Statistical Tools for Maximum Possible Earthquake Magnitude Estimation

Andrzej KIJKO[1] and Mayshree SINGH[2]

[1]Aon-Benfield Natural Hazard Centre, Pretoria University,
Pretoria, South Africa; e-mail: andrzej.kijko@up.a.za

[2]School of Civil Engineering, Surveying and Construction,
University of KwaZulu Natal, Durban, South Africa
e-mail: singhm5@ukzn.ac.za

A b s t r a c t

Several procedures for the statistical estimation of the region-characteristic maximum possible earthquake magnitude, m_{max}, are currently available. This paper aims to introduce and compare the 12 existing procedures. For each of the procedures given, there are notes on its origin, assumptions made in its derivation, condition for validity, weak and strong points, *etc*. The applicability of each particular procedure is determined by the assumptions of the model and/or the available information on seismicity of the area.

Key words: m_{max}, earthquake magnitude distribution, maximum magnitude.

1. INTRODUCTION

The region-characteristic, maximum possible earthquake magnitude, m_{max}, is required by the earthquake engineering community, disaster management agencies and the insurance industry. Although the m_{max} value is important, it is astonishing how little has been done in developing appropriate techniques for estimating this parameter. Presently, there is no universally accepted technique for estimating the value of m_{max}; however, the current procedures for m_{max} can be divided into two main categories: deterministic and probabil-

istic. A presentation and discussion of deterministic techniques for the assessment of m_{max} can be found in, *e.g.*, Wells and Coppersmith (1994), Wheeler (2009), and Mueller (2010). This paper focuses only on the probabilistic techniques.

The maximum regional earthquake magnitude, m_{max}, is defined as the upper limit of earthquake magnitude for a given region and is synonymous with the magnitude of the largest possible earthquake. This definition is used by earthquake engineers (EERI Committee on Seismic Risk 1984) and seismologists (Hamilton 1967, Page 1968, Cosentino *et al*. 1977, Working Group on California Earthquake Probabilities (WGCEP) 1995, Stein and Hanks 1998, Field *et al*. 1999). It assumes 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}.

This paper presents several statistical techniques for the evaluation of m_{max}, which can be used depending on the assumptions about the statistical distribution model and/or the information available about past seismicity. Certain procedures can be applied in extreme cases when no information about the nature of the earthquake magnitude distribution is available. These procedures are capable of generating an equation for m_{max}, which is independent of the assumed frequency-magnitude distribution. Some of the procedures can also be used when the earthquake catalogue is incomplete, *i.e.*, when only a limited number of the largest magnitudes are available.

The described procedures are available in a MATLAB toolbox called MMAX. This toolbox can be obtained from the authors free of charge.

2. THEORETICAL BACKGROUND

The methodology assumes that in the area of concern, within a specified time interval, T, all n of the main earthquakes that occurred with a magnitude greater than or equal to m_{min}, are recorded. The largest observed earthquake magnitude in the area is denoted as m_{max}^{obs}. Next assume that the value of the magnitude m_{min} is known and is denoted as the threshold of completeness. It is further assumed that the magnitudes are independent, identically distributed, random values with probability density function (PDF), $f_M(m)$, and the cumulative distribution function (CDF), $F_M(m)$. The unknown parameter m_{max} is the upper limit of the range of magnitudes and is thus termed the maximum regional earthquake magnitude, $\hat{m}_{max}$, that is to be estimated.

The estimation techniques are organized in three sections: 2.1 Parametric estimators, 2.2 Non-parametric estimators, and 2.3 Fit of CDF of earthquake magnitudes.

2.1 Parametric estimators

Parametric estimators can be used when the parametric models of the frequency-magnitude distributions are known. Five procedures are described. The first procedure, denoted as T-P, is based on complex mathematical considerations (Tate 1959), but is computationally very straightforward and does not require extensive calculations. The next two, known as K-S procedures, are based on the generic equation derived by Cooke (1979) and they differ only in numerical details. In the derivation of the first procedure, the exact distribution of the largest earthquake magnitude is replaced by its Cramér's approximation. The second K-S procedure is based on exact solution of Cooke's (1979) generic equation; it is therefore able to provide a better solution, particularly when the number of observations is limited. Finally, the last two procedures are especially useful when only a rough knowledge of the functional form of earthquake magnitude distribution is available. In all five procedures, the largest observed earthquake magnitude, $m_{\max}^{obs}$, plays a crucial role.

All the procedures presented in this section are based on the underlying principle that the estimated $m_{\max}$ value is equal to $m_{\max}^{obs} + \Delta$, where Δ is a positive correction factor. This principle is similar to the popular deterministic procedure, where the increment Δ varies from 0.25 to 1.0 of a magnitude unit (Wheeler 2009). Despite this similarity, there is a fundamental difference between the two approaches. In the deterministic approach, Wheeler's correction factor is a pure deterministic number, essentially just a guess; however, in the probabilistic approach the correction factor is determined by factors characterizing the seismicity of the area. The correction factor is a function of $m_{\max}^{obs}$ the seismic activity rate, and the ratio between the number of weak and strong events. The correction factor depends on seismic parameters supporting intuitive expectations that it is always positive and its value decreases as the time span of observation increases.

Sections 2.1.1 to 2.1.5 provide a detailed description of the five procedures introduced above.

2.1.1 Tate–Pisarenko procedure

This procedure is very straightforward and does not require extensive calculations. It can be shown that the procedure attempts to correct the bias of the classical maximum likelihood estimator (Pisarenko *et al.* 1996), but it fails to provide an estimator having a smaller mean-squared error.

After the transformation $y = F_M(m)$, the CDF of the largest among $(Y_1, \ldots, Y_n)$, that is Y_n, is equal to y^n, and its expected value $E(Y_n) = n/(n + 1)$. One of the possibilities for obtaining the estimator of $\hat{m}_{\max}$ is to introduce

the condition $E(Y_n) = y_n$, from which we obtain the following equation (Gibowicz and Kijko 1994):

$$F_M(m_{\max}^{obs}) = \frac{n}{n+1} \ . \tag{1}$$

Thus, the estimator of $m_{\max}$ becomes a function of $m_{\max}^{obs}$ and n, and is obtained as a root of eq. (1). The above result is valid for any CDF of earthquake magnitude, $F_M(m)$, and that does not require the fulfilment of the truncation condition. From eq. (1), Kijko and Graham (1998) derived an alternative equation, which is approximate, but which demonstrates the required value of $m_{\max}$ in a more explicit way (Pisarenko *et al.* 1996)

$$m_{\max} = m_{\max}^{obs} + \frac{1}{n \, f_M(m_{\max}^{obs})} \ . \tag{2}$$

It should be noted that in eq. (2), the desired $m_{\max}$ appears on both sides: left, simply as $m_{\max}$, and on the right side as the unknown parameter of the probability density function (PDF) of earthquake magnitude, $f_M(m)$. However, from this equation an estimated value of $m_{\max}$, $\hat{m}_{\max}$, can be obtained through iteration.

The estimator (2) was probably first derived by Tate (1959). If applied to the Gutenberg–Richter magnitude distribution with PDF (Page 1968),

$$f_M(m) = \begin{cases} 0 & \text{for} \quad m < m_{\min} \\ \dfrac{\beta \exp[-\beta(m - m_{\min})]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} & \text{for} \quad m_{\min} \le m \le m_{\max} \ , \\ 0 & \text{for} \quad m > m_{\max} \end{cases} \tag{3}$$

it takes the form

$$m_{\max} = m_{\max}^{obs} + \frac{1 - \exp[-\beta(m_{\max} - m_{\min})]}{n \, \exp[-\beta(m_{\max}^{obs} - m_{\min})]} \ , \tag{4}$$

where $\beta = b \ln(10)$, and b is the parameter of the frequency-magnitude Gutenberg–Richter relation. With small modifications, eq. (4) is equivalent to Tate's (1959) estimator. It was used for the first time by Pisarenko *et al.* (1996). The solution of eq. (4) provides the estimated value of $m_{\max}$, $\hat{m}_{\max}$, and in this paper is referred to as the Tate–Pisarenko estimator of $m_{\max}$ or in short as T-P. The approximate variance of the T-P estimator, which contains both, aleatory and epistemic components, is of the form (Kijko and Graham 1998)

$$\mathrm{Var}(\hat{m}_{\max}) = \sigma_M^2 + \frac{n+1}{n^3}\left[\frac{1 - \exp[-\beta(m_{\max}^{obs} - m_{\min})]}{\beta\exp[-\beta(m_{\max}^{obs} - m_{\min})]}\right]^2, \tag{5}$$

where σ_M denotes the standard error in the determination of the largest observed magnitude, $m_{\max}^{obs}$.

2.1.2 Kijko–Sellevoll procedure (Cramér's approximation)

The largest observed magnitude, m_n, which is also denoted as $m_{\max}^{obs}$, has the cumulative distribution function (CDF)

$$F_{M_n}(m) = \begin{cases} 0 & \text{for} \quad m < m_{\min} \\ \left[F_M(m)\right]^n & \text{for} \quad m_{\min} \leq m \leq m_{\max} \\ 1 & \text{for} \quad m > m_{\max} \end{cases}. \tag{6}$$

After integrating by parts, the expected value of M_n, $E(M_n)$, is

$$E(M_n) = \int_{m_{\min}}^{m_{\max}} m\,\mathrm{d}F_{M_n}(m) = m_{\max} - \int_{m_{\min}}^{m_{\max}} F_{M_n}(m)\,\mathrm{d}m \ . \tag{7}$$

Hence

$$m_{\max} = E(M_n) + \int_{m_{\min}}^{m_{\max}} \left[F_M(m)\right]^n \mathrm{d}m \ . \tag{8}$$

This expression, after replacement of the expected value of the largest observed magnitude, $E(M_n)$, by the largest magnitude already observed, provides the equation

$$m_{\max} = m_{\max}^{obs} + \int_{m_{\min}}^{m_{\max}} \left[F_M(m)\right]^n \mathrm{d}m \ , \tag{9}$$

in which the desired $m_{\max}$ appears on both sides. An estimated value of $m_{\max}$, $\hat{m}_{\max}$, can be obtained through iteration.

Cooke (1979) was probably the first to obtain the estimator of the upper bound of a random variable similar to eq. (9). The difference between eq. (9) and estimator by Cooke (1979) is that the former provides an equation in which the upper limit of integration is $m_{\max}^{obs}$, not $m_{\max}$. Therefore, for large n, the two solutions are asymptotically equivalent. If Cooke's estimator is applied to the assessment of $m_{\max}$, eq. (9) states that the maximum regional earthquake magnitude, $m_{\max}$, is equal to the largest observed magnitude, $m_{\max}^{obs}$, increased by an amount $\Delta = \int_{m_{\min}}^{m_{\max}} \left[F_M(m)\right]^n \mathrm{d}m$. Similar to eq. (2),

eq. (9) is, by its nature, very general and is valid for each CDF, $F_M(m)$. The drawback of the formula is that it requires integration.

For the frequency-magnitude Gutenberg–Richter relation, the respective CDF of magnitudes which are bounded from above by m_{max}, is (Page 1968)

$$F_M(m) = \begin{cases} 0 & \text{for} \quad m < m_{min} \\ \dfrac{1 - \exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} & \text{for} \quad m_{min} \le m \le m_{max} \\ 1 & \text{for} \quad m > m_{max} \end{cases} \quad . \tag{10}$$

Following eqs. (9) and (10), the estimator of m_{max} requires the calculation of the integral

$$\Delta = \int_{m_{min}}^{m_{max}} \left[\frac{1 - \exp\left[-\beta\left(m_{max}^{obs} - m_{min}\right)\right]}{1 - \exp\left[-\beta\left(m_{max} - m_{min}\right)\right]} \right]^n \, dm \, , \tag{11}$$

which does not have a simple solution. Its estimate can be obtained through the application of Cramér's approximation. According to Cramér (1961), for large n, the value of $[F_M(m)]^n$ is approximately equal to $\exp\{-n[1 - F_M(m)]\}$. Simple calculations show that after the replacement of $[F_M(m)]^n$ by its Cramér approximation, integral (11) takes the form

$$\Delta = \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min} \exp(-n) \, , \tag{12}$$

where $n_1 = n/\{1 - \exp[-\beta(m_{max} - m_{min})]\}$, $n_2 = n_1 \exp[-\beta(m_{max} - m_{min})]$, and $E_1(\cdot)$ denotes an exponential integral function (Abramowitz and Stegun 1970), defined as $E_1(z) = \int_z^\infty \exp(-\zeta)/\zeta \, d\zeta$. Hence, following eq. (9), for the Gutenberg–Richter frequency-magnitude relation, the estimator of m_{max} is obtained as an iterative solution of the equation

$$m_{max} = m_{max}^{obs} + \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min} \exp(-n) \, . \tag{13}$$

Numerical tests show that when $m_{max} - m_{min} \le 2$, and $n \ge 100$, the parameter m_{max} in n_1 and n_2 can be replaced by m_{max}^{obs}, thus providing an m_{max} estimator which can be obtained without iterations.

Kijko and Sellevoll (1989) introduced eq. (13) and it has subsequently been used in more than 65 countries around the world. The solution is often

termed the Kijko–Sellevoll estimator of m_{max}, or, in short, K-S. The approximate variance of the estimator for the frequency-magnitude Gutenberg–Richter distribution is of the form (Kijko 2004)

$$\text{Var}(\hat{m}_{max}) = \sigma_M^2 + \left[\frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min} \exp(-n) \right]^2 , \qquad (14)$$

where σ_M denotes the standard error in the determination of the largest observed magnitude, m_{max}^{obs}.

It can be shown (Kijko and Graham 1998) that the procedure provides an estimator of m_{max}, which, in terms of mean-squared error, is substantially better than the respective T-P estimator described in Section 2.1.1. The drawback of the current procedure is that it requires integration, which for some distribution functions can be performed only numerically. Fortunately, for the frequency-magnitude Gutenberg–Richter relation, the analytical solution of the integral is available. Another limitation of the above procedure comes from the fact that it is based on Cramér's approximation. Consequently, the procedure is capable of providing a correct approximation of m_{max} only for a large n. In the following section an alternative approach is presented that is based on the exact solution of the generic equation (9). The proposed exact solution is superior to the approximate one, since it is correct for any number of observations.

2.1.3 Kijko–Sellevoll procedure (exact solution)

It can be shown (Dwight 1961), that if n is a positive integer, the integral (11) can be expressed as

$$\Delta = \frac{m_{max} - m_{min} + \dfrac{1}{\beta} \displaystyle\sum_{i=1}^{n} \frac{(-1)^i}{i} \binom{n}{i} \left(1 - \exp\left[-i\beta \left(m_{max} - m_{min}\right)\right]\right)}{\left(1 - \exp\left[-\beta \left(m_{max} - m_{min}\right)\right]\right)^n} . \qquad (15)$$

It follows from eqs. (9) and (11) that the estimator of m_{max} is obtained as a solution of the equation

$$m_{max} = m_{max}^{obs} + \frac{(m_{max} - m_{min}) + \dfrac{1}{\beta} \displaystyle\sum_{i=1}^{n} \frac{(-1)^i}{i} \binom{n}{i} \left(1 - \exp\left[-i\beta \left(m_{max} - m_{min}\right)\right]\right)}{\left(1 - \exp\left[-\beta \left(m_{max} - m_{min}\right)\right]\right)^n} . \qquad (16)$$

Thus, the solution of eq. (16) provides the exact estimator of m_{max}, when the magnitudes are distributed according to the Gutenberg–Richter relation. The

approximate variance of the exact K-S estimator of m_{max} for the Gutenberg–Richter magnitude distribution is of the form

$$\text{Var}(\hat{m}_{max}) = \sigma_M^2 + \left[\frac{m_{max} - m_{min} + \frac{1}{\beta}\sum_{i=1}^{n} \frac{(-1)^i}{i}\binom{n}{i}\left(1 - \exp\left[-i\beta(m_{max} - m_{min})\right]\right)}{\left(1 - \exp\left[-\beta(m_{max} - m_{min})\right]\right)^n} \right]^2 . \quad (17)$$

A significant shortcoming of both the T-P as well as K-S procedures comes from the implicit assumption that: (i) seismic activity remains constant in time, (ii) the selected functional form of magnitude distribution properly describes the observations, and (iii) the parameters of the assumed distribution functions are known without error. The following sections introduce several alternative techniques that are free from the above limitations.

2.1.4 Tate–Pisarenko–Bayes procedure

As a rule, any seismogenic process is composed of temporal trends, cycles, oscillations and pure random fluctuations. When the variation of seismic activity is a random process, the formalism in which the model parameters are treated as random variables provides the most efficient tool in accounting for the uncertainties considered above (*e.g.*, DeGroot 1970). In this section, a Tate–Pisarenko equation (4) for the assessment of the maximum regional earthquake magnitude is presented when the uncertainty of the Gutenberg–Richter parameter b is taken into account.

Following the assumption that the variation of the b-value in the frequency-magnitude Gutenberg–Richter relation may be represented by a Gamma distribution with parameters p and q, the Bayesian (also known as compound) PDF and CDF of earthquake magnitude takes the form (Campbell 1982)

$$f_M(m) = \begin{cases} 0 & \text{for} \quad m < m_{min} \\ \beta C_\beta \left(\dfrac{p}{p + m - m_{min}} \right)^{q+1} & \text{for} \quad m_{min} \le m \le m_{max} , \\ 0 & \text{for} \quad m > m_{max} \end{cases} \quad (18)$$

and

$$F_M(m) = \begin{cases} 0 & \text{for} \quad m < m_{min} \\ C_\beta \left[1 - \left(\dfrac{p}{p + m - m_{min}} \right)^{q} \right] & \text{for} \quad m_{min} \le m \le m_{max} , \\ 1 & \text{for} \quad m > m_{max} \end{cases} \quad (19)$$

where C_β is a normalizing coefficient equal to $\{1-[p/(p+m_{\max}-m_{\min})]^q\}^{-1}$, $p = \overline{\beta}/(\sigma_\beta)^2$ and $q = (\overline{\beta}/\sigma_\beta)^2$. The symbol $\overline{\beta}$ denotes the known mean value of the parameter β, and σ_β is the known standard deviation of β. By its definition, the Bayesian distribution of $F_M(m)$ is the weighted average of the distribution of M, for all possible values of β.

Knowledge of the Bayesian, Gutenberg–Richter PDF of the earthquake magnitude (18) makes it possible to construct the Bayesian version of the T-P estimator of $m_{\max}$. Following the Tate equation (2), the Bayesian version of the T-P equation becomes

$$m_{\max} = m_{\max}^{obs} + \frac{1}{n\overline{\beta}C_\beta}\left(\frac{p}{p+m_{\max}^{obs}-m_{\min}}\right)^{-(q+1)} . \tag{20}$$

The value of $m_{\max}$ obtained from the iterative solution of eq. (20) will be termed the Tate–Pisarenko–Bayes estimator of $m_{\max}$, or, in short, T-P-B. The approximate variance of the T-P-B estimator of $m_{\max}$, for the frequency-magnitude Gutenberg–Richter relation, is of the form

$$\text{Var}(\hat{m}_{\max}) = \sigma_M^2 + \left[\frac{n+1}{n^3}\right]\frac{1}{(n\beta C_\beta)^2}\left(\frac{p}{p+m_{\max}^{obs}-m_{\min}}\right)^{-2(q+1)} . \tag{21}$$

2.1.5 Kijko–Sellevoll–Bayes procedure

Based on the knowledge of the Bayesian, Gutenberg–Richter CDF of earthquake magnitude (19), it is possible to construct the Bayesian version of the K-S estimator of $m_{\max}$. Following the generic equation (9), the estimation of $m_{\max}$ requires calculation of the integral

$$\Delta = (C_\beta)^n \int_{m_{\min}}^{m_{\max}} \left[1-\left(\frac{p}{p+m-m_{\min}}\right)^q\right]^n dm , \tag{22}$$

which, after application of Cramér's approximation, can be expressed as

$$\Delta = \frac{\delta^{1/q}\exp[nr^q/(1-r^q)]}{\beta}\left[\Gamma(-1/q,\delta r^q) - \Gamma(-1/q,\delta)\right] , \tag{23}$$

where $r = p/(p+m_{\max}-m_{\min})$, $c_1 = \exp[-n(1-C_\beta)]$, $\delta = nC_\beta$, and $\Gamma(\cdot,\cdot)$ is the complementary Incomplete Gamma Function. Following eq. (9), the Bayesian version of the K-S equation $m_{\max}$ takes the form

$$\hat{m}_{\max} = m_{\max}^{obs} + \frac{\delta^{1/q}\exp[nr^q/(1-r^q)]}{\beta}\left[\Gamma(-1/q,\delta r^q) - \Gamma(-1/q,\delta)\right]. \tag{24}$$

Again, as in all previous cases, eq. (24) does not provide an estimator for m_{max} since some terms on the right-hand side also contain m_{max}. The estimator of m_{max}, when the uncertainty of the Gutenberg–Richter parameter b is taken into account, can therefore be calculated only by iteration. Equation (24) is denoted as the Kijko–Sellevoll–Bayes estimator of m_{max}, or, in short, K-S-B. An extensive comparison of the performances of K-S and K-S-B estimators is given by Kijko (2004). The approximate variance of the K-S-B estimator of m_{max} for the frequency-magnitude Gutenberg–Richter distribution is of the form

$$\text{Var}(\hat{m}_{max}) = \sigma_M^2 + \left[\frac{\delta^{1/q} \exp[nr^q/(1-r^q)]}{\beta} \Big[\Gamma(-1/q, \delta r^q) - \Gamma(-1/q, \delta) \Big] \right]^2 . \quad (25)$$

2.2 Non-parametric estimators

Many studies of seismicity show that: (i) the empirical distributions of earthquake magnitudes are of bi- or multi-modal character, (ii) the log-frequency-magnitude relation has a strong non-linear component, or that (iii) the presence of "characteristic" events (Schwartz and Coppersmith 1984) is evident. In order to use the generic equation (9) in such cases, the analytical, parametric models of the frequency-magnitude distributions should be replaced by its non-parametric counterpart, which can be done in several ways. In this study, five non-parametric procedures are presented. In Section 2.2.1 the earthquake magnitude distribution is replaced by its approximation constructed as a sum of the kernel functions using sample data. In Section 2.2.2, the magnitude distribution is replaced by its empirical counterpart based on the formalism of order statistics (David 1981). Sections 2.2.3-2.2.5 describe more general techniques that provide an assessment of m_{max} without any assumption about the nature of the magnitude distributions or their empirical counterparts.

2.2.1 Non-parametric with Gaussian kernel procedure

In the case where the empirical distribution of earthquake magnitudes is complex and cannot be described by a simple analytical model, it can be replaced by its approximate non-parametric counterpart. Such a replacement can be done in several ways. In this section, the earthquake magnitude distribution is replaced by its approximation – the sum of data-based kernel functions. The non-parametric, kernel-based approximation of PDF is an approach that deals with the direct summation of certain types of functions using sample data. Given the sample data $m_i, \quad i = 1, ..., n,$ and the kernel function $K(\cdot)$, the kernel estimator $\hat{f}_M(m)$ of an actual, and unknown PDF $f_M(m)$, is

$$\hat{f}_M(m) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{m - m_i}{h}\right) , \qquad (26)$$

where h is a positive smoothing factor (Parzen 1962). The kernel function $K(\cdot)$ is a PDF, symmetric about zero. The specific choice of it is not so important for the performance of the method; many unimodal distribution functions ensure similar efficiencies. In this work, the Gaussian kernel function, $K(\xi) = (2\pi)^{1/2} \exp(-\xi^2/2)$, is used. However, the choice of the smoothing factor, h, is crucial as it affects the trade-off between random and systematic errors. Several procedures exist for the estimation of the value of this parameter, none of them being distinctly better for all varieties of real data (Silverman 1986). In this work, the least-squares cross-validation (Hall 1983, Stone 1984) was used. The details of the procedure are given by Kijko *et al.* (2001).

Following the functional form of a selected kernel and the fact that the data comes from a finite interval $[m_{\min}, m_{\max}]$, the respective CDF of seismic event magnitude takes the form

$$\hat{F}_M(m) = \begin{cases} 0 & \text{for} \quad m < m_{\min} \\[2ex] \dfrac{\sum_{i=1}^{n}\left[\Phi\left(\dfrac{m - m_i}{h}\right) - \Phi\left(\dfrac{m_{\min} - m_i}{h}\right)\right]}{\sum_{i=1}^{n}\left[\Phi\left(\dfrac{m_{\max} - m_i}{h}\right) - \Phi\left(\dfrac{m_{\min} - m_i}{h}\right)\right]} & \text{for} \quad m_{\min} \leq m \leq m_{\max} , \\[3ex] 1 & \text{for} \quad m > m_{\max} \end{cases} \qquad (27)$$

where $\Phi(\xi)$ denotes the standard Gaussian cumulative distribution function.

Therefore, following the generic eq. (9), the estimator of $m_{\max}$ can be obtained through the iterative solution of the equation

$$m_{\max} = m_{\max}^{obs} + \Delta , \qquad (28)$$

where

$$\Delta \cong \int_{m_{\min}}^{m_{\max}} [\hat{F}_M(m)]^n \, dm = \int_{m_{\min}}^{m_{\max}} \left[\frac{\sum_{i=1}^{n}\left[\Phi\left(\dfrac{m - m_i}{h}\right) - \Phi\left(\dfrac{m_{\min} - m_i}{h}\right)\right]}{\sum_{i=1}^{n}\left[\Phi\left(\dfrac{m_{\max} - m_i}{h}\right) - \Phi\left(\dfrac{m_{\min} - m_i}{h}\right)\right]} \right]^n \, dm . \qquad (29)$$

The value of $m_{\max}$ will then be denoted as the non-parametric, Gaussian-based estimator of $m_{\max}$ or, in short, N-P-G. Properties of the above estimator

are discussed in the following section, where an alternative non-parametric assessment of m_{max} is presented. The approximate variance of the N-P-G estimator of $\hat{m}_{max}$ is given by

$$\text{Var}(\hat{m}_{max}) = \sigma_M^2 + \Delta^2 , \tag{30}$$

where the upper limit of integration, m_{max}, is replaced by its estimate, $\hat{m}_{max}$.

Knowledge of the distribution of earthquake magnitudes makes it possible to construct an upper confidence limit for the desired m_{max}. The number z_α must be selected in such a way that the probability $\text{Pr}[F_M(m_{max}^{obs}; m_{max}) < z_\alpha] = \alpha$, where $F_M(m; m_{max}) \equiv F_M(m)$, and a change of notation indicates that m_{max} remains a parameter of the CDF. Following eq. (6), one obtains $\text{Pr}\{[F_M(m_{max}^{obs}; m_{max})]^n < z_\alpha^n\} = \alpha$, from which $z_\alpha = \alpha^{1/n}$. Hence, after transformations, a $100(1-\alpha)\%$ upper limit on $\hat{m}_{max}$ is provided by $\text{Pr}[m_{max} < F_M^{-1}(m_{max}^{obs}; \alpha^{1/n})] = 1 - \alpha$, where $F_M^{-1}(m)$ denotes inverse of CDF, $F_M(m)$. The above formalism can be applied to assess the upper confidence limit of m_{max} for all the procedures discussed above.

2.2.2 Non-parametric procedure based on order statistics

When the analytical form of the magnitude distribution is not known, it can be replaced by empirical distributions and the formalism of the order statistics can be employed. For n earthquake magnitudes arranged in increasing order, *i.e.*, $m_1 \leq m_2 \leq \ldots \leq m_{n-1} \leq m_n$, any empirical distribution function $\hat{F}_M(m)$ can be approximated as (Cooke 1979)

$$\hat{F}_M(m) = \begin{cases} 0 & \text{for} & m < m_1 \\ \dfrac{i}{n} & \text{for} & m_i \leq m \leq m_{i+1}, \; i = 1, \ldots, n-1 \, . \\ 1 & \text{for} & m > m_n \end{cases} \tag{31}$$

The approximate value of integral Δ is then

$$\Delta \equiv \int_{m_{min}}^{m_{max}^{obs}} \left[\hat{F}_M(m) \right]^n dm = \sum_{i=1}^{n-1} \left(\frac{i}{n} \right)^n (m_{i+1} - m_i)$$

$$= m_{max}^{obs} - \sum_{i=0}^{n-1} \left[\left(1 - \frac{i}{n} \right)^n - \left(1 - \frac{i+1}{n} \right)^n \right] m_{n-i} \, . \tag{32}$$

Since for large n, the value of $(1+1/n)^n \cong e$, the correction factor $\Delta = m_{max}^{obs} - (1-e^{-1})\sum_{i=0}^{n-1} e^{-i} m_{n-i}$, and the order statistics-based estimator of m_{max} takes the form

$$\hat{m}_{max} = m_{max}^{obs} + \left(m_{max}^{obs} - (1-e^{-1})\sum_{i=0}^{n-1} e^{-i} m_{n-i} \right) . \tag{33}$$

The value of m_{max} obtained from eq. (33) is denoted as the non-parametric, order statistics-based estimator or, in short, N-P-OS. Assuming that the standard error in the determination of magnitude $m_1, \ldots, m_n$ is known and equal to σ_M, the approximate variance of the order statistics-based estimator (33) is equal to

$$\mathrm{Var}(\hat{m}_{max}) = c_0 \sigma_M^2 + \Delta^2 \quad , \tag{34}$$

where $c_0 = (1+e^{-1})^2 + e^{-2}(1-e^{-1})/(1+e^{-1}) \cong 1.93$.

The confidence limits for the N-P-OS estimator (33) can be constructed in several ways. For most applications, the most attractive equation was derived by Cooke (1979), which is simple and general. Following the assumption that the number of observations n is large, the approximate $100(1-\alpha)\%$ upper confidence limit for $\hat{m}_{max}$ is given by the formula

$$\Pr\left[m_{max} < m_{max}^{obs} + \frac{m_{max}^{obs} - m_{n-1}}{(1-\alpha)^{-v} - 1} \right] \cong 1 - \alpha \quad , \tag{35}$$

where the parameter v is defined by the equation

$$\lim_{m \to 0^-} \frac{1 - F_M(cm + m_{max})}{1 - F_M(m + m_{max})} = \mathrm{const}^{1/v} \quad . \tag{36}$$

Equation (36) determines the shape of the upper tail of the earthquake magnitude distribution. If the PDF of earthquake magnitude is truncated at m_{max}, (as, *e.g.*, Gutenberg–Richter based PDF, eq. (11)), the parameter v is equal to 1.

The estimator N-P-OS (eq. (33)) is very useful since it can be used when both the functional form of the PDF of earthquake magnitude, $f_M(m)$, and the value of v, describing the behavior of its upper tail are not known. In addition, it does not require the assumption that the upper bound of the magnitude distribution is obtained as a result of truncation. Also, for large n, the knowledge of the functional form of the magnitude distribution is not impor-

tant: the estimator with the distribution of magnitudes approximated by order statistics is as good as the estimator where the exact functional form of the magnitude distribution is applied.

Both of the non-parametric estimators, eqs. (28)-(29) and (33), are very useful. The great attraction of the non-parametric approach is that it does not require one to specify the functional form for the magnitude distribution $F_M(m)$. Therefore, by its nature, it is able to deal with cases with empirical distributions of any complexity: distributions which considerably violate log-linearity, multimodality, and/or account for presence of the "characteristic" earthquakes. The drawback of the estimators is that formally they require knowledge of all events with magnitude above the specified level of completeness m_{min}, though, in practice, this can reduce to the knowledge of a few (*e.g.*, 5) largest events. Such a reduction is possible since the contribution of weak events to the correction factor Δ is insignificant.

2.2.3 Procedure based on a few largest earthquakes

In this section, a simple formula for estimation of the maximum regional magnitude is given which can be applied in the case when no information on analytical form of the magnitude distribution is available and only several largest earthquake magnitudes are available.

In the language of mathematical statistics, the case when a known number of observations are missing from the end of the distribution is known as (single) data censoring (David 1981). The problem of estimating the bounds of random variables when only the m largest (or smallest) observations are available, has been discussed by Cooke (1980). Theoretical results expressed in terms of the problem of determining the maximum regional earthquake magnitude can be summarized as follows.

Suppose that in the area of concern, within a specified time interval, T, from n main seismic events, which occurred with magnitudes $m_i \geq m_{min}$ ($i = 1, \ldots, n$), only the n_0 largest magnitudes are known. Following Gnedenko's condition (Gnedenko 1943), which suggests that for a very broad class of cumulative distribution functions, $F_M(m)$, when m is near to its upper endpoint and $F_M(m)$ is linear in m, it can be assumed that the estimator of m_{max} can be expressed as linear functions of the order statistics

$$\hat{m}_{max} = \sum_{i=1}^{n_0} a_i m_{n-i+1} \ , \qquad (37)$$

where $a_i (i = 1, \ldots, n_0)$ are the coefficients to be determined; namely the distribution of earthquake magnitudes is truncated from the top. Cooke (1980) has found that for truncated distributions the minimization of the mean

squared error of estimator (37) can be obtained when $a_1 = 1 + 1/n_0$, $a_2 = ... = a_{n_0-1} = 0$, and $a_{n_0} = -1/n_0$, that is, $\Delta = \left(m_{max}^{obs} - m_{n-n_0+1}\right)\big/n_0$, and

$$\hat{m}_{max} = m_{max}^{obs} + \frac{1}{n_0}\left(m_{max}^{obs} - m_{n-n_0+1}\right) . \tag{38}$$

The greatest attraction of the estimator (38) is most likely the fact that it relies on its simplicity and that even for a small number of observations n_0 the estimator is nearly optimal (in terms of its mean squared errors). This is a result of the fact that for large n, the few largest observations carry most of the information about its endpoint. It is interesting to note that the value of m_{max}, estimated according to eq. (38), is based only on two observations: the n_0^{th} largest magnitude, m_{n-n_0+1}, and the largest observed magnitude, m_{max}^{obs}. A better estimator of m_{max} can therefore be obtained by including the remaining $n_0 - 2$ largest observations. This can be done by the application of the Quenouille's technique, originally developed for the averaging of the bias of an estimator (Quenouille 1956). An averaging correcting factor, $\Delta = \left(m_{max}^{obs} - m_{n-i+1}\right)\big/n_0$, over the $n_0 - 1$ possible choices of i, produces

$$\Delta = \frac{1}{n_0}\left(m_{max}^{obs} - \frac{1}{n_0-1}\sum_{i=2}^{n_0} m_{n-i+1}\right) , \tag{39}$$

and the estimator of m_{max} equal to

$$\hat{m}_{max} = m_{max}^{obs} + \frac{1}{n_0}\left(m_{max}^{obs} - \frac{1}{n_0-1}\sum_{i=2}^{n_0} m_{n-i+1}\right) . \tag{40}$$

Assuming that the standard errors in the determination of the n_0 largest earthquake magnitudes are the same and are equal to σ_M, the approximate variance of the estimator (40) is

$$\text{Var}(\hat{m}_{max}) = c_0 \sigma_M^2 + \Delta^2 , \tag{41}$$

where $c_0 = (n_0^2 + n_0 - 1)\big/[n_0(n_0 - 1)]$. The greatest attraction of estimator (40) lies in its simplicity and that it requires knowledge of the magnitudes of only a few largest events.

2.2.4 Robson–Whitlock procedure

As with the previous three cases, let us assume that the analytical form of the magnitude distribution is not known and we wish to estimate the right end point of the distribution, namely the maximum earthquake magnitude, m_{max}.

One of the methods to solve this problem is to apply the classic Quenouille (1956) technique of successive bias reduction, modified to fit the factorial series rather than the power series in $1/n$. Robson and Whitlock (1964) showed that, under very general conditions, and when the data are arranged in ascending order of magnitude, namely $m_1 \leq m_2 \leq ... \leq m_{n-1} \leq m_{max}^{obs}$, Quenouille's approach leads to the following rule in estimation of m_{max}:

$$\hat{m}_{max} = m_{max}^{obs} + (m_{max}^{obs} - m_{n-1}) \ . \tag{42}$$

Equation (42) was probably first derived by Robson and Whitlock (1964), and is often called the Robson and Whitlock (R-W) estimator. It can be shown that the above estimator is mean-unbiased to order n^{-2} and asymptotically median-unbiased.

The simplicity of eq. (42) makes it very attractive. It can be applied in cases of limited and/or doubtful seismic data, when quick results, without going into sophisticated analysis, is required. Unfortunately, the reduction of bias of the R-W estimator can be achieved only at the expense of a high value of its mean squared error. In fact, Robson and Whitlock (1964) derived a general formula for an unbiased estimator of truncation point

$$\hat{m}_{max} = \sum_{j=0}^{k} (-1)^j \binom{k+1}{j+1} m_{n-j} \ , \tag{43}$$

where $k = 1, ..., n-1$. Regrettably, this formula does not provide a guarantee that the estimated magnitude, $\hat{m}_{max}$, is equal to, or exceeds, the observed maximum magnitude, m_{max}^{obs}. The approximate variance of the R-W estimator of m_{max} for the frequency-magnitude Gutenberg–Richter distribution is of the form

$$\text{Var}(\hat{m}_{max}) = 5\sigma_M^2 + (m_{max}^{obs} - m_{n-1})^2 \ , \tag{44}$$

where σ_M denotes standard error in the determination of the two largest observed magnitudes, m_{max}^{obs} and m_{n-1}.

In their seminal work Robson and Whitlock (1964) also derived a formula for an approximate $100(1-\alpha)\%$ upper confidence limit for m_{max}, which is given as

$$\Pr\left[m_{max} < m_{max}^{obs} + \frac{1-\alpha}{\alpha}(m_{max}^{obs} - m_{n-1}) \right] \cong 1-\alpha \ . \tag{45}$$

Robson and Whitlock (1964) proved that the approximation (45) becomes exact solution when $f_M(m)$ is uniform. If $f_M(m)$ is an increasing (decreasing) function then the true confidence level is greater (less) than $1-\alpha$,

but for any PDF which is continuous and positive at $m = m_{max}$ the asymptotic confidence level is exactly $1 - \alpha$.

It is interesting to note that making a comparison of the non-parametric estimators N-P-OS and R-W, Cooke (1979) showed that for small values of v, the asymptotical efficiency of the estimator N-P-OS is only slightly lower than the corresponding solution R-W. However, for values of v close to 1, asymptotical efficiency of the estimator N-P-OS is significantly (43%) higher.

The R-W estimator (42) was derived on the basis that nothing is known about the magnitude distribution. If the shape of the tail of the PDF, $f_M(m)$, is known, even though the functional form of the distribution is unknown, a significantly better estimator can be obtained using the methodology considered in the following section.

2.2.5 Robson–Whitlock–Cooke procedure

Cooke (1979) showed that reduction of the mean squared error of the R-W estimator is possible when some information about the shape of the upper tail of the probability distribution function $f_M(m)$ is known. Assuming that the observed magnitudes are sampled from a distribution with a known value of parameter v, eq. (36), the improved version of the R-W estimator (42) is

$$\hat{m}_{max} = m_{max}^{obs} + (2v)^{-1}(m_{max}^{obs} - m_{n-1}) \ . \tag{46}$$

The estimator derived in eq. (46) will be called the Robson–Whitlock–Cooke estimator, or in short, R-W-C. Therefore, for certain earthquake magnitude distributions, the estimator R-W-C can perform better than its counterpart R-W. This situation seems likely to occur if, for instance, the magnitudes are sampled from distributions with a known parameter v, or when the distribution is truncated at m_{max} (as is the Gutenberg–Richter-based PDF – eq. (10)), so that v equals to 1. For truncated magnitude distribution, the improved R-W estimator (46) takes the form

$$\hat{m}_{max} = m_{max}^{obs} + 0.5(m_{max}^{obs} - m_{n-1}) \ , \tag{47}$$

and its variance is

$$\text{Var}(\hat{m}_{max}) = 0.5\,[3\sigma_M^2 + 0.5(m_{max}^{obs} - m_{n-1})^2]. \tag{48}$$

2.3 Fit of CDF of earthquake magnitudes

The next two procedures (Sections 2.3.1 and 2.3.2) have one characteristic in common: they are based on the fit of a CDF for earthquake magnitudes. The first procedure, based on L_1-norm regression analysis, is especially useful when the data are unreliable, contain significant outliers, come from different sources and are a mixture of uncertain historic and recent instrumental

observations. In these cases, the L_1-norm regression analysis is superior to any alternative norm such as L_2, which is equivalent to the classic least-squares procedure. It can be shown (Gentle 1977, Anderson 1982) that the applied L_1 norm is robust in the sense that the estimated parameters are insensitive to large outliers that may be present in the input data.

The second procedure is based on the L_2-norm regression analysis which is equivalent to the popular, classical least-squares procedure. Important facts to remember are that the use of the least-squares regression automatically assumes that the distribution of CDF residuals is of Gaussian (normal) nature, and that the application of the least squares regression, when the residuals are not normally distributed, leads to incorrect estimation of the required parameters (Kijko 1994).

The disadvantage of both procedures is that they cannot be used if the functional form of $F_M(m)$ is unknown.

2.3.1 Procedure based on L_1-norm fit of CDF of earthquake magnitudes

The L_1-norm regression analysis, as applied to the problem of m_{max} determination, is defined as follows: Given the sample of main earthquake magnitudes m_i, $i = 1, ..., n$, all magnitudes are greater than or equal to m_{min}, where the value of the magnitude m_{min} is known as the threshold of completeness. The assumption is made that the magnitudes m_i are independent, identically distributed, random values with a known functional form of CDF, $F_M(m)$. A second assumption is that magnitudes are arranged in increasing order, that is, $m_1 \leq m_2 \leq \ldots \leq m_{n-1} \leq m_n$. Let $\boldsymbol{\theta}$ denote a vector of the unknown parameters of the CDF $F_M(m)$, where one of the parameters is the maximum regional earthquake magnitude, m_{max}. If, for example, the classical frequency-magnitude Gutenberg–Richter relation is used (eq. (10)), the vector is defined as $\boldsymbol{\theta} = (\beta, m_{max})$. The aim is to find the solution vector, $\boldsymbol{\theta}$, denoted as $\hat{\theta}$, that minimizes the misfit function $\Phi(\boldsymbol{\theta})$,

$$\Phi(\boldsymbol{\theta}) = \sum_{i=1}^{n} | F_M(m_i) - \hat{F}_M(m_i) | \, , \tag{49}$$

where $|\cdot|$ denotes the absolute value, and $\hat{F}_M(m_i)$ is the empirical distribution function equal to $i/(n+1)$. The misfit criterion (49) is less sensitive to the outlying observations than the classical least-squares procedure and tends to decrease the effect of large differences between the theoretical $F_M(m_i)$ and the empirical $\hat{F}_M(m_i)$ distribution functions. In the MMAX Toolbox, the minimization of the misfit function (49) is accomplished by a derivative-free Nelder–Mead simplex-based procedure (Press *et al.* 1994).

2.3.2 Procedure based on L_2-norm fit of CDF of earthquake magnitudes

Conceptually this procedure is the same as the previous one, except that the absolute values of CDF residuals, eq. (49), are replaced by respective residuals taken to the power 2. Therefore, the respective misfit function is of the form

$$\Phi(\mathbf{\theta}) = \sum_{i=1}^{n} [F_M(m_i) - \hat{F}_M(m_i)]^2 \ , \tag{50}$$

which is equivalent to the classical least-squares procedure. The use of the least-squares technique is equivalent to the assumption that the distribution of the CDF residuals is of Gaussian nature.

3. EXAMPLE OF APPLICATION: ESTIMATION OF THE m_{max} IN THE VICINITY OF CAPE TOWN, SOUTH AFRICA

The city of Cape Town is located in the Western Cape Province, and is considered one of the regions with the highest level of seismic activity of tectonic origin in South Africa. The city is located only some few kilometers from the epicenter of the Tulbagh–Ceres event, the most destructive earthquake in the modern history of South Africa. The earthquake occurred on 29 September 1969, had a local magnitude, M_L, of 6.3, caused nine deaths and the damage to buildings in the epicentral area was estimated at that time at US \$24 million (Kijko *et al.* 2002).

The location of this city was selected in order to illustrate how the MMAX toolbox could be used. It was assumed that only earthquakes occurring within a radius of 300 km of the epicenter of the Milnerton earthquake contribute to the information on the maximum possible earthquake magnitude, m_{max}, of the area (the Milnerton earthquake of magnitude M_L 6.3 occurred in 1809 and is probably one of the largest earthquakes in the history of the region where ground surface rupture and liquefaction features were observed in the epicentral region). Therefore, only earthquakes that occurred within a radius of 300 km circle with canter coordinates 33.916°S and 18.427°E were used. Seismic events are selected from catalogue of earthquakes in South Africa that occurred between AD 1620 and 2006. The data used in this study were compiled from the three sources: Brandt *et al.* (2005), Seismological Bulletins published annually by the Council for Geoscience, Pretoria, and catalogues provided by the International Seismological Centre in the United Kingdom.

The compiled seismic event catalogue was divided into an incomplete part (historic) and three complete parts, each with a different level of completeness. All earthquake magnitudes were unified and expressed in units of local Richter magnitudes, M_L. The incomplete part of the catalogue spans the

period from 1 January 1801 to 31 December 1970, containing 11 of the largest seismic events that occurred during this period. It was assumed that for all of these events, the standard error in magnitude determination was 0.3. Events prior to 1 January 1801 were not used in the calculations, since they cannot be considered as a reliable source of information for the area. The first complete part of the catalogue, complete from magnitude level 3.8, covers the time interval from 1 January 1971 to 31 December 1990. The second part, with a level of completeness of 3.5, covers the time interval from 1 January 1991 to 31 December 1995. The last part of complete catalogue, complete from magnitude 3.0, spans from 1 January 1996 to 30 September 2006. The standard deviations of magnitude determination of the three complete parts were assumed to be 0.3, 0.2, and 0.1, respectively.

The compiled catalogue was then used as input to the Toolbox, to estimate the value of m_{max} according to the 12 procedures discussed above. The results are shown in Table 1.

Table 1

Summary of m_{max} estimation by application of 12 different procedures

Procedure	Equation describing the applied procedure	Estimated $\hat{m}_{max} \pm SD$
T-P	(4)	7.26 ± 0.61
K-S (Cramér's approximation)	(13)	6.89 ± 0.31
K-S (exact solution)	(16)	6.89 ± 0.31
T-P-B	(20)	6.84 ± 0.28
K-S-B	(24)	6.81 ± 0.27
N-P-G	(28)-(29)	6.82 ± 0.28
N-P-OS	(33)	6.76 ± 0.35
Based on a few largest observations	(40)	6.76 ± 0.31
R-W	(42)	6.80 ± 0.57
R-W-C	(46)	6.75 ± 0.40
Minimization of L_1 norm	(49)	6.82 ± 0.28
Minimization of L_2 norm	(50)	7.74 ± 1.07

The format of the catalogue used for the input is simply a column of magnitudes. The output of the code consists of an ASCII text file and five figures. The output file contains information on the input data, the estimated m_{max} values and their standard deviations. If possible, the confidence limits of the estimated m_{max} are also provided. Figures 1, 2, and 3 show the values of m_{max} and their confidence limits, calculated according to the K-S, K-S-B,
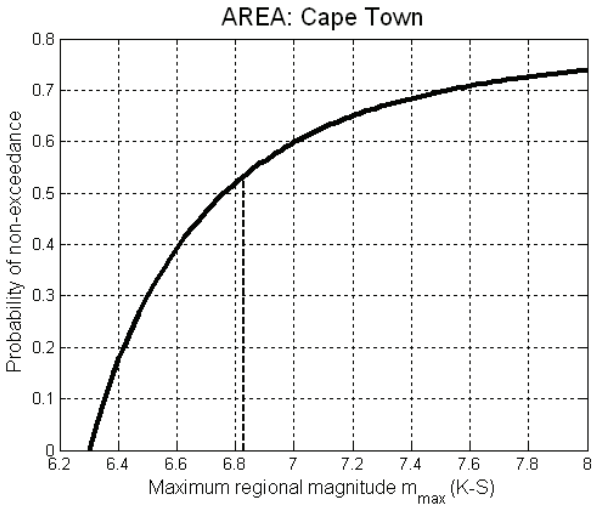
Fig. 1. Graphical output from the MMAX Toolbox, showing the estimated value of m_{max} (dashed line) and its confidence limits (solid line), calculated according to the K-S procedure for the city of Cape Town, South Africa.
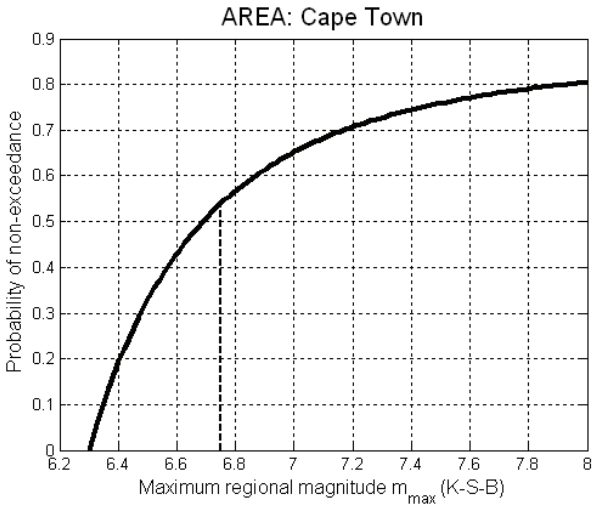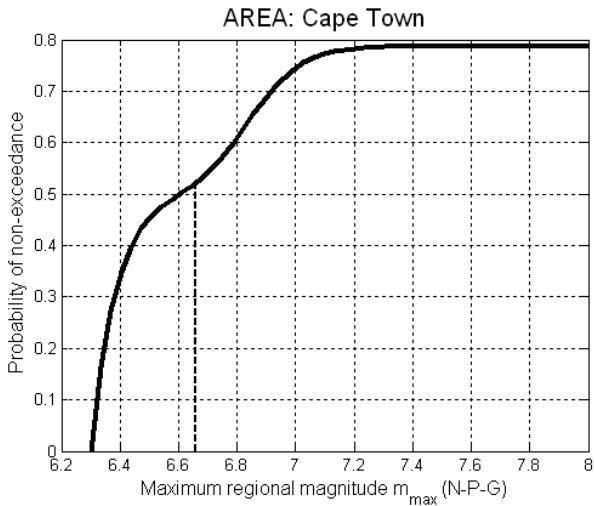

Fig. 2. Graphical output from the MMAX Toolbox, showing the estimated value of m_{max} (dashed line) and its confidence limits (solid line), calculated according to the K-S-B procedure for the city of Cape Town, South Africa.

and N-P-G procedures, respectively. Figures 4 and 5 show the L_1- and L_2-norm fit of the Gutenberg–Richter CDF, eq. (10), and corresponding estimates of m_{max}.

Fig. 3. Graphical output from the MMAX Toolbox, showing the estimated value of m_{max} (dashed line) and its confidence limits (solid line), calculated according to the N-P-G procedure for the city of Cape Town, South Africa.
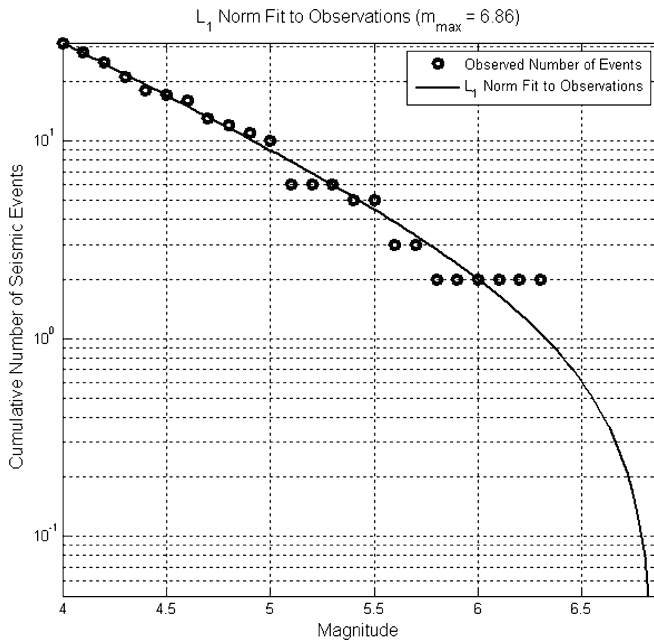

Fig. 4. Graphical output from the MMAX Toolbox showing the L_1-norm fit of the Gutenberg–Richter CDF, eq. (10), to selected data in the vicinity of the city of Cape Town, South Africa.
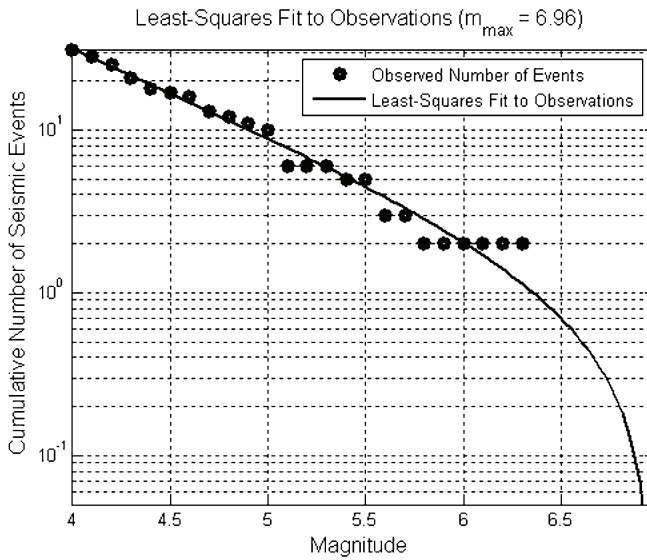
Fig. 5. Graphical output from the MMAX Toolbox showing the L_2-norm (least-squares) fit of the Gutenberg–Richter CDF, eq. (10), to selected data in the vicinity of the city of Cape Town, South Africa.

The Toolbox can also be used when the seismic event catalogue is not available. In such a case, the input data are: the *b*-value of Gutenberg–Richter and its uncertainty, the mean activity rate, λ, and its uncertainty, respective levels of completeness, m_{min}, time span of the catalogue, T, the maximum observed earthquake magnitude in the area, m_{max}^{obs}, and its standard error. Optionally, the second largest observed earthquake magnitude, m_{n-1}, is required. If this information is available, the maximum possible earthquake magnitude is estimated by procedures T-P, K-S (Cramér's approximation), K-S (exact solution), T-P-B, K-S-B, R-B, and R-W-C.

4. DISCUSSION AND CONCLUSIONS

The paper presents 12 statistical procedures for estimation of the region-characteristic, maximum possible earthquake magnitude, m_{max}. All procedures are implemented into the MATLAB program, MMAX, which can be obtained from the authors free of charge. For each of the procedures given, there are notes on its origin, assumptions made in its derivation, condition for validity, weak and strong points, *etc*. The applicability of the particular procedure is determined by the assumptions of the model and/or the available information on the seismicity of the area. It includes cases where:

❑ earthquake magnitudes are distributed according to the truncated frequency-magnitude Gutenberg–Richter relation;

❑ the empirical magnitude distribution deviates moderately from the Gutenberg–Richter model;

❑ when no specific model of the magnitude distribution is assumed; some procedures can also be used when the earthquake catalogue is incomplete and uncertain;

❑ when only a few of the largest magnitudes are available.

All presented procedures, by the nature, are statistical and are driven by seismicity. Therefore, they have certain limitations. Often, statistical procedures provide underestimated m_{max} values, especially in the case of small seismicity samples and/or if applied to areas of low seismic activity (*e.g.*, Chinnery 1979, Bender 1988). It follows that the best approach for evaluation of m_{max} would be to combine all available information on the area-characteristic m_{max} value, *i.e.*: seismicity, seismic activity rate, local geological features, similarities in local and regional tectonic features, constraints based on physical principles, and crustal Lg coda Q attenuation value (Jin and Aki 1988). From a mathematical point of view, the best procedure that allows incorporation of all available information is a Bayesian-based estimation technique (DeGroot 1970). It is interesting to note that the Bayesian-based procedure for m_{max} evaluation is known and occasionally applied in seismic hazard assessment (Wheeler 2009). The foundation of the procedure has been given by Cornell (1994), and details of construction of *a priori* distribution of m_{max} are described in detail by Coppersmith (1994). Unfortunately, the procedure as applied has a mathematical flaw in that it provides values of m_{max} that are systematically underestimated. The underestimation of m_{max} can reach the value of 0.5 magnitude unit (Kijko *et al.* 2009).

While this work presents a spectrum of statistical procedures that can be used for assessment of the area-characteristic maximum possible earthquake magnitude, m_{max}, in different circumstances, it does not provide an answer as to which of the procedures are the best. A systematic investigation of which procedure performs best can be done by means of a Monte-Carlo simulation and will be presented as Part II of this work.

References

Abramowitz, M., and I.A. Stegun (eds.) (1970), *Handbook of Mathematical Functions*: *with Formulas, Graphs and Mathematical Tables*, 9th ed., Dover Publ., New York, 1046 pp.

Anderson, K.R. (1982), Robust earthquake location using *M*-estimates, *Phys. Earth Planet. In.* **30**, 2-3, 119-130, DOI: 10.1016/0031-9201(82)90096-6.

Bender, B. (1988), Reliability of estimates of maximum earthquake magnitudes based on observed maxima, *Seism. Res. Lett.* **59**, 1, 15 (abstract).

Brandt, M.B.C., B. Bejaichund, E.M. Kgaswane, E. Hattingh, and D.L. Roblin (2005), *Seismic History of Southern Africa*, *Seism. Ser., Vol. 37*, Council for Geoscience, Pretoria, 32 pp.

Campbell, K.W. (1982), Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model, *Bull. Seism. Soc. Am.* **72**, 5, 1689-1705.

Chinnery, M.A. (1979), Investigations of the seismological input to the safety design of nuclear power reactors in New England, U.S. Nuclear Regulatory Commission Report NUREG/CR-0563, 72 pp.

Cooke, P. (1979), Statistical inference for bounds of random variables, *Biometrika* **66**, 2, 367-374, DOI: 10.1093/biomet/66.2.367.

Cooke, P. (1980), Optimal linear estimation of bounds of random variables, *Biometrika* **67**, 1, 257-258, DOI: 10.1093/biomet/67.1.257.

Coppersmith, K.J. (1994), Conclusions regarding maximum earthquake assessment. **In:** *The Earthquakes of Stable Continental Regions, Vol. 1. Assessment of Large Earthquake Potential*, Electric Power Research Institute, Palo Alto, 6.1-6.24.

Cornell, C.A. (1994), Statistical analysis of maximum magnitudes. **In:** *The Earthquakes of Stable Continental Regions, Vol. 1. Assessment of Large Earthquake Potential*, Electric Power Research Institute, Palo Alto, 5.1-5.27.

Cosentino, P., V. Ficarra, and D. Luzio (1977), Truncated exponential frequency-magnitude relationship in earthquake statistics, *Bull. Seism. Soc. Am.* **67**, 6, 1615-1623.

Cramér, H. (1961), *Mathematical Methods of Statistics*, 2nd ed., Princeton University Press, Princeton.

David, H.A. (1981), *Order Statistics*, 2nd ed., John Wiley and Sons, New York.

DeGroot, M.H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Dwight, H.B. (1961), *Tables of Integrals and Other Mathematical Data*, 3rd ed., Macmillan Co., New York.

EERI Committee on Seismic Risk (H.C. Shah, chairman) (1984), Glossary of terms for probabilistic seismic risk and hazard analysis, *Earthq. Spectra* **1**, 33-40.

Field, E.H., D.D. Jackson, and J.F. Dolan (1999), A mutually consistent seismic-hazard source model for southern California, *Bull. Seism. Soc. Am.* **89**, 3, 559-578.

Gentle, T.E. (1977), Least absolute values estimation: An introduction, *Commun. Stat*. **B6**, 313-328.

Gibowicz, S.J., and A. Kijko (1994), *An Introduction to Mining Seismology*, Academic Press, San Diego, 399 pp.

Gnedenko, B. (1943), Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math*. **44**, 3, 423-453 (in French).

Hall, P. (1983), Large sample optimality of least squares cross-validation in density estimation, *Ann. Stat*. **11**, 4, 1156-1174.

Hamilton, R.M. (1967), Mean magnitude of an earthquake sequence, *Bull. Seism. Soc. Am*. **57**, 5, 1115-1116.

Jin, A., and K. Aki (1988), Spatial and temporal correlation between coda Q and seismicity in China, *Bull. Seism. Soc. Am.* **78**, 2, 741-769.

Kijko, A. (1994), Seismological outliers: L_1 or adaptive L_p norm application, *Bull. Seism. Soc. Am*. **84**, 2, 473-477.

Kijko, A. (2004), Estimation of the maximum earthquake magnitude, m_{max}, *Pure Appl. Geophys*. **161**, 8, 1655-1681, DOI: 10.1007/s00024-004-2531-4.

Kijko, A., and G. Graham (1998), Parametric-historic procedure for probabilistic seismic hazard analysis. Part I: Estimation of maximum regional magnitude m_{max}, *Pure Appl. Geophys*. **152**, 3, 413-442, DOI: 10.1007/s000240050161.

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**, 3, 645-654.

Kijko, A., S. Lasocki, and G. Graham (2001), Non-parametric seismic hazard in mines, *Pure Appl. Geophys*. **158**, 9-10, 1655-1675, DOI: 10.1007/PL00001238.

Kijko, A., S.J.P. Retief, and G. Graham (2002), Seismic hazard risk assessment for Tulbagh, South Africa: Part I – Assessment of seismic hazard, *Nat. Hazards* **26**, 2, 175-201, DOI: 10.1023/A:1015671813315.

Kijko, A., G. Graham, M. Singh, D. Roblin, and M.B.C. Brandt (2009), Probabilistic PGA and spectral acceleration seismic hazard maps for South Africa. Invited lecture, Workshop R1 "Earthquake Hazard", The IASPEI General Assembly in Cape Town, 11-16 January 2009 (abstract).

Mueller, C.S. (2010), The influence of maximum magnitude on seismic-hazard estimates in the central and eastern United States, *Bull. Seism. Soc. Am.* **100**, 2, 699-711, DOI: 10.1785/0120090114.

Page, R. (1968), Aftershocks and microaftershocks of the great Alaska earthquake of 1964, *Bull. Seism. Soc. Am*. **58**, 3, 1131-1168.

Parzen, E. (1962), On estimation of a probability density function and mode, *Ann. Math. Stat.* **33**, 1065-1076.

Pisarenko, V.F., A.A. Lyubushin, V.B. Lysenko, and T.V. Golubieva (1996), Statistical estimation of seismic hazard parameters: Maximum possible magnitude and related parameters, *Bull. Seism. Soc. Am.* **86**, 3, 691-700.

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling (1994) *Numerical Recipes: The Art of Scientific Programming*, Cambridge University Press, New York.

Quenouille, M.H. (1956), Notes on bias in estimation, *Biometrika* **43**, 3-4, 353-360, DOI: 10.1093/biomet/43.3-4.353.

Robson, D.S., and J.H. Whitlock (1964), Estimation of a truncation point, *Biometrika* **51**, 1-2, 33-39, DOI: 10.1093/biomet/51.1-2.33.

Schwartz, D.P., and K.J. Coppersmith (1984), Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.* **89**, B7, 5681-5698, DOI: 10.1029/JB089iB07p05681.

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Stein, R.S., and T.C. Hanks (1998), $M \geq 6$ earthquakes in southern California during the twentieth century: No evidence for a seismicity or moment deficit, *Bull. Seism. Soc. Am.* **88**, 3, 635-652.

Stone, C.J. (1984), An asymptotically optimal window selection rule for kernel density estimates, *Ann. Stat.* **12**, 4, 1285-1297.

Tate, R.F. (1959), Unbiased estimation: Functions of location and scale parameters, *Ann. Math. Stat.* **30**, 341-366.

Wells, D.L., and K.J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seism. Soc. Am.* **84**, 4, 974-1002.

Wheeler, R.L. (2009), Methods of M_{max} estimation east of the Rocky Mountains, U.S. Geological Survey Open-File Report 2009-1018, 44 pp.

Working Group on California Earthquake Probabilities (WGCEP) (1995), Seismic hazards in southern California: Probable earthquakes, 1994 to 2024, *Bull. Seism. Soc. Am.* **85**, 2, 379-439.