NSF awards 1521855, 1821311, project FISR2017 MIUR, agreement DPC-INGV Annex B2 2019-21

Mapping and managing hazards using Precursory Data, and Analysis

A SDE Framework for Volcanic Precursors

Andrea Bevilacqua (INGV),
Abani Patra (UB), Eric Bruce Pitman (UB), Marcus Bursik (UB),
Augusto Neri (INGV), Barry Voight (PennState), Franco Flandoli (SNS),
Prospero de Martino (INGV), Flora Giudicepietro (INGV),
Giovanni Macedonio (INGV), Stefano Vitale (UNINA)

UNIVERSITY AT BUFFALO

ISTITUTO NAZIONALE
DI GEOFISICA E VULCANOLOGIA

NATIONAL SCIENCE
FOUNDATION

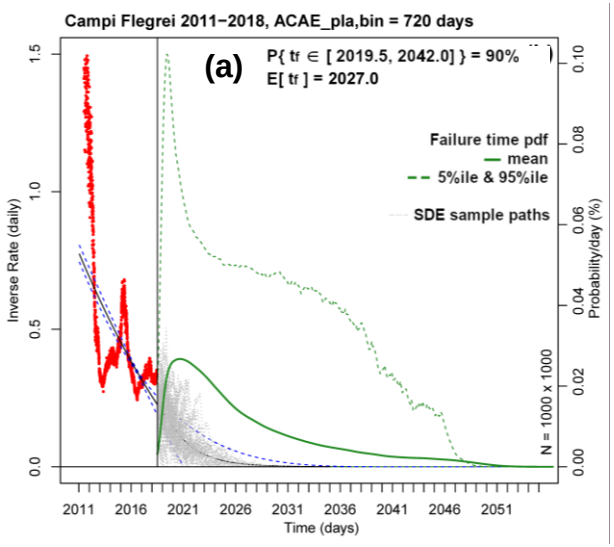
Overview of the talk

Let $(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+$ be the spatial location and onset time of a **target phenomenon**.

A **volcanic precursor** is a real function F defined on a domain $A \times [t_0, t_1] \subseteq \mathbb{R}^3 \times \mathbb{R}_+$, such that F has a local maximum in (x, t).

Our purpose is to obtain **probability forecasts** of (x, t) given the values of F.

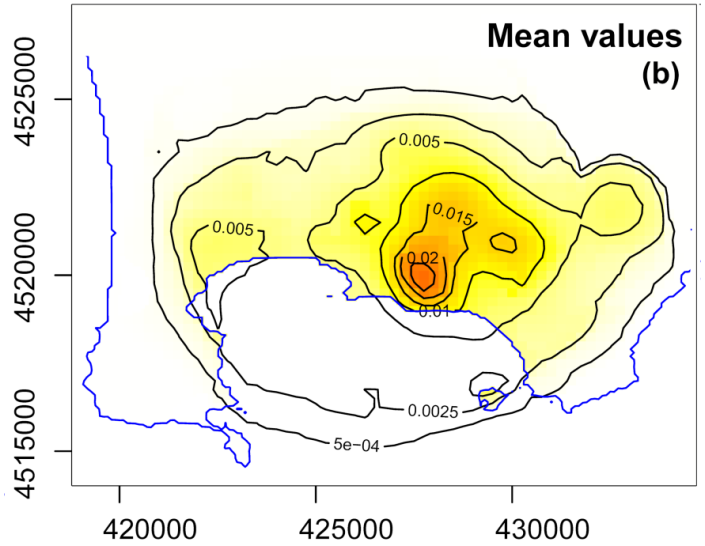
Depending on the additional hypotheses we can make on F, we describe two different cases:



1) Probabilistic enhancement of the Failure Forecast Method.

Let the problem be temporal. Then F: $[t_0, t_1] \to \mathbb{R}$ is determined by a **physical equation** which links the rate $\dot{X} = dF/dt$ and the rate-change $\dot{X}$. This provides estimates of the time t_f such that $F(t_f) = +\infty$.

Figure. Examples of (a) enhanced failure forecast method and (b) posterior spatial map.



2) Update of a prior spatial map through the statistical assimilation of precursor data.

Let the problem be spatial. Then the values of F: $A \to \mathbb{R}$ define a **likelihood map** which is combined with a prior probability distribution of (x,t) using the Bayes' Theorem.

PART 1A

Probabilistic enhancement of the Failure Forecast Method

[Modified from Bevilacqua et al., 2019]

Front. Earth. Sci., 7:135. doi: 10.3389/feart.2019.00135

The Failure Forecast Method (FFM)

The FFM is a well-established tool in the interpretation of monitoring data as possible **precursors**, providing quantitative predictions of a volcanic eruption onset (Voight, 1988).

The model represents the potential **cascade** of precursory signals leading to a significant rupture of materials, with t_f a proxy for the eruption onset t_e.

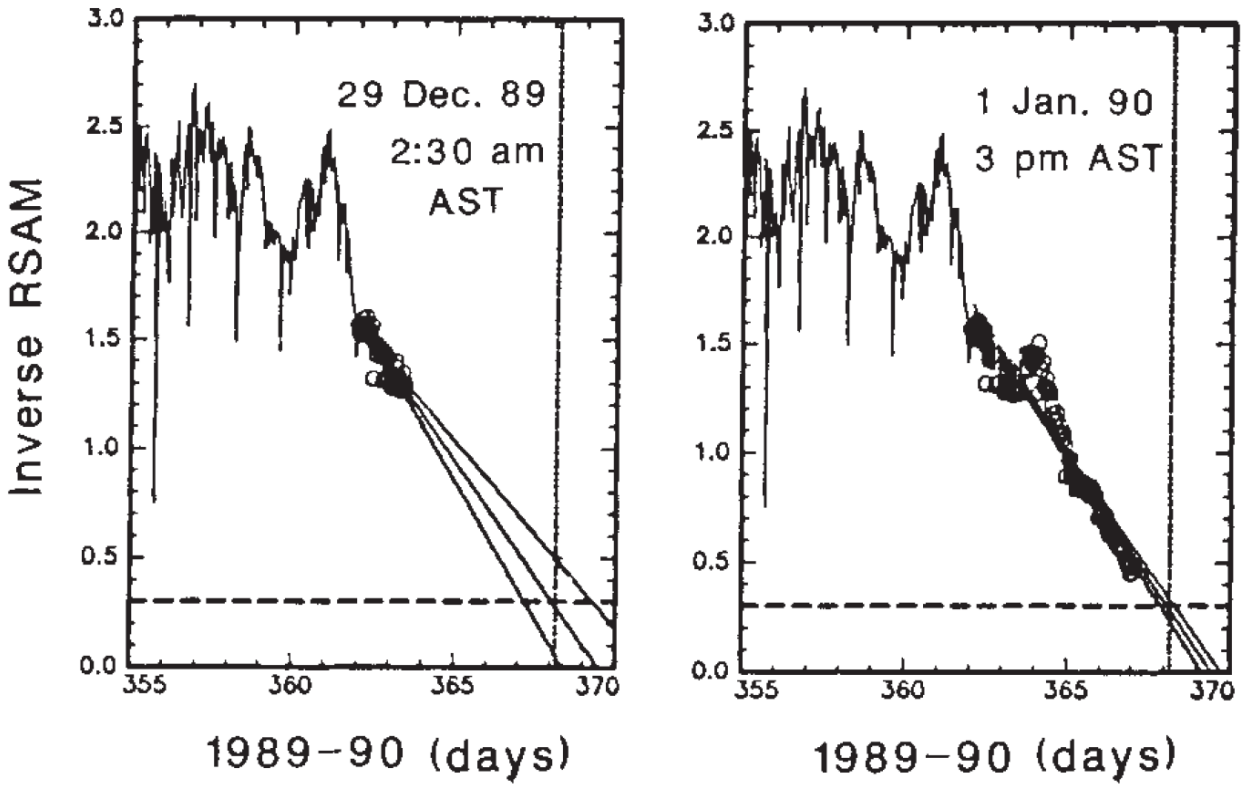


Figure. Examples of linear regression of the inverse rate of cascading seismic signals collected at Redoubt volcano (USA), before a major eruption in 1990 (from Voight & Cornelius, 1991)

The FFM has been retrospectively applied to several volcanic systems, including **explosive** eruptions. **Seismic** and **ground deformation** data are the type of signals most extensively studied with the method.

Uncertainty quantification of the FFM

FFM is known to be affected by sources of **uncertainty**, like:

- the occurrence of multiple phases of acceleration in the signals
- the superposition of signals originating from different causes
- heterogeneity in the breaking material, producing changes in the signals.

In addition, the statistical fitting of model parameters can be poorly constrained.

We perform a **full probability assessment** of FFM, with uncertainty quantification.

In particular, we enhance the classical FFM by:

- systematically characterizing the **uncertainty**, including both aleatoric sources (related to the future forecast) and epistemic sources (related to our current knowledge);

- incorporating a stochastic **noise** in the equations, and a mean-reversion property to constrain it (i.e. a Hull-white model).

We retrospectively test the enhanced FFM over two datasets from Voight, 1988.

These refer to: St. Helens, 1982, Bezymianny, 1960.

Our aim is to produce **probability forecasts** with the FFM, instead of deterministic predictions.

The FFM differential equation (ODE)

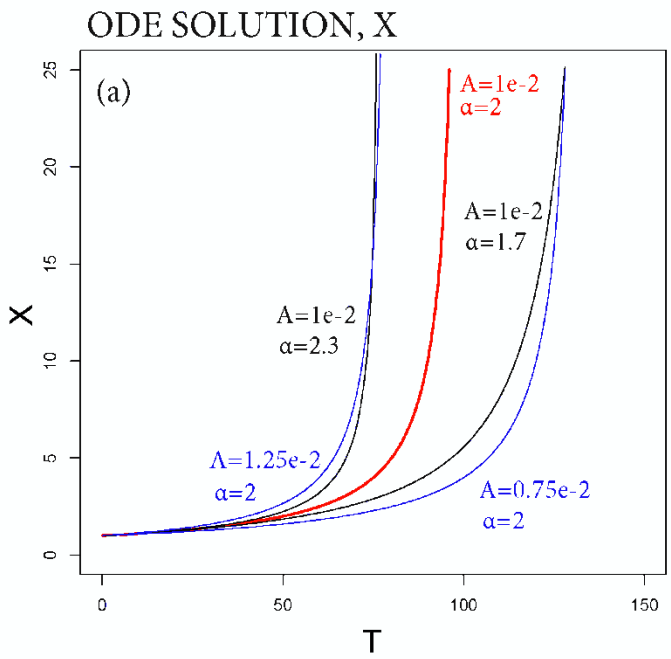
$$dX/dt = AX^{\alpha}$$

where X is the time rate of signals

α - shape
A - slope
t_0 - initial time

$$X(t) = \left[(1-\alpha)A(t-t_0) + X(t_0)^{1-\alpha}\right]^{\frac{1}{1-\alpha}}$$

SOLUTION



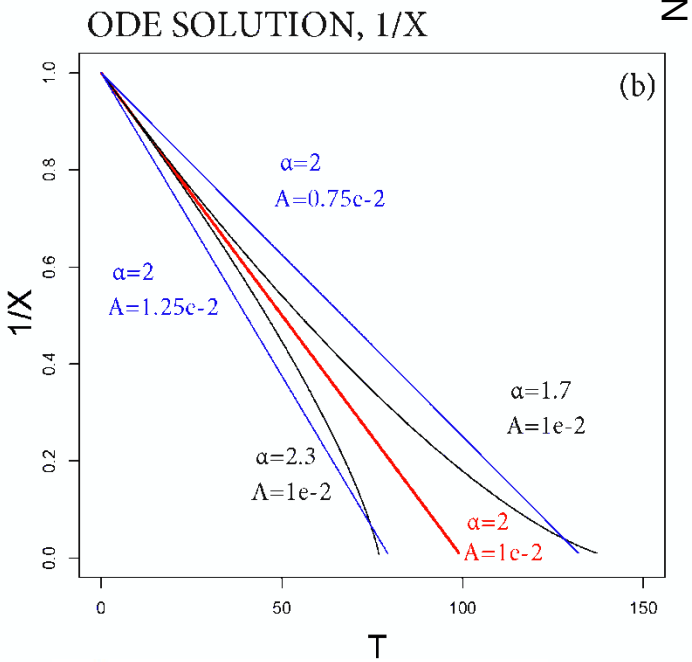


Figure.
ODE solution,
(a) X, and (b) 1/X.

Note the effect of varying slope and convexity parameters.
α, A learnt from data

$$\eta := X^{1-\alpha}$$

change of variables ⟹ *linearization*

$$\frac{d\eta}{dt} = (1-\alpha)A$$

$$\eta(t) = (1-\alpha)A(t-t_0) + \eta(t_0)$$

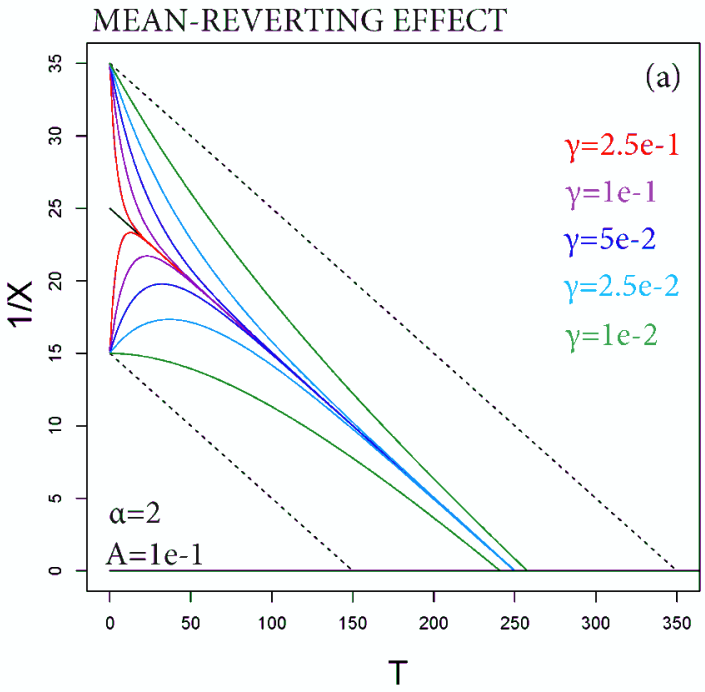
t_f - failure time

$$t_f = \inf\{t : \eta(t) = 0\}$$

New terms: mean-reversion and stochastic noise

$$\eta(t) = (1-\alpha)A(t-t_0) + \underbrace{\beta \exp(-\gamma t)}_{\text{mean-reversion term}} + \eta(t_0)$$

It makes every perturbation decay with time



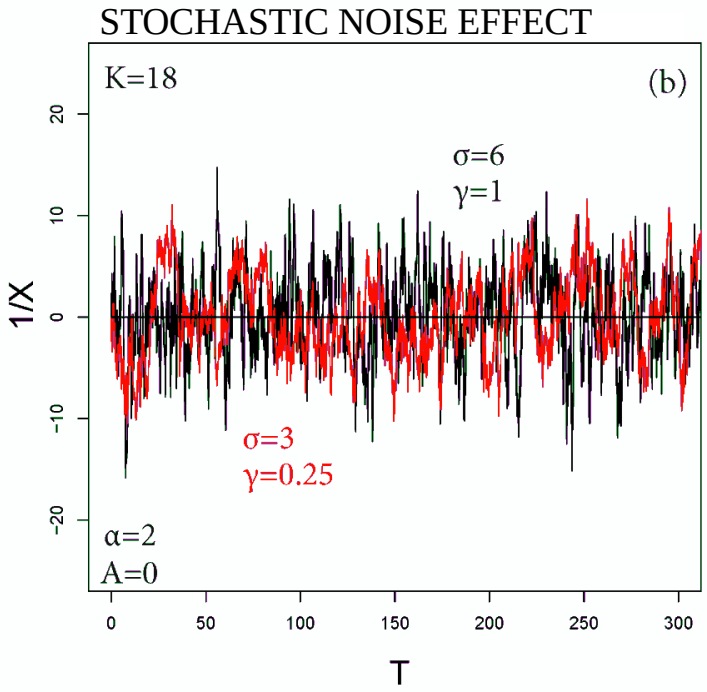


Figure. (a) 1/X with α=2, A=0.1, β=±10. The colors show different γ.

(b) $(\eta_t)_{t>0}$ with A=0. The colors show different (γ ,σ), with equal $K = \sigma^2/\gamma$.

$$d\eta_t = -\gamma \eta_t dt + \underbrace{\sigma dW_t}_{\text{noise term}}$$

(A=0)

SOLUTION

$$\eta_t \sim \mathcal{N}\left(0, \frac{\sigma^2}{2\gamma}\left[1-\exp(-2\gamma t)\right]\right) \simeq \mathcal{N}\left(0, \frac{\sigma^2}{2\gamma}\right)$$

γ - mean-reversion
σ - noise
β - initial perturbation

Parameters are based on the residuals in the linearized problem.

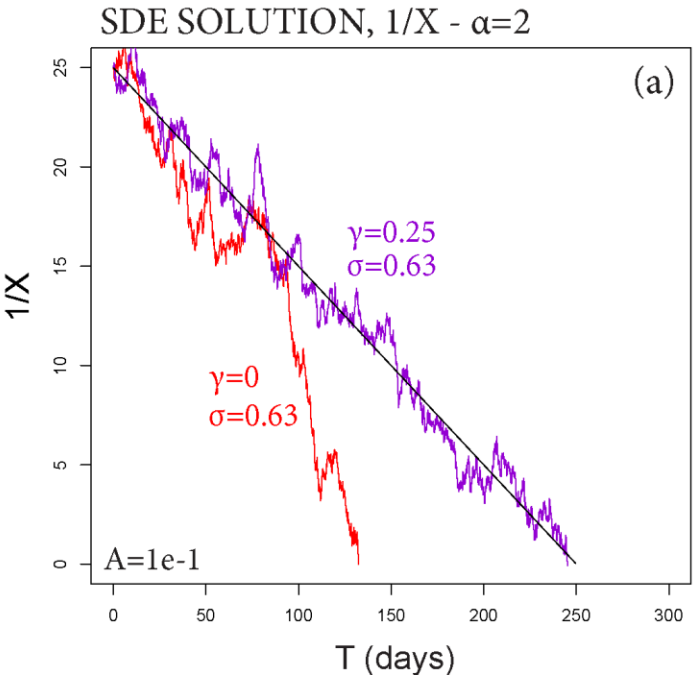
The FFM stochastic differential equation (SDE)

$$d\eta_t = \{\gamma\left[(1-\alpha)A(t-t_0) + \eta_{t_0} - \eta_t\right] + (1-\alpha)A\}\,dt + \sigma dW_t$$

mean-reversion terms | classical FFM | noise term

$$X_t = \left\{ X_{t_0}^{1-\alpha} + \int_{t_0}^{t} \{\gamma\left[(1-\alpha)A(s-t_0) + X_{t_0}^{1-\alpha} - X_s^{1-\alpha}\right] + (1-\alpha)A\}\,dt + \int_{t_0}^{t} \sigma dW_s \right\}^{\frac{1}{1-\alpha}}$$

NONLINEAR FORMULATION



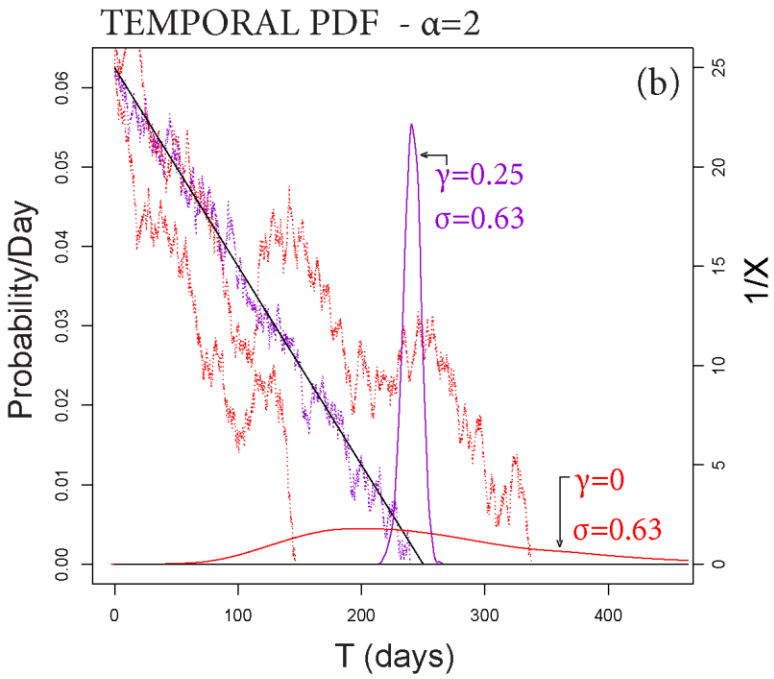


Figure. SDE solutions, with α=2, A=0.1.

The black line is the mean solution.

(a) colored lines are random paths, γ=0 or γ=0.25.

(b) also shows g_{tf}. The solution 1/X is reported again.

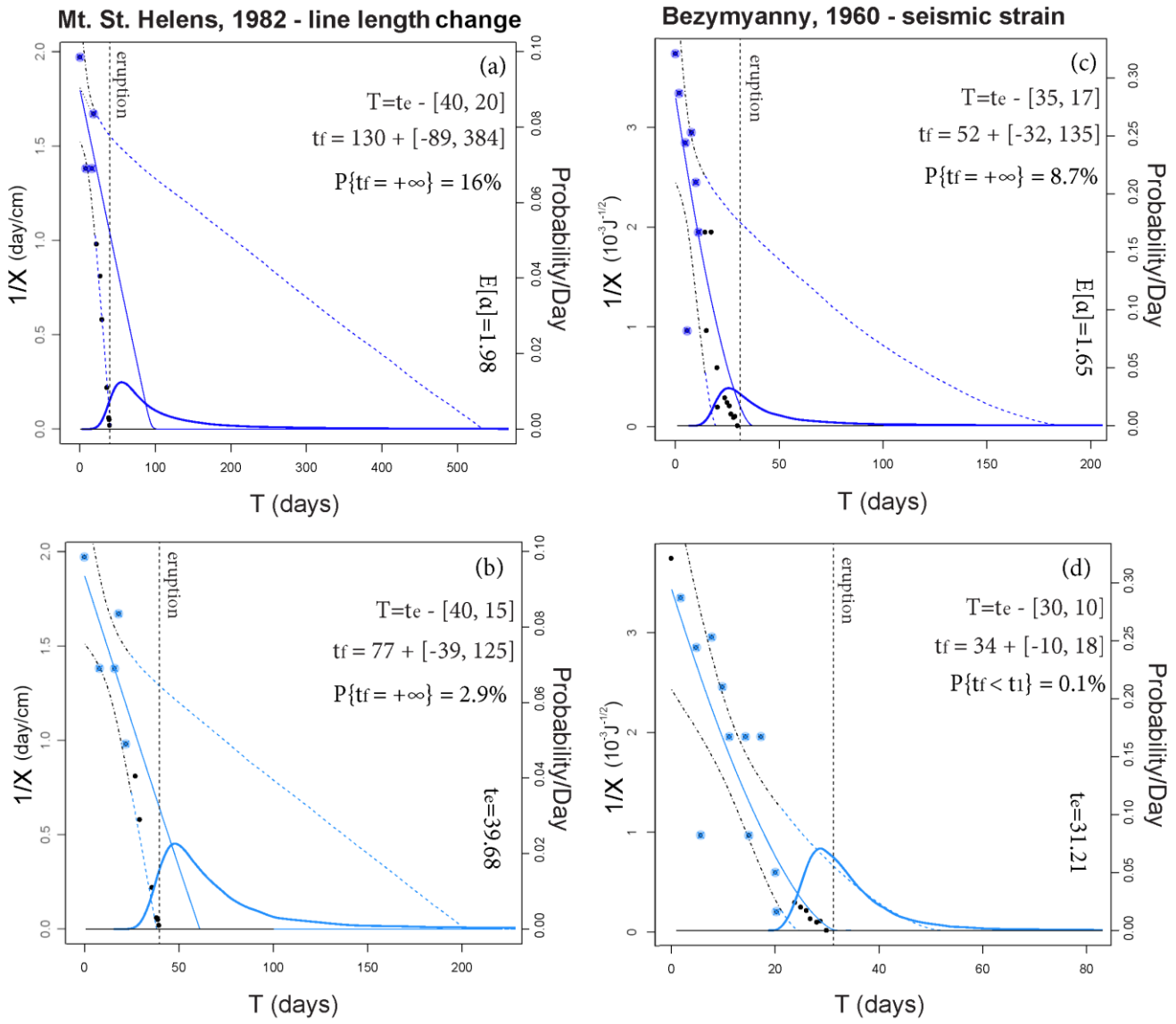
$$t_f(\omega) = \inf\{t : X^{-1}(\omega, t) = 0\}$$

random variable

$$g_{t_f} : \mathbb{R} \to \mathbb{R}+, \qquad \int_0^{\infty} g_{t_f}(x)dx = 1$$

t_f probability density function

Example I - ODE forecasts of t_f



Forecasts can be significantly **uncertain**, especially if based on few data.

We compare two **time windows** with extremes reported in figure. They include different outcomes of data (trends).

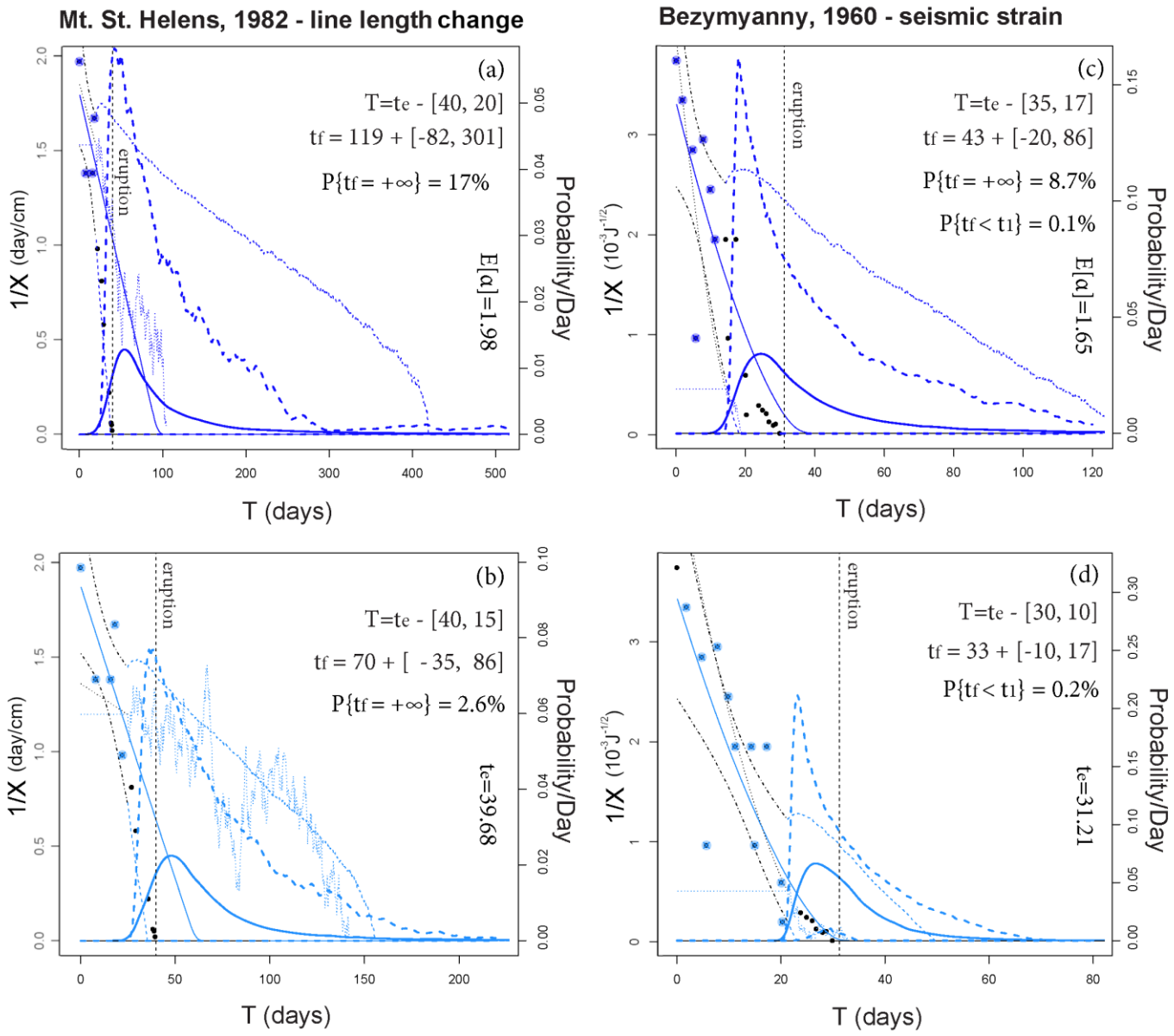
Large colored dots mark the **data points evaluated** with the FFM. Small dots are not considered.

Figure.
Forecasts of t_f based on different time windows $[0, T]$.

The bold line is g_{tf}. Dashed lines bound a 90% confidence of solutions.

A dashed vertical line is t_e.

Example II - SDE forecasts of t_f



If the forecast is poorly constrained, including the noise typically **reduces the uncertainty** affecting t_f. Indeed the noise can push $1/X$ to hit zero earlier, when $1/X$ is small enough.

Large colored dots mark the **data points evaluated** with the FFM. Small dots are not considered.

Figure.
Forecasts of t_f made on different time windows $[0,T]$.

The bold line is g_{tf}.
Bold dashed lines are its 5th and 95th percentile values.

Dashed lines bound a 90% confidence of solutions.
Dotted lines show examples of random paths.
A dashed vertical line is t_e.

Material failure likelihood - $g_{tf}(t_e)$

SDE mean forecasts provide consistent likelihoods with the ODE results.

The 95th percentile values are **significantly higher** than other forecasts, from 5% to 10% in the first and second time windows, and above 15% in the third.

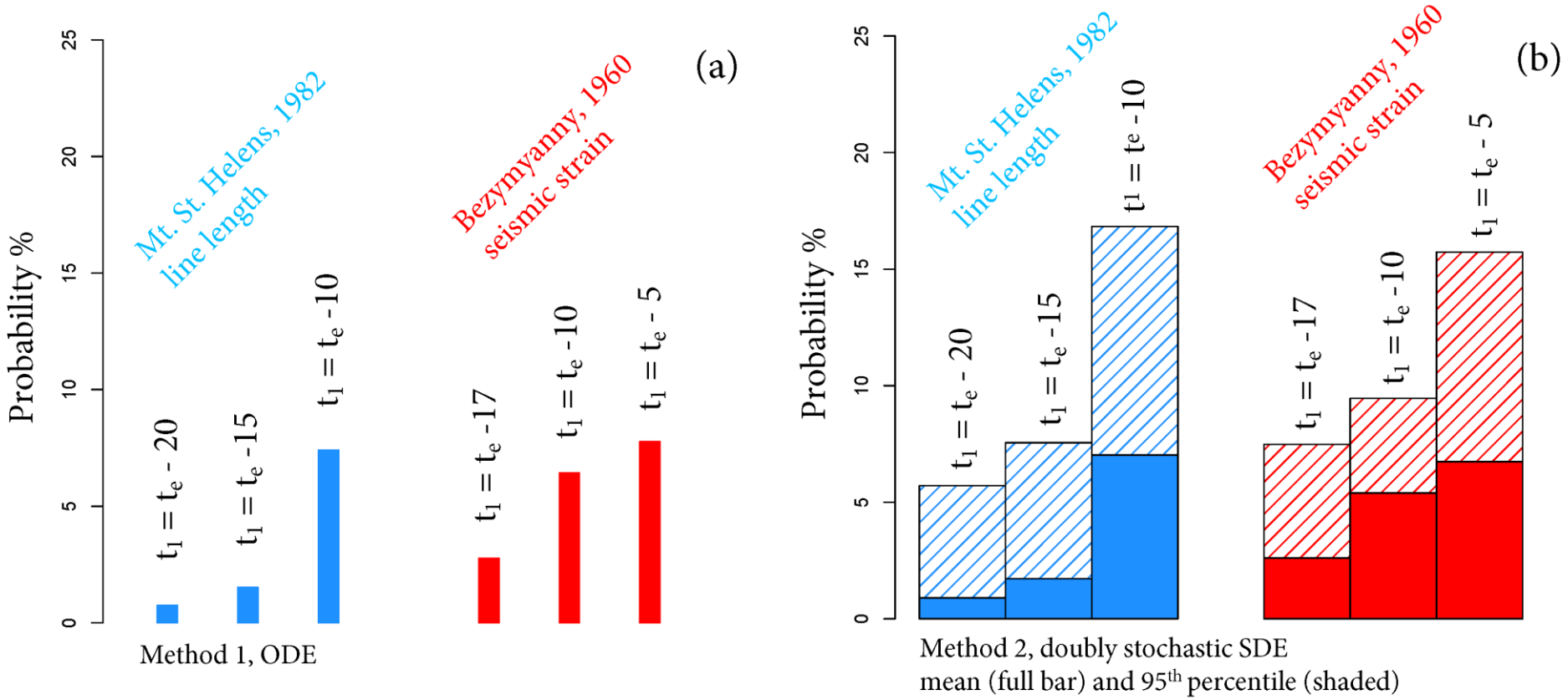


Figure. Barplots of the likelihood $g_{tf}(t_e)$ on three time windows. In (a) the colored bars assume the ODE formulation, in (b) the SDE formulation. Full bars are the mean values, shaded bars are the 95th percentile values.

PART 1B

The FFM applied to the Campi Flegrei caldera

Campi Flegrei caldera

Campi Flegrei (Italy) is a **volcanic field** that has been active in the last 80'000 years.
The depression of Campi Flegrei is generally interpreted as a **calderic structure**.

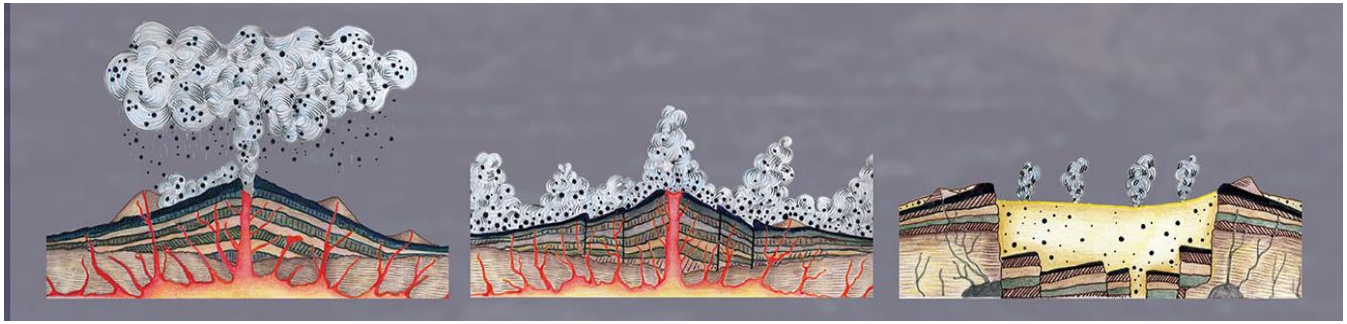


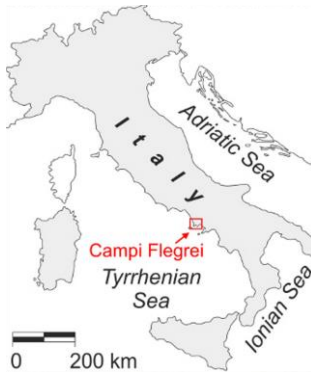
Figure. Schematic of a caldera collapse following the expulsion of magma because of a large scale explosive eruption.

Two large scale collapses are related to the eruptions of:

-Campanian Ignimbrite (40'000 years BP); -Neapolitan Yellow Tuff (15'000 years BP)

The central part of the caldera has been uplifting in the last 10'500 years (a **caldera resurgence** of ~100 m).

Episodes of slow uplift and subsidence of the ground, called **bradyseism**, characterize the recent dynamics of the Campi Flegrei caldera.



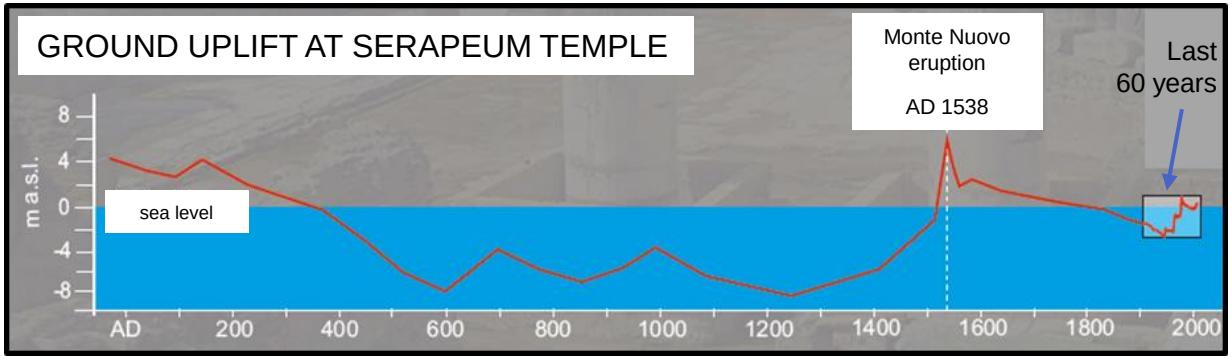
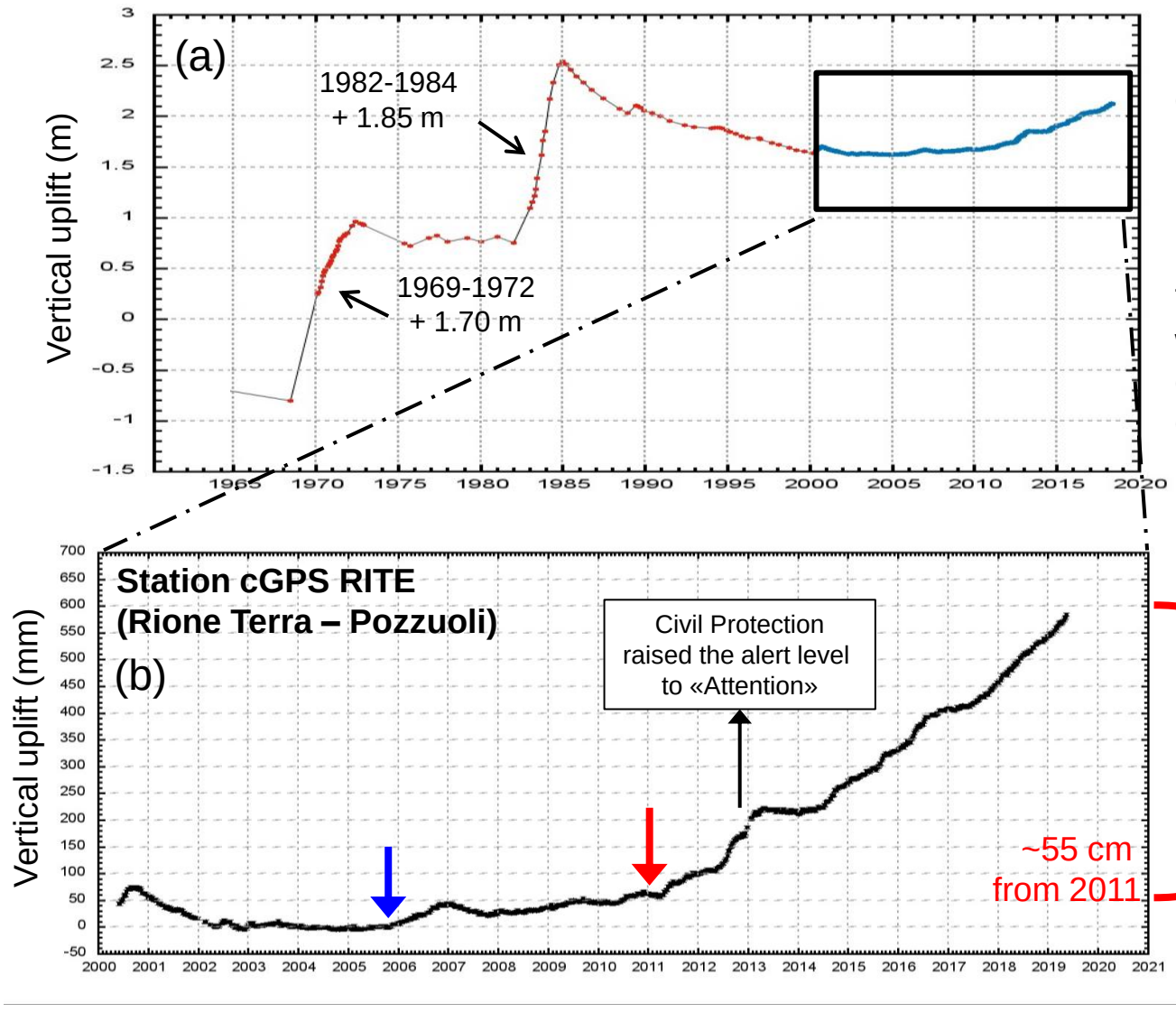


Figure. Plot of the elevation above sea level of the ruins of the Macellum of Pozzuoli (194 BC, called **Serapeum Temple**).

Data reconstructed from the borings of marine organisms.

Bradiseismic crises



In the last decades two major **bradiseismic crises** occurred in 1969/1972 and in 1982/1984, with a ground uplift of 1.70m and 1.85m, respectively.

Thousands of **earthquakes**, with a maximum magnitude of 4.2 caused the partial evacuation of the town of Pozzuoli in October 1983.

Figure. Ground deformation measurements collected at RITE (**vertical uplift**) by means of (a) leveling and (b) GPS.

Ground deformation and seismic data

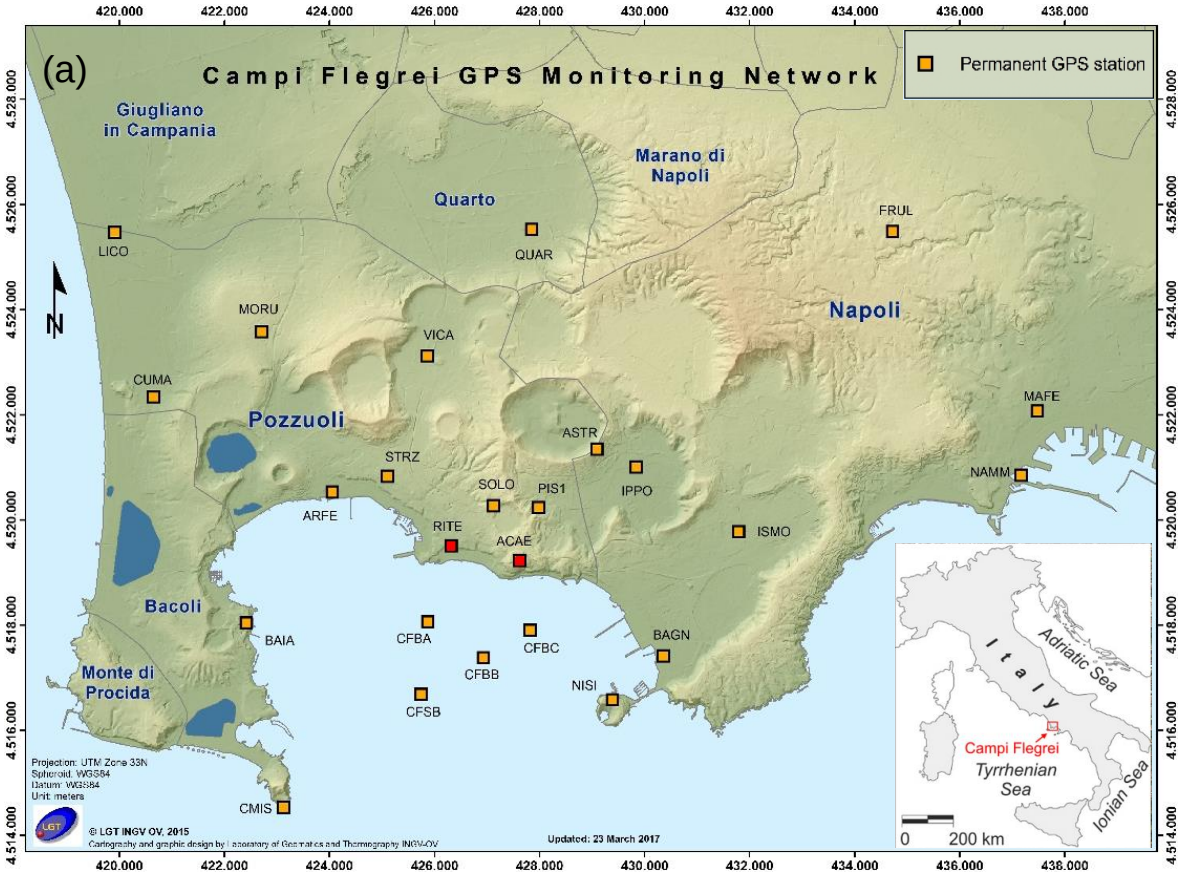
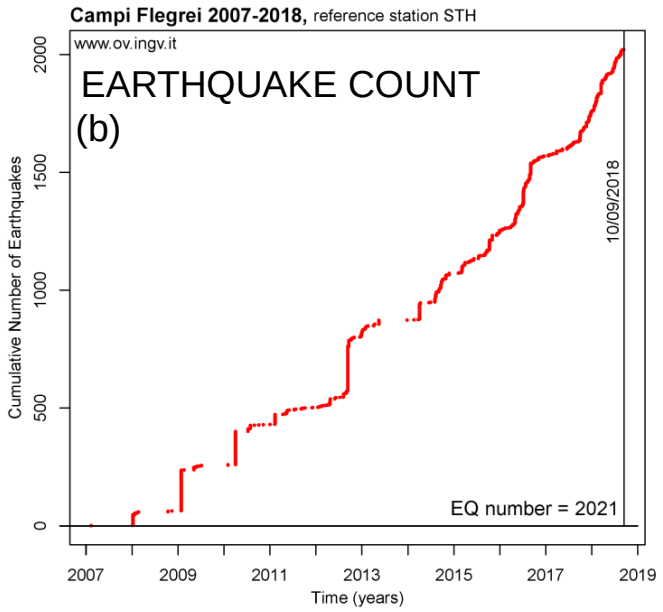


Figure. (a) Campi Flegrei caldera GPS monitoring network (21 stations + 4 bouys). RITE and ACAE stations marked in red.

(b) Cumulative number of EQ measured in CF from 1st Jan 2007 to 10th Sep 2018.

We preliminarily applied our enhanced FFM on Campi Flegrei caldera seismic and ground deformation dataset.

We remark that the time window is much longer than in the classical applications, and **spans over tens of years**.



Seismic data analysis

t_f is the time when **accelerating signals** as observed in the last 10 years would diverge to infinity.
The interpretation of t_f as the onset of a volcanic eruption is **speculative** (Chiodini et al., 2017; Kilburn, 2018).

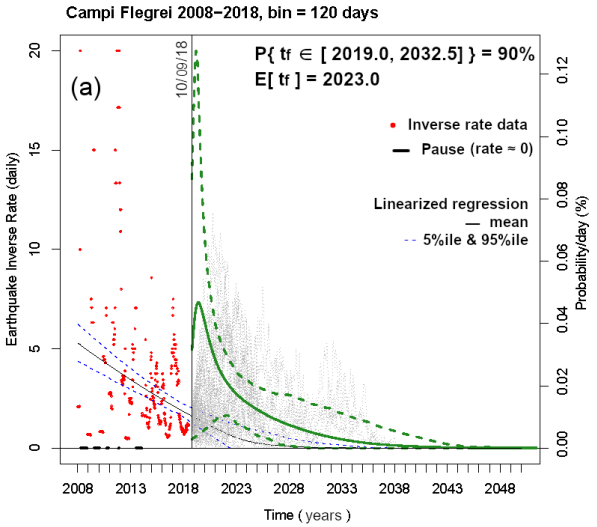
Figure. Probability forecasts of t_f using the seismic data of 2008-2018.
In (a) the inverse rate is obtained on 120 days, in (b-d) on 360 days.

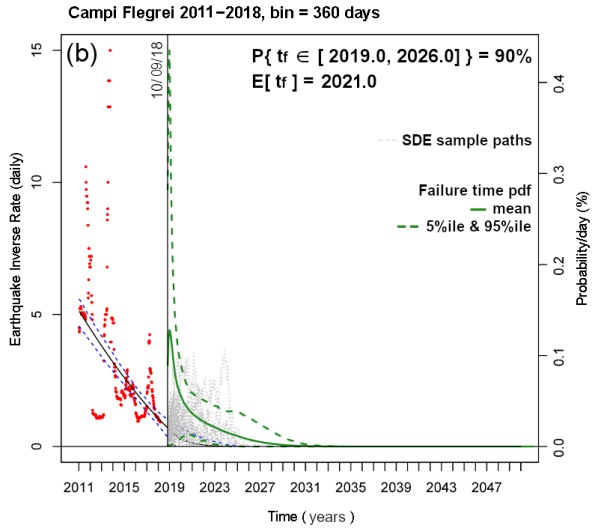
(b, d) are based on the data of 2011-2018, and (c, d) remove the swarms: $\forall i,\ t_{i+1} - t_i > 6$ hours.

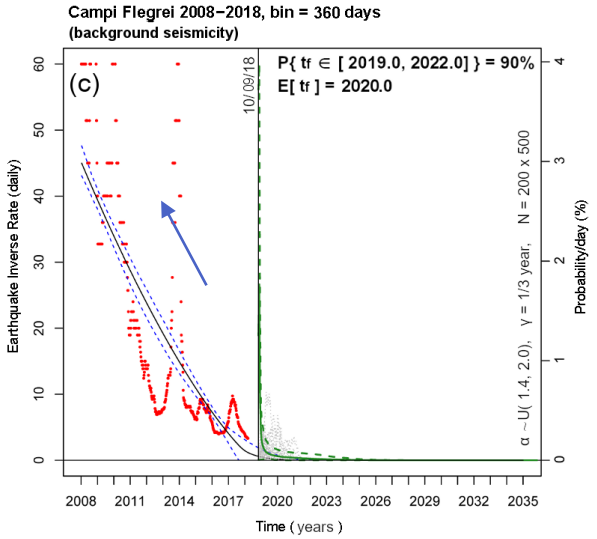
Removing the swarms from the 2008-2018 dataset produces significantly short forecasts.

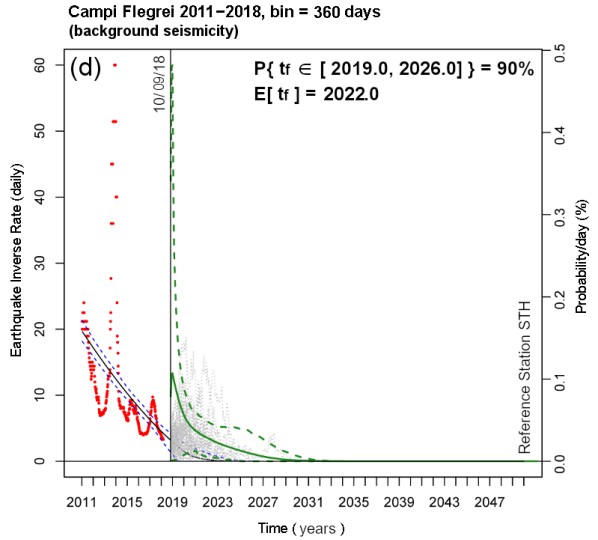
Red points are inverse rate data.
The **green line** is mean value of g_{tf}, the probability/day scale bar is related to it.
Dashed lines mark its 5th and 95th percentiles.

Thin blue dashed lines bound the 90% confidence interval of the ODE paths of 1/X, and a thin line is the mean path.
Grey dotted liner display 50 SDE paths.









Ground deformation data analysis I

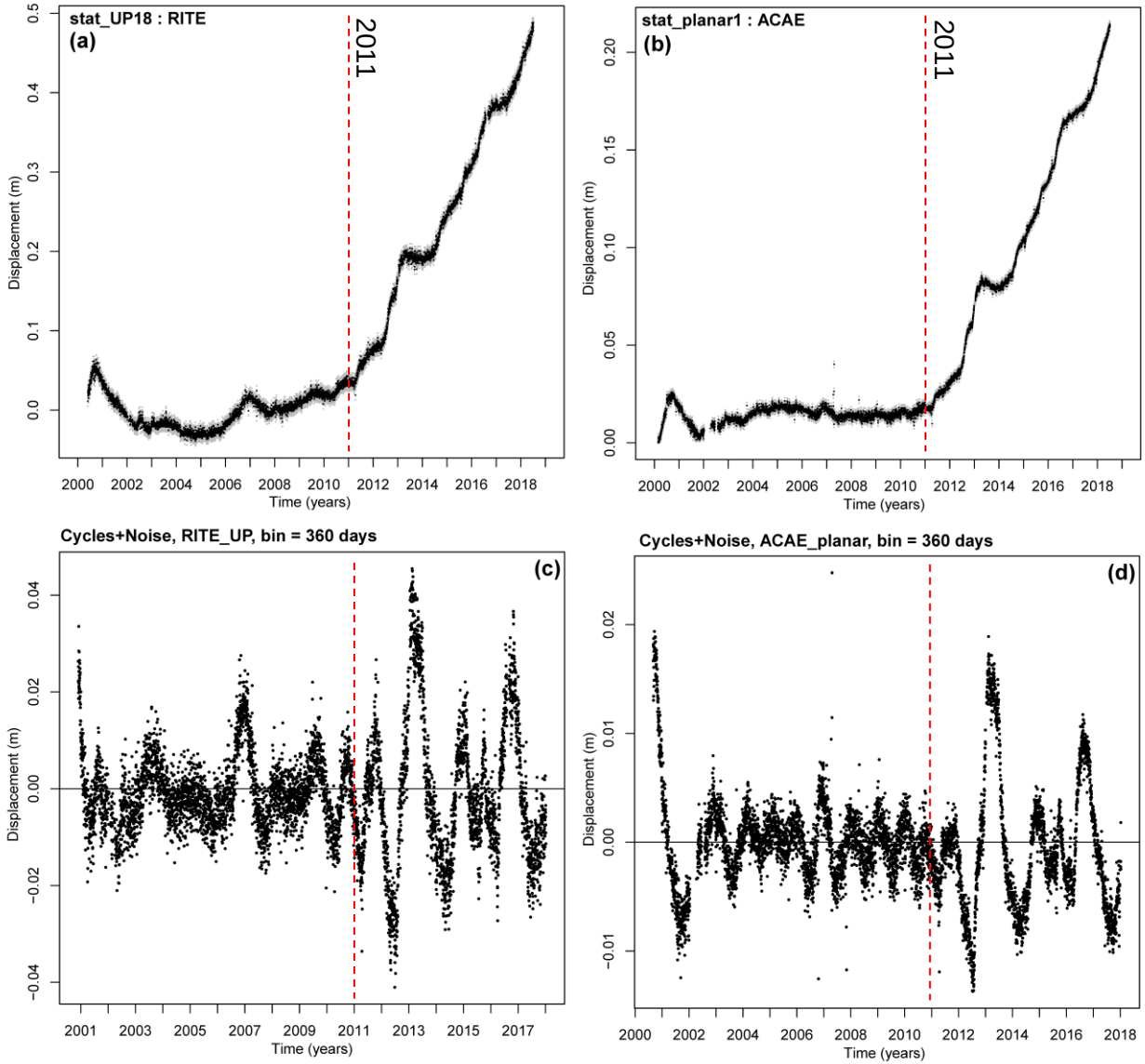


Figure. Ground deformation measurements of 2000-2018 collected at (a) RITE (vertical), and (b) ACAE (planar) GPS stations.

In (c-d) we show the residual signal after subtracting the average measurement over a 360 days moving window.

Seasonal cycles are evident, as well as noise effects.

Period and amplitude of the cycles apparently changed in 2011.

Rate change in 2011 is indicatively reported with a descriptive purpose.

Ground deformation data analysis II

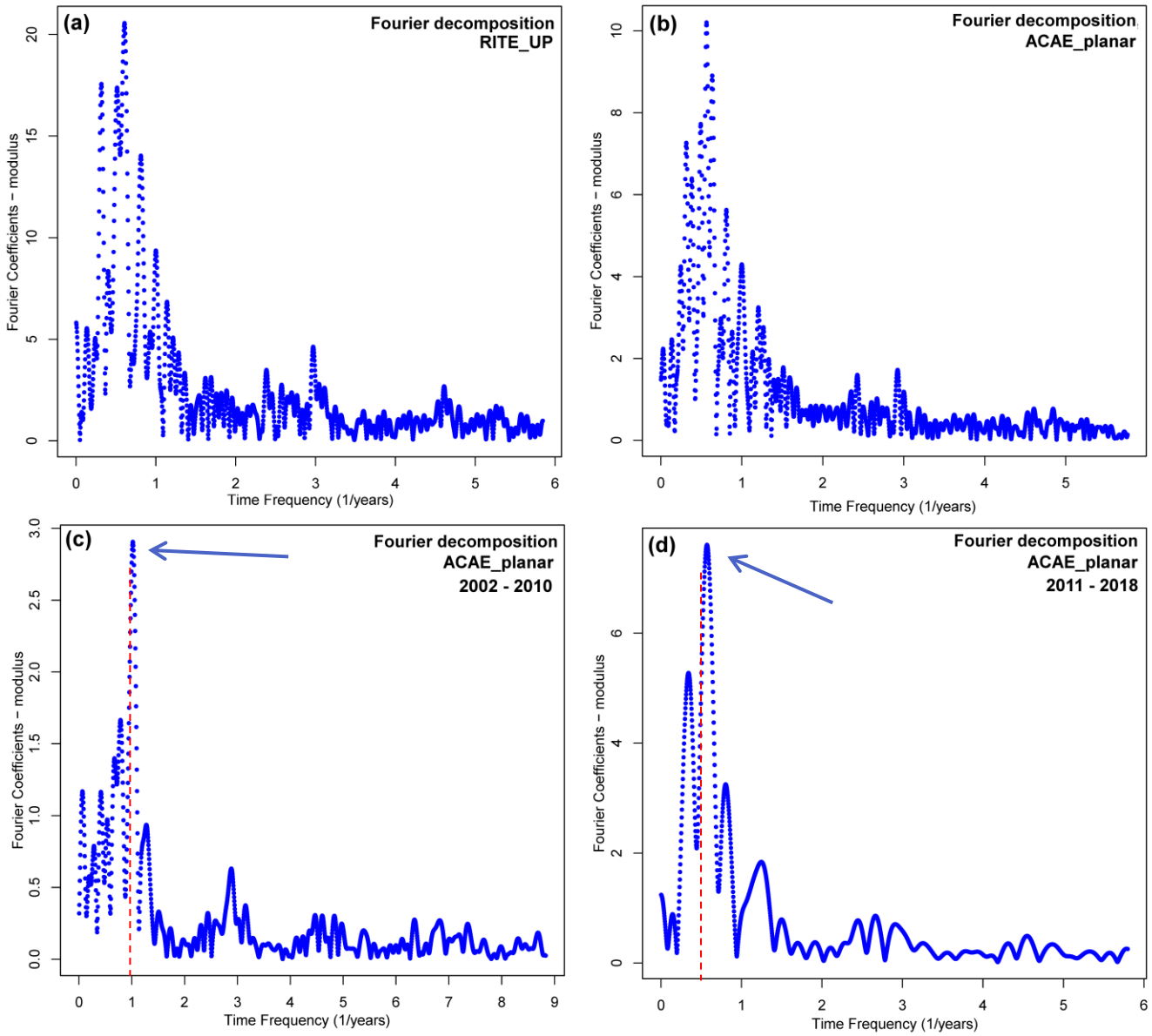


Figure. Fourier analysis of the residual deformation measurements collected at (a) RITE (vertical), and (b-d) ACAE (planar) GPS stations.

In (a-b) we show the $\|\cdot\|_2$ of Fourier coefficients, as obtained over the whole time domain 2000-2018.

In (c) we focus on the subdomain 2002-2010, in (d) on the 2011-2018 subdomain.

The period of the peak mode of the cycles appears to have changed:

~1 year in 2002-2010
~2 years in 2011-2018.

Ground deformation data analysis III

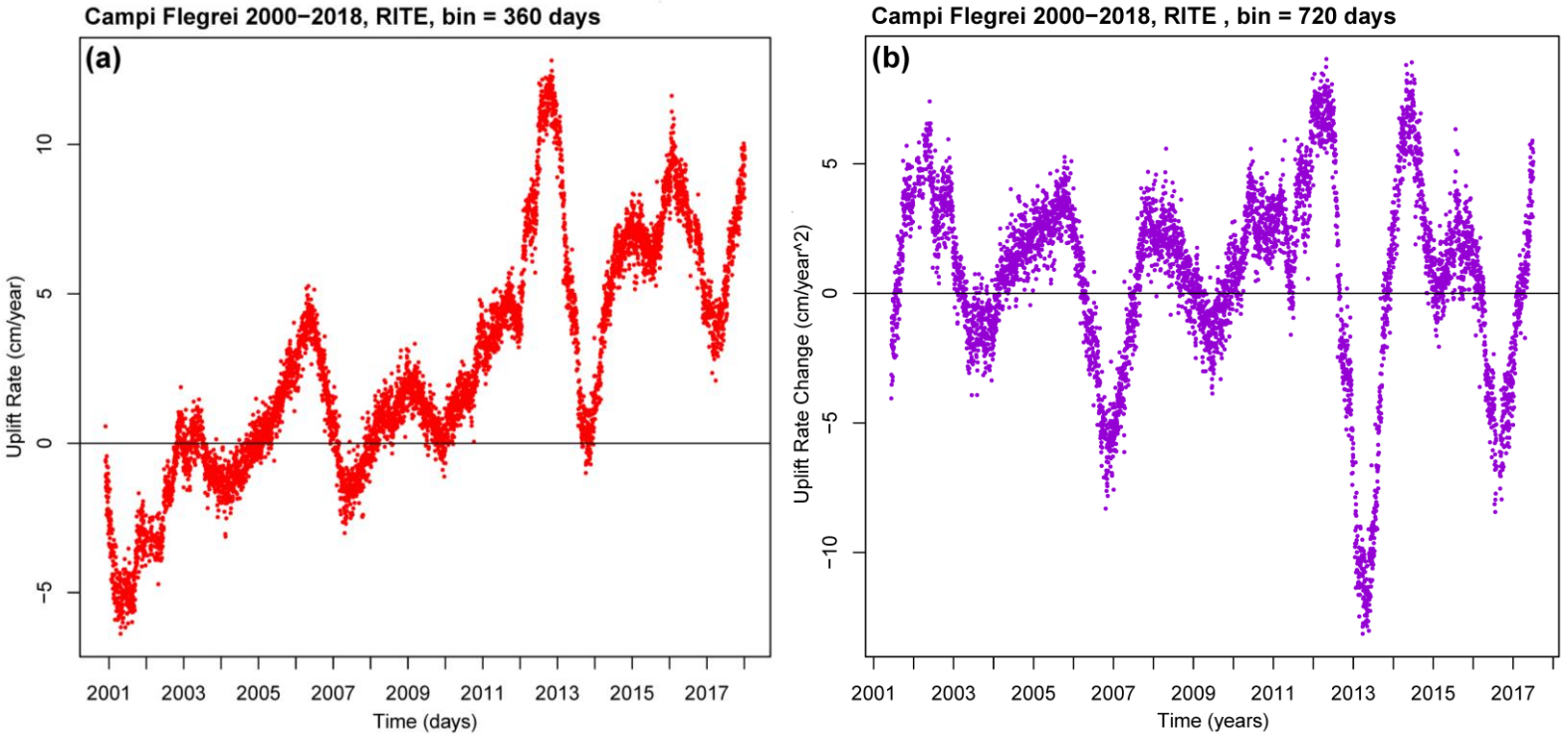


Figure. (a) Time rate of the deformation measurements collected at RITE (vertical) GPS stations, according to finite differences calculated over 360 days. (b) Rate change of the same time series, calculated by second order finite differences over 720 days.

The minor uplift rate increases occurring in 2005-2006 and in 2011-2013 are evident, as well as the average rate, that increased up to above 10 cm/year in 2012-2013 and 2015-2016.

Ground deformation data analysis IV

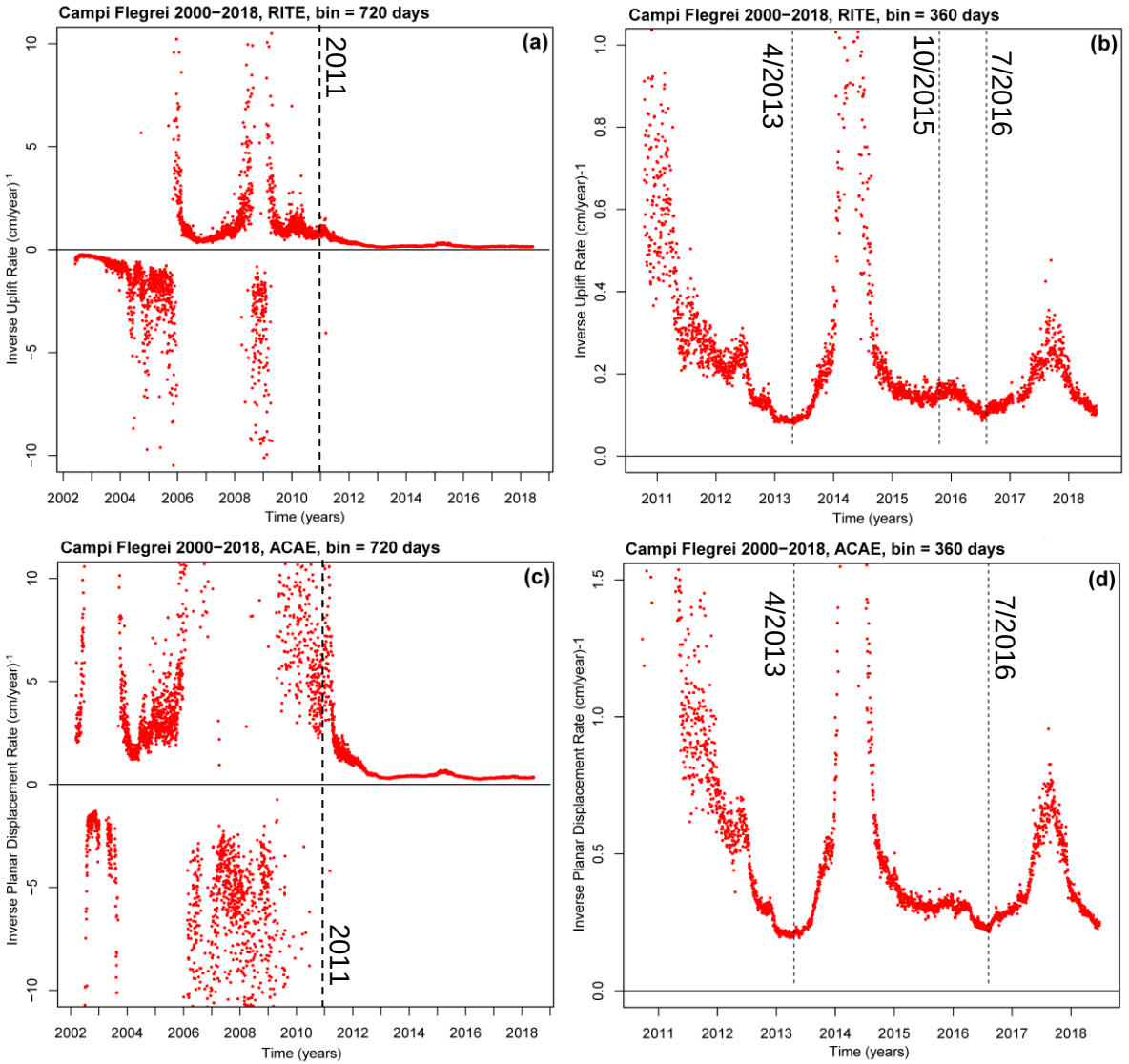


Figure. Inverse rate plots of the deformation measurements collected at (a,b) RITE (vertical), and (c,d) ACAE (planar) GPS stations.

In (a,c) we show the results obtained over the time domain 2002-2018.

In (b,d) we focus on the subdomain 2011-2018.

Multiple phases of inverse rate decrease compatible with the FFM equations occurred in 2011-2013, and 2014-2016.

Ground deformation data analysis V

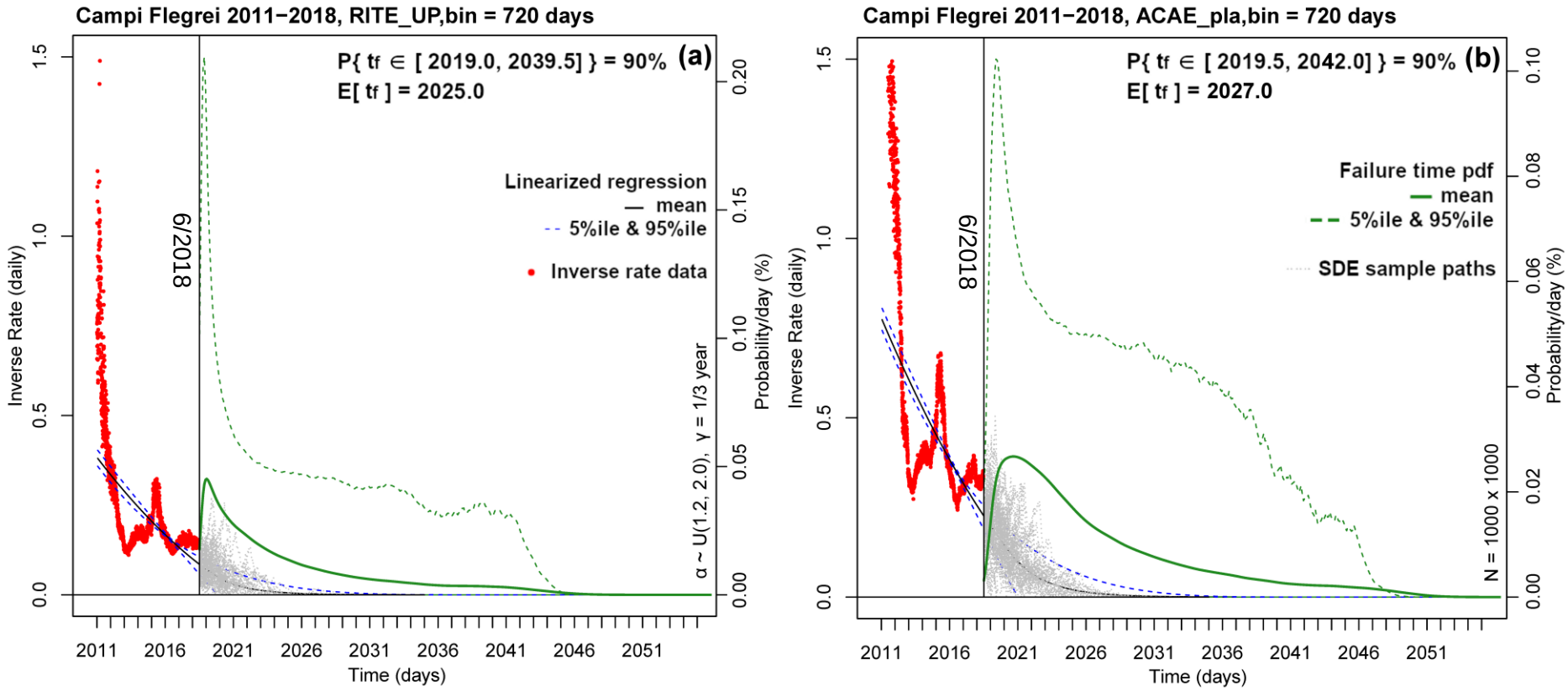


Figure. Probability forecasts of t_f using the deformation data of 2011-2018. In (a) the inverse rate is obtained from RITE (vertical), and in (b) from ACAE (planar) GPS stations.

Red points are inverse rate data. The **green line** is mean value of g_{tf}, the probability/day scale bar is related to it. **Dashed lines** mark its 5th and 95th percentiles.

Thin blue dashed lines bound the 90% confidence interval of the ODE paths of 1/X, and a thin line is the mean path. **Grey dotted** liner display 50 SDE paths.

PART 2

Update of a prior spatial map through the statistical assimilation of precursor data

Long-term vent opening map with uncertainty

A vent opening map is a family of pdfs for the spatial location of a new eruptive vent.
A hierarchical sampling is employed to choose the pdf and then the location.

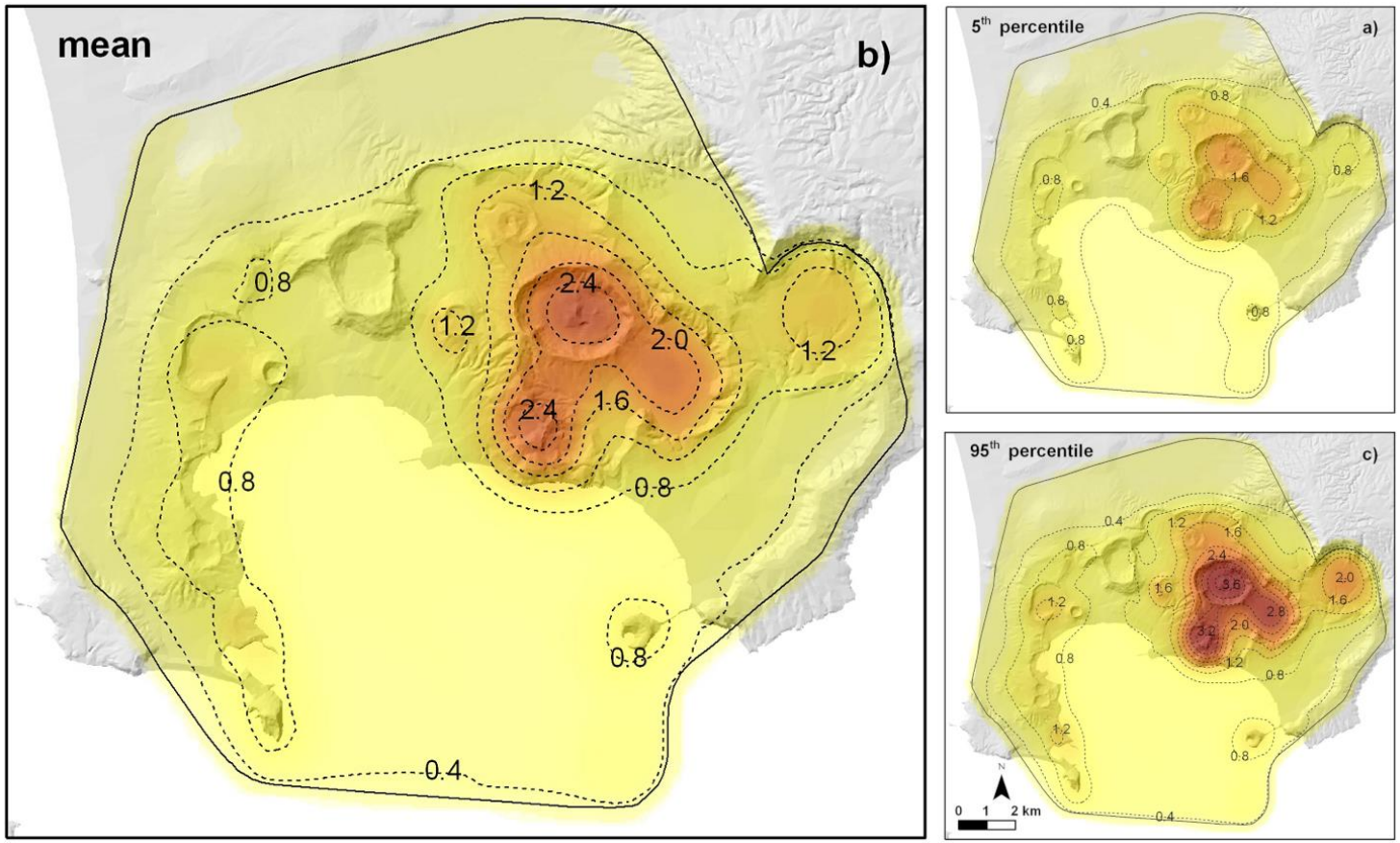


Figure.
Vent opening map based on long-term volcanological data. The values are probabilities over km^2. (Bevilacqua et al., 2015)

A Gaussian random field can be adopted to represent the uncertainty affecting the pdf values.

We choose this field as a prior parameter for the Bayes' theorem.

Gaussian random field of the spatial pdf

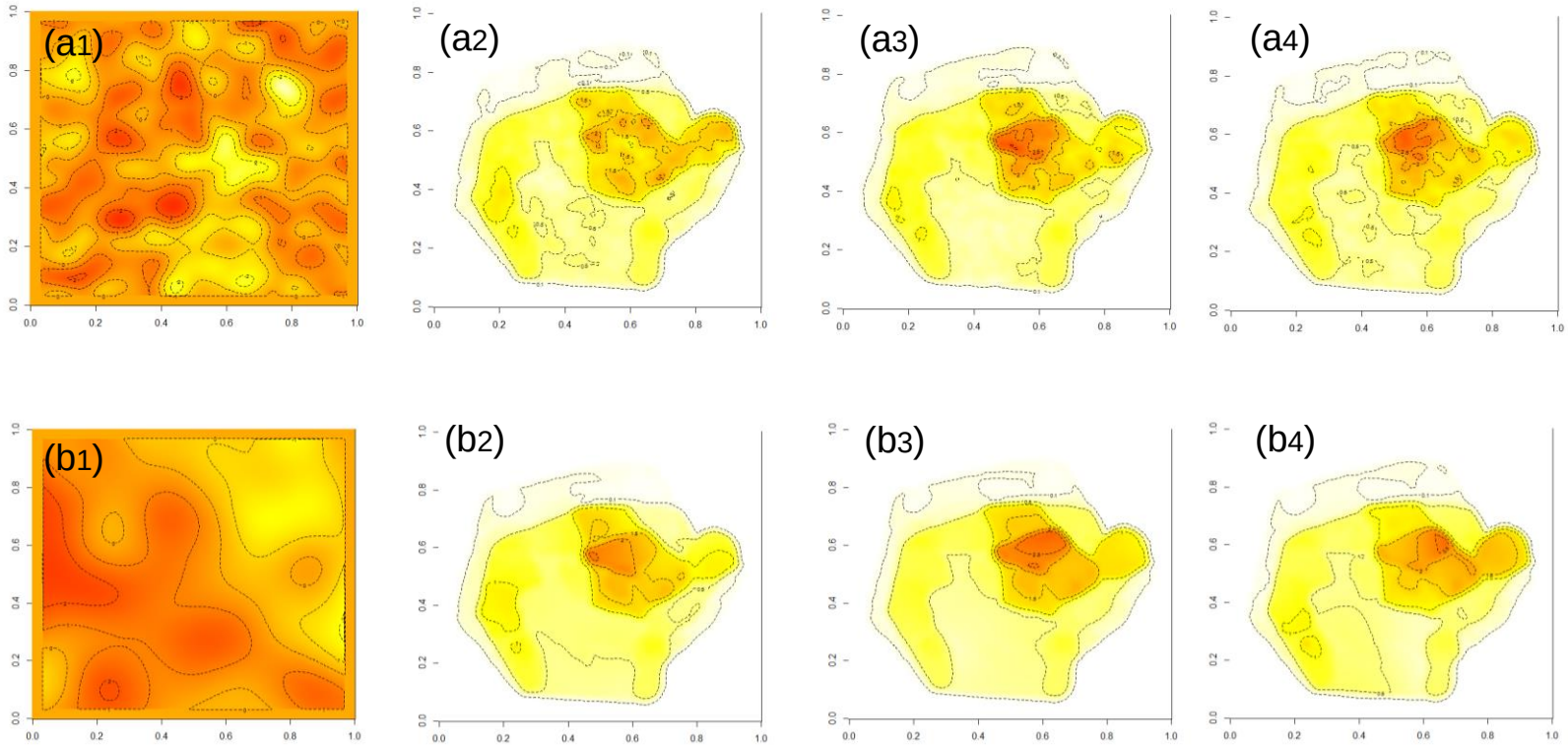


Figure. Example of random fields representing the vent opening map.

The plots assume Gaussian correlation kernels of different bandwidths.

Each sample is normalized to sum to one. Thanks to the law of large numbers, if correlation is not too strong, this is introducing a negligible error.

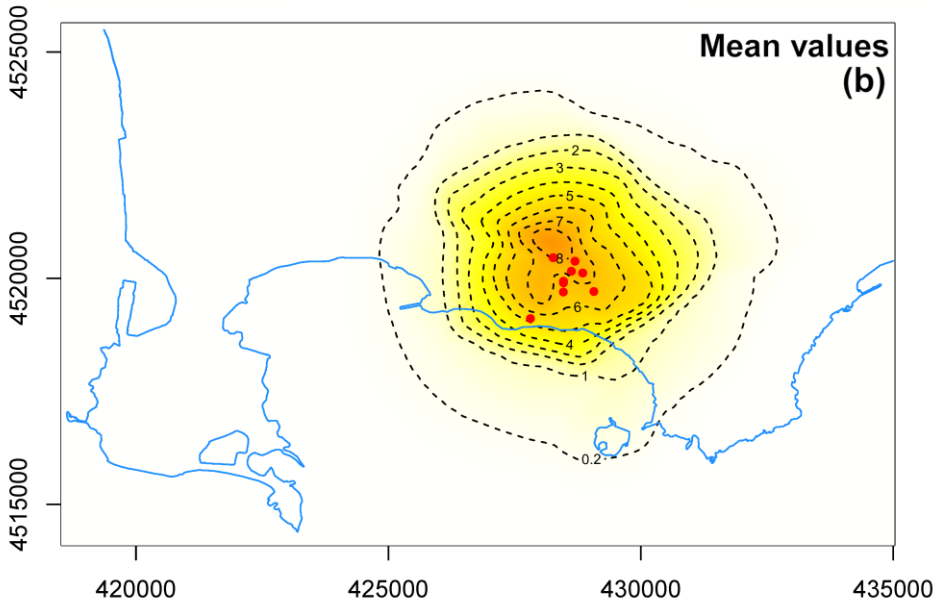
(a1-4) assume a radial limit of 2 km, (b1-4) of 5 km, i.e. beyond this radius the correlation is below 5% of its maximum value.

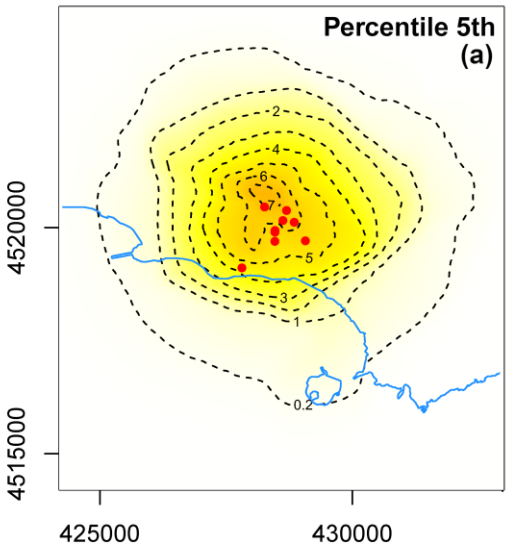
(a1, b1) display random samples of a Gaussian field with constant mean and unitary autocorrelation.

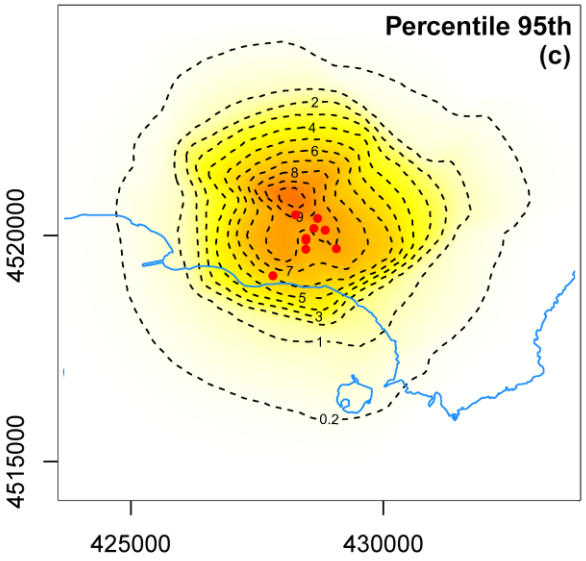
(a2-4, b2-4) show three random samples of a Gaussian field having the same 5^{th} and 95^{th} percentiles of the vent opening map in Bevilacqua et al. 2015.

We assumed a radial limit of 4 km in the following.

Posterior maps based on seismic data - I







Let X be the random variable expressing the new vent opening location according to the prior map g_X

Let F(x) be a likelihood function based on the distance from seismic epicenters. In this example we implement a Gaussian likelihood with 4 km radius (2σ).

Then, let Y be a random variable with values in {0,1}, expressing the event of having an eruption after the observed precursors. That is, P{Y = 1 | X = x} is given by F(x).

The Bayes' theorem states:

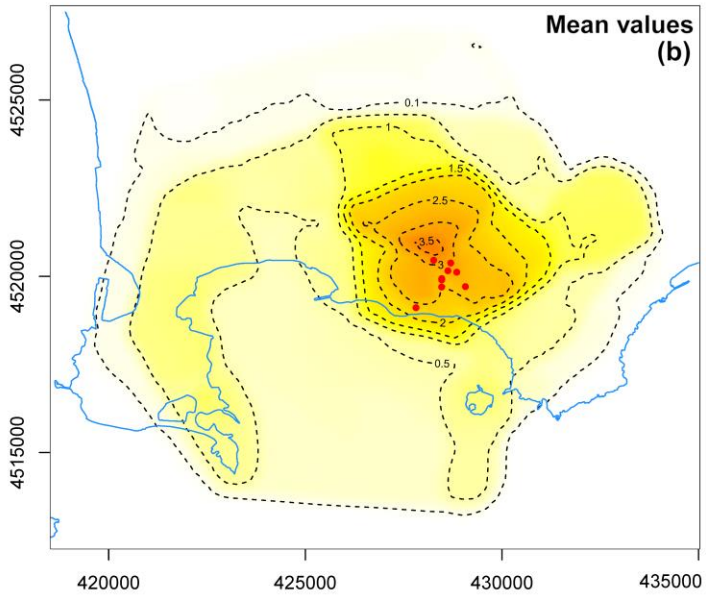
$$g_{X \mid Y=1}(x) = \frac{P\{Y = 1 \mid X = x\}\, g_X(x)}{P\{Y = 1\}} = \frac{F(x)\, g_X(x)}{h}$$

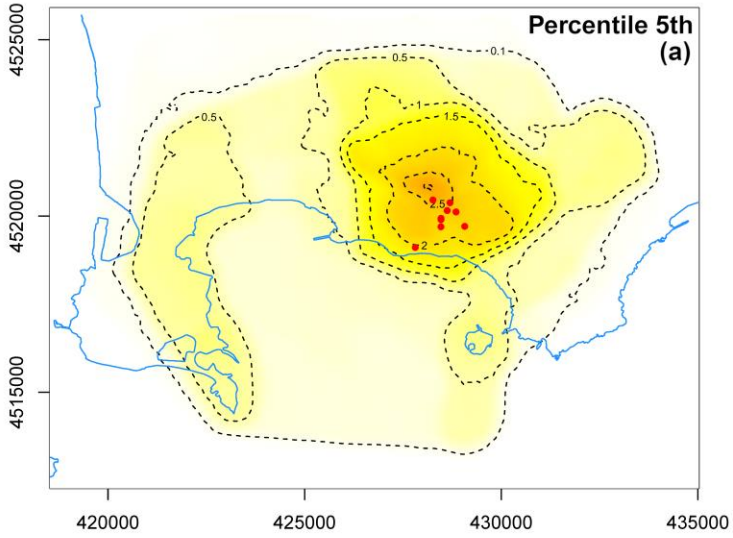
where the constant h is the Bayesian evidence that an eruption is going to happen, given prior and data.

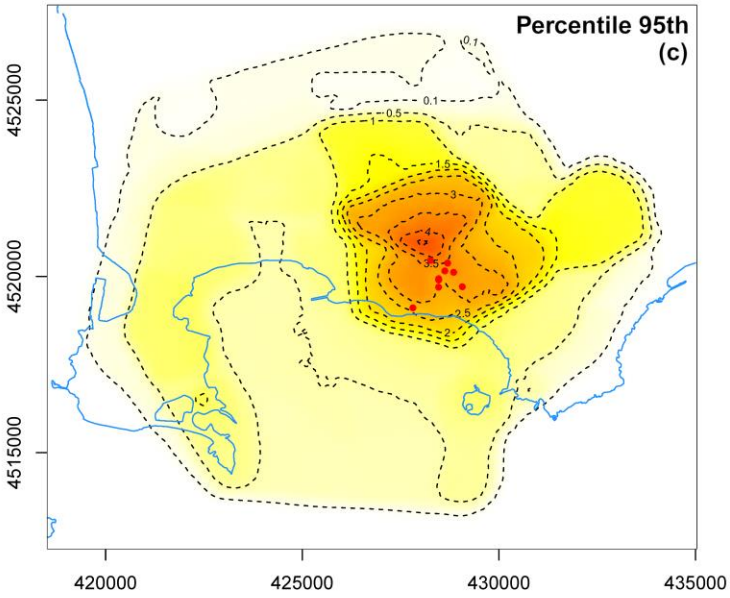
Figure. Vent opening map combining prior information from (Bevilacqua et al., 2015) and the spatial locations of the epicenters of the seismic swarm of 26/09/2017.

Here the map assumes that the swarm was composed of real eruptive precursors.

Posterior maps based on seismic data - II







If the observed signals are not real eruptive precursors, then

$\forall\, x, \quad P\{Y = 1 \mid X = x\} = P\{Y = 1\} \equiv h, \qquad g_{X \mid Y=1}(x) = g_X(x),$

that is, the Bayes' theorem leaves the prior map unchanged.

Let event A have a probability p to happen, then event B has a probability (1-p) to happen.

So, we define a binary logic tree made of two cases:

A. The signals are real eruptive precursors

$$g_{X \mid Y=1}(x) = \frac{F(x)\, g_X(x)}{h}$$

B. The signals are not eruptive precursors

$$g_{X \mid Y=1}(x) = g_X(x)$$

Figure. Vent opening map combining prior information from (Bevilacqua et al., 2015) and the spatial locations of the epicenters of the seismic swarm of 26/09/2017.

Here the map assumes that the swarm was composed of real eruptive precursors with a chance p = 20%.

Multi-polar interpolation method

Our spatial mapping of deformation data relies on a multi-polar interpolation, that is a linear interpolation among the pairs of applied vectors, made in polar coordinates with respect to the intersection of the straight lines defined by the planar displacement components.

Results are independent on arbitrary assumptions on the geometry, the physical properties of the source of deformation and of the elastic medium, but they are only based on a «local» central symmetry.

Centers of symmetry are locally determined for every pair of measurements, enabling the reconstruction of bimodal profiles, deviating from the usual bell-shape.

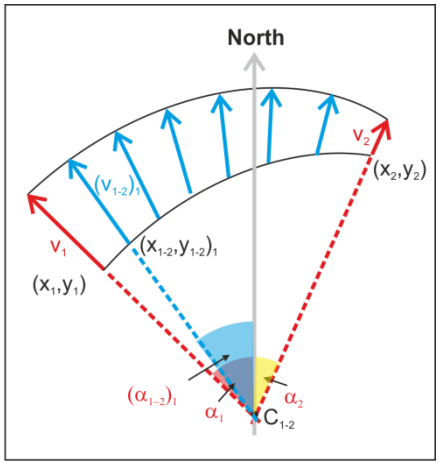


Figure A. schematic of multi-polar interpolation

$$\mathbf{v}^{(ij)\lambda} := [\lambda r^{(j)}+(1-\lambda)r^{(i)},\ \lambda\alpha^{(j)},\ \lambda\mathbf{d}^{(j)}+(1-\lambda)\mathbf{d}^{(i)}]$$

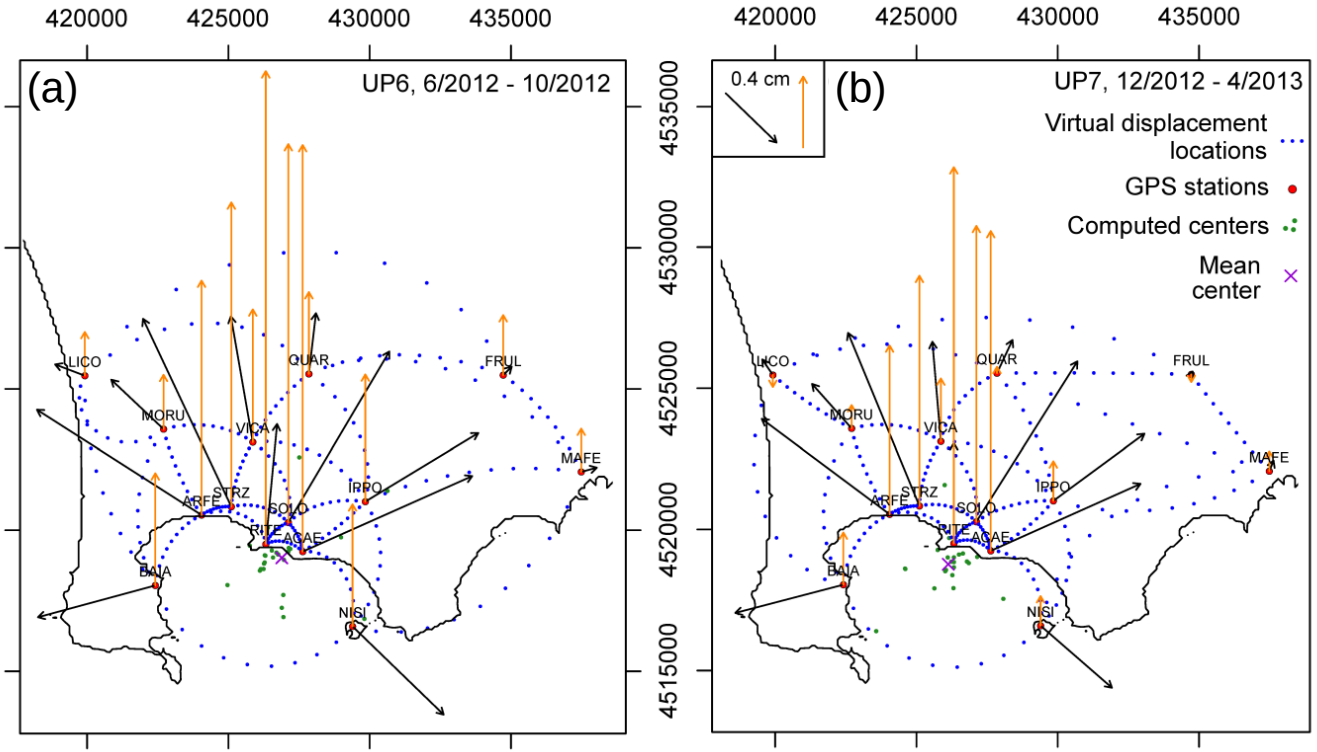


Figure B. (a-d) show examples of GPS data collected in 2012-2013.

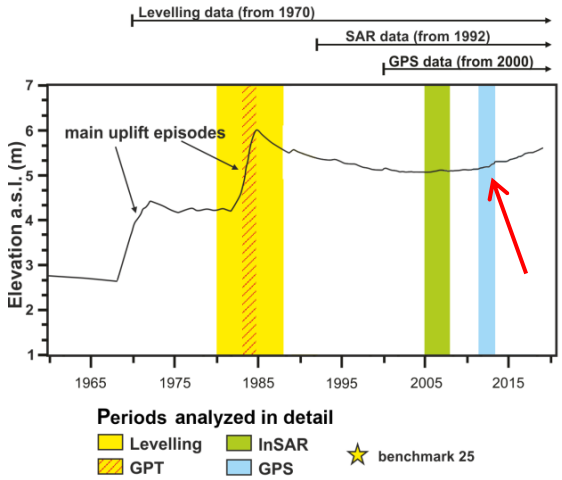
Blue dots mark the virtual displacements, along elliptic trajectories.

Green dots are the centers of symmetry. The **purple cross** is the average of the centers.

Vertical displacement maps

In 1970-1985, Campi Flegrei caldera underwent two episodes of major caldera-wide uplift and seismicity, which raised the central region by 3.5 m, followed by about 20 years of overall subsidence, until 2005.

We focus our analysis on the minor uplifts (cm - scale) occurred in 2011-2013 (De Martino et al. 2014).



UP4 4/2011 - 6/2011

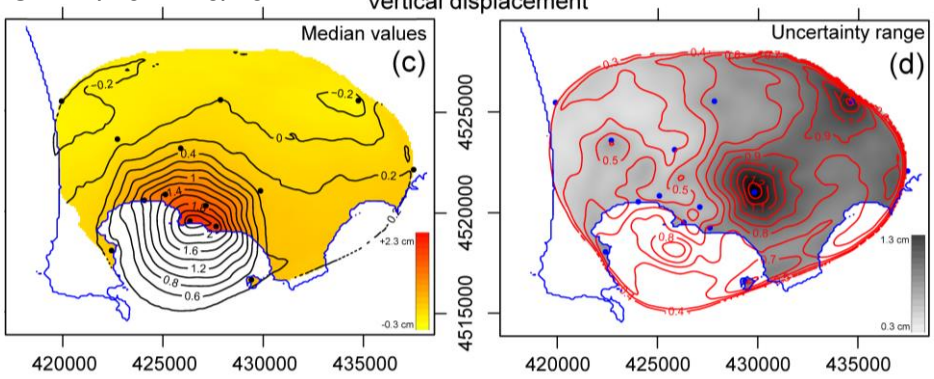
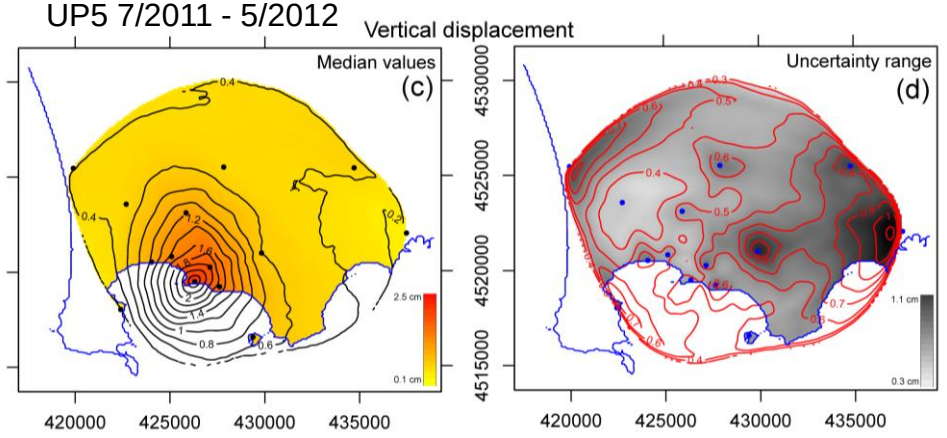
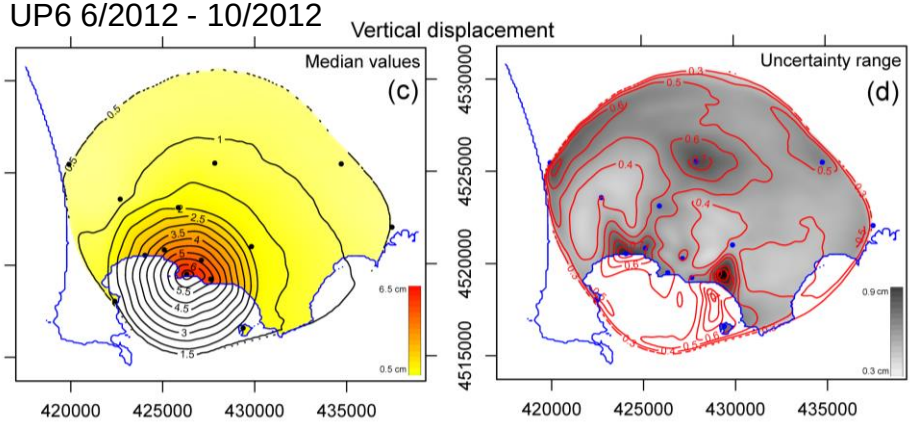


Figure. (a) maps of vertical displacement, based on the GPS data of UP4-UP6. (b) Uncertainty range, given by the difference of the 5th and the 95th percentiles of error.

UP5 7/2011 - 5/2012



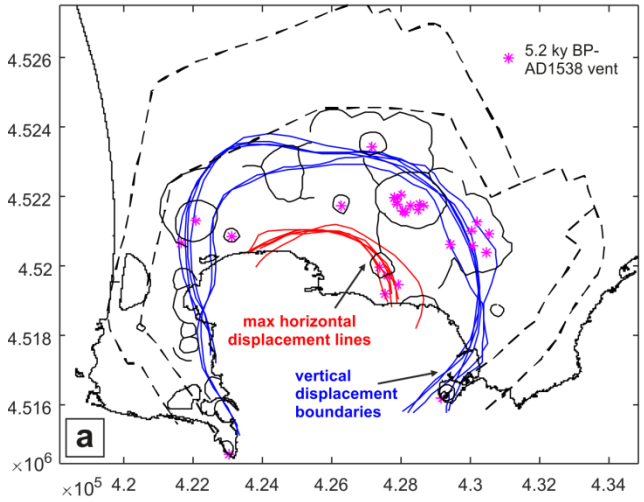
UP6 6/2012 - 10/2012

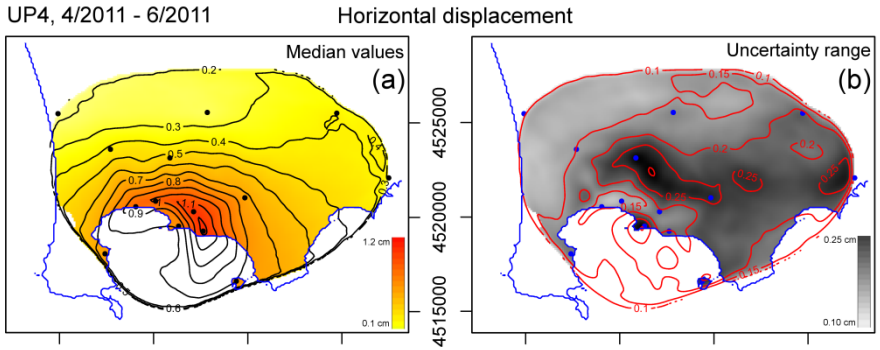


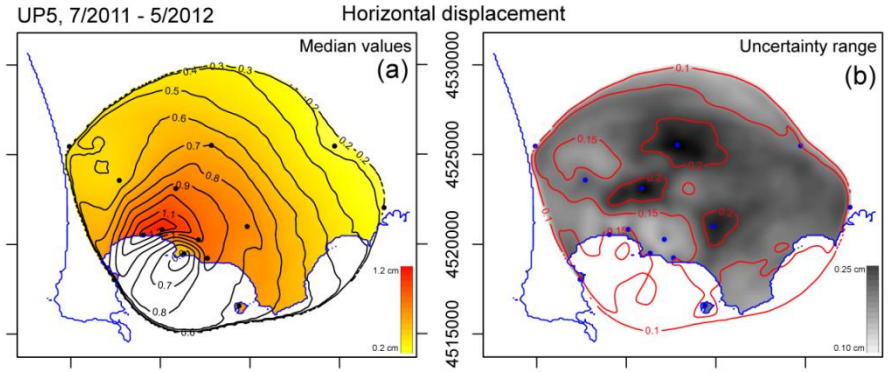
Planar displacement maps

We assume that the planar (horizontal) displacement of the ground is a key input for the construction of a vent opening map, if the deformation source is related to a magmatic intrusion.

Figure. Vertical displacement boundaries and max **horizontal** displacement lines as resulting from the RIM maps above presented. Past vents locations occurred in the last 5.2 ky BP are reported as **pink stars**.







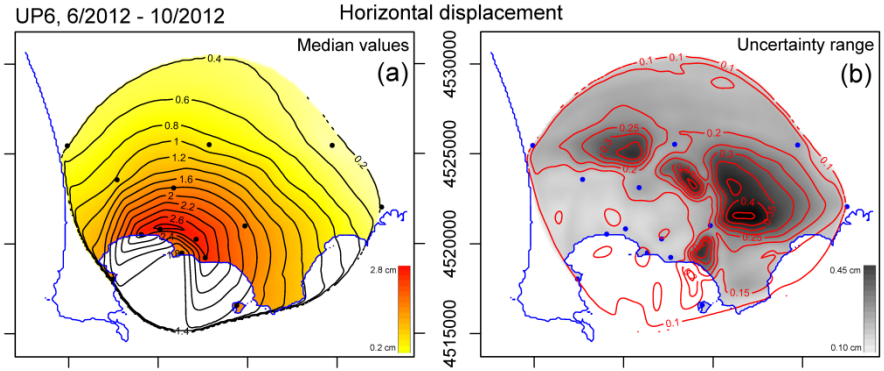


Figure. (a) maps of planar displacement, based on the GPS data of UP4-UP6. (b) Uncertainty range, given by the difference of the 5th and the 95th percentiles of error.

Posterior maps based on deformation data - I

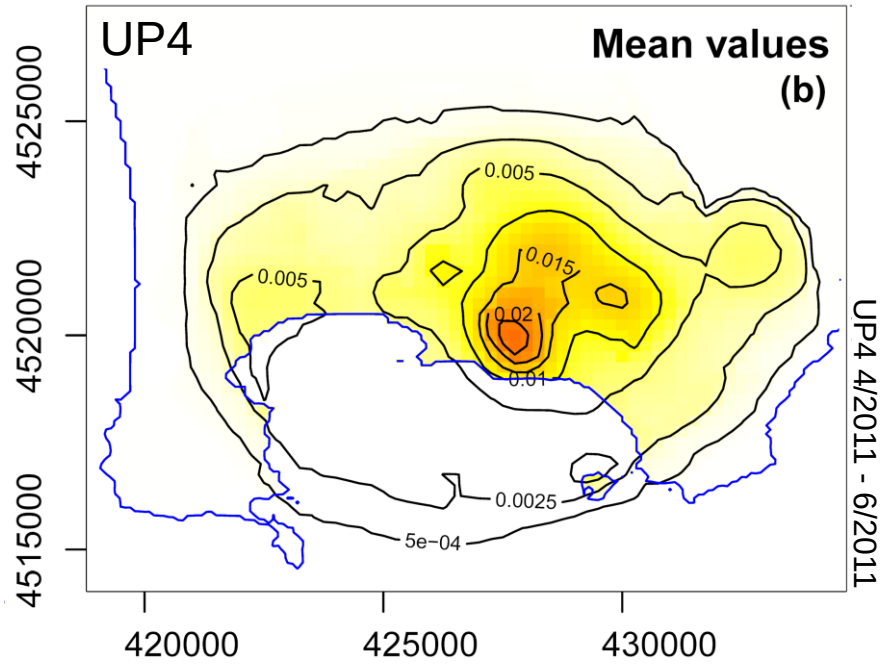
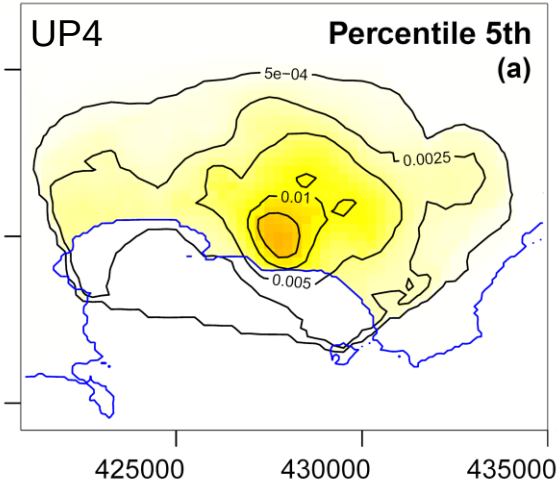
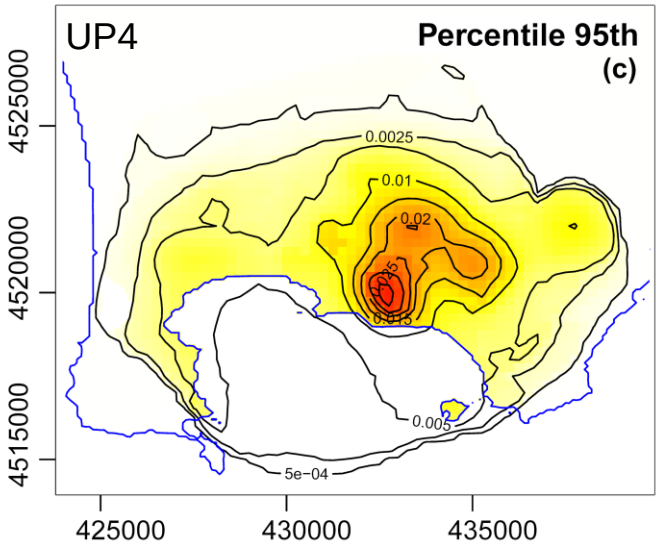


Figure. Vent opening map combining prior information from (Bevilacqua et al., 2015) and the planar displacement map related to UP4.

Here the map assumes that the deformation data consisted of real eruptive precursors.

If compared to the long-term prior, the probability concentrates in the central eastern part of the caldera with a maximum corresponding to the zone of Solfatara crater.





Posterior maps based on deformation data - II

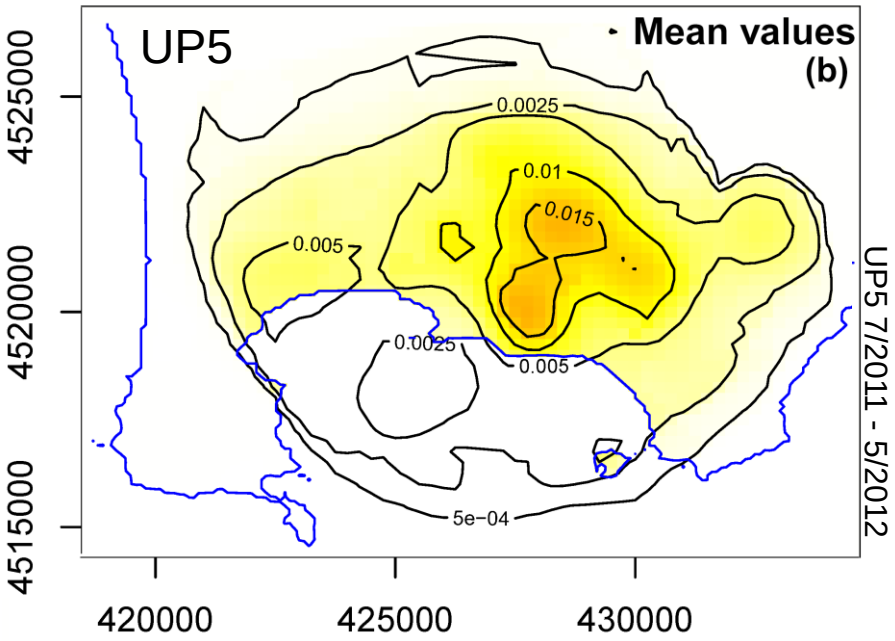
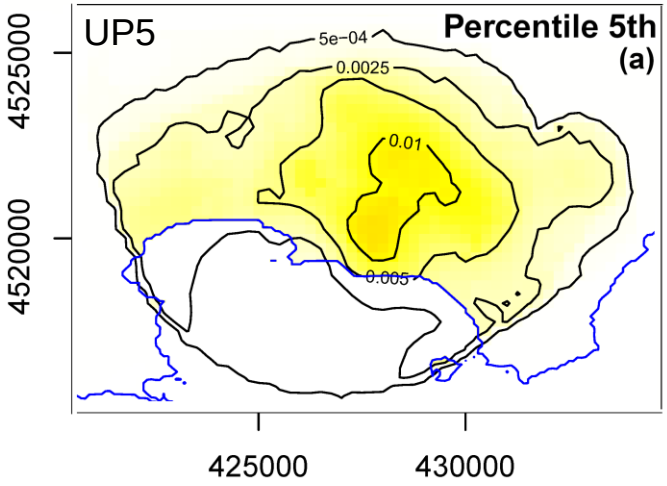
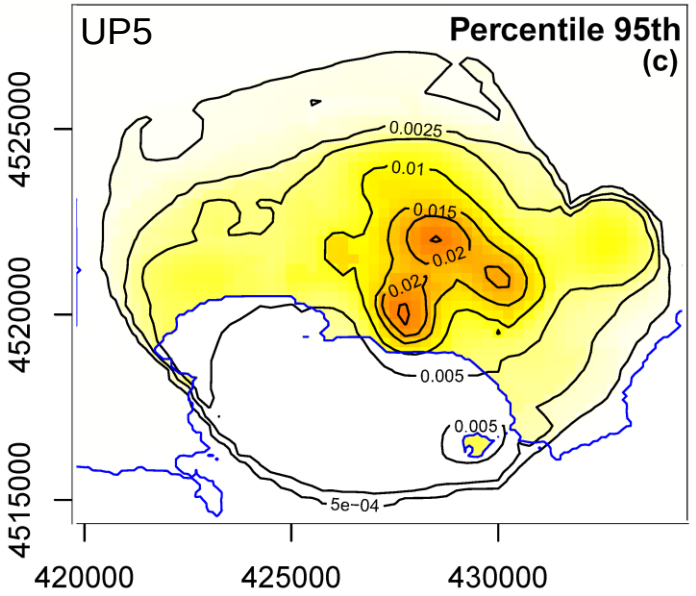


Figure. Vent opening map combining prior information from (Bevilacqua et al., 2015) and the planar displacement map related to UP5.

Here the map assumes that the deformation data consisted of real eruptive precursors.

In this case, the shape is different, with multiple zones of increased vent opening probability of a wide subregion.





Posterior maps based on deformation data - III

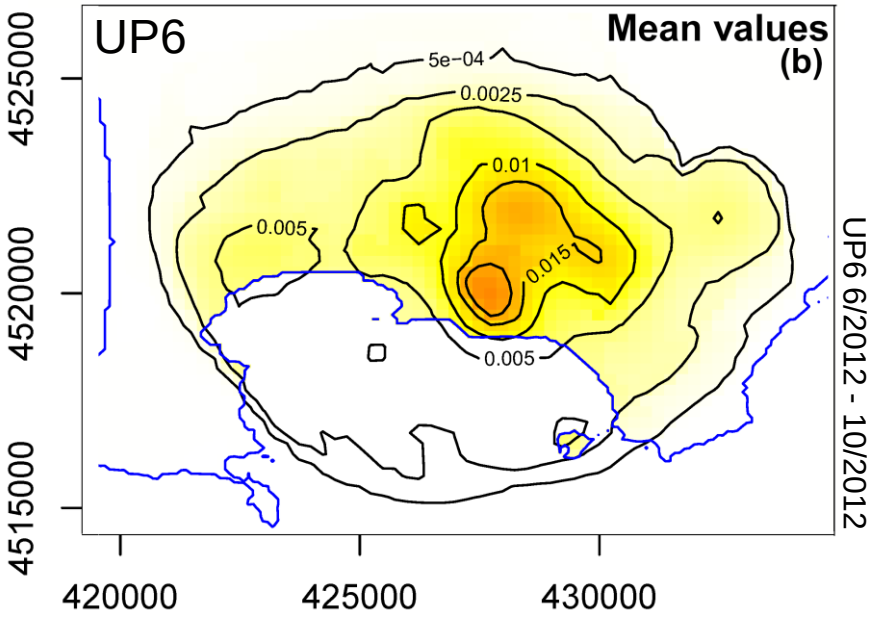
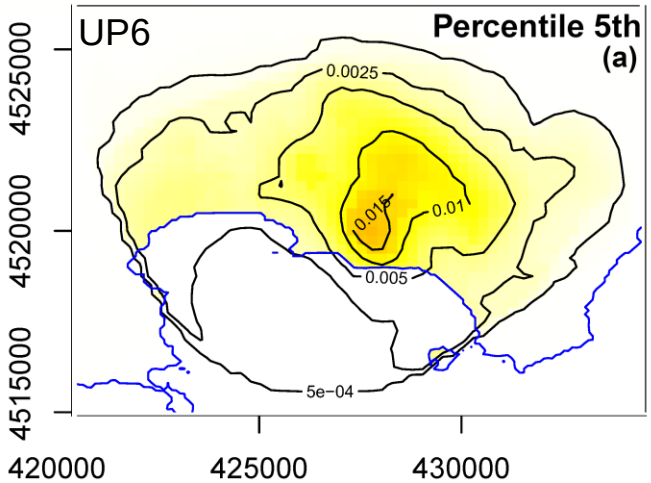
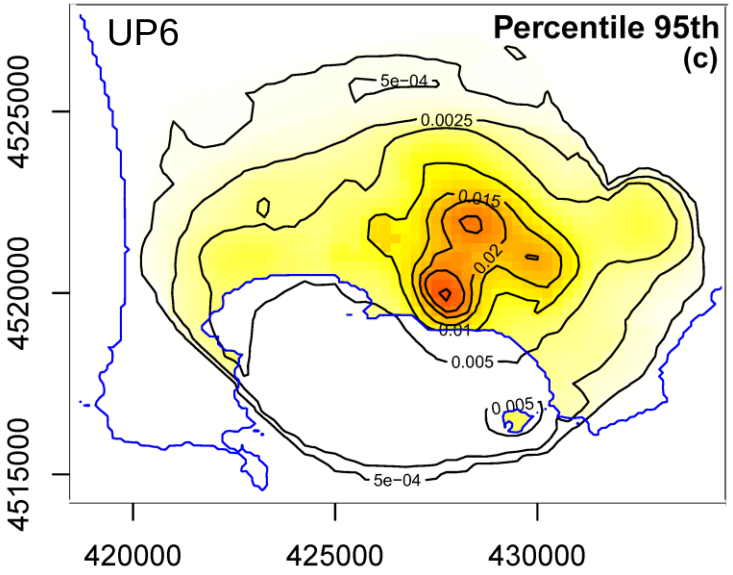


Figure. Vent opening map combining prior information from (Bevilacqua et al., 2015) and the planar displacement map related to UP6.

Here the map assumes that the deformation data consisted of real eruptive precursors.

The probability is concentrated near Solfatara crater, but less sharply than during UP4.





Summary and conclusions

In PART 1 we introduced a new method for performing **short-term eruption timing probability** forecasts, when the eruption onset is well represented by a model of a significant rupture of materials.

- The method enhanced the well known FFM equation. We allowed **random excursions** from the classical solutions. This provided probabilistic forecasts instead of deterministic predictions, giving the user critical insight into a **range of failure or eruption dates**.

- We described an assessment of failure time on **present-day unrest signals** at Campi Flegrei caldera (Italy) using either seismic count and ground deformation data. The new formulation enabled the estimation on **decade-long time windows** of data, locally including the effects of **variable dynamics**.

In PART 2 we introduced a new framework for performing **short-term eruption spatial** forecasts by assimilating precursor signals into a prior (“background”) vent opening map.

- We summarized the uncertainty affecting a **vent opening map pdf** by defining an appropriate Gaussian random field that replicates it.

- We introduced a new interpolation method based on **multiple points of central symmetry**, and we applied it on discrete GPS data collected at Campi Flegrei caldera.

- We described an application of the **Bayes’ theorem** that combines the prior vent opening map and the data-based likelihood product-wise. We provide examples based on either seismic count and interpolated ground deformation data collected at Campi Flegrei caldera.

NSF awards 1521855, 1821311, project FISR2017 MIUR, agreement DPC-INGV Annex B2 2019-21