Bulgarian Chemical Communications, Volume 47, Special Issue B (pp. 309–313) 2015

Biased magnitude estimates – impact on the magnitude-frequency distribution assessment

D. Solakov*

National Institute of Geophysics, Geodesy and Geography, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 3, 1113-Sofia, Bulgaria

"Apparent" distributions of random variables are usually considered in seismology. This is due to the fact that, instead of the actual values of the monitored parameter, the evaluated values are used. The coordinates and magnitude of earthquakes are typical examples. Errors in these estimates can lead to significant shifts in the earthquake relevant parameters distribution (magnitude-frequency distribution, spatial distribution etc). In the present work it is shown that due to the limited capabilities of seismic equipment (seismic networks) and the presence of microseismic noise the magnitude estimates are biased. It is numerically illustrated that the bias is significant for relatively small earthquakes (depending on network density and the noise level at the seismic stations) leading to Gutenberg-Richter $b-value$ biased estimates. A procedure for corrected magnitude estimate considering the noise level at the seismic stations is proposed. The use of corrected magnitudes result in negligible bias of $b-value$ estimates.

Key words: magnitude, Gutenber-Richter b value, statistical methods

INTRODUCTION

Uncertainties are an inherent part of seismic studies. Their evaluation as well as the assessment of their impact on following studies is essential both in theoretical and in practical terms. The recording system also has a significant impact on the assessment of the seismicity. This is substantially important for the assessment of the earthquake magnitude-frequency distributions. The concept of "apparent" magnitude was proposed for the first time in Tinti, Mulargia [1]. As a result of errors in the magnitude determination we have not a sample from the true distribution but a sample from "apparent" (observed) distribution, which is a convolution of two distributions - the real one and determination error distribution. In [1] is considered the case of a double side truncated exponential magnitude-frequency distribution and Gaussian distributed estimation error $-N(0,\sigma^2)$. Thus the "apparent" density function is convolution of the two densities (exponential and normal one):

$$g(M) = \int_{-\infty}^{\infty} f(M-\tau)h(\tau)d\tau$$

where M is the earthquake magnitude and

$$f(M) = \begin{cases} 0, & M < M_0 \\ \beta e^{-\beta M}/(1-e^{-\beta M_1}), & M_0 \le M \le M_1 \\ 0, & M_1 < M \end{cases}$$

is the density of the double side truncated exponential distribution with lower and upper earthquake magnitude bounds M_0 and M_1 respectively and

$$h(\tau) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\tau^2}{2\sigma^2}}$$

is a normal distribution of the estimation error. In case of a sufficiently large difference $M_1 - M_0$ in a wide magnitude interval we will have:

$$g(M) = e^{\frac{\beta^2\sigma^2}{2}} f(M),$$

where $\beta = b\ln 10$ (b – slope of the recurrence graph), i.e. the "apparent" number of earthquakes will be $\exp(\beta^2\sigma^2/2)$ times larger than the real one. In this study it is shown that in real terms the estimate of the magnitude is biased, particularly for small earthquakes and the influence of this bias on the b value estimates is considered.

THEORETICAL JUSTIFICATION

Present-day broadband seismometers measure ground motion velocity over a wide frequency range. The following relation is valid in this range: $A/T = V/2\pi$, where A is the ground displacement amplitude, T is the corresponding period and V is the ground motion velocity. In general, to "recognize" the earthquake on seismograms it is necessary:

$$\log(V_{max})_j - \log(V_n) > 0, \tag{1}$$

where $V_{\max}$ is the maximum velocity for wave type $j = P, S$ etc. and V_n is the level of microseismic noise

* To whom all correspondence should be sent:
dimos@geophys.bas.bg

at the corresponding station. Let us consider the magnitude equation:

$$M_j = \log(\frac{V_{\max}}{2\pi})_j + \delta_j(\Delta) + S_j \pm \varepsilon, \qquad (2)$$

where $\delta_j(\Delta)$ is the calibration function for each wave type j at a distance Δ, S_j – station correction for the corresponding wave type j, ε – random variable, reflecting the uncertainty of the maximum ground motion velocity for the relevant magnitude and distance (a normal distribution with a mean 0 and variance σ^2 for all magnitudes and distances is assumed in this study), due to the complex nature of the physical phenomenon earthquake. For a convenience, it is assumed that the calibration function takes in the constant $-\log(2\pi)$. From equation (1) follows that an earthquake with a magnitude M_j at a distance Δ will generate a maximum ground motion velocity:

$$\log(V_{\max})_j = M_j - \delta_j(\Delta) - S_j \pm \varepsilon, \qquad (3)$$

To recognize the earthquake ground motion on the seismogram the following condition is to be satisfied:

$$\log(V_n) < M_j - \delta_j(\Delta) - S_j \pm \varepsilon. \qquad (4)$$

From (3) follows that the logarithm of the maximum ground motion velocity $\log(V_{\max})_j$, generated by an earthquake M_j at a distance Δ is a random variable that is generally considered as normally distributed (assuming distribution $N(0, \sigma^2)$ of ε) with mean $M_j - \delta_j(\Delta) - S_j$ and variance σ^2. Considering inequality (4), the distribution of the logarithm of the recorded maximum ground motion velocity $\log(V_{\max})_j > \log(V_n)$ will be left side truncated normal distribution. The distribution is similar when a trigger mode is used. In this case, the lower limit is not the noise level but the value of the specified trigger. In practice, for strong earthquakes at short distances, the distribution will be right side truncated, due to the limited capabilities of the seismic equipment. From the foregoing,it follows that for small earthquakes the mean of the logarithm of the reported maximum ground motion velocity will be shifted from the real one – $M_j - \delta_j(\Delta) - S_j$. The bias depends both εdistribution and the difference $M_j - \delta_j(\Delta) - S_j - \log(V_n)$. In Table 1 are given the normal distributions with variance 0.25 and different mean values, as well as the mean of the corresponding left truncated normal distributions with boundary 0 (i.e. $\log(V_n) = 0$).

The table shows that there will be significant differences between the mean of the actual (normal) and the reported (truncated normal) for small mean values of the normal distribution. These results indicate that, theoretically, the magnitude estimate will be biased to higher values. Generally, for magnitudes which generate mean maximum ground motion velocity $10^{3\sigma}$ times greater than the noise level the estimates will be not biased. If this condition is satisfied for all stations used in magnitude determination, the estimates will be practically not biased. If for one or more stations this condition is not satisfied, the estimate will be biased. In general, these considerations are valid for any magnitude, regardless of the used ground motion parameter. This is result of the uncertainties related to the complex nature of earthquakes and seismic wave propagation paths, which form ε error and its variance. It is worth to be noted that in practice are used maximum amplitudes exceeding several times the noise level.

NUMERICAL MODELING

Numerical experiments were carried out to assess the influence of the seismic network, station conditions (geotechnical properties and microseismic noise level), error ε and its variance on the magnitude determination and hence the evaluation of the magnitude-frequency distribution. The following assumptions are accepted in the numerical modeling:

- considered area Q – rectangle with sides X and Y km;
- observation network of N stations uniformly distributed in the area Q. To avoid the case of close stations additionally is accepted the condition that the distance between two stations must be greater than $(X * X/\pi/N)^{0.5}$;
- uniform distribution of the earthquakes in the area Q;

Table 1. Mean values of the normal distribution with $\sigma = 0.25$ and the corresponding left truncated distribution (at value 0)

Mean value of the normal distribution	-1	-0.5	0	0.5	1	1.5
Average value of the truncated distribution (limit 0)	0.187	0.263	0.399	0.644	1.028	1.502
Difference between the average values(normal - truncated normal)	-1.187	-0.763	-0.399	-0.144	-0.028	-0.002

— the magnitude determination is based on the maximum ground motion velocity $V_{\max}$ for P wave (a calibration function $\delta(\Delta)$ that is presented in [2]) is used;
— uniform distribution of station corrections in the interval $[-0.5, 0.5]$;
— double truncated exponential distribution with parameter β in the range $[m, M_{\max}]$ for magnitude, where m is sufficiently small ($m \approx M_0$) – minimal magnitude which could be documented by the network);
— uniform distribution of the logarithm of V_n $[10^{-2.5}\mu m/s, 10^{-0.5}\mu m/s]$. These values are approximately in the range of observed seismic noise at bulgarian stations;
— earthquake with a magnitude M at a distance Δ generates mean maximum ground motion velocity $10^{M_j-\delta_j(\Delta)-S_j\pm\varepsilon}$, where S-station correction, $\varepsilon\ N(0,\sigma^2)$;
— an earthquake is considered reliably documented if at least at k ($0 < k \leq N$) stations is recorded the maximum ground motion velocity of P wave SNR times greater than the noise level. The magnitude is determined as a mean of the estimates for all the stations satisfying this requirement.

The following algorithm (concerning the above presented assumptions) is applied:

1. N random seismic stations that are uniformly distributed in the area Q are generated;
2. Random station corrections S_i and seismic noise level V_{ni} (with distributions as described above), where $i = 1, 2, \dots N$ are generated for each of the seismic stations;
3. An earthquake is generated randomly distributed in the area Q;
4. A magnitude M is generated – random variable with double side truncated exponential distribution in the range $[m, 7]$ with parameter $\beta (\beta = \ln(10) * b)$ is generate;
5. For each station j=1,2,...,N is generated maximum ground motion velocity $V^j_{\max} = M - \delta_j(\Delta) - S_j \pm \varepsilon$, where $\varepsilon\ N(0,\sigma^2)$. If inequality $V^{j_{\max}} > SNR * V_{nj}$ is fulfilled for a station, it is assumed that this station has recorded the earthquake;
6. If the earthquake is eligible for documentation (i.e. the number of stations registered the earthquake is greater than or equal to k), the mean magnitude $\overline{M}$ is calculated from the estimated magnitudes of all stations which have registered the event.
   A corrected magnitude is estimated taking into account the offset $\overline{M}$ of resulting from the truncated distribution of ε (i.e. distribution of the measured maximum ground motion velocity) by the following procedure: For each station which have recorded the earthquake the "corrected" magnitude is calculated:

$$M_{\text{corr}} = \log(V) + \delta(\Delta) + S - E(\varepsilon \mid \overline{M}, \log SNR * V_n)),$$

   where Δ – epicentral distance of the station, $E(\varepsilon \mid \overline{M}, \log SNR * V_n))$ – mathematical mean of ε for the distance Δ, assuming $\overline{M}$ that a magnitude is the real one. M_{corr} accounts the fact that the distribution of the measured maximum ground motion velocity for $\overline{M}$ is truncated with a lower boundary $SNR * V_n$. An averaged corrected magnitude $\overline{M_{corr}}$ is calculated using the corrected magnitudes M_{corr} for each station.
   The procedure is repeated iteratively until the modulus of the difference between two successive iterations is greater than δ.
7. If the earthquake is documented it is included in a catalog as magnitude $\overline{M}$ and $\overline{M_{\text{corr}}}$ and the magnitude values are rounded to 0.01;
8. Steps from 3 to 7 are repeated, while in the catalog 1000 earthquakes with magnitude more than or equal to 4.0 are recorded;
9. For the generated catalogue, the slope b of the magnitude – frequency distribution is estimated for different minimum magnitudes – from m to M_1, where M_1 is the maximum magnitude, for which $N(M \geq M_1) > 25$). Estimates are obtained by using maximum likelihood method [3];
10. 1000 catalogues are generated (steps from 3 to 9) and the resulting estimates of b for different minimum magnitudes are averaged.

The described algorithm is applied for different values of X, Y, N, k, m, b and σ. In Fig. 1 are given the obtained results for a rectangle area Q with sides 600 and 200 km, $N = 50, k = 4, m = 1.5, b = 1.0, \sigma = 0.5, SNR = 6$ and $\delta = 0.01$. In the adopted terms, the minimum magnitude $M_{\min}$, above which all earthquakes are recorded is about 1.7. The figure shows that for a large magnitude interval the b-value estimates are biased and for some minimal magnitudes is more than 20% larger than the real one. The same behavior (uncorrected magnitudes) as it is suggested

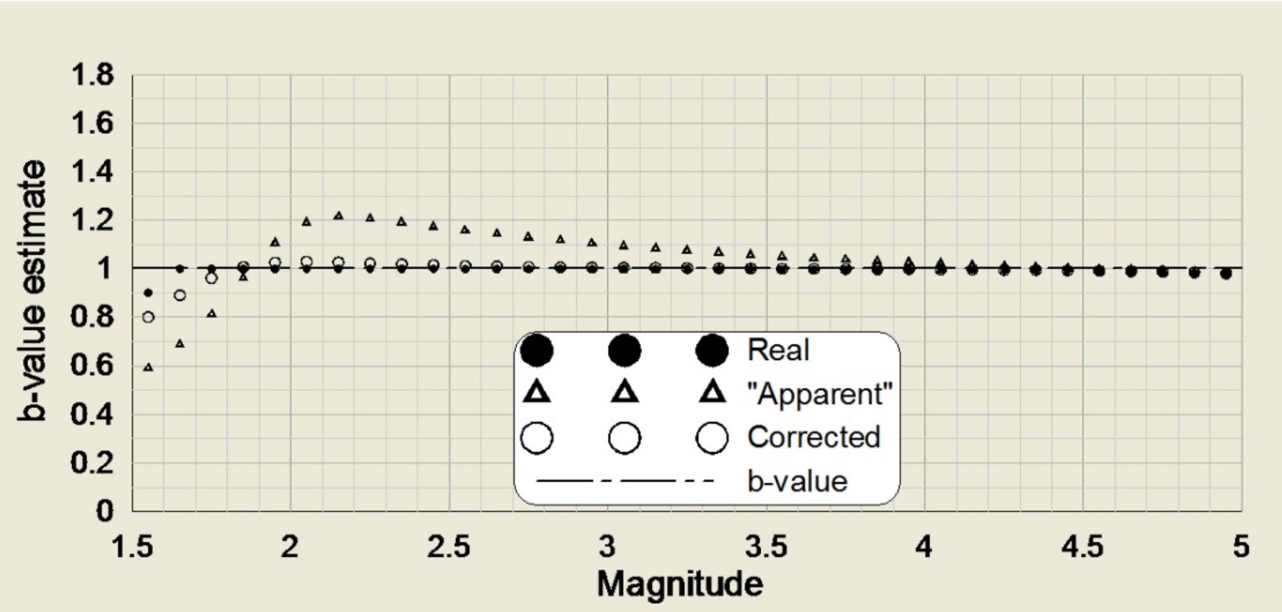

Fig. 1. Calculated b-value versus $M_{\min}$.

by Figure 1 was observed for earthquakes in California between 1984 and 1999 as presented in [4]. The effect is explained [4] by a larger determination error of the small earthquake magnitudes. Corrected magnitudes lead to practically non biased estimates of b. This tendency is not dependant of Q, N, k, m, b and SNR. The bias decreases with decreasing of σ but $\sigma = 0.5$ seems to be a realistic value for scatter of the V_{max} modeling (standard deviations of the PGV attenuation relationships are close to 0.5 (for example [5]).

FINDINGS AND CONCLUSIONS

It has been shown that, the estimation of the earthquake magnitude is biased to higher values due to the level of seismic noise. This offset is considerable for relatively small earthquakes (depending on the density of the network and the level of the noise at the seismic stations) and this leads to a bias of the b estimates (slope of the magnitude frequency distribution). The proposed procedure for the magnitude correction based on the noise level leads on practically not biased assessments. Biases appear in estimates, with assumed normal distribution of the uncertainties that is in reality truncated normal, such as: ground motion attenuation models; models for various parameters of earthquakes in dependence of the magnitude (for example rupture length, fault displacement, size of the rupture etc). In all these cases, for low magnitudes (or for large distances in attenuation models) predominantly observed extreme values are considered. This results in a shift of the parameter estimates.

REFERENCES

[1] S. Tinti and F. Mulargia, *Bull. Seis. Soc. Am.* **75**, 1681-1697 (1985).

[2] E. Samardjieva, “Magnitude classification for near earthquakes based on short-period seismometers and some applications”, *PhD thesis*, Geophysical Institute, Sofia (1986).

[3] R. Page, *Bull. Seismol. Soc. Am.* **58**, 1131-1168 (1968).

[4] K. Felzer, *U. S. Geological Survey, 525 S. Wilson, Pasadena, CA 91106 United States AGU Fall Meeting Abstracts* (2006).

[5] K. Campbell and Y. Bozorgnia, *Earthquake Spectra* **24**, 139-171 (2008).

ОТМЕСТВАНЕ НА МАГНИТУДНИТЕ ОЦЕНКИ – ВЛИЯНИЕ ВЪРХУ ОЦЕНКАТА НА МАГНИТУДНО-ЧЕСТОТНОТО РАЗПРЕДЕЛЕНИЕ

Д. Солаков

Национален институт по геофизика, геодезия и география, Българска академия на науките, ул. "Акад. Г. Бончев" блок 3, 1113 София, България

(Резюме)

За сеизмичните райони земетресенията са неделима част от средата, която ни заобикаля и не съществува друго явление, което да е така "безценно" от научна гледна точка и така катастрофално от социално икономическа и психологична гледна точка като силното земетресение.

Съвременната сеизмология се основава на интерпретацията на огромно количество наблюдателни данни. Развитието й е немислимо без необходимата теоретична база и прилагането на надеждни статистически и изчислителни методи и процедури. Математическите и статистическите подходи играят съществена роля в сеизмологията. Използването на статистическите подходи в сеизмологията трябва да бъде мостът между физически базираните модели (без статистика) и статистически базираните модели (без физика).

В сеизмологията винаги се работи с "привидни" ("видими") разпределения на случайни величини. Това се дължи на факта, че вместо с истинските стойности на наблюдаваните величина, се работи с оценки на тези стойности. Типични примери за това са магнитудът и координатите на земетресенията. Грешките в оценките на тези параметри могат да доведат до значими отмествания в разпределенията на земетресенията по съответните параметри (магнитудно-честотна зависимост, пространствено разпределение).

В настоящата работа е доказано, че в следствие ограничените регистриращи възможности на сеизмичната апаратура (съответно сеизмичните мрежи) и наличието на микросеизмичен шум, оценката на магнитуда на земетресенията е отместена. Числено е показано, че това отместване е значимо за относително по-слаби земетресения (в зависимост от плътността на мрежата и нивото на шума в сеизмичните станции) и води до отместване на оценките за наклона *b*на магнитудно-честотната зависимост. Предложена е процедура за оценка на магнитуда, отчитаща нивото на шума в сеизмичните станции, при която се получават оценки с практически пренебрежимо отместване на оценките на параметъра *b* на магнитудно-честотната зависимост.