

*Third International Symposium on the Effects of Surface Geology on Seismic Motion
Grenoble, France, 30 August - 1 September 2006
Paper Number: xxx*

NOISE BLIND TEST: RETRIEVING DISPERSION CURVES AND INVERSION WITH THE CONDITIONAL NEIGHBORHOOD ALGORITHM

Marc WATHELET

LGIT, University Joseph Fourier, Grenoble, France

ABSTRACT – Within the SESAME European project (Site EffectS assessment using AMbient Excitations, <http://sesame-fp5.obs.ujf-grenoble.fr>), we developed a new software package dedicated to the processing of ambient vibrations. It includes particular implementations of well-known methods like frequency-wavenumber methods and autocorrelation method to derive the dispersion curve of surface waves. Additionally, we propose a new type of inversion procedure based on a direct search (Conditional Neighborhood Algorithm) to retrieve the shear wave velocity profiles with its complete non-linear uncertainty.

These tools are used to solve seven cases proposed during the Noise Blind Test organized within the framework of ESG 2006. At the time of writing this paper, no information is available about the test sites processed during this work.

1. Introduction

Measuring shallow and deep parameters that influence the dynamic behavior of a soil structure at a reasonable cost is an important challenge for seismic risk mitigation. Considering their potential ability to derive the shear wave velocity profiles even at great depths, ambient vibration techniques are becoming more and more popular while their limits of application are still not well defined, lacking validation of user practices and validation of available processing softwares. The Noise Blind Test organized during ESG 2006 offers a very good opportunity to compare and to check the different methods used in the World for processing ambient vibrations.

This paper presents the methodology we used for designing ambient vibration arrays, for computing the dispersion curves and for inverting it in order to obtain the shear wave velocity profiles (Di Giulio et al. In press, Wathelet et al. 2005). Between 20 and 30 arrays are processed during this work but only the arrays recorded for the case N102 are presented here with details.

2. Array design after H/V experiments

The ideal array would contain an infinite number of sensors perfectly sampling the whole space. Unfortunately, for economical and practical reasons, we usually deal with a discrete sampling in space. As the number of sensors decreases, resolution and aliasing restrict the ability of the array to correctly measure the phase velocity. Experience acquired during the SESAME European Project with synthetic datasets (SESAME, Site EffectS using AMbient Excitations, <http://sesame-fp5.obs.ujf-grenoble.fr>) demonstrated that using a dispersion curve outside these limits often leads to completely biased results.

For cases N102, N103, and N104 of this Noise Blind Test, it is the responsibility of the participants to select a maximum of three array layouts with a maximum of ten sensors each. The quality of the velocity estimates is a function of the sampling in space, hence we select the optimum: three arrays with ten sensors.

To orient the choice of the geometries, only the three component signals recorded at one station are given as prior information. Under this constrain, we design arrays so that their expected valid range sufficiently overlap and covers the frequency range of maximum energy. Estimating the valid range of an array and the details of the design strategy is explained in the next paragraphs.

2.1. H/V results

The 1D resonance frequency obtained after H/V measurements gives in general a lower limit on the frequency axis for the correct determination of the dispersion curve from Rayleigh vertical component (Scherbaum et al., 2003). The H/V spectra are calculated for the three cases in figure 1 and the resonance frequencies are given in table 1.

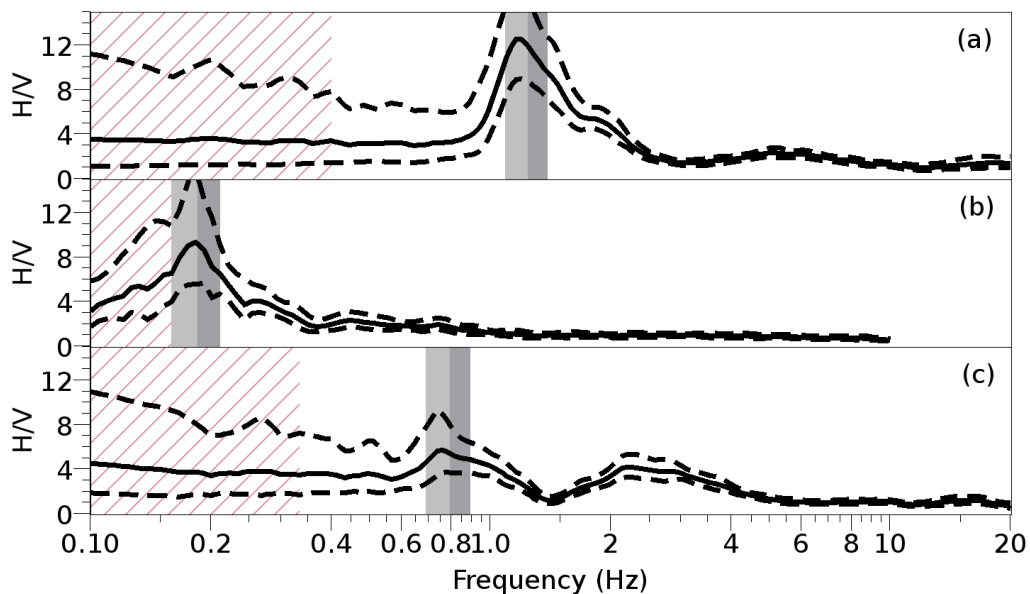


Figure 1. H/V results for cases (a) N102, (b) N103 and (c) N104, average amplitude and standard deviation. The hashed area represents the 10 cycle limit (minimum frequency for a given time window length). The gray area indicate the resonance frequency and its uncertainty.

2.2. Estimating array validity range

Classically, the aperture and the minimum distance between adjacent sensors are considered when designing an array (e.g. Tokimatsu, 1997). The approach proposed here is based on the theoretical array response from which the resolution and aliasing ranges are calculated in the wavenumber space (computed with *build_array* available at <http://www.geopsy.org>). Resolution limit is expected between $K_{min}/2$ and K_{min} where $K_{min}/2$ is measured at the mid-height of the central peak of the theoretical array response. Aliasing limit is expected between $K_{max}/2$ and K_{max} where K_{max} is set when the theoretical array response first exceed 0.5, starting from the origin and in all directions, as show in figures 2b and 2c for the array geometry displayed in figure 2a

(array b, case N102 in table 1). For this example, $K_{min}/2$ is 0.032 rad/m and K_{max} is 0.23 rad/m. All ranges are reported in figure 2d with thin lines on a dispersion curve plot.

2.3. Integrating H/V with array design

The valid range of an array is given in wavenumber space while resonance limits are expressed on a frequency axis. To convert a frequency into a wavenumber, it is mandatory to make assumptions about the expected phase velocity. For real experiments, it is often possible to have a first guess of a simplified soil structure (e.g. sediments over a bed-rock), and to define ranges of velocity for these layers from the literature, or from prior experience with this site.

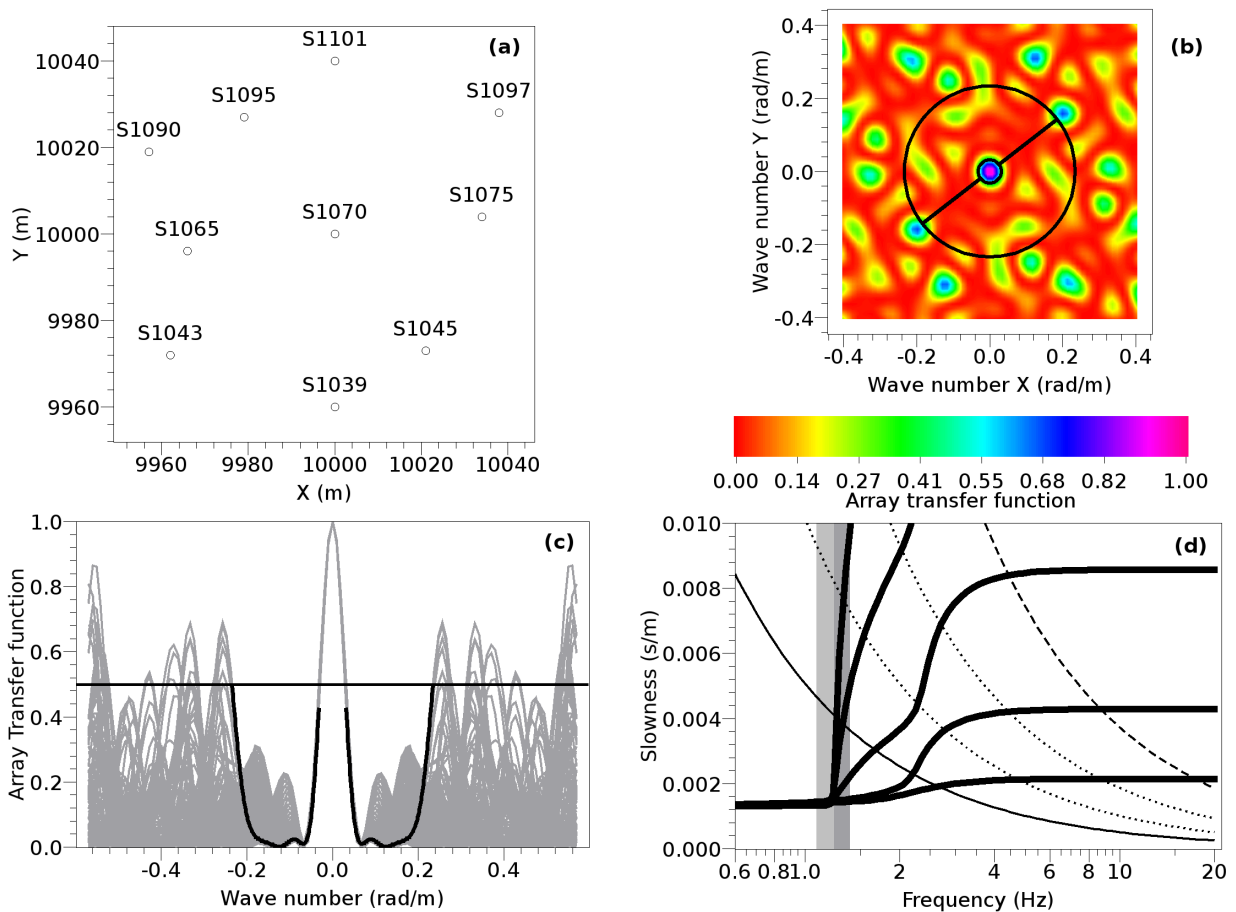


Figure 2. Estimating the array validity range. (a) array geometry used as an example (case N102, array b). (b) Theoretical array response in wavenumber space. The small and large black circles are located at $K_{min}/2$ and K_{max} , respectively. The black line marks the cross section highlighted in figure (c). (c) azimuthal cross-sections of figure (b). (d) Hypothetical dispersion curves (thick lines), $K_{min}/2$ (thin continuous line), K_{min} (lower thin dotted line), $K_{max}/2$ (upper thin dotted line), K_{max} (thin dashed line). The grey area marks the frequency of the H/V calculated in figure 1.

In this case, in the absence of prior knowledge about the possible geology, we assume a reasonable V_s for the sediments, for instance, between 100 and 600 m/s. The thickness is changed to keep $V_s/(4H)$ equal to the resonance frequency obtained from spectral ratios. Poisson's ratio is arbitrary set to 0.4 in the sediments and to 0.25 in the bed-rock. The theoretical dispersion curves are computed for these various cases and are shown in figure 2d (*dc_forward* available at <http://www.geopsy.org>).

The array shown in figure 2d does not require an additional larger array for sites with a very low V_s in the sediments (< 150 m/s), but this is not true for other values. Hence without some rough prior information, it is impossible to assert that the array are properly designed. Hence, we have to assume the phase velocity to be in the expected range mentioned above.

Plots in figure 2 have been computed for the three cases N102, N103 and N104, and the sensor layouts are adjusted in order to define three arrays with overlapping wavenumber ranges (table 1).

Table 1. Selected arrays for cases N102, N103 and N104

| Case | f_0 | Array name | Kmin/2 | Kmax |
|------|----------|------------|--------|-------|
| N102 | 1.25 Hz | a | 0.12 | 1.10 |
| | | b | 0.032 | 0.23 |
| | | c | 0.014 | 0.079 |
| N103 | 0.185 Hz | a | 0.014 | 0.079 |
| | | b | 0.0047 | 0.054 |
| | | c | 0.002 | 0.009 |
| N104 | 0.79 Hz | a | 0.032 | 0.23 |
| | | b | 0.014 | 0.079 |
| | | c | 0.0047 | 0.054 |

3. Determining dispersion curves

Showing all details for all cases of the Noise Blind Test falls beyond the scope of this short paper. Only case N102 is used to illustrate our approach.

Before processing signals, the expected performances of all arrays are estimated with the estimators described in section 2.2 (e.g. table 1). The resonance frequencies of all cases are also of prime importance to target the computations to the right frequency range.

3.1. Signal processing methods

Three methods are used to extract the dispersion curve from ambient vibration signals, the details of which can be found in SESAME deliverable WP06 (<http://sesame-fp5.obs.ujf-grenoble.fr>): frequency-wavenumber method (FK), high resolution frequency-wavenumber method (HRFK), and spatial autocorrelation method (SPAC). All of them are dedicated to vertical components according to their current implementation in *Geopsy* (available at <http://www.geopsy.org>). Extensions to horizontal components are not considered in this work.

Contrary to the implementation of HRFK described in SESAME deliverable WP06 (<http://sesame-fp5.obs.ujf-grenoble.fr>) (block averaging), the cross spectral matrix is evaluated using frequency averaging which allows the estimation of the temporal stability of the computed HRFK velocity. The inverse of the cross spectral matrix is achieved with an eigenvalue decomposition to avoid any smoothing.

SPAC results are checked and selected with *spac2disp* (available at <http://www.geopsy.org>) and they can be directly included in the inversion as described by Wathelet et al. (2005). *Dinver* (available at <http://www.geopsy.org>) allows the simultaneous inversion of FK and SPAC results.

Ten minutes of signal (case N102) are analyzed for each array configuration. No pre-selection (e.g. anti-trigger) is performed before cutting the signals into smaller time

windows. The length of these time windows is variable as a function of frequency keeping 50 cycles within the window (100 for HRFK). Figure 3 shows the FK and HRFK histograms obtained for arrays a to c (case N102). The thin black curves with error bars are the target selected for inversion, shown to ease the comparison between plots of figure 3.

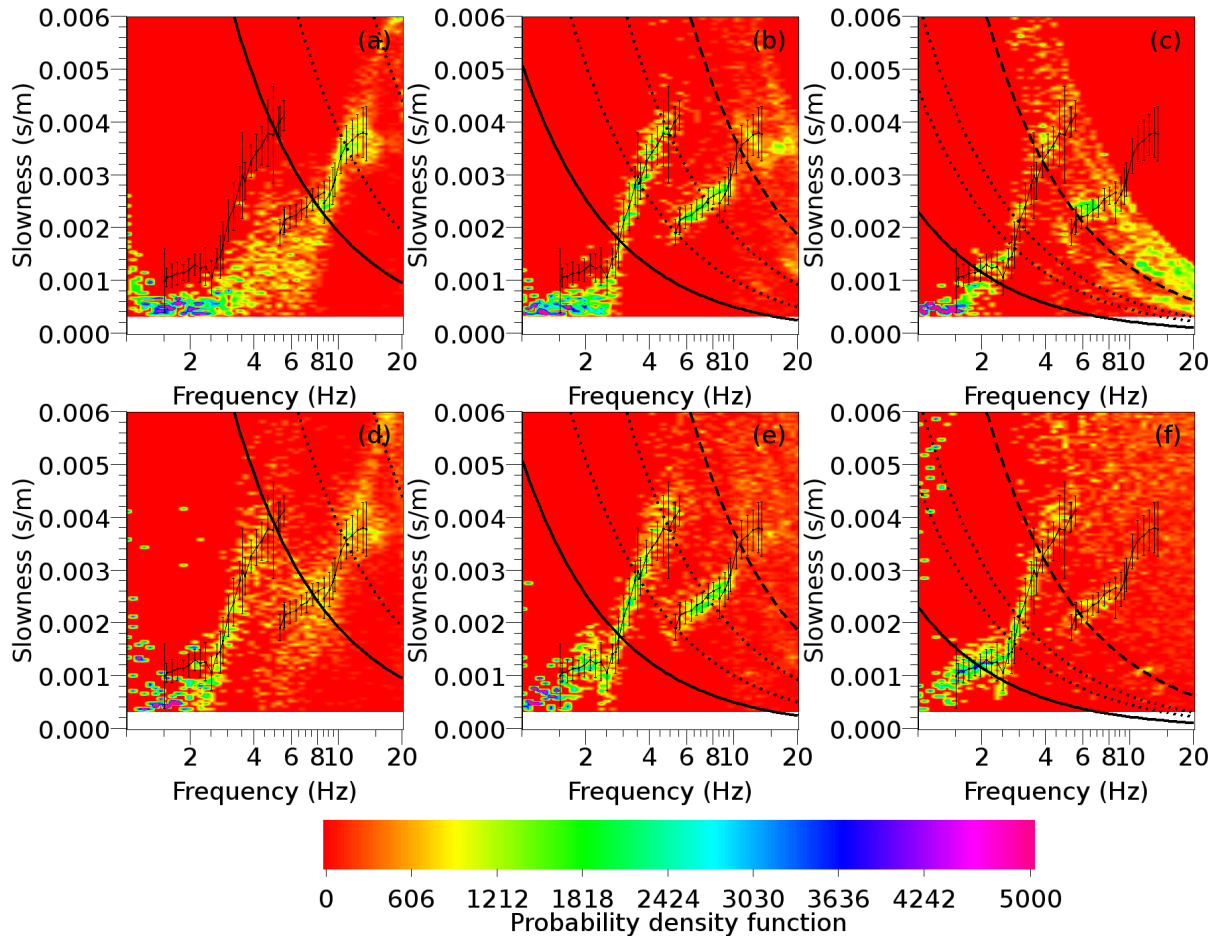


Figure 3. FK (top, a to c) and HRFK (bottom, d to f) statistical results for case N102, arrays a to c (from left to right). The array range is shown in all figures with thin lines. The thin lines with error bars are the selected curves used in the inversion.

3.2. Picking various modes

Obviously, plots of figure 3 contain various modes (at least two). Figure 4 shows a cross-section of figure 3e taken at 5.5 Hz. Additionally, the cumulative probability function is also displayed. The red curves are the Gaussian distributions calculated for the median and the median deviation of the population. These statistical estimates calculated over the raw output of FK or HRFK are totally biased (figure 4, left pane), not representing the complex statistical distribution. When two modes are sufficiently separated on a bi-modal distribution, one can split the original population in two subsets. This is shown in figure 4 (right pane) where the subset with high slowness is removed. The statistical estimators recomputed on the reduced population clearly represent the distribution in a better way as demonstrated by the comparison with the theoretical Gaussian distribution (red curves).

3.3. Merging FK and HRFK results

Considering figure 3a, only one mode can be observed with FK method inside the valid wavenumber range. All estimations outside this range do not correspond to any traveling wave (from the comparison with more resolvent arrays and HRFK). The resolution limit is reached above $K_{min}/2$ (at 9 Hz) because an underestimated slowness is found compared to HRFK and more resolvent arrays. With this array, it is possible to define the velocity until high frequencies (13 Hz). Extensions until 20 Hz might be possible if ambient vibration measurements can be coupled with active source experiments. HRFK method processed for array a confirms that the branch observed with FK method is certainly a higher mode and not the fundamental one. However, the stationarity with time of the HRFK is less good than for FK within the valid wavenumber range, hence FK provides a more precise average dispersion curve.

Comparing figures 3d and 3f, it appears that the HRFK resolution improvement (relative to FK) is very good. It allows, in this case, to extend the valid wavenumber range by a factor 6 which is in agreement with the literature (e.g. Asten and Henstridge, 1984). However, this cannot be taken as a general rule, manual confirm with more resolvent arrays is always mandatory.

From figure 3b, the higher mode can be recognized down to 5 Hz. Aliasing is partially hiding the figure and no cross-check with figure 1a can be done on this mode above 10 Hz. Below 3 Hz, figure 3b clearly demonstrates the danger of using FK as a sole method and outside the valid wavenumber range of the array. This part is partially resolved with HRFK (figure 3e) but with a better precision in figure 3f.

In figure 3f, a good resolution might be achieved down to 1 Hz with HRFK, around the resonance frequency of the site (1.25 Hz). With FK, the resolution limits is probably around 2.5 Hz (between $K_{min}/2$ and K_{min}). In the absence of other means to confirm the velocity at low frequency and to avoid the inclusion of biased information into the inversion, the dispersion curve is cut close to the lower limit of the valid wavenumber range of the array (at 1.5 Hz). Statistics at low frequency are calculated with HRFK results.

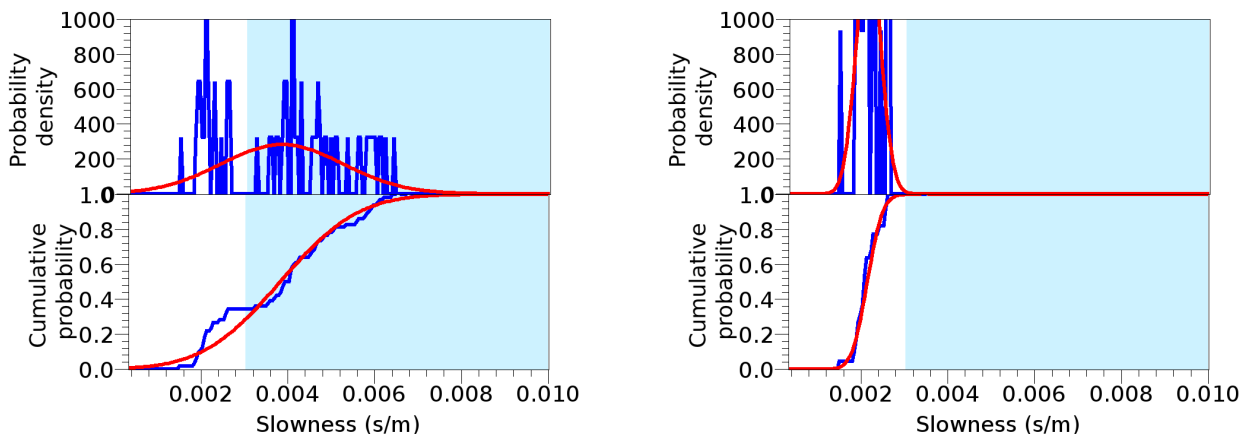


Figure 4. Section across the HRFK histogram of figure 3e taken at 5.5 Hz. Left: unfiltered histogram as produced by HRFK. Right: after selecting the low slowness mode. The red curves represent the Gaussian distribution computed for the median and the median deviation of the remaining population.

From the above discussion, two modes can be observed after FK and HRFK processing. They are identified as the fundamental and the first higher mode but this association is effectively checked during the inversion step (see below).

3.4. Checking merged results with SPAC

SPAC method is applied to all available arrays, i.e. arrays a to c for N102 case. For this case, four narrow rings are defined per array after a manual inspection of the co-array plots, keeping at least 6 to 10 sensor couples and an homogeneous azimuthal distribution within each ring. The results of the autocorrelation computation is given versus frequency in figure 5 (first four rows).

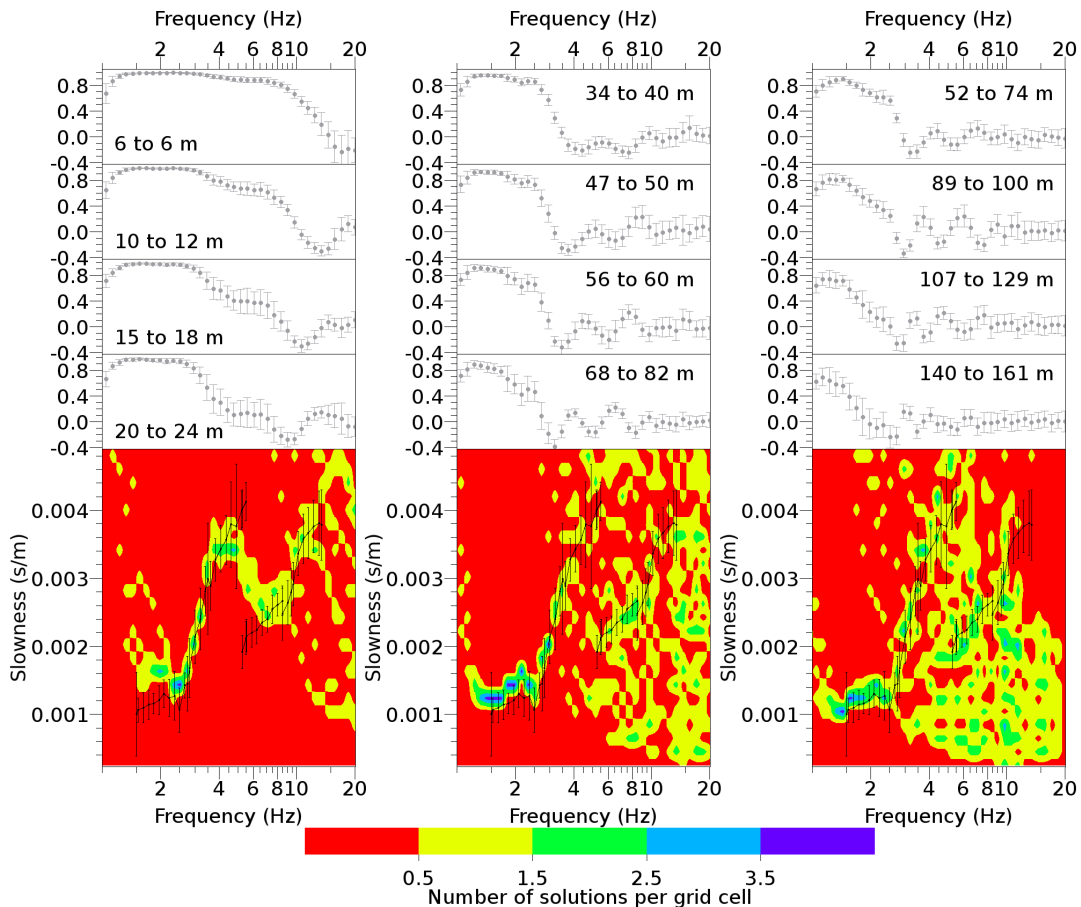


Figure 5. SPAC results for case N102: array a (left), array b (center), and array c (right). The first four rows show the autocorrelation ratio versus frequency for the selected rings. The last row is the dispersion curve computed from the autocorrelation ratio compared to dispersion curve obtained with FK methods (thin lines with error bars).

As described in Wathelet et al. (2005), the frequency-slowness space is divided in cells (25 along slowness, 1 per frequency sample). All the solutions in terms of slowness are calculated for each samples of the autocorrelation curves. The numbers of solutions falling in each cell are counted as shown in figure 5 (bottom row). A dispersion curve can be clearly delineated for each array. They are directly compared to the dispersion curve obtained with FK and HRFK methods (thin line with error bars).

The fundamental mode is visible with all three arrays. With the smallest array (a, figure 5 on the left), it can be followed from 2.5 Hz up to 5 Hz where the shift to the first higher mode occurs. The first higher mode is also visible until 12 Hz in the same way as for FK methods. For array b (figure 5, center), the low frequency part of the dispersion curve matches roughly the dispersion obtained with FK methods. Above 4 Hz, the situation is more confused, and only the low frequency part of higher mode can be confirmed. With the largest array (c, figure 5, right), the dispersion curves fit nicely between 1.5 and 2.5

Hz. The fundamental mode can be followed until 4 Hz but with less precision than for smaller arrays.

In conclusion, SPAC method allows to successfully check the dispersion curve obtained with FK methods, even at low frequency. The autocorrelation ratios are not included in the inversion, because no extension of the frequency range of the dispersion curve can be justified after SPAC analysis. Hence, SPAC results carry almost the same information as the dispersion curve obtained with FK methods.

4. Retrieving the shear wave velocity profile

4.1. Conditional Neighborhood Algorithm (CNA)

The Neighborhood Algorithm proposed by Sambridge (1999) has been applied for the first time by Wathelet et al. (2004) to the inversion of theoretical dispersion curves. The basic principle of such kind of inversion methods (like Genetic Algorithm or Simulated Annealing) is to generate random models into a bounded parameter space. The efficiency of such a method depends on how well the program is able to use the models generated so far to orient the search to the optimum regions of the parameter space. In the case of strong non-uniqueness, these methods can seek for nearly all possible solutions and not only one. Wathelet (2005) demonstrated that the introduction of prior information is natural and easy to implement contrary to linearized and iterative optimization methods where it is difficult to control the inversion paths across the unbounded parameter space.

Wathelet (2005) showed that it is mandatory to include conditions between parameters to overcome the non-uniqueness, specifically for dispersion curve inversion in the presence of low velocity zones. In the original algorithm provided by Sambridge (1999) the parameter space has the shape of a box with fixed ranges for the parameters. Variable changes can help to add conditions between the parameters of the model but Wathelet (2005) proved that the prior probability density function might be affected in an uncontrolled way.

After re-writing the kernel of the Neighborhood Algorithm itself, the solution is provided by the Conditional Neighborhood Algorithm, which currently allows any number of simple conditions between parameters. Written in C++, it is distributed freely under a GPL license at <http://www.geopsy.org> under the software *dinver*. Currently, a module has been written for the inversion of dispersion curves, autocorrelation curves and H/V frequency, either alone or joint. Any other inversion problem can be coded in the same way with any programming language (bash script, fortran, matlab, perl,...).

4.2. Applications

The bi-modal dispersion curves determined with FK methods and crosschecked with SPAC method is inverted to retrieve the shear wave velocity profile (V_s). A parameterized ground model with a sediment layer over a bed-rock is used in a first approach. In a second step a ground model with three layers is tried.

The velocity within the sediments is allowed to vary according to a power law. The parameters are six: V_p in the sediments, V_p in the bed-rock, V_s at the top of the sediments, V_s at the bottom of the sediments, V_s in the bed-rock and the depth of the bed-rock interface. Six simple conditions are introduced between our six parameters to force Poisson's ratio between 0 and 0.5 and to forbid low velocity zones in the ground structure. V_p 's are allowed to vary between 200 and 5000 m/s and V_s ' are allowed to vary from 150 to 3500 m/s (uniform prior probabilities). The density is fixed to 2000 kg/m³ in the whole structure.

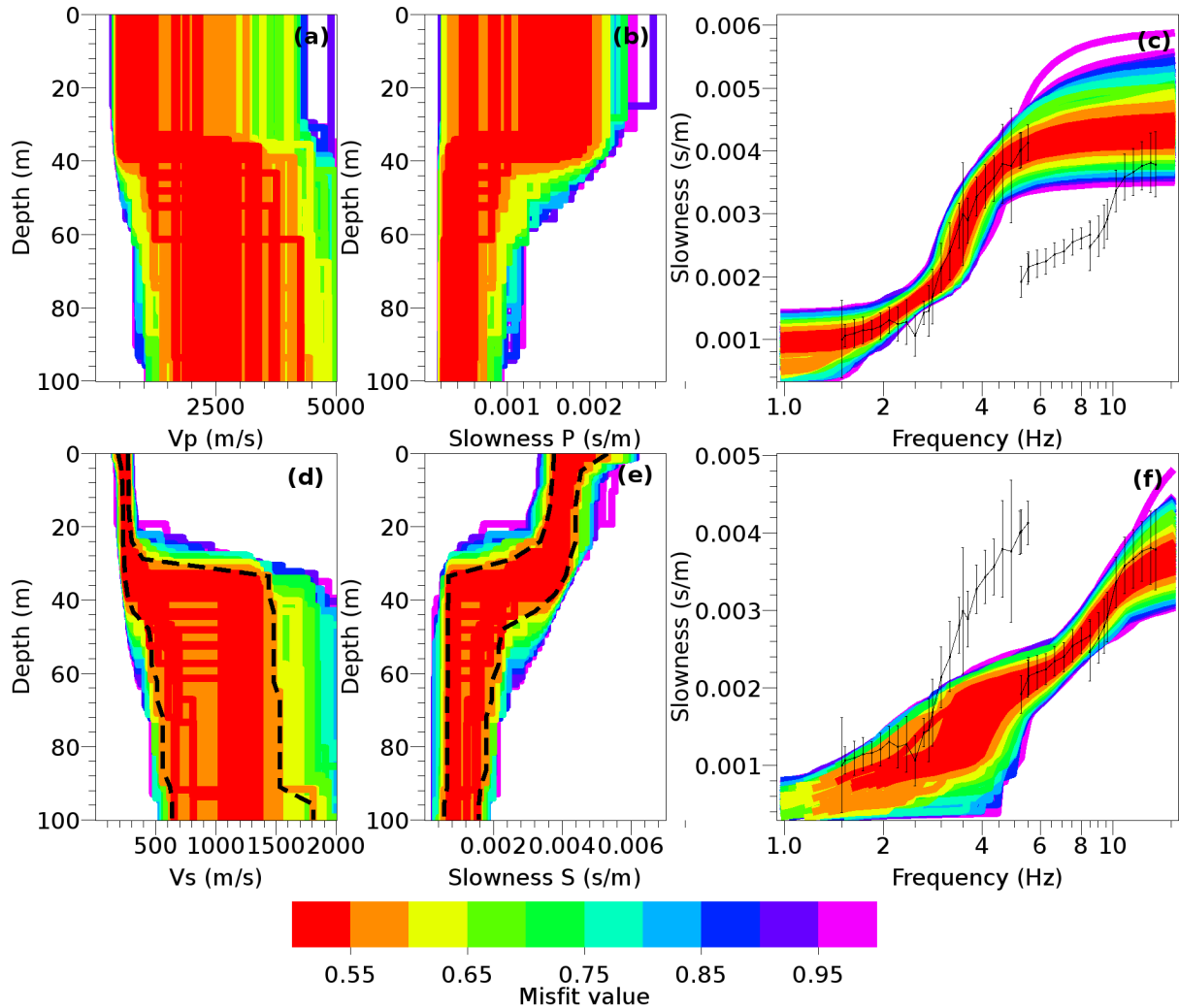


Figure 6. Inversions results for case N102: V_p profiles (a), V_s profiles (d), P slowness profiles (b), S slowness profiles (e), fundamental dispersion curve (c), and first higher mode (f). The dispersion curve used as target during inversion is plotted with thin black lines.

The inversion is processed with five independent runs (distinct random models which imply distinct sampling paths across the parameter space) to test the robustness of the results and to achieve a dense sampling around the solutions. Each run generates 5100 models. All models are plotted in figure 6. The colors represent the misfit values calculated on both modes.

With a relatively simple model, figure 6 show that the two modes of the dispersion can be reasonably fitted on their whole range. The corresponding minimum achieved misfit is 0.51. In this case, adding a new sub-layer in the sediment does not significantly improve the fit of the dispersion curve. The results of the three-layer parameterization are also plotted in figure 6. These high values for the misfit come from the irregularities of the fundamental mode around 2.5 Hz and the lack of fit of the higher mode for frequency greater than 10 Hz. According to the uncertainties of the observed on the dispersion curves, we consider that all models with a misfit lower than 0.6 are equally acceptable. The dotted lines in figure 6(d) and 6(e) delineate the minimum and maximum acceptable velocity profiles

In figure 6 (bottom right), at low frequency (below 4 Hz), the fundamental and the first higher modes can be very close to each other. Hence, we do not exclude that the

slowness observed below 3 Hz is a mix of two modes not resolved by the selected array layouts.

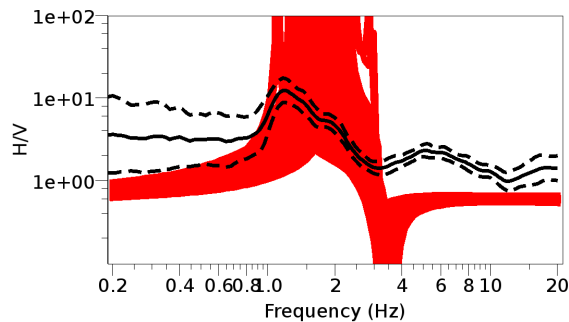


Figure 7. Comparison of fundamental ellipticities computed for all models (misfit less than 0.6) with the H/V spectra.

Finally, figure 7 shows the comparison ellipticities computed for the best models with the H/V spectra. It can be concluded that the ensemble of models generated as a solution has a resonance frequency corresponding to the H/V measurements. The significant width of the red area in figure 7 show that substantial improvement of the uncertainty can be achieved by a joint inversion of H/V peak frequency.

5. Conclusions

For all cases of this Noise Blind Test, dispersion curves are computed from FK, HRFK and SPAC methods with a good mutual agreement inside the valid wavenumber range of each array. The inversion of these curves lead in general to velocity profiles consistent with the resonance frequency observed on H/V spectra. However, no definitive conclusions can be drawn from this experience before any comparison with the true profiles (synthetic cases) or with existing borehole data (real cases).

6. References

- Asten, M. W. and J. D. Henstridge (1984). Array estimators and use of microseisms for reconnaissance of sedimentary basins, *Geophysics* 49, 1828-1837.
- Di Giulio, G., C. Cornou, M. Ohrnberger, M. Wathelet, and A. Rovelli (accepted in Bulletin of Seismological Society of America). 2-D small aperture arrays for velocity profile estimation using ambient seismic noise in a small-size alluvial basin (Colfiorito, Italy).
- Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm: I. Searching a parameter space, *Geophysical Journal International* 138, 479-494.
- Scherbaum, F., K.-G. Hinzen and M. Ohrnberger (2003). Determination of shallow shear wave velocity profiles in the Cologne/Germany area using ambient vibrations, *Geophysical Journal International* 152, 597-612.
- Tokimatsu, K. (1997). Geotechnical site characterization using surface waves, In *Proc. 1st Intl. Conf. Earthquake Geotechnical Engineering*, Ishihara (ed), Balkema, 1333-1368.
- Wathelet, M., D. Jongmans, and M. Ohrnberger (2004). Surface wave inversion using a direct search algorithm and its application to ambient vibration measurements, *Near Surface Geophysics* 2, 211-221
- Wathelet, M., D. Jongmans, and M. Ohrnberger (2005). Direct Inversion of Spatial Autocorrelation Curves with the Neighborhood Algorithm. *Bulletin of the Seismological Society of America*, 95, 1787-1800
- Wathelet, M. (2005). Array recordings of ambient vibrations: surface-wave inversion, Ph.D. Thesis, University of Liège, Belgium