Long-range dependence in low-frequency earthquake catalogs

Ariane Ducellier

University of Washington

SIAM Conference on Mathematics of Planet Earth
July 13th 2022

Definition (1)

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:

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

where $\rho_{X,\tau}$ is the autocorrelation sequence of X_t at time lag τ.

Definition (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:

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

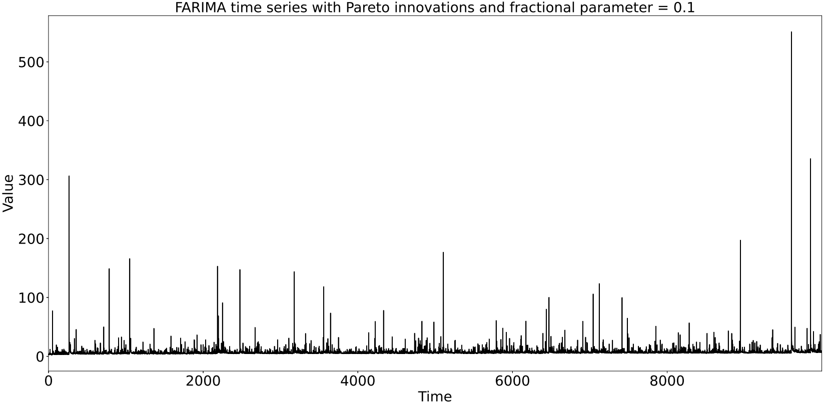
where $S_X(f)$ is the spectral density of X_t at frequency f.

Non summable autocovariance sequence

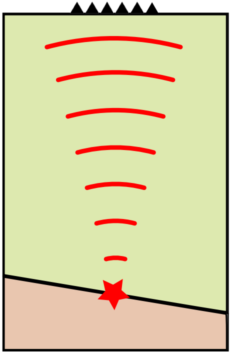
High-lag autocorrelations individually small / Cumulative effect cannot be neglected:

$$\sum_{\tau=-\infty}^{\tau=\infty} \rho_{X,\tau} = \infty$$

Example of synthetic time series: ARFIMA(0, 0.1, 0) with Pareto innovations (10000 points)

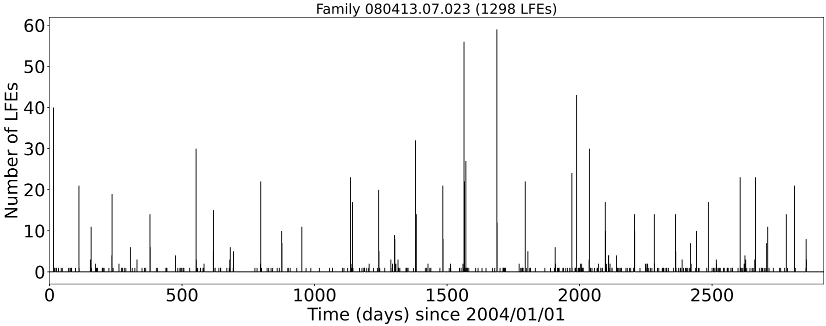


Low-frequency earthquakes (LFEs)



- Small magnitude earthquakes ($M \sim 0 - 2$).
- Reduced amplitudes at frequencies greater than 10 Hz.
- Earthquake source located close to the plate interface.
- Grouped into families: All LFEs from a given family originate from the same small patch.
- Dozens of LFEs within a few hours or days, followed by weeks or months of quiet.

Example of an LFE family



Statistical analysis of LFE occurrence

1. We look at each LFE family independently from the others.
2. We translate the sequence of LFE occurrence times into a discrete time series defined by the number of LFEs per unit of time.

Objective of the study

Look for evidence of long-range dependence by estimating the value of the Hurst parameter H or the fractional differencing parameter d:

$$H = d + \frac{1}{2} \text{ for finite variance processes}$$

$$H = d + \frac{1}{\lambda} \text{ for infinite variance processes}$$

If the time series is a FARIMA process with Pareto innovations, λ is the shape of the Pareto distribution.

Absolute value method

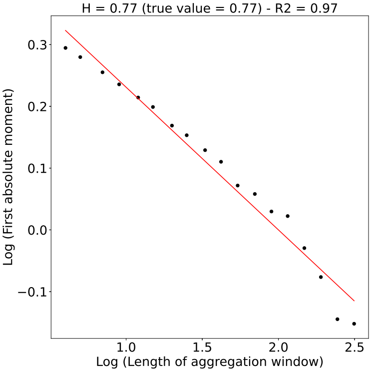
Aggregated time series:

$$X^{(m)}(k) = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i$$

First absolute moment :

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

$AM^{(m)}$ behaves like m^{H-1} for large m

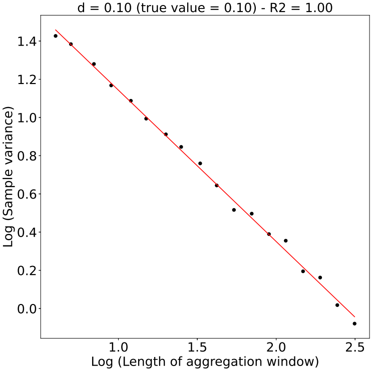


Variance

Sample variance:

$$\hat{Var} X^{(m)} = \frac{m}{N} \sum_{k=1}^{\frac{N}{m}} \left(X^{(m)}(k) - \bar{X} \right)^2$$

$\hat{Var} X^{(m)}$ behaves like m^{2d-1} for large $\frac{N}{m}$ and m



Variance of residuals

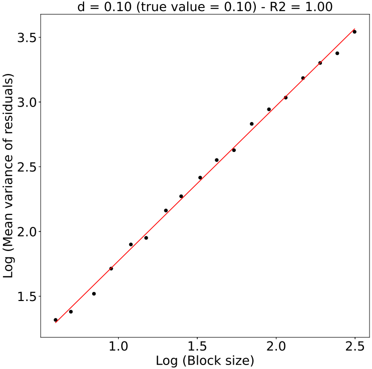
Blocks of size $m \rightarrow$ Compute the partial sums $Y(t) = \sum_{i=1}^{t} X_i$

Fit linear model and compute residuals:

$$\frac{1}{m}\sum_{t=1}^{m}(Y(t) - a - bt)^2$$

Compute variance of residuals.

Mean of variance behaves like m^{2d+1}



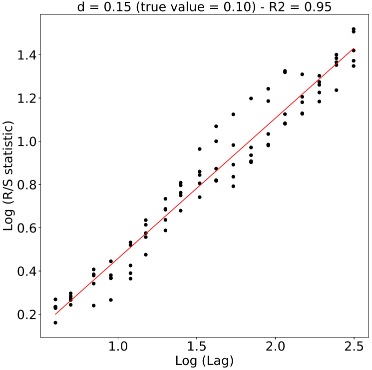
R/S statistic

Divide into K blocks $\rightarrow$ Compute the partial sums $Y_j(t) = \sum_{i=k_j+1}^{k_j+t} X_i$ with $k_j = j\frac{N}{K}$ and $j = 0, \cdots, K-1$

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

$S(j,n)$ = square root of the sample variance of $X(t)$, $t = k_j + 1, \cdots, k_j + n$

$R/S = \frac{R(j,n)}{S(j,n)}$ behaves like $n^{d+\frac{1}{2}}$.

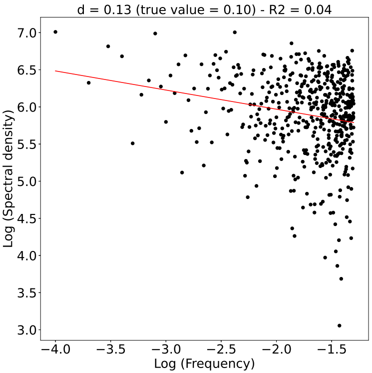


Periodogram

Periodogram:

$$I(f) = \frac{1}{2\pi N} \left| \sum_{j=1}^{N} X_j e^{ijf} \right|^2 \qquad (1)$$

$I(f)$ behaves like $|f|^{-2d}$ for small f



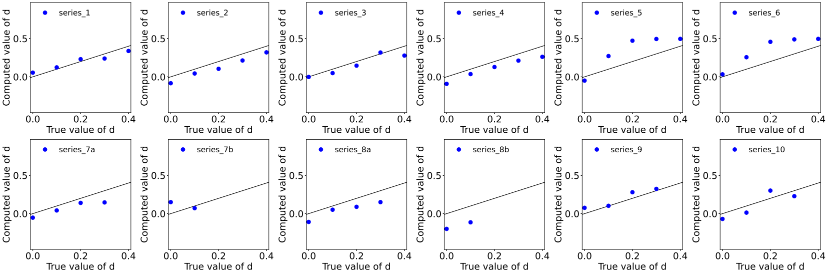
Synthetic times series I

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 $\sigma = 1$.

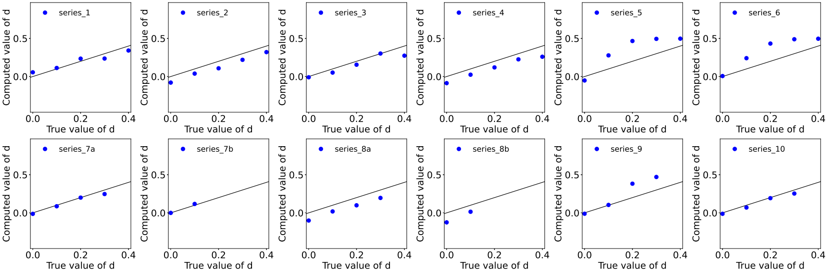
Synthetic times series II

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 $\lambda = 1.5$.

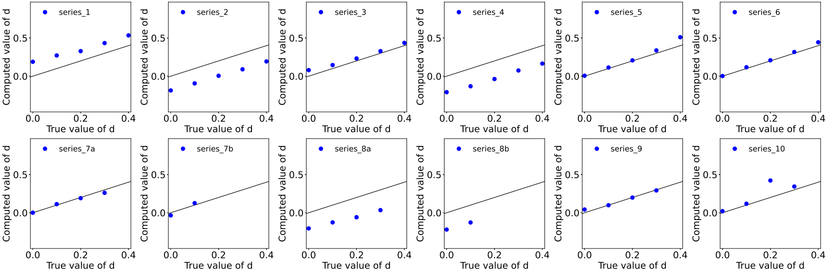
Absolute value method



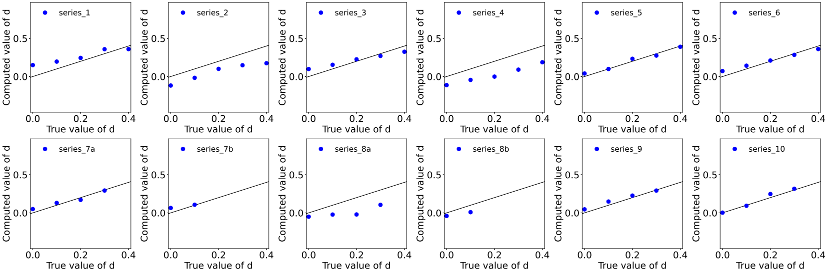
Variance method



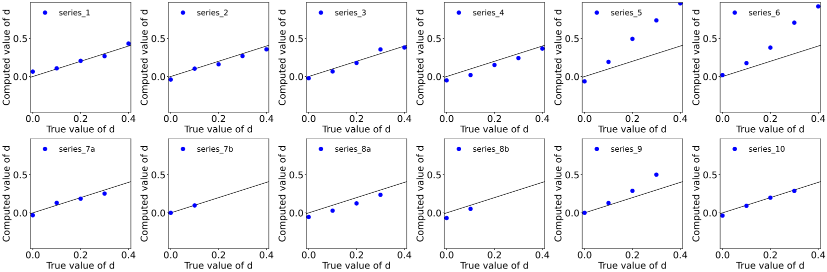
Variance of residuals method



R/S method


series_1
series_2
series_3
series_4
series_5
series_6
series_7a
series_7b
series_8a
series_8b
series_9
series_10
Computed value of d
True value of d
0.0
0.5
0.2
0.4

Periodogram method

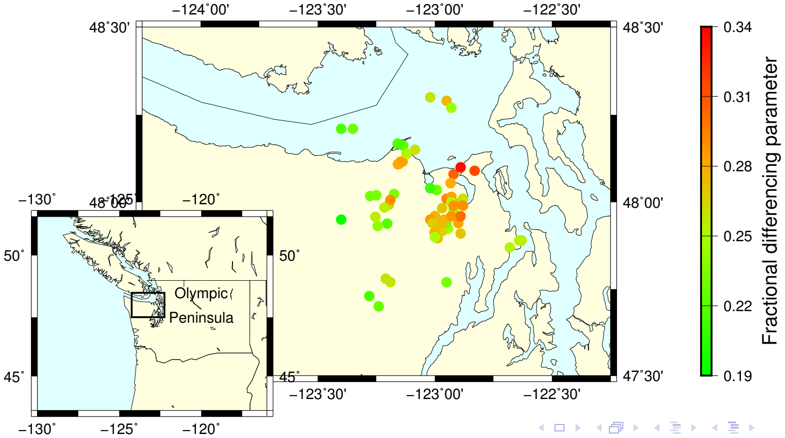


Choice of the method

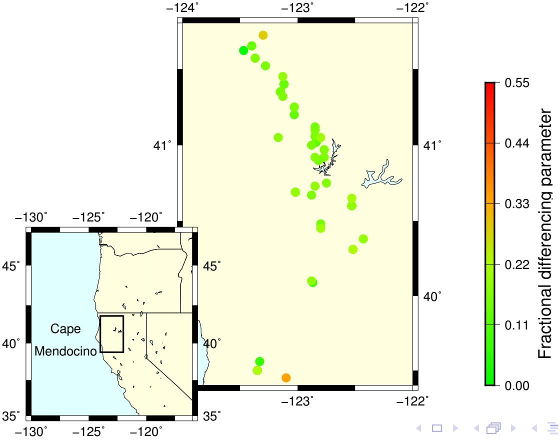
LFE time series look more similar to an ARFIMA process with Pareto innovations (synthetic time series 9).

For this synthetic time series, the variance of residuals method gives the best results.

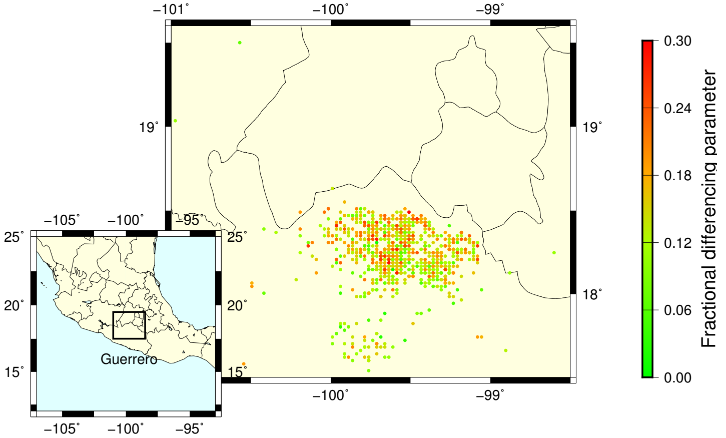
LFE catalog from the Olympic Peninsula (Chestler and Creager, 2017)



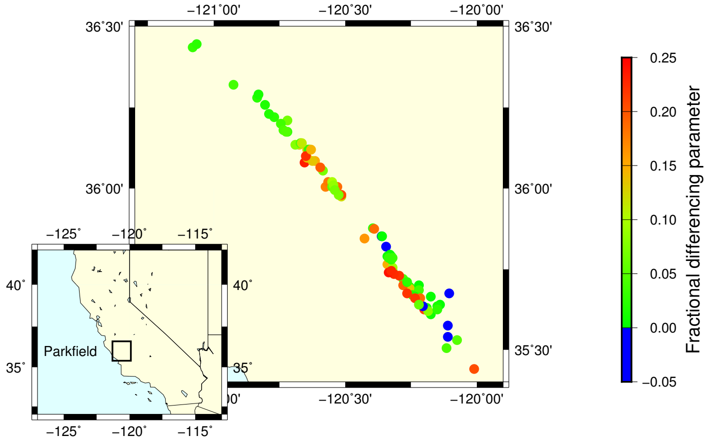
LFE catalogs from southern Cascadia (Ducellier and Creager, 2022)



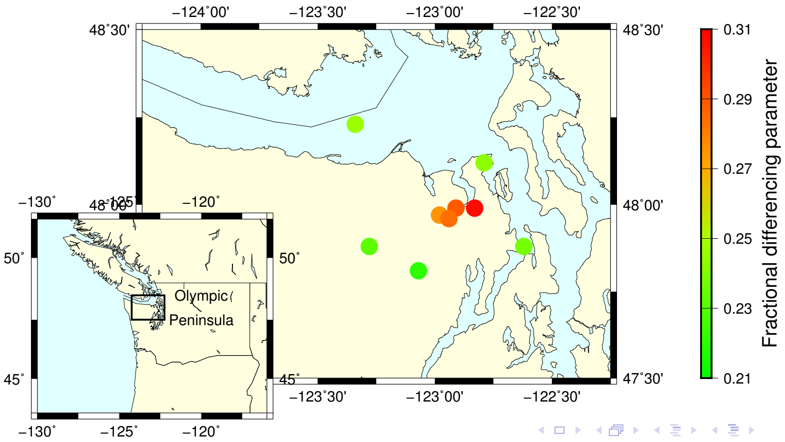
LFE catalogs from Mexico (Frank, 2014)



LFE catalogs from the San Andreas Fault (Shelly, 2017)



LFE catalogs from the Olympic Peninsula (Sweet *et al.*, 2019)



Future work

- Modeling the sequence of LFE occurrence times.
- Epidemic Type Aftershock Sequence (ETAS) model unsuccessful:
  - Model does not fit well the time sequence.
  - Model cannot reproduce the long-range dependence.
- Future work: Use more complex models based on neural networks.

Thank you

Thank you to the authors of the LFE catalogs: Shelley Chestler, William Frank, Alexandre Plourde, David Shelly, Justin Sweet

Questions?