

# Correcting the Gutenberg–Richter $b$ -value for effects of rounding and noise

V. H. Márquez-Ramírez · F. A. Nava ·  
F. R. Zúñiga

Received: 16 December 2014 / Accepted: 12 February 2015 / Published online: 9 April 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The effects of magnitude rounding and of the presence of noise in the rounded magnitudes on the estimation of the Gutenberg–Richter  $b$ -value are explored, and the ways to correct for these effects are proposed. For typical values,  $b = 1$  and rounding interval  $\Delta M = 0.1$ , the rounding error is approximately  $-10^{-3}$  and it can be corrected to a negligible approximately  $-10^{-5}$ . For the same typical values, the effect of noise can be larger, depending on the characteristics of the noise distribution; for normally distributed noise with standard deviation  $\sigma = 0.1$ , the correct  $b$ -value may be underestimated by a factor  $\sim 0.97$ .

**Keywords**  $b$ -Value · Noise · Magnitude

## 1 Introduction

The Gutenberg–Richter magnitude distribution (Gutenberg and Richter 1944; Richter 1958) is possibly the most widely used statistical relationship in seismology, and its slope  $b$ , widely known as the  $b$ -value, characterizes the ratio of small to large magnitudes. The  $b$ -value, hereafter referred to as  $b$ , is considered to be the characteristic for a given region, and many authors (Scholz 1968; Wyss 1973; Smith 1981; Wiemer and Benoit 1996; Enescu and Ito 2001; Nuannin et al. 2005; De Santis et al. 2011; Mallika

et al. 2013) maintain that it decreases slightly before large earthquakes; De Santis et al. (2011) suggest that the knowledge of  $b$  is useful to characterize the information entropy of the system. Thus, knowing the correct value, within uncertainty bounds, of  $b$  is important for seismic hazard studies, because it allows the estimation of the occurrence rate of earthquakes for any given magnitude range, because of its possible role as a precursor and because other information it contains.

Because of the many imponderables affecting the transmission of energy from source to seismograph, including travel path, radiation pattern, site effects, variations in instrument response, and reading errors, in current practice, the magnitudes are usually rounded up to the first decimal place, so that the methods used to evaluate  $b$  experimentally usually work with data binned in classes  $\Delta M = 0.1$  wide.

In the present work, we discuss the effects of rounding and of noise in the determination of  $b$  and propose a way to correct them. The effects of rounding have been previously dealt with by, among others, Utsu (1965), Bender (1983), Tinti and Mulargia (1987), Kijko and Sellevoll (1992), Zúñiga and Wyss (1995), Rhoades (1996), Marzocchi and Sandri (2003), and Bengoubou-Valerius and Gibert (2012); here, we present our estimate of these effects and a new way of correcting the  $b$  measurements. The difference between our version and those of the aforementioned authors is that the correction method we propose is more simple and straightforward and does not involve complicated calculations.

## 2 Correcting the effects of rounding

In what follows, we present our own version of the effects of rounding, in order to introduce the notation that will be used in proposing a correction for them.

V. H. Márquez-Ramírez (✉) · F. R. Zúñiga  
UNAM Campus Querétaro, México, Centro de Geociencias,  
Blvd. Juriquilla No 3001, 76230 México, Querétaro, México  
e-mail: marvh81@gmail.com

F. A. Nava  
Seismology Department, CICESE, Carretera Ensenada-Tijuana  
No 3918, Zona Playitas, 22860 Ensenada, BC, México

The Gutenberg–Richter (G–R) distribution

$$\log_{10} N(M) = a - b(M - M_1); \quad M \geq M_1, \quad (1)$$

where  $N(M)$  is the number of earthquakes with magnitude  $\geq M$  and  $a$  is the  $\log_{10}$  of the total number of events with  $M \geq M_1$  and depends on the overall seismicity rate and on the sampling time; an important issue in the G–R distribution is the magnitude of completeness,  $M_1$  (Wiemer and Wyss 2000); for the magnitude population, we used several values of  $M_1$  with the same results, and it was finally fixed at  $M_1 = 2$ ;  $b$  has been discussed above.

The G–R is a reverse cumulative distribution that implies that magnitudes are distributed exponentially, with probability density function:

$$f(M) = \beta e^{-\beta(M-M_1)}; \quad M \geq M_1, \quad (2)$$

where

$$\beta = \frac{b}{\log_{10}(e)}. \quad (3)$$

From (2) and (3), Aki (1965) determined the maximum-likelihood  $b$  estimate to be

$$b = \frac{\log_{10}(e)}{\bar{M} - M_1}. \quad (4)$$

Equation (4) refers to “exact” magnitudes, but when the  $M$  population, distributed according to (2), is rounded to some  $\Delta M$ , then the class containing magnitude  $M$  in fact contains all magnitudes in the interval  $[M - \Delta M/2, M + \Delta M/2)$ , so that if  $M_1$  is the minimum rounded magnitude, the effective minimum magnitude is  $M_{\min} = M_1 - \Delta M/2$ , and the correct distribution is

$$p(M) = \beta e^{-\beta[M-(M_1-\Delta M/2)]}. \quad (5)$$

Hence, Utsu’s (1965) method of estimating the maximum-likelihood  $b$  is as follows:

$$b = \frac{\log_{10} e}{\bar{M} - (M_1 - \Delta M/2)}. \quad (6)$$

Thus, Aki’s formula is not applicable to practical cases where rounded magnitudes are used, unless  $M_1$  in (4) actually refers to the minimum unrounded magnitude, in which case Aki’s formula coincides with (5). The mistaken application of (4) with the minimum rounded magnitude as  $M_1$  will yield too high estimations of  $b$ .

Figure 1 illustrates the application of (4) and (6) to populations of different sizes and clearly shows the overestimation of  $b$  by Aki’s formula; the horizontal line indicates the  $b$ -value used to generate the magnitude population. Also shown in Fig. 1, the sample sizes larger than 450 are necessary for estimations with standard deviations within  $\pm 0.05$  of the true value (dotted horizontal lines), and the sample sizes proposed as sufficient by Aki

(1965) and Shi and Bolt (1982),  $\sim 50$  and  $\sim 100$  events, respectively, may be too optimistic; we have seen studies with  $b$  estimations based on  $\sim 50$  event samples (Sahu and Saikia 1994; Monterroso and Kulhánek 2003; Spada et al. 2013; Sharma et al. 2013) and can only wonder about the reliability of their results.

Furthermore, working with rounded magnitudes has other consequences that are not completely described by (6). The probability distribution of rounded magnitudes is actually a discrete one, so that the probability of events in the class centered on rounded magnitude  $M$  (class  $M$ ) is

$$\begin{aligned} \Pr(M) &= \int_{M-\Delta M/2}^{M+\Delta M/2} \beta e^{-\beta[m-(M_1-\Delta M/2)]} dm \\ &= e^{-\beta[M-(M_1-\Delta M/2)]} \left( e^{+\beta \Delta M/2} - e^{-\beta \Delta M/2} \right), \end{aligned}$$

and, expanding the exponentials within the parentheses as

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^n}{n!} + \dots$$

(e.g., Dwight (1961), the resulting equation is

$$\begin{aligned} \Pr(M) &= \beta e^{-\beta[M-(M_1-\Delta M/2)]} \Delta M \left[ 1 + \beta^2 \Delta M^2 / (2^2 3!) \right. \\ &\quad \left. + \beta^4 \Delta M^4 / (2^4 5!) + \dots \right]. \quad (7) \end{aligned}$$

The first term in brackets, multiplied by the common factor, is the probability density of magnitude  $M$ , multiplied by  $\Delta M$ ; for  $\Delta M \rightarrow dM$ , all other terms tend rapidly to zero and (7) gives the correct expression for the probability of  $M$ .

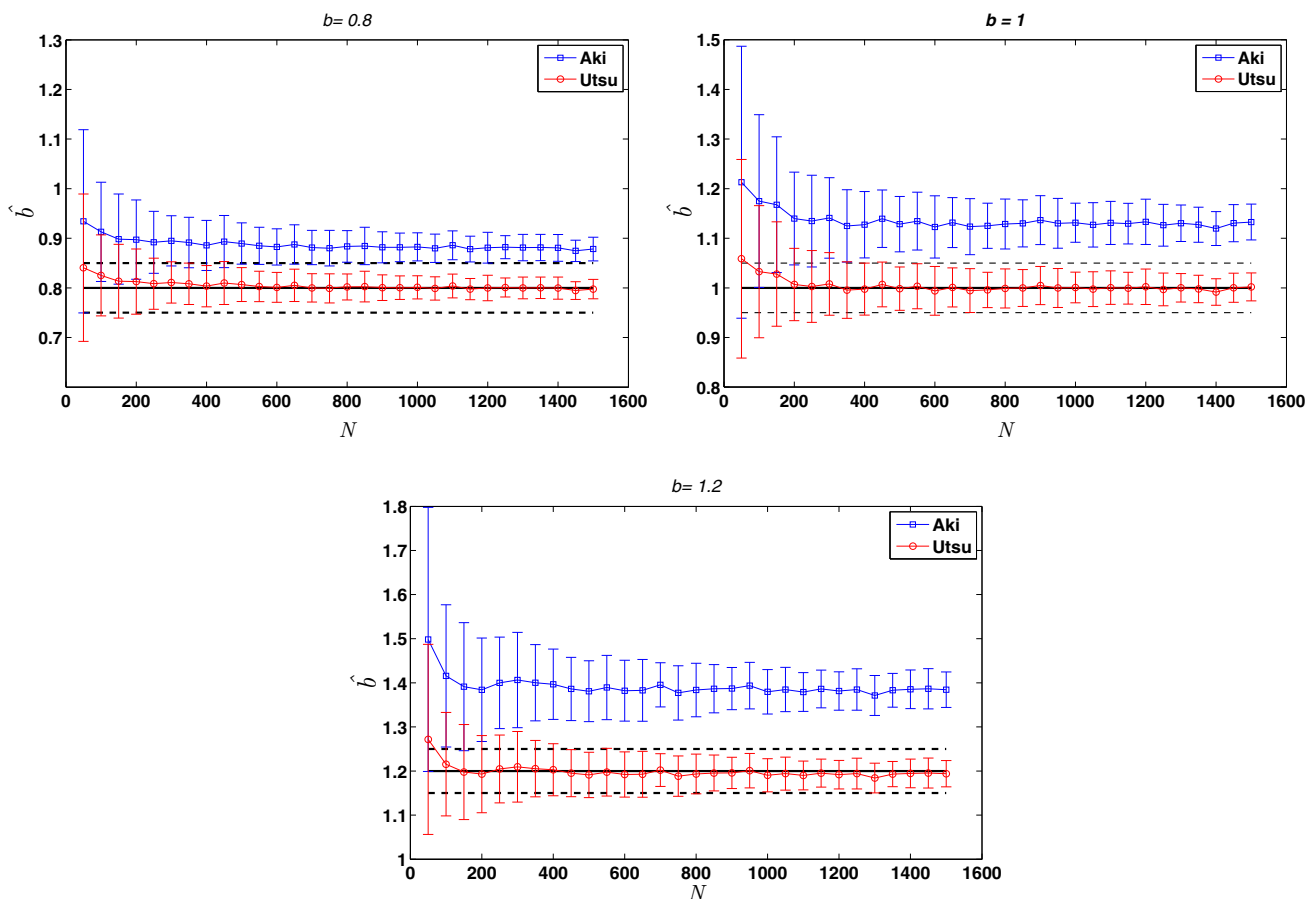
However, for finite  $\Delta M$ , all but the first terms in brackets account for the fact that  $\beta \exp\{-\beta[M - (M_1 - \Delta M/2)]\}$  is not an unbiased estimator of the probability in class  $M$ , and Eq. (7) tells us that, for finite  $\Delta M$ , the rounded magnitudes are not distributed strictly exponentially, but rather as a sum of scaled copies of the exponential distribution.

Consequently, the correct mean of the rounded magnitudes will be

$$\begin{aligned} \bar{M} &= \hat{M} \left[ 1 + \beta^2 \Delta M^2 / (2^2 3!) + \beta^4 \Delta M^4 / (2^4 5!) + \dots \right] \\ &\doteq \hat{M} \left[ 1 + b^2 \Delta M^2 \ln(10)^2 / (2^2 3!) + \dots \right], \quad (8) \end{aligned}$$

where  $\hat{M}$  is the measured mean, so that  $b$  will be overestimated when using the measured mean from (8) in (6), because of the contribution of all but the first terms in brackets. In the bottom line, where the measured mean is expressed in terms of  $b$ , we have kept only the first two terms in the expansion, because, for typical  $b$  and  $\Delta M$  values, contributions from the other terms are of the order of  $10^{-6}$  or less.

Other authors have obtained expressions similar to (8), but they do not specify how to obtain the corrected  $\bar{M}$



**Fig. 1** Estimated  $\hat{b}$ , from Aki’s and Utsu’s formulae, as a function of sample size. The magnitude population was randomly generated with  $b = 0.8$ ,  $b = 1.0$ , and  $b = 1.2$ . Each symbol represents the mean of 100 realizations, and the vertical bars indicate plus/minus one standard deviation

estimation, since the correction depends on the still unknown value of  $b$ . The problem is, however, easily solved by equating (8) with the expression for  $\bar{M}$  as a function of  $b$  obtained from (6), which results in the simple polynomial

$$0.220912 \Delta M^2 b^3 + (\hat{M} - M_{\min}) b - 0.434294 = 0, \quad (9)$$

which can be easily solved for a real root not far from the overestimated value obtained from the measured mean.

Figure 2 illustrates the effects of rounding and of the proposed correction for Monte-Carlo simulation of 50 synthetic random samples of 10,000 magnitudes each, all generated using  $b = 0.8, 1.0$ , and  $1.2$  and rounded to  $\Delta M = 0.1$ ; for ease of view, they have been sorted according to the “true” observed  $\hat{b}$ , i.e., the value actually estimated from the samples before rounding.

From (8), it is clear that the  $b$  overestimation error increases with  $b$  and with  $\Delta M$ ; for the example illustrated in Fig. 2, the mean error  $\hat{b}_{\text{exact}} - \hat{b}_{\text{rounded}}$  is  $-0.00222$ , while the mean error  $\hat{b}_{\text{exact}} - \hat{b}_{\text{corrected}}$  is  $-0.000011$ .

### 3 The effect of noise

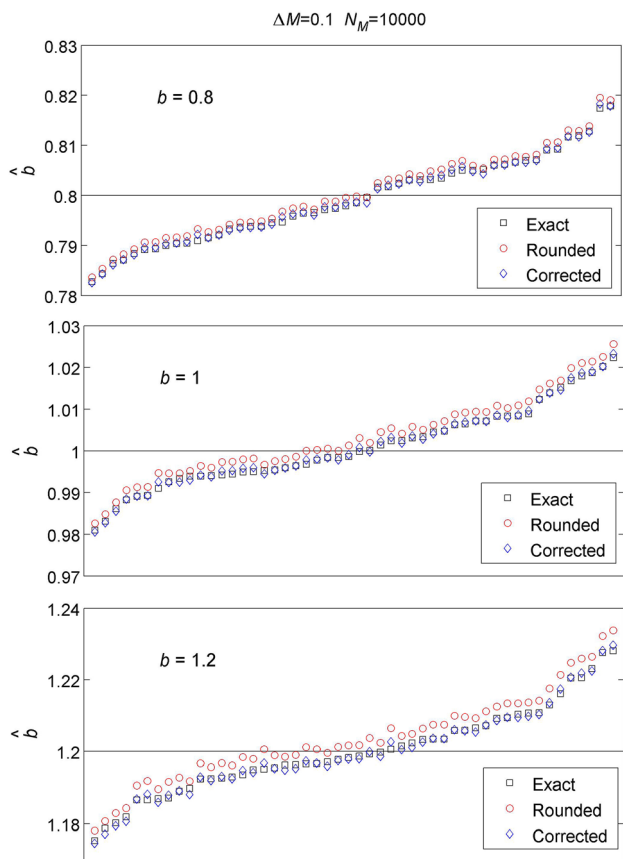
Until now, we have considered rounded magnitudes derived from “exact” magnitudes, i.e., magnitudes generated according to (5) with no error (ignoring computer round-off errors). Actually, as mentioned before, magnitudes are rounded because of uncertainties inherent in their estimation, but the rounding is generally done actually considering the estimates as being exact. Thus, the effect of the uncertainties, which can be considered as noise superimposed on the exact magnitudes, on the estimate rounding should be considered.

For a total population of  $N$  magnitudes distributed exponentially with parameter  $\beta$  and classes of width  $\Delta M$ , the number of events in the  $i$ th class may be approximated by

$$n_i = \beta e^{-\beta M_i} \Delta M N, \quad (10)$$

corrected if desired by (7), and

$$\bar{M} = \frac{1}{N} \sum_i n_i M_i. \quad (11)$$



**Fig. 2** Comparison of  $b$ -values for Monte-Carlo simulations; squares are estimates from “exact” (unrounded) magnitudes, circles are the corresponding estimates after rounding, and rhombs are corrected rounded estimates

The number of events in class  $i - k$  is

$$n_{i-k} = \beta e^{-\beta(M_i-k\Delta M)} \Delta MN = n_i e^{\beta k \Delta M}, \tag{12}$$

and that in class  $i + k$

$$n_{i+k} = \beta e^{-\beta(M_i+k\Delta M)} \Delta MN = n_i e^{-\beta k \Delta M}. \tag{13}$$

Consider now the effect of adding variations (noise)  $v$  to the rounded exact magnitudes in (10); let  $p_0 = \Pr(|v| > \Delta M/2)$  be the probability that a given modified magnitude from a given class  $i$  will be assigned to some other class (after rounding to  $\Delta M$ ). The probabilities that the variation of a magnitude belonging to class  $i - k$  or class  $i + k$  will place it in class  $i$  are  $p_k^- = \Pr(\Delta M[k - 1/2] < v < \Delta M[k + 1/2])$  and  $p_k^+ = \Pr(-\Delta M[k - 1/2] < v < -\Delta M[k + 1/2])$ , respectively.

Thus, the new number of events in class  $i$ ,  $\hat{n}_i$ , will be

$$\hat{n}_i = n_i - p_0 n_i + \sum_{k=1}^{\infty} (p_k^- n_{i-k} + p_k^+ n_{i+k}),$$

from (6) and (7),

$$\hat{n}_i = n_i - p_0 n_i + \sum_{k=1}^{\infty} (p_k^- n_i e^{\beta k \Delta M} + p_k^+ n_i e^{-\beta k \Delta M}),$$

$$\hat{n}_i = n_i \left[ 1 - p_0 + \sum_{k=1}^{\infty} (p_k^- e^{\beta k \Delta M} + p_k^+ e^{-\beta k \Delta M}) \right] \equiv n_i \zeta. \tag{14}$$

The summation in (14) is done over permissible values of  $k$  while contributions are significant. Note that the  $\zeta$  factor does not depend on  $i$  and is the same for all classes; thus, the new distribution will be a scaled version of the original one. The new numbers in each class come from a new exponential distribution

$$\hat{n}_i = \hat{\beta} e^{-\hat{\beta} M_i} \Delta M N,$$

$$\hat{M} = \frac{1}{N} \sum_i \hat{n}_i M_i = \frac{\zeta}{N} \sum_i n_i M_i = \zeta \bar{M},$$

$$\therefore \hat{\beta} = \frac{1}{\hat{M}} = \frac{1}{\zeta \bar{M}} = \frac{\beta}{\zeta},$$

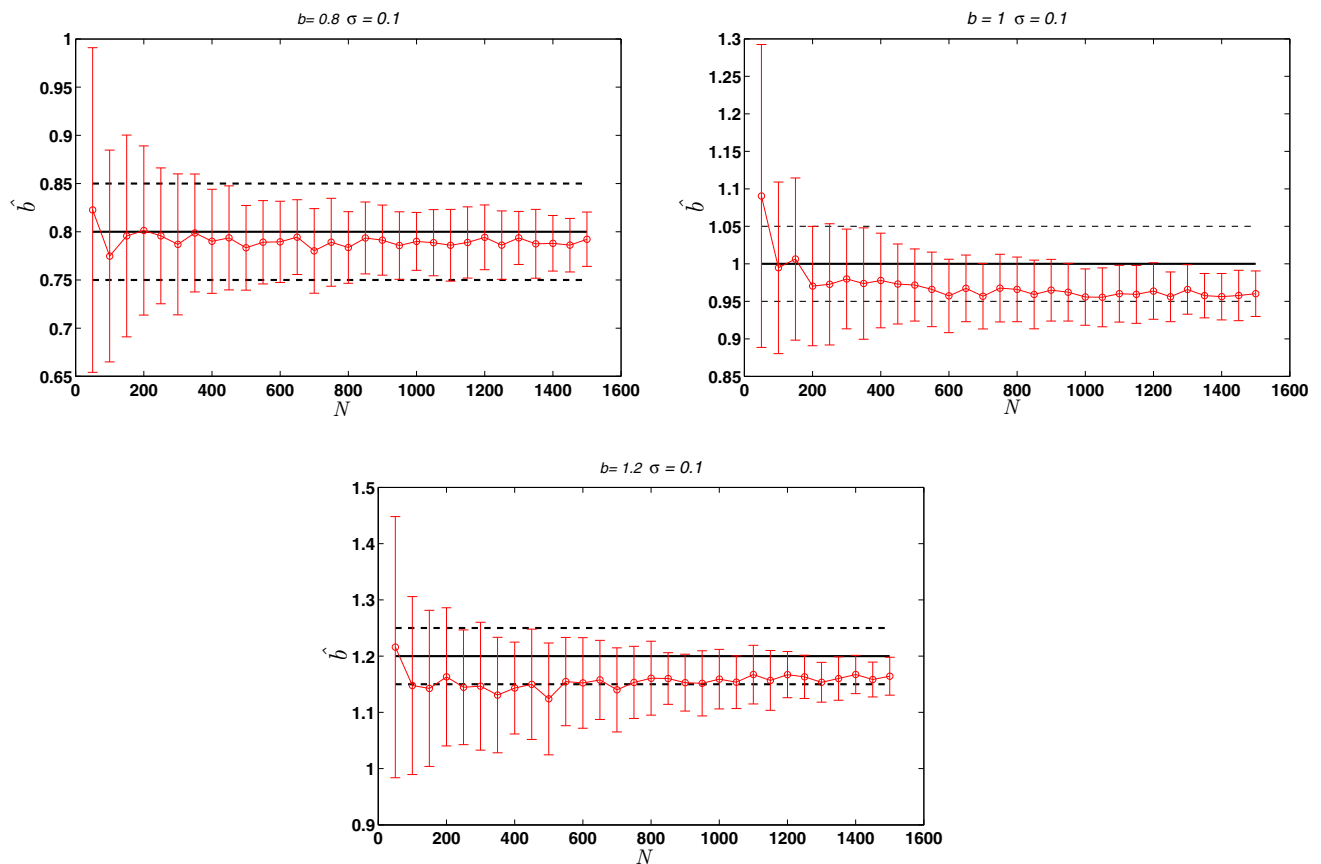
$$\hat{b} = \frac{b}{\zeta}. \tag{15}$$

We have set no constraints to the noise distribution, so that it may have any shape, symmetrical or not, and may even be a function of the magnitude.

Thus, for any estimate about the distribution of noise, the measured estimate  $\hat{b}$  can be corrected by the factor  $\zeta$ , evaluated according to (14), to obtain a better estimate of  $b$ .

On the practical side, since magnitudes are usually rounded up to the first decimal place, data are typically binned in classes  $\Delta M = 0.1$  wide. On the other hand, since magnitudes are customarily considered to have  $\pm 0.1$  uncertainty and since usually the stated uncertainties correspond to 1 or 2 standard deviations, the customary uncertainty could be interpreted as saying that the variations in magnitude have  $\sigma = 0.1$  or  $\sigma = 0.05$  standard deviations, respectively.

As an example, let  $\Delta M = 0.1$  and consider variations distributed as  $N(0, \sigma)$ , so that  $p_k^- = p_k^+$ , with  $\sigma = 0.1$ ; if the original  $b = 1.0$ , then from (14) and (15)  $\zeta = 1.029134$  and  $\hat{b} = 0.971691$ , so that the variations are expected to cause a change  $\delta b \sim -0.028$  in the estimation of  $b$ . This estimation of the effect of noise on rounding is corroborated through Monte-Carlo simulation as shown in Fig. 3, where each point represents the mean of 100 realizations of  $b$  estimation for different population sizes; magnitudes were generated according to (5) with  $b = 0.8, 1.0, \text{ and } 1.2$ , and noise distributed according to  $N(0, \sigma = 0.1)$  was added; then magnitudes were rounded, and  $b$  estimates were obtained using Utsu’s formula. Only multiplying the  $b$  estimates by  $\zeta$ , we reach the actual value.



**Fig. 3** Estimated  $\hat{b}$ , from Utsu's formula, as a function of sample size  $N$ . The magnitude population was randomly generated with  $b = 0.8$ ,  $b = 1.0$ , and  $b = 1.2$ , and noise distributed according to  $N(0, \sigma = 0.1)$  was added. The straight horizontal continuous line represents the true  $b$ -value, and dotted horizontal lines indicate  $\pm 0.05$ . Each symbol represents the mean of 100 realizations, and the vertical bars indicate plus/minus one standard deviation

#### 4 Discussion and conclusions

The effect of rounding on  $b$  estimates made using Utsu's formula is, for typical values of  $b$  and  $\Delta M$ ,  $\sim 10^{-3}$ ; this effect may not be significant for most applications (including those using  $b$  estimates based on small samples), but could be significant when realistic uncertainties in magnitude determination demand using a larger  $\Delta M$ . On the other hand, it is quite easy to bring the error caused by rounding to truly insignificant levels by the correction proposed in (9).

The effect of noise on  $b$  estimates from rounded magnitudes can be, for typical values,  $\sim 10^{-2}$  or larger, so that it could be significant; fortunately, even a rough estimate of the noise level and distribution can be used to approximately correct the estimates through (14) and (15).

**Acknowledgments** This research was partially funded by UNAM-DGAPA postdoctoral scholarship (VH Márquez-Ramírez), by CONACYT grant 222795, and UNAM-DGAPA-PAPIIT grant IN108115. We are grateful for the useful criticism of Angelo De Santis and an anonymous reviewer.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

#### 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
- Bender B (1983) Maximum likelihood estimation of  $b$  values for magnitude grouped data. *Bull Seismol Soc Am* 73:831–851
- Bengoubou-Valerius M, Gibert D (2012) Bootstrap determination of the reliability of  $b$ -values: an assessment of statistical estimators with synthetic magnitude series. *Nat Hazards*. doi:10.1007/s11069-012-0376-1
- De Santis A, Cianchini G, Beranzoli L, Boschi E (2011) The Gutenberg–Richter law and entropy of earthquakes: two case studies in Central Italy. *Bull Seismol Soc Am* 101:1386–1395
- Dwight H (1961) Tables of integrals and other mathematical data. Macmillan, New York

- Enescu B, Ito K (2001) Some premonitory phenomena of the 1995 Hyogo-Ken Nanbu (Kobe) earthquake: seismicity,  $b$ -value and fractal dimension. *Tectonophysics* 338:297–314
- Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. *Bull Seismol Soc Am* 34:185–188
- Kijko A, Sellevoll MA (1992) Estimation of earthquake hazard parameters from incomplete data files. Part II. Incorporation of magnitude heterogeneity. *Bull Seismol Soc Am* 82:120–134
- Mallika K, Gupta H, Shashidhar D, Purnachandra R, Yadav A, Rohilla S, Satyanarayana HVS, Srinagesh D (2013) Temporal variation of  $B$  value associated with  $M \sim 4$  earthquakes in the reservoir-triggered seismic environment of the Koyna–Warna region, Western India. *J Seismol* 17:189–195
- Marzocchi W, Sandri L (2003) A review and new insights on the estimation of the  $b$ -value and its uncertainty. *Ann Geophys* 46(6):1271–1282
- Monterroso DA, Kulhánek O (2003) Spatial variations of  $b$ -values in the subduction zone of Central America. *Geofis Internacional* 4:575–587
- Nuannin P, Kulhanek O, Persson L (2005) Spatial and temporal  $b$  value anomalies preceding the devastating off coast of NW Sumatra earthquake of December 26, 2004. *Geophys Res Lett* 32:L11307. doi:10.1029/2005GL022679
- Rhoades D (1996) Estimation of the Gutenberg–Richter relation allowing for individual earthquake magnitude uncertainties. *Tectonophysics* 258:71–83
- Richter C (1958) *Elementary seismology*. W.H.Freeman, San Francisco, p 768
- Sahu OP, Saikia MM (1994) The  $b$ -value before the 6th August, 1988 India-Myanmar border region earthquake—a case study. *Tectonophysics* 234:349–354
- Scholz CH (1968) The frequency–magnitude relation of micro fracturing in rock and its relation to earthquakes. *Bull Seismol Soc Am* 58:388–415
- Sharma S, Saurabh B, Prakash S, Pabon K, Duarah R (2013) Low  $b$ -value prior to the Indo-Myanmar subduction zone earthquakes and precursory swarm before the May 1995  $M$  6.3 earthquake. *J Asian Earth Sci* 73:176–183
- Shi Y, Bolt B (1982) The standard error of the magnitude-frequency  $b$  value. *Bull Seismol Soc Am* 72:1677–1687
- Smith WD (1981) The  $b$ -value as an earthquake precursor. *Nature* 289:136–139
- Spada M, Tormann T, Wiemer S, Enescu B (2013) Generic dependence of the frequency-size distribution of earthquakes on depth and its relation to the strength profile of the crust. *Geophys Res Lett* 40:709–714
- Tinti S, Mulargia F (1987) Confidence intervals of  $b$ -values for grouped magnitudes. *Bull Seismol Soc Am* 77:2125–2134
- Utsu T (1965) A method for determining the value of  $b$  in a formula  $\log n = a - bM$  showing the magnitude-frequency relation for earthquakes. *Geophys Bull Hokkaido Univ* 13:99–103
- Wiemer S, Benoit J (1996) Mapping the  $b$ -value anomaly at 100 km depth in the Alaska and New Zealand subduction zones. *Geophys Res Lett* 23:1557–1560
- Wiemer S, Wyss M (2000) Minimum magnitude of complete reporting in earthquake catalogs: examples from Alaska, the western United States, and Japan. *Bull Seismol Soc Am* 90:859–869
- Wyss M (1973) Towards a physical understanding of the earthquake frequency distribution. *Geophys J Roy Astron Soc* 31:341–359
- Zúñiga FR, Wyss M (1995) Inadvertent changes in magnitude reported in earthquake catalogs: influence on  $b$ -value estimates. *Bull Seismol Soc Am* 85:1858–1866