Bulletin of the Seismological Society of America, Vol. 82, No. 1, pp. 120–134, February 1992

# ESTIMATION OF EARTHQUAKE HAZARD PARAMETERS FROM INCOMPLETE DATA FILES. PART II. INCORPORATION OF MAGNITUDE HETEROGENEITY

By A. Kijko and M. A. Sellevoll

## Abstract

The maximum likelihood estimation of earthquake hazard parameters (maximum regional magnitude $m_{max}$, activity rate $\lambda$, and the Gutenberg–Richter parameter $b$) from incomplete data files is extended to the case of uncertain magnitude values. Two models of uncertainty are considered. In the first one, earthquake magnitude is specified by two values, the lower and the upper magnitude limit. It is assumed that such an interval contains the real, unknown magnitude. In the second model, uncertainty of earthquake magnitude is defined in the same way as it was proposed by Tinti and Mulargia (1985): the departure of the observed (apparent) magnitude from the true, unknown value is distributed normally. The proposed approach allows the combination of catalog parts of different quality, e.g., those where the assessment of magnitude is questionable and those with magnitudes determined very precisely.

As an illustration, the proposed procedures are applied for the estimation of seismicity parameters in western Norway with adjacent sea areas.

## Introduction

In the first part of our study (Kijko and Sellevoll, 1989: henceforth referred to as KS1), the maximum likelihood estimation of basic earthquake hazard parameters (maximum regional magnitude $m_{max}$, earthquake activity rate $\lambda$, and the $b$ parameter in the Gutenberg–Richter relation) was proposed. The issue addressed in KS1 is how to utilize large historical events and recent complete observations. In addition, the KS1 technique permits several thresholds of completeness as well as gaps in registrations.

However, despite its flexibility, the KS1 approach has an important deficiency: it is not able handle magnitude uncertainties. Earthquake magnitudes are never known exactly. The older (macroseismic) earthquake data recovered from historical records are affected by large uncertainties, due in part to (e.g., Ambraseys *et al.*, 1983; Bender, 1987; Tinti *et al.*, 1987) shortage of documentation, inaccuracy and misunderstanding in the description of the damages, and conversion of macroseismic information to the corresponding magnitude value.

Even instrumentally determined earthquake magnitudes can be very uncertain. Conversion of one type of magnitudes to the single measure common to the whole span of the catalog requires conversion by means of empirical relations. As was pointed out by Chung and Bernreuter (1981), such a procedure is not necessarily valid. In addition, change of characteristics of seismic sensors can cause systematic error in magnitude conversion (see, e.g., the case of magnitude conversion for eastern and western United States, Chung and Bernreuter, 1981; Nuttli and Herrmann, 1982). Correspondingly, a catalog that contains macroseismic and complete data sets is heterogeneous in respect to magnitude determination and requires appropriate handling techniques.

In this paper, we shall present two original approaches to the problem of seismic hazard evaluation and see that incorporation of earthquake magnitude uncertainty entails reconsideration of the estimate technique proposed in KS1.

## Two Models of Magnitude Uncertainty

*Hard Bounds Model.* Uncertainty of earthquake magnitude is specified by two values: $\underline{x}$ and $\bar{x}$. $\underline{x}$ is the lower and $\bar{x}$ is the upper magnitude limit. Introducing an apparent magnitude value equal to $x = 0.5(\underline{x} + \bar{x})$, the lower and the upper magnitude limits are equal to $\underline{x} = x - \delta$ and $\bar{x} = x + \delta$, where $\delta$ is the measure of the magnitude uncertainty equal to $\delta = 0.5(\bar{x} - \underline{x})$ (Fig. 1a).

*Soft Bounds Model.* The second model is based on the concept of apparent magnitude, introduced by Tinti and Mulargia (1985). The apparent magnitude

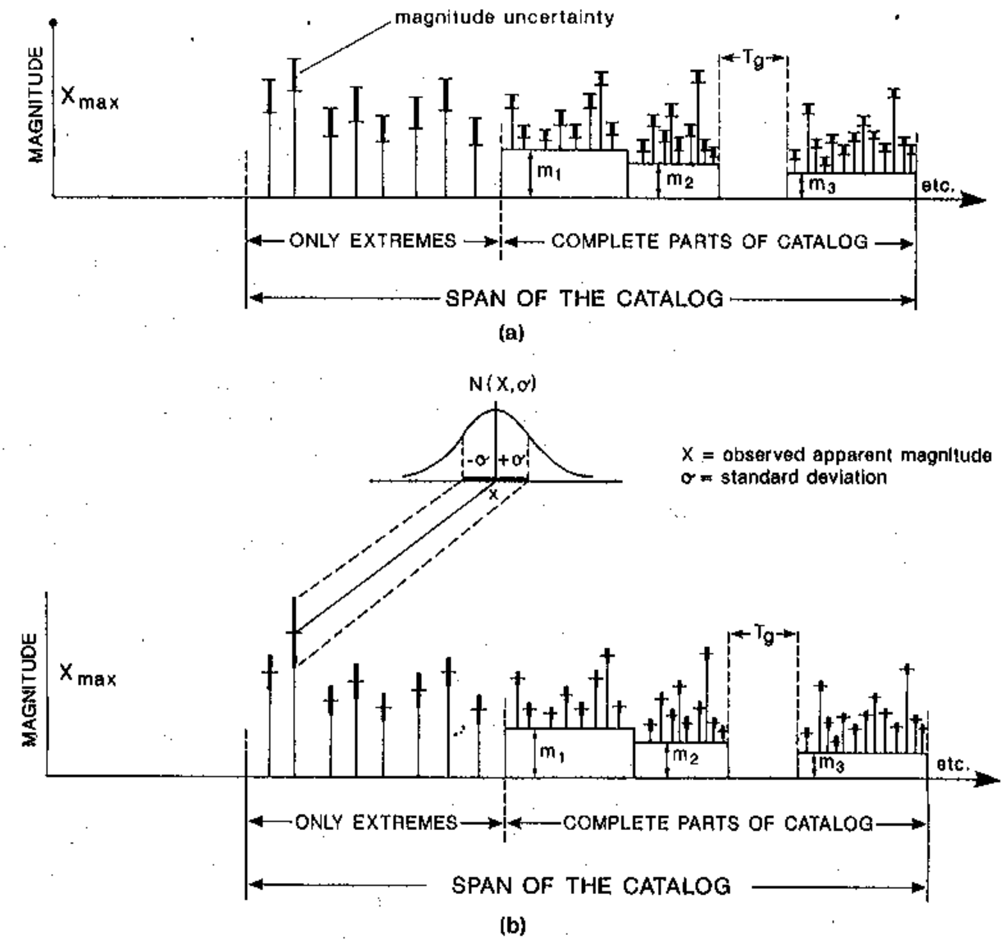


Fig. 1. An illustration of data that can be used to obtain basic seismic hazard parameters by the proposed procedures. Our approach permits the combination of the largest earthquakes with complete data and variable threshold magnitudes. It makes possible to use the largest known historical earthquake ($X_{max}$) that occurred before our catalog begins. It also accepts "gaps" ($T_g$) when records are missing or seismic networks were not in operation. (a) "Hard bounds" model of earthquake magnitude uncertainty. Magnitude of each earthquake is specified by two values: the lower and the upper magnitude limit. It is assumed that such an interval contains the real unknown magnitude. (b) "Soft bounds" model of earthquake magnitude uncertainty. Following Tinti and Mulargia (1985), it is assumed that the observed magnitude is the true magnitude distorted by a random error $\epsilon$. $\epsilon$ is free from systematic errors and follows a Gaussian distribution with zero mean and standard deviation $\sigma$.

of an earthquake is defined as the "observed" magnitude, which differs from the "real" magnitude owing to the random error $\epsilon$. It is assumed that $\epsilon$ follows a Gaussian distribution with zero mean and standard deviation $\sigma$ (Fig. 1b).

The names of our models are given after Backus (1988), who introduced "hard" and "soft" bounds of prior information in inversion of geophysical problems. The choice of the model to work with depends on our knowledge of data collection procedure and catalog preparation. It is clear that such a decision contains a certain amount of subjective judgment.

Assuming the Poisson occurrence of earthquakes with activity rate $\lambda$ and validity of the doubly truncated Gutenberg–Richter magnitude-frequency relation, the density and cumulative magnitude distributions can be written respectively as (e.g., Page, 1968; Cosentino *et al.*, 1977)

$$f(x \mid m) = \beta A(x)/(A_1 - A_2), \tag{1}$$

$$F(x \mid m) = [A_1 - A(x)]/(A_1 - A_2), \tag{2}$$

where $A_1 = \exp(-\beta m)$, $A_2 = \exp(-\beta m_{max})$, $A(x) = \exp(-\beta x)$, and magnitude $x$ belongs to the domain $\langle m, m_{max}\rangle$. $m$ is the threshold magnitude. $\beta$ is related to Gutenberg–Richter parameter $b$ through the relation $\beta = b \ln (10)$. The desired seismicity parameters are $\theta = (\beta, \lambda)$ and $m_{max}$.

The probability that in a time interval $t$ either no earthquake occurs or all occurring earthquakes have apparent magnitude not exceeding $x$ may be expressed as $\exp\{-\lambda(m_0)t[1 - F(x \mid m_0)]\}$ (e.g., Benjamin and Cornell, 1970; Gan and Tung, 1983), where $\lambda(m_0) = [\lambda[1 - F(m_0 \mid m_{min})]$ and $m_0$ is the threshold magnitude for the extreme part of the catalog ($m_0 \geqq m_{min}$). $m_{min}$ plays the role of the "total" threshold magnitude and has rather formal character. The only condition in the choice of its value is that $m_{min}$ cannot exceed the threshold magnitude of any part of the catalog, extreme as well as complete. Hence, the probability distribution function of the strongest earthquake during the time interval $t$, conditional on the earthquake existence, is given by

$$G(x \mid m_0, t) = \frac{\exp\{-\lambda(m_0)t[1 - F(x \mid m_0)]\} - \exp[-\lambda(m_0)t]}{1 - \exp[-\lambda(m_0)t]}. \tag{3}$$

In most practical situations, we deal with enough high activity rate $\lambda(m_0)$ that the term $\exp[-\lambda(m_0)t]$ can therefore be ignored.

Let us discuss the first model of magnitude uncertainty and build the likelihood function of desired seismicity parameters $\theta$. If the uncertainty of earthquake magnitude is specified by the lower and upper magnitude limits $\underline{x}$, $\bar{x}$, the density probability function of the apparent magnitude becomes the convolution of magnitude distribution (1) and uniform distribution in the range $\langle -\delta, \delta\rangle$, where $\delta$ denotes the interval of magnitude uncertainty. After simple calculations, the density probability function of the *apparent* magnitude for the discussed uncertainty model is

$$f(x \mid m, \delta) = (2\delta)^{-1}\begin{cases} F(x+\delta), & m - \delta \leqq x < m + \delta, \\ F(x+\delta) - F(x-\delta), & m + \delta \leqq x < m_{max} - \delta, \\ 1 - F(x-\delta), & m_{max} - \delta < x \leqq m_{max} + \delta, \end{cases} \tag{4}$$