Uncertainties in the estimation of M_{max}

Girish C Joshi and Mukat Lal Sharma*

Indian Institute of Technology Roorkee, Roorkee, Uttarakhand 247 667, India.
**e-mail: mukutfeq@iitr.ernet.in*

In the present paper, the parameters affecting the uncertainties on the estimation of M_{max} have been investigated by exploring different methodologies being used in the analysis of seismicity catalogue and estimation of seismicity parameters. A critical issue to be addressed before any scientific analysis is to assess the quality, consistency, and homogeneity of the data. The empirical relationships between different magnitude scales have been used for conversions for homogenization of seismicity catalogues to be used for further seismic hazard assessment studies. An endeavour has been made to quantify the uncertainties due to magnitude conversions and the seismic hazard parameters are then estimated using different methods to consider the epistemic uncertainty in the process. The study area chosen is around Delhi. The b value and the magnitude of completeness for the four seismogenic sources considered around Delhi varied more than 40% using the three catalogues compiled based on different magnitude conversion relationships. The effect of the uncertainties has been then shown on the estimation of M_{max} and the probabilities of occurrence of different magnitudes. It has been emphasized to consider the uncertainties and their quantification to carry out seismic hazard assessment and in turn the seismic microzonation.

1. Introduction

Seismic hazard assessment provides quantitative evaluations of the nature of ground shaking at a specified location which when combined with the local site conditions and expanded for a region, becomes important input for seismic microzonation. One of the important outputs of the seismic hazard assessment is the maximum probable magnitude of the earthquake, which a source can generate in future. This parameter has important bearing on the rest of the estimations including the strong ground motion for a site or for the seismic microzonation. This parameter depends on the seismicity catalogue, source characteristics and the modeling of uncertainties in the process of seismic hazard assessment. In recent years, the use of probabilistic design of experiments have allowed us to consider uncertainties at various steps of the evaluation of seismic hazard. Treatment of uncertainty estimation depends on the source of the uncertainty.

Uncertainties in the demarcation of source zone boundaries, in the size, location, and the activity rate of earthquakes, and in variation of ground motion characteristics with earthquake size and distance, belong to the most important epistemic uncertainties (McGuire 2004). This type of uncertainty is due to incomplete knowledge of the physics of the earthquake process and/or lack of data and can be reduced by more data and improved theories (Tavakoli 2002). Epistemic uncertainty is due to a lack of knowledge about the behaviour of the system that is conceptually resolvable. Thus epistemic uncertainty can, in principle, be eliminated with sufficient study and, therefore, expert judgements may be useful in its reduction. From a psychological point of view, epistemic (or internal) uncertainty reflects the possibility of errors in our general knowledge.

Keywords. Epistemic; aleatory; uncertainties; hazard assessment; maximum magnitude.

J. Earth Syst. Sci. **117**, S2, November 2008, pp. 671–682
© Printed in India.

Given the uncertainty in earthquake size, location, and associated ground motion, there are several models and parameters which are determined empirically and which therefore have associated uncertainties, referred to as epistemic uncertainties in seismic hazard assessment. These uncertainties result from a finite number of data available on which to base estimates, and would vanish if an infinite number of appropriate data were available (McGuire and Shedlock 1981).

Aleatory uncertainty arises because of natural, unpredictable variation in the performance of the system under study. The knowledge of experts cannot be expected to reduce aleatory uncertainty although their knowledge may be useful in quantifying the uncertainty. Thus, this type of uncertainty is sometimes referred to as irreducible uncertainty. Examples of aleatory uncertainties in seismic hazard analysis are (McGuire 2004) future earthquake locations, future earthquake source properties (e.g., magnitudes), ground motion at a site given median value of motion and the details of the fault rupture process (e.g., direction of rupture). Aleatory uncertainty can be characterized by the standard deviation of the residuals. The probabilistic seismic hazard assessment provides a structure in which the uncertainties can be identified, quantified and combined in a rational approach to provide a more complete depiction of the seismic hazard. Thus, there may be multiple scenarios for a site or for seismic microzonation. In the present paper, the parameters affecting the uncertainties on the estimation of $M_{\max}$ have been investigated by exploring different methodologies being used in the analysis of seismicity catalogue and estimation of seismicity parameters. A case study of estimation of $M_{\max}$ for the area around Delhi region has been considered for the purpose.

2. Delhi – A case study

The Delhi region sprawls over approximately $1500\,\text{km}^2$ and occupies an area between latitude 28°24′ to 28°53′N and longitude 76°50′ to 77°20′E. It is bounded by the Indo-Gangetic alluvial plains in the north as well as east, by the Thar Desert in the west and by Aravalli hill ranges in the south. The terrain of Delhi is flat in general except for a low NNE–SSW trending ridge, which is considered an extension of the Aravalli hill ranges of Rajasthan. Mostly, the seismicity around Delhi appears to be associated with a major geological structure, which is known as Delhi–Haridwar Ridge. This ridge constitutes an important tectonic block between 28°–30°N and 76°–79°E with a NE–SW trend. It coincides with the extension of the Aravalli Mountain belt beneath the alluvium plains of the Ganga basin to the northeast of Delhi towards the Himalayan Mountains. River Yamuna, which is another prominent feature of Delhi, enters the city from north and flows southward with an eastern bend near Okhla. This path forms a near trijunction with the Lahore–Delhi Ridge and the Delhi–Hardwar Ridge. Based on the studies carried out by Sharma *et al* (2003) the sources taken for consideration are Himalayan source, Delhi–Hardwar Ridge, Moradabad Fault and Rajasthan Great Boundary Fault. The seismicity and the source considered are shown in figure 1.

To prepare the earthquake catalogue the earthquake occurrence data have been compiled from India Meteorology Department (IMD, Indian agency), International Seismological Center ISC, UK and United States Geological Survey, USGS, USA (International agencies). The earthquake catalogue has been updated by the analysis of events from historic to present time (1702–2004). The data on aftershock and foreshock have been removed to consider only independent events in the catalogue. Such filtering has been done by analyzing the spatial and temporal distribution of the events based on modified Omori's relationship (Utsu 2002a). The historic events, for which only intensities were assigned for their size, have not been considered for developing relationships between the magnitude scales. After 1960, systematic recording of earthquakes were carried out, but the reporting of events have been done using different magnitude scales. The main reporting agencies for reporting on different scales for an event are ISC and USGS. IMD has reported the size with local magnitude scale (M_L) only. Using the data set from ISC, USGS and IMD only 61, 101 and 30 events could be complied having the magnitudes given on two scales as M_b and M_s, M_L and M_b and M_L and M_s respectively. The magnitudes of the compiled data were verified again from catalogues given by different reporting agencies. The earthquakes used for the present study are shallow earthquakes with depth $\leq 70\,\text{km}$ because of their greater concern in seismic hazard analysis.

3. Homogenization of catalogue

The relationship between different magnitude scales is of paramount importance for conversions to carry out homogenization of seismicity catalogues for further studies related to seismological statistics and for seismicity and amplitude studies. Uncertainties are associated with each magnitude scale which plays an important role while converting one scale magnitude into another scale. Most

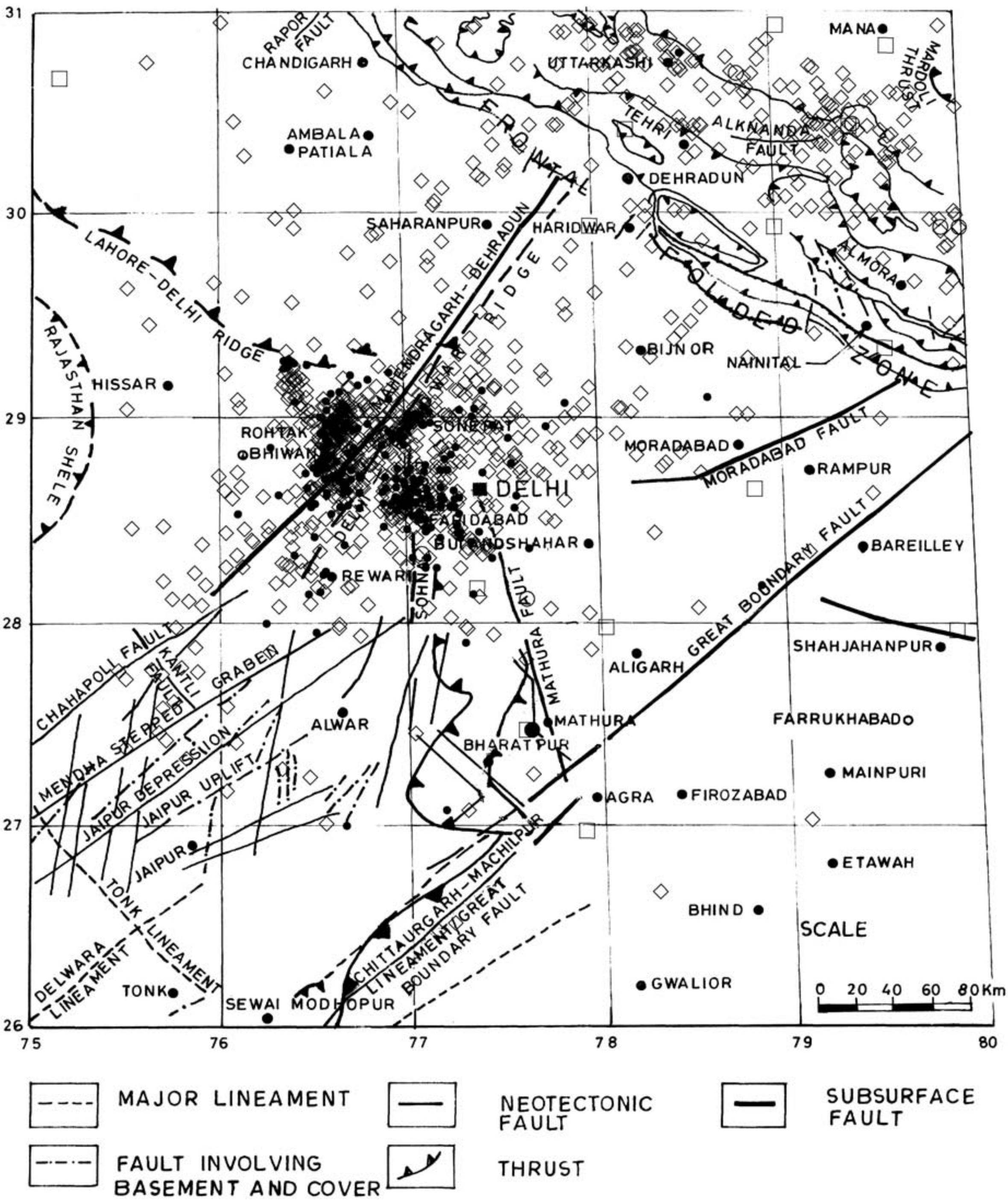

Figure 1. Seismotectonic setup around Delhi.

widely used magnitude scale is local wave magnitude or Richter scale magnitude defined by Richter (1935). The surface wave magnitude and the body wave magnitude were proposed later. Based on empirical and theoretical evidences, the differences in the source mechanism, size and duration produce different relative values for the amplitudes of surface and body waves (Prozorov and Hudson 1974). In general, the ratio of the body to the surface wave amplitude may fluctuate significantly from one earthquake to another and hence it is necessary to use statistical methods treating earthquakes with the same general source condition in big samples for the reliable detection of a systematic trend. The accuracy and reliability of seismic hazard assessment depends on the kind of seismicity data being used which should be homogenous and complete spatially and temporally. While a lot of work has been done for checking the incompleteness of the database, very few efforts have been made for quantifying the uncertainties involved in the heterogeneous magnitude scales and their scaling law for such a database (Ameer *et al* 2005). Also the conversion used to compensate for parameter incompatibilities are usually an aleatory variability since the two parameters of magnitude scales, viz., two scales are generally not perfectly correlated (Bommer *et al* 2005). Many empirical relationships were developed in the past between various magnitude scales for maping one type of magnitude onto the other (Gutenberg and Richter 1956a, 1956b; Bath 1968; Marshal 1970; Gibowicz 1972 and Ameer *et al* 2005) for different regions in the world. Most commonly used method is the least square or linear fitting between two parameters. The empirically obtained curves for the relationship between two magnitude scales represent earthquakes with average source characteristics (Utsu 2002a). The data for developing such relationships between two magnitude scales are generally very less and give a selected sample choice to obtain a relationship between two magnitude scales for the region under study. In the present study three catalogues, C1, C2 and C3 have been prepared using three distinct procedures to homogenize the dataset.

3.1 *Catalogue C1 using linear regression*

For developing the relationship between different magnitudes scales, which would homogenize the compiled catalogue for size, statistical methods have been adopted in the present study. While estimating the magnitude of any earthquake event there is always some inherent variability that the measurement is bias, called aleatory randomness. For establishing a relationship between the two different types of the scales it is necessary, therefore, to consider both scales having different aleatory variables. Considering the two magnitude scales to regressed for relationship as X_i and Y_i the least square fit to a set of n data values ($X_í$, $Y_í$)

$$Y_í = \beta_0 + \beta_1 X_í + \varepsilon_í, \quad í = 1,\ 2, \ldots, n, \qquad (1)$$

where $X_í$ is any of the magnitude scales, i.e., $M_{bı}$ (body wave magnitude), $M_{sí}$ (surface wave magnitude) or $M_{Lí}$ (local magnitude) in the X axis and similarly $Y_í$ is the magnitude scales other than specified on X axis. It is generally assumed that $Y_í$ is subject to the error $\varepsilon_í$ and $X_í$ is not subject to error. If this is true, and if the errors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are independently and normally distributed, $N(0, \sigma^2)$, maximum likelihood estimation, and least squares estimation, namely, $\underset{\beta o, \beta_1}{\text{Minimize}} \ \sum_{i=1}^{n} (Yi - \beta_0 - \beta_1 Xi)^2$ provide the least square estimates of (β_0, β_1). The linearly regressed surface wave magnitude value (M_S) has been analyzed and for linear homogenization, the value in terms of M_S is given as (Joshi and Sharma 2006):

$$M_S = 1.38 M_b - 1.029; \quad r^2 = 0.684, \qquad (2)$$

$$M_S = 0.962 M_L - 0.235; \quad r^2 = 0.528. \qquad (3)$$

3.2 *Catalogue C2 using bivariate analysis for orthogonal curve fitting*

The case of C1, i.e., linear regression with only single scale aleatory variable for making relationship between the two scales, increases the uncertainty while changing one scale magnitude into other. For quantifying variability while making relationships between two scales it is important to give both magnitude scales equal importance and each scale variability should be taken into analysis. Since both types of magnitudes, as reported in the catalogue, are having errors, the usual meaning of dependent and independent variable fails and the variables can be written as $Y_í = \eta_í + \varepsilon_í$ and $X_í = \xi_í + \delta_í$. It is assumed that a straight-line relationship between errors in two magnitudes exists and relationship given as $\eta_í = \beta_0 + \beta_1 \xi_í$ holds between the true but unobserved values $\eta_í$ and the n unknown parameters $\xi_í$. Based on the above and substituting for $\xi_í$, $Y_í$ can be written as $Y_í = \beta_0 + \beta_1 X_í + (\varepsilon_í - \beta_1\ \delta_ı)$. By fitting the two scales using usual least square fit there will be biased estimates in slope parameter of fitted equation (Draper and Smith 2005). There is an identifiable problem to obtain the maximum likelihood estimates of β_0 and $\beta_{1,}$ so that both magnitude scales minimize the errors in both scales and the uncertainty could be modeled under the distributional assumption made above. Four cases are possible as follows (Sprent and Dolby 1980).

Case I. (X, Y) are bivariate normal variables and $E(Y|X) = \beta_0 + \beta_1 X$.

Case II. $Y \sim N\ (\beta_0 + \beta_1 X, \sigma^2)$. The observed X-values are fixed on realizations of a random variable with any (reasonable) distribution.

In both I and II, estimates via maximum likelihood are usual least square estimates $b_1 = S_{XY}/S_{XX}$ and $b_0 = Y_{\text{mean}} - b_1 X_{\text{mean}}$, where

$$S_{XY} = \sum_{i=1}^{n} (X_i - X_{\text{mean}})(Y_i - Y_{\text{mean}}) \qquad (4)$$

$$S_{XX} = \sum_{i=1}^{n} (X_i - X_{\text{mean}})^2.$$

Case III. The estimation cannot be carried through without some additional information being added. The knowledge of the variance ratio $\lambda = \sigma^2/\sigma_\delta^2$ (Barnett 1967; Wang 1989) is a very important parameter to maximum likelihood estimates for fitting two scales. The case of equation defined above if variance ratio (λ) were known, maximum likelihood leads to estimates

$$\hat{\beta}_1 = \frac{\left[\begin{array}{l} S_{YY} - \lambda S_{XX} \\ \quad +\{(S_{YY} - \lambda S_{XX})^2 + 4\lambda S_{XY}^2\}^{1/2} \end{array}\right]}{(2S_{XY})} \quad (5)$$

$$\hat{\beta}_0 = Y_{\text{mean}} - \hat{\beta}_1 X_{\text{mean}},$$

where

$$S_{YY} = \sum_{i=1}^{n} (Y_i - Y_{\text{mean}})^2.$$

Case IV. The fit requires that the ratio of the variance of the error in X to the error in Y be specified. The variance ratio is infinite in standard least squares because variance in $X(\sigma_\delta^2)$ is zero. If an orthogonal fit is done with a large error ratio, the fitted line approaches the standard least squares line of fit.

Finally the regression as per case IV has been applied to data and the relationship thus obtained for M_b–M_s, M_L–M_b and M_L–M_s are as follows (Joshi and Sharma 2006)

$$M_b\text{–}M_s\text{: Variance ratio (2.352)} \quad M_S = 1.34M_b - 1.98, \quad (6)$$

$$M_L\text{–}M_b\text{: Variance ratio (0.001)} \quad M_b = 0.85M_L + 0.81, \quad (7)$$

$$M_L\text{–}M_s\text{: Variance ratio (0.441)} \quad M_S = 1.17M_L - 1.03. \quad (8)$$

3.3 *Catalogue C3 based on magnitude conversion into M_W scale*

The most preferred scale for describing any seismic event is a moment magnitude scale M_W. Since limited data are available of moment magnitude in the study region, for demonstration of the main objective of the study here, other available relationships have been used for homogenization of catalogue into M_W scale. The various relationships developed by Suckale *et al* (2005) have been used in the present analysis. The developed relationships have been derived using Maximum Likelihood approach for regression in each set of magnitude scale conversion. The regressions were carried out assuming that the error of the M_W value and the respective other magnitude type is equal

$$M_W = 1.2690M_S - 1.0436 \quad \text{(converting } M_S \text{ to } M_W\text{)}, \quad (9)$$

$$M_W = 0.7813M_b + 1.5175 \quad \text{(converting } M_b \text{ to } M_W\text{)}, \quad (10)$$

$$M_W = 0.6960M_L + 1.7738 \quad \text{(converting } M_L \text{ to } M_W\text{)}. \quad (11)$$

4. Estimation of *a*, *b* and M_c

A critical issue to be addressed before carrying out seismic hazard analysis is to assess the quality, consistency and homogeneity of the earthquake catalogue. The three catalogues prepared should thus undergo a quality check especially for cut-off magnitude which has direct bearing on the estimation of a and b values of the Gutenberg–Richter relationship. A discussion on the methodologies adopted to estimate the M_c is given below. There are nine methods using which the a, b and M_c values are estimated for the three catalogues C1, C2 and C3 and all the four sources namely, Himalaya (HIM), Delhi–Haridwar Ridge (DHR), Moradabad Fault (MOR) and Rajasthan Great Boundary Fault (RGBF). The nine methods include the estimation of a, b and M_c are Maximum Curvature method (M1), Fixe $M_c = M_{\min}$ (M2), goodness of fit Mc90 (M3) and Mc95 (M4), best combinations of Mc90 and Mc95 and maximum curvature (M5), entire magnitude range (M6), Shi and Bolt (1982) method (M7), Bootstrap method (M8) and Cao and Gao (2002) method (M9).

Woessner and Wiemer (2005) have developed a method to determine the magnitude of completeness M_c and its uncertainty which model the entire magnitude range (EMR method) consisting of the self-similar complete part of frequency-magnitude distribution and incomplete portion, thus providing a comprehensive seismicity model. The approach is similar to that of Ogata and Katsura (1993), and uses a maximum-likelihood

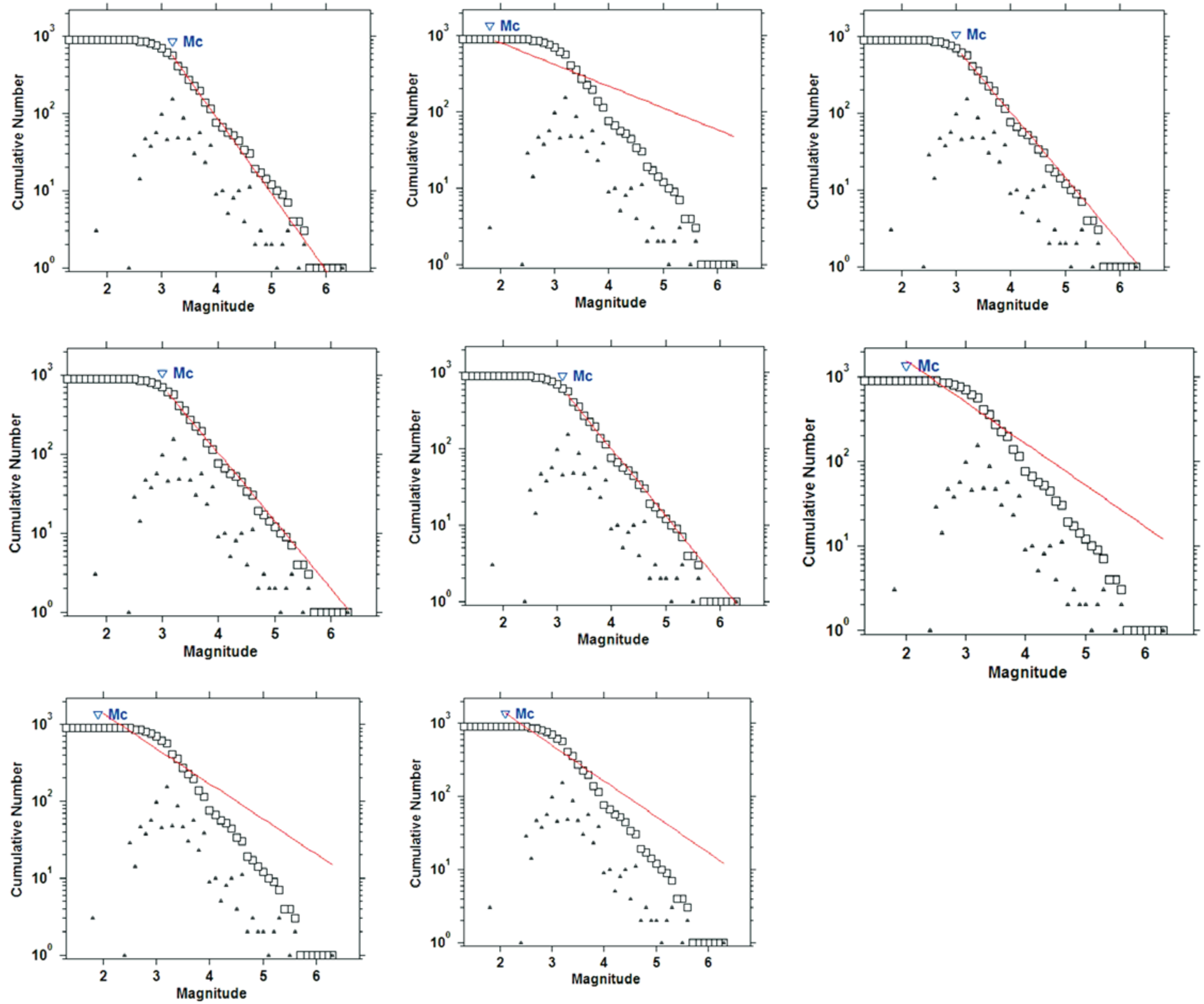

Figure 2. Estimation of a, b and M_c values using **(1)** maximum curvature, **(2)** fixed M_c = minimum magnitude, **(3)** $M_c = 90$, **(4)** best combination, **(5)** EMR, **(6)** Shi and Bolt, **(7)** bootstrapping and **(8)** Cao and Gao methods.

estimate for a model. The method uses the entire magnitude range to obtain a more robust estimate of M_c especially for mapping purpose. For data above an assumed M_c, Woessner and Wiemer (2005) presume a power-law behaviour and compute a and b values using maximum-likelihood estimates for the a and b values (Aki 1965; Utsu 1965). For given data and the assumed M_c, a normal cumulative distribution function $q(M|\mu,\sigma)$ that describes the detection capability as a function of magnitude is fitted to the data. $q(M|\mu,\sigma)$ denotes the probability of seismic network to detect an earthquake of a certain magnitude and can be written as:

$$q(M|\mu,\sigma) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{M_c} \exp\left(-\frac{(M-\mu)^2}{2\sigma^2}\right) dM, & M < M_c \\ 1, & M \geq M_c \end{cases}. \quad (12)$$

Here, μ is the magnitude at which 50% of the earthquakes are detected and σ denotes the standard deviation describing the width of the range where earthquakes are partially detected. The free parameters μ and σ are estimated using maximum-likelihood estimate. The best fitting model is the one that maximizes the log-likelihood function for four parameters: μ and σ, as well as a and b.

Wiemer and Wyss (2000) proposed two methods based on the assumption of self similarity. A fast and reliable estimate of M_c is to define the point of maximum curvature (MAXC) as magnitude of completeness by computing the maximum value of the first derivative of frequency-magnitude curve. In practice, this matches the magnitude bin with highest frequency of events in the non-cumulative frequency-magnitude distribution. Despite the easy applicability and relative robustness of this approach, M_c is often underestimated especially for gradually-curved frequency-magnitude distribution that results from spatial or temporal heterogeneities.

The Goodness of Fit method to calculate M_c compares the observed frequency-magnitude distribution with synthetic ones (Wiemer and Wyss 2000). The goodness-of-fit is computed as the absolute difference of the number of events in the magnitude bins between the observed and synthetic Gutenberg–Richter distribution. Synthetic distributions are calculated using estimated a and b values of the observed dataset for $M \geq M_{\text{cut}}$ as a function of ascending cut-off magnitudes M_{cut}. R defines the fit in percentage to the observed frequency magnitude distribution, and is computed as a function of cut-off magnitude. A model is found at an R-value at which a predefined percentage (90% or 95%) of the observed data is modeled by a straight line.

Cao and Gao (2002) estimate M_C using the satiability of the b-value as a function of cut-off magnitude M_{CO}. This model is based on the assumption that the b-values ascend for $M_{CO} < M_C$, remain constant for $M_{CO} \geq M_C$. If $M_{CO} \leq M_C$, the resulting b-value will be low. M_{CO} approaches its true value and remains constant for $M_{CO} \gg M_C$. Cao and Gao (2002) arbitrarily defined M_C as the magnitude for which the change in b-value, of two successive M_{CO}, is smaller than 0.03. Woessner and Wiemer (2005) have found the criterion to be unstable, since frequency of events in single magnitude bins can vary strongly. To stabilize the results for b-value uncertainty the Shi and Bolt (1982) criterion uses more objective measures as:

$$\partial b = 2.3b^2 \sqrt{\frac{\sum_{i=1}^{N}(M_i - \langle M \rangle)}{N(N-1)}}, \quad (13)$$

with $\langle M \rangle$ being the mean magnitude and N the number of events. Figure 2 shows an example of the estimation of a, b and M_c using these methods for catalogue C3 for the DHR source. The results could not be obtained using Mc95. The results are tabulated in table 1 where a, b and M_c values are shown for the three catalogues and the four sources using all the nine methods. For RGB the values could be estimated for M2 only from catalogue C1. This is because of the lesser data available for the source. The table reveals lot of variation in the values. For example, the b value for DHR has been estimated in the range of 0.45 ± 0.01 (M2) to 0.74 ± 0.05 (M7), 0.55 ± 0.02 (M9) to 1 ± 0.01 (M1), 0.29 ± 0.04 (M2) to 0.99 ± 0.04 (M1) for the catalogue C1, C2 and C3.

5. Estimation of $M_{\max}$

The maximum magnitude is an important variable in the seismic hazard estimation as it reflects maximum potential of strain released in larger earthquakes. The instrumental and historical records of earthquakes are often too short to reflect the full potential of faults or thrusts. The maximum regional magnitude, $M_{\max}$, is defined as the upper limit of magnitude for a given region or it is magnitude of largest possible earthquake. In other words it is a sharp cut-off magnitude at a maximum magnitude $M_{\max}$, so that, by definition, no earthquakes are to be expected with magnitude exceeding $M_{\max}$.

Table 1. *Estimation of a, b and M_c values for all sources using three catalogues from nine methods.*

		C1				C2				C3			
		DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB
M1	a	3.95	3.46	2.8	–	3.58	3.27	2.33	1.56	5.92	4.34	4.13	–
	b	0.71 ± 0.03	0.40 ± 0.06	0.56 ± 0.18	–	0.64 ± 0.02	0.35 ± 0.02	0.45 ± 0.13	0.27 ± 0.01	0.99 ± 0.04	0.50 ± 0.05	0.72 ± 0.19	–
	M_c	1.7 ± 0.05	2.5 ± 0.45	2.4 ± 0.55	–	1 ± 0.01	2.4 ± 0.4	2.0 ± 0.84	2.6 ± 1.9	3.2 ± 0.01	3.8 ± 0.15	3.7 ± 0.37	–
M2	a	3.40	3.0	1.96	1.54	3.59	2.78	1.92	1.50	3.48	3.46	2.25	2.20
	b	0.45 ± 0.01	0.27 ± 0.01	0.28 ± 0.03	0.24 ± 0.05	0.64 ± 0.02	0.21 ± 0.01	0.31 ± 0.05	0.24 ± 0.06	0.29 ± 0.04	0.30 ± 0.04	0.26 ± 0.06	0.33 ± 0.08
	M_c	1 ± 0.01	1.6 ± 0.05	1 ± 0.05	1.5 ± 0.2	1 ± 0.01	1.1 ± 0.11	1 ± 0.01	1 ± 0.05	1.8 ± 0.14	2.9 ± 0.17	2.1 ± 0.37	3.1 ± 0.13
M3	a	3.84	3.83	–	–	3.56	3.84	–	–	5.37	6.02	–	–
	b	0.66 ± 0.04	0.49 ± 0.01	–	–	0.62 ± 0.02	0.49 ± 0.01	–	–	0.84 ± 0.05	0.82 ± 0.01	–	–
	M_c	1.6 ± 0.09	3.2 ± 0.48	–	–	1 ± 0.07	3.4 ± 0.33	–	–	3.0 ± 0.05	4.8 ± 0.17	–	–
M4	a	–	–	–	–	3.67	–	–	–	–	–	–	–
	b	–	–	–	–	0.66 ± 0.02	–	–	–	–	–	–	–
	M_c	–	–	–	–	1.6 ± 0.04	–	–	–	–	–	–	–
M5	a	3.87	3.7	2.80	–	3.59	3.40	2.29	1.56	5.43	5.37	4.13	–
	b	0.67 ± 0.05	0.45 ± 0.08	0.56 ± 0.18	–	0.63 ± 0.01	0.38 ± 0.08	0.44 ± 0.11	0.27 ± 0.01	0.86 ± 0.07	0.70 ± 0.19	0.72 ± 0.19	–
	M_c	1.6 ± 0.11	2.9 ± 0.55	2.4 ± 0.55	–	1.1 ± 0.24	2.6 ± 0.63	2.0 ± 0.92	2.6 ± 1.9	3.0 ± 0.11	4.4 ± 0.5	3.7 ± 0.37	–
M6	a	3.99	3.45	2.75	–	–	3.30	2.25	–	5.42	4.37	4.01	–
	b	0.72 ± 0.03	0.39 ± 0.05	0.55 ± 0.01	–	–	0.36 ± 0.04	0.45 ± 0.01	–	0.88 ± 0.07	0.50 ± 0.07	0.70 ± 0.20	–
	M_c	1.8 ± 0.17	2.6 ± 0.39	2.4 ± 0.23	–	–	2.6 ± 0.32	2.1 ± 0.22	–	3.1 ± 0.18	3.9 ± 0.3	3.7 ± 0.18	–
M7	a	4.03	3.37	–	–	3.63	3.22	–	–	4.18	4.31	–	–
	b	0.74 ± 0.05	0.38 ± 0.02	–	–	0.64 ± 0.05	0.34 ± 0.03	–	–	0.49 ± 0.11	0.49 ± 0.01	–	–
	M_c	1.9 ± 0.19	2.3 ± 0.19	–	–	1.5 ± 0.16	2.1 ± 0.24	–	–	2.0 ± 0.29	3.7 ± 0.38	–	–
M8	a	4.01	3.36	–	–	3.64	3.16	–	–	4.06	4.61	–	–
	b	0.70 ± 0.06	0.37 ± 0.03	–	–	0.65 ± 0.04	0.32 ± 0.03	–	–	0.46 ± 0.11	0.55 ± 0.01	–	–
	M_c	1.8 ± 0.2	2.3 ± 0.13	–	–	1.5 ± 0.17	2.0 ± 0.28	–	–	1.9 ± 0.19	3.8 ± 0.65	–	–
M9	a	3.45	3.05	–	–	3.42	2.85	–	–	4.17	3.7	–	–
	b	0.47 ± 0.02	0.28 ± 0.01	–	–	0.55 ± 0.02	0.23 ± 0.01	–	–	0.49 ± 0.10	0.36 ± 0.02	–	–
	M_c	1.1 ± 0.03	1.7 ± 0.05	–	–	1.1 ± 0.01	1.2 ± 0.11	–	–	2.1 ± 0.19	3.1 ± 0.12	–	–

Table 2. *Estimation of* λ, β *and* $M_{\max}$ *for three catalogues from all the sources using nine different methods.*

		C1				C2				C3			
		DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB
M1	λ	9.93 ± 0.46	10.41 ± 0.66	1.57 ± 0.33		15.33 ± 0.58	9.94 ± 0.64	1.56 ± 0.30	0.30 ± 0.11	105.05 ± 5.5	15.58 ± 1.15	2.32 ± 0.53	
	β	1.80 ± 0.04	1.12 ± 0.05	1.23 ± 0.19		1.55 ± 0.03	0.88 ± 0.03	1.0 ± 0.14	0.62 ± 0.02	2.52 ± 0.05	1.43 ± 0.05	1.62 ± 0.23	
	$M_{\max}$	6.21 ± 0.37	7.0 ± 0.32	6.40 ± 0.36	–	6.90 ± 0.42	6.95 ± 0.30	7.02 ± 0.37	7.45 ± 0.39	6.37 ± 0.31	7.26 ± 0.31	6.51 ± 0.32	–
M2	λ	18.6 ± 0.62	6.26 ± 0.33	1.06 ± 0.15	0.3 ± 0.10	15.33 ± 0.58	6.93 ± 0.36	1.05 ± 0.16	0.46 ± 0.11	20.39 ± 0.70	14.49 ± 0.90	1.77 ± 0.27	0.79 ± 0.23
	β	1.12 ± 0.01	0.67 ± 0.01	0.61 ± 0.05	0.3 ± 0.10	1.55 ± 0.03	0.50 ± 0.01	0.67 ± 0.06	0.62 ± 0.02	0.79 ± 0.02	0.86 ± 0.03	0.44 ± 0.07	0.79 ± 0.23
	$M_{\max}$	6.04 ± 0.30	6.95 ± 0.30	6.26 ± 0.31	6.6 ± 0.32	6.90 ± 0.42	6.92 ± 0.30	6.91 ± 0.32	7.40 ± 0.36	6.32 ± 0.30	7.23 ± 0.30	6.43 ± 0.30	6.67 ± 0.31
M3	λ	11.53 ± 0.50	6.64 ± 0.51			17.76 ± 0.63	5.28 ± 0.48			104.8 ± 5.19	7.40 ± 0.72		
	β	1.73 ± 0.04	1.15 ± 0.02			1.53 ± 0.03	1.12 ± 0.02			2.21 ± 0.05	1.89 ± 0.02		
	$M_{\max}$	6.17 ± 0.34	6.99 ± 0.31	–	–	6.84 ± 0.38	6.99 ± 0.31	–	–	6.37 ± 0.31	7.31 ± 0.32	–	–
M4	λ	–				19.77 ± 0.99							
	β					1.53 ± 0.02							
	$M_{\max}$	–	–	–	–	6.82 ± 0.31	–	–	–	–	–	–	–
M5	λ	11.40 ± 0.5	8.74 ± 0.65	1.57 ± 0.33		12.44 ± 0.53	9.47 ± 0.68	1.56 ± 0.29	0.30 ± 0.11	108.55 ± 5.6	16.77 ± 2.06	2.32 ± 0.53	
	β	1.76 ± 0.05	1.27 ± 0.06	1.23 ± 0.19		1.45 ± 0.03	1.01 ± 0.05	0.99 ± 0.13	0.62 ± 0.02	2.26 ± 0.06	2.14 ± 0.12	1.62 ± 0.23	
	$M_{\max}$	6.19 ± 0.36	7.03 ± 0.33	6.40 ± 0.36	–	6.84 ± 0.38	6.97 ± 0.31	7.07 ± 0.40	7.45 ± 0.39	6.35 ± 0.30	7.36 ± 0.34	6.51 ± 0.32	–
M6	λ	8.33 ± 0.41	9.85 ± 0.65	1.58 ± 0.30			9.23 ± 0.63	1.37 ± 0.27		98.58 ± 5.58	16.24 ± 1.29	2.30 ± 0.53	
	β	1.80 ± 0.04	1.13 ± 0.04	1.26 ± 0.01			0.97 ± 0.04	1.03 ± 0.02		2.36 ± 0.06	1.56 ± 0.06	1.60 ± 0.23	
	$M_{\max}$	6.22 ± 0.37	7.0 ± 0.32	6.42 ± 0.37	–	–	6.96 ± 0.31	7.12 ± 0.44	–	6.36 ± 0.31	7.28 ± 0.31	6.51 ± 0.32	–
M7	λ	6.76 ± 0.39	4.73 ± 0.27			7.88 ± 0.42	11.0 ± 0.67			47.68 ± 2.10	13.97 ± 0.83		
	β	1.91 ± 0.06	1.00 ± 0.03			1.51 ± 0.05	0.85 ± 0.03			0.92 ± 0.03	1.17 ± 0.02		
	$M_{\max}$	6.32 ± 0.44	6.98 ± 0.31	–	–	6.90 ± 0.42	6.95 ± 0.30	–	–	6.30 ± 0.30	7.24 ± 0.30	–	–
M8	λ	6.29 ± 0.42	11.82 ± 0.68			7.88 ± 0.42	11.6 ± 0.68			20.52 ± 0.70	13.83 ± 0.87		
	β	1.89 ± 0.06	1.04 ± 0.03			1.51 ± 0.04	0.83 ± 0.03			0.85 ± 0.03	1.29 ± 0.02		
	$M_{\max}$	6.29 ± 0.42	6.98 ± 0.31	–	–	6.90 ± 0.42	6.95 ± 0.30	–	–	6.31 ± 0.30	7.25 ± 0.30	–	–
M9	λ	17.4 ± 0.61	6.15 ± 0.32			12.99 ± 0.53	6.88 ± 0.36			51.16 ± 2.27	15.15 ± 0.87		
	β	1.24 ± 0.02	0.70 ± 0.01			1.35 ± 0.03	0.54 ± 0.01			1.0 ± 0.03	0.94 ± 0.03		
	$M_{\max}$	6.05 ± 0.30	6.94 ± 0.30	–	–	6.77 ± 0.34	6.93 ± 0.30	–	–	6.31 ± 0.30	7.23 ± 0.30	–	–

Table 3. *Estimation of return periods (RP), probability of occurrence in 50 years (50) and 100 years (100) for three catalogues from all the sources using nine different methods for magnitude 6.0 earthquake.*

		C1				C2				C3			
		DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB	DHR	HIM	MOR	RGB
M1	RP	883	22	470	–	237	10	147	83	225	9	159	–
	50	0.055	0.8943	0.1009	–	0.1902	0.9923	0.2882	0.454	0.1991	0.9968	0.2692	–
	100	0.10705	0.98882	0.19158	–	0.34428	0.99994	0.49335	0.70184	0.35856	0.99999	0.46588	–
M2	RP	362	0	142	119	237	0	62	88	0	0	26	65
	50	0.1289	1	0.2968	0.3441	0.1902	1	0.5512	0.4342	1	1	0.855	0.5342
	100	0.24111	1	0.50550	0.56982	0.34428	1	0.79857	0.67986	1	1	0.97897	0.78305
M3	RP	815	17	–	–	191	16	–	–	117	10	–	–
	50	0.0595	0.9426	–	–	0.23	0.9553	–	–	0.3474	0.9945	–	–
	100	0.11551	0.99671	–		0.40715	0.99800	–	–	0.57407	0.99997	–	–
M4	RP	–	–	–	–	173	–	–	–	–	–	–	–
	50	–	–	–	–	0.2505	–	–	–	–	–	–	–
	100	–	–	–	–	0.43828	–	–	–	–	–	–	–
M5	RP	839	28	470	–	157	14	140	83	144	17	159	–
	50	0.0578	0.8276	0.1009	–	0.2724	0.9705	0.2997	0.454	0.2935	0.9506	0.2692	–
	100	0.11230	0.97028	0.19158	–	0.47060	0.99913	0.50955	0.70184	0.50089	0.99756	0.46588	–
M6	RP	856	21	507	–	–	13	164	–	174	9	153	–
	50	0.0567	0.9047	0.094	–	–	0.9816	0.2621	–	0.2495	0.9955	0.2795	–
	100	0.11024	0.99091	0.17909	–	–	0.99966	0.45554	–	0.43677	0.99998	0.48093	–
M7	RP	976	15	–	–	174	10	–	–	9	0	–	–
	50	0.0499	0.9656	–	–	0.2496	0.9916	–	–	0.9968	1	–	–
	100	0.09734	0.99882	–	–	0.43690	0.99993	–	–	0.99999	1	–	–
M8	RP	981	17	–	–	176	10	–	–	9	0	–	–
	50	0.0497	0.9438	–	–	0.2477	0.9937	–	–	0.9968	1	–	–
	100	0.09688	0.99684	–	–	0.43411	0.99996	–	–	0.99999	1	–	–
M9	RP	471	9	–	–	128	0	–	–	10	0	–	–
	50	0.1007	0.9968	–	–	0.3229	1	–	–	0.9937	1	–	–
	100	0.19118	0.99999	–	–	0.54147	1	–	–	0.99996	1	–	–

The probabilistic approach for estimating the maximum regional magnitude $M_{\max}$ was suggested by Kijko and Sellevoll (1989) based on the doubly truncated G-R relationship. It has been further refined by Kijko and Graham (1998) to consider the uncertainties in the input magnitude data. $M_{\max}$ from Kijko-Sellevoll-Bayes estimator is obtained as a solution of following equation, Kijko and Graham (1998)

$$m_{\max} = m_{\max}^{\text{obs}} + \frac{\delta^{1/q+2}\exp[nr^q/(1-r^q)]}{\beta} \times [\Gamma(-1/q, \delta r^q) - \Gamma(-1/q, \delta)], \quad (14)$$

where $p = \overline{\beta}/(\sigma_\beta)^2$, $q = (\overline{\beta}/\sigma_\beta)^2$ where $\beta = 2.303b$, $\overline{\beta}$ denotes the mean value of β, σ_β is the standard deviation of β and C_β is a normalizing coefficient and which is equal to $\{1-[p/(p+m_{\max}-m_{\min})]^q\}^{-1}$, $r = p/(p + m_{\max} - m_{\min})$, $c_1 = \exp[-n(1 - C_\beta)]$, $\delta = nC_\beta$ and $\Gamma(\cdot,\cdot)$ is the Incomplete Gamma Function. $M_{\max}$ is obtained by iterative solution of equation (14). The results showing the values of λ, β and $M_{\max}$ are given in table 2. The recurrence period and probability of occurrence of magnitude 6.0 earthquakes in 50 years and 100 years in the respective source zones are shown in table 3 for all the three catalogues.

6. Discussions and conclusions

In the present study an endeavour has been made to look for the sources of uncertainties in the process to estimate the $M_{\max}$. Uncertainty is associated with each magnitude scale which plays an important role while converting one scale magnitude into another scale. The relationship between different magnitude scales is of paramount importance for conversions to carry out homogenization of seismicity catalogues for further studies related to seismological statistics and for seismicity and amplitude studies. An endeavour has been made to quantify the uncertainties due to magnitude conversions to be used for seismic hazard analysis. Three catalogues C1, C2 and C3 have been prepared and the seismicity parameters are estimated to look for the effect of the conversions. These parameters are estimated using nine different methods to include the uncertainties. Table 1 shows the variation in the values of a, b and M_c estimated using nine methods and three catalogues. Since the data were less for some of the sources, some of the values could not be estimated (shown as – in the tables). Based on the a, b and M_c the seismicity parameters estimated are tabulated in table 2. The $M_{\max}$ value ranges from 6.04 ± 0.30 (M2) to 6.32 ± 0.44 (M7), 6.77 ± 0.34 (M9) to 6.90 ± 0.42 (M1) and 6.30 ± 0.30 (M7) to 6.37 ± 0.31 (M1) for DHR using the three catalogues C1, C2 and C3, respectively. The variation is much more for a and b values in comparison to the $M_{\max}$ value. The errors shown emphasize the consideration of uncertainties in the estimation of seismic hazard. The epistemic uncertainties can be taken care by the logic tree approach. The final variations in the values have been shown in the recurrence periods and the probabilities of occurrence of different magnitudes. An example of magnitude 6.0 earthquake is taken for the present case. The recurrence period of magnitude 6.0 and the probability of occurrence in 50 and 100 years in the four source zones is shown for the nine methods and the three catalogues. The variations emphasize the consideration of uncertainties while carrying out seismic hazard assessment which in turn has bearing on the seismic microzonation of an area.

References

Aki K 1965 Maximum likelihood estimate of b in the formula log $N = a - bM$ and its confidence limits; *Bull. Earthq. Res. Inst. Tokyo Univ.* **43** 237–239.

Ameer S, Sharma M L, Wason H R and AlSinawi A 2005 Conversion of magnitudes for assessment of recurrence time in revised "Iraq seismicity catalogue"; Symposium on Seismic Hazard analysis and Microzonation, Roorkee India, September 23–24, 2005, **1** 201–211.

Barnett V D 1967 A note on linear structural relationships when both residual variances are known; *Biometika* **54** 670–672.

Bath M 1968 Handbook on earthquake magnitude determination, 2nd edn., Seismological Institute, Uppsala, 131.

Bommer J J, Scherbaum F, Bungum H, Cotton F, Sabetta F and Abrahamson N A 2005 On the use of logic trees for ground-motion prediction equations in seismic hazard analysis; *Bull. Seismol. Soc. Am.* **95(2)** 377–389.

Cao A M and Gao S S 2002 Temporal variation of seismic b-values beneath northeastern Japan island arc; *Geophys. Res. Lett.* **29(9)** 1–3.

Draper N R and Smith H 2005 Applied regression analysis, John Wiley & Sons (Asia) Pvt. Ltd.

Gibowicz S J 1972 The relationship between teleseismic body wave magnitude and local magnitude from New Zealand earthquakes; *Bull. Seismol. Soc. Am.* **62** 1–10.

Gutenberg B and Richter C F 1956a Earthquake magnitude, intensity, energy and acceleration; *Bull. Seismol. Soc. Am.* **46** 105–143.

Gutenberg B and Richter C F 1956b Magnitude and Energy of Earthquakes; *Amm. Geofis.* **9** 1–15.

Joshi G C and Sharma M L 2006 Magnitude Scale conversion relationships for Northern Indian Region using bivariate analysis; 13th Symposium on Earthquake Engineering, Indian Institute of Technology, Roorkee, December 18–20, 2006. Paper 195.

Kijko A and Graham G 1998 "Parametric historic procedure for probabilistic seismic analysis, Part I: Estimation of maximum regional magnitude $M_{\max}$"; *Pure Appl. Geophys.* **152** 413–442.

Kijko A and Sellevol 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. Seismol. Soc. Am.* **40** 1–7.

Kijko A and Sellevoll M A 2001 "Computer program HN2, RELEASE 2.08, 2001", Seismology Unit; URL: http://www.geoscience.org.za/seismo/index.htm.

Marshal P D 1970 Spectral differences between earthquakes and underground explosions; *Geophys. J-R Astr. Soc.* **20** 397–416.

McGuire R K 2004 Seismic Hazard and Risk Analysis; Risk Engineering, Inc. Boulder, Colorado, EERI.

McGuire R K and Shedlock K M 1981 Statistical uncertainties in seismic hazard evaluations in the United States; *Bull. Seismol. Soc. Am.* **71** 1287–1308.

Ogata Y and Katsura K 1993 Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogs; *Geophys. J. Int.* **113** 727–738.

Prozorov and Hudson J A 1974 A study of magnitude difference MS–mb for the earthquakes; *Geophys. J-R Astr. Soc.* **39** 551–564.

Richter C F 1935 An instrumental earthquake magnitude scale; *Bull. Seismol. Soc. Am.* **25** 1–32.

Shi Y and Bolt B A 1982 The standard error of the magnitude-frequency b-value; *Bull. Seismol. Soc. Am.* **72** 1667–1687.

Shrama M L, Wason H R and Dimri R 2003 "Seismic zonation of the Delhi region for Bedrock ground motion"; *Pure Appl. Geophys.* **160** 2381–2398.

Sprent D and Dolby G R 1980 The geometric mean relationships; *Biometrics* **36** 547–550.

Suckale J, Grünthal G and Regnier M C 2005 "Probabilistic Seismic Hazard Assessment for Vanuatu", Scientific Technical Report STR 05/16 GeoForschungsZentrum, Potsdam.

Tavakoli B 2002 Sensitivity of seismic hazard evaluation to uncertainties determined from seismic source characterization; *J. Seismol.* **6** 525–545.

Utsu T 1965 A method for determining the value of b in a formula $\log n = a - bm$ showing the magnitude frequency for earthquakes; *Geophys. Bull. Hokkaido Univ.* **13** 99–103.

Utsu T 2002a Statiscal Features of Seismicity; *International Handbook of Earthquake and Engineering Seismology* **81A** 719–732.

Utsu T 2002b Relationship between magnitude scales; *International Handbook of Earthquake Engineering and Engineering Seismology* **81A** 733–795.

Wang M Y 1989 Likelihood estimation of a simple linear regression model when both variables have error; *Biometrika* **76** 141–148.

Wiemer S and Wyss M 2000 Minimum magnitude of complete reporting in earthquake catalogues: example from Alaska, the Western United States, and Japan; *Bull. Seismol. Soc. Am.* **90** 859–869.

Woessner J and Wiemer S 2005 Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty; *Bull. Seismol. Soc. Am.* **95(2)** 684–698.

MS received 19 September 2007; revised 5 April 2008; accepted 8 April 2008