On the use of the coda of seismic noise autocorrelations to compute H/V spectral ratios

F. N. Tchawe, B. Froment, Michel Campillo, L. Margerin

Geophysical Journal International

Geophys. J. Int. (2020) **220,** 1956–1964
Advance Access publication 2019 December 06
GJI Seismology

doi: 10.1093/gji/ggz553

On the use of the coda of seismic noise autocorrelations to compute H/V spectral ratios

F.N. Tchawe,[1] B. Froment,[1] M. Campillo[2] and L. Margerin[3]

[1] *Institut de Radioprotection et de Sureté Nucléaire, Fontenay-aux-Roses, France. E-mail: tflomin@yahoo.com*
[2] *Institut des Sciences de la Terre, Université Grenoble Alpes, CNRS, IRD, Grenoble, France*
[3] *Institut de Recherche en Astrophysique et Planétologie, Observatoire Midi-Pyrénées, Université Paul Sabatier, CNRS, Toulouse, France*

Accepted 2019 December 4. Received 2019 November 24; in original form 2019 August 2

SUMMARY

The horizontal to vertical spectral ratio (HVSR) of seismic ambient noise has been proven to be a fast and efficient method for characterizing the 1-D resonance frequency of the local subsurface in a practical framework. Over the last decades, theories have been developed in order to extend the exploitation of HVSR beside the frequency of its first peak, notably the diffuse field assumption (DFA) which links the HVSR to the Green's function of the local medium assuming the diffuseness of the seismic ambient noise wavefield. However, the underlying assumption of the seismic ambient noise being a diffuse, equipartitioned field may not be satisfied under certain circumstances. In order to exploit the contribution of scattering in forging diffuse wave fields, we leverage the advantages of coda waves and present a novel procedure for computing the HVSR, using the coda part of ambient noise correlations. We applied this technique to data gathered at the plio-quaternary sedimentary basin of Argostoli, Greece. Results on this data set show the potential of the method to improve the temporal stability of the HVSR measurements compared to the classical computation, and the fit with the theoretical HVSR curve derived from the DFA theory. These results suggest that this procedure could help in extracting physical information from the HVSR and thus could lead to an extended use of these measurements to characterize the mechanical properties of the medium.

Key words: Coda waves; Seismic noise; Site effects.

1 INTRODUCTION

It is well known that the effects of surface geology or local stratigraphy can considerably modify ground motion during earthquakes. This is due to a change in soil properties, which generally become softer in the vicinity of the Earth's surface. In the case of sedimentary basins, seismic waves can be trapped and the characteristics of the soft soil deposits will strongly affect ground motion. As a result, these so-called site effects lead to longer and amplified shaking, which can cause severe damage. Site effect assessment therefore needs to be carefully carried out for preventive measures.

Several geophysical methods have been developed in order to estimate the site amplification characteristics. Amongst them, one of the most appealing technique in an operational framework is the single-station based horizontal to vertical spectral ratio (HVSR) technique which was initially introduced by Nogoshi & Igarashi (1971) and popularized by Nakamura (1989). These authors pointed out the relationship between the HVSR peak frequency computed on seismic ambient noise recordings, and the ground fundamental resonance frequency. The observed correlation between the HVSR peak and the resonance frequency of the soil has been confirmed with both experimental and numerical investigations, using different data sets (e.g. Lermo & Chavez-Garcia 1993; Lachet & Bard 1994; Fäh 1997; Arai & Tokimatsu 2004). However, the lack of sound theoretical basis underlying the physics of this technique has led to various physical interpretations of the HVSR measurements, mostly related to assumptions on the composition of the seismic noise wavefield (see Bonnefoy-Claudet *et al.* 2006; Lunedei & Malischewsky 2015 for a detailed review on the theory of the HVSR technique). More still, to date only the HVSR peak frequency is considered as providing reliable and robust information. Applied and theoretical research on HVSR nowadays aims at exploiting more than just the HVSR peak frequency to provide additional information on seismic amplification effects and to constrain the elastic properties of the shallow geological structure. This requires a model for the full ambient noise wavefield rather than using HVSR interpretations assuming an ambient noise wavefield composed of either surface waves (e.g. Nogoshi & Igarashi 1971; Konno & Ohmachi 1998) or body waves (e.g. Nakamura 1989, 2000). Recently, the DFA proposed by Sánchez-Sesma *et al.* (2011) links the HVSR to the Green's function of the local medium assuming that the seismic ambient noise is a diffuse wavefield. This interpretation can lead to very

attractive applications of HVSR to characterize the medium's mechanical properties beyond the estimate of the resonance frequency. This approach is however based on the quite strong assumption that the ambient noise wavefield is diffuse. In order to get closer to the theoretical requirements in the DFA framework, we propose in this paper an alternative procedure to compute the HVSR that benefits from the diffuse nature of coda waves. The use of the coda of correlations to take advantage of diffusive properties has been implemented for improving Green's function extraction (Stehly *et al.* 2008; Froment 2011) and suggested to compute HVSR (Sánchez Sesma *et al.* 2012).

In Section 2, we first recall the main theoretical aspects behind the DFA before presenting a new method for the HVSR computation. We then present in Section 3 the data set gathered within the Argostoli sedimentary basin (Cephalonia, Greece) and the data processing used in this study. In Section 4, we will show and discuss the results obtained using our approach. We conclude in Section 5 about the potential of the proposed processing.

2 METHOD

2.1 HVSR formulation based on the diffuse field assumption (DFA)

A few studies have considered a theoretical HVSR formulation in the framework of diffuse fields. We can infer a formulation from Margerin (2009) based on the principle of equipartition of elastic waves. Using a detailed model of the subsurface at Pinyon Flats Observatory, Margerin *et al.* (2009) showed that this formulation is in good qualitative and quantitative agreement with the observations. A couple of years later, inspired by the possibility of retrieving the Green's function from the average cross correlations of the displacements within a diffuse field (e.g. Lobkis & Weaver 2001; Sánchez-Sesma & Campillo 2006; Sánchez-Sesma *et al.* 2008), Sánchez-Sesma *et al.* (2011) express the HVSR as a function of the imaginary part of the Green's function. Following Sánchez-Sesma *et al.* (2008), Perton *et al.* (2009), We write the directional energy density at a point $\mathbf{x}$, $E_i(\mathbf{x}, \omega)$, as:

$$E_i(\mathbf{x}, \omega) = \rho\omega^2 \langle u_i(\mathbf{x}, \omega)\, u_i^*(\mathbf{x}, \omega)\rangle \propto \mathrm{Im}[G_{ii}(\mathbf{x}, \mathbf{x}, \omega)], \tag{1}$$

where ρ is the mass density, ω is the circular frequency, $u_i(\mathbf{x}, \omega)$ is the displacement field component in direction i at $\mathbf{x}$ and $G_{ii}(\mathbf{x}, \mathbf{x}, \omega)$ is the Green's function defined as the displacement at $\mathbf{x}$ in the direction i produced by a colocated unit harmonic load in direction i. The asterisk (*) in eq. (1) represents the complex conjugate and the angular brackets denote an average. The middle term of eq. (1) corresponds to the autocorrelation of displacement in direction i at $\mathbf{x}$ ($C_{ii}(\mathbf{x}, \omega)$). Arai & Tokimatsu (2004); Sánchez-Sesma *et al.* (2011) write the energy ratio HVSR as:

$$\begin{aligned} HVSR(\mathbf{x}, \omega) &= \sqrt{\frac{E_1(\mathbf{x}, \omega) + E_2(\mathbf{x}, \omega)}{E_3(\mathbf{x}, \omega)}} \\ &= \sqrt{\frac{\langle |u_1(\mathbf{x}, \omega)|^2\rangle + \langle |u_2(\mathbf{x}, \omega)|^2\rangle}{\langle |u_3(\mathbf{x}, \omega)|^2\rangle}}, \end{aligned} \tag{2}$$

where the subscripts 1 and 2 refer to horizontal, and 3 to vertical degrees of freedom. In what follows, we will consider the eastern (E), northern (N) and vertical (Z) components of displacement to denote these 3 degrees of freedom, respectively. Eq. (2) thus expresses the HVSR in terms of the autocorrelation of the displacement signals. For a diffuse noise field, Sánchez-Sesma *et al.* (2011) express HVSR as a function of the Green's function by using eqs (1) and (2):

$$HVSR(\mathbf{x}, \omega) = \sqrt{\frac{\mathrm{Im}[G_{EE}(\mathbf{x}, \mathbf{x}, \omega)] + \mathrm{Im}[G_{NN}(\mathbf{x}, \mathbf{x}, \omega)]}{\mathrm{Im}[G_{ZZ}(\mathbf{x}, \mathbf{x}, \omega)]}}. \tag{3}$$

For a detailed explanation of the above formulations, the reader is called to read the papers by Margerin (2009) and Sánchez-Sesma *et al.* (2011).

Eq. (3) directly relates the easy-to-implement HVSR to the local medium's intrinsic properties through the ratio of noise autocorrelations. By expressing the Green's function under the diffuse field assumptions, Sánchez-Sesma *et al.* (2011), García-Jerez *et al.* (2016) formulated an algorithm for forward calculation of HVSR. This has led to the inversion of HVSR measurements to get the local 1-D structure (e.g. Spica *et al.* 2015; García-Jerez *et al.* 2016; Piña-Flores *et al.* 2017). It is worth noting that this formulation of HVSR entails that the link between the amplitude of HVSR and the level of amplification due to the local subsurface geology is not straightforward. Instead, the HVSR amplitude reflects the characteristics of the medium through the partition of energy in the signal considered.

The underlying assumption that the seismic ambient noise is a diffuse, equipartitioned field may be delusive under certain circumstances. Equipartition in the seismic noise wavefield is expected to arise when the distribution of noise sources is spatially homogeneous, which is highly unlikely in practice. Non-isotropy in noise illumination remains one of the main challenges and limitations of noise-based applications (e.g. Stehly *et al.* 2006; Weaver *et al.* 2009; Weaver 2010; Hillers *et al.* 2012; Mordret *et al.* 2013). Another important process contributing to randomization is the scattering of seismic waves by heterogeneities during the propagation in the medium. However, the seismic noise is a mixture of different contributions made of ballistic, singly scattered or multiply scattered waves. Our aim here is to extract a multiply scattered wavefield from the seismic noise in order to get closer to the theoretical requirements of the DFA framework, hence leading to more reliable characterizations of sites. Therefore, we will consider the coda of noise autocorrelation functions in our procedure as the multiply scattered part of the Green's functions below the station of interest. The exploitation of the coda part of noise correlation functions is supported by studies that have highlighted that late arrivals of the ambient noise correlation signals carry physical information (through the so-called C3 method (e.g. Stehly *et al.* 2008; Froment *et al.* 2011)).

2.2 HVSR Formulation on the coda of autocorrelation (HVSRc)

Our procedure first consists in computing autocorrelations of noise signals recorded at a given point $\mathbf{x}$ in order to reconstruct the local Green's function (GF) beneath $\mathbf{x}$. This first step allows us to extract the scattered part of the reconstructed deterministic signal, which is impossible on the original raw ambient noise record due to the mix of direct and scattered waves from various sources. We then select the coda part of the autocorrelation and perform the HVSR on this signal. For the sake of clarity, we will call this coda-based measurement HVSRc in contrast with the usual noise-based HVSRn which is defined as:

$$HVSR_n = \sqrt{\frac{u_E^2(\mathbf{x}, \omega) + u_N^2(\mathbf{x}, \omega)}{u_Z^2(\mathbf{x}, \omega)}}. \tag{4}$$

Considering a unit source at a point $\mathbf{x}$ in the direction i, the 3-component noise correlation at $\mathbf{x}$ is defined as:

$$C_{iZ}(\mathbf{x}, \omega) = \langle u_i(\mathbf{x}, \omega) u_Z^*(\mathbf{x}, \omega) \rangle \tag{5}$$

$$C_{iE}(\mathbf{x}, \omega) = \langle u_i(\mathbf{x}, \omega) u_E^*(\mathbf{x}, \omega) \rangle \tag{6}$$

$$C_{iN}(\mathbf{x}, \omega) = \langle u_i(\mathbf{x}, \omega) u_N^*(\mathbf{x}, \omega) \rangle, \tag{7}$$

where Z, E and N correspond to the vertical, east and north directions, respectively. $C_{ij}(\mathbf{x}, \omega)$ is the reconstructed signal in the direction j at point $\mathbf{x}$ due to a colocated unit source in the direction i, with $i, j \in \{Z, E, N\}$. It is worth noting that in an ideal 1-D configuration and an adequate distribution of noise sources, the cross terms $C_{ij|i \neq j}$ are null, but can be made non-zero with the presence of scattering. We then select the coda part of each temporal correlation signals and we perform the HVSR measurement as follows:

$$HVSR_{ci} = \sqrt{\frac{|\boldsymbol{c}_{iE}(\mathbf{x}, \omega)|^2 + |\boldsymbol{c}_{iN}(\mathbf{x}, \omega)|^2}{|\boldsymbol{c}_{iZ}(\mathbf{x}, \omega)|^2}}, \tag{8}$$

where $HVSR_{ci}$ stands for the HVSRc obtained for a unit virtual source in direction i. $\boldsymbol{c}_{iZ}$, $\boldsymbol{c}_{iE}$ and $\boldsymbol{c}_{iN}$ are the Fourier transforms of the coda part of the time equivalent correlation signals of C_{iZ}, C_{iE}, C_{iN}, respectively. Note that, unlike most HVSR applications, we do not average horizontal components here but we sum them up to be consistent with the DFA computation expressed in eqs (2) and (3). The resulting HVSRc can then be obtained by averaging the contribution from all three virtual sources:

$$HVSR_c = \frac{1}{3}\left(HVSR_{cZ} + HVSR_{cE} + HVSR_{cN}\right). \tag{9}$$

It is worth noting that in a 1-D assumption, a horizontal force generates both Love and Rayleigh waves in contrast to a vertical force which do not generate Love waves. In this regard, a horizontal force is expected to produce a more complete wavefield which would favor the emergence of equipartition. This hypothesis is confirmed by the results presented in Section 4.2, which will illustrate the fact that our formulation brings us closer to a diffuse wavefield. As a result, the final formulation used in the paper is obtained by averaging only the $HVSR_{ci}$ for the horizontal virtual sources:

$$HVSR_c = \frac{1}{2}\left(HVSR_{cE} + HVSR_{cN}\right). \tag{10}$$

In contrast to the DFA formulation by Sánchez-Sesma *et al.* (2011), we use the cross terms of the correlation tensor that are supposed to be null in the DFA interpretation derived for a 1-D tabular medium. Our approach consists in exploiting the scattering process and the resulting field (the coda of the complete correlation tensor including cross-terms), in order to compute the HVSR on a field that gets as close as possible to a diffuse field.

3 DATA AND PROCESSING

3.1 Data

We used data from the plio-quaternary Argostoli basin in Greece. The stratigraphy of the Argostoli basin shows a cretaceous substratum formed of massive limestone covered by plio-quaternary sediments that make this basin prone to site effects. The basin is about 60-to-100-m deep and 2-km wide (Fig. 1) and has been widely studied in the context of the European NERA and the French SINAPS@ projects (e.g. Cultrera 2014; Imtiaz 2015; Imtiaz *et al.* 2018; Perron *et al.* 2018a,b; Theodoulidis *et al.* 2018a,b) to understand its seismic response since it's located in an active seismic region. A linear network of 62 stations crossing the basin was deployed for 7 months between September 2011 and April 2012 as part of the FP7 EU-NERA 2010-2014 project (Cultrera 2014; Theodoulidis *et al.* 2018a). As a part of this deployment, a small array was also installed at the basin centre (Array A, see Fig. 1). The 21 stations of array A were equipped with CMG-40 30 s sensors of the Guralp Systems sensors, which continuously recorded the three component of ground motion. This array was used, in particular for characterizing the 1-D velocity profile within the basin (Imtiaz *et al.* 2018). Due to its central location, we used data from array A in this study, in order to assume a 1-D configuration which is the basis of HVSR interpretations in terms of resonance and energy partition. In this paper, we present the results for the central station KEA00 using 3 months of data (January 2012–April 2012). The data was downloaded from the French Seismologic and Geodetic Network datacentre (RESIF, http://www.resif.fr/, network 4C (2011–2014): NERA-JRA1-A) (Cornou & Bard 2018).

3.2 Noise processing

We first removed the mean and the trend of the continuous ambient noise recordings. The records were then resampled at 40 Hz and the instrument response was removed. We cut the daily records into 4-hr segments so as to consider durations similar to those used in HVSR campaign in an operational context. The procedure in the HVSR computations normally considers hourly signals divided in short windows that are selected so as to keep those containing the most stationary noise, using for example the STA/LTA criterion (e.g. Koller *et al.* 2004). The objective of such processing is to remove transients related to specific sources. However, in order to fully test the ability of our procedure to stabilize the measurements by using the scattered coda waves, we decided to use the complete data set and not to introduce a pre-selection that would require some subjective parametrization. We only removed 4-hr windows that contained glitches and more than 10 per cent of zeros. Data was filtered between 0.6 and 15 Hz and correlations were computed following eqs (4)–(7) on 15-min-long windows. We then averaged the positive and negative (i.e. causal and acausal) parts of the correlation signals. We also computed the HVSRn (see eq. 4) on 15-min segments in order to limit differences in processing between both computations. For both procedures (HVSRc and HVSRn), 16 15-min HVSR estimations were averaged to get the HVSR measurement associated with each 4-hr segment. The detailed numerical scheme is shown in the supplementary material (Fig. S1). It highlights the difference between the processing used for our computation of the 'classical' noise-based measurement [HVSRn, eq. (4)] and our coda-based approach [HVSRc, eqs (8) and (10)].

3.3 Selection of the coda signal

We choose to select the beginning of the coda window at a lag time of 20 s in order to get rid of the influence of the first arrivals in the signal and favor the use of multiply scattered waves in our computation. Regarding the end of the window, we use a criterion based on the energy level in the signal. Fig. 2(b) shows the 2-s-smoothed energy envelope of the noise autocorrelations C_{NN} stacked over 3.5 months (Fig. 2a). Based on previous studies on the Argostoli basin, we know that the resonant frequency of the basin is between 1.5 and 2 Hz (e.g. Imtiaz 2015; Theodoulidis *et al.* 2018a). We thus computed the energy envelope in the frequency range of interest, that is between

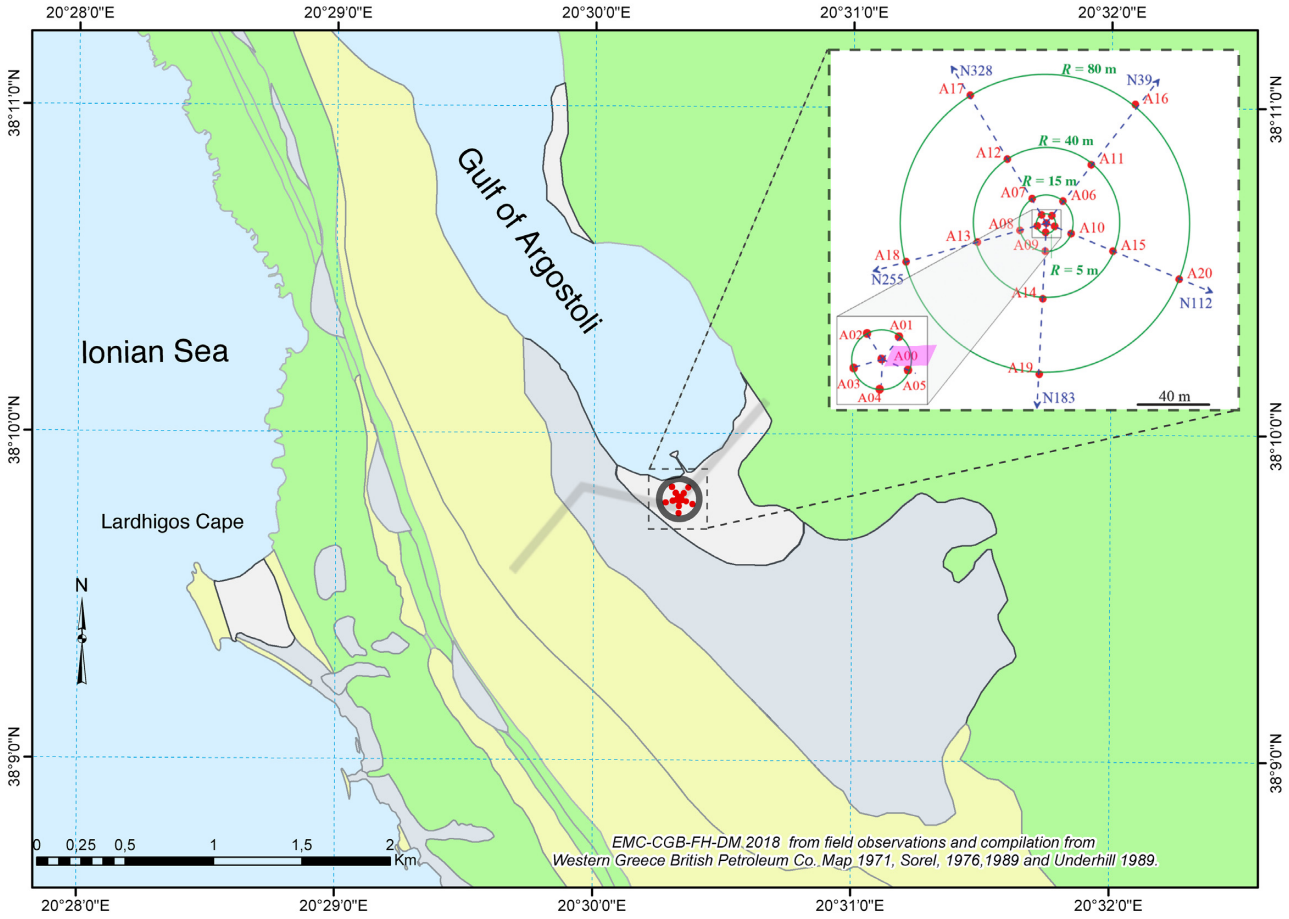

Figure 1. Simplified geological map of the Argostoli basin (from Cushing *et al.* 2019). The cretaceous substratum is shown in green. This substratum is covered by pliocene and quaternary fillings shown in yellow and grey, respectively. The location of the linear network composed of 62 stations is delineated by the grey thick line crossing the basin. Array A is represented by the black circle with its stations in red. The zoom details the layout of array A (from Imtiaz *et al.* (2018)). The purple tag highlights the station KEA00 whose results are discussed in this paper.

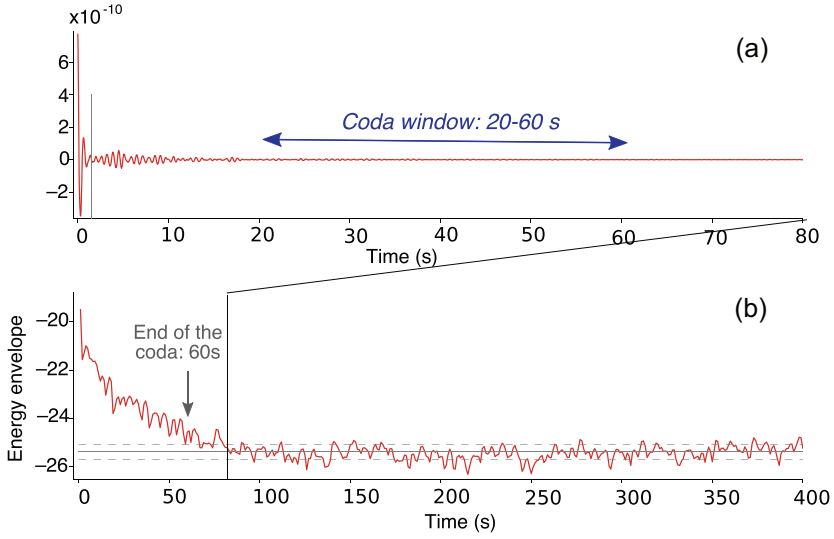

Figure 2. (a) First 80 seconds of the 3.5-month-stacked NN autocorrelation filtered between 1 and 6 Hz. The arrows highlight the extension of the coda window selected (20–60s). (b) Energy envelope of the signal shown in (a). The level of late fluctuations is shown by the plain and dashed horizontal lines standing for the mean (and mean ± 1 standard deviation) signal amplitude between 200 and 500 s. This level is used to choose the end of the coda window (i.e. 60 s in this case).

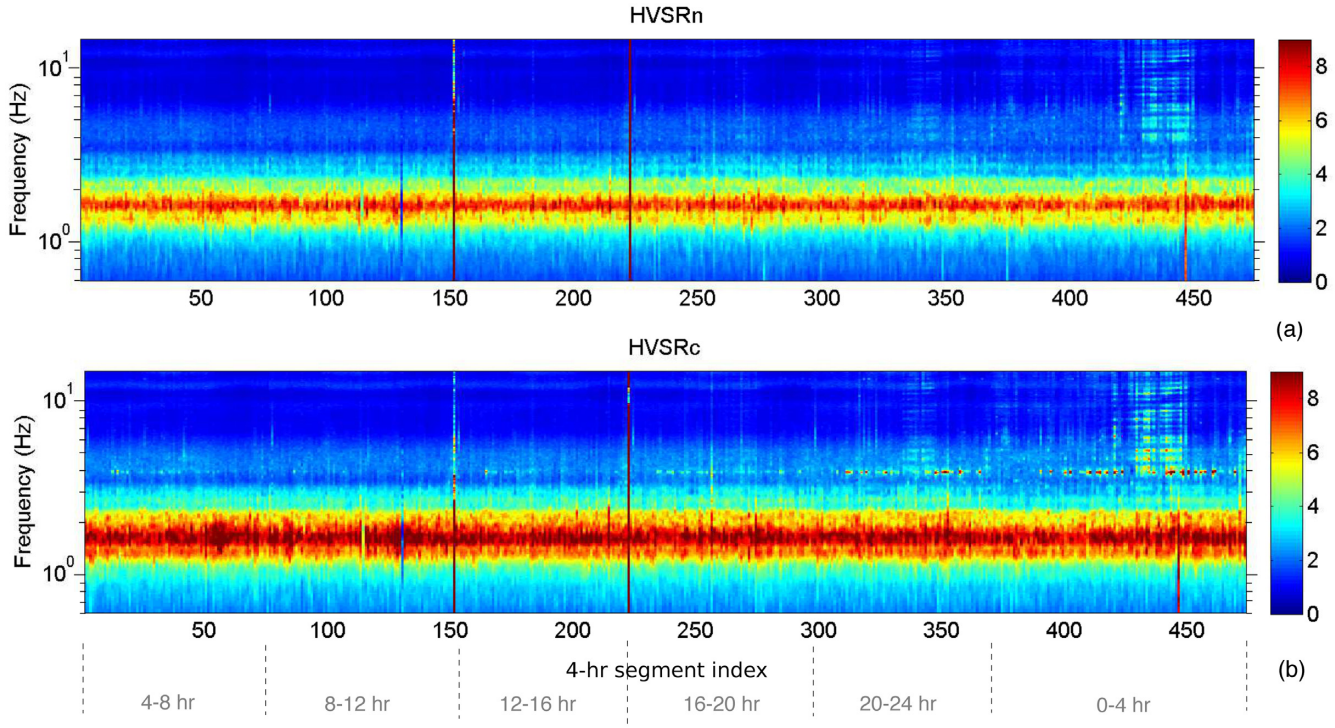

Figure 3. Temporal evolution of 4-hr noise based HVSRn (top panel) and HVSRc (bottom panel). The colour represents the amplitude of the HVSR and the number on the *x*-axis stands for the index of 4-hr segments. Measurements have been sorted out as a function of time slots.

1 and 6 Hz. The energy decays in the first tens of seconds, before starting to fluctuate steadily at about 60–80 s (the interval between the dashed grey lines). Such a decay is characteristic of coda waves. The onset of this steady fluctuations (i.e. 60 s) was chosen as the end of the coda window (see Fig. 2b).

4 RESULTS

Fig. 3 presents the results of both the HVSRn and HVSRc computation for the whole time period considered. For the sake of clarity, we have gathered results for similar time slots (i.e. 0–4 hr; 4–8 hr; 8–12 hr; 12–16 hr; 16–20 hr; and 20–24 hr). This allows us to view our results as a function of day period. Note that the number of 4 hr segments may differ for the different time slots due to the selection procedure discussed in Section 2. Fig. 3 shows that both computations clearly display a peak in the HVSR around 1.7 Hz for each 4-hr measurement that is consistent with estimations of resonance frequency in the area. However, the amplitude of this peak is larger for the HVSRc computation. We observe that the HVSRc shows a narrow peak at 3.8 Hz. This peak is also present on the HVSRn, but it is narrower and it disappears after any smoothing necessary to extract the main information from HVSR curves. This peak is stronger at night and its narrowness made us investigate its origin to get an opinion about its physical meaning. The standard spectral ratio (SSR) reported by previous studies in the area (e.g. Imtiaz 2015; Theodoulidis *et al.* 2018a), computed using earthquake recordings, show an amplification just below 4 Hz at A00 compared to the nearby rock. The consistency between this SSR results and our HVSR results opens up the possibility that the narrow peak at 3.8 Hz observed in our results might have a physical significance, and may reflect an actual amplification characteristic of the local subsurface. Nevertheless, from our analysis, it turns out that a strong narrow peak is observed at the same frequency on the spectrum of the northern component of the concerned 4-hr segments (which is also present on the eastern component but at a much lower level, see Fig. S2). In view of the spectral signature of this signal, we associate this peak with a monochromatic source. The dominance of this source at night reveals either that it is active only at night or it becomes dominant relatively to other sources during quieter periods. This source seems to last for a few tens of minutes polluting the noise correlation signal at 3.8 Hz for durations longer than the propagation time. Moreover, Fig. S3 shows that the NN noise correlation is completely dominated by the source autocorrelation, which makes the NN correlation function unusable at this frequency. It is worth noting that a central water pump is located at approximately 500 m northeast of station A00. This pump daily supplies water to the urban area of Argostoli. Although the spectral content of the noise generated by this water pump has not yet been studied, it might be a good candidate to generate the observed 3.8-Hz signal. The effect of this 3.8-Hz peak being limited to a very narrow frequency band, it does not prevent us from discussing the HVSRc at other frequencies, especially changes implied by the HVSRc procedure on the main HVSR peak.

4.1 Temporal stability

Short-term variations on HVSR measurements usually related to changes in noise source excitation are ordinarily observed, making the amplitude of HVSR unreliable. By expressing the HVSR as a function of the Green's function of the medium, eq. (3) implies that HVSR curves are stable (except in specific cases where the medium experiences specific changes leading to individual variations on spectral characteristics of the three components). Since the coda waves of reconstructed Green's functions are expected to be less affected by sources' signatures than their ballistic part, we

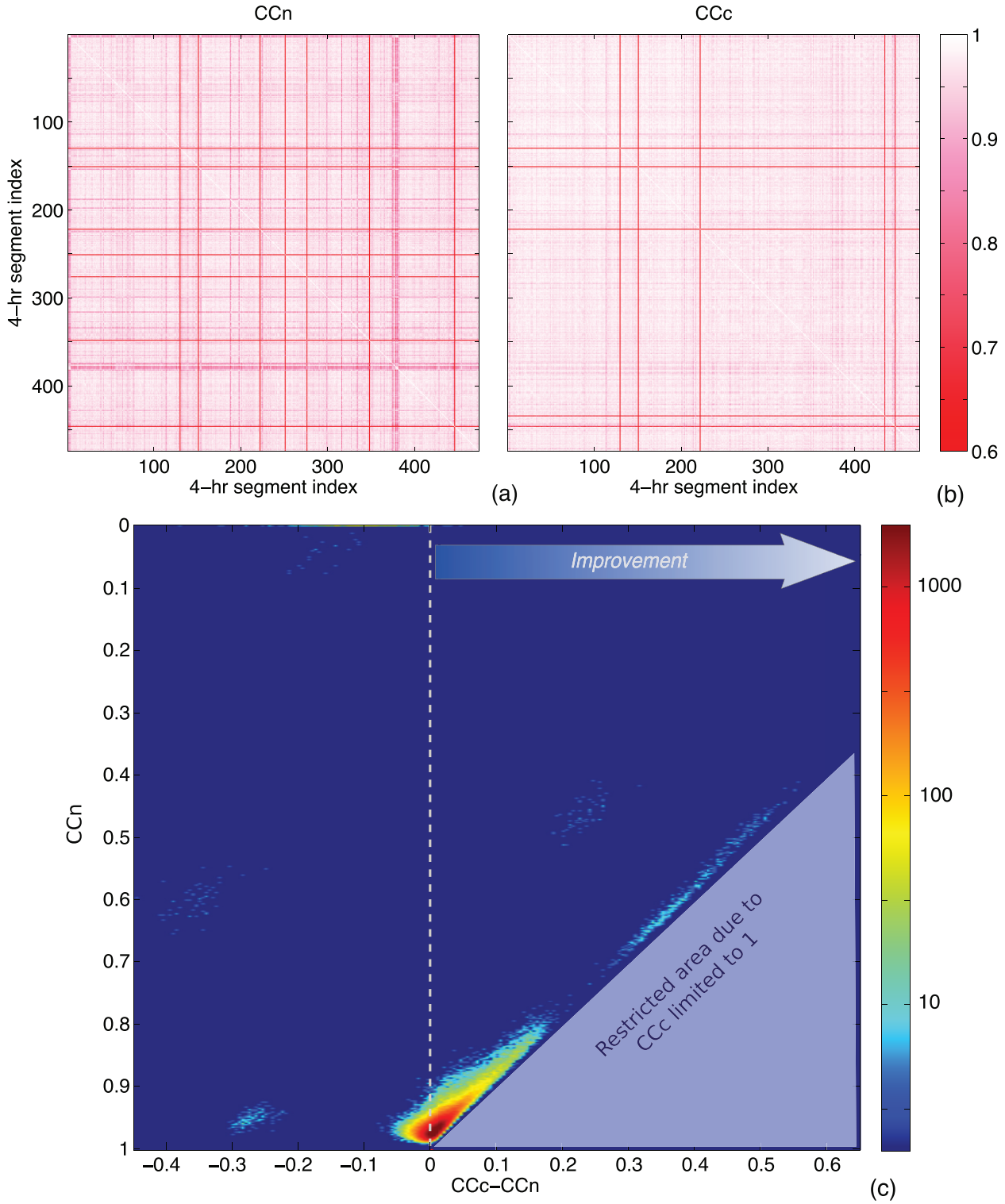

Figure 4. (a and b) Matrices of correlation coefficients (CC) for each pair of single HVSRn (a) or HVSRc (b) measurements shown in Fig. 3. The diagonal elements of each matrix are all equal to 1. (c) Change in CC provided by the coda-based approach (CCc–CCn) relative to the CC for the usual HVSRn procedure (i.e. CCn). The colour represents the number of CC measurements (i.e. pairs of 4-hr segments).

suggest that our procedure could help extracting the physical information from HVSR measurements by getting rid of the source signature that leads to these highly varying HVSR measurements. In turn, an increased temporal stability of HVSR measurements could be seen as an indicator of their better ability to reflect intrinsic characteristics of the medium. To quantify the level of stability from one HVSR measurement to another, we compute the correlation coefficient between every 4-hr-based HVSR measurements.

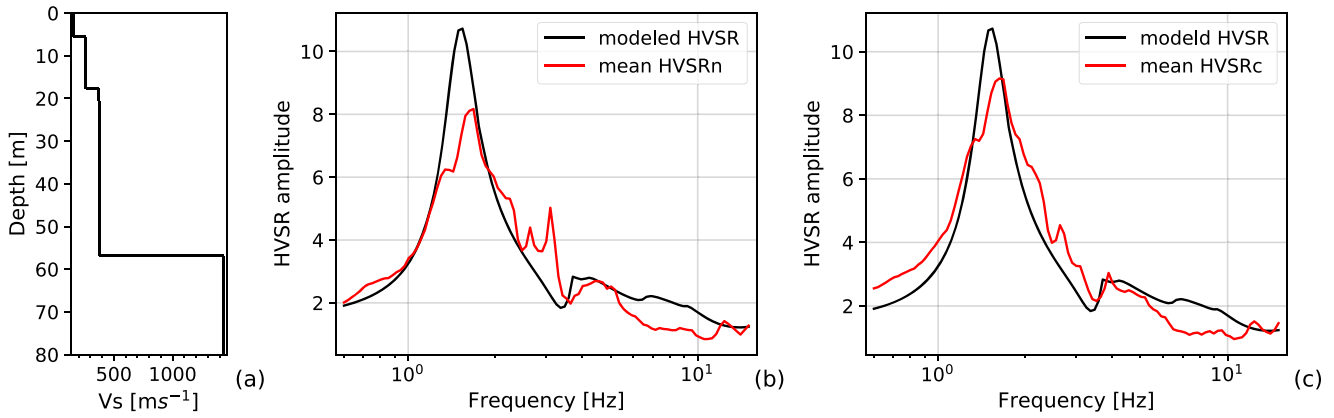

Figure 5. (a) 1-D velocity profile used for the computation of the modelled HVSR; (b) black curve: modelled HVSR curve estimated from forward modelling using the HV-inv code by García-Jerez *et al.* (2016) (see text), red curve: mean HVSRn over all 4-hr segments; (c) black curve : modelled HVSR as in (b), red curve: mean HVSRc over all 4-hr segments.

We focus here on the stability of the main peak that carries the first order information on the medium. To do that, we compute the correlation coefficient (CC) between 0.6 and 3.5 Hz, thereby avoiding the problematic 3.8-Hz frequency. Figs 4(a) and (b) show respectively, the matrices of CC for the two procedures, that is, HVSRn (CCn) and HVSRc (CCc). Lighter colours standing for higher CC, these figures illustrate that individual 4-hr-based HVSRc measurements present a higher similarity, and thus a higher stability in time. Fig. 4(c) further represents this improvement (via the difference in CC associated with each procedure, CCc-CCn). This improvement is represented relative to the CC for the usual HVSRn procedure (i.e. CCn). For this data set, CCn is pretty high for a majority of measurements but we see that the HVSRc procedure clearly improves the stability for CCn $<$ 0.95 (as indicated by CCc-CCn$>$0). For moderate original CC (0.5 $<$ CCn $<$ 0.9) we see that our procedure improves significantly the similarity between each HVSR measurements by reaching CCc close to 1, as delineated by the distribution of our observations.

4.2 Comparison with the theory

In this section, we compare the resulting HVSRc and HVSRn curves averaged over all 4-hr segments, with the theoretical HVSR curve derived from the DFA. Based on the DFA, García-Jerez *et al.* (2016) developed a computer code for the forward calculation and inversion of HVSR for a horizontally layered medium (*HV-Inv*). This *HV-Inv* code was used in this study for the forward computation of what we refer to as the theoretical HVSR for station KEA00. The velocity profile considered in the calculation (Fig. 5a) comes from previous studies carried on the basin (e.g. Imtiaz *et al.* 2018 and Theodoulidis *et al.* 2018b). This 1-D *Vs* model was obtained from array analysis of ambient noise recorded at array A (see Fig. 1), and corresponds to a 4-layered medium over a half-space. In our forward computation, we considered an equipartitioned wavefield composed of surface (Love and Rayleigh) waves and body waves. We considered the contribution of the fundamental mode and the first five higher modes in the calculation of the surface wave part of the Green's function. Note that the velocity model retrieved from array A can model adequately the wavefield between 1 and 20 Hz (Imtiaz *et al.* 2018). This gives an estimate of the validity range of the modelled HVSR curve. Amplitude-wise, the mean HVSRc curve (red curve on Fig. 5c) has a better fit with the modelled HVSR (black curve on Figs 5b and c) compared to the mean HVSRn curve (red curve on Fig. 5b), especially at the peak frequency. Moreover, the mean HVSRn curve also exhibits secondary peaks at around 3 Hz which are due to instabilities in the ambient noise record in certain windows. The attenuation of these peaks on our HVSRc curve reflects the ability of our procedure in handling this kind of outliers. Fig. S4 shows a similar comparison using the median and enables an analysis without the influence of outliers. We computed the root mean square error (RMSE (eq. 11)) to get a first order estimate of the overall agreement between the data (HVSRc and HVSRn) and the model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i^{model} - Y_i^{data})^2}, \tag{11}$$

where Y_i^{model} (resp. Y_i^{data}) corresponds to the amplitude of the modelled HVSR curve (resp. data HVSR curves) at frequency i. n stands for the discrete frequencies of the HVSR curves. The RMSE was computed between 1 and 15 Hz in conformity with the resolvable frequency range of the model. The RMSE computed between the HVSRc curve and the model is lower (0.6) than the RMSE computed using the HVSRn curve (0.75). Similarly, the RMSE remains lower for the median HVSRc (0.59) than the median HVSRn (0.88). This confirms that our procedure helps improve the fit with modelled HVSR curve. We must bear in mind that the modelled HVSR must be interpreted cautiously due to uncertainties in the velocity model considered and the hypothesis assumed in theoretical formulation, for example regarding the attenuation in the medium.

5 CONCLUSION

We proposed an alternative procedure to compute the single station horizontal-to-vertical spectral ratios. The idea is to extract a diffuse field from raw ambient noise in order to get closer to the theoretical requirements of the DFA, that directly relates HVSR to intrinsic properties of the medium. Results presented here show that our procedure helps in stabilizing the HVSR measurement, by reducing short-term fluctuations from one date to another. It also leads to a HVSR curve closer to a theoretical estimate derived from

the DFA. These results suggest that this procedure could help in extracting physical information from HVSR measurements and thus lead to an extended use of HVSR measurements to characterize the mechanical properties of the medium. Note that it does not mean that the amplitude of HVSR will give the level of amplification due to the local subsurface geology. Instead, this amplitude will reflect the characteristics of the medium through the partition of energy in the extracted coda. We noted that our processing cannot get rid of the influence of all the noise sources, as revealed by the 3.8-Hz artefact, which probably reflects the signature of the source's autocorrelation, for a long-lasting monochromatic dominant source. This suggests that this approach does not restrict careful analysis throughout the procedure that is anyhow recommended while processing ambient noise recordings, especially at frequencies higher than 1 Hz. These first results on the use of coda of noise correlations to compute HVSR are encouraging. The method will have to be tested on other configurations and data sets to confirm its interest and generalize its application. This approach can help extend the utility of HVSR to characterize the medium beyond the estimate of the soil's fundamental resonance frequency, especially in constraining the mechanical properties of the medium beneath the station (e.g. the 1-D velocity profile). In particular, HVSR curves are sensitive to the existence and depth of sharp velocity contrasts, and thus can complement the information provided by surface waves dispersion curves in a joint inversion. We could also imagine to use HVSR to monitor temporal variations of the medium properties since this procedure may help cleaning up measurements from unwanted variations due to the noise origin.

ACKNOWLEDGEMENTS

The seismic data used in this study was downloaded from the data centre of the French Seismologic and Geodetic Network (RESIF, http://www.resif.fr/). The HV-Inv tool is available upon request at: https://w3.ual.es/GruposInv/hv-inv/. We thank Cécile Cornou for the communication of the Vs profile used in Section 4.2. We thank Edward Cushing for the geological map used in Fig. 1. We thank Céline Gélis, Cécile Cornou, Sylvette Bonnefoy-Claudet, Edward Cushing, Laurent Stehly for helpful discussions during the completion of this work. We thank the editor Dr Margarita Segou and two reviewers, José Piña-Flores and Nikos Theodoulidis for their helpful comments to improve the manuscript.

Michel Campillo and Ludovic Margerin acknowledge support by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 742335, F-IMAGE).

REFERENCES

Arai, H. & Tokimatsu, K., 2004. S-wave profiling by inversion of microtremor H/V spectrum, *Bull. seism. Soc. Am.,* **94,** 53–63.

Bonnefoy-Claudet, S., Cornou, C., Bard, P.-Y., Cotton, F., Moczo, P., Kristek, J. & Fäh, D., 2006. H/V ratio: a tool for site effects evaluation. Results from 1D noise simulations, *Geophys. J. Int.,* **167,** 827–837.

Cornou, C. & Bard, P.-Y., 2018. Seismic network 4C: French and Greek part of the NERA-JRA1 experiment (RESIF-SISMOB, ITSAK, GFZ, INGV). RESIF - Réseau Sismologique et géodésique Français, Seismic Network.

Cultrera, G., 2014. The Argostoli (Cephalonia, Greece) Experiment, in *Proceedings of the 2nd European Conference on Earthquake Engineering and Seismology,* Istanbul, Turkey, doi:10.13140/2.1.2353.0888.

Cushing, E. *et al.*, 2019. Building of a 3D model of the active Plio-Quaternary Basin of Argostoli (Cephalonia Island, Greece): an integrated geophysical and geological approach, *Eng Geol,* **265,** doi:10.1016/j.enggeo.2019.105441.

Fäh, D., 1997. Microzonation of the city of Basel, *J. Seism.,* **1,** 87–102.

Froment, B., 2011. Utilisation du bruit sismique ambiant dans le suivi temporel de structures géologiques, Sciences de la Terre. Université de Grenoble, 2011. Français. ffNNT: 2011GRENU035ff.fftel-00648796.

Froment, B., Campillo, M. & Roux, P., 2011. Reconstructing the Green's function through iteration of correlations, *Compt. Rendus Geosci.,* **343,** 623–632.

García-Jerez, A., Piña-Flores, J., Sánchez-Sesma, F., Luzón, F. & Perton, M., 2016. A computer code for forward calculation and inversion of the H/V spectral ratio under the diffuse field assumption, *Comp. Geosci.,* **97,** 67–78.

Hillers, G., Campillo, M., Lin, Y., Ma, K. & Roux, P., 2012. Anatomy of the high-frequency ambient seismic wave field at the TCDP borehole, *J. geophys. Res.,* **117,** B06301.

Imtiaz, A., 2015. Seimic wave field, spatial variability and coherency of ground motion over short distances: near source and alluvial valley effects, *PhD thesis*, Université Joseph Fourier, Grenoble.

Imtiaz, A., Perron, V., Hollender, F., Bard, P.-Y., Cornou, C., Svay, A. & Theodoulidis, N., 2018. Wavefield characteristics and spatial incoherency: a comparative study from Argostoli Rock- and Soil-site dense seismic arrays, *Bull. seism. Soc. Am.,* **108**(5A), 2839–2853.

Koller, M., Chatelain, J.-L., Guillier, B., Duval, A.-M., Atakan, K., Lacave, C. & Bard, P.-Y., SESAME-Team, 2004. Practical user guidelines and software for the implementation of the H/V ratio technique: measuring conditions, processing methods and results interpretations, in *Proceedings of the 13th World Conference on Earthquake Engineering,* Vancouver, Canada, V ol. Paper 3132.

Konno, K. & Ohmachi, T., 1998. Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. seism. Soc. Am.,* **88(1)**, 228–241.

Lachet, C. & Bard, P.-Y., 1994. Numerical and theoretical investigations on the possibilities and limitations of Nakamura's technique, *J. Phys. Earth,* **42**(5), 377–397.

Lermo, J. & Chavez-Garcia, F., 1993. Site effect evaluation using spectral ratio with only one station, *Bull. seism. Soc. Am,* **83**(5), 1574–1594.

Lobkis, O. & Weaver, R., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.,* **110,** 3011–3017.

Lunedei, E. & Malischewsky, E., 2015. A review and some new issues on the theory of the H/V technique for ambient vibrations, in *Perspectives on European Earthquake Engineering and Seismology,* Vol. **39,** pp. 371–394, Spinger.

Margerin, L., 2009. Generalized eigenfunctions of layered elastic media and application to diffuse fields, . *acoust. Soc. Am.,* **125**(1), 164–174.

Margerin, L., Campillo, M., Van Tiggelen, B. & Hennino, R., 2009. Energy partition of seismic coda waves in layered media: theory and application to Pinyon Flats Observatory, *Geophys. J. Int.,* **177,** 571–585.

Mordret, A., Landès, M., Shapiro, N., Singh, S., Roux, P. & Barkved, O., 2013. Near-Surface study at the Valhall oil field from ambient noise surface wave tomography, *Geophys. J. Int.,* **193,** 1627–1643.

Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface using microtremor on ground surface, *Railway Technical Research Institute, Quarterly Reports,* **30**(1), 25–33.

Nakamura, Y., 2000. Clear identification of fundamental idea of Nakamura's technique and its applications, in *Proceedings of the 12th World Conference on Earthquake Engineering,* **24,** 25–30.

Nogoshi, M. & Igarashi, T., 1971. On the amplitude characteristics of microtremor (part 2), *J. Seismol. Soc. Japan,* **24** (1), 26–40.

Perron, V., Gélis, C., Froment, B., Hollender, F., Bard, P.-Y., Cultrera, G. & Cushing, E., 2018a. Can broad-band earthquake site responses be preicted by the ambient noise spectral ratio? Insight from observations at two sedimentary basins, *Geophys. J. Int.,* **215,** 1442–1454.

Perron, V. *et al.*, 2018b. Accelerometer, Velocimeter dense-array, and rotation sensor datasets from the Sinaps@ Postseismic Survey (Cephalonia 2014-2015 Aftershock Sequence), *Seismol. Res. Lett.,* **89**(2), 678–687.

Perton, M., Sánchez-Sesma, F., Rodríguez-Castellanos, A., Campillo, M. & Weaver, R., 2009. Two perspectives on equipartition in diffuse elastic fields in three dimensions, *Acoust. Soc. Am.,* **126**(3), 1125–1130.

Piña-Flores, J., Perton, M., García-Jerez, A., Carmona, E., Luzón, F., Molina-Villegas, J. & Sánchez-Sesma, F., 2017. The inversion of spectral ratio H/V in a layered system using the diffuse field assumption (DFA), *Geophys. J. Int.,* **208,** 577–588.

Sánchez-Sesma, F. & Campillo, M., 2006. Retrieval of the Green's function from cross correlation: the canonical elastic problem, *Bull. seism. Soc. Am.,* **96,** 1182–1191.

Sánchez-Sesma, F., Perez-Ruiz, J., Luzon, F., Campillo, M. & Rodríguez-Castellanos, A., 2008. Diffuse fields in dynamic elasticity, *Wave Motion,* **45,** 641–654.

Sánchez-Sesma, F. *et al.*, 2011. A theory for microtremor H/V spectral ratio: application for a layered medium, *Geophys. J. Int.,* **186**(1), 221–225.

Sánchez Sesma, F., Piña, J., Campillo, M., Luzón, F., García Jerez, A., Albarello, D. & Lenedei, E., 2012. Seismic ambient noise H/V spectral ratio using the ACA (autocorrelations of coda of autocorrelations) approach, in *AGU Fall Meeting Abstracts,* S52C–04.

Spica, Z. *et al.*, 2015. Velocity models and site effects at Kawah Ljen Volcano and Ljen Caldera (Indonesia) determined from ambient noise cross-correlations and directional energy density spectral ratios, *J. Volc. Geotherm. Res.,* **302,** 173–189.

Stehly, L., Campillo, M. & Shapiro, N., 2006. A study of the seismic noise from its long-range correlation properties, *J. geophys. Res.,* **111,** B10306.

Stehly, L., Campillo, M., Froment, B. & Weaver, R., 2008. Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise, *J. geophys. Res.,* **113**(B11), doi:10.1029/2008JB005693.

Theodoulidis, N. *et al.*, 2018a. Basin effects on ground motion: the case of a high-resolution experiment in Cephalonia (Greece), *Bull. Earthq. Eng.,* **16**(2), 529–560.

Theodoulidis, N. *et al.*, 2018b. The ARGONET (Greece) seismic observatory: an accelerometric vertical array and its data, *Seismol. Res. Lett.,* **89**(4), doi:10.1785/0220180042.

Weaver, R., 2010. Equipartition and retrieval of Green's function, *Earthq. Sci.,* **23,** 397–402.

Weaver, R., Froment, B. & Campillo, M., 2009. On the correlation of non-isotropically distributed ballistic scalar diffuse waves, *J. acoust. Soc. Am.,* **126**(4), doi:10.1121/1.3203359.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Numerical workflow for HVSRn and HVSRc computation on a 4-hr noise signal.

Figure S2. (a) Northern component of the noise recorded between 0 and 4 hr on 21 March 2012. This segment has been chosen since the 3.8-Hz peak is particularly pronounced on the resulting HVSR. Coloured windows show the 15-min segments. N-component spectra of grey windows do not show a pronounced peak compared to the spectra of coloured ones. This figure highlights that there is no clear evidence in the unfiltered temporal signal allowing us to perform a selection to remove the windows affected by the dominant monochromatic source at 3.8 Hz. The red box around the second window points to the signal used in (b) and Fig. S3. (b) Three-component Fourier spectra of the 15-min segment indicated by the red box in (a). A narrow peak is observed just below 4 Hz on the horizontal components, and is particularly pronounced on the northern component.

Figure S3. (a) HVSRc computed using a N-oriented virtual source (HVSRcN) on the 15-minute segment indicated by the red box in Fig. S2(a). This shows that the 3.8-Hz peak dominates the HVSR for this segment. (b–d) NN component of the noise correlation (C_NN) computed on the same 15-min segment. Panel (b) shows the unfiltered correlation signal, (c) and (d) show the correlations filtered between 1 and 2 Hz (i.e. around the resonance frequency), and between 2.5 and 5 Hz respectively. Panel (c) indicates that the noise correlation filtered around the 3.8-Hz peak is dominated by the autocorrelation of this self-entertained monochromatic source. This prevents assuming the emergence of the Green's function in the correlation signal at this frequency.

Figure S4. Same as Fig. 5 using the median HVSR rather than the mean.

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.