Geophys. J. Int. (2007) doi: 10.1111/j.1365-246X.2006.03282.x

Combining genetic and linearized algorithms for a two-step joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves

Matteo Picozzi and Dario Albarello

Department of Sciences of the Earth, University of Siena, 53100 *Siena, Italy. E-mail: picozzi@unisi.it*

Accepted 2006 November 2. Received 2006 November 2; in original form 2006 April 4

SUMMARY

The joint inversion of Rayleigh wave dispersion and H/V curves from environmental noise measurements allows the retrieval of *S*-wave velocity profiles for the shallow subsoil. For this purpose, genetic and linearized algorithm have been combined in a two-step inversion procedure, that allows the principal drawbacks typical of the application of each algorithm separately to be overcome. In the first step, a genetic algorithm procedure is used to constrain the subvolume of the parameter space where the absolute minimum of the misfit function is located. In the second step, a linearized inversion algorithm, having as an initial guess the minimum misfit model deduced from the first step, is applied to force the inversion towards the optimal solution. To evaluate the feasibility and effectiveness of this approach, seismic noise recordings at a test site in the Po river valley (North Italy) have been analysed. Here, detailed geophysical and geological information is available along with earthquake recordings, which allow a well constrained definition of both the local shear wave profile and transfer function. Comparisons between theoretical and experimental *S*-wave velocity profiles and, above all, between the theoretical and experimental site response functions shows that this combination of inversion procedures can very efficiently to manage the extreme non-linearity of the problem.

Key words: inversion, Love waves, Rayleigh waves, seismic array, seismic noise, surface waves.

INTRODUCTION

The local shear wave velocity structure and total thickness of the overlying sedimentary cover are essential parameters controlling the local amplification of ground motion during earthquakes. A number of procedures suitable for mapping the mechanical properties of the shallow subsoil (microzonation) have so far provided mainly based on bore-hole measurements and active seismic prospecting (see e.g. Kramer 1996).

In the last decade, very promising results have been obtained from the analysis of environmental-noise recordings, both using single-stations to estimate *horizontal-to-vertical* (H/V) spectral ratio curves (e.g. Fäh *et al.* 2001; Arai & Tokimatsu 2004) and array configurations to estimate surface wave dispersion curves (e.g. Scherbaum *et al.* 2003). These techniques can be successfully applied where other approaches are not feasible (e.g. due to the high level of environmental seismic noise) and allow good lateral coverage at reasonable costs. This last feature becomes particularly important when limited budgets are available for seismic microzonation (e.g. in low seismicity areas and developing countries). Therefore, passive seismic techniques based on the measurement of environmental noise have been the subject of extensive research (see e.g. SESAME European project 2005).

The key problem of these techniques concerns the accuracy of the soil characterization they provide. In fact, both dispersion and H/V ratio curves must be inverted to retrieve the local *S*-wave profile, but this inversion is hampered by the extreme non-linearity of the data-model parameters relationship. Scherbaum *et al.* (2003) showed that the H/V ratio and the dispersion curves display different sensitivity to the *S*-wave velocity and thickness of the sedimentary layers. In particular, the dispersion curve represents the main constraint in the definition of the *S*-wave velocity of the soft sediments, while the fundamental frequency estimated from the ellipticity peak constrains the total thickness of the sediment cover. Thus, when the inversion is applied to these curves separately, there is an un-resolvable trade-off between the model parameters that hampers the analysis results. To overcome this drawback, Parolai *et al.* (2005) and Arai & Tokimatsu (2005) proposed a joint inversion of phase velocity and H/V ratio curves. They showed that with this approach, the trade-off between the model parameters can be reduced and a reliable evaluation of the local *S*-wave velocity structure can be obtained. Two different algorithms have been proposed for the joint inversions problem: genetic algorithms (hereafter GA) applied by Parolai *et al.* (2005) and linearized algorithms (hereafter LIN) adopted by Arai & Tokimatsu (2005).

GA (e.g. Goldberg 1989) require the computation of the forward problem and of the cost function only, avoiding any potential

numerical instabilities that may arise from, for example, matrix inversion and partial derivative estimates. For this reason, they are considered inherently stable methods. Moreover, direct-search methods guarantee the exploration of a wide portion of the solution space. Therefore, since different minima of the cost function are explored, the convergence towards an optimum solution is highly probable, even if this is computationally extensive. In addition, since GA allows the exploration of a large parameter space, both the best-fitting model data and general information about the parameter trade-off can be retrieved.

LIN strategies are based on the generalized least-squares method (e.g. Marquardt 1963; Wiggins 1972). These have the advantage of converging rapidly towards an acceptable minimum of the cost function, starting from a suitable initial guess model. Moreover, by the estimation of the resolution and model covariance matrix, one can estimate confidence limits for the unknown parameters.

Despite the fact that both GA and LIN approaches can be effective, in many cases they lead towards non-optimal parametrization. Indeed, in highly ill-posed inversion problems, GA could be stalled by a complicated fitness landscape and be unable to exactly single out the global optimum solution (Mosegaard & Sambridge 2002). On the other hand, when LIN methods are used, poor starting models are likely to result in low-quality or undesired parameter estimation (Menke 1989).

In order to overcome these difficulties, a possible strategy consists in combining both kinds of inversion methods to benefit from the advantages each (e.g. Hernandez *et al.* 1999; Alecu *et al.* 2004; Bourova *et al.* 2005; Santos *et al.* 2005). In the present work, an attempt is made to introduce a two-step inversion scheme by combining GA and LIN techniques: GA is used to provide the LIN with an initial model reasonably near to the optimal solution to increase the likelihood of convergence towards the best-fitting model. This approach is applied for the joint inversion of H/V ratios and Rayleigh wave dispersion curves deduced from environmental noise measurements carried out at a well-known test site located near the Casaglia Village in Northern Italy. Here, high quality geological and geophysical data are available (Cocco *et al.* 2001), which will allow a satisfactory check of the two-step inversion results.

TEST SITE AND DATA

The Casaglia test site is located in the Po Valley (Northern Italy) (Fig. 1). The geological setting of the whole area is well known, as several exploratory wells have been drilled by the Italian petroleum company (AGIP) because of the presence of a low-enthalpy geothermal field, which is exploited for industrial activities and house heating in the nearby city of Ferrara.

Geologically, the Po Valley is a large subsident basin filled by Quaternary alluvium. In the eastern part of the basin recent sediments can reach a thickness of several kilometres, and overlies a SW–NE aligned thrust-fault system (Cocco *et al.* 2001). The site under study is located on top of the thrust called the 'Ferrarese uprise' where the total thickness of the Quaternary deposits (constituting an alternation of sands and silts) is about 130 m. Within the sedimentary cover, the stratification can be considered as being planar with horizontal interfaces, at least for the wavelengths of interest in the present study (hundreds of metres).

Among the geophysical studies carried out at the Casaglia site, Malagnini *et al.* (1997) used cross-hole and up-hole measurements to constrain the body-wave velocity to 80 m depth. They also estimated the *S*-wave velocity profile to the base of the Quaternary alluvium from the inversion of phase and group velocity dispersion curves relative to Rayleigh waves generated by artificial sources. Moreover, a direct estimate of the local transfer function was obtained by Cocco *et al.* (2001) from the analysis of eight seismic events (with a magnitude range between 2.2 and 3.6) recorded using a borehole broad-band seismometer (Guralp CMG3T) at the base of the alluvial layers (135 m of depth) and a second broad-band seismometer at the surface.

All these pieces of information have been considered to evaluate the reliability of the inversion results presented in this work. However, no initial model has been deduced from these data to feed the

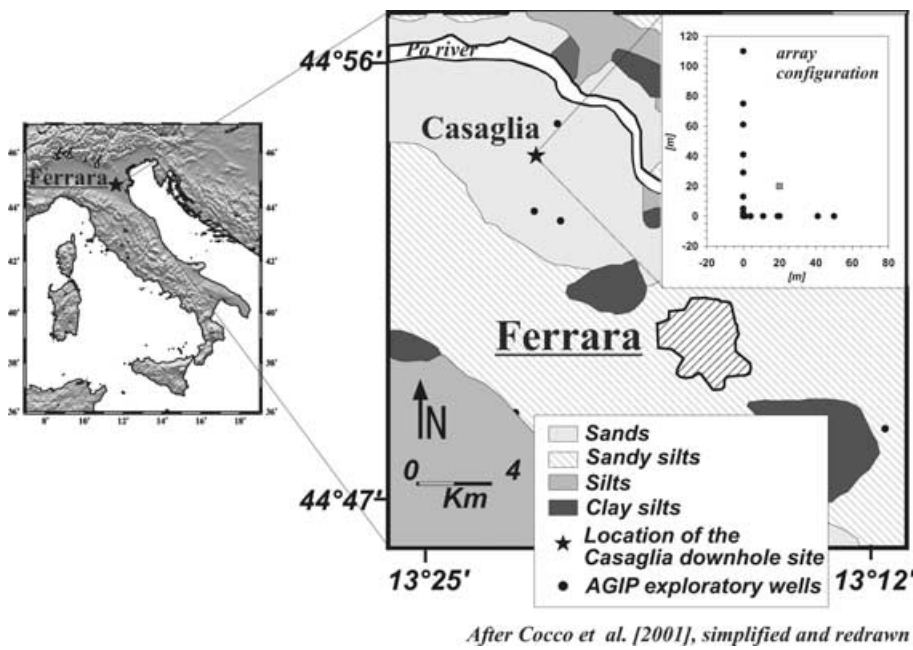

Figure 1. Geographical location and schematic geological map showing the Casaglia test site (modified from Cocco *et al.* 2001). The asterisk indicates the position of the borehole site and also corresponds to the location of the seismic array. Black dots indicate the position of exploratory wells made by the Italian petroleum company (AGIP). In the inset, the array configuration (dots) is reported along with the location of the 3-D component sensor (grey square).

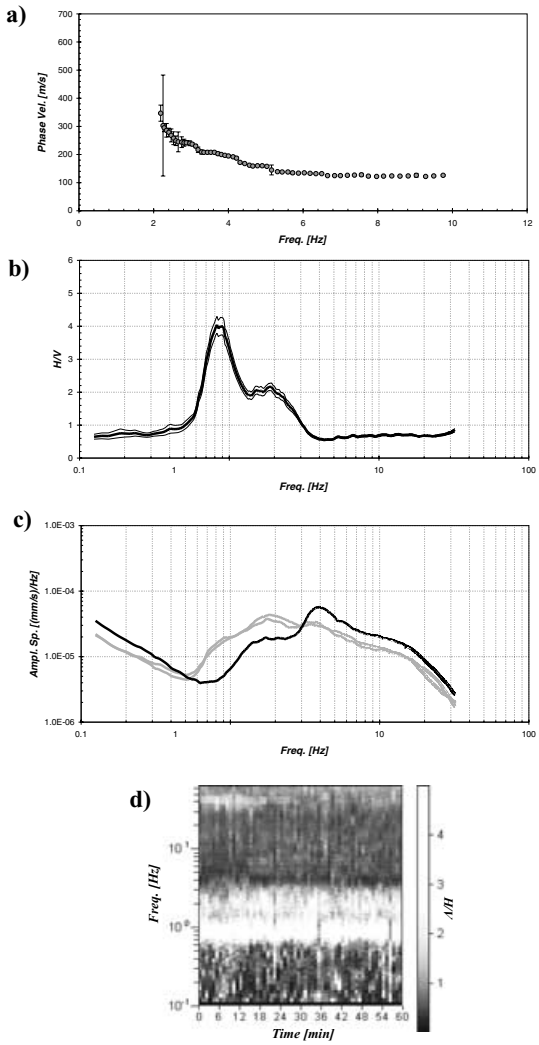

Figure 2. Experimental results from environmental noise measurements. (a) Rayleigh wave dispersion curves (grey dots). The error bars around the dispersion curve have been computed following (Menke 1989). (b) H/V ratios curve (thick line) and relative 95 per cent confidence interval (thin lines). (c) Horizontal (grey lines) and vertical (black line) velocity spectra of the recorded noise. (d) Contour plot of the H/V curves computed for each window. The scale-bar on the right indicates H/V amplitudes.

inversion procedure. This will allow us to check the robustness of the joint inversion strategy proposed here.

At the test site, 1 hr long recordings of ambient seismic noise have been performed with a sampling rate of 128 Hz by using both single station and array configurations. The small-scale array (Fig. 1) consisted of 16 vertical geophones (4.5 Hz) deployed in an L-shape configuration with a maximum dimension ca. 120 m. A digital acquisition system produced by Micromed (http://micromed-it.com/brainspy1.htm) has been used. Single station three-directional measurements have been performed by using a portable instrument specifically designed to measure seismic noise (www.tromino.it). The Rayleigh wave dispersion curve (Fig. 2a) was estimated using the *ESAC* procedure (Ohori *et al.* 2002). Uncertainty associated with the phase velocity values are estimated following Parolai *et al.* (2006). The fundamental frequency of the available geophones limits our capability to estimate reliable Rayleigh wave velocities for frequencies lower than about 2 Hz (i.e. for the deeper portion of the sedimentary cover), with realistic values of phase velocities obtained in the range 2–10 Hz. The maximum exploration depth can be estimated as half of the longest monitored wavelength, and is of the order of 80 m.

The H/V ratio curve (Fig. 2b) has been computed using 60 s time windows of noise recording signal, tapered with a 5 per cent cosine function before the computation of the spectrogram for each noise component. Spectrograms were smoothed using the Konno & Ohmachi (1998) window ($b = 40$). This ensures the reduction of numerical instabilities while preserving the major features of the spectra, especially the shape of the H/V ratios maxima. The resulting spectral ordinates relative to horizontal components have been geometrically averaged and divided by the vertical spectral ordinate to compute the H/V function. Results obtained by considering 60 time windows have then been averaged to compute the final H/V curve. Azimuthal and temporal stability and confidence intervals relative to H/V ordinates have been estimated by following D'Amico *et al.* (2004) and by using the code 'Grilla®' (www.tromino.it). The H/V curve exhibits two different resonant peaks. The fundamental resonance frequency has been estimated to be 0.86 ± 0.01 Hz, while the second peak corresponds to a frequency of 1.91 ± 0.04 Hz, with both estimated resonance frequencies in good agreement with those obtained by Cocco *et al.* (2001). Despite the second peak being partially merged with the first, both H/V maxima are related to a minimum in the amplitude of the vertical velocity spectra (Fig. 2c), as expected if the H/V peaks resulted from the polarization of the Rayleigh waves (Scherbaum *et al.* 2003). Moreover, they are stable in time (Fig. 2d) and can be considered reliable according to the SESAME criteria (SESAME European project 2005).

FORWARD MODELLING

Modelling Rayleigh wave phase velocities and H/V curves requires the choice of a suitable formalization of the background physical problem and of a reference structural model. Both of these aspects are discussed below.

Physical background

The formulations proposed by Tokimatsu *et al.* (1992) and Arai & Tokimatsu (2004) have been followed, based on the assumption that the noise wavefield is dominated by surface waves, at least with regards to its most coherent and statistically persistent component. Previous studies (e.g. Gucunski & Woods 1991; Tokimatsu *et al.* 1992) showed that when *S*-wave velocity varies irregularly with depth, several propagation modes can contribute to the surface wave field. The contribution of each mode m at each frequency f is modulated by a medium response factor $A_m(f)$, which is representative of the energy carried by each mode during the propagation and only depends on the mechanical layering of the subsoil. In the following, the presence of several propagation modes have been taken into account when modelling the dispersion and H/V curves.

The modelling of the dispersion curve is performed in terms of Rayleigh waves only, since only vertical motion measurements are used. The frequency dependence of H/V ratios is considered to depend upon the Rayleigh wave ellipticity, while their amplitude depends upon both Love and Rayleigh waves (Arai & Tokimatsu 2000, 2004). Further details about the formulation can be found in the references listed.

Structural model

In this study, the above formulations have been applied to compute the apparent phase velocities and H/V ratios for the case of vertically heterogeneous 1-D models using the modified Thomson–Haskell method proposed by Herrmann (2002). Here, the medium is replaced by a stack of horizontal uniform layers overlaying a homogeneous infinite half-space (hereafter *IHS*), the latter corresponding to rigid bedrock.

As clearly illustrated by Scherbaum *et al.* (2003), the filtering role of the medium on the frequency range around the fundamental resonance frequency, f_o, corresponding to the *IHS* resonance, makes dispersion curves nearly insensitive to choice of *IHS*. However, Wathelet (2005) pointed out that positioning the *IHS* at the bottom of the soft sedimentary cover could result in an artificial and physically unrealistic truncation of higher modes of both Love and Rayleigh waves in the low frequency band.

Thus, the introduction of a wrong *IHS* could potentially bias the modelling, and therefore, the inversion results. The magnitude of this bias is expected to be negligible in the presence of a sharp impedance contrast at the bottom of a soft layer overlying hard bedrock, since, in this case, the surface waves are naturally constrained in the soft sediments. However, when this impedance contrast is weaker, higher modes are not constrained to propagate in the shallow subsurface only. In this case, the lack of higher modes could determine a considerable bias in the velocity profile estimated from the inversion of H/V and dispersion data. The strong sensitivity of the H/V peak on *IHS* features has been clearly shown by Picozzi *et al.* (2005). Some numerical tests illustrating this aspect are also presented in Appendix A. These tests indicate that the impedance contrast (*IC*) magnitude at the bottom of the soft sedimentary layer is well constrained by the shape and amplitude of the whole H/V peak. On the other hand, the velocity and thickness of the layer located below this *IC* can be poorly resolved since these only affect a small portion of the H/V curve (i.e. just below f_o).

These considerations have induced Arai & Tokimatsu (2004, 2005) to include in their inversion analysis of H/V curves the deep ground structure inferred from available geological and geophysical information. Unfortunately, this information is not available for the test site in this study, where the available geophysical information is limited to the soft sedimentary layers. However, due to the presence of a geothermal field in the study area, one can guess that the underlying bedrock (i.e. Miocene flysch units, Cocco *et al.* 2001) is probably weathered. This suggests that the *IC* at the bottom of the alluvial sedimentary layer could be quite low (Malagnini *et al.* 1997). Therefore, higher modes are needed to correctly reproduce the observed H/V peak, and thus, a careful characterization of *IHS* is required.

IMPLEMENTATION OF THE TWO-STEP JOINT INVERSION PROCEDURE

The inversion algorithm GA proposed by Yamanaka & Ishida (1996) and the LIN proposed by Arai & Tokimatsu (2005) have been implemented in a MATLAB code. Since a complete description of the methods can be found in the reference papers, in the following they will be here only briefly outlined.

In view of the fact that the aim of this work is to combine GA and LIN in a two-step inversion scheme, a single cost function was adopted during the two inversion stages. In particular, in order to obtain a fair weighting of both the data set at hand (i.e. the dispersion and the H/V curves), both inversions have been performed using a cost function in the form

$$\text{cost} = \left[\sqrt{\sum_{j=1}^{N}\left(\frac{c_o(f)-c_(f)}{c_o(f)}\right)^2 \Big/ N}\right] + \left[\sqrt{\sum_{j=1}^{K}\left(\frac{hv_o(f)-hv(f)}{hv_o(f)}\right)^2 \Big/ K}\right], \quad (1)$$

where the symbol $c(f)$ indicates the Rayleigh wave phase velocity, $hv(f)$ indicates the H/V spectral ratios as a function of the frequency f, where the subscript o indicates the observed data with N and K being the number of data points in the dispersion and H/V ratio curves, respectively.

The first step: GA inversion scheme

Similarly to other direct-search (or non-linearized) inversion methods (e.g. Simulated Annealing Kirkpatrick *et al.* 1983, or neighbourhood algorithm (NA) Sambridge 1999), the GA scheme (Goldberg 1989) requires a series of appropriate tuning parameters. The parameter values used in the present study take advantage of previous experience using this algorithm by Yamanaka & Ishida (1996) and Parolai *et al.* (2005).

The initial step is to define the parameter space to be explored. Sensitivity analyses (e.g. Tokimatsu 1997; Xia *et al.* 1999; Arai & Tokimatsu 2004) show that the dependence of dispersion and of H/V curves on *P*-wave velocity, V_P, and mass density per unit volume is smaller than that on *S*-wave velocity, V_S, and the thickness, H, of the layers. However, as shown by Boore & Toksöz (1969), the *Poisson's ratio* significantly affects the Rayleigh dispersion curves in the low frequency range. Numerical test (see Appendix B) show that this also holds for H/V curves. This implies that *Poisson's ratio* can be potentially retrieved from the joint inversion of Rayleigh wave dispersion and H/V ratios curves. To reduce the number of free parameters, density set to vary between 1800 kg m^{-3} at the surface and 2300 kg m^{-3} at depth, in agreement with *a priori* geological information. Therefore, the inversion concerned V_S, *Poisson's ratio* and H values relative to each layer.

Parolai *et al.* (2006) showed that setting the number of layers to about eight produces reliable results for the kind of velocity structure off concern (i.e. sedimentary cover thickness of few hundred meters). Therefore, the inversions have been carried out considering models with eight layers for retrieving the *S*-wave velocity profile in the soft sedimentary cover. A couple of deep layers have been introduced below the bottom of this cover to avoid the possible biases in the modelling (see above). Both velocity and thickness of the first deep layer have been freely varied during the inversion as far as it has been necessary to extend the higher modes information until the lowest frequency of interest. The second deep layer is the actual *IHS*. Actually, the distinction between soft sediments and deep layers only relies on the larger variation intervals of the relevant parameters (thickness, *Vs*, *Poisson's ratios*) that during the inversion have been attributed to the deep layers. Since features of these layers are poorly resolved by the inversion of available data, they will not be considered in the interpretation of results.

During the inversion, an initial population of 50 individuals has been randomly generated from the search hyper-volume of V_S, H and *Poisson's ratios*. Then, a series of genetic operations (selection, uniform crossover, mutation, dynamic mutation and elite selection) have been applied in order to generate a new population with the

same size. This population has been reproduced based on the cost function values (1).

The crossover probability has been set to 0.8, while the mutation probability has been initially set to 0.02, however, this value is dynamically varied during the inversion. Iterations are terminated after the 100th generation, as we have found that no further significant reduction of the misfit is observed. Therefore, this inversion scheme involved the analysis of 5000 models. The procedure was repeated three times by varying the seed of the random number generator responsible for the generation of the trial populations. At the end of the inversion procedure, the optimal model (i.e. corresponding to the minimal cost function value) is selected. The efficiency of this strategy in generating models near to the global optimal solution has been demonstrated in previous numerical tests (Parolai *et al.* 2005).

The second step: LIN inversion scheme

Following Arai & Tokimatsu (2005), the second inversion step has been performed using a combination of the *SVD* method (Golub & Reinsch 1970; Press *et al.* 1992) with the modified Marquardt's technique (Marquardt 1963; Fletcher 1971). This LIN inversion strategy has been considered in order to retrieve the parameters V_S and H, starting from the initial model provided by the GA procedure. The cost function (1) is then minimized by an iterative procedure until a sufficiently small value is found. During the inversion, in agreement with the strategy proposed by Marquardt (1963), different damping parameters have been tested. In all cases, 50 iterations are used as the stopping criteria.

INVERSION RESULTS AND DISCUSSION

GA inversion

Following the procedure outlined in the previous section, the joint inversion has been performed with the curves evaluated from the Casaglia data set. The whole dispersion curve (2.2–9.75 Hz) is considered in the inversion. Concerning the H/V curve, data relative to the frequency range 0.5–2.9 Hz has been taken into account, since in this frequency band the role of Rayleigh wave ellipticity is assumed to be dominant (Scherbaum *et al.* 2003).

The wide range of variations for the mechanical parameters of the soft sedimentary layers can be properly appreciated by looking at the tested models presented in Fig. 3(a). In particular, the *S*-wave velocity in the layer that determines the magnitude of the *IC* at the bottom of the sedimentary layers has been varied in the range 300–2000 m s^{-1}, while the total thickness of the soft sedimentary cover is allowed to vary between 100 and 200 m. The minimum misfit model of Fig. 3(a) indicates that the local *S*-wave velocities are quite low (<500 m s^{-1}) until about 110 m depth, below which a sharp

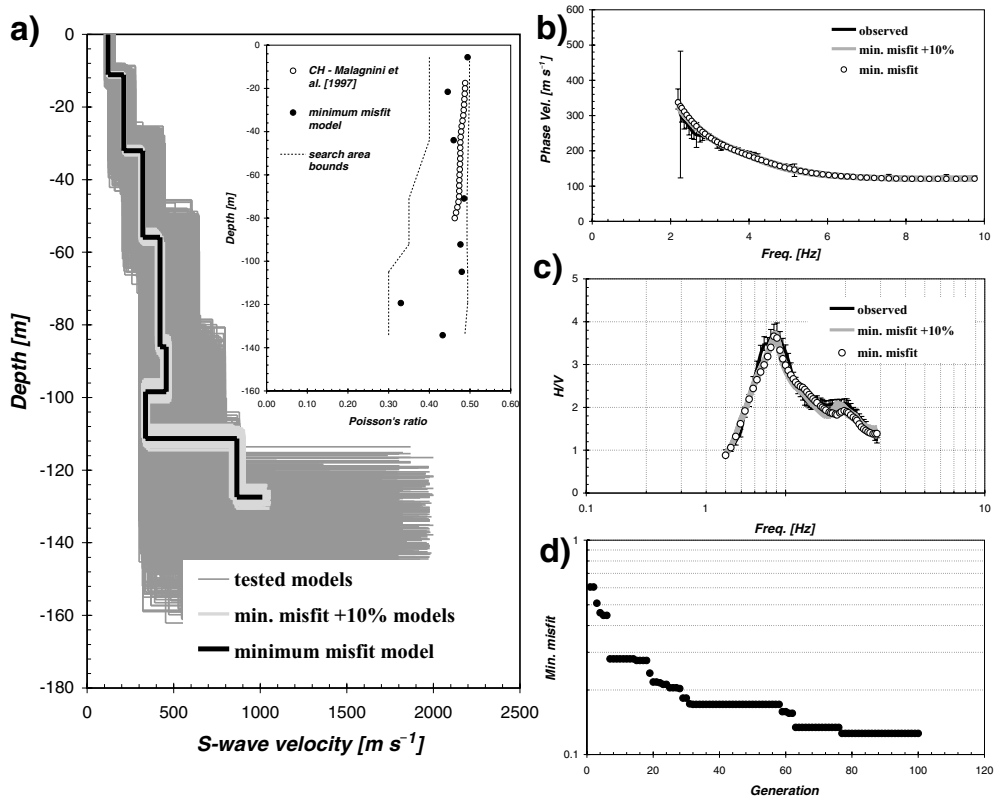

Figure 3. Results of the joint inversion carried out with GA. (a) Tested S-wave velocity models (dark grey lines), minimum misfit model (black line), models lying inside the minimum misfit +10 per cent range (light grey lines). Inner inset: Poisson's ratio from Malagnini *et al.* (1997) (white dots), for the minimum misfit model GA (black dots) along with lower and upper limits of the parameters search area (dashed lines). (b) Observed phase velocities (black line) and relative error bars (black), phase velocities for the minimum cost model (dots) and phase velocities for the models lying inside the minimum misfit +10 per cent range (grey lines). (c) Observed H/V ratio (black line) and relative error bars (black), theoretical H/V ratios for the minimum cost model (white dots) and H/V ratios for the models lying inside the minimum misfit +10 per cent range (grey lines). (d) Minimum misfit value (cost function in eq. 1) as a function of model generation.

increase of velocity is observed. The comparison between the best-fit and the minimum misfit +10 per cent models allows us to judge the confidence level of the results. Until about 110 m, the *S*-wave velocity structure is very well constrained and all models selected are close to each other. On the contrary, for the deeper part of the model, the uncertainty level associated with V_S increases, mainly because the results of inversion are constrained by the H/V values only. However, in spite of this uncertainty, the H/V curve provides a reliable estimate about the *IC* magnitude at the interface between the soft sedimentary layer and the deep ground structure. Finally, Fig. (3a, inner inset) shows that the *Poisson's ratios* of the minimum misfit model are also in good agreement with those estimated by Malagnini *et al.* (1997) from cross-hole measurements.

Figs 3(b) and (c) compare the experimental and theoretical curves relative to the best-fitting model and the models inside the minimum cost +10 per cent range. It can be observed that both the experimental dispersion and H/V curves are fairly well retrieved, even if locally they are slightly over/under estimated (e.g. the H/V at about 2 Hz and the dispersion curve under about 3 Hz.) This suggests that the GA provides a best-fitting model that is near the global minimum solution. This is also indicated by the decreasing trend of the minimum misfit relative to each population as a function of generation number (Fig. 3d). It is evident that the misfit value decreases rapidly in the first part of the process, and then quite slowly in the following iteration. The slow improvement in misfit with increasing number of generations suggests the convergence towards the global minimum (Yamanaka & Ishida 1996). Therefore, in spite of GA not being able to single out the model that best reproduces the data set at hand, still it allows a wide portion of the parameter space to be explored and the identification of the hypervolume where the global minimum lies.

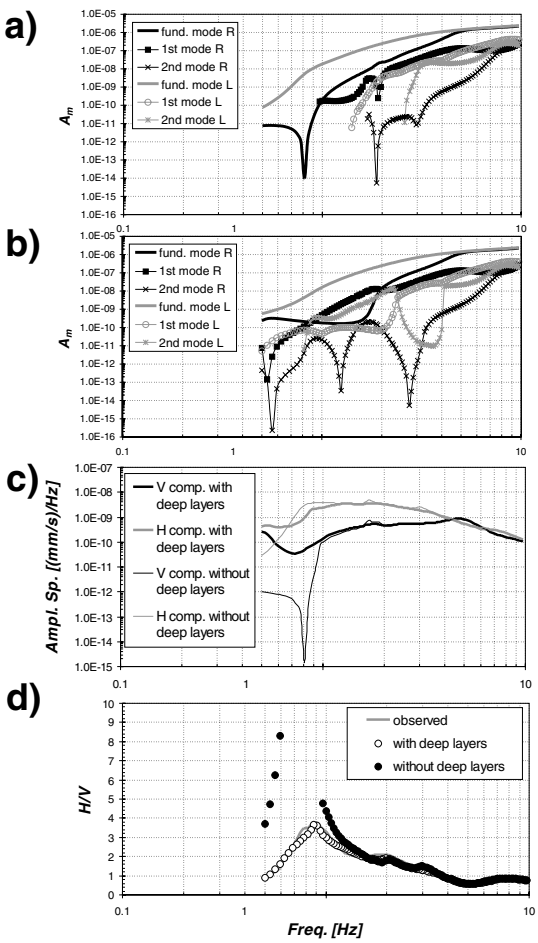

Figure 4. Main features of the best-fitting model provided by GA. (a), (b) Medium response function $A_m(f)$ (c) Amplitude spectra of the vertical, V, and horizontal, H, components. (d) H/V curves.

Fig. 4 depicts the theoretical medium response functions, $A_m(f)$, the theoretical spectra and the H/V curves computed for the GA best-fitting model without and including the two deep layers, respectively. For all the models inside the minimum misfit +10 per cent range, the first of these deep layers has a thickness varying in the range of 300–600 m, while the *S*-wave velocity is close to the velocity of the layer that determines the main *IC* at the bottom of the sedimentary cover. From Fig. 4(b), it is clear that, when the deep layers are present, higher modes extend their influence up to 0.5 Hz. Moreover, without the presence of higher modes the relevant vertical component spectrum exhibits a pronounced hole around the fundamental resonant frequency (Fig. 4c), which results in unrealistic H/V peaks (Fig. 4d).

LIN inversion

The best-fitting model obtained by GA has been used as the initial guess for the inversion performed using LIN methods. However, since the original LIN scheme (Arai & Tokimatsu 2005) is devoted to the identification of V_S and H only, the relevant *Poisson's ratio* has been fixed *a priori* on the basis of the GA results.

Fig. 5(d) shows that after only six iteration, the cost function is reduced by about the 25 per cent. Then, since no further improvement of the cost function has been observed in the next iterations, the inversion is terminated. The *S*-wave velocities of the best-fitting LIN model are similar to the GA model in the upper 100 m (Fig. 5a), While below this depth, an increase of the *S*-wave velocities is obtained, along with the deepening of the main *IC*. However, the LIN model allows theoretical curves to be obtained that better reproduce the input data than those from *GA* (Figs 5b and c). The lower value of the cost function reached in the LIN inversion with respect to that provided by GA (Fig. 5d) indicates that, a better global minimum is found for the data set available. The standard error ratios of the LIN model (Wiggins 1972; Arai & Tokimatsu 2004) are close to 0.2, indicating that the estimated result is reliable.

Nevertheless, it is interesting to verify the capability of the LIN inversion to detect the global minimum misfit alone. In order to perform this test, three models, among those tested by GA (with a misfit equal to the minimum misfit plus 10, 25 and 50 per cent, respectively), have been selected and used as initial guess for the LIN inversion. Fig. 6 shows the results of this analysis. When the input model is close to the optimal model (i.e. minimum misfit +10 per cent), LIN is still able to identify the global minimum, with only a little discrepancy concerning the soft sediments total thickness (Fig. 6a). However, when the input model is more distant from the optimal solution (i.e. minimum misfit +25 per cent and, especially, +50 per cent), LIN is still able to reduce the misfit, but it is not able to identify the global minimum of the cost function.

Comparison with other geophysical data

Fig. 7(a) shows the very good agreement between results obtained from the joint inversion of dispersion and H/V curves from noise

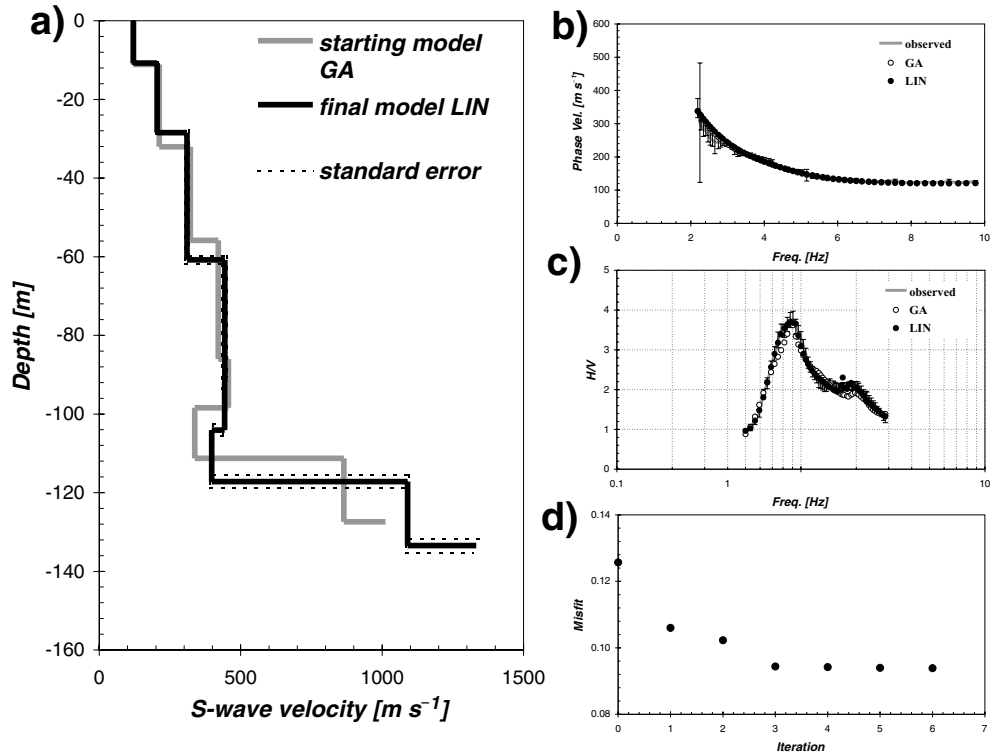

Figure 5. Results of the joint inversion carried out with the LIN algorithm. (a) *S*-wave velocity profiles. (b) Rayleigh wave dispersion curves. (c) H/V curves. (d) Minimum misfit value (cost function eq. 1) as a function of iteration (see caption of Fig. 3).

data and those provided by Malagnini *et al.* (1997). However, for the deeper portion of the soft sedimentary cover, our models tend towards higher velocity values due to the information provided by the two H/V peaks concerning the main *ICs*. Although the model space explored is very large (Fig. 3a), the thickness of the soft sedimentary cover provided by the GA and LIN joint inversions (Fig. 7a) is in good agreement with that provided by drilling (130 m) (Cocco *et al.* 2001).

Finally, as suggested by Boore & Brown (1998) and Ohrnberger *et al.* (2004), the comparison of our results with those obtained by other authors has also been performed in terms of local transfer functions. Theoretical transfer functions have been evaluated using

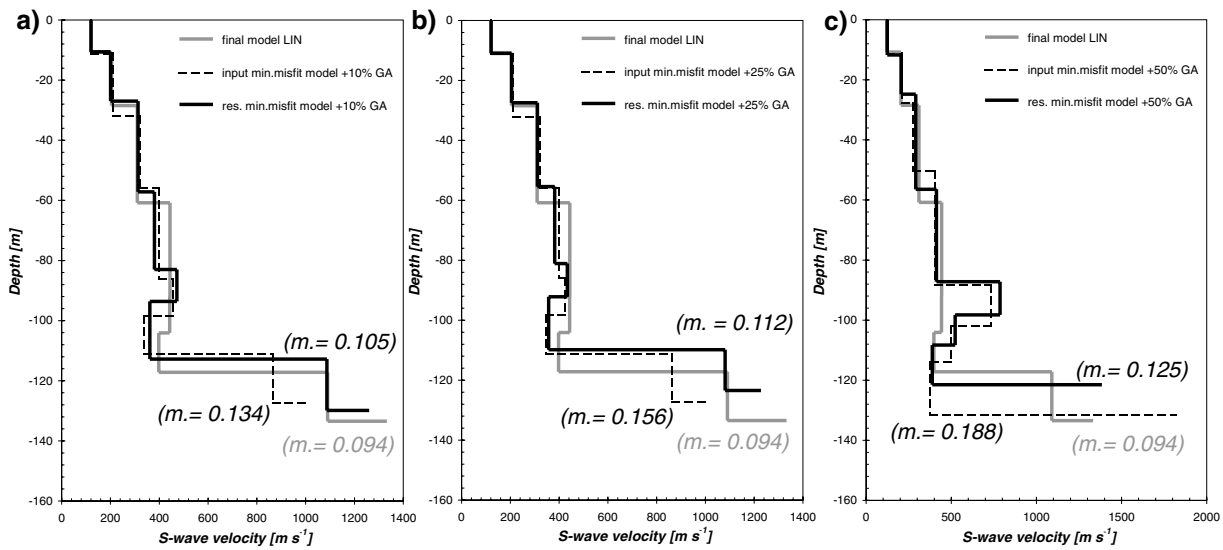

Figure 6. Results of the joint inversion carried out with the LIN algorithm having as starting models GA models located at different distance from the global optimum solution: (a) input minimum misfit +10 per cent range, (b) +25 per cent range, and (c) +50 per cent range. The misfit of each model is also indicated.

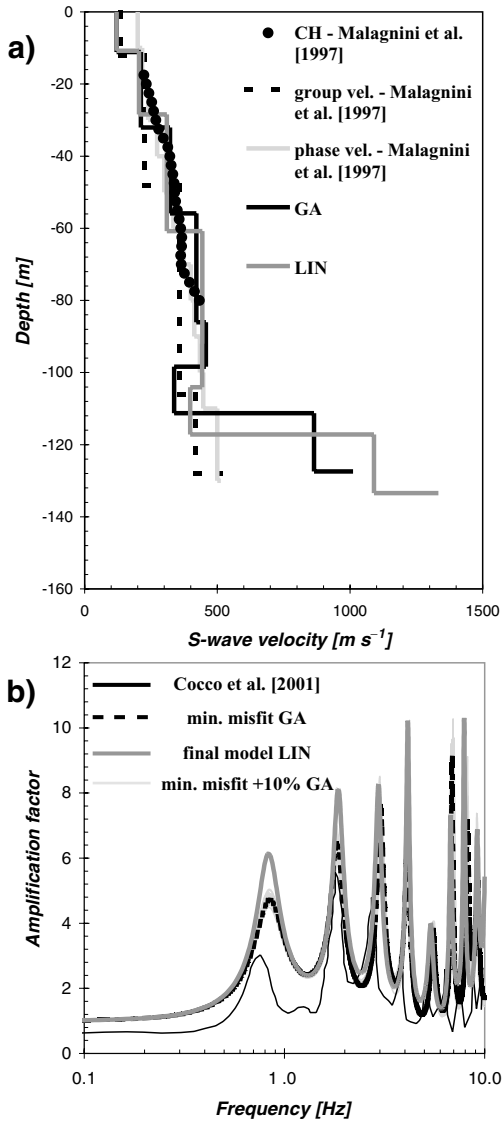

Figure 7. Comparison between joint inversion results of this study and available geophysical and geological information (from Cocco *et al.* 2001). (a) *S*-wave velocity models. (b) Experimental and theoretical local transfer functions.

the EERA code (Bardet *et al.* 2000) for the final LIN model, the best-fitting GA model and those GA models inside the minimum misfit +10 per cent range. These computations have been performed adopting the values of the quality factor Q proposed for the study area by Malagnini *et al.* (1997). These theoretical local transfer functions have been compared with those directly measured from recordings of local earthquakes by Cocco *et al.* (2001). Fig. 7(b) shows that the models obtained by the joint inversion analysis of seismic noise can reproduce satisfactorily the general trend of the experimental local transfer function in the frequency range of engineering interest (0.1–10 Hz), especially with regards to the location of both fundamental and higher resonance frequencies.

CONCLUSION

A two-step inversion strategy has been proposed for estimating *S*-wave velocity profiles. In particular, GA and LIN algorithms have been applied for the joint inversion of dispersion and H/V curves deduced from seismic noise measurements. The reliability of the inversion results has been assessed by considering the well-constrained geophysical and geological information available at a test site located in Northern Italy.

In the first step of the analysis, the use of GA allows the performing of a non-linear inversion analysis that does not depend upon an explicit starting model. This is the most straightforward property of GA, considering that site-effect investigations are often required in regions where there is little or no knowledge about the subsurface available. The best-fitting model of GA is then used as the starting model for the LIN inversion that has been able to drive the inversion to the global optimal minimum of the cost function. This results in a model that satisfactorily reproduces, within relevant errors, all the experimental data. It has been shown that, on the other hand, if the starting model is too far from the global minimum area, LIN can be trapped in some local minima and an incorrect *S*-wave velocity profile is retrieved.

Using the best-fitting GA and LIN models, it has been shown that the relevant local transfer function can be reasonably well determined (i.e. the location of fundamental and higher frequencies is fairly well retrieved). Moreover, even if the minimum misfit GA model does not coincide with the optimal absolute minimum of the cost function, the relevant transfer function estimate that result is unbiased. Thus, if the site response function is only of concern, GA represents a self-consistent tool for seismic microzonation studies.

Our results also indicate that the joint inversion of Rayleigh wave dispersion and H/V spectral ratio curves could also provide a first-order estimate of the *Poisson's ratio*. Furthermore, it has also been shown that, when a relatively low impedance contrast exists at the bottom of the soft sedimentary layer, a proper forward modelling of H/V curves requires at least some information about the deep ground structure.

It is worth noting that the experimental H/V curve, despite the fact that it is an essential piece of information for retrieving the structural characteristics of the local subsoil, does not correspond to the local transfer function deduced from experimental data and theoretical computations. This implies that results from the use of H/V to evaluate the amplification of earthquake ground shaking (especially for the higher modes) must be considered with caution.

ACKNOWLEDGMENTS

We are grateful to S. Richwalski and an anonymous reviewer for useful suggestions and comments that improved the manuscript. K. Fleming kindly improved our English.

REFERENCES

Alecu, T.I., Voloshynovskiy, S. & Pun, T., 2004. Regularized two-step brain activity reconstruction from spatiotemporal EEG data, *Image Reconstruction from Incomplete Data III, SPIE International Symposium on Optical Science and Technology,* Denver, Colorado, USA.

Arai, H. & Tokimatsu, K., 2000. Effects of Rayleigh and Love waves on microtremor H/V spectra, *Proc. 12th World Conf. On Earthquake Engineering*, 2232/4/A.

Arai, H. & Tokimatsu, K., 2004. S-wave velocity profiling by inversion of microtremor H/V spectrum, *Bull. seism. Soc. Am.*, **94**(1), 53–63.

Arai, H. & Tokimatsu, K., 2005. *S*-wave Velocity profiling by joint inversion of microtremor dispersion curve and horizontal-to-vertical (H/V) spectrum, *Bull. seism. Soc. Am.*, **95**(5), 1766–1778.

Bardet, J.P., Ichii, K. & Lin, C.H., 2000. EERA, A Computer Program for Equivalent-linear Earthquake site Response Analyses of Layered Soil Deposits. Department of Civil Engineering, University of Southern California.

Boore, D.M. & Toksöz, M.N., 1969. Rayleigh wave particle motion and crustal structure, *Bull. seism. Soc. Am.*, **59,** 331, 346.

Boore, D.M. & Brown, L.T., 1998. Comparing shear-wave velocity profiles from inversion of surface-wave phase velocities with downhole measurements: systematic differences between CXW method and downhole measurements at six USC strong-motion sites, *Seism. Res. Lett.*, **69,** 222–229.

Bourova, E., Kassaras, I., Pedersen, H.A., Yanovskaya, T., Hatzfeld, D. & Kiratzi, A., 2005. Constraints on absolute *S* velocities beneath the Aegean Sea from surface wave analysis, *Geophys. J. Int.*, **160,** 1006–1019.

Cocco, M. *et al.,* 2001. Broadband waveforms and site effects at a borehole seismometer in the Po alluvial basin (Italy), *Annali di Geofisica*, **44,** 137–154.

D'Amico, V., Picozzi, M., Albarello, D., Naso, G. e Tropenscovino, S., 2004. Quick estimate of soft sediments thickness from ambient noise horizontal to vertical spectral ratios: a case study in southern Italy, *J. Earthquack Eng.*, **8**(6), 895–908.

Fäh, D., Kind, F. & Giardini, D., 2001. A theoretical investigation of average H/V ratios, *Geophys. J. Int.*, **145,** 535–549.

Fletcher, R., 1971. A modified Marquardt subroutine for nonlinear least squares, Harwell Report, AERE-R 6799.

Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning,* Addison-Wesley, Reading, Mass.

Golub, G.H. & Reinsch, C., 1970. Singular Value Decomposition and Least Square Solutions, in *Linear Algebra, volume II of Handbook for Automatic Computations,* chapter I/10, pp. 134–151, eds Wilkinson, J.H. & Reinsch, C., Springer Verlag, Berlin.

Gucunski, N. & Woods, R.D., 1991. Use of Rayleigh modes in interpretation of SASW test, *Proc. 2th int. Conf. Recent Advances in Geot. Earthq. Eng. Soil Dyn.-S.Louis*, 1399–1408.

Hernandez, B., Cotton, F. & Campillo, M., 1999. Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake, *J. Geophys. Res. Solid Earth*, **104**(B6), 13 083–13 099.

Herrmann, R.B., 2002. *Computer Programs in Seismology,* Version 3.2, Saint Louis University.

Kirkpatrick, S.C., Gelatt, D. & Vecchi, M.P., 1983. Optimization by simulated annealing, *Science*, **220,** 671–680.

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

Kramer, S.L., 1996. *Geotechnical Earthquake Engineering,* Prentice Hall, Upper Saddle River, New Jersey 07458.

Malagnini, L., Hermann, R.B., Mercuri, A., Opice, S., Biella, G. & De Franco, R., 1997. Shear-wave velocity structure of sediments from the inversion of explosion-induced Rayleigh waves: comparison with cross-hole measurements, *Bull. seism. Soc. Am.*, **87,** 1413–1421.

Marquardt, D.W., 1963. An algorithm for least squares estimation on non-linear parameters, *J. Soc. Ind. Appl. Math.,* **11,** 431–441.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory,* rev. ed. Academic, San Diego, CA.

Mosegaard, K. & Sambridge, M., 2002. Monte Carlo analysis of inverse problems, *Inverse Problems*, **18,** R29–R54.

Ohori, M., Nobata, A. & Wakamatsu, K., 2002. A comparison of ESAC and FK methods of estimating phase velocity using arbitrarily shaped microtremor analysis, *Bull. seism. Soc. Am.*, **92,** 2323–2332.

Ohrnberger, M., Scherbaum, F., Krüger, F., Pelzing, R. & Reamer, S.-K., 2004. How good are shear-wave velocity models obtained from inversion of ambient vibrations in the lower Rhine embayment (N.W. Germany)?, *Boll. Geofis. Teor. Appl.*, **45,** 215–232.

Parolai, S., Picozzi, M., Richwalski, S.M. & Milkereit, C., 2005. Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes, *Geoph. Res. Lett.*, **32**. doi:10.1029/2004GL021115.

Parolai, S., Richwalski, S.M., Milkereit, C. & Fäh, D., 2006. *S*-wave Velocity Profile for Earthquake Engineering Purposes for the Cologne Area (Germany), *Bull Earthq. Eng.*, 65–94, doi:10.1007/s10518-005-5758-2.

Picozzi, M., Parolai, S. & Richwalski, S.M., 2005. Joint inversion of H/V ratios and dispersion curves from seismic noise: Estimating the *S*-wave velocity of bedrock, *Geophys. Res. Lett.*, **32,** No. 11 doi:10.1029/2005GL022878.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in Fortran,* 2nd edn., Cambridge University Press, New York.

Sambridge, M., 1999. Geophysical inversion with a Neighbourhood algorithm, I, Searching a parameter space, *Geophys. J. Int.*, **138,** 479–494.

Santos, A.B., Sampaio, E.E.S. & Porsani, M.J., 2005. A Robust Two-Step Inversion of Complex Magnetotelluric Apparent Resistivity Data, *Studia Geophysica et Geodetica*, **49**(1), 109–125.

Scherbaum, F., Hinzen, K.-G. & Ohrnberger, M., 2003. Determination of shallow shear-wave velocity profiles in Cologne, Germany area using ambient vibrations, *Geophys. J. Int.*, **152,** 597–612.

SESAME European project, 2005. Guidelines for the implementation of the H/V spectral ratio technique on ambient vibrations measurements, processing and interpretation. Deliverable D23.12. http://sesame-fp5.obs.ujf-grenoble.fr/SES_TechnicalDoc.htm

Tokimatsu, K., 1997. Geotechnical site characterization using surface waves, in *Earthquake Geotechnical Engineering,* pp. 1333–1368, ed. Ishihara, Balkema, Rotterdam.

Tokimatsu, K., Tamura, S. & Kojima, H., 1992. Effects of multiple modes on Rayleigh wave dispersion characteristics, *J. Geotech. Eng.*, **118,** 1529–1543.

Wathelet, M., 2005. Array Recordings of Ambient Vibrations: Surface-Waves Inversion. *PhD thesis*, Faculté des Sciences Appliquées, Université de Liège, France.

Wiggins, R.A., 1972. The general inverse problem: Implication of surface waves and free oscillations for Earth Structure, *Rev. Geophys.*, **10,** 251–285.

Xia, J., Miller, R.D. & Park, C., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics*, **64**(3), 691–700.

Yamanaka, H. & Ishida, H., 1996. Application of GENERIC algorithms to an inversion of surface-wave dispersion data, *Bull. seism. Soc. Am.*, **86,** 436–444.

APPENDIX A: DEEP GROUND STRUCTURE AND *H/V* SPECTRAL RATIO MODELLING

To demonstrate the importance of including deeper structure in H/V ratio modelling during the inversion analysis, a model (hereafter *original model*) that consists of six layers has been used. Layer thicknesses are: 20, 20, 30, 200, 300 and 400 m, respectively. The corresponding *S*-wave velocities are: 200, 300, 400, 800, 900 and 1200 m s^{-1}, respectively. The *Poisson's ratio* and density were fixed as 0.48, 0.45, 0.4, 0.35, 0.3 and 0.3 and 1700, 1700, 1800, 2200, 2300 and 2300 kg m^{-3}, respectively. The velocity of the infinite half-space (hereafter *IHS*) was fixed to 1300 m s^{-1}, the *Poisson's ratio* to 0.28, and the density 2400 kg m^{-3}. In order to obtain a data set consistent with the one from the test site at the Casaglia Village, the Rayleigh wave phase velocities and the H/V ratio were computed for the frequency band 2–10 and 0.2–10 Hz, respectively. This information was used as input in the inversion process. While the use of the same forward modelling algorithm for generating the data and performing the inversion does not allow an independent evaluation of the final model quality, in any case, the aim of this test is simply to highlight the drawbacks related to the use of an improper forward modelling of the H/V curve. The joint inversion analysis of dispersion and H/V curves was carried out by GA. However, in order to reduce the number of free

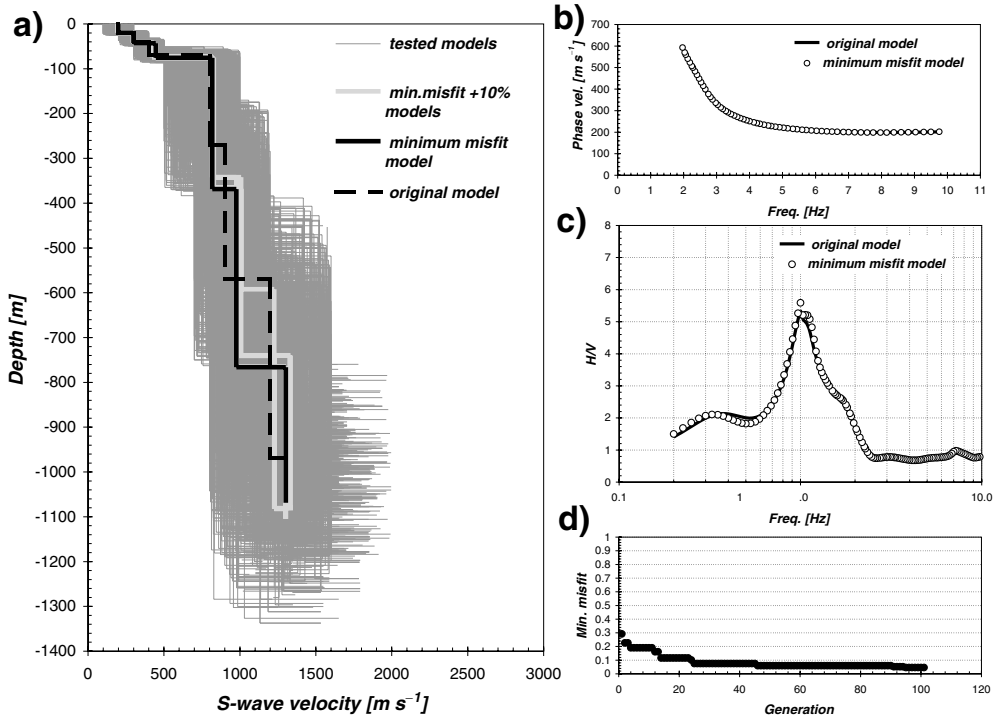

Figure A1. Joint inversion with GA. (a) *S*-wave velocity models. (b) Rayleigh wave dispersion curves. (c) H/V spectral ratio curves. (d) Minimum misfit value (cost function eq. 1) as a function of generation.

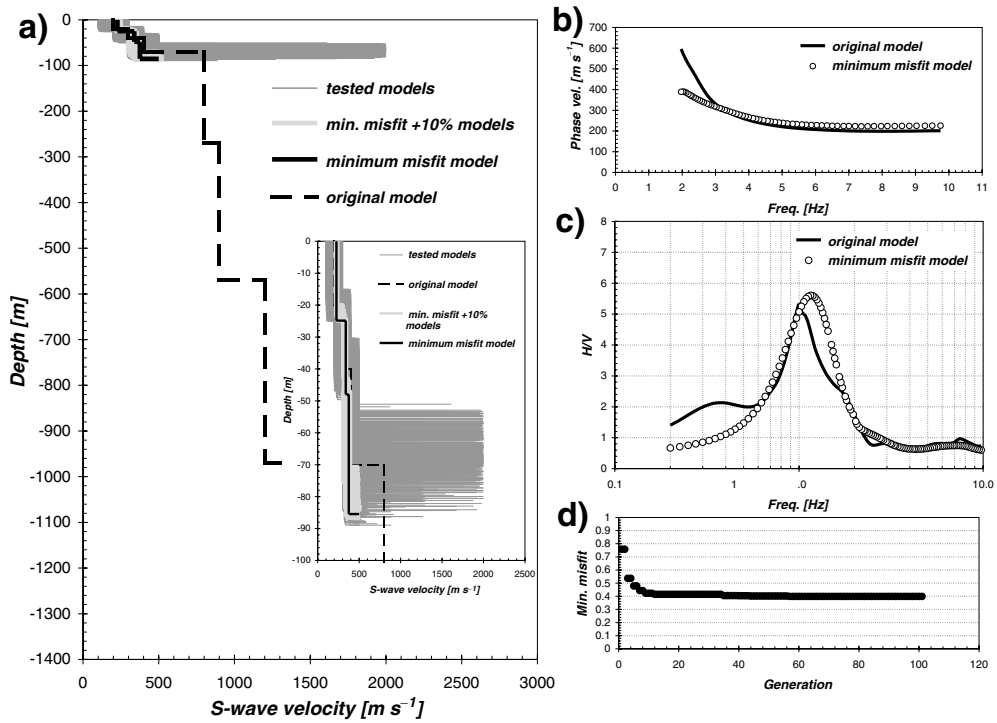

Figure A2. Joint inversion with GA. (a) *S*-wave velocity models. Inner inset: view of results at enlarged scale (b) Rayleigh wave dispersion curves. (c) H/V spectral ratio curves. (d) Minimum misfit value (cost function eq. 1) as a function of generation.

parameters, only *S*-wave velocities and thicknesses were considered as free parameters.

Fig. A1 shows the results of the joint inversion performed using the same number of layers as the *original model*. This choice allows the inclusion of higher modes until the lowest frequency of interest. Therefore, the velocity structure and, especially, the main *IC* at the base of the soft sedimentary cover (responsible for the H/V peak) are fairly well retrieved (Fig. A1a). This is

shown also by the satisfactorily fit of the dispersion and the H/V curves (Figs A1b and c), and by the minimum misfit reached at the last generation. Moreover, it is worth noting that even if the deep layers are not well resolved, they provide an estimate of the *S*-wave velocity trend in the deeper portion of the minimum misfit model that is in good agreement with the original model.

Fig. A2 shows the results of the joint inversion performed using tested models with only four layers (i.e. three sedimentary layers overlying an infinite half-space). Since in this case the higher modes cannot exist under the fundamental frequency, it is not possible for the GA (as well as for any other inversion algorithm) to retrieve a reliable final model (Fig. A2a). In fact, even if the *S*-wave velocity of the sedimentary layers is well estimated (Fig. A2a, inner inset), the total thickness of the model and the *S*-wave velocity of the *IHS* are poorly retrieved. The absence of higher modes determines a peculiar contrast in the reproduction of the dispersion and H/V data sets (Figs A2b and c). In fact, at low frequencies, the dispersion curve could be fitted only by models with higher velocities in the *IHS*. On the contrary, those *IHS* higher velocities are not tolerated by the relative theoretical H/V curves, which around and below the fundamental frequency are computed using the fundamental mode only. Finally, the poor quality of the final result is indicated also by trend of the minimum misfit values vs. the generation number, which shows the tendency to assume constant values during the inversion process (Fig. A2d).

APPENDIX B: ESTIMATION OF POISSON'S RATIO BY JOINT INVERSION OF DISPERSION AND H/V SPECTRAL RATIO CURVE

The joint inversion of dispersion and H/V curves can provide an estimate of the *Poisson's ratios* that is fairly well retrieved. To show this, similarly to the previous section, we used a reference model with the characteristics of the sedimentary cover (*S*-wave velocity, thickness and density) fixed, and only *Poisson's ratios* being allowed to vary. The model (hereafter *original model*) consists of four layers with densities of 1700, 1700, 1800 and 2200 kg m^{-3}, thicknesses of 20, 20, 30 and 800 m, and *S*-wave velocities of 110, 250, 500 and 800 m s^{-1}. In the first test, the *Poisson's ratio* was fixed to 0.48 for all layers, while in the second it was fixed to 0.35. In both cases, the Rayleigh wave phase velocities and the H/V ratio were computed for the frequency band 2–10 and 0.2–10 Hz, respectively, and used as input for the inversion stage. As expected, the computed curves (Figs B1c, d and B2c, d) clearly indicate that both the dispersion curve and the H/V peak are influenced by *Poisson's ratio* variations. Then, the joint inversion analyses were carried out by GA on both data sets, allowing the model parameters, *S*-wave velocity, thickness and *Poisson's ratio* of each layer, to vary over a wide range.

Figs B1 and B2 show that in both cases all characteristics of the model are fairly well retrieved and a good fit with the dispersion and H/V curves is obtained.

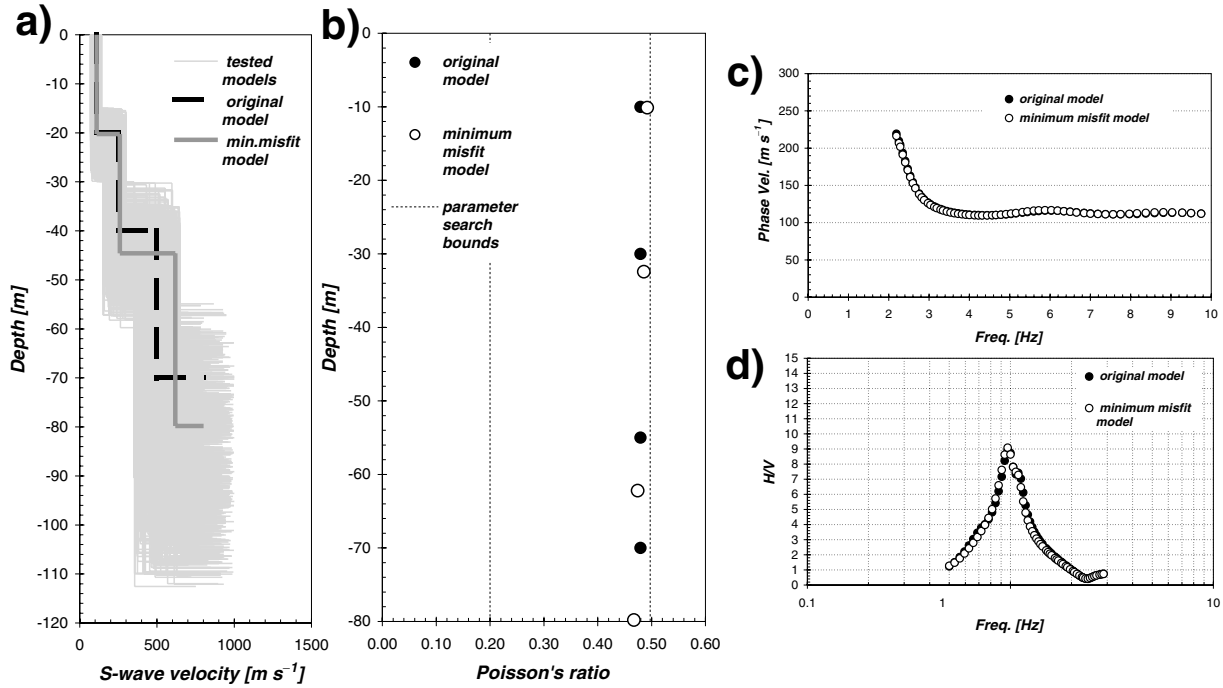

Figure B1. Joint inversion with GA, with Poisson's ratio fixed to 0.48. (a) *S*-wave velocity models. (b) Poisson's ratio of each layer. (c) Rayleigh wave dispersion curves. (d) H/V spectral ratio curves.

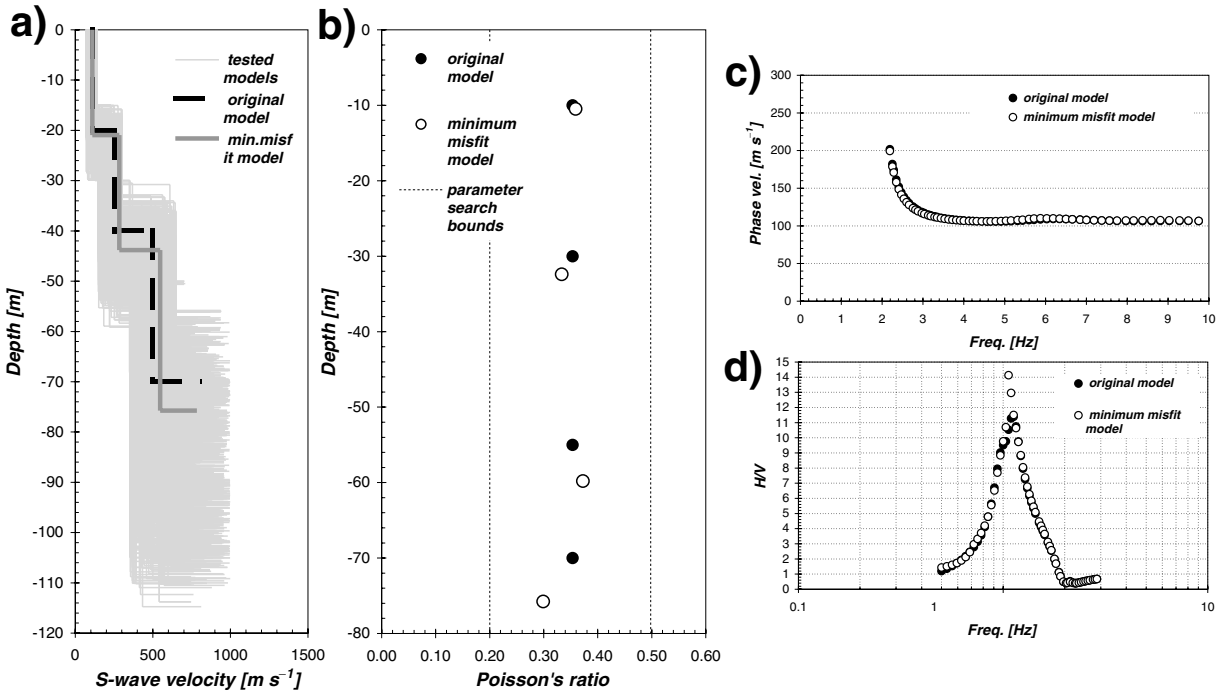

Figure B2. Joint inversion with GA, with Poisson's ratio fixed to 0.35. (a) *S*-wave velocity models. (b) Poisson's ratio of each layer. (c) Rayleigh wave dispersion curves. (d) H/V spectral ratio curves.