

# Detection, classification, and location of seismovolcanic tremors with multi-component seismic data, example from the Piton de la Fournaise volcano (La Réunion, France)

Cyril Journeau<sup>1,1</sup>, Nikolai M. Shapiro<sup>2,2</sup>, Leonard Seydoux<sup>3,3</sup>, Jean Soubestre<sup>4,4</sup>, Valérie Ferrazzini<sup>5,5</sup>, and Aline Peltier<sup>3,3</sup>

<sup>1</sup>Institut des Sciences de la Terre (ISTerre)

<sup>2</sup>Institut Physique du Globe de Paris

<sup>3</sup>Institut De Physique Du Globe De Paris

<sup>4</sup>Instituto Volcanológico de Canarias

<sup>5</sup>IPGP

November 30, 2022

## Abstract

We apply three different methods based on the analysis of the multi-component seismic data to detect seismovolcanic tremors and other seismovolcanic signals, to propose an approach to classify them and to locate their sources. We use continuous seismograms recorded during one year by 21 stations at the Piton de la Fournaise volcano (La Réunion, France). The first method allows to detect seismovolcanic signals based on stability in time of the inter-components cross-correlations function. Two other methods based on the simultaneous analysis of the whole network can be used to detect seismovolcanic signals and to locate their sources. In a second approach, the seismic wavefield is analyzed by calculating the width of the network covariance matrix eigenvalue distribution. The third method consists in performing the 3D back-projection of the inter-stations crosscorrelations in order to calculate the network response function. Simultaneous analysis of the parameters measured by the three different methods can be used to classify different types of seismovolcanic tremors. Our results demonstrate that all three methods efficiently detect seismovolcanic tremors accompanying the 2010 eruptions and the preceding pre-eruptive seismic swarms. Furthermore, methods 2 and 3 based on simultaneous analysis of the whole network detect a large number of volcanic earthquakes. Our location results show that each seismovolcanic tremor is located in a distinct region of the volcano, close to the eruptive site at a shallow depth and the preceding seismic crisis is located deeper at about the sea level under the summit crater.

# Detection, classification, and location of seismovolcanic signals with multi-component seismic data, example from the Piton de la Fournaise volcano (La Réunion, France)

5 **Journeau, C.<sup>1,2</sup>, N.M. Shapiro<sup>1,3</sup>, L. Seydoux<sup>1</sup>, J. Soubestre<sup>4</sup>, V. Ferrazzini<sup>5,6</sup>  
and A. Peltier<sup>5,6</sup>**

<sup>1</sup>Institut des Sciences de la Terre, Université Grenoble Alpes, CNRS (UMR5275), Grenoble, France.

<sup>2</sup>Département de Géosciences, Ecole Normale Supérieure, PSL Res. Univ., Paris, France.

<sup>3</sup>Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, Russia.

10 <sup>4</sup>Instituto Volcanológico de Canarias (INVOLCAN), Granadilla de Abona, Tenerife, Canary Islands, Spain.

<sup>5</sup>Université de Paris, Institut de physique du globe de Paris, CNRS, F-75005 Paris, France.

<sup>6</sup>Observatoire Volcanologique du Piton de la Fournaise, Institut de Physique du Globe de Paris, Université  
de Paris, CNRS (UMR 7154), La Réunion, Paris, France.

## Key Points:

- 15
- Use of inter-components cross-correlations methods to analyze seismovolcanic tremors at Piton de la Fournaise volcano
  - Detection and classification of the three seismovolcanic eruptive tremors occurring in 2010
  - Location of seismovolcanic tremor and VT sources and highlighting of a pre-eruptive
- 20 seismic sources migration from the sea level to the eruptive site

---

Corresponding author: Cyril Journeau, [cyril.journeau@univ-grenoble-alpes.fr](mailto:cyril.journeau@univ-grenoble-alpes.fr)

**Abstract**

We apply three different methods based on the analysis of the multi-component seismic data to detect seismovolcanic tremors and other seismovolcanic signals, to propose an approach to classify them and to locate their sources. We use continuous seismograms recorded during one year by 21 stations at the Piton de la Fournaise volcano (La Réunion, France). The first method allows to detect seismovolcanic signals based on stability in time of the inter-components cross-correlations function. Two other methods based on the simultaneous analysis of the whole network can be used to detect seismovolcanic signals and to locate their sources. In a second approach, the seismic wavefield is analyzed by calculating the width of the network covariance matrix eigenvalue distribution. The third method consists in performing the 3D back-projection of the inter-stations cross-correlations in order to calculate the network response function. Simultaneous analysis of the parameters measured by the three different methods can be used to classify different types of seismovolcanic tremors. Our results demonstrate that all three methods efficiently detect seismovolcanic tremors accompanying the 2010 eruptions and the preceding pre-eruptive seismic swarms. Furthermore, methods 2 and 3 based on simultaneous analysis of the whole network detect a large number of volcanic earthquakes. Our location results show that each seismovolcanic tremor is located in a distinct region of the volcano, close to the eruptive site at a shallow depth and the preceding seismic crisis is located deeper at about the sea level under the summit crater.

**Index terms:**

- 0520 - Data analysis: algorithms and implementation
- 4302 - Geological natural hazards
- 7280 - Volcano seismology
- 8419 - Volcano monitoring

**Keywords:**

- Geophysics
- Volcano seismology
- Seismovolcanic tremors
- Cross-correlations methods
- Network-based methods

- Piton de la Fournaise volcano

## 1 Introduction

Volcanic activity is generally accompanied by a large spectrum of seismic signals from pre-eruptive seismic swarms to seismovolcanic co-eruptive tremors (e.g., McNutt & Roman, 2015). The generation of seismic waves in volcanoes might be related either to the release of mechanical stresses within the volcanic edifices and underlying crust (e.g., Roman & Cashman, 2006) or to the movement of magma and related hydrothermal fluids and to their complex interactions with solid rock (e.g., Chouet, 1996). The interpretation of these signals and the quantification of their source mechanisms allow us to better understand the oscillatory behavior of volcanic systems and provide information about the underlying physical processes. Thus, seismovolcanic data play a fundamental role in observing and understanding ongoing processes within active volcanic systems and are critical for monitoring volcanoes and anticipating their eruptive behavior.

Collection of the seismological data on volcanoes is rapidly advancing because of the modern technological developments. More and more volcanoes are equipped with networks composed of several very sensitive and broadband seismographs recording continuous data and transmitting them in real time to volcano observatories and more broadly. Archived continuous data sets and real-time data streams from several volcanoes are available through main international data centers and data collecting systems such as IRIS (<https://ds.iris.edu/ds/nodes/dmc/>), OFFEUS/EIDA (<http://www.orfeus-eu.org/data/eida/>), etc. This rapid increase in data quality, quantity, and availability opens new opportunities for the seismovolcanic monitoring.

At the same time, a full exploration of information contained in seismovolcanic data is a very challenging problem. Many of the methods used in today's volcano seismology are inherited from the past and were initially developed for a manual or weakly automated analysis of very light data streams. Single-station and/or event counting approaches remain dominantly used in the seismic volcano monitoring. These approaches have a significant advantage of being simple and robust making their implementation possible in the majority of volcano observatories suffering from the lack of resources and infrastructure. At the same time, even with massive implementation of these simple methods, the information contained in the seismovolcanic data remains very poorly used.

A full and coherent exploration of the information provided by modern seismovolcanic monitoring facilities requires a more systematic usage of advanced network-based methods. Most advanced volcanic observatories are developing computing infrastructure that allows "in situ" implementation of such methods. In a case of less well equipped observatories, the data streams can be transmitted to larger regional or national (and eventually international) data analysis centers where advanced algorithms required significant computing power can be applied. The results of such advanced seismological analysis can be then re-transmitted back to the observatories to be incorporated with other collected data and information for evaluation of the state of the volcanoes.

Network-based seismological methods are recognized as having a large potential for improving volcano monitoring because of their capacity to detect signals different from "regular" earthquakes and to locate/characterize their sources. The main idea of the network-based analysis is, instead of trying to detect signals at individual records, to characterize the level of spatial coherence of a wavefield as it is recorded by the ensemble of instruments. As a result, they do not generate discrete catalogs of events but continuous functions representing the properties of the recorded wavefields and, in favorable conditions, can be "back-propagated" to characterize the emitting sources.

Network-based methods are particularly well adapted for analysis of seismovolcanic tremors (e.g., McNutt & Nishimura, 2008) that are part of the long-period (LP) seismicity (Chouet, 1996). These ground vibrations with a sustained amplitude lasting from minutes to months, characterized by a low-frequency content (0.5 - 5 Hz) accompany eruptions and are often synonymous with the arrival of magma at the subsurface and the first signs of degassing. Therefore, they are a key element in volcano monitoring and a thorough study of these signals could allow us to bring new information about volcanic systems and to develop new methods of eruption forecasting. They can be caused by several volcanic processes including magma-hydrothermal interactions and magmatic degassing (e.g., Chouet & Matoza, 2013) but their physical origin is poorly understood and the study of their source is part of the current challenges of the modern volcanic seismology.

The network-based methods can be also useful in characterizing volcanic earthquakes. In some cases, their time resolution can be enhanced to resolve individual events. Alternatively, smoothed network-response functions can be considered as proxies to the properties of the swarms composed of many earthquakes and to infer their approximate source locations.

115 Network signal analysis is a very large family of methods (Rost & Thomas, 2002) and several of them have been used in volcano seismology. Different methods are based on their own sets of assumptions, resulting in different network-response functions with different detection capabilities and resolutions (in space, time, and frequency domains). Simultaneous application of different network-based methods can provide us with complementary  
120 information necessary to detect and to identify different types of volcanic activity. Testing this multi-method approach was the main motivation for our present work. We focused on three different methods and looked to applied them to a dataset from a volcano with well known activity. For this reason, we selected Piton de la Fournaise (PdF) volcano (La Reunion island, France). This volcano is one of the most active in the world and is systematic-  
125 ally observed by the Observatoire Volcanologique du Piton de la Fournaise (OVPF) with different monitoring techniques and, in particular, with a modern network of broadband seismographs.

The first of the tested methods is new (has been developed in this study) and explores the stability of inter-components cross-correlations to detect tremors. In a simplest case,  
130 it can be used in a single-station mode. The two other methods are existing inter-stations cross-correlations network-based approaches described by Seydoux, Shapiro, De Rosny, et al. (2016) and by Droznin et al. (2015). The latter has been modified here to be computed on a 3D grid. We compare these three methods in term of volcano-seismic signals detection. We also show how the simultaneous analysis of the parameters measured by the three different  
135 methods can be used to separate the different types of detected signals, which suggests a potential application to classify them. Moreover we will also present locations of sources detected by the network-based methods. The location method is from Soubestre et al. (2019) but here the temporal resolution is enhanced to study the pre-eruptive seismic crises. Thus, another motivation of our study is to show how these methods can be applied to follow the  
140 pre-eruptive seismic processes and seismovolcanic tremor activity through time and space.

We analyze the data recorded by the PdF seismic network during the 2010 year when three eruptions occurred (see Figure 1 for the positions of eruptive fissures and associated lava flows):

- **the January eruption** that lasted 11 days from 02.01.2010 to 12.01.2010 with the  
145 eruptive fissure located on the west part of the summit cone, inside the Dolomieu crater

- **the October eruption** that lasted 17 days from 14.10.2010 to 31.10.2010 with the eruptive site located in the south flank of the volcano
- **the December eruption** that lasted about 15 hours between the 09.12.2010 and the 10.12.2010 with the eruptive site located in the northwestern part of the caldera.

150

These three eruptions were preceded by pre-eruptive seismic crises and accompanied by seismovolcanic tremors starting at the onset of the eruption. Examples of seismograms and of real-time seismic amplitude measurements (RSAM) (Endo & Murray, 1991) during the January and October eruptions are shown in Figure S1.

155

Piton de la Fournaise volcano and its seismic network are introduced in section 2. Then, a rapid description of its seismovolcanic tremors is made in section 3. The methods applied in this study are described in section 4, with the obtained results presented in section 5 and discussed in section 6.

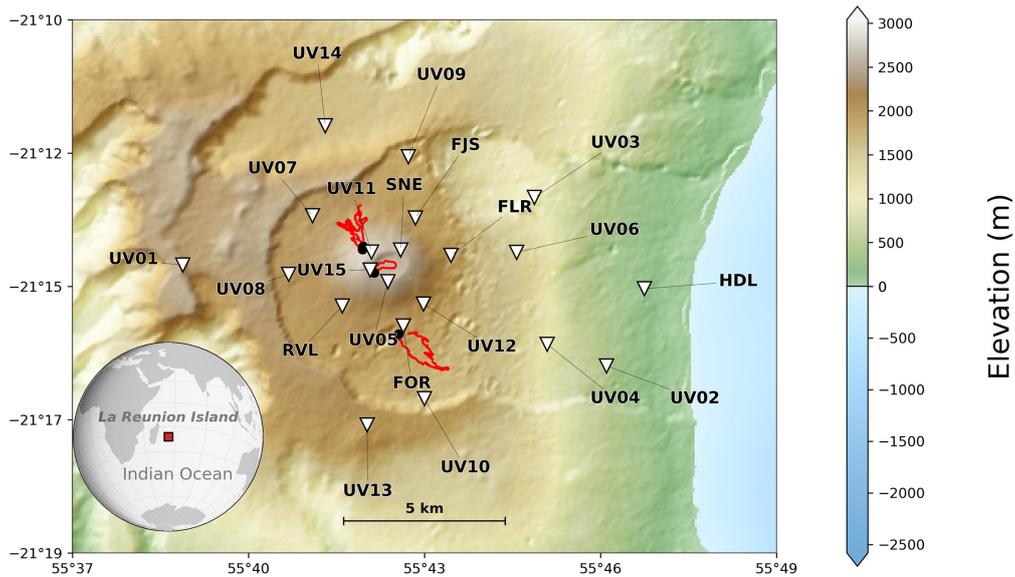


Figure 1: Color shaded relief map of the Piton de la Fournaise volcano, located on the southeast part of La Réunion island. The bottom left inset shows the location of La Réunion island in the Indian Ocean. Seismic stations used in this study are represented with white inverted triangles. Black dots stand for the position of the eruptive fissures corresponding to the January, October and December 2010 eruptions, located at the west part of the summit cone, south flank and northwest flank respectively. Red lines stand for the lava flows associated with these eruptions.

## 2 Piton de la Fournaise volcano and the seismic monitoring system

160 La Réunion is a volcanic island located in the southwest of the Indian ocean in the  
Marscarene Basin (Figure 1). Based on dating and geodynamic reconstruction, its volcanism  
is thought to have a hot spot origin and being generated by the same mantle plume that  
produced the Deccan Trapp about 65 Ma ago (Courtilot et al., 1986). The head of this  
mantle plume would have been passed beneath La Réunion about 5 Ma and would be now  
165 located 300 km southwest of the island (Bonneville, 1990). Tsekhmistrenko et al. (2018)  
present a high resolution body-wave tomography of the whole mantle column beneath the  
western Indian Ocean providing a more complex view of the plume beneath La Réunion.  
Their model reveals that the upwelling shows considerable tilt in the lower mantle instead  
of being near-vertical and splits into branches near the surface.

170 The Piton de la Fournaise volcano (PdF) is an intraplate shield basaltic volcano of  
Hawaiian type located in the southeast part of the island. It is 2631 m high and active since  
about 530 ka. Throughout its existence, this volcano has been marked by the formation of  
successive concentric calderas, the latest being the Enclos Fouqué, formed about 4500 years  
ago (Bachèlery, 1981), in which there is a summit cone of 400 m high and 3 km wide at its  
175 base with two craters: the Bory crater and the Dolomieu crater, this latter being the most  
active one.

The PdF is considered as one of the most active volcanoes in the world, with approxi-  
mately one eruption occurring every 9 months in average. Thus, its very frequent eruptions  
weakly dangerous for the population (97% of the recent eruptions occurred within the unin-  
habited caldera) make it an ideal natural laboratory to study active volcanic processes.  
180

Geodetic and seismic data suggest the presence of a shallow magma reservoir around  
2500  $\pm$  1000 m below the PdF summit craters (Staudacher et al., 2016). However, the exact  
geometry of storage zones and their possible inter-connections are still debated. While Lénat  
and Bachèlery (1990) suggested the presence of small magma pockets scattered between the  
summit and the sea level, Peltier et al. (2008) inferred from geodetic data a single magma  
185 reservoir located above the sea level, regularly recharged by a deeper source.

Since the installation of the volcano observatory of PdF (Observatoire Volcanologique  
du Piton de la Fournaise - OVPF) in 1979, the seismic network has evolved from a small  
number of short period stations recording on a triggered basis, to a modern network com-

190 bining broadband and short period sensors and recording continuously (Battaglia et al.,  
2016). In 1999, there were 19 short period stations in the network, mostly installed near  
the summit and around the volcano. In May 1999, the continuous recording appeared with  
storage on CD-ROMs and the network was completed after the 2007 Dolomieu crater col-  
lapse, including up to six broadband 3-component stations by the middle of 2009. Since  
195 October 2009, 15 broadband stations located on the volcano were installed as part of the  
Understanding Volcano project (UnderVolc) (Brenquier et al., 2012; Brenquier, 2014), in  
addition to the existing seismic network. Figure 1 shows the 21 broadband 3-component  
stations present in 2010 that we use in this study.

### 3 seismovolcanic tremor of Piton de la Fournaise volcano

200 The PdF seismovolcanic tremors are generally co-eruptive, starting when the eruption  
begins, and disappearing at the end. It is characterized by a relatively broad spectrum  
around 0.5-10 Hz (Battaglia & Aki, 2003).

The signals generated by seismovolcanic tremors are generally emergent, without any  
clear onset. This implies a lack of seismic phase identification and prevents the use of tra-  
205 ditional strategies for seismovolcanic tremor source location. Therefore this is necessary  
to develop new methods better suited for its analysis. Battaglia and Aki (2003) proposed  
a method based on seismic amplitudes for locating events. They showed that once cor-  
rected for station sites effects using coda site amplification factors, the spatial distribution  
of amplitudes shows smooth and simple contours for events including LP events and seismo-  
210 vulcanic tremor. Following this approach, Battaglia et al. (2005b) located sources of several  
co-eruptive seismovolcanic tremors at the PdF and found that they better correlated with  
the position of the eruptive vents for the 5-10 Hz band, suggesting a generation directly  
located at the eruptive site. While for frequency above 1.5 Hz, the seismovolcanic tremor  
sources are generally found at shallow depth, they suggest that lower frequencies could be  
215 related to deeper processes rather than directly by the eruption observed at the surface.  
Taisne et al. (2011) used a similar method to study the PdF pre-eruptive seismic crises.

## 4 Methods

Most of volcano observatories monitor seismovolcanic tremor in real-time based on  
measuring their amplitudes at a single reference station located near the volcano (e.g., Endo  
220 & Murray, 1991). The idea of this approach is that the position of the seismovolcanic tremor

source remains more-or-less stable in time and that the amplitude of the seismic signal can be considered as a proxy of the intensity of the seismovolcanic tremor generating processes. This single station/component approach can be developed further with using the spectral shape of the signal (instead of its averaged amplitude) leading to a better discrimination of different phases of seismovolcanic tremor (Unglert & Jellinek, 2015; Unglert et al., 2016). One of the main difficulties of single component approaches is that they cannot be used to get information about the spatial location and extent of the source. In particular, this might result in difficulties with distinguishing seismovolcanic tremors from other types of processes leading to increased seismic amplitudes such as, for example, meteorological and anthropogenic noises.

In this paper we systematically use and compare methods based on correlations of signals to perform detection and location of seismovolcanic signals. We first briefly describe the data pre-processing used in all analyses considered in this paper. Then, we introduce a single-station method aimed at detecting long-acting sources based on stability of inter-components cross-correlations. In a next step, we consider methods based on simultaneous analysis of the whole network of receivers. Our approach is different from the antenna-type detection of seismovolcanic tremors based on small-aperture arrays and plane-wave approximation for measuring slowness and azimuth of incoming wave (e.g., Goldstein & Chouet, 1994; Métaixian et al., 2002; Haney, 2014). We rather consider a geometry where the source of seismic radiation is located within the network. It implies that the moveouts are not described with plane waves but with a wave propagation from a point source (cylindrical wave in a case of surface waves and ray-predicted times for body waves). In this case, it has a strong imprint on inter-stations cross-correlations (e.g., Shapiro et al., 2006; Ballmer et al., 2013; Droznin et al., 2015). In volcanic environments these cross-correlations can be used to detect seismovolcanic signals and to locate their source. Here we apply two network-based methods. First, we use the representation of ensemble of inter-stations cross-correlations in terms of network covariance matrix suggested by Seydoux, Shapiro, De Rosny, et al. (2016); Seydoux, Shapiro, de Rosny, and Landes (2016); Seydoux et al. (2017) and used by Soubestre et al. (2018, 2019) to study seismovolcanic tremors. Second, we adapt on a 3D grid the approach of Droznin et al. (2015) that backproject the cross-correlations envelopes in time using a 1D S-wave velocity model to detect seismovolcanic signals and we follow the approach described in Soubestre et al. (2019) to locate their sources.

We are focusing on the information contained in the time-variations of the signal phase differences between stations. We also consider the instrument response of the seismic stations to be stable over time and, therefore, do not consider instrument correction. We then further apply the pre-processing usually performed in ambient-noise interferometry (e.g., Shapiro & Campillo, 2004; Shapiro et al., 2005; Bensen et al., 2007).

The data are stored in 24-hours-long time-series in SAC down-sampled from 100 to 20 samples per second. We apply demeaning, linear detrending and bandpass filtering between 1 and 10 Hz. Then, a spectral whitening is applied followed by a temporal normalization according to Bensen et al. (2007).

#### 4.1 Single-station inter-components cross-correlations

We present here a new method to detect some seismovolcanic signals including seismovolcanic tremors based on the stability of the inter-components cross-correlations at a single station. Let us consider a wavefield generated by a seismovolcanic tremor source that remains at the same location and with a constant source mechanism. As argued by Droznin et al. (2015); Soubestre et al. (2018), the cross-correlation of such a wavefield between two receivers will remain stable in time and can be used as a "fingerprint" of this particular source. This principle also holds for a cross-correlation between two components of the same receiver and we use this property to design a simple single-station seismovolcanic tremor detector. We start with computing inter-components cross-correlations  $CC_{i,j}$  ( $i, j = E, N, Z$ ) in moving windows of length  $\Delta t$  and shifted by  $\Delta t/2$ . We then compute correlation coefficients between cross-correlation waveforms from consecutive time windows and take the mean value of 6 consecutive computations. The obtained functions  $cc_{i,j}^6$  provides us with estimation of the stability of the cross-correlation waveforms that, in turn, is related to existence of stable seismovolcanic tremor source. In our case we use moving windows of length  $\Delta t = 200$  s to compute inter-components cross-correlations and therefore the  $cc_{i,j}^6$  functions correspond to a short-term estimation of the cross-correlation waveforms stability over 700 s. The  $cc_{i,j}^6$  function is represented in the part 5.1. This method allows us to detect seismovolcanic tremor activity by using only one components pair of a single seismic station. In order to improve the seismovolcanic tremor detection results, we can first average the three  $cc_{i,j}^6$  functions obtained from the three components pairs of one single station, and then average the previous functions on several stations.

## 4.2 Inter-station single-component cross-correlations: covariance matrix spectral width

285

290

295

300

We use the method proposed by Seydoux, Shapiro, De Rosny, et al. (2016) where an ensemble of inter-stations cross-correlations is computed in the frequency domain (also called network covariance matrix). This matrix is decomposed on the basis of its eigenvectors associated with real positive eigenvalues. The main idea behind this method is to consider that the rank of the covariance matrix (number of non-zero eigenvalues) is related to the number of independent signals composing the wavefield and then of independent seismic sources. A single source would generate a rank 1 covariance matrix, and this rank would increase with the number of independent sources. However, because of the presence of seismic noise (generated by ocean micro-seismic, wind, etc) and instrumental electronic noise in the data, it is complicated to estimate the effective rank of the covariance matrix and the number of independent sources can not be directly inferred. In addition, as the covariance matrix is estimated from a finite number of windows, this may reduce the ability to capture the independent seismic signals. Nevertheless, the width of the covariance matrix eigenvalues spectrum ( $\sigma(f)$  called spectral width) can be used as a proxy for the number of independent seismic sources composing the wavefield (Seydoux, Shapiro, De Rosny, et al., 2016). Thus, this spectral width is high when it corresponds to ambient seismic noise produced by distributed sources, while it is low when it corresponds to a spatially coherent signal produced by a single localized source, such as seismovolcanic tremor and volcanotectonic earthquakes (VTs).

Following Seydoux, Shapiro, De Rosny, et al. (2016), we start by cutting the daily traces into a number of overlapping time windows that we call averaging windows. Then, each averaging window is cut into  $M$  overlapping subwindows of length  $\delta t$  such that  $\Delta t = Mr\delta t$ , where  $\Delta t$  is the length of the averaging time window and  $r$  the overlapping ratio that is 0.5 on our case. We apply a taper on the data prior to calculate the Fourier transform on the subwindow to avoid sharp windows, that is why the signal segments are overlapped with a factor of 50 % in order not to lose any information located in consecutive windows. We then collect the array data vector from the  $N$  seismic stations  $\mathbf{u}(f)$  as  $\mathbf{u}(f) = [u_1(f), u_2(f), \dots, u_N(f)]^T$  and compute the outer product  $\mathbf{u}(f)\mathbf{u}^\dagger(f)$  within each subwindow, where  $\dagger$  stands for the Hermitian transpose. Then the covariance matrix  $\mathbf{C}(f)$

is estimated on any averaging window from  $M$  consecutive subwindows

$$\mathbf{C}(f) = \langle \mathbf{u}(f)\mathbf{u}^\dagger(f) \rangle_{\Delta t} = \frac{1}{M} \sum_{m=1}^M \mathbf{u}_m(f)\mathbf{u}_m^\dagger(f), \quad (1)$$

305 with  $\mathbf{u}_m(f)$  the data Fourier spectra vector in the subwindow  $m$ . Note that the inverse Fourier transform of equation 1 is the inter-station single-component cross-correlations estimated over the  $M$  subwindows.

This calculation can be extended on consecutive averaging window, over the whole year 2010 with  $M = 50$  and  $\delta t = 40$  s. We note that  $\delta t$  must be long enough for the seismovolcanic tremor seismic wave to have time to propagate through the entire network. We thus obtain 1000 s long windows that overlap with a 50 % factor. The obtained covariance matrix is Hermitian and positive semi-definite, so it can be decomposed on the basis of its complex eigenvectors  $\mathbf{v}_n$  associated with real positive eigenvalues  $\lambda_n$ :

$$\mathbf{C}(f) = \sum_{n=1}^N \lambda_n(f)\mathbf{v}_n(f)\mathbf{v}_n^\dagger(f). \quad (2)$$

Then eigenvalues are arranged in decreasing order and the covariance matrix spectral width is computed as a function of frequency as:

$$\sigma(f) = \frac{\sum_{i=1}^N (i-1)\lambda_i(f)}{\sum_{i=1}^N \lambda_i(f)}. \quad (3)$$

### 4.3 Inter-station single-component cross-correlations: network response and source location

#### Detection of seismovolcanic signals from full rank inter-stations

310

**cross-correlations matrix:** We follow the approach of (Droznin et al., 2015) who developed a 2D source-scanning algorithm based on the stacks of inter-stations cross-correlations envelopes, in order to detect seismovolcanic tremors and to locate their sources. This method is based on an assumption that seismovolcanic tremor source is relatively shallow and that seismic energy propagates toward the stations along the surface. The differential travel times of signals emerging in inter-stations cross-correlations are then interpreted in terms of surface wave (2D) propagation in order to find the geographical location of the seismovolcanic tremor source. Here we apply this method in a 3D grid in order to locate also deeper seismovolcanic sources such as VTs. The location is done by performing a 3D grid search with tested sources spaced 200 meters apart. We first compute the inter-stations cross-correlations (we only use the vertical component seismograms) between all pairs of available stations. We therefore end up with 210 pairs of cross-correlations when data from

320

the 21 stations are available. The next step consists in computing the smoothed envelopes of the cross-correlation waveforms. This is done by computing the absolute value of the analytic signal derived from the Hilbert transform and by performing a convolution with a 3 s long Gaussian window. Then, for every tested source position in our 3D grid we compute travel times for all stations using a 1D S-wave velocity model of the PdF volcano from Mordret et al. (2015) (shown in Figure S2) by considering that seismovolcanic tremors and their cross-correlations are dominated by S waves (Soubestre et al., 2019). We note that the weak S-wave velocities observed in the near-surface layers constituted of poorly consolidated volcanic materials (Figure S2) are consistent with the horizontal wave propagation velocity that we can estimate as seen in Figure S3 by performing a linear regression between the maxima of the smoothed cross-correlations envelopes. The Figure S3 represents the smoothed cross-correlations calculated between the UV05 station (the closest available station to the January eruptive center) and all the others available stations for the day 03.01.2010. We chose the UV05 station because we assume that the source of the seismovolcanic tremor is located close to the January eruptive center and its position is stable. The red line is the result of the linear regression whose slope corresponds to a horizontal wave propagation velocity of 700 m/s. We also observe similar velocities for other days with seismovolcanic tremor activity.

Then, for every point  $\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$ , each cross-correlation envelop  $S^{i,j}$  is shifted by the time difference needed for the wave to travel from the tested source to the two stations  $i$  and  $j$  and we finally compute the network response  $R(\mathbf{r})$  by stacking at zero lag time the value of the shifted envelopes for all stations pairs as:

$$R(\mathbf{r}) = \sum_{i=1}^N \sum_{j=i+1}^N S^{i,j} [t^i(\mathbf{r}) - t^j(\mathbf{r})] \quad (4)$$

with  $N$  the number of stations and  $t^i(\mathbf{r})$  the travel time between the tested source and the station  $i$ . This network response (called likelihood afterwards) characterizes the likelihood of location of a seismic source in a particular position. Finally, we calculate the normalized maximum network response called NRF afterwards for Network Response Function

$$NRF(n) = 100 \times \frac{R^{max}(n) - R^{min}(n)}{R^{ref}} \quad (5)$$

for every day of the 2010 year, where  $n$  is the day number and  $R^{ref}$  a normalization coefficient chosen as an average of  $R^{max} - R^{min}$  over the October 2010 eruption, the longest period of seismovolcanic tremor activity during the year. The function defined in such a way maximizes when most of its values are concentrated in the vicinity of the main maximum,

345 i.e., in a case of a strong localized source. The NRF is then used as a proxy for signal detection. The obtained detection results are shown in section 5.1.

### Location of seismovolcanic sources from low rank inter-stations

**cross-correlations matrix:** This location method (described in Soubestre et al. (2019)) is similar to the 3D source-scanning algorithm described in the previous part but here we focus on the dominant wavefield. To do that, we focus on the first eigenvector of the network covariance matrix containing information about the location and the mechanism of the dominating seismovolcanic tremor source. The first eigenvector represents the most coherent part of the wavefield and acts as a de-noising operator. We show in the Figure S4 a comparison between the cross-correlation functions (CCFs) from the full recorded seismic wavefield (left pannel) and the CCFs extracted from the first eigenvector filtered covariance matrix (right pannel). We note that this filtering based on covariance matrix decomposition increases the signal-to-noise ratio and we therefore use it to improve our location results as it can be seen in Soubestre et al. (2019).

The filtered covariance matrix  $\tilde{\mathbf{C}}(f)$  is obtained from the complex outer product of the first eigenvector  $v_1(f)$  with himself :

$$\tilde{\mathbf{C}}(f) = v_1(f)v_1^\dagger(f). \quad (6)$$

360 Then, the time-domain filtered cross-correlations are retrieved by performing the inverse Fourier transform of  $\tilde{\mathbf{C}}(f)$ , and smoothed to obtain their envelopes, as explained in the previous section.

After the smoothing, each cross-correlation envelope is shifted by the time difference between travel times from every tested point of the 3D grid to the two considered stations. The network responses are then obtained by stacking the value at zero lag-time of shifted cross-correlations envelopes using equation (4). In order to better visualize the location result, a normalized network response  $\tilde{R}(\mathbf{r})$  is computed

$$\tilde{R}(\mathbf{r}) = \frac{R(\mathbf{r}) - R^{min}}{R^{max} - R^{min}} \quad (7)$$

with  $R^{max}$  and  $R^{min}$  the absolute maximum and minimum of  $R(\mathbf{r})$ . The obtained location results are shown in section 5.2.

365 We tested different values of the Gaussian window width used to smooth the cross-correlations reconstructed from the first covariance matrix eigenvector. We make our tests on four hours of the day 15.10.2010 during the October seismovolcanic tremor between 15h and

19h, when inter-components cross-correlations remain stable. We thus stack the obtained cross-correlations computed in 1000 s windows during these four hours. We compute their envelopes by testing different values of smoothing Gaussian window width and normalize them by their maxima. We then perform a 3D back-projection of these obtained cross-correlations envelopes and compute the maximum likelihood value for each tested value of the smoothing Gaussian window width. We show in Figure S5 the maximum likelihood value as a function of the smoothing Gaussian window width. We chose a Gaussian window width value of 1.5 s, corresponding to the corner of the curve in Figure S5 pointed by the red dashed lines, beyond which the maximum likelihood value no longer increases significantly.

## 5 Results

### 5.1 Detection of seismovolcanic tremors and other types of seismic activity

Figure 2a shows the result of the  $cc_{i,j}^6$  function (described in section 4.3) for the N - Z components pair of station UV12. We note that we can detect the three eruptive seismovolcanic tremors (highlighted in red in the figure) only with one components pair of a single station. We can then improve these detection results by averaging the three  $cc_{i,j}^6$  functions obtained from the three components pairs and we show an example for the UV12 station in the Figure 2b. In order to improve the signal-to-noise ratio we perform an average over several stations. We show in Figure 2c the result averaged over 12 stations (36 components) whose continuous records without significant interruption are available through the whole year 2010. We represent some zoom of this station-averaged  $cc_{i,j}^6$  function in Figure 6a, 7a and 8a in which we can observe the high values associated to the seismic crises preceding the three eruptions. In addition to the seismovolcanic tremors and to their preceding seismic swarms, we can also observe high values of the function (highlighted in orange in Figure 2) at the beginning of September when an increase of the seismovolcanic activity is detected and the 23.09.2010 when a dyke intrusion occurred without leading to an eruption (Roult et al., 2012).

Figure 3a shows the covariance matrix spectral width (equation 3) computed on 1000 s long overlapping time windows for every frequency, this can be seen as a time-frequency representation of the wavefield recorded by the whole network highlighting the coherent signals. Figure 3b represents the average of this spectral width on the frequency-band 1-5

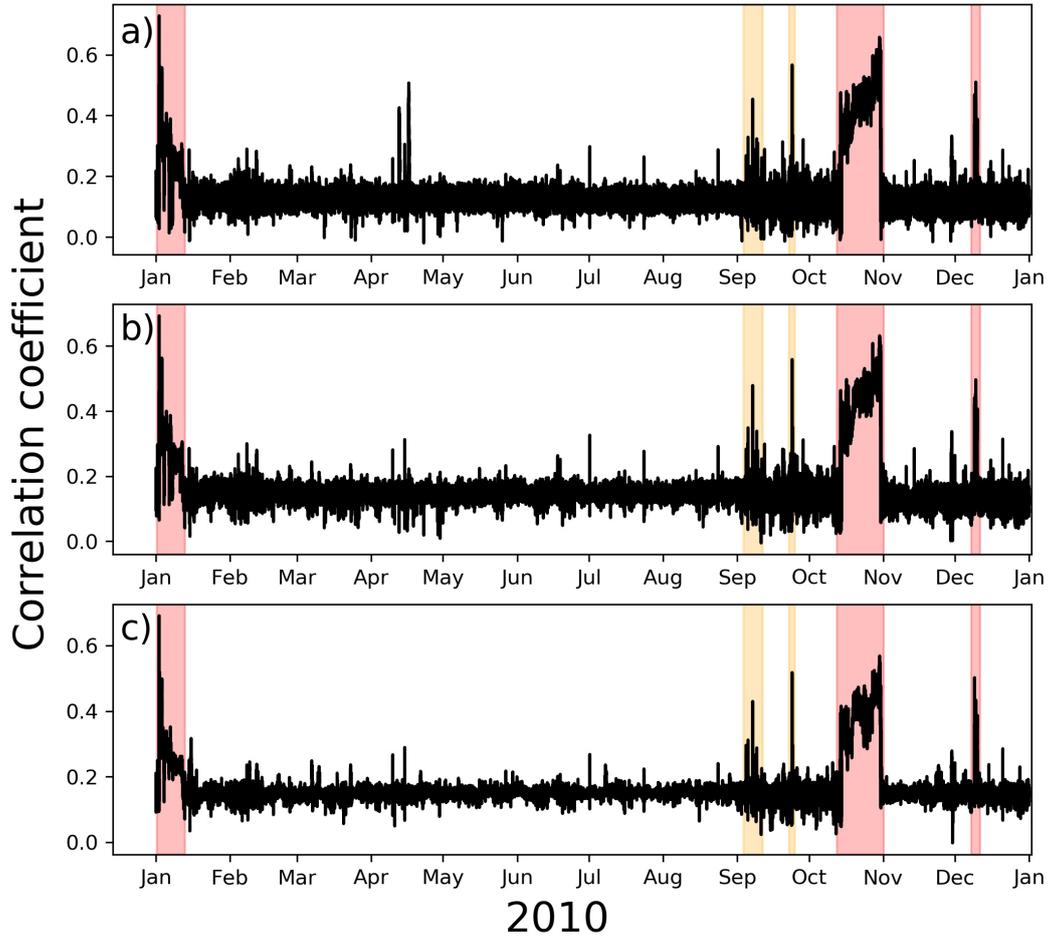


Figure 2: Seismovolcanic activity detection from inter-components cross-correlations at a single station. The three red areas highlight the January, October and December eruptions while the two orange areas highlight an increase of seismovolcanic activity at the beginning of September and seismovolcanic activity linked to a dyke intrusion the 23.09.2010. (a) Correlation coefficient function  $cc_{i,j}^6$  computed for the N - Z components pair of the UV12 seismic station. (b) Average of the three correlation coefficient functions  $cc_{i,j}^6$  obtained for the three components pairs of the UV12 seismic station. (c) Average of the averaged function represented in (b) on 12 seismic stations whose data availability is stable and identical over the 2010 year (UV01 to UV14, without UV04 and UV06). High values of this correlation coefficient function are also observed during the seismic crises preceding the three eruptions and zoom are provided in Figures 6a, 7a and 8a.

Hz. This function represents another network-based detector and all three seismovolcanic tremors can be identified with lower values of the spectral width.

Figure 3c shows the NRF (equation 5) computed every day (time resolution similar  
 400 to Droznin et al. (2015)). This allows us to detect the January and October seismovolcanic tremors (red areas in the figure). However we can not detect the December seismovolcanic tremor here because of the temporal resolution. Figure 3d shows the same function computed with a finer time resolution, i.e., in 1000 s long time windows. We can now detect the three seismovolcanic tremors and we also observe that the NRF becomes more sensitive to the  
 405 high-seismicity periods.

In addition to the seismovolcanic tremors highlighted in red, we can also detect two other periods of high seismic activity highlighted in orange in the figure. A first one from the beginning of September corresponding to an increase in the seismicity prior to the October eruption and the second one at the end of December that are followed by an increase of  
 410 the micro seismicity under the summit since the beginning of the year 2011 (OVPF, 2011). The pre-eruptive seismic crises corresponding to each eruption are also well detected (see the zoom in the Figures 6b-c, 7b-c and 8b-c).

We note that most of the high values observed in the correlation coefficient function and NRF and low values for the spectral width correspond to known volcanic activities  
 415 highlighted in red for eruptions or orange for earthquakes and injection. However the origin of others remains unidentified to date (during April for example) and it can correspond to noise or some previously undetected volcanic activity. While the computation of the NRF requires an a-priori on the medium velocity, there is no a-priori for the spectral width calculation. Despite this, the two network-based methods result in a similar first-order  
 420 information and we observe an anti-correlation when seismovolcanic tremor sources produce high value of the NRF and low value of  $\sigma(f)$ . Also, by comparing the Figures 2 and 3 we note that while the two network-based detectors catch the same seismovolcanic signals, the correlation coefficient function doesn't detect the seismicity during September (except at the beginning and during the dyke intrusion the 23.09.2010) and at the end of December.

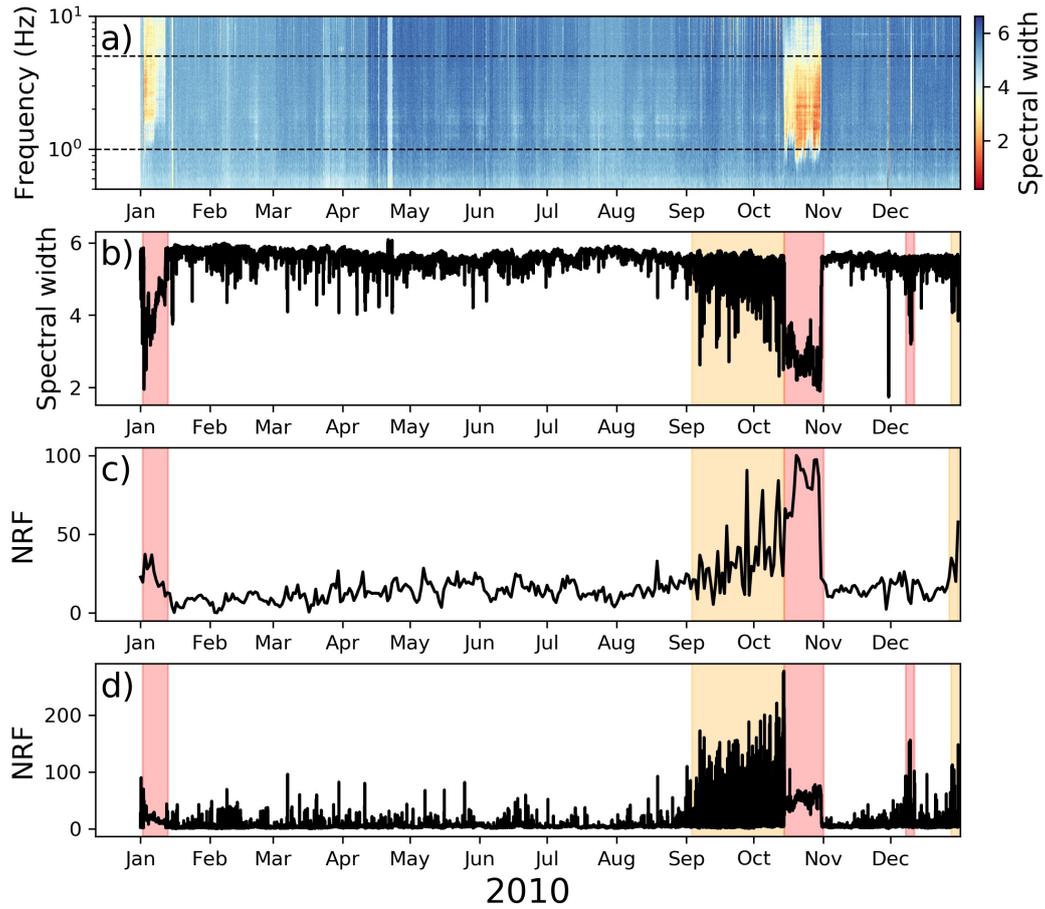


Figure 3: Seismovolcanic activity detection from inter-stations cross-correlations. The red areas highlight periods of detected seismovolcanic tremors while orange areas highlight periods of high seismic activity in September and at the end of December. (a) Covariance matrix spectral width computed for every frequency on overlapping 1000 s long time windows. The two horizontal dashed black lines stand for the lower and upper limits of the covariance matrix spectral width averaging. (b) Average of the covariance matrix spectral width in the frequency band 1-5 Hz. (c) Network response function (equation 5) computed every day. (d) Network response function computed on 1000 s long time windows. In order to limit the impact of the missing stations on the functions (a), (b) and (d), we normalize them by the number of functional stations for each day.

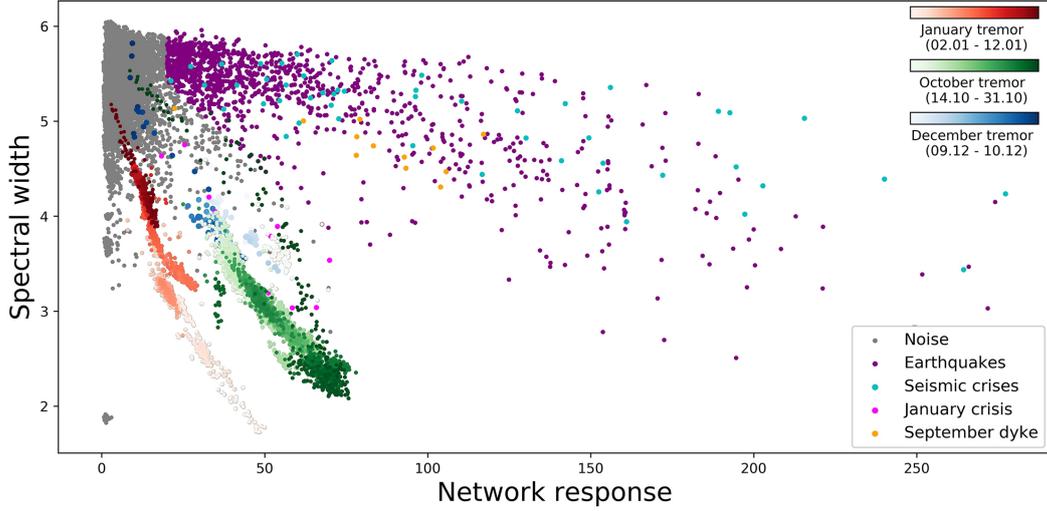


Figure 4: Average spectral width as a function of the network response function. Colors correspond to different clusters of events. The three seismovolcanic tremors are represented with color gradations from lightest to darkest in order to observe their evolution over time. Red colors stand for the January seismovolcanic tremor with corresponding spectral width and network response represented in Figures 6b and 6c respectively. Green colors stand for the October seismovolcanic tremor as represented in Figures 7b and 7c. Blue colors stand for the December seismovolcanic tremor (see Figures 8b and 8c). Grey points correspond to noise. Purple colors stand for the earthquakes occurring during the 2010 year. Cyan colors correspond to the October and December seismic crises while the January seismic crisis is represented with fuchsia colors. And the September dyke intrusion is represented by orange points.

425

A more detailed comparison of two network-based methods is shown in Figure 4 that represents the average spectral width as a function of the NRF. We can see different trends corresponding to clusters of different seismic events. This allows to clearly differentiate the January and October seismovolcanic tremors represented by red and green colors respectively, as the January seismovolcanic tremor is characterized by lower values of the network response function (see also Figures 6b, 6c, 7b and 7c). December seismovolcanic tremor can also be seen in Figure 4 with blue colors, with similar values of the network response function but higher values of the average spectral width than the October seismovolcanic tremor (see also Figure 8b and 8c). We notice that we can thus separate the summit seismovolcanic tremor (occurring in January) from the two flank seismovolcanic tremors occurring in the South and North-West flanks in October and December, respectively. We also observe a different temporal evolution for the January and October seismovolcanic tremor. January

430

435

seismovolcanic tremor is characterized by very low values of spectral width at the beginning of the eruption (very light red in Figure 4) and it evolves until reaching values equivalent to that of seismic noise (dark red in Figure 4). Conversely the October seismovolcanic tremor begins with intermediate values of spectral width (very light green in Figure 4), then it ends with very low spectral width values (dark green in Figure 4) and finally the spectral width becomes higher and higher until the eruption is completely over.

Another trend is distinguishable in Figure 4 with more scattered points represented in purple which mostly correspond to the earthquakes occurring during the 2010 year. The September dyke intrusion is represented by orange points. Cyan colors correspond to the October and December seismic crises while the January seismic crisis is represented with fuchsia colors because we see that it is characterized by different values of NRF and spectral width. Grey points in the top left corner of Figure 4, characterized by low values of network response and high values of average spectral width, correspond to seismic noise during the whole year. We also note a few points on the bottom left of the Figure 4, characterized by very low values of the average spectral width, that correspond to a few hours in the end of November where there was a technical problem and several stations were down.

Finally, we can also make a 3-D comparison between the station-averaged  $cc_{i,j}^6$  function (whose zoom during eruptions are shown in Figures 6a, 7a and 8a), the NRF and the spectral width as seen in Figure 5. Before comparing all the three functions, we re-calculated the NRF and the spectral width with 200 s long overlapping windows and we computed a moving average between 6 consecutive windows. In addition to the separation between summit and flank seismovolcanic tremors, we can also distinguish the intrusion occurring for a few hours on 23.09.2010 ((Roult et al., 2012)) represented with orange points colored in orange. Moreover, we can on the one hand distinguish the pre-eruptive seismic swarms from the others VTs colored in purple and on the other hand distinguish the January seismic swarm colored in fuchsia from the October and December seismic swarms colored in cyan.

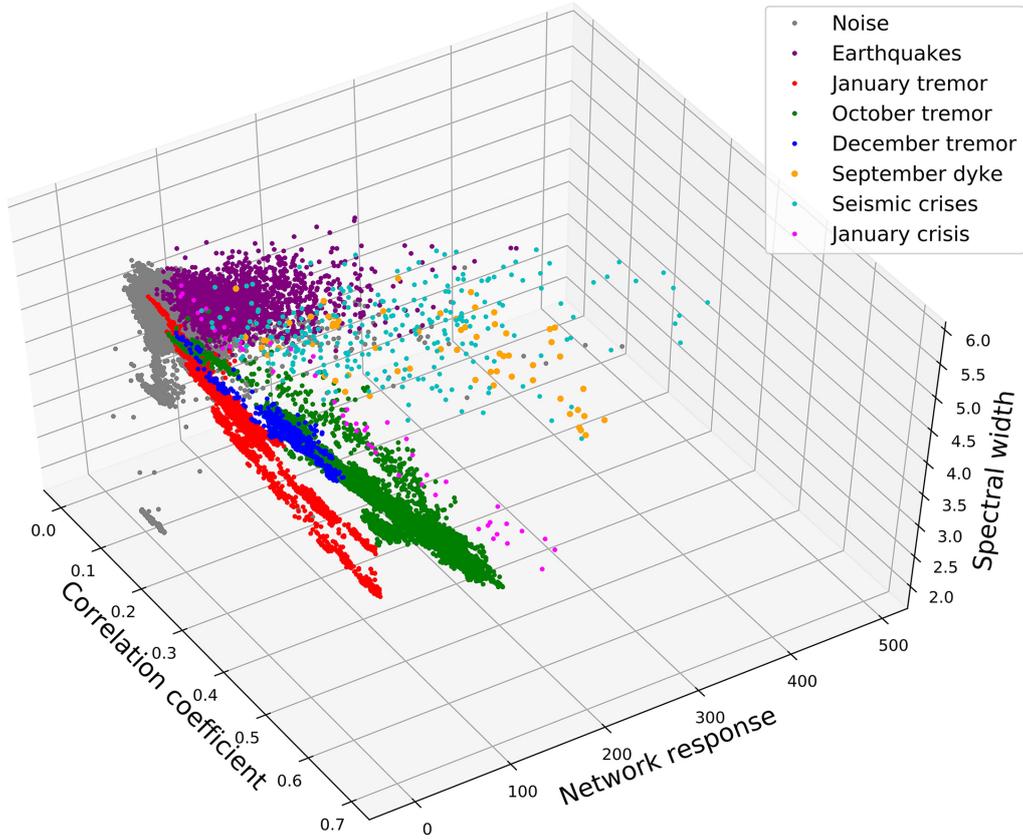


Figure 5: 3D comparison between the station-averaged  $cc_{i,j}^6$  function labeled as "correlation coefficient", the NRF and the average spectral width. Grey colors stand for the seismic noise, purple colors stand for the earthquakes. The January, October and December seismovolcanic tremors are represented by red, green and blue colors respectively. The orange color stands for the dyke intrusion occurring in September, Fuchsia colors stand for the January pre-eruptive seismic crisis and Cyan colors stand for the October and December pre-eruptive seismic crises.

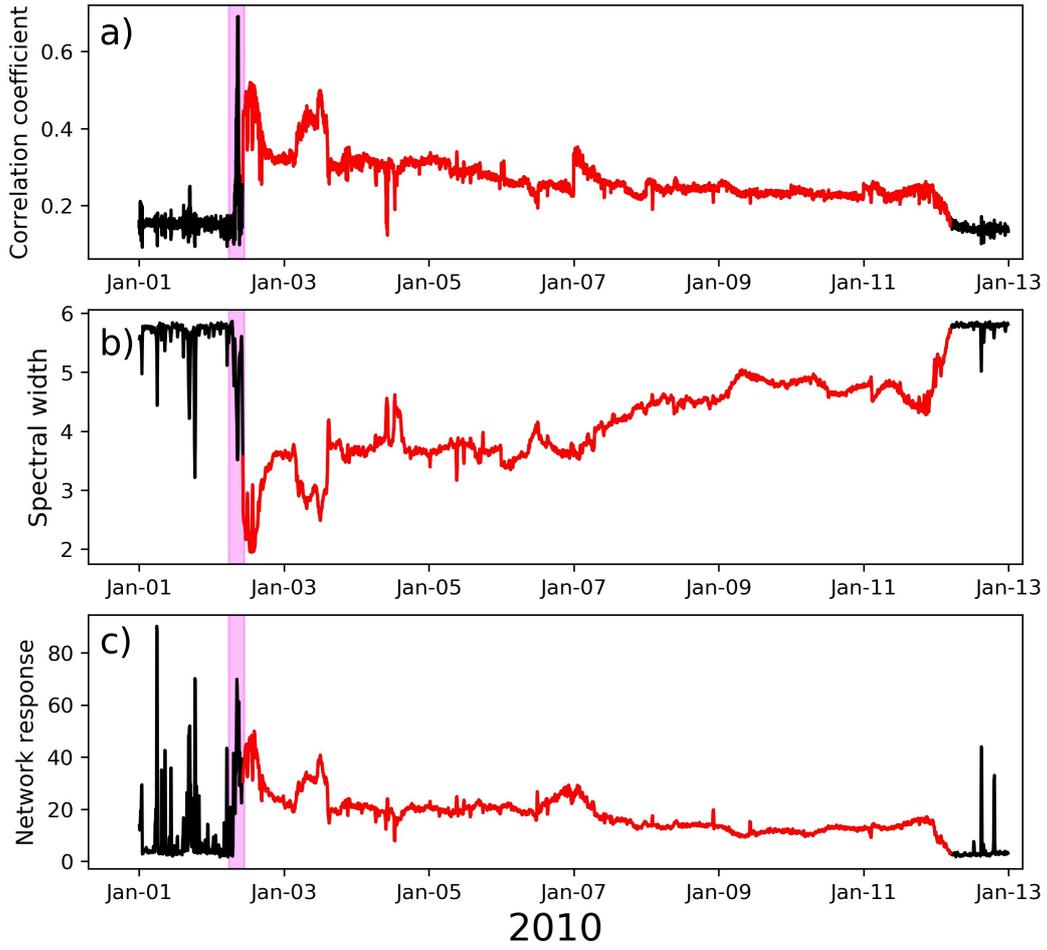


Figure 6: Zoom on the January eruption and its pre-eruptive seismic swarm for (a) the station-averaged  $c_{i,j}^6$  coefficient correlation function, (b) the spectral width and (c) the NRF. The red color stands for the January seismovolcanic tremor activity and the fuchsia color stand for the corresponding pre-eruptive seismic crisis, following the color code of the Figures 4 and 5.

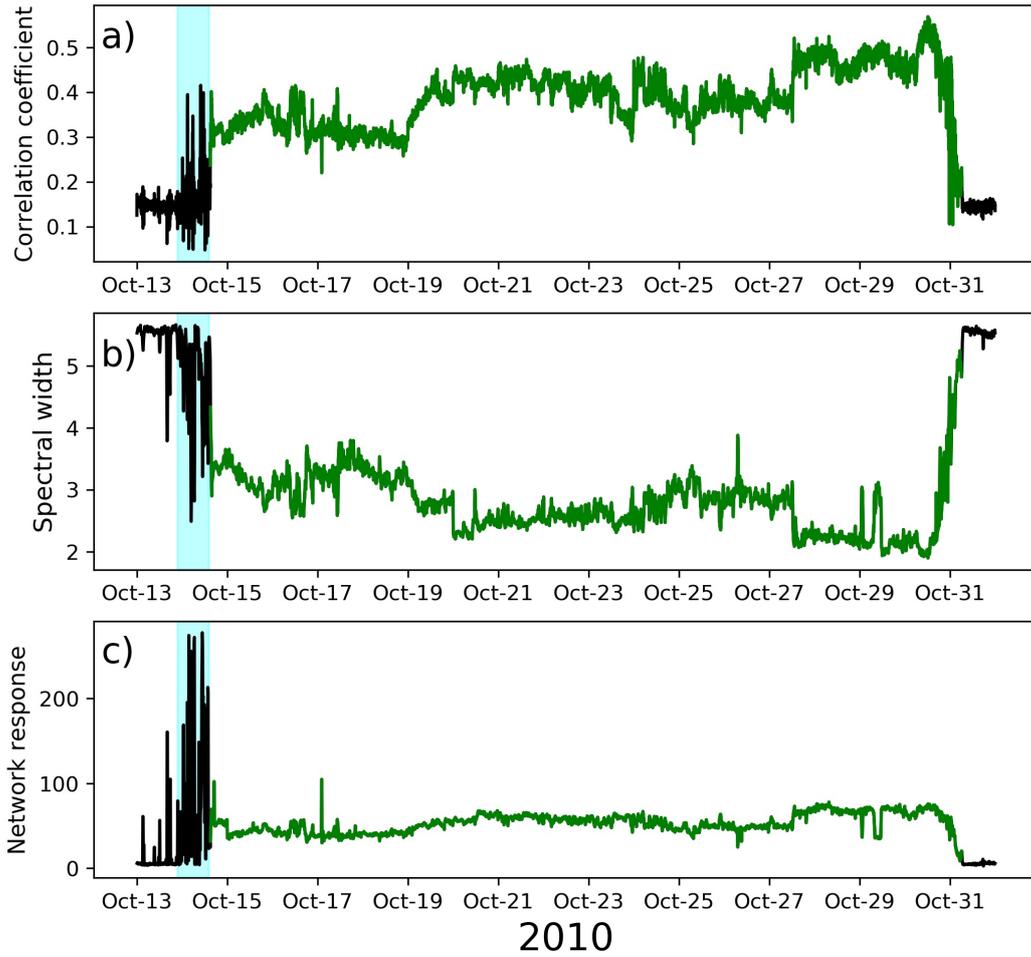


Figure 7: Zoom on the October eruption and its pre-eruptive seismic swarm for (a) the station-averaged  $c_{i,j}^6$  coefficient correlation function, (b) the spectral width and (c) the NRF. The green color stands for the October seismovolcanic tremor activity and the cyan color stands for the corresponding pre-eruptive seismic crisis, following the color code of the Figures 4 and 5.

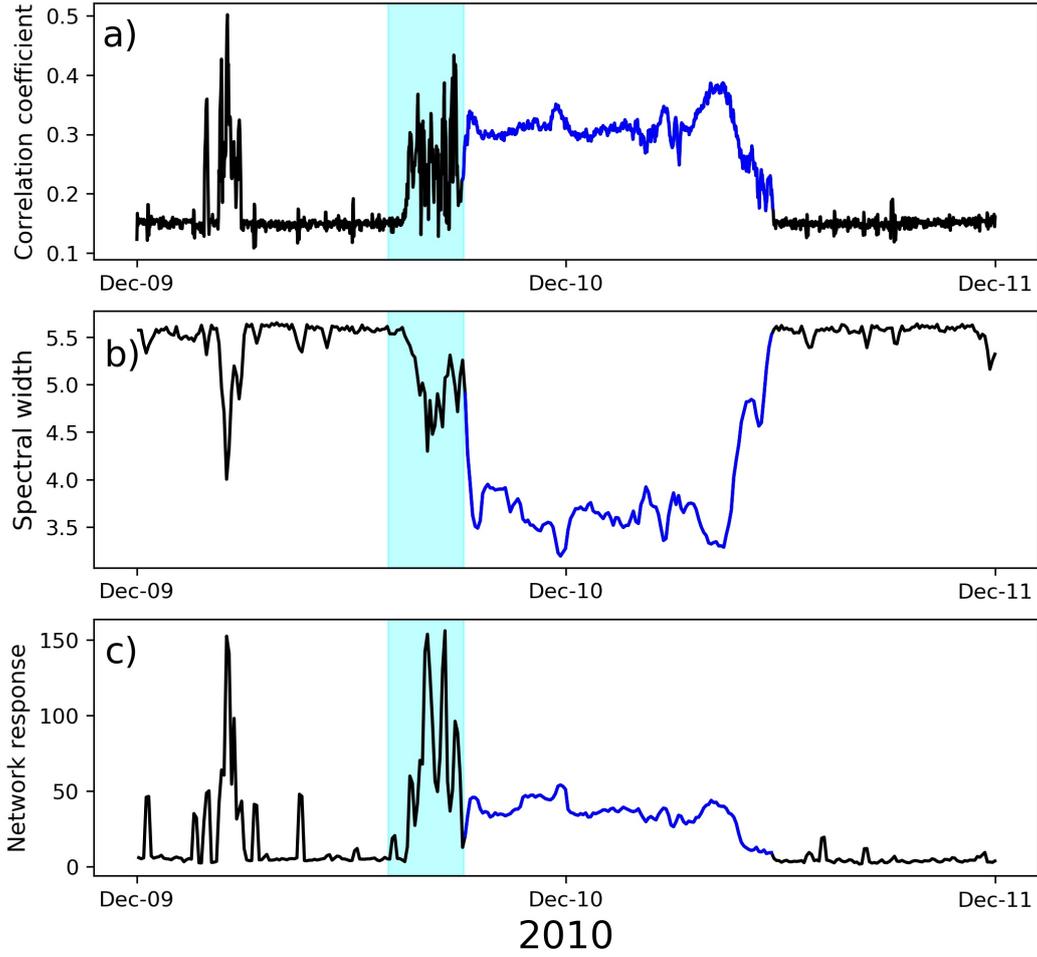


Figure 8: Zoom on the December eruption and its pre-eruptive seismic swarm for (a) the station-averaged  $c_{i,j}^6$  coefficient correlation function, (b) the spectral width and (c) the NRF. The blue color stands for the December seismovolcanic tremor activity and the cyan color stands for the corresponding pre-eruptive seismic crisis, following the color code of the Figures 4 and 5.

## 5.2 Location of seismovolcanic tremor sources

465 Figure 9 shows the seismovolcanic tremor sources location results in 3D based on the extraction of the dominant wavefield (as explained in the part 4.3, using a Gaussian window width of 1.5 s to smooth the CCFs and a 1D S-wave velocity model from Mordret et al. (2015). We calculate the daily covariance matrix by setting  $M = 50$  and  $\delta t = 40$  s to end up with 1000 s long windows that overlap with a 50% factor. Figure 9 shows the seismovolcanic tremor sources location results corresponding to the January eruption (9a), the October

470 eruption (9b) and the December eruption (9c). We define a 3D grid with points spaced  
every 200 m in order to compute the likelihood function that is then normalized using the  
equation (7). For Figures (9a) and (9b) we stacked the 1000 s overlapping windows for the  
entire days 03.01.2010 and 15.10.2010, respectively, while we stacked them only for the 15  
hours of eruption between the day 09.12.2010 and 10.12.2010 for Figure (9c). We observe  
475 that the seismovolcanic tremors seem to be located directly under the eruptive fissure for  
these three eruptions.

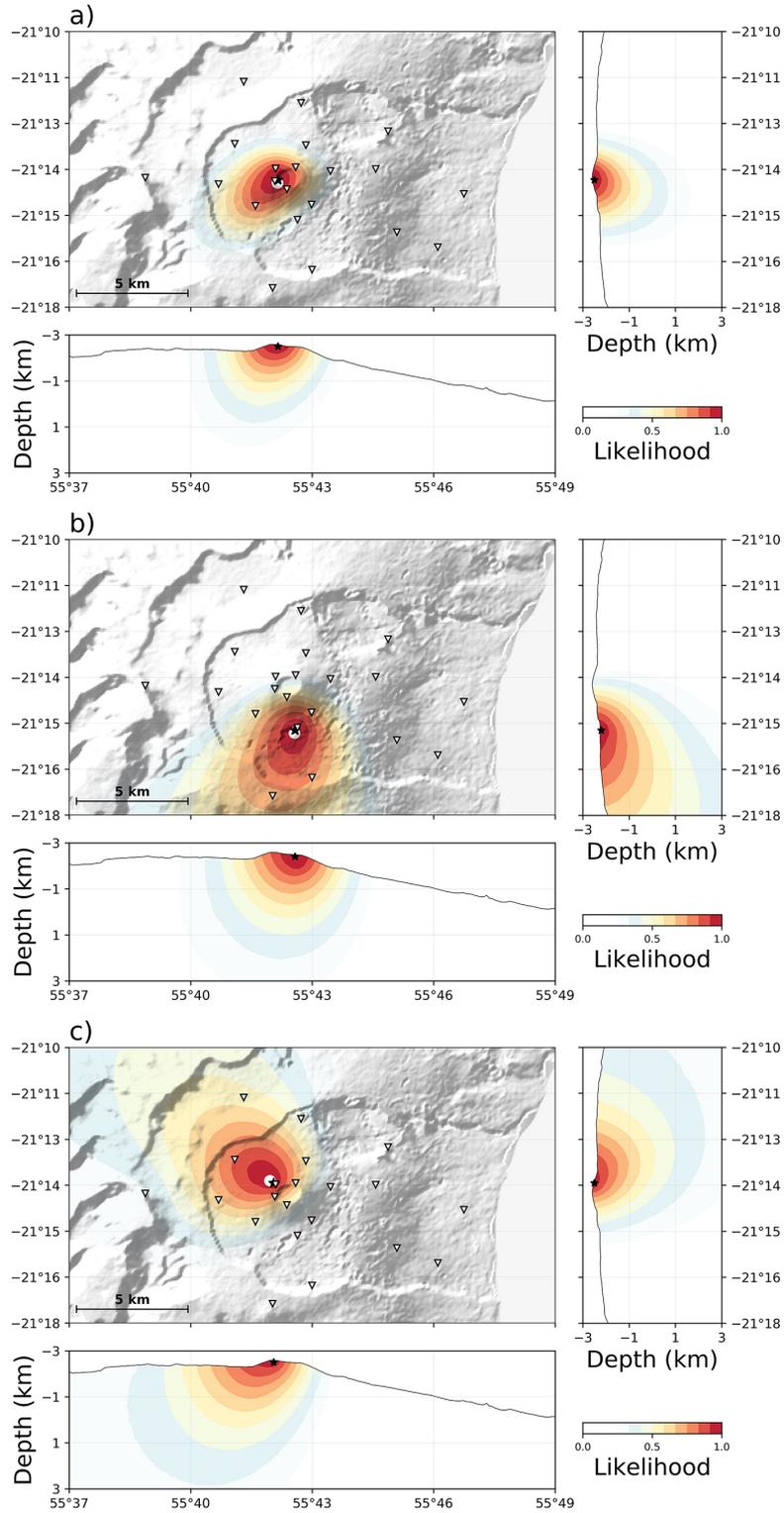


Figure 9: Normalized likelihood functions  $\tilde{R}$  calculated from the covariance matrix first eigenvectors for (a) the day 03.01.2010, (b) the day 15.10.2010 and (c) the 15 hours of the December eruption between days 09.12.2010 and 10.12.2010. The black stars stand for the location of the source with the maximal likelihood value and the white dot points for the eruptive sites. 3D location are performed using a 1D S-wave velocity of the PdF volcano from (Mordret et al., 2015).

### 5.3 Pre-eruptive swarms

To focus on details of volcano-tectonic (VT) events occurring in the observed pre-eruptive seismic swarms we modify the temporal normalization of the data and instead of applying the running absolute mean normalization (following Bensen et al. (2007); Seydoux, Shapiro, De Rosny, et al. (2016)), we divide the traces (once spectral whitening is applied) by their mean absolute deviation (MAD). We also enhance the temporal resolution of the covariance matrix analysis and modified the parameters introduced in the part 4.2, by setting  $M = 25$  and  $\delta t = 16$  s. As a result, the covariance matrix spectral width is computed in 200 s long overlapping time windows. Figure 10 shows the result of such enhanced seismic event detection and location during the first day of the October eruption (14.10.2010) with pre-eruptive seismic crisis and the seismovolcanic tremor starting at 15:00. We can clearly distinguish discrete VT events occurring in the seismic crisis in Figure 10a in difference with the continuous following seismovolcanic tremor.

Then we perform a 3D location by focusing on the first eigenvector of the covariance matrix computed for this day. We stack 10 consecutive 200 s long overlapping windows, to end up with a location point every 1000 s. Figure 10b shows the source depth as a function of time and the Figure 10c represents the source depth as a function of the latitude with the symbol colors corresponding to the windows central times.

We observe that VT events occurring since the beginning of the day and during the seismic crisis are located under the summit dome between  $\pm 1$  km around the sea level which is quite in agreement with what is typically observed for pre-eruptive swarms at PdF. For example Schmid (2011) located VT events from the 14.10.2010 seismic swarms beneath summit crater between 500 m below sea level and 1000 m above sea level. Then the seismicity seems to migrate toward the eruption site in the south flank for October eruption. Using template matching and relocation techniques, Duputel et al. (2019) show a clear correlation between pre-eruptive earthquake locations and the azimuth of eruptive sites which is consistent in our case with the presence of VT earthquakes in the southern branch of the summit cone before the migration.

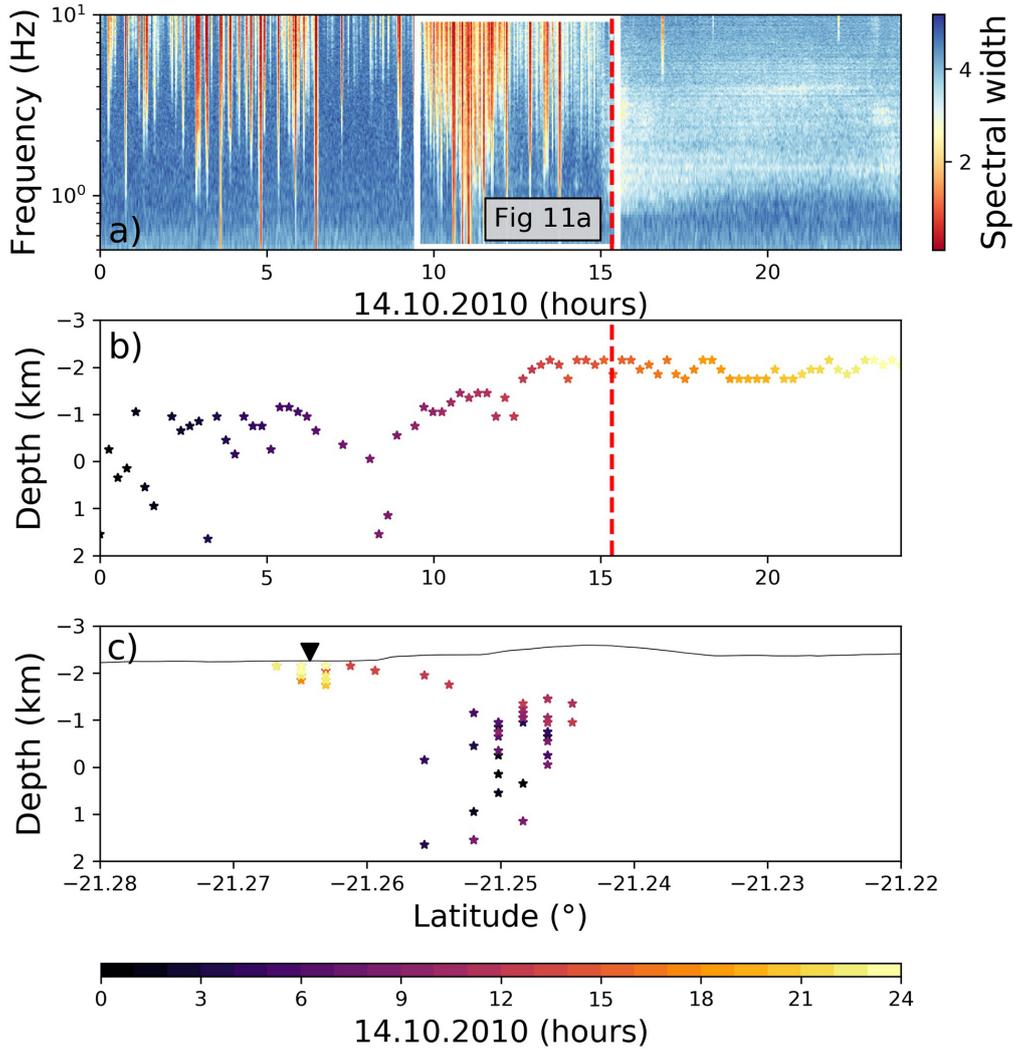


Figure 10: (a) Zoom on the covariance matrix spectral width calculated on 200 s overlapping time windows during the day 14.10.2010, first day of the October eruption. Red line stands for the beginning of the eruption. The white frame corresponds to the zoom represented in the Figure 11a. (b) 3D location of the detected seismic signals on 1000 s time window during the day 14.10.2010, the depth is represented as a function of time and the red vertical dashed line stands for the eruption beginning. (c) Same location with the depth as a function of latitude with the color representing the windows central times. The black triangle at the surface stands for the eruptive site position.

505

In Figure 11a, a zoom of the spectral width is represented for the day 14.10.2010 between 09:30 (beginning of the seismic swarm main part) and 15:30 (10 minutes after the eruption onset). Figure 11b shows an example of a 3D location on a 200 s long window of

a VT event occurring at 10:30, situated 1 km above the sea level under the summit cone. In addition to the seismovolcanic tremor and VT events constituting the pre-eruptive seismic swarm, the computation of the covariance matrix spectral width allows also to detect  
510 another type of recently discovered seismic swarm of events characterized by a very high frequency content (called HF swarm) (De Barros et al., 2013). As stated by Roult et al. (2012), the classical sequence of events preceding an eruption at PdF is: a first increase of VT events, a swarm of VT with larger magnitudes, then a seismically quiet or low in-  
515 tensity sequence with small number of VT events linked to magma migration followed by the seismovolcanic tremor onset and the eruptive fissure opening. During the quiescent period between the swarm and the seismovolcanic tremor, De Barros et al. (2013) identified a swarm of HF events the 14.10.2010 since 13:00, located in the area of the eruptive fracture in the south flank and interpreted as a response of the shallow part of the edifice to the  
520 stress changes due to the dike propagation. We can observe in the Figure 11a some coherent signals between the end of the VT swarm and the seismovolcanic tremor, in particular from 14:00 to 15:00, for frequencies larger than 3 Hz. We represent in Figure 11c an example of a 3D location on a 200 s long window of a HF event occurring at 14:10. There is therefore a migration of the seismicity from the pre-eruptive seismic swarm under the summit cone at  
525 the sea level toward a HF swarm superficially near the future eruptive site more than one hour before the eruption.

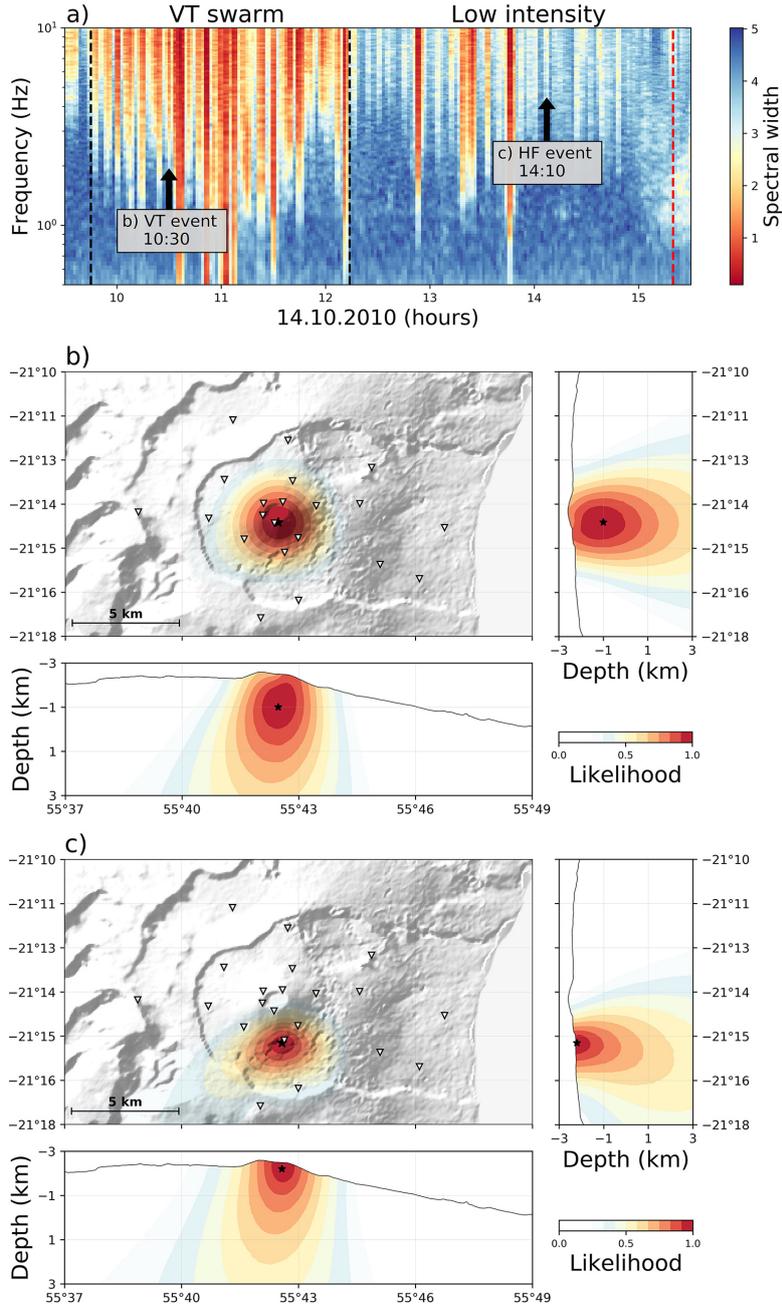


Figure 11: (a) Zoom of the covariance matrix spectral width the day 14.10.2010 between the beginning of the seismic swarm main part (09:30) and 10 minutes after the eruption onset. The two black arrows show the timing of the two events with locations shown in (b) and (c). The two vertical dashed black lines stand for the beginning and the end of the VT swarm main part. Then there is a phase of low intensity seismicity with a small number of VT events. During this phase the lateral migration of the dike begins and we can observe a coherent high-frequency swarm. The red vertical dashed line marks the onset of the seismovolcanic tremor and associated eruption. (b) Likelihood functions  $\tilde{R}$  calculated from the covariance matrix first eigenvectors, the day 14.10.2010 for a VT event at 10:30 and (c) for a HF event that is part to the high-frequency seismic swarm.

## 6 Discussion and conclusions

In this paper, we tested methods that use the phase information contained in continuous seismic signals to study seismovolcanic tremors. First, we introduced a new method of  
530 detection of seismovolcanic tremors based on stability of inter-components cross-correlations at a single station. In a next step, we systematically compared this new method with two previously proposed network-based methods by applying all three methods to continuous data recorded by stations of the Piton de la Fournaise (PdF) volcano observatory seismic  
535 network during 2010. The second method based on the network covariance matrix has been previously applied to the dataset used in this paper by Seydoux, Shapiro, De Rosny, et al. (2016). However, in this previous paper the time resolution of the analysis was relatively low (one day), and here we increased it to be able to detect individual volcanic earthquakes. The third method called "network response function" (Droznin et al., 2015) has been applied for  
540 the first time to the PdF data and in a 3D grid considering a 1D S-wave velocity model.

All three methods have been demonstrated to be very sensitive to seismovolcanic tremors and detected all three eruptive seismovolcanic tremors that occurred during 2010. They are also capable to detect strong pre-eruptive seismic swarms. In addition to this, the two network-based methods detect a large number of volcanic earthquakes and, in  
545 particular, the significant increase of seismicity during the month preceding the October 2010 PdF eruption as seen in Figure 3. This seismicity increase is well correlated with the measured decrease of seismic velocity that is interpreted to be linked to the volcanic edifice deformation induced by magma pressure build-up and injection (Brenguier et al., 2012).

The "single-station inter-components cross-correlations" method can be implemented  
550 on volcanoes monitored by a small number of stations. Among the two network-based methods, the "covariance matrix" detector has some advantage because it does not require a-priori information about the wave propagation velocity. Additional advantage of network based-methods is their ability to locate seismovolcanic tremor sources. In the case of a network with significant number of well distributed stations, the 3D location based on "spatially  
555 filtered" cross-correlations (reconstructed from dominant covariance matrix eigenvector) can be done. Apart from the seismic data quality, the accuracy of such location depends on the knowledge of the internal velocity structure of the volcano. In this paper, the tests were limited to location with using a 1D S-wave velocity model. In principle, wave propagation in 3D velocity models can be easily incorporated. Permana et al. (2019) used a similar

560 method to locate seismovolcanic tremor using a 1-D S wave model as in Soubestre et al. (2019) (except that they do not focus on a dominant seismovolcanic tremor source by filtering by the first eigenvector of the covariance matrix). They simulate seismovolcanic tremor sources by combining multiple VTs and show that they are determined with location errors of approximately 1 km or less.

565 Moreover, with pushing their time resolution, the discussed method could bring insight on the complexity of the eruption dynamics as shown in Figures 6,7,8. For example, we observe in Figures 4 and 6b that for the January seismovolcanic tremor, the spectral width is very low at the beginning and becomes larger and larger over time until it reaches a value equivalent to that corresponding to seismic noise. Conversely, for the October eruption we see in Figure 7b a more complicated temporal evolution for the spectral width that could 570 reveal a more complex dynamic. By comparing with Figures 7a and 7c, we observe two main changes occurring at the same time from days 19.10.2010 and 27.10.2010, that are linked with intensification of the seismovolcanic tremor amplitude visible on the RSAM measurement in Figure S1. Thus such methods can also be useful to monitor the seismovolcanic tremor 575 intensity.

Another interesting prospective is to use the results of the multi-component analysis as input for algorithms of machine learning (ML), to build advanced approaches for detection and classification of different seismovolcanic signals to be used in monitoring the activity and eruptive behavior of volcanoes. Figures 4 and 5 show that, in the space of parameters 580 determined with the multi-component methods, different types of seismovolcanic activity are separated and form relatively well defined clusters. It suggests that ML could be applied with this kind of "features" determined from the data of the PdF seismic network. Previous applications of machine learning to the PdF seismological data used features derived from individual sensors to study volcanic earthquakes (Hibert et al., 2017; Maggi et al., 2017) and 585 seismovolcanic tremors (Ren et al., 2019). The eventual advantage of multi-station features is that they incorporate inter-locations phase differences and are, therefore, sensitive to the location and the mechanism of the signal sources.

## Acknowledgments

This study was financially supported by the European Research Council (ERC) under the European Union Horizon 2020 Research and Innovation Programme (grant agree- 590

ment 787399-SEISMAZE) and by the Russian Ministry of Education and Science (grant N 14.W03.31.0033) for NS. Computations were performed using the IPGP High-Performance Computing infrastructure S-CAPAD (supported by the Ile-de-France region via the SEASAME programme, by France-Grille, and by the CNRS MASTODONS programme). The data used are publicly available and can be found online. The seismic data used for this study were provided by the Observatoire Volcanologique du Piton de la Fournaise (OVPF). The permanent network data (<https://www.fdsn.org/networks/detail/PF/>) can be downloaded from the IPGP Data Center (<http://ws.ipgp.fr/fdsnws/dataselect/1/>), the RESIF data center (<http://ws.resif.fr/fdsnws/dataselect/1/>), or the IRIS Data Management Center (IRIS-DMC, <http://service.iris.edu/fdsnws/dataselect/1/>). The UnderVolc data (<https://doi.org/10.15778/RESIF.YA2009>) are available from the RESIF data center. This is IPGP contribution number XXXX.

## References

- Bachèlery, P. (1981). Le Piton de la Fournaise (Ile de la Réunion): Etude volcanologique, structurale et pétrologique (Thèse de doctorat).
- Ballmer, S., Wolfe, C. J., Okubo, P. G., Haney, M. M., & Thurber, C. H. (2013). Ambient seismic noise interferometry in Hawai'i reveals long-range observability of volcanic tremor. Geophysical Journal International.
- Battaglia, J., & Aki, K. (2003). Location of seismic events and eruptive fissures on the Piton de la Fournaise volcano using seismic amplitudes. Journal of Geophysical Research, 108(B8).
- Battaglia, J., Aki, K., & Ferrazzini, V. (2005b). Location of tremor sources and estimation of lava output using tremor source amplitude on the Piton de la Fournaise volcano : 1. Location of tremor sources. Journal of Volcanology and Geothermal Research, 147, 268–290.
- Battaglia, J., Brenguier, F., & Roult, G. (2016). Seismic Monitoring at Piton de la Fournaise. In P. Bachèlery, J.-F. Lénat, A. Di Muro, & L. Michon (Eds.), Active Volcanoes of the South-west Indian Ocean (pp. 223–242).
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., . . . Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophysical Journal International, 169(3), 1239–1260.
- Bonneville, A. (1990). Structure de la lithosphère. In L. J.-F. (Ed.), Le volcanisme de la

- réunion, monographie (pp. 1–18).
- Brenguier, F. (2014). Undervolc experiment, 2009-2011, code ya (understanding volcanic processes). resif - rseau sismologique et godsique franais. terrestrial seismic network. doi: 10.15778/RESIF.YA2009
- Brenguier, F., Kowalski, P., Staudacher, T., Ferrazzini, V., Lauret, F., Boissier, P., ... Di Muro, A. (2012). First Results from the UnderVolc High Resolution Seismic and GPS Network Deployed on Piton de la Fournaise Volcano. Seismological Research Letters, 83.
- Chouet, B. A. (1996). Long-period volcano seismicity: its source and use in eruption forecasting. Nature, 380, 309-316. doi: 10.1038/380309a0
- Chouet, B. A., & Matoza, R. S. (2013). A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. Journal of Volcanology and Geothermal Research, 252, 108–175.
- Courtillot, V., Besse, J., Vandamme, D., Montigny, R., Jaeger, J., & Cappetta, H. (1986). Deccan flood basalts at the Cretaceous/Tertiary boundary? Earth and Planetary Science Letters, 80, 361–374.
- De Barros, L., Bean, C. J., Zecevic, M., Brenguier, F., & Peltier, A. (2013). Eruptive fracture location forecasts from highfrequency events on Piton de la Fournaise Volcano. Geophysical Research Letters, 40, 4599–4603.
- Droznin, D., Shapiro, N. M., Droznina, S. Y., Senyukov, S., Chebrov, V., & Gordeev, E. (2015). Detecting and locating volcanic tremors on the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records. Geophysical Journal International, 203(2), 1001–1010.
- Duputel, Z., Lengline, O., & Ferrazzini, V. (2019). Constraining Spatiotemporal Characteristics of Magma Migration at Piton De La Fournaise Volcano From Pre-eruptive Seismicity. Geophysical Research Letters, 46(1), 119-127. doi: 10.1029/2018GL080895
- Endo, E. T., & Murray, T. (1991). Real-time seismic amplitude measurement (rsam): a volcano monitoring and prediction tool. Bulletin of Volcanology, 53(7), 533–545. doi: 10.1007/BF00298154
- Goldstein, P., & Chouet, B. (1994). Array measurements and modeling of sources of shallow volcanic tremor at Kilauea volcano, Hawaii. Journal of Geophysical Research: Solid Earth, 99(B2), 2637–2652.
- Haney, M. M. (2014). Backprojection of volcanic tremor. Geophysical Research Letters,

- 41(6), 1923-1928. doi: 10.1002/2013GL058836
- Hibert, C., Provost, F., Malet, J.-P., Maggi, A., Stumpf, A., & Ferrazzini, V. (2017). Automatic identification of rockfalls and volcano-tectonic earthquakes at the piton de la fournaise volcano using a random forest algorithm. Journal of Volcanology and Geothermal Research, 340, 130 - 142. doi: <https://doi.org/10.1016/j.jvolgeores.2017.04.015>
- Lénat, J.-F., & Bachèlery, P. (1990). Structure et fonctionnement de la zone centrale du Piton de la Fournaise. In J.-F. Lénat (Ed.), Le volcanisme de La Réunion (pp. 257-296).
- Maggi, A., Ferrazzini, V., Hibert, C., Beauducel, F., Boissier, P., & Amemoutou, A. (2017). Implementation of a Multistation Approach for Automated Event Classification at Piton de la Fournaise Volcano. Seismological Research Letters, 88(3), 878-891. doi: 10.1785/0220160189
- McNutt, S. R., & Nishimura, T. (2008). Volcanic tremor during eruptions: Temporal characteristics, scaling and constraints on conduit size and processes. Journal of Volcanology and Geothermal Research, 178(1), 10 - 18. doi: <https://doi.org/10.1016/j.jvolgeores.2008.03.010>
- McNutt, S. R., & Roman, D. C. (2015). Chapter 59 - volcanic seismicity. In H. Sigurdsson (Ed.), The encyclopedia of volcanoes (second edition) (Second Edition ed., p. 1011 - 1034). Amsterdam: Academic Press. doi: <https://doi.org/10.1016/B978-0-12-385938-9.00059-6>
- Métaxian, J.-P., Lesage, P., & Valette, B. (2002). Locating sources of volcanic tremor and emergent events by seismic triangulation: Application to Arenal volcano, Costa Rica. Journal of Geophysical Research: Solid Earth, 107(B10).
- Mordret, A., Rivet, D., Lands, M., & Shapiro, N. M. (2015). Three-dimensional shear velocity anisotropic model of piton de la fournaise volcano (la reunion island) from ambient seismic noise. Journal of Geophysical Research: Solid Earth, 120(1), 406-427. doi: 10.1002/2014JB011654
- OVPF. (2011). Activity bulletin 08/01/2011, ISSN 2610-5101.
- Peltier, A., Famin, V., Bachlery, P., Cayol, V., Fukushima, Y., & Staudacher, T. (2008). Cyclic magma storages and transfers at Piton de La Fournaise volcano (la Réunion hotspot) inferred from deformation and geochemical data. Earth Planet. Sci. Lett., 270, 180-188. doi: 10.1016/j.epsl.2008.02.042

- Permana, T., Nishimura, T., Nakahara, H., Fujita, E., & Ueda, H. (2019). Reliability  
 690 evaluation of volcanic tremor source location determination using cross-correlation  
 functions. *Geophysical Journal International*, 220(2), 1300-1315. doi: 10.1093/gji/  
 ggz523
- Ren, C. X., Peltier, A., Ferrazzini, V., Rouet-Leduc, B., Johnson, P. A., & Brenguier, F.  
 (2019). Machine learning reveals the seismic signature of eruptive behavior at piton  
 695 de la fournaise volcano.
- Roman, D. C., & Cashman, K. V. (2006). The origin of volcano-tectonic earthquake swarms.  
*Geology*, 34(6), 457-460. doi: 10.1130/G22269.1
- Rost, S., & Thomas, C. (2002). Array seismology: Methods and applications. *Reviews of  
 700 Geophysics*, 40(3), 2-1-2-27. Retrieved from <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000RG000100> doi: 10.1029/2000RG000100
- Roult, G., Peltier, A., Taisne, B., Staudacher, T., Ferrazzini, V., & Di Muro, A. (2012).  
 A new comprehensive classification of the Piton de la Fournaise activity spanning the  
 1985-2010 period. Search and analysis of short-term precursors from a broad-band  
 seismological station. *J. Volcanol. Geotherm. Res.*, 241, 78–104.
- Schmid, A. (2011). Quelle prédictibilité pour les éruptions volcaniques ? de l'échelle  
 705 mondiale au piton de la fournaise (Thèse de doctorat).
- Seydoux, L., de Rosny, J., & Shapiro, N. M. (2017). Pre-processing ambient noise cross-  
 correlations with equalizing the covariance matrix eigenspectrum. *Geophysical Journal  
 International*, 210(3), 1432-1449. doi: 10.1093/gji/ggx250
- Seydoux, L., Shapiro, N. M., De Rosny, J., Brenguier, F., & Landes, M. (2016). Detecting  
 710 seismic activity with a covariance matrix analysis of data recorded on seismic arrays.  
*Geophysical Journal International*, 204(3), 1430–1442.
- Seydoux, L., Shapiro, N. M., de Rosny, J., & Landes, M. (2016). Spatial coherence of the  
 seismic wavefield continuously recorded by the usarray. *Geophysical Research Letters*,  
 715 43(18), 9644-9652. doi: 10.1002/2016GL070320
- Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband rayleigh waves from  
 correlations of the ambient seismic noise. *Geophysical Research Letters*, 31(7).
- Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-resolution  
 surface-wave tomography from ambient seismic noise. *Science*, 307(5715), 1615–1618.  
 720 doi: 10.1126/science.1108339
- Shapiro, N. M., Ritzwoller, M. H., & Bensen, G. D. (2006). Source location of the 26

- sec microseism from cross-correlations of ambient seismic noise. Geophysical Research Letters, **33**(18), L18310. doi: 10.1029/2006GL027010
- 725 Soubestre, J., Seydoux, L., Shapiro, N. M., de Rosny, J., Droznin, D. V., Droznina, S. Y.,  
 ... Gordeev, E. I. (2019). Depth migration of seismovolcanic tremor sources below the  
 klyuchevskoy volcanic group (kamchatka) determined from a network-based analysis.  
Geophysical Research Letters, **46**(14), 8018-8030. doi: 10.1029/2019GL083465
- 730 Soubestre, J., Shapiro, N. M., Seydoux, L., de Rosny, J., Droznin, D. V., Droznina, S. Y., ...  
 Gordeev, E. I. (2018). Network-based detection and classification of seismovolcanic  
 tremors: Example from the klyuchevskoy volcanic group in kamchatka. Journal of  
 Geophysical Research: Solid Earth, **123**(1), 564-582. doi: 10.1002/2017JB014726
- 735 Staudacher, T., Peltier, A., Ferrazzini, V., Di Muro, A., Boissier, P., Catherine, P., ...  
 Lebreton, J. (2016). Fifteen Years of Intense Eruptive Activity (1998-2013) at Piton  
 de la Fournaise volcano: a Review. In P. Bachèlery, J.-F. Lénat, A. Di Muro, &  
 L. Michon (Eds.), Active Volcanoes of the South-west Indian Ocean (pp. 139–170).  
 doi: 10.1007/978-3-642-31395-0\_9
- Taisne, B., Brenguier, F., Shapiro, N. M., & Ferrazzini, V. (2011). Imaging the dynamics of  
 magma propagation using radiated seismic intensity. Geophys. Res. Lett., **38**(4), 2–6.  
 doi: 10.1029/2010GL046068
- 740 Tsekhmistrenko, M., Sigloch, K., & Hosseini, K. (2018). Whole-mantle structure under the  
 Reunion hotspot in the western Indian Ocean from multifrequency P-wave tomogra-  
 phy. In EGU General Assembly Conference Abstracts (p. 1076).
- Unglert, K., & Jellinek, A. M. (2015). Volcanic tremor and frequency gliding during dike  
 intrusions at kilauea tale of three eruptions. Journal of Geophysical Research: Solid  
 Earth, **120**(2), 1142-1158. doi: 10.1002/2014JB011596
- 745 Unglert, K., Radi, V., & Jellinek, A. (2016). Principal component analysis vs. self-organizing  
 maps combined with hierarchical clustering for pattern recognition in volcano seismic  
 spectra. Journal of Volcanology and Geothermal Research, **320**, 58 - 74. doi: [https://  
 doi.org/10.1016/j.jvolgeores.2016.04.014](https://doi.org/10.1016/j.jvolgeores.2016.04.014)