Project B2 – Task 7

(Agreement DPC – INGV Att. B2, 2019-21)

SUBTASK 7.2 – STATISTICAL ANALYSIS

Previous studies and analysis techniques

Andrea Bevilacqua (PI), Augusto Neri (PI), Franco Flandoli (SNS),
Raffaele Azzaro (OE), Stefano Branca (OE)

ISTITUTO NAZIONALE
DI GEOFISICA E VULCANOLOGIA

Previous studies

Study	Method and conclusion	Dataset
Sharp, Lombardo, Davis (1981) Earthquakes in the time interval 1600 - 1978 with $I_0 \geq$ V (620 events - 146 primary) Eruptions in the time interval 1600 - 1978 (132 events – of which 49 flank)	Statistical test of independence between Poisson processes, based on *Cox (1955)*, generalized to a case with rate changes. Conclusion – (i) Poissonian distribution of flank eruptions and primary earthquakes. (ii) Abnormal number of flank eruptions after summit eruptions and after primary earthquakes.	DATASET 479 YRS ~ 10^2 vs 10^2 events
Nercessian, Hirn, Sapin (1991) Modification of the empirical method of aftershock removal (620 events – of which 180 primary)	Test of *Cox (1955)* assuming the eruptions as precursors of the earthquakes. Conclusion – Abnormal number of earthquakes after the onset of flank eruptions and after the end of flank eruptions.	
Gasperini, Gresta, Mulargia (1990) Earthquakes in the time interval 1978 - 1987 magnitude > 2.8 (1458 events) Eruptions (18 events - 9 flank)	Earthquake clusters recognition and modeling. Conclusion – correlation not calculated because 'insufficient data', and not qualitatively apparent to the authors.	DATASET 10 -17 YRS ~10^3 vs 10^1 ~10^1 vs 10^1
Mulargia (1992) Seismic sequences in 1974-1991 (12 events) Flank eruptions (11 events)	Statistical test of Poisson independence from *Brillinger (1976)*. Conclusion – flank eruptions as precursors of seismic sequences and not the contrary.	
Gresta, Marzocchi, Mulargia (1994) Earthquakes in the time interval 1600 - 1989. with $I_0 \geq$ IX (7 events) Eruptions with volume $\geq 10^7$ m^3 (40 events)	Correlation test according to the Spearman ranking coefficient. Conclusion – correlation between the end of major eruptions and the major earthquakes, and not with the eruption onsets.	DATASET 490 YRS ~ 10^1 vs 10^2

Overview of historical eruptions

Preliminary dataset from _Sharp et al. (1981)_.

FLANK ERUPTIONS
Mulargia, Tinti, Boschi (1985)
Kolmogorov-Smirnov test confirms Poisson distribution.

Linear correlation between volume and duration.

RATE CHANGES (Poisson)
Mulargia, Gasperini, Tinti (1987)
Flank eruptions 1600-1981
[1 change determined in 1865]

Other rate changes are indicatively reported with a descriptive purpose.

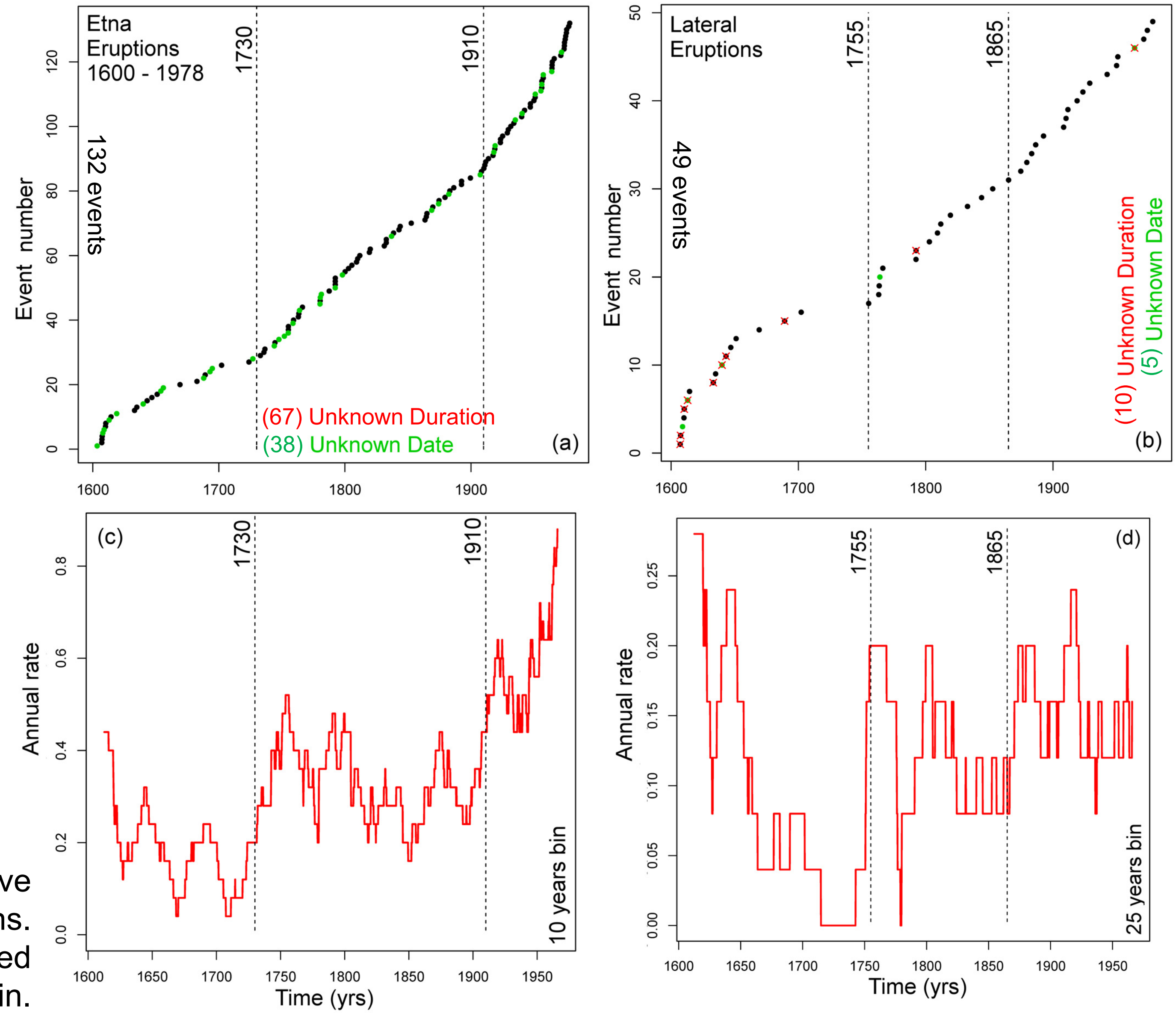


Figure. (a, b) cumulative number of eruptions. (c, d) annual rate calculated over 10 and 25 years bin.

Overview of historical earthquakes

Preliminary dataset
from Nercessian et al. (1991).

Nonhomogeneous distribution:
(i) underrecording effects
(e.g. before 1875)
(ii) rate changes in the activity
(e.g. after 1935)

Empirical algorithm

Sharp et al., (1981)
$\forall t_j$ terremoto, t_i aftershock if:
$t_i - t_j < f[I_0(t_j)]$
f(V) = 38 days
f(VII) = 88 days
f(IX) = 207 days
+
Nercessian et al. (1991)
deleted t_i re-inserted as primary if:
$I_0(t_i) \geq I_0(t_j)$ and
$I_0(t_i) = \max\{I_0(t_k) : t_i - t_j < f[I_0(t_j)]\}$

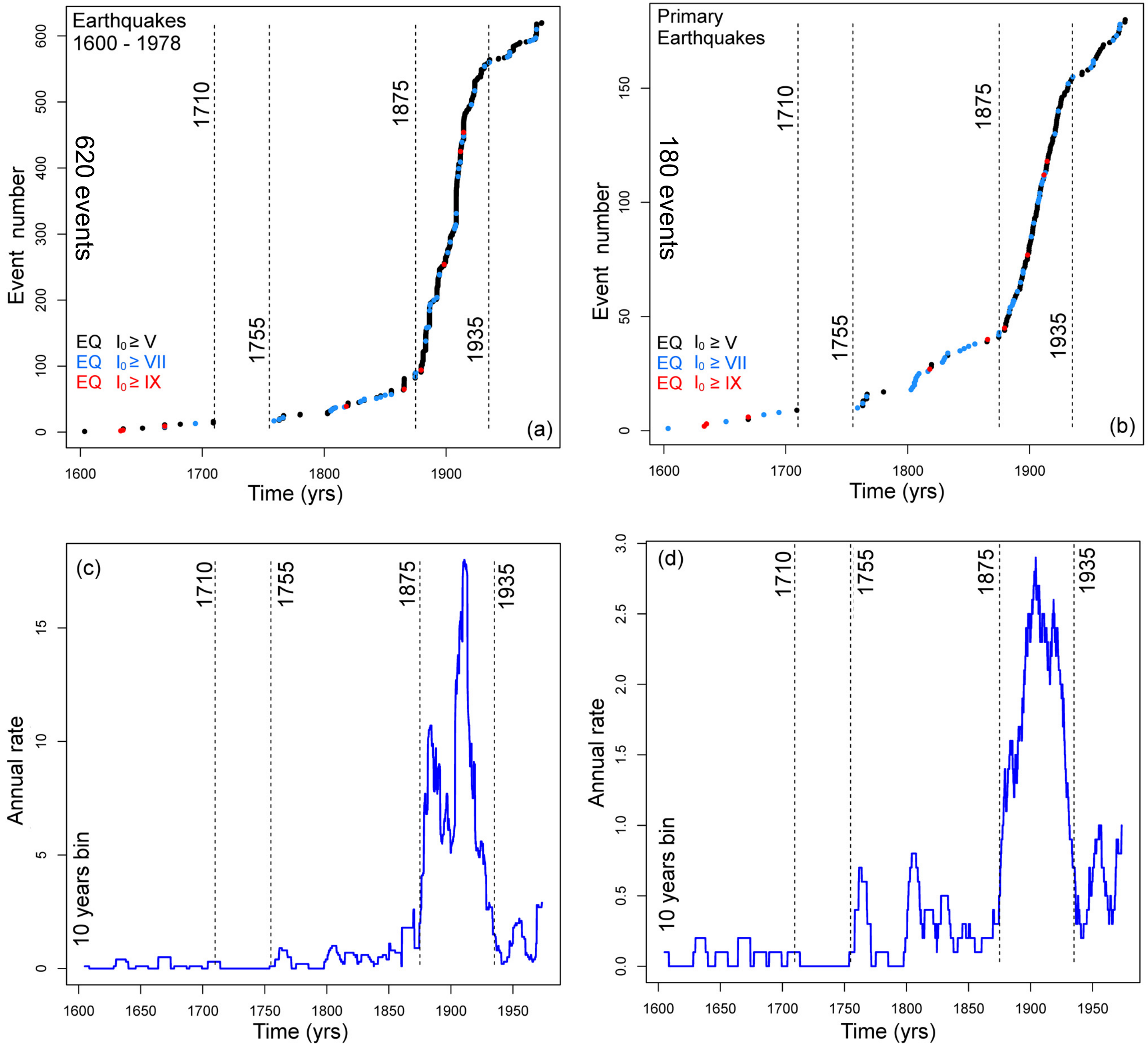


Figure. (a, b) cumulative number of earthquakes. (c, d) annual rate calulated over 10 years bin.

Epicentral Intensity threshold I_0

Underrecording indicatively reduced after 1800 AD if $I_0 \geq$ VII.

A relation of $I_0 \geq$ IX
with flank erupions
Is observed in 6/9 events.

EQ. MAGNITUDE = f (I_0)
<u>*Azzaro, D'Amico, Tuvè (2011)*</u>
I_0 & Magnitudo 1971-2010

I_0	M_d	M_L	M_W
V	3.0	2.6	3.2
VII	3.8	3.4	3.8
IX	> 4.0	> 4.3	> 4.7

Mean data.

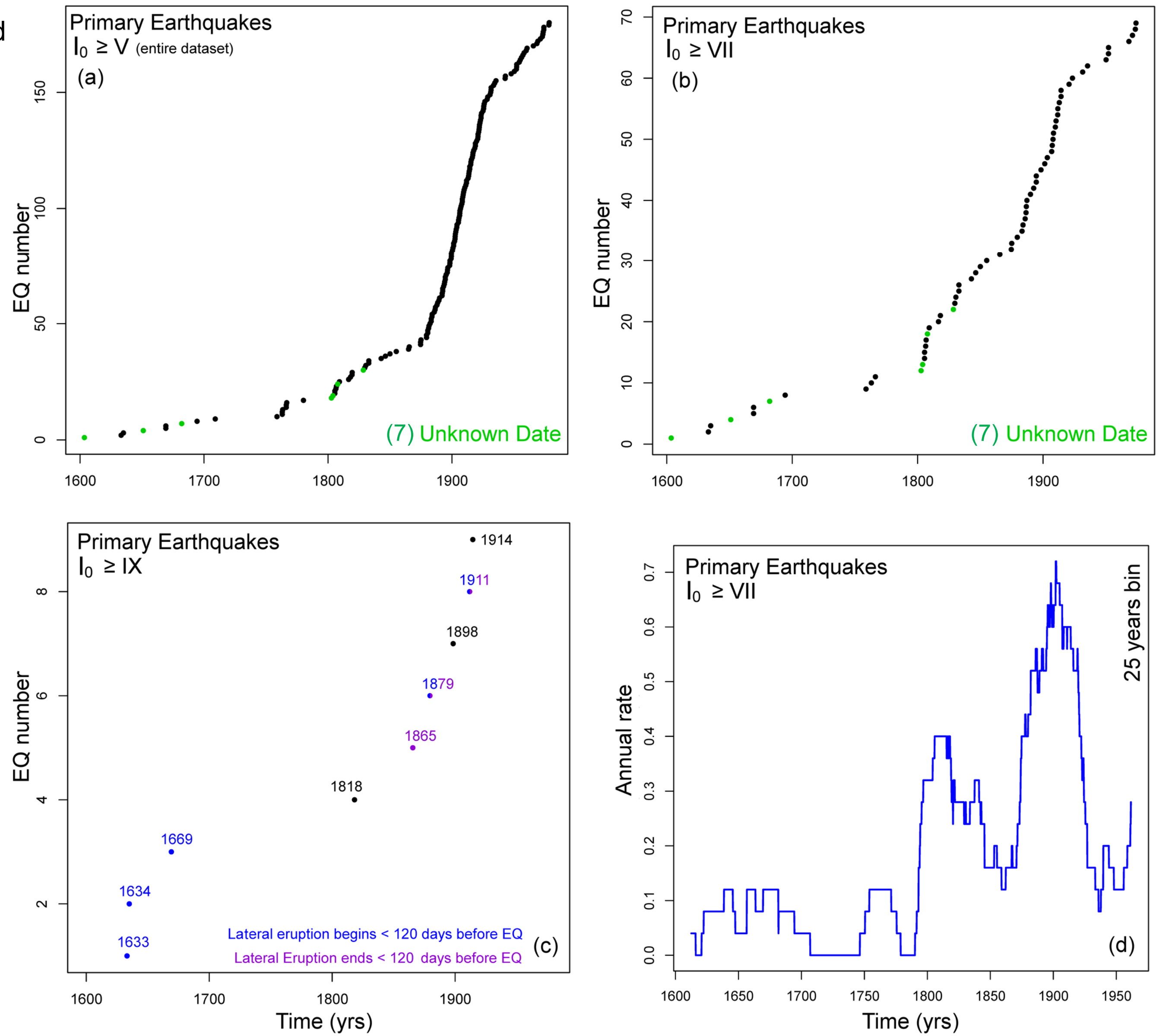


Figure. (a, b, c) cumulative number of earthquakes. (d) Annual rate $I_0 \geq$ VII calculated over 25 years bin.

Time difference histograms

Figure. Histograms based on sampling the difference between earthquakes times t_i and flank eruption times e_j.

We consider all the possible pairs with a difference lower than 120 days:
$C = \{t_i - e_j : |t_i - e_j| < 120 \text{ days}\}$

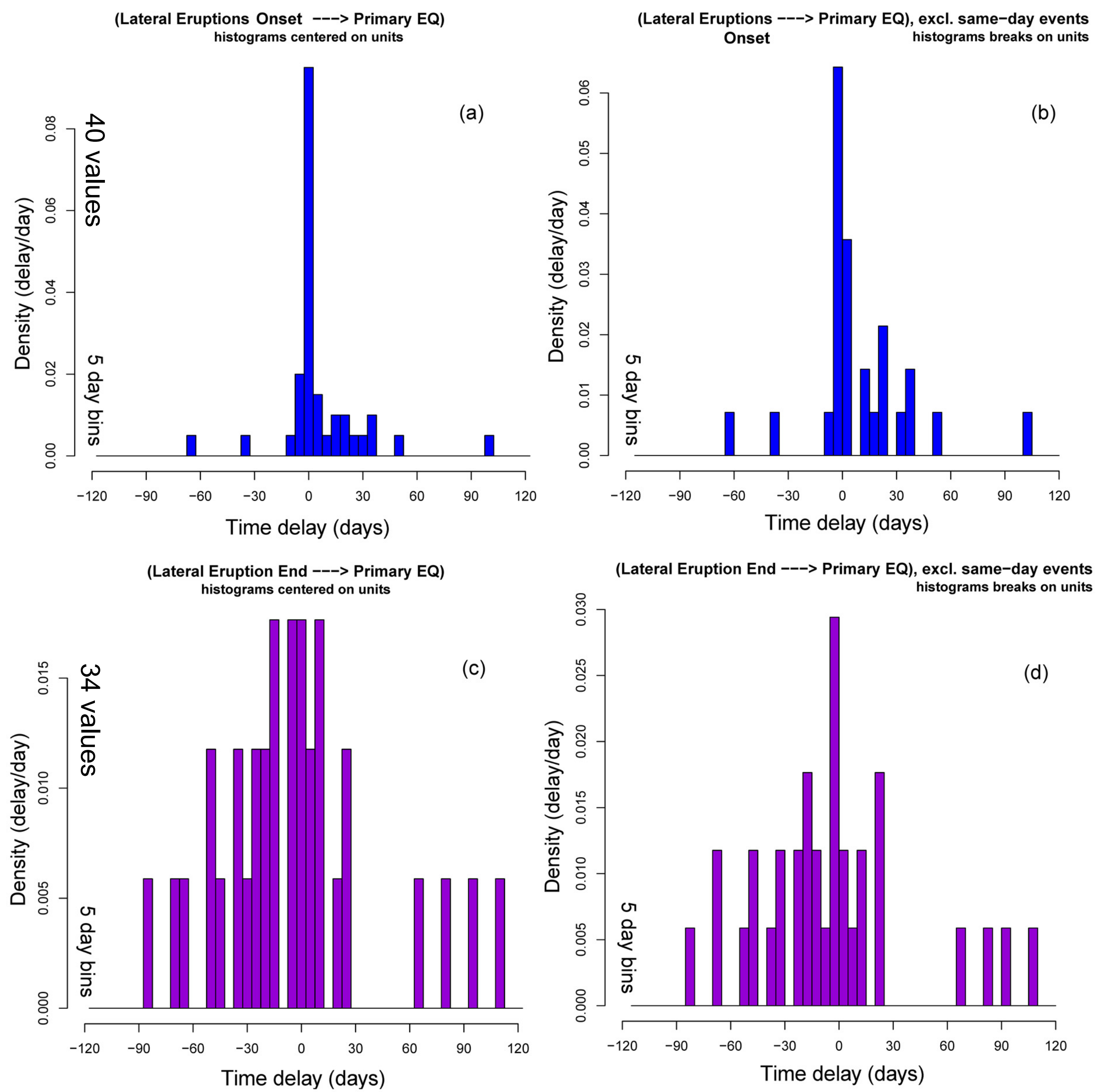
(a, b) – onset of flank eruptions
30% of EQ occur before the eruption
30% of EQ on the same day
40% of EQ after the eruption onset

If excluding differences < 5 days
21% of EQ occur before the eruption
79% of EQ after the eruption onset

Of those EQ after the eruption onset:
87.5% is within 45 days
12.5% is within 45 days to 120 days.

(c, d) – end of flank eruptions
65% of EQ occur before the eruption end
35% of EQ after the eruption end

Of those EQ after the eruption end:
67% is within 30 days
33% within 60 days and 120 days,
with a gap bewteen 30 days and 60 days.



The events occured in an unknown date are excluded.

Time difference histograms, $I_0 \geq$ VII

Task 7 – Etna

Figure. Histograms based on sampling the difference between earthquakes times t_i with $I_0 \geq$ VII and flank eruption times e_j.

$C = \{t_i - e_j : |t_i - e_j| < 120 \text{ days}\}$

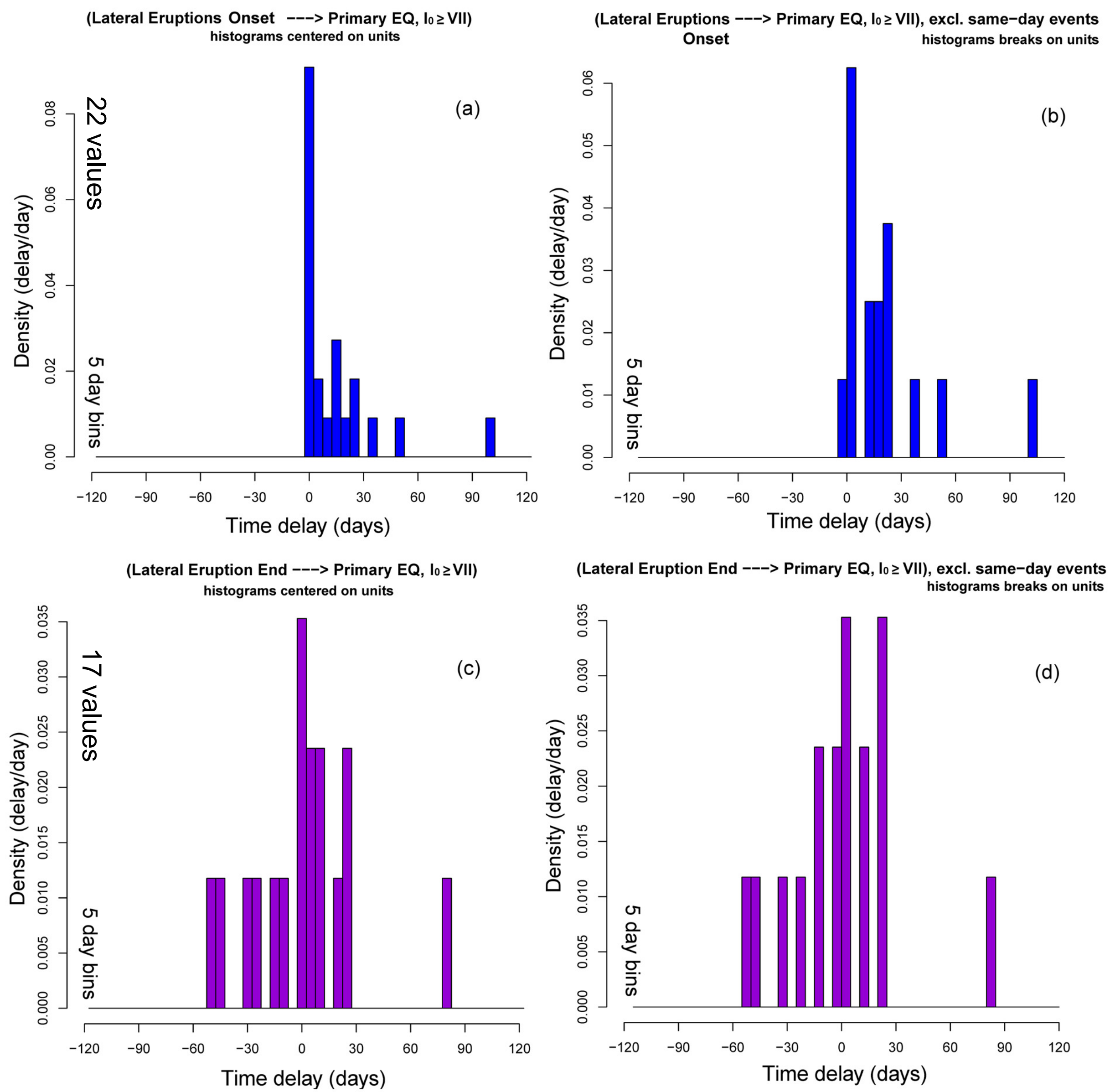
(a, b) - onset of flank eruptions
5% of EQ occur before the eruption
27% of EQ on the same day
68% of EQ after the eruption onset

If excluding differences < 5 days
100% of EQ after the eruption onset

Of those EQ after the eruption onset:
87 % is within 45 days
13% is within 45 days to 120 days.
Like in the previous dataset

(c, d) - end of flank eruptions
47% of EQ occur before the eruption end
53% of EQ after the eruption end

Of those EQ after the eruption end:
89% is within 30 days
11% tra 60gg e 120gg.
One single event



The events occured in an unknown date are excluded.

Time difference histograms, $I_0 \geq IX$

Figure. Histograms based on sampling the difference between earthquakes times t_i with $I_0 \geq IX$ and flank eruption times e_j.

$C = \{t_i - e_j : |t_i - e_j| < 120 \text{ days}\}$

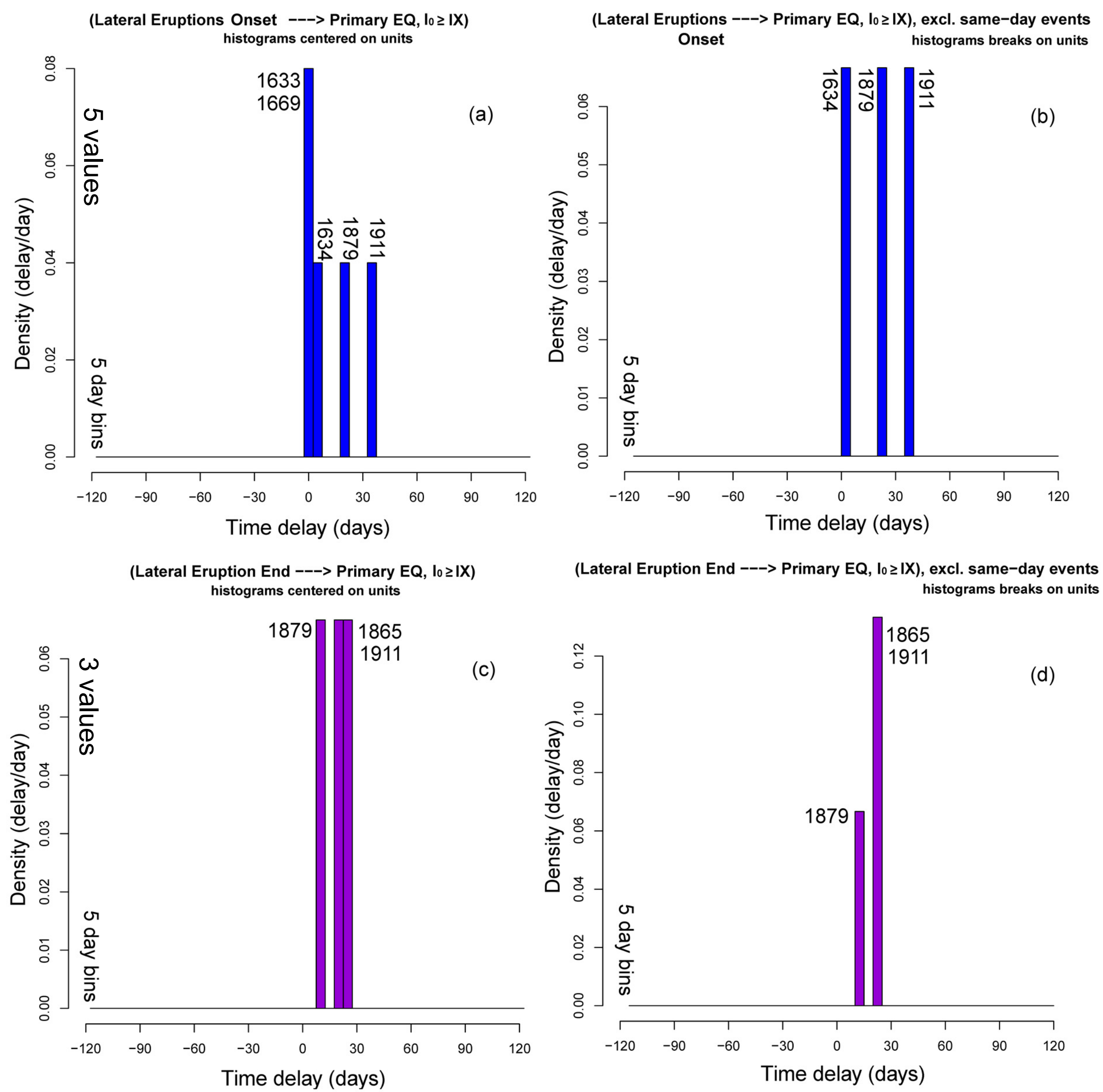
(a, b) - onset of flank eruptions
40% of EQ on the same day
60% of EQ after the eruption onset

If excluding differences < 5 days
100% of EQ after the eruption onset

Of those EQ after the eruption onset:
100% is within 45 days

(c, d) - end of flank eruptions
100% of EQ after the eruption end

Of those EQ after the eruption end:
100% is within 30 days



Poisson Processes Testing - I

Cox (1955)

Let N: $\mathscr{B}([0, +\infty]) \to \mathbb{N}$ be a Poisson process with intensity λ, and let
Ñ: $\mathscr{B}([0, +\infty]) \to \mathbb{N}$ be an arbitrary point process (simple) independent from N.
Let $(A_i)_{i=1\ldots,n,}$ and $(B_j)_{j=1\ldots,m}$ be their points in the interval [0, T].

$\forall\ (t_1, t_2)$, $N([t_1, t_2]) = |\, (A_i)_{i=1\ldots,n} \cap [t_1, t_2]\, |$, $\tilde{N}([t_1, t_2]) = |\, (B_i)_{i=1\ldots,n} \cap [t_1, t_2]\, |$

Classical notation:
$N(t) := N([0, t])$

For all *t* > 0, let ξ(t) be the random variabile defined as

$$\xi(t) = \sum_{j=1}^{m} N(\,(Bj,\, Bj + t]\,).$$

If:
(i) t << T, so that the boundary effects are negligible, and
(ii) $|\, \{ j : \tilde{N}(\,(Bj,\, Bj + t]\,) > 1 \}\, | << m$, that is, the effect of the overlapping of the intervals is negligible,

then:
ξ(t) is approximated by the sum of *m* independent Poisson variables, and so, by a new Poisson random variable. Moreover, we have $E[\xi(t)] = m\lambda t \approx \frac{nmt}{T}$.

The calculation is generalized to the case of N being a nonhomogeneous Poisson process, assuming λ costant over appropriate subintervals of [0, T]. In the following we adopt subintervals of 25 yrs duration.

As a practical rule that implies (i) and (ii), we specified that t < E [ΔB].

Poisson Processes Testing - II

We want to test if the events occurred at the times $(B_j)_{j=1\ldots,m}$ are precursors to those occurred in $(A_i)_{i=1\ldots,n}$. Under the assumptions (i) and (ii) of the previous slide.

H0 - null hypothesis

$\xi(t)$ is a Poisson random variable and $E[\xi(t)] = \frac{nmt}{T}$.

H1 - alternative hypothesis

$\xi(t)$ is not a Poisson random variable with $E[\xi(t)] = \frac{nmt}{T}$, and so the events in $(A_i)_{i=1\ldots,n}$ are not independent of those in $(B_j)_{j=1\ldots,n}$.

The test is performed by comparing $\xi(t)$ to the 5th and 95th percentiles of a Poisson random variable of intensity $\frac{nmt}{T}$.

The result of the test as a function of t is a step graph, marking the times t at which H0 is rejected with a level of confidence α = 90%.

If we apply the test on $(B_j)_{j=1\ldots,m} = (A_i)_{i=1\ldots,n-1}$, it verifies the lack of memory in the process (total randomness). In that case we use the notation $\zeta(t)$ to express the summative random variable.

H0 - null hypothesis

$\zeta(t)$ is a Poisson random variable and $E[\zeta(t)] = \frac{n(n-1)t}{T}$.

H1 - alternative hypothesis

$\zeta(t)$ is not a Poisson random variable with $E[\zeta(t)] = \frac{n(n-1)t}{T}$, and so the events in $(A_i)_{i=1\ldots,n}$ are not totally random.

The test is performed by comparing $\zeta(t)$ to the 5th and 95th percentiles of a Poisson random variable of intensity $\frac{n(n-1)t}{T}$.

Test of total randomness (Poisson) - EQ

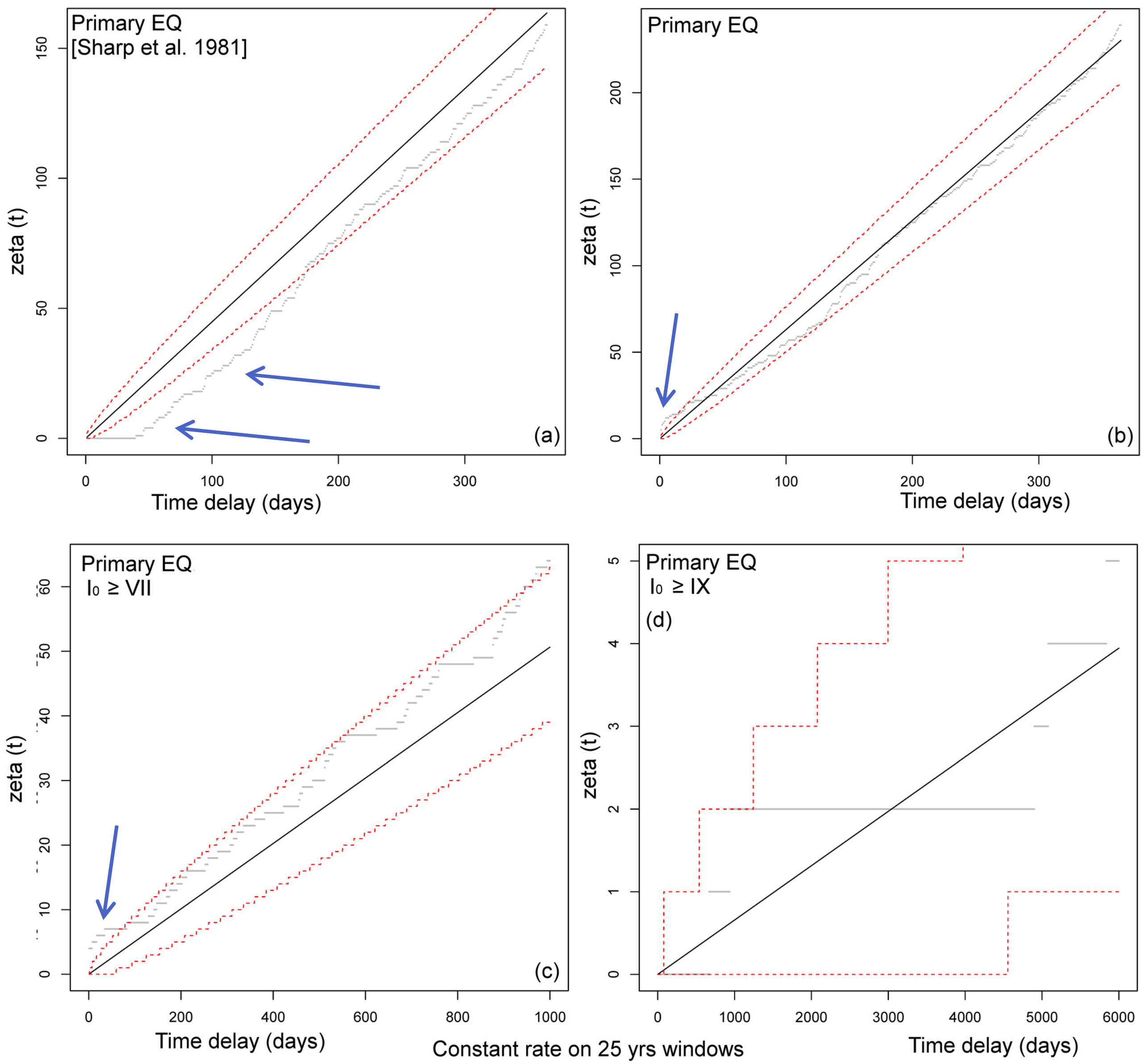
Figure. Step graphs related to the test of total randomness.

(a) primary EQ from *Sharp et al.* (1981). H0 rejected if t < 180 dd.

(b) primary EQ in this study. H0 rejected if t < 15 dd.

(c) primary EQ with $I_0 \geq$ VII. H0 rejected if t < 70 dd.

(d) primary EQ with $I_0 \geq$ IX. H0 not rejected.



Test of total randomness (Poisson) - Eruptions

Task 7 – Etna

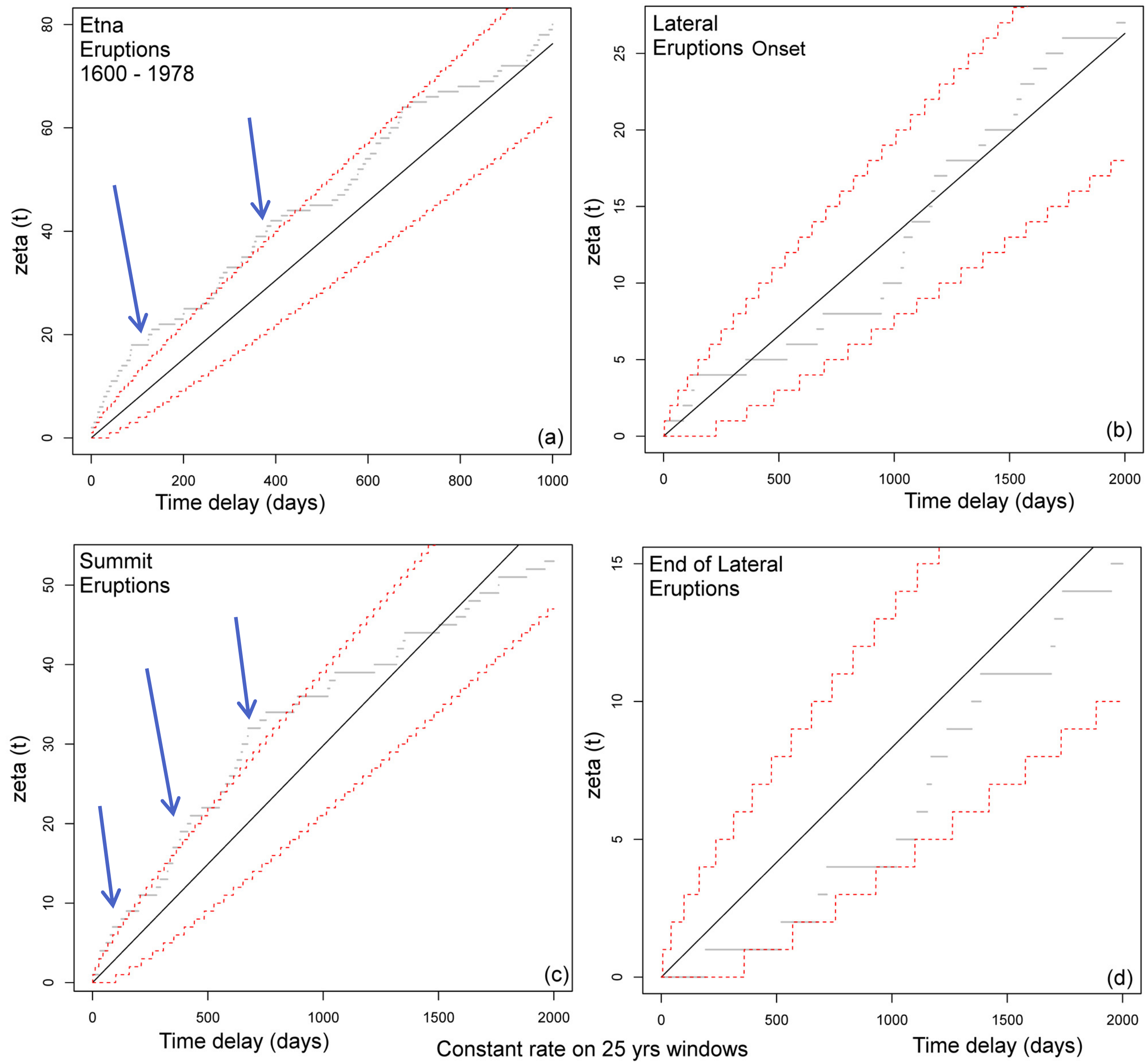
Figure. Step graphs related to the test of total randomness.

(a) eruptions 1600 - 1978.
H0 rejected if t < 450 dd.

(b) onset of flank eruptions.
H0 not rejected.

(c) summit eruptions.
H0 rejected if t < 900 dd.

(d) end of flank eruptions.
H0 not rejected.



Test of independence (Poisson) - I

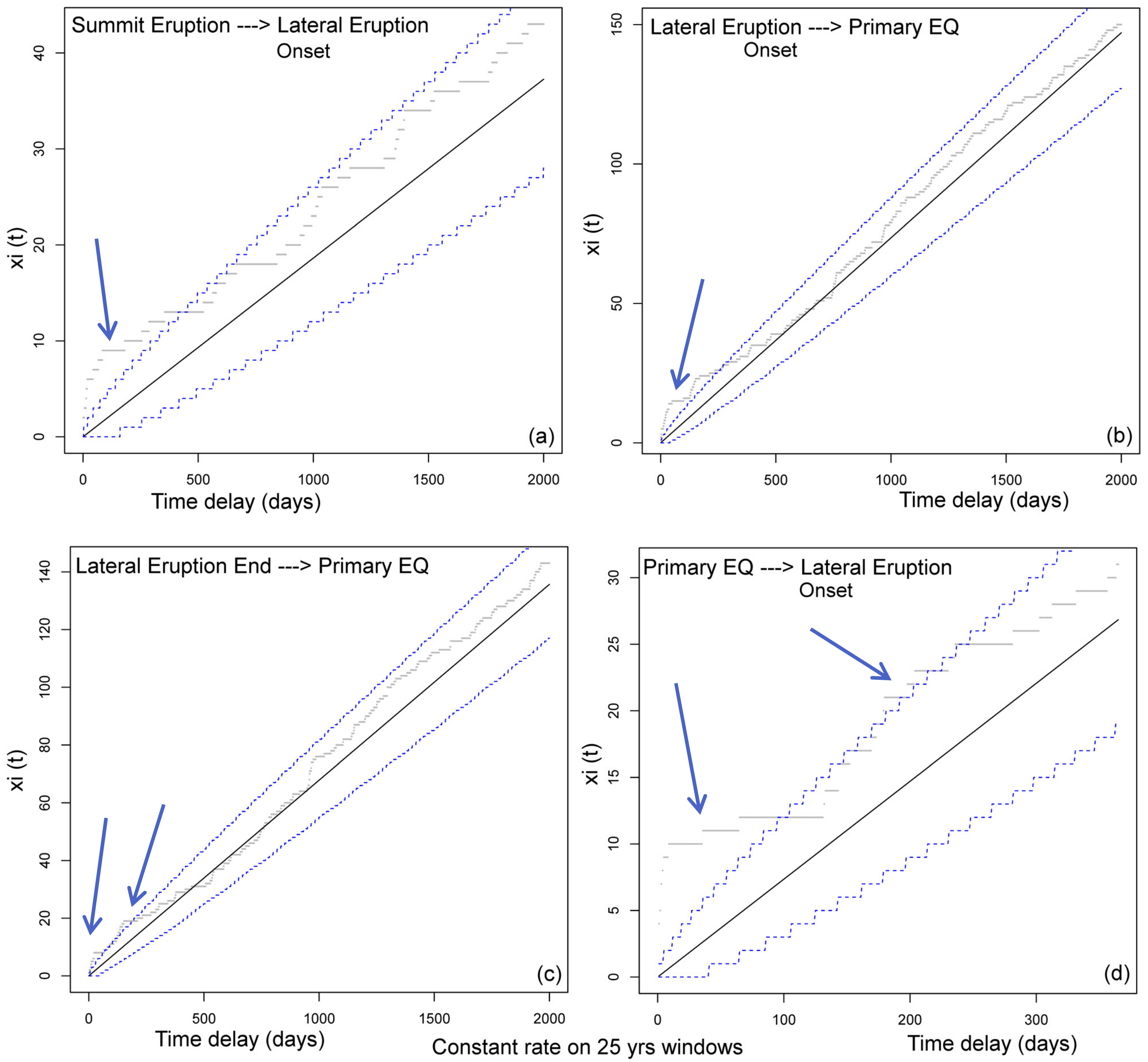
Figure. Step graphs related to the test of independence.

(a) summit eruption ---> flank eruption.
H1 accepted if t < 400 dd.

(b) flank eruption onset ---> primary EQ
H1 accepted* if t < 250 dd.

(c) flank eruption end ---> primary EQ
H1 accepted* if t < 200 dd.

(d) primary EQ ---> flank eruption onset
H1 accepted if t < 100 dd. and if t ≈ 200 dd.



*(b, c) not reliable if t < 15 dd.

Test of independence (Poisson) - II

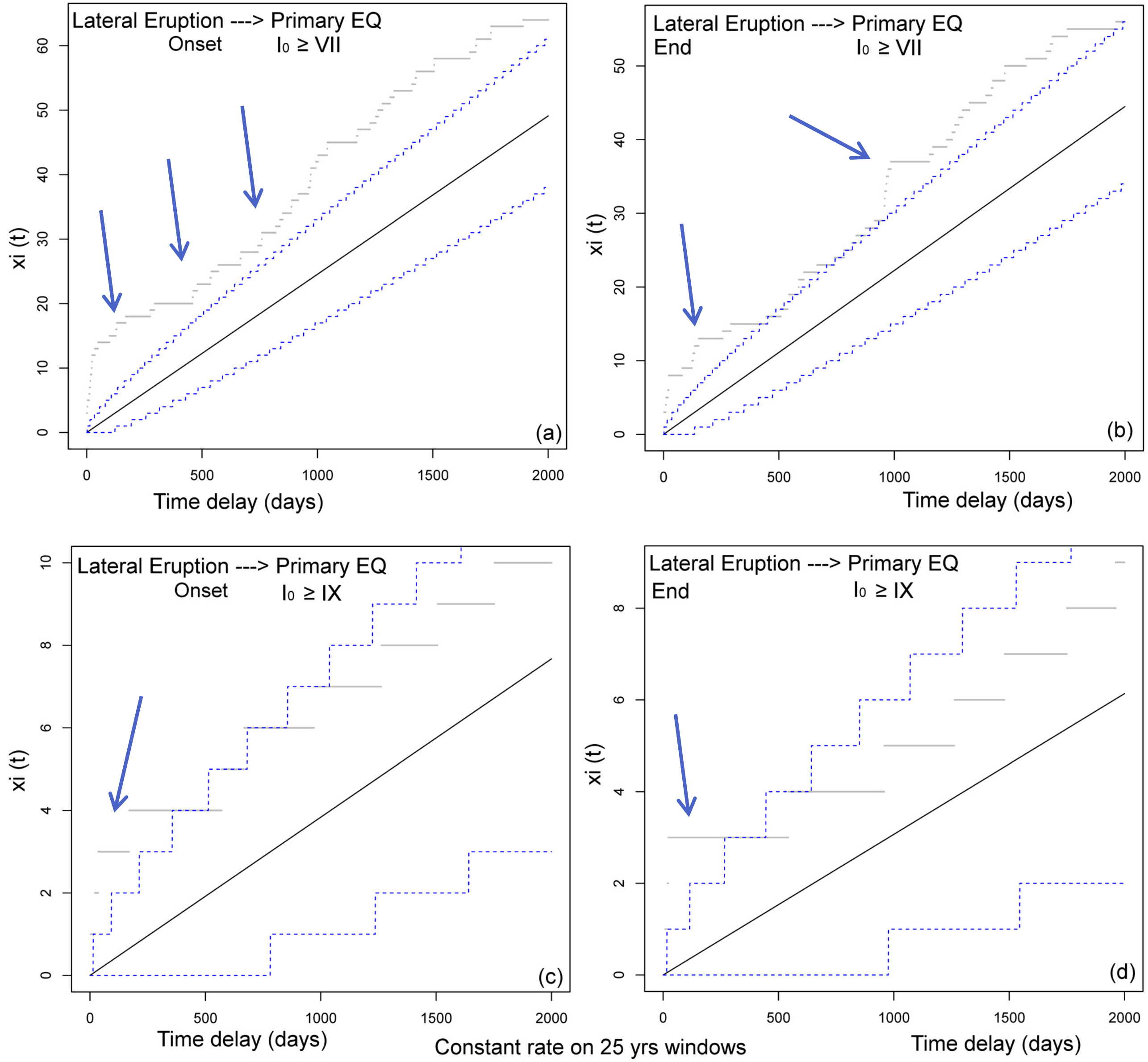
Figure. Step graphs related to the test of independence.

(a) flank eruption onset ---> primary EQ $I_0 \geq$ VII.
H1 accepted* for every t.

(b) flank eruption end ---> primary EQ $I_0 \geq$ VII.
H1 accepted* if ogni t.
Remarkable peak at t ≈ 900 dd.

(c) flank eruption onset ---> primary EQ $I_0 \geq$ IX.
H1 accepted if t < 400 dd.

(d) flank eruption end ---> primary EQ $I_0 \geq$ IX.
H1 accepted if t < 300 dd.



*(a, b) not reliable if t < 70 dd.

Test of independence (Poisson) - III

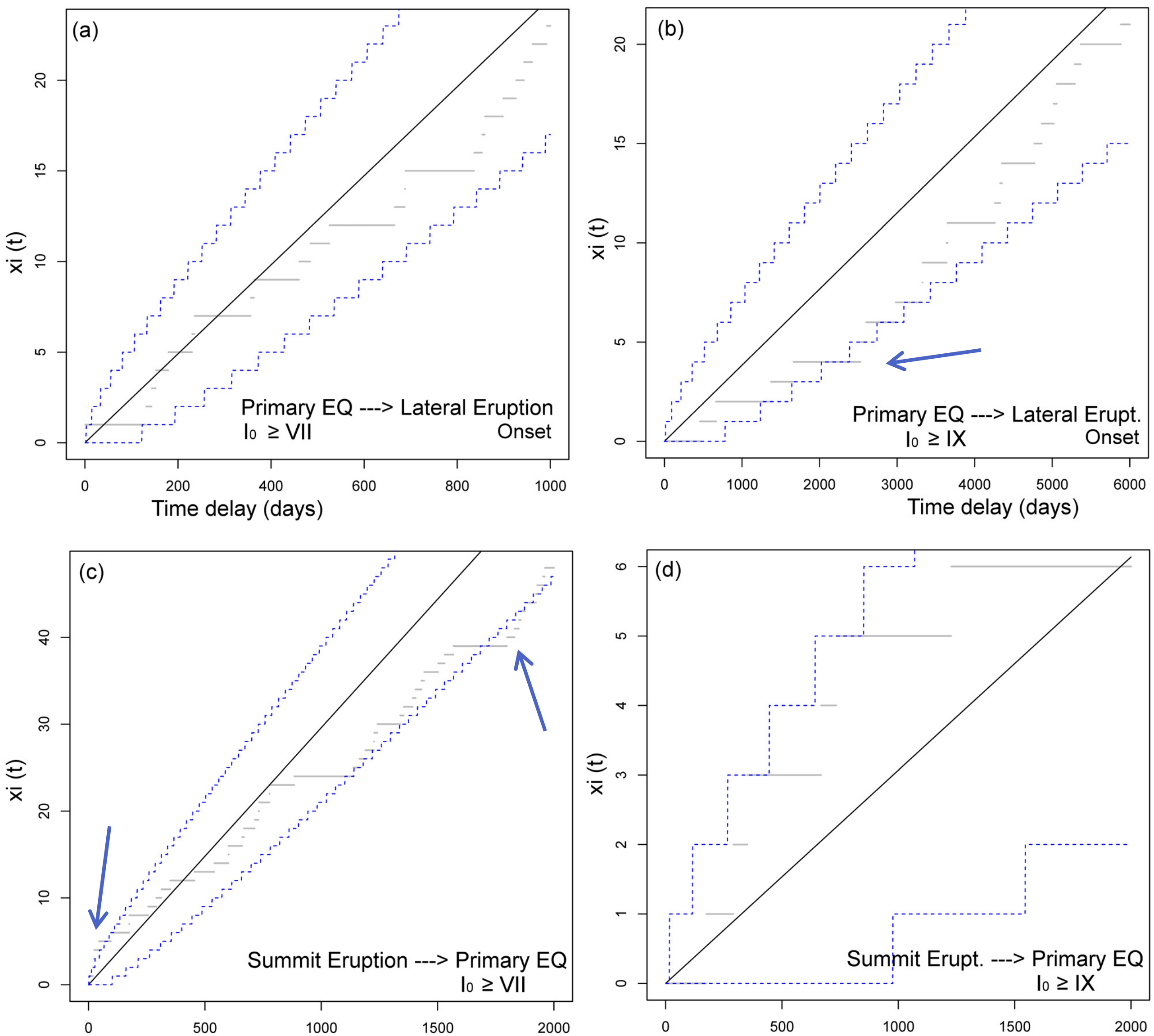
Figure. Step graphs related to the test of independence.

(a) primary EQ $I_0 \geq VII$ ---> flank eruption onset. H0 not rejected.

(b) primary EQ $I_0 \geq IX$ ---> flank eruption onset. H0 not rejected.

(c) summit eruption ---> primary EQ $I_0 \geq VII$. H1 accepted* if t < 70 dd.

(d) summit eruption ---> primary EQ $I_0 \geq IX$. H0 not rejected.



Constant rate on 25 yrs windows

*(c) not reliable if t < 70 dd.

Future research activities

- Update and comparison with extendend and more recent time datasets, with an improvement in UQ. E.g. *Branca, Del Carlo (2005)*; *Azzaro, D'Amico et al. (2015)* http://www.ct.ingv.it/macro/etna/html_index.php

- Additional analysis with the local Test of *Brillinger (1976)*, focusing on the duration in the correlation and on its quantitive estimate through the calculation of the product density of the two point processes.

- Exploration of the correlation structures in a context of time-dependent modeling (not Poissonian, but based on renewal models on the major faults
  E.g. *Azzaro et al. (2012); Azzaro et al. (2013); Azzaro et al. (2017)*
  models based on Brownian passage time - *Matthews, Ellsworth, Reasenberg (2002).*

- (?) Application of other methodologies to distinguish aftershock from primary events. Test modification to allow for residual clustering.

- (?) Study and comparison with methods developed in the modeling of precursors pattern, albeit those are classically applied to monitoring data, and not to the historical records.
  E.g. Mulargia, Gasperini ,Marzocchi (1991); Mulargia, Marzocchi, Gasperini (1992); Cardaci et al. (1993); Vinciguerra et al. (2001); Sandri, Marzocchi, Gasperini (2005)