K 3/4

EXPERT OPINION ENCODING IN SEISMIC HAZARD ANALYSIS

C. P. MORTGAT, K. W. CAMPBELL
TERA Corporation, 2150 Shattuck Avenue, Berkeley, California 94704, U.S.A.
D. L. BERNREUTER
Lawrence Livermore Laboratory, University of California,
P.O. Box 808, Livermore, California 94550, U.S.A.

Abstract

This paper presents general results of a survey of ten expert seismologists and geologists conducted as part of a Bayesian seismic hazard analysis of the northeastern United States. The survey, conducted by a written questionnaire, proved very useful in providing the most up-to-date information on Eastern seismotectonics within a limited budget and time frame. The questionnaire was composed of five sections: source zone configuration, maximum earthquakes, earthquake occurrence, attenuation, and overall level of confidence. The respondents were encouraged to answer in terms of probability distributions, but most responded with best estimates and ranges. It was found that a great deal of scatter prevailed among the responses of the experts. This posed an interesting problem with regard to forming a consensus of opinion. An example in terms of earthquake occurrence rates is presented to show the different methods and effects of combining the expert subjective information with historical earthquake occurrence data.

INTRODUCTION

Probability techniques have been used extensively in recent years to assess the risks to life and property associated with the occurrence of earthquakes (Epstein and Lomnitz 1966, Cornell 1968, Donovan 1973, Lomnitz 1974, Kiremidjian and Shah 1975, Algermissen and Perkins 1976). A major shortcoming of these procedures is their complete reliance on the historical record of earthquakes. This record is often plagued by incompleteness and inaccuracies, and is generally too short to serve as an adequate statistical sample for analysis.

Recently, models based on Bayesian probability methods have been introduced in seismic hazard analyses as a way to increase the usable information and reduce the large amount of uncertainty encountered at several steps in the methodology (Cornell 1972, Mortgat 1976, Campbell 1977, Hasselman et al. 1977). One can take full advantage of the capabilities of these models only if consistent and rational ways are developed to extract subjective seismic input from qualified experts and introduce it into the analysis.

This paper summarizes the development and results of a survey to elicit expert opinion on seismicity and intensity attenuation in the northeastern region of the United States. Because it is difficult, or perhaps impossible, to precisely quantify such factors given the sparse historical record, expert judgment was considered crucial. The responses of ten expert seismologists and geologists were used to estimate the seismic hazard at several sites within the region from a probabilistic point of view. The combined opinions were used in a quasi-Bayesian analysis of seismic hazard. The analysis used subjectively-modified historical data as a posterior Bayesian estimate of the earthquake occurrence and attenuation characteristics of the region.

A formal Bayesian analysis would consider, independently, both subjective opinion and historical data. It would then rigorously combine them, each with their corresponding weight, to provide a "posterior" input to be used in the analysis. However, such an analysis implies independence between subjective opinion and data. Due to their inherent knowledge of historical seismicity and attenuation in the East, it was considered unreasonable to expect the experts to divorce themselves from these data while forming an opinion. Therefore, such an opinion is necessarily a posterior estimate and cannot be used in a formal Bayesian analysis without double weighting the data. What was asked, then, was that each expert consider the available seismic data in the eastern United States and temper this by his general experience in the region, possible similarities between the East and other regions, geologic and tectonic considerations, expert judgment and similar types of information. In other words, we asked that each expert be a "Bayesian processor."

ENCODING TECHNIQUES

The available techniques for subjective opinion encoding can be grouped into three categories as a function of the effort and level of sophistication they require.

- A session between the expert whose judgment is being encoded and an analyst who conducts the interview
- A computer-aided interrogation where an interactive computer terminal with graphic display asks qualified questions as a function of the expert's answers. Advantage is taken of the computational capability and the vast memory storage of the computer to provide quick cross-checking between several modes of interrogations, immediate inconsistency resolution through additional questioning and direct input from light pen.
- A written questionnaire where the expert is expected to answer in terms of probabilistic values and/or graphs. Questions are presented in such a manner that they provide sufficient cross-checking. After the answers have been interpreted, a follow-up questionnaire or interview may be needed to resolve inconsistencies.

In view of the level of effort required by each of the methods and the scope of the study, the questionnaire method was used to elicit the opinion of the ten Eastern U.S. seismic experts.

QUESTIONNAIRE

The questionnaire was divided into the following five sections:

- Source Zone Configuration
- Maximum Earthquakes
- Earthquake Occurrence
- Attenuation
- Overall Level of Confidence

In the Source Zone Configuration section we were concerned with the specification of various areas or regions that appear to be unique in their potential to generate earthquakes. In particular, we were seeking in this section the definition of regions within which the experts felt future earthquake activity would be homogeneous. As a point of reference, we provided maps giving two possible seismic zonations of the Eastern United States. We asked the experts to carefully review these figures and to indicate where they thought they might be inadequate by modifying, deleting and adding zones. The experts were asked to indicate their "degree-of-belief" in each source zone and source zone alternative by estimating the chances of the seismicity within these zones being part of the background seismicity of the entire region. We also asked them to identify any localized tectonic structures that might be important to the seismic hazard of nearby sites and to indicate their "degree-of-belief" in their activity.

In the Maximum Earthquake section, we addressed the question of determining the size of the largest event that, in the experts' opinions, could be expected to occur in each of the source zones for a given time period in the future. Since extrapolation of results from short time periods to very long ones is controversial due to possible long-term variations in seismicity and other parameters, we explicitly considered two distinct time periods. The first one was chosen to be 150 years since it was generally on the order of our time period of interest and approximately equivalent to the length of recorded history in the East. The second time period was chosen to be 1,000 years since such a period covers most non-catastrophic perturbations in seismic activity and leaves out the uncertainties associated with the extremely long-term geological variations which were outside the scope of the questionnaire. They were also asked to consider the largest event that they might expect to occur within the current tectonic framework in each source zone without specifying any time period.

The Earthquake Occurrence section considered the occurrence of earthquakes within the next 150 years for each source zone. Occurrences were expressed either in terms of the number of earthquakes expected to occur within that period (for example: 47 in 150 years) or as the mean rate of occurrence per year (i.e., 0.313 per year). The experts were asked to subjectively assess the future seismicity in the East based on the available data and their judgment as to the validity, quality and completeness of these data to represent the true seismicity in the East. To aid in their decision-making, we presented in an accompanying seismicity booklet earthquake occurrence data for the source zones presented in the zonation maps. These data included (1) a listing, in descending order of intensity, of all earthquakes having epicentral intensities IV or greater and (2) a table giving the number of occurrences of earthquakes of each MMI unit from IV through XII. These data were not "corrected" for completeness, but rather represented the latest generally available information on locations and sizes of recorded or felt events.

An attractive approach to supplement the limited strong motion data in the East is to infer, based on theoretical or experimental considerations, the difference in peak acceleration and velocity ground motion between the Eastern United States and the Western United States and to modify correspondingly the Western attenuation relations and intensity-ground motion correlations in order to make them applicable in the East. For this purpose, the section on Attenuation provided general information concerning the validity of existing attenuation relationships and ground-motion correlation for use in the Eastern United States. Attenuation data were not specifically provided for this task; rather, we asked the experts to rely on their inherent knowledge of Eastern U.S. attenuation.

In order to obtain a measure of the overall confidence the experts had in their answers, in the final section they were asked to rate on a scale of 1 to 10 (10 being the highest) the confidence they had in their responses to the different sections of the questionnaire and the various source zones. In this way a consensus or partial consensus could be reached among the experts through weighted average procedures based on self-assigned levels of confidence.

The responses to each question could be made in one of several ways, all of which could be converted to a usable format for analysis. Several formats were provided, and a typical answer sheet is shown in Figure 1. These formats were:

- A best estimate only (fixed quantity)
- A range of values defined by lower and upper bounds and associated with a uniform distribution
- A range of values defined by lower and upper bounds and associated with a non-uniform distribution
- A written discussion

The experts were advised not to spend more than three days on the questionnaire and to concentrate on the areas with which they were most familiar.

RESPONSE

It is impossible to summarize the responses within the space of this paper. Moreover, at the time of this writing, the answers were still being processed for use in the hazard analysis. However, the following qualitative remarks are of interest.

- The questionnaire proved extremely useful in the sense that it provided in a relatively usable format the most up-to-date state of knowledge regarding the seismicity of the Eastern United States.
- Experts were given the choice of answering in terms of magnitude or intensity. This introduced some difficulty when consensus was to be obtained since those units are not closely correlated.
- Consensus was often made difficult by the large scatter encountered in many answers. Figure 2 shows a typical density function describing the size of the largest event to be expected in a given region. In this plot, all the experts have been equally weighted.
- The weights obtained from self-ranking proved to be less useful than expected. The experts tended to rate themselves high for the questions they answered, and not answer the ones for which they rated themselves low.
- All experts provided their answers in terms of either a best estimate only or best estimate together with a range of uncertainty. Only one expert sketched probability density functions.
- The experts answered about 50% of the questions and usually skipped redundant questions, therefore eliminating any cross-checking possibility. When cross-checking was available it usually proved disturbing with answers varying by orders of magnitudes for redundant questions. In such cases the cross checks were disregarded rather than aggregated.
- Questions expressed in terms of probability were usually avoided emphasizing the point that many people still feel uncomfortable when dealing directly with probabilities.

Several weighting schemes are being considered to aggregate the responses. At one extreme the input from each expert may be used in a separate hazard analysis leading to ten hazard curves to which the weights are applied. At the other extreme, all the experts' responses may be aggregated in a simple weighted input to the analysis. A number of intermediate schemes are also being considered. At the time of this writing, no final conclusions can be made regarding the different weighting procedures.

WEIGHT OF SUBJECTIVE INPUT

In order to demonstrate how the subjective input may influence the final results of the analysis, a simplified hypothetical case study is presented.

Consider a site (Figure 3) for which the only seismic source is the host region. The past activity of this source for the last 100 years is plotted in Figure 4. Figure 5 presents the results of the analysis conducted using only the data without introducing any subjective input.

Figure 6 shows three possible responses associated with the yearly rate of occurrence of earthquakes of magnitude equal to or greater than 4.0. Note that the mean value of this gamma type distribution is 1.0 in all three cases, but that the uncertainty varies greatly from curve to curve (σ = 1.0, 0.2 and 0.01, respectively). These plots would be the result of three experts having the same feeling about the mean rate of occurrence of earthquakes but different levels of confidence in their opinion.

Similarly, Figure 7 presents possible subjective input concerning the probability that, given an earthquake has occurred, it is of some fixed magnitude M. Here the distribution is of the beta type bounded between 0 and 1.0 since this is the complete range of values any probability can assume. Again the mean value is the same in all three cases (0.16) but the confidence varies greatly. A similar distribution must be obtained for each magnitude under consideration.

To show the influence of uncertainty in the subjective input, Figure 8 presents the results of the analysis using only this input (i.e., excluding the data) for the lowest and highest levels of confidence. As expected, the variation is noticeable, but not dramatic, since the mean value remains the same. Four additional examples are given to show the weight that the subjective input has when used with the data (Figure 5):

- Low-confidence subjective input confirming the data
- High-confidence subjective input confirming the data
- Low-confidence subjective input noticeably different from the data
- High-confidence subjective input noticeably different from the data

Figure 5 emphasized the point that only high-confidence subjective input different from the data introduces variations in the results and that low-confidence subjective input is relatively irrelevant in the analysis.

CONCLUSION

The use of a written questionnaire proved to be very useful in the sense that within a reasonable time and budget frame it provided the most up-to-date state of knowledge regarding the seismicity and seismotectonics of the Eastern United States. The problems arising from obvious inconsistencies within an expert's responses and broad scatter among experts are being worked on. The use of a follow-up questionnaire to resolve some of these inconsistencies is being considered.

REFERENCES

1. Algermissen, S. T. and D. M. Perkins (1976). "A Probabilistic Estimate of Maximum Acceleration in Rock in the Contiguous United States," U.S. Geologic Survey, Open File Report 76-416.

2. Campbell, K. W. (1977). "The Use of Seismotectonics in the Bayesian Estimation of Seismic Risk," School of Engineering and Applied Science, University of California, Los Angeles, UCLA-ENG-7744.

3. Cornell, C. A. (1972). "Bayesian Statistical Decision Theory and Reliability - Based Design," Proc. Int'l. Conf. on Structural Safety and Reliability, edited by A. M. Freudenthal, Smithsonian Institute, pp. 47-66.

4. Cornell, C. A. (1968). "Engineering Seismic Risk Analysis," Bull. Seism. Soc. Am., 58:1583-1606.

5. Donovan, N. C. (1973). "Earthquake Hazards for Buildings," Building Practices for Disaster Mitigation, edited by R. Wright et al., National Bureau of Standards, Building Science Series 46, U.S. Government Printing Office, Washington, D.C., pp. 82-111.

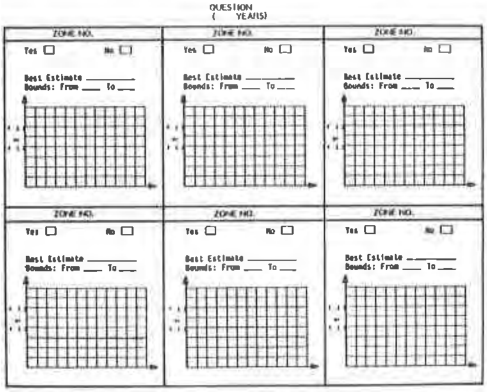
6. Epstein, B. and C. Lomnitz (1966). "A Model for the Occurrence of Large Earthquakes," Nature, 211:954-956.

7. Hasselman, T. K., J. D. Chrostowski, and J. H. Wiggins, "A Bayesian Approach for Mapping Earthquake Intensity," presented at ASCE National Convention, Oct. 17-21, 1977, San Francisco.

8. Kiremidjiian, A. S., and H. C. Shah (1975). "Seismic Hazard Mapping of California," J. A. Blume Earthquake Engineering Center, Stanford University, Report 21.

9. Lomnitz, C. (1974). Global Tectonics and Earthquake Risk, Elsevier Scientific Publishing Company, New York.

10. Mortgat, C. P. (1976). "A Bayesian Approach to Seismic Hazard Mapping; Development of Stable Design Parameters," Ph.D. Dissertation in Engineering, Dept. of Civil Engineering, Stanford University.



QUESTION
(YEARS)

ZONE NO.	ZONE NO.	ZONE NO.
Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___	Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___	Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___
ZONE NO.	ZONE NO.	ZONE NO.
Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___	Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___	Yes ☐ No ☐ Best Estimate ________ Bounds: From ___ To ___

FIGURE I

TYPICAL ANSWER SHEET

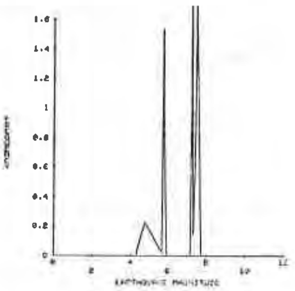


FIGURE 2

DISTRIBUTION OF UPPER MAGNITUDE CUTOFF FOR A REGION

FIGURE 3

HYPOTHETICAL SITE AND HOST REGION

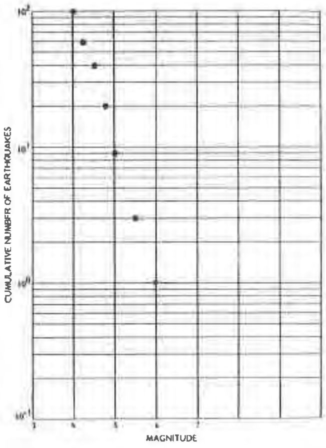


FIGURE 4

CUMULATIVE NUMBER OF EARTHQUAKES (100 YEARS)

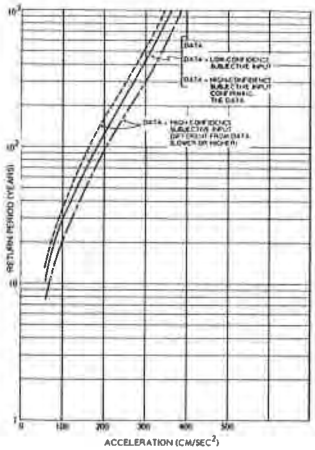


FIGURE 5

HAZARD AT THE SITE DUE TO
DATA AND DIFFERENT SUBJECTIVE INPUTS

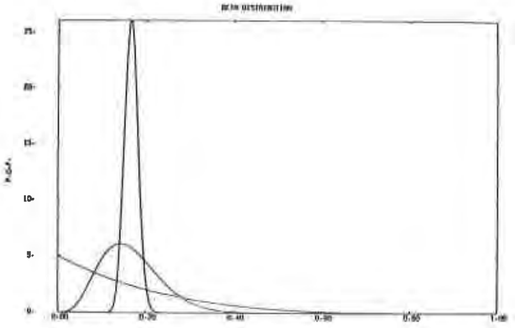


FIGURE 6

DIFFERENT LEVELS OF CONFIDENCE IN SUBJECTIVE INPUT
RELATED TO PROBABILITY THAT ANY EARTHQUAKE
IS OF A GIVEN MAGNITUDE M_i

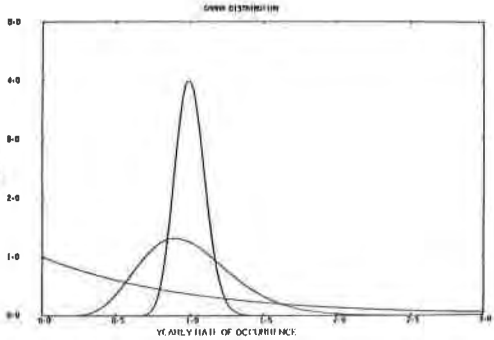


FIGURE 7

HAZARD AT THE SITE DUE TO
SUBJECTIVE INPUT CONFIRMING THE DATA

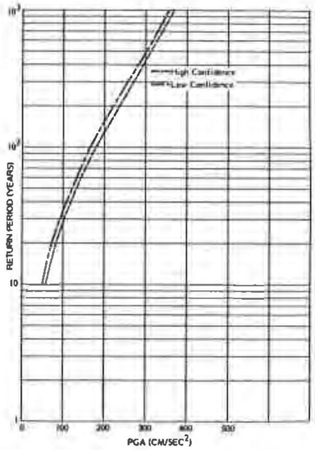


FIGURE 8

DIFFERENT LEVELS OF CONFIDENCE IN SUBJECTIVE
INPUT RELATED TO YEARLY RATE OF
EARTHQUAKE OCCURRENCE