Network-Based Detection and Classification of Seismovolcanic Tremors: Example From the Klyuchevskoy Volcanic Group in Kamchatka

Jean Soubestre, Nikolai M. Shapiro, Léonard Seydoux, Julien de Rosny, Dmitry V. Droznin, Svetlana Ya. Droznina, Sergey L. Senyukov, Evgeniy I. Gordeev

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1002/2017JB014726

Key Points:
- A network-based method for the automatic detection and classification of seismovolcanic tremors is developed
- The developed clustering algorithm automatically identifies seven clusters associated with different periods of activity of four volcanoes
- The developed method does not require a priori knowledge, is fully automatic, and can be updated continuously with newly available data

Supporting Information:
- Supporting Information S1

Correspondence to:
J. Soubestre,
soubestre@ipgp.fr

Citation:
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*, 564–582. https://doi.org/10.1002/2017JB014726

Received 18 JUL 2017
Accepted 5 JAN 2018
Accepted article online 8 JAN 2018
Published online 26 JAN 2018

Network-Based Detection and Classification of Seismovolcanic Tremors: Example From the Klyuchevskoy Volcanic Group in Kamchatka

Jean Soubestre[1], Nikolai M. Shapiro[1,2], Léonard Seydoux[1], Julien de Rosny[3], Dmitry V. Droznin[4], Svetlana Ya. Droznina[4], Sergey L. Senyukov[4], and Evgeniy I. Gordeev[4,5]

[1]Institut de Physique du Globe de Paris, UMR CNRS 7154, Paris, France, [2]Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, Russia, [3]Institut Langevin, CNRS PSL Research University, France, [4]Kamchatka Branch of the Geophysical Survey, Russian Academy of Sciences, Petropavlovsk-Kamchatsky, Russia, [5]Institute of Volcanology and Seismology, FEB RAS, Petropavlovsk-Kamchatsky, Russia

Abstract We develop a network-based method for detecting and classifying seismovolcanic tremors. The proposed approach exploits the coherence of tremor signals across the network that is estimated from the array covariance matrix. The method is applied to four and a half years of continuous seismic data recorded by 19 permanent seismic stations in the vicinity of the Klyuchevskoy volcanic group in Kamchatka (Russia), where five volcanoes were erupting during the considered time period. We compute and analyze daily covariance matrices together with their eigenvalues and eigenvectors. As a first step, most coherent signals corresponding to dominating tremor sources are detected based on the width of the covariance matrix eigenvalues distribution. Thus, volcanic tremors of the two volcanoes known as most active during the considered period, Klyuchevskoy and Tolbachik, are efficiently detected. As a next step, we consider the daily array covariance matrix's first eigenvector. Our main hypothesis is that these eigenvectors represent the principal components of the daily seismic wavefield and, for days with tremor activity, characterize dominant tremor sources. Those daily first eigenvectors, which can be used as network-based fingerprints of tremor sources, are then grouped into clusters using correlation coefficient as a measure of the vector similarity. As a result, we identify seven clusters associated with different periods of activity of four volcanoes: Tolbachik, Klyuchevskoy, Shiveluch, and Kizimen. The developed method does not require a priori knowledge and is fully automatic; and the database of the network-based tremor fingerprints can be continuously enriched with newly available data.

1. Introduction

Detecting and characterizing seismicity is a key aspect of volcano monitoring (Chouet & Matoza, 2013; McNutt, 1992; Sparks et al., 2012). Most preeruptive and coeruptive processes within volcanic systems are accompanied by seismic emission that can be used to detect and identify different phases of volcanic activity. Seismic signals associated with volcanic unrest are very variable. Stresses induced by the ascending magma are released in form of volcano-tectonic earthquakes (Roman & Cashman, 2006). Another large class of activity known as long-period volcanic seismicity is more directly related to processes induced by fluid movement within the volcano plumbing or hydrothermal systems (Chouet, 1996b; Iverson et al., 2006) that may result in long-period earthquakes (e.g., Shapiro, Droznin, et al., 2017) and highly irregular signals known as volcanic tremors.

Many approaches currently used for seismovolcanic monitoring originate from the classical earthquake seismology and are, therefore, poorly adapted for the analysis of seismic tremors. Another difficulty of traditional approaches for seismic monitoring is that they are largely based on visual inspection of seismic records for detecting events and identifying characteristic patterns. With continuously growing monitoring networks and data fluxes, this way of analyzing data becomes more and more problematic. Therefore, the seismovolcanic monitoring (similarly to many other areas) must rely more on data-intensive automatic methods for analysis and classification of signals leading to the idea of applying methods from the area of machine learning (e.g., Carniel, 1996, 2014; Orozco-Alzate et al., 2012). Machine learning approaches related to seismovolcanic data are most of the time applied on single time series from individual stations to perform blind source separation

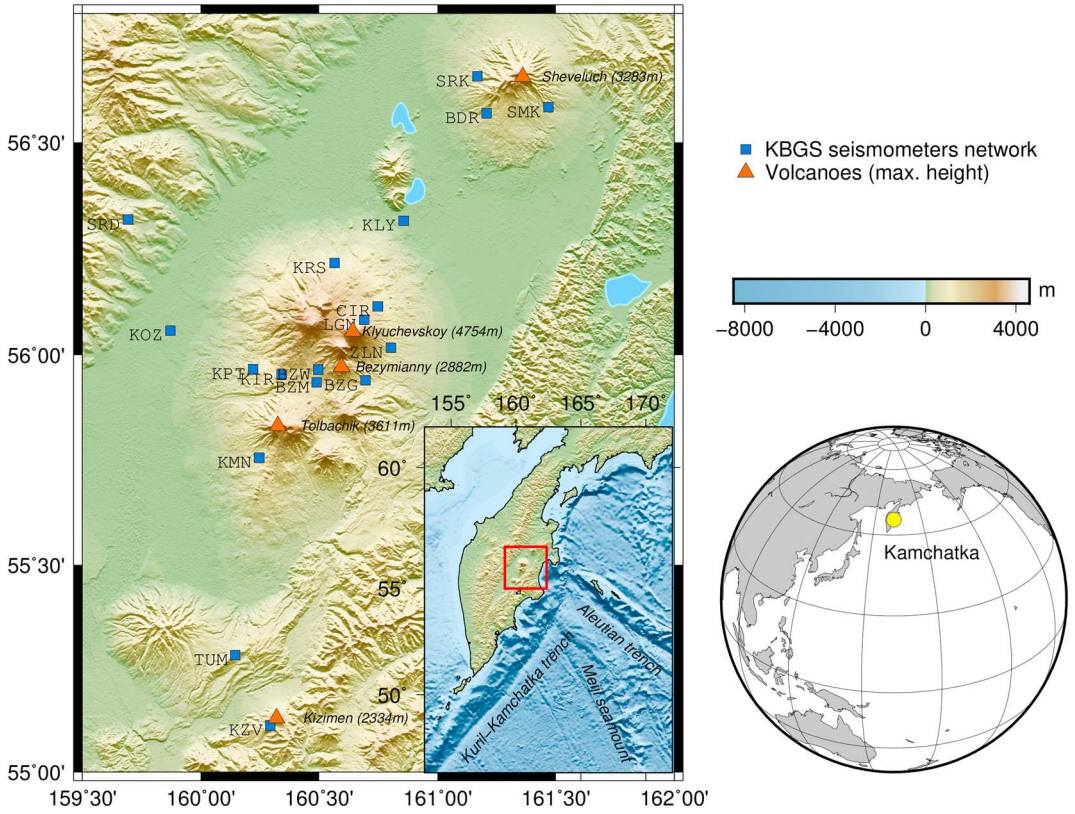

Figure 1. Seismic network (stations: blue squares) monitoring the Klyuchevskoy volcanic group (KVG) and surrounding volcanoes (red triangles). The Russian Kamchatcka peninsula is located by a yellow circle on the Earth globe and the insert shows the particular location of the KVG region (red rectangle in the insert) in front of the intersection of the Meiji seamount with the edge of the Pacific Plate's subduction beneath the North American Plate along the Kuril-Kamchatka trench on the one hand and along the Aleutian trench on the other hand.

with dimension reduction methods such as Independant Component Analysis (Acernese et al., 2003; Cabras et al., 2008, 2010; Ciaramella et al., 2011; Capuano et al., 2016), Nonnegative Matrix Factorization (Cabras et al., 2012, 2014), or Degenerate Unmixing Estimation Technique (Moni et al., 2012). A few machine learning studies of seismovolcanic signals adopt a multistation approach but dealing with features (a machine learning terminology that stands for any attribute of the waveform) derived from individual stations of the network (e.g., waveform, spectrum, spectrogram, …). Thus, Hibert et al. (2017) use Random Forest algorithm to classify rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using waveform, spectrum, spectrogram, and polarization as features. At the same volcano and using also Random Forest algorithm, Maggi et al. (2017) classify eight classes of seismovolcanic signals automatically selecting the best waveform and spectrum features from a combination of stations. Unglert and Jellinek (2017) use Principal Component Analysis and Hierarchical Clustering algorithms to do pattern recognition of tremor spectra from multiple stations of one volcano and of different volcanoes (Kilauea, Okmok, Redoubt, and Pavlof).

Most of the previously proposed approaches for automatic classification of seismovolcanic signals are based on analysis of features derived from individual sensors (e.g., Carniel, 1996; Hibert et al., 2017; Maggi et al., 2017; Orozco-Alzate et al., 2012; Unglert & Jellinek, 2017). In this paper, we present a method that automatically classifies volcanic tremors with their characteristic patterns defined from an ensemble of signals recorded by a network composed of multiples sensors. The feature we use being derived from the sensors all together, we call the method we developed a network-based method. So far, array methods have been used to determine the location of tremor sources with small-aperture arrays (e.g., Almendros et al., 1997; Goldstein & Chouet, 1994, Haney, 2010; Métaxian et al., 2002). The method presented here is aimed at classifying tremors without locating their sources. The main idea is to explore the coherence of the tremor wavefield

between different locations. This coherence is expressed in interstation cross correlations of continuous seismic records with tremors resulting in waveforms with characteristic arrival times and phases that can be used as their "fingerprints" (Droznin et al., 2015). An ensemble of interstation cross correlations represented in Fourier domain forms a covariance matrix (Seydoux, Shapiro, de Rosny, Brenguier, & Landés, 2016; Seydoux, Shapiro, de Rosny, & Landès 2016). The first eigenvector of this matrix characterizes a dominant component of the recorded wavefield (Anderson, 1963), and we propose to use it as characteristic patterns for the tremor classification. For this classification, we employ a clustering algorithm where the similarity between different wavefield realizations is measured from correlation coefficients between the covariance matrix first eigenvectors.

We apply the developed method to continuous records by permanent monitoring network in the region of the Klyuchevskoy group of volcanoes (Kamchatka, Russia) (Figure 1) during a four and a half years long time period between January 2009 and June 2013. During this period of time, the Klyuchevskoy, Tolbachik, Kizimen, and Shiveluch volcanoes experienced very long episodes of volcanic tremor (Droznin et al., 2015) and abundant amount of long-period earthquakes (Shapiro, Droznin, et al., 2017). The two most important tremors lead to a summit eruption at the end of 2010 at Klyuchevskoy (e.g., Senyukov, 2013) and a fissure eruption beginning at the end of 2012 at Tolbachik (e.g., Gordeev et al., 2013). Short explosive eruptions also occurred at Bezymianny.

The studied volcanic group and the monitoring seismic network are presented in section 2. Section 3 details the specificity of volcanic tremors monitoring. The estimation of the array covariance matrix and the clustering algorithm based on its eigenvectors are described in section 4. The different detected and classified volcanic tremor sources are presented in section 5.

2. Klyuchevskoy Volcanic Group Monitoring

The Klyuchevskoy volcanic group (KVG) consists of 13 stratovolcanoes (the three most active being Klyuchevskoy, Bezymianny, and Tolbachik) located in an approximatively 70 km diameter zone on the Russian Kamchatka peninsula (Shapiro, Sens-Schönfelder, et al., 2017). Its very particular tectonic setting in front of the intersection of the Meiji seamount with the edge of the Pacific Plate's subduction beneath the North American Plate along the Kuril-Kamchatka trench on the one hand and along the Aleutian trench on the other hand (cf. red rectangle in Figure 1 insert) turns it into one of the most active subduction zone volcanic group in the world. Different geodynamic models proposed to explain this important volcanism include the following: geochemical evidence for fluid release from the thick and highly hydrated Meiji seamount crust (Dorendorf et al., 2000); geochemical evidence for a mantle flow around the corner of the Pacific Plate (Yogodzinski et al., 2001); seismic, tectonic, and petrological evidence for a recent (2 Myr ago) detachment of a portion of the subducting slab (Levin et al., 2002, 2005; Park et al., 2002).

2.1. Recent Activity

Three volcanoes of the KVG (Klyuchevskoy, Bezymianny, and Tolbachik) were active during recent decades (e.g., Gordeev et al., 1989; Ivanov, 2008; Ozerov et al., 2007; Senyukov et al., 2009, 2013). In addition, two other very active volcanoes, Shiveluch and Kizimen, are respectively located north and south of KVG (Figure 1). The volcano earthquakes together with tectonic regional and teleseismic events were used to study the internal structure of the KVG with the seismic tomography (e.g., Balesta et al., 1991; Gorbatov et al., 1999; Gontovaya et al., 2004; Ivanov et al., 2016; Koulakov et al., 2011, 2013, 2016; Lees et al., 2007; Slavina et al., 2001, 2012) and with receiver functions (Levin et al., 2014; Nikulin et al., 2010). The sustained volcanic activity of the KVG and closely located volcanoes requires permanent monitoring and makes this region particularly suitable for testing volcanic monitoring methods.

2.2. Monitoring Seismic Network

The Kamchatka Branch of the Geophysical Survey (KBGS) of the Russian Academy of Sciences operates the permanent seismic network monitoring the KVG and surrounding volcanoes (Chebrov et al., 2013; Gordeev et al., 2006). In this paper, we use continuous records during four and a half years between January 2009 and June 2013 by 18 stations in 2009 and 19 stations from 2010 onward. The stations and the five monitored volcanoes are shown in Figure 1. The seismic stations have three components, each one being equipped by a CM-3 short period seismometer with a corner frequency of around 0.8 Hz. Continuous records are sampled at 128 Hz, and only the vertical component is analyzed in this article.

3. Volcanic Tremors Monitoring

Volcanic tremors are continuous phenomena that can last from a few hours to several months, as in the case of the KVG which experienced very long volcanic tremors at Klyuchevskoy and Tolbachik during the studied period. They have different natures (harmonic, monochromatic, banded, spasmodic, etc.) and might be generated by different mechanisms such as magma moving through narrow cracks, fragmentation and pulsation of pressurized fluids within the volcano, or escape of pressurized steam and gases from fumaroles (Konstantinou & Schlindwein, 2003). Volcanic tremors might be generated within different parts of volcanoes and might characterize different types of volcanic activity. They can occur either during eruptions and are, therefore, called eruptive tremor or before eruptions and are called noneruptive tremors (McNutt & Nishimura, 2008). In many volcanoes, noneruptive volcanic tremors constitute an important attribute of volcanic unrest and their detection and characterization is used for volcano monitoring and eruption forecasting (Chouet, 1996a; McNutt, 1992).

3.1. Single-Station Approach

As many volcanological observatories do all over the world, the KBGS monitors volcanic tremors using a simple single-station approach. For example, stations LGN and KMN are used for Klyuchevskoy and Tolbachik tremors, respectively (cf. Figure 1). This simple approach suffers from three main limitations. First, the tremor monitoring of one volcano is no longer operational if one of those reference stations (LGN or KMN) breaks down, which is often the case and it can take a long time to reestablish stations due to their locations in remote areas. Second, the very irregular trace of electronic noise possibly emitted by sensors or by the transmitting system can be confused with tremor signals. Third and most important, the use of a single station does not permit to spatially localize the tremor origin and to distinguish different tremor sources in case of simultaneous functioning. To overcome those limitations, network-based methods (or array-based methods) can be used.

3.2. Network-Based Approach

Network-based seismic methods have been originally developed in mid-1960s with the installation of the two first large-scale seismic arrays LASA (Frosch & Murray, 1966) and NORSAR (Bungum et al., 1971), in the context of the monitoring of nuclear explosions. Indeed, benefiting from signals simultaneously registered by many receivers, network-based seismic methods are able to detect and locate weak sources embedded in the background noise. So far, small-aperture arrays have been used to determine the location of tremor sources (e.g., Almendros et al., 1997; Goldstein & Chouet, 1994; Haney, 2010; Métaxian et al., 2002). Another approach is using the Real-time Seismic Amplitude Method (Endo & Murray, 1991) simultaneously with several receivers. However, it can be unstable in the presence of noise and it is therefore more efficient for eruptive tremors, when the signal is stronger. Therefore, here we use a method based on the array covariance matrix that enhances coherent signals within a network by reducing local noise. The method and its application for the characterization of both eruptive and noneruptive volcanic tremors are described in the following sections.

4. Method: Array Covariance Matrix Eigenvalues and Eigenvectors

The proposed method is based on the analysis of eigenvalues and eigenvectors of the seismic array covariance matrix, of which estimation is detailed in section 4.1. First, following the approach developed by Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016), the width of the covariance matrix eigenvalues distribution is analyzed in section 4.2 to detect time periods with strong volcanic tremors. Such tremors have already been evidenced with this method at the Piton de la Fournaise volcano in La Reunion Island, together with other types of seismic sources from oceanic (microseismic noise) and tectonic (local, regional, and teleseismic earthquakes) origins (Seydoux, Shapiro, de Rosny, Brenguier, & Landés, 2016). In a next step, the frequency-dependent first eigenvector of the covariance matrix is analyzed in section 4.3 considering that this first eigenvector represents the principal component of the daily seismic wavefield and, for days with tremor activity, characterizes the dominant tremor sources. Those first eigenvectors can therefore be used as network-based fingerprints of tremor sources. This approach constitutes a generalization to the whole network of the phase-matching detection algorithm derived on cross correlations of specific stations pairs by Droznin et al. (2015). The focus of our study being on long-duration tremors and not on short transient events, all inferences are obtained from daylong records.

4.1. Array Covariance Matrix Definition

In the following analysis, we consider only vertical component records. Let us consider the array data vector $\mathbf{u}(t) = \left[u_1(t)\, u_2(t)\, \ldots\, u_N(t)\right]^T$, where $u_i(t)$ is the seismic trace registered by the seismometer i and N the number of sensors ($N = 18$ in 2009 and $N = 19$ from 2010 onward in this article), and its Fourier transform in the frequency domain $\mathbf{u}(f) = \left[u_1(f)\, u_2(f)\, \ldots\, u_N(f)\right]^T$. From a mathematical point of view, the array covariance matrix $\mathbf{C}(f)$ is defined as the expected value of the cross-spectra product between $\mathbf{u}(f)$ and its transposed complex conjugate:

$$\mathbf{C}(f) = \mathbf{E}\left[\mathbf{u}(f)\, \mathbf{u}^{\dagger}(f)\right], \tag{1}$$

where **E** stands for the Expected Value and † denotes the Hermitian transpose (complex conjugate transpose).

When working with real data, the estimation of the array covariance matrix requires several steps. First, the used seismic records must be preprocessed to enhance certain types of signals that are targeted in the analysis and to diminish a possible influence of "nondesired" signals. In a second step, interstation covariances of preprocessed records are statistically estimated.

4.1.1. Preprocessing of Seismograms

In this study, we are interested in detecting volcanic tremors. Therefore, in the first step seismic traces are filtered between 10^{-1} Hz and 10^{1} Hz (we intentionally filter in a frequency band wider than the tremor spectral band) and downsampled from 128 Hz to 25.6 Hz to accelerate computations. Instrument responses are not removed because all sensors are identical. The next step of the preprocessing is the amplitude normalization. The idea of this normalization is to emphasize the network-wide coherence of the wavefield formed by long-duration tremors.

The simplest idea is to completely ignore the signal amplitude, as done in many applications based on cross correlation of ambient seismic noise (Bensen et al., 2007). This approach, referenced as the "classical" normalization in the following text, combines a spectral whitening and a temporal equalization of the signal amplitude. The spectral whitening consists of smoothing the signal spectrum with a df-long running average. The temporal equalization consists of smoothing the temporal signal with a dt-long running average. We used $df = 0.33$ Hz and $dt = 1.25$ s in this study.

The described classical normalization removes almost completely the amplitude information and is necessary to reduce the influence of strong impulsive signals (earthquakes). This approach is appropriate when analyzing the nearly continuous and stationary seismic tremors. However, in many cases the seismovolcanic tremors are not fully stationary in time and rather represent sequences of many impulsive events. In this case, the amplitude modulation of tremor signals contains useful information and it might be better to preserve it. At the same time, this is still important to preprocess the records to diminish the influence of strong earthquakes and of local noise. The solution is to normalize the signals with spectral whitening but without time equalization. The idea is that the information about the time amplitude modulation of tremors remains preserved. Since the normalization is achieved only with spectral whitening, this approach is referenced as "spectral" normalization in the following text.

4.1.2. Daily Covariance Matrix

The array covariance matrix is estimated according to the method proposed by Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016). Every filtered and downsampled trace is divided into time windows with overlapping of 25% (windows overlapping ratio $R = 1/4$) on which normalization (classical or spectral normalization) is applied. Then, every preprocessed window is subdivided into M overlapping subwindows of duration δ_t (subwindows overlapping ratio $r = 1/2$), on which $\mathbf{u}(f)\,\mathbf{u}^{\dagger}(f)$ cross-spectra matrices are computed. The array covariance matrix $\mathbf{C}(f)$ is then statistically estimated on each window of duration $\Delta_t = M\, r\, \delta_t$ as the average of those cross-spectra matrices:

$$\mathbf{C}(f) = \langle \mathbf{u}(f)\, \mathbf{u}^{\dagger}(f) \rangle_{\Delta t} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{u}_m(f)\, \mathbf{u}_m^{\dagger}(f) \tag{2}$$

The methodology to choose optimal parameters M and δ_t is detailed in Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016). The rank of the array covariance matrix is equal to the minimum between the number of sensors N ($N = 18$ or 19 in this article) and the number of subwindows M (Seydoux, Shapiro, de Rosny, & Landès 2016). We use $M = 50$ in this article, to ensure the covariance matrix estimation to be statistically robust. Likewise, the subwindows duration δ_t depends on the time scale of the phenomenon one wants to detect.

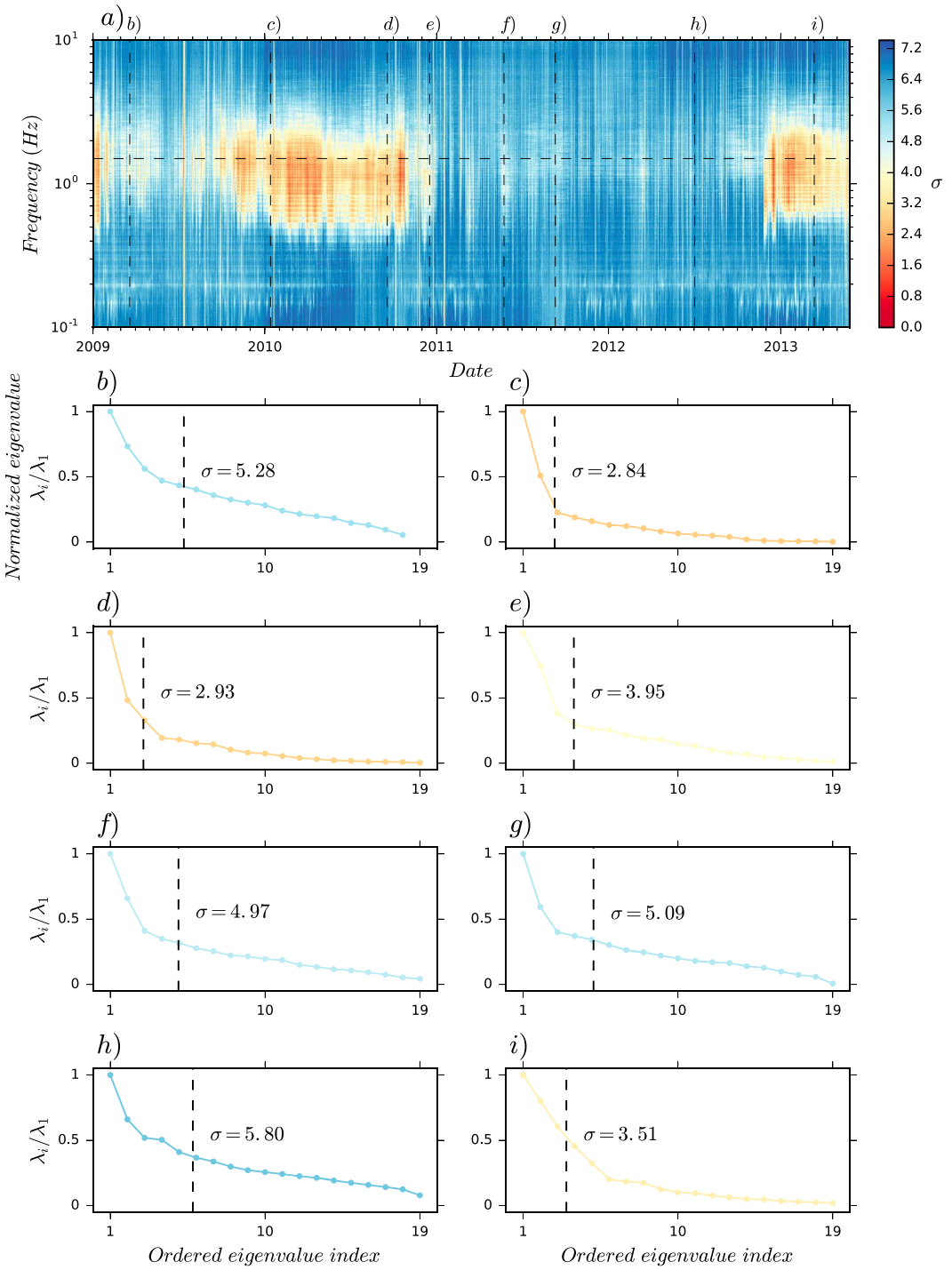

Figure 2. (a) Spectral width of the array covariance matrix eigenvalues distribution for the spectral normalization and the following parameters explained in the text: $M = 50$; $r = 1/2$; $\delta_t = 1{,}000$ s; $\Delta_t = M\,r\,\delta_t = 25{,}000$ s; and $R = 1/4$. This method developed by Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016) is efficient to detect the volcanic tremors at Klyuchevskoy (2009–2010) and Tolbachik (2013) but is not able to distinguish them. (b–g and i) Spectral width and eigenvalues distribution corresponding to the central days of clusters derived in section 5. (h) Calm period of seismic noise characterized by an high spectral width and a flat eigenvalues distribution.

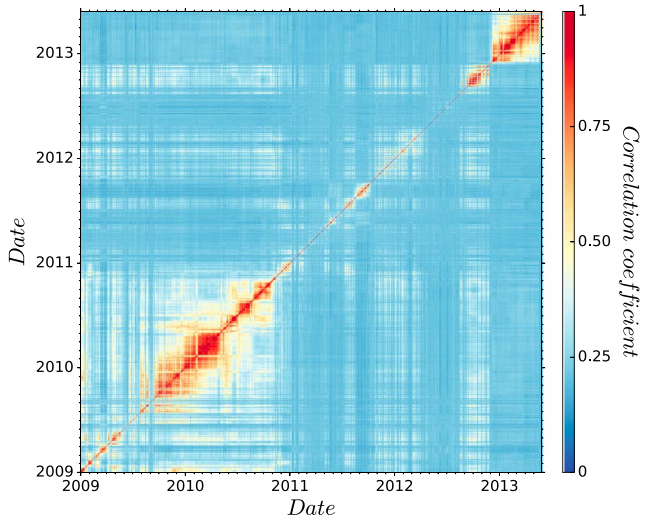

Figure 3. Matrix of correlation coefficients between all daily array covariance matrix's first eigenvector, for the spectral normalization. The correlation matrix is obtained by stacking such matrices in the frequency band 1.0 – 2.0 Hz where volcanic tremors are the more energetic as it appears on Figure 2. Correlation coefficients are here equal to the absolute value of the complex scalar product between the first eigenvectors of two different days normalized by their respective norm, according to equations (5) and (6).

In this study, we are interested in long-duration events and are not very concerned by resolving short-time duration signals. Therefore, we consider δ_t=1, 000 s, which is the optimal duration to detect both continuous tremors and time-modulated sequences of many impulsive events described in section 4.1.1. This duration results in a very detailed resolution in frequency and also insures that all possible simultaneously seismic waves are contained within one subwindow. Those parameters result in a window of duration $\Delta_t = M\ r\ \delta_t = 50 \times 1/2 \times 1,000 = 25,000$ s. One daily array covariance matrix is finally computed as the average of the $t_{\text{day}}/(\Delta_t \times R) = 86,400/(25,000 \times 1/4) = 13.8$ array covariance matrices of 1 day. This daily array covariance matrix is called "array covariance matrix" in the following text.

4.2. Array Covariance Matrix's Eigenvalues — Detection of Tremors

The array covariance matrix is inherently Hermitian and positive semidefinite (Seydoux, Shapiro, de Rosny, Brenguier, & Landés, 2016). Therefore, it can be decomposed on the basis of its eigenvectors $\mathbf{V}_i$ (analyzed in section 4.3) associated with real positive eigenvalues. According to Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016), the number of nonzero eigenvalues is related to the number of incoherent signals composing the wavefield and then of independent seismic sources. One localized source generates an array covariance matrix of rank one, and the rank value increases with the number of independent sources. This rank value could therefore theoretically be used to detect the number of sources recorded by the array.

However, real seismic data contain seismic noise (whose sources are multiple: hum, oceanic microseisms, wind, anthropic activity, scatterers acting like secondary sources, …) and electronic noise. It makes impossible using the covariance matrix rank to directly infer the number of independent sources. To solve this issue, Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016) introduced the scalar value of "spectral width" defined as the width of the covariance matrix eigenvalues distribution (with eigenvalues sorted in decreasing order):

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

This spectral width can be seen as a proxy for the number of independent seismic sources. Thus, the spectral width corresponding to ambient seismic noise produced by distributed noise sources is high, whereas the spectral width of a signal spatially coherent at the network scale generated by a single localized source is low.

The approach based on the covariance matrix spectral width has been successfully applied by Seydoux, Shapiro, de Rosny, Brenguier, and Landés (2016) to detect the 2010 volcanic tremor at the Piton de la Fournaise volcano in La Reunion Island. Here we apply this method on the KVG data. The frequency-dependent spectral width computed from daily covariance matrices during the studied period (4.5 years) is shown in Figure 2 for the spectral normalization. For comparison, the corresponding figure for classical normalization appears in Figure S1 of the supporting information S1. Those figures are different from spectrograms of individual stations and can be seen as "spectrograms of the whole network." The difference lies in the fact that spectrograms computed from one station can be affected by local noise, while Figures 2 and S1 just emphasize signals that are coherent across the whole network. Both figures clearly show three main long-duration episodes when the spectral width was reduced: (1) in the beginning of 2009, (2) from the end of 2009 to the end of 2010, and (3) from the end of 2012 to mid-2013. The two first episodes correspond to periods of strong noneruptive tremors at Klyuchevskoy and the last one to a strong eruptive tremor at Tolbachik (Droznin et al., 2015). This result confirms that the distribution of the network covariance matrix eigenvalues can be used to detect strong volcanic tremors. In addition to this, many short episodes with elevated wavefield coherence (decreases in the spectral width) are observed.

To overcome the limitation of this approach based on eigenvalues to distinguish different sources of volcanic tremor, the section 4.3 focuses on the eigenvectors of the daily array covariance matrix. Particularly the first

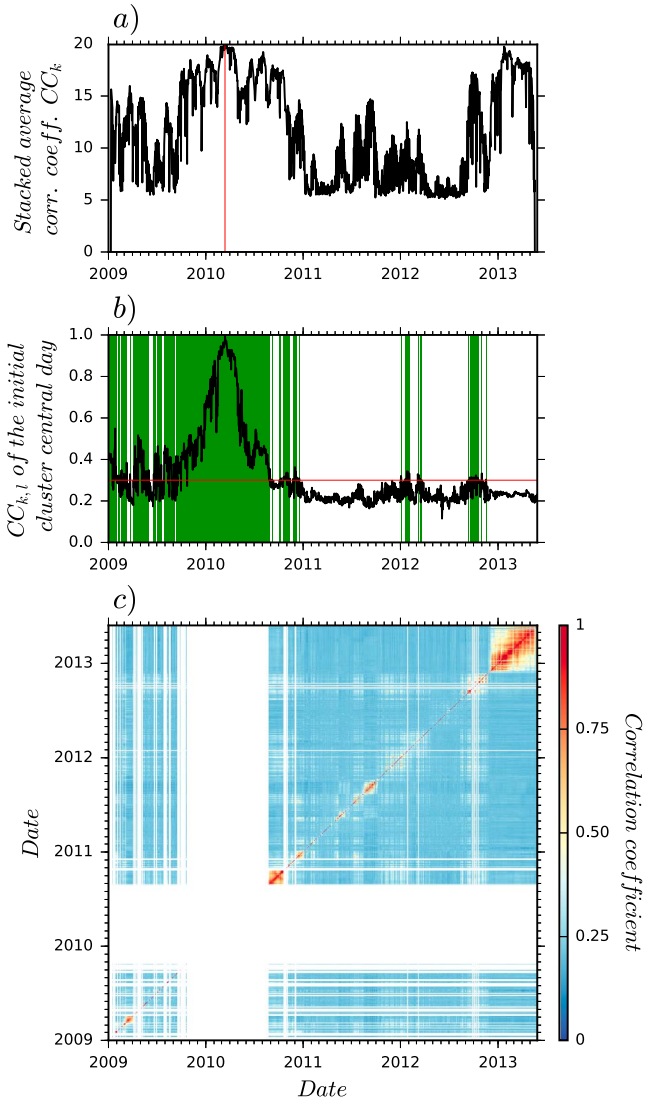

Figure 4. First step of the clustering process: creation of the initial clusters. Example of creation of the first initial cluster for the spectral normalization. (a) n_{days}-long moving window time stack of average correlation coefficients $CC_{k,l}$ to obtain stacked average correlation coefficient CC_k (black curve), and the initial cluster central point (vertical red line). (b) Line of the matrix of average correlation coefficients $CC_{k,l}$ corresponding to the initial cluster central day determined in Figure 4a (black curve), to determine days forming the initial cluster (green rectangles) according to a threshold (horizontal red line). (c) Days corresponding to the found cluster are removed from the average correlation coefficients matrix (equation (6) and Figure 3) before continuing up to the determination of the n_{clus} initial clusters.

eigenvector, with the hypothesis that it represents the principal component of the daily seismic wavefield and, for days with tremor activity, it characterizes the dominant tremor sources.

4.3. Array Covariance Matrix's First Eigenvector—Classification of Tremors

The array covariance matrix eigenvalues prove efficient to detect volcanic tremors, but it has to be noticed that it is not sufficient to distinguish different sources of tremor. So far, the distinction between Klyuchevskoy and Tolbachik tremors has been inferred from a priori knowledge (e.g., Droznin et al., 2015). A further step in characterizing tremors is to analyze the eigenvectors of the covariance matrix. In particular, the first eigenvector corresponding to the maximum eigenvalue represents the principal component of the analyzed seismic wavefield and, for periods with tremor activity, characterizes the dominant tremor source.

Let us consider a wavefield generated by a tremor source that remains at the same location and with a constant source mechanism. In this case, the elements $C_{i,j}(f)$ of the covariance matrix can be written as follows:

$$C_{i,j}(f) = |S(f)|^2 P_s(r_i, f) P_s^*(r_j, f) \tag{4}$$

where $|S(f)|^2$ is the tremor source power spectrum, asterisk stands for conjugate, r_i and r_j are positions of corresponding stations, and terms P_s describe the radiation from the tremor source to stations (convolution of the source mechanism with the media Green's functions). In the case of a wavefield dominated by such a single tremor source, the first eigenvector of its covariance matrix at a given frequency will be very close to $P_s(f)$ and can, therefore, be used as a network-based fingerprint of the source position and mechanism, or in other words, of a particular type of seismovolcanic activity. Following this idea, we compute daily network covariance matrices and extract their first eigenvectors that are then used to classify the volcanic tremor sources.

4.3.1. Measuring Similarity Between Covariance Matrix Eigenvectors—Multifrequency Correlation Coefficients

We use correlation coefficient as a measure of similarity between first eigenvectors from days k and l: $\mathbf{V}_k^1(f)$ and $\mathbf{V}_l^1(f)$. These latter being complex, the correlation coefficient at frequency f, $cc_{k,l}(f)$, is computed as the absolute value of the complex scalar product normalized by the respective norms:

$$cc_{k,l}(f) = \frac{|\mathbf{V}_k^1(f) \cdot \mathbf{V}_l^1(f)|}{\|\mathbf{V}_k^1(f)\| \|\mathbf{V}_l^1(f)\|}, \tag{5}$$

where the complex scalar product of two complex vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as the inner product of $\mathbf{x}$ and $\mathbf{y}^*$. We then compute average correlation coefficients, as the average of the n correlation coefficients $cc_{k,l}(f)$ calculated at frequencies between 1.0 Hz and 2.0 Hz (typical spectral band of volcanic tremors):

$$CC_{k,l} = \frac{1}{n} \sum_{f=1Hz}^{f=2Hz} cc_{k,l}(f) \tag{6}$$

An ensemble of such average correlation coefficients computed between all days during the considered period forms a large matrix shown in Figure 3 for the spectral normalization. For comparison, the corresponding figure for classical normalization appears in Figure S2 of the supporting information S1. This matrix has a specific structure and reveals several periods of elevated similarity between consecutive days, implying that the dominating tremor source remained stable during these periods. In the following section, we describe

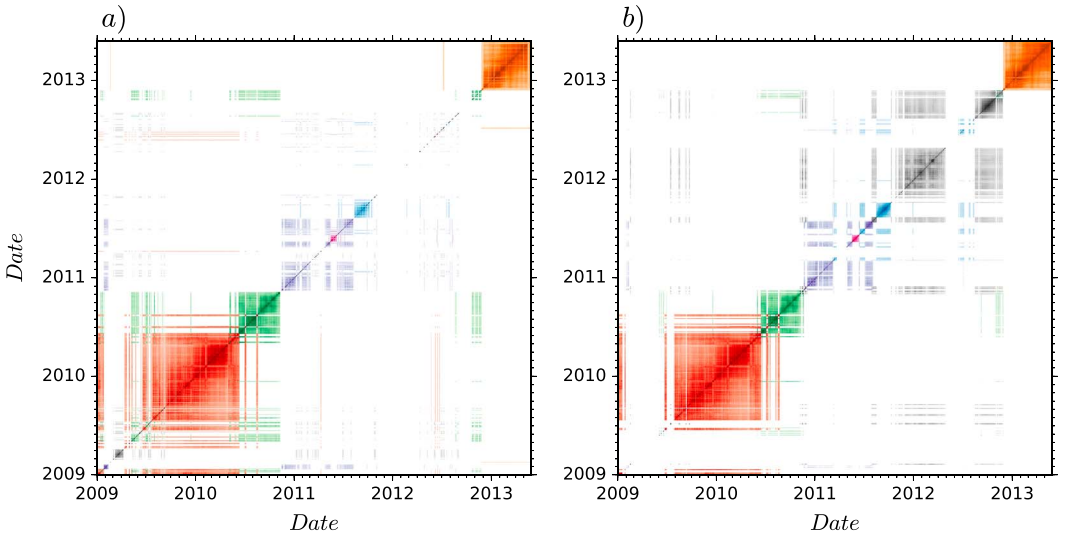

Figure 5. Time-time representation of the seven first clusters, associated with volcanic activity, of the $n_{clus} = 10$ clusters determined by the clustering process for two different normalizations. Similar clusters obtained by the two different normalizations are colored with the same color, and clusters obtained by only one normalization are colored in gray (cf. Table 1). (a) Spectral normalization. (b) Classical normalization (spectral whitening and temporal equalization).

how the interday correlation coefficients can be used to automatically identify with a clustering algorithm the periods when similar tremor sources were acting.

4.3.2. Clustering of Tremor Sources

The clustering process works in two steps: the first step is the creation of the initial clusters, and the second step entails their iterative resorting to obtain the final clusters.

Step 1. The different stages of the first step of creation of the initial clusters, which are disjointed in terms of days, are detailed in Figure 4 for the case of spectral normalization. First, the average correlation coefficients $CC_{k,l}$ (equation (6) and Figure 3) are stacked in time through a n_{days}-long moving window, to obtain stacked average correlation coefficients of day k, CC_k (black curve in Figure 4a). The duration of the stacking moving window is chosen equal to 20 days ($n_{days} = 20$), which approximately corresponds to the lower limit of duration of tremor episodes that we want to classify:

$$CC_k = \sum_{l=k-10\,\text{days}}^{l=k+10\,\text{days}} CC_{k,l} \tag{7}$$

The maximum of this time stack determines the initial cluster central point (vertical red line in Figure 4a). Then the line (or column) of the matrix of average correlation coefficients corresponding to this initial central day is considered (black curve in Figure 4b), and only the days presenting a correlation higher than a threshold s_{cc} (horizontal red line in Figure 4b) are kept in the cluster (green rectangles in Figure 4b). Days corresponding to the found cluster are then removed from the average correlation coefficients matrix (Figure 4c), and the process is continued up to the determination of the n_{clus} initial clusters.

Step 2. Once the n_{clus} initial clusters are created, the second step of their resorting is realized in an iterative way. For each iteration, the resorting contains two stages. First, the central day of each of the n_{clus} initial clusters is redetermined by time stacking the average correlation coefficients $CC_{k,l}$ not any more through a n_{days}-long moving window, but through a window containing only the days of the concerned cluster. Second, each day of the studied period (between January 2009 and June 2013) is compared to the n_{clus} cluster's redefined central days and placed in the cluster where the correlation is the highest. With the parameters $f_{band} = (1.0–2.0)$ Hz, $n_{clus} = 10$, $n_{days} = 20$, and $s_{cc} = 30\%$ used in this article, three iterations are needed for this iterative process to converge, in the sense that the n_{clus} cluster's redefined central days are no more changing between two consecutive iterations.

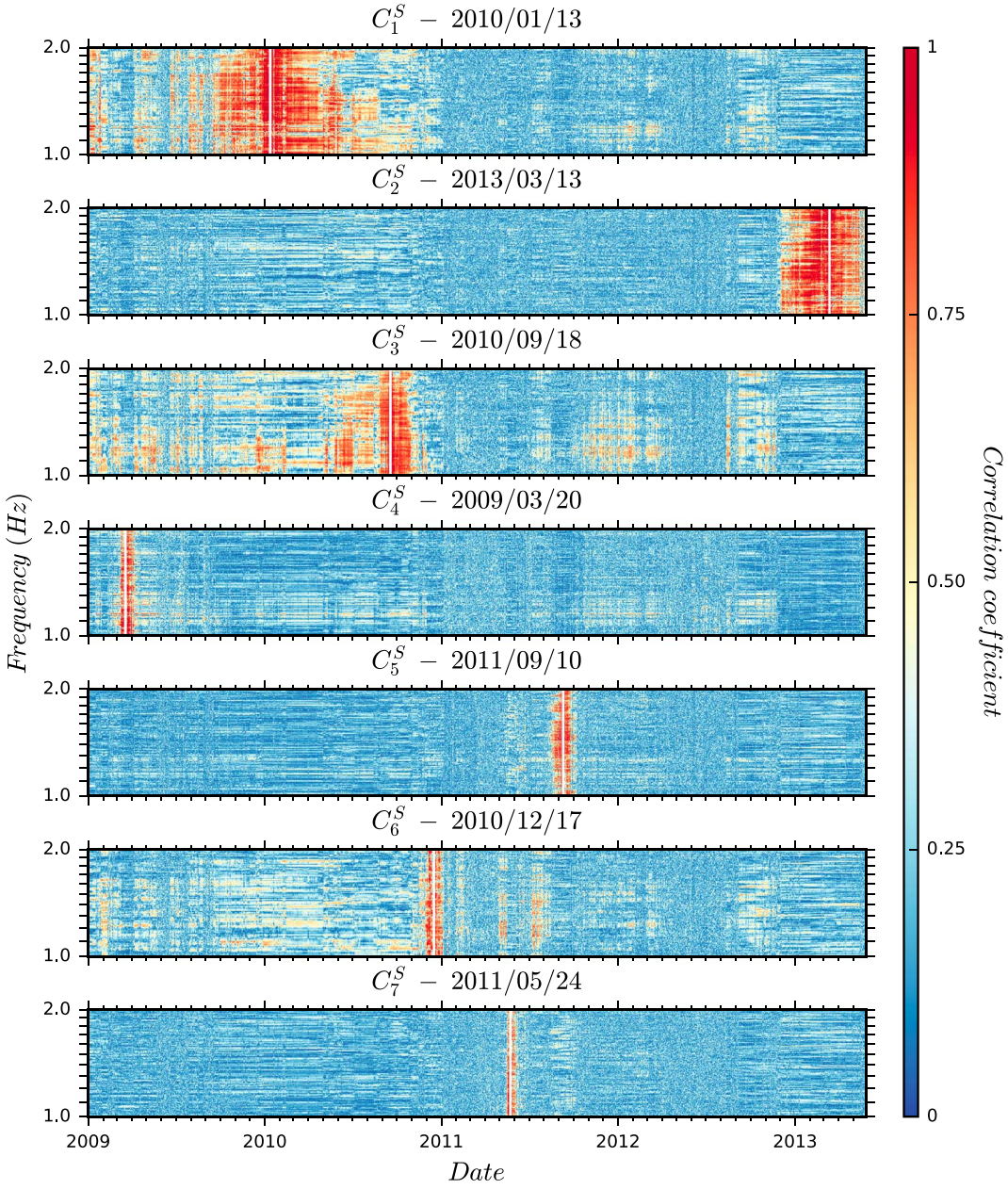

Figure 6. Time-frequency representation of the seven first clusters, associated with volcanic activity, of the $n_{clus} = 10$ clusters for the spectral normalization. The color represents the degree of correlation (projection in the sense of complex scalar product) of the characteristic eigenvectors (white lines) on the array covariance matrix's first eigenvector at every day and every frequency. Cluster C_1^S is related to the Klyuchevskoy volcano which was very active in 2009 and 2010, with an increase of activity leading to a summit eruption in October 2010 corresponding to cluster C_3^S. Cluster C_2^S is clearly associated with the Tolbachik volcano which experienced an important volcanic tremor from December 2012 when all the other volcanoes were quite quiet. Cluster C_4^S is related to the Shiveluch volcano that was very active during quasi all the studied time period, with a peak of activity characterized by really dense series of long-period events at the beginning of 2009. And clusters C_5^S, C_6^S, and C_7^S are associated with the Kizimen volcano.

5. Results: Detection and Classification of Volcanic Tremor Sources

We apply the clustering algorithm described in the previous section to the daily eigenvectors of network covariance matrices computed from the records after the classical and spectral normalizations to obtain two sets of 10 clusters. The first seven clusters, possibly related to seismovolcanic activity, are shown in Figure 5 for the spectral and classical normalizations. For simplicity, in the following text we name clusters as C_K^C

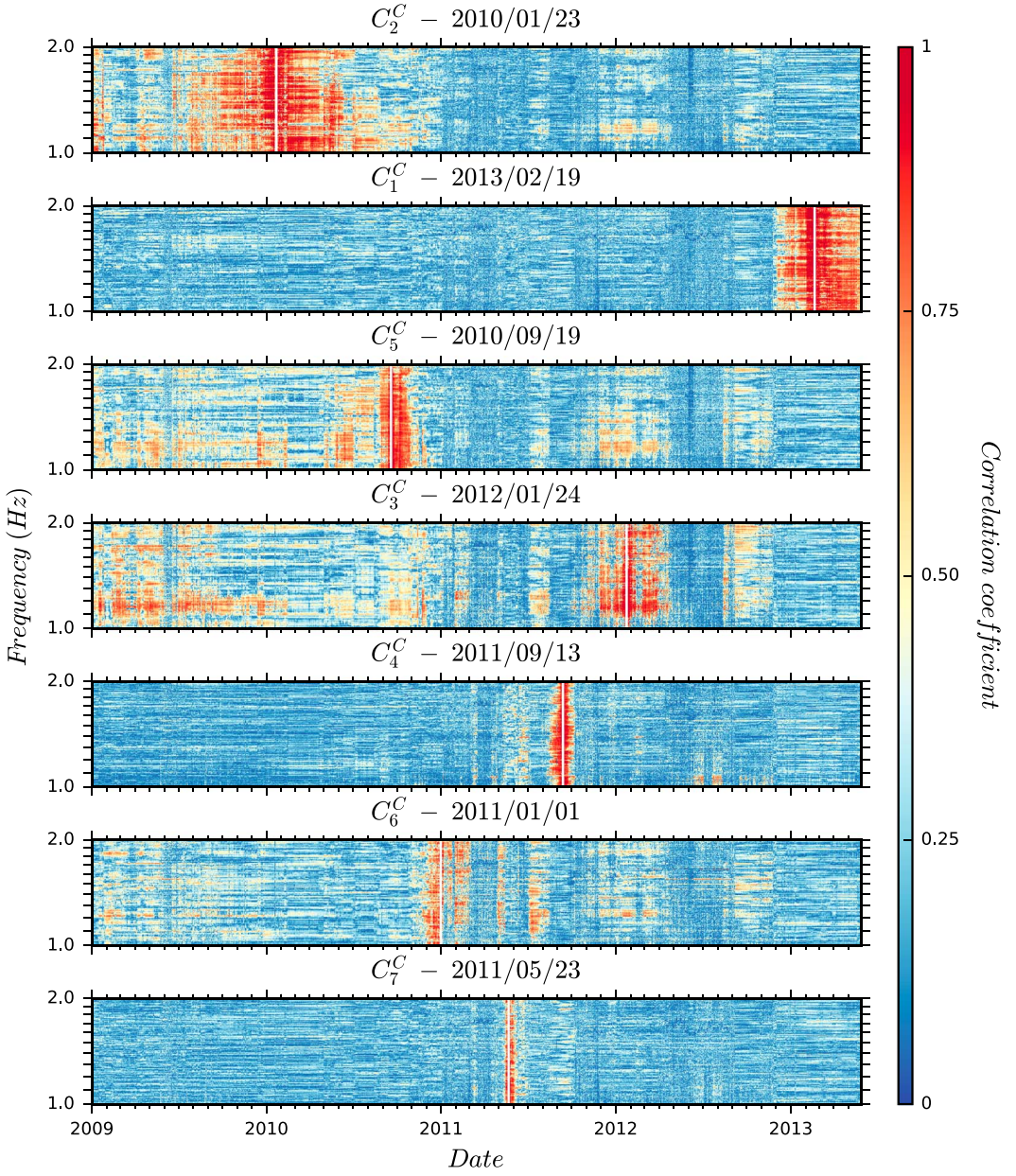

Figure 7. Identical to Figure 6 but for preprocessing corresponding to classical normalization. The order of the clusters has been modified for an easier comparison with Figure 6 (cf. Table 1). Here cluster C_1^C is associated with the Tolbachik volcano which experienced an important volcanic tremor from December 2012 when all the other volcanoes were quite quiet. Cluster C_2^C is related to the Klyuchevskoy volcano which was very active in 2009 and 2010, with an increase of activity leading to a summit eruption in October 2010 corresponding to cluster C_5^C. Clusters C_4^C, C_6^C, and C_7^C are associated with the Kizimen volcano.

and C_K^S, where superscripted letters ^{C} and ^{S} correspond to the classical and spectral normalizations and K indicates the respective cluster number.

Each cluster central day, iteratively determined as described in section 4.3.2, corresponds to the most characteristic first eigenvector of each cluster. Those characteristic eigenvectors can be used as fingerprints for the matching detection of tremor sources, comparing them with the daily eigenvectors from the continuous records. Figures 6 and 7 show the projection (complex scalar product) in the time-frequency domain of the array covariance matrix's first eigenvector at every day and every frequency onto those characteristic eigenvectors (white lines in Figures 6 and 7), respectively for spectral and classical normalization. This representation shows the time and frequency extent of each classified tremor source.

Table 1
Summary of Covariance Matrix Eigenvector Clusters Corresponding to the Seismovolcanic Sources

Spectral		Classical		
Cluster	Date	Cluster	Date	Origin
C_1^S	2010/01/13	C_2^C	2010/01/23	Klyuchevskoy
C_2^S	2013/03/13	C_1^C	2013/02/19	Tolbachik
C_3^S	2010/09/18	C_5^C	2010/09/19	Klyuchevskoy
C_4^S	2009/03/20	-	-	Shiveluch
C_5^S	2011/09/10	C_4^C	2011/09/13	Kizimen
C_6^S	2010/12/17	C_6^C	2011/01/01	Kizimen
C_7^S	2011/05/24	C_7^C	2011/05/23	Kizimen
-	-	C_3^C	2012/01/24	-

Note. Clusters obtained after the spectral and classical normalizations are shown in left and right columns, respectively. Each line shows clusters corresponding to similar source (origin).

First, we note that most of the clusters obtained by the two different data preprocessing methods are very similar, even if their respective order is not the same. The similarity between these clusters is summarized in Table 1. The first three clusters in Table 1 are clearly related to the activity of Klyuchevskoy and Tolbachik volcanoes already detected with the covariance matrix spectral width analysis (Figures 2 and S1). The spatial origin of tremors forming these clusters is based on a priori knowledge from Droznin et al. (2015) about the activity periods of those two volcanoes.

Another approach is directly using the moveout information contained in the covariance matrix eigenvectors. According to equation (4), the first eigenvector approximately corresponds to the wavefield emitted by the dominating tremor source at a given frequency. For each one of the ensemble of such first eigenvectors corresponding to the central day of each cluster, we compute the outer product at frequencies within the tremor frequency band (1 – 2 Hz)

$$C_{i,j}^1(f) = V_i^1(f)V_j^{1*}(f), \tag{8}$$

and take the inverse Fourier transform to estimate the time cross-correlation function corresponding to this first eigenvector (i.e., to the dominant tremor source at the central date of the considered cluster). Seismic waves are expected to propagate with some average velocity from the tremor source. Therefore, in cross correlations we expect the arrival time of maximal amplitude to approximately align with difference in distances between the respective stations and the source location (Ballmer et al., 2013). This property can be used to test different possible source locations. If the tested location is elected close to the correct one, the maximal amplitudes in the cross-correlation waveforms are expected to align with respect to the distance difference. Such a test for clusters C_2^S (13 March 2013) is illustrated in Figure 8. It clearly shows that a good alignment is obtained when selecting a correct reference source location, that is, the Tolbachik volcanic edifice, and conversely an incorrect source location at station SRD (positioned at the north-west of the network cf. Figure 1) leads to a bad alignment. Similar figures corresponding to clusters C_1^S and C_3^S related to Klyuchevskoy volcano appear in supporting information S1 as Figures S3 and S4.

An interesting result is that tremor activity at Klyuchevskoy is clearly separated in two clusters. Cluster C_1^S (13 January 2010) can be classified as noneruptive tremor (McNutt & Nishimura, 2008) and cluster C_3^S (18 September 2010) corresponds to most energetic eruptive final phase. This implies that the source of dominating tremor was modified getting closer to the eruption. Possible explanation of this modification is either that the tremor-generating activity migrated toward the surface getting closer to the eruption or that the parameters of the tremor-generating process changed.

A part the activity of Klyuchevskoy and Tolbachik, our method detected several clusters during periods when these volcanoes were not erupting. The cross-correlation moveout analysis illustrated in supporting information S1 shows that signals corresponding to clusters C_5^S, C_6^S, and C_7^S originate from the Kizimen volcano that was erupting starting in December 2010 and till the end of 2013 (Figures S6 – S8). Seismic activity

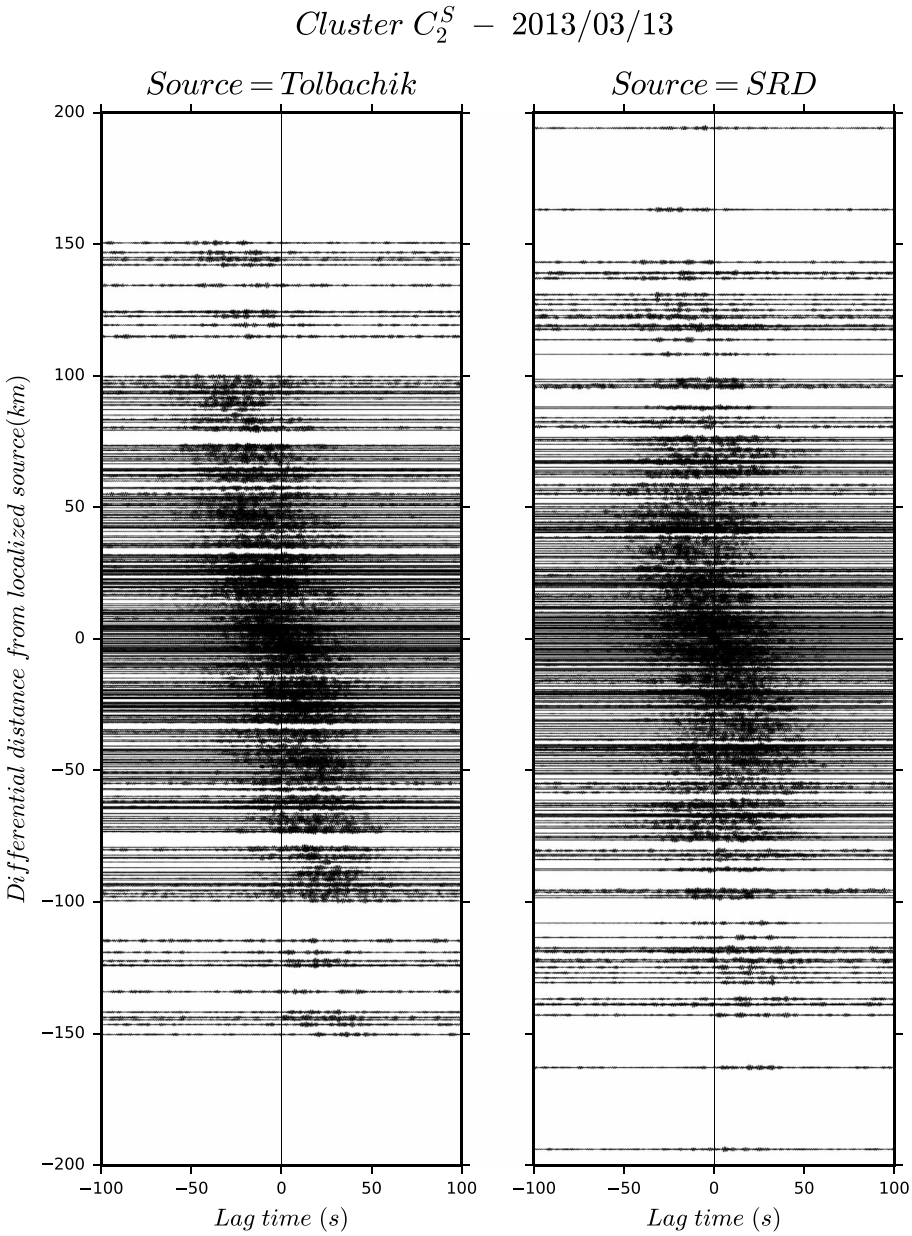

Figure 8. Comparison of cross correlations of cluster C_2^S (13 March 2013) plotted against differential distance of station pairs to a localized source, for two different source positions. (left) Tremor source located in Tolbachik. (right) Source located at station SRD positioned at the north-west of the network (cf. Figure 1). Only the good source location (left case) shows an alignment of cross-correlations characteristics of a localized source.

associated with this long eruption was very irregular (Firstov & Shakirova, 2014). The initiation of the eruption in December 2010 was accompanied by a very intense seismic crisis with most of events located within the volcanic edifice. This crisis is likely detected by our method as cluster C_6^S (17 December 2010). An example of daily seismogram recorded during this crisis is shown in Figure 9c. Several episodes of very frequent long-period earthquakes sometimes called as "drumbeats" occurred in Kizimen in 2011–2012 (Firstov & Shakirova, 2014). Two strongest periods of this drumbeats correspond to clusters C_5^S (10 September 2011) and C_7^S (24 May 2011), as illustrated with seismograms recorded at station KZV (the closest to Kizimen) in Figures 9b–9d.

Finally, the cross-correlation moveout analysis shows that cluster C_4^S (20 March 2009) originated from Shiveluch (Figure S5 in the supporting information S1). A close look at seismograms from closely located

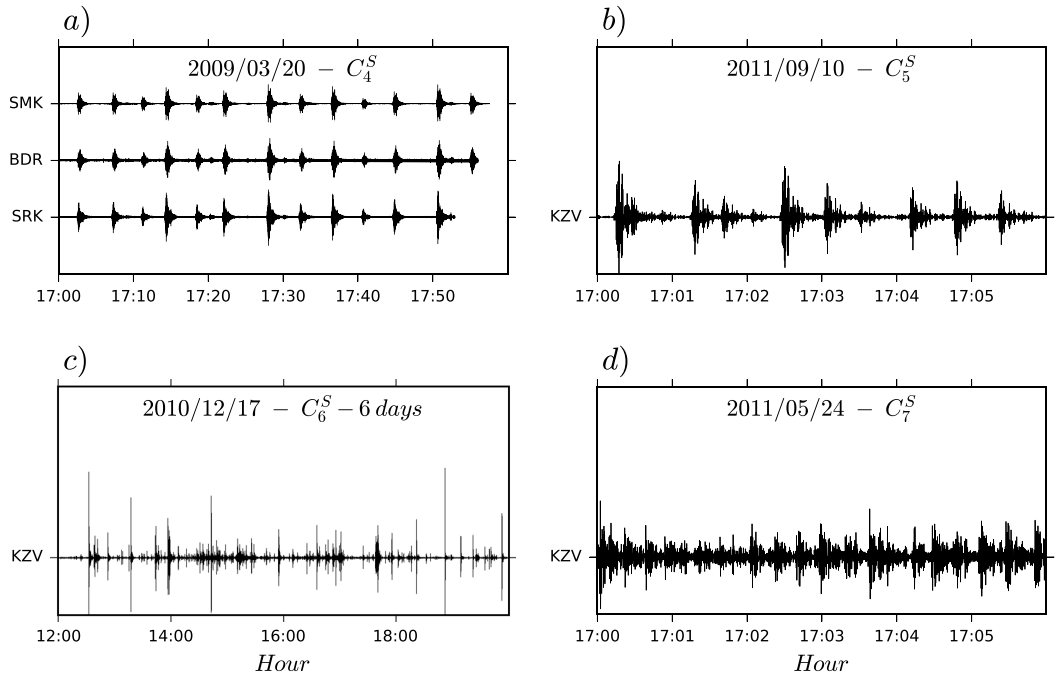

Figure 9. Raw seismic vertical traces for the central day of different clusters (for the spectral normalization) at stations located close to the volcanoes of the studied area (cf. Figure 1 for the precise location of stations). (a) Traces of stations SMK, BDR, and SRK located close to the Shiveluch volcano for the central day of cluster C_4^S (20 March 2009). (b) Trace of station KZV located close to the Kizimen volcano for the central day of cluster C_5^S (10 September 2011). (c) Trace of station KZV located close to the Kizimen volcano corresponding to 6 days before the central day of cluster C_6^S (11 Decmber 2010). (d) Trace of station KZV located close to the Kizimen volcano for the central day of cluster C_7^S (24 May 2011).

stations (Figure 9a) shows that this cluster corresponds to an episode of frequent long-period earthquakes at Shiveluch similar to those at Kizimen that produced clusters C_5^S and C_7^S. This cluster has been only detected with using the spectral normalization. The contribution of this sequence of long-period events on the covariance matrix was almost removed when applying the temporal equalization of seismograms in the classical preprocessing. The reason is that the seismovolcanic signal emitted from Shiveluch are not continuous in time with the long-period earthquakes separated by 1 or 2 min of noise. In this case, temporal equalization of the signal amplitude results in the dominant contribution of the noise in the covariances. This example shows that the data preprocessing without temporal equalization might be more adequate for the detection of tremors based on interstation covariances (cross correlations).

6. Discussion and Conclusions

The results presented in this study show that the analysis based on the network covariance matrix can be successfully used to detect and to classify seismovolcanic tremors. The distribution of eigenvalues of this matrix can be used to detect the episodes of strongest tremor activity. More advanced analysis of the covariance matrix eigenvectors detects weaker signals even from sources located on the periphery of the seismic network, as in the case of Kizimen and Shiveluch volcanoes.

The results of the analysis depend on the type of normalization used for preprocessing (classical or spectral normalization) and on the choice of the parameters used for the covariance matrix computing (the number of subwindows M and their duration δ_t) and in the clustering algorithm (the frequency band for stacking correlation coefficients matrices f_{band}, the number of searched clusters n_{clus}, the length of summing moving windows n_{days}, and the cluster's correlation threshold s_{cc}). These parameters can be chosen in order to emphasize certain type of signals and based on the properties of the used seismic network. In this paper, we focused on detection of relatively long episodes of tremors in a frequency range between 1 and 2 Hz, typical for the Klyuchevskoy volcano group. Therefore, we analyzed the covariance matrices computed from daylong time

series. The time resolution of the analysis could be significantly enhanced with using shorter time windows. This could be useful when focusing on details of individual tremor episodes, which is out of scope of the present work.

The choice of the number of searched clusters n_{clus} depends on the potential number of detected sources. In this study of the volcanic system of the KVG with several erupting volcano and several episodes of activity, we used a relatively high number ($n_{clus} = 10$) and retrieved seven separate clusters of seismovolcanic activity. In many other volcanic systems, the possible number of existing sources might be not as large as in the Klyuchevskoy volcanic group and in such cases this could be preferable to use smaller values of n_{clus} in order to make the detection more robust. The choice of the length of summing moving windows n_{days} depends of the lower limit of duration on events that one wants to classify. We used $n_{days} = 20$ because it corresponds to the shortest duration of long-period seismic events that occurred in the KVG during the time period of the study. Choosing a smaller length n_{days} has no influence on the detected clusters but on the order on which they appear. But choosing a length n_{days} too high results in loosing shortest clusters.

A very important aspect of the developed method is that, once all parameters selected, the data analysis and following detection and classification of sources becomes fully automatic without need of additional a priori information. Therefore, our method belongs to the family of machine learning methods. The algorithm includes three main steps of machine learning-based signal analysis: (1) data preprocessing, (2) feature extraction (characteristic first eigenvectors), and (3) learning (clustering) (Goodfellow et al., 2016). The main difference of our method comparing to previous implementations of the machine learning for the analysis of the seismovolcanic signals is that we extract "features" not from single-component signals but from an ensemble of records by a network of stations. We use as signal features the eigenvectors of the networks covariance matrix. The advantage of such features is that these eigenvectors contain an information about the location and the mechanism of the dominating tremor source (equation (4)).

More exactly, the additional information from the network (comparing to scalar signal properties) is contained in the phase of the interstation covariances which is related to a combination of the moveout of waves traveling from the tremor source, and of the source mechanism. Therefore, an ensemble of these covariances from a network could be also used as a feature for future classification. However, the dimension of this object is significantly larger than the dimensions of the eigenvectors, which make the following classification less robust. In addition, determination of the first covariance matrix eigenvector is an additional step in signal processing that extracts the main component of the recorded wavefield corresponding to the dominant source. Figure 10 shows a comparison of time cross-correlations obtained from inverse Fourier transforms of a full covariance matrix with those estimated from the first eigenvector (equation (8)) during the centroid date of cluster C_2^S. We can see that with using the first eigenvector instead of the whole matrix, we remove some noise at differential distances above 50 km. As a result, cross correlations obtained from the first eigenvector reflect slightly better the propagation from the dominant source (Tolbachik volcano).

With the described method, we detected seven separate episodes with different sources corresponding to tremor activity at four volcanoes: Klyuchevskoy, Tolbachik, Kizimen, and Shiveluch. While some of these sources have been previously described in the literature, the burst of long-period events at Shiveluch (cluster C_4^S) is reported for the first time here. Future studies of volcanic tremors at volcanoes of the Klyuchevskoy group might focus on details of individual tremor episodes with applying the analysis with a better time resolution. The periods when more than one volcano emitted tremors simultaneously might be particularly interesting and challenging for the analysis. This could possibly occur in 2010, when Klyuchevskoy and Shiveluch were simultaneously active and in 2012 when some activity at Klyuchevskoy continued during the Tolbachik eruption. Separation of such simultaneous tremor sources might require that additional eigenvectors of the network covariance matrix be analyzed. Another option could be to include horizontal components in the analysis in order to bring polarization constraint in the classification process (Vidale, 1986).

The detection and classification of volcanic tremors is also important for improving the noise-based imaging and monitoring of the Klyuchevskoy group of volcanoes. Previous studies (e.g., Droznina et al., 2017) have shown that using standard data preprocessing used in the noise-based seismology (Bensen et al., 2007), the stable noise cross correlations could not be extracted at frequencies higher than 0.5 Hz because of the "contamination" with tremor signals. More advanced data preprocessing might include two additional steps. First, we can select for analysis only "calm" periods without tremor when simple noise-based techniques

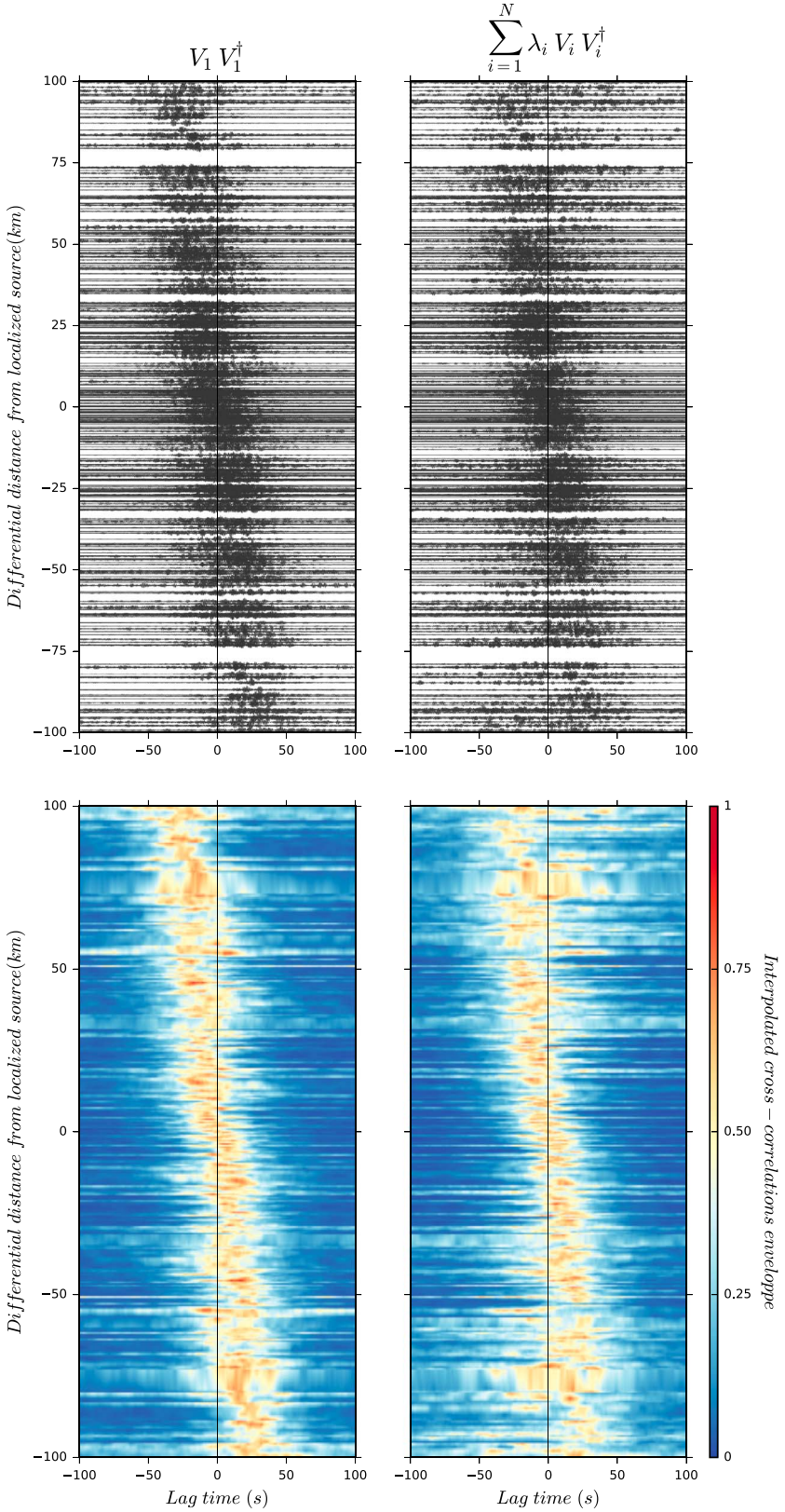

Figure 10. Comparison of cross correlations of cluster C_2^S (13 March 2013) obtained from the first eigenvector (left) or from the full covariance matrix (right). (top) Individual cross correlations. (bottom) Two-dimensional interpolated cross correlations corresponding to the smoothed envelope of the individual cross correlations linearly interpolated.

can be directly applied. A second and potentially more powerful step would be filtering out the detected tremors with the network-based approach proposed by Seydoux et al. (2017) to expand the periods on which noise-based tomography and monitoring could be realized.

Acknowledgments
We thank Kathi Unglert and Mark Jellinek for helpful discussions. This study was supported by the Russian Ministry of Education and Science (grant N 14.W03.31.0033) and by the French project "Labex UnivEarth" and the Université Sorbonne Paris Cité project "VolcanoDynamics". Computations were performed using the infrastructure of the Institute of Volcanology and Seismology and of the Kamchatka Branch of the Geophysical Survey, as well as the IPGP High-Performance Computing infrastructure S-CAPAD (supported by the Ile-de-France region via the SEASAME program, by France-Grille, and by the CNRS MASTODONS program). Seismological time series used for the analysis are available from the Kamchatka Branch of the Geophysical Survey of Russian Academy of Sciences (http://www.emsd.ru) on request.

References

Acernese, F., Ciaramella, A., De Martino, S., De Rosa, R., Falanga, M., & Tagliaferri, R. (2003). Neural networks for blind-source separation of Stromboli explosion quakes. *IEEE Transactions on Neural Networks*, *14*(1), 167–175. https://doi.org/10.1109/TNN.2002.806649

Almendros, J., Ibáñez, J., Alguacil, G., Del Pezzo, E., & Ortiz, R. (1997). Array tracking of the volcanic tremor source at Deception Island, Antarctica. *Geophysical Research Letters*, *24*(23), 3069–3072. https://doi.org/10.1029/97GL03096

Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *The Annals of Mathematical Statistics*, *34*(1), 122–148.

Balesta, S., Gontovaya, L., & Kargapol'tsev, A. (1991). Results of seismic investigations of the Earth's crust near Klyuchevskoi volcano, vulkanol.

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*, *194*, 512–523. https://doi.org/10.1093/gji/ggt112

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.

Bungum, H., Husebye, ES, & Ringdal, F. (1971). The NORSAR array and preliminary results of data analysis. *Geophysical Journal International*, *25*(1–3), 115–126.

Cabras, G., Carniel, R., & Jones, J. (2012). Non-negative matrix factorization: An application to Erta 'Ale volcano, Ethiopia. *Bollettino di Geofisica Teorica ed Applicata*, *53*(2), 231–242. https://doi.org/10.4430/bgta0056

Cabras, G., Carniel, R., Jones, J. P., & Takeo, M. (2014). Reducing wind noise in seismic data using non-negative matrix factorization: An application to Villarrica volcano, Chile. *Geofísica Internacional*, *53*(1), 77–85. https://doi.org/10.1016/S0016-7169(14)71491-6

Cabras, G., Carniel, R., & Wassermann, J. (2008). Blind source separation: An application to the Mt. Merapi volcano, Indonesia. *Fluctuation and Noise Letters*, *8*(03n04), L249–L260. https://doi.org/10.1142/S0219477508005124

Cabras, G., Carniel, R., & Wassermann, J. (2010). Signal enhancement with generalized ICA applied to Mt. Etna volcano, Italy. *Bollettino di Geofisica Teorica ed Applicata*, *51*(1), 57–73.

Capuano, P., De Lauro, E., De Martino, S., & Falanga, M. (2016). Detailed investigation of long-period activity at Campi Flegrei by convolutive independent component analysis. *Physics of the Earth and Planetary Interiors*, *253*, 48–57. https://doi.org/10.1016/j.pepi.2016.02.003

Carniel, R. (1996). Neural networks and dynamical system techniques for volcanic tremor analysis. *Annals of Geophysics*, *39*(2), 241–252. https://doi.org/10.4401/ag-3967

Carniel, R. (2014). Characterization of volcanic regimes and identification of significant transitions using geophysical data: A review. *Bulletin of Volcanology*, *76*(8), 848. https://doi.org/10.1007/s00445-014-0848-0

Chebrov, V., Droznin, D., Kugaenko, Y. A., Levina, V., Senyukov, S., Sergeev, V., … Yashchuk, V. (2013). The system of detailed seismological observations in Kamchatka in 2011. *Journal of Volcanology and Seismology*, *7*(1), 16–36. https://doi.org/10.1134/S0742046313010028

Chouet, B. A. (1996a). New methods and future trends in seismological volcano monitoring, *Monitoring and mitigation of volcano hazards* (pp. 23–97). New York: Springer.

Chouet, B. A. (1996b). Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature*, *380*(6572), 309–316.

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. https://doi.org/10.1016/j. jvolgeores.2012.11.013

Ciaramella, A., De Lauro, E., Falanga, M., & Petrosino, S. (2011). Automatic detection of long-period events at Campi Flegrei caldera (Italy). *Geophysical Research Letters*, *38*, L18302. https://doi.org/10.1029/2011GL049065

Dorendorf, F., Wiechert, U., & Wörner, G. (2000). Hydrated sub-arc mantle: A source for the Kluchevskoy volcano, Kamchatka/Russia. *Earth and Planetary Science Letters*, *175*(1), 69–86. https://doi.org/10.1016/S0012-821X(99)00288-5

Droznin, D., Shapiro, N., 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. https://doi.org/10.1093/gji/ggv342

Droznina, S. Y., Shapiro, N. M., Droznin, D. V., Senyukov, S. L., Chebrov, V. N., & Gordeev, E. I. (2017). *S*-wave velocity model for several regions of the Kamchatka peninsula from the cross correlations of ambient seismic noise. *Physics of the Solid Earth*, *53*(3), 341–352.

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.

Firstov, P., & Shakirova, A. (2014). Seismicity observed during the precursory process and the actual eruption of Kizimen volcano, Kamchatka in 2009–2013. *Journal of Volcanology and Seismology*, *8*(4), 203–217. https://doi.org/10.1134/S0742046314040022

Frosch, RA, & Murray Jr., P. E. (1966). The concept of a large aperture seismic array. In *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* (pp. 368–384).

Goldstein, P., & Chouet, B. (1994). Array measurements and modeling of sources of shallow volcanic tremor at Kilauea volcano, Hawaii. *Journal of Geophysical Research*, *99*(B2), 2637–2652.

Gontovaya, L., Stepanova, M., Khrenov, A., & Senyukov, S. (2004). Depth model of the lithosphere in the region of the Klyuchevskoy group of volcanoes (Kamchatka), vulkanol.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. Retrieved from http://www.deeplearningbook.org

Gorbatov, A., Dominguez, J., Suarez, G., Kostoglodov, V., Zhao, D., & Gordeev, E. (1999). Tomographic imaging of the *p*-wave velocity structure beneath the Kamchatka peninsula. *Geophysical Journal International*, *137*(2), 269–279. https://doi.org/10.1046/j.1365-246X.1999.t01-1-00801.x

Gordeev, E., Chebrov, V., Levina, V., Senyukov, S., Shevchenko, Y. V., & Yashchuk, V. (2006). The system of seismological observation in Kamchatka, vulkanol.

Gordeev, E., Melnikov, Y. Y., Sinitsyn, V., & Chebrov, V. (1989). Volcanic tremor of Kliuchevskoi volcano (1984 eruption), *Volcanic Hazards* (pp. 486–503). Berlin: Springer.

Gordeev, E., Murav'ev, Y. D., Samoilenko, S., Volynets, A., Mel'nikov, D., & Dvigalo V. (2013). The Tolbachik fissure eruption of 2012–2013: Preliminary results. *Doklady Earth Sciences*, *452*, 1046–1050. https://doi.org/10.1134/S1028334X13100103

Haney, M. M. (2010). Location and mechanism of very long period tremor during the 2008 eruption of Okmok volcano from interstation arrival times. *Journal of Geophysical Research*, *115*, B00B05. https://doi.org/10.1029/2010JB007440

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. https://doi.org/10.1016/j.jvolgeores.2017.04.015

Ivanov, V. (2008). Current cycle of the Kluchevskoy volcano activity in 1995–2008 based on seismological, photo, video and visual data. In *Proceedings of Conference, Petropavlovsk-Kamchatsky* (pp. 27–29).

Ivanov, A., Koulakov, I. Y., West, M., Jakovlev, A., Gordeev, E., Senyukov, S., & Chebrov, V. (2016). Magma source beneath the Bezymianny volcano and its interconnection with Klyuchevskoy inferred from local earthquake seismic tomography. *Journal of Volcanology and Geothermal Research*, *323*, 62–71. https://doi.org/10.1016/j.jvolgeores.2016.04.010

Iverson, R. M., Dzurisin, D., Gardner, C. A., Gerlach, T. M., LaHusen, R. G., Lisowski, M., … Vallance, J. W. (2006). Dynamics of seismogenic volcanic extrusion at Mount St Helens in 2004–05. *Nature*, *444*(7118), 439–443. https://doi.org/10.1038/nature0532

Konstantinou, K. I., & Schlindwein, V. (2003). Nature, wavefield properties and source mechanism of volcanic tremor: A review. *Journal of Volcanology and Geothermal Research*, *119*(1), 161–187.

Koulakov, I., Gordeev, E. I., Dobretsov, N. L., Vernikovsky, V. A., Senyukov, S., & Jakovlev, A. (2011). Feeding volcanoes of the Kluchevskoy group from the results of local earthquake tomography. *Geophysical Research Letters*, *38*, L09305. https://doi.org/10.1029/2011GL046957

Koulakov, I., Gordeev, E. I., Dobretsov, N. L., Vernikovsky, V. A., Senyukov, S., Jakovlev, A., & Jaxybulatov, K. (2013). Rapid changes in magma storage beneath the Klyuchevskoy group of volcanoes inferred from time-dependent seismic tomography. *Journal of Volcanology and Geothermal Research*, *263*, 75–91. https://doi.org/10.1016/j.jvolgeores.2012.10.014

Koulakov, I. Y., Kukarina, E., Gordeev, E., Chebrov, V., & Vernikovsky, V. (2016). Magma sources in the mantle wedge beneath the volcanoes of the Klyuchevskoy group and Kizimen based on seismic tomography modeling. *Russian Geology and Geophysics*, *57*(1), 82–94. https://doi.org/10.1016/j.rgg.2016.01.006

Lees, J. M., Symons, N., Chubarova, O., Gorelchik, V., & Ozerov, A. (2007). Tomographic images of Klyuchevskoy volcano p-wave velocity. In J. Eichelberger et al. (Eds.), *Volcanism and Subduction: The Kamchatka region* (Vol. 172, pp. 293–302). Washington, DC: American Geophysical Union. https://doi.org/10.1029/172GM21

Levin, V., Droznina, S., Gavrilenko, M., Carr, M. J., & Senyukov, S. (2014). Seismically active subcrustal magma source of the Klyuchevskoy volcano in Kamchatka, Russia. *Geology*, *42*(11), 983–986. https://doi.org/10.1130/G35972.1

Levin, V., Shapiro, N., Park, J., & Ritzwoller, M. (2002). Seismic evidence for catastrophic slab loss beneath Kamchatka. *Nature*, *418*(6899), 763–767. https://doi.org/10.1038/nature00973

Levin, V., Shapiro, N. M., Park, J., & Ritzwoller, M. H. (2005). Slab portal beneath the western Aleutians. *Geology*, *33*(4), 253–256. https://doi.org/10.1130/G20863.1

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. https://doi.org/10.1785/0220160189

McNutt, S. R. (1992). Volcanic tremor. *Encyclopedia of Earth System Science*, *4*, 417–425.

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.

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*, *107*(B10), 2243. https://doi.org/10.1029/2001JB000559

Moni, A., Bean, C. J., Lokmer, I., & Rickard, S. (2012). Source separation on seismic data: Application in a geophysical setting. *IEEE Signal Processing Magazine*, *29*(3), 16–28. https://doi.org/10.1109/MSP.2012.2184229

Nikulin, A., Levin, V., Shuler, A., & West, M. (2010). Anomalous seismic structure beneath the Klyuchevskoy group, Kamchatka. *Geophysical Research Letters*, *37*, L14311. https://doi.org/10.1029/2010GL043904

Orozco-Alzate, M., Acosta-Muñoz, C., & Londoño-Bonilla J. M. (2012). *The automated identification of volcanic earthquakes: Concepts, applications and challenges*: INTECH Open Access.

Ozerov, A. Y., Firstov, P., & Gavrilov, V. (2007). Periodicities in the dynamics of eruptions of Klyuchevskoi volcano, Kamchatka. *Volcanism and Subduction: The Kamchatka Region*, *172*, 283–291. https://doi.org/10.1029/172GM20

Park, J., Levin, V., Brandon, M., Lees, J., Peyton, V., Gordeev, E., & Ozerov, A. (2002). A dangling slab, amplified arc volcanism, mantle flow and seismic anisotropy in the Kamchatka plate corner. In J. Eichelberger et al. (Eds.), *Plate Boundary Zones* (pp. 295–324). Washington, DC: AGU. https://doi.org/10.1029/GD030p0295

Roman, D. C., & Cashman, K. V. (2006). The origin of volcano-tectonic earthquake swarms. *Geology*, *34*(6), 457–460. https://doi.org/10.1130/G22269.1

Senyukov, S. (2013). Monitoring and prediction of volcanic activity in Kamchatka from seismological data: 2000–2010. *Journal of Volcanology and Seismology*, *7*(1), 86–97. https://doi.org/10.1134/S0742046313010077

Senyukov, S., Droznina, S. Y., Nuzhdina, I., Garbuzova, V., & Kozhevnikova, T. Y. (2009). Studies in the activity of Klyuchevskoi volcano by remote sensing techniques between January 1, 2001 and July 31, 2005. *Journal of Volcanology and Seismology*, *3*(3), 191–199. https://doi.org/10.1134/S0742046309030051

Seydoux, L., de Rosny, J., & Shapiro, N. M. (2017). Preprocessing ambient noise cross-correlations with equalizing the covariance matrix eigen-spectrum. *Geophysical Journal International*, *210*, 1432–1449. https://doi.org/10.1093/gji/ggx250

Seydoux, L., Shapiro, N., de Rosny, J., Brenguier, F., & Landès, M. (2016). Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays. *Geophysical Journal International*, *204*(3), 1430–1442. https://doi.org/10.1093/gji/ggv531

Seydoux, L., Shapiro, N., de Rosny, J., & Landès, M. (2016). Spatial coherence of the seismic wavefield continuously recorded by the USarray. *Geophysical Research Letters*, *43*, 9644–9652. https://doi.org/10.1002/2016GL070320

Shapiro, N., Droznin, D. V., Droznina, S. Y., Senyukov, S. L., Gusev, A. A., & Gordeev, E. I. (2017). Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer. *Nature Geosciences*, *10*, 442–445. https://doi.org/10.1038/NGEO2952

Shapiro, N. M., Sens-Schönfelder, C., Lühr, B. G., Weber, M., Abkadyrov, I., Gordeev, E. I., … Saltykov, V. A. (2017). Understanding Kamchatka's extraordinary volcano cluster. *Eos*, 98. https://doi.org/10.1029/2017EO071351

Slavina, L., Garagi, I., Gorelchik, V., Ivanov, B., & Belyankin, B. (2001). Velocity structure and stress-deformation state of the crust in the area of the Kluchevskoy volcano group in Kamchatka. *Journal of Volcanology and Seismology*, *1*, 49–59.

Slavina, L., Pivovarova, N., & Senyukov, S. (2012). The dynamics of the velocity field beneath the north group of volcanoes, Kamchatka, and its relationship to the evolution of volcanic activity. *Journal of Volcanology and Seismology*, *6*(2), 100–115. https://doi.org/10.1134/S0742046312020066

Sparks, R., Biggs, J., & Neuberg, J. (2012). Monitoring volcanoes. *Science*, *335*(6074), 1310–1311. https://doi.org/10.1126/science.1219485

Unglert, K., & Jellinek, A. (2017). Feasibility study of spectral pattern recognition reveals distinct classes of volcanic tremor. *Journal of Volcanology and Geothermal Research*, *336*, 219–244. https://doi.org/10.1016/j.jvolgeores.2017.03.006

Vidale, J. E. (1986). Complex polarization analysis of particle motion. *Bulletin of the Seismological society of America*, *76*(5), 1393–1405.

Yogodzinski, G., Lees, J. M., Churikova, T., Dorendorf, F., Wöerner, G., & Volynets, O. (2001). Geochemical evidence for the melting of subducting oceanic lithosphere at plate edges. *Nature*, *409*(6819), 500–504.