Statistical analysis of low-frequency earthquake catalogs

Ariane Ducellier

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2021

Committee:

Donald B. Percival, Chair

Yen-Chi Chen

Thomas S. Richardson

Program Authorized to Offer Degree:

Statistics

University of Washington

Abstract

Statistical analysis of low-frequency earthquake catalogs

Ariane Ducellier

Chair of the Supervisory Committee:
Professor Donald B. Percival
Department of Statistics

Low-frequency earthquakes (LFEs) are small magnitude (less than 2) earthquakes, with reduced amplitudes at frequencies greater than 10 Hz relative to ordinary small earthquakes. They are usually grouped into families of events, with all the earthquakes of a given family originating from the same small patch on the plate interface and recurring more or less episodically in a bursty manner. In this thesis, I analyze catalogs of events from several LFE families located in Cascadia, Mexico, and the San Andreas Fault. First, for each given family of LFEs, I translate the catalog into a continuous time series defined by the number of events per unit of time. Long-range dependence is a phenomenon that may arise in the statistical analysis of time series data. It relates to the slow rate of decay of the statistical dependence between two points with increasing time interval between the points. I look for evidence of long-range dependence in the time series

by analyzing them with several graphical methods. Second, I consider all the LFEs associated with a single family as a point process (a marked point process if magnitude is available). I then model this point process with an Epidemic-Type Aftershock Sequence (ETAS) model and fit the parameters of the model for each LFE family. Finally, I generate synthetic ETAS models using the fitted parameters, in order to verify whether this type of models can produce apparent long-range dependence as observed in the first part of this thesis. The statistical characterization of LFE occurrence could provide important constraints on future mechanical models of LFE generation.

Acknowledgements

I would like to give many thanks to Professor Donald Percival who has kindly accepted to be my advisor for this project. I first learned about long-range dependence in the class on wavelet analysis that he taught at University of Washington in Spring 2018, and I have been very interested in this topic ever since.

I also would like to thank Professor Yen-Chi Chen and Professor Thomas Richardson who have kindly accepted to be members of my Master's thesis committee.

I would like to thank my PhD advisor Professor Kenneth Creager for encouraging me in pursuing this research project in addition to my PhD studies.

The workshops and the computing resources provided by the Research Computing Club at University of Washington have been very helpful in getting me started on this project.

I would also like to thank all the seismology researchers who have provided me the datasets analyzed in this thesis: Shelley Chestler, William Frank, Alexandre Plourde, David Shelly, and Justin Sweet.

Finally, I would like to thank my husband for his constant enthusiasm and support for this project.

Contents

List of Figures

To my husband...

Chapter 1

Introduction

Low-frequency earthquakes (LFEs) are small magnitude earthquakes, with typical magnitude less than 2, and reduced amplitudes at frequencies greater than 10 Hz relative to ordinary small earthquakes. Their occurrence is often associated with tectonic tremor and slow slip events along the plate boundary in subduction zones and occasionally transform fault zones. They are usually grouped into families of events, with all the earthquakes of a given family originating from the same small patch on the plate interface, and recurring more or less episodically in a bursty manner.

LFE datasets from the Olympic Peninsula (Washington), Northern California, Mexico, and the San Andreas Fault are available. Each dataset contains the geographic locations of the LFE families, and, for each family, the timing (and sometimes the magnitude) of each of the LFE events. Each LFE family is thus associated with a catalog of earthquake events. In this thesis, I will analyze each LFE catalog independently from the others, and I am not interested in the possible correlation between two catalogs associated with two different LFE families.

There are two ways to make a statistical analysis of these catalogs:

- First, we can consider all the LFEs associated with a single family as a point process (a marked point process if magnitude is available).
- Second, for each family of LFEs, it is possible to translate the catalog into a discrete time series defined by the number of events per unit of time.

In the second chapter of this thesis, I consider each LFE catalog as a time series, and I study evidence of long-range dependence in these time series.

In the third chapter, I consider each LFE catalog as a point process, and I try to fit an Epidemic Type Aftershock Sequence (ETAS) model on them.

In the fourth chapter, I generate synthetic catalogs using the ETAS models fitted in the third chapter, and look for evidence of long-range dependence in the resulting synthetic time series.

Chapter 2

Long-range dependence in low-frequency earthquake catalogs

Long-range dependence is a phenomenon that may arise in the statistical analysis of time series data. It relates to the slow rate of decay of the statistical dependence between two points with increasing time interval between the points. Evidence of long-range dependence has been found in regional earthquake catalogs (Telesca et al., 2000; Li, Chen, and Mi, 2002; Telesca, Lapenna, and Macchiato, 2004; Enescu, Ito, and Struzik, 2006; Jiménez et al., 2006; Xu and Burton, 2006; Bunde and Lennartz, 2012; Wang, 2013; Gkarlaouni et al., 2017; Barani et al., 2018; Matcharashvili et al., 2018; Mukhopadhyay and Sengupta, 2018), global earthquake catalogs (Ogata and Abe, 1991; Sarlis et al., 2018), and in catalogs of mining-induced seismicity (Węglarczyk and Lasocki, 2009). We can investigate whether there is long-range dependence in a catalog by computing the fractional differencing parameter d, which represents how fast the variance in the number of LFEs in a time window of a given length increases with the length of the time window considered. A fractional differencing parameter such that $0 < d < 0.5$ is characteristic of long-range dependence in the time series, whereas an homogeneous Poisson process has a fractional differencing parameter equal to 0.

Some evidence of long-range dependence has been found in low-frequency earthquake catalogs. For instance, Frank et al. (2016) carried out a statistical analysis of a catalog of low-frequency earthquakes recorded between January 2005 and April 2007 in the Guerrero, Mexico, subduction zone. They computed the autocorrelation sequence and the spectral density function of the low-frequency earthquake rate for two 4-month-long windows, one corresponding to an inter-slow slip period, and one corresponding to the 2006 slow slip event. They observed that low-frequency earthquakes behave as an homogeneous Poisson's process during the inter-slow slip period, and as a process with long-range dependence during the co-slow slip period.

In the first section of this chapter, I explain in more detail what long-range dependence is. In the second section, I describe several methods to compute the fractional differencing parameter d of a time series with long-range dependence. In the third section, I apply these methods to synthetic time series. In the fourth section, I describe the low-frequency earthquake catalogs to which I will apply the method. Finally, in the last section, I present the results of the analysis.

2.1 Long-range dependence

A time series is any sequence of observations associated with an ordered independent variable t. For the analyses carried out in this thesis, I assume that the time series is defined essentially over a range of integers (usually $t = 1, ..., N$, where N denotes the number of values in the time series).

A real-valued random variable is a function, or mapping, from the sample space of possible outcomes of a random experiment to the real line.

A discrete parameter real-valued stochastic process $\{X_t : t = ..., -1, 0, 1, ...\}$ is a sequence of random variables indexed over the integers. A process such as $\{X_t\}$ can serve as a stochastic model for a sequence of observations of some physical phenomenon. We assume that these observations are recorded at a sampling interval of Δt.

The process $\{X_t\}$ is said to be (second order) stationary if:

1. $E\{X_t\} = \mu_X$ for all integers t (i.e. μ_X does not depend on t).
2. $cov\{X_t, X_{t+\tau}\} = s_{X,\tau}$ for all integers t and τ (i.e. $s_{X,\tau}$ depends only on τ and does not depend on t).

The sequence $\{s_{X,\tau} : \tau = ..., -1, 0, 1, ...\}$ is called the autocovariance sequence (ACVS) of the stochastic process. The autocorrelation sequence (ACS) is $\rho_{X,\tau} = s_{X,\tau}/s_{X,0}$.

The spectral density $S_X(f)$ is related to the ACVS $s_{X,\tau}$ by the relationship:

$$s_{X,\tau} = \int_{-f_N}^{f_N} S_X(f)\, e^{i2\pi f \tau \Delta t} df \text{ where } f_N = \frac{1}{2\Delta t} \text{ is the Nyquist frequency.} \tag{2.1}$$

Long-range dependence can then be defined by imposing conditions on the autocorrelation $\rho_{X,\tau}$ or on the spectral density $S(f)$ (Beran, 1994).

Condition on the autocorrelation X_t is called a process with long-range dependence if there exists a real number $\delta \in (0, 1)$ and a constant c_ρ such that:

$$\lim_{\tau \to \infty} \frac{\rho_{X,\tau}}{c_\rho \tau^{-\delta}} = 1 \tag{2.2}$$

Condition on the spectral density X_t is called a process with long-range dependence if there exists a real number $\delta \in (0, 1)$ and a constant c_S such that:

$$\lim_{f \to 0} \frac{S_X(f)}{c_S |f|^{-\delta}} = 1 \tag{2.3}$$

2.2 Methods to measure long-range dependence

In this thesis, I use several different graphic methods to compute either the fractional differencing parameter d, or the Hurst parameter. For finite variance processes, the fractional differencing parameter d and the Hurst parameter H are related by:

$$H = d + \frac{1}{2} \tag{2.4}$$

For infinite variance processes, the fractional differencing parameter d and the Hurst parameter H are related by:

$$H = d + \frac{1}{\lambda} \tag{2.5}$$

where λ is a parameter of the distribution of the innovations of a FARIMA process (e.g. the shape of the Pareto distribution, or the stability parameter of the λ-stable distribution). The graphical methods to compute d or H methods are described in Taqqu and Teverovsky (1998). Here is a brief summary of the methods.

Let us define a time series X_i, of length N. We define the corresponding aggregated series:

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i \text{ for } k = 1, 2, ..., \left[\frac{N}{m}\right] \tag{2.6}$$

The first absolute moment of the time series is then:

$$AM^{(m)} = \frac{m}{N} \sum_{k=1}^{\frac{N}{m}} \left| X^{(m)}(k) - \overline{X} \right| \tag{2.7}$$

and $AM^{(m)}$ behaves as m^{H-1}. If we make a log-log plot of the absolute moment $AM^{(m)}$ versus the length m of the aggregation window, the result should be a straight line of slope $H - 1$.

The sample variance of the time series is:

$$\widehat{Var} X^{(m)} = \frac{m}{N} \sum_{k=1}^{\frac{N}{m}} \left(X^{(m)}(k) - \overline{X} \right)^2 \tag{2.8}$$

and $\widehat{Var} X^{(m)}$ behaves as m^{2d-1}. If we make a log-log plot of the sample variance $\widehat{Var} X^{(m)}$ versus the length m of the aggregation window, the result should be a straight line of slope $2d - 1$.

To define the variance of residuals, we divide the time series into blocks of size m. We then compute the partial sums of the time series $Y(t) = \sum_{i=1}^{t} X_i$, and fit the partial sums within each block by a least-squares line $a + bt$. The sample variance of the residuals $\frac{1}{m} \sum_{t=1}^{m} (Y(t) - a - bt)^2$ is then computed. Its median over the blocks is proportional to m^{2H} and its average over the blocks is proportional to m^{2d+1}. In the case of the low-frequency earthquake time series, it is better to use the average rather than the median, because the median may often be equal to 0 as there are long periods of time without any earthquake events. However, this may not work very well in the case of infinite variance series. If we make a log-log plot of the average of variance of residuals versus the length m of the aggregation window, the result should be a straight line of slope $2d + 1$.

For the R/S method, the original time series is divided into K blocks. Then we define the partial sum of the time series for the block j by:

$$Y_j(t) = \sum_{i=k_j+1}^{k_j+t} X_i \text{ with } k_j = j\frac{N}{K} \text{ and } j = 0, \cdots, K-1 \tag{2.9}$$

We then define for each lag n:

$$R(j,n) = \max_{1\le t\le n}\left[Y_j(t) - \frac{t-1}{n}Y_j(n)\right] - \min_{1\le t\le n}\left[Y_j(t) - \frac{t-1}{n}Y_j(n)\right] \tag{2.10}$$

and $S(j,n)$ is the square root of the sample variance of $X(t)$, $t = k_j+1, \cdots, k_j+n$. The R/S statistic is defined by: $R/S = \frac{R(j,n)}{S(j,n)}$ and behaves as $n^{d+\frac{1}{2}}$.

The periodogram is defined as:

$$I(f) = \frac{1}{2\pi N}\left|\sum_{j=1}^{N} X_j e^{ijf}\right|^2 \text{ where } f \text{ is the frequency} \tag{2.11}$$

and $I(f)$ behaves as $|f|^{-2d}$ close to the origin for small f. If we make a log-log plot of the periodogram $I(f)$ versus the frequency f, the result should be a straight line of slope $-2d$.

2.3 Tests on synthetic data

To verify how well these methods compute the fractional differencing parameter d or the Hurst parameter H, I generated synthetic time series with long-range dependence using the same parameters as in Taqqu and Teverovsky (1998) and using the R package *fracdiff*. All the time series are Fractional Autoregressive Integrated Moving Average (FARIMA) with parameters (p, d, q) where the ARMA part is described by:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \tag{2.12}$$

where the e_t are the innovations. As in Taqqu and Teverovsky (1998), there are ten types of time series:

1. Gaussian FARIMA (1, d, 0) with $\phi_1 = 0.5$ and the innovations follow a normal distribution $\mathcal{N}(0,1)$.
2. Gaussian FARIMA (0, d, 1) with $\theta_1 = 0.5$ and the innovations follow a normal distribution $\mathcal{N}(0,1)$.
3. Gaussian FARIMA (1, d, 1) with $\phi_1 = -0.3$, $\theta_1 = -0.7$ and the innovations follow a normal distribution $\mathcal{N}(0,1)$.
4. Gaussian FARIMA (1, d, 1) with $\phi_1 = 0.3$, $\theta_1 = 0.7$ and the innovations follow a normal distribution $\mathcal{N}(0,1)$.
5. FARIMA (0, d, 0) where the innovations follow an exponential distribution with rate $\lambda = 1$.
6. FARIMA (0, d, 0) where the innovations follow a lognormal distribution with parameters $\mu = 0$ and $\sigma = 1$.

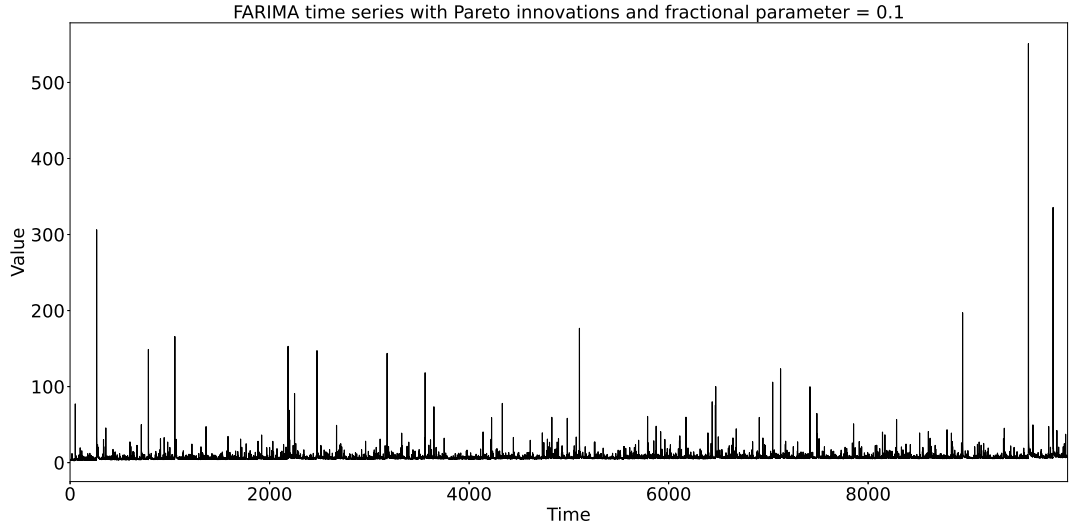

FIGURE 2.1: Example of a FARIMA time series with Pareto innovations and fractional differencing parameter $d = 0.1$.

7. FARIMA(0, d, 0) where the innovations follow a symmetric stable distribution with scale equal to 1 and stability parameter $\lambda = 1.5, 1, 2$.

8. FARIMA(1, d, 1) with $\phi_1 = 0.3$, $\theta_1 = 0.7$ and the innovations follow a symmetric stable distribution with scale equal to 1 and stability parameter $\lambda = 1.5, 1, 2$.

9. FARIMA(0, d, 0) where the innovations follow a Pareto distribution with shape $\lambda = 1.5$.

10. FARIMA(0, d, 0) where the innovations follow a skewed stable distribution with scale equal to 1, skewness equal to 1 and stability parameter $\lambda = 1.5$.

For each type of FARIMA model, I generate a single time series. The time series thus used in this thesis are just an example of realization of a FARIMA process.

As an example, I show here the 9th time series (Pareto innovations) with $d = 0.1$ as this time series looks a lot like the low-frequency earthquake time series with many values close to 0 and a burst of events from time to time. Figure 2.1 shows the time series, and Figures 2.2 to 2.6 show the graph corresponding to each method used to compute the fractional differencing parameter: absolute value method, variance method, variance of residuals methods, R/S statistic, and periodogram.

2.4 Data: The low-frequency earthquake catalogs

The Chestler catalog The low-frequency earthquake catalog from Chestler and Creager (2017a) and Chestler and Creager (2017b) contains event times for 69 families of low-frequency earthquakes located in the Olympic Peninsula, Washington. The catalogs start in June 2009 and end in September 2011. I converted each of the 69 catalogs into time series by counting the number of earthquakes per one-minute-long time window.

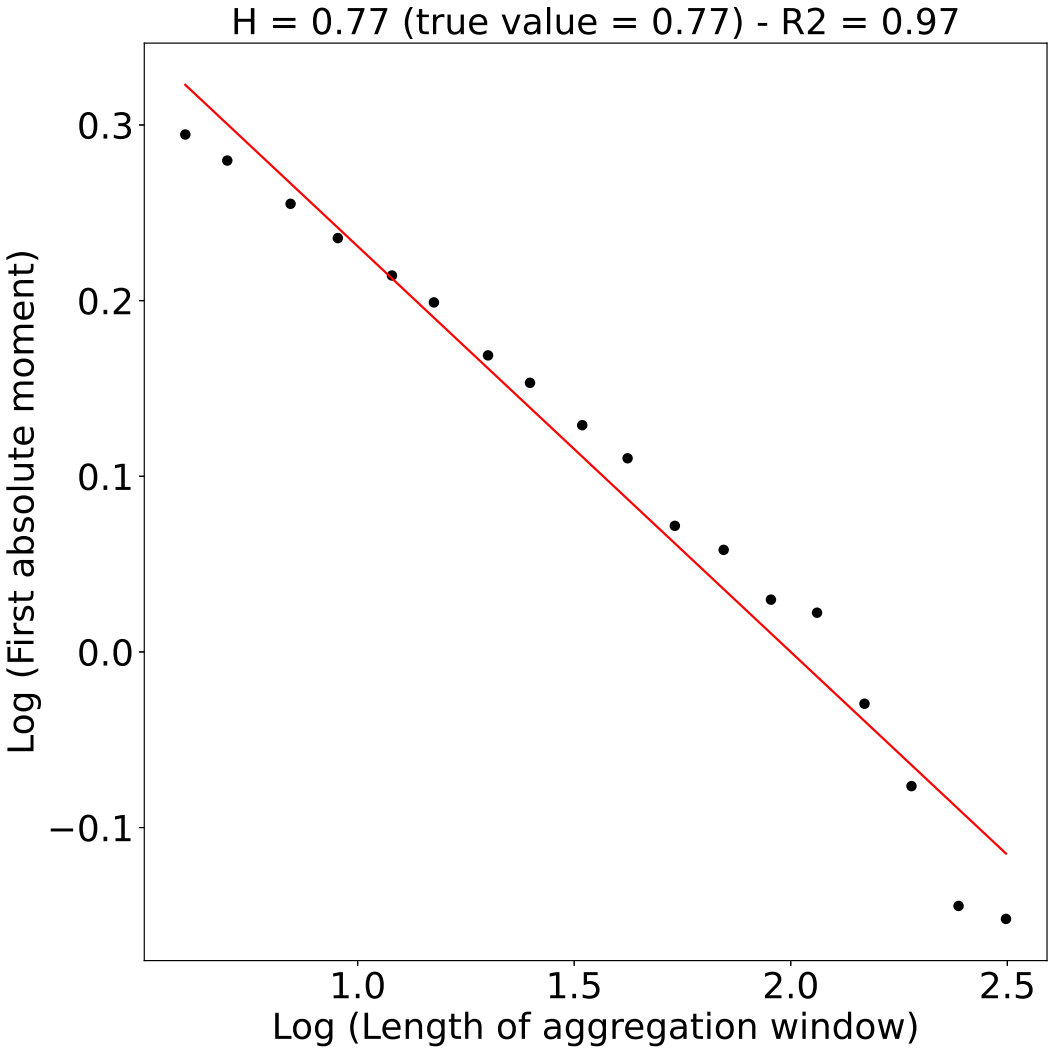

FIGURE 2.2: Absolute value method applied to the FARIMA time series from Figure 2.1. The horizontal axis is the logarithm of the length m of the aggregation window and the vertical axis is the logarithm of the first absolute moment $AM^{(m)}$.

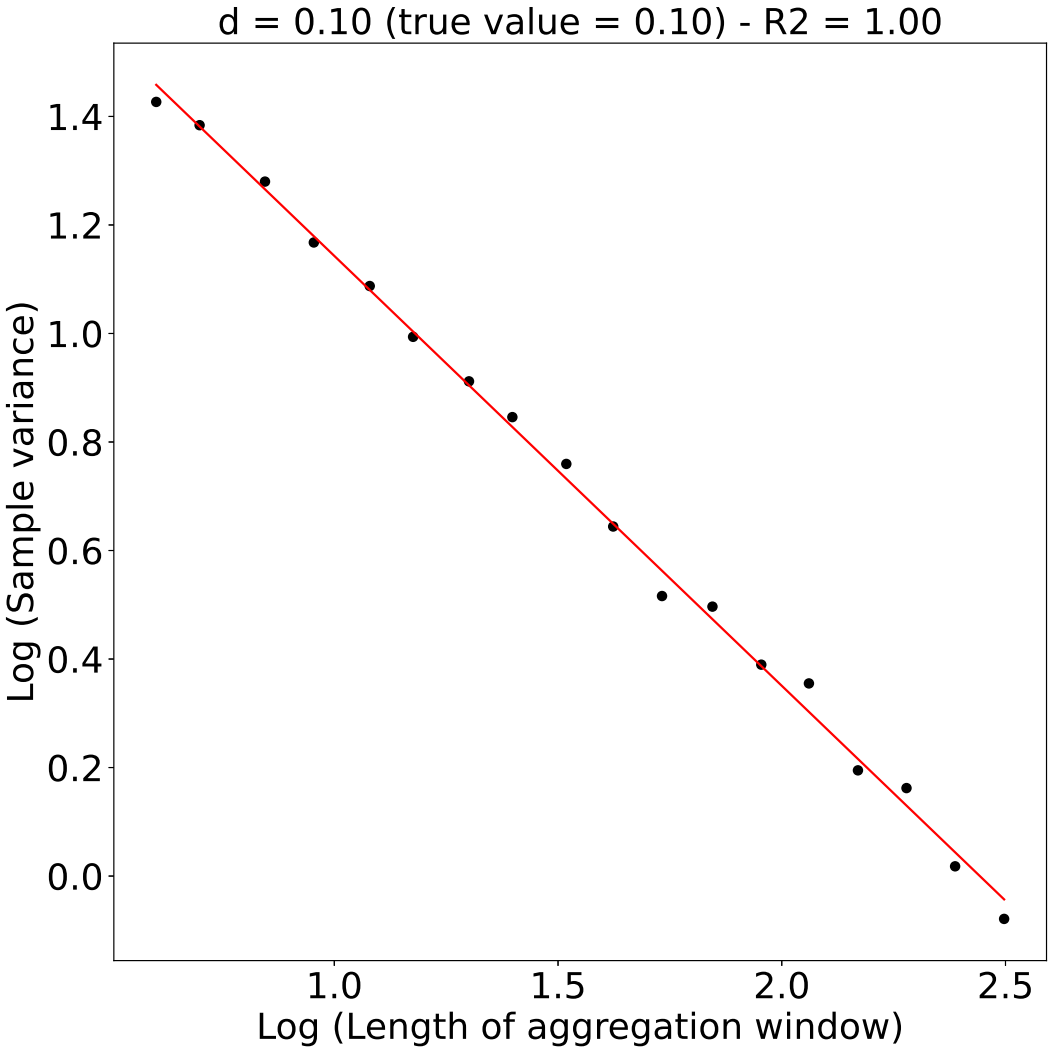

FIGURE 2.3: Variance method applied to the FARIMA time series from Figure 2.1. The horizontal axis is the logarithm of the length m of the aggregation window and the vertical axis is the logarithm of the sample variance of the aggregated time series $X^{(m)}$.

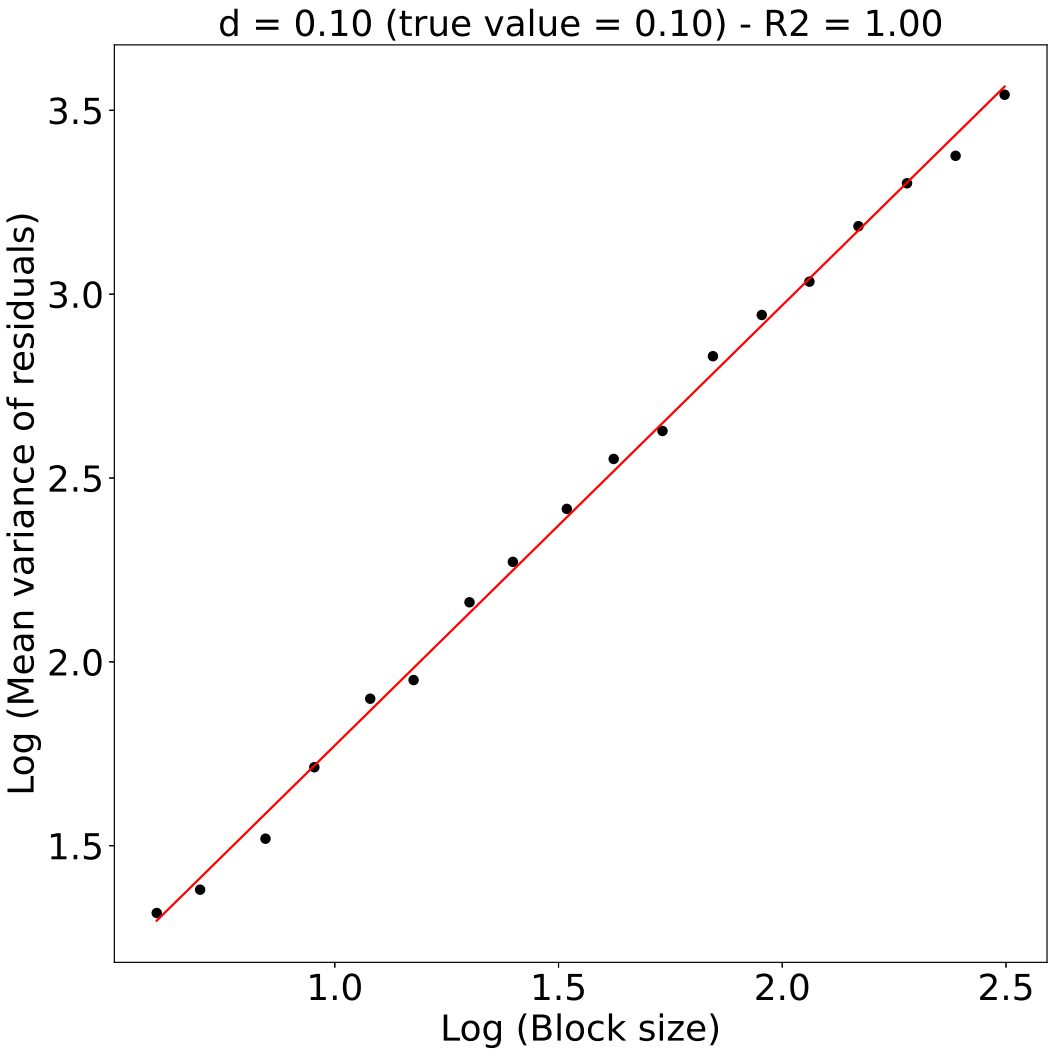

FIGURE 2.4: Variance of residuals method applied to the FARIMA time series from Figure 2.1. The horizontal axis is the logarithm of the size of the blocks m and the vertical axis is the logarithm of the average of the variance of the residuals.

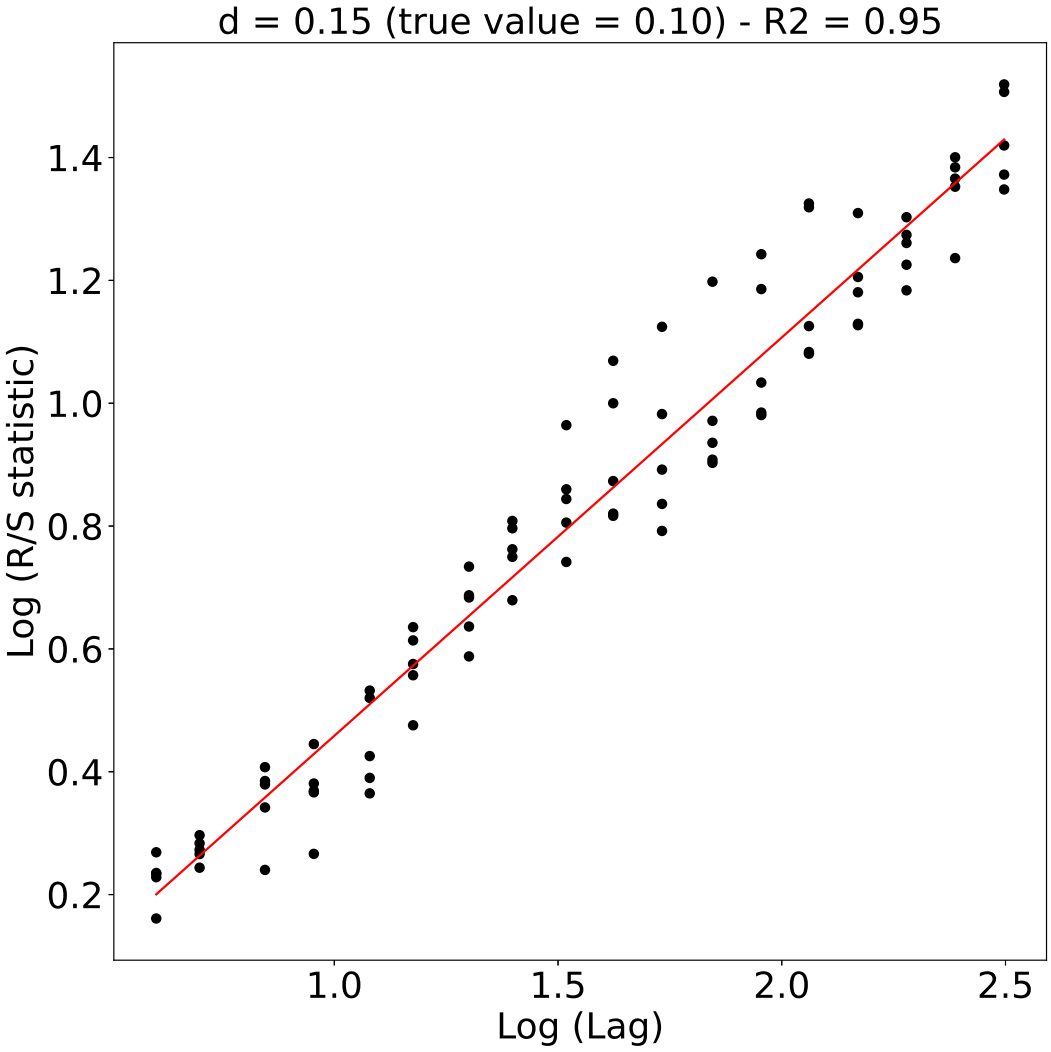

FIGURE 2.5: R/S method applied to the FARIMA time series from Figure 2.1. The horizontal axis is the logarithm of the lag time n and the vertical axis is the logarithm of the ratio $\frac{R(j,n)}{S(j,n)}$.

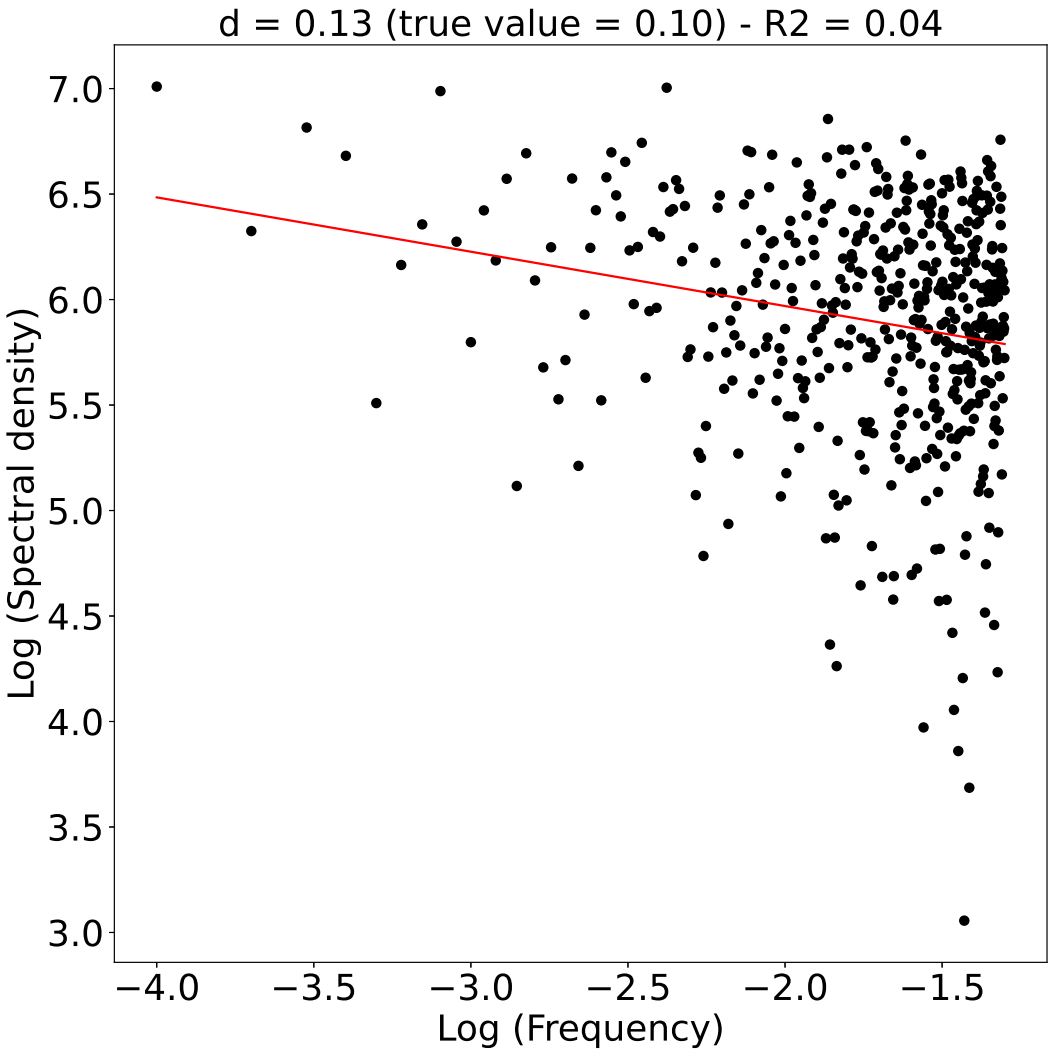

FIGURE 2.6: Periodogram method applied to the FARIMA time series from Figure 2.1. The horizontal axis is the logarithm of the frequency f and the vertical axis is the logarithm of the estimator of the spectral density $I(f)$.

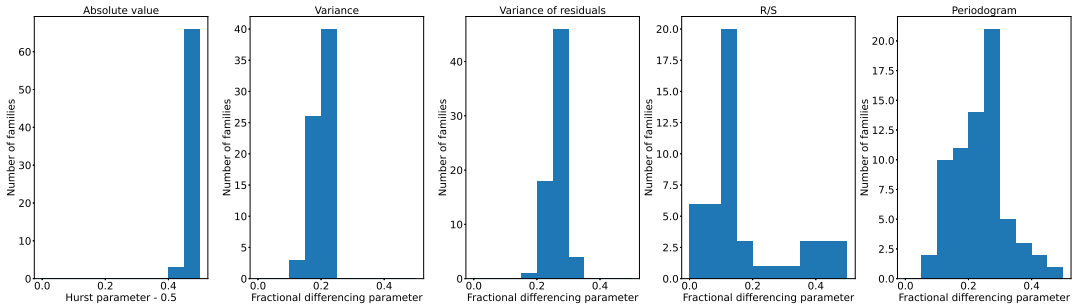

FIGURE 2.7: Distribution of the value of the Hurst parameter H or the fractional differencing parameter d for the catalogs from Chestler and Creager (2017a) and Chestler and Creager (2017b) for the five methods of estimation. For better comparison between the distributions of H and d, I plotted $H - 0.5$ instead of H.

The Northern California catalog The low-frequency earthquake catalog from Ducellier and Creager (in revision) contains event times for 66 families of low-frequency earthquakes located in Northern California. The catalogs start in January 2004 and end in December 2011.

The Frank catalog The low-frequency earthquake catalog from Frank et al. (2014) contains event times for 1120 families of low-frequency earthquakes located in Guerrero, Mexico. The catalogs start in January 2005 and end in April 2007.

The Shelly catalog The low-frequency earthquake catalog from Shelly (2017) contains event times for 88 families of low-frequency earthquakes located on the San Andreas Fault, California. The catalogs start in April 2001 and end in September 2016.

The Sweet catalog The low-frequency earthquake catalog from Sweet et al. (2019) contains event times for 9 families of low-frequency earthquakes located in the Olympic Peninsula, Washington. The catalogs start in October 2006 and end in October 2011.

2.5 Results

I first compared the true value of the fractional differencing parameter d with the computed value for all the synthetic time series and all the graphical methods. The results are shown in Figures 6.1 to 6.5. It is difficult to say which one works best as it may vary depending on the type of innovations used and the value of the fractional differencing parameter. For the time series with Pareto innovations, the variance of residuals method and the R/S method seem to work well. The R/S method has also been widely used to estimate long-range dependence in earthquake catalogs. Figures 2.7 to 2.11 show the value of the Hurst parameter H and the fractional differencing parameter d for all the families of low-frequency earthquakes for the five catalogs studied in this thesis and the five methods used for the estimation.

The R/S method leads to several values of the fractional differencing parameter d lower than 0 or much higher than 0.5 and does not seem to work very well on these datasets. The values of d can be quite different from one method to another.

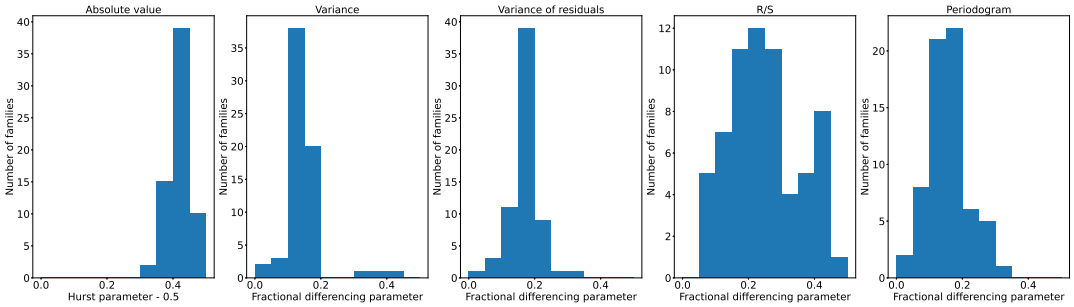

FIGURE 2.8: Distribution of the value of the Hurst parameter H or the fractional differencing parameter d for the catalogs from Ducellier and Creager (in revision) for the five methods of estimation. For better comparison between the distributions of H and d, I plotted $H - 0.5$ instead of H.

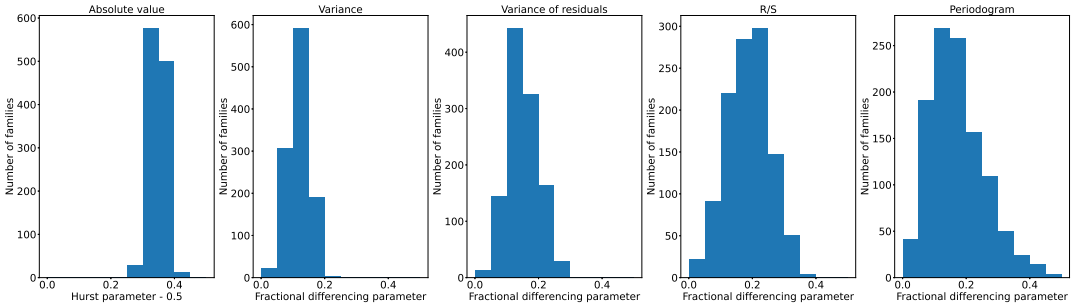

FIGURE 2.9: Distribution of the value of the Hurst parameter H or the fractional differencing parameter d for the catalogs from Frank et al. (2014) for the five methods of estimation. For better comparison between the distributions of H and d, I plotted $H - 0.5$ instead of H.

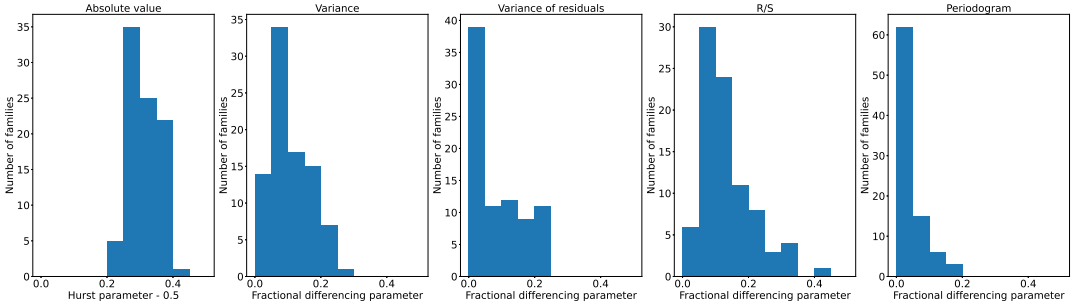

FIGURE 2.10: Distribution of the value of the Hurst parameter H or the fractional differencing parameter d for the catalogs from Shelly (2017) for the five methods of estimation. For better comparison between the distributions of H and d, I plotted $H - 0.5$ instead of H.

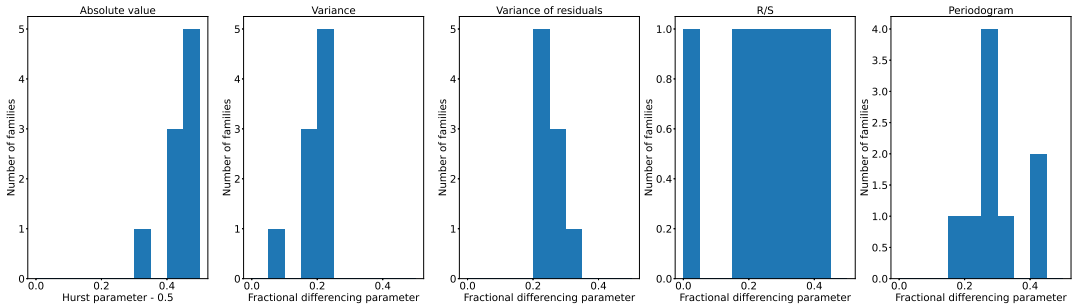

FIGURE 2.11: Distribution of the value of the Hurst parameter H or the fractional differencing parameter d for the catalogs from Sweet et al. (2019) for the five methods of estimation. For better comparison between the distributions of H and d, I plotted $H - 0.5$ instead of H.

However, many low-frequency earthquake families seem to show evidence of long-range dependence. Maps of the values of the fractional differencing parameter are shown in Figures 6.6 to 6.10, but it does not seem that there is any spatial pattern in the variations on the values of d. The fractional differencing parameter can also be quite different from one low-frequency earthquake family to another.

In this chapter, I have analyzed the low-frequency earthquake catalogs as time series datasets, and I have found some evidence of long-range dependence. In the following chapter, I will analyze the low-frequency earthquake catalogs as a point process and estimate the parameters of a statistical model that could generate the low-frequency earthquake events.

Chapter 3

Earthquake occurrence modeling with ETAS

Temporal models of seismicity have been developed to describe, analyze and forecast the probabilities of regular earthquake occurrences (Zhuang et al., 2012). To model long-term seismicity, the most basic model is a stationary Poisson model. Stress-release models assume that the stress level in a certain region linearly increases with time and then drop when there is an earthquake (Zheng and Vere-Jones, 1991). To model short-term seismicity, the Omori-Utsu formula (Omori, 1894; Utsu, 1957) describes the decay of the number of aftershocks after a mainshock as an inverse power law $K(t+c)^{-p}$, where t is the time from the occurrence of the main shock, c is a constant introduced to better fit the model, and the decay rate p for regular earthquake catalogs can vary between 0.6 and 2.5 (Utsu, Ogata, and Matsu'ura, 1995). Short-term exciting, long-term correcting (SELC) models (Schoenberg and Bolt, 2000) combine both the stress release and the clustering aspects of earthquake activity. In this thesis, I am more interested in short-term seismicity and I chose to model earthquake activity with an Epidemic-Type Aftershock Sequence (ETAS) model. The ETAS model was first introduced by Ogata (1988). It takes into account the magnitude frequency distribution law of Gutenberg and Richter (Gutenberg and Richter, 1944) and the Omori-Utsu law of aftershock decay, but it allows each event, irrespective of whether it is a small or a big event, to trigger its own offspring.

In the first section of this chapter, I describe the ETAS model. In the second section of this chapter, I describe the two R packages that I have used to fit the models. Finally, in the third section of this chapter, I fit ETAS models on several LFE families and compare the results from the two R packages.

3.1 Epidemic-Type Aftershock Sequence

In this model, the probability of earthquake occurrence depends on the history of the process. We denote $\mathcal{H}_t$ the history up to but not including time t:

$$\mathcal{H}_t = \{(t_i, m_i)\, \forall i : t_i < t\} \tag{3.1}$$

where t_i is the time of the ith event and m_i is its magnitude. If we denote by $N_\delta(t)$ the number of events occurring in $[t, t+\delta)$, the ground intensity function is defined by:

$$\lambda_g(t \mid \mathcal{H}_t) = \lim_{\delta \to 0} \frac{1}{\delta} Pr\{N_\delta(t) > 0 \mid \mathcal{H}_t\} \tag{3.2}$$

In the ETAS model, the ground intensity function is written as:

$$\lambda_g\left(t \mid \mathcal{H}_t\right) = \mu + \sum_{i:t_i<t} A e^{\alpha(m_i - m_C)} \left(1 + \frac{t - t_i}{c}\right)^{-p} \tag{3.3}$$

where m_C is a threshold magnitude. The term μ represents the background event rate. The term $e^{\alpha(m_i - m_C)}$ indicates that bigger earthquakes generate more aftershocks. The last term corresponds to the Omori-Utsu law that says that the number of aftershocks decays as $K\left(t + c\right)^{-p}$.

Each of the magnitudes of the events m_i follows an exponential distribution. It is assumed that the magnitudes of the events are independent and identically distributed, and that they are also independent from the history of the process:

$$Pr\left(m_i > m\right) = e^{-\beta(m - m_C)} \text{ for } m > m_C \tag{3.4}$$

where m_C is the magnitude of completeness of the catalog. For regular earthquake catalogs, we have β close to 2.3. For low-frequency earthquake catalogs, Sweet, Creager, and Houston (2014) and Bostock et al. (2015) found much bigger values equal to 9.2 and 14.5 respectively. The seismic moment M is related to the magnitude m by the relationship:

$$M = 10^{1.5m + 9.1} \tag{3.5}$$

As the magnitude follows an exponential distribution, the seismic moment follows a Pareto distribution. However, Sweet, Creager, and Houston (2014) suggested that for low-frequency earthquakes an exponential distribution of the seismic moment could be more appropriate than a Pareto distribution.

3.2 Fitting the model

To fit the model, I have used two R packages: PtProcess by Harte (2010) and bayesianETAS by Ross (2016).

To compute the likelihood, we denote τ the time of the last event before time t, and $\emptyset_{(\tau,t)}$ the null outcome, that is there is no event between time τ and time t. The conditional distribution of the time of the next event, that is the probability that the time t of the next event is lower or equal to T given the history until time τ ($\mathcal{H}_\tau$) and the fact that there is no event between time τ and time t, is denoted:

$$H\left(t \mid \mathcal{H}_\tau \cap \emptyset_{(\tau,t)}\right) = Pr\left\{T \geq t \mid \mathcal{H}_\tau \cap \emptyset_{(\tau,t)}\right\} \tag{3.6}$$

and the corresponding density function is denoted $h\left(t \mid \mathcal{H}_\tau \cap \emptyset_{(\tau,t)}\right)$. Then we have:

$$\lambda_g\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right) = \frac{h\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right)}{1 - H\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right)} \tag{3.7}$$

Solving the differential equation gives:

$$H\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right) = 1 - \exp\left\{-\int_\tau^t \lambda_g\left(u \mid \mathcal{H} \cap \emptyset_{(\tau,u)}\right) du\right\} \tag{3.8}$$

and:

$$h\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right) = \lambda_g\left(t \mid \mathcal{H} \cap \emptyset_{(\tau,t)}\right) \exp\left\{-\int_\tau^t \lambda_g\left(u \mid \mathcal{H} \cap \emptyset_{(\tau,u)}\right) du\right\} \quad (3.9)$$

The PtProcess package fits the model by maximizing the likelihood of the ground intensity function. If T_1 and T_2 are the beginning and end time of the catalog and we only explicitly include in the likelihood the events which time t_i are such that $T_1 \leq t_i \leq T_2$, the log-likelihood is equal to:

$$\begin{aligned}
\log L &= \log h\left(t_1 \mid \mathcal{H}_{T_1} \cap \emptyset_{(T_1,t_1)}\right) + \sum_{i=2}^{n} \log h\left(t_i \mid \mathcal{H}_{t_i-1} \cap \emptyset_{(t_{i-1},t_i)}\right) \\
&\quad + \log\left(1 - H\left(T_2 \mid \mathcal{H}_{t_n} \cap \emptyset_{(t_n,T_2)}\right)\right) \\
&= \sum_{i=1}^{n} \log \lambda_g\left(t_i \mid \mathcal{H}_{t_i}\right) - \int_{T_1}^{t_1} \lambda_g\left(u \mid \mathcal{H}_{T_1} \cap \emptyset_{(T_1,u)}\right) du \\
&\quad - \sum_{i=2}^{n} \int_{t_{i-1}}^{t_i} \lambda_g\left(u \mid \mathcal{H}_{t_{i-1}} \cap \emptyset_{(t_{i-1},u)}\right) du - \int_{t_n}^{T_2} \lambda_g\left(u \mid \mathcal{H}_{t_n} \cap \emptyset_{(t_n,u)}\right) du \\
&= \sum_{i:T_1 \leq t_i \leq T_2} \log \lambda_g\left(t_i \mid \mathcal{H}_{t_i}\right) - \int_{T_1}^{T_2} \lambda_g\left(t \mid \mathcal{H}_t\right) dt
\end{aligned} \quad (3.10)$$

For the distribution of the magnitude, the value of β is estimated independently by subtracting the value of the magnitude of completeness m_C to all the magnitudes m_i and then computing the value of the maximum likelihood estimator for an exponential distribution. For the LFE catalogs, I used a magnitude of completeness equal to 0.

The bayesianETAS package fits the model using a Bayesian framework. The ground intensity is written a little differently as:

$$\lambda_g\left(t \mid \mathcal{H}_t\right) = \mu + \sum_{i:t_i<t} K\left(p-1\right) c^{p-1} e^{\alpha(m_i - m_C)} \frac{1}{\left(c + t - t_i\right)^p} \quad (3.11)$$

which is equivalent to Equation 3.3 if $A = K\frac{p-1}{c}$. Priors are chosen for the five parameters: a Gamma distribution for μ, and uniform distributions for $\log K$, $\log \alpha$, c and p. Samples of the posterior distribution are then drawn using a MCMC scheme.

3.3 Results

I focused on the catalogs by Chestler and Creager (2017a) and Chestler and Creager (2017b) as they are the only ones for which magnitude is provided. For each of the families in the catalog, I fitted an ETAS model with the two R packages using the maximum likelihood estimator for both cases. Figure 3.1 shows a comparison of the results from both packages for the 5 parameters of the ETAS model.

I obtained very different values with the two packages. For the PtProcess package, Harte (2010) suggests to choose initial values for the parameters, take the logarithm, then maximize the likelihood using the R function optim, then use the results of the first optimization process as an input for a second optimization carried

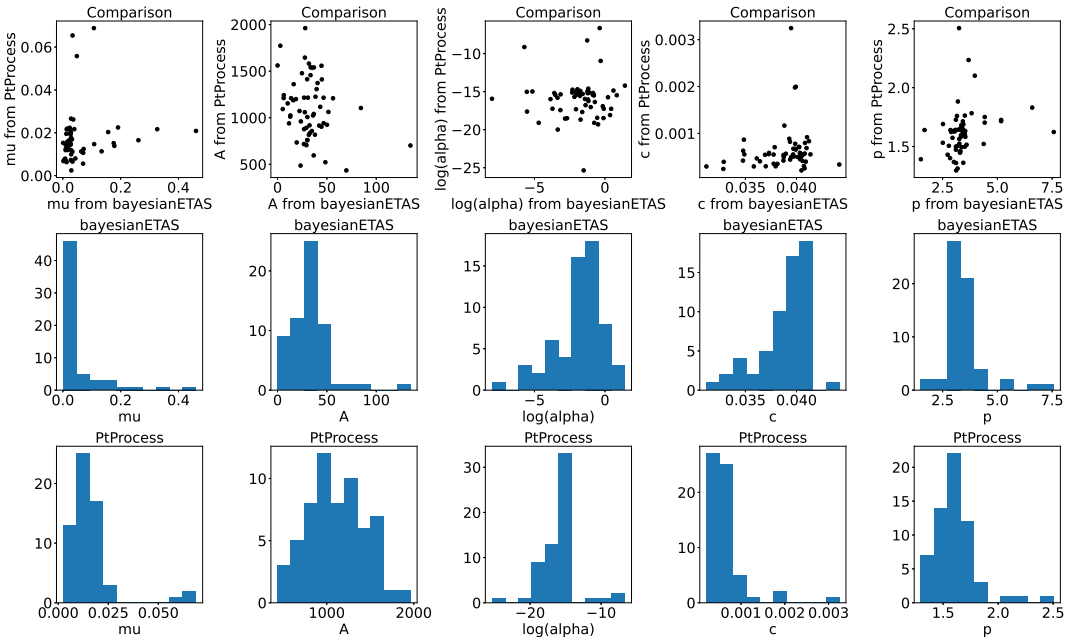

FIGURE 3.1: Top: Comparison between the maximum likelihood estimate obtained with the bayesianETAS package and the maximum likelihood estimate obtained with the PtProcess package for the five parameters of the ETAS model. Middle: Distribution of the values obtained with the bayesianETAS package for all the LFE families. Bottom: Distribution of the values obtained with the PtProcess package for all the LFE families.

out with the R function nlm. We then obtain the maximum likelihood estimator by taking the exponential of the result. It is also possible to compute the maximum likelihood estimator with the package bayesianETAS although it is unclear how the optimization is done. A first reason for the difference in the results from the two packages is that the ground intensity function is not written exactly in the same way in both packages: bayesianETAS uses the parameter K and PtProcess uses the parameter $A = K\frac{p-1}{c}$. Another problem is that the 5 parameters of the ground intensity function function are not independent of each other. Finally, the user has to choose starting values for the package PtProcess, but it is unclear what the starting values are for the package bayesianETAS.

In the MCMC method from the package bayesianETAS, the parameters sets $\{\mu\}$, $\{K, \alpha\}$ and $\{p, c\}$ are independent of each other conditional on some latent variables, and it may improve the results. Indeed, the results from the bayesian method can be quite different from the results from the maximum likelihood method. Figures 7.1 to 7.5 show a comparison of the values of μ, A, α, c, p obtained with the maximum likelihood method of the package bayesianETAS with the sample mean of the posterior distributions obtained with the package bayesianETAS for all families. In order to improve the results from the optimization with PtProcess, it might be interesting to use the parameter estimates from bayesianETAS (for example the sample mean of the posterior distribution) as the starting values.

I now focus one one LFE family (2009.11.12.6.12.19). The ground intensity function parameters obtained with the package PtProcess for this family are:

$$\begin{cases} \mu = 0.0139 \\ A = 836.4 \\ c = 0.000587 \\ p = 1.504 \\ \alpha = 2.49510^{-7} \end{cases} \tag{3.12}$$

while the parameters obtained with the maximum likelihood method from the package bayesianETAS for this family are:

$$\begin{cases} \mu = 0.178 \\ A = 32.76 \\ c = 0.0403 \\ p = 3.165 \\ \alpha = 0.296 \end{cases} \tag{3.13}$$

Figure 3.2 shows the posterior distribution of the parameters obtained from the package bayesianETAS. For parameters c and p, the posterior distribution is limited to a single point. The cause of this anomaly is unclear. This may be due to the prior distribution chosen for these parameters, but its is not possible to change the prior with the package bayesianETAS.

The goodness of fit of the model can be evaluated by computing the transformed times. If we denote t_i the times of the observed events, the transformed times are:

$$\tau_i = \int_0^{t_i} \widehat{\lambda}_g \left(t \mid \mathcal{H}_t \right) dt \tag{3.14}$$

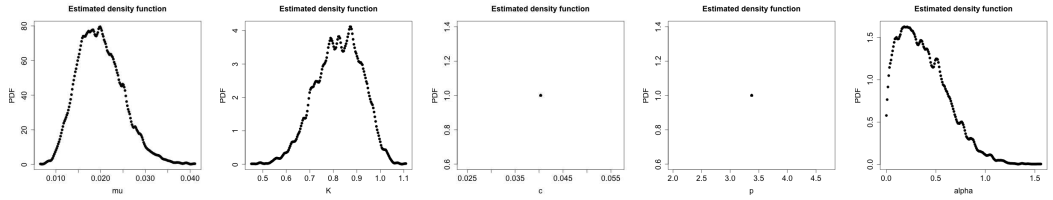

FIGURE 3.2: Posterior distribution of the parameters μ, K, c, p and α obtained with package bayesianETAS for family 2009.11.12.6.12.19.

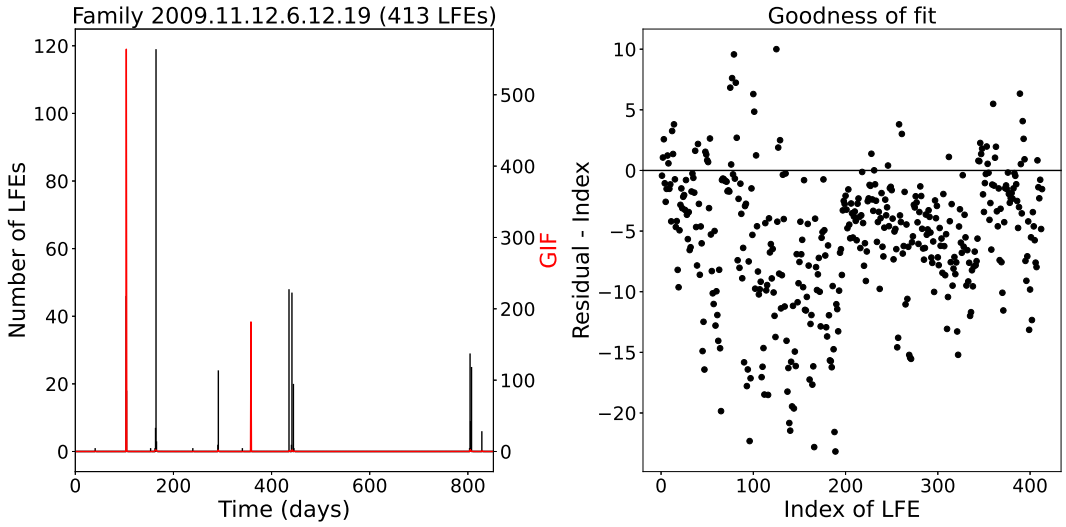

FIGURE 3.3: Left: Number of LFEs per day (black) for family 2009.11.12.6.12.19 and fitted ground intensity function (red). Right: Residual process $\tau_i - i, i = 1, \cdots, n$ as a function of $i, i = 1, \cdots, n$ for family 2009.11.12.6.12.19.

where $\widehat{\lambda}_g(t \mid \mathcal{H}_t)$ is the fitted ground intensity function. The sequence of transformed times is called the residual process. If the data were sampled from the fitted ground intensity function, then the transform times should form a stationary Poisson process with rate parameter equal to 1. To verify if this is the case, we can plot the $\tau_i, i = 1, \cdots, n$ as a function of $i = 1, \cdots, n$ and check if they follow the first diagonal. Another way of representing it is to plot $\tau_i - i, i = 1, \cdots, n$ as a function of $i = 1, \cdots, n$ and check if it has a zero slope.

Figure 3.3 shows the fitted ground intensity function and the $\tau_i - i, i = 1, \cdots, n$ for family 2009.11.12.6.12.19. There is a peak (at about 350 days) in the ground intensity function that does not correspond to a burst in LFE activity. There are also a lot of deviations from a horizontal line for the $\tau_i - i, i = 1, \cdots, n$ which indicates that the fitted model may not fit the data very well.

In this chapter, I have analyzed the low-frequency earthquake catalogs as a point process and I have fitted an Epidemic-Type Aftershock Sequence model for each family of the catalog by Chestler and Creager (2017a) and Chestler and Creager (2017b). In the next chapter, I will explore whether this kind of model can generate earthquake catalogs that show some evidence of long-range dependence.

Chapter 4

Stochastic modeling

In the second chapter of this thesis, I showed that several families of low-frequency earthquakes show evidence of apparent long-range dependence. In the third chapter of this thesis, I modeled the low-frequency earthquake catalogs with an Epidemic-Type Aftershock Sequence model and fitted the parameters of these models. I now verify whether it is possible to observe apparent long-range dependence with this type of model. In the first section of this chapter, I explain how to generate synthetic earthquake catalogs using the thinning algorithm. In the second section of this chapter, I use the parameters from the model obtained in Chapter 3 to generate synthetic catalogs and I try to fit FARIMA models on them.

4.1 The thinning algorithm

I use the same method as in the R package PtProcess to simulate earthquake catalogs. The ground intensity function $\lambda_g(t \mid \mathcal{H}_t)$ decreases with time t. If we generate an event during a simulation interval starting at time τ using the maximum event rate over this interval $\lambda_{max} = \lambda_g(\tau \mid \mathcal{H}_\tau)$, the inter-event time will be too small and we will generate too much events. The idea of the thinning algorithm is thus to first generate too much events and then to "thin", that is eliminate some of the events generated. The procedure is the following:

1. Start at time τ.
2. Calculate the ground intensity function at time τ: $\lambda_g(\tau \mid \mathcal{H}_\tau)$.
3. Simulate a random number ξ that follows an exponential distribution with rate $\lambda_g(\tau \mid \mathcal{H}_\tau)$.
4. Simulate a random number U that follows a uniform distribution on the interval $(0, 1)$. If:

$$U \leq \frac{\lambda_g(\tau + \xi)}{\lambda_g(\tau)}$$

   then a new event occurs at time $t = \tau + \xi$. Simulate the associated magnitude of the event using an exponential distribution with rate β as in equation 3.4.
5. Increment τ for the simulation of the next event: $\tau \leftarrow \tau + \xi$
6. Return to step 2.

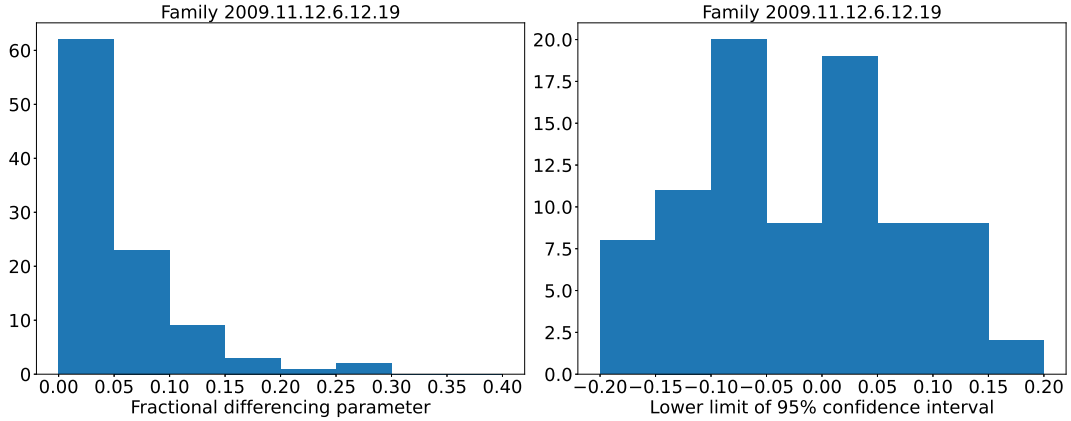

FIGURE 4.1: Left panel: Distribution of the fractional differencing parameter obtained from 100 simulated earthquake sequences using the ETAS model parameters obtained for family 2009.11.12.6.12.19 with package bayesianETAS. Right panel: Distribution of the lower boundary of the 95 % confidence interval for the parameter d of the fitted models.

4.2 Numerical simulations

I first used the results from the package bayesianETAS. For each family of low-frequency earthquakes, I used the posterior distribution of the ETAS model parameters (μ, K, c, p, α) and I randomly generated a sample. I then used the thinning algorithm to generate a sequence of earthquake events with associated magnitude. I then transformed the earthquake sequence into a time series by counting the number of earthquakes per day. I fitted a FARIMA model on this time series using the R package fracdiff and obtained the corresponding value of the fractional differencing parameter d. I then repeated this process 100 times. Figure 4.1 shows as an example the distribution of the fractional differencing parameter d for the family 2009.11.12.6.12.19. To verify whether there is long-range dependence in the time series, I also computed the 95 % (Wald) confidence interval of the fitted model for the parameter d using the function confint.fracdiff of the package fracdiff. If this interval contains 0, we cannot reject the hypothesis that there is no long-range dependence in the time series. The distribution of the lower boundary of the 95 % confidence interval for the fitted fractional differencing parameter d and for family 2009.11.12.6.12.19 is also shown on Figure 4.1.

For the package PtProcess, we only have the parameters estimated using the maximum likelihood. I generated 100 earthquake sequences with the same set of parameters and computed the distribution of the fractional differencing parameter d for the family 2009.11.12.6.12.19 (Figure 4.2).

Finally, Figure 4.3 shows the distribution of the fractional differencing parameter d for all the families of the catalog from Chestler and Creager (2017a) and Chestler and Creager (2017b) and the distribution of the lower boundary of the 95 % confidence interval for the parameter d of the fitted models. In both cases, about half of the simulated time series obtained using the parameters from the fitted ETAS model do not show evidence of long-range dependence. This may be due because the fitted

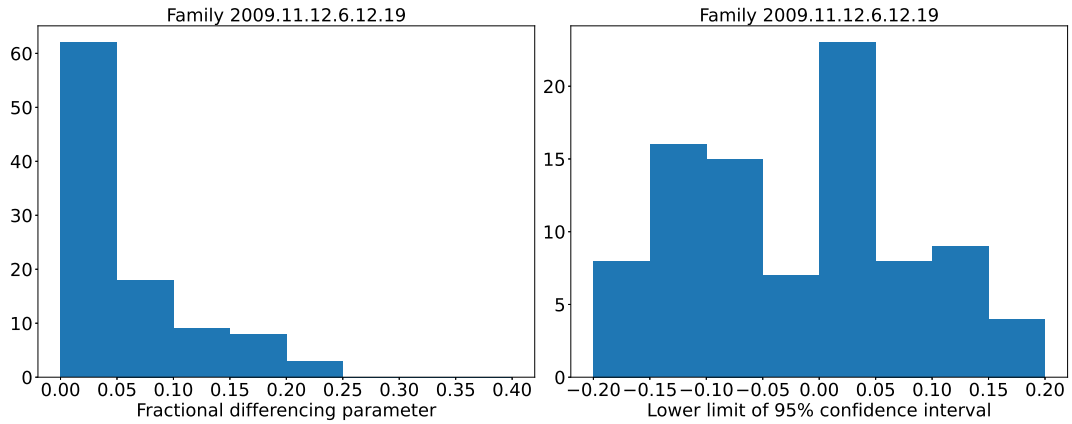

FIGURE 4.2: Left panel: Distribution of the fractional differencing parameter obtained from 100 simulated earthquake sequences using the ETAS model parameters obtained for family 2009.11.12.6.12.19 with package PtProcess. Right panel: Distribution of the lower boundary of the 95 % confidence interval for the parameter d of the fitted models.

models did not seem to fit well the initial earthquake sequences. Another problem is that the simulated earthquake sequences did not show characteristic features of a low-frequency earthquake catalog, with bursts of earthquake activity for several days followed by weeks or months of much lower earthquake activity.

For about half of the simulations, the lower boundary of the 95 % confidence interval is negative, which means than the confidence interval either contains 0, or contains only negative values. This means that for about half of the simulations, we cannot reject the hypothesis that there is no long-range dependence in the simulated time series. Therefore, the ETAS models obtained in Chapter 3 are not very successful at reproducing the long-range dependence observed in Chapter 2.

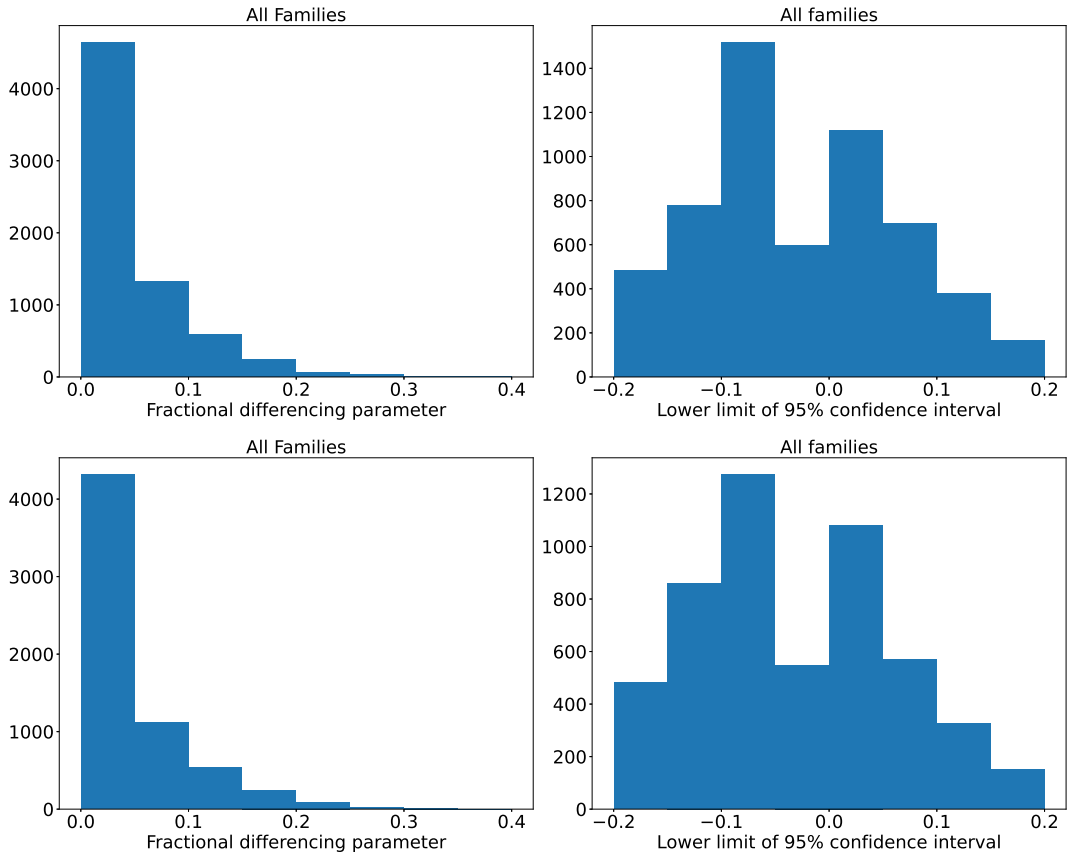

FIGURE 4.3: Left panels: Distribution of the fractional differencing parameter obtained from simulated earthquake sequences using the ETAS model parameters obtained for all families with packages bayesianETAS (top) and PtProcess (bottom). Right panels: Distribution of the lower boundary of the 95 % confidence interval for the parameter d of the fitted models obtained with packages bayesianETAS (top) and PtProcess (bottom).

Chapter 5

Perspectives

For research beyond the scope of this thesis, I first intend to improve the methods to estimate the fractional differencing parameter d. In the periodogram method, the regression is done over the whole range of frequencies. However, as the exact form of the spectral density is known for the synthetic time series generated with a FARIMA model, it is possible to compare the values of the spectral density with the values of the periodogram and determine for which interval of frequencies the linear relationship holds and the linear regression can be used to estimate the value of d.

I only generated one possible realization of each type of FARIMA model studied. It would be more interesting to generate several realizations of the same FARIMA model to be able to see the distribution of the estimated value of the fractional differencing parameter. Additionally, the synthetic time series have length $N = 10000$ as in Taqqu and Teverovsky (1998), which is different from the length of the LFE time series. The length of the LFE time series depends only on the value of the unit of time chosen for the transformation of the catalog into a time series and any value can be chosen. In the future, I intend to test several values of the length of the synthetic time series to verify whether some methods to detect long-range dependence work better than others depending on the length of the time series and the value of the fractional differencing parameter. Then, the value of the unit of time used to transform the LFE catalogs into time series can be chosen in such a way that the final time series has the same length as the synthetic time series.

By estimating both the value of the Hurst parameter H and the fractional differencing parameter d for a single LFE time series, it is possible to deduce what could be the value of the corresponding shape of the Pareto distribution α. Then, I intend to make new synthetic time series using this value of α to evaluate how well the different methods work on a synthetic time series as close as possible to the actual time series.

Low-frequency earthquakes are small magnitude earthquakes and the signal-to-noise ratios of the seismic recordings are low. As a result, many low-magnitude LFEs may be missing from the datasets, and many LFEs may actually be false detections. It is reasonable to assume that the false detections is a homogeneous Poisson process and thus the presence of false detections in the earthquake catalog will lead to an overestimate of the value of the background seismicity rate μ in the ETAS model. However, one important parameter of the ETAS model is the value of the magnitude of completeness m_C, that is the magnitude above which 100% of the earthquakes are detected and there is no missing data. When fitting the ETAS model, only earthquakes whose magnitudes are higher than m_C should be kept and the others should be discarded. In a future work, it would be useful to try to establish what could be

the magnitude of completeness of the LFE catalogs used in this thesis.

The package bayesianETAS does not currently contain methods to compute the fitted ground intensity function and the corresponding residuals. In a future work, I intend to extend the goodness-of-fit analysis done for the models obtained with the package PtProcess and do the same goodness-of-fit analysis with the models obtained with the package bayesianETAS.

Chapter 6

Supplementary figures for Chapter 2

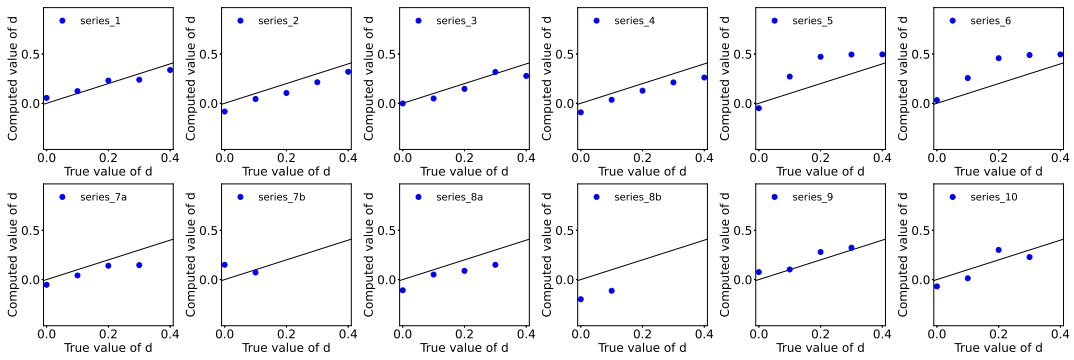

FIGURE 6.1: Comparison between the true value of the fractional differencing parameter and the computed value of the fractional differencing parameter with the absolute value method for the ten synthetic time series.

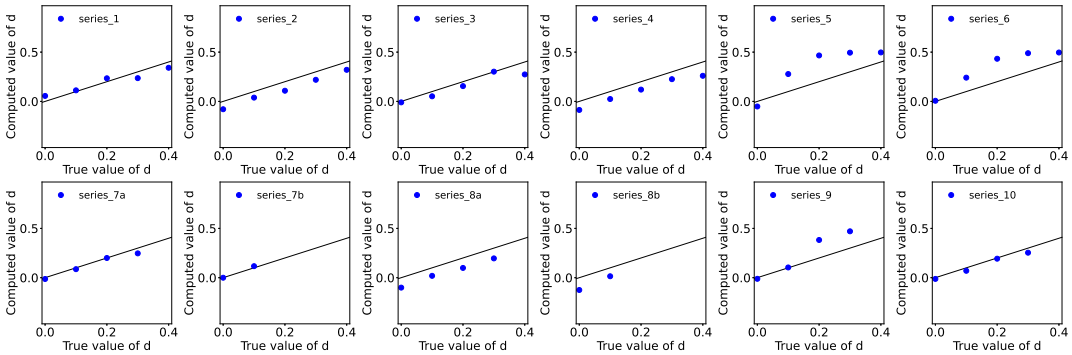

FIGURE 6.2: Comparison between the true value of the fractional differencing parameter and the computed value of the fractional differencing parameter with the variance method for the ten synthetic time series.

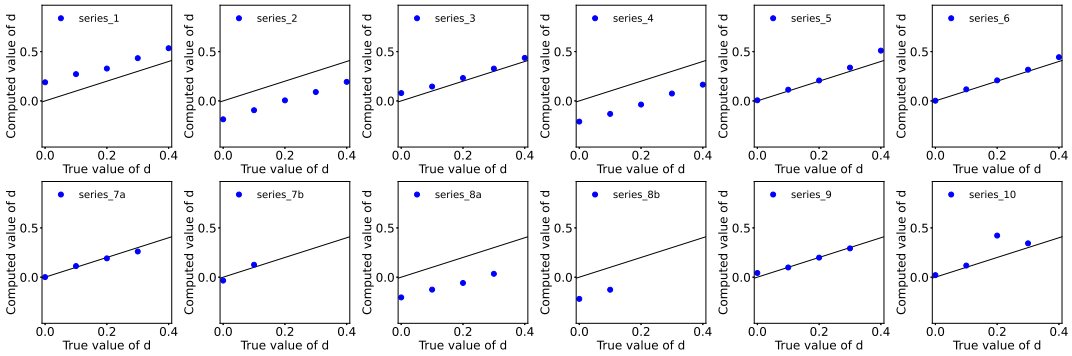

FIGURE 6.3: Comparison between the true value of the fractional differencing parameter and the computed value of the fractional differencing parameter with the variance of residuals method for the ten synthetic time series.

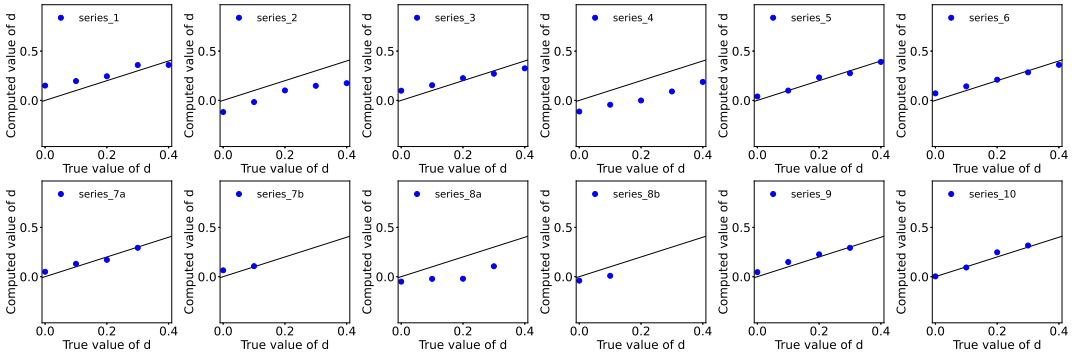

FIGURE 6.4: Comparison between the true value of the fractional differencing parameter and the computed value of the fractional differencing parameter with the R/S method for the ten synthetic time series.

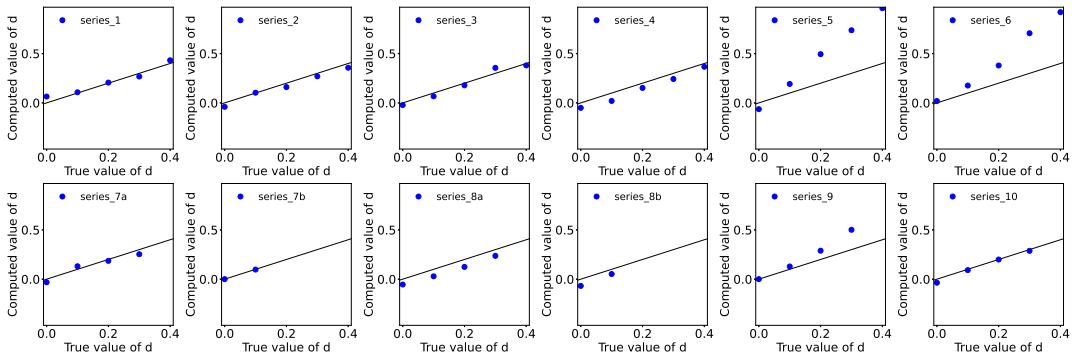

FIGURE 6.5: Comparison between the true value of the fractional differencing parameter and the computed value of the fractional differencing parameter with the periodogram method for the ten synthetic time series.

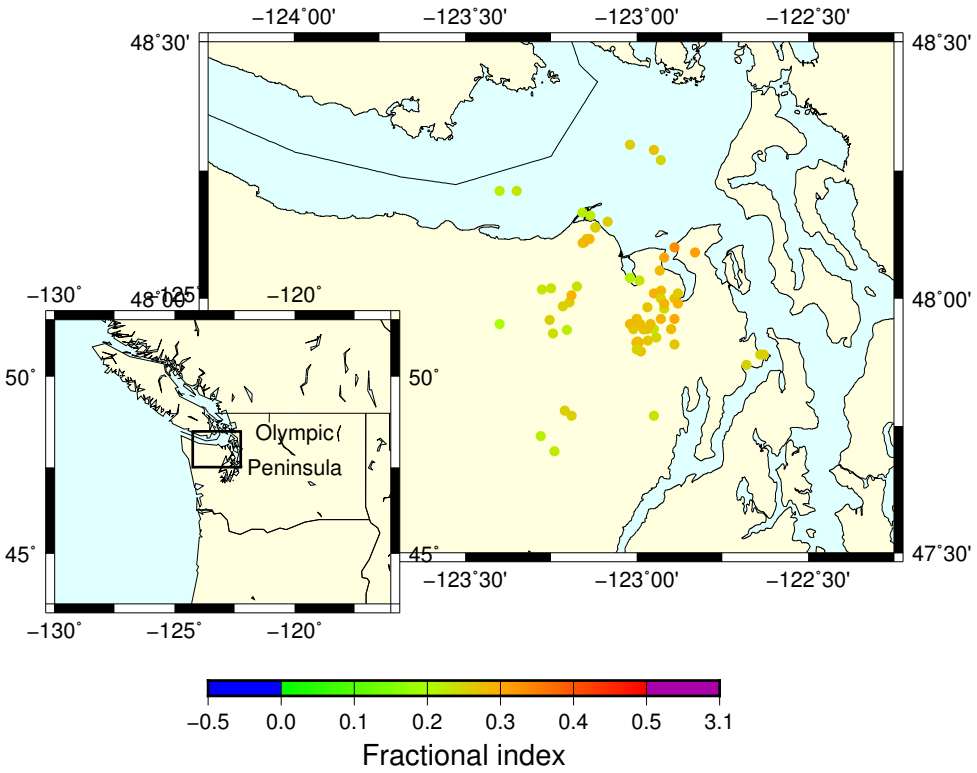

FIGURE 6.6: Map of the fractional differencing parameters computed with the variance of residuals method for the Chestler catalogs.

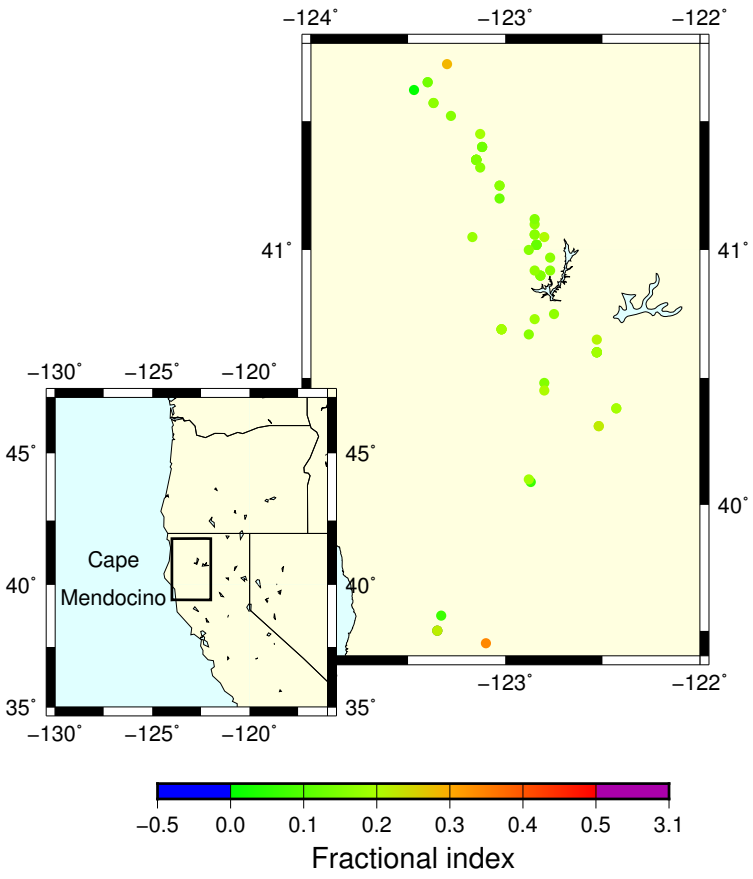

FIGURE 6.7: Map of the fractional differencing parameters computed with the variance of residuals method for the Northern California catalogs.

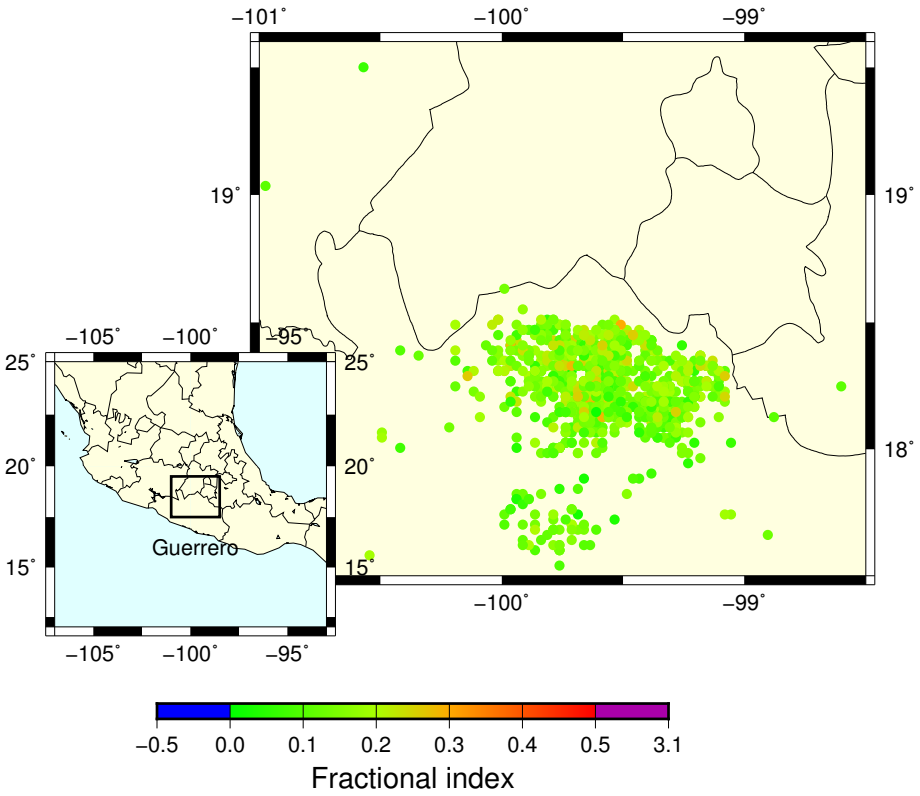

FIGURE 6.8: Map of the fractional differencing parameters computed with the variance of residuals method for the Frank catalogs.

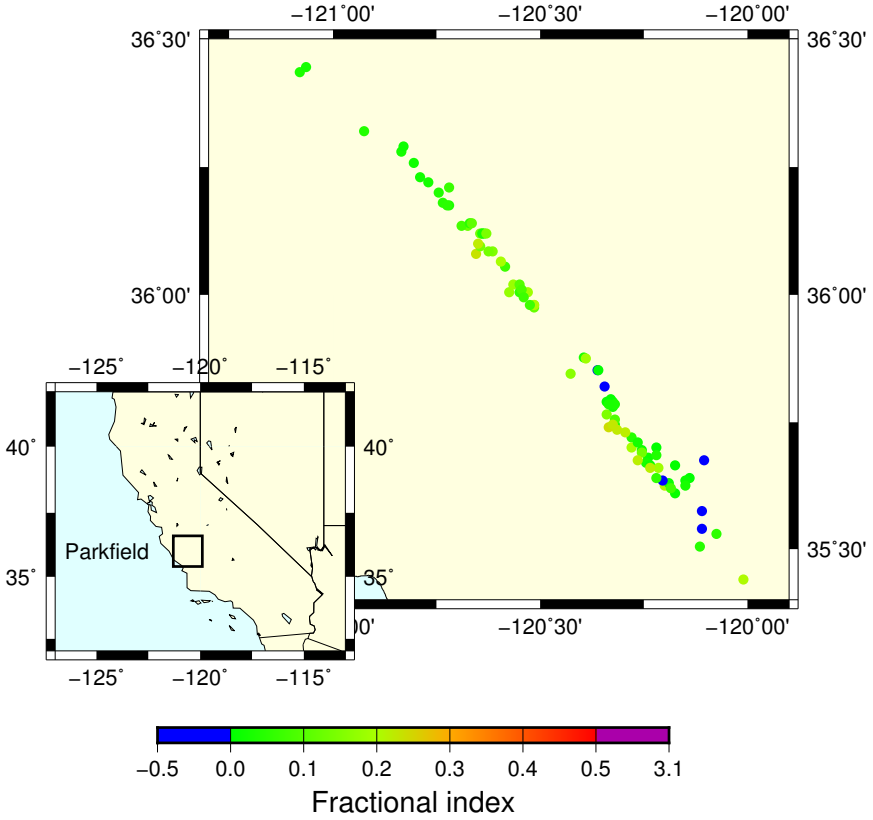

FIGURE 6.9: Map of the fractional differencing parameters computed with the variance of residuals method for the Shelly catalogs.

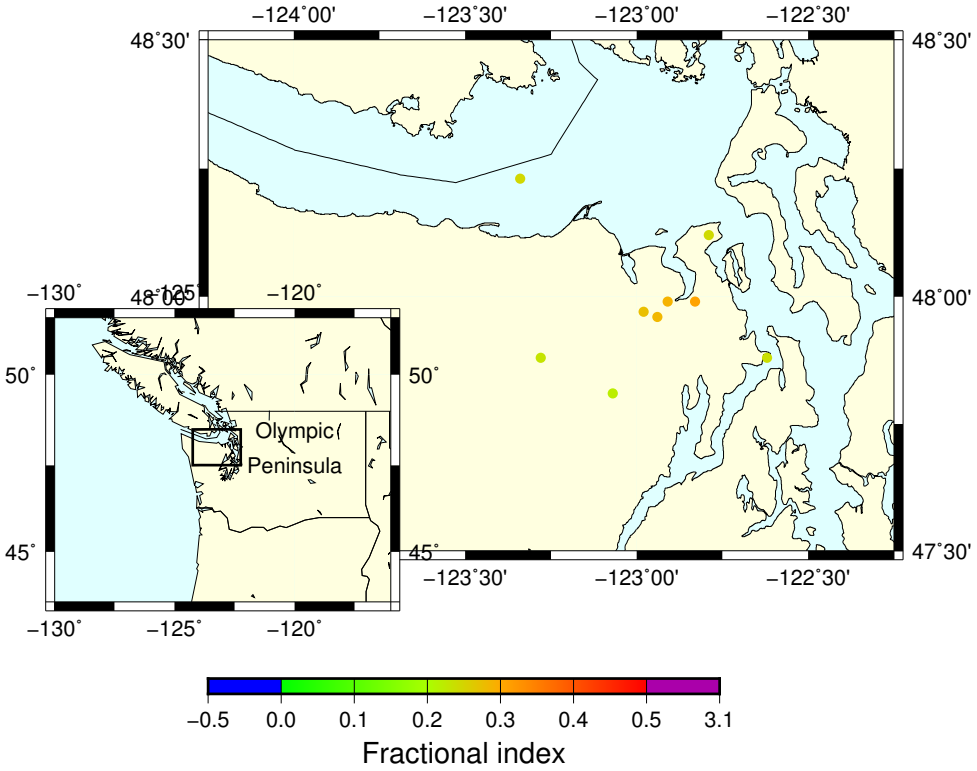

FIGURE 6.10: Map of the fractional differencing parameters computed with the variance of residuals method for the Sweet catalogs.

Chapter 7

Supplementary figures for Chapter 3

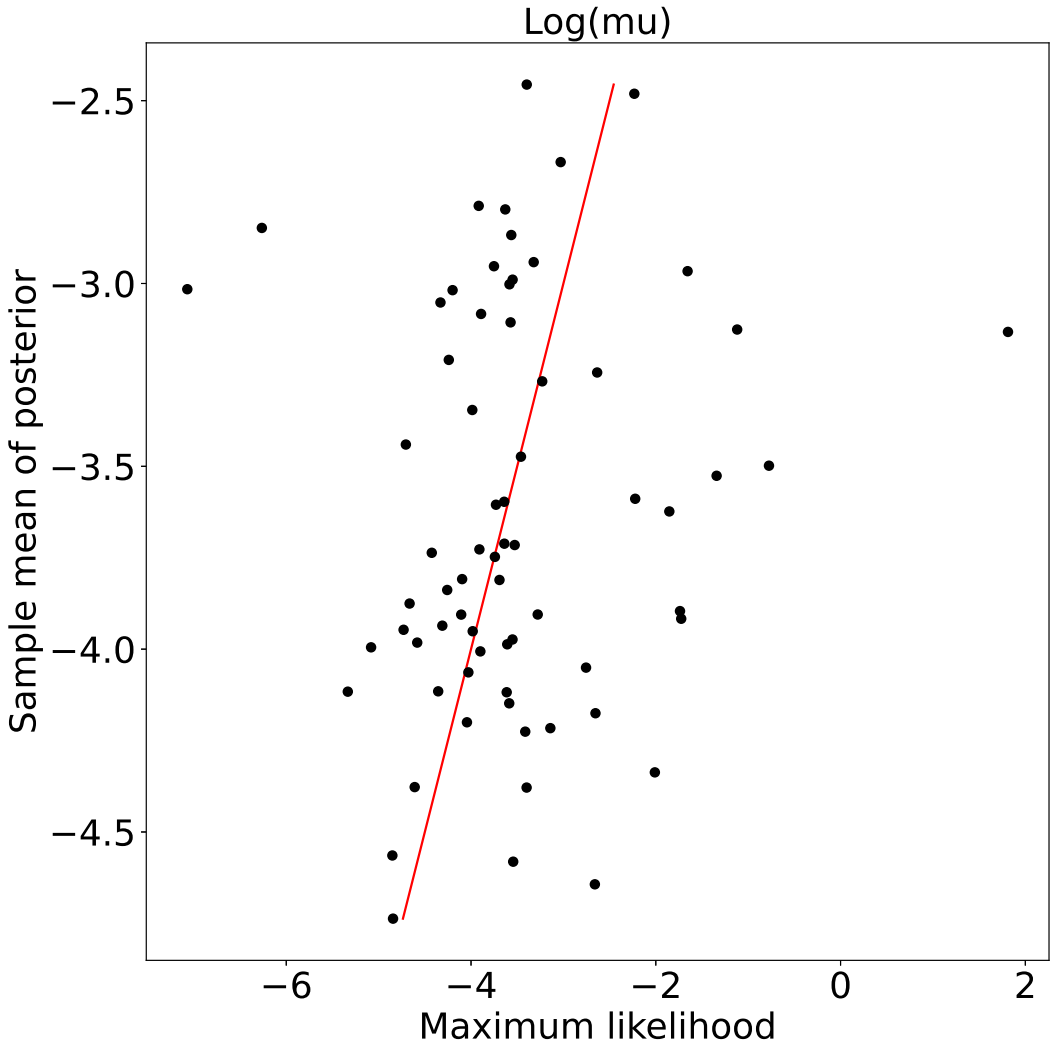

FIGURE 7.1: Comparison of the values of $\log(\mu)$ obtained with the maximum likelihood method with the logarithm of the sample mean of the posterior distribution of μ obtained with the package bayesianETAS.

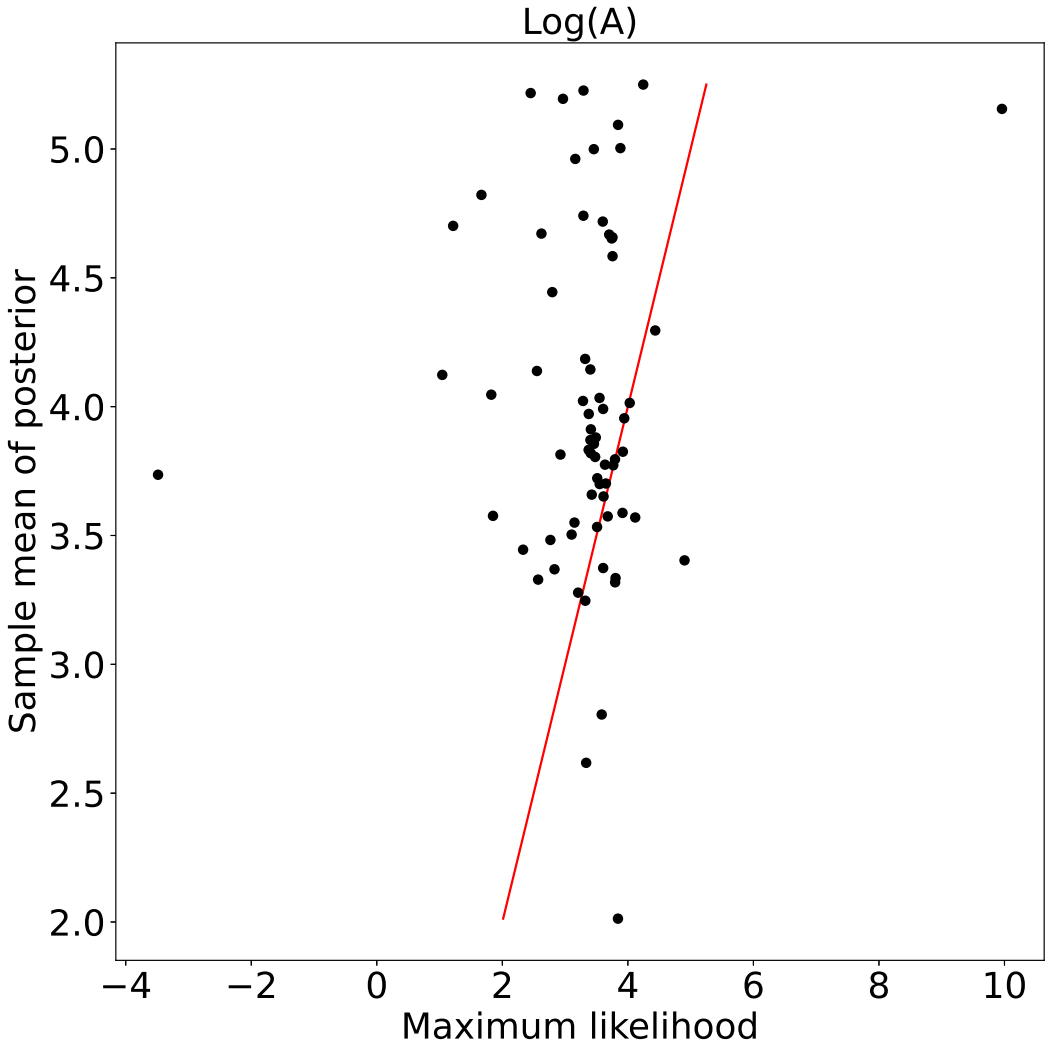

FIGURE 7.2: Comparison of the values of $\log(A)$ obtained with the maximum likelihood method with the logarithm of the sample mean of the posterior distribution of A obtained with the package bayesianETAS.

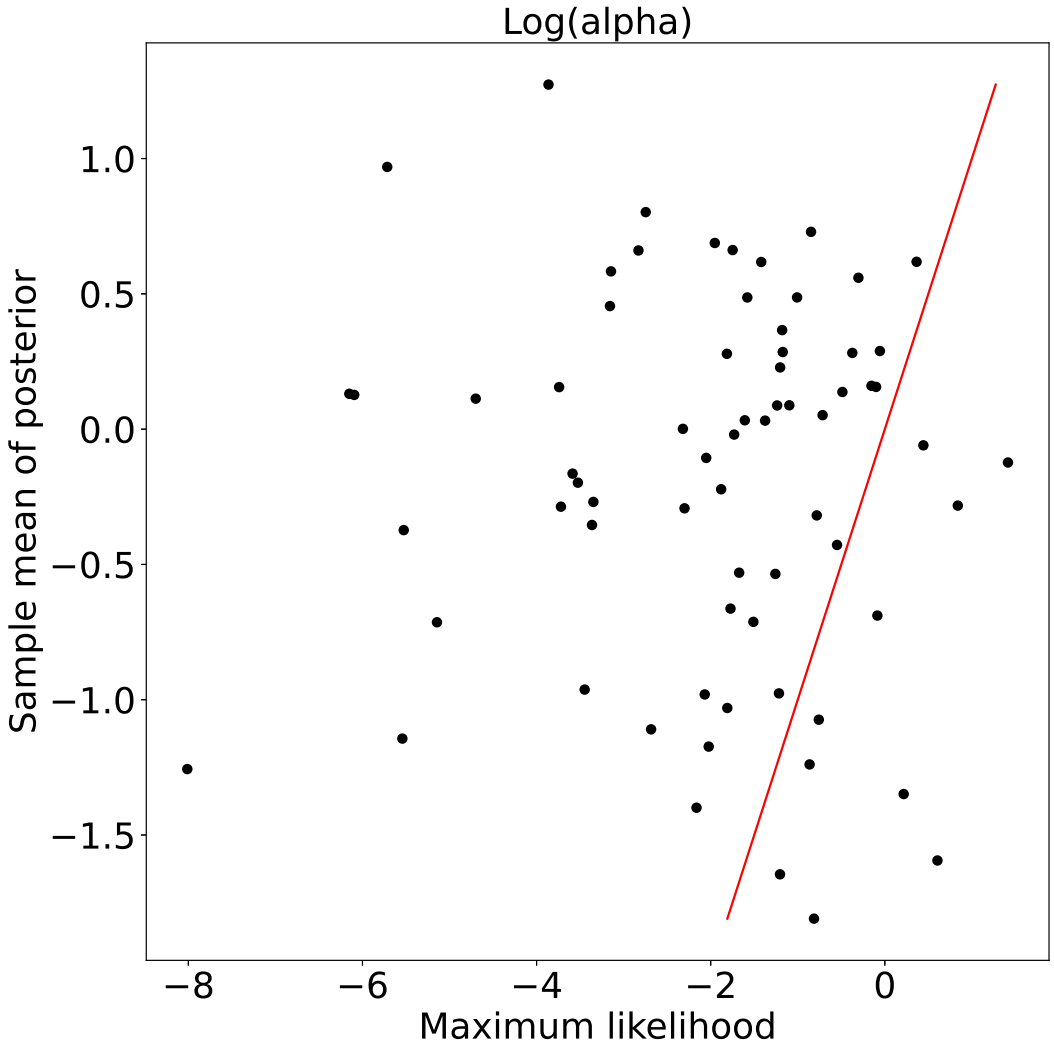

FIGURE 7.3: Comparison of the values of $\log(\alpha)$ obtained with the maximum likelihood method with the logarithm of the sample mean of the posterior distribution of α obtained with the package bayesianETAS.

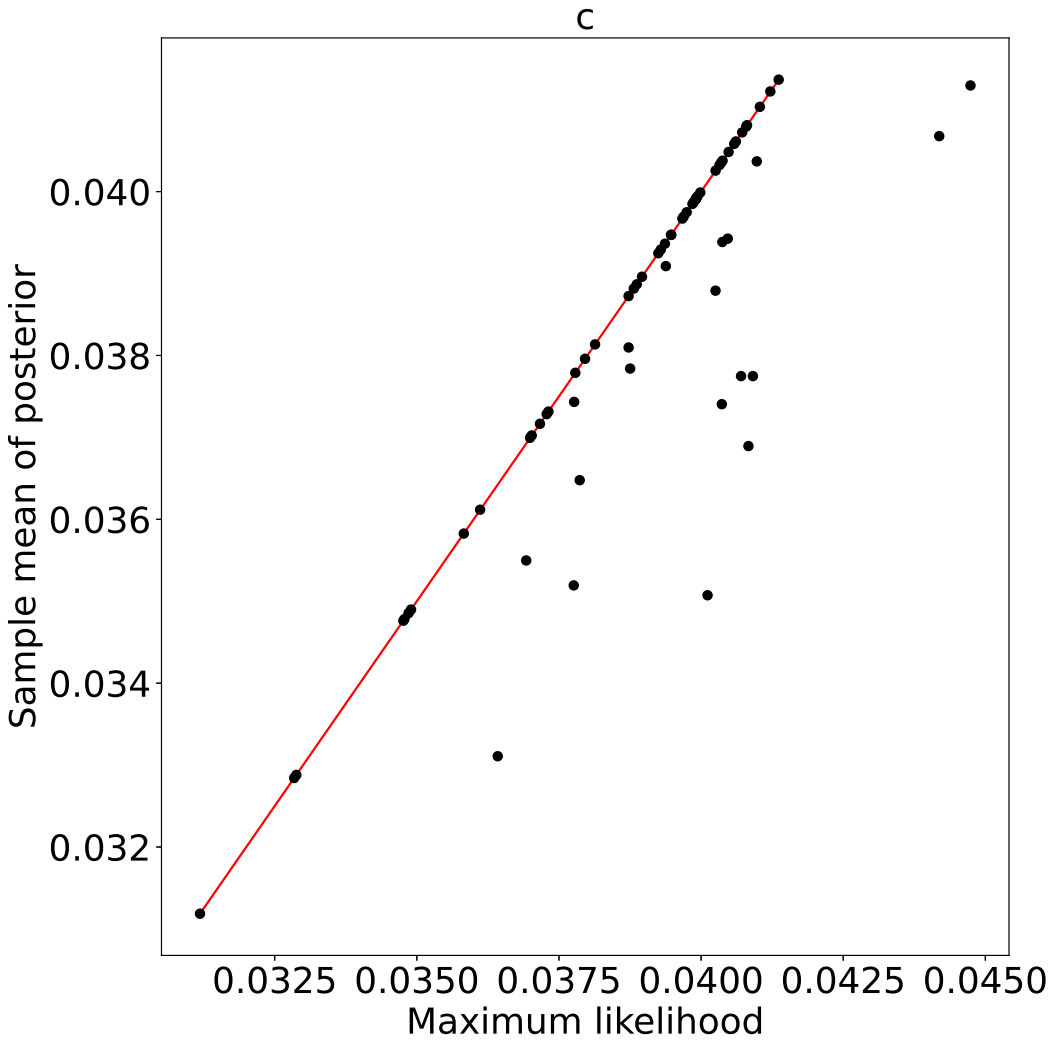

FIGURE 7.4: Comparison of the values of c obtained with the maximum likelihood method with the sample mean of the posterior distribution of μ obtained with the package bayesianETAS. The maximum likelihood estimate of c for family 2010.8.16.15.12.2 is not plotted here because it is much larger than the other values.

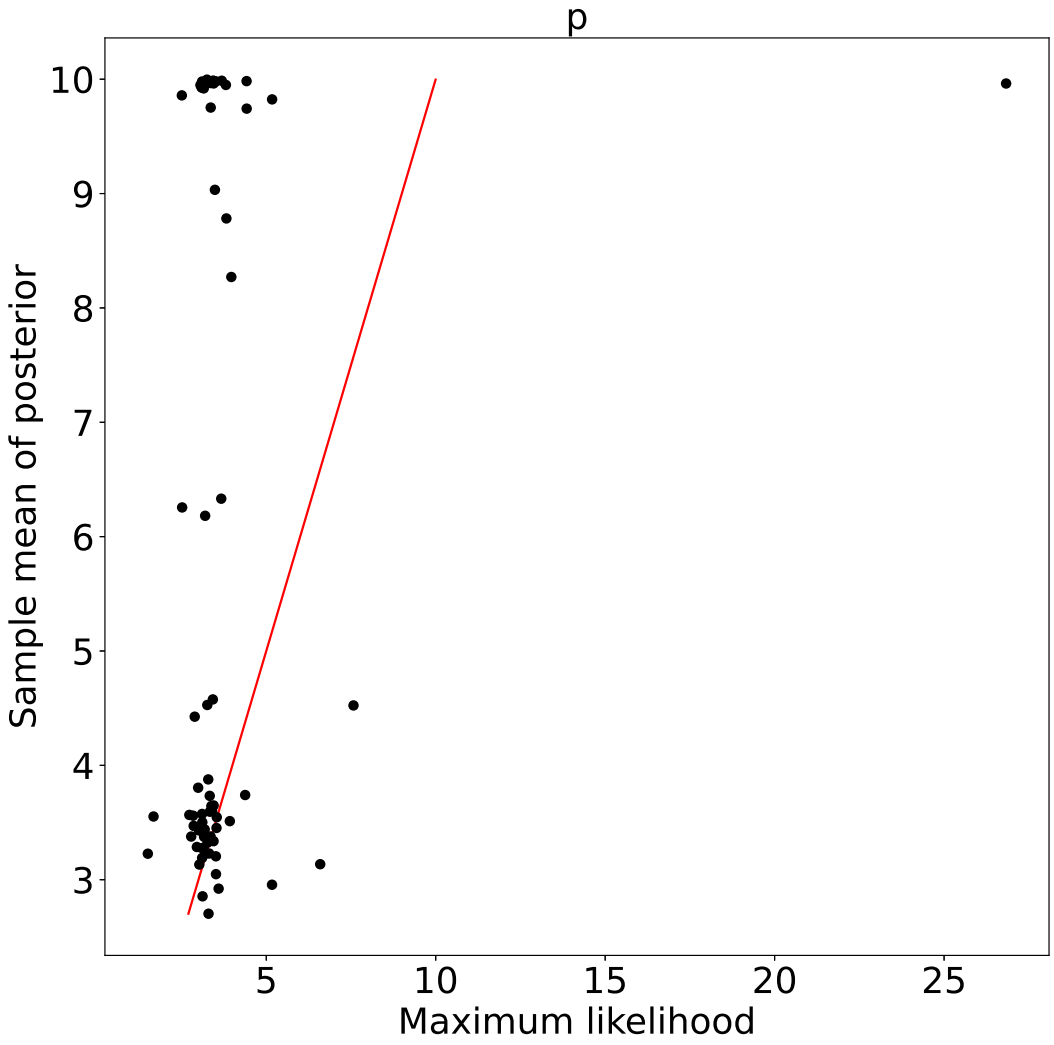

FIGURE 7.5: Comparison of the values of p obtained with the maximum likelihood method with the sample mean of the posterior distribution of μ obtained with the package bayesianETAS. The maximum likelihood estimate of p for family 2010.8.9.6.10.55 is not plotted here because it is much larger than the other values.

Bibliography

Barani, S. et al. (2018). "Long-range dependence in earthquake-moment release and implications for earthquake occurrence probability". In: *Scientific Reports* 8.1, pp. 1–11.

Beran, J. (1994). *Statistics for Long-Memory Processes*. New York, NY, USA: Chapman and Hall.

Bostock, M.G. et al. (2015). "Magnitudes and moment-duration scaling of low-frequency earthquakes beneath southern Vancouver Island". In: *Journal of Geophysical Research: Solid Earth* 120, pp. 6329–6350.

Bunde, A. and S. Lennartz (2012). "Long-term correlations in Earth sciences". In: *Acta Geophysica* 60.3, pp. 562–588.

Chestler, S.R. and K.C. Creager (2017a). "Evidence for a scale-limited low-frequecny earthquake source process". In: *Journal of Geophysical Research. Solid Earth* 122, pp. 3099–3114.

— (2017b). "A model for low-frequency earthquake slip". In: *Geochemistry, Geophysics, Geosystems* 18, pp. 4690–4708.

Ducellier, A. and K.C. Creager (in revision). "An eight-year-long low-frequency earthquake catalog for Southern Cascadia". In: *under revision for Journal of Geophysical Research: Solid Earth*.

Enescu, B., K. Ito, and Z.R. Struzik (2006). "Wavelet-based multiscale resolution analysis of real and simulated time-series of earthquakes". In: *Geophysical Journal International* 64, pp. 63–74.

Frank, W.B. et al. (2014). "Using systematically characterized low-frequency earthquakes as a fault probe in Guerrero, Mexico". In: *Journal of Geophysical Research Solid Earth* 119, pp. 7686–7700.

Frank, W.B. et al. (2016). "The evolving interaction of low-frequency earthquakes during transient slip". In: *Science Advances* 2, e1501616.

Gkarlaouni, C. et al. (2017). "Hurst analysis of seismicity in Corinth rift and Mygdonia graben (Greece)". In: *Chaos, Solitons and Fractals* 96, pp. 30–42.

Gutenberg, B. and C.F. Richter (1944). "Frequency of earthquakes in California". In: *Bulletin of the Seismological Society of America* 34.4, pp. 185–188.

Harte, D. (2010). "PtProcess: An R package for modelling marked point processes indexed by time". In: *Journal of Statistical Software* 35.8, pp. 1–32.

Jiménez, A. et al. (2006). "Testing the persistence in earthquake catalogs: The Iberian Peninsula". In: *Europhysics Letters* 73.2, pp. 171–177.

Li, J., Y. Chen, and H. Mi (2002). "$1/f^{\beta}$ temporal fluctuation: Detecting scale-invariance properties of seismic activity in North China". In: *Chaos, Solitons and Fractals* 14, pp. 1487–1494.

Matcharashvili, T. et al. (2018). "Simple statistics for complex Earthquake time distributions". In: *Nonlinear Processes in Geophysics* 25, pp. 497–510.

Mukhopadhyay, B. and D. Sengupta (2018). "Seismic moment release data in earthquake catalogue: Application of Hurst statistics in delineating temporal clustering and seismic vulnerability". In: *Journal Geological Society of India* 91, pp. 15–24.

Ogata, Y. (1988). "Statistical models for earthquake occurrences and residual analysis for point processes". In: *Journal of the American Statistical Association* 83.401, pp. 9–27.

Ogata, Y. and K. Abe (1991). "Some statistical features of the long-term variation of the global and regional seismic activity". In: *International Statistical Review* 59.2, pp. 139–161.

Omori, F. (1894). "On the aftershocks of earthquakes". In: *Journal of the College of Science, Imperial University of Tokyo* 7, pp. 111–200.

Ross, G.J. (2016). "Bayesian estimation of the ETAS model for earthquake occurrences". Available at http://www.gordonjross.co.uk/bayesianetas.pdf.

Sarlis, N. et al. (2018). "Micro-scale, mid-scale, and macroscale in global seismicity identified by empirical mode decomposition and their multifractal characteristics". In: *Scientific Reports* 8, pp. 1–15.

Schoenberg, F. and B. Bolt (2000). "Short-term exciting, long-term correcting models for earthquake catalogs". In: *Bulletin of the Seismological Society of America* 90.4, pp. 849–858.

Shelly, D.R. (2017). "A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep San Andreas fault". In: *Journal of Geophysical Research. Solid Earth* 122, pp. 3739–3753.

Sweet, J.R., K.C. Creager, and H. Houston (2014). "A family of repeating low-frequency earthquakes at the downdip edge of tremor and slip". In: *Geochemistry, Geophysics, Geosystems* 15, pp. 3713–3721.

Sweet, J.R. et al. (2019). "Variations in Cascadia low-frequency earthquake behavior with downdip distance". In: *Geochemistry, Geophysics, Geosystems* 20, pp. 1202–1217.

Taqqu, M. and V. Teverovsky (1998). "On estimating the intensity of long-range dependence in finite and infinite variance time series". In: *A Practical Guide to Heavy Tails: Statistical Techniques and Application*. Ed. by R.J. Adler, R.E. Feldman, and M.S. Taqqu. Boston, MA, USA: Birkhäuser.

Telesca, L., V. Lapenna, and M. Macchiato (2004). "Mono- and multi-fractal investigation of scaling properties in temporal patterns of seismic sequences". In: *Chaos, Solitons and Fractals* 19, pp. 1–15.

Telesca, L. et al. (2000). "Self-similarity properties of seismicity in the Southern Aegean area". In: *Tectonophysics* 321, pp. 179–188.

Utsu, T (1957). "Magnitude of earthquakes and occurrence of their aftershocks". In: *Zisin (Journal of the Seismological Society of Japan)* 10, pp. 35–45.

Utsu, T., Y. Ogata, and R. Matsu'ura (1995). "The centenary of the Omori formula for a decay law of aftershock activity". In: *Journal of Physics of the Earth* 43, pp. 1–33.

Wang, J.-H. (2013). "Memory effect in $M \geq 6$ earthquakes of South-North Seismic Belt, Mainland China". In: *Journal of Seismology* 17, pp. 913–924.

Węglarczyk, S. and S. Lasocki (2009). "Studies of short and long memory in mining-induced seismic processes". In: *Acta Geophysica* 57.3, pp. 696–715.

Xu, Y. and P.W. Burton (2006). "Time varying seismicity in Greece: Hurst's analysis and Monte Carlo simulation applied to a new earthquake catalogue for Greece". In: *Tectonophysics* 423, pp. 125–136.

Zheng, X. and D. Vere-Jones (1991). "Application of stress release models to historical earthquakes from North China". In: *Pure and Applied Geophysics* 135.4, pp. 559–576.

Zhuang, J. et al. (2012). *Basic models of seismicity: temporal models*. Community Online Resource for Statistical Seismicity Analysis. Available at http://www.corssa.org.