Near Surface Geophysics, 2004, 211-221

Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements

M. Wathelet[1, 2*], D. Jongmans[1] and M. Ohrnberger[3]

[1] *LIRIGM, Université Joseph Fourier, BP 53, 38041 Grenoble cedex 9, France*

[2] *GEOMAC, Université de Liège, 1 Chemin des Chevreuils, Bât. B52, 4000 Liège, Belgium*

[3] *Institut für Geowissenschaften der Universität Potsdam, POB 601553, D-14415 Potsdam, Germany*

Received January 2004, revision accepted August 2004

ABSTRACT

Passive recordings of seismic noise are increasingly used in earthquake engineering to measure *in situ* the shear-wave velocity profile at a given site. Ambient vibrations, which are assumed to be mainly composed of surface waves, can be used to determine the Rayleigh-wave dispersion curve, with the advantage of not requiring artificial sources. Due to the data uncertainties and the non-linearity of the problem itself, the solution of the dispersion-curve inversion is generally non-unique. Stochastic search methods such as the neighbourhood algorithm allow searches for minima of the misfit function by investigating the whole parameter space. Due to the limited number of parameters in surface-wave inversion, they constitute an attractive alternative to linearized methods. An efficient tool using the neighbourhood algorithm was developed to invert the one-dimensional V_s profile from passive or active source experiments. As the number of generated models is usually high in stochastic techniques, special attention was paid to the optimization of the forward computations. Also, the possibility of inserting *a priori* information into the parametrization was introduced in the code.

This new numerical tool was successfully tested on synthetic data, with and without *a priori* information. We also present an application to real-array data measured at a site in Brussels (Belgium), the geology of which consists of about 115 m of sand and clay layers overlying a Palaeozoic basement. On this site, active and passive source data proved to be complementary and the method allowed the retrieval of a V_s profile consistent with borehole data available at the same location.

INTRODUCTION

For the majority of seismic prospecting methods, natural or cultural ambient vibrations constitute an undesired part of the signal, which has to be eliminated as much as possible. However, the noise field is influenced by the subsurface structure and the use of array records of seismic noise has been recognized as a method for deriving the S-wave velocity profile at a given site (e.g. Aki 1957; Asten 1978; Tokimatsu 1995). The hypothesis behind the method is that ambient vibrations mainly consist of surface waves, whose dispersion characteristics depend primarily on the body-wave velocities (V_p for compressional-wave velocities and V_s for shear-wave velocities), the density and the thickness of the different layers (Murphy and Shah 1988; Aki and Richards 2002). Noise energy depends upon the source locations and upon the impedance contrast between the rocky basement and the overlying soft sediments (Chouet *et al*. 1998; Milana *et al*. 1996). A knowledge of the shear-wave velocity (V_s) profile at a given site is of major importance in earthquake engineering, and ambient vibrations measured by an array of vertical sensors are increasingly applied for determining V_s profiles (e.g. Horike 1985; Tokimatsu 1995; Ishida *et al*. 1998; Miyakoshi *et al*. 1998; Yamamoto 1998; Satoh *et al*. 2001; Scherbaum *et al*. 2003).

In a first step, the Rayleigh phase-velocity dispersion curve is derived from the processing of simultaneous ground-motion recordings at various stations. The recording time is usually greater than or equal to half an hour and the number of stations is generally between 6 and 10, depending upon the available equipment (sensors, synchronized or multichannel stations) and time (the set-up may take quite a long time for a large number of sensors). The geometry of the station layout is not strictly imposed by the processing method itself, but a circular shape ensures an equal response of the array for waves coming from all azimuths. The common approaches used to derive the dispersion curve from the raw signals can be classified into two main families: frequency–wavenumber (Lacoss *et al*. 1969; Capon 1969; Kvaerna and Ringdahl 1986; Ohrnberger 2001) and spatial auto-

* marc.wathelet@ujf-grenoble.fr

correlation (Aki 1957; Roberts and Asten 2004). The first methods are best suited for plane wavefields with a single dominant source of noise but may be also used in more complex situations, averaging the apparent velocity over longer periods of time. The output of a basic frequency–wavenumber processing consists of semblance maps which indicate the azimuth and the velocity (or slowness) of the waves travelling with the highest energy. The grid maps are obtained by shifting and stacking the recorded signals over small time windows. The former class of methods assumes stationary-wave arrivals both in time and space and hence an infinite number of simultaneous sources. Spatial autocorrelation methods are considered as more efficient by some authors (e.g. Ohori *et al*. 2002), but the relative performances of each method have not been rigorously investigated so far. Here, we applied the frequency–wavenumber method which has been widely used (e.g. Asten and Henstridge 1984; Ohrnberger 2001).

At the second stage, the dispersion curve is inverted to obtain the V_s (and eventually the V_p) vertical profile, as in the classical active-source methods (Stokoe *et al*. 1989; Malagnini *et al*. 1995). Compared with these latter methods, noise-based techniques offer the following advantages (Satoh *et al*. 2001): (i) they can be easily applied in urban areas; (ii) they do not require artificial seismic sources; (iii) they allow greater depths to be reached (from tens of metres to hundreds of metres according to the array aperture and the noise-frequency content). Like all surface-wave methods, the geometry obtained is purely one-dimensional and is averaged within the array, implying that the technique is not suitable when strong lateral variations are present.

The derivation of 1D S-wave velocity profiles from surface-wave dispersion curves is a classical inversion problem in geophysics, usually solved using linearized methods (Nolet 1981; Tarantola 1987). In his computer program, Herrmann (1987) implemented a damped least-squares method that uses an analytical formulation for derivatives and a starting model. At each iteration, a better estimate of the model is calculated by linearizing the problem and the best solution, minimizing a misfit function, is obtained after a few iterations. If the misfit function exhibits several minima, which is usually the case when uncertainties on the dispersion curve are high, the derivative-based methods give a single optimal model which strongly depends upon the starting model. For active-source measurements, some authors proposed inverting the complete waveforms or particular wavefield transforms (Yoshizawa and Kennett 2002; Forbriger 2003) to get a better constraint on the solution. This is not applicable to ambient vibrations for which no information about the source properties is available. In geophysics, a new class of methods, based on uniform pseudo-random sampling of a parameter space (Monte-Carlo type), has emerged during the last 15 years: they are simulated annealing (Sen and Stoffa 1991), genetic algorithms (Lomax and Snieder 1994) and more recently the neighbourhood algorithm developed by Sambridge (1999). The objective of these algorithms is to investigate the whole parameter space, looking for good data-fitting sets of parameters.

In this work we have developed a new code using the neighbourhood algorithm for inverting dispersion curves. The software allows the inclusion of *a priori* information on the different parameters and a major effort has been made to optimize the computation time at the different stages of inversion. In particular, we have re-implemented the dispersion-curve computation in C++ language using Dunkin's (1965) formalism. The code is tested on synthetic cases as well as on one real data set, combining ambient vibrations and active-source data. In both cases, the role of *a priori* information for constraining the solution is emphasized.

INVERSION METHOD

The neighbourhood algorithm

The neighbourhood algorithm is a stochastic direct-search method for finding models of acceptable data fit inside a multi-dimensional parameter space (Sambridge 1999). For surface-wave inversion, the main parameters are the S-wave velocity, the P-wave velocity, the density and the thickness of each layer. Like other direct-search methods, the neighbourhood algorithm generates pseudo-random samples (one sample is one set of parameters corresponding to one ground model) in the parameter space and the dispersion curves are computed (forward problem) for all these models. The *a priori* density of probability is set as uniform over the whole parameter space, the limits of which are defined by the *a priori* ranges of all chosen parameters. The comparison of the computation results with the measured dispersion curve provides one misfit value that indicates how far the generated model is from the true solution. The originality of the neighbourhood algorithm is to use previous samples for guiding the search for improved models. Once the data misfit function is known at all previous samples (forward computations), the neighbourhood algorithm provides a simple way of interpolating an irregular distribution of points, making use of Voronoi geometry to find and investigate the most promising parts of the parameter space. For satisfactory investigation of the parameter space, the number of dispersion-curve computations can be very high (a few thousands to a few tens of thousands). The computation time has then to be optimized in order to obtain an efficient dispersion-curve inversion tool. Compared to other stochastic search methods (genetic algorithm, simulating annealing) the neighbourhood algorithm has fewer tuning parameters (only 2) and seems to achieve comparable or better results (Sambridge 1999). For poorly constrained parameters, the results may differ when starting two separate inversions. Hence, the robustness of the final results is generally checked by running the same inversion several times with different random seeds, an integer value that initializes the pseudo-random generator.

Dispersion-curve computation (forward problem)

The theoretical elastic computation of the dispersion curve for a stack of horizontal and homogeneous layers has been studied by Thomson (1950) and Haskell (1953) and has been modified by

Dunkin (1965) and Knopoff (1964). Only the Rayleigh phase velocities are considered here as the experimental dispersion curve is generally obtained from processing the vertical components of noise. As ambient vibrations may contain waves travelling in all directions, Love dispersion-curve computation requires the measurement of the two horizontal components and is much more difficult because records contain both Rayleigh and Love waves.

The dispersion-curve computation was carefully designed in order to reduce the computation time and to avoid misinterpretation of the different modes in particular cases. Together with a rewriting of Dunkin's (1965) formulae, we use an efficient root search, based on the Lagrange polynomial and constructed by iteration with Neville's method (Press *et al*. 1992). On a Pentium 1.7 GHz, the code that we have developed is able to compute the fundamental-mode dispersion curve of a single layer over a half-space with 30 samples in 850 microseconds (more than 1000 computations per second).

Parametrization of the model

The parametrization of the model (i.e. choosing the number of layers to invert) is not a straightforward problem. On the one hand, to avoid ill-posed problems, the number of parameters should be as low as possible; on the other hand, the parametrized model should include all possible classes of 1D structure able to match the complexity of the measured dispersion curve. Probably the best compromise is to start with the simplest model and progressively add new layers if the data are not sufficiently matched (Scherbaum *et al*. 2003). Obviously, the depth interval of the chosen parametrization should be consistent with the available frequency range of the dispersion curve. Estimations of the penetration depth based on one-third of the wavelength (Tokimatsu 1995) are useful but probably too restrictive. We prefer a trial-and-error approach, starting with large parameter ranges and focusing on the zones where the dispersion curve provides information.

Material density has a very small influence on dispersion curves and is generally fixed at a constant value in each layer, based on geotechnical information. V_p and V_s are linked together by Poisson's ratio that must lie between 0 and 0.5. We add an option to specify either the thickness or the depth of the bottom of each layer. This is a useful means of introducing some constraint on depth values (*a priori* information). The dispersion-curve computation is designed for layers with homogeneous properties. However, soft sediment compaction may induce a regular increase in V_p and V_s values as a function of depth (Bachrach *et al*. 2000). Scherbaum *et al*. (2003) assumed a power-law variation within the sedimentary column in the Lower Rhine Embayment (Germany) to reduce the number of parameters. The velocity (either V_p or V_s) at depth z_i is given by

$$V_i = V_0\left((z_n+1)^{\alpha} - (z_0+1)^{\alpha} + 1\right), \quad (1)$$

where z_0 is the top of the layer considered, V_0 is the velocity at z_0 and α is the power-law exponent, generally varying between 0 and 1.

For dispersion-curve computations, the function $V_i(z)$ (equation (1)) is discretized into a fixed number of homogeneous sublayers. Their number is generally kept as low as possible (between 5 and 10) to avoid a drastic increase in the inversion computation time.

Misfit definition

Once the theoretical dispersion curve has been calculated from the random parameters given by the neighbourhood algorithm, the misfit value must be evaluated. If the data curves are given with an uncertainty estimate, the misfit is given by ,

$$misfit = \sqrt{\sum_{i=0}^{n_F} \frac{(x_{di} - x_{ci})^2}{\sigma_i^2 n_F}}, \quad (2)$$

where x_{di} is the velocity of the data curve at frequency f_i, x_{ci} is the velocity of the calculated curve at frequency f_i, σ_i is the uncertainty of the frequency samples considered and n_F is the number of frequency samples considered. If no uncertainty is provided, σ_i is replaced by x_{di} in equation (2).

SYNTHETIC DATA

The code was tested on many synthetic cases, of which two are presented here. The first case is a homogeneous single layer (V_s=500 m/s) overlying a half-space. The V_p and V_s profiles are shown in Figs 1(a) and 1(b), while the dispersion curve for the fundamental mode and the first higher mode are shown in Fig. 1(c). The period range is between 0.04 s (25 Hz) and 0.3 s (3 Hz), covering the major part of the dispersion curve for the fundamental mode. As mentioned above, even for this simple model, the two modes are very close around 0.12 s, which is likely to create problems for computing the dispersion curve correctly (risk of jumping modes). The second model Figs 1(d) and 1(e) has two layers overlying a half-space, one of which exhibits a power-law variation of the velocity. The corresponding dispersion curves are plotted in Fig. 1(f).

Figures 1(g) and 1(h) show the fundamental Rayleigh ellipticity functions, i.e. the spectral ratio of the horizontal over the vertical eigenfunctions (Aki and Richards 2002) for the two models. In most situations, ellipticities are similar to the well-known *H*/*V* ratios (Fäh *et al*. 2001), which are increasingly used in earthquake engineering for determining site response properties. This *H*/*V* ratio frequently exhibits a dominant peak that experimentally was often found to coincide with the resonance frequency of the site (Tokimatsu 1995; Bard 1998). The theoretical relationship between these two parameters is still a matter of fundamental research (Malischewsky and Scherbaum 2004). On the other hand, Scherbaum *et al*. (2003) recently demonstrated that the soil structure acts as a high-pass filter on the vertical component of the motion, with a threshold frequency close to the frequency of the ellipticity peak. This means that, in practice, it will be very difficult to obtain the Rayleigh dispersion curve below the fre-

quency of the peak. For the two cases, the ellipticiy exhibits a peak around 6 Hz Figs 1(g) and 1(h) and the dispersion curves will be cut below 6 Hz (0.16 s) in the following inversions. This is a severe limitation for constraining the model parameters at greater depths.

The dispersion curve of the fundamental mode for the first case white dots on Fig. 2(c) was inverted with eight distinct and independent runs, simultaneously started with different random seeds, and generating 8×5050 models. The parameters are the V_p and V_s values within each layer and the thickness of the first layer. The V_p and V_s profiles resulting from the inversion, as well as the corresponding dispersion curves, are plotted in Figs 2(a) and 2(b), with a grey scale indicating the misfit value. V_s is better retrieved than V_p because the dispersion curve is more sensitive to V_s than V_p, especially for the bottom half-space where V_p has negligible influence. For the best models (lowest misfit values), the dispersion-curve data are very well fitted and the five parameters are correctly retrieved, although the properties of the basement are less constrained. These greater uncertainties for the bottom layer are due to the lower limit imposed on the frequency range, usual in real data. Figures 2(d)–2(f) present the results for the same case but with an extension of the dispersion curve towards lower frequencies. We observe a far better retrieval of the bottom V_s whose relative uncertainty is similar to the superficial uncertainties (about 6% on a velocity scale). At 30 m, Fig. 2(b) shows an irregular velocity distribution with a trough at 2050 m/s. Additional computations (not shown here) show that the misfit increases around the best model (V_s =1750 m/s) is regular and that the irregularities are due to a lack of sampling within the parameter space. Numerical tests on various models demonstrate that the inversion process works for a large range of V_s and V_p values (Poisson's ratio from 0 to 0.49).

The dispersion curve for the second case (three layers, Fig. 1) was also inverted with eight distinct and independent runs. The parameters are the bottom depth, V_p, V_s, and optionally the power-

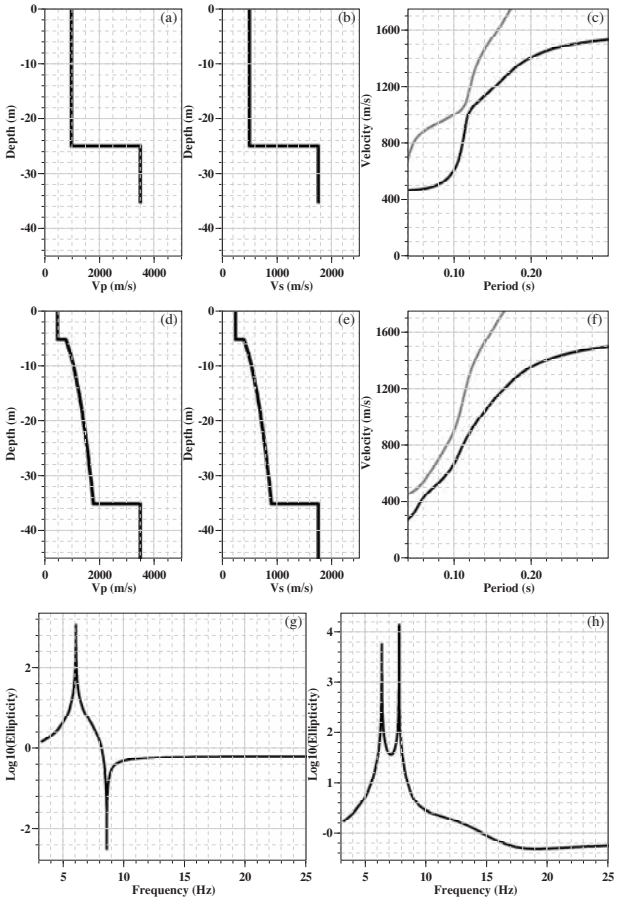

FIGURE 1

Synthetic data computed with the implemented forward algorithm for two cases: a single layer over a half-space and two layers over a half-space. (a) V_p and (b) V_s profiles for the single-layer model. (c) The corresponding dispersion curves for the fundamental mode (black line) and the first higher mode (grey line). (d) V_p and (e) V_s profiles for the two-layer model. (f) The corresponding dispersion curves for the fundamental mode (black line) and the first higher mode (grey line). (g) Ellipticity of the fundamental Rayleigh mode for the single layer and (h) for two layers.

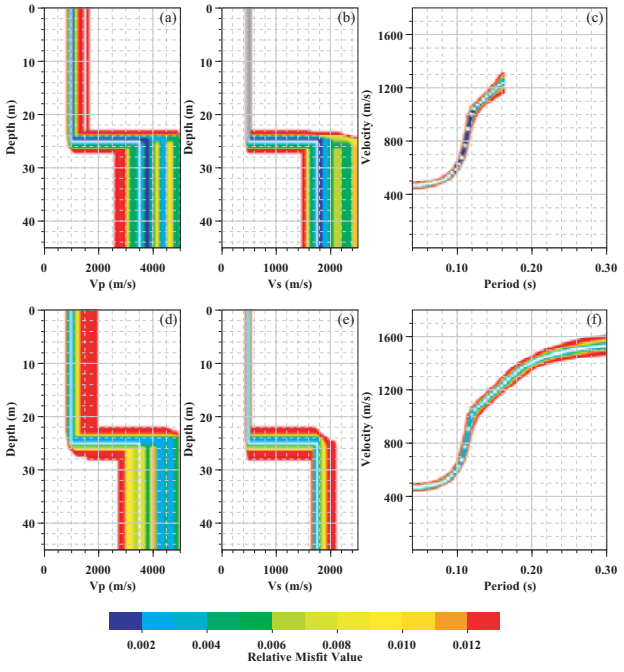

FIGURE 2

Inversion of the dispersion curve (between 0.04 s and 0.16 s) for two-layer synthetic data (fundamental-mode dispersion curve from Fig. 1f) with the neighbourhood algorithm. The white lines on (a) and (b) are the true velocity profiles; the resulting (a) V_p and (b) V_s profiles are obtained with free parameters. (c) Dispersion curves calculated for each individual model plotted on (a) and (b). The white dots refer to the original true dispersion curve. (d), (e) and (f).

law exponent for each layer (common for V_p and V_s), adding four new parameters compared to the preceding case. Due to the larger parameter space (dimension 9), the number of generated models is increased to 8×15050. The inversion results are shown in Figs 3–5. Figures 3(a)–3(c) show the V_p and V_s profiles, as well as the comparison between the calculated curves and the dispersion-curve data. A good fit is observed between the data and many of the models, indicating a problem of non-uniqueness of the inversion. The result is that, except for the first-layer characteristics, the other parameters are poorly constrained. In particular, for the best models, the depth of the deeper layer can range from 20 to 60 m, according to the velocity profile.

Figure 4(a) shows some 2D projections of the misfit function within the parameter space. Instead of a distinct minimum, the graphs show patches characterized by the same misfit value, with clear correlations between some parameters (V_s and thickness of the first layer, for instance). On the other hand, the misfit values obtained are real minima, as shown by the evolution of the misfit value as a function of the number of generated models (Fig. 5).

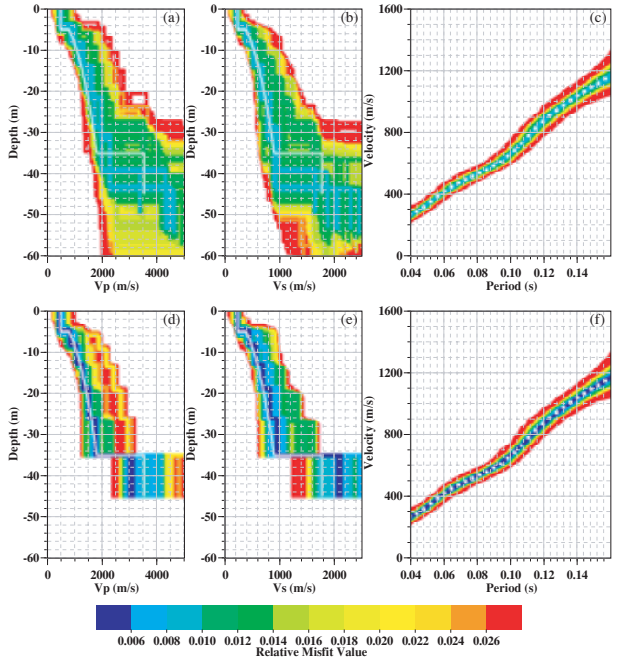

FIGURE 3

Inversion of the three-layer synthetic data (fundamental-mode dispersion curve from Fig. 1(f) with the neighbourhood algorithm. The white lines on plots (a), (b), (d) and (e) are the true velocity profiles; the resulting (a) V_p and (b) V_s profiles are obtained with all thicknesses, gradients and velocities left as free parameters. (c) Dispersion curves calculated for each individual model plotted on (a) and (b). The dots refer to the original true dispersion curve that the inversion tends to fit. The resulting (d) V_p and (e) V_s profiles are obtained when the depth of the top of the half-space is fixed at its true value (*a priori* information). (f) As (c) but for models from (d) and (e).

Convergence was obtained for all the eight seeds, although the investigation paths through the parameter space were not the same. These results show that the dispersion curve alone is unable to constrain the velocity profiles. The possibility of introducing *a priori* information was then tested by fixing the depth (35 m) of the bottom layer (given by borehole data, for instance). The inversion was re-run with the same characteristics and the results are plotted in Figs 3(d)–3(f) and 4(b). The reduction of the dimension of the parameter space and the introduction of reliable *a priori* information allow both the achievement of a better fit of the dispersion-curve data and good confinement of the velocity profiles in the two upper layers. Other tests (not shown here), performed with a weaker constraint on the depth of the bottom layer (depth

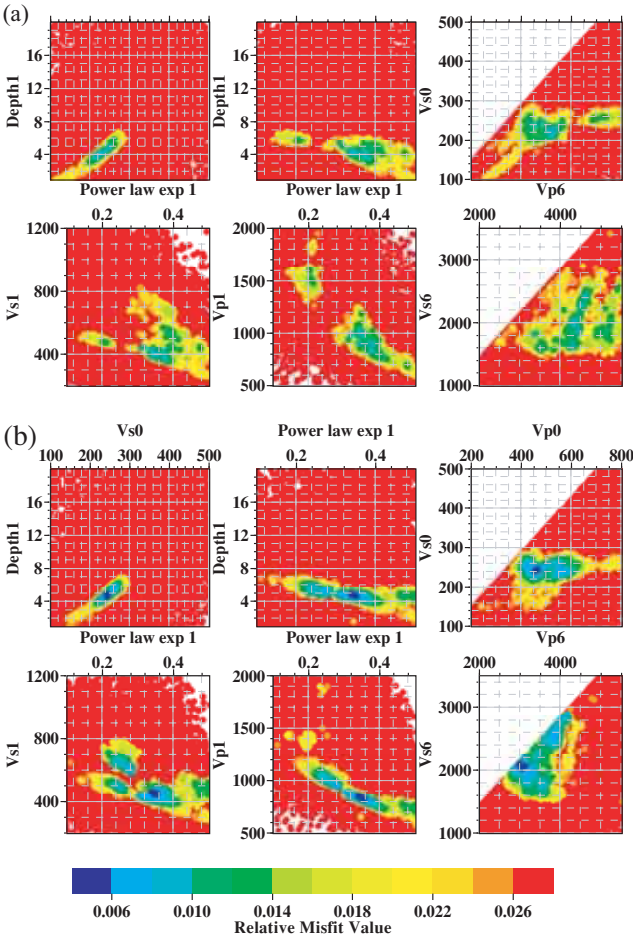

FIGURE 4

Misfit value as a function of two parameters for the two inversions of Fig. 3: (a) all free parameters; (b) *a priori* information about the depth of the top of the half-space. 'Depth1' denotes the depth of the top of the second layer, 'Depth6' denotes the top of the half-space, 'Power law exp 1' denotes the exponent α in equation (1) for the second layer, 'Vp0' and 'Vs0' denote the velocities of the first layer, 'Vp1' and 'Vs1' denote the velocities of the second layer, and 'Vp6' and 'Vs6' denote the velocities of the half-space.

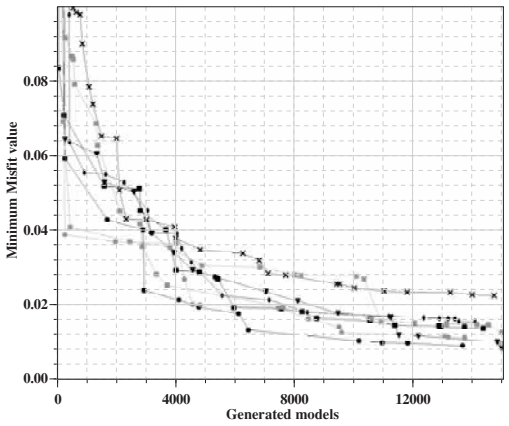

FIGURE 5

Convergence history of the inversions of Figs 3(a)–3(c) and 4(a). The parameter space representation has been constructed with eight independent runs (distinct random seeds).

range between 32 m and 38 m), led to similar results. Looking at Fig. 4(b), the introduction of the depth constraint permits a clear minimum to appear in the general shape of the misfit function. Also, good-fitting models with V_p values exceeding 4000 m/s were removed by the *a priori* information.

REAL DATA

The whole process of deriving velocity profiles from ambient vibration recordings was applied at a site located in the south of Brussels, Belgium, inside the park of the Royal Observatory of Belgium (50°47'56''N-04°21'33''E; Fig. 6(a). The topography is almost flat and the soil structure mainly consists of a succession of sand and clayey-sand horizontal layers overlying a Palaeozoic bedrock (the so-called Brabant Massif). The same structure extends to the north-west towards the North Sea with a regular increase in the total thickness of the sediments corresponding to deepening of the Palaeozoic substratum (Nguyen *et*

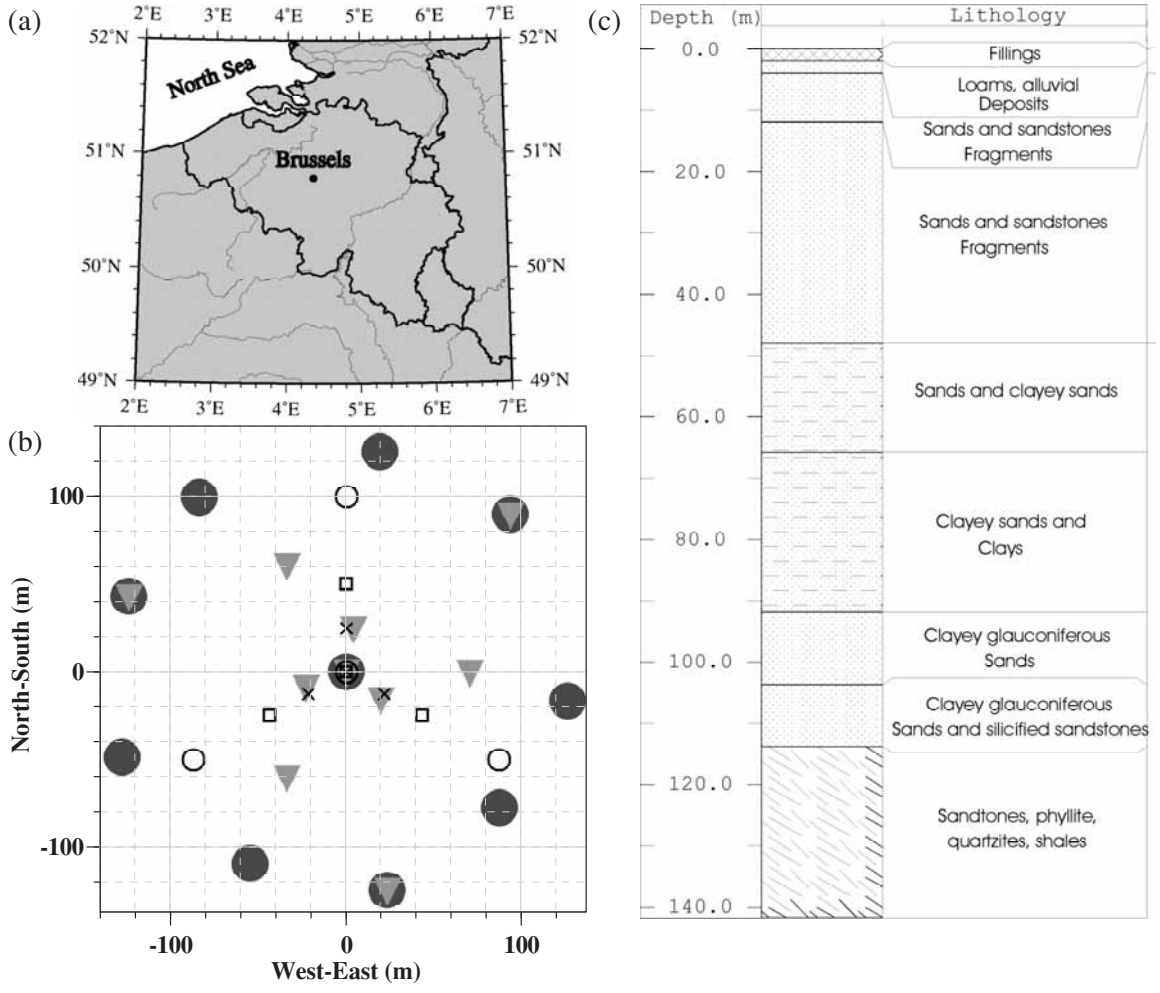

FIGURE 6

Location map for the real case, park of the Royal Observatory of Belgium in Uccle, Brussels (Belgium). (a) Regional map. (b) Site map including the location of all the seismic stations. Six non-simultaneous arrays were recorded: 'radius 130' (large dark-grey circles), 'radii 25-75-130' (large light-grey triangles), 'radius 100' (empty circles), 'radius 50' (empty squares) and 'radius 25' (crosses) (c) Borehole description.

TABLE 1
Description of noise arrays recorded at the Brussels site

Array name	Number of stations	Geometry	Recorded time	Processed time
Radius 130	10	1 central station and 9 stations equally distributed on a circle (radius 130 m)	40 min	30 min
Radii 25-75-130	10	1 central station and 3 stations on each of the circles with radii 25, 75 and 130 m	1h 57 min	30 min
Radius 100	4	1 central station and 3 stations equally distributed on a circle (radius 100 m)	1h 01 min	25 min
Radius 50	4	1 central station and 3 stations equally distributed on a circle (radius 50 m)	59 min	30 min
Radius 25	4	1 central station and 3 equally distributed on a circle (radius 25 m)	1h 02 min	20 min

TABLE 2
Minimum and maximum distances between stations for the arrays described in Table 1. Minimum and maximum frequencies of the valid range as observed in the resulting dispersion curves. Minimum and maximum frequencies calculated from the wavelength criteria (equation (3)) for a velocity of 400 m/s

Array name	Minimum distance	Maximum distance	Min. obs. frequency	Max. obs. frequency	Min. theo. frequency	Max. theo. frequency
Radius 130	69 m	260 m	1.04 Hz	2.78 Hz	0.51 Hz	2.90 Hz
Radii 25-75-130	25 m	223 m	1.19 Hz	3.13 Hz	0.90 Hz	8 Hz
Radius 100	100 m	173 m	1.04 Hz	2.38 Hz	0.77 Hz	2 Hz
Radius 50	50 m	87 m	1.32 Hz	3.13 Hz	1.53 Hz	4 Hz
Radius 25	25 m	43 m	2.27 Hz	4.16 Hz	3.10 Hz	8 Hz

al. 2004). A deep borehole, located in the park, provides a good estimate of the bedrock depth (115 m) and shows that the basal part of the soil column (between 70 and 115 m) is more clayey than the upper part Fig. 6(c).

Figure 6(b) shows a site map with the five arrays of sensors (Lennartz, 5 s resonance period) that were measured and whose characteristics are detailed listed in Table 1. The performance of an array in estimating phase velocities depends on the ratio of the sensor spacing and the wavelength. Tokimatsu (1995) provided the following rules:

$$2D_{\min} < \lambda_{\min} < \lambda_{\max} < 3D_{\max}, \quad (3)$$

where $D_{\min}$ and $D_{\max}$ are the minimum and maximum distances between stations and $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum wavelengths. The relationship between $D_{\min}$ and $\lambda_{\min}$ is to avoid an aliasing effect while the other one is derived empirically and could partially result from the filtering effect of the site. The estimated valid frequency ranges for the five arrays are given in Table 2 and they are compared to the limits of equation (3). In addition to theoretical limits, we estimated the valid range of experimental dispersion curves by taking into account two criteria: (i) the phase velocity must be approximately the same for every non-overlapping time window; (ii) large (small) aperture arrays are more reliable at lower (higher) frequencies. From Table 2, the theoretical limits approximately follow the observed values but they cannot be taken routinely to define the valid range of an experimental dispersion curve.

For the five arrays, a frequency–wavenumber analysis was applied to the recordings. Figure 7 details intermediate results for array 'radii 25-75-130'. For each frequency, the signals are divided into small time windows and the apparent velocity is estimated from each of them. From these statistical samples, histograms (one per period, normalized in the slowness domain) are drawn in Fig. 7(a). The median dispersion curve is plotted with its median deviation (calculated with the same statistical samples). Figures 7(b) and 7(c) show slowness maps stacked over all available time windows for two separate frequencies (1.8 and

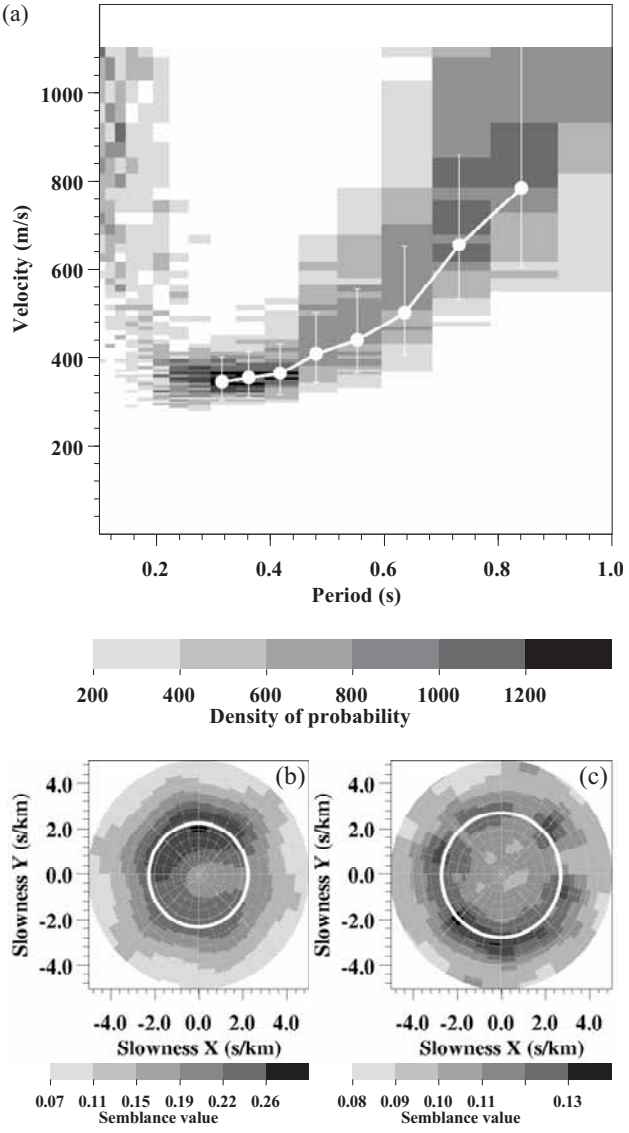

FIGURE 7

Dispersion curve from frequency–wavenumber processing for array 'radii 25-75-130'. (a) Histograms of observed apparent velocities for each period. They are normalized in the slowness domain (surface below curve = 1). The white curve indicates the median with the median deviation. (b) and (c).

4.2 Hz). The maxima are located at the velocity of the more energetic wavetrain travelling across the array (white circles). When maps are stacked over long periods of time, they provide the average velocity and the azimuth distribution. The computed dispersion curves with their error bars are shown in Fig. 8 inside their validity ranges. The combination of five arrays leads to a good definition of the dispersion curve from 1.04 Hz to 4.17 Hz (or 0.24 to 0.96 s). The grey line with circles is the average of the five curves and was used for the inversion. Dispersion curve results from active-source experiments (explosions, 24 vertical

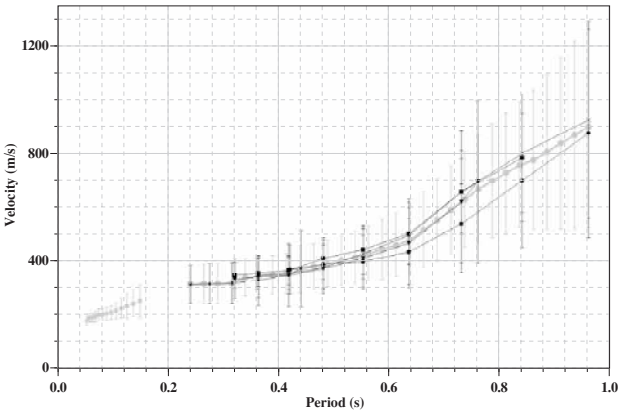

FIGURE 8

Dispersion curves calculated from recorded signals using the frequency–wavenumber technique: 'radius 130' (crosses), 'radii 25-75-130' (squares), 'radius 100' (black circles), 'radius 50' (triangles), 'radius 25' (diamonds). The grey dots constitute the average curve used for inversion. The grey squares were obtained with active-source experiments on the same site.

receivers with a threshold frequency of 4.5 Hz, placed every 5 m) performed on the same site (Nguyen *et al*. 2004) are also shown in Fig. 8. The frequency ranges are clearly distinct and the two methods seem to complement each other.

First, the dispersion curve obtained from ambient vibrations was inverted alone, using the neighbourhood algorithm. The measured phase-velocity dispersion curve Fig. 9(c) exhibits a regular shape and a two-layer model (one soil layer with a power-law variation of the velocity over the substratum) was used. The velocity profiles after inversion show that only the shear-wave velocity down to 80 m is well resolved, with an increase from 200 m/s at the surface to 500 m/s at 80 m. With regard to the substratum depth, the best models show a velocity increase between depths of 130 m and 150 m, although other models cannot be disregarded. V_p values in the first 40 m previously obtained during refraction lines (Nguyen *et al*. 2004) compare well with the dispersion-curve inversion results Fig. 9(a). In a second step, we introduced the substratum depth (115 m) given by the borehole as *a priori* information in the inversion process. As in the synthetic case, it results in a significant improvement in the constraint on the V_s profile, while a larger variation is still possible on V_p values. Even if the results were satisfactory, we also tested the inversion process with a three-layer structure, imposing the bedrock depth and the presence of a shallow velocity contrast between 10 m and 100 m. Inversion Figs 10(a)–10(c) led to similar velocity profiles, except at shallow depths where velocity values exhibited large variations. This lack of resolution results from the absence of data at short periods. In the preceding case Fig. 9(e), velocity values close to the surface were constrained by the power-law relationship over the whole layer. If we now combine the active-source measurements with the low-fre-

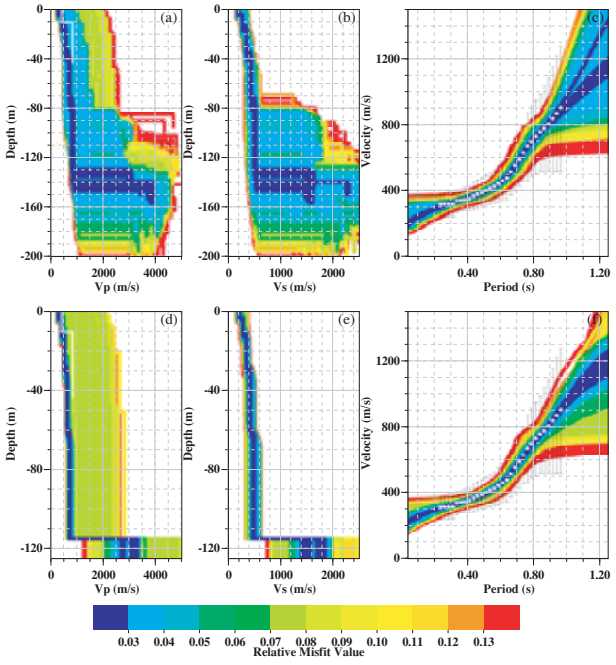

FIGURE 9

Inversion of the real case: (a), (b) and (c) without *a priori* information, and (d), (e) and (f) taking the depth measured in a deep borehole (around 115 m) as a constraint on the top of the half-space. (a) and (d) depict the V_p profiles, (b) and (e) the V_s profiles. For each case, (c) and (f) contain the dispersion curves for the fundamental Rayleigh mode corresponding to the models of (a) and (b), or (d) and (e), together with the experimental dispersion curve and its uncertainty (as Fig. 8). The white line represents the V_p profile measured by surface-refraction experiments (on (a) and (d)).

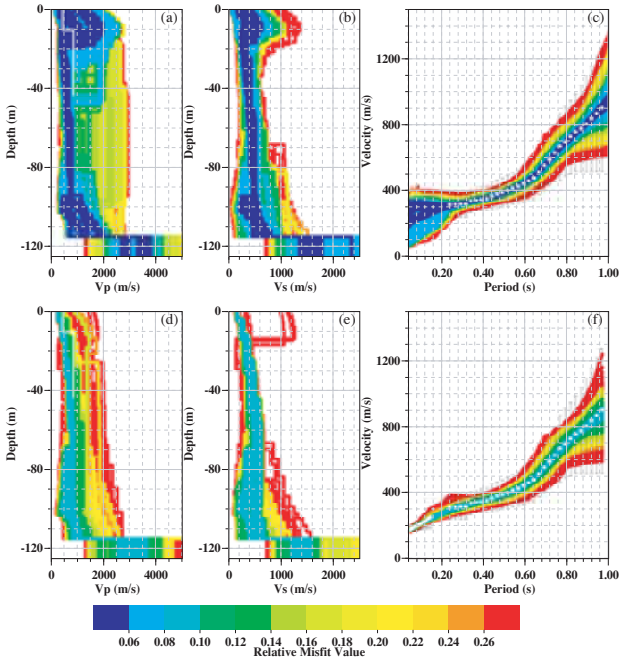

FIGURE 10

Inversion of the noise array data with and without active-source experiments: (a) and (d) the resulting V_p profiles; (b) and (e) the resulting V_s profiles; (c) and (f) the computed fundamental Rayleigh dispersion curves for models from (a), (d) and (b), (e). The dots on (c) and (f) represent the experimental curves. The white line represents the V_p profile measured by surface refraction experiments (on (a) and (d)).

quency information from the noise array Figs 10(d)–10(f), a relatively well-constrained V_s profile is found, very similar to the one shown in Fig. 9(e). The V_s profile shows a constant velocity (or even a slight inversion) in the soil column below 60 m, a limit which corresponds to the presence of clayey layers. This application shows that ambient vibrations and active-source recordings complement one another in the derivation of Rayleigh phase-velocity dispersion curves over a large period range.

Finally, for each preceding inversion Figs 9(a), 9(b), 10(a), 10(b), the fundamental Rayleigh ellipticity curve was calculated for the model with the lowest misfit. The results are shown in Fig. 11, together with the measured H/V spectral ratios, which have a peak frequency at about 1 Hz, in accordance with the results from Nguyen *et al*. (2004). The inverted models all show a frequency peak around 1 Hz, in agreement with the available H/V data.

CONCLUSIONS

A new computer code based on the neighbourhood algorithm was developed for the inversion of Rayleigh dispersion curves, with the aim of integrating any data resulting from the process-

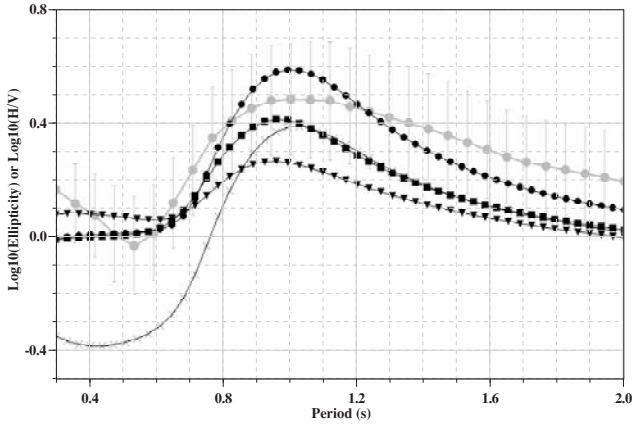

FIGURE 11

Comparison of the measured H/V spectral ratio (grey circles, including standard deviation) with the calculated ellipticities from the best model (with the lowest misfit) of Figs 9(a)-9(c) (black circle), Figs 9(d)-9(f) (black squares), 10(a)-10(c) (crosses) and Figs 10(d)-10(f) (black triangles).

ing of active-source experiments or ambient-noise recordings. Much effort was devoted to the optimization of the computation time, particularly in the calculation of the dispersion curve, as thousands of models have to be computed. A flexible parametrization, including a velocity variation inside the layers, has contributed to the reduction in the number of parameters, allowing a better investigation of the parameter space with a direct-search algorithm such as the neighbourhood algorithm. The software was also designed to allow the introduction of *a priori* information.

The method was successfully tested on several synthetic data sets, two of which have been presented here. These tests showed the efficiency of the developed tool and the limits of the dispersion-curve inversion alone. The introduction of *a priori* information when available is of major importance in constraining the solution.

The real-case analysis also proved the applicability and the reliability of the method. The introduction of borehole data (depth of the substratum) also considerably improved the results.

Combining active and passive seismic sources proved to be very helpful at the Brussels site, allowing V_s values to be obtained down to the bedrock, located at more than 100 m depth. The inversion led to a velocity profile which agrees with the borehole log.

In the future, we plan to introduce into the inversion the peak frequencies of the ellipticity, which were found to constrain the layer thickness in the case of the single layer over a half-space (Scherbaum *et al*. 2003) and could supplement borehole data when they are not available.

ACKNOWLEDGEMENTS

We are grateful to Frank Scherbaum for sharing his experience and his software for noise processing. Daniel Vollmer, Hans Havenith, Thierry Camelbeek and the team of the Royal Observatory of Belgium are acknowledged for their support during field experiments. We also thank M. Sambridge who allows the free distribution of his inversion code for research purposes. This study was made possible through the support of the SESAME European project ("Site EffectS assessment using AMbient Excitation", Project EVG1-CT-2000-00026).

REFERENCES

Aki K. 1957. Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute* **35**, 415–456.

Aki K. and Richards P.G. 2002. *Quantitative Seismology*. 2nd edition, University Science Books.

Asten M.W. 1978. Geological control on the three-component spectra of Rayleigh-wave microseism. *Bulletin of the Seismological Society of America* **68**, 1623–1636.

Asten M.W. and Henstridge J.D.1984. Array estimators and use of microseisms for reconnaissance of sedimentary basins. *Geophysics* **49**, 1828-1837.

Bachrach R., Dvorkin J. and Nur M.A. 2000. Seismic velocities and Poisson's ratio of shallow unconsolidated sands. *Geophysics* **65**, 559–564.

Bard P.-Y. 1998. Microtremor measurements: A tool for site effect estimation? In: *The Effect of Surface Geology on Seismic Motion* (eds Irikura, Kudo, Osaka and Sasatani). Balkema.

Capon J. 1969. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE* **57**, 1408–1418.

Chouet B., De Luca G., Milana G., Dawson P., Martini M. and Scarpa R. 1998. Shallow velocity structure of Stromboli Volcano, Italy, derived from small-aperture array measurements of strombolian tremor. *Bulletin of the Seismological Society of America* **88**, 653–666.

Dunkin J.W. 1965. Computation of modal solutions in layered, elastic media at high frequencies. *Bulletin of the Seismological Society of America* **55**, 335–358.

Fäh D., Kind F. and Giardini D. 2001. A theoretical investigation of average H/V ratios, *Geophysical Journal International* **145**, 535–549.

Forbriger T. 2003. Inversion of shallow-seismic wavefields. Part 2: Infering subsurface properties from wavefield transforms. *Geophysical Journal International* **153**, 735–752.

Haskell N.A. 1953. The dispersion of surface waves on a multi-layered medium. *Bulletin of the Seismological Society of America* **43**, 17–34.

Herrmann R.B. 1987. *Computer Programs in Seismology*. St Louis University.

Horike M. 1985. Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas. *Journal of Physics of the Earth* **33**, 59–96.

Ishida H., Nozawa T. and Niwa M. 1998. Estimation of deep surface structure based on phase velocities and spectral ratios of long-period microtremors. 2nd International Symposium on the Effect of Surface Geology on Seismic Motion, Yokohama, Japan, **2**, pp. 697–704.

Knopoff L. 1964. A matrix method for elastic wave problems. *Bulletin of the Seismological Society of America* **54**, 431–438.

Kvaerna T. and Ringdahl F. 1986. Stability of various fk-estimation techniques. In: Semiannual Technical Summary, 1 October 1985 – 31 March 1986, NORSAR Scientific Report, 1-86/87, Kjeller, Norway, pp. 29–40.

Lacoss R.T., Kelly E.J. and Toksöz M.N. 1969. Estimation of seismic noise structure using arrays. *Geophysics* **34**, 21–38.

Lomax A.J. and Snieder R. 1994. Finding sets of acceptable solutions with a genetic algorithm with application to surface wave group dispersion in Europe. *Geophysical Research Letters* **21**, 2617–2620.

Malagnini L., Herrmann R.B., Biella G. and de Franco R. 1995. Rayleigh waves in Quaternary alluvium from explosive sources: determination of shear-wave velocity and Q structure. *Bulletin of the Seismological Society of America* **85**, 900–922.

Malischewsky P.G. and F. Scherbaum. 2004. Love's formula and H/V-ratio (ellipticity) of Rayleigh waves, Wave Motion **40**, 50–67

Milana G., Barba S., Del Pezzo E. and Zambonelli E. 1996. Site response from ambient noise measurements: new perspectives from an array study in Central Italy. *Bulletin of the Seismological Society of America* **86**, 320–328.

Miyakoshi K., Kagawa T. and Kinoshita S. 1998. Estimation of geological structures under the Kobe area using the array recordings of microtremors. 2nd International Symposium on the Effect of Surface Geology on Seismic Motion, Yokohama, Japan, **2**, pp. 691–696.

Murphy J.R. and Shah H.K. 1988. An analysis of the effects of site geology on the characteristics of near-field Rayleigh waves. *Bulletin of the Seismological Society of America* **78**, 64–82.

Nguyen F., Van Rompaey G., Teerlynck H., van Camp M., Jongmans D. and Camelbeeck T. 2004. Use of microtremor measurement for assessing site effects in Northern Belgium interpretation of the observed intensity during the M_S= 5.0 June 11 1938 earthquake. *Journal of Seismology* **8**, 41–56.

Nolet G. 1981. Linearized inversion of (teleseismic) data. In: *The Solution of the Inverse Problem in Geophysical Interpretation* (ed. R. Cassinis), pp. 9–37. Plenum Press.

Ohori M., Nobata A. and Wakamatsu K. 2002. A comparison of ESAC and FK methods of estimating phase velocity using arbitrarily shaped microtremor arrays. *Bulletin of the Seismological Society of America* **92**, 2323–2332.

Ohrnberger M. 2001. *Continuous automatic classification of seismic signals of volcanic origin at Mt Merapi, Java, Indonesia*. Dissertation, University of Potsdam.

Press W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P. 1992. *Numerical Recipes in Fortran*, 2nd edition. Cambridge University Press.

Roberts J.C. and Asten M.W. 2004. Resolving a velocity inversion at the geotechnical scale using the microtremor (passive seismic) survey method. *Exploration Geophysics* **35**, 14–18.

Sambridge M. 1999. Geophysical inversion with a neighbourhood algorithm I. Searching a parameter space. *Geophysical Journal International* **103**, 4839–4878.

Satoh T., Kawase H. and Matsushima S.I. 2001. Differences between site characteristics obtained from microtremors, S-waves, P-waves, and codas. *Bulletin of the Seismological Society of America* **91**, 313–334.

Scherbaum F., Hinzen K.-G. and Ohrnberger M. 2003. Determination of shallow shear wave velocity profiles in the Cologne/Germany area using ambient vibrations. *Geophysical Journal International* **152**, 597–612.

Sen M.K. and Stoffa P.L. 1991. Nonlinear one-dimensional seismic waveform inversion using simulated annealing. *Geophysics* **56**, 1624–1638.

Stokoe K.H.II, Rix G.J. and Nazarian S. 1989. In situ seismic testing with surface waves. Proceedings of the XII International Conference on Soil Mechanics and Foundation Engineering, pp. 331–334.

Tarantola A. 1987. *Inverse Problem Theory*. Elsevier Science Publishing Co.

Thomson W.T. 1950. Transmission of elastic waves through a stratified solid medium. *Journal of Applied Physics* **21**, 89–93.

Tokimatsu K. 1995. Geotechnical site characterization using surface waves. In: *Earthquake Geotechnical Engineering* (ed.Ishihara), pp. 1333–1368. Balkema, Rotterdam.

Yamamoto H. 1998. An experiment for estimating S-wave velocity structure from phase velocities of Love and Rayleigh waves in microtremors. 2nd International Symposium on the Effect of Surface Geology on Seismic Motion, Yokohama, Japan, **2**, pp. 705–710.

Yoshizawa K. and Kennett B.L.N. 2002. Non-linear waveform inversion for surface waves with a neighbourhood algorithm – application to multimode dispersion measurements. *Geophysical Journal International* **149**, 118–133.