Seismic Hazard Assessment for Specified Area

ESTIMATION OF MAXIMUM REGIONAL MAGNITUDE m_{max}

At present there is no generally accepted method for estimating the value of the maximum regional magnitude m_{max}. The methods for evaluating m_{max} fall into two main categories: deterministic and probabilistic.

The deterministic procedure most often applied is based on the empirical relationships between the magnitude and various tectonic and fault parameters. The relationships are variously developed for different seismic areas and different types of faults. In most cases, unfortunately, the value of the parameter m_{max} determined by means of any deterministic procedure is very uncertain. The value of m_{max} can also be estimated purely on the basis of the seismological history of the area, viz. by utilizing seismic event catalogues and appropriate statistical estimation procedures. In this section the authors present a statistical procedure which can be used for the evaluation of the maximum regional magnitude m_{max}. It is assumed that both the analytical form and the parameters of the distribution functions of earthquake magnitude are known.

Suppose that in the area of concern, within a specified time interval T, there are n main seismic events with magnitudes M_1, M_2,..., M_n. Each magnitude $M_i \geq m_{min}$ ($i = 1,\ldots, n$), where m_{min} is a known threshold of completeness (i.e. all events having magnitude greater than or equal to m_{min} are recorded). It is further assumed that the seismic event magnitudes are independent, identically distributed, random values with probability density function (PDF), $f_M(m \mid m_{min}, m_{max})$ and cumulative distribution function (CDF), $F_M(m \mid m_{min}, m_{max})$ respectively. Parameter m_{max} is the upper limit of the range of magnitudes and thus denotes the unknown maximum regional magnitude which is to be estimated. For the Gutenberg-Richter relation, which is a frequency-magnitude relation, the respective CDF of earthquake magnitudes which are bounded from above by m_{max}, is

$$F_M\left(m \mid m_{\min}, m_{\max}\right)=\begin{cases}0 & \text{for } m<m_{\min}, \\ \dfrac{1-\exp\left[-\beta\left(m-m_{\min}\right)\right]}{1-\exp\left[-\beta\left(m_{\max}-m_{\min}\right)\right]}, & \text{for } m_{\min} \leq m \leq m_{\max}, \\ 1, & \text{for } m>m_{\max}.\end{cases} \tag{1}$$

where $\beta = b\ln(10)$, and b is the b-parameter of the Gutenberg-Richter relation.

From the condition that compares the largest observed magnitude $m_{\max}^{obs}$ and the maximum expected magnitude during a specified time interval T, we obtain, after integration by parts and simple transformations, the maximum regional magnitude $m_{\max}$ (Kijko and Graham, 1998)

$$\hat{m}_{\max} = m_{\max}^{obs} + \int_{m_{\min}}^{m_{\max}^{obs}} \left[F_M\left(m \mid m_{\min}, m_{\max}^{obs}\right)\right]^n dm. \tag{2}$$

In the exact version of this formula, as used in our computer program, the upper limit of integration $m_{\max}^{obs}$ is replaced by $m_{\max}$.

Further modifications of estimator (2) are straightforward. For example, following the assumption that the number of earthquakes occurring in unit time within a specified area obeys the Poisson distribution with parameter λ, after replacing n by λT, estimator (2) becomes

$$\hat{m}_{\max} = m_{\max}^{obs} + \int_{m_{\min}}^{m_{\max}^{obs}} \left[F_M\left(m \mid m_{\min}, m_{\max}^{obs}\right)\right]^{\lambda T} dm. \tag{3}$$

It is not difficult to show that for the Gutenberg-Richter-based magnitude CDF (1), the estimator (2) takes the form

$$\hat{m}_{\max} = m_{\max}^{obs} + \frac{E_1\left(Tz_2\right) - E_1\left(Tz_1\right)}{\beta \exp\left(-Tz_2\right)} + m_{\min} \exp\left(-\lambda T\right) \tag{4}$$

where $z_1 = -\lambda A_1 / (A_2 - A_1)$, $z_2 = -\lambda A_2 / (A_2 - A_1)$, $A_1 = \exp(-\beta m_{\min})$, $A_2 = \exp\left(-\beta m_{\max}^{obs}\right)$, and $E_1(\cdot)$ denotes an exponential integral function. The above estimator of $m_{\max}$ for the doubly truncated Gutenberg-Richter relation was first obtained by Kijko (1983), who was inspired by discussions with M.A. Sellevoll. Consequently we refer to equation (4) as the Kijko-Sellevoll estimator or, in short, K-S. From equations (3) and (4), the approximate variance of the maximum regional magnitude $\hat{m}_{\max}$, estimated according to the K-S procedure, is

$$Var\left(\hat{m}_{\max}\right) = \sigma_M^2 + \left[\frac{E_1(Tz_2) - E_1(Tz_1)}{\beta \exp(-Tz_2)} + m_{\min} \exp(-\lambda T)\right]^2 \tag{5}$$

where σ_M^2 is the variance in the determination of the largest observed magnitude $m_{\max}^{obs}$.

It should be noted that the above approach has several limitations. One of these is the assumption that the β-value in the CDF $F_M(m \mid m_{\min}, m_{\max})$ is known without error. However, it is possible for uncertainties in the β-value to be taken into account. Following the assumption that the variation of the β-value in the Gutenberg-Richter-based CDF (1) may be represented by a gamma CDF having parameters p and q, the compound CDF of earthquake magnitudes takes the form

$$F_M(m|m_{\min} m_{\max}) = \begin{cases} 0 & for \;\; m < m_{\min}, \\ C_\beta \left\{1 - \left[p / (p + m - m_{\min})\right]^q\right\} & for \;\; m_{\min} \le m \le m_{\max}, \\ 1, & for \;\; m > m_{\max}. \end{cases} \tag{6}$$

where C_β is a normalizing coefficient. It is not difficult to show that p and q can be expressed through the mean and variance of the β-value, where $p = \bar{\beta} / (\sigma_\beta)^2$ and $q = (\bar{\beta} / \sigma_\beta)^2$. The symbol $\bar{\beta}$ denotes the mean value of the parameter β, σ_β is the known standard deviation of β and describes its uncertainty, and C_β is equal to $1/\{1-[p/(p+m_{\max}-m_{\min})]^q\}$. Equation (6) is known as the Bayesian Exponential-Gamma CDF of earthquake magnitude. Knowledge of equation (6) makes it possible to construct the Bayesian version of estimator K-S. Thus, following (3) and (5), and through an application of Cramer's approximation, the Bayesian extension of the K-S estimator and its approximate variance becomes (Kijko and Graham, 1998):

$$\hat{m}_{\max} = m_{\max}^{obs} + \frac{\delta^{1/q+2} \exp\left[nr^{q} / \left(1 - r^{q}\right)\right]}{\beta} \left[\Gamma\left(-1/q, \delta r^{q}\right) - \Gamma\left(-1/q, \delta\right)\right] \tag{7}$$

$$Var\left(\hat{m}_{\max}\right) \cong \sigma_{M}^{2} + \left\{\frac{\delta^{1/q+2} \exp\left[nr^{q} / \left(1 - r^{q}\right)\right]}{\beta} \left[\Gamma\left(-1/q, \delta r^{q}\right) - \Gamma\left(-1/q, \delta\right)\right]\right\}^{2} \tag{8}$$

where $\delta = nC_{\beta}$ and $\Gamma(\cdot,\cdot)$ is the Incomplete Gamma Function. The Bayesian version of the K-S estimator will be denoted as K-S-B. From the above relations it follows that the assessment of the maximum regional magnitude $m_{\max}$ requires knowledge of the area-specific mean seismic activity rate λ and the Gutenberg-Richter parameter b. The maximum likelihood procedure for the assessment of these two parameters is presented in the following sections. Extensive comparison of performances of K-S and K-S-B estimators is given by Kijko and Graham (1998).

ASSESSMENT OF AREA-SPECIFIC PARAMETERS IN CASE OF INCOMPLETE DATA SETS

Since the technique applied for assessment of area-specific seismic hazard parameters is similar to the procedure already described by Kijko and Sellevoll (1989, 1992), only the main points of the procedure are presented. Let us assume that in the vicinity of the specified site

(i) the occurrence of the main seismic events in time can be described by a Poissonian process with an area-specific mean activity rate λ, and

(ii) earthquake magnitudes are distributed according to the doubly truncated Gutenberg-Richter-based relation (1).

Then the CDF of the largest magnitudes occurring during the time interval t, is (Kijko and Sellevoll, 1992)

$$F_M^{\max}\left(m|m_0, m_{\max}, t\right) = \frac{\exp\{-\lambda_0 t c F_M\left(m|m_{0,} m_{\max}\right)\} - \exp(-\lambda_0 t)}{1 - \exp(-\lambda_0 t)} \quad (9)$$

where $cF_M(m\,|\,m_0,m_{\max}) = 1\text{-}F_M(m\,|\,m_0,m_{\max})$, $\lambda_0 \equiv \lambda(m_0) = \lambda cF_M(m_0\,|\,m_{\min},m_{\max})$ is the mean activity rate of earthquake occurrence within the specified area with magnitude m_0 and above, m_0 is the lower earthquake magnitude in the extreme part of the catalogue, and $m_0 \geq m_{\min}$. The parameter $\lambda \equiv \lambda(m_{\min})$ is the mean activity rate of earthquakes with magnitude $m_{\min}$ and above. Magnitude $m_{\min}$ is the minimum threshold magnitude for the entire catalogue (Figure 1).

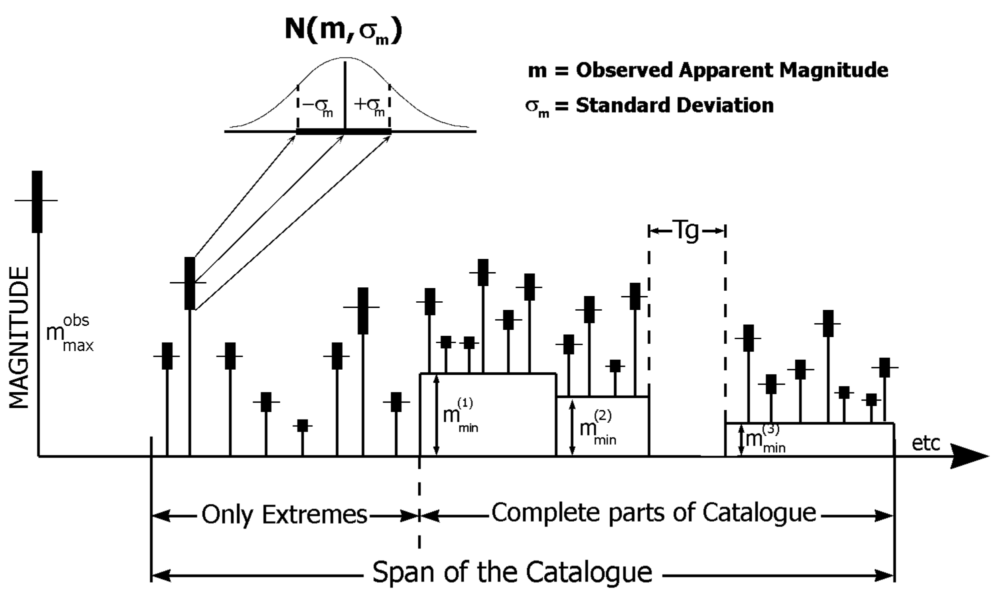

Figure 1. Illustration of data which can be used to obtain basic seismic hazard parameters for the area in the vicinity of the selected site by the procedure used. The approach permits the combination of largest earthquake data and complete data having variable threshold magnitudes. It allows the use of the largest known historical earthquake ($m_{\max}^{obs}$) which occurred before the catalogue began. It also accepts "gaps" (T_g) when records were missing or the seismic networks were out of operation. Uncertainty in earthquake magnitude is also taken into account in that an assumption is made that the observed magnitude is true magnitude subjected to a random error that follows a Gaussian distribution having zero mean and a known standard deviation. After Kijko and Sellevoll (1992).

If the error in the determination of magnitude is assumed to be normally distributed with standard deviation σ_M (Tinti and Mulargia, 1985a,b), the CDF of the apparent magnitude becomes (Gibowicz and Kijko, 1994)

$$F_M\left(m \middle| m_{\min}, m_{\max}, \sigma_M\right) = F_M\left(m \middle| m_{\min}, m_{\max}\right) D(m, \sigma_M) \tag{10}$$

where $D(m,\sigma_M) = \{A_1[\mathrm{erf}(y_1) + 1] + A_2[\mathrm{erf}(y_2) - 1] - 2C(m,\sigma_M)A(m)\}/2[A_1\text{-}A(m)]$, $C(m,\sigma_M) = 0.5\exp(\chi^2)[\mathrm{erf}(y_1 + \chi) + \mathrm{erf}((y_2 - \chi)]$, $y_1 = (m\text{-}m_{\min})/\sqrt{2}\,\sigma_M$, $y_2 = (m_{\max}\text{-}m)/\sqrt{2}\,\sigma_M$, σ_M denotes the standard error of earthquake magnitude determination, $A(m) = \exp(-\beta m)$, $A_1 = \exp(-\beta m_{\min})$, $A_2 = \exp(-\beta m_{\max})$, erf(·) is the error function, $\chi = \beta\sigma_M/\sqrt{2}$, and the magnitude m is unbounded from both ends. Further application of CDF (10) requires additional renormalizations. If m_C is the cut-off value of apparent magnitude, at and above which the earthquakes are complete, then its normalized CDF $\tilde{F}_M\left(m \middle| m_C, m_{\max}, \sigma_M\right)$ is zero up to m_C, and is equal to $F_M(m \mid m_{\min},m_{\max},\sigma_M)/[1 - F_M(m_C \mid m_{\min},m_{\max},\sigma_M)]$ for $m \geq m_C$. Finally, from the assumed model of apparent magnitude it follows that the "true" mean activity rate $\lambda(m)$ must be replaced by its "apparent" counterpart, $\tilde{\lambda}(m)$, according to the approximate relation $\tilde{\lambda}(m) = \lambda(m)\exp(\chi^2)$.

From the definition of PDF and from relations (9) and (10) it follows that the PDF of the strongest earthquake within a period t, with apparent magnitude $m \geq m_0$ and standard deviation σ_M, is

$$f_M^{\max}\left(m \middle| m_0, m_{\max}, t, \sigma_M\right) = \frac{\tilde{\lambda}_0 t \tilde{f}_M\left(m \middle| m_0, m_{\max}, t, \sigma_M\right) \cdot \exp\left[-\tilde{\lambda}_0 t c\tilde{F}_M\left(m \middle| m_0, m_{\max}, t, \sigma_M\right)\right]}{1 - \exp\left(-\tilde{\lambda}_0 t\right)}, \tag{11}$$

where $c\tilde{F}_M\left(m \middle| m_0, m_{\max}, t, \sigma_M\right) = 1 - \tilde{F}_M\left(m \middle| m_0, m_{\max}, t, \sigma_M\right)$.

After introducing the PDF (11), one can construct the likelihood function of the strongest earthquake magnitudes from the extreme part of the catalogue. Such a function depends on the unknown area-characteristic parameters (λ,β), and becomes

$$L_0(\lambda,\beta) = const \prod_{j=1}^{n_0} f_M^{\max}\left(m_{0j} \middle| m_0, m_{\max}, t_{0j}, \sigma_{M0j}\right) \tag{12}$$

In relation (12), the m_{0j} is the apparent magnitude of the strongest earthquake occurring during the time interval t_j, σ_{M0j} is the value of its standard deviation, $j = 1,..., n_0$, and n_0 is the number of earthquakes in the extreme part of the catalogue.

It is assumed that the second, complete part of the catalogue can be divided into s sub-catalogues (Figure 1). Each of them has a time span T_i and is complete, starting from the known magnitude $m_{\min}^{(i)}$. For each sub-catalogue i, m_{ij} is the apparent magnitude, $m_{ij} \geq m_{\min}^{(i)}$, and σ_{Mij} is its standard deviation, $j = 1,..., n_i$, where n_i denotes the number of earthquakes in each complete sub-catalogue and $i = 1,...,s$. If the size of seismic events is independent of their number, the likelihood function of earthquake magnitudes present in each complete sub-catalogue i, is equal to $L_i(\lambda,\beta) = L_i(\beta)\, L_i(\lambda)$, which is the product of the function of β, $L_i(\beta)$, and the function of λ, $L_i(\lambda)$. Following the definition of the likelihood function of a set of independent observations, the function $L_i(\lambda)$ is given by $const \prod_{j=1}^{n_i} \tilde{f}_M\left(m_{ij} \middle| m_{\min}^{(i)}, m_{\max}, \sigma_{Mij}\right)$. The assumption that the number of earthquakes per unit time is a Poisson random variable gives a form of $L_i(\lambda)$ equal to $const\left(\tilde{\lambda}_i t_i\right)^{n_i} ex\left(-\tilde{\lambda}_i t_i\right)$, where const is a normalizing factor and $\tilde{\lambda}_1$ is the apparent, mean activity rate for the complete sub-catalogue i. For the ith complete sub-catalogue the true mean activity rate is equal to $\lambda_i \equiv \lambda\left(m_{\min}^{(i)}\right) = \lambda c F_M\left(m_{\min}^{(i)} \middle| m_{\min}, m_{\max}\right)$. Functions $L_i(\beta)$ and $L_i(\lambda)$, for $i = 1,...,s$, define the likelihood functions for each complete sub-catalogue. Finally, the joint likelihood function of all data in the catalogue, extreme and complete, is given by:

$$L(\lambda,\beta) = \prod_{i=0}^{n_s} L_i(\lambda,\beta) \tag{13}$$

The maximum likelihood estimates λ and β are the values of $\hat{\lambda}$ and $\hat{\beta}$ that maximize the likelihood function (13). From a formal point of view, the maximum likelihood estimate of $m_{\max}$ is simply the largest observed earthquake magnitude $m_{\max}^{obs}$. This follows from the fact that the likelihood function (13) decreases monotonically for $m_{\max} \rightarrow \infty$.

Therefore, by including one of the formulae for $m_{\max}$ [the K-S estimator (4) or its Bayesian version (eq. 7)] and by putting $\partial \ln L(\lambda,\beta)/\partial\lambda = 0$ and $\partial \ln L(\lambda,\beta)/\partial\lambda = 0$, we obtain a set of equations determining the maximum likelihood estimate of the area-specific parameters $\hat{\lambda}$, $\hat{\beta}$ and $\hat{m}_{\max}$. Such a set of equations is given by Kijko and Sellevoll (1989, 1992), and can be solved by an iterative scheme.

REFERENCES

Gibowicz, S.J. and Kijko, A. (1994). *An Introduction to Mining Seismology (*Academic Press), San Diego.

Kijko, A. (1983), A modified form of the first Gumbel distribution: model for the occurrence of large earthquakes, Part II: Estimation of parameters, *Acta Geophys. Pol.*, **31**, 27-39.

Kijko, A. and Graham, G. (1998), "Parametric-Historic" procedure for probabilistic seismic hazard analysis. Part I: Assessment of maximum regional magnitude $m_{\max}$, *Pure Appl. Geophys*, **152**, 413-442.

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

Kijko, A. and Sellevoll, M.A. (1992), Estimation of earthquake hazard parameters from incomplete data files, Part II, Incorporation of magnitude heterogeneity, *Bull. Seism. Soc. Am*. **82**, 120-134.

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

Tinti, S. and Mulargia, F. (1985b), Application of the extreme value approaches to the apparent magnitude distribution of the earthquakes, *Pure Appl. Geophys*. **123**, 199-220.