

Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays

L. Seydoux, N. M. Shapiro, J. de Rosny, F. Brenguier, M. Landès

▶ **To cite this version:**

L. Seydoux, N. M. Shapiro, J. de Rosny, F. Brenguier, M. Landès. Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays. Geophysical Journal International, 2016, 204, pp.1430-1442. 10.1093/gji/ggv531 . insu-03581673

HAL Id: insu-03581673

https://insu.hal.science/insu-03581673

Submitted on 20 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Geophysical Journal International

Geophys. J. Int. (2016) **204,** 1430–1442 doi: 10.1093/gji/ggv531
GJI Seismology

Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays

L. Seydoux,[1] N.M. Shapiro,[1,2] J. de Rosny,[3] F. Brenguier[4] and M. Landès[1]

[1]*Institut de Physique du Globe de Paris, UMR CNRS* 7154, *F*-75005 *Paris, France. E-mail: seydoux@ipgp.fr*
[2]*Institute of Volcanology and Seismology FEB RAS,* 9 *Piip Boulevard, Petropavlovsk-Kamchatsky* 683006, *Russia*
[3]*ESPCI ParisTech, CNRS, Institut Langevin, PSL Research University,* 1 *rue Jussieu, F*-75005 *Paris, France*
[4]*Institut des Sciences de la Terre, UMR CNRS* 5275, *F*-38041 *Grenoble, France*

Accepted 2015 December 10. Received 2015 December 9; in original form 2015 April 20

SUMMARY

Modern seismic networks are recording the ground motion continuously at the Earth's surface, providing dense spatial samples of the seismic wavefield. The aim of our study is to analyse these records with statistical array-based approaches to identify coherent time-series as a function of time and frequency. Using ideas mainly brought from the random matrix theory, we analyse the spatial coherence of the seismic wavefield from the width of the covariance matrix eigenvalue distribution. We propose a robust detection method that could be used for the analysis of weak and emergent signals embedded in background noise, such as the volcanic or tectonic tremors and local microseismicity, without any prior knowledge about the studied wavefields. We apply our algorithm to the records of the seismic monitoring network of the Piton de la Fournaise volcano located at La Réunion Island and composed of 21 receivers with an aperture of ∼15 km. This array recorded many teleseismic earthquakes as well as seismovolcanic events during the year 2010. We show that the analysis of the wavefield at frequencies smaller than ∼0.1 Hz results in detection of the majority of teleseismic events from the Global Centroid Moment Tensor database. The seismic activity related to the Piton de la Fournaise volcano is well detected at frequencies above 1 Hz.

Key words: Time-series analysis; Volcano seismology; Statistical seismology; Volcano monitoring.

1 INTRODUCTION

The fast development of large seismic networks deployed worldwide and providing continuous records of the ground motion at the Earth's surface every day has challenged seismologists to develop automatic processes to detect seismic events, such as tectonic earthquakes, volcanic tremors or surface explosions, in continuous seismic waveforms. Traditional passive seismological methods are based on signals generated by earthquakes or explosions, namely, on sources localized in space and time. Many algorithms for detection and characterization of such signals exist and are successfully applied for automatic data processing and analysis. They usually assume that the studied sources generate wave fronts that can be characterized by reasonably well-defined changes of signal properties at individual receivers. Some well-known methods for exploring these rapid signal changes compare the Short-Term Average with the Long-Term Average introduced by Allen (1982) or some higher order statistical functions, such as the kurtosis and the skewness functions (e.g. Saragiotis *et al.* 2002). The accuracy of the event detection can also be improved by combining multiple approaches (e.g. Baillard *et al.* 2014).

However, many types of signals produced by some natural seismic sources or by anthropogenic activity have an emergent character and do not present any clear onset. Among examples of such emergent signals we may mention the volcanic and the tectonic tremors and the 'microseismic noise' generated by storms and waves in the oceans. Therefore, these signals cannot be efficiently identified with the aforementioned methods based on detection of the sharp onset or pulse-like feature in the waveform. The spatial coherence of emergent signals remains however higher than the 'incoherent' seismic noise, and can be thus quantified when records at several spatial locations are available. In other words, detection of these types of signals can be done with array-based methods.

The development of array-based methods in seismology was mainly motivated by the installation of first two large-scale seismic arrays LASA (Frosch & Green 1966) and NORSAR (Bungum *et al.* 1971) in mid-1960s in the context of the monitoring of nuclear explosions. Arrays of sensors provide dense spatial samples of the wavefield in a small-space region and their 2-D configuration can be used to access the apparent traveltimes of the seismic waves between the receivers (apparent because we only access the projection of the seismic wavefield onto the array plane), hence the apparent

© The Authors 2016. Published by Oxford University Press on behalf of The Royal Astronomical Society.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

slowness. The well-known method generally used as first approach to estimate the apparent slowness and backazimuth of the recorded waves is the plane-wave beamforming, that assumes plane wave fronts. A good review of this class of methods can be found in Rost & Thomas (2002). This technique stands well for far-fields seismic sources, but fails to correctly estimate slowness and backazimuth of sources lying within the array. Similarly, the strongly scattered fields that cannot be simply characterized by the beamforming.

The method here proposed focuses on the array covariance matrix, that contains information on the spatial coherence of the recorded wavefield. The main idea comes from Bartlett (1954) and Lawley (1956) who proposed that the number of independent signals can be deduced from the rank of the covariance matrix with the largest eigenvalues corresponding to these signals and the smaller ones to the background noise. Estimating the effective rank of the covariance matrix from its eigenvalues spectrum in a presence of noise is not a simple problem, and several approaches were proposed to solve it. Wax & Kailath (1985) proposed to use a criteria from the information theory by comparing the observed eigenvalues with the distributions obtained from a signal model that consists in a sum of an unknown number of independent signals embed in a background noise (usually considered a white noise). The number of independent signals was then frequently used to define the coherent and the incoherent subspaces (the spaces formed by the signal- and the noise-related eigenvectors, respectively). The coherent subspace can then be injected into the beamforming analysis to obtain better slowness and backazimuth estimates (Goldstein & Archuleta 1987; Chiou & Bolt 1993; Wagner & Owens 1996).

The random matrix theory, which also finds application in telecommunication (Müller 2002; Tulino & Verdú 2004) and in acoustics (Aubry & Derode 2009a), can help to understand the eigenvalue distribution of covariance matrices. Marchenko & Pastur (1967) provided a thorough description of the probability of the eigenvalue distribution of some sets of particular random matrices. Sengupta & Mitra (1999) and Mestre (2008) applied this description to the particular case of covariance matrices. In underwater acoustics, Gerstoft *et al.* (2012) and Menon *et al.* (2012a) used these statements to describe the probability of distribution of the ocean noise covariance matrices, recorded by linear arrays. These results then have been used in Menon *et al.* (2012b) to filter directional sources in order to extract stable Green's function from cross-correlation (Roux *et al.* 2004) of ambient noise between the receivers.

The purpose of this study is to develop a robust detector of signals based on their coherence across an array of sensors. The exact separation between the coherent and the incoherent parts of the wavefield is not necessary for this goal. Instead, we use a simple and robust characteristic of the covariance matrix eigenvalue distribution. An example of such approach is the one by Wagner & Owens (1996) who used the magnitude of the largest eigenvalue as a signal detector. We follow this idea, but instead of focusing on the largest eigenvalue, we calculate the width of the eigenvalue distribution of the covariance matrix in a sliding time window, at different frequencies. We apply a time-running normalization and a spectral whitening to the data as described in Bensen *et al.* (2007) in order to discard the signal amplitude information, and to focus on the spatial coherence. The main advantages of our method are the simplicity of its implementation and the fact that no *a priori* knowledge on the recorded signals is needed.

In this paper, we mainly focus on describing the proposed signal detection method. We start with explaining basic elements of the covariance matrix theory. Then, we discuss the practical implementation of the method to the data recorded by an array of seismic receivers. We perform a set of tests with synthetic data to optimize the choice of the method parameters. Finally, we illustrate our method with an application to the real data from the UnderVolc broad-band seismic array (Brenguier *et al.* 2012) deployed around the Piton de la Fournaise volcano located on the eastern side of La Réunion Island, in the Indian Ocean (Brenguier *et al.* 2012). We show that at relatively long periods (above 10 s) the covariance matrix analysis of the UnderVolc network records is mainly sensitive to teleseismic earthquakes. At high frequencies (0.5–5 Hz), the method is sensitive to local seismic sources and in particular to those related to volcanic activity, such as the pre-eruptive swarms of seismovolcanic events and the coeruptive volcanic tremors.

2 COVARIANCE MATRIX: THEORY

We define the array data vector at frequency f as

$$\mathbf{u}(f) = \begin{bmatrix} u_1(f) \\ u_2(f) \\ \vdots \\ u_N(f) \end{bmatrix}, \tag{1}$$

where $u_i(f)$ is the complex Fourier spectra of the record by sensor i and N is the number of sensors. The covariance matrix is defined as

$$\Gamma(f) = \mathrm{E}\left[\mathbf{u}(f)\mathbf{u}^{\dagger}(f)\right], \tag{2}$$

where E represents the expected value and † denotes Hermitian transpose so that the outer product $\mathbf{u}(f)\mathbf{u}^{\dagger}(f)$ is an $N \times N$ matrix. Following Gerstoft *et al.* (2012), we consider the data being composed of three main parts:

$$\mathbf{u}(f) = \sum_{k=1}^{K} \alpha_k \mathbf{s}_k(f) + \mathbf{n}^p(f) + \mathbf{n}^n(f), \tag{3}$$

where $\mathbf{s}_k(f)$ is the signal produced by the kth coherent source with complex amplitude α_k, K is the number of independent sources, $\mathbf{n}^p(f)$ is the propagating seismic noise and $\mathbf{n}^n(f)$ is the non-propagating noise (sensor-self noise, weather perturbation, etc.). The covariance matrix obtained by injecting eq. (3) in eq. (2) is

$$\mathbf{\Gamma}(f) = \mathbf{\Gamma}^s(f) + \mathbf{\Gamma}^p(f) + \mathbf{\Gamma}^n(f), \tag{4}$$

where $\mathbf{\Gamma}^s(f)$ is the covariance matrix of the coherent sources, $\mathbf{\Gamma}^p(f)$ is the covariance matrix of the incoherent propagating noise and $\mathbf{\Gamma}^n(f)$ is the covariance matrix of the non-propagating noise. The three terms in eq. (3) are not correlated and, therefore, no cross-terms appear in eq. (4) because their expected values vanish.

We assume the sources to be independent, that mathematically leads to $\mathrm{E}[\alpha_k \alpha_{k'}] = \mathrm{E}\left[|\alpha_k|^2\right]\delta_{kk'}$. To simplify, we consider $\mathrm{E}[|\alpha_k|^2] = 1$, so the first term in eq. (4), that is, the covariance matrix of K independent sources with is given by

$$\mathbf{\Gamma}^s(f) = \sum_{k=1}^{K} \mathbf{s}_k(f)\mathbf{s}_k^{\dagger}(f). \tag{5}$$

One can easily demonstrate that this matrix is of rank K and thus contains K non-zero eigenvalues, and $N - K$ zero eigenvalues.

We consider the non-propagating noise to be uncorrelated between the sensors. The third term in (4) can thus be written as

$$\mathbf{\Gamma}^n(f) = \nu^2(f)\mathbf{I}, \tag{6}$$

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

where $\nu^2(f)$ is the spectral density of the non-propagating noise and $\mathbf{I}$ is the $N \times N$ identity matrix. Therefore, this term is of rank N, that is, the eigenvalue distribution is flat, with all eigenvalues equal to $\nu^2(f)$.

We know that the ambient seismic noise is dominated by surface waves (Friedrich *et al.* 1998). Therefore, we assume the propagating noise to be 2-D, and for simplicity, to be isotropic. If, in addition, we consider the media to be homogeneous, we can use at any given frequency f the analytic solution of Cox (1973) for the propagating noise covariance matrix :

$$\Gamma^p_{ij}(f) = \mathrm{J}_0(2\pi f \gamma \|\mathbf{r}_j - \mathbf{r}_i\|)\,, \tag{7}$$

where γ is the slowness of the waves, $\mathbf{r}_i$ is the planar coordinates of the sensor i at the Earth's surface, such as $\|\mathbf{r}_j - \mathbf{r}_i\|$ is the distance between the sensors i and j and J_0 represent the Bessel function of the first kind. An asymptotical form of the eigenvalue distribution of the 2-D isotropic noise covariance matrix is given for the case of a linear array by Gerstoft *et al.* (2012) who shown that the eigenvalue distribution of this matrix depends on the typical sensor-spacing d and the wavelength $\lambda_\varphi = c_\varphi/f$ (where c_φ is the phase velocity). Indeed, when $2d/\lambda_\varphi > 1$, the covariance matrix is full rank, and with the increasing value of this ratio the eigenvalue distribution becomes flatter. Inversely, $2d/\lambda_\varphi < 1$ implies a rank deficiency, so that the number of non-zero eigenvalues decreases when this ratio tends to zero. A simple explanation of this deficiency is that at low frequencies the array aperture is too small compared to the wavelength to measure the wavefield oscillations, and the seismic noise is seen as coherent.

3 METHOD: COVARIANCE MATRIX ESTIMATED FROM THE DATA

The different steps of the computation and the analysis of the array covariance matrix are presented in Fig. 1. First, the signals are divided into overlapping time windows (averaging windows) in which energy normalization is applied (temporal normalization and spectral whitening). Then, we compute the Fourier transform of the pre-processed signal within smaller subwindows and calculate the array cross-spectra matrices in each of the subwindows. The array covariance matrix is obtained from the average of the cross-spectra matrices. The covariance matrix is thus computed over a set of M subwindows within each averaging window. We finally compute the spectral width of all the obtained covariance matrices, in different averaging windows as a function of the frequency.

3.1 Array covariance matrix estimation

We estimate the covariance matrix $\mathbf{C}(f)$ from the time average of the Fourier cross-spectra matrices computed over a set of M overlapping subwindows (see Fig. 1) of length δt:

$$\mathbf{C}(f) = \left\langle \mathbf{u}(f)\mathbf{u}^\dagger(f) \right\rangle_{\Delta t} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{u}_m(f)\mathbf{u}_m^\dagger(f)\,, \tag{8}$$

where $\mathbf{u}_m(f)$ is the vector of data Fourier spectra in the subwindow m. The total duration of the window in which the array covariance matrix is estimated is thus $\Delta t \approx Mr\delta t$, where r is the overlapping ratio. Note that the array covariance matrix is inherently Hermitian, and therefore is, diagonalizable.

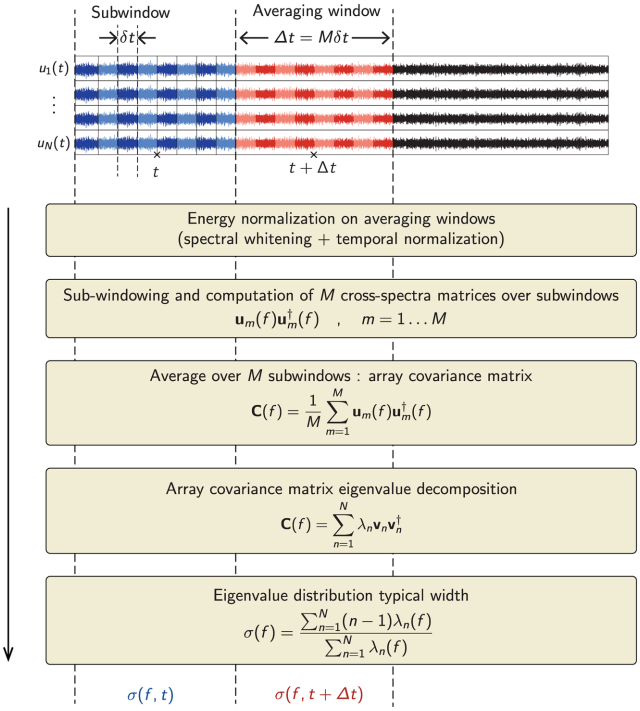


Figure 1. Sequence of operations used for estimating the array covariance matrix from continuous seismic records, and computing the spectral width. See the text for details.

3.2 Array covariance matrix eigenvalue distribution

We obtain the eigenvalues of the covariance matrix with using algorithms from the `Eigen 3.2.6` C++ library. Because we estimate the covariance matrix from a finite set of subwindows, its eigenvalue distribution differs from the theoretical one described in Section 2. Marchenko & Pastur (1967) proposed an description of the eigenvalue probability of distribution of estimated covariance matrices. These results were applied to underwater acoustics in Gerstoft *et al.* (2012) in order to qualify the effect of the estimation on the covariance matrix of ocean propagating and non-propagating noises recorded by a linear array of hydrophones. A main result is that the non-propagating noise eigenvalue spectrum estimated from a finite number of subwindows is not anymore constant but decays steadily. It slowly converges to a flat distribution with an increasing number of subwindows.

3.3 Array covariance matrix spectral width

The eigenvalues λ_i of the Hermitian matrix $\mathbf{C}(f)$ are real and positive numbers. We arrange them in decreasing order. The covariance matrix spectral width is then computed as a function of frequency as:

$$\sigma(f) = \frac{\sum_{i=1}^{N}(i-1)\lambda_i(f)}{\sum_{i=1}^{N}\lambda_i(f)}\,. \tag{9}$$

The values of σ estimated within different time windows become nearly zero when the analysed records are dominated by a single strong source and they approach some maximal value $\sigma_{\max}(f)$ when a window containing only non-propagating noise is analysed. The value of $\sigma_{\max}(f)$ depends on the relative amplitudes of the non-propagating and the propagating noises and on the wavelength-to-interstation-distance ratio $2d/\lambda_\varphi$.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

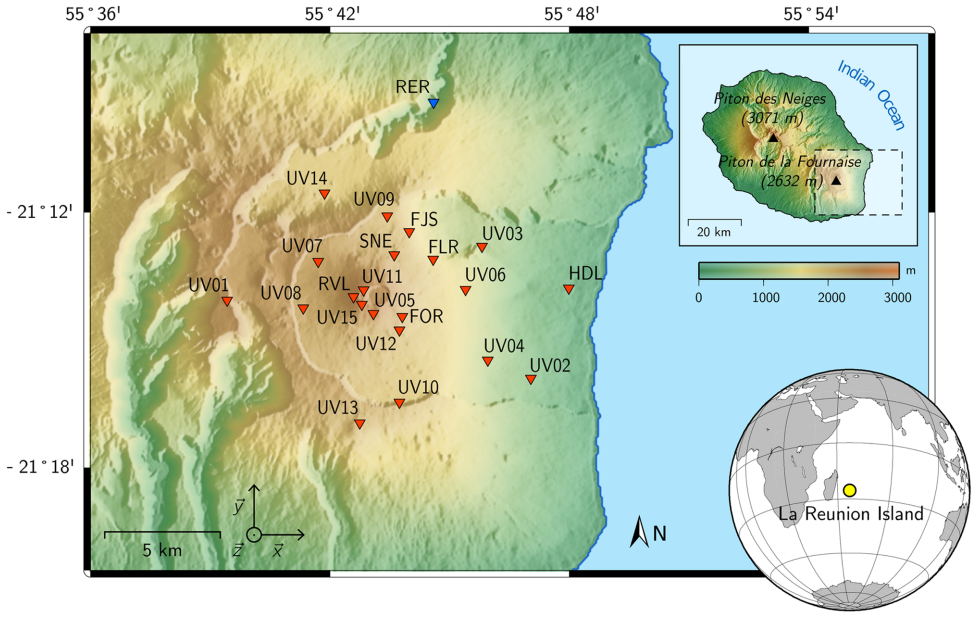


Figure 2. Map of the Piton de la Fournaise volcano. The top right inset shows the location of the Piton de la Fournaise volcano on La Réunion Island and the bottom right inset shows the location of La Réunion island in the Indian Ocean. The locations of UnderVolc stations are shown with red triangles and the location of the RER seismic station of the Geoscope seismic network is shown in blue triangle.

4 PITON DE LA FOURNAISE SEISMIC NETWORK

We applied our method to the data from the UnderVolc experiment (Brenguier *et al.* 2012) when an array of 21 broad-band seismographs was installed around the Piton de la Fournaise volcano on La Réunion island as shown in Fig. 2. The network consists of 15 Gurlap 30-s CMG-40T broad-band three-component seismometers and of 6 additional Kinemetrics Q330 associated with Gurlap 3ESPC three-components sensors. This array recorded many teleseismic earthquakes as well as seismovolcanic events associated with the volcanic activity. This study focuses on the vertical-component records. An extension of our method to a full analysis of three-component seismograms in under way. We analyse seismic data recorded during 2010 (from January 1 to December 31) when all 21 stations were recording simultaneously.

5 SYNTHETIC TESTS

We performed synthetic tests to characterize the covariance matrix spectral width as a function of the wavefield coherence, for propagating signals only. These tests were applied to the geometry of the UnderVolc network system (Fig. 2) and we ignored the topography.

5.1 Building synthetic data

We built the data as described in Fig. 3: we considered a number of K vertically polarized plane surface waves propagating isotropically in the network plane, with the same slowness and frequency (respectively, γ and f). The expression of the kth plane wave at coordinate $\mathbf{r} = (x, y)$ is given by

$$\Phi_k(\mathbf{r}, f) = \exp(-2i\pi f \gamma \mathbf{e}_k \cdot \mathbf{r} - i\varphi_k), \quad k = 1 \dots K, \tag{10}$$

where φ_k is the phase, $\mathbf{e}_k = \cos(\theta_k)\mathbf{e}_x + \sin(\theta_k)\mathbf{e}_y$ is the direction vector of the wave with θ_k homogeneously distributed between 0 and 2π. The synthetic data vector $\tilde{\mathbf{u}}$ is then obtained from the sum of all recorded plane waves at $\mathbf{r}_i$, the coordinates of the sensor i:

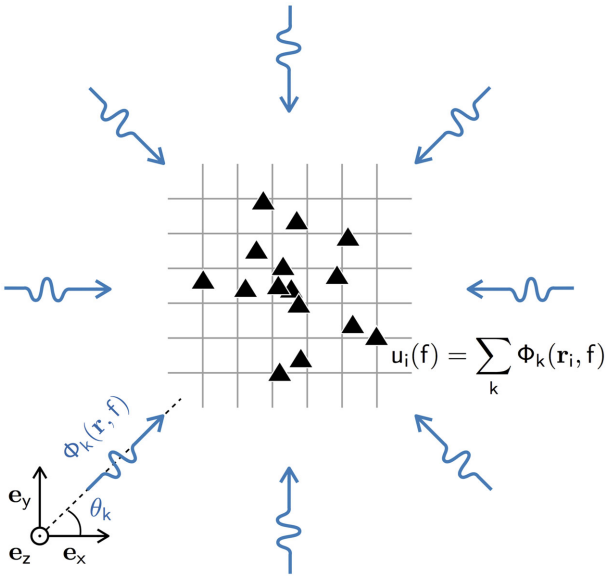


Figure 3. Construction of the wavefield used in synthetic tests: a sum of plane waves recorded by a seismic array. The black triangles show the location of the seismic stations and the blue arrows represent the incoming plane waves.

$$\tilde{u}_i(f) = \sum_{k=1}^{K} \Phi_k(\mathbf{r}_i, f). \tag{11}$$

The method requires averaging with time subwindows over a given duration Δt. Therefore, we generated M different synthetic records $\tilde{\mathbf{u}}_m(f), m = 1 \dots M$ corresponding to different time subwindows in the real data. The seismic noise can be seen as an incoherent wavefield and we simulate it as a sum of plane waves with randomly varying phase from a given subwindow to another. A coherent

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

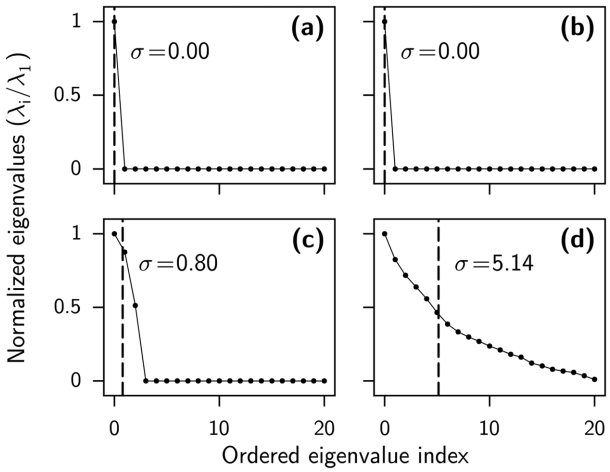


Figure 4. Effect of the coherence or incoherence nature of the wavefield on the eigenvalue distribution of covariance matrices computed from synthetic wavefields, for the four cases considered in Section 5. Values of the covariance matrix spectral width σ are indicated with vertical dashed lines.

wavefield is also simulated as a sum of plane waves. In this case, however, the phase of each wave does not change in the subwindows. We then built the synthetic covariance matrix using eq. (8) from M different sets of synthetic signals in order to perform the ensemble average (combining eqs 10 and 11)

$$\tilde{\mathbf{C}}(f) = \sum_{m=1}^{M} \tilde{\mathbf{u}}_m(f)\tilde{\mathbf{u}}_m^{\dagger}(f)\,. \tag{12}$$

5.2 Example of the covariance matrix spectra

We set the frequency to 0.2 Hz and the slowness to 1/2000 s m^{-1}. These parameters are approximatively those of the Rayleigh wave propagation within La Réunion crust (Brenguier *et al.* 2008). We investigated the effect of the number of sources and the coherence of these sources within four cases. We used $M = 100$ generations of the wavefield to perform the ensemble average, and we compared the results obtained when $K = 3$ and 100 plane waves, for the coherent and the incoherent cases (constant and random phases, respectively). The eigenvalue distributions of the covariance matrices obtained in all four cases are presented in Fig. 4. Not surprisingly, when the plane waves are coherent, the number of non-zero eigenvalue remains equal to 1 regardless of the number of sources (Figs 4a and b). However, there are 3 non-zero eigenvalues when the wavefield contains 3 incoherent plane waves (Fig. 4c), and the covariance matrix spectrum is wider for a wavefield composed of 100 incoherent plane waves (Fig. 4d).

These simulations illustrate the main idea used in our analysis, namely that the number of non-zero eigenvalues is related to the number of independent signals. Fig. 4(b) illustrates that waves generated by the same source and scattered by the media will produce a covariance matrix of rank 1, because they remain coherent. This is in agreement with the theoretical results provided in, for example, Aubry & Derode (2009b).

5.3 Selecting optimal parameters of the data analysis

Results of application of the proposed method to continuous seismic records depend on the choice of the parameters used in the

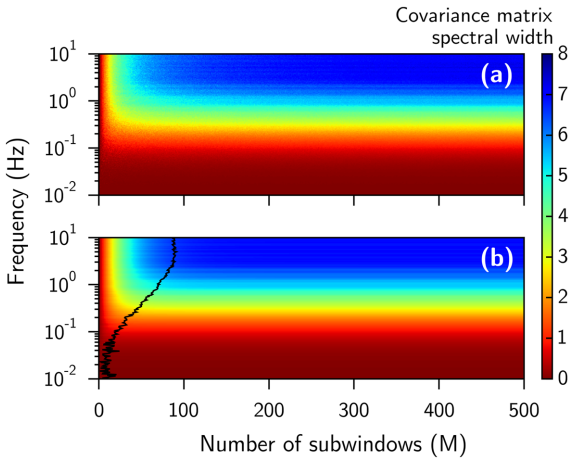


Figure 5. Convergence of the covariance matrix spectral width $\sigma(f)$ as function of the number of subwindows M (eq. 8). The synthetic wavefield composed of 100 incoherent plane waves was used for these tests. (a) Direct results of synthetic tests. (b) Fit obtained with eq. (13). The value $3M_0(f)$ is shown with the black solid line, and shows the limit where $\sigma(f)$ reaches 95 per cent of its maximum value.

covariance matrix computation, namely of the number of subwindows M and of their length δt. This leads to a trade-off between a statistically robust estimation of the covariance matrix, and a good time resolution.

5.3.1 *Length of the subwindows δt*

The length of subwindows δt is directly linked to the lowest frequency that can be resolved. In our case, the longest period that we aim to study is 20 s (0.05 Hz). Therefore, we can use 48s-long subwindows in order to have at least two oscillations of the signal in every subwindow.

5.3.2 *Number of subwindows M*

When $M = 1$, the estimated covariance matrix is degenerated and its rank equals 1. With increasing M the estimation of the covariance matrix becomes more robust and its eigenvalue spectrum becomes wider. The covariance matrix spectral width $\sigma(f)$ converges to a maximum value $\sigma_{\max}(f)$. We used our synthetic signals to test the speed of this convergence at different frequencies between 0.01 and 10 Hz. We computed the covariance matrix spectral width as function of frequency and of subwindows number $\sigma(f, M)$ as shown in Fig. 5(a). This function can be reasonably fitted as proposed in Fig. 5(b) with the following expression:

$$\tilde{\sigma}(f, M) = \sigma_{\max}(f)(1 - e^{-M/M_0(f)})\,. \tag{13}$$

Eq. (13) can be used to evaluate the maximal value reachable by the covariance matrix spectral width and, more importantly, to estimate the number of subwindows required to approach this maximal value with a certain accuracy. Eq. (13) indicates that 95 per cent of the maximal value of the covariance matrix spectral width is reached after stacking $3M_0(f)$ subwindows. We display $3M_0(f)$ with a black solid line on Fig. 5(b). We can see that with considered network configuration a minimal number of 100 subwindows are required to reach the 95 per cent threshold at all considered frequencies.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

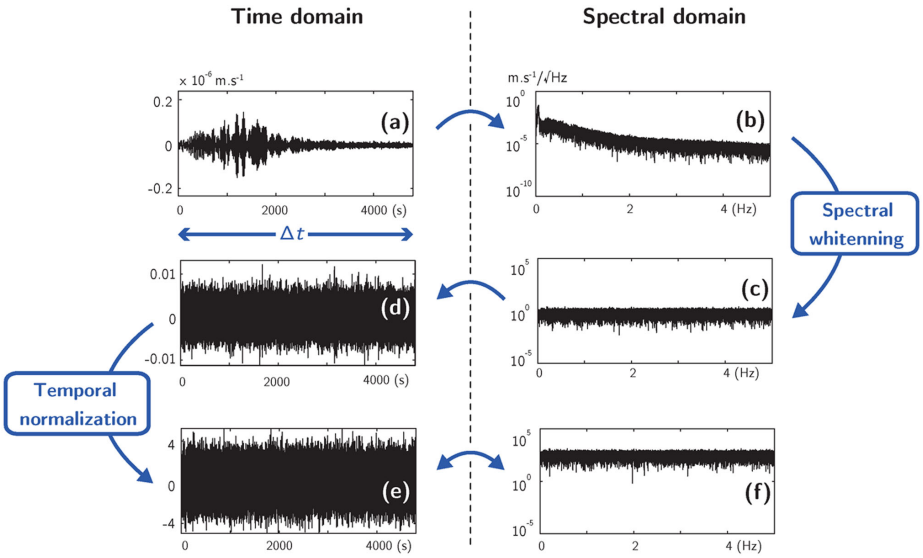


Figure 6. Illustration of the pre-processing of seismic records. (a) Raw data. (b) Raw data amplitude Fourier spectra. (c) Whitened spectral amplitude. (d) Real part of the inverse Fourier transform of the of whitened spectrum. (e) Temporally normalized time-series. (f) Fourier amplitude spectra of (e).

6 ANALYSIS OF THE UNDERVOLC DATA

6.1 Data preparation

First, all seismograms are bandpassed between 0.01 and 10 Hz and downsampled from 100 to 20 Hz to accelerate the computations. Then, we apply the spectral and temporal normalizations to compensate the non-stationarity of seismic signals and to disregard any amplitude information. Bensen *et al.* (2007) present several ways to normalize the data in this way. We apply the spectral whitening and the temporal normalization as described in Fig. 6, according to Bensen *et al.* (2007).

Spectral whitening consists in dividing the signal spectrum by a smooth version of its amplitude

$$u^W(f) = \frac{u(f)}{\langle\langle |u(f)| \rangle\rangle_{df}}, \tag{14}$$

where $|u(f)|$ represents the real absolute value of the spectrum and $\langle\langle\cdot\rangle\rangle_{df}$ stands for the df-long running average used to smooth the spectral amplitude. In our case, we used $df = 0.33$ Hz. The operator in eq. (14) preserves the phase, and the amplitude of the whitened spectrum is close to 1 for all frequencies.

We then apply a temporal normalization in the time domain:

$$u^N(t) = \frac{u^W(t)}{\langle\langle |u^W(t)| \rangle\rangle_{dt}} \tag{15}$$

where $u^W(t)$ is the real part of the inverse Fourier transform of $u^W(f)$ and $\langle\langle\cdot\rangle\rangle_{dt}$ is the dt-long running average used to smooth the temporal normalization. In our case, we use a 1.25 s smoothing window.

6.2 Difference between the background noise and an earthquake

We applied our technique to the records of the $M \approx 7.5$ Nicobar Island earthquake which occurred around 7:00 p.m. UTC on June, 12. We analysed the signals over two time periods shown in Fig. 7(a): a *noise* window where no significant event occurred and an *event* window where the Nicobar Island earthquake is recorded by the

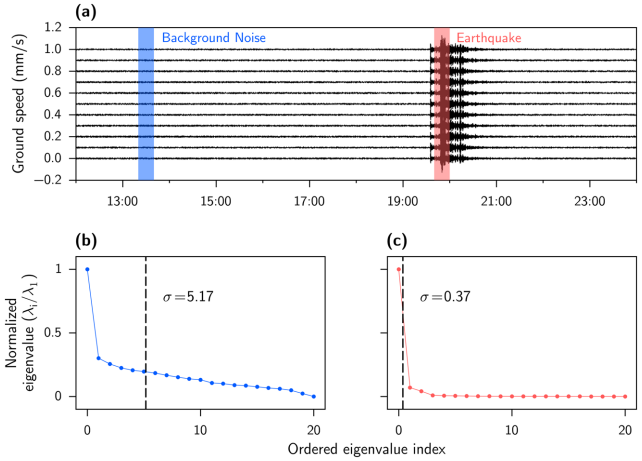


Figure 7. Example of covariance matrix spectra computed from real records (shown in (a) after bandpassing filtered around 0.07 Hz). Covariance matrices were computed at $f = 0.07$ Hz. (b) Covariance matrix spectrum for noise. (c) Covariance matrix spectrum for signal. Values of the covariance matrix spectral width σ are indicated with vertical dashed lines.

UnderVolc array. We use a hundred of 48s-long 50 per cent overlapping subwindows to compute the two covariance matrices. Therefore, the full window durations are $\Delta t \approx 48 \times 100/2 = 2400$ s. We compute the covariance matrix eigenvalues at a frequency $f = 0.07$ Hz.

The covariance matrix spectrum computed from the noise window shown in Fig. 7(b) contains no null eigenvalues and its width is large, indicating the low level of the signal coherence. On the contrary, the signal generated by the Nicobar Island earthquake is strongly coherent across the network and the corresponding covariance matrix spectrum shown in Fig. 7(c) has a single eigenvalue that is significantly greater than zero. In this case, the spectral width σ is very small, as expected.

6.3 Analysis of continuous records in 2010

We apply the method over the whole year 2010. Following the results of the synthetic tests, we compute covariance matrix on

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

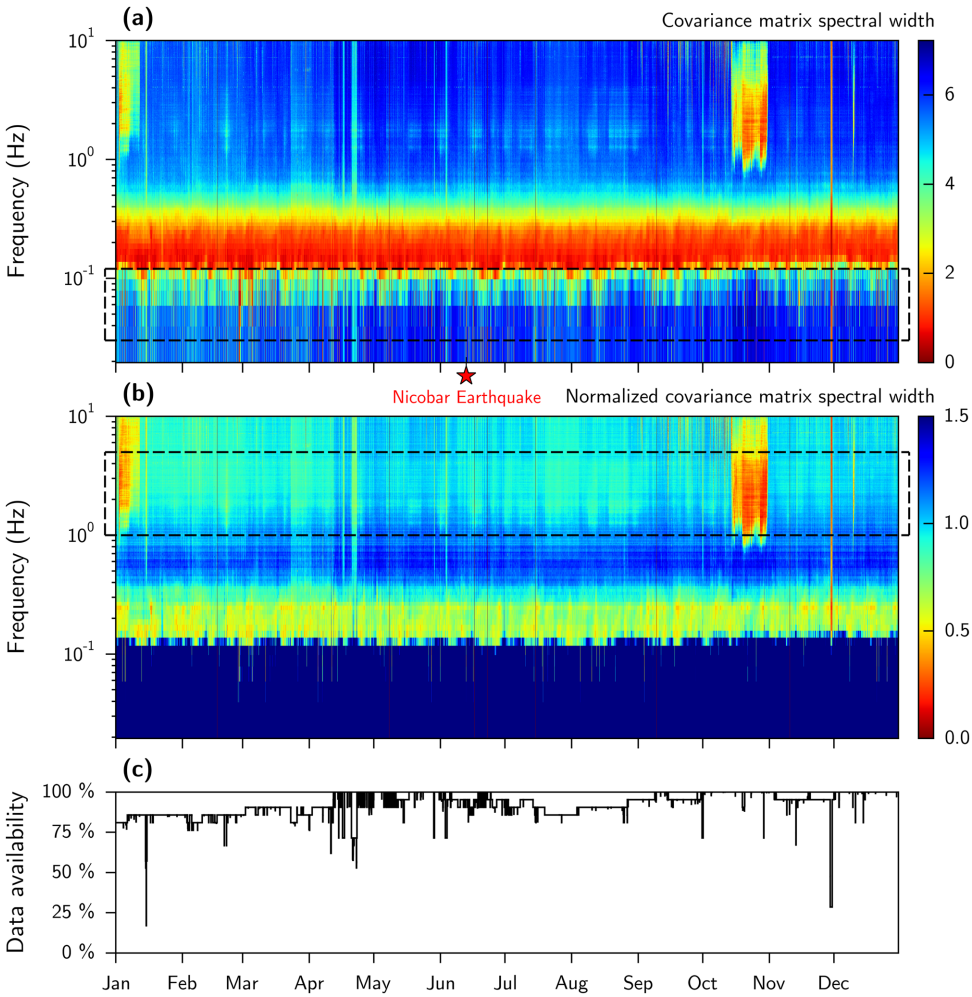


Figure 8. Results of analysis of continuous records of the Undervolc seismic network in 2010. (a) Spectral width of covariance matrix σ as function of time and frequency. (b) Normalized covariance matrix spectral width obtained after dividing the values shown in (a) by the maximal expected width estimated in Section 5. (c) Amount of working station as a function of time.

2400s-long windows. Each window contains 100 of half overlapping 48s-long subwindows. The spectral width of the covariance matrix as a function of time and frequency is displayed in Fig. 8(a).

As discussed in Section 5, the covariance matrix spectral width does not converge to the same value at all frequencies. Therefore, we normalize the raw result shown in Fig. 8(a) by using the maximal value $\sigma_{\max}(f)$ obtained in Section 5. The normalized covariance matrix spectral width shown in Fig. 8(b) exhibits strong variability in time and frequency. While most of this variability is related to variations of the level of coherence of the recorded wavefield, some of the anomalies can be caused by misfunctioning stations. For this reason, we show in Fig. 8(c) the percentage of working stations as a function of time and can see that some of the apparent increases of coherence, in particularly those in April and November, are caused by missing data.

At frequencies below 0.15 Hz, we observe that the normalized spectral width of the covariance matrix is higher than 1, highlighting that the coherence level is much lower than the synthetic results. This indicates that the noise recorded at these relatively long periods does not correspond to propagating long-period seismic waves that would be expected to be coherent across the network. Indeed, because of the non-perfect condition of temporary installation of seismic sensors on the Piton de la Fournaise volcano during the UnderVolc experiment, the long period noise is dominated by local perturbations likely of meteorological origin.

To estimate the amplitude of this non-propagating noise, we compare in Fig. 9(b) the spectral density of the records provided by the UnderVolc sensors over one day with the RER seismic station from the Geoscope seismic network located at la Réunion Island (see Fig. 2). We see that below 0.1 Hz, the spectral amplitude of the noise recorded by the UnderVolc stations is high compared to the one recorded by RER. The two microseismic peaks are clearly visible in the RER data (at 0.06 and 0.14 Hz) whereas only the secondary microseismic peak is recorded by the UnderVolc stations. We also compare the spectra of the Nicobar Island earthquake from record of UnderVolc and RER stations in Fig. 9(c), and observe that it does not differ from UnderVolc stations to RER, at frequencies higher than 0.02 Hz. We thus conclude that earthquake detection is possible. This additional information is clearly in agreement with the hypothesis of an non-propagating noise recorded by the UnderVolc stations, and with the results obtained in Fig. 8.

Around 0.2 Hz, many coherent arrivals are also visible all year long around 0.2 Hz corresponding to the secondary microseismic peak. This is in agreement with the standard model considering that this peak consists of waves emitted by strong sources originating from the interaction between the oceanic gravity waves and

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

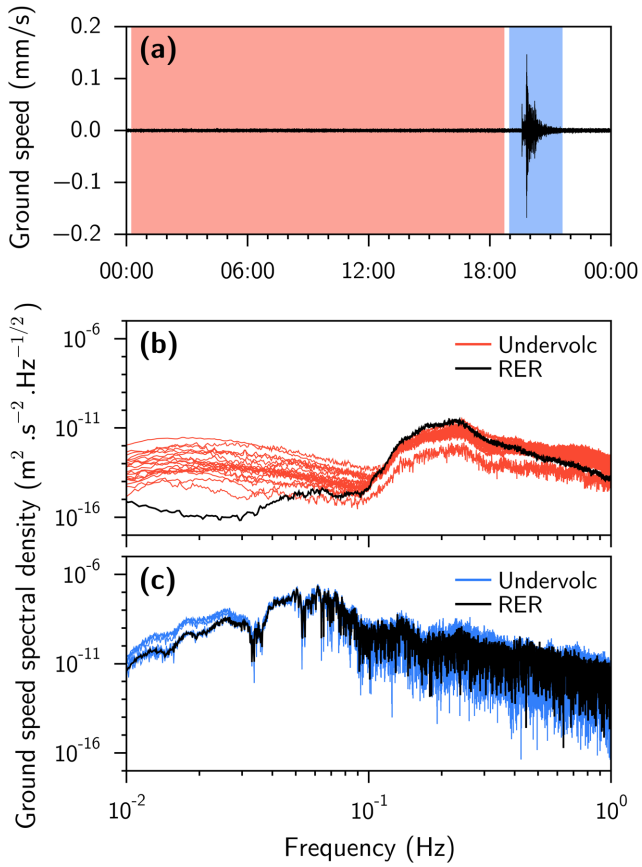


Figure 9. Comparison of amplitude spectra from records of the Undervolc stations and from the broad-band GEOSCOPE station RER. We analyse the records of July 12, when the Nicobar Island earthquake occurred around 7:00 p.m. The power spectral densities of the noise records (collected from 0:00 to 6:20 p.m.) were computed using Welch's method (Welch 1967). The event (signal recorded from 6:20 to 8:20 p.m.) amplitude spectra were computed using the Fast Fourier Transform.

the seafloor (Longuet-Higgins 1950; Friedrich *et al.* 1998). Waves from every individual microseismic sources are coherent across the network.

Finally, in the frequency range between 0.5 and 10 Hz, the normalized spectral width of the covariance matrix is most of the time close to 1, indicating that incoherent signals are recorded by the network. Strongly coherent signals are appearing a few times in the year in this frequency range and are linked to seismovolcanic activity of the Piton de la Fournaise volcano as described later in this paper.

6.4 Example of coherent signals from teleseismic earthquakes

Strong peaks of coherence are clearly visible on Fig. 8(a), at frequencies below 0.15 Hz. Most of these peaks coincide in time with expected arrival of surface waves generated by teleseismic earthquakes. This indicates that the analysis of the wavefield coherence could be used for earthquake detection. Therefore, we establish a simple detection criteria based on our method and compare its results with the Global Centroid Moment Tensor (GCMT) catalogue (Ekström *et al.* 2012). We use the covariance matrix spectral width shown in Fig. 8(a), because the normalized results obtained in Fig. 8(b) are saturated by the non-propagating noise, as discussed in the previous section.

Observation and detection of teleseismic earthquakes during 8 d in 2010 June is illustrated in Fig. 10. Signals from strongest events (such as Nicobar earthquakes occurred on June 12) are clearly seen on seismograms (Fig. 10d). In addition, the analysis of the covariance matrix width (Figs 10a and b) reveals peaks from weaker earthquakes not directly observable on seismograms. Fig. 10(b) shows the covariance matrix width $\hat{\sigma}(t)$ averaged between 0.03 and 0.12 Hz (delimited with a black frame in Fig. 10a) that we use in our simple detection algorithm. First, we identify all intervals when the averaged spectral width is below the median of $\hat{\sigma}(t)$ (shown with the horizontal dotted line in Fig. 10b). To reduce the influence of the small noisy peaks, we establish a detection threshold σ_{th} (shown with the horizontal dashed line in Fig. 10b) and keep only the intervals whose absolute minima are below this threshold. We choose $\sigma_{\text{th}} = 3.3$ in this example. The retained intervals are considered as *alarms* and are shown with coloured boxes in Figs 10(b)–(d).

Finally, we compare the obtained detections with the GCMT catalogue. For every earthquake from this catalogue, we compute the arrival time of Rayleigh waves at the position of the Undervolc array. An earthquake is considered as *detected* when its arrival time at La Réunion Island is within an *alarm* window (coloured boxes in Figs 10b–d).

Even after removing the amplitude information with the preprocessing technique present in Section 6, the spatial coherence of the recorded signals related to the earthquake and the expected detection efficiency depend on the signal-to-noise ratio. This ratio depends in turn on the earthquake magnitude and on the epicentral distance between the earthquake and La Réunion Island. To simplify this relationship, we normalize the magnitudes to compensate the differences in the epicentral distances Δ. The corrected magnitudes M_{eff} correspond to distance of 90 deg (derived from Aki & Richards 2002, app. 2):

$$M_{\text{eff}} = M_s + 1.656\log(\Delta) + 1.656\log(90)\,. \tag{16}$$

Corrected times and magnitudes of earthquakes from the GCMT catalogue are shown with coloured circles in Fig. 10(c). During the considered period from June 9 to June 17, the GCMT catalogue lists 32 earthquakes with effective magnitude greater than 5. Earthquakes with arrival times that are not present in a detection window are considered as *undetected*, and are shown in red circles. Otherwise, they are considered as *detected*, and are shown in blue circles. In a similar way, time segments of $\hat{\sigma}(t)$ that contain no earthquakes are considered as *false alarms* and are represented in red colour, while those that contain one or more earthquakes are considered as *detections* and are represented in blue.

6.5 Observation of the teleseismic activity during 2010

We applied the method described in the previous subsection to the continuous records of 2010. The relative amounts of *detections* and *false alarms* and of *detected* and *undetected* events depends on the choice of the detection threshold σ_{th} and of the minimal effective magnitude of earthquakes taken into account. We define two ratios R^{real} and R^{succ}. The first one, R^{real}, is defined as the ratio between the number of *detections* and the total number of *alarms*. The second one, R^{succ}, is the ratio of the *detected* earthquakes over the total number of earthquakes listed in the GCMT catalogue. We vary the detection threshold from 2 to 3.5 and the minimal effective magnitude from 4 to 7.5, and observe the variation of R^{real} and

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

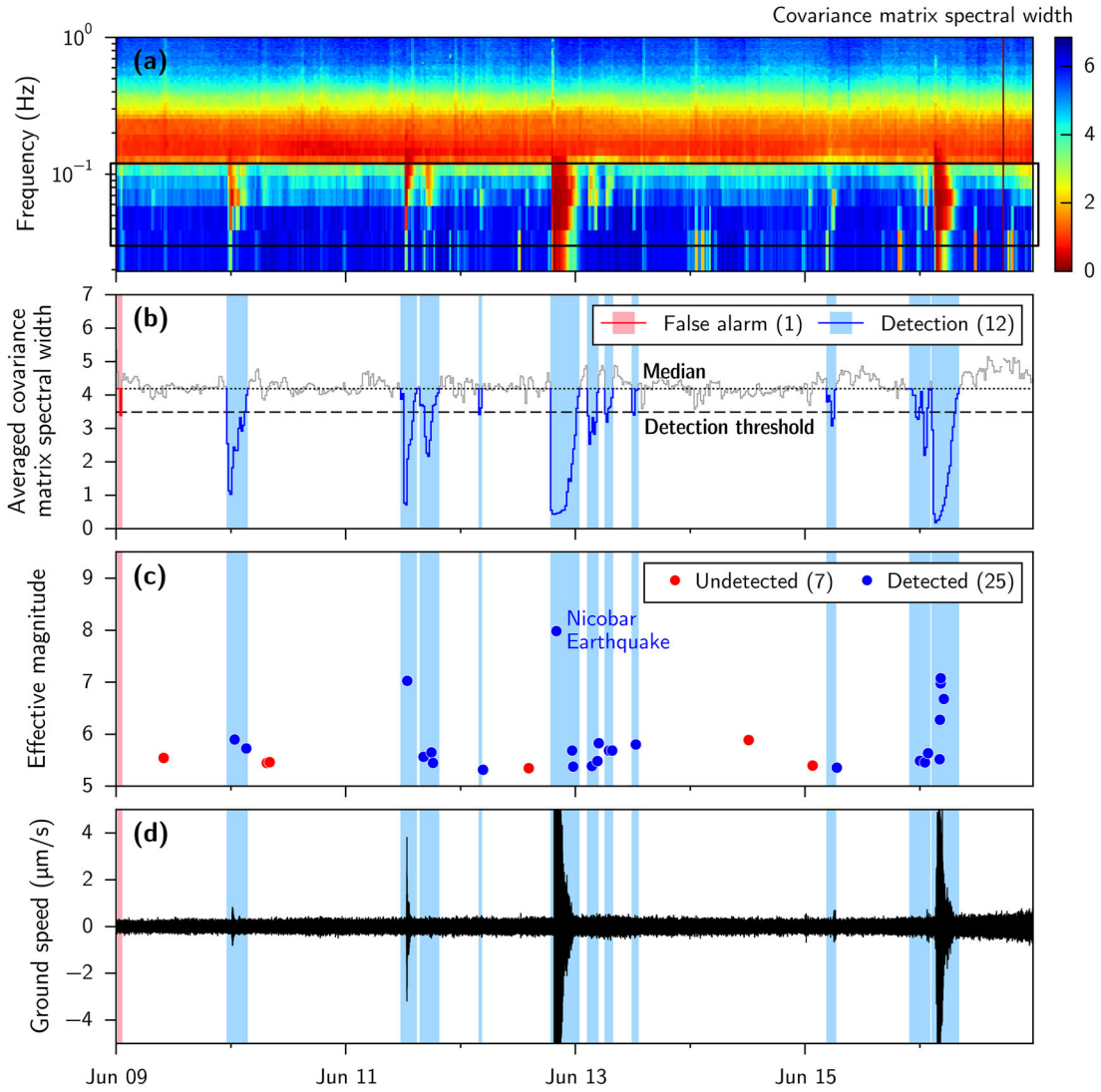


Figure 10. Example of earthquake detection around 2010 June 12. Observation of teleseismic earthquakes at frequencies below 0.12 Hz. (a) Raw covariance matrix spectral width. The black frame indicates the zone where the covariance matrix spectral width is averaged. (b) Covariance matrix spectral width averaged between 0.03 and 0.12 Hz. Coloured segments indicate time period where a detection is declared. Blue segment are *detections* that correspond to at least one event from the GCMT catalogue, red segments are *false alarms*. (c) Arrival times of earthquakes from the GCMT catalogue as a function of time and effective magnitude (eq. 16). Blue circles show the *detected* events and red circles show the *undetected* ones. (d) Seismic trace recorded by the Undervolc station UV14 as a function of time, bandpass filtered between 0.03 and 0.12 Hz.

R^{succ} in Figs 11(a) and (b), respectively. We found that our simple detection algorithm is 'optimized' with the detection threshold of 3.3 and with the minimal effective magnitude of 5.3. With these parameters, 75 per cent of *alarms* are successful *detections* and we detect 75 per cent of earthquakes (951 *detections* out of 1246 total events). The final detection results during the year 2010 are shown in Fig. 12.

6.6 Observation of the seismovolcanic activity

We use the covariance matrix spectral width at high frequencies to detect the seismovolcanic activity because the volcanic sources are close to the network and the high-frequency radiation is not strongly attenuated as was the case for the teleseismic earthquakes. We average the normalize covariance matrix spectral width between 1 and 5 Hz (frequency region delimited with a black dashed frame in Fig. 8b), and present it as a function of time in Fig. 13(a). We see that the signal is not coherent (the average covariance matrix spectral width is close to 1) most of the time. A set of coherent events are detected as minima of spectral width. We compare their occurrences and durations with the catalogue of Piton de la Fournaise activity described in Roult *et al.* (2012). Figs 13(b)–(d) are zoomed in periods of known volcanic eruptions.

In the three cases presented in Figs 13(b)–(d), the occurrence and duration of volcanic tremors are well observed. In these three cases, the yellow windows indicate the durations of the volcanic tremors, and show great agreement with the periods of increased coherence. The green windows show the seismic crises recorded before the eruptions. The eruption of October is particularly visible because it was the most energetic in 2010. The smaller eruptions of January and December are also clearly distinguishable.

Numerous minima of the covariance matrix spectral with averaged between 1 and 5 Hz are also observed outside the catalogued

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

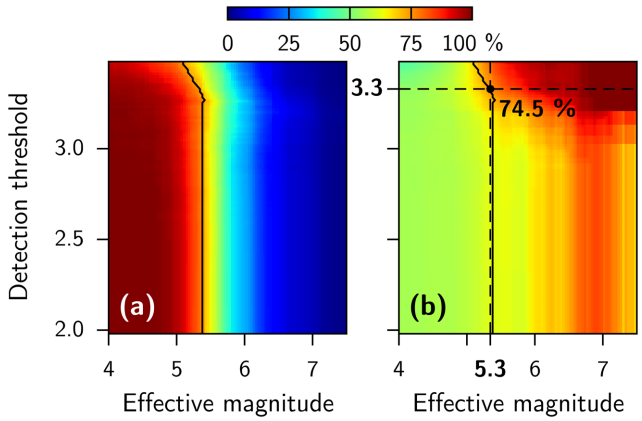


Figure 11. Analysis of observation of teleseismic earthquakes. (a) R^{real} as function of the detection threshold and the minimal effective magnitude. The black line indicates points where $R^{\text{real}} = 75$ per cent. (b) R^{succ} as function of the detection threshold and the minimal effective magnitude. The solid black line corresponds to the coordinates where $R^{\text{real}} = 75$ per cent. The maximal value of R^{succ} on these coordinates is 74.5 per cent and is reached when the detection threshold is 3.3 and the minimal effective magnitude is 5.3.

periods of volcanic activity. A few of them correspond to days with missing data (Fig. 8c). However, the large majority of these minima correspond to small volcanic earthquakes. In particular, we observe a long period (∼1.5 months) of increased seismic activity preceding the eruption in 2010 October. A detailed comparison with seismograms shows that the some small earthquakes are not reflected in the presented results of the covariance matrix analysis. This is related to the choice of the duration of the time windows δt and Δt in eq. (8). In this study, we selected relatively long windows to be able to analyse the long-period part of the signal, while a more detailed study of the seismovolcanic activity would require focusing at frequencies above 1 Hz and, therefore, running the covariance matrix analysis in shorter windows.

7 DISCUSSION AND CONCLUSIONS

We developed an approach that allows to quantify the level of spatial coherence of a wavefield recorded by a network of sensors with a single parameter that is the spectral width of the covariance matrix. We first tested this method with a set of synthetic signals and illustrated it with an application to the real data, namely, to one year of continuous vertical-component records from 21 broad-band seismographs installed around the Piton de la Fournaise volcano at La Réunion island. The main result of our analysis is presented in Fig. 8 and shows that the level of the seismic wavefield coherence varies strongly with respect to the time and frequency, and that clearly distinguishable coherent events are seen in different frequency ranges.

We distinguish three frequency bands with different characteristic behaviour of the wavefield coherence. Between 1 and 5 Hz, the wavefield is dominated most of the time by the propagating noise, or diffuse seismic field and the coherent signals are emitted by sources located very close to the network, mainly related to the seismovolcanic activity of the Piton de la Fournaise volcano. For such local sources, the high-frequency radiation is not attenuated. In the secondary microseismic band around 0.2 Hz, the wavefield is on average more coherent that at lower and higher frequencies. A detailed analysis of Fig. 8(b) shows that the microseismic noise in this frequency range is composed of many separated arrivals

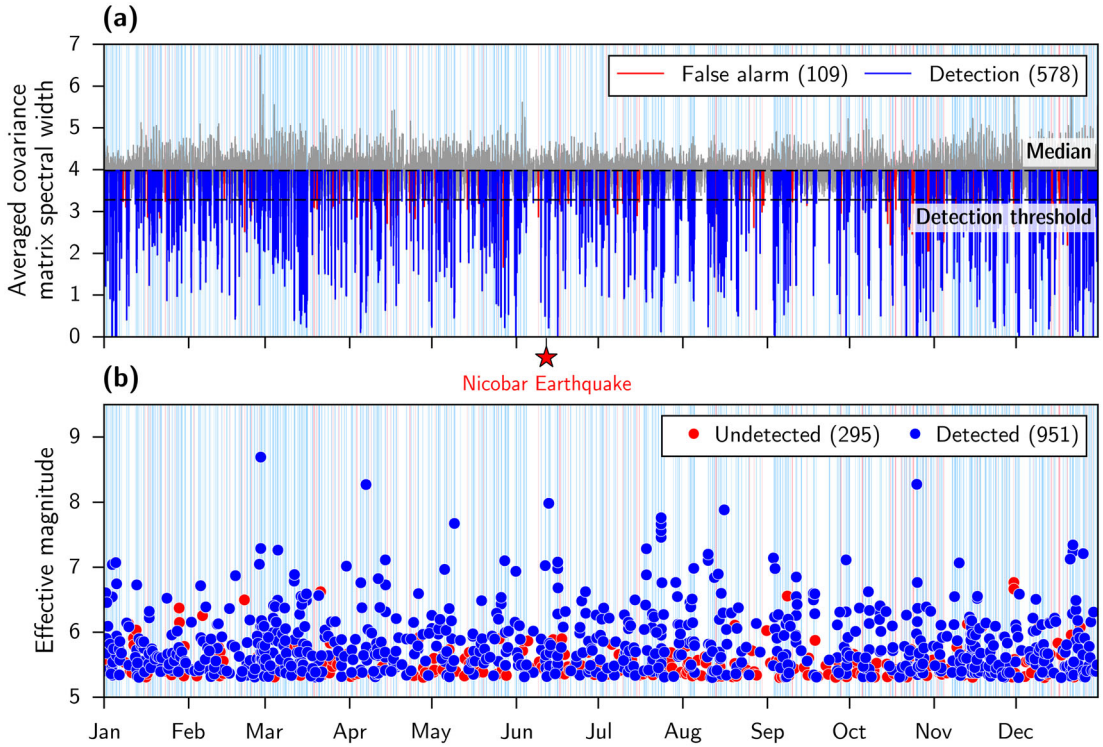


Figure 12. Earthquake detector. Observation of teleseismic earthquakes at frequencies below 0.1 Hz. (a) Normalized covariance matrix spectral width σ averaged between 0.03 and 0.12 Hz. Blue segments indicate time periods when the averaged value of σ that correspond to at least one event from the GCMT catalogue and red segments indicate time periods with no matching events from the GCMT catalogue. (b) Earthquakes from the GCMT catalogue as a function of time and effective magnitude (eq. 16). Blue and red circles show 'detected' and 'undetected' events, respectively.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

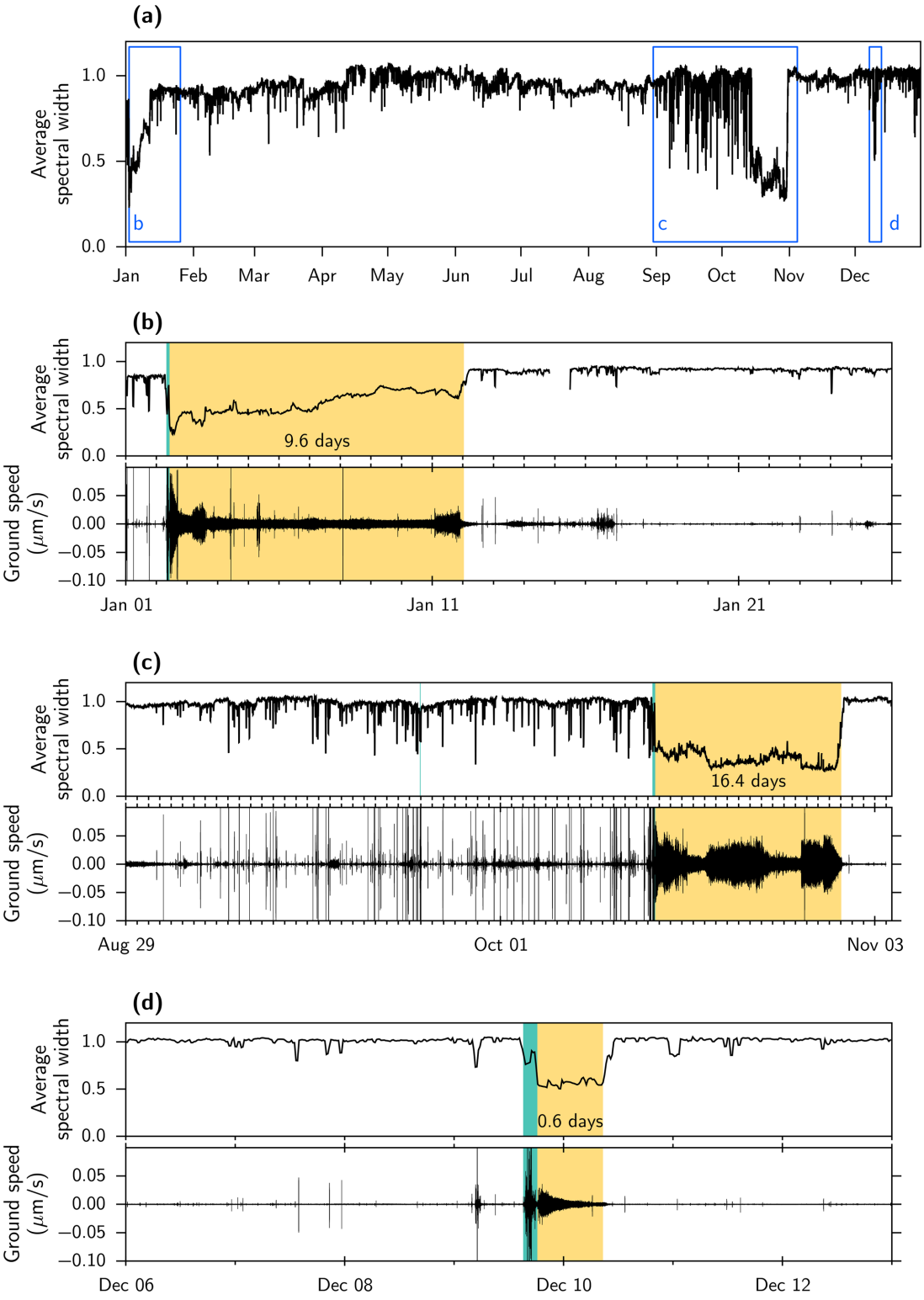


Figure 13. Observation of the seismovolcanic activity at frequencies above 1 Hz. (a) Covariance matrix spectral width σ averaged between 1 and 5 Hz and shown for the whole year 2010. Blue frames indicate periods corresponding to three eruptive crises occurred in 2010. (b) Zoom between January 1 and January 20. (c) Zoom between August 28 and November 16. (d) Zoom between December 3 and December 14. In each case, green boxes indicate pre-eruptive seismic crisis and yellow boxes indicate starting and ending dates of eruptions from Roult *et al.* (2012). Records from the UV05 seismic station filtered between 1 and 5 Hz are presented in each case.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

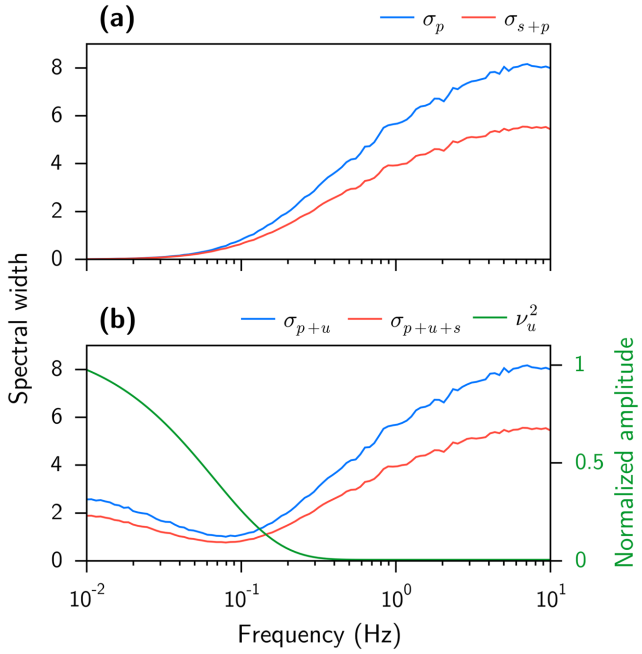


Figure 14. (a) Covariance matrix spectral width as function of frequency estimated for a combination of two first terms from eq. (3): with an propagating noise only (σ_p) and adding one coherent signal (σ_{s+p}). (b) Similar to (a) but after adding the third term: the non-propagating noise (whose spectral intensity $\nu_n^2(f)$ is schematically illustrated with the green line).

characterized by the increased values of the coherence. The duration of these microseismic coherent arrivals varies between a few hours and a few tens of hours and they likely correspond to signals generated by separated storms or swells. At lowest frequencies, below 0.1 Hz, the records are dominated by the non-propagating noise because of the non-perfect conditions of temporary installation of seismic sensors on a volcano. The coherent arrivals from the teleseismic earthquakes are clearly distinguishable in this frequency in contrast with the incoherent noise leading to a possibility to detect them based on the analysis of the covariance matrix spectra.

The presented example clearly shows that the array covariance matrix is very sensitive to the coherent signals generated by strong earthquakes and by the volcano. At the same time, the simple approach discussed in our paper does not aim to outperform more classical earthquake detection algorithms. The UndeVolc array is not and well suited for the detection of teleseismic earthquakes. This temporary seismic network was installed on an active volcano where most of the sites were affected by strong environmental noise generated either by volcanic or by meteorological activity. The detection based on the covariance matrix spectral width can strongly improve the analysis of emergent signals such as, for example, volcanic tremors illustrated in Fig. 13. The detection of such signals could be significantly enhanced with a choice of parameters used in the covariance matrix computation (the number of subwindows M and of their length δt) specifically optimized for high frequencies and not for a broad-band analysis presented in this paper. Such a more detailed study is, however, out of the scope of this initial paper mainly aimed at presenting the theory and the method of the covariance matrix analysis.

Fig. 14 illustrates relative contributions of different types of signals described in eq. (3) within these bands. We see in Fig. 14(a) that covariance matrix spectral width of the propagating noise (in blue line) decreases with the decreasing frequency. This indicates that the spatial coherence of the propagating noise increases when the corresponding wavelength becomes too wide in comparison with the interstation spacing and the array aperture. The red line shows the same propagating noise, with an additional 10 times stronger coherent source. The detection of this coherent source is therefore hard to accomplish because the difference between the propagating noise and the source embedded in propagating noise spectral width is weak a frequencies below 0.15 Hz. The covariance matrix spectral width of the data shows that the wavefield contains incoherent signal at frequencies below 0.1 Hz which can be explained by the presence of non-propagating noise at the sensors, most probably due to the weather perturbations. The detectability of the seismic source in the presence of this additional non-propagative noise is presented in Fig. 14(b). The green line shows the non-propagating noise spectral amplitude. The noise amplitude is chosen to be the same as the signal's one and rapidly decays with the increasing frequency. The source is now detected, and this model explains the detection capacity of our method for this data set.

The results of the covariance matrix analysis are not unique and depend on the choice of the time windows δt and Δt in eq. (8). In the presented example of the application of the method to the real data we decided, for simplicity, to use the constant set of these windows for all frequencies and selected the long windows lengths that allowed us to analyse the longest periods. As a consequence of this choice, the results of the analysis of the seismovolcanic activity shown in Fig. 13 are not optimal, as discussed in the previous section. This shows that in the future studies the choice of the windows δt and Δt should be optimized considering the specific properties of the studied signals and, in particular, their dominant frequency range.

The analysed seismic records were pre-processed with applying a spectral whitening and a temporal normalization to compensate for inhomogeneity of amplitudes, which is a standard practice in many applications based on correlations of the ambient seismic noise (e.g. Bensen *et al.* 2007). Our results show that, while these procedures compensate for the effect of strongest amplitudes, they are not sufficient to fully 'randomize' the records and that a clear imprint of different seismic sources remains in the wavefield coherence.

ACKNOWLEDGEMENTS

This work was supported by the City of Paris under the programme Emergence', by the European Research Council under the contract FP7 ERC Advanced grant 227507 (WHISPER) and by the Russian Science Foundation (grant 14-47-00002). Computations in this study were performed using the High-Performance Computing infrastructure S-CAPAD at the Institut de Physique du Globe de Paris, which is supported by the Île-de-France region (via the SEASAME programme), France-Grille (http://www.france-grilles.fr), and the CNRS MASTODONS programme.

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* University Science Books, 704 pp.

Allen, R., 1982. Automatic phase pickers: their present use and future prospects, *Bull. seism. Soc. Am.,* **72**(6B), S225–S242.

Aubry, A. & Derode, A., 2009a. Random matrix theory applied to acoustic backscattering and imaging in complex media, *Phys. Rev. Lett.,* **102**(8), doi:10.1103/PhysRevLett.102.084301.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022

Aubry, A. & Derode, A., 2009b. Detection and imaging in a random medium: a matrix method to overcome multiple scattering and aberration, *J. Appl. Phys.,* **106**(4), 044903, doi:10.1063/1.3200962.

Baillard, C., Crawford, W.C., Ballu, V., Hibert, C. & Mangeney, A., 2014. An automatic Kurtosis-based P-and S-phase picker designed for local seismic networks, *Bull. seism. Soc. Am.,* **104**(1), 394–409.

Bartlett, M.S., 1954. A note on the multiplying factors for various χ^2 approximations, *J. R. Stat. Soc. B (Met.),* **16**(2), 296–298.

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., Shapiro, N. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169**(3), 1239–1260.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N., Nadeau, R.M. & Larose, E., 2008. Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations, *Science,* **321**(5895), 1478–1481.

Brenguier, F. *et al.*, 2012. First results from the UnderVolc high resolution seismic and GPS network deployed on Piton de la Fournaise volcano, *Seism. Res. Lett.,* **83**(1), 97–102.

Bungum, H., Husebye, E. & Ringdal, F., 1971. The NORSAR array and preliminary results of data analysis, *Geophys. J. Int.,* **25**(1–3), 115–126.

Chiou, S.-J. & Bolt, B.A., 1993. Seismic wave slowness-vector estimation from broad-band array data, *Geophys. J. Int.,* **114**(2), 234–248.

Cox, H., 1973. Spatial correlation in arbitrary noise fields with application to ambient sea noise, *J. acoust. Soc. Am.,* **54**(5), 1289–1301.

Ekström, G., Nettles, M. & Dziewoński, A., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200,** 1–9.

Friedrich, A., Krüger, F. & Klinge, K., 1998. Ocean-generated microseismic noise located with the Gräfenberg array, *J. Seismol.,* **2**(1), 47–64.

Frosch, R.A. & Green, P.E., 1966. The concept of a large aperture seismic array, *Proc. R. Soc. Lond., A: Math. Phys. Eng. Sci.,* **290**(1422), 368–384.

Gerstoft, P., Menon, R., Hodgkiss, W.S. & Mecklenbräuker, C.F., 2012. Eigenvalues of the sample covariance matrix for a towed array, *J. acoust. Soc. Am.,* **132**(4), 2388–2396.

Goldstein, P. & Archuleta, R.J., 1987. Array analysis of seismic signals, *Geophys. Res. Lett.,* **14**(1), 13–16.

Lawley, D., 1956. Tests of significance for the latent roots of covariance and correlation matrices, *Biometrika,* **43**(1–2), 128–136.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc. Lond., A.: Math. Phys. Eng. Sci.,* **243**(857), 1–35.

Marchenko, V.A. & Pastur, L.A., 1967. Distribution of eigenvalues for some sets of random matrices, *Mat. Sb.,* **114**(4), 507–536.

Menon, R., Gerstoft, P. & Hodgkiss, W.S., 2012a. Asymptotic eigenvalue density of noise covariance matrices, *IEEE Trans. Signal Process.,* **60**(7), 3415–3424.

Menon, R., Gerstoft, P. & Hodgkiss, W.S., 2012b. Cross-correlations of diffuse ocean noise using eigenvalue based statistical inference, *J. acoust. Soc. Am.,* **132**(5), 3213–3224.

Mestre, X., 2008. On the asymptotic behavior of the sample estimates of eigenvalues and eigenvectors of covariance matrices, *IEEE Trans. Signal Process.,* **56**(11), 5353–5368.

Müller, R.R., 2002. A random matrix model of communication via antenna arrays, *IEEE Trans. Inform. Theory,* **48**(9), 2495–2506.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 2-1–2-27.

Roult, G., Peltier, A., Taisne, B., Staudacher, T., Ferrazzini, V., Di Muro, A. & the OVPF team, 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. Volc. Geotherm. Res.,* **241,** 78–104.

Roux, P., Kuperman, W.A. & the NPAL Group, 2004. Extracting coherent wave fronts from acoustic ambient noise in the ocean, *J. acoust. Soc. Am.,* **116**(4), 1995–2003.

Saragiotis, C.D., Hadjileontiadis, L.J. & Panas, S.M., 2002. PAI-S/K: a robust automatic seismic P phase arrival identification scheme, *IEEE Trans. Geosci. Remote Sens.,* **40**(6), 1395–1404.

Sengupta, A.M. & Mitra, P.P., 1999. Distributions of singular values for some random matrices, *Phys. Rev. E,* **60**(3), doi:10.1103/PhysRevE.60.3389.

Tulino, A.M. & Verdú, S., 2004. *Random Matrix Theory and Wireless Communications*, Vol. 1, Now Publishers Inc.

Wagner, G.S. & Owens, T.J., 1996. Signal detection using multi-channel seismic data, *Bull. seism. Soc. Am.,* **86**(1A), 221–231.

Wax, M. & Kailath, T., 1985. Detection of signals by information theoretic criteria, *IEEE Trans. Acoust. Speech Signal Process.,* **33**(2), 387–392.

Welch, P.D., 1967. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.,* **15**(2), 70–73.

Downloaded from https://academic.oup.com/gji/article/204/3/1430/677678 by guest on 20 February 2022