1

Seismicity and Earthquake Catalogues Described as Point Processes

Earthquakes have been observed in recent centuries using seismograph networks or simply visual assessment of their effects. All these observations, concerning specific geographical regions and time periods, are generally reported in earthquake catalogues. The simplest structure of an earthquake catalogue is composed by strings of parameters representing each event as a point in a multi-dimensional space. Typically, these parameters include the occurrence time, location and size of each event. The location may or may not include the hypocentral depth, besides the geographical coordinates of the epicenter, while the size is most commonly given through the earthquake magnitude or, alternatively, its seismic moment. All these parameters may also be characterized by their uncertainties.

We refer the reader to any standard seismology text for a detailed description of the earthquake parameters included in usual catalogues, as well as a description of the methodologies used for their determination. In this book, unless explicitly stated otherwise, we assume that every i-th seismic event in a catalogue is represented in a five-dimensional (5-D) space of parameters constituted by its occurrence time (t_i), geographical coordinates (x_i, y_i), depth (z_i) and magnitude (m_i). A catalogue is characterized by a specific volume of this space. Depth is often neglected, reducing the 5-D space to a 4-D space. In the latter case, every earthquake is represented by the vector (t_i, x_i, y_i, m_i).

1.1. The Gutenberg–Richter law

The empirical Gutenberg–Richter (G–R) law, in agreement with observations in many different areas of the world, states that the magnitude density distribution of a complete sample of earthquakes is a negative exponential function ([ISH 39, GUT 54]):

$$dN(M) = e^{\alpha-\beta M}dM \qquad [1.1]$$

The cumulative magnitude distribution (i.e. the number of earthquakes with magnitude < M) is:

$$N(M) = \int_M^\infty dN(m) = \int_M^\infty e^{\alpha-\beta m}dm = \frac{e^\alpha}{\beta}e^{-\beta M} = N(0)e^{-\beta M} \qquad [1.2]$$

where $N(0)$ is the total number of events with $M \geq 0$.

The cumulative number can also be written as:

$$N(M) = N(0)e^{-\beta M}e^{-\beta M_0}e^{\beta M_0} = N_0 e^{-\beta(M-M_0)} \qquad [1.3]$$

where $N_0 = N(M_0)$.

The G–R law is more commonly expressed in decimal logarithms as:

$$log_{10}(N(M)) = a - bM, \qquad [1.4]$$

where $a = \log_{10}(N(0))$ and $b = \beta \cdot \log_{10}e$.

The average magnitude of a complete set of N_0 events with magnitude $M > M_0$ is

$$\bar{M} = \frac{1}{N_0}\int_{M_0}^\infty M \cdot \frac{dN(M)}{dM}dM = \beta\int_{M_0}^\infty M \cdot e^{-\beta(M-M_0)}dM = \frac{1}{\beta} + M_0 \qquad [1.5]$$

(independent of N_0).

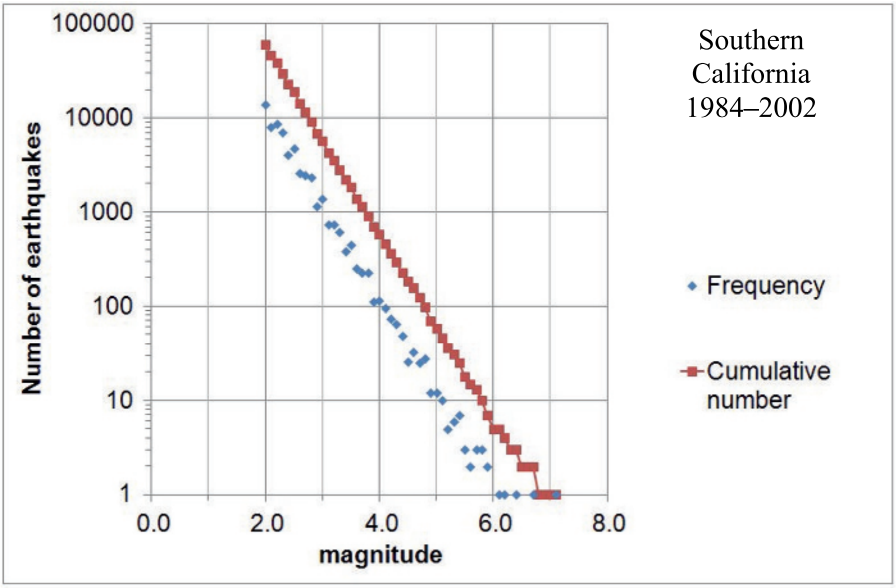


Figure 1.1. *An example of cumulative number of earthquakes versus the magnitude threshold*

If, as commonly observed, the value of b is approximately equal to 1, then $\beta \approx 2.303$ and $1/\beta \approx 0.4343$.

From above, a formula used for the determination of β (or more commonly b) from a complete set of magnitude observations can be obtained:

$$\beta = \frac{1}{\bar{M} - M_0} \qquad [1.6]$$

and

$$b = \frac{log_{10}e}{\bar{M} - M_0} \approx \frac{0.4343}{\bar{M} - M_0}. \qquad [1.7]$$

We now write a relationship to be used for determining the expected value of the largest magnitude of a set of N earthquakes with $M \geq M_0$.

The probability that the largest event in a sample has a magnitude $\leq M$ is equal to the probability that no event has a magnitude $\geq M$. Applying the G–R law and the Poisson statistical rule:

$$P(0|M) = e^{-N(M)} = e^{-N_0 e^{-\beta(M-M_0)}}. \quad [1.8]$$

Assuming that the largest magnitude of the sample is approximately equal to M, we can give an estimate of the density distribution of the largest magnitude, deriving this equation and changing its sign:

$$f(M) = \frac{dP(0|M)}{dM} = N_0 \beta e^{\left[\beta(M_0-M) - N_0 e^{\beta(M_0-M)}\right]}. \quad [1.9]$$

Figure 1.2 shows three examples of density distribution, for samples containing different numbers of events. The modal value of M, the value corresponding to the maximum of its density distribution M^*, is obtained by equaling its derivative to zero:

$$\frac{df(M)}{dM} = N_0 \beta^2 e^{\left[\beta(M_0-M) - N_0 e^{\beta(M_0-M)}\right]}\left[N_0 e^{\beta(M_0-M)} - 1\right] \quad [1.10]$$

from which,

$$M^* = M_0 + \frac{\ln N_0}{\beta} \quad [1.11]$$

The expected value of the largest magnitude is computed from the average of the magnitude distribution:

$$\bar{M}_{max} = \frac{1}{N_0} \int_{M_0}^{\infty} M N_0 \beta e^{\left[\beta(M_0-M) - N_0 e^{\beta(M_0-M)}\right]} dM \quad [1.12]$$

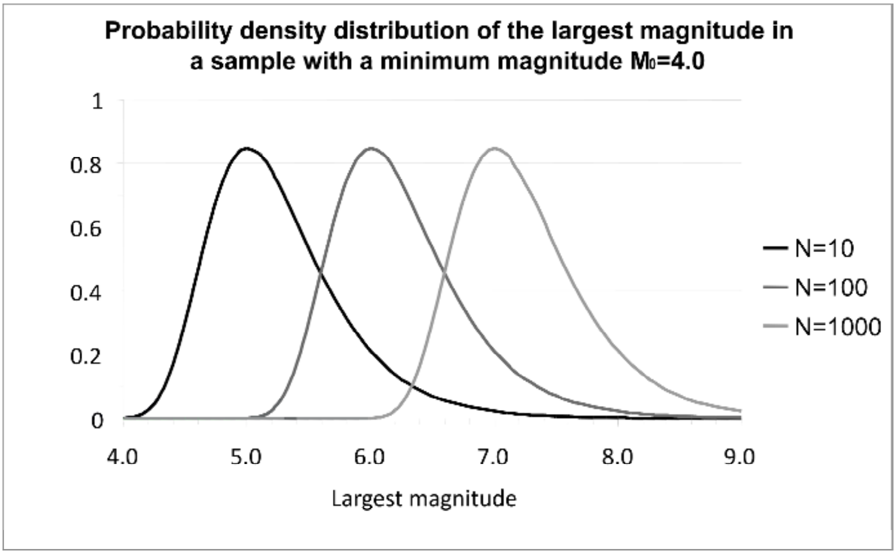


Figure 1.2. *Probability density distribution of the largest magnitude in a sample with a minimum magnitude M0=4.0. The different colors represent samples with different numbers of events. For a color version of this figure, see www.iste.co.uk/console/earthquake.zip*

1.2. The time-independent Poisson model

Under the hypothesis that earthquakes behave as a stationary stochastic process and with a uniform rate, the expected number N_0 of events with magnitude equal to or larger than M_0 in the observation time T_0 is a constant proportional to T_0. The Poisson formula gives the probability that in the same time period T_0 are N observed events:

$$P(N|\lambda) = \frac{e^{-\lambda}\lambda^N}{N!} \qquad [1.13]$$

where λ is the expected rate of a Poisson distribution.

From [1.13] it follows that the probability that no events occur in the same time period and with the same lower magnitude threshold is:

$$P(0|\lambda) = e^{-\lambda} \quad [1.14]$$

from which, the probability that, in the time interval T_0, at least one event occurs with magnitude equal to or larger than M_0:

$$P(N \geq 1|\lambda) = 1 - e^{-\lambda} \quad [1.15]$$

By the term "recurrence time", we mean the expected inter-event time between the occurrence of any event and the following. The recurrence time $T_r(M_0)$ is obviously a function of the threshold magnitude M_0 adopted for the analysis. An estimation of the recurrence time is given by the ratio between the observation time T and the number of observed events N_0:

$$T_r(M) = \frac{T}{N(M)} = \frac{T}{N_0 e^{-\beta(M-M_0)}} = \frac{T}{N_0} e^{\beta(M-M_0)} \quad [1.16]$$

The more accurate this estimation is, the larger the number N_0 is. It must be noted that the average value of the inter-event times generally provides underestimation of $T_r(M_0)$.

Once the recurrence time is known, it is possible to obtain from it the expected number of events, for a given threshold magnitude, for any observation time T:

$$\lambda = N(M|T) = \frac{T}{T_r(M)} \quad [1.17]$$

and, by means of the Poisson formula, the probability that such magnitude is exceeded during the time T:

$$P(M|T) = 1 - e^{-T/T_{r(M)}} \quad [1.18]$$

from which it is possible to compute the time T within which there is a given probability P of observing at least one earthquake of magnitude equal to or larger than M:

$$T(M) = -T_r(M)\ln(1-P) \tag{1.19}$$

The density distribution of the inter-event times, under the hypothesis of a constant occurrence rate, can be obtained by applying Poisson statistics with simple reasoning. Starting from the occurrence time t of any earthquake, the probability that the next earthquake occurs within the time $t+\Delta t$ is the complement to one of the probabilities that no earthquake occurs in the time interval Δt. From the Poisson formula, this probability is:

$$1 - P(0|\Delta t) = 1 - e^{-\Delta t/T_r(M)} \tag{1.20}$$

This probability, multiplied by the total number of earthquakes N_0, represents the cumulative distribution of the inter-event times that occur within a time interval Δt after the previous one. The density distribution of the inter-event times that occur at the time $t+\Delta t$ is given by the derivative of the cumulative distribution:

$$f(\Delta t) = \frac{dN(\Delta t)}{d\Delta t} = N_0 e^{-\Delta t/T_r(M)} \tag{1.21}$$

It can be noted that, unlike the intuitive idea that the density distribution of the inter-event times for events occurring randomly in time should be constant, Poisson statistics allows us to state that such a density distribution decreases as a negative exponential versus the duration of the time interval Δt.

1.3. Occurrence rate density as a space–time continuous variable

We introduce the definition of *occurrence rate density*:

$$\lambda(x,y,t,m) = \lim_{\Delta x,\, \Delta y,\, \Delta t,\, \Delta m \to 0} \frac{P_{\Delta x,\, \Delta y,\, \Delta t,\, \Delta m}(x,y,t,m)}{\Delta x\, \Delta y\, \Delta t\, \Delta m} \tag{1.22}$$

Here, Δx, Δy, Δt and Δm are increments of longitude, latitude, time and magnitude, $P_{\Delta x\,\Delta y\,\Delta t\,\Delta m}(x,y,t,m)$ is the probability of occurrence of an event in this volume. According to many authors such as [OGA 88, OGA 98, VER 95, VER 11, EVI 99, EVI 01, LOM 10a, GOS 06], the occurrence rate density is called *conditional intensity function or intensity function*. We prefer to follow the terminology adopted by [JAC 96] who calls it *conditional rate density* in order to avoid confusion with the concept of seismic intensity, as a measure of the strength of an earthquake.

Assuming that the point process does not have memory and the probability of occurrence of future events is constant irrespective of past activity, and adopting the *Gutenberg–Richter* law:

$$\mu(x,y,m)=e^{\alpha(x,y)-\beta(x,y)m}, \qquad [1.23]$$

where $\mu(x,y,m)$ is the space density of earthquakes of magnitude equal to or larger than m.

If $\beta(x,y)$ is assumed to be independent of the space coordinates ($\beta = b\ln(10) = 2.30259b$) (while α is allowed to change from point to point), we obtain the decomposition

$$\mu(x,y,m)=\mu_0(x,y)e^{\beta(m-m_0)}, \qquad [1.24]$$

where m_0 is the minimum magnitude of completeness and $\mu_0(x,y)=\mu(x,y,m_0)$ is the space density of earthquakes of magnitude equal to or larger than m_0.

By derivation of equation [1.24] with respect to m, in reverse order, the density magnitude distribution of earthquakes is obtained:

$$\mu(x,y,m)=\mu_0(x,y)\beta e^{-\beta(m-m_0)}. \qquad [1.25]$$

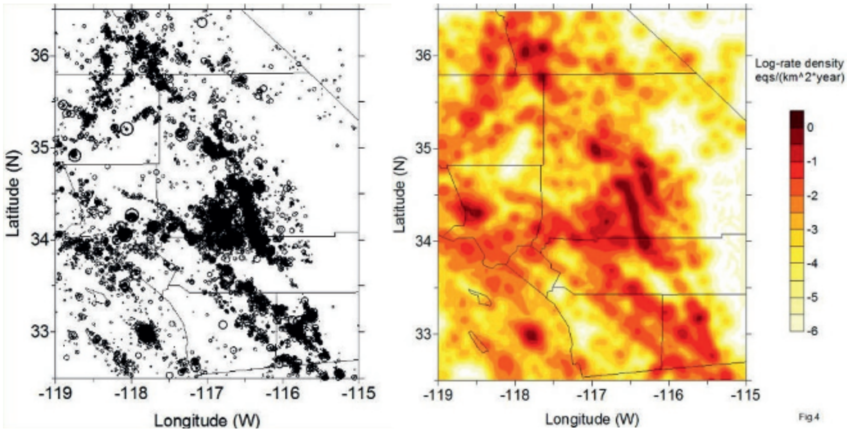
1.4. Time-independent spatial distribution

In the context of the present book, the estimate of the special density $\mu_0(x,y)$ is based entirely on instrumental catalogues of past seismicity, ignoring additional information such as intensity observations of historical earthquakes or surface faulting from prehistoric earthquakes. Each point source of the catalogue is replaced with a type of density function $f_i(x,y)$ (smoothing kernel function), which is greatest at the epicenter but tails off over the region surrounding it. The $\mu_0(x,y)$ distribution is obtained by summing all $f_i(x,y)$ for the earthquakes in the catalogue and scaling appropriately.

Among various methods introduced in the literature (e.g. [KAG 94], [KAG 00], [JAC 99], [HEL 07]), here, the method presented by [FRA 95] is adopted. First, the region to be analyzed is subdivided into square cells of appropriate size. The number of earthquakes N_k in each cell k represents the maximum likelihood estimate of the number of earthquakes above m_0 for that cell. The grid of N_k values is then smoothed spatially by multiplying by a Gaussian function with correlation distance c. For each cell k, the smoothed value $\tilde{N}_k$ is obtained from:

$$\widetilde{N_k} = \frac{\sum_l N_l e^{-\Delta_{kl}^2/c^2}}{\sum_l e^{-\Delta_{kl}^2/c^2}} \qquad [1.26]$$

where Δ_{kl} is the distance between the centers of the k-th and l-th cells. In equation [1.26], $\tilde{N}_k$ is normalized to preserve the total number of events. In order to implement equation [1.26], we have to choose c: the larger the value of c, the smoother will be the value of $\tilde{N}_k$, and the larger will be the contamination between different areas. In this work, the optimal choice of c is made by trial and error, adjusting c until the maximum likelihood is obtained from one half of the catalogue and a seismicity grid derived from the other half. The values of $\tilde{N}_k$ obtained so far still change abruptly from one cell to another. In order to obtain a continuous function, the single value of $\mu_0(x,y)$ is computed by interpolation between the four cells whose centers surround the point (x,y). The result is given as number of events per cell size, per total time spanned by the catalogue. In order to obtain the result in units of number of events per unit time and per unit area, it must be divided by the total time of the input catalogue and the area of a cell. Figure 1.3 gives an example of a map of earthquake epicenters and the corresponding map of the spatial density distribution obtained by equation [1.26].



a) b)

Figure 1.3. *a) Epicentral distribution of the earthquakes with magnitude M ≥ 2.0 reported by the Southern California Earthquake Data Center in the time period 1984–2002. b) Smoothed distribution of the Southern California (1984–2002) seismicity obtained by the smoothing algorithm of equation [1.26] with a correlation distance c = 5.0 km. For a color version of this figure, see www.iste.co.uk/console/earthquake.zip*

1.5. Clustered seismicity

As commonly observed, the occurrence rate of seismic events in a particular region increases soon after a strong earthquake, and then it goes down again, according to the so-called *modified Omori law* [OGA 83] expressed, for events exceeding the threshold magnitude m_0 by:

$$\lambda(t, m \geq m_0) = K(t+c)^{-p} \qquad [1.27]$$

with K, c and p constants that are characteristic of any particular case. The expected total number of events in the period of time [0, T] is obtained by integrating [1.27] over the time as:

$$N(T,\ m \geq m_0) = \begin{cases} K\left[\log(T+c) - logc\right] & if\ p=1 \\ K\left[(\mathrm{T}+\mathrm{c})^{1-p} - c^{1-p}\right]/(1-p) & if\ p \neq 1, \end{cases} \qquad [1.28]$$

which allows the computation of K from the observation of seismic events of magnitude equal to or larger than m_0 over a given time window.

The earthquakes that trigger further seismicity can be a single mainshock (in the usual subjective meaning), as in the case of a typical aftershock sequence, or more than one mainshock, as in the case of a multiple aftershock sequence, or even all the previous events, as in the epidemic model [OGA 88]. The resultant expected rate density of seismic events, taking into account the contribution of the previous inducing earthquakes, will be given by:

$$\lambda(x,y,t,m) = f_r \lambda_0(x,y,m) + \sum_{i=1}^{N} H(t-t_i) \cdot \lambda_i(x,y,t,m), \qquad [1.29]$$

where f_r is a factor called failure rate, $H(t)$ is the step function and $\lambda_i(x,y,t,m)$ is the single contribution, the time dependence of which is expressed by [1.27]. The first and second terms to the right of the sign of equality in [1.29] represent the independent and the triggered seismicity, respectively. The rate density corresponding to any earthquake is, in general, constituted by the superposition of these two components.

Following the [REA 89] hypothesis, the contribution to the occurrence rate density of any previous earthquake (x_i,y_i,t_i,m_i) is decomposable (for $t>t_i$) into three terms, respectively, representing the time, magnitude and space distribution, as:

$$\lambda_i(x,y,t,m) = K(t-t_i+c)^{-p}\, \beta e^{-\beta(m-m_i)} f(x-x_i, y-y_i), \qquad [1.30]$$

where $f(x,y)$ is the spatial distribution of the triggered seismicity with respect to the coordinates of the inducing event.

The decomposition given in [1.30] implies several important hypotheses. First, we assume that β, K, c and p are independent of the geographical coordinates (x, y) and (x_i, y_i). This is, of course, an initial approximation, that can be refined if there is enough experience for formulating a more complex situation.

Modeling the spatial distribution $f(x,y)$ independently of the magnitude of the inducing earthquake is another strong approximation. We shall introduce, in the following, different forms of the spatial term (e.g. [OGA 98, CON 03]).

Finally, it must be noted that putting the magnitude m_i of the inducing events in the same term of the magnitude m of the triggered events and with the same coefficient β is a strongly subjective hypothesis that implies the self-similarity of the process over all the ranges of magnitude. With this assumption, the probability of an earthquake of magnitude 3.0 after an earthquake of magnitude 4.0 is the same as that of an earthquake of magnitude 5.0 after an earthquake of magnitude 6.0, and so on. While it is not proven that this assumption is true, we use it because of its simplicity, in the context of the parsimony of hypotheses that is a rule all through this work.

We model the spatial distribution of the triggered seismicity using a function $f\,(x\text{-}x_i,\ y\text{-}y_i)$ that has circular symmetry around the point of coordinates $(x_i,\ y_i)$ and is normalized to one, so as to preserve the total expected number of triggered events of magnitude equal to or larger than m_0, related to K through [1.28] [CON 01a, CON 01b]. Defining the distance from the point $(x_i,\ y_i)$ as: $r=\sqrt{\left(x-x_i\right)^2+\left(y-y_i\right)^2}$, the first condition is written as:

$$f\left(r\right)=ce^{(-r^2/2\sigma^2)}, \qquad [1.31]$$

where σ is a suitable constant that defines the "width" of the distribution and c is a normalization constant. The normalization is obtained by imposing that the integral over the whole plane is equal to 1:

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f\left(x-x_i,y-y_i\right)dxdy=2\pi\int_{-\infty}^{+\infty} f\left(r\right)dr=1 \qquad [1.32]$$

This, by substitution of [1.31], gives

$$c=\frac{1}{2\mu\sigma^2}. \qquad [1.33]$$

Then, for the space distribution, we will assume:

$$f\left(r\right)=\frac{1}{2\pi\sigma^2}e^{(-r^2/2\sigma^2)}. \qquad [1.34]$$

1.6. Epidemic models

We shall now build the full occurrence rate density for the class of models referred to by [OGA 88] and [OGA 98] as *epidemic model*. Although many different algorithms have been introduced by different authors in their respective versions, all these models are characterized by the common feature that *all* the events are likely to have influence on the occurrence rate of the following ones and *all* the events are likely to be influenced by the previous ones. The simplicity of this class of clustering models allows us to avoid the use of arbitrary and subjective definitions such as "foreshocks", "mainshocks", "aftershocks", independent and triggered events, etc., that would introduce further free parameters in the formulation: all the earthquakes are considered identical from this view point.

Let us first consider the model adopted by [CON 01a] and [CON 01b]. By substitution of [1.34] into [1.30], and of the latter into [1.29], together with [1.25], we have the complete expression of the rate density:

$$\lambda(x,y,t,m) = f_r \mu_0(x,y)\beta e^{-\beta(m-m_0)} + \frac{\beta K}{2\pi\sigma^2}\sum_{i=1}^{N}\frac{H(t-t_i)}{(t-t_i+c)^p} e^{-\beta(m-m_i)} e^{(-r^2/2\sigma^2)} \qquad [1.35]$$

with $r = \sqrt{(x-x_i)^2 + (y-y_i)^2}$.

Looking at [1.35], it is easily recognizable that the magnitude distribution of both "spontaneous" and "triggered" events follows the Gutenberg–Richter law [1.23] with the same coefficient β. Though it has been frequently claimed that this hypothesis is contradicted by the so-called "Bath's law" (i.e. the difference in magnitudes of the "main shock" and the largest aftershock in a sequence is roughly 1.2), it was shown by [VER 69] that the contradiction appears to be largely, if not entirely, due to an accidental bias in the selection of data.

Console [CON 03] introduced a different expression of the spatial density distribution. In polar coordinates (r,θ), this distribution is:

$$f(r,\theta) = \frac{(q-1)}{\pi}\frac{d^{2(q-1)}}{(r^2+d^2)^q} \qquad [1.36]$$

where d and q are two free parameters of the process. On the right side of equation [1.36], θ does not appear explicitly because of the isotropy of the model, but the normalization has been done by integration of this parameter from 0 to 2π. In their study, the spatial parameters σ and d are independent of the magnitude of the inducing earthquake.

The spatial distribution of the triggered seismicity can also be modeled by a function that depends on the magnitude of the triggering earthquake. For instance, [CON 10b] adopted a spatial distribution that in polar coordinates (r,θ) can be written as:

$$f(r,\theta)=\left(\frac{d_j^2}{r^2+d_j^2}\right)^q \quad [1.37]$$

where r is the distance from the epicenter of the triggering event, q is a free parameter modeling the decay with distance and d_j is the characteristic triggering distance. The difference between [1.36] and [1.37] is that in the latter d_j is related to the magnitude m_j of the j-th triggering earthquake:

$$d_j = d_0 10^{\alpha(m_j - m_0)/2} \quad [1.38]$$

where d_0 is a constant characteristic triggering distance of an earthquake of magnitude m_0 and α is a free parameter, that is, the spatial distribution is scaled with magnitude.

We may note that the spatial kernel distribution [1.37] is not normalized. This implies that the expected number of events triggered by an earthquake depends on its magnitude through the value of the parameter α.