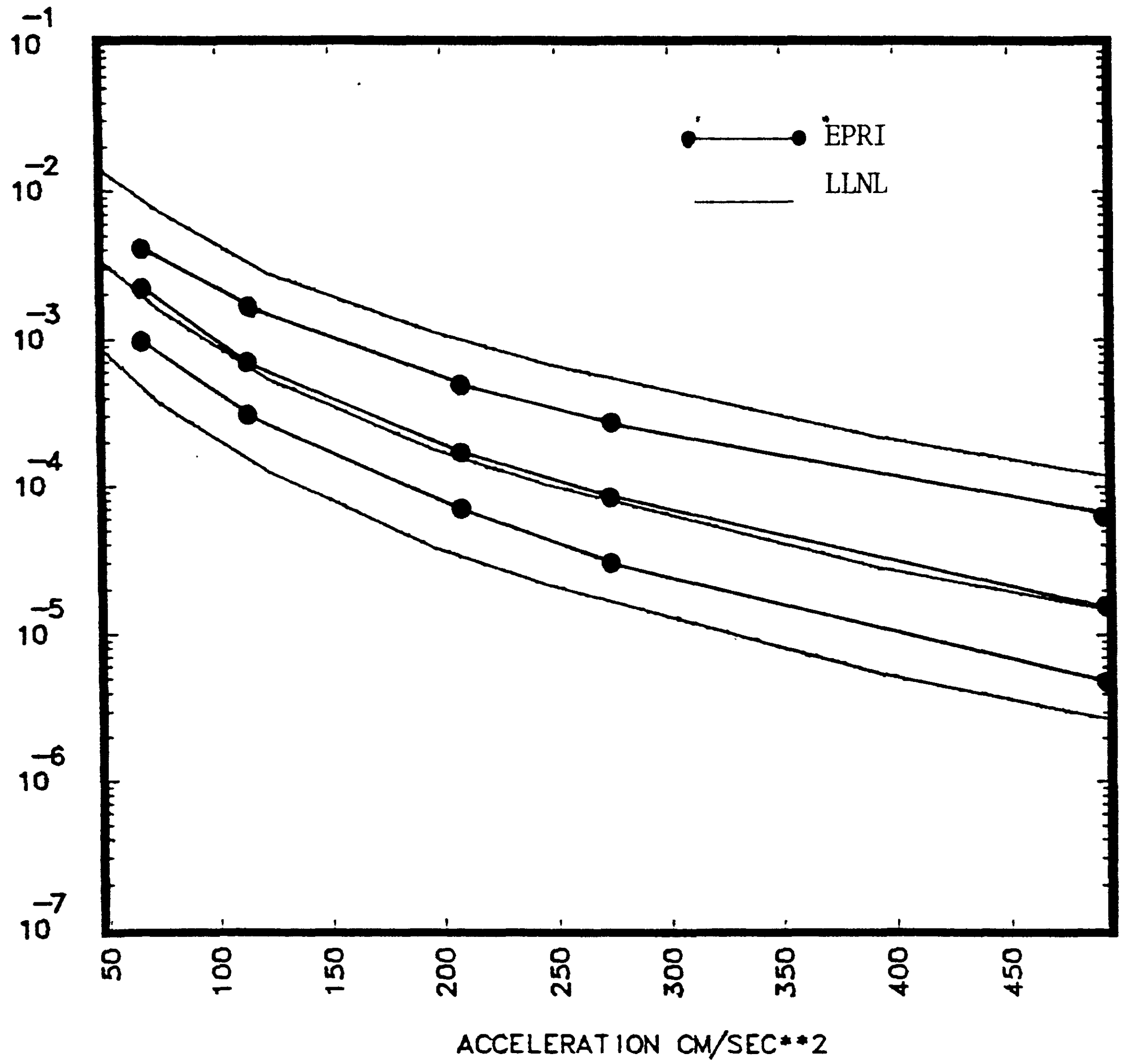
5.4 Comparisons to Other Studies

Seismic hazard analyses have been developed for the Maine Yankee site by Yankee Atomic Electric Company (1983 and 1984), the Limerick site by ERTEC 1982, and the Millstone site by Dames and Moore 1983. The ERTEC and the Dames

EUS SHC. SENSITIVITY TASK LLNL-EPRI COMPARISON
GMM=NUTTLI 1984, MO=5.0, NO SITE CORRECTION

PERCENTILES = 15.0,50.0 AND 85.0

HAZARD CURVE USING ALL EXPERTS



LIMERICK

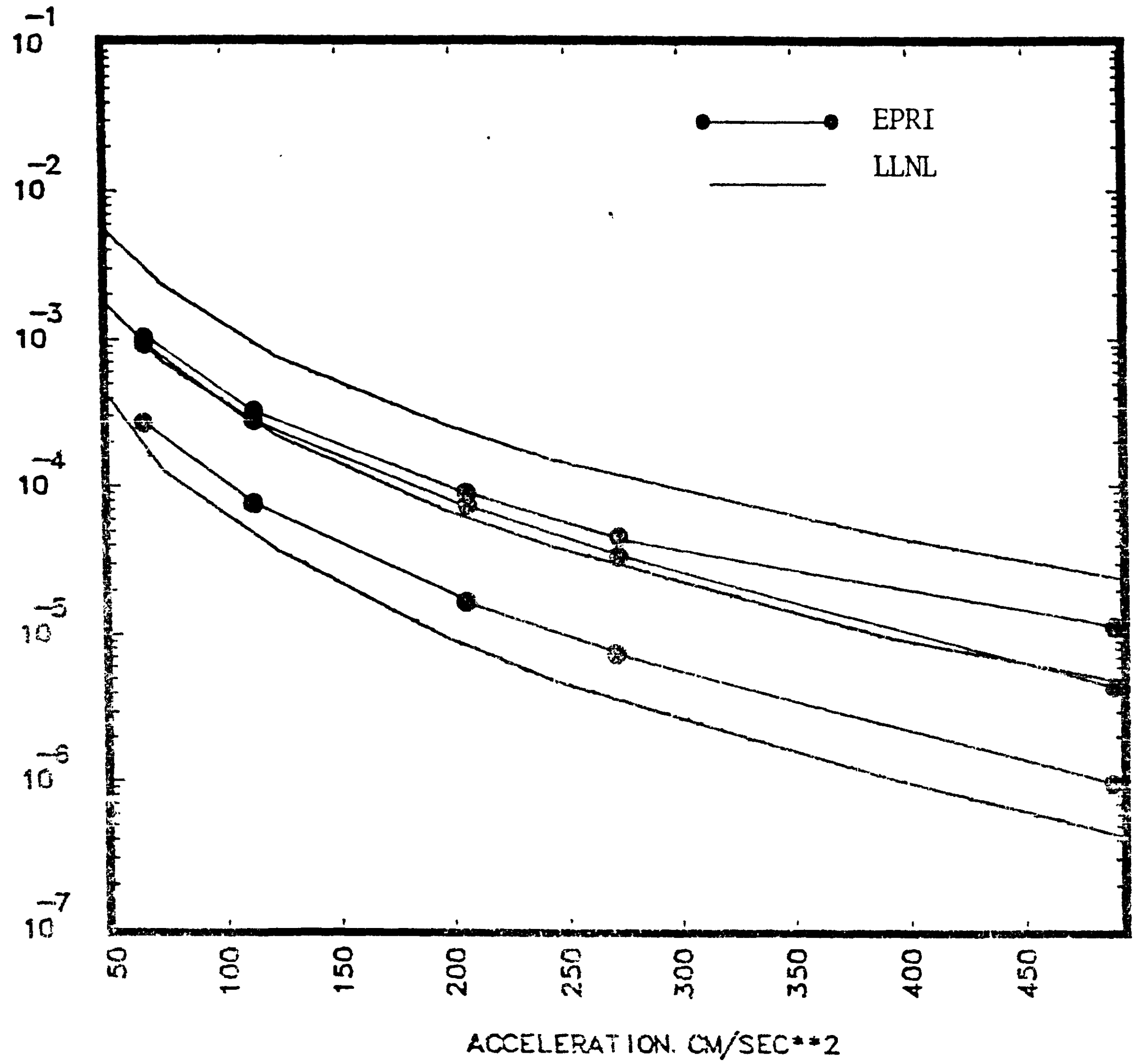
Fig. 5.3 Comparison Between EPRI's CPHC and LLNL's CPHC. Same Ground Motion Model and Lower Bound of Integration Used to Generate the Hazard Curves.

2287 40008

EUS SHC. SENSITIVITY TASK LLNL-EPRI COMPARISON
GMM=NUTTLI 1984, MO=5.0, NO SITE CORRECTION

PERCENTILES = 15.0,50.0 AND 85.0

HAZARD CURVE USING ALL EXPERTS



BRAIDWOOD

Fig. 5.4 Comparison Between EPRI's CPHC and LLNL's CPHC. Same Ground Motion Model and Lower Bound of Interation Used to Generate the Hazard Curves.

2286 40008

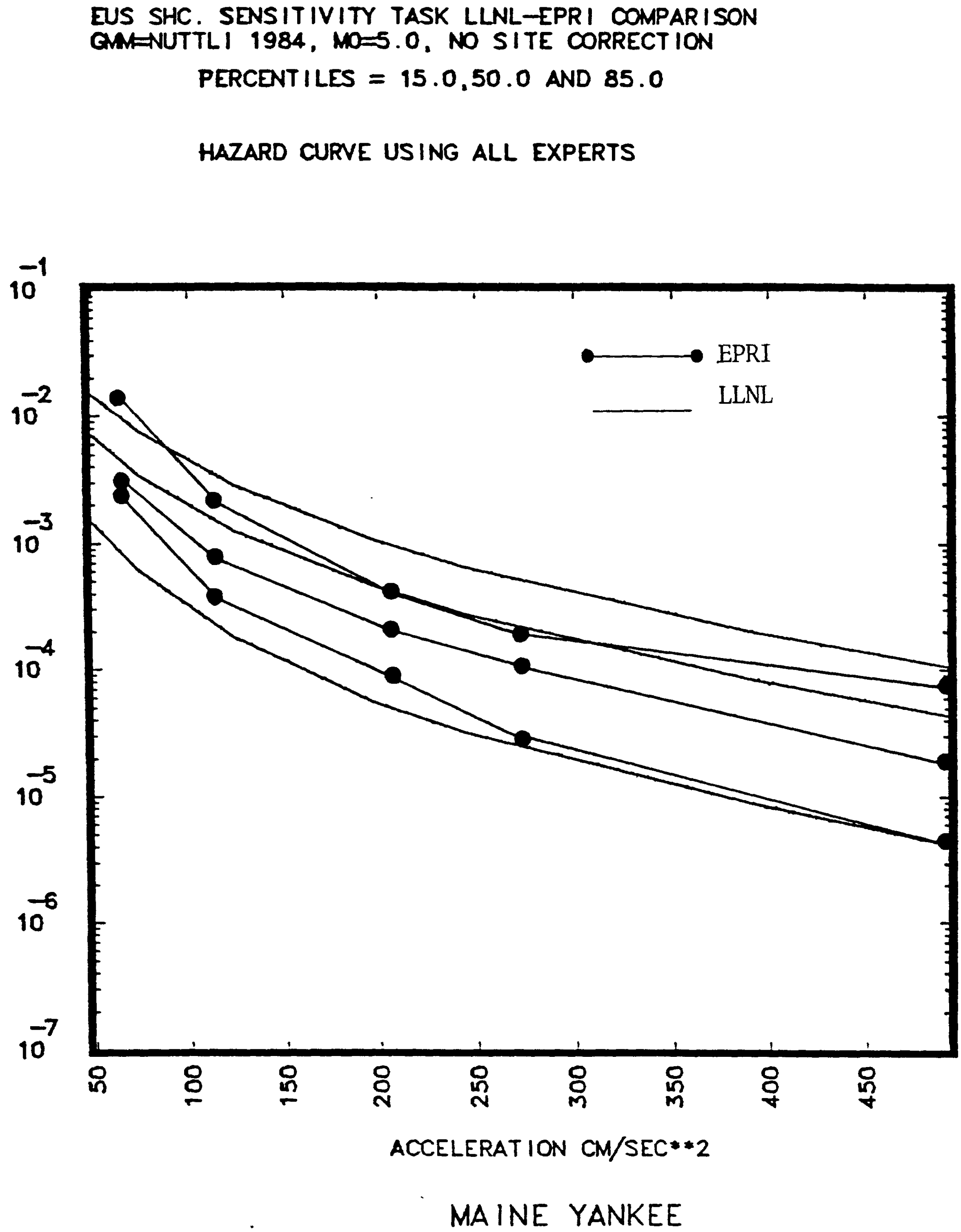


Fig. 5.5 Comparison Between EPRI's CPHC and LLNL's CPHC. Same Ground Motion Model and Lower Bound of Integration Used to Generate the Hazard Curves.

2252 40008

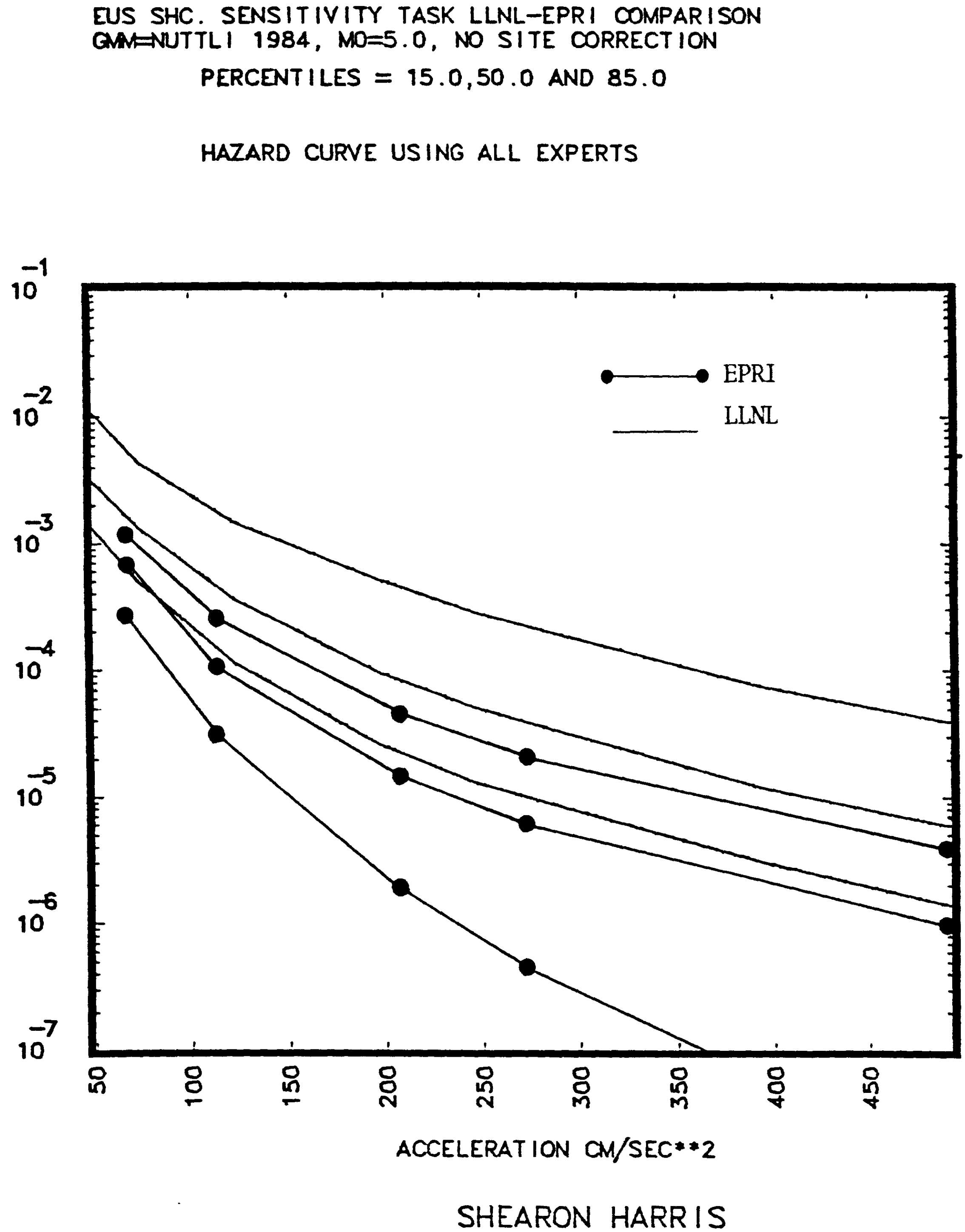


Fig. 5.6 Comparison Between EPRI's CPHC and LLNL's CPHC. Same Ground Motion Model and Lower Bound of Integration Used to Generate the Hazard Curves

2351 40008

and Moore studies were performed to provide seismic hazard estimates for PRA studies for the Millstone and Limerick nuclear power plants. The Yankee Atomic study is the most complete and a full uncertainty analysis was performed and CPHC's were developed. Thus, it is possible to directly compare Yankee Atomic's results to the LLNL results. Yankee Atomic's CPHC for the Maine Yankee site and the LLNL CPHC are compared in Fig. 5.7. It is observed that the two median hazard curves are in reasonable agreement although the LLNL bounds are much wider than Yankee Atomic's bounds. In Section 6 of Bernreuter et al. 1985 similar comparisons between the results of the SHCP and ERTEC's results at Limerick and Dames and Moore's results at Millstone are discussed.

6.0 SUMMARY AND CONCLUSIONS

The results of the LLNL study provide the NRC with the tools for characterizing the seismicity of the EUS and for describing the hazard at any location within that region. These tools are:

(a) A data base of estimates of the seismicity of the EUS, based on expert opinions of the seismicity, in the form of

- A catalog of maps of zonation of the EUS along with estimates of the seismicity of each zone, including a measure of uncertainty.

- A catalog of ground motion models for propagating the motion to any location within the EUS.

(b) A hazard methodology which uses the estimates in (a) to develop an estimate of the seismic hazard at any location in the EUS. The seismic hazard is described in terms of a hazard curve or a uniform hazard spectrum.

In using the data base it must be recognized that the results are based on information which was available to the experts at the time of the elicitation in 1984. As additional events occur and more data become available there may

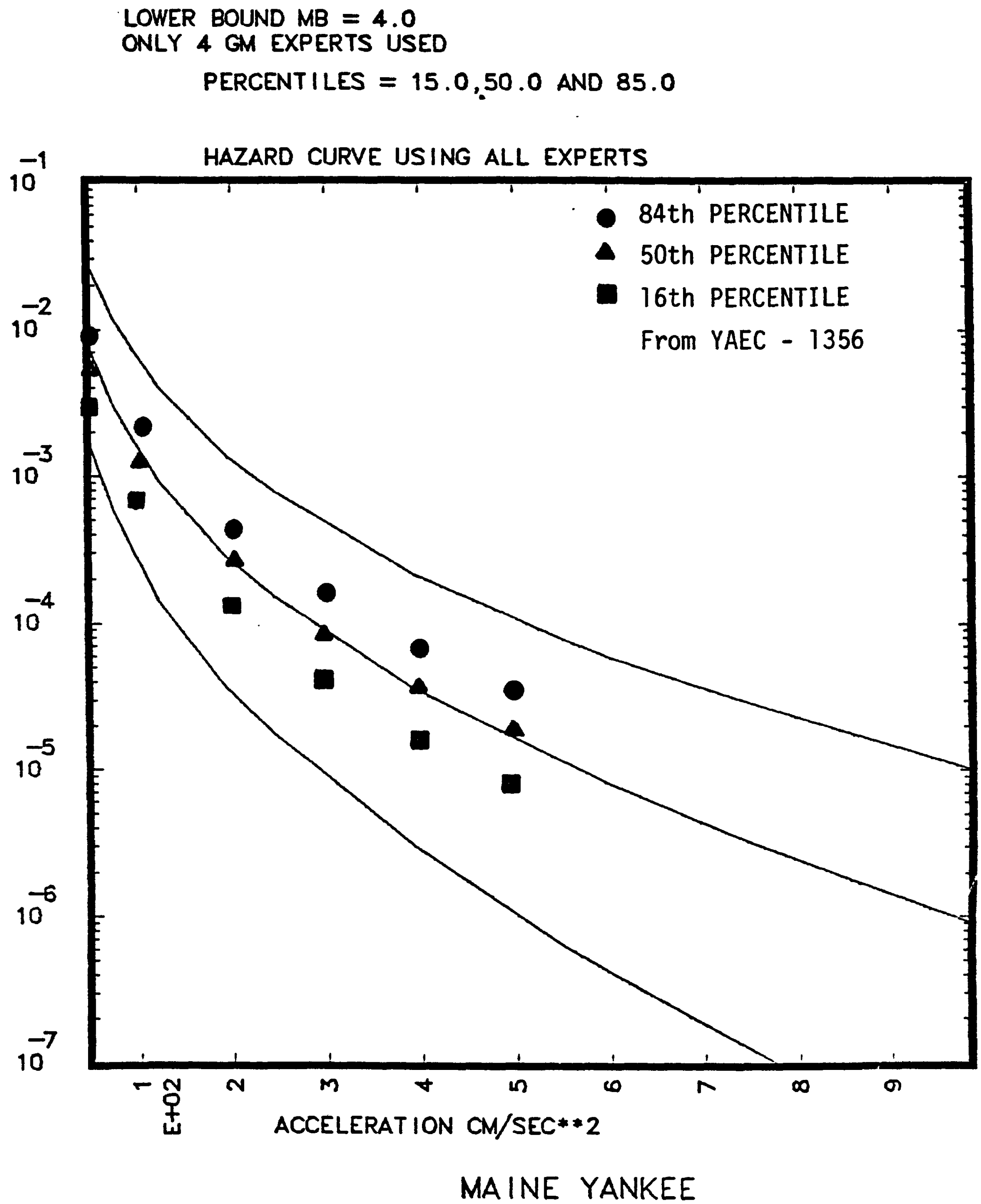


Fig. 5.7 Comparison of CPHC Obtain by using all the LLNL S-Experts and the four G-Experts with Ground Motion Models Similar to the one used by YAEC (1985) and the CPHC from YAEC (1985). A Lower Bound of m_b = 4.0 was used consistent with the YAEC study.

be a basis for a change in opinion. This is particularly true of the ground motion models where there is considerable activity in development of new and improved models.

The detailed conclusions reached in Bernreuter et al. 1985 and illustrated in this paper are:

(1) There is substantial uncertainty in the estimated hazard. The typical range in the value of the probability of exceedance between the 15th and 85th percentile curves for the PGA is on the order of 40 times for low PGA; it is more than 100 times at high PGA values.

The range between the 15th and the 85th percentile hazard curves represents the uncertainty in estimating the seismic hazard at a site due to two sources of uncertainty:

- o The uncertainty of each expert in the zonation, models and values of the parameters of the analyses
- o The variation in the hazard estimates due to the diversity of opinions between experts.

The latter inter-expert variation is an important contributor to the overall uncertainty in the estimated hazard. Specifically, the magnitude of uncertainty introduced by the diversity of opinions between experts is of the same order, on the average, as the uncertainty in the hazard due to the uncertainty of an individual expert in the value of the parameters.

(2) The 50th percentile CPHC appears to be a stable estimator for the seismic hazard at a site. That is, it is the least sensitive to changes in the parameters, when compared to other estimators considered in this study.

(3) The process is reasonably stable. Given the large uncertainties there is reasonable agreement between various studies.

REFERENCES

Algermissen, S.T., Perkins, P.C., Thenhaus, S.L. Hanson, and Bender, B.L., Probabilistic Estimates of Maximum Acceleration and Velocity in Rock in the Contiguous United States, U.S. Geological Survey Open-File Report 82-1033, 1982, 99 pages

Bernreuter, D.L. (1981) Seismic Hazard Analysis: Application of Methodology, Results and Sensitivity Studies, NUREG/CR-1582 (UCRL-53030, Vol. 4).

Bernreuter D.L., Savy, J.B., Mensing, R.W., Chen, H.C., Davis, B.C. (1985) Seismic Hazard Characterization of the Eastern United States: Vol. 1 and Vol.2, LLNL Report UCID 20421 Vol. 1 and Vol. 2.

Cornell, C.A, (1968) "Engineering Seismic Risk Analysis" Bull. Seismic. Soc. Am., Vol. 58, pp. 1583-1606.

Dames and Moore (1983), Revised Report- Seismic Hazard Assessment Millstone Nuclear Power Plant, Unit III, Millstone Point, Connecticut. For Westinghouse, Inc. Dames and Moore, Cranford, N.J., also Appendix 1-B of the Millstone Unit 3 Probabilistic Safety Study, Northeast Utilities, (1983).

Der Kiureghian, A. and A. Ang.(1977) "A Fault-Rupture Model for Seismic Risk Analysis" Bull. Seismic. Soc. Am., Vol. 67, pp. 1173-1194.

EPRI, (1985a), Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States: Volumes 1,2,3, Research Project Number P101-29 (Draft).

EPRI, (1985b) Seismic Hazard Methodology for Nuclear Facilities in the Eastern United States: Preliminary Seismic Hazard Test Computations for Parametric Analysis and Comparative Evaluations EPRI Research Project Number P101-29 (Draft).

ERTEC (1982), Seismic Ground Motion Hazard at Limerick Generating Station, ERTEC Rocky Mountain, Inc., also in Vol. II, Appendix A, Severe Accident Risk Assessment Limerick Generating Station, NUS report 4161,(1983)

McGuire, R.K., FORTRAN Computer Program for Seismic Risk Analysis, U.S. Geological Survey, Open-File Report 76-67, 1976, 90 pp.

Mortgat, C.P. and Shah, H.C. (1979), "A Bayesian Model for Seismic Hazard Mapping", Bull. Seismic. Soc. Am. Vol. 69, pp. 1237-1251.

Savy, J.B., Bernreuter, D.L., and Chen, J.C., (1986) "A Generic Method for Modeling Site Effects in Seismic Hazard Analyses". This publication.

Veneziano, D., Cornell, C.A. and O'Hara, T. (1984) Historical Method of Seismic Hazard Analysis, EPRI NP-3438, Project 1233-10, Interim Report, May 1984.

2358 40008

Yankee Atomic Electric Company, (1982) Supplemental Seismic Probabilistic Study, Report YAEC-1331).

Yankee Atomic Electric Company (1983), Maine Yankee Seismic Hazard Analysis, Report YAEC-1356

Yankee Atomic Electric Company (1984), Review and Comments on NUREG/CR3756 Seismic Hazard Characterization of the Eastern United States: Methodology and Interim Results for Ten Sites, Report YAEC-1455.

AN OVERVIEW OF EPRI's SEISMIC HAZARD METHODOLOGY DEVELOPMENT PROGRAM

J. Carl Stepp
Electric Power Research Institute
Palo Alto, CA 94303

Introduction

The purpose of a probabilistic seismic hazard-analysis is to provide a documented basis for informed decision making about ground-motions appropriate for seismic design of a specific facility at a given site. Seismic hazard is usually depicted as probabilities that given levels of ground shaking will be exceeded annually. Uncertainty on the hazard at any site may be large; therefore, a hazard methodology should include procedures to quantify uncertainty that properly reflect uncertain input parameters and computational models.

Recently, developments of probabilistic seismic hazard methodology specifically to characterize seismic hazard at low probabilities of ($<10^{-3}$ per year) at locations in the central and eastern United States have been based on input interpretations by multiple experts (Bernreuter and Minichino 1983; Bernreuter, et al, 1985). In these studies a number of individual scientists provide interpretations of seismic source zones and their associated seismicity parameters. To express an individual's uncertainty, multiple alternative interpretations are elicited. Uncertainty about seismic wave attenuation is treated similarly by eliciting weights on potentially applicable attenuation relationships from multiple experts.

An ongoing seismic hazard methodology development program at the Electric Power Research Institute (EPRI) follows this general approach (EPRI, 1986). However, a number of modifications have been incorporated to minimize bias and unquantified uncertainty, to provide fully trackable interpretations of input parameters based on state-of-the-art earth science practice, to specifically distinguish scientific and information uncertainty, and to make maximum use of historic earthquake data. The goal of the program is to develop procedure that are consistent with earth science practice, that facilitate expressions of uncertainty in seismic hazard input interpretations, and are general applicable.

Motivation

During the past ten years expanded studies of the seismotectonics of the central and eastern United States have been undertaken in programs funded by the U.S. Geological Survey (USGS) and the U.S. Nuclear Regulatory Commission (USNRC). Notwithstanding the substantial effort, with the exception of the New Madrid seismic zone, specific tectonic structural explanations of earthquake activity have not been identified. Indeed, extensive focused studies in the region of the Charleston Earthquake of 1886 (Rankin, 1977; Gohn, 1983) have identified alternative hypotheses to explain the tectonic structural cause of the event, but no single preferred explanation has emerged. Moreover, no unique tectonic feature has been identified that would suggest that the Charleston area has unique seismotectonic characteristics relative to other regions of the central and eastern United States. These findings together with evolving general scientific understanding of seismicity and tectonic strain release have lead to two important conclusions that impact modeling of seismic hazard.

- Earthquakes may occur in the central and eastern United States wherever favorable tectonic features and conditions exist.

- The historic record of earthquake activity in the central and eastern United States is likely too short to itself define sources of future moderate and large earthquakes.

These conclusions suggest that a major source of uncertainty in seismic hazard estimates, that is perhaps irreducible in the short term, is related to seismic source zone interpretations. The development of criteria to assess activity (to quantify favorable tectonic features and conditions) and a structured approach to use the criteria to interpret alternative seismic source zone geometries and source seismicity parameters have been the major focus of the EPRI seismic hazard program. The procedures rest on a number of assumptions.

Assumptions

It is assumed that alternative sources of future occurrences of moderate and large earthquakes in the central and eastern United States can be deduced from tectonic hypotheses and geological, geophysical, and tectonic data, and historic seismicity. Earthquakes are assumed to occur as a result of frictional failure processes and criteria to assess the relative probabilities that tectonic features are active can be established based on current understanding of the tectonic stress regime, failure mechanisms, association with seismicity, and tectonic history and characteristics of tectonic features (Coppersmith and Youngs, 1986; EPRI, 1986).

A basic premise has been that the methodology should not constrain tectonic bases for seismic source zones. Only the mild constraints are imposed that each source zone has a constant probability of activity and a single distribution on maximum magnitude throughout. The distribution of earthquakes in a source is assumed to be derivable from the existing earthquake catalog, seismotectonic characteristics of the source and analogy with other sources having similar characteristics. No earthquake occurrence model is assumed, although a single recurrence distribution applies to each source zone, and the usual constraint of spatially homogeneous seismicity within a source is avoided; that is it is assumed that the rate of activity may vary spatially within a source zone. Choices of the applicable occurrence distribution and the degree of variation in rate of activity within seismic sources are left to the interpreter. The primary emphasis has been placed on obtaining interpretations of the tectonic framework, sources of future moderate and large earthquakes, and source seismicity parameters consistent with state-of-the-art earth science practice and available geoscience data.

Methodology Developments

The EPRI seismic hazard methodology has incorporated a number of new developments and approaches, both scientific and procedural. These include formation of Earth Science Teams to perform all input interpretations; development and utilization of a common, uniform data base for interpretations; a structured approach to develop interpretations and to characterize uncertainty on interpretations that specifically includes separate considerations of scientific and data uncertainty; extensive analysis and upgrading of the historic earthquake catalog; and new procedures to estimate seismicity parameters.

Input interpretations for hazard assessment have been developed by six Earth Science Teams rather than by individual scientists. Each team constituted minimally a geologist, a seismologist, and a tectonophysicist; an attempt was made to strike a balance between practicing professionals and academic researchers. Since an indicator of activity is the extent of tectonic features in the frittle crust, most teams also included a geophysicist. Typically earth scientists specialize along disciplines and a single scientist rarely is equally proficient in the range of discipline expertise needed to interpret sources of future earthquakes. Teams were constrained to reach within-team consensus on all interpretations. By taking the team approach it is believed that bias in the input interpretations due to uneven discipline knowledge is minimized.

A strong emphasis of the EPRI program has been the development of a uniform data base for the entire region of the central and eastern United States (King and Stepp, 1985). Typically source interpretations have drawn on available, published data and any additional information that individuals providing interpretations might have available. For the EPRI program a specific effort has been made to identify data requirements to assess competing hypotheses for earthquakes in the central and eastern United States (Table 1). The data were analyzed specifically to emphasize anomalies in the brittle zone of the Earths' crust. Uniform data base products as shown in Table 2 were provided to all participants in the program on maps at scales of 1:2.5 M and 1:1 M. By this approach it is believed that bias due to uneven knowledge of data among teams has been minimized.

Table 1

CENTRAL AND EASTERN UNITED STATES EARTHQUAKE HAZARD DATA BASE

Reference Information

a. Bibliographic reference system

b. Map reference system

c. Crustal structure of the eastern United States

Data Catalogs

a. Earthquake catalog

b. Stress catalog

c. Source parameter catalog

Digital Data Sets

a. Magnetic anomaly

b. Bouguer gravity

c. Free-air gravity

d. Isostatic gravity

e. Topography

f. Basement drillhole data

Table 2

DATA BASE MAPS FOR SEISMIC HAZARD INTERPRETATIONS

Seismicity 1534-1984

Magnitudes $\geq$ 4.0, intensities $\geq$ V

Seismicity 1534-1984

All magnitudes and intensities

Seismicity 1811-1984, New Madrid Area

All magnitudes and intensities

Cumulative Seismic Moment

Earthquake Source Mechanisms and Crustal Stress Orientations

Free Air Gravity Anomalies

Unfiltered
125 km high-pass filter
250 km high-pass filter
Horizontal gradient

Bouguer Gravity Anomalies

125 km high-pass filter
250 km high-pass filter
Horizontal gradient
Vertical gradient

Isostatic Gravity Anomalies

Topography

125 km high-pass filter
Horizontal gradient

Magnetic Intensity Anomalies

125 km high-pass filter
Horizontal gradient

Gravity and magnetic data were compiled to aid in mapping potential earthquake sources. These data proved to be valuable for mapping the geometry and geographic extent of tectonic features as well as for mapping lateral changes in material properties in the Earths' crust that are potential sources of earthquakes. Thus, these data were a fundamental basis for the Earth Science Teams identification and mapping of potentially active tectonic features as well as their tectonic framework interpretations. To be consistent, all data were gridded at 4 km, high-pass filtered to enhance anomalous sources within the brittle crust, and mapped on common scales of 1:2.5 M and 1:1.0 M. Vertical and horizontal derivatives were taken to aid in delineating boundaries of tectonic features.

Development of the seismic hazard methodology and interpretations took place in four distinct activities, each designed to accomplish development of a specific product as follows:

- Definition of data needs, compilation of data and construction of a data base, analysis of data and development of data base products;
- Evaluation of tectonic processes and development tectonic stress regime interpretations;
- Evaluation of geomechanical processes of failure and development of tectonic framework and seismic source zone procedures and interpretations; and
- Analyses and upgrading of the earthquake catalog and development of source zone seismicity parameter procedures and interpretations.

All of these developments took place through a process of intensive interactions among participants in the program. The structure selected to implement these interactions was a series of seven workshops as shown in Table 3.

Table 3

SEISMIC HAZARD DEVELOPMENT WORKSHOPS

Workshop	Objective
1. Data Needs	Determine data compilation and processing needed during program.
2. Tectonic Processes and Tectonic Stress Regime	Identify tectonic processes active in the central and eastern United States and their contribution to crustal stresses.
3. Tectonic Stress Regime Interpretations	Review Earth Science Team's Interpretations of Tectonic Stress Regime
4. Methods for Tectonic framework and Seismic Source zone interpretations	Identify geomechanical of processes, failure mechanics and tectonic characteristics diagnostic of earthquake potential.
5. Tectonic Framework and Seismic Source Zone Interpretations	Evaluate Earth Science Team's tectonic frameworks and seismic sources.
6. Methods for Seismicity Parameters Interpretations	Evaluate recurrence models and establish methods of estimating seismicity parameters.
7. Seismicity Parameters Interpretations	Review interpretations of seismicity parameters for seismic sources.

Workshop 1 accomplished definition of data needs for the program including the types and scope of data to be compiled and the types of analyses and manner of data presentation to be accomplished. The starting point to achieve these objectives was a compilation of hypotheses for earthquake causes in the central and eastern United States. Included were hypotheses known in the published literature and those additional ones that were identified by the program participants as having any level credibility. The composite data needs consisted of an assessment of data requirements to evaluate the credibility of each hypothesis as a local earthquake cause and to map associated tectonic features in the Earths' crust.

Workshops 2 through 7 were structured in pairs to accomplish development of approaches and interpretations of tectonic stress regime, tectonic framework and seismic source zones, and source seismicity parameters. The first of each workshop pair was devoted to developing of a state-of-knowledge information base and interpretation procedures. Working papers for these workshops were prepared by the program participants and with additional input and participation of key researchers. Procedures were explored in depth during each workshop to establish a common understanding among participants of the state-of-knowledge about processes and the weight of available data to perform interpretations. The Earth Science Teams proceeded with this information to make their independent interpretations. The second workshop of each pair was devoted to reviewing the Earth Science Teams interpretations. Each Earth Science Team shared with the program participants its rationale and the strength of theory and data to support its interpretation. No effort was made to reach a consensus among teams on any interpretation element, but teams were asked to reach internal consensus on all within-team interpretations. Follow-on meetings were held with each team to clarify procedures and facilitate their application.

In addition to the workshops and interactive meetings, two seminars were held to develop state-of-knowledge perspectives of:

- Methods to define tectonic mechanisms causing earthquakes in the central and eastern United States; and

- Stress concentration mechanisms, geomechanical processes of failure, and earthquake potential in the central and eastern United States.

These seminars each brought together a small number of researchers working on these subjects. They were asked to share their latest research, to identify reasonably resolved as well as unresolved issues, and to indicate the weight of data for application interpretations. Proceedings of these seminars formed the bases for working papers used in Workshops 2 and 4.

Together the seminars, workshops, and interactive meetings and the focused step-by-step development of methodology and interpretations provided a powerful structure for developing state-of-knowledge information, obtaining a common understanding of its strengths and limitations, and applying the knowledge base to develop procedures and interpretations. By taking this approach, it is believed that unquantified uncertainty due to different data bases and variations in understanding of processes has been reduced and that the uncertainty in hazard results reflects the state of the scientific community's uncertainty about earthquake causes and processes in the central and eastern United States.

Procedures to interpret seismic source zones emphasize evaluations of the causes of earthquakes (Coppersmith and Youngs, 1986; EPRI, 1986). The structured approach involves compilation of hypotheses of earthquake causes, identification candidate tectonic features, development of criteria to assess the probability that each candidate tectonic feature is active and procedures to aggregate active tectonic features into seismic source zones. Documentation and justification of each step in the interpretation has been emphasized. By this approach the basis for seismic source zone interpretations in theory and in geological, geophysical, and seismological data is made clear. The entire interpretation is amenable to peer review.

Procedures to assess activity of candidate tectonic features and to develop a tectonic framework for seismic source zone interpretations (Coppersmith and Youngs, 1986; EPRI, 1986) has been a major development of the EPRI program. The procedure facilitates assessment of uncertainty associated with the interpretation by separately considering the resolution of the criteria as indicators of activity and the resolution of existing data to resolve whether or not a combination of criteria is associated with a given tectonic feature. Weights assigned to combinations of criteria by a term are statements of the teams' scientific uncertainty about the degree to which the criterion or combination can resolve activity, scientific

2352 40008

uncertainty. A teams' numerical weight given to the observation of a combination of criteria associated with a tectonic feature, given the existing data, represents its assessment of the data uncertainty. All candidate tectonic features are evaluated by this procedure, separately by each Earth Science Team, and each is assessed a marginal probability of activity. The resulting interpretation is a map of tectonic features each having an assessed marginal probability of activity which is a measure of the teams' consensus that it is active in the present tectonic stress regime. This interpretation has been referred to as the tectonic framework and it forms the basis for subsequent alternative seismic source zone interpretations.

The procedure to interpret seismic source zones consists of two parts: definition of source zone geometry and derivation of probability that the source is active (Coppersmith and Youngs, 1986; EPRI, 1986). The starting bases for defining source geometries are the tectonic features contained in the tectonic framework interpretation; the marginal probabilities of activity assessed for elements of the tectonic framework form the basis for deriving the probability that each source zone is active. Line sources and area sources can be accommodated equally well. Seismic source zones may represent single elements of the tectonic framework, classes of features (e.g., traissic basins, families of faults), or groups of features (e.g., a group of features that are not easily separable, but are interpreted to be active in the tectonic stress regimes). Feature specific sources and sources representing classes of features can be treated with equal ease; the probability that the source is active is the marginal probability that the feature or class of features is active. For sources representing groups of features, the probability that the seismic source is active is equal to the probability that at least one of the features is active. These together with default and background zones that are also incorporated into the procedure, give flexibility to completely specify seismic source zone geometries and their probabilities of being active that is needed to capture complex alternative tectonic interpretations. Alternative interpretations result in multiple combinations of active sources that might affect a site, depending on interpretations of dependencies among them. Depiction of alternative interpretations is facilitated by use of logic tree structures.

A number of innovations in the assessment temporal and spatial rates of earthquake occurrence have been made in the EPRI program (EPRI, 1986, Veneziano, 1986). The primary objective has been to maximize use of the available earthquake data set.

24#1 40008

The earthquake catalog itself has been carefully evaluated by the Earth Science Teams. Beginning with available national and local network catalogs a consensus catalog has been developed that is believed to be free to duplications and extraneous events such as explosions for all events larger than magnitude m_b4 1/2. The seismic hazard computational methodology used is sufficiently general to accept different seismicity models. However, the Earth Science Teams unanimously preferred the classic exponential distribution on earthquake magnitude. The applicability of this model for low strain rate regions such as the central and eastern United States has been demonstrated by Carnell and Winterstein (1986). Therefore, for each seismic source zone earthquake occurrence is determined by estimating the three parameters of the magnitude distribution, (i.e., the activity rate or a-value and the slope of the distribution, or b-value) and the maximum magnitude. The body wave magnitude scale m_b was selected as the most useful for central and eastern North America earthquakes. Procedures were developed to make maximum use of the earthquake catalog to estimate the activity rate and distribution.

The earthquake catalog was subjected to extensive analysis to: 1) establish a single magnitude measure m_b with error estimate, 2) analyze clustering (classify earthquakes as main or dependent), and 3) estimate incompleteness (Veneziano, 1986). The procedure used to obtain uniform magnitude estimates uses all measures of size (maximum intensity, felt area, magnitude) for each earthquake in the catalog and yields an estimate of error. The estimate is based on correlations between m_b and other measures of earthquake size developed by establishing appropriate correlations using recent events in the catalog that have multiple size observations reported. The procedure to identify clusters and aftershocks is general and accommodates spatially non-homogeneous catalogs with incompleteness-induced non-stationarity. By this procedure main earthquakes are distinguished from dependent events in the master catalog for recurrence estimation. Incompleteness is assessed using a probability-of-detection model which is a function of time, magnitude, and geographic position. Inputs for this assessment are the earthquake catalog and designated completeness regions that are similar with respect to demographic trends and seismograph coverage. The usual period of completeness for a given magnitude and geographic position is replaced by an **"equivalent period of completeness"**. The

24#2 40008

latter is the time interval by which the total number of events in a magnitude range must be divided to obtain and unbiased estimate of the recurrence rate (Veneziano, 1986). Together these analyses permit maximum use of the earthquake catalog to estimate source zone seismicity parameters.

The seismicity model developed and used in the EPRI program is the exponential recurrence model, with procedures to account for catalog incompleteness, clustering (aftershocks), non-stationarity and error in magnitude estimation (EPRI, 1986, Veneziano, 1986). The model assumes that earthquakes occur in clusters and that the clusters themselves are distributed in time and space according to a Poisson process. This is a relaxation of the usual assumption that all earthquakes occur independently according to a Poisson process. Estimation of the occurrence parameters is accomplished simultaneously with the estimation of incompleteness.

The procedure allows the interpretation of spatially non-homogeneous seismicity within a source zone. This option is controlled by the user by specifying the amount of smoothing of the rate and distribution parameters within a source zone. Total smoothing represents the usual assumption of homogeneous seismicity and no smoothing represents the interpretation that future earthquakes will be distributed in space as past events contained in the earthquake catalog. These innovations provide for flexibility in interpretations while maximizing use of the available earthquake information.

Assessment of uncertainty in hazard estimates is facilitated by use of a logic tree structure to weight alternative input interpretations (Coppersmith and Youngs, 1986; EPRI, 1986). Following usual practice, each node of the logic tree represents an element (source zone geometry, maximum magnitude, etc.) of the input interpretation and each branch of a node represents a permissionable alternative interpretation of the element. The probability given each branch the assessed weight on that alternative relative to other permissible interpretations. Each end branch of the logic tree structure represents an weighted alternative interpretation of all input parameters and can be used for a single seismic hazard computation. The computation has a weight which is the product of the assessed probability values of all intermediate branches. The uncertainty in seismic hazard at a site based on an Earth Science Teams' alternative interpretations of all input parameters is then the range of hazard results obtained from the end branches for all alternative source

24#3 40008

zone combinations relevant to that site. Total uncertainty can be obtained through a weighted combination of the six Earth Science Teams' interpretations (EPRI, 1986).

The seismic hazard methodology developed in the EPRI program is believed to be completely general and to be applicable to intraplate regions such as eastern North America. With minor modifications (tectonic activity criteria, manner of specifying source dimensions in the attenuation function) it is believed to be equally applicable to other tectonic environments. The method has been found to be easily applicable as a structure for interpretations input to seismic hazard analysis and has the power of specifically requiring weighted specification of all alternative interpretations. The implementing interpretations developed for the central and eastern United States (Litehiser, et al, 1986; McWhorter, et al, 1986; Barstow, et al, 1986; Holt, et al, 1986; Statton, et al, 1986; White, et al, 1986) are considered to be state-of-the-art, given current uncertainty in our understanding of earthquake processes and limitations of available data. Sensitivity analysis obtained by conditioning the hazard results on various input parameters should emphasize areas for future research that could significantly reduce uncertainty in the results.

References

Barstow, Noel, William Hinze, Pradeep Talwani, and Barry Voight, 1986, "Seismic Hazard Methodology for the Central and Eastern United States, Volume 10: Seismic Source Zones and Seismicity Parameters"; NP-4661, Vol. 10, Electric Power Research Institute.

Bernreuter, D. L., and C. Minichino, 1983, "Seismic Hazard Analysis Overview and Executive Summary"; NUREG/CR-1582, Vol. 1, U.S. Nuclear Regulatory Commission.

Bernreuter, D. L., J. B. Savy, R. W. Mensing, J. C. Chen, and B. C. Davis, 1985, "Seismic Hazard Characterization of the Eastern United States, Volume 1: Methodology and Results for Ten Sites"; UCID-20421 Vol. 1, Lawrence Livermore National Laboratory.

Coppersmith, K., and R. Youngs, 1986, this volume.

Cornell, C. A., and S. R. Winterstein, 1986, "Seismic Hazard Implications of Non-Poissonian Recurrence Models"; abs, Earthquake Notes, Vol. 57, No. 1, p. 25.

Gohn, Gregory S., ed, 1983, "Studies Related to the Charleston, South Carolina, Earthquake of 1886 - Tectonics and Seismicity"; Geological Survey Professional Paper 1313, U.S. Geological Survey.

EPRI, 1986, "Seismic Hazard Methodology for the Central and Eastern United States, Volume 1: Probabilistic Seismic Hazard Methodology"; NP-4661, Vol. 1, Electric Power Research Institute.

King, J. L., and J. C. Stepp, 1984, "National Geophysical Data Sets for Assessing Earthquake Potential in the Central and Eastern United States"; Proceedings, Conference on Pathways and Future Directions for Environmental Data and Information Users, NOAA.

Klimkiewiez, G., R. Holt, G. Leblanc, and D. Wise, 1986, "Seismic Hazard Methodology for the Central and Eastern United States Volume 5: Seismic Source Zones and Seismicity Parameters"; NP-4661, Vol. 5, Electric Power Research Institute.

Litchiser, J., T. Buschback, R. D. Hatcher, R. Stanley and I. Zietz, 1986, "Seismic Hazard Methodology for the Central and Eastern United States, Volume 9: Seismic Source Zones and Seismicity Parameters"; NP-4661, Vol. 9, Electric Power Research Institute.

McWhorter, J., C. Fairhurst, R. Herrmann, L. McGinnis, R. Rodriguez, 1986; "Seismic Hazard Methodology for the Central and Eastern Unites States, Volume 6: Seismic Source Zones and Seismicity Parameters"; NP-4661, Vol. 6, Electric Power Research Institute.

Rankin, D. W., ed, 1977, "Studies Related to the Charleston, South Carolina, Earthquake of 1886 - A Preliminary Report"; Geological Survey Professional Paper 1028, U.S. Geological Survey.

Statton, C. T., T. Engelder, J. Kelleher, R. Quittmeyer, and T. Turcotte, 1986, "Seismic Hazard Methodology for the Central and Eastern United States, Volume 8: Seismic Source Zones and Seismicity Parameters": NP-4661, Vol. 8, Electric Power Research Institute.

White, R., R. Buttler, M. Chapman, J. Chulick, J. Dwyer, A. Johnson, L. Long, M. Schaeffer, and W. Seay, 1986, "Seismic Hazard Methodology for the Central and Eastern United States, Volume 7: Seismic Source Zones and Seismicity Parameters"; NP-4661, Vol. 7, Electric Power Research Institute.

Veneziano, D., 1986, this volume.

APPENDIX A

WORKSHOP ON PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENTS
San Francisco, California
November 25-27, 1985

Professor Keitti Aki
Department of Geological Sciences
University of Southern California
Los Angeles, California 90089

Dr. Shelton S. Alexander
Head, Dept. of Geosciences
Professor of Geophysics
403 Deike Building
Penn State University
University Park, Pennsylvania 16802

Dr. Ted Algermissen
U.S. Geological Survey
Box 25046, MS 966
Denver Federal Center
Denver, Colorado 80225

Dr. Walter Arabasz
Department of Geology and Geophysics
University of Utah
717 Mineral Science Building
Salt Lake City, Utah 84112

Mr. Rodrigo Araya
834 Riley Dr. #84
Albany, California 94706

Ms. Bernice Bender
U.S. Geological Survey
Box 25046, MS 966
Denver Federal Center
Denver, Colorado 80225

Mr. Leon Beratun
Mail Stop: 1130-55
U.S. N.R.C.
Washington, D.C. 20555

Mr. Donald Bernreuter
Earth Sciences Department
Lawrence Livermore National Laboratory
University of California
Livermore, California 94550

Mr. Rafael Blazquez
1009 Ninth St., Apt. 52
Albany, California 94710

Dr. David Boore
U.S. Geological Survey
Branch of Engineering Seismology
345 Middlefield Road, MS 977
Menlo Park, California 94025

Dr. Stephan Brocoum
Mail Stop P-514
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. Ken Campbell
U.S. Geological Survey
Box 25046, MS 967
Denver Federal Center
Denver, Colorado 80225

Dr. Kevin Coppersmith
Geomatrix Consultants
One Market Plaza
Spear Street Tower, Suite 717
San Francisco, California 94105

Professor Allin Cornell
Dept. of Civil Engineering
Terman Engineering Center
Stanford University
Stanford, California 94305

John J. Dwyer
Law Engineering
2749 Delk Drive
Marietta, California 30067

Mr. Gus Giese-Koch
Mail Stop P-514
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. Walter Hays
U.S. Geological Survey
905 National Center
Reston, Virginia 22092

Mr. Buck Ibrahim
U.S. Nuclear Regulatory Commission
MS 623-SS
Washington, D.C. 20555

Mr. John Jacobson
Yankee Atomic Electric Co.
1671 Worcester Road
Framingham, Massachusetts 01701

Dr. William Joyner
U.S. Geological Survey
Branch of Engineering Seismology
345 Middlefield Road, MS 977
Menlo Park, California 94025

Dr. Jerry King
Electric Power Research Institute
3412 Hillview Avenue
P.O. Box 10412
Palo Alto, California 94303

Professor Armen Der Kiuregian
University of California, Berkeley
Department of Civil Engineering
725 Davis Hall
Berkeley, California 94720

Dr. Ram Kulkarni
Woodward Clyde Consultants
One Walnut Creek Center
100 Pringle Avenue
Walnut Creek, California 94596

Mr. Cole R. McClure
Manager of Engineering Geology
Bechtel
P.O. Box 3965
San Francisco, California 94022

Dr. Robin McGuire
Risk Engineering Inc.
5255 Pine Ridge Road
Golden, Colorado 80403

Dr. Richard Mensing
Nuclear Systems Safety Program
Lawrence Livermore National Laboratory
P.O. Box 808
Livermore, California 94550

Mr. Givliano Milana
Via Altino 16
00183 Roma, Italy

Dr. Christian Mortgat
TERA Corp.
2150 Shattuck Avenue
Berkeley, California 94701

Dr. Andrew Murphy
Mail Stop: 1130-SS
Nuclear Regulatory Commission
Washington, D.C. 20555

Mr. David Perkins
U.S. Geological Survey
Box 25046, MS 966
Denver Federal Center
Denver, Colorado 80225

Dr. Paul Pomeroy
Rondout Associates Inc.
P.O. Box 224
Stone Ridge, New York 12484

Mr. Maurice Power
Geomatrix Consultants
One Market Plaza
Spear Street Tower, Suite 717
San Francisco, California 94105

Dr. Leon Reiter
P-514
Office of Nuclear Reactor Regulation
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. Albert Rogers
U.S. Geological Survey
Denver Federal Center
P.O. Box 20456, MS 966
Denver, Colorado 80224

Dr. Robert Rothman
Mail Stop P-514
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. Jean Savy
Nuclear Systems Safety Program
Lawrence Livermore National Laboratory
P.O. Box 808
Livermore, California 94550

Mr. Leonells Serva
ENEA-DISP
Via Vitaliano Brancati 48
00143 Roma, Italy

Professor Haresh Shah
Stanford University
Dept. of Civil Engineering
Stanford, California 94305

2554 40008

Dr. Jogeshwar P. Singh
Harding-Lawson Associates
666 Howard Street, 3rd Floor
San Francisco, California 94105

Professor Burton Slemmons
Geological Sciences Department
Mackay School of Mines
University of Nevada
Reno, Nevada 89557

Dr. Phyllis Sobel
Mail Stop P-514
Nuclear Regulatory Commission
Washington, D.C. 20555

Mr. Tom Statton
Woodward Clyde Consultants
201 Willowbrook Blvd.
P.O. Box 290
Wayne, New Jersey 07470

Dr. Carl Stepp
Electric Power Research Institute
3412 Hillview Avenue
P.O. Box 10412
Palo Alto, California 94303

Mr. Paul Thenhaus
U.S. Geological Survey
Box 25046, MS 966
Denver Federal Center
Denver, Colorado 80225

Dr. Nafi Toksoz
Department of Earth and Planetary Sciences
Room E 34-440
Massachusetts Institute of Technology
Cambridge, Massachusetts 01239

Professor Danielle Veneziano
Massachusetts Institute of Technology
Room I-382
Department of Civil Engineering
Cambridge, Massachusetts 02139

Dr. Ian Wall
Electric Power Research Institute
P.O. Box 10412
3412 Hillview Avenue
Palo Alto, California 94303

2558 40008

APPENDIX B

GLOSSARY OF TERMS USED IN PROBABILISTIC EARTHQUAKE HAZARDS ASSESSMENTS

Accelerogram. The record from an accelerometer showing acceleration as a function of time. The peak acceleration is the largest value of acceleration on the accelerogram.

Acceptable Risk. A probability of occurrences of social or economic consequences due to earthquakes that is sufficiently low (for example in comparison to other natural or manmade risks) as to be judged by appropriate authorities to represent a realistic basis for determining design requirements for engineered structures, or for taking certain social or economic actions.

Active fault. A fault is active if, because of its present tectonic setting, it can undergo movement from time to time in the immediate geologic future. This active state exists independently of the geologists' ability to recognize it. Geologists have used a number of characteristics to identify active faults, such as historic seismicity or surface faulting, geologically recent displacement inferred from topography or stratigraphy, or physical connection with an active fault. However, not enough is known of the behavior of faults to assure identification of all active faults by such characteristics. Selection of the criteria used to identify active faults for a particular purpose must be influenced by the consequences of fault movement on the engineering structures involved.

Asthenosphere. The worldwide layer below the lithosphere which is marked by low seismic wave velocities. It is a soft layer, probably partially molten.

Attenuation law. A description of the average behavior of one or more characteristics of earthquake ground motion as a function of distance from the source of energy.

Attenuation. A decrease in seismic signal strength with distance which depends not only on geometrical spreading, but also may be related to the physical characteristics of the transmitting medium that cause absorption and scattering.

b-value. A parameter indicating the relative frequency of earthquakes of different sizes derived from historical seismicity data.

Capable fault. A fault along which future surface displacement is possible, especially during the lifetime of the engineering project under consideration.

Convection. A mechanism of heat transfer through a liquid in which hot material from the bottom rises because of its lesser density, while cool surface materials sinks.

Convergence Zone. A band along which moving plates collide and area is lost either by shortening and crustal thickening or subduction and destruction of crust. The site of volcanism, earthquakes, trenches, and mountain building.

Design earthquake. A specification of the ground motion at a site based on integrated studies of historic seismicity and structural geology used for the earthquake-resistant design of a structure.

Design spectra. Spectra used in earthquake-resistant design which correlate with design earthquake ground motion values. Design spectra typically are smooth curves that take into account features peculiar to a geographic region and a particular site.

Design time history. One of a family of time histories used in earthquake-resistant design which produces a response spectrum enveloping the smooth design spectrum, for a selected value of damping.

Duration. A qualitative or quantitative description of the length of time during which ground motion at a site exhibits certain characteristics such as being equal to or exceeding a specified level of acceleration such as 0.05g.

Earthquake hazards. The probability that natural events accompanying an earthquake such as ground shaking, ground failure, surface faulting, tectonic deformation, and inundation, which may cause damage and loss of life, will occur at a site during a specified exposure time. See earthquake risk.

Earthquake risk. The probability that social or economic consequences of earthquakes, expressed in dollars or casualties, will equal or exceed specified values at a site during a specified exposure time.

Earthquake waves. Elastic waves (P, S, Love, Rayleigh) propagating in the Earth, set in motion by faulting of a portion of the Earth.

Effective peak acceleration. The peak ground acceleration after the ground-motion record has been filtered to remove the very high frequencies that have little or no influence upon structural response.

Elastic rebound theory. A theory of fault movement and earthquake generation that holds that faults remain lock while strain energy accumulates in the rock, and then suddenly slip and release this energy.

Epicenter. The point on the Earth's surface vertically above the point where the first fault rupture and the first earthquake motion occur.

Exceedance probability. The probability (for example, 10 percent) over some period of time that an event will generate a level of ground shaking greater than some specified level.

Exposure time. The period of time (for example, 50 years) that a structure is exposed to the earthquake threat. The exposure time is sometimes related to the design lifetime of the structure and is used in seismic risk calculations.

Fault. A fracture or fracture zone in the Earth along which displacement of the two sides relative to one another has occurred parallel to the fracture. See Active and Capable faults.

Focal depth. The vertical distance between the hypocenter and the Earth's surface in an earthquake.

Ground motion. A general term including all aspects of motion; for example, particle acceleration, velocity, or displacement; stress and strain; duration; and spectral content generated by a nuclear explosion, an earthquake, or another energy source.

Intensity. A numerical index describing the effects of an earthquake on the Earth's surface, on man, and on structures built by him. The scale in common use in the United States today is the Modified Mercalli scale of 1931 with intensity values indicated by Roman numerals from I to XII. The narrative descriptions of each intensity value are summarized below.

I. Not felt--or, except rarely under especially favorable circumstances. Under certain conditions, at and outside the boundary of the area in which a great shock is felt: sometimes birds and animals reported uneasy or disturbed; sometimes dizziness or nausea experienced; sometimes trees, structures, liquids, bodies of water, may sway--doors may swing, very slowly.

II. Felt indoors by few, especially on upper floors, or by sensitive, or nervous persons. Also, as in grade I, but often more noticeably: sometimes hanging objects may swing, especially when delicately suspended; sometimes trees, structures, liquids, bodies of water, may sway, doors may swing, very slowly; sometimes birds and animals reported uneasy or disturbed; sometimes dizziness or nausea experienced.

III. Felt indoors by several, motion usually rapid vibration. Sometimes not recognized to be an earthquake at first. Duration estimated in some cases. Vibration like that due to passing of light, or lightly loaded trucks, or heavy trucks some distance away. Hanging objects may swing slightly. Movements may be appreciable on upper levels of tall structures. Rocked standing motor cars slightly.

IV. Felt indoors by many, outdoors by few. Awakened few, especially light sleepers. Frightened no one, unless apprehensive from previous experience. Vibration like that due to passing of heavy or heavily loaded trucks. Sensation like heavy body of striking building or falling of heavy objects inside. Rattling of dishes, windows, doors; glassware and crockery clink or clash. Creaking of walls, frame, especially in the upper range of this grade. Hanging objects swung, in numerous instances. Disturbed liquids in open vessels slightly. Rocked standing motor cars noticeably.

V. Felt indoors by practially all, outdoors by many or most; outdoors direction estimated. Awakened many or most. Frightened few--slight excitement, a few ran outdoors. Buildings trembled throughout. Broke dishes and glassware to some extent. Cracked windows--in some cases, but not generally. Overturned vases, small or unstable objects, in many instances, with occasional fall. Hanging objects, doors, swing generally or considerably. Knocked pictures against walls, or swung them out of place. Opened, or closed, doors and shutters abruptly. Pendulum clocks stopped, started or ran fast, or slow. Move small

2556 40018

objects, furnishings, the latter to slight extent. Spilled liquids in small amounts from well-filled open containers. Trees and bushes shaken slightly.

VI. Felt by all, indoors and outdoors. Frightened many, excitement general, some alarm, many ran outdoors. Awakened all. Persons made to move unsteadily. Trees and bushes shaken slightly to moderately. Liquid set in strong motion. Small bells rang--church, chapel, school, etc. Damage slight in poorly built buildings. Fall of plaster in small amount. Cracked plaster somewhat, especially fine cracks chimneys in some instances. Broke dishes, glassware, in considerable quantity, also some windows. Fall of knickknacks, books, pictures. Overturned furniture in many instances. Move furnishings of moderately heavy kind.

VII. Frightened all--general alarm, all ran outdoors. Some, or many, found it difficult to stand. Noticed by persons driving motor cars. Trees and bushes shaken moderately to strongly. Waves on ponds, lakes, and running water. Water turbid from mud stirred up. Incaving to some extent of sand or gravel stream banks. Rang large church bells, etc. Suspended objects made to quiver. Damage negligible in buildings of good design and construction, slight to moderate in well-built ordinary buildings, considerable in poorly built or badly designed buildings, adobe houses, old walls (especially where laid up without mortar), spires, etc. Cracked chimneys to considerable extent, walls to some extent. Fall of plaster in considerable to large amount, also some stucco. Broke numerous windows and furniture to some extent. Shook down loosened brickwork and tiles. Broke weak chimneys at the roof-line (sometimes damaging roofs). Fall of cornices from towers and high buildings. Dislodged bricks and stones. Overturned heavy furniture, with damage from breaking. Damage considerable to concrete irrigation ditches.

VIII. Fright general--alarm approaches panic. Disturbed persons driving motor cars. Trees shaken strongly--branches and trunks broken off, especially palm trees. Ejected sand and mud in small amounts. Changes: temporary, permanent; in flow of springs and wells; dry wells renewed flow; in temperature of spring and well waters. Damage slight in structures (brick) built especially to withstand earthquakes. Considerable in ordinary substantial buildings, partial collapse, racked, tumbled down, wooden houses in some cases; threw out panel walls in frame structures, broke off decayed piling. Fall of walls, cracked, broke, solid stone walls seriously. Wet ground to some extent, also ground on steep slopes. Twisting, fall, of chimneys, columns, monuments, also factory stacks, towers. Moved conspicuously, overturned, very heavy furniture.

IX. Panic general. Cracked ground conspicuously. Damage considerable in (masonry) buildings, some collapse in large part; or wholly shifted frame buildings off foundations, racked frames; serious to reservoirs; underground pipes sometimes broken.

X. Cracked ground, especially when loose and wet, up to widths of several inches; fissures up to a yard in width ran parallel to canal and stream banks. Landslides considerable from river banks and steep coasts.

2582 40008

Shifted sand and mud horizontally on beaches and flat land. Changes level of water in wells. Threw water on banks of canals, lakes, rivers, etc. Damage serious to dams, dikes, embankments. Severe to well-built wooden structures and bridges, some destroyed. Developed dangerous cracks in excellent brick walls. Destroyed most masonry and frame structures, also their foundations. Bent railroad rails slightly. Tore apart, or crushed endwise, pipelines buried in earth. Open cracks and broad wavy folds in cement pavements and asphalt road surfaces.

XI. Disturbances in ground many and widespread, varying with ground material. Broad fissures, earth slumps, and land slips in soft, wet ground. Ejected water in large amounts charged with sand and mud. Caused sea-waves ("tidal" waves) of significant magnitude. Damage severe to wood-frame structures, especially near shock centers. Great to dams, dikes, embankments often for long distances. Few, if any (masonry) structures, remained standing. Destroyed large well-built bridges by the wrecking of supporting piers or pillars. Affected yielding wooden bridges less. Bent railroad rails greatly, and thrust them endwise. Put pipelines buried in each completely out of service.

XII. Damage total--practically all works of construction damaged greatly or destroyed. Disturbances in ground great and varied, numerous shearing cracks. Landslides, falls of rock of significant character, slumping of river banks, etc., numerous and extensive. Wrenched loose, tore off, large rock masses. Fault slips in firm rock, with notable horizontal and vertical offset displacements. Water channels, surface and underground, disturbed and modified greatly. Dammed lakes, produced waterfalls, deflected rivers, etc. Waves seen on ground surfaces (actually seen, probably, in some cases). Distorted lines of sight and level. Threw objects upward into the air.

Liquefaction. Temporary transformation of unconsolidated materials into a fluid mass.

Lithosophere. The outer, rigid shell of the earth, situated above the asthenosphere containing the crust, continents, and plates.

Magnitude. A quantity characteristic of the total energy released by an earthquake, as contrasted to intensity that describes its effects at a particular place. Professor C. F. Richter devised the logarithmic scale for local magnitude (M_L) in 1935. Magnitude is expressed in terms of the motion that would be measured by a standard type of seismograph located 100 km from the epicenter of an earthquake. Several other magnitude scales in addition to M_L are in use; for example, body-wave magnitude (m_b) and surface-wave magnitude (M_s), which utilize body waves and surface waves, and local magnitude (M_L). The scale is open ended, but the largest known earthquake have had M_s magnitudes near 8.9.

Mantle. The main bulk of earth between the crust and core, ranging from depths of about 40 to 2900 kilometers.

Mid-oceanridge. Characteristic type of plate boundary occurring in a divergence zone, a site where two plates are being pulled apart and new oceanic lithosphere is being created.

Plate tectonics. The theory and study of plate formation, movement, interaction, and destruction.

Plate. One of the dozen or more segments of the lithosphere that are internally rigid and move independently over the interior, meeting in convergence zones and separating in divergence zones.

Region. A geographical area, surrounding and including the construction site, which is sufficiently large to contain all the geologic features related to the evaluation of earthquake hazards at the site.

Response spectrum. The peak response of a series of simple harmonic oscillators having different natural periods when subjected mathematically to a particular earthquake ground motion. The response spectrum may be plotted as a curve on tripartite logarithmic graph paper showing the variations of the peak spectral acceleration, displacement, and velocity of the oscillators as a function of vibration period and damping.

Return period. For ground shaking, return period denotes the average period of time or recurrence interval between events causing ground shaking that exceeds a particular level at a site; the reciprocal of annual probability of exceedance. A return period of 475 years means that, on the average, a particular level of ground motion will be exceeded once in 475 years.

Risk. See earthquake risk.

Rock. Any solid rock either at the surface or underlying soil having a shear-wave velocity 2,500 ft/sec (765 m/s) at small (0.0001 percent) strains.

Sea-floor spreading. The mechanism by which new sea floor crust is created at ridges in divergence zones and adjacent plates are moved apart to make room.

Seismic Microzoning. The division of a region into geographic areas having a similar relative response to a particular earthquake hazard (for example, ground shaking, surface fault rupture, etc.). Microzoning requires an integrated study of: 1) the frequency of earthquake occurrence in the region, 2) the source parameters and mechanics of faulting for historical and recent earthquakes affecting the region, 3) the filtering characteristics of the crust and mantle constituting the regional paths along which the seismic waves travel, and 4) the filtering characteristics of the near-surface column of rock and soil.

Seismic zone. A generally large area within which seismic design requirements for structures are uniform.

Seismotectonic province. A geographic area characterized by similarity of geological structure and earthquake characteristics. The tectonic processes causing earthquakes have been identified in a seismotectonic province.

Source. The source of energy release causing an earthquake. The source is characterized by one or more variables, for example, magnitude stress drop, seismic moment. Regions can be divided into areas having spatially homogeneous source characteristics.

Strain. A quantity describing the exact deformation of each point in a body. Roughly the change in a dimension or volume divided by the original dimension or volume.

Stress. A quantity describing the forces acting on each part of a body in units of force per unit area.

Strong motion. Ground motion of sufficient amplitude to be of engineering interest in the evaluation of damage due to earthquakes or in earthquake-resistant design of structures.

Subduction zone. A dipping planar zone descending away from a trench and defined by high seismicity, interpreted as the shear zone between a sinking oceanic plate and an overriding plate.

Transform fault. A strike-slip fault connecting the ends of an offset in a mid-ocean ridge. Some pairs of plates slide past each other along transform faults.

Trench. A long and narrow deep trough in the sea floor; interpreted as marking the line along which a plate bends down into a subduction zone.

Triple junction. A point that is common to three plates and which must be the meeting place of three boundary features, such as convergence zones, divergence zones, or transform faults.

APPENDIX C

CONFERENCES TO DATE

Conference I	Abnormal Animal Behavior Prior to Earthquakes, I Not Open-Filed
Conference II	Experimental Studies of Rock Friction with Application to Earthquake Prediction Not Open-Filed
Conference III	Fault Mechanics and Its Relation to Earthquake Prediction Open-File No. 78-380
Conference IV	Use of Volunteers in the Earthquake Hazards Reduction Program Open-File No. 78-336
Conference V	Communicating Earthquake Hazard Reduction Information Open-File No. 78-933
Conference VI	Methodology for Identifying Seismic Gaps and Soon-to-Break Gaps Open-File No. 78-943
Conference VII	Stress and Strain Measurements Related to Earthquake Prediction Open-File No. 79-370
Conference VIII	Analysis of Actual Fault Zones in Bedrock Open-File No. 79-1239
Conference IX	Magnitude of Deviatoric Stresses in the Earth's Crust and Upper Mantle Open-File No. 80-625
Conference X	Earthquake Hazards Along the Wasatch and Sierra-Nevada Frontal Fault Zones Open-File No. 80-801
Conference XI	Abnormal Animal Behavior Prior to Earthquakes, II Open-File No. 80-453
Conference XII	Earthquake Prediction Information Open-File No. 80-843
Conference XIII	Evaluation of Regional Seismic Hazards and Risk Open-File No. 81-437
Conference XIV	Earthquake Hazards of the Puget Sound Region, Washington Open-File No. 83-19
Conference XV	A Workshop on "Preparing for and Responding to a Damaging Earthquake in the Eastern United States" Open-File No. 82-220
Conference XVI	The Dynamic Characteristics of Faulting Inferred from Recording of Strong Ground Motion Open-File No. 82-591
Conference XVII	Hydraulic Fracturing Stress Measurements Open-File No. 82-1075
Conference XVIII	A Workshop on "Continuing Actions to Reduce Losses from Earthquakes in the Mississippi Valley Area Open-File No. 83-157

Open-File Services Section
Branch of Distribution
U.S. Geological Survey
Box 25425, Federal Center
Denver, Colorado 80225

CONFERENCES TO DATE

Conference XIX — Active Tectonic and Magmatic Processes Beneath Long Valley Caldera, Eastern California
Open-File No. 84-939

Conference XX — A Workshop on "The 1886 Charleston, South Carolina, Earthquake and its Implications for Today"
Open-File No. 83-843

Conference XXI — A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in the Northeastern United States"
Open-File No. 83-844

Conference XXII — A Workshop on "Site-Specific Effects of Soil and Rock on Ground Motion and the Implications for Earthquake-Resistant Design"
Open-File No. 83-845

Conference XXIII — A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in Arkansas and Nearby States"
Open-File No. 83-846

Conference XXIV — A Workshop on "Geologic Hazards in Puerto Rico"
Open-File No. 84-761

Conference XXV — A Workshop on "Earthquake Hazards in the Virgin Islands Region"
Open-File No. 84-762

Conference XXVI — A Workshop on "Evaluation of the Regional and Urban Earthquake Hazards in Utah"
Open-File No. 84-763

Conference XXVII — Mechanics of the May 2, 1983 Coalinga Earthquake
Open-File No. 85-44

Conference XXVIII — A Workshop on "The Borah Peak, Idaho, Earthquake"
Open-File No. 85-290

Conference XXIX — A Workshop on "Continuing Actions to Reduce Potential Losses from Future Earthquakes in New York and Nearby States"
Open-File No. 85-386

Conference XXX — A Workshop on "Reducing Potential Losses From Earthquake Hazards in Puerto Rico
Open File No. 85-731

Conference XXXI — A Workshop on "Evaluation of Regional and Urban Earthquake Hazards and Risk in Alaska"
Open File No. 86-79

Conference XXXII — A Conference on "Future Directions in Evaluating Earthquake Hazards of Southern California
Open-File No.

Conference XXXIII — A Workshop on "Earthquake Hazards in the Puget Sound, Washington Area"
Open-File No. 86-253

Conference XXXIV — A Workshop on "Probabilistic Earthquake-Hazards Assessments"
Open-File No 86-185

Open-File Services Section, Branch of Distribution, U.S. Geological Survey, Box 25425, Federal Center, Denver, Colorado 80225