Supplementary Material

Slowness vector estimation using the Continuous Wavelet Transform (CWT): Application to a sparse Ocean Bottom Seismometer Array

by

R. Cabieces, F. Krüger, A. García-Yeguas, A. Villaseñor, E. Buforn, A. Pazos, A. Olivar, and J. Barco

1. OBSs Array Clock Correction

Clock synchronization is the major step in the preprocessing procedure because it is critical to correct the drift of the OBSs clock (Sens-Schönfelder, 2008), to apply any array method. Typically, a linear drift correction calculated from the recovery skew is applied to the data set. However, it has already been shown that there is a significant nonlinear component, especially relevant at the beginning of the deployment (Gardner and Collins, 2012). Thus, to calculate the correct slowness vector it is crucial for the instrument clocks to be synchronized, otherwise it may have phase misalignments during the beamforming. To fulfill the synchronization premise, we apply Time Symmetry Analysis (TSA) (Gouedard et al., 2014). This method is based on the travel-time delay measurements of Rayleigh waves reconstructed from seismic ambient noise (Fig. S1). In practice, the Rayleigh waves are found from the partial reconstruction of the Green's Function (GF) of the medium between two stations. The GF are formed through noise cross correlation among receivers giving rise to the noise cross-correlation green functions (NCCFs).

FIGURE S1

We have followed Bensen et al. (2007), in order to obtain the daily NCCFs and the reference. In this procedure, we compute the stack of all daily NCCFs between a land station and the

OBS we want to synchronize, to form the reference. Then, we can achieve the daily lag time, from the cross correlation of the reference and a daily NCCF.

Indeed, the daily cross correlation is the result of a stack of 30 days, which is a necessary time span to reduce the effects uncorrelated noise.

It is also important to emphasize that it is essential to denoise the surface wave before the cross correlation. In this sense, the daily NCCF and the reference has been bandpass filtered from 5-10 s, which are frequencies enclosing the microseismic peaks (Stehly et al., 2007).

The linear stack procedure followed to retrieve the reference and the daily NCCFs has been replaced by the time-scale phase-weighted stack (ts-PWS) to enhance the quality and the SNR of the GF reconstruction (Ventosa et al., 2017). The time-scale phase-weighted stack (ts-PWS) uses complex frames of wavelets to improve the computational efficiency compared with the conventional PWS and considers the instantaneous phase coherence (Fig. S2).

FIGURE S2

To optimize the fit, we invert by the least mean square procedure up to four curves for each OBS, each curve corresponding to a selected land station. The land stations were selected in a distance range of [180 to 250] km, located in southern Portugal. In this paper, we have inverted four curves for every OBS, whose final skew was bigger than 0.04 s (Table S1).

TABLE S1

The fit of all daily cross correlations (reference-daily NCCGFs) with the OBS02 is shown in Fig. S3.

FIGURE S3

2. Horizontal OBS Components Orientation

In this paper, the horizontal components orientation is carried out following (Stachnik et al., 2012). This method estimates the seismometer orientation based on the elliptical particle motion of Rayleigh waves. With this concept in mind, the polarization analysis is effectively measured by the normalized zero-lag cross correlation between the Hilbert transformed vertical component and the radial component for different rotated test angles by

In which, denotes the zero-lag cross-correlation (ZLCC) of the Hilbert transformed vertical component and the radial component respectively and are the autocorrelations. , is the normalized ZLCC.

Rotating the radial component for each test angle to find the maximum correlation, will give us the possible backazimuth (angle of the horizontal component measure from the North).

The earthquakes chosen to orientate the OBSs had magnitudes Mw > 6.0 and depth < 100 km (Table S2) to assure the high SNR of the Rayleigh wave fundamental mode (Fig. S4). Moreover, we found the best ZLCC results applying a band-pass filtered between [0.01-0.06] Hz corner frequencies. The final sensor orientation is estimated from the mean of the orientations with and .

TABLE S2

FIGURE S4

3. Beamforming

Once the CFs are built, the slowness vector must be calculated from the beamforming of all traces. Beamforming is just a delay-stack process of each trace according to a test slowness vector. For this purpose, we implement a Broad Band Frequency-Wavenumber (BB-FK) beamforming algorithm (Kværna and Doornbos, 1986; Kvaerna and Ringdahl, 1986) in which the beam power is assessed by using

where, is the Fourier Transform of the trace, is the number of sensors in the array, the vector position of the sensor from the coordinate origin, the wavenumber vector and are the lower and upper limits of the frequency band, respectively.

The semblance function $S(f, \mathbf{k})$ (Neidell and Taner, 1971) is defined following the same notation as BB-FK

where, is computed to analyze the coherence of the wave propagating through the array giving the normalized power of the trace stack for a test slowness vector. Thus, to find the right slowness vector for the P- and S-waves, the maximum power and semblance among the test slowness vectors are estimated for a short sliding time window. Then, this time window moves throughout the entire period of interest (Fig. S5).

Fig. S5 shows the beamforming of the CFs' sliding time windows of 24 s, with the CFs derived from the application of CWT methodology to the transverse component of Earthquake

T13 (Table 1) in the frequency range [0.05 – 0.1] Hz. Fig S5a shows the maximum semblance (Eq. 14) computed for the sliding time windows. Fig S5b shows the maximum time-varying absolute power (Eq. 13). The temporal evolution of the back azimuth that corresponds to the maximum semblance is shown in Fig. S5c and the temporal evolution of the slowness that corresponds to the maximum semblance is shown in Fig. S5d. On the bottom of Fig. S5, we show the slowness maps that correspond to the maximum semblance for the P-wave and the S-wave.

As a reminder, our methodology is designed to find the slowness vector based on a plane wavefront model. For this reason, it is limited to epicenter distances larger than twice the array aperture (Almendros et al., 1999).

FIGURE S5

4. Parameter settings

The parameters selected for the slowness vector estimation are critical. In this section, we summarize the most decisive ones. The main parameters are the number of cycles of the CWT, the time windows of the STA/LTA algorithm, the corner frequencies of the filter applied previously to the envelope computation and the array beamforming parameters. The following parameters are found empirically for our array and they must be considered merely as guidelines, due to each array configuration needing a different parameter optimization.

Wavelet analysis is based on one-octave band decomposition, so the width of the wavelet contracts as the frequency increases in the time-frequency plane. This decomposition allows for overall frequency resolution smoothing, ideally designed for common signal conditions in which low-frequency energy is persistent, while high-frequency components exist for a short duration and rapidly change the instantaneous frequency within the signal. That is a valid approach for the majority of seismic signals and, conventionally, the scalogram is computed

by contracting the wavelet toward higher frequencies while setting a constant number of cycles.

Eight was the value we established for the number of cycles over the vast majority of earthquakes analyzed in this paper. The selection was made empirically but also to fulfill the minimum requirements regarding time-frequency resolution. The time-frequency resolution of the wavelet can be estimated from the width of the Fourier transform of the Morlet wavelet at -3 dB, namely Full Width at Half-Maximum (FWHM):

where is the width of the Morlet wavelet in the frequency domain. Eq. S5 gives a quantitative value for the frequency resolution estimation, which conversely can be obtained in the time domain from the envelope of the real and complex part of the Morlet wavelet.

However, the evident multiresolution time-frequency technique drawback is that the seismic signals may contain a wide range of time-varying frequency content at higher frequencies and long-term frequency energy content at lower frequencies. For these situations, it is preferable to use a low number of cycles to detect and separate the components of the low frequencies and a high number of cycles to avoid vanishing signal energy at higher frequencies. Then, it would be reasonable to adapt the frequency smoothing to the specific signal requirements by increasing the number of cycles with the frequency of the wavelet (Cohen, 2014). Fig. S6 shows the wavelet analysis for the vertical component of Earthquake T2 (Table1) recorded at OBS02. An important feature of this signal is its low signal-to-noise ratio and that the P-wave is an emergent signal (Fig. S6a). We conducted several tests with different numbers of cycles and with a gradual variation of the number of cycles with frequency. In Fig. S6b the number of cycles is relatively low (w = 5), to highlight the temporal dynamics. In Fig. S6c the number

of cycles is high (w = 16) to highlight frequency dynamics. But, since this is a signal with persistent energy content over the higher frequency range, it would be more convenient in this specific case to slightly increase the number of cycles with the frequency of the wavelet (Fig. S6d). Despite this feature, the wavelet analysis method with w = 5 can better detect the onset of the P-wave and S-wave (Fig. S6b), but the spectral information vanishes. On the contrary, the onset of the P and S-wave is slightly diffuse if the analysis is made with w = 16, but we have a better estimation of the frequency content. Fig. S6d shows a time-frequency resolution trade-off designed to allow high temporal resolution for higher frequencies but also to separate the frequency oscillations. An example of bank of atoms generated to build the scalogram with a constant number of cycles and varying the number of cycles is shown in Figure S7a and S7b.

FIGURE S7

FIGURE S7

Concerning the STA/LTA algorithm, we set the ratio of the STA and LTA window length to 1/40 to find the CFs. For this dataset, we have found that 1 s was the best choice for the STA time window. Choosing 1 s the sensibility is high enough to detect the seismic signal onset and at the same time avoids false triggers around spikes. In terms of LTA, we have selected 40 s to encapsulate a few periods of typical ocean noise. Moreover the 40 s LTA duration allows us to detect weak P-waves compared to S-waves as well as a high improvement of the S-wave detection when there is P-wave coda contamination.

The waveform of the envelope relies strongly on the frequency content of the signal. For this reason it is necessary to filter the raw signal previous to estimating the envelope. However, a priori selection of the corner frequencies is complicated due to the large variety of frequencies with significant energy content. In this paper, we apply a high pass filter of order 3 to the raw

seismograms with corner frequencies at 0.5 Hz, because both the STA/LTA and the envelope method in general need pre-filtering to suppress long-period noise.

With respect to the frequency domain array beamforming, the time window and the bandwidth are the main parameters. From our experience, the time window must range from 24 – 26 s to encompass cycles of the CFs and the period that it takes for the seismic wave to cross the aperture of the array. The optimum frequency bandwidth for the analyzed array ranges from [0.05 – 0.15] Hz.

5. Algorithm for Slowness Vector Estimation

We outline in the pseudo-code the workflow of the main program in "Main Program", the algorithm of the wavelet processing to compute the scalogram and the CFs in Algorithm 1 and the function to compute the Continuous Wavelet Transform in Algorithm 2. ObsPy functions

are written in blue. The entire code was written in Python 3 and uses parallelization process to boost the performance of the algorithm.

Main Program

```
begin
  raw <- Read(day)
  rawH <- Read(day)

onset <- scan(raw)

# Vertical Component Analysis
 rtiempo(in:raw,out:removedtime)
 deconv(in:removedtime,in:dataless,out:Velocity_path,in:onset)
 wavelet(in:Velocity_path,out:CF)
 SlownessP,BAZ_P,ErrorSlownessS,ErrorAngleP <-
FK(in:CF_path,in:path,in:coordinates,in:parameters)

# Horizontal Component Analysis
 rotate(in:rawH,out:rotated,out:AZMAX)
 rtiempo(in:rotated,out:removedtimeH)
 deconv(in:removedtimeH,in:dataless,out:VelocityH_path,onset)
 wavelet(in:VelocityH,out:CFH)
 SlownessS,BAZ_S,ErrorSlownessS,ErrorAngleS <- FK(in:CFH_path,in:path,
in:coordinates,in:parameters)
end
```

Algorithm 1: wavelet(**in:**velocity_path, **out:**CF)

```
begin

  for files in listdir[velocity_path)] do
      tr <- read file
    dt <- tr.stats.starttime
    tr.detrend()
    tr.taper()
    tr.filter('highpass',freq=1,corners=3,zerophase=True)
    npts <- tr.stats.npts
    delta <- tr.stats.delta
    ## Parameters Morlet Wavelet
    tmin <- 2
    fmin <- 2
    fmax <- 8
    wmin <- 6
    wmax <- 10
    nf <- 40
    scalogram<-ccwt(in:tr.data,in:1/delta,in:fmin,in:fmax,in:wmin,
                    in: wmax,in: nf,in : tmin)
    scalogram <- np.abs(scalogram)**2
    CF <- np.log10(scalogram)
    CF <- np.diff(CF)
    for j <- 1 to npts do
          row <- np.mean(pdif[:,j])
          Mat.append(row)
    end for
    st <- Stream([Trace(data=Mat, header=stats)])
    st <- st.filter('lowpass', freq=0.1, corners=3, zerophase=True)
    st.taper(max_percentage=0.1)
    st.plot()
    st.write(ficheros_procesados_path,format='MSEED')

  end for
end
```

Algorithm 2: ccwt(data,srate,fmin,fmax,wmin,wmax,nf,tmin)

```
begin

    #Continuous Wavelet Transform (Morlet Wavelet)

    #Wavelet parameters
    dt <- 1/srate
```

```
    npts <- len(data)

    #Logarithmically space central frequencies
    frex <- np.logspace(np.log10(fmin), np.log10(fmax), nf)

    #Kernel of the Mother Morlet Wavelet
    wtime <- np.arange(-tmin, tmin+dt, dt)
    half_wave <- (len(wtime)-1)/2
    nCycles <- np.linpace(wmin,wmax,nf)

    ###FFT parameters
    nKern <- len(wtime)
    nConv <- nKern+npts-1
    tf <- np.zeros((len(frex), npts-1), dtype<-np.complex)
    ##FFT data
    dataX<-np.fft.fft(data,nConv)

    ##loop over frequencies
    for each fi in [0, len(frex)-1] do
       s <- nCycles[fi]/(2*np.pi*frex[fi])
       A <- 1/(np.pi*s**2)**0.25
       cmw <- np.multiply(np.exp(np.multiply(1j*2*(np.pi)*frex
        [fi],wtime)),np.exp(-1*np.divide(np.power(wtime,2),2*s**2)))
       cmw <-cmw.conjugate()
       cmw <- A*cmw
       cmwX <- np.fft.fft(cmw,nConv)
        #Convolution
       cwt <- np.fft.ifft(np.multiply(cmwX,dataX))
       end <- len(cwt)
       cwt <- cwt[int(half_wave+1):int(end-half_wave)]
       tf[:,fi] <- cwt
    end for
end
```

Figure Captions

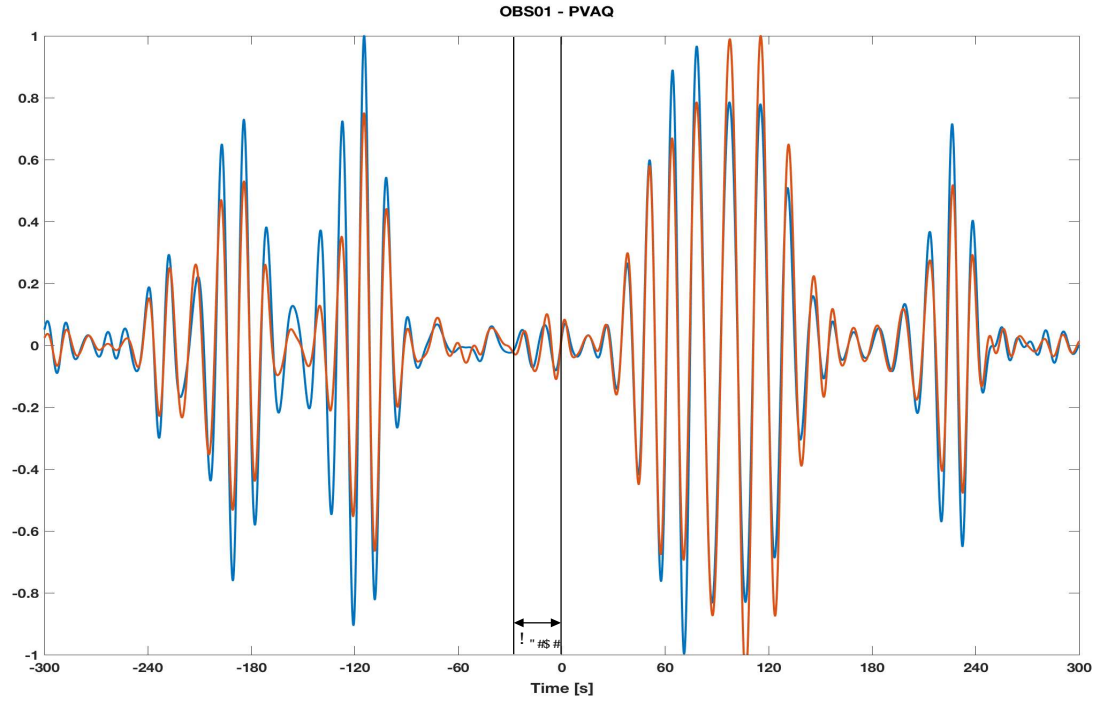

Figure S1. The red line is the reference build from the stack of all daily NCCFs, while the blue line is an example of NCCF reconstructed between the OBS01 and the station PVAQ. , initial correction is applied to the reference point in order to achieve an absolute clock correction. The blue boxes are the time window on the Rayleigh wave. It was estimated 4.0 km/s to start the windowing over the surface wave.

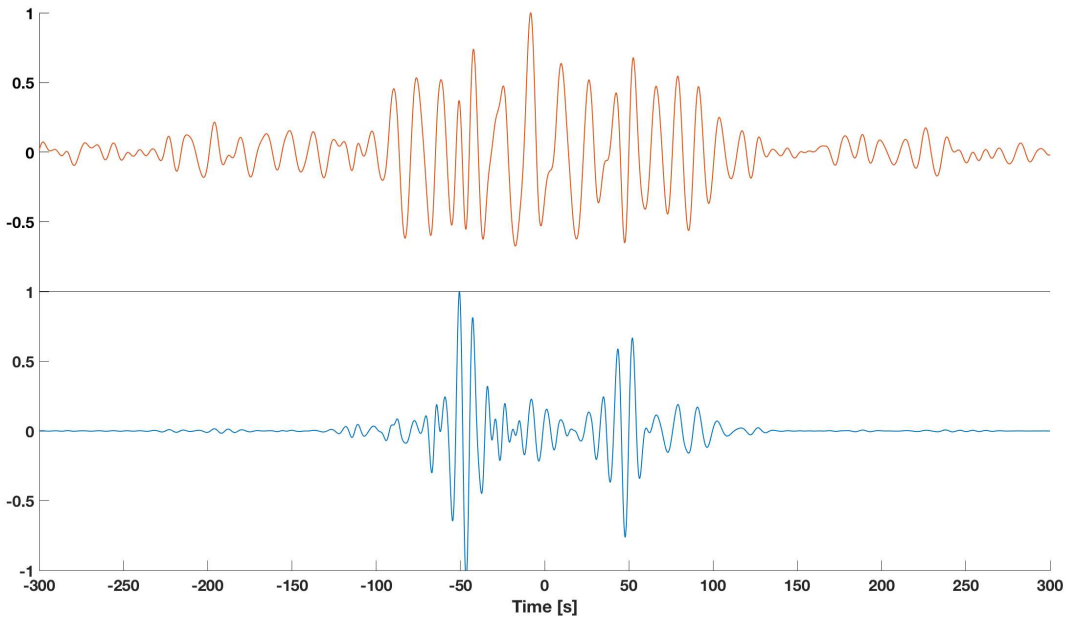

Figure S2. Comparison between the reference (OBS01-0BS02) build from a linear stack (upper panel, red line) and the ts-PWS (Ventosa *et al.* 2017) (lower panel, blue line). The blue line enhances the predominant periods of the signal.

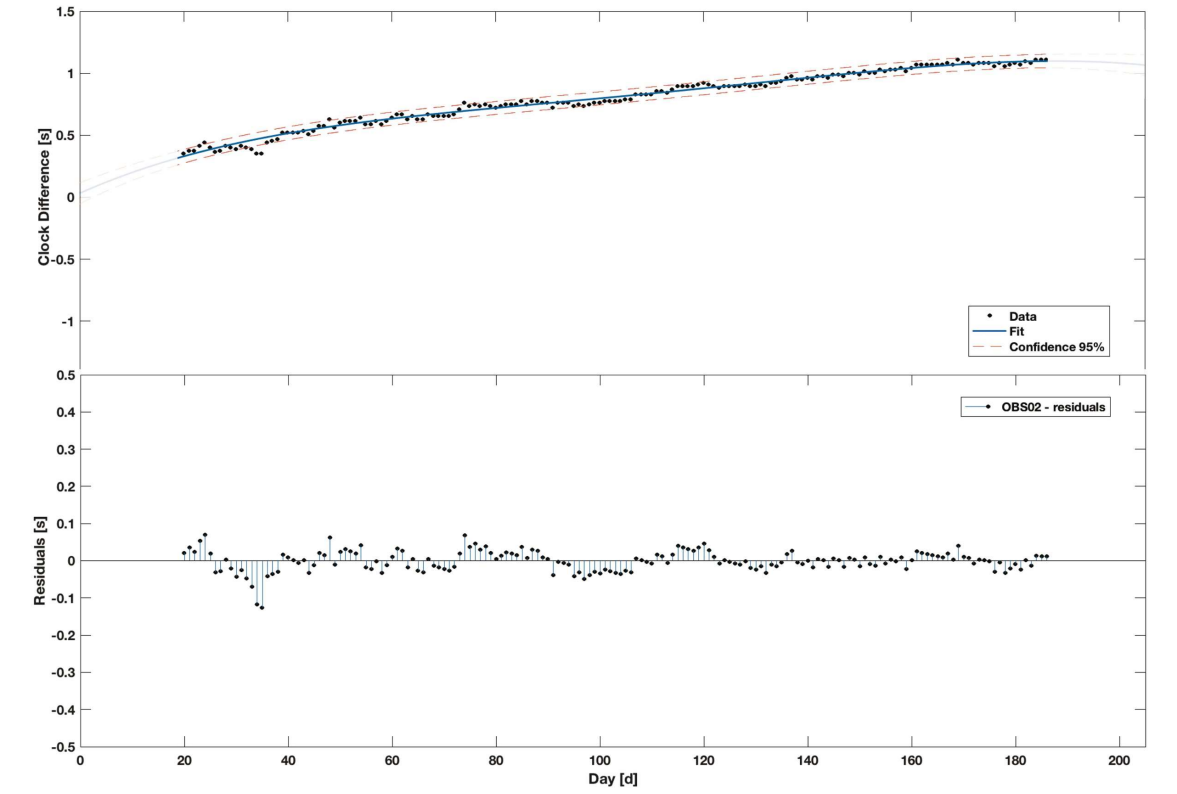

Figure S3. Curve fitting of the daily time difference OBS02. The upper panel black dots are the results of the cross correlation between the reference of the OBS02 and land stations (South Portugal, Fig. 7) NCCFs, the blue line is the curve fit and the red dashed line is the confidence bound 95%. The lower panel displays the residual fit plot between the curve fitting and the results of the cross correlations.

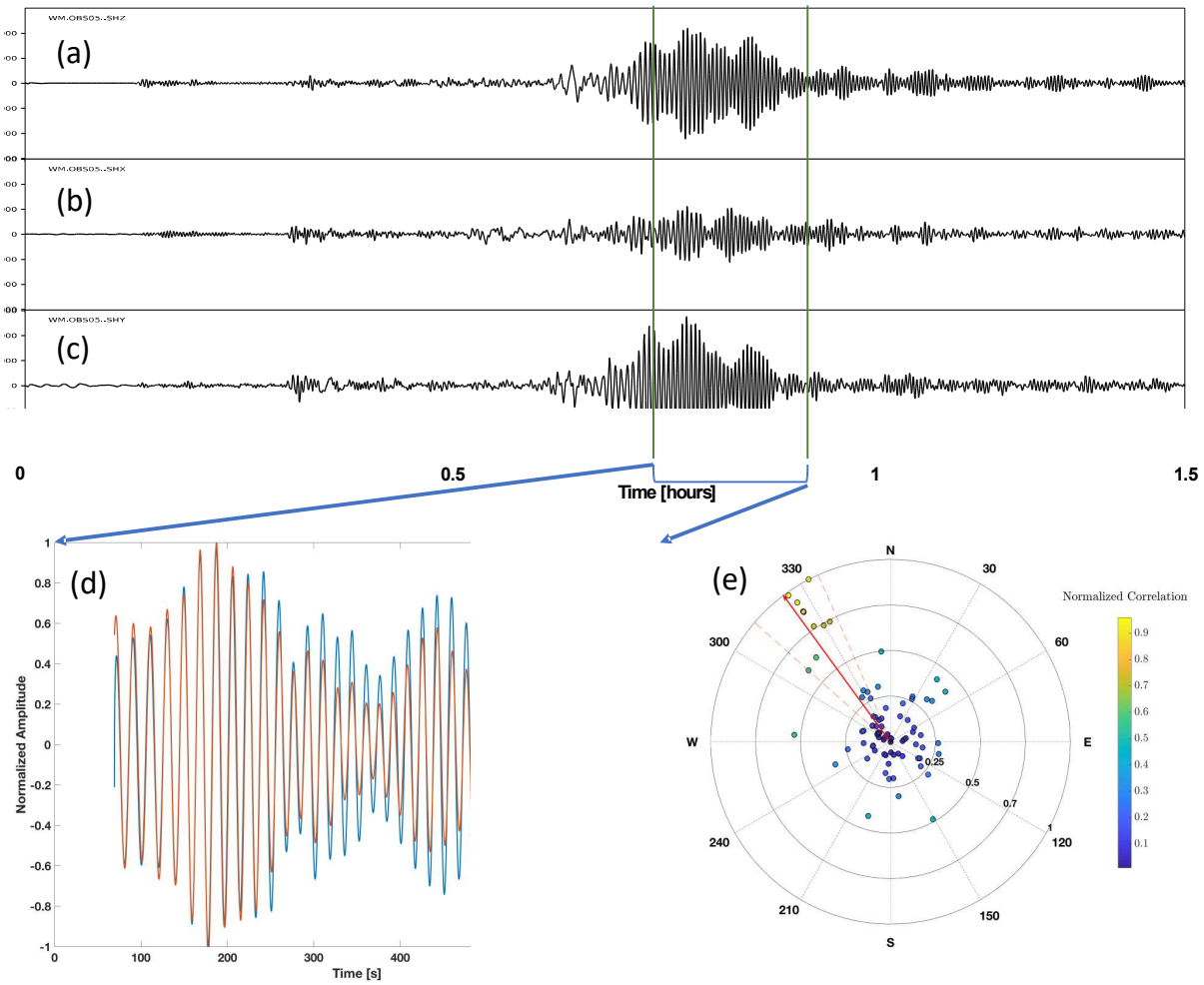

Figure S4. Horizontal component orientation from the Chile earthquake in 2015-09-16 (Mw 8.3). a), b) and c) are the vertical and horizontal components respectively of the OBS05. d) shows band-pass filtered [0.01-0.06] Hz of the earthquake, blue represents the radial component and the red line the Hilbert transform of the vertical component (Correlation 0.92). e) is an example of polar plot orientation OBS05, in which radius and color are directly related to normalized correlation. The red line is the azimuth horizontal orientation mean value after the culling of the orientations results and the dashed red line plots is the standard deviation .

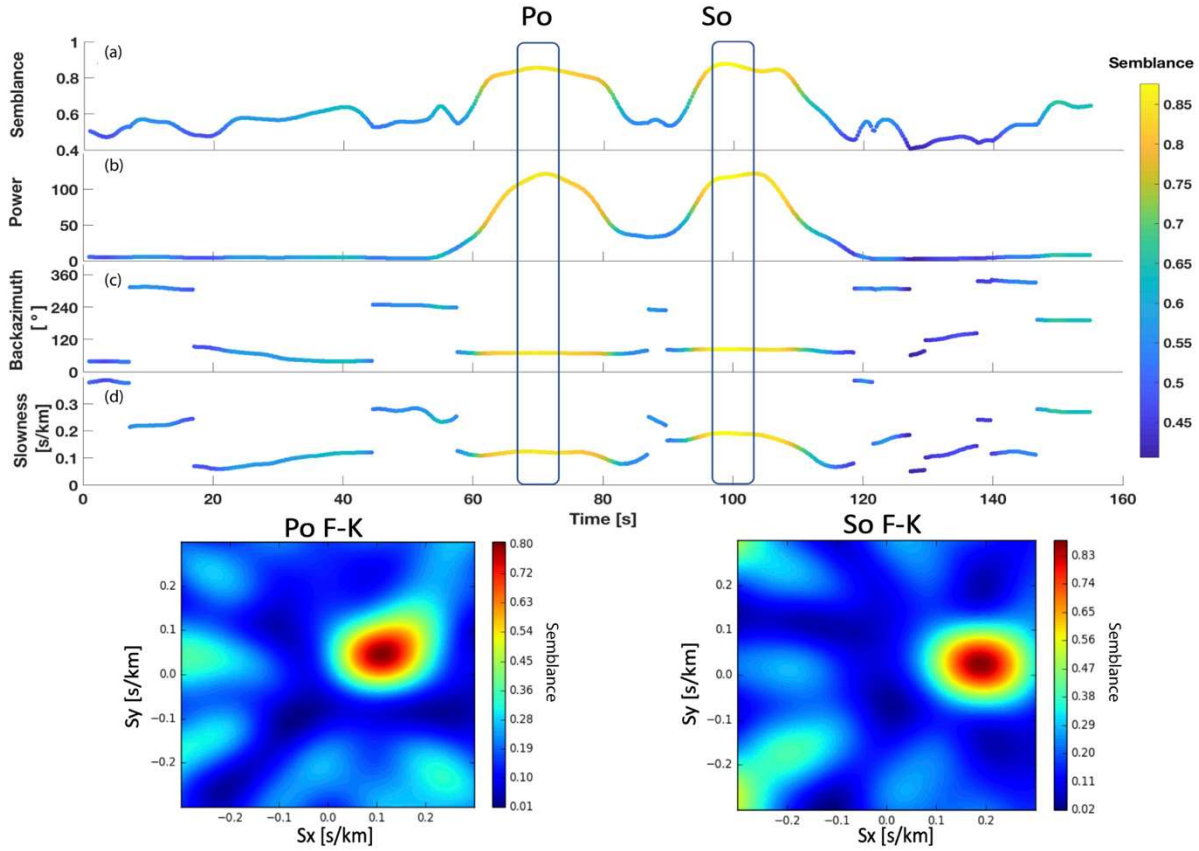

Figure S5. Earthquake T13, BB-FK analysis. CWT method upper panels, a) Semblance *S(f,**k**)*. b) *P(f,**k**)*. c) and d), temporal representation of the back azimuth and slowness in accordance with the maximum *S(f,**k**)*. Lower panels: left, Po wave slowness map, right, So wave slowness map.

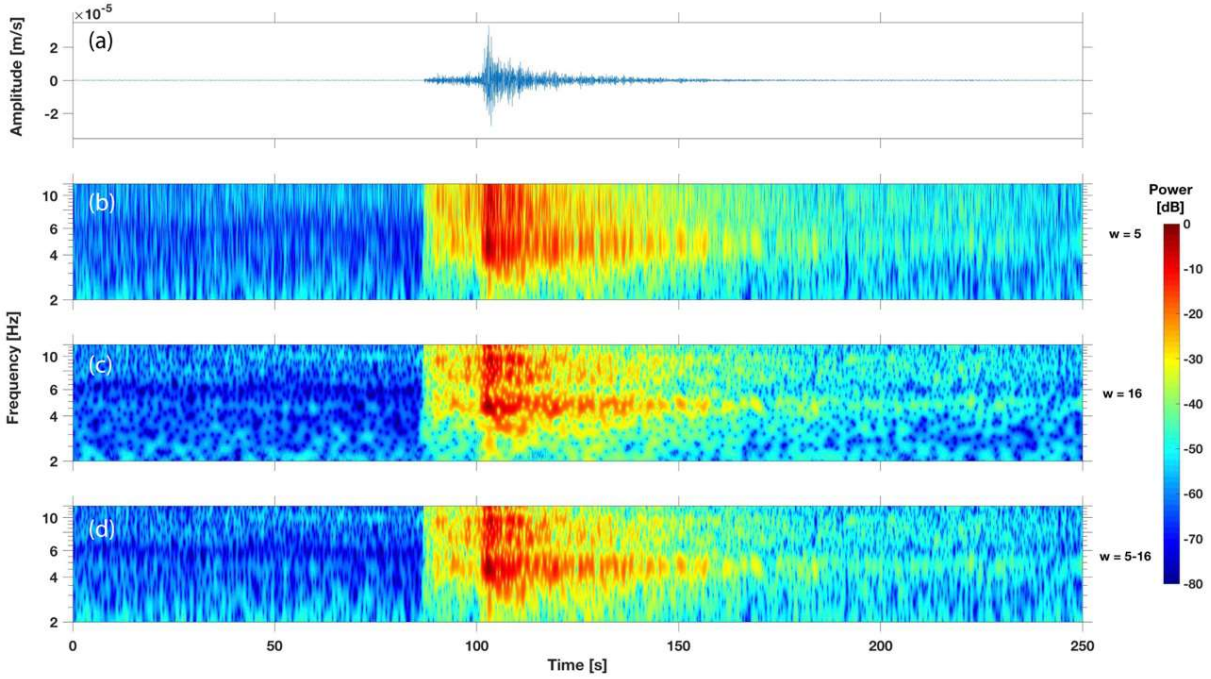

Figure S6. a) Seismogram of Earthquake T1 recorded on the vertical component of OBS06, b) Scalogram computed with number of cycles equal to 5. c) scalogram computed with number of cycles equal to 16. d) scalogram computed with number of cycles varying from 5 – 16 in the frequency band 2 – 12 Hz.

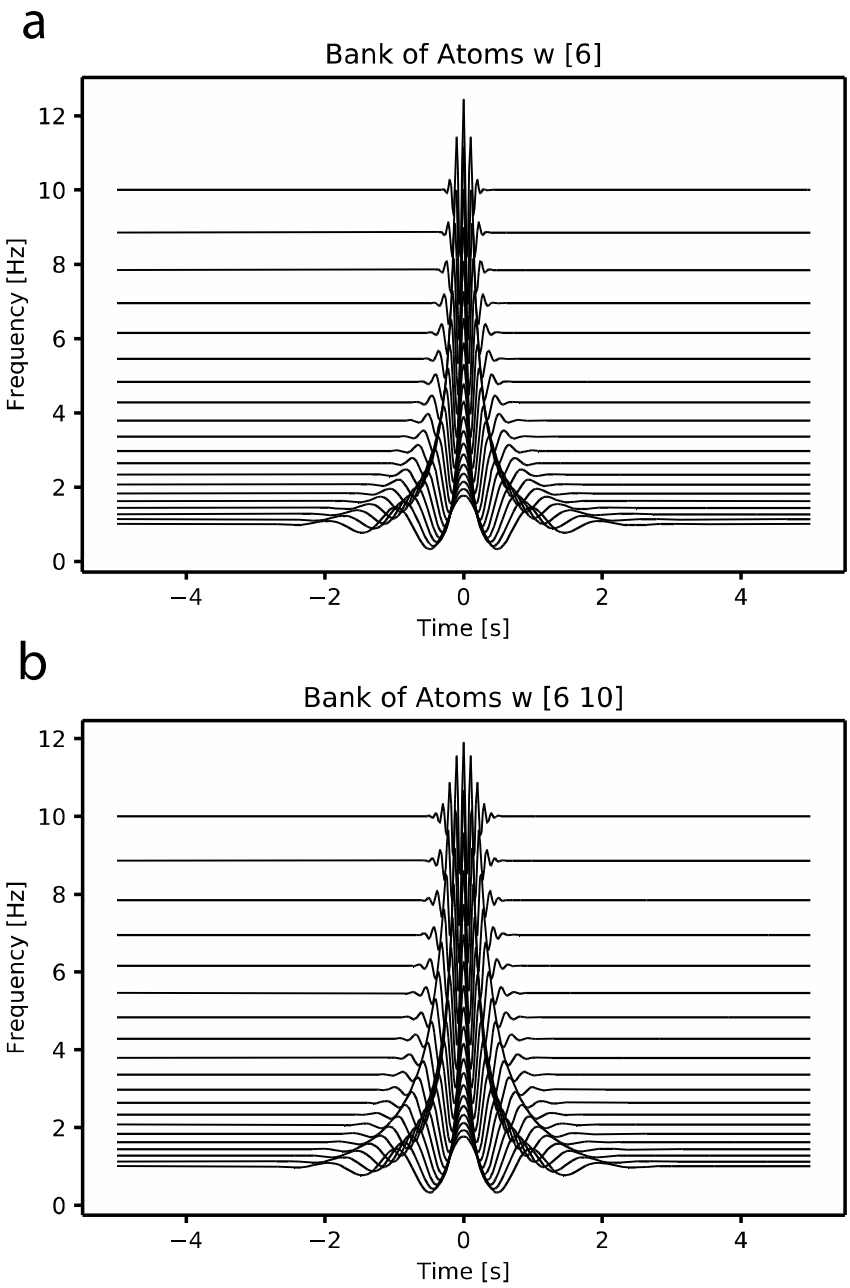
a
Bank of Atoms w [6]
Frequency [Hz]
Time [s]
b
Bank of Atoms w [6 10]
Frequency [Hz]
Time [s]

Table S1. Fit Results, first column is the skew of the OBSs at the recovery

	Skew (s)	*RMS (s)*	*R-Square*
OBS01	*0.219*	*0.041*	*0.730*
OBS02	*1.235*	*0.026*	*0.985*
OBS05	*1.108*	*0.072*	*0.868*

Table S2. Horizontal Orientation results. Statistics applied after discarding correlation values < 0.4 and mean values from orientations inside

	Mean (º)	*Variance*	*Standard deviation*	*Number of Orientations*
OBS01	*14.5*	*26.3*	*5.1*	*14*
OBS02	*211.5*	*37.4*	*6.1*	*12*
OBS03	*8.7*	*18.7*	*4.3*	*11*
OBS05	*325.8*	*27.8*	*5.3*	*14*

OBS06	*188.7*	*105.3*	*10.2*	*12*

REFERENCES

Almendros, J., Ibáñez, J.M., Alguacil, G., Del Pezzo, E., 1999. Array analysis using circular-wave-front geometry: an application to locate the nearby seismo-volcanic source. Geophysical Journal International 136, 159-170.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F.-C., Moschetti, M.P., Shapiro, N.M., Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophysical Journal International 169, 1239-1260.

Cohen, M.X., 2014. Analyzing neural time series data: theory and practice. MIT press.

Gardner, A.T. & Collins, J.A., 2012. Advancements in high-performance timing for long term underwater experiments: a comparison of chip scale atomic clocks to traditional

microprocessor-compensated crystal oscilla- tors, in 2012 Oceans, pp. 1–8, IEEE, Hampton Roads, VA, USA.

Gouedard, P., Seher, T., McGuire, J.J., Collins, J.A., van der Hilst, R.D., 2014. Correction of Ocean-Bottom Seismometer Instrumental Clock Errors Using Ambient Seismic Noise. Bulletin of the Seismological Society of America 104, 1276-1288.

Kværna, T., Doornbos, D., 1986. An integrated approach to slowness analysis with arrays and three-component stations. NORSAR Semiannual Technucal Summary 1, 2-85.

Kvaerna, T., Ringdahl, F., 1986. Stability of various fk estimation techniques. NORSAR Semiannual technical summary 1, 1-86.

Neidell, N.S., Taner, M.T., 1971. Semblance and other coherency measures for multichannel data. Geophysics 36, 482-497.

Sens-Schönfelder, C., 2008. Synchronizing seismic networks with ambient noise. Geophysical Journal International 174, 966-970.

Stachnik, J.C., Sheehan, A.F., Zietlow, D.W., Yang, Z., Collins, J., Ferris, A., 2012. Determination of New Zealand Ocean Bottom Seismometer Orientation via Rayleigh-Wave Polarization. Seismological Research Letters 83, 704-713.

Stehly, L., Campillo, M., Shapiro, N.M., 2007. Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts. Geophysical Journal International 171, 223-230.

Ventosa, S., Schimmel, M., Stutzmann, E., 2017. Extracting surface waves, hum and normal modes: time-scale phase-weighted stack and beyond. Geophysical Journal International 211, 30-44.