Sampling uncertainties and source b likelihood for the Gutenberg-Richter b value from the Aki-Utsu method

F. A. Nava, L. Ávila-Barrientos, V. H. Márquez-Ramírez, I. Torres & F. R. Zúñiga

Springer

Springer

CrossMark

ORIGINAL ARTICLE

Sampling uncertainties and source *b* likelihood for the Gutenberg-Richter *b* value from the Aki-Utsu method

F. A. Nava · L. Ávila-Barrientos ·
V. H. Márquez-Ramírez · I. Torres · F. R. Zúñiga

Abstract The Aki-Utsu method of Gutenberg-Richter (G-R) *b* value estimation is often misapplied so that estimations not using the G-R histogram are often meaningless because they are not based on adequate samples. We propose a method to estimate the likelihood $\Pr(b\ b_m, N, M_1, M_2)$ that an observed b_m estimate, based on a sample of N magnitudes within an $[M_1 - \leq \Delta M/2, M_2 + \Delta M/2)$ range, where $\Delta M = 0.1$ is the usual rounding applied to magnitudes, is due to a "true" source *b* value, *b*, and use these likelihoods to estimate source *b* ranges corresponding to various confidence levels. As an example of application of the method, we estimate the *b* values before and after the occurrence of a 7.4-magnitude earthquake in the Mexican subduction zone, and find a difference of 0.82 between them with 100% confidence that the *b* values are different.

F. A. Nava (✉) · L. Ávila-Barrientos
Sismología, Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, B.C., Mexico
e-mail: fnava@cicese.mx

V. H. Márquez-Ramírez · F. R. Zúñiga
Centro de Geociencias, Universidad Nacional Autónoma de México, Juriquilla, Qro, Mexico

I. Torres
Fac. Ingeniería, Benemérita Universidad Autónoma de Puebla, Puebla, Mexico

1 Introduction

Large earthquakes can have dire consequences for society, including loss of lives, financial losses, and disruptions in important activities. Since earthquake occurrence cannot be prevented or predicted deterministically, probabilistic forecasting is of great importance for risk reduction. Among the statistical/probabilistical tools used for forecasting, one of the oldest and most widely used is the Gutenberg-Richter (G-R) relation (Ishimoto and Iida, 1939; Gutenberg and Richter, 1944; Richter, 1958).

The G-R relation states that the number of earthquakes with magnitudes greater than or equal to a given magnitude M is distributed as

$$\log_{10} N(M) = a_1 - b(M - M_1); \quad M \geq M_1 \qquad (1)$$

where M_1 is a lower threshold below which the number of observed magnitudes does not behave according to Eq. (1) because of lack of coverage (Fig. 1). The parameter a_1 is the logarithm of the total number of recorded earthquakes with magnitude $M \geq M_1$ and depends on the seismicity rate and the length of the observation time. The parameter b relates the relative numbers of small to large magnitudes and is typically ~1.

The G-R relation has been widely used for seismic hazard and risk estimations (e.g., Bender, 1983) because it gives direct estimates of the occurrence ratio of earthquakes with magnitudes above a given value. An inverse relationship between the stress level in a given region and the local *b* value has been proposed (e.g., Scholz, 1968; Ghosh et al., 2008), and several studies
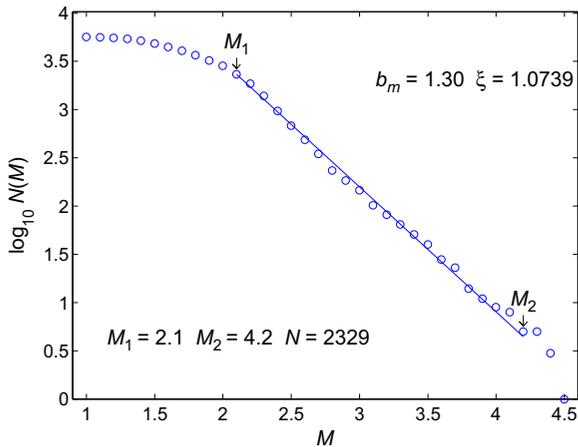
Fig. 1 Example of a G-R distribution histogram showing the approximate locations of the linear range limits M_1 and M_2. The data come from the Southern California Earthquake Center (SCEC) catalog for a rectangle with vertices (33.37° N, 117.04° W), (33.77° N, 116.75° W), (33.12° N, 115.86° W), and (32.72° N, 116.15° W), from October 7, 1985, to June 5, 1997

report possibly premonitory changes in the values of b before large earthquakes (e.g., Shaw et al., 1992; Wyss and Wiemer, 2000; Enescu and Ito, 2001; Márquez-Ramirez, 2012;, Smith, 1981, 1986). Regional variations of b have been used to identify zones where large ruptures could be expected (e.g., Wiemer and Wyss, 1997; Zúñiga and Wyss, 2001, Montuori et al., 2010). The b value has been employed in seismic fractality studies since Aki (1981) proposed a relation between b and the fractal dimension of fault planes (e.g., Singh et al., 2009).

The applications of b in seismology, particularly in the field of seismic risk estimation, indicate the necessity of estimating b correctly in order to obtain significant and reliable results. Unfortunately, it is almost impossible to determine just how reliable most reported b values are. In what follows, we will address the problem of reliability associated with sampling size and propose a way to make b value estimates more useful.

2 The Aki-Utsu maximum likelihood estimate and the problem

Equation (1) implies that magnitudes are distributed exponentially as

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

where $\beta = b\ln(10)$ (c.f. Lomnitz, 1974), and it is a well-known property of the exponential distribution that β is related to the mean of the distribution, μ, as $\beta = (\mu - M_1)^{-1}$ (c.f. Parzen, 1960), so that

$$b = \frac{\log_{10}e}{\mu - M_1}. \tag{3}$$

Aki (1965) showed that the maximum likelihood estimate of b is given by

$$b_m = \frac{\log_{10}e}{\overline{M} - M_1}. \tag{4}$$

In practice, magnitudes are rounded, usually to $\Delta M = 0.1$, so that the actual minimum magnitude is $M_1^U = M_1 - \Delta M/2$ and, instead of formula (4), Utsu's (1965) formula

$$b_m = \frac{\log_{10}e}{\overline{M} - M_1^U}. \tag{5}$$

should be used. We will consider magnitudes rounded to $\Delta M = 0.1$ and assume that the observed data will be likewise rounded.

Formula (5), which we will refer to as the Aki-Utsu method, has been widely used as a simple and straightforward way of estimating b directly from the magnitude mean, without the need of constructing and fitting a G-R magnitude histogram. Some people take a sample of magnitudes, which may be quite small, evaluate the average magnitude without asserting whether the magnitudes belong to the linear range where Eq. (1) is valid, and somehow selecting a value for M_1 (or simply ignoring it), obtain an estimate b_m.

From Eqs. (3) and (5), it is clear that b_m will approximate the actual b value only if $\overline{M}$ is reasonably close to the true μ. This means that $\overline{M}$ should be estimated from an adequate sample, i.e., one that includes enough data. Although this requirement should be obvious, in too many cases, $\overline{M}$ is estimated from too small samples, either because data are scarce or because no attention is paid to the matter of representativity. Nava et al. (2017a), Kramer (2014), and Felzer (2006), among others, have shown that small samples can result in b value estimates that differ widely from the true b values, and their results indicate that a very large number of reported b value estimates are not trustworthy because they are based on too few data. Other estimates do not

state how many data were used, so that it is impossible to evaluate how trustworthy they are.

A second factor that is commonly a source of error is that for magnitudes above some M_2 magnitude the G-R histogram is not linear (Fig. 1) so that Eq. (1) does not apply; magnitudes slightly above or below $\log_{10}N(M)\sim 0$ are either under- or over-sampled (Nava et al., 2017b), so that including them in the $\overline{M}$ estimation is another source of error.

A further complication of estimating $\overline{M}$ without recourse to the G-R histogram is that it is impossible without it to make sure that the sampled magnitudes do distribute according to Eq. (1), i.e., belong to the $M_1^U \leq M \leq M_2^U$ range (where $M_2^U = M_2 + \Delta M/2$). Without the G-R histogram of each particular sample, M_1 cannot be properly estimated, and the effects of the M_2 limit are not considered, so that the result of the direct, indiscriminate, use of formula (5) is, to say the least, questionable.

Nava et al. (2017b) suggest that the problem of over- or under-representation of large magnitudes in experimental samples can be corrected by identifying the M_2 magnitude, eliminating magnitudes above it from the sample, and adding to the measured mean the expected contribution from all magnitudes above M_2. This correction is needed by formula (5), which assumes magnitudes to be distributed as in Eq. (2), independently of whether large magnitudes do, or do not, behave like that in reality (e.g., Kagan, 2002; Lomnitz-Adler and Lomnitz, 1979; Kijko, 1982, 2004; Sornette and Sornette, 1999; Sornette, 2009). However, this correction requires that the measured mean magnitude, $\overline{M}_{<M_2}$, from $M \leq M_2$ be correct, and for small- or medium-sized samples, both $\overline{M}$ and $\overline{M}_{<M_2}$ vary wildly, as noted by Nava et al. (2017a).

In order to apply correctly the Aki-Utsu method, samples should be representative of the linear range of the G-R histogram and the limit magnitudes M_1 and M_2 should be correctly identified. However, both samples and the linear range seldom are as large as would be desired; in what follows, we propose a way to improve the b value estimation and to quantify its uncertainty.

3 Source likelihood

Nava et al. (2017a) used a Monte Carlo method to obtain the probabilities that a given b would result in some measured value b_m, $\Pr(b_m | b)$; here we will use Monte Carlo methods, in a different way, to estimate the likelihood, $\Pr(b_m | b)$, that a given measured b_m results from a sample taken from a population that has an actual (or *true*) or *source* b value, which will be henceforward denoted simply as b.

Now, suppose that a given b_m has been obtained from a sample of size N corresponding to the linear part of an observed G-R distribution between magnitude limits M_2^U and M_2^U. To obtain the likelihood of this being a sample from a population with a given source b, a large number N_r of realizations of N synthetic exponentially distributed magnitudes in the $[M_1^U, M_2^U)$ range are generated for different values of b as

$$M = M_1^U - \ln\left[r(1-r_2) + r_2\right]/\beta, \qquad (6)$$

where $r_2 = \exp[-(M_2 - M_1)/\beta]$ and r is a uniformly distributed pseudo-random number in the (0,1) interval from the Matlab *rand* function using the *Twister* generator. For each realization, the synthetic magnitudes are rounded to $\Delta M = 0.1$, to correspond to the usual practice for observed data, and then the b_m value for each realization is estimated using (5). This process is repeated for all source b values for which some sample realizations result in b_m. A histogram of the number of times that b_m is the result of a realization for each b value is built, and all b values for which at least one b_m value is observed are considered; on normalizing this distribution by the total number of b_m appearances, the distribution of source b likelihoods is obtained.

We use source b values starting at b_m and increasing or decreasing by some Δb; evaluations in each direction are stopped after realizations with no b_m are obtained.

Notice that, since the b likelihoods are obtained using the same magnitude range that was used for b_m, no correction for unsampled large magnitudes is needed and that, since the observed sample size is used throughout, the uncertainty associated with sample size is automatically taken into account.

Of course, in order to appropriately estimate the likelihoods, it is essential to know M_1 and M_2, as well as N; the source likelihoods, strictly speaking, are $\Pr(b | b_m, N, M_1, M_2)$, but we will continue to denote them simply as $\Pr(b | b_m)$.

Figure 2 shows an isoprobability contour plot of $\Pr(b | b_m)$ likelihoods for a representative 0.8 to 1.2 b_m range, considering a sample length $N=100$ and $\Delta b = 0.02$ precision, for a magnitude range from $M_1 = 3.0$ to an upper limit estimated from Eq. (6) as $M_2 = \ln$

Δb=0.02 N=100 M_1=3 N_{M_2}=1 N_r = 100000
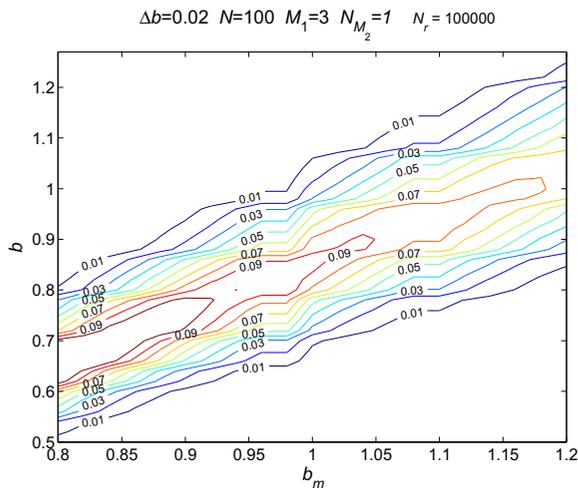
Fig. 2 Isoprobability contour plot of Pr(b b_m) source likelihoods

$(N_{M_2}/\beta) + M_1^U$, where $N_{M_2} = 1$ is the minimum number of occurrences; $N_r = 100,000$ realizations were done for each b. Since M_2 depends on b, for large b values large magnitudes area scarce, source b likelihoods have a larger spread for large b_m values.

The contour plot gives a general idea of how likelihoods distribute, but of course, in practice, the b likelihoods should be estimated for exactly the same sample length and magnitude range used for estimating b_m. Notice that this implies doing the G-R histogram of all the available data to select, as best as possibly, the magnitude limits, and then applying Eq. (5) to the magnitudes within these limits.

Figure 3 shows an example of the b likelihood distribution for synthetic magnitudes generated using Eq. (6) for a source value $b = 1.0$; these magnitudes yielded

a sample of $N = 1494$ data in the $M_1 = 3.0$ to $M_2 = 5.3$ linear G-R range, which resulted in a measured $b_m = 1.02$. The likelihoods were estimated from $N_r = 25,000$ realizations for each b, considering a precision of $\Delta b = 0.01$, and those classes in the distribution with likelihoods larger than 0.075 the maximum are listed in Table 1.

The measured $b_m = 1.02$ is indicated in Fig. 3 by a thin dashed vertical line, and from the likelihood distribution, the following likelihood probabilities for $\Delta b = 0.01$ are obtained: $\Pr(b = b_m | b_m) = 0.114$, $\Pr(b \neq b_m | b_m) = 0.886$, $\Pr(b < b_m | b_m) = 0.674$, and $\Pr(b > b_m | b_m) = 0.211$.

However, the most important feature is that of the maximum source likelihood, corresponding to a b value that we will denote by b_x. It is this most likely value, shown as a thick dash-dot vertical line in Fig. 3, which does not always coincide with the measured b_m, which should be used as an estimate of the real b. In this example, although $b_m = 1.02$, the most likely source b value is $b_x = 1.00$ (which happens to be the b value used to generate the synthetic sample), and for $\Delta b = 0.01$ precision $\Pr(b = b_x b_m) = 0.142$, $\Pr(b \neq b_x | b_m) = 0.858$, $\Pr(b < b_x b_m) = 0.395$, and $\Pr(b > b_x b_m) = 0.462$.

Centered around b_x, the *confidence intervals*, i.e., the b ranges corresponding to given likelihoods of including the true b value, are $\Pr(b$ $b_m) = 0.50^+$ for $0.99 \leq b \leq 1.02$, $\Pr(b$ $b_m) = 0.75^+$ for $0.97 \leq b \leq 1.03$, and $\Pr(b$ $b_m) = 0.90^+$ for $0.96 \leq b \leq 1.05$ where the $+$ sign indicates that the likelihood contained in the interval is equal to or a little larger than the number it follows. Hence, we can say that, based on the measured b_m, the sample size, and the magnitude range, for this example

Fig. 3 Illustration of likelihood histogram (light blue line) for $\Delta b = 0.01$. The observed $b_m = 1.02$ is shown as a thin dashed vertical red line, and the most likely b value, $b_x = 1.00$, is shown as a thick dash-dot vertical black line; the limits of the 0.75 and 0.90 likelihood confidence ranges are indicated by short, thick vertical black lines with crosses and asterisks, respectively

b_m=1.02 N=1494 Δb=0.01 M_1=3 M_2=5.3 N_r = 25000

Table 1 Values of the source b likelihoods for a measured $b_m = 1.02$ estimated from $N_r = 25{,}000$ realizations of $N = 1494$ magnitudes in the $M_1 = 3.0$ to $M_2 = 5.3$ magnitude range with $\Delta b = 0.01$. Only likelihoods larger than 0.075 times the maximum are listed

b	0.94	0.95	0.96	0.97	0.98	0.99	1.00
$Pr(b\|b_m)$	0.011	0.025	0.045	0.073	0.103	0.131	0.142
b	1.01	1.02	1.03	1.04	1.05	1.06	
$Pr(b\|b_m)$	0.137	0.114	0.089	0.059	0.033	0.018	

and with $\Delta b = 0.01$ precision we can estimate the true source b with 50% confidence within a 0.03 range, with 75% confidence within a 0.06 range, and with 90% confidence within a 0.09 range. The 0.75^+ and 0.90^+ intervals are indicated in Fig. 3 by short vertical lines with crosses and asterisks in the extremes, respectively. Although any confidence level may be considered, we present these three representative ones and no lower ones because it would be absurd to state estimates having less than an even chance of being true.

A word about Δb: we can work with any precision we like; the smaller Δb is, the more detail it furnishes, but more detail than necessary to adequately define the confidence intervals may be useless, because detail does not significantly alter the actual length of the confidence intervals, and it requires much more computation because many more b values have to be tried and the number of realizations for each of these values must be larger to achieve sufficient numbers in each histogram class. Usually, $\Delta b \approx 0.01$ is appropriate for samples larger than about 1000, $\Delta b \approx 0.02$ for samples around 500, and $\Delta b \approx 0.05$ for samples around 300.

Figure 4 shows an example of a smaller $N = 498$ synthetic sample in the $M_1 = 3.0$ to $M_2 = 4.9$ range,

which resulted in $b_m = 1.08$ with $\Delta b = 0.02$. The figure conventions are the same as for the previous figure. Significant likelihoods are listed in Table 2.

In this example, for $\Delta b = 0.02$, although $b_m = 1.08$, the most likely source b value is $b_x = 1.04$ (a value closer to the actual $b = 1.00$) and $Pr(b = b_x|b_m) = 0.151$, $Pr(b \neq b_x | b_m) = 0.849$, $Pr(b < b_x|b_m) = 0.402$, and $Pr(b > b_x|b_m) = 0.447$. The *confidence intervals* are $Pr(b|b_m) = 0.50^+$ for $1.02 \leq b \leq 1.08$, $Pr(b|b_m) = 0.75^+$ for $0.98 \leq b \leq 1.10$, and $Pr(b|b_m) = 0.90^+$ for $0.96 \leq b \leq 1.12$, considerably wider than the corresponding intervals for the larger sample shown above.

Finally, we will illustrate how to use the likelihood estimates to interpret observations and evaluate their significance. Suppose we are exploring the possibility of b changing over time in some region of interest; if our data consists of the two first columns of Table 3, as would be reported in most papers, the data would plot versus time as indicated by the circles in Fig. 5 (two of the circles are obscured by the diamonds that coincide with them), and it would appear that b is indeed increasing monotonically, conclusions would be drawn, and perhaps a seismic hazard estimate would be made on the strength of these observations.

Fig. 4 Illustration of likelihood determination histogram (light blue line) for $\Delta b = 0.02$. The observed $b_m = 1.08$ is shown as a thin dashed vertical line, and the most likely b value, $b_x = 1.04$, is shown as a thick dash-dot vertical line; the limits of the 0.75 and 0.90 likelihood confidence ranges are indicated by short, thick vertical lines with crosses and asterisks, respectively

Table 2 Values of the source b likelihoods for a measured $b_m = 1.08$ estimated from $N_r = 25,000$ realizations of $N = 494$ magnitudes in the $M_1 = 3.0$ to $M_2 = 4.9$ magnitude range with $\Delta b = 0.02$. Only likelihoods larger than 0.075 times the maximum are listed

b	0.94	0.96	0.98	01.00	1.02	1.04
$Pr(b \mid b_m)$	0.021	0.043	0.075	0.110	0.139	0.151
b	1.06	1.08	1.10	1.12	1.14	1.16
$Pr(b \mid b_m)$	0.141	0.115	0.084	0.053	0.030	0.014

However, if we take into account the information about N, M_1, and M_2, the resulting likelihood distributions would tell us a somewhat different story. The most likely values b_x (shown as filled diamonds) from each measurement do not, except for the two cases that had the largest samples, coincide with the corresponding b_m (shown as circles). The latest three b_x values appear to be definitely higher than the first two, although the last datum could indicate that b is no longer increasing. Be that as it may, the question is how significant is the information in these data. Two horizontal (green) lines indicate the greatest lower bound (inf) and the least upper bound (sup) of the 90% confidence intervals, and the band between them is a range of values that could be common to all data; hence, we cannot discard the possibility that b is stationary (or perhaps even slowly decreasing!) at this confidence level. A similar exercise with the inf and sup for the 75% confidence level bands tells us that we can (barely) be sure that the fourth (highest) value is indeed higher than the first one, but we cannot tell, at this confidence level, whether it effectively differs from any of the other three data closer to it in time. Finally, we can tell that the earlier data are indeed lower than the later ones, but only at the 50% confidence level. Thus, we have information about the level of significance of the differences between

Table 3 List of b measurements made over time and of their characteristics

T (years)	b_m	N	M_1	M_2
3.0	0.94	489	3.0	5.4
6.0	0.96	495	3.1	5.2
8.5	0.98	848	3.0	6.0
12.0	1.00	700	3.0	5.8
14.0	1.02	477	3.2	5.3
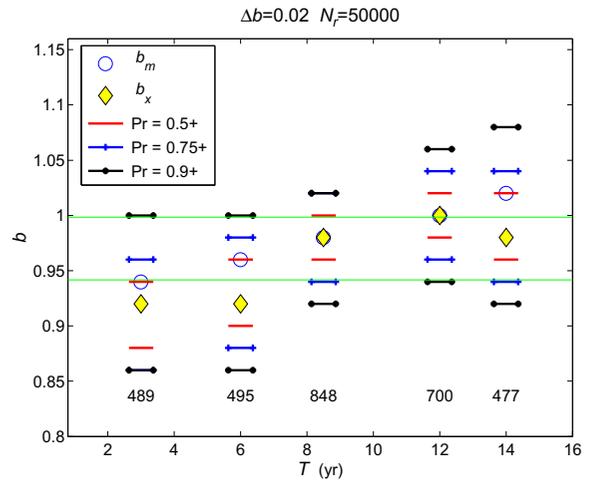
Fig. 5 Significance application of confidence intervals. Circles indicate the measured values, diamonds are the more likely source b values, and horizontal short lines indicate three different confidence intervals. Horizontal green lines show the inf and the sup of the 90% confidence intervals

observations, which we can use to accept or reject a conclusion depending on the required confidence.

As is apparent in the shown figures, the likelihood confidence intervals are larger for smaller samples; an overview of interval lengths for different sample sizes corroborates the results of Nava et al. (2017a) that indicate that estimates from very small samples are useless, because interval lengths are so large that the real value associated with the measurement could be almost anything. Hence, reported estimates of b that do not give at least sample size information are completely useless, and it is impossible to quantify the significance of the conclusions based on them. Estimates should always be accompanied by data about sample sizes and, if possible, about the magnitude range used in the determination, which means that no estimates should be obtained directly from the data mean without making sure that they are based on data that do conform to the G-R distribution.

4 Application to the seismicity of the Mexican subduction zone

To illustrate the application of the proposed method, we selected the seismicity occurring in a region containing the March 20, 2012, $M = 7.4$ earthquake that occurred near the border of the Guerrero and Oaxaca states (rectangle, Fig. 6). The area was chosen perpendicular to the

Fig. 6 Mexican trench. Dots are epicenters for $M \geq 3.3$ from 1988 to 2016, diamonds indicate earthquakes with $M \geq 7.0$, and the star is the 2012 $M = 7.4$ earthquake. The rectangular box contains the study area where it can be seen that most of the events are clustered near the main event epicenter
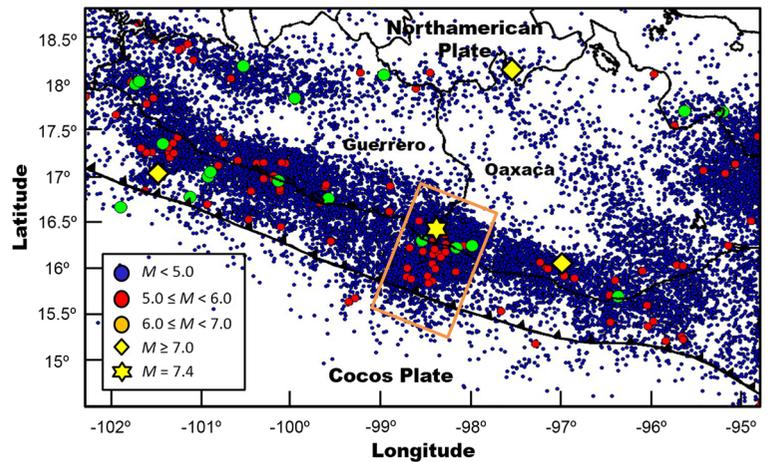
trench so as to include seismicity due to the stresses that eventually caused the $M = 7.4$ earthquake.

The seismicity in most of the Mexican Trench is due to the subduction of the Cocos plate beneath the North American one, and this subduction zone is the site of most of the largest earthquakes in Mexico.

The database employed for the b value determinations was obtained from Mexico's National Seismological Service (*Servicio Sismológico Nacional*, SSN, http://www2.ssn.unam.mx:8080/catalogo/) and spans from 1988 to 2016, for $M \geq 3.3$; the chosen earthquake is the largest one in the catalog.

The area within the rectangle of Fig. 6 was chosen as the one most likely to reflect the effects on b

of the stress accumulation and release associated with the $M = 7.4$ earthquake. Figure 7 shows the cumulative number of earthquakes and the corresponding cumulative seismic moment release; periods in which the slope of the cumulative number remains constant are periods where coverage was probably also constant, so that samples taken within these periods should not reflect effects that are artifacts of changes in coverage.

To see whether the b value changed from before to after the earthquake, we analyzed separately the seismicity before the earthquake (September 13, 1995, to March 20, 2012) and that beginning after the majority of aftershocks, and extending to the end of 2015 (August

Fig. 7 Cumulative number of earthquakes (top) and cumulative seismic moment release vs. time for the earthquakes shown in Fig. 6. The occurrence of the 2012 $M = 7.4$ earthquake is indicated by a thick vertical line (top) and by the step in moment release (bottom). The pre-earthquake time window extends from the thin vertical solid line at 1995.7 to just before the main event, and the post earthquake time window extends from the dashed vertical line to the end of the sample
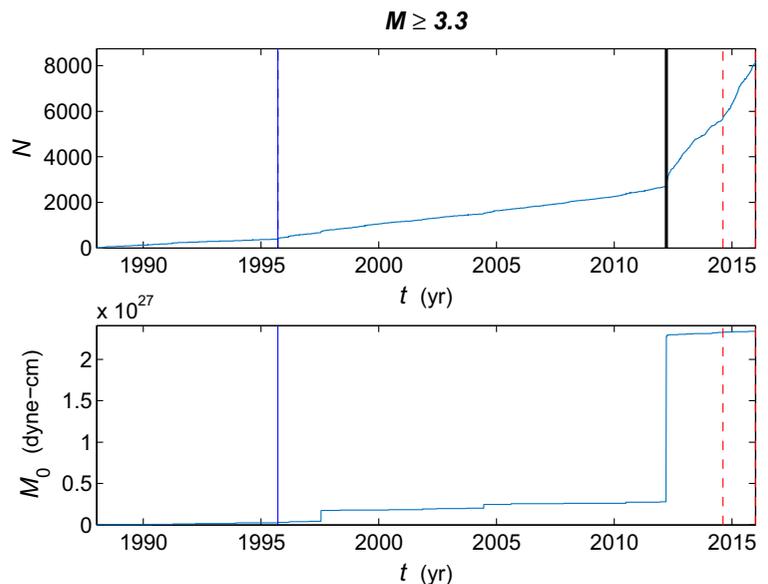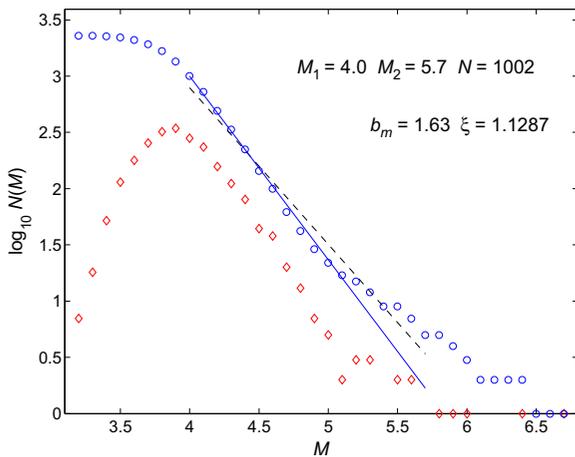
Fig. 8 G-R histogram for the seismicity before the earthquake (circles) and non-cumulative histogram (diamonds). The solid straight line corresponds to the $b_m = 1.63$ from the Aki-Utsu method, and the dotted line is the least squares fit

13, 2014, to December 31, 2015); these time windows are shown in Fig. 7.

Figure 8 shows the G-R histogram for the seismicity before the earthquake and shows the straight line fit by the Aki-Utsu method and by least squares. Also shown are the limits of the magnitude range $M_1 = 4.0$ and $M_2 = 5.7$, and the number of magnitudes $N = 1002$ used to determine $b_m = 1.63$, as well as the ratio of the measured standard deviation to the magnitude mean, $\xi = 1.1287$, which should be unity for an exponential distribution.

Figure 9 shows the source likelihoods $\Pr(b\,b_m)$ determined from b_m, M_1, M_2, and N. The most likely source b, $b_x = 1.64$, is slightly higher than b_m. Also shown are the 75% $1.58 \leq b \leq 1.70$ and 90% $1.55 \leq b \leq 1.72$ confidence intervals.

The G-R histogram for the seismicity after the earthquake, contained within the time window shown at the extreme right of Fig. 8, August 13, 2014, to December 31, 2015, is shown in Fig. 10. Also shown are the limits of the magnitude range $M_1 = 3.7$ and $M_2 = 4.7$, the number of magnitudes $N = 792$ used to determine $b_m = 2.42$, and $\xi = 1.1934$ that indicates that after the earthquake the magnitude distribution was less exponentially distributed than before the earthquake.

The source likelihood distribution $\Pr(b\,b_m)$ for events after the earthquake is shown in Fig. 11. The most likely source b, $b_x = 2.46$, is slightly higher than b_m. Also shown are the 75% $2.36 \leq b \leq 2.56$ and 90% $2.30 \leq b \leq 2.60$ confidence intervals.

Thus, we found that the b value changed drastically after the occurrence of the $M = 7.4$ earthquake, from 1.50 to 2.40, and since the distributions shown in Figs. 9 and 11 do not overlap at all, we can tell that the change is significant with 100% confidence. This behavior agrees with the hypothesis that b should be lower for regions where stress is high, than for regions where stress has been depleted by the occurrence of a large earthquake (e.g., Wyss, 1973).

It may be objected that the b values found above are too large, since they are greater than the 1.5 maximum value (Olsson 1999), but it should be taken into account that the said limit applies to moment magnitudes; other magnitude scales, such as the one used by the SSN, are not necessarily related to the seismic moment in the same way, so that the b values based on them can have a different maximum value (note that the b estimate for southern California in Fig. 1 does have a value below 1.5). The important facts are the following: the SSN

Fig. 9 Source likelihoods $\Pr(b\,b_m)$ for seismicity before the main event. The thin, dashed, vertical line indicates the measured b_m, and the thick dash-dot vertical line indicates the most probable source value b_x. The short, vertical lines with crosses or asterisks in the extremes are the limits of the 75 and 90% confidence intervals, respectively
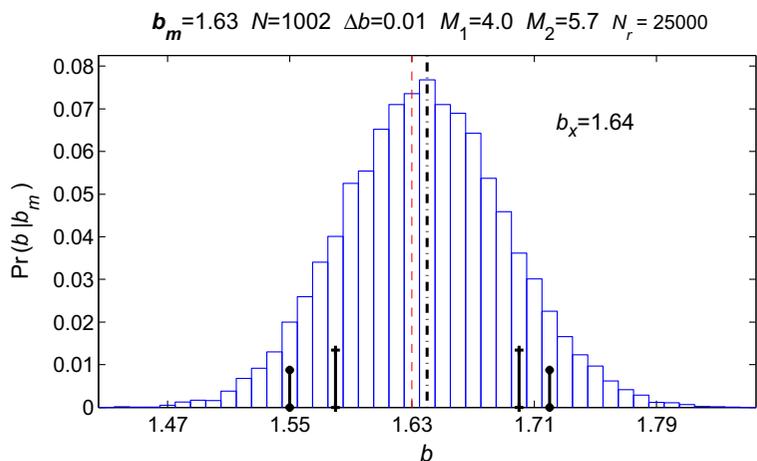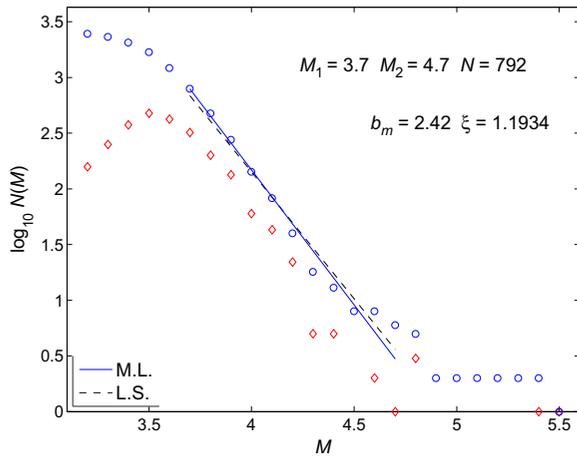
Fig. 10 G-R histogram for the seismicity after the earthquake and the majority of aftershocks (circles) and non-cumulative histogram (diamonds). The solid straight line corresponds to the $b_m = 2.42$ from the Aki-Utsu method, and the dotted line is the least squares fit

scale is consistent, in that it distributes according to the G-R relation, and the estimated b_x values do change significantly from before to after the $M = 7.4$ earthquake, so that the data from the SSN catalog can be useful for hazard estimation purposes.

5 Discussion

The apparent ease of b value estimation using the Aki-Utsu formula has resulted in many estimates the validity of which is impossible to evaluate; probably, most of them are erroneous. It is absolutely necessary to know at least the number of magnitudes used to estimate a given b_m in order to have some idea of how reliable the estimate is, so we urge that all reports of b value estimates should ensure that the sampled magnitudes do correspond to the linear range of the observed G-R distribution and should state how many samples were used for each, together with the limits of the above-mentioned range.

We present a method to find, based on the number of magnitudes in the sample, N, and on the limits of the magnitude range, M_1 and M_2, the likelihood that the measured b_m comes from a population having a true source b value b, $\Pr(b \, b_m, N, M_1, M_2)$. Along with the estimated likelihoods comes the possibility of identifying the most likely source b, b_x, and of estimating b ranges corresponding to various confidence levels.

The Monte Carlo approach to source b likelihoods estimation is necessary since a Bayesian approach is not possible because prior b probabilities are not known and cannot be assumed to be uniform.

Since the source b likelihoods are obtained using the same sample length and magnitude range used to get the original estimate, they automatically incorporate effects associated with these parameter values, and so need no further corrections.

The source b likelihoods permit knowing the confidence with which differences in estimated b values can be considered to be significant.

Application of the source likelihoods method to the estimation of b values both before and after the occurrence of an $M = 7.5$ in the Mexican subduction zone allows us to affirm, with 100% confidence, that there

Fig. 11 Source likelihoods $\Pr(b \, b_m)$ for seismicity after the main event. The thin, dashed, vertical line indicates the measured b_m, and the thick dash-dot vertical line indicates the most probable source value b_x. The short, vertical lines with crosses or asterisks in the extremes are the limits of the 75 and 90% confidence intervals, respectively
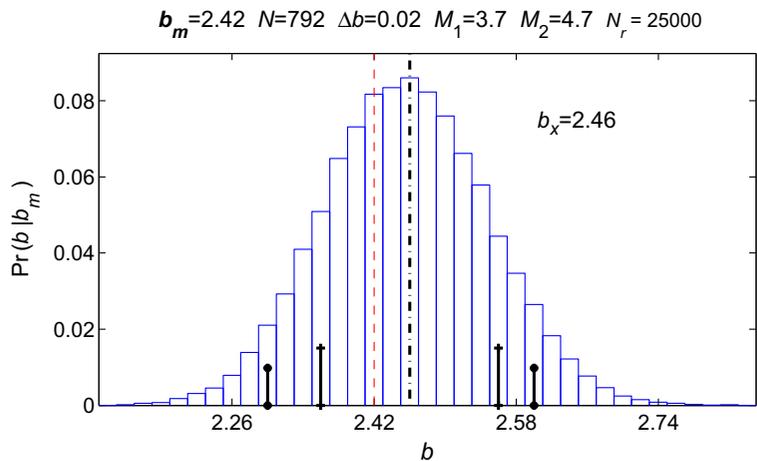
was a significant change ≈0.82 from before to after the earthquake. Corroboration of this result and its implications for seismic hazard assessment are beyond the scope of this work and are to be explored in the future.

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

Aki K (1981) A probabilistic synthesis of precursory phenomena. In: Earthquake prediction and international review, Simpson D and Richards P (eds), Am Geophys U, pp 566–574

Bender B (1983) Maximum likelihood estimation of b values for magnitude grouped data. Bull Seism Soc Am 73:831–851

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

Felzer K (2006) The Gutenberg-Richter *b* value http://pasadena. wr.usgs.gov/office/kfelzer/AGU2006Talk.pdf

Ghosh A, Newman A, Amanda M, Thomas A, Farmer G (2008) Interface locking along the subduction megathrust from b-value mapping near Nicoya Peninsula, Costa Rica. Geophys Res Lett 35:L01301. https://doi.org/10.1029/2007GL031617

Gutenberg B, Richter C (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 34:185–188

Ishimoto M, Iida K (1939) Observations sur les seisms enregistrés par le microseismograph construit dernièrement (I). Bull Earthquake Res Inst Univ of Tokyo 17:443–478

Kagan Y (2002) Seismic moment distribution revisited: I. Statistical results. Geophys J 148(3):520–541

Kijko A (1982) A comment on "A modified form of the Gutenberg-Richter magnitude-frequency relation" by J. Lomnitz-Adler and C. Lomnitz. Bull Seism Soc Am 72, 1759-1762

Kijko A (2004) Estimation of the maximum magnitude earthquake mmax. Pure Appl Geophys 161:1–27. https://doi.org/10.1007/s00024-004-2531-4

Kramer Y (2014) Minimum sample size for detection of Gutenberg-Richter's b-value. arXiv:1410.1815 [physics. geo-ph]

Lomnitz C (1974) Global tectonics and earthquake risk. Elsevier Sc. Pub. Co., CH

Lomnitz-Adler J, Lomnitz C (1979) A modified form of the Gutenberg-Richter magnitude-frequency relation. Bull Seism Soc Am 69:1209–1214

Márquez-Ramírez V (2012) Patrones espacio-temporales de sismicidad previos al sismo de Armería, Colima, 22 enero 2003, M_w=7.4. MSc. Earth Sciences, CICESE

Montuori C, Falcone G, Murru M, Thurber C, Reyners M, Eberhart-Phillips D (2010) Crustal heterogeneity highlighted by spatial b-value map in the Wellington region of New Zealand. Geophys J Int 183:451–460. https://doi.org/10.1111/j.1365-246X.2010.04750.x

Nava F, Márquez-Ramírez V, Zúñiga F, Ávila-Barrientos L, Quinteros-Cartaya C (2017a) Gutenberg-Richter *b*-value maximum likelihood estimation and sample size. J Seismol 21: 127-135. Published online 02 June 2016. doi: https://doi.org/10.1007/s10950-016-9589-1

Nava F, Márquez-Ramírez V, Zúñiga F, Lomnitz C (2017b) Gutenberg-Richter *b*-value determination and large magnitudes sampling. Natural Hazards, DOI: https://doi.org/10.1007/s11069-017-2750-5, published online January 21, 2017

Olsson R (1999) An estimation of the maximum *b*-value in the Gutenberg-Richter relation. Geodynamics 27:547–552

Parzen E (1960) Modern probability theory and its applications. J Wiley & Sons Inc, Japan 464pp

Richter C (1958) Elementary seismology. W H Freeman and Co, USA

Scholz C (1968) The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. Bull Seismol Soc Am 58:399–415

Shaw E, Carlson J, Langer J (1992) Patterns of seismic activity preceding large earthquakes. J Geophys Res 97(B1):479–488

Singh C, Singh A, Chadha R (2009) Fractal and b-value mapping in eastern Himalaya and southern Tibet. Bull Seismol Soc Am 99:3529–3533. https://doi.org/10.1785/0120090041

Smith W (1981) The b-value as an earthquake precursor. Nature 289:136–139

Smith W (1986) Evidence for precursory changes in the frequency-magnitude b-value. Geophys J R Astr Soc 86: 815–838

Sornette D, Sornette A (1999) General theory of the modified Gutenberg-Richter law for large seismic moments. Bull Seismol Soc Am 89:1121–1130

Sornette D (2009) Dragon-kings, black swans and the prediction of crises. Swiss Finance Inst Res Paper 09-36:1–18

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, Wyss M (1997) Mapping the frequency-magnitude distribution in asperities: an improved technique to calculate recurrence times? J Geophys Res 102(B7):15115–15128

Wyss M (1973) Towards a physical understanding of the earthquake frequency distribution. Geophys J Royal Astr Soc 31: 341–359

Wyss M, Wiemer S (2000) Change in the probability for earthquakes in southern California due to the Landers magnitude 7.3 earthquake. Science 290:1334

Zúñiga R, Wyss M (2001) Most-and least-Likely locations of large to great earthquakes along the Pacific coast of Mexico estimated from local recurrence times based on *b*-values. Bull Seismol Soc Am 91:1717–1728