Geophysical Journal International

Geophys. J. Int. (2018) **214,** 1218–1236
Advance Access publication 2018 May 18
GJI Seismology

doi: 10.1093/gji/ggy190

Retrieving robust noise-based seismic velocity changes from sparse data sets: synthetic tests and application to Klyuchevskoy volcanic group (Kamchatka)

C. Gómez-García,[1,*] F. Brenguier,[2] P. Boué,[2] N.M. Shapiro,[3,4] D.V. Droznin,[5] S.Ya. Droznina,[5] S.L. Senyukov[5] and E.I. Gordeev[6]

[1]*Instituto de Ciencias de la Tierra Jaume Almera, CSIC, Lluís Solé i Sabarís s/n,* E-08028 *Barcelona, Spain. E-mail: cgomez@cp.dias.ie*
[2]*Institut des Sciences de la Terre, Université Grenoble Alpes,* F-38041 *Grenoble cedex* 9*, France*
[3]*Institut de Physique du Globe de Paris, Paris Sorbonne Cité, CNRS,* 1 *rue Jussieu,* F-75238 *Paris cedex 5, France*
[4]*Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, ul. B. Gruzinskaya* 10*, Moscow* 123995*, Russia*
[5]*Kamchatka Branch of the Geophysical Service, Russian Academy of Sciences,* 9 *Piip Boulevard, Petropavlovsk-Kamchatsky, Kamchatsky Region* 683006*, Russia*
[6]*Institute of Volcanology and Seismology FEB RAS,* 9 *Piip Boulevard, Petropavlovsk-Kamchatsky, Kamchatsky Region* 683006*, Russia*

Accepted 2018 May 14. Received 2018 May 11; in original form 2018 May 9

SUMMARY

Continuous noise-based monitoring of seismic velocity changes provides insights into volcanic unrest, earthquake mechanisms and fluid injection in the subsurface. The standard monitoring approach relies on measuring traveltime changes of late coda arrivals between daily and reference noise cross-correlations, usually chosen as stacks of daily cross-correlations. The main assumption of this method is that the shape of the noise correlations does not change over time or, in other terms, that the ambient-noise sources are stationary through time. These conditions are not fulfilled when a strong episodic source of noise, such as a volcanic tremor, for example, perturbs the reconstructed Green's function. In this paper, we propose a general formulation for retrieving continuous time-series of noise-based seismic velocity changes without the requirement of any arbitrary reference cross-correlation function (CCF). Instead, we measure the changes between all possible pairs of daily cross-correlations and invert them using different smoothing parameters to obtain the final velocity change curve. We perform synthetic tests in order to establish a general framework for future applications of this technique. In particular, we study the reliability of velocity change measurements versus the stability of noise CCFs. We apply this approach to a complex data set of noise cross-correlations at Klyuchevskoy volcanic group (Kamchatka), hampered by loss of data and the presence of highly non-stationary seismic tremors.

Key words: Time-series analysis; Coda waves; Seismic interferometry; Seismic noise; Volcano monitoring.

1 INTRODUCTION

Noise-based monitoring techniques have been used extensively in the past decade for different applications. The observation of continuous seismic velocity changes proved to be useful for detecting crustal seasonal changes (e.g. Sens-Schönfelder & Wegler 2006; Meier *et al.* 2010; Ugalde *et al.* 2014), co- and post-seismic evolution of stress in fault areas (e.g. Brenguier *et al.* 2008a; Hobiger *et al.* 2012) and, more recently, for studying the effects of fluid injection (e.g. Zhou *et al.* 2010; Ugalde *et al.* 2013) and aseismic deformation transients (Hillers *et al.* 2015).

Estimation of temporal velocity changes in volcano interiors using seismic noise cross-correlation has been shown to be an efficient method for early detection of volcanic unrest prior to eruptions at Piton de la Fournaise Volcano, La Réunion (e.g. Brenguier *et al.* 2008b; Duputel *et al.* 2009). Although precise eruption and eruptive intensity forecasting is still a challenge, it has been demonstrated that this method provides meaningful constraints on the location of oncoming eruptions (Obermann *et al.* 2013).

The most important step in noise-based monitoring is the Green's function (GF) reconstruction between two receivers from the correlation of ambient seismic noise (e.g. Shapiro & Campillo 2004;

* Now at: Dublin Institute for Advanced Studies, Geophysics Section, 5 Merrion Square, Dublin 2, Ireland.

© The Author(s) 2018. Published by Oxford University Press on behalf of The Royal Astronomical Society.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

Shapiro *et al.* 2005; Larose *et al.* 2006; Wapenaar *et al.* 2010; Campillo *et al.* 2011). If the noise sources are evenly distributed over the Earth's surface, leading to an isotropic and equipartioned wavefield at the two station locations, the cross-correlation function (CCF) between these two stations converges towards the GF (e.g. Roux *et al.* 2005; Wapenaar & Fokkema 2006). This is an ideal situation but, in practice, noise sources are distributed irregularly leading to a partial reconstruction of the GF (Shapiro *et al.* 2006).

For monitoring purposes, it is possible to retrieve temporal seismic velocity changes over a set of repetitive in-time noise cross-correlations, even with anisotropic distributions of noise sources, as long as this distribution does not change too much over time (Hadziioannou *et al.* 2009). Moreover, measuring traveltime changes in the coda part of the noise cross-correlations makes velocity change measurements less sensitive to noise source temporal changes (Sens-Schönfelder & Wegler 2006; Wegler & Sens-Schönfelder 2007; Colombi *et al.* 2014). The standard monitoring approach relies on measuring traveltime changes of late coda arrivals between a daily and a reference noise-cross-correlation, usually chosen as a stack of all daily cross-correlations. There are two main techniques for the retrieval of the relative velocity changes. One is the stretching technique for which the relative velocity variation is the best-fitting factor by which the time axis of the current CCF is stretched or compressed to obtain the best correlation with the reference trace (e.g. Obermann *et al.* 2013; Sens-Schönfelder *et al.* 2014; Hillers *et al.* 2015). The second technique is the moving window cross-spectral (MWCS) analysis where we obtain the relative velocity change by adjusting, in the frequency domain, the phase differences between the current and the reference CCF in each time window (e.g. Brenguier *et al.* 2008b; Clarke *et al.* 2011). Advantages and disadvantages of both techniques have been discussed in the literature (e.g. Hadziioannou *et al.* 2009; Hillers *et al.* 2015). In this study, we use the MWCS technique because we consider that this technique is less sensitive to source variability. We assume that the measured time delay from the coda waveform of noise cross-correlations ($\mathrm{d}\tau$) is caused by a spatially homogeneous relative velocity change, $\mathrm{d}v/v$. Under this assumption, the relative delay time ($\mathrm{d}\tau/\tau$) is constant and independent of the lapse time at which it is measured: $\mathrm{d}\tau/\tau = -\mathrm{d}v/v$.

In different environments, and especially on volcanoes, the noise correlations can be altered by strong episodic sources of noise, such as a volcanic tremor, for example, that overlaps in frequency with more stable microseismic noise. Thus, there is a problem with the definition of the reference function if the sources are non-stationary (Sens-Schönfelder *et al.* 2014). Very strong non-stationary noise has been described by Ballmer *et al.* (2013) and Droznin *et al.* (2015) in case of emission of low frequency volcanic tremor, a typical feature of the unrest of many volcanoes and an important seismic source for monitoring plumbing systems (e.g. Chouet 1996).

In this paper, we describe a generalized approach for retrieving robust noise-based seismic velocity changes, where the final time-series is obtained by measuring the changes between all possible pairs of CCFs and inverting them (Brenguier *et al.* 2014, Section 2), that is, without the definition of an arbitrary reference CCF. We detail the method carrying out synthetic tests that allow us to evaluate the reliability of measured velocity changes in regard to the level of stability of noise CCFs and the influence of temporary changes (Section 3). Finally, we apply our procedure to a data set from the Klyuchevskoy volcanic group (Kamchatka) as a case study (Section 4), where the recorded wavefield is dominated by strongly localized volcanic tremor sources and is characterized by loss of data.

2 METHOD

We retrieve continuous time-series of velocity changes without the requirement of a reference stacked CCF. The procedure relies on measuring seismic velocity changes between all possible pairs of daily CCFs. An inversion step is further required to retrieve a continuous time-series of daily seismic velocity changes (Brenguier *et al.* 2014). By considering (ccf_i) as a CCF that corresponds to day i, we can thus estimate a seismic velocity change between day i and day j (δv_{ij}) by applying the MWCS analysis to ccf_i and ccf_j:

$$\delta v_{ij} = \frac{v_j - v_i}{v_i} = \mathrm{MWCS}(ccf_i, ccf_j), \tag{1}$$

where δv_{ij} is referred as a doublet measurement. This concept was used, initially, in pairs of microearthquakes (Poupinet *et al.* 1984). In a systematic manner, we can then estimate a velocity change between all of the pairs of daily CCFs for one given station pair. This constitutes the data vector of eq. (2):

$$\mathbf{d} = \begin{bmatrix} \delta v_{12} \\ \delta v_{13} \\ \delta v_{14} \\ \vdots \\ \delta v_{n-1n} \end{bmatrix}, \tag{2}$$

where $\mathbf{d}$ is of length $\frac{n\cdot(n-1)}{2}$, with n the number of daily CCFs.

Our final goal is to reconstruct the time-series of daily velocity changes. We can define these velocity changes as $\delta v_i = \frac{v_i - v_{\mathrm{ref}}}{v_{\mathrm{ref}}}$, with v_{ref} the reference velocity averaged along the entire time period considered. The series of velocity changes constitutes our model vector, $\mathbf{m}$, of eq. (3):

$$\mathbf{m} = \begin{bmatrix} \delta v_1 \\ \delta v_2 \\ \delta v_3 \\ \vdots \\ \delta v_n \end{bmatrix}, \tag{3}$$

where $\mathbf{m}$ is of length n, the number of daily CCFs.

The relation between $\mathbf{d}$ and $\mathbf{m}$ is given by

$$\delta v_j - \delta v_i = \frac{v_j - v_i}{v_{\mathrm{ref}}} = \frac{v_j - v_i}{v_i} \cdot \frac{v_i}{v_{\mathrm{ref}}} = \delta v_{ij} \cdot \frac{v_i}{v_{\mathrm{ref}}} = \delta v_{ij} \cdot (1 + \delta v_i). \tag{4}$$

Under the assumption that δv_i and δv_{ij} are small compared to 1 ($<0.1\%$), we can write at the first order the direct linear relationship between $\mathbf{d}$ and $\mathbf{m}$ as $\delta v_{ij} = \delta v_j - \delta v_i$ or $\mathbf{d} = \mathbf{G}\mathbf{m}$, with $\mathbf{G}$ being a sparse matrix of dimension $\left[\frac{n\cdot(n-1)}{2}, n\right]$:

$$\mathbf{G} = \begin{bmatrix} -1 & 1 & 0 & \dots & & & \dots & 0 \\ -1 & 0 & 1 & 0 & \dots & & & \vdots \\ -1 & 0 & 0 & 1 & 0 & \dots & & \\ \vdots & & & & & \ddots & & \vdots \\ 0 & \dots & & & & \dots & 0 -1 & 1 \end{bmatrix}. \tag{5}$$

The assumption made above (δv_i and $\delta v_{ij} <$ 0.1%) is necessary to apply our method. Temporal velocity changes (δv_i) are sensitive to transient stress changes (e.g. Niu *et al.* 2008) and the magnitude order of the seismic velocity changes depends on the level of applied stress in the medium. Some examples of typical magnitude orders of δv_i estimations are $\sim -0.1\%$ in the Piton de la Fournaise volcano (Brenguier *et al.* 2008b; Obermann *et al.* 2013), $\sim -0.12\%$ due to the Tohoku-Oki earthquake (Brenguier *et al.* 2014), $\sim -0.15\%$

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

due to the Parkfield earthquake (Schaff 2012), $\sim -0.5\%$ due to the Nicoya Peninsula earthquake (Chaves & Schwartz 2016) or $\sim -0.8\%$ in Ruapehu volcano (Mordret *et al.* 2010).

The final time-series of velocity changes ($\mathbf{m}$) is obtained by further inversion, using a classical Bayesian linear least-squares formulation (Tarantola 2005; details in Brenguier *et al.* 2014):

$$\mathbf{m} = \left(\mathbf{G}^t\mathbf{C}_\mathbf{d}^{-1}\mathbf{G} + \alpha\mathbf{C}_\mathbf{m}^{-1}\right)^{-1}\mathbf{G}^t\mathbf{C}_\mathbf{d}^{-1}\mathbf{d}, \tag{6}$$

where $\mathbf{C_d}$ is a covariance matrix of dimension $\left[\frac{n\cdot(n-1)}{2}, \frac{n\cdot(n-1)}{2}\right]$ that describes the Gaussian uncertainties of the data vector $\mathbf{d}$. These values correspond to the estimated uncertainties of each $\delta\nu_{ij}$ value, using the MWCS analysis. $\mathbf{C_m}$ is an *a priori* covariance matrix of dimension $[n, n]$ for model vector $\mathbf{m}$. The parameter α is a weighting coefficient: it is determined in a way that matrix $\left(\mathbf{G}^t\mathbf{C}_\mathbf{d}^{-1}\mathbf{G}\right)$ and $\left(\alpha\mathbf{C}_\mathbf{m}^{-1}\right)$ have approximately the same weight. As α behaves as the amplitude of the inverse of the distribution $\mathbf{C_m}$, the higher the α, the less the model can change from one point to another. Therefore, the amplitude of the final time-series becomes lower and smoother.

The values of $\mathbf{C_m}$ describe for day i how $\delta\nu_i$ is correlated to $\delta\nu_j$ at day j:

$$C_{m_{ij}} = \mathrm{e}^{\frac{-|i-j|}{2\beta}}, \tag{7}$$

where β is the characteristic correlation length between the model parameters $\delta\nu_\mathrm{i}$. A day, i, is more correlated with the β days before and after than with any others. For this reason, high values of β correspond to long-term variations (LTV), whereas low β values represent short-term variations (STV), the opposite situation.

We compute the difference data − misfit to have an estimation of how well-constrained the inversion is. The misfit values are the differences between all pairs of points of the reconstructed time-series of velocity changes. To assess the data − misfit, we subtract each doublet measurement to the corresponding misfit and average the absolute value of the final result.

In Fig. 1 we compare the standard and the general approach. Even though the computing cost of the general formulation is higher than that of the standard approach, this formulation manifests several advantages. We can manage irregular sampling in time of noise correlations; therefore, this technique is more efficient when the data set is complex. Also, long-term or short-term trends are obtained directly from the inversion process rather than fitting the velocity changes with polynomial functions, as in the standard approach (Brenguier *et al.* 2008b).

In this work, we consider station pairs independently to obtain single time-series of velocity fluctuations but we can also invert several ray paths at the same time to achieve a more homogeneous and general trend of seismic velocity variations rather than averaging over different time-series from different station pairs. By concatenating doublet measurements from different station pairs for a global inversion, the robustness of retrieved velocity changes improves as the effect of missing data is minimized.

In the following, we describe synthetic tests to state the advantages and limits of that novel approach.

3 SYNTHETIC TESTS

We analyse how the stability of noise-correlations influences the reconstruction of velocity change time-series for different cases:

(1)Seasonal-type trends which produce long-term periodic-type fluctuations, that is, long-term velocity changes.

(2)Rapid transient changes similar to those produced as a result of an earthquake or a volcanic eruption. The effect of those changes in the noise correlations is the retrieval of a sudden velocity drop (STV), corresponding to a permanent or almost permanent velocity change.

(3)Transient noise perturbations due to a local source emission, such as the perturbation induced by an episodic volcanic tremor (Droznin *et al.* 2015). The consequence is a sudden velocity drop and a sudden recovery, producing short- and medium-term velocity fluctuations.

We use a synthetic test approach by artificially stretching noise cross-correlations in order to simulate synthetic velocity changes. We further degrade the quality of the data set of noise cross-correlations by adding different levels of random noise in order to simulate unstable to stable noise cross-correlations. We then apply our novel method for reconstructing velocity changes and finally compare the 'expected' and the 'reconstructed' time-series of velocity changes. We also study the improvement of averaging the reconstructed time-series of velocity changes for different station pairs.

The Pearson correlation coefficient (coherence) between two synthetic noise cross-correlations is used as a proxy for the quality of the associated doublet measurement and used to build the $\mathbf{C_d}$ matrix of data weighting. The average of all Pearson correlation coefficients between all pairs of noise cross-correlations (CCFs) is referred as the coherence level (*coh*). This value describes the level of added random noise by varying from 0 (totally incoherent noise CCFs) to 1 (no random noise added).

We refer to velocity change measurements at a crustal scale using microseismic noise correlations in the frequency range from 0.1 to 1 Hz. However, this approach can be extended to other frequency domains and sources of seismic noise.

3.1 Long-term periodic-type fluctuation test

By stretching a single arbitrary CCF with different daily velocity changes (referred as expected velocity changes henceforth), we simulate daily synthetic CCFs. Fig. 2, right-hand panel, shows the expected velocity changes that we apply and the simulated long-term periodic-type velocity changes to retrieve. The rest of the panels of Fig. 2 are examples of synthetic CCFs with different levels of noise. The different panels of synthetic CCFs are associated with a coherence level (*coh*), which is a measure of the level of added random noise. By adding random noise, we are 'hiding' the original time-series of velocity changes that we want to reconstruct after inversion, that is, the 'expected' velocity changes.

We obtain the data vector of velocity changes, $\mathbf{d}$, by applying an MWCS analysis between all possible pairs of CCFs. For n daily CCFs, we estimate $\frac{n(n-1)}{2}$ doublet measurements. We measure doublets in windows of 10 s centred between the direct surface wave arrival time and a lapse time of 70 s in the coda. Moving windows are overlapped by 80%. We finally perform the inversion to retrieve daily velocity changes (vector $\mathbf{m}$). As we are studying LTVs, we use high β values to retrieve $\mathrm{d}\nu/\nu$ series, $\beta = 1000$, while α decreases with *coh*, from $\alpha = 5000$ to $\alpha = 100$, to fit better the expected velocity change curve.

We compare the reconstructed time-series of velocity changes obtained from the synthetic CCFs of Fig. 2 with the expected one (Fig. 3). The more noise we add, the less the coherence level and the more the reconstructed time-series of velocity changes differs from the expected velocity changes.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

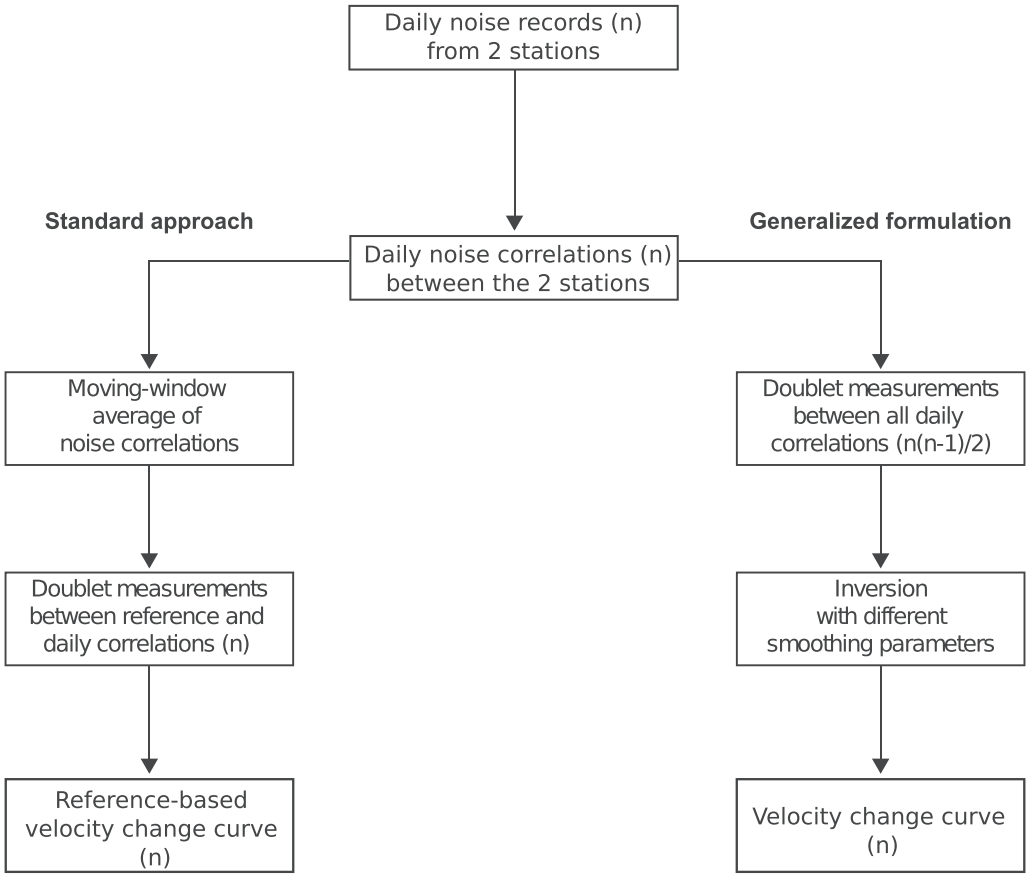


Figure 1. Workflow diagram showing the main steps of the standard approach and the general one. *n* is the number of days.

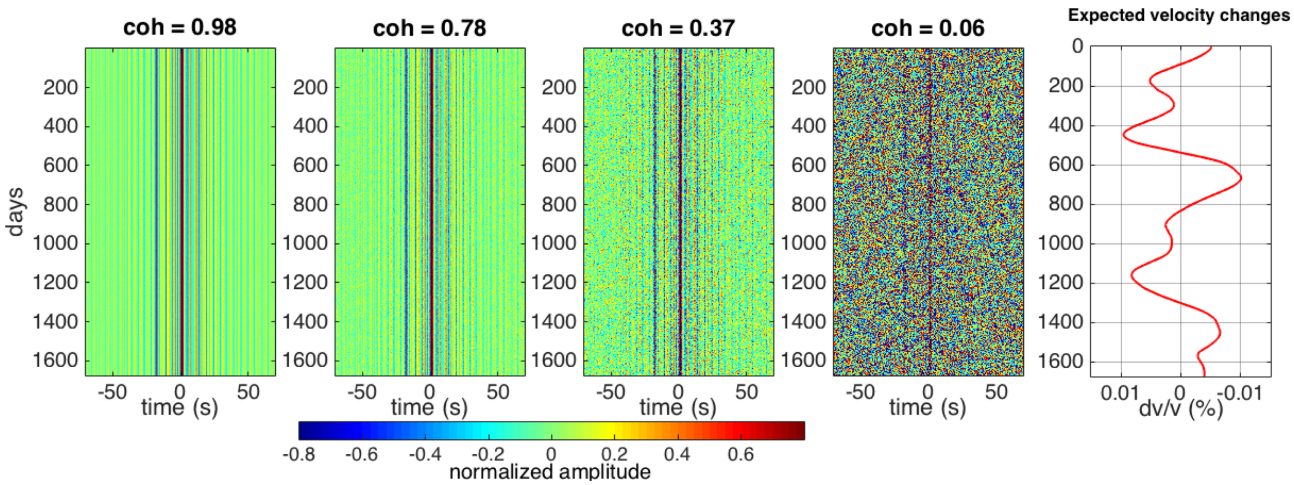


Figure 2. Examples of synthetic stretched CCFs with different levels of random noise. The coherence level (*coh*) is on top of each figure. On the right, expected velocity changes (red curve) that are applied to stretch the CCFs.

We test three different levels of expected velocity changes (Fig. 4a) to achieve the final one. The peak amplitude of the expected velocity change curve 1 is 0.001%, while expected velocity change curves 2 and 3 present peak amplitudes of 0.005 % and 0.01%, respectively. For Figs2 and 3, we use the expected velocity curve 3. By considering higher velocity change amplitudes (expected velocity change curve 3), we achieve higher similarity between the reconstructed time-series of velocity changes and the expected ones, for the same level of noise (Fig.4b).

To simulate the averaging of inverted time-series of velocity changes over different station pairs, we compute different station pairs with synthetic cross-correlation data: we apply the same velocity change stretching procedure but with different random noise to simulate different synthetic station pairs. We use the expected velocity change curve 3 and a fixed high level of noise ($coh = 0.06$)

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

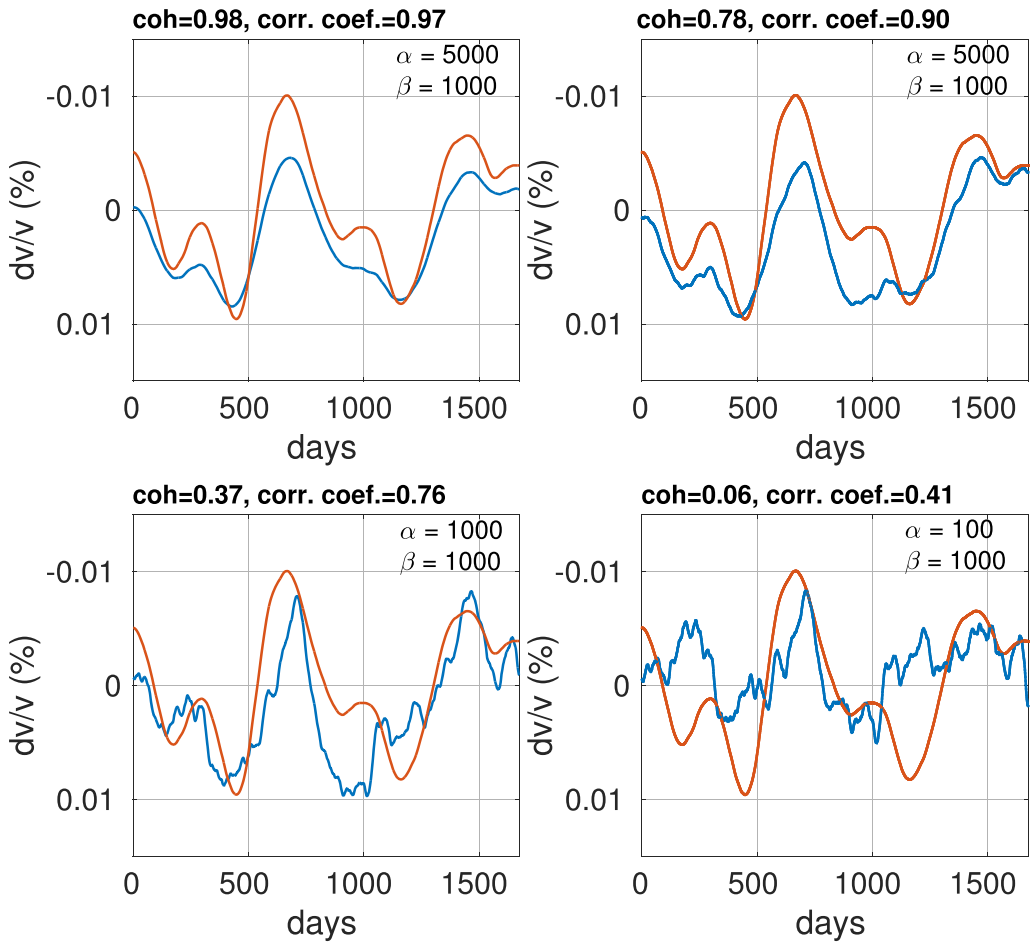


Figure 3. Reconstructed velocity change time-series (blue curves) versus the expected velocity changes (red curve) for different coherence levels. Coherence levels and correlation coefficients between both curves on top of each figure.

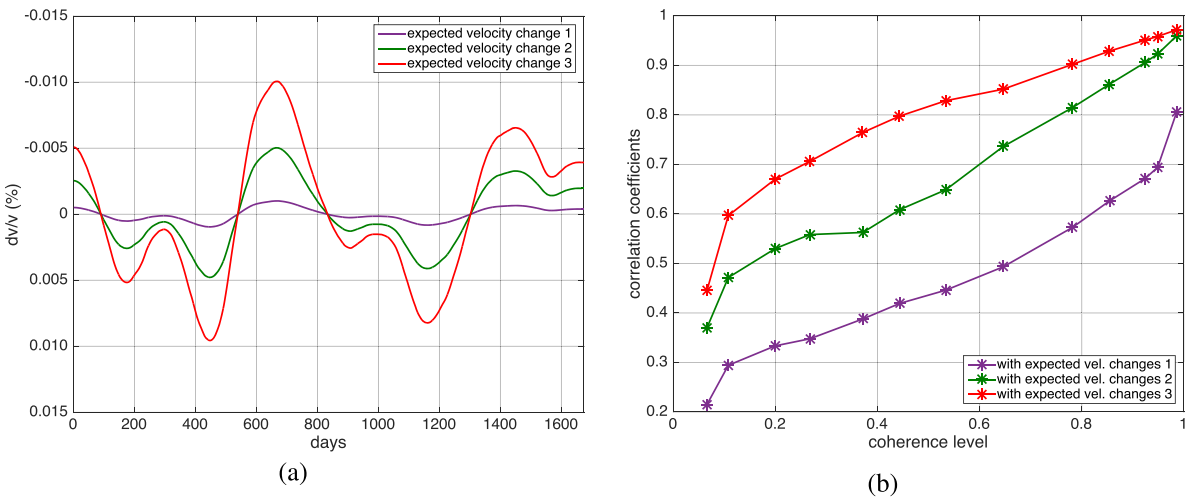


Figure 4. (a) Expected velocity change curves used in the long-term periodic-type fulctuation test. (b) Convergence curves of the coherence levels and the correlation coefficients between the reconstructed velocity change time-series and the different expected velocity changes.

to simulate up to 50 different synthetic station pairs. After obtaining the 50 reconstructed velocity change curves, we average them to study the improvement. N_{sta} is the number of averaged curves of reconstructed velocity changes.

Even with such a low *coh* considered (*coh* = 0.06), we see the improvement when averaging over different station pairs (Fig. 5a): from a correlation of 0.22 for $N_{\text{sta}} = 1$ to 0.87 for $N_{\text{sta}} = 50$ (correlation increased by a factor of 3.9), whereas for $N_{\text{sta}} = 20$ we already

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

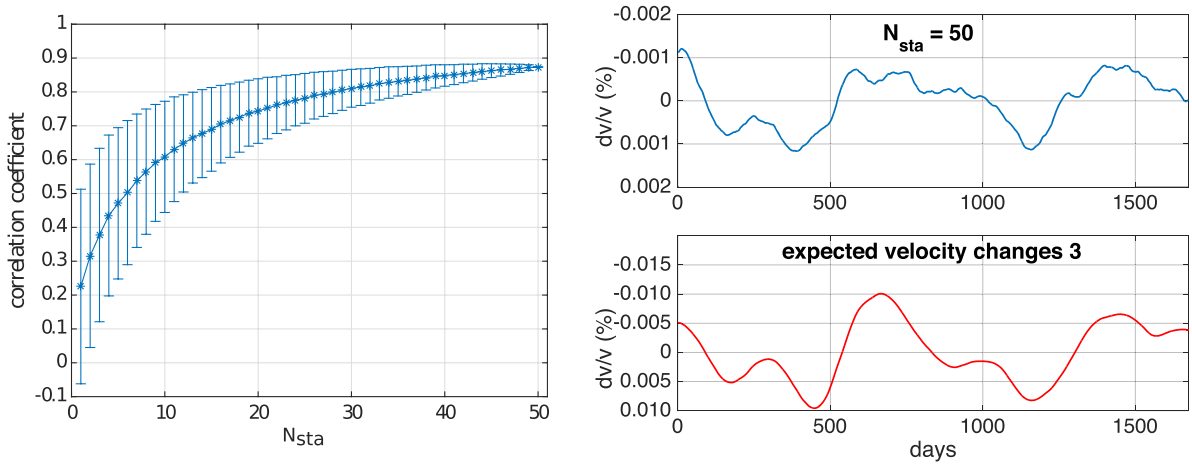


Figure 5. For a coherence level = 0.06 and the expected velocity change curve 3: (a) correlation coefficients between the reconstructed velocity change time-series and the expected velocity changes 3 as a function of the number of averaged curves of reconstructed velocity changes, N_{sta}. Associated standard deviations in blue bars. (b) Reconstructed velocity change time-series for $N_{\text{sta}} = 50$ (blue) and the expected velocity change curve 3 (red).

reach a correlation coefficient of 0.7. In general, it is thus recommended to average seismic velocity changes over at least 20 station pairs when the noise cross-correlations are so unstable. Although the correlation coefficient is 0.22 for $N_{\text{sta}} = 1$ in Fig. 5(a), we see a higher correlation coefficient, 0.41, for the same coherence level, *coh*= 0.06. This is because we picked one of the best examples to show. The standard deviation bars for each N_{sta} of Fig. 5(a) are the variations associated to the average of different combinations of station pairs. The maximum number of combinations used is 10 000. The same is applied for the following tests. Fig.5(b) shows the averaged curve for $N_{\text{sta}} = 50$ and the expected velocity change curve 3. The amplitude of the reconstructed time-series of velocity changes for $N_{\text{sta}} = 50$ is one magnitude order smaller than the expected velocity changes, probably due to an edge effect of the time-series.

3.2 Velocity drop test

To test the reconstruction of an abrupt, rapid change of velocity, similar to the effect of an earthquake (e.g. Brenguier *et al.* 2008a), we add a Heaviside step function with a velocity change of 0.05 % to the previous expected velocity change curve 3 (Fig.6, red curve), referred as the drop curve.

As we are interested in recovering the drop, we use another coefficient to study the similarity between the reconstructed time-series and the drop curve instead of using the Pearson correlation coefficient. To estimate the quality of the reconstructed drop, we measure the difference between the mean velocity changes after and before the drop:

$$\text{diff} = \left(\overline{\frac{\mathrm{d}\nu}{\nu}}\right)_{\text{after drop}} - \left(\overline{\frac{\mathrm{d}\nu}{\nu}}\right)_{\text{before drop}}. \tag{8}$$

We compute diff for both the reconstructed velocity change curve and the expected drop curve. We then estimate the quality of the reconstructed drop with the ratio:

$$Q_{\text{drop}} = \left|\frac{\text{diff}_{\text{reconstructed velocity change curve}}}{\text{diff}_{\text{drop curve}}}\right|. \tag{9}$$

Here Q_{drop} is 1 when perfectly reconstructed and <1 otherwise. In this test, we invert for time-series of velocity changes using a low β to obtain STV, $\beta = 5$, and we avoid a smoothing factor ($\alpha \approx 0$), as we want to study just the effect of the velocity drop.

As the level of noise increases (*coh* decreases), the drop in the reconstructed time-series of velocity changes becomes smaller until it almost disappears (when the coherence level is nearly zero; Fig. 6). We observe the convergence of Q_{drop} for different coherence values of the synthetic cross-correlations in Fig. 7.

We also study the improvement of averaging the reconstructed velocity change curves over different station pairs. For a fixed coherence level of 0.37, we study the convergence of the retrieved drop by increasing N_{sta} (Fig. 8a). Interestingly, by averaging more reconstructed velocity changes, we smooth the sharp STV while the recovered drop remains the same. We also estimate the increasing signal-to-noise ratio (SNR) associated with the larger number of averaged synthetic functions, N_{sta}, as

$$\text{SNR} = \frac{\text{level of recovered drop}}{\text{rms(averaged}\,\frac{\mathrm{d}\nu}{\nu}\,\text{curve)}}, \tag{10}$$

with rms(averaged $\frac{\mathrm{d}\nu}{\nu}$ curve) being the root mean square of the velocity change mean curve of each N_{sta} (Fig. 8a).

A way to increase the coherence between CCFs and, therefore, to improve the temporal resolution of the velocity change measurements, is the use of denoising methods such as the Curvelet filtering (Stehly *et al.* 2015) or the Wiener filtering. We applied a FIR Wiener filter to our CCFs without obtaining a great improvement in the reconstructed velocity changes, probably because this technique only has an effect on the amplitude of the frequency spectrum whereas the method presented in this paper only uses the phase of the signal.

For a coherence level of 0.37 and $N_{\text{sta}} = 50$, we obtain a Q_{drop} of 0.6 and a SNR of 38 (Figs 8a and b). Again, it is interesting to note that, for highly unstable correlations (e.g. coh = 0.37), averaging over different station pairs will not improve the value of the level of the velocity drop that will remain underestimated. Averaging over different receiver pairs will, however, improve the SNR of the recovered velocity changes and thus, will allow a better estimate of the timing of the velocity drop.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

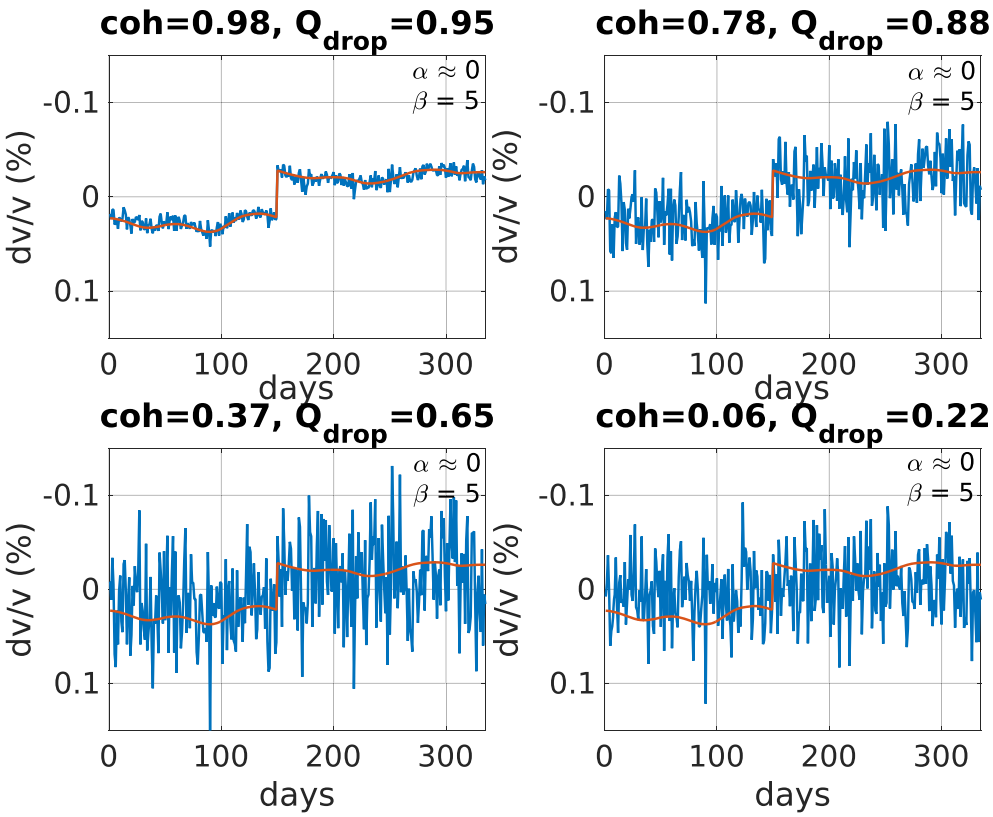


Figure 6. Reconstructed velocity change time-series (blue curves) versus the drop curve (red curve) for different coherence levels. Coherence levels and Q_{drop} on top of each figure.

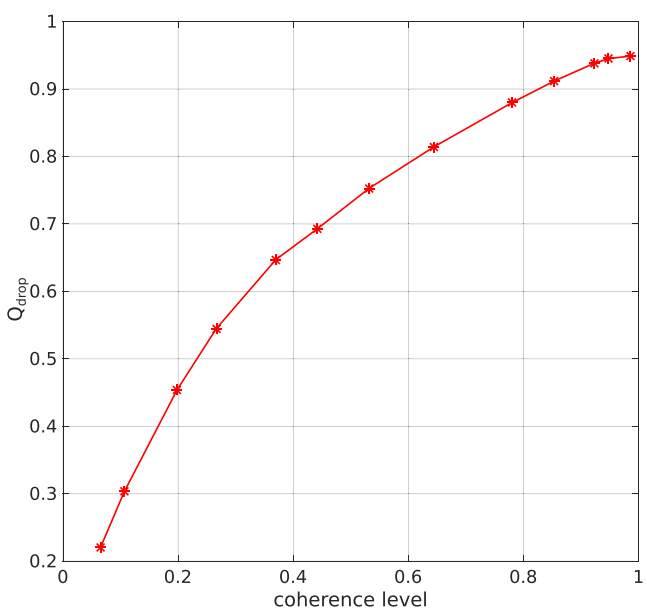


Figure 7. Convergence curve between the coherence levels and Q_{drop}.

3.3 Transient noise perturbation test

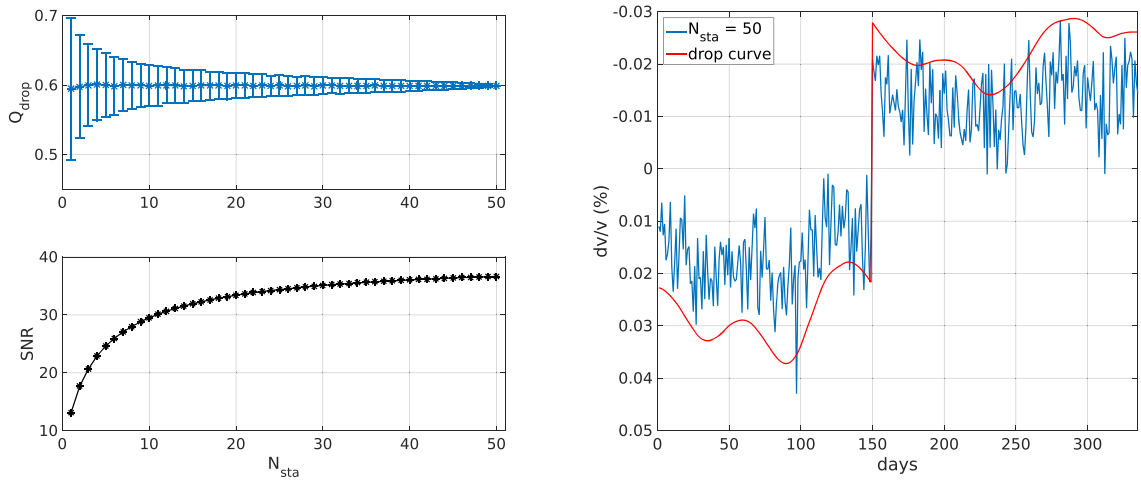
In this test we study the effect of an episodic strong change in the noise-correlation shape induced by a pronounced variation of a noise source, for example, a passing storm or an episodic volcanic tremor. This last situation has been described by Ballmer *et al.* (2013) and Droznin *et al.* (2015) in case of noise-correlations affected by the occurrence of low-frequency volcanic tremor. We herewith test the ability of our method to recover robust short- to medium-term fluctuations.

To compute the synthetic stretched CCFs, we consider two real normalized CCFs, one corresponding to a non-tremor period (C_1) and the other to a tremor period (C_2). Basically, we consider C_1 as the true GF and C_2 as a pure tremor-related bias. With both, we compute two new averaged correlations: $C_3 = 0.8 \times C_1 + 0.2 \times C_2$ and $C_4 = 0.8 \times C_2 + 0.2 \times C_1$, corresponding to a calm period (C_3) and to a tremor period (C_4), respectively. We concatenate N_1 correlations C_3, N_2 correlations C_4 and again N_1 correlations C_3, N_1 and N_2 being random numbers of daily CCFs. Then, the same way as previous tests, we stretch the CCFs and add different levels of random noise to these correlations.

Fig. 9(a) is an example of synthetic stretched CCFs with a certain level of random noise ($coh = 0.54$). We see clearly the differences in the shape of CCFs corresponding to the calm periods, C_3 (from day 1 to 30 and from day 90 to 120 in Fig. 9a), and to the tremor period, C_4 (from day 30 to 90 in Fig. 9a). Fig. 9(b) is the associated correlation coefficient matrix of Fig. 9(a) that represents all Pearson correlation coefficients between all pairs of CCFs. We observe the lower correlation between CCFs of the tremor period comparing with the calm periods.

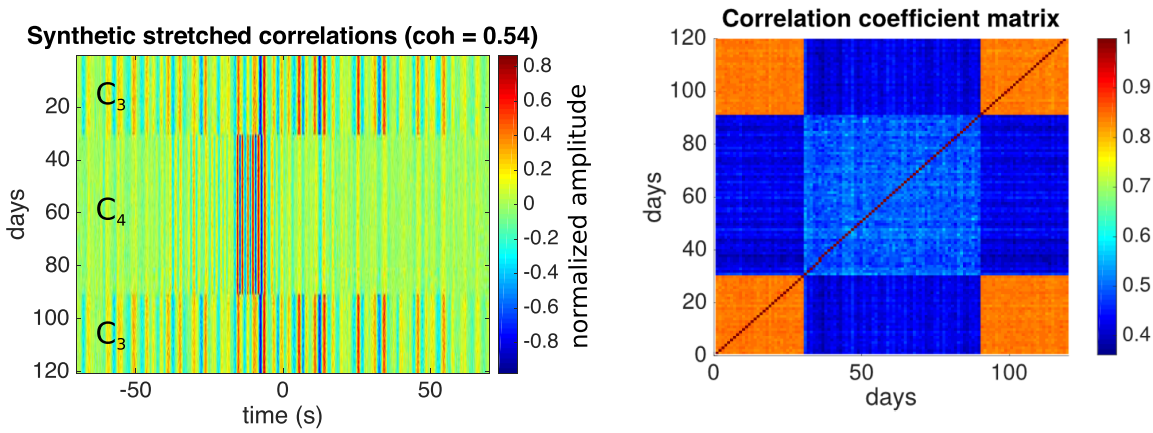
Fig. 10 shows some examples of the resulting reconstructed time-series of velocity changes for the maximum coherence level of 0.85 and for some lower ones, where the coherence level decreases due to the increased level of random noise in the synthetic CCFs. We also plot the expected velocity curve for comparison. As we are interested in evaluating the sudden velocity drop and sudden recovery in the reconstructed time-series of velocity variations, we consider

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018



(a) (b)

Figure 8. For a coherence level = 0.37: (a) Q_{drop} (blue curve with associated standard deviations in blue bars) and signal-to-noise ratio (SNR; black curve) as a function of the number of synthetic averaged functions (N_{sta}). (b) Reconstructed velocity change time-series for $N_{\text{sta}} = 50$ and drop curve.



(a) (b)

Figure 9. (a) Example of normalized synthetic stretched CCFs with a random level of noise (shown for a coherence level between CCFs of 0.54). (b) Correlation coefficient matrix associated to the doublets. C_3 and C_4 refer to calm and tremor periods, respectively.

$\beta = 5$ and $\alpha \approx 0$, as in the previous test. In cases of high *coh*, we observe a double velocity drop in the recovered synthetic velocity change curves (between days 30 and 90) due to the concatenation of different synthetic CCFs, that is, the first N_1 days (calm period), the next N_2 days (tremor period) and the last N_1 days (calm period again) (Fig. 9). We explain this double velocity drop by looking at the correlation coefficient matrix (Fig. 9b). As the correlation coefficients of the CCFs between the calm and the tremor period are very low (Fig. 9b), our method treats these data segments separately and, thus, generates this baseline difference between the two periods. Therefore, these artificial velocity drops are artefacts from our method. The double velocity drop observed in the reconstructed time-series is hidden when the level of noise increases.

Even more interesting, when we increase the number of inverted synthetic time-series of velocity changes for a low *coh* to study the improvement associated with averaging over different station pairs (Fig. 11a), we see clearly the increased similarity between the inverted curves and the expected one (Fig. 11b). This is because only C_1, the medium, is coherent and the noise source perturbation is not seen the same way by all receiver pairs. This means that for some station pairs, the double velocity drop induced by the tremor has, sometime, opposite sign which, simply, cancels out while summing over different receiver pairs. We show, as well, the improvement of the generalized formulation compared to the standard approach when averaging over different station pairs (Figs 11c and d). For the same synthetic data and coherence level, the double velocity drop does not cancel out (Fig. 11d).

3.4 Summary

To conclude, the synthetic tests have shown the behaviour of three different realistic scenarios that have their imprints in the noise correlations and, therefore, we have to handle them in the reconstructed velocity changes. The scenarios are (1) long-term periodic-type fluctuations produced by a seasonal-type trend, (2) sudden velocity drops as effects of sudden changes in the structure, such as earthquakes or volcanic eruptions and (3) transient noise perturbations

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

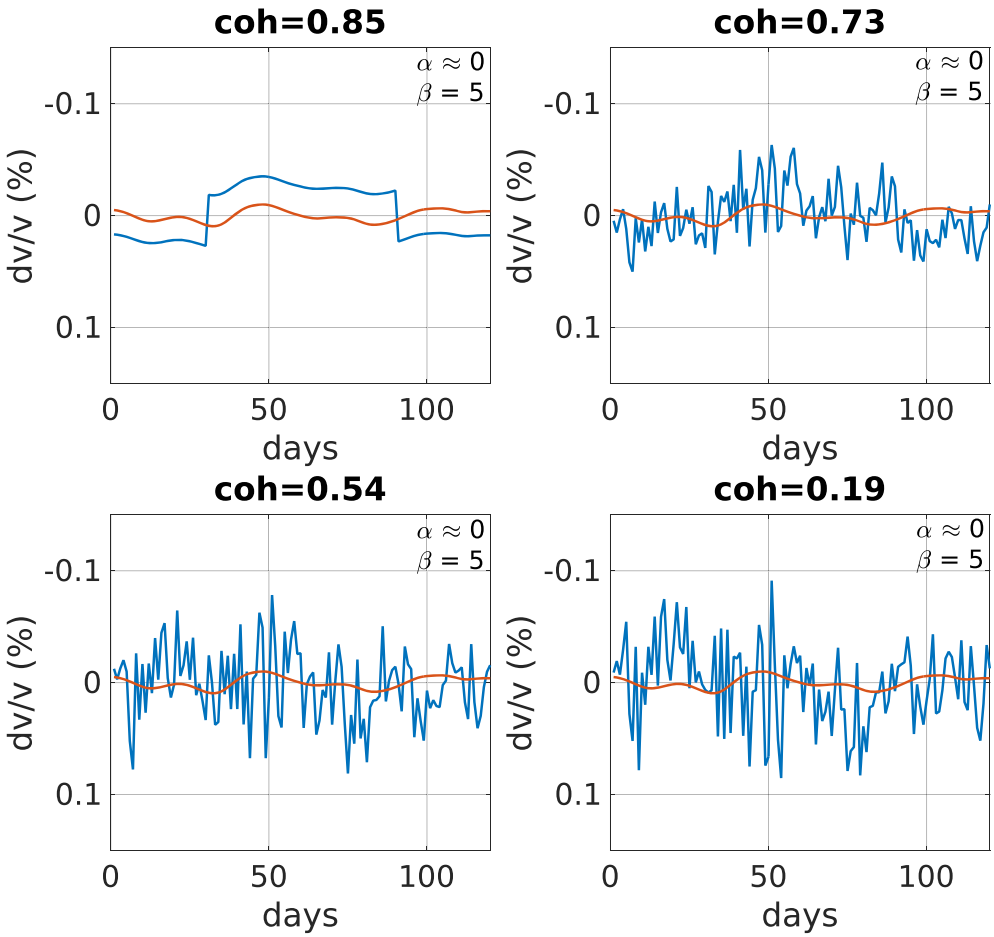


Figure 10. Synthetic velocity change time-series (blue curves) versus the expected velocity changes (red curve) for different coherence levels. Coherence levels on top of each figure.

due to the effect of a transient local source emission, such as a volcanic tremor. In general, we choose the inversion parameters with respect to the type of velocity changes we are looking for, STV or LTV, respectively. In Table 1 we summarize the parameters used in the synthetics. These values respond to our particular tests and should be changed according to the characteristics of the data set, but the magnitude orders of Table 1 can be used as a guidance. In general, it is easier to choose first the value of β, then adjust α, which works as a smoothing factor. We choose lower β values for STV, as we have done in the velocity drop and transient noise perturbation tests, than for LTV (long-term periodic-type fluctuation test case). The parameter α is directly proportional to the *coh* of the reconstructed time-series of velocity changes (Fig. 3 and Table 1): higher values of α ($\alpha > 1000$) are recommended with high *coh*, as the higher the α value, the lower the amplitude and smoother the final time-series of velocity changes. In the case of the velocity drop and the transient noise perturbation tests, we avoided the smoothing ($\alpha \approx 0$), in order to study just the effect of the sudden velocity drops (and recoveries in case of the last test), but, to process real data to retrieve STV, we can use medium to low values of α.

We have also investigated the effect of the amplitude of the expected velocity change curves: time-series of velocity changes are better reconstructed when the amplitudes of the expected changes are higher (Figs3 and 4). In case of medium to low *coh*, the difference in retrieving the amplitudes of the different expected changes is considerable (Fig. 4a). According to these results, in case of real data with a low *coh*, we can count on reconstructing STV and LTV with amplitudes of the same magnitude order than the peak-to-peak amplitude of the reconstructed velocity changes but not much smaller. We have also explored the improvement of averaging the reconstructed time-series of velocity changes for different station pairs even with very low *coh* (Figs 5, 8 and 11). However, although there is a substantial increase in the correlation coefficient between the reconstructed velocity change time-series and the expected velocity changes or in the SNR of the reconstructed velocity changes (for the velocity drop test), the amplitudes of the final velocity change time-series (both STV and LTV) are underestimated. This is a drawback of our method that needs to be further studied. One reason might be that the covariance matrix associated to the data vector is, actually, damping the data during the inversion (eq. 6). Leaving the amplitude aside, these synthetic tests give us an estimation of the number of station pairs of real data that we need to average in order to retrieve a proper velocity change curve when *coh* is low.

Finally, in the situation of strong noise perturbations, we have observed artificial velocity drops produced by our method, visible only when the *coh* of the CCFs is high. There are two approaches to handle this scenario: in case the coherence level between the noise CCFs is high, it might be worth correcting for the artificial baseline difference after the inversion to retrieve proper velocity changes. Otherwise, when *coh* is low, the only way to retrieve a

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

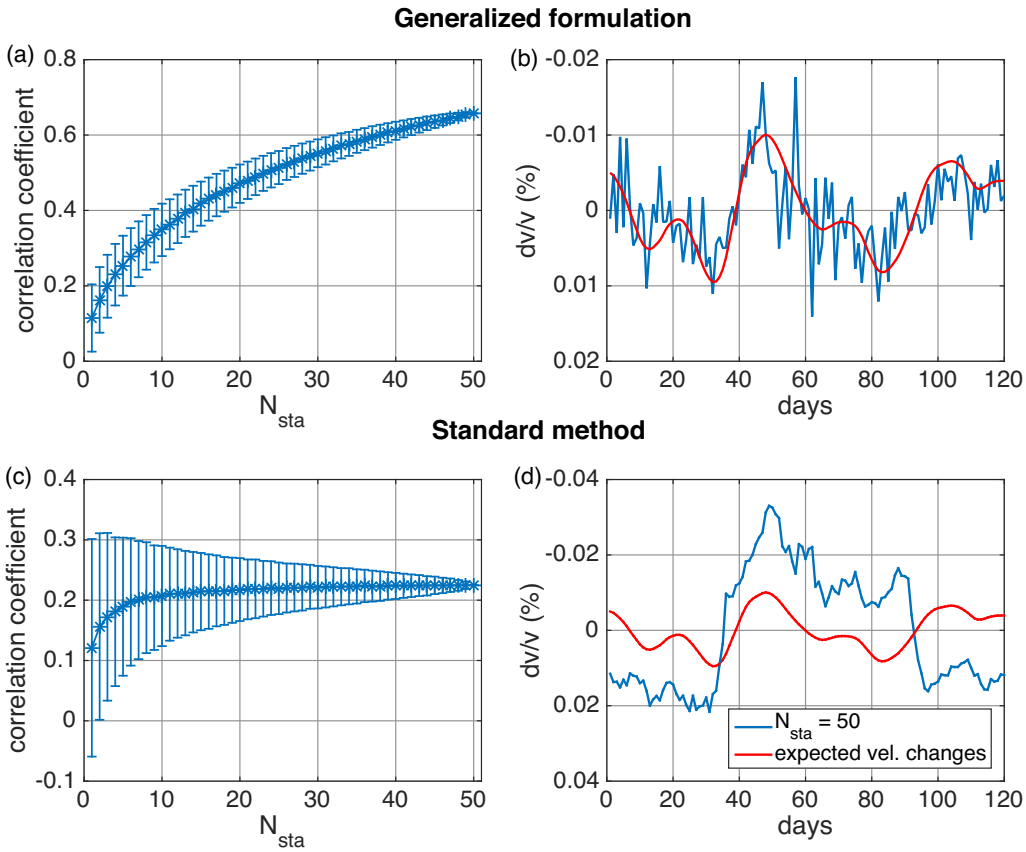


Figure 11. For a coherence level = 0.54: correlation coefficients between synthetic velocity change curves and the expected velocity change curve as a function of the number of synthetic averaged functions (N_{sta}) for (a) the generalized formulation and for (c) the standard method. Reconstructed velocity change time-series for $N_{sta} = 50$ (blue curve) compared to the expected velocity changes (red curve) for (b) the generalized formulation and for (d) the standard method. In the standard method we have used a moving window of 5 d.

Table 1. Inversion parameters used in the synthetic tests.

Synthetic tests	α	β
Long-term periodic-type fluctuation test	high *coh* $\not\Rightarrow$ high α; low *coh* $\not\Rightarrow$ low α	1000
Velocity drop test	≈ 0	5
Transient noise perturbation test	≈ 0	5

proper velocity change curve is to average over sufficient station pairs.

4 APPLICATION TO REAL DATA

With synthetic tests, we have established a general framework to identify and interpret long-term periodic-type velocity changes from seasonal-type trends, rapid velocity drops (due to transient changes) and sudden velocity drops and recoveries as an effect of transient and sudden local source emissions. We have analysed the effect of the regularization parameters and the averaging over station pairs for the three different cases. Now, we apply the method to a complex data set of noise cross-correlations at Klyuchevskoy volcanic group (Kamchatka), hampered by loss of data and the presence of highly non-stationary seismic tremors.

4.1 Klyuchevskoy volcanic group

The Klyuchevskoy volcanic group (KVG), located in Kamchatka, is one of the most active clusters of subduction-zone volcanoes in the world, where the annual rate of explosive eruptions is three to five (Schneider *et al.* 2000). The KVG has an averaged extension of 70 km and 13 stratovolcanoes. It includes active volcanoes such as Klyuchevskoy, Krestovsky, Ushkovsky, Bezymianny and Tolbachik. The Klyuchevskoy volcano, the most outstanding volcano which is 4750 m high, is associated with the emission of basaltic and basaltic–andesitic lavas and it has a mean eruptive rate of 1 m^3 s^{-1} over the last 10 kyr (Fedotov *et al.* 1987). Two other active volcanoes, Shiveluch and Kizimen, are located only 60 km north and south of KVG, respectively. This cluster of volcanoes is located off the edge of a tectonic junction: the Pacific Plate is subducting down the Aleutian Trench and also moving under the Okhotsk Plate. The high volcanic activity is also a consequence of the Hawaii–Emperor Seamount chain that terminates in the Kuril–Kamchatka Trench. Geodynamic models that have been proposed to explain the exceptional activity of the KVG include fluid being released from the thick, highly hydrated Hawaii–Emperor Seamount crust (Dorendorf *et al.* 2000), mantle flow around the corner of the Pacific plate (Yogodzinski *et al.* 2001) and recent detachment of a portion of the subducting slab (Levin *et al.* 2002; Levin *et al.* 2005).

The volcanic activity of the KVG leads to the generation of strong volcanic tremors (Gordeev *et al.* 1990) with sources located very close to the surface and at depth near the crust mantle boundary

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

(Shapiro *et al.* 2017a). These tremors spoil the ambient noise cross-correlations. We use the information of Droznin *et al.* (2015) and Soubestre *et al.* (2018) about detection of these signals and about location of their sources in Klyuchevskoy volcanic group to recover seismic velocity fluctuations in this region, as we use the same data set of noise cross-correlations.

The particular tectonic settings surrounding KVG and its strong eruptions with high seismic activity (e.g. Senyukov *et al.* 2009; Zharinov & Demyanchuk 2009; Ozerov *et al.* 2013) enable many seismic tomographic surveys (e.g. Slavina *et al.* 2012; Koulakov *et al.* 2013; Lees *et al.* 2013) and receiver function analysis to study the internal structure of the KVG (Nikulin *et al.* 2010).

Tomographic studies on the KVG reveal an extremely high Vp/Vs ratio (up to 2.2), below 25 km depth. This feature can act as a channel that brings deep mantle materials to the bottom of the crust. It is also responsible for all volcanic activity in the KVG (Koulakov *et al.* 2013).

Our study covers January 2009 to July 2013 when both the Klyuchevskoy and the Tolbachik volcanoes erupted. Both volcanoes are characterized by effusive eruptions with basaltic to basal-andesitic lavas (e.g. Churikova *et al.* 2013, 2015; Belousov *et al.* 2015). Two eruptions took place on the Klyuchevskoy volcano. The first one started in June 2008 and the volcanic activity ceased at the end of January 2009. The second Klyuchevskoy eruption goes from July 2009 to 2010 December 7. Spatterings of hot magma started on 2009 August 2. The summit eruption activity was characterized by weak ash emissions (less than 300 m of height), although in 2010 the ash emissions were stronger (9 km of height). The eruption intensity decreased at the end of 2010. All the recorded Klyuchevskoy summit eruptions are characterized by a gradual growth of activity (Senyukov 2013). A detailed analysis of records of volcanic tremors has been used by Soubestre *et al.* (2018) to identify two different stages of the 2009–2010 Klyuchevskoy eruption with the stronger second stage starting approximately in June 2010.

The last eruption is the fissure eruption of the Tolbachik volcano (2012–2013). The 2012–2013 Tolbachik eruption started on 2012 November 27 corresponding to an eruptive tremor (Fig. 16) due to a first magma migration (Caudron *et al.* 2015). The Tolbachik regional zone of cinder cones is 900 km^2 in size and 70 km long. Before last eruption (2012–2013), historical eruptions in Tolbachik zone occurred in 1740, 1941 and 1975–1976 (Gordeev *et al.* 2013).

The three eruptions that take place during our study are characterized by the generation of seismic tremors (Gordeev *et al.* 1990; Droznin *et al.* 2015; Shapiro *et al.* 2017a).

4.2 Data

We use continuous records from a total of 18 three-component seismic stations (Fig. 12) of the seismic network deployed by the Kamchatka Branch of the Geophysical Service (KBGS) of the Russian Academy of Sciences (Chebrov *et al.* 2013). Each station component has a CM-3 short-period sensor. We analyse data recorded continuously between 2009 January 1 and 2013 July 7.

Records are digitized at 128 samples per second and downsampled to 8 samples per second. Cross-correlations are calculated in 24-hr long segments. We pre-process the continuous records following the method described by Bensen *et al.* (2007). We choose a spectral band between 0.08 and 0.7 Hz because, after 0.7 Hz, the correlations are too much affected by volcanic tremor correlation signals. After whitening, 1-bit normalization suppresses high-amplitude data, such as earthquake signals, and emphasizes low-amplitude data, such as ambient seismic vibrations. Even after reducing persistent signals from localized sources with pre-processing, volcanic tremors still act as potential biasing signals perturbing the reconstructed GF. Then, we compute daily CCFs for all possible station pairs. We work with coda waves of daily CCFs between the vertical-component records of the station pairs (Rivet *et al.* 2014).

For passive monitoring techniques, both the continuity of the records and the good quality of data are important. For this reason, we do first a quality check of the daily CCFs for each possible seismic station pair, 209 pairs in total. We visually inspect all CCFs of each station pair to rank them in different groups according to the quality of the recordings. Taking into account the continuity and regularity over time of the CCFs, where coda waves are clearly distinguished, we consider three quality groups, from best to worst: A, B and C. We can apply our method to the CCFs of the station pairs ranked in groups A and B but not to those of group C.

We work with station pairs ranked in group A, there are 23 in total. Fig. 13 shows an example of daily CCFs computed for a station pair ranked in group A with its associated correlation coefficient matrix. The periods with highest correlation coefficients correspond to the first two-thirds of 2010 and to 2013. While most of the station pairs of the group A are in the vicinity of Klyuchevskoy and Tolbachik volcanoes, three station pairs (from stations BDR, SMK and SRK) are farther away from the rest, in the vicinity of the Shiveluch volcano (Fig. 12). Because of this, in our study we separate these three pairs near Shiveluch from the others. We compute all the doublets for the 23 station pairs with the MWCS analysis.

Correlation coefficient matrices for each station pair ranked with A are in Figs 14 and 15. We can see different patterns in correlation coefficients if we compare the main group of station pairs (Fig. 14) with the northern group (Fig. 15). All pairs show a strong correlation in the second half of 2010 and in 2013, matching with the ongoing Klychevskoy and Tolbachik eruptions [Droznin *et al.* (2015), fig. 5], respectively. Highest correlation values are observed between the stations of the main group (Fig. 14).

Daily averaged levels of tremors are shown in Fig. 16, determined by the KBGS operators. The strongest tremor activities of both volcanoes also match with the highest correlation coefficients between CCFs (Fig. 14), which means that tremors are the main sources.

Before the inversion, we reject the doublets where the associated correlation coefficients (Figs 14 and 15) are smaller than 0.3. Thereby, we ensure the recovered temporal velocity variation curves tend towards zero for days with bad quality recordings.

4.3 Results

We compute the whole relative velocity changes (STV + LTV) for all station pairs of the quality group A and then, we average, independently, the stations near Shiveluch (three station pairs; Fig. 17) and the main group of station pairs (20 pairs; Fig. 18) near Klyuchevskoy because the velocity changes associated with these two volcanoes can be very different. Both STV and LTV are not independent measurements but result directly from inversion, applying different regularization parameters. The choice of the regularization parameters for the inversion is based on the conclusions of our synthetic tests. The parameters used for the whole velocity variations are $\alpha = 100$ and $\beta = 5$ to highlight STV (Figs 17 and 18, black curves).

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

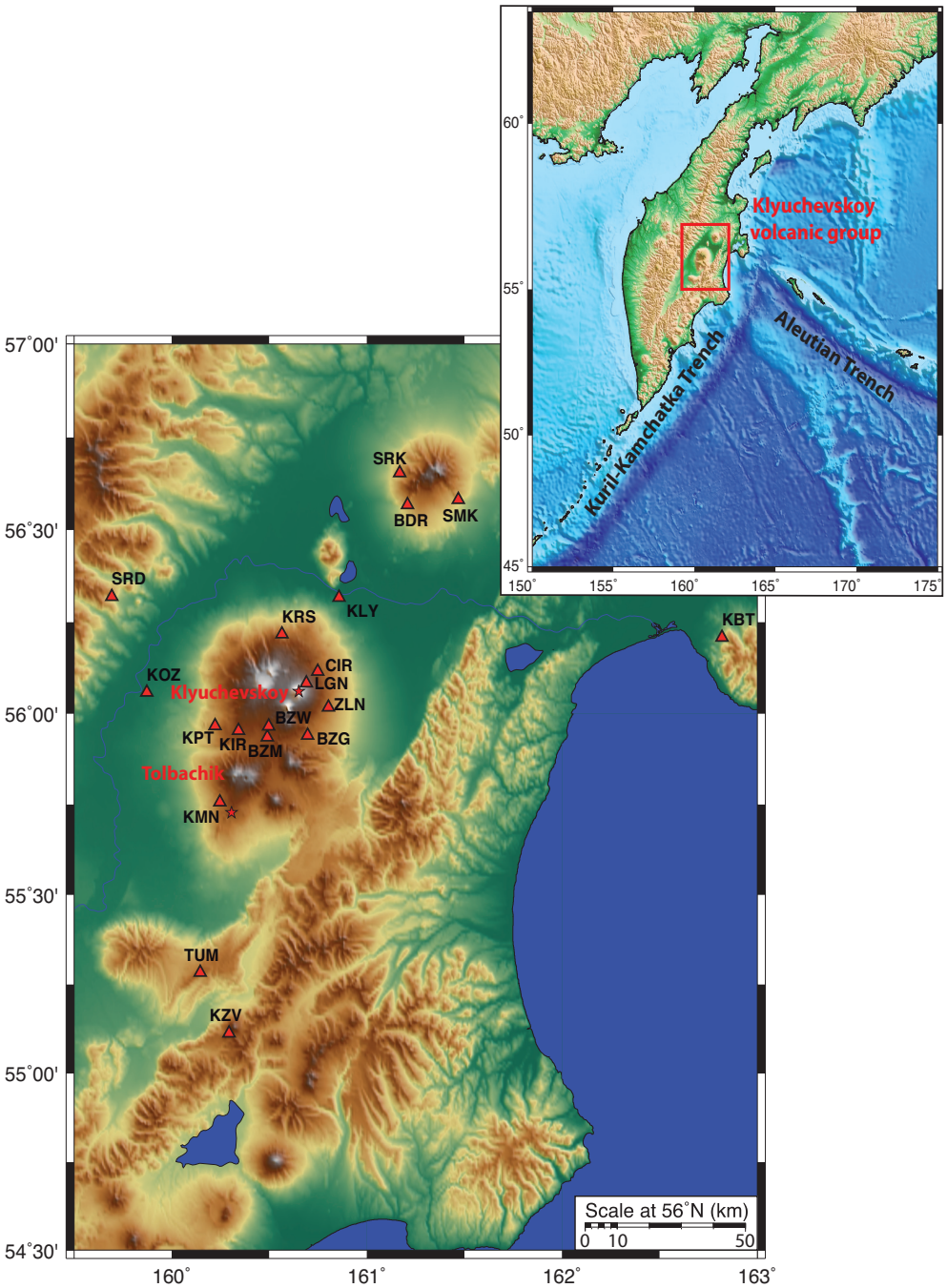


Figure 12. Topographic map of the Klyuchevskoy group of volcanoes in Kamchatka peninsula with positions of seismic stations. Red stars are the eruptive centres of the 2009–2010 Klyuchevskoy and of the 2012–2013 Tolbachik volcanoes.

However, to converge towards the actual relative velocity changes of the medium, we need to retrieve a stable trend due to LTV. Thus, we compute reconstructed velocity change time-series from all considered station pairs with a high β value ($\beta = 1000$) to obtain precise velocity change curves that avoid STV. The smoothing parameter is the same than before, $\alpha = 100$. After obtaining all the individual LTV, we average them all to get the general trend (Figs 17 and 18, red curves).

We compute the data − misfit for each individual time-series of velocity changes retrieved from the different station pairs. Then, we average all the values obtained for the main group of station pairs and for the three station pairs near Shiveluch (data − misfit = 0.074% for both groups).

To establish a relation between the results of the real data and the synthetics, we need to know the amplitude of the retrieved velocity changes and the coherence level of the real CCFs. Concerning the amplitude, the maximum peak-to-peak amplitudes of the retrieved LTV and the whole relative velocity changes, that is, STV + LTV, are about 0.02 % and 0.05%, respectively, for 20 station pairs (Fig.18), which correspond to the magnitude orders of the amplitudes of the expected velocity change curves used in Sections 3.1 [Fig. 4(a), expected velocity curve 3], 3.2 and 3.3. On the other hand, the mean

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

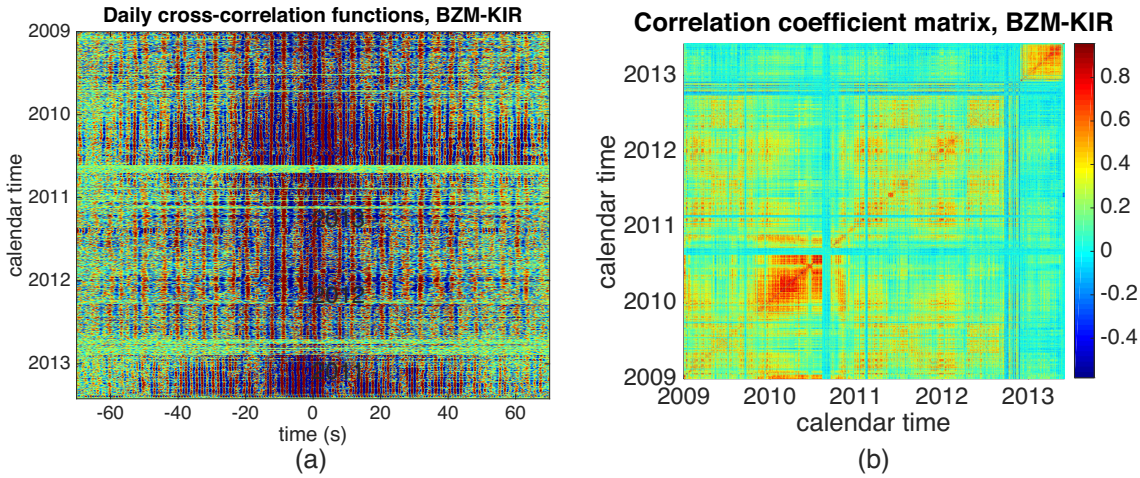


Figure 13. (a) Daily CCF computed from station pair BZM–KIR. (b) Correlation coeffient matrix associated to the doublets of the station pair BZM–KIR.

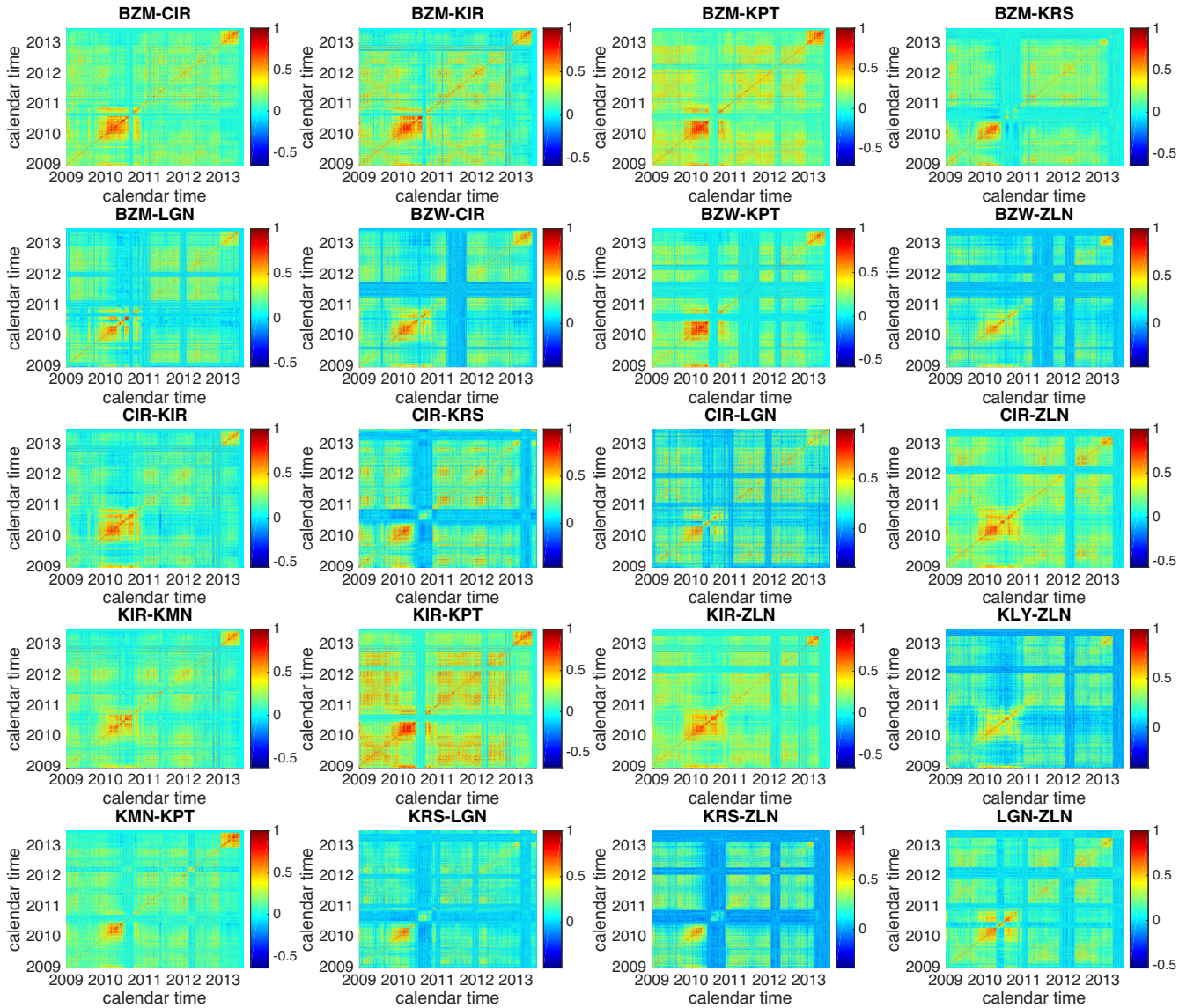


Figure 14. Correlation coefficient matrices between all daily CCFs from January 2009 to July 2013 associated to 20 station pairs of group A located in the vicinity of Klyuchevskoy and Tolbachik volcanoes.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

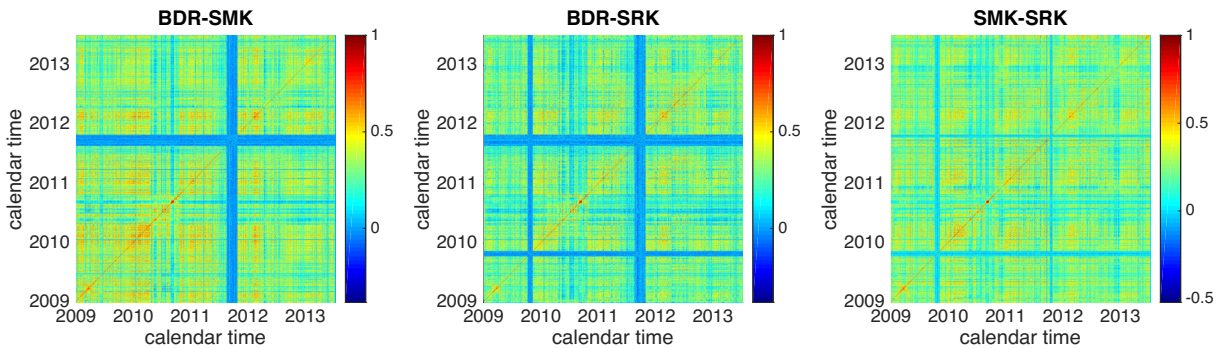


Figure 15. Correlation coefficient matrices between all daily CCF from January 2009 to July 2013 associated to the station pairs of group A located in the vicinity of Shiveluch.

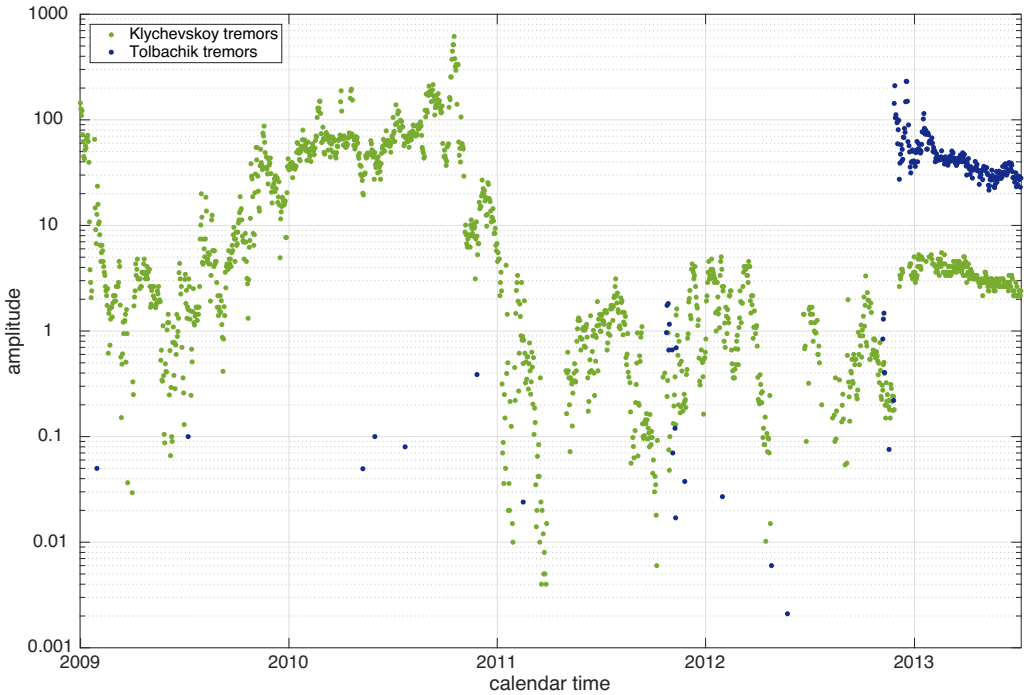


Figure 16. Normalized tremor amplitudes for Klychevskoy (green) and Tolbachik (blue) volcanoes.

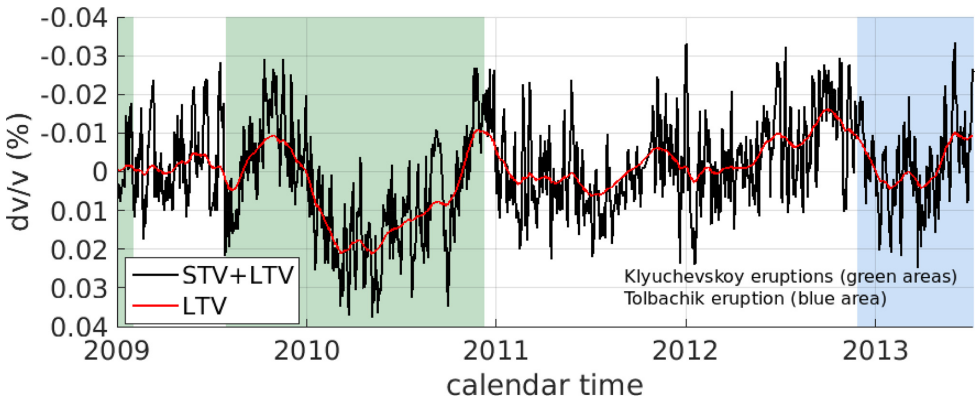


Figure 17. Evolution of relative velocity changes measured from three pairs of stations located near Shiveluch from January 2009 to July 2013. The whole relative velocity changes (STV+LTV in black) and long-term velocity variations (red curve) are overlaid. Klyuchevskoy and Tolbachik eruptive periods are shown with green and blue rectangles, respectively.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

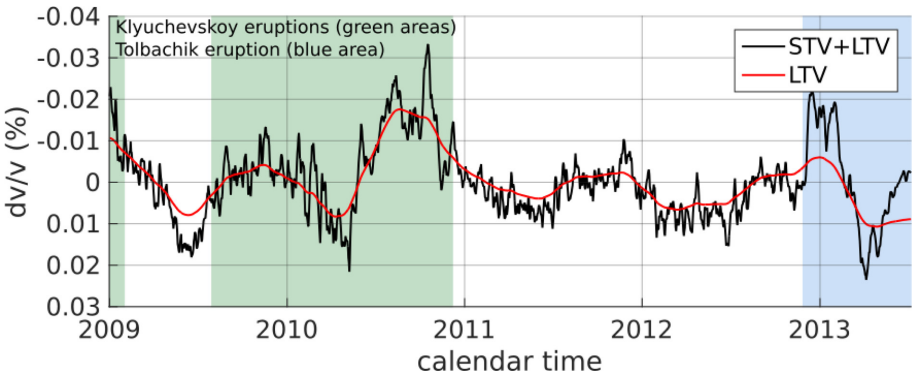


Figure 18. Evolution of relative velocity changes on Klyuchevskoy volcanic group from January 2009 to July 2013 (averaging of time-series of velocity changes over 20 station pairs). The whole relative velocity changes (STV+LTV in black) and long-term velocity variations (red curve) are overlaid. Klyuchevskoy and Tolbachik eruptive periods are shown with green and blue rectangles, respectively.

coherence level of the real CCFs, after rejecting correlation coefficients smaller than 0.3 (Figs 14 and 15), is $coh = 0.41$ for both groups of station pairs. Then, we can compare the similarity of the LTV and STV of the real data with the synthetics: for long-term periodic-type fluctuations the correlation with the expected velocity change curve is 0.77 (Fig. 4b, $coh = 0.41$) and, for short-term fluctuations, $Q_{\text{drop}} = 0.67$ (Fig. 7, $coh = 0.41$). It is important to note that, for STV and LTV, averaging over different pairs keeps these changes underestimated (Figs 5a and 8a), even though the coherence level of the real CCFs, $coh = 0.41$, is higher than those coh values considered in the synthetics. Nevertheless, the SNR increases by a factor of 1.6 when considering three station pairs instead of only one, and up to 2.5 with 20 station pairs, in case of short-term fluctuations (Fig. 8a) and, for long-term periodic-type variations, the correlation with the expected velocity curves of the reconstructed time-series of velocity changes goes from 0.22 ± 0.28, in case of only one station pair considered, to 0.38 ± 0.25, averaging over three station pairs, and to 0.74 ± 0.10 with 20 station pairs (Fig. 5a). Extrapolating to our results with real data ($coh = 0.41$), the correlation is close to 1 when averaging over 20 station pairs. Therefore, we achieve stable LTV and STV with the averaged time-series of velocity changes of the main group of 20 station pairs near Klyuchevskoy (Fig. 18) while the final time-series of velocity changes of the three station pairs near Shiveluch is still very noisy (Fig. 17). In this case it would be necessary to average more station pairs in order to obtain cleaner velocity changes. We also improve the ability of our method to recover velocity changes during the occurrence of low-frequency volcanic tremors by averaging different synthetic station pairs (Fig. 11a). Although there are high correlations between daily CCFs when strong tremor activities take place (around 0.8 during 2010 and 2013 periods in Figs 14 and 16), the high instability of correlations keeps low the mean coherence level of the final inversion ($coh = 0.41$). Under these circumstances, we need to average over enough station pairs. By averaging over 20 station pairs, the correlation of the reconstructed time-series of velocity changes with the expected velocity curve increases by a factor between 2.2 and 17.3, with regard to a single station pair (Fig. 11a). However, we would retrieve more proper short- to medium-term velocity changes due to episodic volcanic tremors by averaging over more than 40 station pairs, to interpret these velocity drops and retrievals without ambiguity (Fig. 11a).

4.4 Interpretation of the results

The seismic velocity variations measured near Shiveluch (Fig. 15) are difficult to interpret because this measurement was done only with three station pairs and is, therefore, very noisy. Besides, the measurements made with 20 station pairs surrounding the Klyuchevskoy group of volcanoes show velocity variations that can be interpreted in relationship of eruptive history of the two most active volcanoes of this group: Klyuchevskoy and Tolbachik. The whole velocity variations (STV+LTV) are controlled by the combination of two main mechanisms: (1) the variations of the media mechanical properties caused by the magma motion and pressurization within the volcano plumbing systems and (2) the environmental effects. These two mechanisms cannot be simply separated as STV and LTV computed during the data analysis because the long-duration eruptions of Klyuchevskoy and Tolbachik have their imprints on both STV and LTV.

The environmental contribution to the seismic velocity variations is expected to be controlled by seasonal changes in temperature, in hydrological loads and in snow cover. These seasonal effects are particularly strong in Kamchatka and, therefore, we decided to estimate and remove them from the whole time-series, expecting that the remaining velocity variations mainly reflect the dynamics of the volcano plumbing system. To estimate the average long-term seasonal component from the velocity variation time-series shown in Fig. 18, we computed median $\frac{dv}{v}$ values for every Julian day. Then, the obtained 1-yr periodic function has been smoothed in a 3-month long moving window. The resulted seasonal variations are shown with a thick grey line in Fig. 19(a). The seasonality is very clear with a very pronounced velocity increase during winter (between end of December and end of April) and a pronounced velocity decrease during summer (between end of May and end of August).

After removing this seasonal trend, the velocity variations exhibit three significant periods with decrease over 0.01 % (Fig19b). The first of this velocity drops corresponds to the end of the 2008–2009 Klyuchevskoy eruption. The second drop starts at the end of May 2010 and terminates simultaneously with the 2009–2010 Klyuchevskoy eruption. The third velocity decrease starts, approximately, simultaneously with the 2012–2013 Tolbachik eruption. Therefore, all detected decreases in seismic velocity are observed during eruptions and most likely reflect the inflation-caused dila-

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

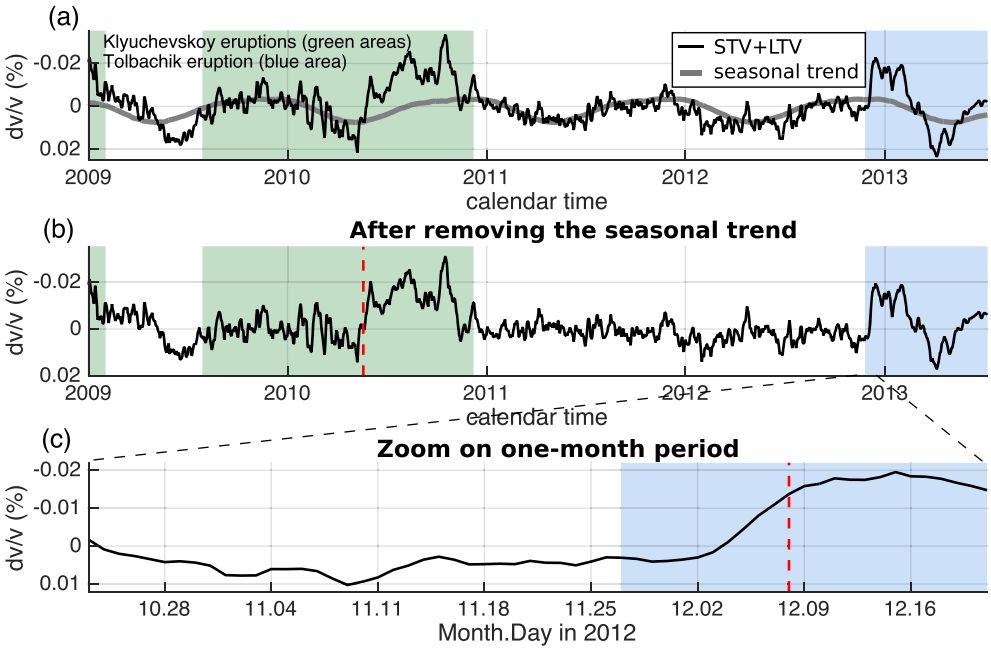


Figure 19. Evolution of relative velocity changes on Klyuchevskoy volcanic group from January 2009 to July 2013 (averaging of time-series of velocity changes over 20 station pairs). (a) The whole relative velocity changes (in black) and average seasonal variations (thick grey curve) are overlaid. (b) Velocity variations after removing the seasonal component. Periods of the Klyuchevskoy and Tolbachik eruptions are shown with green rectangles, respectively. The vertical red dashed line indicates the onset of the second stage of the 2009–2010 Klychevskoy eruption (Soubestre *et al.* 2018). (c) Zoom on one-month period including the beginning of the Tolbachik eruption (blue rectangle). The vertical red dashed line indicates the onset of the main eruption stage.

tion of the shallow crustal layers. Nevertheless, the durations of the observed velocity drops do not exactly coincide with the known periods of eruptive activity. A possible explanation for this is that during the long-duration of Kamchatka volcanoes, the state of the plumbing system exhibits significant changes.

The detailed source analysis of coeruptive tremors by Soubestre *et al.* (2018) has identified two separate stages of activity during the 2009–2010 Klyuchevskoy eruption. The second stage that started approximately in May 2010 (indicated with vertical dashed line in Fig. 19b) was more intensive with magma likely moving closer to the surface. The observed velocity drop coincides in time with the second stage and confirms that the large-scale magma migration occurred between the two stages of eruption.

The level of seismic velocity changes also strongly varied during the 2012–2013 Tolbachik eruption. We observe, in particular, that the onset of the strong velocity drop does not coincide with the beginning of the eruption (Fig. 19c) but rather with the beginning of its main stage, when the outpouring of lava concentrated in a single vent where the main eruptive Naboko cone started to grow (Belousov *et al.* 2015). The later variations in seismic velocities are consistent with changes in tremor sources identified based on correlations of continuous seismic records (Shapiro *et al.* 2017b).

5 CONCLUSIONS

A general framework has been established to provide insights into volcanic unrest using continuous noise-based seismic velocity change observations. Particular care is required to recover temporal velocity variations from CCFs where the noise field recordings are affected by transient tremor signals. The generalized formulation can also be used to study crustal earthquake relaxations and the effects of fluid injections in the subsurface, regardless of the seismic activity. This approach will be useful for improving noise-based seismic monitoring at all scales in cases where noise sources are not stable in time and a localization of the changes is not attempted. To summarize, we classify the principal ideas of this work in three itemized sections.

5.1 Methodology

A general formulation for retrieving velocity changes is applied avoiding the definition of an arbitrary reference CCF. The final time-series of velocity changes is obtained by inversion, using a classical Bayesian linear least-squares formulation. The role of α and β, the regularization parameters, is essential and further studied with the synthetic tests. STV and LTV are sorted after the inversion: for retrieving LTV, a high β is needed. STV is obtained afterwards, subtracting the LTV from the whole relative velocity changes, which are computed with a lower β.

5.2 Synthetic tests

The choice of the inversion parameters depends on the type of velocity changes to retrieve (STV or LTV) and on the *coh* of the CCFs (Table 1). α, the smoothing parameter, and the *coh* level are directly proportional.

Artificial velocity drops might appear in the reconstructed velocity change time-series when the data are affected by strong noise perturbations and in case of strong differences in the *coh* of CCFs between calm and noise-perturbed periods. If the *coh* is low we have to average over sufficient station pairs to retrieve a proper velocity change curve. In general, the number of station pairs needed to average depends on the *coh* level of the CCFs. This method produces an underestimation of the amplitude of the final averaged time-series of velocity changes that might have a relation with the

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

Table 2. Inversion parameters used with real data.

Real data	α	β
Whole relative velocity changes (STV + LTV)	100	5
LTV	100	1000

covariance matrix of the data vector, which could decrease the doublet measurements. In case of low *coh* we can retrieve STV and LTV of a similar magnitude order than the peak-to-peak amplitude of the reconstructed time-series of velocity changes. Despite the drawback, averaging over different station pairs still improves the reconstructed time-series of velocity changes.

5.3 Real data

The method was applied to the Klyuchevskoy volcanic group data set of noise cross-correlations, interfered with strong and localized volcanic tremors and the loss of data. Two groups of station pairs were treated separately because the surrounding volcanoes produce different behaviours in the CCFs: three station pairs located in the vicinity of the Shiveluch volcano and 20 station pairs, the main group of stations, in the KVG area. The parameters used in the inversion are summarized in Table 2 considering the results of the synthetics (Table 1). Stable LTV and STV are obtained for the main group of 20 station pairs. Regarding the three station pairs near Shiveluch, more pairs to average would be necessary in order to have cleaner velocity changes. To interpret velocity drops during the occurrence of volcanic tremors without ambiguity, it would be necessary to average over, at least, twice the number of station pairs used (20 receiver pairs).

Long-term eruptions of Klyuchevskoy and Tolbachik are controlled by the fluctuations of the media mechanical properties and by environmental effects. Therefore, both STV and LTV are affected by the two mechanisms and cannot be separated. Three velocity decrease periods over 0.01 % are observed after removing the seasonal trend due to the environmental effects to the whole velocity variations. The decreases are related with the inflation-caused dilation of the shallow crustal layers. The first decrease occurs at the end of the 2008–2009 Klyuchevskoy eruption, the second corresponds to the second stage of the 2009–2010 Klyuchevskoy eruption (Soubestre *et al.* 2018) and the third coincides with the beginning of the main stage of the 2012–2013 Tolbachik eruption (Belousov *et al.* 2015). The duration of these velocity decrease periods does not exactly coincide with the eruptive activity, probably because of the continuous and significant changes of the plumbing system in the Kamchatka volcanoes.

ACKNOWLEDGEMENTS

All the data used in this study were provided by the Kamchatka Branch of Geophysical Survey of Russian Academy of Sciences (http://www.emsd.ru). This study was supported by REPSOL CO-DOS and REPSOL CO-DOS Phase 2 projects, by the European COST action TIDES (ES1401), by the Russian Ministry of Education and Science (grant number 14.W03.31.0033), by the French projects 'Labex UnivEarth' and Université Sorbonne Paris Cité project 'VolcanoDynamics', by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant agreement number 742335, F-IMAGE) and by a research grant from Science Foundation Ireland (SFI) under grant number 13/RC/2092 and co-funded under the European Regional Development Fund and by iCRAG industry partners. Computations were performed using the IPGP High-Performance Computing infrastructure S-CAPADE (supported by the Île-de-France region via the SEASAME programme, by France-Grille, and by the CNRS MASTODONS programme).

REFERENCES

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, *Geophys. J. Int.,* **194**(1), 512–523.

Belousov, A., Belousova, M., Edwards, B., Volynets, A. & Melnikov, D., 2015. Overview of the precursors and dynamics of the 2012–13 basaltic fissure eruption of Tolbachik Volcano, Kamchatka, Russia, *J. Volc. Geotherm. Res.,* **307,** 22–37.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & 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., 2008a. Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations, *Science,* **321**(5895), 1478–1481.

Brenguier, F., Shapiro, N.M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O. & Nercessian, A., 2008b. Towards forecasting volcanic eruptions using seismic noise, *Nat. Geosci.,* **1**(2), 126–130.

Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N.M., Briand, X., Emoto, K. & Miyake, H., 2014. Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science,* **345,** 80–82.

Campillo, M., Sato, H., Shapiro, N.M. & Van Der Hilst, R.D., 2011. New developments on imaging and monitoring with seismic noise, *C. R. Geosci.,* **343**(8–9), 487–495.

Caudron, C., Taisne, B., Kugaenko, Y. & Saltykov, V., 2015. Magma migration at the onset of the 2012–13 Tolbachik eruption revealed by Seismic Amplitude Ratio Analysis, *J. Volc. Geotherm. Res.,* **307,** 60–67.

Chaves, E.J. & Schwartz, S.Y., 2016. Monitoring transient changes within overpressured regions of subduction zones using ambient seismic noise, *Sci. Adv.,* **2**(1), e1501289, doi:10.1126/sciadv.1501289.

Chebrov, V.N., Droznin, D.V., Kugaenko, Y., Levina, V.I., Senyukov, S.L., Sergeev, V.A., Shevchenko, Y. & Yashchuk, V.V., 2013. The system of detailed seismological observations in Kamchatka in 2011, *J. Volcanol. Seismol.,* **7**(1), 16–36.

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

Churikova, T., Gordeychik, B., Edwards, B.R., Ponomareva, V. & Zelenin, E., 2015. The Tolbachik volcanic massif: a review of the petrology, volcanology and eruption history prior to the 2012–2013 eruption, *J. Volc. Geotherm. Res.,* **307,** 3–21.

Churikova, T.G., Gordeychik, B.N., Ivanov, B.V. & Wörner, G., 2013. Relationship between Kamen volcano and the Klyuchevskaya group of volcanoes (Kamchatka), *J. Volc. Geotherm. Res.,* **263,** 3–21.

Clarke, D., Zaccarelli, L., Shapiro, N.M. & Brenguier, F., 2011. Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise, *Geophys. J. Int.,* **186**(2), 867–882.

Colombi, A., Chaput, J., Brenguier, F., Hillers, G., Roux, P. & Campillo, M., 2014. On the temporal stability of the coda of ambient noise correlations, *C. R. Geosci.,* **346**(11–12), 307–316.

Dorendorf, F., Wiechert, U. & Wörner, G., 2000. Hydrated sub-arc mantle: a source for the Kluchevskoy volcano, Kamchatka/Russia, *Earth planet. Sci. Lett.,* **175**(1), 69–86.

Droznin, D.V., Shapiro, N.M., Droznina, S.Y., Senyukov, S.L., Chebrov, V.N. & Gordeev, E.I., 2015. Detecting and locating volcanic tremors on

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

the Klyuchevskoy group of volcanoes (Kamchatka) based on correlations of continuous seismic records, *Geophys. J. Int.,* **203**(2), 1001–1010.

Duputel, Z., Ferrazzini, V., Brenguier, F., Shapiro, N., Campillo, M. & Nercessian, A., 2009. Real time monitoring of relative velocity changes using ambient seismic noise at the Piton de la Fournaise volcano (La Réunion) from January 2006 to June 2007, *J. Volc. Geotherm. Res.,* **184**(1), 164–173.

Fedotov, S.A., Khrenov, A.P. & Jarinov, N.A., 1987. Klyuchevskoy volcano, its activity in 1932-1986 and possible development, *Volcanol. Seismol.,* **4,** 3–16.

Gordeev, E., Murav'ev, Y., Samoilenko, S., Volynets, A., Mel'nikov, D. & Dvigalo, V.N., 2013. The Tolbachik fissure eruption of 2012-2013: preliminary results, *Dokl. Earth Sci.,* **452**(2), 1046–1050.

Gordeev, E.I., Saltykov, V.A., Sinitsyn, V.I. & Chebrov, V.N., 1990. Temporal and spatial characteristics of volcanic tremor wave fields, *J. Volc. Geotherm. Res.,* **40**(1), 89–101.

Hadziioannou, C., Larose, E., Coutant, O., Roux, P. & Campillo, M., 2009. Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: laboratory experiments., *J. acoust. Soc. Am.,* **125**(6), 3688–3695.

Hillers, G., Husen, S., Obermann, A., Planès, T., Larose, E. & Campillo, M., 2015. Noise-based monitoring and imaging of aseismic transient deformation induced by the 2006 Basel reservoir stimulation, *Geophysics,* **80**(4), KS51–KS68.

Hobiger, M., Wegler, U., Shiomi, K. & Nakahara, H., 2012. Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan, *J. geophys. Res.,* **117**(B9), doi:10.1029/2012JB009402.

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, *J. Volc. Geotherm. Res.,* **263,** 75–91.

Larose, E., *et al.*., 2006. Correlation of random wavefields: an interdisciplinary review, *Geophysics,* **71**(4), SI11–SI21.

Lees, J.M., Symons, N., Chubarova, O., Gorelchik, V. & Ozerov, A., 2007. Tomographic images of Klyuchevskoy volcano P-wave velocity, in *Volcanism and Subduction: The Kamchatka Region; Geophysical Monograph Series,* pp. 293–302, eds Eichelberger , J., E.Gordeev , Kasahara , M., Izbekov , P. & Lees, J., American Geophysical Union, doi:10.1029/172GM21.

Levin, V., Shapiro, N., Park, J. & Ritzwoller, M., 2002. Seismic evidence for catastropic slab loss beneath Kamchatka, *Nature,* **418**(6899), 763–767.

Levin, V., Shapiro, N.M., Park, J. & Ritzwoller, M.H., 2005. Slab portal beneath the western Aleutians, *Geology,* **33**(4), 253–256.

Meier, U., Shapiro, N.M. & Brenguier, F., 2010. Detecting seasonal variations in seismic velocities within Los Angeles basin from correlations of ambient seismic noise, *Geophys. J. Int.,* **181**(2), 985–996.

Mordret, A., Jolly, A., Duputel, Z. & Fournier, N., 2010. Monitoring of phreatic eruptions using interferometry on retrieved cross-correlation function from ambient seismic noise: results from Mt. Ruapehu, New Zealand, *J. Volc. Geotherm. Res.,* **191**(1), 46–59.

Nikulin, A., Levin, V., Shuler, A. & West, M., 2010. Anomalous seismic structure beneath the Klyuchevskoy Group, Kamchatka, *Geophys. Res. Lett.,* **37**(14), doi:10.1029/2010GL043904.

Niu, F., Silver, P.G., Daley, T.M., Cheng, X. & Majer, E.L., 2008. Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site, *Nature,* **454**(7201), 204–208.

Obermann, A., Planes, T., Larose, E. & Campillo, M., 2013. Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise, *J. geophys. Res.,* **118**(12), 6285–6294.

Ozerov, A.Y., Firstov, P.P. & Gavrilov, V.A., 2007. Periodicities in the dynamics of eruptions of Klyuchevskoi volcano, Kamchatka, in *Volcanism and Subduction: The Kamchatka Region; Geophysical Monograph Series,* **172,** pp. 283–291, eds Eichelberger, J., Gordeev, E., Izbekov, P., Kasahara, M. & Lees, J., Americal Geophysical Union, doi:10.1029/172GM20.

Poupinet, G., Ellsworth, W.L. & Frechet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. geophys. Res.,* **89**(B7), 5719–5731.

Rivet, D., Brenguier, F., Clarke, D., Shapiro, N.M. & Peltier, A., 2014. Long-term dynamics of Piton de la Fournaise volcano from 13 years of seismic velocity change measurements and GPS observations, *J. geophys. Res.,* **119**(10), 7654–7666.

Roux, P., Sabra, K.G., Gerstoft, P., Kuperman, W.A. & Fehler, M.C., 2005. P-waves from cross-correlation of seismic noise, *Geophys. Res. Lett.,* **32**(19), 1–4.

Schaff, D.P., 2012. Placing an upper bound on preseismic velocity changes measured by ambient noise monitoring for the 2004 Mw 6.0 Parkfield earthquake (California), *Bull. seism. Soc. Am.,* **102**(4), 1400–1416.

J.Schneider, D., , G.Dean, K., Dehn, J., P.Miller, T. & Kirianov, V.Y., 2000. Monitoring and analyses of volcanic activity using remote sensing data at the Alaska Volcano Observatory: case study for Kamchatka, Russia, December 1997, in *Remote Sensing of Active Volcanism; Geophysical Monograph Series,* **116,** pp. 65–85, eds Mouginis-Mark, P.J., Crisp, J.A. & Fink, J.H., American Geophysical Union, doi:10.1029/GM116p0065 .

Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferemetry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.,* **33**(21), 1–5.

Sens-Schönfelder, C., Pomponi, E. & Peltier, A., 2014. Dynamics of Piton de la Fournaise volcano observed by passive image interferometry with multiple references, *J. Volc. Geotherm. Res.,* **276,** 32–45.

Senyukov, S.L., 2013. Monitoring and prediction of volcanic activity in Kamchatka from seismological data: 2000-2010, *J. Volcanol. Seismol.,* **7**(1), 86–97.

Senyukov, S.L., Droznina, S.Y., Nuzhdina, I.N., Garbuzova, V.T. & Kozhevnikova, T.Y., 2009. Studies in the activity of Klyuchevskoi volcano by remote sensing techniques between January 1, 2001 and July 31, 2005, *J. Volcanol. Seismol.,* **3**(3), 191–199.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31**(7), doi:10.1029/2004GL019491.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307**(5715), 1615–1618.

Shapiro, N.M., Ritzwoller, M. & Bensen, G., 2006. Source location of the 26 sec microseism from cross-correlations of ambient seismic noise, *Geophys. Res. Lett.,* **33**(18), doi:10.1029/2006GL027010.

Shapiro, N.M., Droznin, D., Droznina, S.Y., Senyukov, S., Gusev, A. & Gordeev, E., 2017a. Deep and shallow long-period volcanic seismicity linked by fluid-pressure transfer, *Nat. Geosci.,* **10**(6), 442–445.

Shapiro, N.M., Droznin, D., Droznina, S.Y., Senyukov, S., Gusev, A. & Gordeev, E., 2017b. Variations of properties of seismic tremor during the 2012-2013 eruption of the Tolbachik volcano based on correlations of continuous seismic records in the Tolbachik fissure eruption of 2012-2013: results of investigations, probability model, role for understanding of basalt eruption in subduction zones, *Tolbachik fissure eruption in 2012-2013,* eds Gordeev, E.I. & Dobratsov, N.L., Siberian branch of the Russian Academy of Sciences publishing house, ISBN 978-5-7692-1555-1.

Slavina, L.B., Pivovarova, N.B. & Senyukov, S.L., 2012. Velocity structure of the crust and upper mantle at the northern group of Kamchatka volcanoes (based on the travel time of P-waves from volcanic earthquakes), *Izv. Atmos. Ocean. Phys.,* **48**(7), 696–705.

Soubestre, J., Shapiro, N.M., Seydoux, L., de Rosny, J., Droznin, D., Droznina, S.Y., Senyukov, S.L. & Gordeev, E.I., 2018. Network-based detection and classification of seismo-volcanic tremors: example from the Klyuchevskoy volcanic group in Kamchatka, *J. geophys. Res.,* **123,** 564–582.

Stehly, L., Froment, B., Campillo, M., Liu, Q.Y. & Chen, J.H., 2015. Monitoring seismic wave velocity changes associated with the M_w 7.9 Wenchuan earthquake: increasing the temporal resolution using curvelet filters, *Geophys. J. Int.,* **201**(3), 1939–1949.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation,* SIAM.

Ugalde, A., Villaseñor, A., Gaite, B., Casquero, S., Martí, D., Calahorrano, A., Marzán, I. & Carbonell, R., 2013. Passive seismic monitoring of an experimental CO_2 geological storage site in Hontomín, *Seismol. Res. Lett.,* **84**(1), 75–84.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018

Ugalde, A., Gaite, B. & Villaseñor, A., 2014. Temporal variations of seismic velocity at Paradox Valley, Colorado, using passive image interferometry, *Bull. seism. Soc. Am.,* **104**(3), 1088–1099.

Wapenaar, K. & Fokkema, J., 2006. Green's function representations for seismic interferometry, *Geophysics,* **71**(4), SI33–SI46.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X. & Verdel, A., 2010. Tutorial on seismic interferometry: part 1-basic principles and applications, *Geophysics,* **75**(5), 75A195–75A209.

Wegler, U. & Sens-Schönfelder, C., 2007. Fault zone monitoring with passive image interferometry, *Geophys. J. Int.,* **168**(3), 1029–1033.

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

Zharinov, N.A. & Demyanchuk, Y.V., 2009. The February–July 2007 eruption of the summit crater of Klyuchevskoi Volcano, Kamchatka, *J. Volcanol. Seismol.,* **3**(3), 179–190.

Zhou, R., Huang, L., Rutledge, J.T., Fehler, M., Daley, T.M. & Majer, E.L., 2010. Coda-wave interferometry analysis of time-lapse VSP data for monitoring geological carbon sequestration, *Int. J. Greenhouse Gas Control,* **4**(4), 679–686.

Downloaded from https://academic.oup.com/gji/article-abstract/214/2/1218/4999900
by CNRS user
on 09 July 2018