Proper *b*-value Estimation Method in Rock Acoustic Emission Testing

Xi. Ling. Liu[1], Meng. Si. Han[1], Wei. He[2], Xi. Bing. Li[1], and Dao. Long. Chen[1]

[1]School of Resources and Safety Engineering, Central South University, Changsha 410083, People's Republic of China.

[2]School of Mathematics and Statistics, Central South University, Changsha 410083, People's Republic of China.

Corresponding author: Xi. Ling. Liu (lxlenglish@163.com)

Key Points:

- The cumulative data counting will inevitably result in deviation of doubly truncated size distribution from the power law.
- Attenuation will not modify the size distribution in a certain interval and *b*-value is theoretically verified to be unchanged.
- A new *b*-value estimation procedure is proposed and applied to a dilation rupturing test, it shows good performance.

Abstract

The power-law relationship has been regarded as the fundamental description of the size distribution from large-scale-earthquake to small-scale-laboratory rock ruptures. However, deviation from the power-law relationship has often been reported, especially when the amplitude distribution is used for *b*-value estimation in rock acoustic emission testing, the effect of attenuation should be considered. Here, we perform a detailed analysis on the deviation of the size distribution from a power law and prove that the cumulative frequency distribution will inevitably result in deviation from the power law. We also discuss modification of the attenuation on the doubly truncated size distribution from a more general perspective. We find that in a certain interval, the attenuation will not modify the size distribution and the *b*-value is theoretically verified to be unchanged. Based on these discussions, we propose a new *b*-value estimation procedure for rock acoustic emission testing and apply it to a dilation rupturing test, the procedure exhibits good performance.

Plain Language Summary

The power-law relationship is the intrinsic characteristic of the size distribution from earthquake to laboratory rock ruptures, indicating that the number of small-scale ruptures is much greater than the number of large-scale ones. Nevertheless, due to many reasons such as the statistical methods, inadequate data acquisition and data truncation, deviation from a power-law of the size distribution is inevitable in practice. In order to understand the deviation of the size distribution in the laboratory rock fracture testing, we perform a detailed analysis of the statistical methods and the attenuation effect on the power-law relationship of the size distribution, and propose a new procedure for scaling parameter estimation of power-law relationship. This newly proposed procedure shows its reliability in a dilation rupturing test and can well ensure the robustness of the scaling parameter estimation.

1 Introduction

The size distribution of earthquakes commonly follows a power-law relationship given by $\log_{10}(N) = a - bM$, where N is the number of earthquakes with a magnitude $\geq M$ and a and b are constants (Gutenberg & Richter, 1944). Here, the parameter b describes the size distribution scaling, which is often referred to as the *b*-value, and the spatial and temporal variation of the *b*-value is always regarded as an essential clue for earthquake precursors.

As *b*-value is a statistical value, its result is influenced by many factors, and the errors in estimation can result in potentially misleading *b*-value variations. One of the most important reason for errors in *b*-value estimation is the fact that the size distribution is not usually log linear in the whole range of magnitudes, and sometimes it follows neither the logarithmic linear nor any smoothly nonlogarithmic linear law but is more complex (Lasocki and Papadimitriou, 2006; Vorobieva et al., 2016). In fact, it is often observed that changes in the size distribution for larger earthquakes can be explained by finite size effects (Casertano, 1982; Clauset, 2007) and those for smaller events are caused by incomplete reporting in earthquake catalogs (Wiemer & Wyss, 2000; Cao & Gao, 2002; Mignan, 2012; Kwiatek et al., 2014; Raub et al., 2017). Meanwhile, errors in *b*-value estimation are also related to the data volume. It is generally considered that a size of 50 events can be adopted as the minimum number of events for stable *b*-value estimation (Schorlemmer et al., 2004; Kurz et al., 2006; Amorese et al., 2010). Other factors such as bin width, the cumulative or incremental data counting, and the estimation algorithm will all influence the *b*-value result (Weichert, 1980; Bender, 1983; Amorese et al., 2010, Nava et al., 2016).

Acoustic emission (AE) generated by fracturing of rock has been found to obey the same form of size distribution, and there have been numerous studies indicating that in many ways the mechanism of earthquake foreshock sequences could therefore be reproduced through AE from the statistical behavior of micro-fracture activity observed in laboratory fracture experiments (Mogi, 1962; Scholz, 1968, 2015; Lockner, 1993; Goebel et al., 2014; Vorobieva et al., 2016). Also the variation of *b*-value obtained in rock AE deformation tests and earthquakes has always been compared to establish analogies in the damage process and precursory analysis (Scholz, 1968; Lockner et al., 1991; Goebel et al., 2013). Because of the acquisition threshold and the limitation of the maximum output voltage in some AE equipment, a similar deviation from the power-law relationship of the size distribution will also appear in laboratory rock AE tests, and the influence of data volume, cumulative or incremental data counting, bin width and estimation algorithm on the *b*-value result is obvious (Cosentino et al., 1977; Lockner, 1993; Liakopoulou et al., 1994).

Furthermore, it is worth noting that unlike the *b*-value in earthquakes which uses the frequency–magnitude distribution for estimation, the apparent frequency–amplitude distribution is mostly used for *b*-value estimation in rock AE tests (Cox & Meredith, 1993; Weiss, 1997; Liakopoulou et al., 1994), and this apparent amplitude is directly derived from the sensor-collected signal that is attenuated from the source. Consequently, the distribution of these apparent amplitudes is not the physical size distribution of the sources, so attenuation could also modify the AE *b*-value (Lockner et al., 1991; Unander, 1993; Weiss, 1997; Lavrov, 2005). Although some researchers have made provisions when using apparent amplitude to estimate *b*-value by adjusting the AE amplitude for $\frac{1}{r}$ geometric spreading and averaging over transducers to give an equivalent event amplitude (Lockner et al., 1991), or by using the root-mean-square principle to obtain a relative AE magnitude (Kwiatek et al., 2014), the amplitude distribution actually cannot fully represent the crack size distribution. The attenuation effect was firstly proposed by Lockner et al. (1991), and then the attenuation effect on *b*-value has been theoretically discussed by Unander (1993) and Weiss (1997). They assumed a constant attenuation coefficient and uniformly distributed sources in the medium, and they considered that the attenuation is composed of a geometric term and a “dissipative” exponential term. In fact, the attenuation coefficient is not constant in the rock material. Not only is it a frequency-dependent parameter but it also varies with the direction of elastic wave propagation because of the anisotropy of the rock, making it extremely difficult to theoretically describe the attenuation in rock.

Since estimation result of *b*-value can be influenced by many factors especially the deviation of size distribution from power-law relationship. Therefore, a reliable *b*-value estimation procedure is critical in rock AE testing which can well ensure the robustness of comparative precursory analysis between laboratory AE tests and earthquakes. Here we perform a detailed analysis of the deviation of the size distribution from power-law relationship, and intend to examine the influence of statistical method on size distribution. Meanwhile we investigate the modification of attenuation on doubly truncated size distribution from a more general perspective which considering that the randomly distributed sources are attenuated to a certain extent and collected by the sensors independent of the composition of the attenuation terms. Based on the deviation discussion, we propose a new *b*-value estimation procedure named

FGS which specify the minimum data volume and statistical method, and employ Fisher optimal split and global search algorithm to determine the logarithmic linear segment in the frequency–amplitude distribution. In order to verify the reliability of FGS method, we design a static dilation rock rupturing AE test by injecting nonexplosive cracking agent into three predrilled boreholes to form a specific fracture surface in cubic rock specimen (as shown in Figure 1). This experimental design is to ensure that the sensor-collected AE signals are all generated by expansion rupturing in rock specimen, and does not rely on source location to identify valid rupturing data.

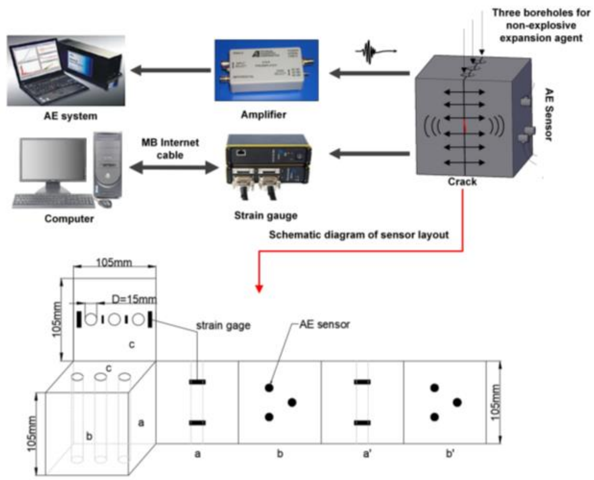


Figure 1. Schematic of the expansion rupturing tests. A nonexplosive expansion agent is injected into three boreholes of each rock specimen, and a fracture surface along the center of the three boreholes is formed. The specimen surfaces labeled by a’ and b’ are the surfaces parallel to a and b, respectively.

2 Deviation from the power-law size distribution

The power-law relationship of the size distribution is an intrinsic characteristic of earthquakes and rock AE events that indicates that the number of small-scale ruptures is much greater than the number of large-scale ones. The power-law relationship can be interpreted as a manifestation of the self-organized critical behavior of earth dynamics (Bak & Tang, 1989; Christensen et al., 2002; Newman, 2005; Clauset, 2007). However, deviation from the power-law size distribution has always been observed in earthquake and rock AE experiments (Lockner, 1993; Cox & Meredith, 1993; Liakopoulou et al., 1994; Weiss, 1997; Main, 2000). One of the reasons for this deviation is inadequate data acquisition. At the lower magnitude end, this is apparent from the difficulty in determining the magnitude of completeness, *M*c (Cao & Gao, 2002; Mignan, 2012; Godano et al., 2014; Raub et al., 2017; Radziminovich et al., 2019). Meanwhile, small-scale ruptures can be masked by larger ones, which is also a very significant factor for inadequate data acquisition in rock AE experiments when avalanche destruction occurs during the final loading stage or in high-strain-rate loading tests (Meredith & Main, 1990; Cox and Meredith, 1993; Liakopoulou et al., 1994).

Another reason for this deviation is due to the statistical methods used in *b*-value estimation. In fact, because of the accuracy of equipment acquisition and the finite size of earthquake and AE amplitudes, the *b*-value is estimated in a doubly truncated exponential distribution (Page, 1968; Cosentino et al., 1977). Generally, the density function of a doubly truncated distribution generated by an underlying exponential one can be expressed as

$$f(x) = \begin{cases} \alpha e^{-\beta x} & a_0 \leq x \leq a_x \\ \text{other} & \text{otherwise} \end{cases}, \tag{1}$$

where α and β are constants, $\beta = b\ln(10)$, and "other" is any other distribution that is excluded from consideration. Then, the cumulative frequency distribution in the interval $[a_0, a_x]$ is $N(x) = N_{total}\int_x^{a_x} \alpha\, e^{-\beta x} dx = N_{total}\frac{\alpha}{\beta}\left(e^{-\beta x} - e^{-\beta a_x}\right)$, where N_{total} is the total number of events in the catalog. If the cumulative frequency distribution is expressed again on a logarithmic scale, the probability function becomes $\log_{10}N(x) = \log_{10}N_{total} + \log_{10}\frac{\alpha}{\beta} - \frac{\beta x}{\ln 10} + \log_{10}\left(1 - e^{-\beta(a_x - x)}\right)$. Accordingly, it is found from this equation that, when x is closer to the upper limit of the magnitude a_x, $\log_{10}\left(1 - e^{-\beta(a_x - x)}\right) \to -\infty$, x is not linearly related to $\log_{10}N(x)$, and the logarithmic cumulative frequency distribution behavior can be significantly altered. However, if the incremental statistical method is used, the incremental frequency distribution in the interval is $\mathrm{N}(x_{i-1}) = \mathrm{N}_{\text{total}}\int_{x_{i-1}}^{x_i} \alpha\, e^{-\beta x} dx = \mathrm{N}_{\text{total}}\frac{\alpha}{\beta}\left(e^{-\beta x_{i-1}} - e^{-\beta x_i}\right) = \mathrm{N}_{\text{total}}\frac{\alpha}{\beta}e^{-\beta x_{i-1}}\left(1 - e^{-\beta \Delta x}\right)$, and again the probability function on a logarithmic scale will be $\log_{10}N(x_{i-1}) = \log_{10}N_{\text{total}} + \log_{10}\frac{\alpha}{\beta} + \log_{10}\left(1 - e^{-\beta\Delta x}\right) - \frac{\beta x_{i-1}}{\ln 10}$, which shows that, for a certain bin width, x_{i-1} is linearly related to $\log_{10}N(x_{i-1})$. The above analysis means that, theoretically, the cumulative frequency distribution used for *b*-value estimation will inevitably result in deviation from a power law while the incremental frequency distribution will not.

In addition, owing to the natural smoothing effect of plotting cumulative frequency data, a regression analysis based on the cumulative frequency distribution will systematically increase the goodness of fit and affect the computation of the magnitude of completeness, *Mc* (Main, 2000; Wiemer & Wyss, 2000; Schorlemmer et al., 2004; Amorese et al., 2010). Moreover, the inaccurate magnitude determinations and breaks in power-law slope will possibly be smoothed out in cumulative frequency distribution, which cause the estimated *b*-value to incorrectly describe the size distribution scaling in practice.

3 Effect of attenuation on AE frequency–amplitude distribution

Since the attenuation of elastic wave in rock material is hard to be theoretically described, it is therefore difficult to analyze the modification of attenuation on size distribution. Here we perform an analysis from another more general perspective to investigate the effect of attenuation on *b*-value by considering that the randomly distributed sources are attenuated to a certain extent and collected by the sensors independent of the composition of the attenuation terms.

3.1 Continuous probability density function of the attenuated source amplitude

Suppose the AE source Z is randomly distributed in a region Ω. We define the origin (0, 0, 0) and $\left(x_1, y_1, z_1\right)$ in Ω as the coordinates of the sensor and source, respectively. Therefore, the continuous source amplitude probability density function in exponential distribution which truncated at an interval $[a_0, a_x]$ will be:

$$Z \sim f(z) = \begin{cases} \alpha e^{-\beta z} & a_0 \le x \le a_x \\ \text{other} & \text{otherwise} \end{cases}, \tag{2}$$

where α and β are constants, and "other" is any other distribution that is excluded from consideration. In addition, the attenuation amplitude can be summarized by the relation $A = A_0 g(x_1, y_1, z_1)$, where $0 < g(x_1, y_1, z_1) < 1$, A and A_0 are the amplitude of the AE source after and before attenuation, respectively, and the attenuation relation on a logarithmic scale is

$$\log_{10}(A) = \log_{10}(A_0) + \log_{10} g(x_1, y_1, z_1), \tag{3}$$

Letting $X = \log_{10} A$ and $Z = \log_{10} A_0$, we can simplify Equation (3) as $X = Z + \log_{10} g(x_1, y_1, z_1)$.

The density function of the coordinate of the source is $f(x_1, y_1, z_1)$, which is any random distribution. Suppose that the source amplitude Z and its location (x_1, y_1, z_1) are independent of each other, so that $f_Z(z) f(x_1, y_1, z_1)$ is the joint density function. The probability of the attenuated amplitude X can be given by

$$\begin{aligned} F_X(x) &= P(X \le x) = P\left(Z + \log_{10} g(x_1, y_1, z_1) \le x\right) \\ &= \iint_{Z + \log_{10} g(x_1, y_1, z_1) \le x} f_Z(z) f(x_1, y_1, z_1) dz dv \\ &= \iiint_\Omega dv \int_{a_0}^{x - \log_{10} g(x_1, y_1, z_1)} f_Z(z) f(x_1, y_1, z_1) dz + \iiint_\Omega dv \int_{-\infty}^{a_0} f_Z(z) f(x_1, y_1, z_1) dz \end{aligned}$$

$$= \begin{cases} \frac{\alpha}{\beta} e^{-\beta a_0} - \frac{\alpha}{\beta} e^{-\beta x} \iiint_\Omega f(x_1, y_1, z_1) \left(g(x_1, y_1, z_1)\right)^{\frac{\beta}{\ln 10}} dv + \iiint_\Omega dv \int_{-\infty}^{a_0} f_Z(z) f(x_1, y_1, z_1) dz & a_0 - m \le x \le a_x - M \\ \text{other} & \text{otherwise} \end{cases}, \tag{4}$$

where $M = \max_{(x_1, y_1, z_1) \in \Omega} \left(-\log_{10} g(x_1, y_1, z_1)\right)$ and $m = \min_{(x_1, y_1, z_1) \in \Omega} \left(-\log_{10} g(x_1, y_1, z_1)\right)$ are the maximum and minimum attenuation values from the source to sensor, respectively. The probability of the attenuated amplitude X in interval $[a_0 - m, a_x - M]$ can be given by

$$\begin{aligned} P(a_0 - m \le x \le a_x - M) &= F(a_x - M) - F(a_0 - m) \\ &= \frac{\alpha}{\beta}\left(e^{-\beta(a_0 - m)} - e^{-\beta(a_x - M)}\right) \iiint_\Omega f(x_1, y_1, z_1) \left(g(x_1, y_1, z_1)\right)^{\frac{\beta}{\ln 10}} dv, \end{aligned} \tag{5}$$

Let $X_0 = \left[Z + \log_{10} g(x_1, y_1, z_1)\right] | [a_0 - m, a_x - M]$. Then, the distribution function of probability for the attenuated amplitude X_0 can be given by

$$F_{X_0}(x) = P(X_0 \le x) = \frac{P(X_0 \le x, a_0 - m \le X \le a_x - M)}{P(a_0 - m \le X \le a_x - M)} = \begin{cases} 1 & x > a_x - M \\ \frac{e^{-\beta(a_0 - m)} - e^{-\beta x}}{e^{-\beta(a_0 - m)} - e^{-\beta(a_x - M)}} & a_0 - m \le x \le a_x - M, \\ 0 & x < a_0 - m \end{cases} \tag{6}$$

and the density function is

$$f_{X_0}(x) = \begin{cases} \frac{\beta e^{-\beta x}}{e^{-\beta(a_0 - m)} - e^{-\beta(a_x - M)}} & a_0 - m \le x \le a_x - M \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

3.2 Discrete probability density function of the attenuated source amplitude

Similarly, suppose there is a randomly distributed AE source Z_2 in a region Ω with exponentially distributed amplitude. If Y is the attenuation amplitude from source to sensor, then its probability can be given as

$$P(Y=a_i)=p_i,\ i=1, 2, \dots, n, \text{ here } \sum_{i=1}^{n} p_i=1, \tag{8}$$

where a_i is a discrete attenuation value and $a_1 < a_2 < \cdots < a_n$. Then, the amplitude X after attenuation will be

$$X=Z_2-Y,\ Z_2 \sim f_{Z_2}(z)=\begin{cases}\alpha e^{-\beta z} & a_0 \le x \le a_x \\ \text{other} & \text{otherwise}\end{cases}, \tag{9}$$

Suppose that the source amplitude Z_2 and Y are independent of each other; then, the probability of the attenuated amplitude X can be given by

$$F_X(x)=P(X\le x)=P(Z_2-Y\le x)=\left[p_1F_{Z_2}(x+a_1)+p_2F_{Z_2}(x+a_2)+\cdots+p_nF_{Z_2}(x+a_n)\right], \tag{10}$$

The density function is

$$f_X(x)=\left[p_1f_{Z_2}(x+a_1)+p_2f_{Z_2}(x+a_2)+\cdots+p_nf_{Z_2}(x+a_n)\right]$$

$$=\begin{cases}\left[p_1\alpha e^{-\beta a_1}+p_2\alpha e^{-\beta a_2}+\cdots+p_n\alpha e^{-\beta a_n}\right]e^{-\beta x} & a_0-\min\limits_{1\le i\le n}\{a_i\}\le x\le a_x-\max\limits_{1\le i\le n}\{a_i\} \\ \text{other} & \text{otherwise}\end{cases}, \tag{11}$$

Let $X_1=X\Big|a_0-\min\limits_{1\le i\le n}\{a_i\}\le X\le a_x-\max\limits_{1\le i\le n}\{a_i\}$. Then the distribution function of probability for the attenuated amplitude X_1 can be given by

$$F_{X_1}(x)=P(X_1\le x)=\frac{P\left(X_0\le x,\ a_0-\min\limits_{1\le i\le n}\{a_i\}\le X\le a_x-\max\limits_{1\le i\le n}\{a_i\}\right)}{P\left(a_0-\min\limits_{1\le i\le n}\{a_i\}\le X\le a_x-\max\limits_{1\le i\le n}\{a_i\}\right)}$$

$$=\begin{cases}0 & x<a_0-\min\limits_{1\le i\le n}\{a_i\} \\ \dfrac{\int_{a_0-\min_{1\le i\le n}\{a_i\}}^{x}\beta e^{-\beta x}dx}{\int_{a_0-\min_{1\le i\le n}\{a_i\}}^{a_x-\max_{1\le i\le n}\{a_i\}}\beta e^{-\beta x}dx} & a_0-\min\limits_{1\le i\le n}\{a_i\}\le x\le a_x-\max\limits_{1\le i\le n}\{a_i\}, \\ 1 & x>a_x-\max\limits_{1\le i\le n}\{a_i\}\end{cases} \tag{12}$$

and the density function is

$$f_{X_1}(x)=\begin{cases}0 & \text{otherwise} \\ \dfrac{\beta e^{-\beta x}}{\left(e^{-\beta\left(a_0-\min_{1\le i\le n}\{a_i\}\right)}-e^{-\beta\left(a_x-\max_{1\le i\le n}\{a_i\}\right)}\right)} & a_0-\min\limits_{1\le i\le n}\{a_i\}\le x\le a_x-\max\limits_{1\le i\le n}\{a_i\},\end{cases} \tag{13}$$

As shown in Equations (2) and (7) or Equations (9) and (13), when the exponential source distribution is doubly truncated at a finite interval, the distribution function of the attenuated amplitudes obeys the same exponential distribution at a certain interval, and the β value is theoretically verified to be unchanged. This means that, if the AE source obeys an exponential distribution, attenuation will not modify the source distribution within a certain interval. That is, if the attenuated amplitude distribution can be found in a power-law relationship in a certain interval, its *b*-value is the same as that of the AE source amplitude distribution.

4 Proper *b*-value estimation procedure for rock AE and its application

Based on the discussion above, we propose a procedure when using apparent amplitude for rock AE *b*-value estimation with the following steps:

Step 1: Minimum number of events for *b*-value estimation

From the statistical point of view, the minimum amount of data for estimation does not have a universal set value. It can be selected after the level of significance is specified in practical applications. A commonly recognized axiom is that the power of a statistical test increases with the increase in sample size (Siegel, 1956). For logarithmic linear fitting of a power-law distribution, the average error of the estimated scaling parameter becomes <1% when the sample size exceeds 50 (Clauset, 2007), and 50 events was also adopted as the minimum number of events for stable *b*-value estimation in various studies (Schorlemmer et al., 2004; Kurz et al., 2006; Amorese et al., 2010), while the stability estimation of *b*-value needs to include the space-time window of 100 earthquakes, as illustrated by Shi and Bolt (1982). Amorese et al. (2010) also stated that the computation of *b*-value includes at least ~200 events. In comparison to the cumulative frequency distribution, larger data dispersion will occur if the incremental frequency distribution is used for *b*-value estimation. Therefore, it is necessary to increase the data volume to ensure estimation robustness, so a minimum of 200 events should be utilized in the incremental frequency distribution.

Step 2: Perform incremental data counting and choose the bin width

As discussed in Section 2, the cumulative frequency distribution can inevitably deviate from a power law compared with the incremental frequency distribution, and the cumulative frequency distribution will also smooth the distribution breaks during the statistics process. Therefore, to display frequency distribution characteristics of AE data more realistically, the AE data should be counted incrementally.

Another parameter used for AE frequency–amplitude distribution statistics is bin width. In seismology, the selection of bin width is critical to grouping magnitudes. Improper choice of the bin width may introduce a significant bias in the magnitude of completeness and *b*-value estimation (Wiemer & Wyss, 2000; Marzocchi & Sandri, 2009; Schorlemmer et al., 2004). Because earthquake magnitudes are given with one digit after the decimal point, many researches have demonstrated that it is reasonable to set the bin width as the magnitude round-off interval of 0.1 (Bender, 1983; Main, 2000; Lasocki & Papadimitriou, 2006). As AE amplitudes were measured in decibels, and the amplitudes are usually divided by 20 to produce a *b*-value comparable to that reported in the seismic literature (Cox & Meredith, 1993; Liakopoulou, 1994, Weiss, 1997; Sagar et al., 2012), the divided value will have two digits after the decimal point with a 0.05 round-off interval. Therefore, the better bin width for amplitude grouping needs to be set to 0.05; this means that the number of amplitudes should be counted in steps of 1 dB. In fact, owing to the large amount of AE data collected in rock deformation tests, the coarser bin width will smooth the distribution breaks, while the finer bin width can demonstrate the frequency–amplitude distribution characteristics in more detail. Therefore, no matter what kind of equivalent amplitude or AE magnitude is used for *b*-value estimation in the rock AE test, the bin width should be set to the size of the round-off interval at most.

Step 3: Determine the logarithmic linear segment in the frequency–amplitude distribution

As discussed earlier, the deviation of the size distribution from the intrinsic power-law relationship will appear at both lower and upper amplitude ends. However, the analysis in Section 3 demonstrated that, if the attenuated amplitude distribution can be found in a power-law relationship in a certain interval, its *b*-value is the same as that of the AE source amplitude distribution. Therefore, a specific method is needed for determining the logarithmic linear segment in the incremental frequency–amplitude distribution. Here, we propose a method based on the Fisher optimal split and global search algorithm with the following procedures.

(1) Suppose that there are *n* points in the logarithmic frequency–amplitude distribution, A_{i-1} and A_i are the amplitudes of two successive points, and the segment slope for $A = A_i$ is defined as

$$S(A_i) = \frac{\log(N_i) - \log(N_{i-1})}{A_i - A_{i-1}} \tag{14}$$

Then, the $n - 1$ segment slopes are computed for each amplitude increment and the corresponding standard deviation is expressed by $std0$. Similarly, the standard deviation of all $S(A_i) < 0$ is expressed by $std1$.

(2) We define an interval $[S(A_i) - r_0 std1, S(A_i) + r_0 std1]$, where r_0 is a scaling parameter that is initially set to 0.1. For each segment slope where $S(A_i) < 0$, if the most number of slopes fall within that defined interval at a certain point *i*, then $S(A_i)$ is replaced by $S(A_{i0})$, which is defined as the benchmark slope, and the number of segment slopes falling within that defined interval at point *i* is marked as *s*. A step size *h* is defined as $h = \frac{\max\{S(A_{i0}) - \min(S(A_i)),\ \max(S(A_i)) - S(A_{i0})\}}{u\times s / std0}$, and *u* is another scaling parameter that is initially set to 10.

(3) Then, the interval defined in (2) is adjusted to $[S(A_{i0}) - kh, S(A_{i0}) + kh]$, k=1, 2, 3, 4, …, *m*. The number of slopes falling in each interval at various *k* can form an *m*-tuple sequence, marked as $\{temp1_m\}$, and the first-order difference of $\{temp1_m\}$ can form an $(m-1)$-tuple sequence, marked as $\left\{temp2_{m\text{-}1}\right\}$.

(4) The Fisher optimum split method is used to divide $\left\{temp2_{m\text{-}1}\right\}$ into two categories, each with a minimum sum of squared deviation. If $k = k_i$ is the optimal split point, the corresponding interval $[S(A_{i0}) - k_i h, S(A_{i0}) + k_i h]$ will be obtained. Because slopes less than zero are the ones to be considered, the interval can be adjusted to $[S(A_{i0}) - k_i h, 0)$.

(5) To further optimize the benchmark slope $S(A_{i0})$, a finer interval $[S(A_{i0}) - std2, 0)$, where $std2$ is the standard deviation of slopes falling in the interval $[S(A_{i0}) - k_i h, 0)$, is used to remove the abnormal slopes at both ends. Then, a new benchmark slope for points within interval $[S(A_{i0}) - std2, 0)$ can be obtained using least squares regression.

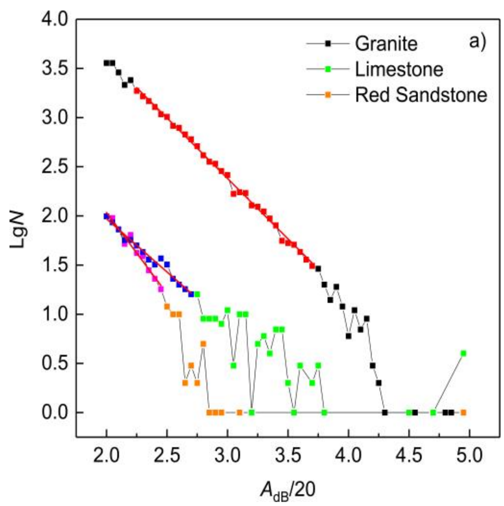
(6) The new benchmark slope obtained in (5) is assigned to $S(A_{i0})$, and the calculations in (3) and (4) are repeated to get the final optimal split point. Then the smallest amplitude and the largest amplitude in the interval $[S(A_{i0}) - k_i h, 0)$ are the left and right endpoints, respectively, of the logarithmic linear segment.

(7) A global search algorithm is used to run 1000 calculations to determine the optimal value of the initially set r_0 and u by finding the minimum error variance of the linear segment regression, and finally, the logarithmic linear segment of the amplitude distribution is screened out.

Here we call this newly proposed *b*-value estimation procedure as FGS method, to investigate the performance of FGS, a rupturing scheme using a nonexplosive expansion agent was designed (as shown in Figure 1) to conduct an AE test on granite, limestone, and red sandstone, which is intended to ensure that the sensor-collected AE signals are all generated by expansion rupturing in rock specimens. The AE activities were recorded by six sensors with a resonant frequency of 140 kHz at a 5 MHz sampling frequency. The estimated *b*-value using FGS method and corresponding goodness-of-fit R^2 and error variance of each rock specimen are given in Table 1 and shown in Figure 2(a). The temporal variations of the *b*-values of granite, limestone, and red sandstone are shown in Figure 2(b - d). These were estimated using a running window of 4916, 200, and 280 events and a running step of 2458, 100, and 100 events, respectively.

Table 1. *Estimated b-value and the Corresponding Goodness-of-Fit, R^2, Standard Error, and Error Variance of Each Rock Specimen Obtained by Using the FGS Method*

	Granite	Limestone	Red sandstone
b-value	1.2365±0.0247	1.0819±0.0467	1.6502±0.1151
R^2	0.9966	0.9763	0.9625
Error Variance	0.0011	0.0015	0.0027



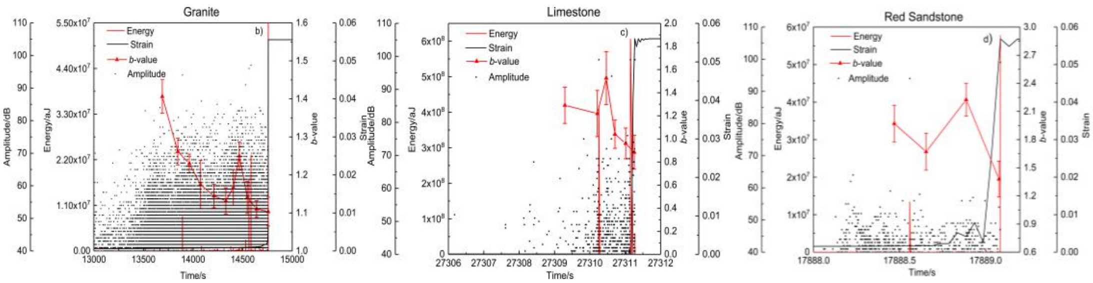


Figure 2. Estimated *b*-value obtained by using the FGS method of each rock specimen. Linear segment in logarithmic incremental frequency–amplitude distribution of three rock specimens (a). Temporal variation of *b*-value, strain data, amplitude, and energy of the AE signal with respect to time (b - d). The vertical bars are the standard error of the *b*-value.

Figure 2 (a) shows that the FGS method can well screen out the linear segment in the logarithmic incremental frequency–amplitude distribution, and the *b*-values of the three rock specimens as listed in Table 1 are distinct. As these three kinds of rock specimens are bonded by mineral grains of different sizes and different types of discontinuities, their rupture scale after deformation will also be different. Red sandstone is composed of fine-grained particles and seldom has large discontinuities, which accounts for a high proportion of small-scale ruptures. Granite has various large mineral grains and defects or voids. Limestone contains numerous of joints created during deposition, and these generate more large-scale ruptures. Therefore, the *b*-value of red sandstone is the largest, followed by granite, and the *b*-value of limestone is the smallest. Furthermore, Figure 2 (b - d) shows that the temporal variations of the *b*-value correspond well with the energy and amplitude of AE signals. The appearance of a high-energy and high-amplitude signal is usually triggered by large-scale ruptures, which in turn leads to a decrease in the *b*-value. The experimental results in this research show that the *b*-value estimated by using the FGS method can well demonstrate the size distribution characteristic in rock AE tests.

5 Conclusions

(1) The power-law relationship is the intrinsic characteristic of the size distribution from earthquake to laboratory rock ruptures, indicating that the number of small-scale ruptures is much greater than the number of large-scale ones. Nevertheless, deviation from a power-law size distribution is inevitable in practice. In fact, whether small-scale events cannot be fully recorded owing to limited equipment accuracy or small-scale ruptures are masked by larger ones that mainly occur in the avalanche destruction stage, the cause of the deviation of the size distribution from a power law can be ultimately attributed to inadequate data acquisition. Furthermore, from our discussion of the attenuation effect on the amplitude distribution, if the infinite distribution interval is selected, attenuation will not modify the amplitude distribution; however, once the size distribution is truncated at a certain interval, modification will occur at both ends. This indicates that data truncation actually is the fundamental cause of attenuation modification of the size distribution. Inadequate data acquisition and data truncation will both cause deviation at both ends of the size distribution; therefore, the data catalog is critical, and the *b*-value should be estimated at a proper interval. The proposed *b*-value estimation procedure (FGS) in this study is designed to search that proper interval, and the application also shows its reliability.

(2) Because the *b*-value is a statistical value, its estimation result is significantly influenced by the data volume. Generally, a sufficiently large number of events can improve the robustness of the *b*-value. However, when spatial or temporal variations in *b*-value are investigated, the pursuit of high resolution of the variations in *b*-values will reduce the number of events, which correspondingly leads to a decrease in reliability and robustness of the estimation. In other words, the more we want to increase the resolution of the variations in *b*-values, the less precisely we are able to simultaneously estimate the *b*-value because fewer events are used. Nevertheless, a reliable estimation is crucial in *b*-value analysis. Therefore, the robustness of *b*-value estimation should be guaranteed in priority even at the sacrifice of resolution of the spatial or temporal variation.

Acknowledgments

- This work was supported by the National Natural Science Foundation of China (Grant No. 41630642).
- The original AE data of dilation rupturing tests used in this work will be uploaded to a public repository called Digital Rocks Portal. The data archiving is underway now, we temporarily provide a copy of these AE data as Supporting Information named Datasets ds01 - ds03 for review purposes.

References

Amorese, D., Grasso, J. R., & Rydelek, P. A. (2010). On varying b-values with depth: results from computer-intensive tests for Southern California. *Geophysical Journal International*, *180*(1), 347–360. https://doi.org/10.1111/j.1365-246X.2009.04414.x

Bak, P., & Tang, C. (1989). Earthquakes as a Self-Organized Critical Phenomenon. *Journal of Geophysical Research*, *94*(B11), 15635-15637. https://doi.org/10.1029/JB094iB11p15635

Bender, B. (1983). Maximum likelihood estimation of b values for magnitude-grouped data. *Bulletin of the Seismological Society of America*, *73*, 831–851. https://doi.org/10.1016/0040-1951(84) 90 122 -7

Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical Research Letters*, *29*(9), 48-1–48-3. https://doi.org/10.1029/2001GL013775

Casertano, L. (1982). A statistical analysis of upper and lower limits of earthquake magnitude. *Pure & Applied Geophysics*, *120*(4):840-849. https://doi.org/10.1007/bf00876665

Christensen, K., Danon, L., Scanlon, T., & Bak, P. (2002). Unified scaling law for earthquakes. *Proceedings of the National Academy of Sciences, 99*(Supplement 1), 2509-2513. https://doi.org/10.1073/pnas.012581099

Clauset. (2007). Power-Law Distributions in Empirical Data. *Siam Review, 51*(4), 661-703. https://doi.org/10.1137/070710111

Cosentino, P., Ficarra, V., & Luzio, D. (1977). Truncated exponential frequency-magnitude relationship in earthquake statistics. *Bulletin of the Seismological Society of America, 67*(6), 1615-1623.

Cox, S. J. D., & Meredith, P. G. (1993). Microcrack formation and material softening in rock measured by monitoring acoustic emissions. *International Journal of Rock Mechanics & Mining Sciences & Geomechanics Abstracts*, *30* (l), 11-24. https://doi.org/ 10.1016/0148-9062(93)90172-A

Godano, C., Lippiello, E., & De Arcangelis, L. (2014). Variability of the b value in the Gutenberg–Richter distribution. *Geophysical Journal International*, *199*(3):1765-1771. https://doi.org/10.1093/gji/ggu359

Goebel, T. H. W., Becker, T. W., Sammis, C. G., Dresen, G., & Schorlemmer, D. (2014). Off-fault damage and acoustic emission distributions during the evolution of structurally complex faults over series of stick-slip events. *Geophysical Journal International*, *197*(3), 1705-1718. https://doi.org/10.1093/gji/ggu074

Goebel, T. H. W., Schorlemmer, D., Becker, T. W., Dresen, G., & Sammis, C. G. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophysical Research Letters*, *40*(10), 2049–2054. https://doi.org/10.1002/grl.50507

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, *34*, 185-188. https://doi.org/10.1038/156371a0

Kurz, J. H., Finck, F., Grosse, C. U., & Reinhardt, H. W. (2006). Stress drop and stress redistribution in concrete quantified over time by the b-value analysis. *Structural Health Monitoring*, *5*(1), 69-81. https://doi.org/10.1177/1475921706057983

Kwiatek, G., Goebel, T. H. W., & Dresen, G. (2014). Seismic moment tensor and\r, *b*\r, value variations over successive seismic cycles in laboratory stick-slip experiments. *Geophysical Research Letters*, *41*(16), 5838-5846. https://doi.org/10.1002/2014GL060159

Lasocki, S., & Papadimitriou, E. E. (2006). Magnitude distribution complexity revealed in seismicity from greece. *Journal of Geophysical Research Solid Earth*, *111*(B11), -. https://doi.org/10.1029/2005JB003794

Lavrov, A. V. (2005). Fracture localization in space as a cause of variations in the amplitude distribution of acoustic emission signals. *Acoustical Physics*, *51*(3), 321-327. https://doi.org/10.1134/1.1922546

Liakopoulou. F., Main, I. G., Crawford, B. R., & Smart, B. G. D. (1994). Microseismic properties of a homogeneous sandstone during fault nucleation and frictional sliding. *Geophysical. Journal. International*, *119*(1), 219-230. https://doi.org/10.1111/j.1365-246X.1994.tb00923.x

Lockner, D. A. (1993). The Role of Acoustic Emission in the Study of Rock Fracture, *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *30*(7), 883~899. https://doi.org/10.1016/0148-9062(93)90041-B

Lockner, D. A., Byerlee, J. D., Kuksenko, V., Ponomarev, A., & Sidorin, A. (1991). Quasi-static fault growth and shear fracture energy in granite. *Nature*, *350*, 39–42. https://doi.org/10.1038/350039a0

Main, I. (2000). Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: fact or artifact. *Bulletin of the Seismological Society of America*, *90*(1), 86-97. https://doi.org/10.1785/0119990086

Marzocchi, W., & Sandri, L. (2009). A review and new insights on the estimation of the b-value and its uncertainty. *Annals of geophysics = Annali di geofisica*, *46*(6), 1271-1282. https://doi.org/10.4401/ag-3472

Meredith, P. G., & Main, I. G. (1990). Temporal variations in seismicity during quasi-static and dynamic rock failure. *Tectonophysics*, *175*, 249-268. https://doi.org/10.1016/0040-1951(90)90141-T

Mignan, A. (2012). Functional shape of the earthquake frequency-magnitude distribution and completeness magnitude. *Journal of Geophysical Research Solid Earth*, *117*(B8). https://doi.org/10.1029/2012JB009347

Mogi, K. (1962). Study of the elastic shocks caused by the fracture of heterogeneous materials and its relations to earthquake phenomena. *Bulletin Earthquake Research Institute*, *40*, 125-173.

Nava, F. A. , Márquez-Ramírez, V. H., Zúñiga, F. R., Ávila-Barrientos, L., & Quinteros, C. B. (2016). Gutenberg-richter b-value maximum likelihood estimation and sample size. *Journal of Seismology*, *21*(1), 1-9. https://doi.org/10.1007/s10950-016-9589-1

Newman, & M. E. J. (2005). Power laws, pareto distributions and zipf\"s law. *Contemporary Physics*, *46*(5), 323-351. https://doi.org/10.1080/00107510500052444

Page, R. (1968). Aftershocks and microaftershcks of the Great Alaska Earthquake of 1964. *Bulletin of The Seismological Society of America*, *58*, 323-335.

Radziminovich, N. A., Miroshnichenko, A. I., & Zuev, F. L. (2019). Magnitude of completeness, b-value, and spatial correlation dimension of earthquakes in the South Baikal Basin, Baikal Rift System. *Tectonophysics*, *759*, 44-57. https://doi.org/10.1016/j.tecto.2019.04.002.

Raub, C., Garzón, P. M., Kwiatek, G., Bohnhoff, M., & Dresen, G. (2017). Variations of seismic b-value at different stages of the seismic cycle along the North Anatolian Fault Zone in northwestern Turkey. *Tectonophysics*, 712–713, 232–248. https://doi.org/10.1016/j.tecto.2017.05.028

Sagar, R. V., Prasad, B. K. R., & Kumar, S. S. (2012). An experimental study on cracking evolution in concrete and cement mortar by the b-value analysis of acoustic emission technique. *Cement and Concrete Research*, 42(8), 1094-1104. https://doi.org/10.1016/j.cemconres.2012.05.003

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the Seismological Society of America*, *58*(1), 399-415. https://doi.org/10.1109/IGARSS.2013.6723499

Scholz, C. H. (2015). On the stress dependence of the earthquake b-value. *Geophysical Research Letters*, *42*(5), 1399-1402. https://doi.org/10.1002/2014GL062863

Schorlemmer, D., Wiemer, S., & Wyss, M. (2004). Earthquake statistics at Parkfield: 1. Stationarity of b values. *Journal of Geophysical Research*, 109. https://doi.org/10.1029/2004JB003234.

Shi, Y. L., & Blot, B. A. (1982). The standard error of the **mag**nitude-frequency b value, *Bulletin of the Seismological Society of America*, *72*(5), 1677-1687. https://doi.org/10.5194/angeo-22-3221-2004

Siegel, S. (1956). Nonparametric statistics for the behavioral sciences, McGraw-Hill. https://doi.org/10.2307/2551606

Unander, T. E. (1993). The effect of attenuation on b-values in acoustic emission measurements-A Theoretical Investigation. *International Journal of Rock Mechanics and Mining Science & Geomechanics Abstracts.*, *30*(7), 947-950. https://doi.org/10.1016/0148-9062(93)90050-N

Vorobieva, I., Shebalin, P., & Clément Narteau. (2016). Break of slope in earthquake-size distribution and creep rate along the San Andreas fault system. *Geophysical Research Letters, 43*(13), 6869-6875. https://doi.org/10.1002/2016GL069636

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bulletin of the Seismological Society of America. 70*, 1337–1346. https://doi.org/10.1007/BF02245436

Weiss, J. (1997). The role of attenuation on acoustic emission amplitude distributions and b-values. *Bulletin of the Seismological Society of America*, *87*(5), 1362-1367. https://doi.org/10.1029/96JB03805

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: examples from alaska, the western united states, and japan. *Bulletin of the Seismological Society of America*, *90*(4), 859-869. https://doi.org/10.1785/0119990114