A new autoregressive moving average modeling of H/V spectral ratios to
estimate the ground resonance frequency

Arantza Ugalde⁽¹⁾, Juan José Egozcue⁽²⁾, César R. Ranero^(1,3)

⁽¹⁾ Institute of Marine Sciences - CSIC, Pg. Marítim de la Barceloneta 37-49, 08003
Barcelona, Spain. E-mail: a.ugalde@icm.csic.es

⁽²⁾ Technical University of Catalonia, Dept. Civil and Environmental Engineering, Jordi
Girona, 1-3, 08034 Barcelona, Spain. E-mail: juan.jose.egozcue@upc.edu

⁽³⁾ Institució Catalana de Recerca i Estudis Avançats, ICREA, Passeig de Lluís
Companys, 23, 08010 Barcelona, Spain. E-mail: cranero@cmima.csic.es

13 Highlights

- 14 • A new method to estimate the H/V spectral ratio using microtremors
- 15 • An Autoregressive Moving Average filter models the H/V transfer function
- 16 • The ARMA approach is a robust method for estimating soil resonance frequencies
- 17 • Engineering bedrock can be more precisely estimated using the ARMA method

18

19 Abstract

20 We propose a new method to estimate the horizontal-to-vertical (H/V) spectral ratio
21 using microtremor measurements. The technique is based on modeling the H/V transfer
22 function by means of an AutoRegressive Moving Average (ARMA) filter. As compared
23 with the conventional, Fourier-based spectra processing routines, this method is
24 efficient in extracting the fundamental resonant frequency with a higher spectral
25 resolution. We demonstrate its usefulness by an application to several sites in
26 northeastern Iberian Peninsula. For the data sets examined, our approach shows
27 improvement over the Fourier transform-based method under adverse experimental
28 conditions. We conclude that the new technique provides robust estimates of the
29 fundamental resonance frequency and can be effective to characterize empirically the
30 unconsolidated sediment thickness and bedrock geometry.

31 Keywords

32 Microtremor, H/V spectral ratio, ARMA model

33 1. Introduction

34 The horizontal-to-vertical (H/V) spectral ratio analysis using the background
35 microtremor (ambient-vibration) wavefield (Nakamura, 1989; 2019) is a widely used
36 technique to estimate the fundamental resonance frequency of the ground. This
37 parameter is used in seismic microzonation studies to map areas that exhibit local site
38 effects, or amplification of seismic ground motions due to the surface geology.
39 Microtremor wavefield results from a combination of natural (sea waves, wind, rainfall)
40 and anthropogenic (traffic, industrial activities) sources that show temporal and spatial
41 changes. Long period (greater than 2 s) ambient-vibration signals, also called
42 microseisms, are associated with ocean waves at deep ocean and coastal regions, and
43 propagate as Rg and Lg phases over continental paths. Short-period microtremors
44 originate mainly from human activities and consist of body (P and S) and surface
45 (Rayleigh and Love) waves (e.g., Okada, 2003). The ground motion amplitude spectrum
46 at a site will depend on the vibration sources and on the thickness and physical
47 properties of the uppermost soil layers, which act like a filter that may increase the
48 duration and amplitude of the recorded signal in certain frequencies (e.g., Molnar and
49 Cassidy, 2006).

50 The H/V method involves computing the ratio between the amplitude spectrum of the
51 horizontal and vertical components of microtremor measurements at a single seismic
52 station. The main peak of the H/V curve has proven, empirically, to provide the
53 fundamental horizontal-component shear-wave (SH) resonance frequency (f_0) of soft
54 sedimentary layers (e.g., Bard, 2008). As for the amplification level, it may depend on
55 the local geological conditions and is used as a lower-bound estimate of earthquake site
56 response for engineering purposes (e.g., Lermo and Chávez-García, 1994; Molnar and
57 Cassidy, 2006). The spectral ratio is sensitive to strong acoustic impedance contrasts

between the bedrock and the overlying sediments and can be used to determine the geotechnical parameters (depth to bedrock and shear-wave velocities) of the subsurface layers. Considering a simple model of one layer over a half-space, an impedance contrast is theoretically expected at

$$h = \beta / 4f_0, \quad (1)$$

where β is the average S -wave velocity of the soil layer (Haskell, 1960; Ibs-von Seht and Wohlenberg, 1999). The H/V spectral ratio curve has been modeled theoretically based on the diffuse field approach that takes into account the full ambient-vibration wavefield (García-Jerez et al., 2016; Sánchez-Sesma, 2017), and by Rayleigh wave ellipticity. One reason of the lack of agreement regarding the correct physical model of the H/V spectral ratio is that the composition of the noise wavefield at the H/V peak frequencies is not known (see Lunedei and Malischewsky, 2015, for a review). All the theories proposed are able to reproduce the main peak frequency of the H/V spectral ratio; however, the soil profile inverted from experimental data depends on the modeling of the whole H/V spectral ratio curve. Within this theoretical framework, the H/V spectral ratios have been inverted to estimate the subsurface shear wave velocity structure. The estimated velocity models can be improved by combining information of other geophysical and geotechnical data during the inversion (e.g. Piña-Flores et al., 2017; Rosa-Cintas et al., 2017; Pertou et al., 2018).

Due to its lower cost, environmental friendliness and relative ease of computing with respect to other geotechnical and geophysical methods, the H/V technique has been used extensively for geologic and engineering studies (e.g., Leyton et al., 2013; Molnar et al., 2018; Famiani et al., 2020; Molnar et al., 2020; Rohmer et al., 2020). Motivated by the increasing use of the H/V method, standard criteria for data acquisition, processing and interpretation were provided in the frame of the European project SESAME (Site

EffectS assessment using AMbient Excitations) (SESAME, 2004; Bard, 2008). Although standardized recording and processing routines are used for estimating the H/V ratio, it has been observed that measurement conditions can produce misleading results (e.g., Castellaro and Mulargia, 2009). Moreover, the spectral processing methodologies used may also influence the final estimates (e.g., Diagourtas et al., 2001; Anthony et al., 2020). In this study, we devised a computational approach that overcomes adverse measurement conditions and spectral processing limitations. The method is based on AutoRegressive Moving Average (ARMA) models that are high-resolution spectral analysis methods used to model rational transfer function processes. It is not a physical model of the process but a mathematical approach to enhance the spectral estimation of the H/V ratio. This parametric modeling of data is an alternative to the conventional use of non-parametric spectral techniques based on the fast Fourier Transform (FFT). It improves the inherent performance limitations of the FFT approach, such as frequency resolution and data windowing (Kay and Marple, 1981). We proposed the original idea behind this approach in Ugalde et al. (1998), and it is further developed in this work. Here we describe the modeling used to estimate the fundamental resonance frequency of soils through H/V spectral ratios and illustrate the new method using data from the permanent Catalan Seismic Network (north-eastern Iberian peninsula; ICGC, 2000).

2. The problem

The spectral ratio of the horizontal (H) and vertical (V) components of microtremor recordings at a point $\mathbf{x}$ can be expressed as:

$$\frac{H}{V}(\mathbf{x}, \omega) = \sqrt{\frac{E_1(\mathbf{x}, \omega) + E_2(\mathbf{x}, \omega)}{E_3(\mathbf{x}, \omega)}}, \quad (2)$$

106 where $E_i(\mathbf{x}, \omega)$ represent the spectral energies of the horizontal ($i=1, 2$) and vertical
 107 ($i=3$) components and ω is the angular frequency. According to the standard procedures
 108 (e.g., SESAME, 2004), Eq. (2) is evaluated by computing, and then smoothing, the
 109 Fourier power spectra of the vertical and horizontal recordings over several overlapping
 110 time windows. The H/V ratio is calculated for every single time window and is finally
 111 averaged, so that the source deconvolution is performed in each time window. Instead,
 112 Eq. (2) can be evaluated through first averaging the horizontal and vertical spectra for
 113 all windows and then computing the ratio of the averaged spectra. Using this technique,
 114 the effect of the source deconvolution is balanced over many time windows, which
 115 increases stability (e.g., Pertou et al., 2018; Spica et al., 2018).

116 Finite duration local nonstationarity or isolated outliers (point defects) in the data
 117 segments may affect the power spectra and, thus, the statistical behavior of the spectral
 118 ratios estimates. Some transient nonstationary signals of natural or anthropogenic origin
 119 can be recognized in the recordings as an increase of all the three-component power
 120 spectra at certain frequencies (e.g., Castellaro and Mulargia, 2009). However, other
 121 artifacts can be difficult to detect.

122 Assuming the H/V spectral ratio function in Eq. (2) can be modeled as an ARMA
 123 process, the resonance frequencies of the soil still can be recovered under these
 124 unfavorable conditions. ARMA models are suitable representations of processes
 125 involving rational spectra and have been applied successfully for a long time in many
 126 scientific fields, including geophysics and engineering (e.g., Kozin, 1988; Li and
 127 Kareem, 1990; Glaser, 1996; Akintug and Rasmussen, 2005; Ólafsson and
 128 Sigbjörnsson, 2011).

129 3. The autoregressive moving average model

130 A stationary process affected by additive nonstationary random noise can describe the
 131 horizontal (h) and vertical (v) components of ambient-vibration recordings at a point $\mathbf{x}$.
 132 The model can be written in the time domain as:

$$\begin{aligned} 133 \quad h[n] &= u_h[n] + w_h[n] \\ v[n] &= u_v[n] + w_v[n] \end{aligned} \quad (3)$$

134 where $u_h[n] = u_1[n] + ju_2[n]$ is a complex signal that combines the two orthogonal
 135 horizontal components, $u_3[n]$ is the vertical component signal, $0 \leq n \leq N - 1$, N is the
 136 length of the data record, j is the imaginary unit, and $\mathbf{x}$ is implicit to simplify the
 137 notation. The signals are assumed to be uncorrelated with the additive noise in the
 138 horizontal w_h and vertical w_v components.

139 The horizontal-to-vertical ratio of Eq. (3) can be expressed in polynomial form as:

$$U_h(z) = S(z)U_v(z), \quad (4)$$

140 where z is the shift operator. Mathematically, $S(z)$ can be considered a filter transfer
 141 function between the input $u_v[n]$ and output $u_h[n]$. Linear filters with rational transfer
 142 functions are related to general ARMA models through the parameterization:

$$S(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^q b[k]z^{-k}}{1 + \sum_{k=1}^p a[k]z^{-k}}, \quad (5)$$

143 where $B(z)$ and $A(z)$ are two polynomials having, in this case, complex ($b[k]$) and real
 144 ($a[k]$) coefficients, (p, q) represent the order of the model, $z = e^{j\omega\Delta t}$ and Δt is the
 145 sampling interval. In view of these considerations, we devised that the horizontal-to-
 146 vertical ratio can be actually estimated using ARMA spectral techniques.

147 By parameterizing the horizontal and vertical components of the signal with the same
 148 number of coefficients, we assume that our H/V process can be modeled with a general
 149 ARMA (p, p) model. Among the different approaches proposed to estimate the model

150 coefficients, we use the modified covariance method. It consists of minimizing the
 151 forward and backward prediction errors in the least squares sense and is thought to
 152 enhance the spectral peaks. Considering a forward linear prediction, for which the
 153 estimate at time index n is based on p samples indexed earlier in time, the general
 154 ARMA system equation for the output $u_h[n]$ and input $u_3[n]$ stationary signals is given
 155 by:

$$\sum_{k=0}^p a[k](h[n-k] - w_h[n-k]) = \sum_{k=0}^p b[k](v[n-k] - w_v[n-k]). \quad (6)$$

156 Multiplying each side of Eq. (6) by $v[n-m]$, taking expectations, and after some
 157 algebra, we get the forward prediction error as:

$$e_f[m] = \sum_{k=0}^p -b[k]r_{vv}[m-k] + \sum_{k=0}^p a[k]r_{hv}[m-k]. \quad (7)$$

158 Here $m = -(N-1) + p, \dots, N-1$ and $r_{il}[s]$ are the autocovariance ($i = l$) or
 159 crosscovariance ($i \neq l$) functions at time lag s , estimated as:

$$\begin{aligned} r_{hh}[s] &= \frac{1}{N} \sum_{n=0}^{N-s} h[n+s] h^*[n] \\ r_{vv}[s] &= \frac{1}{N} \sum_{n=0}^{N-s} v[n+s] v[n] \\ r_{hv}[s] &= \frac{1}{N} \sum_{n=0}^{N-s} h[n+s] v[n], \end{aligned} \quad (8)$$

160 where $*$ denotes the complex conjugate and $s = 0, \pm 1, \dots, \pm(N-1)$. The relations
 161 $r_{hh}[-s] = r_{hh}^*[s]$, $r_{vv}[-s] = r_{vv}[s]$ and $r_{hv}[-s] = r_{vh}^*[s]$ yield results for negative
 162 lags.

163 Similarly, we consider the general ARMA system equation for backward linear
 164 prediction, in the sense that the estimate at time index n is based on p samples indexed
 165 later in time. Multiplying by $h^*[n+m]$, taking expectations, and after simplification,
 166 we obtain the backward prediction error as

$$e_b[m] = \sum_{k=0}^p b[k]r_{hv}^*[m-k] - \sum_{k=0}^p a[k]r_{hh}^*[m-k]. \quad (9)$$

Model parameters $a[k]$ (real) and $b[k] = b_1[k] + jb_2[k]$ (complex) can be estimated by minimizing the forward and backward prediction errors:

$$e_{fb} = \sum_{m=-(N-1)+p}^{N-1} \mu |e_f[m]|^2 + \eta |e_b[m]|^2, \quad (10)$$

where μ and η are the error weights. To find the filter coefficients that minimize Eq. (10), we set the partial derivatives of e_{fb} with respect to $a[k]$, $b_1[k]$ and $b_2[k]$ equal to zero for $k = 0, \dots, p$, and $a[0] = 1$. Thus, we obtain a system of $3p + 2$ linear equations for $3p + 2$ unknowns that we solve using the Gauss-Jordan elimination algorithm (see Appendix A). Once the model parameters are estimated, the transfer function $(H/V)(\omega)$ can be computed using Eq. (5).

The critical problem of ARMA modeling is the determination of the best model order p , which would be equivalent to set the amount of smoothing. Too small values of p may cause data underfitting, in which case the obtained spectral ratio is too smooth to allow high spectral resolution. On the contrary, a too high order may cause data overfitting, thus leading to an unstable model. Our compromise criterion between underfit and overfit is to compute the transfer function in Eq. (5) for increasing model order p until the frequency of the H/V peak stabilizes.

182

183 4. The H/V spectral ratio

Because we were interested in testing carefully the new methodology, we focused on high quality seismic data recorded in good seismometer vaults. For this purpose we considered several permanent, broadband seismic stations of the Catalan Seismic

Network (ICGC, 2000). We used 45 minutes of data recorded during quiet-weather, night-time hours on June 30, 2019. In the model-based ARMA analysis the length of the data sequence does not affect the frequency resolution; therefore in this approach we used 512-sample (5.12 s) length, 50% overlapping windows. For each window, we computed the auto and crosscovariance functions of the vertical and horizontal components according to Eq. (8), estimated the model coefficients by solving the system of Eq. (A.3), and computed the transfer function using Eq. (5). Finally, we geometrically averaged the H/V spectral ratios from all time windows. We repeated this procedure for increasing values of p . As a result of the ARMA complex-valued approach, negative and positive frequencies were calculated. According to Eq. (5), negative frequencies represent a rotation in the complex plane. Therefore, the asymmetry of the curves indicates the possible directional dependence of the H/V spectral ratio, because the real and imaginary parts of $b[k]$ parameterize the two horizontal components of motion.

The complex coherence function between the horizontal and vertical components was also computed following the procedure described in Bendat and Piersol (1986), where the autospectral and cross-spectral density functions were estimated according to Lagunas and Gasull (1984). The estimated coherences from all time windows were geometrically averaged. Coherence values can help to interpret the H/V spectral ratio curve.

Since we propose a method based on different data processing, we compared our results with the popular, well-established, open-source Geopsy software package (e.g., Wathelet et al., 2020), based on SESAME's standard approach. We calculated the power spectra using the FFT for Hanning-tapered, 50% overlapping time windows of 80 s and smoothed the power spectra by means of the Konno and Ohmachi (1998)

212 weighting function using a smooth parameter of $2\pi/b = 0.4$, where b is the width of the
 213 smoothing window.

Fig. 1. (a) 50% percentile of the H/V spectral ratio for station CFON (Catalan Seismic Network) using the ARMA model for order $p=60$. The grey area encloses the 25% and 75% percentile regions. The dashed line represents the coherence between the horizontal and vertical components. The frequency of the main peak for negative (f^-) and positive (f^+) frequencies is plotted. (b) H/V spectral ratio using Geopsy. The dashed lines indicate the standard deviation.

214 4.1 CFON site

215 Fig. 1 shows the H/V spectral ratio obtained for station CFON. At this site, a clear peak
 216 emerges at $f_0^{ARMA} = \pm 3.75 \pm 0.10$ Hz that stabilizes for model order $p=60$. The spectral
 217 ratios for positive and negative frequencies show a difference in amplitude of about
 218 13% at f_0 . CFON is located in a mountain area at an elevation of 973 m, which suggests
 219 the influence of topographic or subsurface structure effects on the amplitude at the
 220 resonant frequency (e.g., Kawase et al., 2019; Matsushima et al., 2014). In this case, the
 221 main frequency peaks show low coherence values. Previous geophysical
 222 characterization studies at CFON (available at <https://www.icgc.cat/>; last accessed April

2020) found an upper layer of recent, low compacted soil of quaternary colluvial materials with S -wave velocities $\beta=281$ m/s extending to ~ 10 m depth. An underlying second layer of completely to moderately altered substrate with $\beta=536$ m/s extends to a depth $H\sim 40$ m where overlays bedrock. Using these values, the fundamental frequency of the profile can be estimated by means of the simplified equation (method 7 of Dobry et al., 1976):

$$f = \frac{1}{2\pi} \left(\frac{4 \sum_{i=1}^n \frac{(H - z_m(i))^2}{v(i)^2} h(i)}{\sum_{i=1}^n (x(i) + x(i+1))^2 h(i)} \right)^{1/2} \quad (11)$$

where n is the number of layers, $H - z_m$ is the midpoint depth of layer i , h is the thickness of each layer, and $x(i+1) = x(i) + h(i)(H - z_m(i))/\beta(i)^2$ is the estimated fundamental mode of vibration at the upper boundary of layer i . Eq. (11) gives a value of $f = 3.73$ Hz, which is in accordance with the resonance frequency of 3.75 Hz estimated with the ARMA-based approach.

Fig. 1(b) shows the H/V spectral ratio obtained for station CFON using Geopsy. The computed resonance frequency is $f_0^{FFT} = 3.87 \pm 0.13$ Hz. We note that, taking into account the error intervals, results of the ARMA approach and the FFT-based method are in agreement for this site.

4.2 CORI site

A comparison between ARMA and the FFT-based method performance is also presented in Fig. 2 for CORI seismic station. No resonance frequency can be identified

in this case with either method, which agrees with the installation of CORI on class A ground type site (bedrock outcrop) according to Eurocode 8 (EC8, 2004).

Fig. 2. (a) H/V spectral ratio for station CORI using the ARMA model for order $p=60$. The grey area encloses the 25% and 75% percentile regions. (b) H/V spectral ratio for station CORI using Geopsy. The dashed lines indicate the standard deviation.

4.3 ICJA site

To further assess the ARMA model performance, we analyzed data recorded at station ICJA, which is deployed in an urban area at the Institute of Earth Sciences Jaume Almera in the city of Barcelona. Fig. 3(a) displays the spectral ratio curve obtained for model order $p=100$, which presents the maximum amplification at $f_0^{ARMA}=\pm 2.66$ Hz. The coherence is low for the main peak. A very similar value of $f_0^{FFT}=2.69$ Hz is obtained using the FFT-based method (Fig. 3b). Considering a mean shear-wave velocity of about 422 m/s from previous studies in this zone (Cadet et al., 2011) using Eq. (1) we infer a theoretical sediment thickness of about 40 m at this site. This result is in agreement with the lithology found at the adjacent Almera-1 borehole (Jurado and Salvany, 2016) that sampled Quaternary sediment to a depth of 41 m overlying Paleozoic basement rocks.

Fig. 3. (a) H/V spectral ratio for station ICJA using the ARMA model for order $p=100$. The grey area encloses the 25% and 75% percentile regions. (b) H/V spectral ratio for station ICJA using Geopsy. The dashed lines in indicate the standard deviation.

257 4.4 CARA site

258 We found that the ARMA approach and the FFT-based method yield significantly
 259 different results for particular datasets. Fig. 4(a, b) compares the resonance frequency
 260 obtained using data from CARA borehole seismic station, which is deployed at 27
 261 meters depth on class A ground. Using data recorded during night-time hours, the
 262 ARMA method detects a main peak frequency $f_0^{ARMA} = \pm 12.51 \pm 0.04$ Hz that stabilizes
 263 for model order $p=70$; however, no significant peak can be identified using Geopsy for
 264 the same data.

Fig. 4. H/V spectral ratio for station CARA using the ARMA model (a, c) and Geopsy (b, d). The grey area in the left column encloses the 25% and 75% percentile regions. The dashed lines in the right column indicate the standard deviation. Results for night-time (a, b) and day-time hours (c, d) are plotted. See the text for discussion.

265 Nevertheless, using data recorded during day-time hours, the FFT-based method yields
 266 a main peak frequency $f_0^{FFT}=11.92\pm1.02$ Hz (Fig. 4d). For the same day-time data, the
 267 ARMA results stabilize earlier (model order $p=30$) and the frequency pattern is
 268 unaltered (Fig. 4c); however, the H/V spectral ratio curve presents higher (more than 4
 269 times) amplification levels than for night-time data. An increase of the coherence is not
 270 observed. Energy directionality of the main peak is also noted as an increase of about
 271 17% in the peak amplitude between negative and positive frequencies (Fig. 4c).

Fig. 5. Vertical (Z), East-West (E) and North-South (N) power spectral density (PSD) at station CARA for day-time hours in dB related to $(\text{m/s}^2)^2/\text{Hz}$ before (a) and after (b) notch-filtering with a quality factor Q (centre frequency/bandwidth) of 150.

Fig. 6. (a) H/V spectral ratio for station CARA using the ARMA model for order $p=60$. The grey area in (a) encloses the 25% and 75% percentile regions. (b) H/V spectral ratio for station CARA using Geopsy. The dashed lines indicate the standard deviation. Day-time, notch-filtered data at center frequencies 10.0 Hz and 12.5 Hz are used.

Fig. 7. Azimuth variation of the H/V spectral ratio for station CARA using Geopsy.

272 To test the possible influence of high-frequency noise sources on these observations, we
 273 computed the power spectral density of ambient noise at station CARA (Fig. 5a). Two
 274 persistent, narrow spectral peaks are clearly recognized at 10 Hz and 12.5 Hz, which are
 275 compatible with continuous quasi-harmonic signals of industrial origin (e.g.,
 276 Antonovskaya et al., 2019). CARA seismic station is surrounded by several
 277 hydroelectric power facilities. In particular, Vielha hydropower station is located at
 278 about 1.8 km from CARA site. This facility is equipped with water turbines rotating at
 279 750 rpm that explain the observed spectral peak at 12.5 Hz. The spectral peak at 10 Hz
 280 may be generated by six hydroelectric facilities with turbine rotation velocities of 600
 281 rpm located at distances less than 7.5 km from CARA. These observations indicate that
 282 the main H/V spectral ratio peak observed is induced by the 12.5 Hz frequency
 283 generated at the hydroelectric power station. To filter out the signals excited by the
 284 industrial facilities we applied a notch filter to the data and recomputed the H/V spectral
 285 ratio. The resulting power spectral density is plotted in Fig. 5 (b). Fig. 6 shows that no
 286 spectral peak of stratigraphic origin can be inferred from the ARMA model at CARA

287 site using the notch-filtered data. But a frequency peak still emerges at $f_0^{FFT}=10.44$ Hz
 288 using the method based on the FFT. Checking H/V in all directions of the horizontal
 289 plane, we note a strong azimuthal variation that presents a maximum around 18° ($+180^\circ$)
 290 (Fig. 7). These results will be discussed later.

Fig. 8. (a) H/V spectral ratio for station ARBS using ARMA for model order $p=50$. The grey area encloses the 25% and 75% percentile regions. (b) H/V spectral ratio for station ARBS using Geopsy. The dashed lines indicate the standard deviation.

291 4.5 ARBS site

292 We also applied the ARMA method to data recorded at ARBS, which is a 30-meters-
 293 depth borehole seismic station deployed on rock at an elevation of 2,145 m in the
 294 Pyrenean range. Fig. 8(a) shows coherence values around 0.5 and a relatively flat
 295 spectral ratio curve, which may indicate a lack of high acoustic impedance at depth.
 296 However, a peak at $f_0^{FFT}=2.25\pm0.34$ Hz emerges using the FFT-based method for the
 297 same data set (Fig. 8b). The frequency of the main peak changes during the day in the
 298 range 2-3 Hz, having the minimum values at night and the maximum ones during the 8-
 299 15 h time interval. These observations support a relationship of the main peak with
 300 anthropogenic noise sources, given that ARBS is located at ~ 8 km distance from

Andorra la Vella (capital of the Principality of Andorra). Moreover, by analyzing the H/V variation as a function of azimuth we note that the spectral ratio presents the maximum values around 69° (Fig. 9). The directivity observed may be related to the direction of the noise sources. However, ARBS is deployed at the crest of a ridge, and the obtained preferential direction N69E corresponds to the direction of the largest slope, that presents maximum height drops of $>60\%$. Therefore, if the H/V spectral ratio peak observed with the FFT-based method in this site was interpreted as the resonance frequency of the ground, the directionality might be related to the effects of surface topography that is generally observed to cause larger amplification on the horizontal components preferentially along the direction perpendicular to the ridge axis (e.g., Chávez-García et al., 1996; Massa et al., 2014).

Fig. 9. Azimuth variation of the H/V spectral ratio for station ARBS using Geopsy.

5. Discussion and conclusions

5.1 The spectral processing

The H/V spectral ratio of microtremor recordings is conventionally computed by means of non-parametric spectral techniques based on the Fourier Transform. Several parameters actually used for the analysis such as the size of the data windows, percentage of overlap and the level of smoothing, may bias the spectral estimates and lead to misleading results. The smoothing process reduces the variance of the spectral estimates, at the cost of compromising the frequency resolution. Moreover, the spectral estimates contain the contribution of the properties from several neighboring frequencies over the interval of smoothing. These amplitudes may be discordant, thus distorting features of interest. Therefore, considerable care must be exercised in choosing the right parameters depending on the nature of the input data (e.g., Diagourtas et al., 2001; Anthony et al., 2020). For this reason, previous experience is often required when applying these techniques for computing and interpreting the H/V spectral ratio.

Here we proposed a new technique that is expected to overcome the above-mentioned difficulties. It is a parametric technique based on modeling the H/V time series by an ARMA filter in which the system is represented by a polynomial transfer function expressed in terms of the model parameters. The method for estimating the spectral ratio consists of inverting for the optimum filter coefficients that model the process. The strongest contribution of this new method is its robustness for estimating the natural frequency of the ground with high spectral resolution. Other intrinsic advantages of the method are its applicability to short data lengths and the avoidance of window-biasing effects. The technique is less of an art than the FFT-based method and it can be

automated. However, it requires an extensive investigation of the model that represents the data through the determination of the appropriate order of the model p . The computational complexity of the method depends on the model order. When p is high, then the computational complexity will be large compared to the FFT-based methods. Several criteria and empirical methods have been proposed over the years to determine the model order, but they do not always work in practical cases (including our data). Here we devised an empirical approach adapted to our problem, which consists in running the code for increasing values of p until the H/V spectral ratio stabilizes. The algorithm performance is mostly affected by the empirical procedure chosen to determine the model order; however, high performance computing may speed up considerably this process.

5.2 ARMA approach performance

We tested the new technique with seismic data acquired by several permanent stations of the Catalan seismic network and compared the obtained results with those computed using the widely recognized Geopsy software based on classical spectral analysis methods. The predominant frequencies obtained with the ARMA approach and Geopsy show good correspondence for measurements at CFON, CORI and ICJA sites. CFON station shows one distinctive resonant frequency at 3.75 Hz that is related to a lithological interface at a depth of 30 m. CORI presents a relatively flat spectral ratio curve as theoretically expected at bedrock sites. ICJA site shows a peak of stratigraphic origin at 2.66 Hz that is related to a high impedance contrast at a depth of 40 m. Conversely, results for CARA and ARBS present some discrepancies caused by the processing methodologies due to some singular features of the underlying seismic data. At CARA site, two predominant frequencies at 10 Hz and 12.5 Hz are clearly identified using the ARMA approach. They are caused by the activity at the existing hydroelectric

facilities close to the site that generates continuous quasi-harmonic mechanical oscillations (Antonovskaya et al., 2019). The main peak at 12.5 Hz is caused by the closest facility, having water turbines rotating at 750 rpm and located at 1.8 km distance from the seismic station. The rest of the closest hydropower facilities in the area (six), which are located at straight-line distances of less than 7.5 km from the site, have turbine rotation velocities of 600 rpm and are responsible for the spectral peak around 10 Hz. The order of the ARMA model defines the complexity of the signal (i.e., the number of harmonic components that can be modeled). For this reason, the high signal-to-noise ratio signals at CARA for day-time hours can be modeled with a lower number of parameters than signals recorded during night-time hours, which require a more complex model to enhance the spectral peaks. Whereas the exact frequencies of the disturbances emerge with the ARMA technique no matter at what time the measurements are taken, the FFT-based method fails to detect the peaks using data from night-time hours. However, two broad H/V peaks emerge around these frequencies when the spectral ratio is computed using day-time data (when the demand of electricity is high and so is the signal-to-noise ratio). Although the spectral peaks detected are not related to stratigraphic features, this is an example of the superior performance of the ARMA approach to extract predominant frequencies from stationary data. After processing the data with a notch filter, no resonance frequency peak can be identified in the resulting spectral ratio using the ARMA technique, but the FFT-based method still detects a broad peak around 10.44 Hz, which presents a strong directivity with a maximum around 18° ($+180^\circ$) (Fig. 7). If we interpret this frequency as the resonance frequency of the ground, the directivity might be connected to the topography of the site, as CARA is deployed at a mountain ridge, and the direction 198°N points to the direction of the maximum slope, having an average height drop of $\sim 46\%$. This would

indicate that surface topographic effects might affect the results in this site, causing large amplifications on the horizontal components preferentially along the direction perpendicular to the ridge axis, as it has been observed before (e.g., Massa et al., 2014). However, if the frequency peak observed at CARA is related to anthropogenic noise, the directivity observed could be related to the direction of the noise sources. The fact that the ARMA approach does not identify the spectral peak favors this second interpretation. The same behavior is observed at ARBS. In this case, the ARMA approach yields a near-flat spectral ratio curve, whereas the FFT-based method detects a peak at 2.25 Hz. The predominant frequency at ARBS changes during the day, which suggests an anthropogenic origin of the noise sources. Also, it presents the maximum H/V amplification at an azimuth of 69° , which again may corresponds to the direction of the largest average slope or to the direction of the noise sources. More detailed analyses would be required at CARA and CARBS to reach a conclusion.

In conclusion, the proposed ARMA method is a powerful algorithm to estimate the resonance frequency of soils using microtremor records. We showed that it yields more robust results with a higher spectral resolution than the classical methods,. Because of the complex-valued approach, the asymmetry of the curves between negative and positive frequencies indicates the possible directional dependence of noise sources. The empirical interpretation of the H/V spectral peaks obtained with the ARMA approach in terms of the depth of the impedance contrast agrees well with the available geophysical and geotechnical information. The proposed method may be useful to estimate with improved precision the depth of the engineering bedrock, which can be used to optimize the distribution of geotechnical campaigns in engineering projects.

413 Acknowledgements

414 This research was partially supported by Euskontrol, S. A., through the ISURI project.
415 The activity of J. J. Egozcue was supported by the project METHods for COMpositional
416 analysis of DATA (CODAMET), Ministerio de Ciencia, Innovación y Universidades
417 (Ref: RTI2018-095518-B-C22, 2019-2021). This is a contribution of the Barcelona
418 Center for Subsurface Imaging that is a Grup de Recerca de la Generalitat de Catalunya.
419 Geopsy package can be downloaded from <http://www.geopsy.org/> (last accessed April
420 2020). Results of the ARMA method were plotted using the Generic Mapping Tools
421 (GMT) software (Wessel et al., 2013).

422

423 Data availability

424 Seismic data (waveforms and instrument response) have been obtained through the
425 Federation of Digital Seismograph Networks (FDSN) web service available at
426 <https://www.icgc.cat> (last accessed April 2020). ARBS station is deployed at La
427 Rabassa, Andorra, in collaboration between the Institut Cartogràfic i Geològic de
428 Catalunya (ICGC) and the Institut d'Estudis Andorrans (IEA). ICJA station belongs to
429 the Institute of Earth Sciences Jaume Almera (<http://labsis.ictja.csic.es/edtarray>, last
430 accessed April 2020).

431

432 Appendix A

433 To minimize the forward and backward prediction errors in Eq. (10), we set:

$$\frac{\partial e_{fb}}{\partial a[l]} = 0 \ (l \neq 0), a[0] = 1; \frac{\partial e_{fb}}{\partial b_1[l]} = 0; \frac{\partial e_{fb}}{\partial b_2[l]} = 0 \quad (\text{A.1})$$

434 This leads to the set of equations:

$$\begin{aligned}
\frac{\partial e_{fb}}{\partial a[l]} &= \sum_k \sum_m a[k] (r_{hv}[m-l]r_{hv}^*[m-k] + r_{hh}[m-l]r_{hh}^*[m-k]) \\
&\quad - b_1[k] (r_{hv}[m-l]r_{vv}[m-k] + r_{hh}[m-l]r_{hv}^*[m-k]) \\
&\quad - j b_2[k] (r_{hv}[m-l]r_{vv}[m-k] + r_{hh}[m-l]r_{hv}^*[m-k]) = 0 \\
\frac{\partial e_{fb}}{\partial b_1[l]} &= \sum_k \sum_m -a[k] (r_{vv}[m-l]r_{hv}[m-k] + r_{hv}[m-l]r_{hh}^*[m-k]) \\
&\quad + b_1[k] (r_{vv}[m-l]r_{vv}[m-k] + r_{hv}[m-l]r_{hv}^*[m-k]) \\
&\quad + j b_2[k] (r_{hv}[m-l]r_{hv}^*[m-k]) = 0 \\
\frac{\partial e_{fb}}{\partial b_2[l]} &= \sum_k \sum_m -j a[k] (r_{vv}[m-l]r_{hv}[m-k] + r_{hv}[m-l]r_{hh}^*[m-k]) \\
&\quad + j b_1[k] (r_{hv}[m-l]r_{hv}^*[m-k]) \\
&\quad + b_2[k] (r_{vv}[m-l]r_{vv}[m-k] + r_{hv}[m-l]r_{hv}^*[m-k]) = 0,
\end{aligned} \tag{A.2}$$

435 where $m = -(N-1) + p, \dots, N-1$; $k = 0, \dots, p$; $l = 0, \dots, p$; and we have considered

436 $\mu = \eta = 1/2$.

437 Eq. (A.2) can be expressed in matrix form as:

$$\mathbf{R} \cdot \mathbf{f} = \mathbf{R}_0, \tag{A.3}$$

438 where $\mathbf{f} = (a[1] \dots a[p] b_1[0] \dots b_1[p] b_2[0] \dots b_2[p])^T$ and

$$\mathbf{R}_0 = \begin{pmatrix} \sum_m -(r_{hv}[m-1]r_{hv}^*[m] + r_{hh}[m-1]r_{hh}^*[m]) \\ \vdots \\ \sum_m -(r_{hv}[m-p]r_{hv}^*[m] + r_{hh}[m-p]r_{hh}^*[m]) \\ \sum_m (r_{vv}[m]r_{hv}[m] + r_{hv}[m]r_{hh}^*[m]) \\ \vdots \\ \sum_m (r_{vv}[m-p]r_{hv}[m] + r_{hv}[m-p]r_{hh}^*[m]) \\ \sum_m j(r_{vv}[m]r_{hv}[m] + r_{hv}[m]r_{hh}^*[m]) \\ \vdots \\ \sum_m j(r_{vv}[m-p]r_{hv}[m] + r_{hv}[m-p]r_{hh}^*[m]) \end{pmatrix}$$

439 References

- 440 Akintug, B., Rasmussen, P. F., 2005. A Markov switching model for annual
441 hydrologic time series. *Water Resour. Res.* 41, W09424,
442 doi:10.1029/2004WR003605.
- 443 Anthony, R. E., Ringler, A. T., Wilson, D. C., Bahavar, M., Koper, K. D., 2020. How
444 processing methodologies can distort and bias power spectral density estimates
445 of seismic background noise. *Seismol. Res. Lett.* XX, 1–13, doi:
446 10.1785/0220190212.
- 447 Antonovskaya, G., Kapustian, N., Basakina, I., Afonin, N., Moshkunov, K., 2019.
448 Hydropower Dam State and Its Foundation Soil Survey Using Industrial Seismic
449 Oscillations. *Geosciences*, 9, 187, doi:10.3390/geosciences9040187.
- 450 Bard, P.-Y. (ed), 2008. The *H/V* technique: results of the SESAME Project. *Bull.*
451 *Earthq. Eng.* 6(1), 1-147 (special issue).
- 452 Bendat, J. S., and A. G. Piersol (1986). *Random Data: Analysis and Measurement*
453 *Procedures*, Wiley, New York, 407 pp.
- 454 Cadet, H., Macau, A., Benjumea, B., Bellmund, F., Figueras, S., 2011. From ambient
455 noise recordings to site effect assessment: the case study of Barcelona
456 microzonation. *Soil Dynam. Earthq. Eng.* 31, 271–281, doi:
457 10.1016/j.soildyn.2010.07.005.
- 458 Castellaro, S., Mulargia, F., 2009. The effect of velocity inversions on H/V. *Pure Appl.*
459 *Geophys.* 166, 567–592, doi: 10.1007/s00024-009-0474-5.

460 Chávez-García, E J., Sáchez, L. R., Hatzfeld, D., 1996. Topographic site effects and
 461 HVSR. A comparison between observations and theory. *Bull. Seism. Soc. Am.*
 462 86, 1559-1573.

463 Diagourtas, D., Tzanis, A., Makropoulos, K., 2001. Comparative study of microtremor
 464 analysis methods. *Pure Appl. Geophys.* 158, 2463–2479, doi:
 465 10.1007/PL00001180.

466 Dobry, R., Oweis, I., Urzua, A., 1976. Simplified procedures for estimating the
 467 fundamental period of a soil profile. *Bull. Seism. Soc. Am.* 66(4), 1293-1321.

468 EC8 (2004) Eurocode 8: Design of structures for earthquake resistance. Part 1: General
 469 rules, seismic actions and rules for buildings, European Norm, European
 470 Committee for Standardisation, European Committee for Standardisation Central
 471 Secretariat, Rue de Stassart 36, B-1050 Brussels, Belgium.

472 Famiani, D., Brunori, C. A., Pizzimenti, L., Cara, F., Caciagli, M., Melelli, L.,
 473 Mirabella, F., Barchi, M. R., 2020. Geophysical reconstruction of buried
 474 geological features and site effects estimation of the Middle Valle Umbra basin
 475 (central Italy). *Engineering Geology* 269, 105543,
 476 doi:10.1016/j.enggeo.2020.105543.

477 García-Jerez, A., Piña-Flores, J., Sánchez-Sesma, F. J., Luzón, F., Pertón, M., 2016. A
 478 computer code for forward calculation and inversion of the H/V spectral ratio
 479 under the diffuse field assumption. *Comput. Geosci.* 97, 67–78,
 480 doi:10.1016/j.cageo.2016.06.016.

481 Glaser, S., 1996. Insight into liquefaction by system identification. *Géotechnique* 46,
 482 641-655, doi: 10.1680/geot.1996.46.4.641.

483 Haskell, N. A., 1960. Crustal reflection of plane SH waves. *J. Geophys. Res.* 65, 4147-
484 4150, doi: 10.1029/JZ065i012p04147.

485 Ibs-von Seht, M., Wohlenberg, J., 1999. Microtremor measurements used to map
486 thickness of soft sediments. *Bull. Seism. Soc. Am.* 89(1), 250-259.

487 Institut Cartogràfic i Geològic de Catalunya (ICGC), 2000. Catalan Seismic Network.
488 International Federation of Digital Seismograph Networks. Other/Seismic
489 Network. doi:10.7914/SN/CA.

490 Jurado, M. J., Salvany, J. M., 2016. Scientific drilling in the campus: Almera-1
491 borehole, unraveling urban subsurface geology in Barcelona (Spain). *Geotemas*
492 16, 617-620.

493 Kawase, H., Nagashima, F., Nakano, K., Mori, Y., 2019. Direct evaluation of S-wave
494 amplification factors from microtremor H/V ratios: Double empirical corrections
495 to “Nakamura” method. *Soil dynamics and Earthquake Engineering* 126,
496 105067, doi: 10.1016/j.soildyn.2018.01.049.

497 Kay, S. M., Marple, S. L., 1981. Spectrum analysis – A modern perspective.
498 *Proceedings of the IEEE* 69 (11), 1380-1419, doi: 10.1109/PROC.1981.12184.

499 Konno, K., Ohmachi, T., 1998. Ground-motion characteristics estimated from spectral
500 ratio between horizontal and vertical components of microtremor. *Bull Seismol*
501 *Soc Am* 88, 228–241.

502 Kozin, F., 1988. Autoregressive moving average models of earthquake records.
503 *Probabilistic Engineering Mechanics* 3, 58–63, doi:10.1016/0266-
504 8920(88)90016-1.

505 Lagunas, M., Gasull, A., 1984. An improved maximum likelihood method for power
506 spectral density estimation. IEEE transactions on acoustics, speech, and signal
507 processing 32(1), 170-173, doi: [10.1109/TASSP.1984.1164292](https://doi.org/10.1109/TASSP.1984.1164292).

508 Lermo, J., Chávez-García, F. J., 1994. Are microtremors useful in site response
509 evaluation? Bull. Seismol. Soc. Am. 84, 1350–1364.

510 Leyton, F., Ruiz, S., Sepúlveda, S. A., Contreras, J. P., Rebolledo, S., Astroza, M.,
511 2013. Microtremors' HVSR and its correlation with surface geology and damage
512 observed after the 2010 Maule earthquake (Mw 8.8) at Talca and Curicó, Central
513 Chile, Engineering Geology 161, 26-33, doi:/10.1016/j.enggeo.2013.04.009.

514 Li, Y., Kareem, A., 1990. ARMA systems in wind engineering. Probabilistic
515 Engineering Mechanics 5, 49-59, doi:10.1016/S0266-8920(08)80001-X.

516 Lunedei, E., Malischewsky, P., 2015. A review and some new issues on the theory of
517 the H/V technique for ambient vibrations. In: A. Ansal (ed.), Perspectives on
518 European Earthquake Engineering and Seismology, Geotechnical, Geological
519 and Earthquake Engineering 39, 371-394, doi:10.1007/978-3-319-16964-4_15.

520 Massa, M., Lovati, S., 2014. Overview of topographic effects based on experimental
521 observations: meaning, causes and possible interpretations. Geophys. J. Int. 197,
522 1537–1550, doi: 10.1093/gji/ggt341.

523 Matsushima, S., Hirokawa, T., De Martin, F., Kawase, H., Sánchez-Sesma, F. J., 2014.
524 The effect of lateral heterogeneity on horizontal-to-vertical spectral ratio of
525 microtremors inferred from observation and synthetics. Bull. Seism. Soc. Am.
526 104, 381-393, doi: 10.1785/0120120321.

527 Molnar, S., Assaf, J., Sirohey, A., Adhikari, S. R., 2020. Overview of local site effects
528 and seismic microzonation mapping in Metropolitan Vancouver, British

529 Columbia, Canada. Engineering Geology 270, 105568,
530 doi:10.1016/j.enggeo.2020.105568.

531 Molnar, S., Cassidy, J. F., Castellaro, S., Cornou, C., Crow, H., Hunter, J. A.,
532 Matsushima, S., Sánchez-Sesma, F. J., Yong, A., 2018. Application of
533 microtremor horizontal-to-vertical spectral ratio (MHVSR) analysis for site
534 characterization: state of the art. Surv. Geophys. 39, 613-631,
535 doi:10.1007/s10712-018-9464-4.

536 Molnar, S., Cassidy, J. F., 2006. A comparison of site response techniques using weak-
537 motion earthquakes and microtremors. Earthquake Spectra 22(1), 169-188, doi:
538 10.1193/1.2160525.

539 Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface
540 using microtremor on the ground surface. Q. Rep. Railway Tech. Res. Inst. 30,
541 no. 1, 25–33.

542 Nakamura, Y., 2019. What is the Nakamura method?. Seismol. Res. Lett., doi:
543 10.1785/0220180376.

544 Okada, H., 2003. The Microseismic Survey Method. Society of Exploration
545 Geophysicists of Japan, translated by Koya Suto. Geophysical Monograph
546 Series No. 12, Society of Exploration Geophysicists, Tulsa.

547 Ólafsson, S., Sigbjörnsson, R., 2011. Digital filters for simulation of seismic ground
548 motion and structural response. J. Earthquake Engineering 15, 1212-1237, doi:
549 10.1080/13632469.2011.565862.

550 Pertou, M, Spica, Z., Caudron, C., 2018. Inversion of the horizontal-to-vertical spectral
551 ratio in presence of strong lateral heterogeneity. Geophys. J. Int. 212, 930–941,
552 doi: 10.1093/gji/ggx458.

553 Piña-Flores (2018)

554 Rohmer, O., Bertrand, E., Mercerat, E.D., Régnier, J., Pernoud, M., Langlaude, P.,
555 Alvarez, M., 2020. Combining borehole log-stratigraphies and ambient vibration
556 data to build a 3D Model of the Lower Var Valley, Nice (France), Engineering
557 Geology 270, 105588, doi:10.1016/j.enggeo.2020.105588.

558 Rosa-Cintas (2017)

559 Sánchez-Sesma, F. J., 2017. Modeling and inversion of the microtremor H/V spectral
560 ratio: physical basis behind the diffuse field approach. Earth Planets Space, 69-
561 92, doi:10.1186/s40623-017-0667-6.

562 SESAME, 2004. Guidelines for the implementation of the *H/V* spectral ratio technique
563 on ambient vibrations. Measurements, processing and interpretation. WP12-
564 Deliverable D23.12, p 62. Available as supplementary material at:
565 <http://dx.doi.org/10.1007/s10518-008-9059-4> (last accessed April 2020).

566 Spica, Z. J., Pertou, M., Nakata, N., Liu, X., Beroza, G. C., 2018. Site characterization
567 at Groningen gas field area through joint surface-borehole H/V analysis.
568 Geophys. J. Int. 212, 412–421, doi: 10.1093/gji/ggx426.

569 Ugalde, A., Egozcue, J. J., Alfaro, A., Pujades, L. G., Canas, J. A., 1998. Estimation of
570 the system function of soils using microtremors. Annales Geophysicae, part 4,
571 Nonlinear Geophysics & Natural Hazards, Supplement 4 to Volume 16, C 1208.

572 Wathelet, M., Chatelain, J.-L., Cornou, C., Giulio, G. D., Guillier, B., Ohrnberger, M.,
573 Savvaidis, A., 2020. Geopsy: A User-Friendly Open-Source Tool Set for
574 Ambient Vibration Processing, Seismol. Res. Lett. XX, 1–12, doi:
575 10.1785/0220190360.

576 Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. F., Wobbe, F., 2013. Generic
577 Mapping Tools: Improved version released. EOS Trans. AGU 94 409-410, doi:
578 10.1002/2013EO450001.