

EXPRESS LETTER

How much averaging is necessary to cancel out cross-terms in noise correlation studies?

Walter E. Medeiros,^{1,2} Martin Schimmel³ and Aderson F. do Nascimento^{1,2}

¹*Departamento de Geofísica - Universidade Federal do Rio Grande do Norte, Natal, Brazil. E-mail: walter@geofisica.ufrn.br*

²*INCT-GP/CNPq-Instituto Nacional de Ciências e Tecnologia em Geofísica do Petróleo-CNPq, Brazil*

³*Instituto de Ciencias de la Tierra Jaume Almera-CSIC, Barcelona, Spain*

Accepted 2015 August 14. Received 2015 August 13; in original form 2015 June 19

SUMMARY

We present an analytical approach to jointly estimate the correlation window length and number of correlograms to stack in ambient noise correlation studies to statistically ensure that noise cross-terms cancel out to within a chosen threshold. These estimates provide the minimum amount of data necessary to extract coherent signals in ambient noise studies using noise sequences filtered in a given frequency bandwidth. The inputs for the estimation process are (1) the variance of the cross-correlation energy density calculated over an elementary time length equal to the largest period present in the filtered data and (2) the threshold below which the noise cross-terms will be in the final stacked correlograms. The presented theory explains how to adjust the required correlation window length and number of stacks when changing from one frequency bandwidth to another. In addition, this theory provides a simple way to monitor stationarity in the noise. The validity of the deduced expressions have been confirmed with numerical cross-correlation tests using both synthetic and field data.

Key words: Time-series analysis; Interferometry.

1 INTRODUCTION

In the past 15 yr, seismic ambient noise studies for monitoring and imaging purposes have gained increasing importance in seismology and surrounding research fields (e.g. Lobkis & Weaver 2001; Derode *et al.* 2003; Shapiro & Campillo 2004; Snieder 2004; Wapenaar 2004; Roux *et al.* 2005; Curtis *et al.* 2006; Sens-Schönfelder & Wegler 2006; Draganov *et al.* 2013). All of these noise studies are based on interferometric principles in which empirical Green functions (EGFs) are extracted based on different signal processing strategies (e.g. Bensen *et al.* 2007; Schimmel *et al.* 2011). The strategies to extract coherent signals from ambient noise are commonly based on cross-correlation and stacking approaches. The processing, however, is complicated by certain decisions taken to improve the signal extraction. One of the main decisions, particularly important in monitoring problems, is the choice of the amount of data to be used in the averaging process. Usually this choice is based on some empirical or pre-established experience values (e.g. Seats *et al.* 2012). The time series need to be sufficiently long to assure a balanced coverage of noise sources for a complete build-up of the EGF and to reduce the impact of noise cross-terms (e.g. Snieder 2004), which in all theoretical derivations are supposed to vanish. The emergence of EGFs has been studied from statistical view points by analysing the variance of cross-correlation studies (e.g. Sabra *et al.* 2005; Weaver & Lobkis 2005) as function of data

length and frequency bandwidth (FBW). However, the length of data which needs to be averaged for cross-term cancellation is still unknown (Weemstra *et al.* 2014).

Using a statistical approach, we analyse here how much data averaging is needed to restrain the cross-terms variance to be below a pre-defined value in the final stacked correlograms. We obtain expressions that provide the starting correlation window length and number of stacks for ambient noise correlation studies, taking also in consideration the dependence on the FBW. Our expressions can also be used to detect non-stationarity in data caused by fluctuations in the noise field. Note that the results say nothing about how efficiently EGFs emerge which means that analysing more data might be necessary, depending on source characteristics (Sabra *et al.* 2005; Weaver & Lobkis 2005) and abundance of scatterers (Larose *et al.* 2006). However, fulfilment of the minimum averaging condition makes SI viable, as exemplified by Weemstra *et al.* (2014) in their analysis of amplitude attenuation.

2 METHODS

Let $u(t)$ and $z(t)$ be two simultaneous noise records from two different locations. Both records are filtered in the FBW $[f_{\min}, f_{\max}]$, where we expect to exist coherent signals (s) and random noise (n and r) so that $u(t) = s(t) + n(t)$ and $z(t) = s(t + \alpha) + r(t)$. We assume that

$s(t)$ and $s(t + \alpha)$ are the same signal lagged by α , and that $n(t)$ and $r(t)$ are both stationary, spatially non-correlated zero-mean random noises. The signals can be of arbitrary shape, but we assume without loss of generality that the amplitudes are not attenuated.

The geometrically normalized cross-correlation of $u(t)$ and $z(t)$ for time lag τ in a correlation window length L is given by

$$c_{uz}(\tau) = c_{ss}(\tau, \alpha) + c_{sr}(\tau) + c_{ns}(\tau, \alpha) + c_{nr}(\tau), \quad (1)$$

where c_{ab} is the correlation of $a(t)$ and $b(t)$, defined by

$$c_{ab}(\tau) = \frac{\int_{t=t_0}^{t_0+L} a(t+\tau)b(t)dt}{\sqrt{\int_{t=t_0}^{t_0+L} [a(t+\tau)]^2 dt \cdot \int_{t=t_0}^{t_0+L} [b(t)]^2 dt}}, \quad (2)$$

where t_0 is the start time of the correlation window length. The denominator $\Omega(\tau)$ is the geometric mean of the energy within the correlation window length L . Note that all cross-correlation terms in eq. (1) involve noise. In turn, if one uses $z(t) = u(t)$ in eq. (1), the noise auto-correlation c_{nn} is peaked at zero lag and decays approximately as $e^{-D\tau}$ (Beran 1992), where D is a diffusion coefficient. Seismic noise has very short-range temporal correlation (that is, D is very high) so that $c_{nn}(\tau)$ is considered henceforth as a noise cross-correlation term for $\tau \neq 0$.

Usually, noise is assumed to be uncorrelated with signals or with other noise. That is, all noise terms in eq. (1) are assumed to cancel out after averaging. However, we will not take this assumption as being *a priori* warranted; instead of this, we will focus on the noise cross-terms to find statistical conditions which ensure that they cancel out to within an ascribed confidence level.

Let us express the already defined correlation window length L as function of f_{\min} by the equations

$$L = KL_0, \quad \text{where } L_0 = 1/f_{\min}. \quad (3)$$

Note that L_0 is the largest noise period contained in the data and that K is an integer number. Now we define the minimum total record length R to be used in the averaging process as

$$R = N(L + \tau_{\max}), \quad (4)$$

where N is the number of stacks and τ_{\max} is the maximum lag to be used (that is, $-\tau_{\max} \leq \tau \leq \tau_{\max}$). Next, we estimate K and N to statistically ensure that the noise cross-terms cancel out to within a pre-specified threshold.

2.1 Cancellation of noise cross-terms

The noise cross-term $c_{ns}^i(\tau)$ (eq. 1) is given by

$$c_{ns}^i(\tau) = \frac{\int_{t=t_0}^{t_0+L} s^i(t+\tau)n^i(t)dt}{\Omega^i(\tau)}, \quad (5)$$

where the superscript i stands for the i th summand of the stacking and $\Omega^i(\tau)$ is defined in association with eq. (2). The integrand is a signal-random noise product which results in a new sequence of random numbers $m^i(t, \tau)$. So

$$c_{ns}^i(\tau) = \int_{t=t_0}^{t_0+L} m^i(t, \tau)dt. \quad (6)$$

Using the relation between L and L_0 (eq. 3), eq. (6) can be expressed as the sum

$$c_{ns}^i(\tau) = \sum_{k=1}^K \int_{t=t_k}^{t_k+L_0} m_k^i(t, \tau)dt = \sum_{k=1}^K \mu_k^i(\tau), \quad (7)$$

where k stands for the k th segment of L , each one with time length L_0 (eq. 3). $\mu_k^i(\tau)$ is proportional to the mean of $m_k^i(t, \tau)$ over L_0 , that is, $\mu_k^i(\tau) = L_0 \langle m_k^i(\tau) \rangle$, where $\langle \cdot \rangle$ stands for the mean.

Let σ_k^{i2} be the time-domain variance of $\mu_k^i(\tau)$ over L_0 . Thus, σ_k^{i2} is proportional to the variance of the mean $\langle m_k^i(\tau) \rangle$. Observe that σ_k^{i2} is a standardized energy density variance because $\mu_k^i(\tau)$ is normalized relative to $\Omega^i(\tau)$ (eq. 5). For the summation process, σ_k^{i2} is the input variance. After K summations along the correlation window length L , $c_{sn}^i(\tau)$ (eq. 7) has its variance reduced to $\sigma_K^{i2} = \sigma_k^{i2}/K$, according to the expected decay of the sample mean variance with the number of measurements (Navidi 2011). The stacking operation further reduces this variance, so that:

$$c_{ns}(\tau) = \sum_{i=1}^N \sum_{k=1}^K \mu_k^i(\tau) \quad \text{and} \quad \sigma_K^{N2} = \sigma_k^{i2}/(NK). \quad (8)$$

The output variance σ_K^{N2} thus decreases with increasing values for K or number of stacks N . Given a confidence level ϵ^2 to be satisfied by the output variance of the stacked cross-terms (that is, by imposing $\sigma_K^{N2} \leq \epsilon^2$), NK can be estimated as

$$NK \geq \sigma_k^{i2}/\epsilon^2, \quad (9)$$

ensuring that $c_{sn}(\tau)$ integrates to zero within this confidence level.

The same value NK is sufficient to cancel out the other noise cross-terms in eq. (1). Moreover, using eqs (3) and (4), this statistical condition will be attained for any lag τ . Altogether, the above arguments justify the use of eqs (3) and (4). Note that σ_k^{i2} may be estimated from the correlated series.

The order of magnitude of NK can be estimated from eq. (9). Because $c_{sn}(\tau)$ is a standardized cross-correlation (eqs 2 and 5), $\mu_k^i(\tau)$ in eq. (7) satisfies $-1 \leq \mu_k^i(\tau) \leq 1$. So, in the worst case the input variance $\sigma_k^{i2} \approx 1$, given that $\langle \mu_k^i(\tau) \rangle \approx 0$. Using $\epsilon \approx 0.01$, then $NK \approx 10^4$. Computational cost of correlation operations is usually higher than for stacking, which may justify $N > K$. To simplify, we use $N \approx K$ so $K \approx 100$. For example, if $f_{\min} = 0.05$ Hz then $L \approx 2000$ s (eq. 3). Therefore, for zero lag, the minimum total record length R (eq. 4), ensuring that the noise cross-terms cancel out within the confidence level $\epsilon \approx 0.01$, must be approximately equal to 2.3 days.

2.2 Variance dependence on the FBW

So far, the dependence of NK on the FBW does not explicitly appear. In fact, this dependence is already implicitly incorporated in σ_k^{i2} . Nevertheless, the presented approach can be adjusted to relate NK for different FBW data. Consider two cross-correlations A and B which result from records filtered in the FBWs $[f_{\min}^A, f_{\max}^A]$ and $[f_{\min}^B, f_{\max}^B]$, having standardized time-domain energy variances σ_A^2 and σ_B^2 , respectively, as input variances. The respective products $N_A K_A$ and $N_B K_B$ are then related, as deduced in the Additional Supporting Information, by the following expression:

$$N_B K_B / N_A K_A = \sigma_B^2 / \sigma_A^2 = (n_A^2 - 1) / (n_B^2 - 1), \quad (10)$$

where $n_A = f_{\max}^A / f_{\min}^A$ and $n_B = f_{\max}^B / f_{\min}^B$. Note that $(n_A^2 - 1) / (n_B^2 - 1)$ is a measure of the ratio between relative FBWs. Eq. (10) thus shows that a cross-correlation, calculated with band-passed white noise sequences with relatively wide FBW, needs a smaller product NK to satisfy eq. (9) than a cross-correlation based on relatively narrow FBW data.

The white noise hypotheses is not an unrealistic assumption nor a limiting factor in our approach. Commonly, noise data are

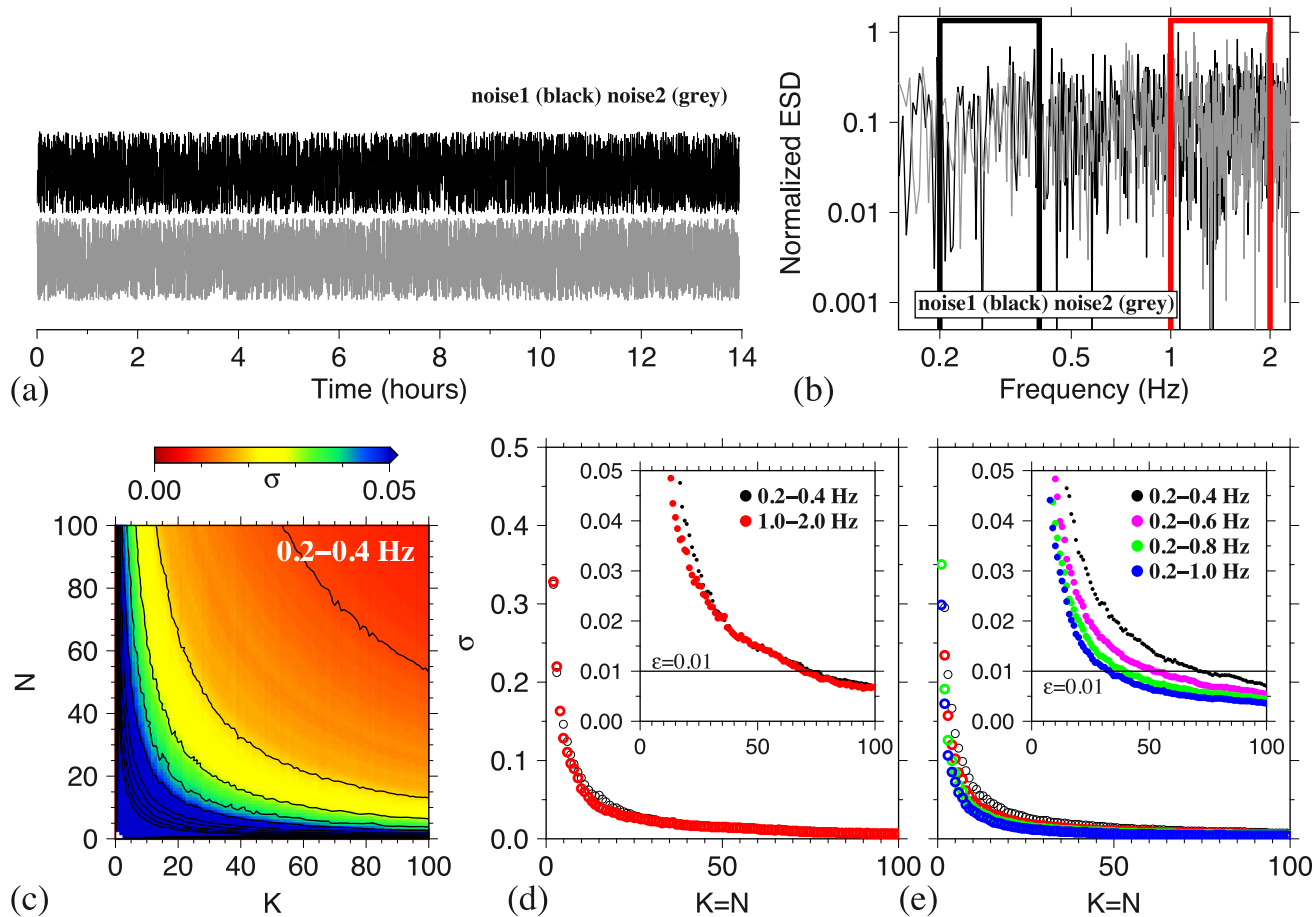


Figure 1. Synthetic data: (a) Noise sequences. (b) Energy density spectra of the noise sequences. Coloured boxes indicate FBWs. (c) Contour plot of σ_K^N for zero-lag cross-correlation of the sequences filtered in the FBW [0.2, 0.4] Hz. (d) σ_K^N decay curves along the line $N = K$ in the (K, N) plane and their zooms in the interval $0 < \sigma < 0.05$ (inset) for the FBWs shown in (b). (e) The same as in (d) for different FBWs. Data in (a) and (b) have been decimated for visual purposes to every 50th and 100th sample, respectively. Data filtering was done using a zero-phase Butterworth bandpass filter.

approximately whitened within the considered FBW before computation of cross-correlations (Bensen *et al.* 2007). Nevertheless, in the Additional Supporting Information (available in the online version of the article) we show also how to apply eq. (10) for non-white noise.

3 NUMERICAL TESTS

3.1 Synthetic data

Two synthetic noise sequences (Fig. 1a) were generated using pseudo-random white noise generators (Press *et al.* 1996). For zero-lag cross-correlation of the two sequences filtered in the FBW [0.2–0.4] Hz (Fig. 1b), σ_K^N values were calculated (Fig. 1c) using eq. (8) and visiting 100 times (each one for a different realization of the noise sequences) every point of the mesh (K, N) , $1 \leq K(\text{or } N) \leq 100$, with a step size equal to 1. As predicted by eq. (10), σ_K^N (for $N = K$) decay curves associated with different FBWs, but having the same ratio $n = f_{\max}/f_{\min}$, practically overlap (Fig. 1d). In addition, correlograms calculated with higher ratio n attain the tolerance threshold at smaller values of $K = N$ (Fig. 1e). Along the line $N = K$, we can use the K -ratio K_B/K_A , associated with the respective values of K where the decay curves cross the same threshold ϵ to measure the relative decay rate. Table S1 (Additional Supporting Information)

shows that there is good agreement among predicted (eq. 10) and observed (Fig. 1e) K -ratios associated with $\epsilon = 0.01$.

3.2 Field data

We use now vertical component ambient noises recorded over 22 days starting on 2010 April 14 at two broad-band stations separated by 477 km (Fig. 2a). These sequences are representative of field data because they contain amplitude variations due to variant ambient noise excitation and earthquakes (the spikes in Fig. 2a). Energy density spectra (normalized to amplitude 1) are clearly non-white showing two maxima at 0.05 and 0.2 Hz (Fig. 2b).

For zero-lag cross-correlations of the two sequences filtered in the FBW [0.1, 0.2] Hz (blue box in Fig. 2b), σ_K^N values were calculated (Fig. 2c) using 25 randomly chosen data segments for each (K, N) mesh point; that is, using 25 zero-lag cross-correlation values per (K, N) and window length $L = K \times L_0 = K \times 10$ s. The FBW [0.1, 0.2] Hz was deliberately chosen so that the filtered sequences severely depart from the white noise assumption (Fig. 2b). Although there exist perturbations in the σ_K^N curves, the overall hyperbolic appearance remains visible (Fig. 2c).

Nonetheless the noises are non-white, correlograms calculated with higher ratios n attain the tolerance threshold at smaller values of $N = K$ (Fig. 2d). After the data preprocessing (1 bit + whitening), σ_K^N decay curves associated with different FBWs, but having the

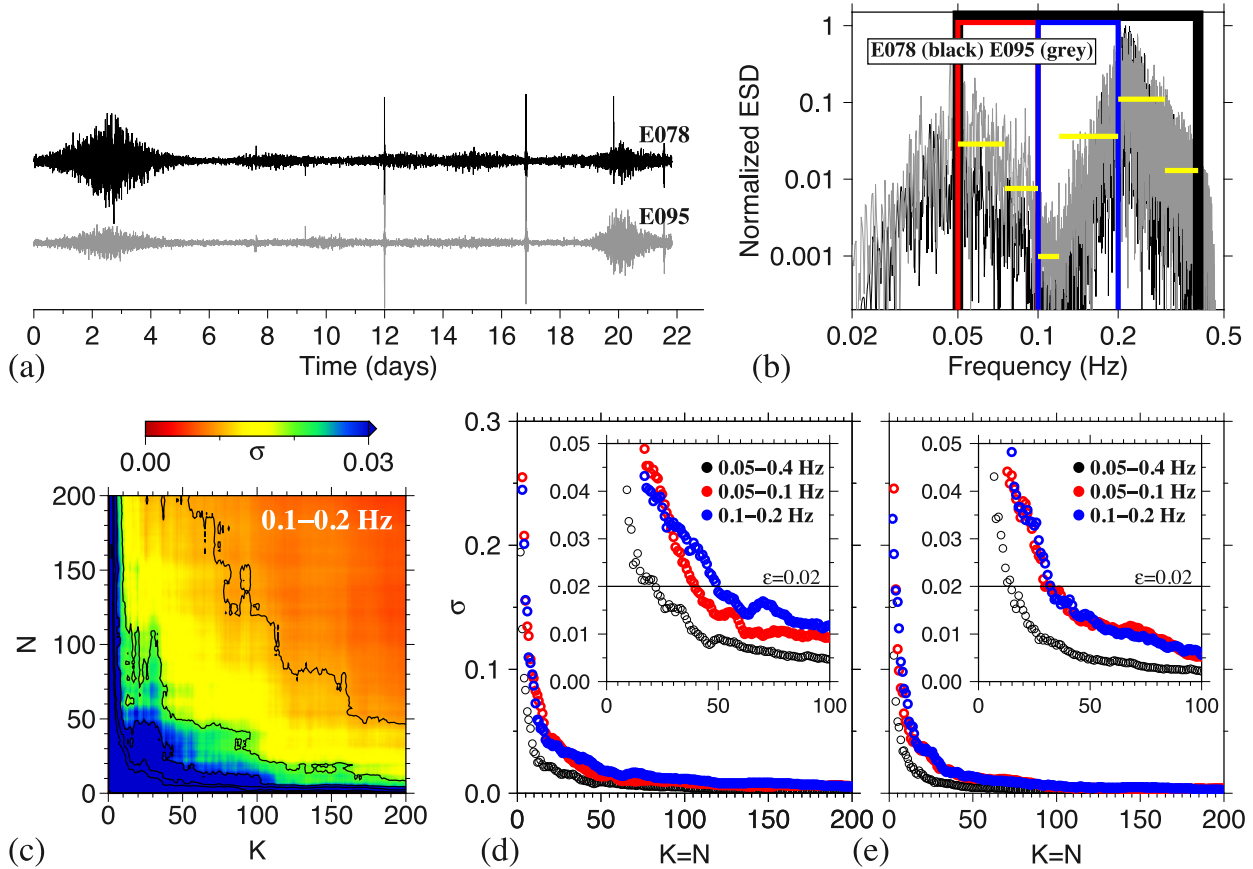


Figure 2. Field data from two broad-band stations of the Spanish TopoIberia project: (a,b) The same as in Figs 1(a) and (b). (c) The same as in Fig. 1(c), but using the FBW [0.1–0.2] Hz. (d) The same as in Fig. 1(d) using non-preprocessed data. (e) The same as in (d) but using preprocessed data (1 bit + whitening). In (b), yellow lines show staircase approximations of the spectra, which are used to calculate the predicted K -ratio values in Table S1 for non-preprocessed data (see Additional Supporting Information).

same ratio n , almost coincide (Fig. 2e) as predicted by eq. (10). In addition, reasonable ratios among the predicted and observed K -ratios are obtained in most cases (Table S1 Additional Supporting Information).

There is an overall agreement between synthetic and field data results. However, a striking difference is present: for field data σ_N^N curves are non-monotonically decreasing (Figs 2d and e), evidencing non-stationarity, particularly when using non-preprocessed data. A non-stationary process is one whose statistical properties change over time, in particular the mean and variance (Nason 2006). So, a departure from the expected time decay behaviour of the sample mean variance must involve non-stationarity. Indeed, monitoring the σ_K^N decay curves is a simple strategy to detect non-stationarity. When dealing with field data, care must be taken to face data non-stationarity. The 1 bit normalization (Cupillard *et al.* 2011), the use of short-time windows (Seats *et al.* 2012), and noise classification and selection approaches (Groos & Ritter 2009) are some of the used procedures to reduce data non-stationarity. Note that the concept of non-stationarity depends on the time scale and/or number of visits in a long series one uses. For non-preprocessed field data, we verified that the oscillations in the $N = K$ decay curves are attenuated with increasing sample number.

The dependence of σ_K^N with $(KN)^{-1/2}$ (eq. 8) for cross-term cancellation is consistent with the results associated with the emergence of EGFs presented by (1) Snieder (2004) (his eq. A12, where the ratio of the standard deviation of the cross-terms to the diagonal terms is proportional to $T^{-1/2}$), (2) Weaver & Lobkis (2005) (their

eq. 43, where the signal-to-noise ratio variance Var is proportional to T) and (3) Sabra *et al.* (2005) (their eq. 20, where the noise cross-correlation function variance is proportional to $1/T$). In addition, the dependence of σ_A^2 with $1/(n_A^2 - 1)$ (eq. 10) is also consistent with the results of Weaver & Lobkis (2005) ($Var \propto 1/\bar{w}^2$ in their same eq. 43, where \bar{w} is a central frequency).

Considering separately the effect of f_{\min} on R (eqs 3 and 4) provides us with a geometry to describe the σ_K^N behaviour in the (K, N) plane through eqs (8) and (9). This geometry shows that, to reduce the output variance, averaging through K (correlation window length) is equivalent to averaging through N (stacking). In addition, because the summation in eq. (8) does not depend on how the summands are grouped, the final output variance depends on the initial and final points in the (K, N) plane but not on the trajectory in this plane. Thus, σ_K^N is a potential function of K and N .

4 CONCLUSIONS

The correlation window length and number of stacks can be jointly estimated to statistically guarantee that noise cross-terms cancel out in stacked correlograms under the assumption that the noise is stationary. For practical use, one should estimate the cross-correlation energy density variance contained in the time window equal to the largest period present in the filtered FBW. Then correlation window length and number of stacks are estimated so that the output variance is reduced to attain a threshold in the final stacked

correlograms. An expression was also deduced to reschedule the required correlation window length and number of stacks when changing from one FBW to another.

Stationarity is the most severe assumption of our deductions. However, it is possible to monitor noise non-stationarity by detecting a non-monotonically decreasing behaviour of the energy density variance. The presence of non-stationarity and the absence of coherent signals in ambient noise are factors which are not controlled by the theory described here. Therefore, our approach should be considered as a starting point to guide computation trials in EGF extractions. The final length of the correlation window and number of stacks depends on the abundance, quality and duration of coherent signals. However, the minimum amount of data is controlled by the cancellation of noise cross-terms, which is a necessary condition in SI.

ACKNOWLEDGEMENTS

We thank CNPq for the research grants (WEM—No. 304301/2011-6; MS—Brazilian Science Without Border Program No. 40.2174/2012-7; AFDN—No. 302316/2011-6 and No. 484441/2012-4). MS also thanks the Spanish MISTERIOS (CGL2013-48601-C2-1-R) and TopoIberia projects. Figures were produced with GMT (Wessel & Smith 1998). We are very grateful to the editors J. Trampert and M. Ritzwoller, and to D. Mikesell and anonymous reviewers for their careful and encouraging revisions which helped to improve our manuscript.

REFERENCES

- 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**, 1239–1260.
- Beran, J., 1992. Statistical methods for data with long-range dependence, *Statistical Science*, **7**, 404–427.
- Cupillard, P., Stehly, L. & Romanowicz, B., 2011. The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise, *Geophys. J. Int.*, **184**, 1397–1414.
- Curtis, A., Gerstoft, P., Sato, H., Snieder, R. & Wapenaar, K., 2006. Seismic interferometry—turning noise into signal, *Leading Edge*, **25**, 1082–1091.
- Derode, A., Larose, E., Tanter, M., de Rosny, J., Tourin, A., Campillo, M. & Fink, M., 2003. Recovering the Green's function from field-field correlations in an open scattering medium, *J. acoust. Soc. Am.*, **113**, 2973–2976.
- Draganov, D., Campman, X., Thobcke, J., Verdel, A. & Wapenaar, K., 2013. Seismic exploration-scale velocities and structure from ambient noise seismic (<1 Hz), *J. geophys. Res.*, **118**, 4354–4360.
- Groos, J.C. & Ritter, J.R.R., 2009. Time domain classification and quantification of seismic noise in an urban environment, *Geophys. J. Int.*, **179**, 1213–1231.
- Larose, E. *et al.*, 2006. Correlation of random wavefields: an interdisciplinary review, *Geophysics*, **71**(4), SI11–SI21.
- Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.*, **110**, 3011–3017.
- Nason, G.P., 2006. Stationary and non-stationary time series. in *Statistics in Volcanology*, pp. 129–142, eds Mader, H. & Coles, S.C., The Geological Society.
- Navidi, W., 2011. *Statistics for Engineers and Scientists*, 3rd edn, McGraw-Hill.

- Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1996. *Numerical Recipes in Fortran 77: The Art of Computing*, 2nd edn, Cambridge Univ. Press.
- Roux, P., Sabra, K.G., Kuperman, W.A. & Roux, A., 2005. Ambient noise cross correlation in free space: theoretical approach, *J. acoust. Soc. Am.*, **117**, 79–84.
- Sabra, K.G., Roux, P. & Kuperman, W.A., 2005. Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function, *J. acoust. Soc. Am.*, **118**(6), 3524–3531.
- Schimmel, M., Stutzmann, E. & Gallart, J., 2011. Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale, *Geophys. J. Int.*, **184**, 494–506.
- Seats, K.J., Lawrence, J.F. & Prieto, G.A., 2012. Improved ambient noise correlation functions using Welch's method, *Geophys. J. Int.*, **188**, 513–523.
- Sens-Schönfelder, C. & Wegler, U., 2006. Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophys. Res. Lett.*, **33**, L21302, doi:10.1029/2006GL027797.
- Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient noise, *Geophys. Res. Lett.*, **31**, L07614, doi:10.1029/2004GL019491.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E*, **69**(4), doi:10.1103/PhysRevE.69.046610.
- Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.*, **93**, doi:10.1103/PhysRevLett.93.254301.
- Weaver, R.L. & Lobkis, O.I., 2005. Fluctuations in diffuse field-field correlations and the emergence of the Green's function in open systems, *J. acoust. Soc. Am.*, **117**(6), 3432–3439.
- Weemstra, C., Westra, W., Snieder, R. & Boschi, L., 2014. On estimating attenuation from the amplitude of the spectrally whitened ambient seismic field, *Geophys. J. Int.*, **197** (3), 1770–1788.
- Wessel, P. & Smith, W.H.F., 1998. New, improved version of the Generic Mapping Tools released, *EOS, Trans. Am. geophys. Un.*, **79**, 579, doi:10.1029/98EO00426.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this paper.

- (1) The deduction of eq. (10),
- (2) How to use it for non-white noise, and
- (3) The predicted and observed K -ratios for both synthetic and field data.

Table S1. Predicted and observed K -ratios (definition in text) associated with the tolerance thresholds $\epsilon = 0.01$ for synthetic data (Fig. 1e) and $\epsilon = 0.02$ for field data (Figs 2d and e). Double horizontal lines separate synthetic (above) from field data (below). For field data, P and R stand for *with* and *without* preprocessing (1 bit + whitening), respectively.

(<http://gji.oxfordjournals.org/lookup/suppl/doi:10.1093/gji/ggv336/-/DC1>).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.