# A. Integrated analysis of seismicity

## Contents

A. Integrated analysis of seismicity	1
A.1 Event rate, energy release, and seismic b-value	1
A.2 Nonextensive Tsallis statistics and q-value	2
A.3 Time-domain earthquake distribution	3
A.4 Fractal structures and characteristic scales	3
A.5 Band-limited fractal	4
A.6 Single-link cluster analysis and spatial correlation length	5
A.7 ETAS modeling	5
A.8 ETAS model with time-dependent forcing rate	6
A.9 Extended ETAS model	8
A.10 Nonparametric analysis	8
A.11 Seasonality and tidal response of seismic activity	10
A.12 Stress diffusion derived from seismic activity	10

Statistical parameters of earthquake occurrence, including event rate (n), cumulative event number (N), seismic *b*-value in the magnitude–frequency relation, spatial correlation length (*SCL*), and the fractal dimension  $(D_2)$  of the hypocenter distribution, as an integrated set, were found in a previous study to be useful as an indicator of the critical point behavior of rock fracture in stressed rock samples (Lei & Satoh, 2007). In addition, the epidemic-type aftershock sequence (ETAS) model, self-exciting model, 4D correlation model were also included. Follows describe statistical models often used for characterize seismic activity.

## A.1 Event rate, energy release, and seismic b-value

Energy released by an earthquake relates the magnitude by

$$E_i \propto 10^{CM_i} \tag{A1-1}$$

where *C* is a constant. The most important cases are C = 0.75 and 1.5, which yield the Benioff strain and the classic energy, respectively. The energy release rate can be estimated by summing (1) within a given unit time interval ( $\Delta t$ ):

$$\dot{E}(t) = \sum E_i / \Delta t \tag{A.1-2}$$

The cumulative energy release can be simply defined by:

$$\sum E = \int \dot{E} dt \tag{A.1-3}$$

The cumulative event number (n) of magnitude M or greater follows the Gutenberg–Richter magnitude–frequency relationship (Gutenberg and Richter, 1954) given by

$$\log_{10} n = a - bM \tag{A.1-4}$$

In Eq. (4), both a and b are constant, and b is referred to as the b-value.

In general, *b*-value is a function of rock property (e.g density and size distribution of pre-existing cracks/fractures) and stress level (Lei et al., 2006; 2007). However, for a given limited area b-value may reflect stress evolution: decreasing b-value is potentially a good indicator of increasing stress (Scholz 1968; Schorlemmer et al., 2005) or diffusion of pore pressure (Lei and Satoh, 2007; Hainzl and Fischer, 2002).

#### A.2 Nonextensive Tsallis statistics and q-value

Abe S. (2002) suggested that among the Rényi, Tsallis, and normalized Tsallis entropies, only the Tsallis entropy is stable and can give rise to experimentally observable quantities. More recently, a very interesting model for earthquake dynamics related to the Tsallis nonentensive framework has been proposed by the Sotolongo-Costa and Posad (2004) as SCP model. Such a model consists basically of two rough profiles interacting via fragments filling the gap between them where the fragments are produced by local breakage of the local plates.

$$\log(N_{>m}) = \log N + \left(\frac{2-q}{1-q}\right) \log\left[1 - a(1-q)(2-q)^{(1-q)/(q-2)} 10^{2m}\right]$$
(A.2-1)

This is not a trivial result, and incorporates the characteristics of nonextensivity into the distribution of earthquakes by magnitude. The parameter a is the constant of proportionality between e and r.

Later, the SCP model has been revisited by Slva et al. (2006) by considering a different definition for mean values in the context of Tsallis nonextensive statistics and introducing a scale between the earthquake energy and the size of fragment  $e \sim r^3$  rather than  $e \sim r$ .

$$\log(N_{>m}) = \log N + \left(\frac{2-q}{1-q}\right) \log \left[1 - \left(\frac{1-q}{2-q}\right) \left(\frac{10^{2m}}{a^{2/3}}\right)\right]$$
(A.2-2)

Vilar et al. have tested the viability of the nonextensive models for earthquakes from the best studied major fault zone in the world, i.e., the San Andreas fault. By using 6188 earthquake events (in the interval 2~8) taken from the Network Earthquake International Center catalogs and Bulletin of the International Seismological Centre they have shown, in agreement with other similar analyzes, that for values of the nonextensive parameter of the order of q = 1.6-1.7.

## A.3 Time-domain earthquake distribution

Will be included in later version.

#### A.4 Fractal structures and characteristic scales

The spatial distribution of many geological systems such as earthquake hypocenters and active faults show complexity at all scales. Self-similarity and fractal distribution are concepts for describing such complexities. For spatially distributed points such as earthquake hypocenters, correlation integral C (r) defined by following equation is used to examine the self-similarity of the distribution :

$$C(r) = \frac{2N_r(R < r)}{N(N-1)}$$
(A.4-1)

Where, Nr(R < r) is the number of pars of points having a distance R less than r. If the distribution is fractal C(r) would be a power of ,  $C(r) \propto r^{D_2}$ , and  $D_2$  is the so-called fractal dimension. In other words, the N(r)-r relation in a double logarithm coordinate will be a line statistically, and the slope of the line gives a capacity fractal dimension.

The box-counting method (BCM) is also often used to evaluate the similar structure of a system. The BCM is particularly useful for such as trace of active faults. Figure 3-5 is a conceptual diagram for the box counting method applied to points and lines. By changing the size (r) of a square box and counting the minimum number N(r) of boxes necessary to cover all the data, the relation of N(r) to r is obtained.

More generally,



Figure A4-1. A conceptual plot showing a box-counting method applied to line extended object (a)

and point distributed data (b). By changing the size (r) of the square box and counting the minimum number N(r) (Lei & Kusunose, 1999)

## A.5 Band-limited fractal

Ouillon *et al.* [1995, 1996] applied a multifractal analysis to joints and faults from 1 cm to 100 km scales, and pointed out that different geometries and scaling laws hold for different scale ranges, separated by boundaries that correlate well with the thickness of lithological units that constitute the continental crust. Lei *et al.* (1993), proposed that the distribution of the earthquake epicenters in and around China, range from several km to several thousand km, is not a strict fractal. There are two characteristic scales, which divided the whole scale range into three bands. Strictly speaking, the distribution of earthquake epicenters is not a fractal. But, since within each band a power law distribution is established quite well, it was therefore referenced as a "band-limited" fractal. Band-limited fractal, which is a special case of non homogeneous fractal (multifractal), is used to emphasize the existence of characteristic scales that are band boundaries. In each band, the measure may be non homogeneous or homogeneous. For instance, in a large scale, earthquakes are concentrated on the boundary of plates and in seismic zones, having a fractal dimension approximately equal to 1.0. On the other hand, in a scale from tens to hundreds km, we would see many clusters (or in other words, swarms) distributed along active fault zones, with a fractal dimension less than 1.0. When an individual cluster is magnified, the epicenters are observed to be distributed in a fault zone with a higher fractal dimension.



Figure A5-1. Generalized-correlation-integral functions of the epicenter distribution of earthquakes in China versus distance in double logarithm coordinates. The data shows band-limited multifractal structure with three bands. The fractal dimension for each band and for q = 2 to 15 are measured by the least-square fitting and also specified in the right side.(Lei et al., 1993)

### A.6 Single-link cluster analysis and spatial correlation length

The concept of critical earthquake has been proposed to model earthquake process as a critical phenomenon, where the largest or catastrophic earthquake is considered as analogous to a critical phenomenon happening at a second-order phase transition in analogy to percolation phenomenon (Yamashita & Knopoff, 1989; Sornette A. & Sornette D., 1990). Under such a consideration, if the system (spatially distributed earthquakes) approaches a critical point, the spatial correlation length (hereafter abbreviated as SCL) is expected to grow according to a power law (Bruce & Wallace, 1989):

$$\xi(t) \propto (t_f - t)^{-k} \tag{A.6-1}$$

where *k* is a positive constant.

The SCL of a set of N consecutive events can be estimated using single-link cluster analysis (Frohlich and Davis, 1990). Initially, each individual hypocenter is linked with its nearest neighbor hypocenter to form a set of clusters. Then, every cluster is linked with its nearest cluster. This process is repeated until N events are connected with N-1 links. Following previous work (Zöller et al., 2001; Tyupkin and Giovambattista, 2005), the SCL is here defined as the median of the distance distribution of the N-1 links. Where, the distance between any two clusters is calculated based on their geometric centers.

#### A.7 ETAS modeling

It is known that smaller changes in shear stress in a fault system, transferred from (seismic or aseismic) fault slip elsewhere or any other kind of loading, can lead to changes in seismicity (Reasenberg & Simpson, 1992; Stein, 1999; Toda et al., 2002). In some sensitive cases stress change in the order of 0.1 bar can cause change in seismic activity (King et al., 1994, Cochran et al., 2004). Because the triggering process for a given earthquake sequence is complex and because it is difficult to precisely describe the transfer of stress in a fault system of fractal complexity, statistical approaches such as the epidemic-type aftershock sequence (ETAS) model, which incorporates the Omori law by assuming that each earthquake has a magnitude-dependent ability to trigger its own Omori-law-type aftershocks, have received significant attention (Ogata et al., 2003; Ogata, 2005; Sornette & Sornette, 1999; Helmstetter & Sornette, 2003). The ETAS model is an appropriate tool for testing the significance of changes in seismic patterns, such as the relative quiescence that occurs prior to a large earthquake or large aftershock (Ogata, 1992, 2001). The ETAS model is also helpful in detecting minor stress changes (Helmstetter & Sornette, 2003) and extracting a fluid signal from seismicity data (Hainzl & Ogata, 2005; Lei at al., 2008).

We provide a summary of this model below; a more detailed account can be found in the cited references. The ETAS model incorporates the Omori law by assuming that each earthquake has a magnitude-dependent ability to trigger its own Omori-law-type aftershocks (Ogata, 1992; Helmstetter & Sornette, 2003). The total occurrence rate is described as the sum of the rate of all preceding earthquakes and a constant rate  $p_0$  that represents the stable Poisson process; in other words, random background activity:

$$\lambda(t) = p_0 + \sum_{\{i:t_i < t\}} K_0 e^{\alpha(M_i - M_{\min})} (t - t_i + c)^{-p}$$
(A.7-1)

where  $M_{min}$  is the low-end magnitude cut-off of the catalog and  $\alpha$  is a constant that specifies the degree of magnitude dependence. We also tested the Poisson model ( $\lambda(t) = p_0$ ) and a self-exciting model that employs an exponential decay function instead of the Omori law (*Lei et al.*, 2000).

$$\lambda(t) = p_0 + \sum_{\{i:t_i < t\}} a_0 e^{\beta(t-t_i)}$$
(A.7-2)

Model parameters in Eq. A7-1 and A7-2 were estimated by minimizing the Akaike information criterion (AIC). As a result, ETAS models provide the smallest value of AIC of all the models, indicating that ETAS models provide the best fit to the seismic data. The parameter ( $p_0$ ,  $K_0$ , c,  $\alpha$ , p) is assumed to take the same values for all events, thus representing regional characteristics of seismicity. Typical values of ETAS parameters obtained for various earthquake data sets are p = [0.9, 1.4], c = [0.03, 0.3 day],  $\alpha = [0.2, 3.0]$  (Guo & Ogata, 1997).

The parameter is useful in characterizing earthquake sequences qualitatively in relation to the classification into seismic types(Ogata, 1992). For example, earthquake swarms have  $\alpha$  values less than 1, and clear and simple mainshock-aftershock activity in Japan has  $\alpha > 2$ .

As for the ETAS model, the likelihood function is

$$L(\mu, K_0, \alpha, c, p) = \left\{ \prod_{j=1}^N \lambda(t_j) \right\} \exp\left\{ -\int_0^T \lambda(t) dt \right\}$$
(A-7-3)

The log-likelihood function is

$$LL = \sum_{j=1}^{N} \ln \left( \mu + K_0 \sum_{i:t_i < t_j} e^{\alpha (M_i - M_c)} (t_j - t_i + c)^{-p} \right) - \mu T - K_0 \int_0^T \sum_{i:t_i < t_j} e^{\alpha (M_i - M_c)} (t_j - t_i + c)^{-p} dt \quad (A-7-4)$$

The parameter is useful in characterizing earthquake sequences qualitatively in relation to the classification into seismic types (Ogata, 1992). For example, earthquake swarms have  $\alpha$  values less than 1, and clear and simple mainshock-aftershock activity in Japan has  $\alpha > 2$ .

#### A.8 ETAS model with time-dependent forcing rate

$$\lambda(t) = \lambda_0(t) + \sum_{\{i:t_i < t\}} K_0 e^{\alpha(M_i - M_{\min})} (t - t_i + c)^{-p}$$
(A.8-1)

In some cases such as injection-induced seismicity the forcing rate is strongly dependent on injection factors and thus could not be treated as constant (Lei et al., 2013). Otherwise incorrect ETAS parameters (superior low value of  $\alpha$ ) could be estimated (Marson et al. 2013). For such cases it is required to assume that the forcing rate  $\mu$  can vary with time, while the parameters A,  $\alpha$ , p, and c are constant. Marson proposed a practically useful method for the estimation of the forcing rate which consists of:

1. Initially assuming a constant forcing rate  $\lambda_0(t) = \lambda_0$ , using the Gamma distribution of waiting times between consecutive earthquakes.

$$f(\tau) = C\tau^{\gamma - 1} e^{-\tau/\beta}$$
(A.8-2)

Here,  $\tau$  is the normalized interevent time that is obtained by multiplying the interevent time  $\Delta t$  with the earthquake rate  $\lambda$ . Based on the assumption that the seismicity consists of a Poissonian background activity and triggered aftershocks following the Omori law, Molchan (2005) showed that the value  $1/\beta$  is the fraction of mainshocks among all seismic events, in other words, the forcing rate.

2. Computing the MLE ETAS parameters  $\theta(K, \alpha, c, p)$  knowing  $\lambda_0(t)$ .

3. Updating the estimate of the forcing rate based on these parameters. This step first requires the computation of the probability  $\omega_i$  that the i<sup>th</sup> earthquake is a background earthquake, for all i. This probability is defined as [Zhuang, 2002]

$$\varpi_i = \frac{\lambda_0(t_i)}{\lambda_0(t_i) + \nu(t_i)} \tag{A.8-3}$$

The forcing rate is then obtained by smoothing these probabilities over time:

$$\lambda_0(t) = \frac{1}{t_b - t_a} \sum_{i=a}^b \overline{\varpi}_i$$
(A.8-4)

where a and b are the indices of the sequence number earthquakes which define a window of  $n_e$  earthquakes centered on t.

The appropriate smoothing parameter ne can be determined by minimizing following Akaike Information Criterion (AIC):

$$AIC(n_e) = AIC(ETAS) + 2N/(2n_e + 1)$$
(A.8-5)

where N is the number of earthquakes in the data set.

It is convenient to plot the theoretical/observed cumulative number of events as a function of the transformed time, i.e., the integral of the occurrence rate over the time interval (S, T). The time series on the transformed time axis is thus a Poisson process, which provides a visual confirmation of goodness of fit and allows examination of change points.

### A.9 Extended ETAS model

Consider a space-time occurrence rate lq(t, x, y) at the time and location (t, x, y) as an extension of the ETAS model. Having compared the three typical space-time extensions of the ETAS model, Ogata [1998] eventually recommends the following model expressed by

$$\lambda(t, x, y) = p_0 + \sum_{\{i:t_i < t\}} \frac{K_0 e^{\alpha(M_i - M_c)}}{(t - t_i + c)^p} \left[ \frac{(x - x_j, y - y_j)S_j(x - x_j, y - y_j)^{\wedge}}{e^{\alpha(M_j - M_c)}} + d \right]^{-q}$$
(A.9-1)

where  $S_j$  is a dimensional 2x2 positive definite symmetric matrix, and (. , .)<sup>^</sup> indicates transpose of the vector. The square brackets have the dimension of degree<sup>2</sup> where "degree" corresponds to the global distance in latitude. The quadratic form within the brackets indicates that the aftershocks are spatially distributed with ellipsoidal contours.

## A.10 Nonparametric analysis

As a point process in time, space, and magnitude, the observed (dressed) seismicity rate density  $\lambda(x, t)$ , defined as the number of earthquakes per unit time and unit volume at position x and time t, is modeled as

$$\lambda(x,t) = \lambda_0 + \sum_{\{i:t_i < t\}} \lambda_i(x,t)$$
(A.10-1)

where  $\lambda_0$  is the uniform background rate density, and  $\lambda_i(x, t)$  is the (bare) contribution of earthquake i that occurred at {xi, ti}, representing the aftershocks directly caused by this earthquake. By assuming only that (i) the triggering process is linear [i.e., the bare contributions  $\lambda_i(x, t)$  sum up], and (ii) a mean-field response to the occurrence of an earthquake can be estimated that depends only on its magnitude,  $\lambda_i(x, t) = \lambda(|x - x_i|, t - t_i, m_i)$ , Marsan and Lengline (2008) proposed an algorithm for estimating the bare kernel:

1. Knowing an a priori bare kernel  $\lambda(r, t, m)$  and  $\lambda_0$ , the triggering weights are calculates by

$$w_{i,j} = \alpha_j \lambda (|r_j - r_i|, t_j - t_i, m_i), j > i$$
  

$$w_{i,j} = 0, j \le i$$
  

$$\alpha_j = 1 / \left( \sum_{i=1}^N \lambda (|r_j - r_i|, t_j - t_i, m_i) + \lambda_0 \right)$$
(A.10-2)

2. The updated bare rates are then computed as

$$\lambda(r,\delta t,m) = \sum_{i,j\in A} w_{i,j} / (N_m \times \delta t \times V(r,\delta r))$$
(A.10-3)

where A is the set of pairs such that  $|x_i - x_j| = |\Delta x| \pm \delta r$ ,  $m_i = m \pm \delta m$ , and  $t_j - t_i = t \pm \delta t$  ( $\delta r$ ,  $\delta t$ , and  $\delta m$  are discretization parameters), Nm is the number of earthquakes such that  $m_i = m \pm \delta m$ , and  $S(|\Delta x|, \delta r)$  is the surface covered by the disk with radii  $|\Delta x| \pm \delta r$ . The a posteriori background rate is

$$\lambda_0 = \frac{1}{T \times V} \sum_{j=1}^N w_{0,j}$$
(A.10-4)

where T is the duration of the time series (containing N earthquakes) and V is the space volume (3D) or surface area (2D) analyzed. This corresponds to stacking all the aftershocks following mainshocks i of similar magnitudes, but counting an aftershock j according to its weight  $w_{i,j}$ .

These two steps are repeated until convergence is reached. The convergence is equivalent to maximizing the logarithm likelihood of the model:

The log-likelihood function is

$$LL = \sum_{i=1}^{N} \ln \lambda(t_i) - \int_0^T \lambda(t) dt$$

$$= \sum_{i=1}^{N} \ln \left( \lambda_0 + \sum_{j=1}^{i-1} \lambda(|r_j - r_i|), t_j - t_i, M_i) \right) - \lambda_0 VT - \sum_{i=1}^{N} \int_0^{T-t_i} \lambda_t (t - t_i, M_i) dt$$
(A.10-5)

It is helpful to decouple the spatial and the temporal dependence of the kernel into spatial density  $\lambda_s$  (number of triggered earthquakes per unit volume) and temporal rate  $\lambda_t$  (number of triggered earthquakes per unit time):

$$\lambda_i(r, t, m_i) = \lambda_s(r, m_i) \times \lambda_t(t, m_i)$$
(A.10-6)

By constraining the spatial density to be normalized:

$$\int_{0}^{\infty} A\lambda_{s}(r,m_{i})dr = 1$$
(A.10-7)

where A= $2\pi r$  (2D) or  $4\pi r^2$  (3D), the temporal rate  $\lambda_t$  is obtained from:

$$\lambda_t(t,m_i) = \int_0^\infty A\lambda_i(r,t,m_i)dr$$
(A.10-8)

#### The kernel

The a posteriori background temporal rate is given by:

$$\lambda_0 = \frac{\sum_{j=1}^N W_{0,j}}{T * V}$$
(A.10-9)

The background temporal rate is given by:

$$\lambda_{0,t} = \frac{\sum_{j=1}^{N} W_{0,j}}{T}$$
(A.10-10)

Marsan D. and O. Lengline (2008), Extending earthquakes' reach through cascading, Science, 319, 1076-1079, DOI: 10.1126/science.1148783.

## A.11 Seasonality and tidal response of seismic activity

Utilities associated with seasonality and tidal response of seismic activity is now going for thorough examination.

#### A.12 Stress diffusion derived from seismic activity

Following Marsan & Bean (1999) a 4D correlation approach was used for analyzing stress diffusion derived from seismic activity. For a space time distributed points set  $(x_i, t_i)$  of N events and cover a period of duration T, we firstly calculate the average rate of events N(r, t) occurring in a space-time window of size  $(L_W, T_W)$  at a distance r and time delay t after any event:

$$N(r,t) = \frac{c}{T_{W}L_{W}^{3}N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \Theta(t_{j} - t_{i} \in [t;t+T_{W}]) \cdot \Theta(|x_{j} - x_{i}| \in [r;r+L_{W}])$$
(A.12-1)

where  $\Theta(p)$  is 1/0 if p is true/false, and c is a artificial factor. For convenience we call any event i as

"main" event and all events *j* of  $t_j < t_i$  as it's "after events".

N(r, t) will change with t if the data includes temporally correlated events. N(r, t) will change distance r and reflects the spatial structure of the series. For a spatially random series, N(r, t) is proportional with  $r^{D-1}$  where D is the spatial dimension. We consider the case that the time-space point series includes both random and correlated events. Since correlation distance in time is limited, N(r, t) is expected be convergence with t. To remove the temporally uncorrelated events from N(r, t), we define a new function C(r, t) as

$$C(r,t) = N(r,t) - N(r,t = t_{\max})$$
(A.12-2)

Where,  $t_{max}$  must be selected carefully to be great than the correlation distance in time. In addition,  $t_{max}$  must be << T to avoid edge effects. Since the series has not only limited duration but also limited space, a  $r_{max}$  is need to be introduced to avoid the effect of spatial edge. Now C(r, t) gives the probability of after events triggered by the main event after a delay t, at a distance r away from this main event. The mean distance of C(r,t) at time given by

$$R(t) = \frac{\int r \cdot C(r,t)dt}{\int C(r,t)dt}$$
(A.12-3)

can be used to derive stress diffusivity.