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 λ, 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 λ, 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 δ 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

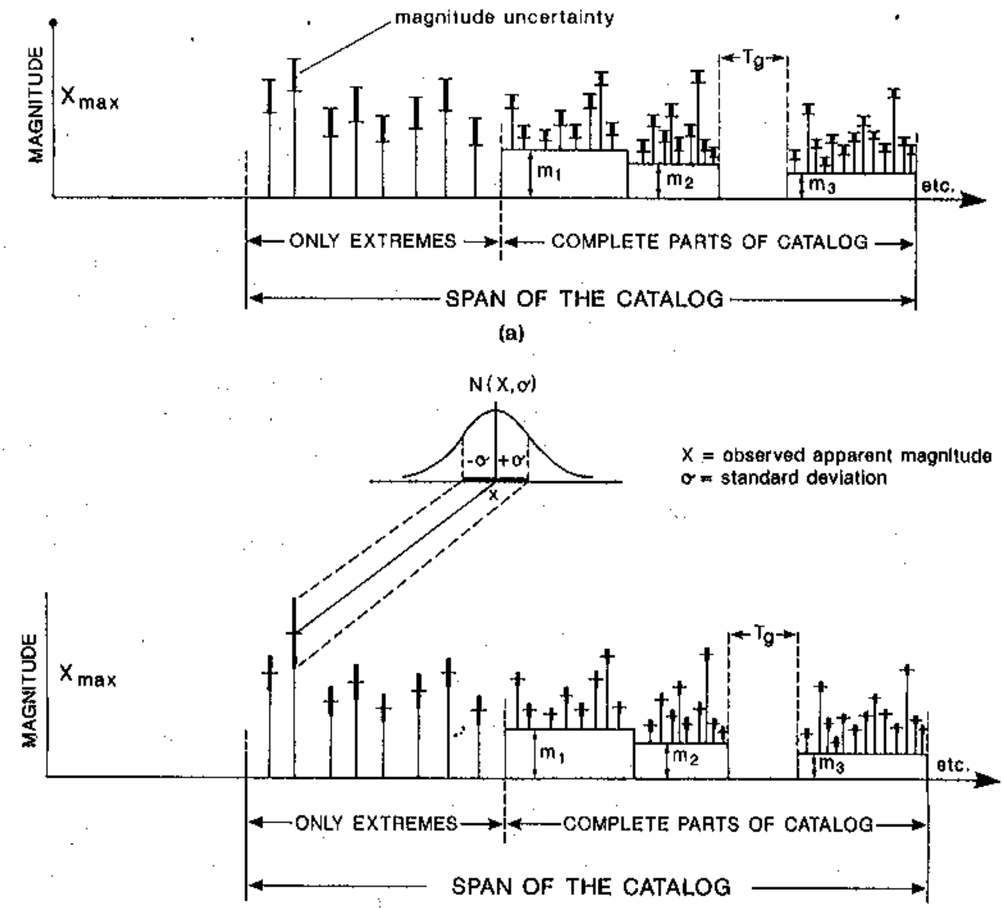

Fig. 1. An illustration of data that can be used to obtain basic seismic hazard parameters by the proposed procedures. Our approach permits the combination of the largest earthquakes with complete data and variable threshold magnitudes. It makes possible to use the largest known historical earthquake (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 ϵ. ϵ is free from systematic errors and follows a Gaussian distribution with zero mean and standard deviation σ.

of an earthquake is defined as the "observed" magnitude, which differs from the "real" magnitude owing to the random error ϵ. It is assumed that ϵ follows a Gaussian distribution with zero mean and standard deviation σ (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 λ 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. β 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 θ. 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 δ denotes the interval of magnitude uncertainty. After simple calculations, the density probability function of the *apparent* magnitude for the discussed uncertainty model is

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

or equivalently

$$f(x \mid m, \delta) = C_\delta(x \mid m, \delta)\frac{\beta A(x)}{A_1 - A_2}, \tag{5}$$

where the correction function $C_\delta(x \mid m, \delta)$ is given by

$$\begin{cases} \{\exp[\beta(x-m)] - \exp(-\beta\delta)\}/2\beta\delta, & \text{for } m-\delta \leqq x < m+\delta, \\ c_f, & \text{for } m+\delta \leqq x < m_{max}-\delta, \\ \{\exp(\beta\delta) - \exp[-\beta(m_{max}-x)]\}/2\beta\delta, & \text{for } m_{max}-\delta < x \leqq m_{max}+\delta \end{cases}, \tag{6}$$

where $c_f = [\exp(\beta\delta) - \exp(-\beta\delta)]/2\beta\delta$.

It is clear that apparent distribution (5) progressively deviates from the classical Gutenberg–Richter one as magnitude uncertainty δ increases. In the larger part of the magnitude domain $\langle m+\delta, m_{max}-\delta\rangle$, the apparent magnitude distribution is proportional to "true" distribution (1). Since for any positive magnitude uncertainty δ, the correction factor $c_f > 1$, the apparent magnitude distribution (5) within the interval $\langle m+\delta, m_{max}-\delta\rangle$ overestimates the number of earthquakes.

The further application of apparent magnitude (4) requires its renormalization. In our original model (1), there is a sudden transition between the magnitude range where we are capable of recording all earthquakes (for $x \geqq m$) and the range (for $x < m$) where no earthquake can be recorded. Such an assumption seems to be unrealistic, since in practice the transition occurs gradually. Let us assume therefore that the cutoff magnitude m is chosen in such a way that all earthquakes with "true" magnitude in the range $\langle m-\delta, m\rangle$ and with apparent magnitude not smaller than m are recorded. Then, the normalized density and cumulative probability functions of the apparent magnitude for hard bound model become, respectively,

$$\tilde{f}(x \mid m, \delta) = (c_f A_1 - A_2)^{-1} \times \begin{cases} c_f \beta A(x), & m \leqq x < m_{max}-\delta \\ [A(x-\delta) - A_2]/(2\delta), & m_{max}-\delta \leqq x \leqq m_{max}+\delta \end{cases} \tag{7}$$

and

$$\tilde{F}(x \mid m, \delta) = (c_f A_1 - A_2)^{-1} \times \begin{cases} c_f[A_1 - A(x)], & m \leqq x < m_{max}-\delta, \\ [A_1 - A(m_{max}-\delta)] - A_2(x - m_{max}+\delta)/(2\delta) & \\ \quad -[A(x) - A(m_{max}-\delta)]\exp(\beta\delta)/(\beta\delta), & m_{max}-\delta \leqq x \leqq m_{max}+\delta. \end{cases} \tag{8}$$

The application of standard maximum likelihood technique at such a stage leads to correct evaluation of the parameter β only. The estimated activity rate λ will be still "apparent." In order to obtain the "true" λ evaluation, respective corrections must be introduced.

The assumption on the restricted range of earthquake magnitude to the interval $\langle m, m_{max}\rangle$ and the application of the technique proposed by Bender (1987), Appendix B) leads to the following relation between the apparent activity rate $\tilde{\lambda}(x)$ and the "true" rate $\lambda(x)$.

$$\tilde{\lambda}(x) = \lambda(x) C_\delta(x \mid m, \delta), \tag{9}$$

where $C_\delta(x \mid m, \delta)$ is defined by relation (6) and $m \leqq x \leqq m_{max}$. Taking into account our assumption regarding the nature of lower end of the magnitude range m, relation (9) takes the following form:

$$\tilde{\lambda}(x) = \lambda(x) \begin{cases} c_f, & m \leqq x < m_{max} - \delta \\ \dfrac{\exp(\beta\delta) - \exp[-\beta(m_{max} - x)]}{2\beta\delta}, & m_{max} - \delta < x \leqq m_{max}. \end{cases} \tag{10}$$

From relation (3) it follows that $g(x \mid m_0, t, \delta)$, the density probability function of the strongest earthquake within a period t with the apparent magnitude x and uncertainty δ, is

$$\tilde{\lambda}(m_0) t \tilde{f}(x \mid m_0, \delta) \exp[-\tilde{\lambda}(m_0) t [1 - \tilde{F}(x \mid m_0, \delta)] / [1 - \exp(-\tilde{\lambda}(m_0) t]. \tag{11}$$

After introducing probability (11), the likelihood function of θ, following from the extreme part of the catalog, becomes

$$L_0(\theta \mid x_0) = \text{const} \prod_{i=1}^{n_0} g(x_{0i} \mid m_0, t_{0i}, \delta_{0i}). \tag{12}$$

In relation (12) for each earthquake i, from the first (extreme) part of the catalog, the input data are the apparent magnitude x_{0i} of the strongest earthquake occurring during the time interval t_i and the value of its magnitude uncertainty δ. $i = 1, \ldots, n_0$, and n_0 is the number of earthquakes in the extreme part of the catalog. The time intervals t_i are calculated according to the simple formula:

$$t_i = \begin{cases} \tau_1 - t_{01}, & \text{for } i = 1 \\ \tau_i - \tau_{i-1}, & \text{for } i = 2, \ldots, n_0 - 1 \\ t_{02} - \tau_{n0-1}, & \text{for } i = n_0, \end{cases} \tag{13}$$

where t_{01} and t_{02} mean the beginning and the end of the extreme part of the catalog, and $\tau_1, \ldots, \tau_{n0}$ are the extreme earthquake origin times. For compactness and convenience of notation, magnitudes and its uncertainties are grouped into $\mathbf{x}_0 = | x_{0i}, \delta_{0i} |$, $i = 1, \ldots, n_0$. From the same reason, time intervals t_i are grouped into $\mathbf{t} = (t_1, \ldots, t_{n0})$. Const is a normalization factor independent of θ.

Let us assume that the second, complete part of the catalog can be divided into s subcatalogs (Fig. 1). Each of them has its span T_i and is complete starting from the known magnitude m_i. Let for each subcatalog i, $\mathbf{x}_i = | x_{ij}, \delta_{ij} |$ be apparent magnitude and its uncertainty with $j = 1, \ldots, n_i$, where n_i denotes the number of earthquakes in each complete subcatalog, and $i = 1, \ldots, s$.

If the size of seismic events is independent of their number, the likelihood function of θ, $L_i(\theta \mid \mathbf{x}_i)$, is the product of two functions $L_\beta(\beta \mid \mathbf{x}_i)$ and $L_\lambda(\lambda \mid \mathbf{x}_i)$:

Relation (7) generates $L_\beta(\beta\,|\,\mathbf{x}_i)$ of the following form

$$L_\beta(\beta\,|\,\mathbf{x}_i) = \text{const}\cdot\prod_{j=1}^{n_i}\tilde{f}(x_{ij}\,|\,m_i,\delta_{ij}). \tag{14}$$

The assumption that the number of earthquakes per unit time is a Poisson random variable gives the similar form of $L_\lambda(\lambda\,|\,\mathbf{x}_i)$ as in KS1: const $\exp[-\tilde{\lambda}(m_i)t_i][\lambda(m_i)t_i]^{n_i}$, where const is a normalizing factor, the apparent activity rate is defined by relation (10), and $\lambda(m_i) = \lambda[1 - F(m_i\,|\,m_{min})]$. Relations (2), (7) and (8) and $L_\lambda(\lambda\,|\,\mathbf{x}_i)$, for $i = 1,\ldots s$, define the likelihood function of the parameters sought for each complete subcatalog. Finally, the joint likelihood function based on all data is given by

$$L(\theta\,|\,\mathbf{x}) = \prod_{i=0} L_i(\theta\,|\,\mathbf{x}_i). \tag{15}$$

A certain approximation of likelihood function (15) for the hard bounds model, including analytical forms of the derivatives $\partial L(\theta\,|\,\mathbf{x})/\partial\beta$, $\partial L(\theta\,|\,\mathbf{x})/\partial\lambda$, analysis of main implications and discussion of special cases can be found in our recent paper (Kijko and Sellevoll, 1990).

Let us now turn our attention to the soft bounds model of magnitude uncertainty, introduced by Tinti and Mulargia (1985). If the error of magnitude determination is assumed to be normally distributed with a standard deviation σ, the density and cumulative probability functions of the apparent magnitude become respectively:

$$f(x\,|\,m,\sigma) = \beta A(x)/(A_1 - A_2)C_\sigma(x\,|\,m,\sigma), \tag{16}$$

$$F(x\,|\,m,\sigma) = [A_1 - A(x)]/(A_1 - A_2)D_\sigma(x\,|\,m,\sigma), \tag{17}$$

where

$$C_\sigma(x\,|\,m,\sigma) = \frac{e^{\gamma^2}}{2}\left[\text{erf}\left(\frac{m_{max} - x}{\sqrt{2}\,\sigma} + \gamma\right) + \text{erf}\left(\frac{x - m}{\sqrt{2}\,\sigma} - \gamma\right)\right],$$

$$D_\sigma(x\,|\,m,\sigma) = \left\{A_1\left[\text{erf}\left(\frac{x - m}{\sqrt{2}\,\sigma}\right) + 1\right] + A_2\left[\text{erf}\left(\frac{m_{max} - x}{\sqrt{2}\,\sigma}\right) - 1\right]\right.$$
$$\left. - 2C_\sigma(x\,|\,m,\sigma)A(x)\right\}/2[A_1 - A(x)],$$

erf(·) is the error function (Abramowitz and Stegun, 1970), $\gamma = \beta\sigma/\sqrt{2}$, and x is unlimited from both ends.

It should be noticed that for the magnitude interval $\langle m, m_{max}\rangle$, *apparent* magnitude distributions (16) and (17) can be expressed by *true* magnitude distributions (1) and (2) in simple forms:

$$f(x \mid m, \sigma) = f(x \mid m) C_\sigma(x \mid m, \sigma), \tag{18}$$

$$F(x \mid m, \sigma) = F(x \mid m) D_\sigma(x \mid m, \sigma). \tag{19}$$

It may be verified that, for x inside the interval $\langle m, m_{max} \rangle$, the correction function $C_\sigma(x \mid m, \sigma)$ may be well approximated by a constant equal to $\exp(\gamma^2)$, and

$$\begin{cases} \lim_{\sigma \to} C_\sigma(x \mid m, \sigma) = 1, \\ \lim_{\sigma \to} D_\sigma(x \mid m, \sigma) = 1. \end{cases} \tag{20}$$

Relations (20) are in full agreement with our intuitive expectations: the less the random errors perturb the real magnitude, the more the apparent magnitude distributions $f(x \mid m, \sigma)$ and $F(x \mid m, \sigma)$ appear to be correspond to $f(x \mid m)$ and $F(x \mid m)$. It is interesting to notice that the apparent magnitude distributions (16) and (17) may assume values even outside of the original domain of $\langle m, m_{max} \rangle$. From a formal point of view, apparent magnitudes range between $\pm\infty$.

The further application of distributions (16) and (17) requires additional renormalizations. If m is the lowest magnitude at and above which the observations are complete, then its normalized density probability function $\tilde{f}(x \mid m, \sigma)$ is zero up to m and is equal to $f(x \mid m, \sigma)/[1 - F(m \mid m, \sigma)]$ for $x \geqq m$. In a similar way, the normalized cumulative probability function of apparent magnitude is $\tilde{F}(x \mid m, \sigma) = [F(x \mid m, \sigma) - F(m \mid m, \sigma)]/[1 - F(m \mid m, \sigma)]$. In fact, $\tilde{f}(x \mid m, \sigma)$ and $\tilde{F}(x \mid m, \sigma)$ are conditional distributions of x given that $x \geqq m$.

Finally, assuming that the model in which density function (1) vanishes below the cutoff magnitude m is unrealistic, and in practice the transition occurs gradually, the relation between the apparent activity rate $\tilde{\lambda}(x)$ and the "true" one takes the form

$$\tilde{\lambda}(x) = \lambda(x) \frac{e^{\gamma^2}}{2} \left[1 + \operatorname{erf}\left(\frac{m_{max} - x}{\sqrt{2}\,\sigma} + \gamma \right) \right]. \tag{21}$$

The likelihood function of the parameter θ is designed in the similar way as for the hard bounds model. For an extreme part of the catalog, for each earthquake i, two input values are required: the apparent magnitude x_{0i} of the strongest earthquake occurring during the time interval t_i and the value of its standard deviation σ_{0i} ($i = 1, \ldots, n_0$). For compactness of notation, earthquake magnitudes and their standard deviations are denoted as $\mathbf{x}_0$, ($\mathbf{x}_0 = | x_{0i}, \sigma_{0i} |$, $i = 1, \ldots, n_0$). For the same reason, time intervals t_i are grouped into $\mathbf{t} = (t_1, \ldots, t_{n0})$. Similarly, let for each complete part of the catalog $\mathbf{x}_i = | x_{ij}, \sigma_{ij} |$, ($j = 1, \ldots, n_i$) to denote the apparent magnitude values and their standard deviations.

In order to estimate the parameters $\theta = (\beta, \lambda)$, the maximum likelihood procedure is used. The maximum likelihood estimate $\hat{\theta}$ is the value of θ that maximizes likelihood function (8). Our likelihood function does not provide a satisfactory evaluation of m_{max}. Following Kijko and Dessokey (1987) and the

KS1 approach, the condition that the largest *observed* magnitude X_{max} is equal to $EXPECT\{x_{max}|T\}$, i.e., the largest *expected* magnitude in the span of the catalog T, often provides satisfactory evaluation of m_{max}. The formula for $EXPECT\{x_{max}|T\}$ is given by Kijko (1988) and KS1.

An Example: Seismic Hazard Parameters for the Western Norway Coastal Area

As an illustration, the described estimation procedures have been used to estimate seismicity parameters for western Norway and the adjacent sea area (Fig. 2), limited by 58° to 64°N, 4° to 8°E. The data used originate from Båth (1956), Muir Wood *et al.*, (1989), and Sellevoll *et al.*, (1982) with an appendix of December 1989. The catalog compiled has been divided into four parts. The first part contains six of the largest earthquakes for 1 January 1831 to 31 December 1890 studied by Muir Wood *et al.* (1989). The second part (from 1 January 1891 to 31 December 1950) includes data from Båth's (1956) catalog for Fennoscandian earthquakes for the years 1891 to 1950. The third part (from 1 January 1951 to 31 December 1979) and fourth part (from 1 January 1980 to 31 December 1989) include data presented by Sellevoll *et al.* (1982) with appendix of December 1989.

Before 1891, the macroseismic data are generally much less complete than during the period from 1891 to 1950, which is a time period characterized by fairly complete macroseismic observations, but with instrumental data available in a few cases. The systematic collection of macroseismic data by help of questionnaires began in Norway in 1887 and has continued without break since then. The establishment of a seismic network for Scandinavia started about 1950. Engineering work in the North Sea for oil production required much more detailed information than available and local seismic network were established in western Norway during the 1980s in order to meet these requirements.

The macroseismic magnitudes for the period from 1951 to 1989 have been determined in such a way as to obtain best agreement with Båth's catalog, for the period from 1891 to 1950. The M_s magnitudes calculated by Muir Wood *et al.* (1989) have been converted to our macroseismic magnitudes by establishing a relationship between their and our catalog from 1891 to 1987 (51 pairs of magnitudes). Table 1 shows the first part of the catalog. The second part contains the complete catalog of $n_1 = 40$ earthquakes with threshold magnitude $m_1 = 3.8$. The third part contains $n_2 = 37$ events with threshold magnitude $m_2 = 3.6$. The last part contains $n_3 = 27$ earthquakes completed from $m_3 = 3.0$. Based on experience of our earthquake catalog preparation, it was assumed that for each data set hard bounds of magnitude uncertainty are equal to $\delta_0 = 0.3$, $\delta_1 = 0.25$, $\delta_2 = 0.2$, and $\delta_3 = 0.15$.

We did not investigate theoretically the relation between the hard bounds model with given parameters and the parameters of a corresponding soft bounds model. It is clear that σ for the soft bounds model depends on the subjective choice of confidence level. Fortunately, as we shall see, estimated hazard parameters are not very sensitive to changes in the magnitude uncertainty characteristics δ and σ as long they remain in the reasonable ranges. Let us illustrate this with two examples.

In the first example, let us assume that we underestimate errors in magnitude determination and for some cases the "true" magnitude X lies outside the specified interval $\langle \underline{x}, \bar{x} \rangle$. If no additional information is available we could

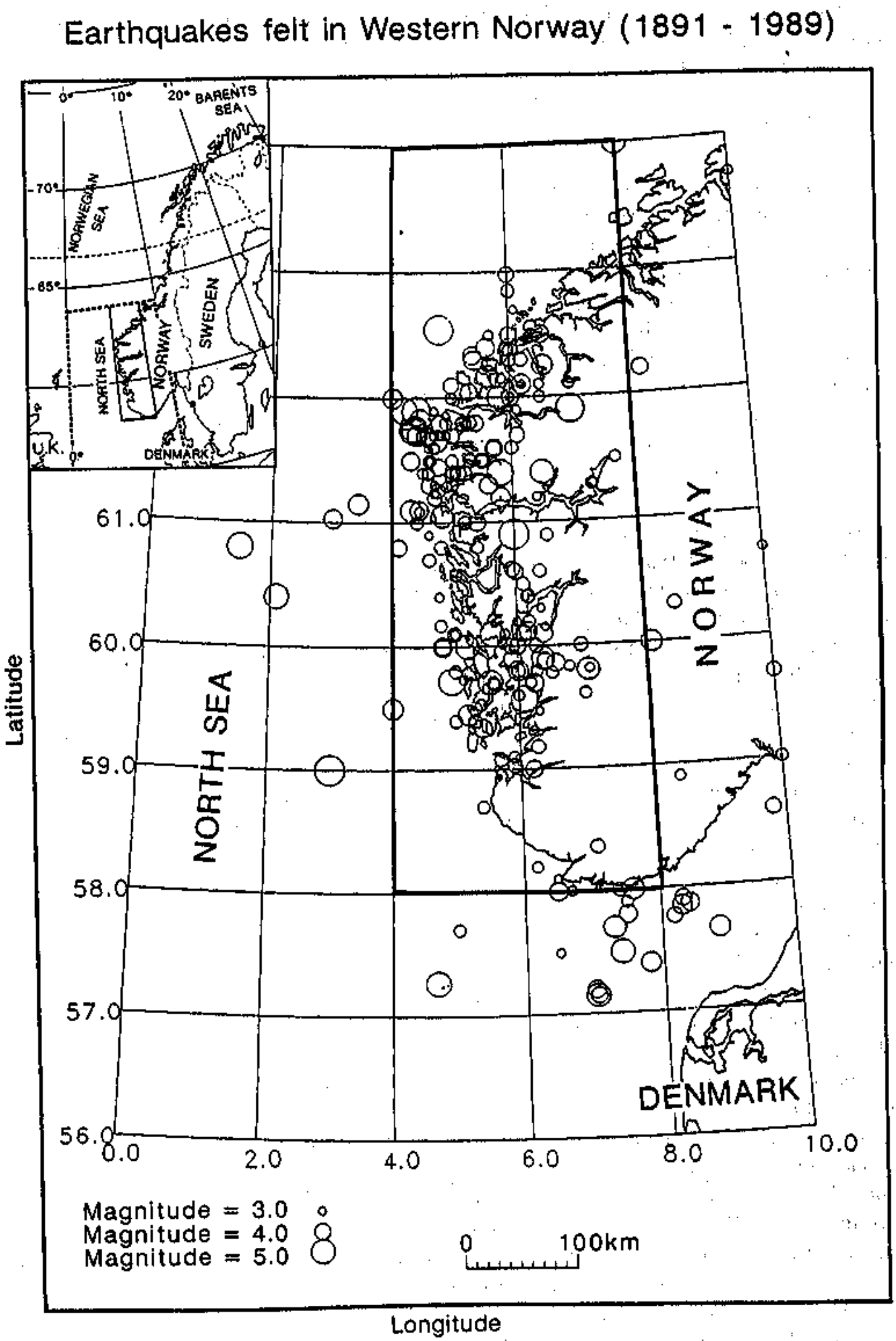

FIG. 2. Seismicity map of western Norway and adjacent area for the period from 1891 to 1989.

TABLE 1

MAXIMUM MACROSEISMIC MAGNITUDES OF EARTHQUAKES FELT IN WESTERN NORWAY AND ADJACENT AREA FOR THE PERIOD FROM 01/01/1831 TO 12/31/1890*

Event	Date	Magnitude
1	08–17 1834	5.2
2	09–03 1834	5.3
3	05–07 1865	5.2
4	06–13 1883	4.3
5	09–05 1886	4.2
6	10–25 1886	5.1

*After Muir Wood *et al.* (1989).

construct a corresponding soft bound model of observed magnitude with mean located at the center of the interval $\langle \underline{x}, \bar{x} \rangle$ and subjectively choose the standard deviation σ of soft bound model equal to the hard bound magnitude uncertainty δ. This choice of σ corresponds to a confidence level of 0.683.

In the second example, we are "almost sure" that the true magnitude X lies within our interval $\langle \underline{x}, \bar{x} \rangle$. Here we subjectively choose a confidence level of 0.997, which requires that the soft bound model's $\sigma = \delta/3$.

The results of applying the two approaches based on hard bounds and soft bounds uncertainty models to our data are given in Table 3. For comparison, in the last column the results of the "standard" approach (KS1), where magnitude errors are not taken into account, are also included. Return periods estimated by means of the hard bounds model, the soft bounds model with $\sigma = \delta$ and the "standard approach" are shown in Figure 3. The case with small magnitude uncertainties ($\sigma = \delta/3$) is not included. According to our experience, this case does not reflect reality; the uncertainties with $\sigma = \delta/3$ are too small to be correct.

The obtained results (Fig. 3, Table 3) indicate that the magnitude uncertainty, although very important in many respects, does not play a significant role in seismic hazard evaluation as long as it remains within a physically defensible range. The three curves of mean return periods presented in Figure 3, although calculated by three different procedures, do not significantly differ.

Significant difference in return period evaluation was observed particularly for large soft bounds magnitude uncertainties. For example, Table 3, where $\sigma = \delta$, gives a calculated return period of 9.6 years at magnitude 5 for the soft bounds procedure. Doubling the uncertainty for this procedure ($\sigma = 2\delta$) yields a calculated return period of 10.8 years, an increase of 12.5%. On the other hand, the hard bounds procedure is much less sensitive to the assumed large magnitude uncertainty. At magnitude 5, Table 3 shows a return period of 9.1 years for the hard bounds procedure. Doubling δ gives a return period of 9.7 years, a modest increase.

The fact that the hard bounds procedure is much less sensitive to the assumed magnitude uncertainty can be readily explained. In practice, the deviation of the two discussed models from the "standard" one (KS1) is determined by the value of the correction factors c_f. For the hard bounds model c_f is equal to $[\exp(\beta\delta) - \exp(-\beta\delta)]/(2\beta\delta)$, and for the soft bounds model $c_f = \exp(\beta^2\sigma^2/2)$. By definition, for the "standard" approach $c_f = 1$. Figure 4 shows the value of the correction factors c_f as a function of magnitude uncertainty: δ for the hard bounds and σ for the soft bounds model. The calculations were performed for $\beta = 2.0$. For the both discussed models, for a relatively small magnitude uncertainty (up to $\cong 0.2$) the correction factors c_f do not differ significantly from unity and the presence of errors can be ignored. The deviation of c_f factors from unity increases when magnitude uncertainty increases. In addition, the value of c_f for the soft bounds model increases much faster then the respective value for the hard bounds model. Such facts have an obvious physical meaning: large magnitude errors play an important role in the process of seismic hazard

TABLE 2

SUMMARY OF THE COMPLETE PARTS OF CATALOG

Magnitude	Frequency
Complete Part No. 1 Time Period: 01/01/1891 to 12/31/1950 Number of Earthquakes: 40 Threshold Magnitude: 3.8	
3.8	2
3.9	6
4.0	6
4.1	4
4.2	5
4.3	2
4.4	2
4.5	1
4.6	1
4.7	2
4.8	1
4.9	3
5.0	1
5.2	2
5.4	1
5.7	1
Complete Part No. 2 Time Period: 01/01/1951 to 12/31/1979 Number of Earthquakes: 37 Threshold Magnitude: 3.6	
3.6	3
3.7	3
3.8	4
3.9	4
4.0	2
4.1	1
4.2	4
4.3	3
4.4	2
4.5	2
4.6	1
4.7	2
4.8	1
4.9	1
5.2	1
5.3	2
5.5	1

TABLE 2

(*Continued*)

Magnitude	Frequency
Complete Part No. 3 Time Period: 01/01/1980 to 12/31/1989 Number of Earthquakes: 27 Threshold Magnitude: 3.0	
3.0	3
3.2	2
3.3	1
3.5	2
3.6	2
3.7	2
3.9	2
4.0	2
4.2	1
4.3	4
4.5	1
4.6	1
5.0	2
5.5	1
5.6	1

TABLE 3

ESTIMATION OF EARTHQUAKE HAZARD PARAMETERS AND RETURN PERIODS BY THE THREE DESCRIBED PROCEDURES

	Return Periods (Yr)		
Magnitude	Hard Bound Model ($\hat{\beta} = 1.29$, $\hat{\lambda}_{2.0} = 8.38$, $\hat{m}_{max} = 5.77$)	Soft Bound Model ($\hat{\beta} = 1.32$, $\hat{\lambda}_{2.0} = 8.51$, $\hat{m}_{max} = 5.77$)	Magnitude Errors Ignored ($\hat{\beta} = 1.29$, $\hat{\lambda}_{2.0} = 8.46$, $\hat{m}_{max} = 5.77$)
3.0	0.4	0.4	0.4
4.0	1.8	1.8	1.7
4.5	3.7	3.9	3.7
5.0	9.1	9.6	9.0
5.2	14.3	15.1	14.1
5.4	25.4	26.9	25.0
5.6	63.3	67.3	62.5
5.7	164.2	174.9	162.1

Calculations performed for the "hard bounds" magnitude uncertainties: $\delta_0 = 0.3$, $\delta_1 = 0.25$, $\delta_2 = 0.2$, and $\delta_3 = 0.15$. Results of application of the same data with the "soft bounds" magnitude uncertainties equal to $\sigma_0 = 0.3$, $\sigma_1 = 0.25$, $\sigma_2 = 0.2$, and $\sigma_3 = 0.15$. Results of application of the "standard approach," when magnitude uncertainty is not taken into account.

parameters evaluation, and the soft bounds model is much more sensitive to the large magnitude errors than the hard bounds model.

In order to examine how general our conclusions are, the approaches described above have been applied to the number of synthetic catalogs and to two real data sets. These comprise a catalog of events in Vrancea, Romania, extending from 984 to 1986, and a set of the largest events in Egypt between

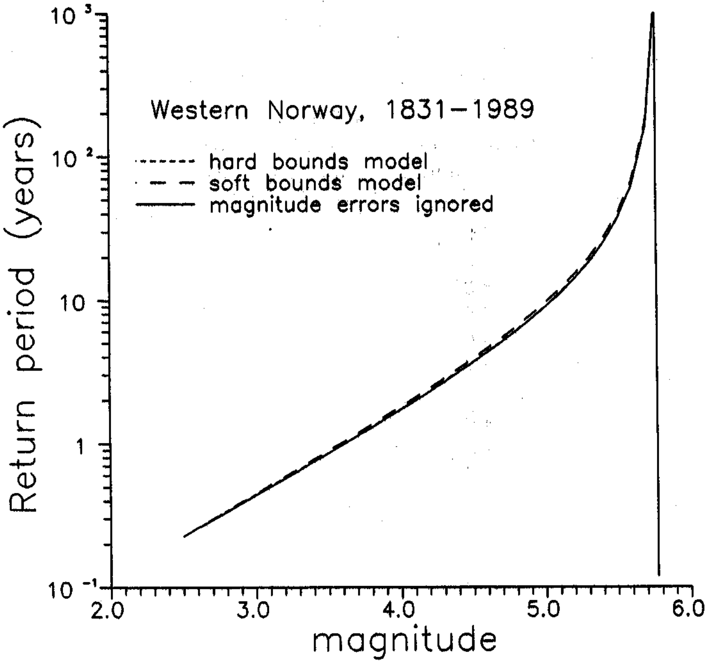

FIG. 3. Return periods estimated by the hard bounds model (curve 1), soft bounds model (curve 2), and standard approach when magnitude uncertainties are ignored (curve 3). The three curves corresponding to the three different models are practically the same, confirming our theoretical considerations.

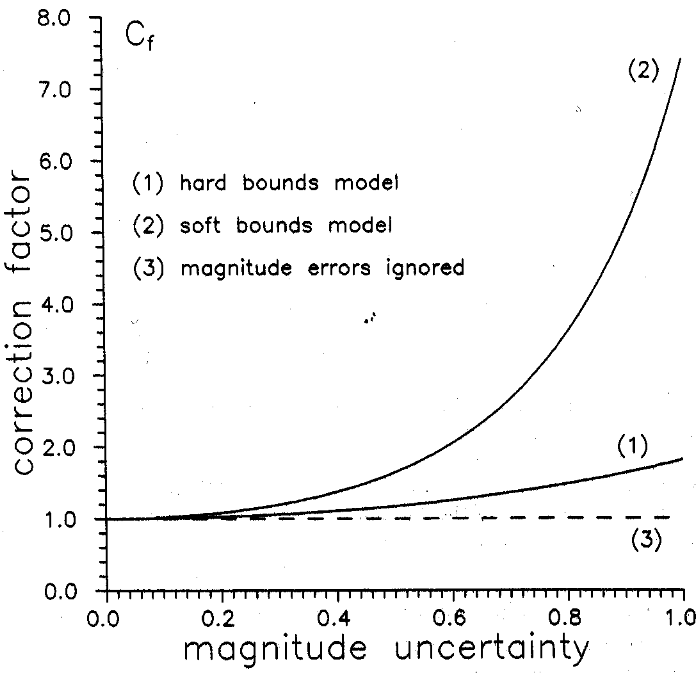

FIG. 4. The effect of magnitude uncertainty. The deviation of the two discussed models from the standard one is determined by the values of the correction factor c_f. Curve (1): hard bounds model, c_f as a function of magnitude uncertainty δ. Curve (2): soft bounds model, c_f as a function of magnitude uncertainty δ. Curve (3): standard KS1 model, in which magnitude errors are not taken into account; for such a case $c_f = 1.0$.

600 B.C. to A.D. 1989. The main characteristics of the obtained results are similar, although the data sets correspond to areas with different patterns of seismicity and vary in number of events and time span. In general, the evaluation of hazard parameters is stable and the results are in concordance with observational data. If insignificant magnitude errors are used, the three procedures return the values that are close to each other, although they were computed using entirely different techniques.

It is difficult to speak of the superiority of any of the three procedures discussed. The choice of one of them is largely subjective, depending on personal judgment regarding the nature of available earthquake data. For really large magnitude errors the standard procedure, which ignores them, gives always too high values of hazard. In terms of return periods, it produces return periods that are too short. In this sense the standard procedure could be characterized as a conservative or pessimistic one.

Remarks and Conclusions

A procedure for the use of incomplete seismological data is here extended to take account of magnitude uncertainty.

Two different models of earthquake magnitude uncertainty are described. In the first model, the uncertainty of earthquake magnitude is determined by hard bounds. It is assumes that such an interval contains the real unknown magnitude. In the second soft bound model, the real unknown magnitude differs from the observed magnitude by a Gaussian random error with zero mean and known standard deviations.

Two procedures, based on the two uncertainty models, as well as the standard procedure, were applied to a set of earthquake data felt and recorded in western Norway and adjacent sea area during the last 160 years. The obtained results suggest that any assumption regarding the *nature* of errors of earthquake magnitude distribution is not significant. Both procedures, for a reasonable range of the assumed magnitude uncertainty, give comparable results. A disregard of significant magnitude uncertainties leads to an overestimation of seismic hazard.

The described procedures are especially useful for seismic hazard evaluation based on incomplete historical data, when magnitude uncertainty is much greater than that from instrumental data.

A computer program was used to calculate the values in Table 3. It is written in FORTRAN 77 for IBM PC and compatible computers and can be provided if a blank 5.25" floppy disk is sent to one of the authors.

Acknowledgments

The authors sincerely thank Mr. Bruce Taylor for helpful suggestions, advice, and comments that improved an earlier version of this manuscript. We are also very grateful to the anonymous reviewer for helpful and stimulating comments.

References

Abramowitz, M. and I. A. Stegun (1970). *Handbook of Mathematical Functions*, 9th ed., Dover, New York.

Ambraseys, N., E. Banda, J. Irving, S. Mallard, C. Melville, T. Morse, R. Muir Wood, D. Mundoz, L. Serva, D. Shilston, and J. Vogt (1983). Notes on historical seismicity, *Bull. Seism. Soc. Am.* **73**, 1917–1920.

Backus, G. E. (1988). Comparing hard and soft prior bounds in geophysical inverse problems, *Geophys. J.* **94**, 249–261.

Båth, M. (1956). *An Earthquake Catalogue for Fennoscandia for the Years 1891-1950*. Sveriges Geologiska Undersøkning, Ser. C, No. 545, 52.

Bender, B. (1987). Effects of observational errors in relating magnitude scales and fitting the Gutenberg-Richter parameter β, *Bull. Seism. Soc. Am.* **77**, 1400-1428.

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

Chung, D. H. and D. L. Bernreuter (1981). Regional relationships among earthquake magnitude scales. *Rev. Geophys. Space Phys.* **19**, 649-663.

Cosentino, P., V. Ficara, and D. Luzio (1977). Truncated exponential frequency-magnitude relationship in the earthquake statistics, *Bull. Seism. Soc. Am.* **67**, 1615-1623.

Gan, Z. J. and C. C. Tung (1983). Extreme value distribution of earthquake magnitude, *Phys. Earth Planet. Interiors* **32**, 325-330.

Kijko, A. (1988). Maximum likelihood estimation of Gutenberg-Richter b parameter for uncertain magnitude values, *Pageoph* **127**, 573-579.

Kijko, A. and M. Dessokey (1987). Application of extreme magnitude distribution to incomplete earthquake files, *Bull. Seism. Soc. Am.* **77**, 1429-1436.

Kijko, A. and M. A. Sellevoll (1989). Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes, *Bull. Seism. Soc. Am.* **79**, 645-654.

Kijko, A. and M. A. Sellevoll (1990). Estimation of earthquake hazard parameters from incomplete and uncertain data files, *Natural Hazards* **3**, 1-13.

Muir Wood, R., G. Woo, and H. Bungum (1989). The history of earthquakes in the Northern North Sea, in *Earthquakes at North-Atlantic Passive Margins: Neotectonics and Postglacial Rebound*, S. Gregersen and P. W. Basham (Editors), Kluwer Academic Publishers, 297-306.

Nuttli, O. W. and R. B. Herrmann (1982). Earthquake magnitude scales, *J. Geotech. Eng. Div. ASCE* **108**, 783-786.

Page, R. (1968). Aftershocks and microaftershocks, *Bull. Seism. Soc. Am.* **58**, 323-335.

Sellevoll, M. A., J. Almas, and A. Kijko (1982). *A Catalogue of Earthquakes Felt in Norway 1953-1975 and a Statistical Analysis Related to the Completeness of Catalogues for Earthquakes felt in Norway 1899-1979, with Appendix of December 1989*, Seismological Observatory, University of Bergen, 1982.

Tinti, S. and F. Mulargia (1985). Effects of magnitude uncertainties on estimating the parameters in the Gutenberg-Richter frequency-magnitude law, *Bull. Seism. Soc. Am.* **75**, 1681-1697.

INSTITUTE OF GEOPHYSICS
POLISH ACADEMY OF SCIENCES
KS. JANUSZA 64
01-452 WARSAW, POLAND
(A.K.)

SEISMOLOGICAL OBSERVATORY
UNIVERSITY OF BERGEN
5007 BERGEN, NORWAY
(M.A.S.)

Manuscript received 12 February 1990