

# Identification of lateral discontinuities via multi-offset phase analysis of surface wave data

Giulio Vignoli<sup>1,2\*</sup> and Giorgio Cassiani<sup>1</sup>

<sup>1</sup>Università di Padova, Dipartimento di Geoscienze, 35122 Padova, Italy, and <sup>2</sup>Università di Ferrara, Math4Tech Center, 44100 Ferrara, Italy

Received February 2009, revision accepted September 2009

## ABSTRACT

Surface wave methods are based on the inversion of observed Rayleigh wave phase-velocity dispersion curves. The goal is to estimate mainly the shear-wave velocity profile of the investigated site. The model used for the interpretation is 1D, hence results obtained wherever lateral variations are present cannot be considered reliable.

In this paper, we study four synthetic models, all with a lateral heterogeneity. When we process the entire corresponding seismograms with traditional  $f$ - $k$  approach, the resulting 1D profiles are representative of the subsurface properties averaged over the whole length of the receivers lines. These results show that classical analysis disregards evidences of sharp lateral velocity changes even when they show up in the raw seismograms.

In our research, we implement and test over the same synthetic models, a novel robust automated method to check the appropriateness of 1D model assumption and locate the discontinuities. This new approach is a development of the recent multi-offset phase analysis with the following further advantages: it does not need previous noise evaluation and more than one shot.

Only once the discontinuities are clearly identified, we confidently perform classical  $f$ - $k$  dispersion curve extraction and inversion separately on both sides of the discontinuity. Thus the final results, obtained by putting side by side the 1D profiles, are correct 2D reconstructions of the discontinuous S-wave distributions obtained without any additional ad-hoc hypotheses.

## INTRODUCTION

In surface wave inversion, observed dispersion (i.e., variation of phase velocity with frequency) is used to infer the soil material properties. The experimental information is extracted from seismograms as multi-modal curves representing the dependence of phase-velocity on frequency (e.g., Socco and Strobbia 2004). Different approaches to describe and handle surface wave propagation in non-1D media have been described and investigated in the literature (Keilis-Borok *et al.* 1989; Abraham *et al.* 2004; Socco *et al.* 2006) but the existing

practice is strongly limited to the inversion of data in terms of 1D shear velocity profiles.

The use of Rayleigh waves to determine the dynamic properties of soil deposits and structures started more than sixty years ago (Bergström and Linderholm 1946). Early studies led to the so-called steady-state method based on constant frequency excitation (Van der Pol 1951; Jones 1955). This approach evolved later to become what is commonly known as the continuous surface wave method (Matthews, Hope and Clayton 1996). In the early 1980s, a new method to estimate the S-wave velocity profile, called spectral analysis of surface waves, was introduced (Heisey, Stokoe and Meyer 1982; Nazarian, Stokoe and Hudson 1983). Spectral analysis of surface waves consists in recording the surface waves generated

---

\*E-mail: giulio.vignoli@unipd.it

by an impact source by means of only two sensors. The measure is repeated for different spacing between the geophones and, because of that, it usually takes several hours to complete the whole data acquisition. The dispersion curve is first determined by calculating a cross power spectrum that defines the phase difference, as a function of frequency, between the waves measured by the two sensors. Since the distance between the sensors is known, the dispersion curve is determined for each spacing. The curve of total dispersion, corresponding to the average of the dispersion curves determined for the various spacing, is then inverted in order to determine the velocity profile of the shear waves as a function of depth. The generated 1D S-wave velocity profiles can be aligned one next to the other, each at a location corresponding to the centre of its geophone pair, to build a pseudo-2D section. However, the traditional spectral analysis of surface waves (because its use of only two receivers) is limited by: poor resolution, difficulty to properly distinguish multiple modes and an imprecise understanding of near-field effects (Zywicki 1999; Yoon 2005).

With the availability of multi-channel data acquisition systems, most modern surface wave procedures utilize a spatial array of multiple receivers. A multi-offset estimate averages the errors of the single traces and gives more reliable results compared with the usual spectral analysis of surface waves, even with the same array length. Moreover, multi-station approaches can detect easily multi-modal propagation (Strobbia 2003). For all these reasons multichannel analysis of surface waves (MASW) has become popular over the past 10 years (Park, Miller and Xia 1999; Xia, Miller and Park 1999). The price to be paid for this enhanced data quality is spatial averaging and loss of lateral resolution in the reconstruction of shear-wave velocity cross-section, effectively producing a single 1D shear velocity profile.

The need to overcome these limitations has led recently to attempts to produce (pseudo-)2D sections using *a priori* continuity assumptions (with some arbitrary parameter choice, e.g., Socco *et al.* 2008) or phase shifts between neighbouring traces (going effectively back to spectral analysis of surface waves and its limitations, e.g., Luo *et al.* 2008). With the aim of facing these issues, we describe in this paper a possible approach to maximize the number of channels without compromising lateral resolution and with no introduction of arbitrary continuity constrains. The prior data analysis suggested here allows the selection of the appropriate number of adjacent traces to perform the standard MASW.

In the present paper we deal with multi-station gathers: a source shot is recorded by multiple receivers placed along a straight line (McMechan and Yedlin 1981; Xia *et al.* 1999;

Park *et al.* 1999). The aim is mainly to estimate the shear-wave velocity profile of the subsurface by inverting the modal curves. Recently, several attempts have been made to invert the entire multi-modal plot. The traditional (Beatty and Schmitt 2003; Xia *et al.* 2003; Song and Gu 2007) simultaneous inversion of the fundamental and higher mode data has nothing special in terms of algorithms except for including higher mode data in the inversion process as an extra set of data with equal weighting or different weighting dependent on data accuracy. A possible advantage of multi-modal inversion is in its capability to reduce the ill-posedness of the problem by including extra independent data (i.e., higher modes). In this paper, we focus on the inversion of the fundamental mode of the dispersion curves, as in this case the analysis of phase versus offset is straightforward. Extensions to multi-modal are possible in order to improve stability of the inversion process and increase resolution of the final result but only when one mode is predominant, otherwise features due to discontinuities could be mistaken with those generated by mode jumps.

The underlying hypothesis to perform the fundamental mode or multi-mode inversion is that a 1D elastic layered half-space is representative of the subsoil under investigation. Thus, even with a perfect, noise-free, (unrealistic) data set and an exact forward modelling, errors can arise if the earth portion to be reconstructed is not appropriately parameterized by a 1D model.

The most usual approach to the processing of multi-offset surface wave data uses integral transformations (McMechan and Yedlin 1981; Gabriel *et al.* 1987; Park, Miller and Xia 1998): data are transformed from time-space domain into another domain (e.g., frequency-wavenumber) where the points of dispersion curve can be found as the maximizers of amplitude spectrum. This method is very robust but because taking into account mainly amplitude, disregards the information carried by the phase spectrum; especially this approach does not allow checking consistency of the one-dimensional parameterization. Recently, an approach named multi-offset phase analysis has been developed (Lin and Chang 2004; Strobbia and Foti 2006). Multi-offset phase analysis is a procedure for dispersion curve evaluation by linear regression of phase-offset data. In their original paper, Strobbia and Foti (2006) suggested to use multi-offset phase analysis in order to recognize lateral velocity variations by studying, for different shots and at some appropriate frequencies, changes of slope in phase-offset plot that occur at the same spatial position. In principle, in presence of a 1D elastic layered half-space, the phase-versus-offset curve at any frequency shall be a straight line.

In the present work the multi-offset phase analysis idea is pushed further: we propose an automated search of knee points at all frequencies, in order to derive a quantitative indication of offsets where lateral discontinuities occur. This novel procedure, unlike the original multi-offset phase analysis formulation, does not need: an estimation of noise level (in fact we do not perform any statistical test assuming the knowledge of data variance; e.g.,  $\chi^2$  test), nor multiple shots (considering all frequencies at the same time is enough to detect lateral discontinuities evidences).

The paper is organized as follows. We initially summarize the mathematical aspects of the data processing required for our approach to surface wave analysis. Second, we describe the adopted data inversion procedure. Then a section is dedicated to four numerical tests: the proposed procedure is shown to yield robust results in presence of lateral discontinuities and different levels of noise. Finally we draw our conclusions and considerations for future work.

## DATA ANALYSIS PROCEDURE

We consider situations where the subsoil is made of horizontal, dissipative, isotropic and homogeneous layers, bounded above by a free surface and below by a homogeneous isotropic homogeneous half-space; in this situation the surface waves response of the system can be described in the frequency domain by the superposition of modal contributions (Aki and Richards 2002):

$$s(\omega, x) = \sum_m s_m(\omega, x) = \sum_m A_m(\omega, x) e^{j(\omega t - k_m(\omega)x)}, \quad (1)$$

where  $x$  is the source-receiver distance and  $\omega$  is the angular frequency, while  $k_m(\omega)$  and  $A_m(\omega, x)$  are the wavenumber and the amplitude of the  $m$ -th mode;  $j$  is the complex unit.

Amplitude can be split into two parts: one dependent on the source only and the other function of the path only; consequently the contribution of the  $m$ -th mode to the total displacement can be expressed as follows (Strobbia and Foti 2006):

$$s_m(\omega, x) = I(\omega) R_m(\omega) \frac{e^{-\alpha_m(\omega)x}}{\sqrt{x}} e^{j(\omega t - k_m(\omega)x + \varphi_0(\omega))}, \quad (2)$$

$s_m(\omega, x)$  is a combination of: the source spectrum  $I(\omega)e^{j\varphi_0(\omega)}$ , the path response  $R_m(\omega)$  with geometric (represented by the factor  $1/\sqrt{x}$ ) and material (coefficient  $\alpha_m(\omega)$ ) attenuation and the modal phase:

$$\phi(\omega, x) = -k_m(\omega)x + \varphi_0(\omega). \quad (3)$$

The experimental phase  $\phi(\omega, x)$  is extracted from time-to-frequency Fourier transform applied on the recorded seismogram. Consequently, under the assumption that one mode (the  $m$ -th one) is predominant in the propagation, the wavenumber  $k_m(\omega)$  can be recovered by using a simple linear regression, as shown in equation (3).

Strobbia and Foti (2006) pointed out how to improve the quality of regression procedure by using data from multiple shots: several acquisitions with the same setup allow the evaluation of the statistical distribution of data; hence, not only the mean of the phase but also its standard deviation can be estimated. Eventually this information surplus can be used to propagate uncertainty from the phase to the wavenumber domain and then from the latter to the phase velocity. Ultimately, phase velocity uncertainty can lead to an estimation of the uncertainty associated with shear-wave velocity values and thicknesses resulting from the inversion process.

In the present study, we do not focus on such error propagation through the acquisition chain, so the regression process is performed with no weights.

Equation (3) holds for each receiver. Thus, considering  $N$  geophones:

$$\phi_i(\omega, x_i) = -k_m(\omega)x_i + \varphi_0(\omega), \quad i = 1, \dots, N. \quad (4)$$

The system of linear equations can be written in matrix notation:

$$\Phi = \mathbf{G}^* \mathbf{m}, \quad (5)$$

where  $\Phi$  is the vector of known observed phases  $\phi_1, \phi_2, \dots, \phi_N$ , while the components of vector  $\mathbf{m} = [k_m \ \varphi_0]^T$  are the coefficients of the interpolating straight line to be found and  $\mathbf{G} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ 1 & 1 & \dots & 1 \end{bmatrix}^T$  is the Vandermonde data matrix (Macon and Spitzbart 1958; Horn and Johnson 1991).

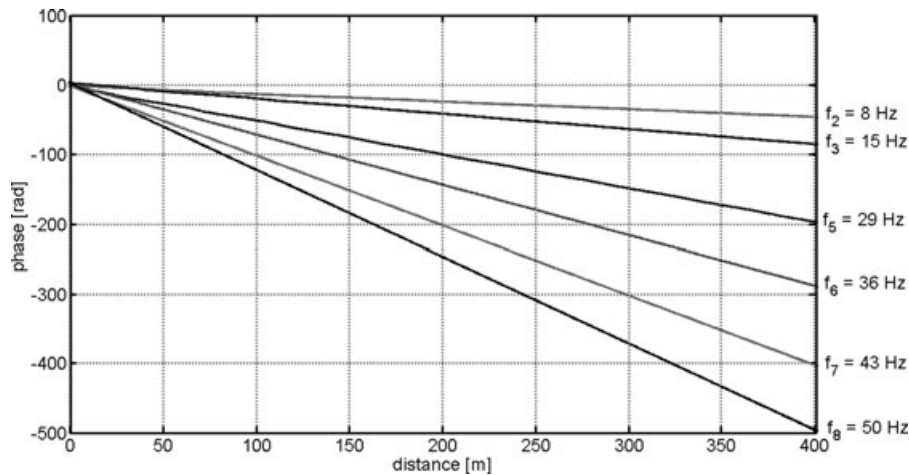
The matrix  $\mathbf{G}^T \mathbf{G}$  is always positive definite (at least anytime the  $x_i$  are distinct from each other) so it is invertible and the pseudo-inverse of  $\mathbf{G}$  is:

$$\mathbf{G}^{-g} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T, \quad (6)$$

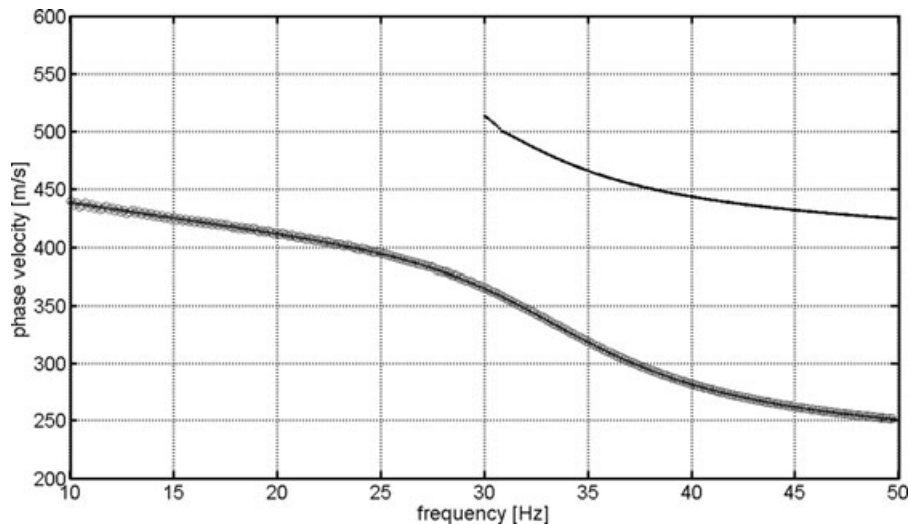
so that the least-squares solution for the coefficients is:

$$\mathbf{m} = \mathbf{G}^{-g} \Phi. \quad (7)$$

If propagating the data error estimation to evaluate the polynomial coefficient variance is not important, approaches not involving expensive inverse matrix calculation are to be



**Figure 1** Result of application of multi-offset phase analysis to a synthetic data set: phase response as a function of offset is plotted for different frequencies. The synthetic model consists of a layer (3 m thick) over a half-space, both homogeneous, having S-wave velocities equal respectively to 250 m/s and 500 m/s.



**Figure 2** Dispersion curve for the same model of Fig. 1: dot-lines represent the first two modes calculated by using transfer matrix algorithm as described by Dunkin (1975), while circles show the most energetic propagation mode extracted from the synthetic seismogram calculated with a spectral element code (Ampuero 2008) for that model. The fact that the circles coming from the numerical model lie precisely on top of the lower dotted line confirms that most of the energy propagates only through the fundamental mode, so that equation (4) and the general multi-offset phase analysis approach is easily applicable.

preferred because they are faster; for instance, exactly the same result with less effort can be reached by means of QR factorization with column pivoting (for example, as implemented in the backslash operator in Matlab; in this case, simply:  $m = G \setminus \Phi$ ).

Figure 1 shows an example of multi-offset phase analysis application to a synthetic data set: the simulated seismogram is transformed into a frequency-offset spectrum with a Fourier transform and the phase is extracted. Because the model rep-

resents a laterally homogeneous soil with a thin layer over a half-space and the propagation mode is the principal fundamental one (Fig. 2), phase-offset points lie on a straight line with slope coefficient equal to the wavenumber  $k(\omega)$  related to frequency. In principle, from these recovered wavenumbers it is possible to derive the phase velocities for each frequency, hence the dispersion curve.

Theoretically, multi-offset phase analysis can be used in place of standard frequency-wavenumber analysis (Strobbia

and Foti 2006). The latter approach extracts the dispersion curve from a double Fourier transform of the observed signal: collected data are mapped from time-offset ( $t-x$ ) domain to frequency-wavenumber ( $f-k$ ) domain; the maxima of the  $f-k$  spectrum amplitude single out the phase velocities  $v_r = \omega/k(\omega)$  at which excitation can propagate. In the following, the multi-offset phase analysis is used to verify the consistency between the observed data and the 1D parameterization, at least on selected portions of the subsurface. The traditional  $f-k$  analysis is used to invert the data in the regions where the 1D approach is deemed acceptable by the multi-offset phase analysis.

The obtained dispersion data are generally inverted using a forward model simulating far-field plane Rayleigh waves in laterally homogeneous media. When inverting real field data, attention must be paid to two aspects that are not contained in this forward modelling: near-field effects and lack of lateral homogeneity.

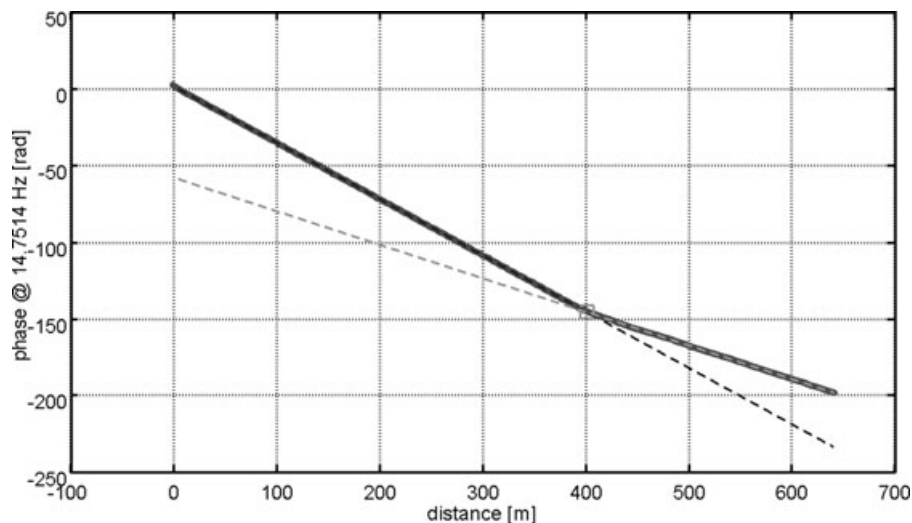
$f-k$  analysis has been used for many years and has demonstrated its capabilities (Gabriels, Sneider and Nolet 1987, Xianhai and Hanming, 2007) and limitations (Al Hunaidi and Rainer 1995; Foti 2000; Lu and Zhang 2004). In particular  $f-k$  analysis, as it ignores phase information, does not allow for consistency check with respect to the 1D subsurface hypothesis.

Multi-offset phase analysis can be effectively used for detecting such errors connected with wrong model assumptions. When one mode of Rayleigh waves dominates the recorded

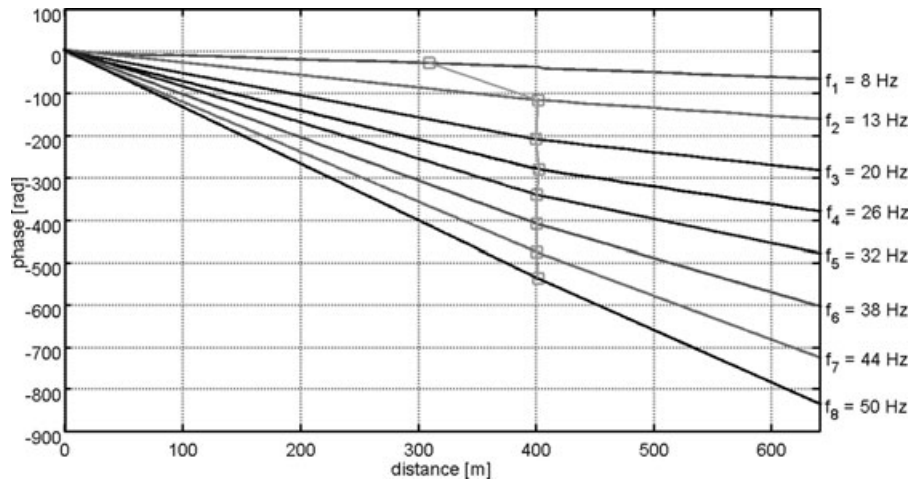
wavefield (so anytime multi-offset phase analysis is easily applicable) near-field effects and lateral heterogeneities can be detected by studying the phase-offset data at all frequency. Lateral contacts between formations manifest themselves as knees in the phase-versus-offset graph: in presence of lateral heterogeneities equation (4) is only piecewise valid and, for different subsets of contiguous geophone locations ( $x_i$ ), different values of  $k_m(\omega, x_i)$  are obtained for the same frequency  $\omega = 2\pi f$ . The location of slope discontinuities in the phase-versus-offset piecewise linear function is independent of frequency in case of lateral heterogeneities in the subsurface structure, while it is frequency-dependent when they represent effects of the near field.

In the framework of the multi-offset phase analysis procedure, Strobbia and Foti (2006) proposed a statistical test to evaluate the level of discrepancy between the model and the investigated site. On the path of Strobbia and Foti research, we propose in this paper an automatic analysis for detecting the position of lateral variation. Different from Strobbia and Foti (2006), we do not use a  $\chi^2$  test for phase-versus-offset fitting and so we neglect prior knowledge of data noise level; more important, we take into account changes of slope in the phase-versus-offset plots for all frequencies in order to locate the discontinuities.

In presence of a slope change (be it induced by the near-field or by geological discontinuities) the phase-versus-offset curve can be fitted with a piecewise function  $\phi^{(piece)}$  consisting of two straight lines as shown in Fig. 3. These two-part



**Figure 3** Best fitting of extracted phase-offset curve (dotted lines) with a piecewise linear function made up of two straight lines (dashed lines). The square indicates the knee point where the two straight lines intersect and the slope change occurs. This point, if occurring at the same location for different frequencies, locates the lateral variation of subsurface properties in the investigated site.



**Figure 4** Extracted phase-versus-offset curves for different frequencies. Data are generated by a synthetic model with an abrupt velocity change at 400 m from the source. In the vicinity of the correct offset value the slope changes are fairly well detected by the described piecewise linear function fitting (square marks) for all considered frequencies.

functions are completely characterized by four parameters: two slope coefficients,  $k_{sx}$  and  $k_{dx}$  and two intersections with the phase-axis,  $\varphi_{sx}$  and  $\varphi_{dx}$ . The fitting procedure is the same as in equation (5) but limited to a subset of the seismic profile.

For each  $x_i$  with  $i = 1, 2, \dots, N$ , one has:

$$\phi^{(piece)}(\omega, x_i) = \begin{cases} -k_{sx}(\omega)x_k + \varphi_{sx}(\omega), & \forall x_k \leq x_i \\ -k_{dx}(\omega)x_k + \varphi_{dx}(\omega), & \forall x_k > x_i \end{cases}, \quad (8)$$

while  $\mathbf{m}_{sx} = [k_{sx} \ \varphi_{sx}]^T$  can be found by using the pseudo-inverse of  $\mathbf{G}_{sx} = \begin{bmatrix} x_1 & x_2 & \dots & x_i \\ 1 & 1 & \dots & 1 \end{bmatrix}^T$  applied to  $\Phi_{sx} = [\phi_1 \ \phi_2 \ \dots \ \phi_i]^T$ :

$$\mathbf{m}_{sx} = \left( \mathbf{G}_{sx}^T \mathbf{G}_{sx} \right)^{-1} \mathbf{G}_{sx}^T \Phi_{sx} \quad (9)$$

or again, avoiding matrix inversion and using the Matlab notation, by:  $\mathbf{m}_{sx} = \mathbf{G}_{sx} \backslash \Phi_{sx}$ . An analogue procedure can be used for  $\mathbf{m}_{dx} = [k_{dx} \ \varphi_{dx}]^T$ ,  $\mathbf{G}_{dx} = \begin{bmatrix} x_{i+1} & x_{i+2} & \dots & x_N \\ 1 & 1 & \dots & 1 \end{bmatrix}^T$  and  $\Phi_{dx} = [\phi_{i+1} \ \phi_{i+2} \ \dots \ \phi_N]^T$ .

For every frequency  $\omega$  there exists at least one  $x_0 = (\varphi_{sx} - \varphi_{dx}) / (k_{dx} - k_{sx})$  minimizing the misfit functional  $M$ :

$$\begin{aligned} M(\omega, x_i) &= \left\| \Phi^{(piece)}(\omega, x_i) - \Phi(\omega) \right\|_{L_2}^2 \\ &= \sum_{i=1}^N \left( \phi_i^{(piece)}(\omega, x_i) - \phi_i(\omega) \right)^2. \end{aligned} \quad (10)$$

Such  $x_0$  is the horizontal location of a possible lateral discontinuity (Fig. 4). If consistent values of  $x_0$  are found at a

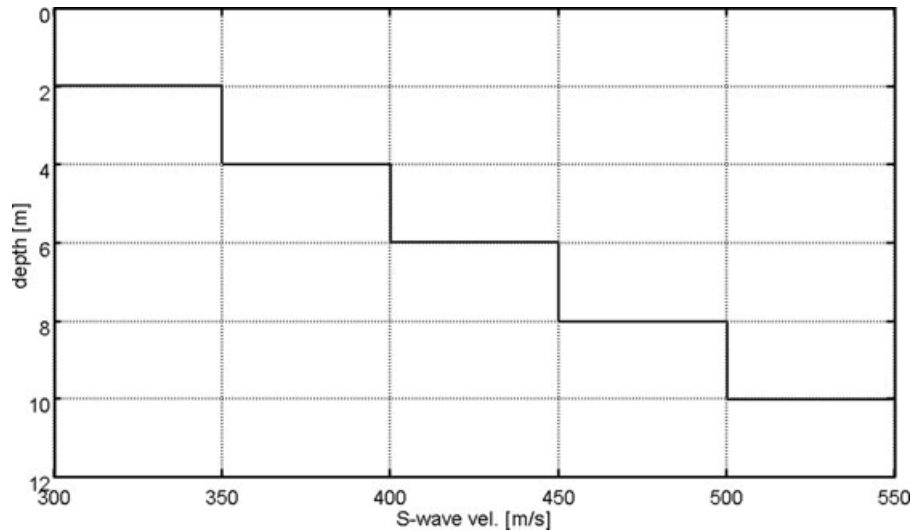
large number of sampled frequencies,  $x_0$  can be assumed as the location of a lateral discontinuity in the subsurface and it is reasonable to split the original profile into two parts. The subdividing process can continue until truly 1D sub-domains are identified. In this process it should be kept in mind that the longer the acquisition profile, the better the spectral resolution along  $k$ -direction; having a good resolution in the  $f$ - $k$  domain ensures an easier identification of individual modes (Socco and Strobbia 2004). Potentially this requirement may conflict with the subdivision of the data into shorter profiles.

The multi-offset phase analysis procedure described in this paper can be used to detect truly 1D ground portion and, anytime the profile is long enough, to invert them one next to the other without any additional constraint.

## DATA INVERSION

Data processing must be followed by inversion, i.e., the identification of the parameters of the conceptual model of the subsurface from the observed data. As shown elsewhere (Ewing, Jardetzky and Press 1957; Fowler 1990) the governing parameter for Rayleigh wave propagation is the shear-wave velocity distribution in the subsurface.

The aim of this research is not to develop a new technique for an effective inversion but to demonstrate how a better analysis of data allows a more reliable final result and to propose a strategy to reach this goal. For this reason, here we discuss only a straightforward least square solution, i.e., we seek the model  $\mathbf{q}$  with the smallest misfit between the



**Figure 5** Starting model for all linearized least square inversions described in the tests of this paper. It consists of a ramp with 5 steps of 50 m/s every 2 metres of depth from 300–550 m/s.

calculated and observed Rayleigh wave velocity:

$$\min_q \|v_r^{calc}(q) - v_r^{obs}\|, \quad (11)$$

the components of vector  $q = \{v_{s1}, \dots, v_{sN}; v_{p1}, \dots, v_{pN}; \rho_1, \dots, \rho_N; h_1, \dots, h_N\}$  are the model parameters.  $v_{s_i}$ ,  $v_{p_i}$ ,  $\rho_i$  and  $h_i$  are respectively: the S-wave and P-wave velocities, the density and thickness of the  $i$ -th layer; they are the quantities on which the dispersion characteristics of surface waves primarily depend (Murphy and Shah 1988; Aki and Richards 2002). For the sake of simplicity, in the synthetic models considered in what follows: density is taken constant ( $\rho_i = 2000 \text{ kg/m}^3$  for all  $i = 1, \dots, N$ ) and the Poisson ratio is set in all cases equal to 0.33. Hence, P-wave velocity is completely determined by S-velocity ( $v_{p_i} \cong 2v_{s_i}$ ). In this way the  $2N - 1$  variables to be reconstructed are the shear-wave velocities  $v_{s_i}$  and the layer thicknesses  $h_i$ . Any consideration about the (important) issue of ill-posedness of the inverse problem described by equation (11) is considered beyond the scope of this work and no discussion of the possible regularization will be given.

As a part of the inversion procedure, the choice of a forward model is essential. In this study, a dispersion curve computation has been implemented following Dunkin (1965) formalism; zeros of the non-linear secular function are found using Wathelet (2005) suggestions together with *fzero* Matlab function. In order to minimize the objective function (equation (11)) a linearized method (Nolet 1981, Tarantola 1987; Herrmann 1987; Zhdanov 2002) has been adopted: a sub-

space trust region method (Coleman and Li 1996) is used and, at each iteration, an approximated linear solution is calculated with a preconditioned conjugate gradients method (Mordecai 2003). This kind of solver is effectively implemented in the Matlab's function *lsqnonlin*.

The non-linearity of the problem, in addition to its intrinsic ill-posedness, makes the minimizer of equation (11) very sensitive to the choice of the starting point; for all the following synthetic models the starting model is the same and is shown in Fig. 5.

## NUMERICAL TESTS AND DISCUSSION

In this section four synthetic models are discussed to exemplify the proposed enhanced multi-offset phase analysis procedure: these are simplified earth models chosen to illustrate the effects of lateral heterogeneities and to show the capabilities of the above-described method; they are characterized by abrupt horizontal changes of S-wave velocity and increasing complexity. For each model the synthetic seismogram, corresponding to a point source and a linear acquisition array, has been calculated using the SEM2DPACK software: a 2D spectral element code for seismic wave propagation originally designed for earthquake dynamics studies (Ampuero 2008).

The excitation source in all cases is a Ricker wavelet (with frequency peak at 20 Hz and onset at 0.055 s) corresponding to a vertical motion; simulated recorded data are considered reliable up to 50 Hz. All profiles are 640 m long and data are

collected by vertical geophones spaced 1 m; some inaccuracy of receiver locations may arise case by case because of the correspondence to the needed mesh nodes.

In spite of the inaccuracy occurring in the receiver location (occasionally as large as 0.3 m), the multi-offset phase analysis procedure shows its robustness by providing excellent results. Moreover, for the first example, a sensitivity analysis is conducted with respect to data noise.

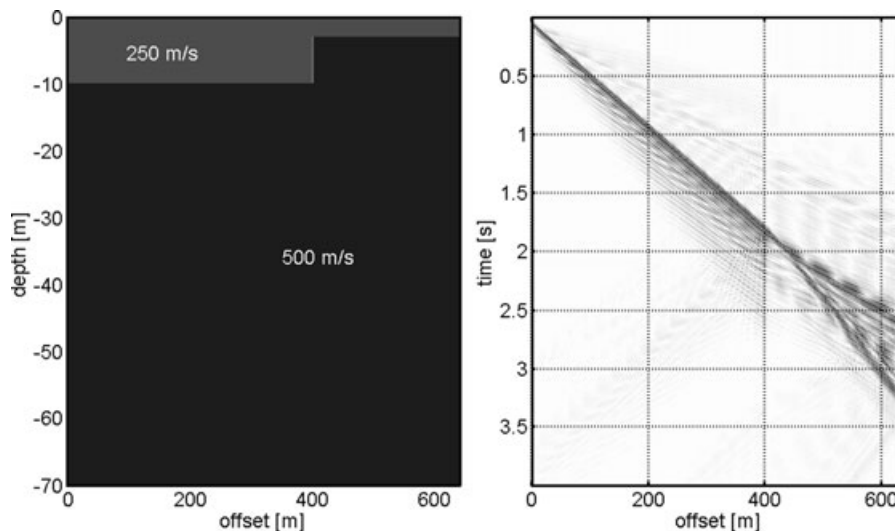
### Model A

As a first example a simple model is considered (Fig. 6), consisting of 3 layers: the top one is laterally homogeneous, is 3 m thick and has a S-wave velocity of 250 m/s; the second, from 3–10 m depth, presents a sharp velocity change at 400 m offset: the first part has an S-wave velocity equal to 250 m/s, while the second part has a velocity equal to 500 m/s; the bottom half-space has an S-wave velocity equal to 500 m/s. In summary, example A mimics the presence of a fault with a vertical throw equal to 7 m.

The corresponding synthetic seismogram is shown in Fig. 6. Note how the presence of a lateral discontinuity at 400 m offset manifests itself as a clear change of apparent velocity of the dispersive wave train.

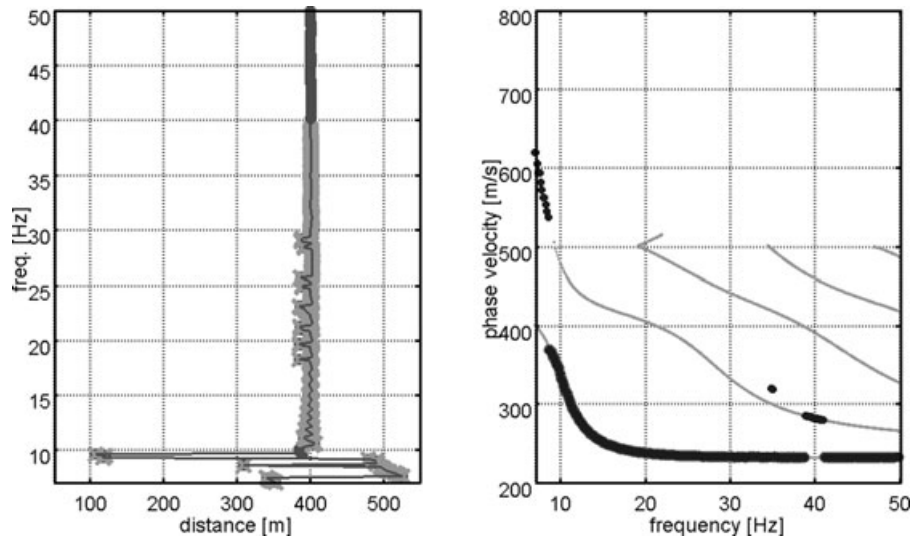
As discussed in the section ‘Data analysis procedure’, a one-dimensional Fourier transform from the time to the frequency

domain is sufficient in order to extract the phase information: thus obtaining  $\Phi$  (equation (5)) is straightforward. For each frequency, the minimization of functional  $\text{Min}$  equation (10) provides an estimation of the spatial location (offset) where the slope change occurs in the phase-versus-offset plot (Fig. 7, left panel). In this case the detected knee point locations lie at 400 m for nearly all frequencies but the lowest ones below 10 Hz. The lower frequencies correspond to Rayleigh waves with larger wavelengths (at 10 Hz the wavelength in this example goes from 25–50 m) that penetrate deeper in the subsurface (about one wavelength); it is no surprise that in this example these frequencies are not sensitive to the conditions in the shallow zone. On the contrary all frequencies higher than 10 Hz are very sensitive to the presence of a lateral heterogeneity in the first 10 m depth and Fig. 7 clearly marks the correct location of this lateral contrast. First, we neglected the presence of this lateral contrast and used the entire synthetic seismogram (Fig. 6, right panel) in a classical dispersion curve  $f$ - $k$  analysis: the corresponding energy peaks are in excellent agreement with the modal curves corresponding to the 1D velocity profile between 0–400 m. In particular, most of the energy peaks are concentrated on the fundamental mode, with some local jump to the second mode. In essence, the dispersion curve derived from the entire seismogram is insensitive to the presence of the lateral heterogeneity at 400 m offset. It is therefore not surprising if the inversion of this dispersion curve (Fig. 8) leads essentially to recovering only the

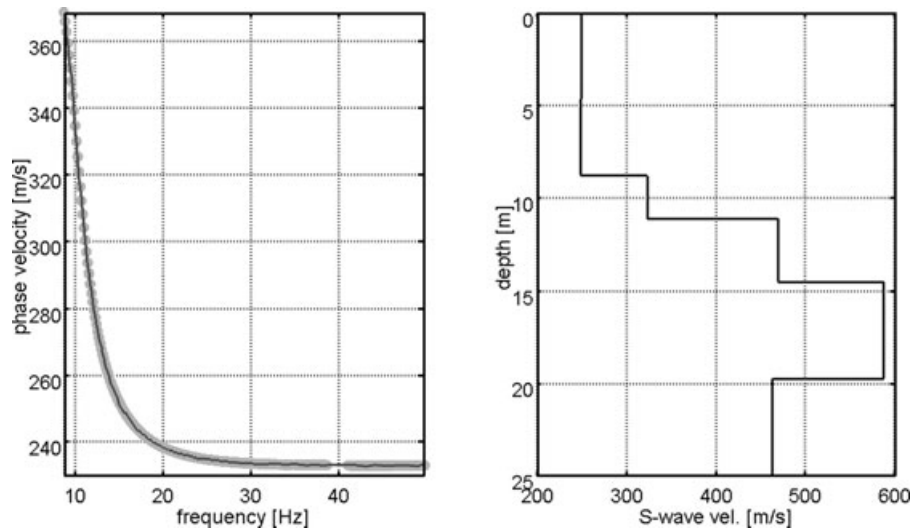


**Figure 6** Model A. The left-hand panel shows the geometry of the system consisting of 3 horizontal layers. The top layer is 3 m thick with S-wave velocity = 250 m/s; the second layer is 7 m thick and has a sharp discontinuity of S-wave velocity at 400 m offset: on the left its velocity equals 250 m/s, on the right 500 m/s; the bottom half-space is characterized by a velocity = 500 m/s. The corresponding synthetic seismogram obtained using SEM2DPACK (Ampuero 2008) is shown in the right-hand panel.





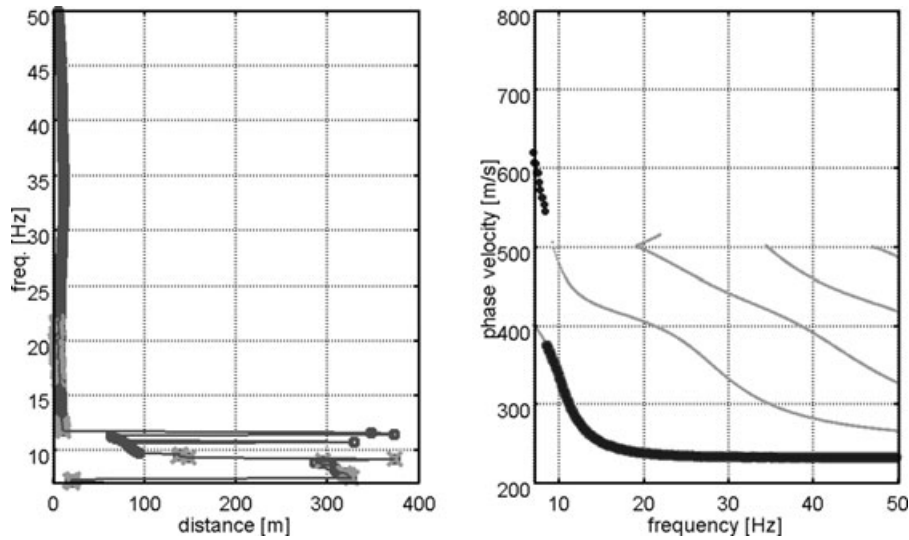
**Figure 7** Model A. In the left-hand panel the location is shown of the knee points  $x_0$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of slope coefficient greater than 20%. In the right-hand panel the dispersion curve extracted from the synthetic seismogram (Fig. 6, right-hand panel) is compared with the first three modal curves calculated for a velocity profile equal to the left part (0–400 m) of Model A. Apparently, the dispersion curve derived from the measured phase velocity is insensitive to the presence of the lateral heterogeneity at 400 m offset: propagation occurs mainly through the fundamental mode of a perfectly layered earth (with some jumps of the main energy peak to the second mode), with parameters corresponding to the left-hand side of the model (Fig. 6).



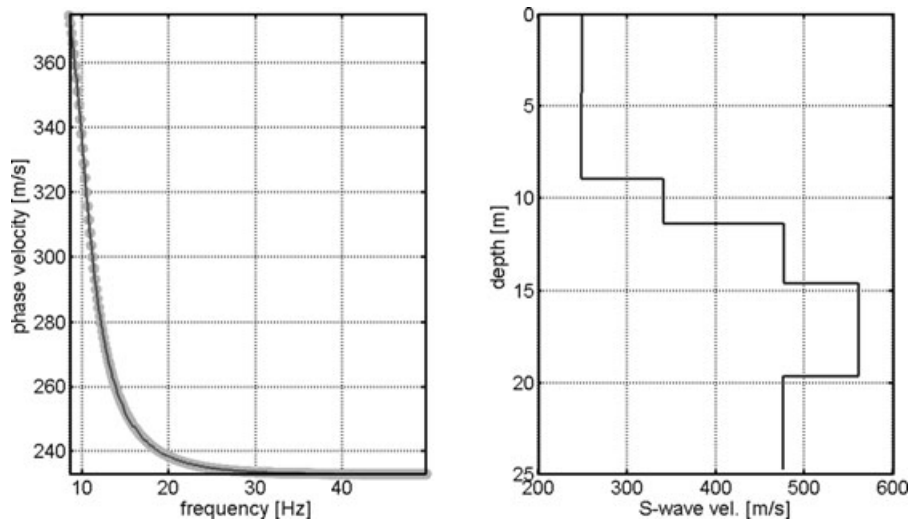
**Figure 8** Model A. In the left-hand panel, calculated (light line) and extracted (dark line) dispersion curves for Model A (Figs 6 and 7) are compared: the percentage cumulative data misfit is as small as 0.1%. On the right-hand panel the inversion result is shown. These results show that inversion of the entire seismogram (Fig. 6) essentially leads to a good identification of the subsurface structure between 0–400 m offset, while the velocity step at 400 m is basically invisible to this analysis.

velocity profile in the left-hand side of the acquisition array. The prevalence of this part is probably due to the larger extent of this portion (400 m out of 640 m total). However the result is disturbing, as it demonstrates that the classical analysis is perfectly capable of overriding a sharp velocity contrast that shows itself even in the raw data (Fig. 6, left panel).

We attribute this strong weakness essentially to the fact that the traditional  $f$ - $k$  approach neglects the phase information and puts all emphasis on the amplitude, in this case thus neglecting the smaller right-hand side of the profile that corresponds only to a smaller percentage of the total energy in the seismogram.



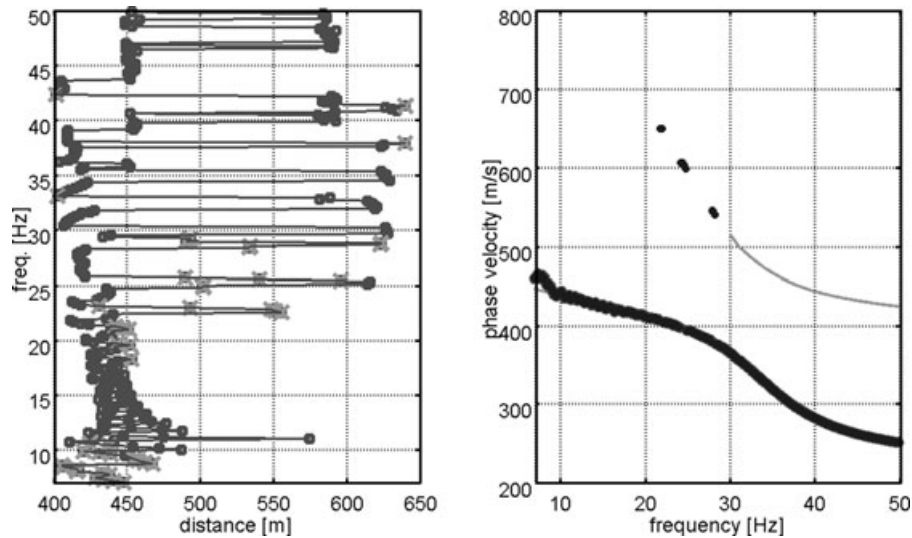
**Figure 9** Model A, analysis restricted to the left part (0–400 m). In the left-hand panel the location is shown of the knee points  $x_0$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset (0–400 m) is compared with the first three modal curves calculated for corresponding velocity profile.



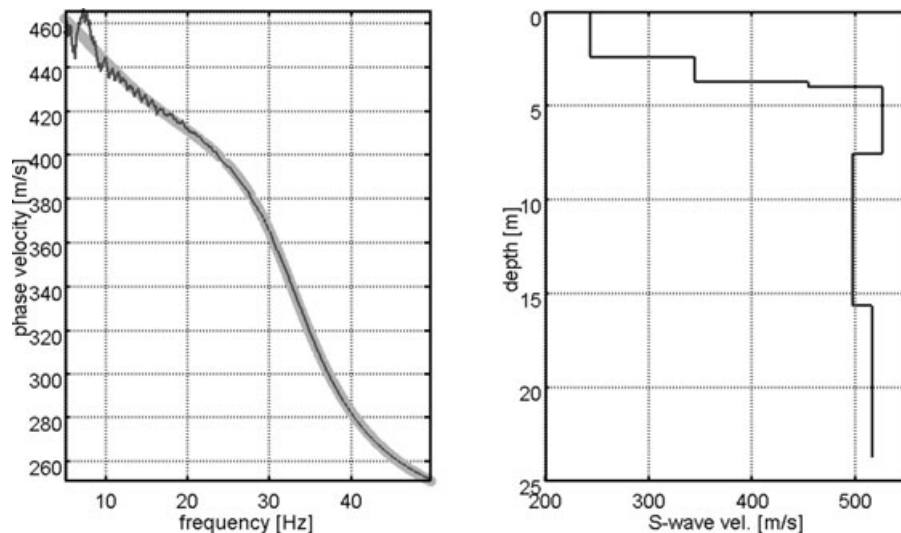
**Figure 10** Model A, analysis restricted to the left part (0–400 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.1%. On the right-hand panel the inversion result is shown. These results are basically indistinguishable from those in Fig. 8, where the entire seismogram (0–640 m) is used for inversion.

On the contrary, given the results of the multi-offset phase analysis in this example, it is natural to split the data set into two ensembles: 0–400 m and 401–640 m offset. We repeated the multi-offset phase analysis on the first subset (0–400 m, Fig. 9): the left panel shows that very few knee points correspond to a meaningful (greater than 20%) variation of slope coefficient and these changes occur very close to the source location. This may be indicative of the identification of a near-

field effect (Strobbia and Foti 2006) but it is more likely to be a numerical noise effect, as the first of the two fitting lines interpolates too few receiver locations to be significant. The inversion of the synthetic data on the 0–400 m offset range (Fig. 10) provides a good estimate of the corresponding velocity profile, practically indistinguishable from the result obtained using the entire seismogram from 0–640 m (compare Figs 8 and 10).



**Figure 11** Model A, analysis restricted to the right part (401–640 m). In the left-hand panel the locations of the knee points  $x_0$  are shown; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset is compared with the first two modal curves calculated for 1D velocity profile corresponding to the 401–640 m interval of Model A (Fig. 6): the excellent correspondence confirms that the data subset can be treated as representative of the subsurface structure in this portion of the profile.

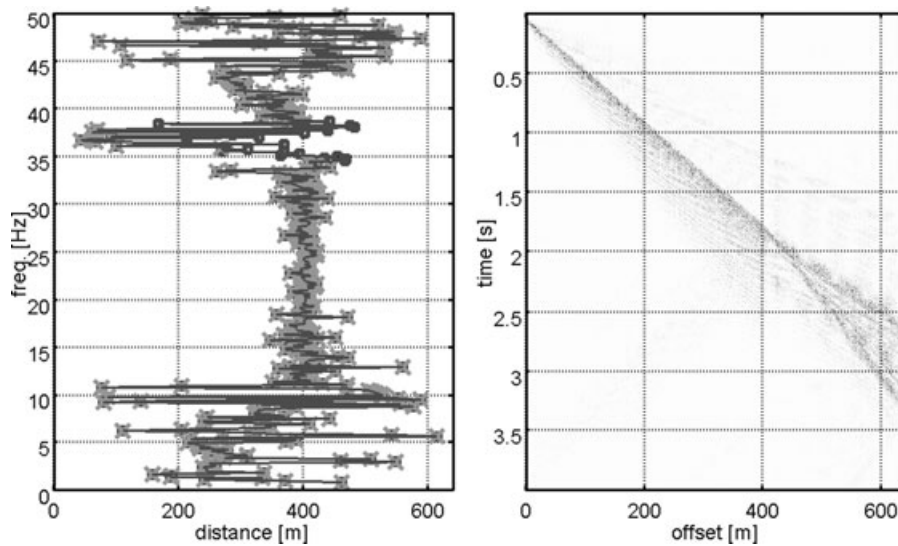


**Figure 12** Model A, analysis restricted to the right part (401–640 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.6%. On the right-hand panel the inversion result is shown. The estimated S-wave velocity 1D-profile is in good agreement with the subsurface structure in the 401–640 m offset range (Fig. 6).

The multi-offset phase analysis performed on the second subset of Model A (401–640 m) shows that the found knee points are scattered irregularly across the 420 m offset range (Fig. 11), thus indicating that no clear lateral discontinuity can be inferred from the data (indeed there is no discontinuity in this range). The energy peak locations in the  $f$ - $k$  spectrum lie comfortably on top of the fundamental modal curve cor-

responding to the S-wave velocity profile in the 401–640 m range (Fig. 11, right panel) and the inversion of the same data subset leads to identifying correctly the velocity profile in this range (Fig. 12).

In summary, the application of multi-offset phase analysis to Model A leads to the correct identification of the location of the lateral discontinuity, as well as to the correct



**Figure 13** Model A. Noise sensitivity analysis: a normally distributed noise (with mean 0 and standard deviation equal to twice the trace amplitude) is added to the original data. In the left panel the locations of the knee points  $x_0$  are shown; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. The corresponding noisy synthetic seismogram is shown in the right-hand panel.

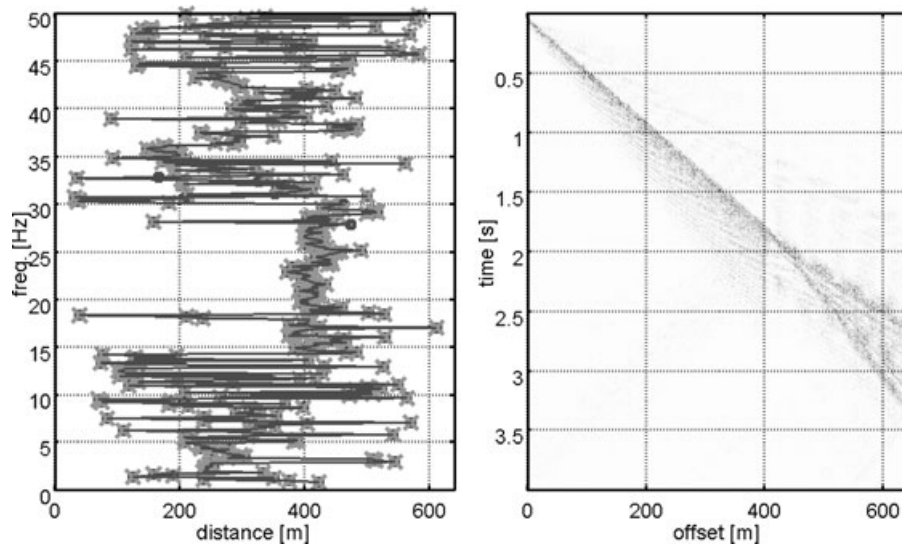
reconstruction of the velocity profiles on the right and the left of such discontinuity. On the contrary, a blind application of the traditional  $f$ - $k$  analysis and inversion on the entire seismogram leads to a 1D velocity profile that in essence corresponds to the structure of the largest part of the subsurface along the acquisition line, in this case the portion to the left side of the discontinuity.

As mentioned above, data are already corrupted by noise due to inaccuracy in the receiver locations. However, we wanted to study the method robustness with respect to specific different levels of random noise. In Fig. 13 we added noise to the original synthetic seismogram generated from Model A (Fig. 6): the left panel shows the multi-offset phase analysis result when we sum to every data point a normally distributed random noise with zero mean and standard deviation equal to twice the local signal amplitude. We can see that, even with this relatively high level of noise, the result is still stable and the location of the lateral discontinuity is clearly and correctly detected. In an attempt to find when the noise level becomes critical we perform the enhanced multi-offset phase analysis in presence of noise with progressively larger standard deviation. Multi-offset phase analysis result is still good in the presence of a noise characterized by a standard deviation equal to four times the amplitude (Fig. 14) and starts to fail only when the noise level reaches a standard deviation of five times the signal amplitude (Fig. 15). This confirms how stable the multi-offset phase analysis result is with respect to increasing noise level.

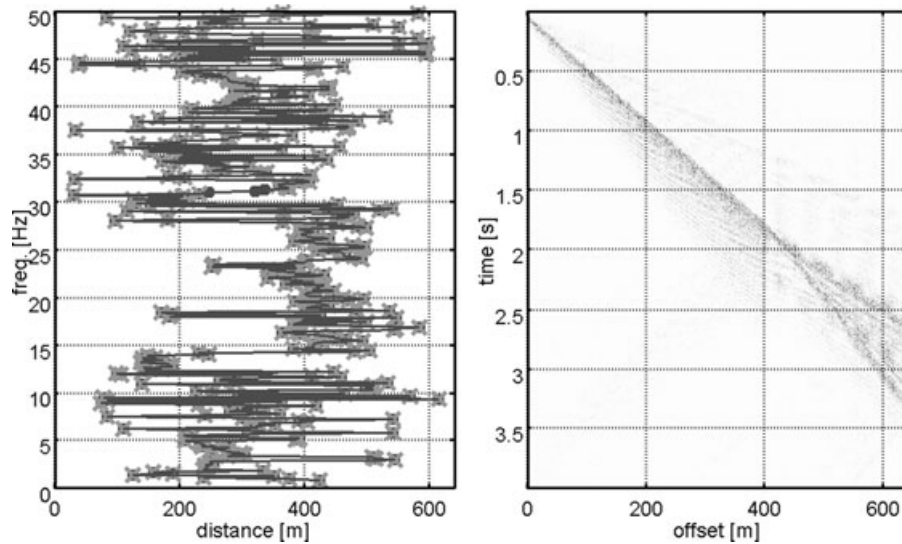
### Model B

This model is more complex than Model A, in that it is built with three different velocity values. The structure (Fig. 16) consists of two layers (respectively 3 m and 7 m thick) over the bottom half-space. The S-wave velocity of the top layer is 250 m/s; the second layer velocity changes sharply from 500 m/s to 250 m/s at 400 m offset. The half-space has a shear-wave velocity equal to 650 m/s. The MOPA analysis applied to the entire seismogram leads to the results in Fig. 17. Nearly all knee points are well aligned at 400 m, where the true discontinuity is located. The dispersion curve obtained from the classical  $f$ - $k$  analysis applied to the entire seismogram, as for Model A, is consistent with the modal curves relevant to the velocity model on the left side of the profile (0–400 m), i.e., to the prevalent part in terms of extent, even though some energy peaks jump apparently to a lower curve, consistent with the fundamental mode relevant to the right side (401–640 m) of the profile (Fig. 21). Note that in this case the second layer has on the left its portion with higher velocity, while the opposite was true in Model A.

Using the main dispersion curve extracted in Fig. 17 in the inversion procedure leads essentially to recovering the velocity profile of the left side (Fig. 18). This is consistent with the results for Model A: once again, a classical processing and inversion procedure applied to the entire profile leads to the identification of a 1D velocity profile that in essence corresponds to the structure of the largest part of the subsurface



**Figure 14** Model A. Noise sensitivity analysis: a normally distributed noise (with mean 0 and standard deviation equal to 4 times the trace amplitude) is added to the original data. In the left-hand panel the locations of the knee points  $x_0$  are shown; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. The corresponding noisy synthetic seismogram is shown in the right-hand panel.

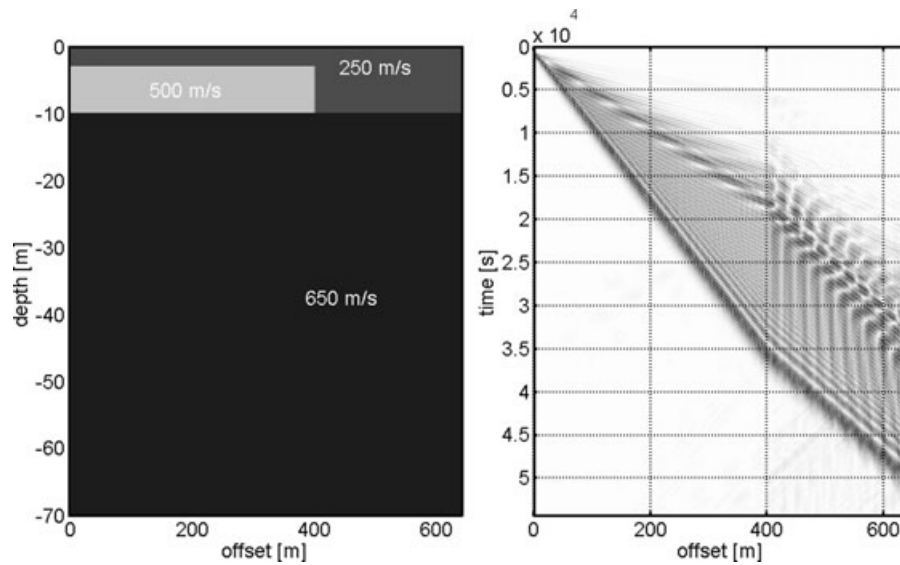


**Figure 15** Model A. Noise sensitivity analysis: a normally distributed noise (with mean 0 and standard deviation equal to 5 times the trace amplitude) is added to the original data. In the left-hand panel the locations of the knee points  $x_0$  are shown; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. The corresponding noisy synthetic seismogram is shown in the right-hand panel.

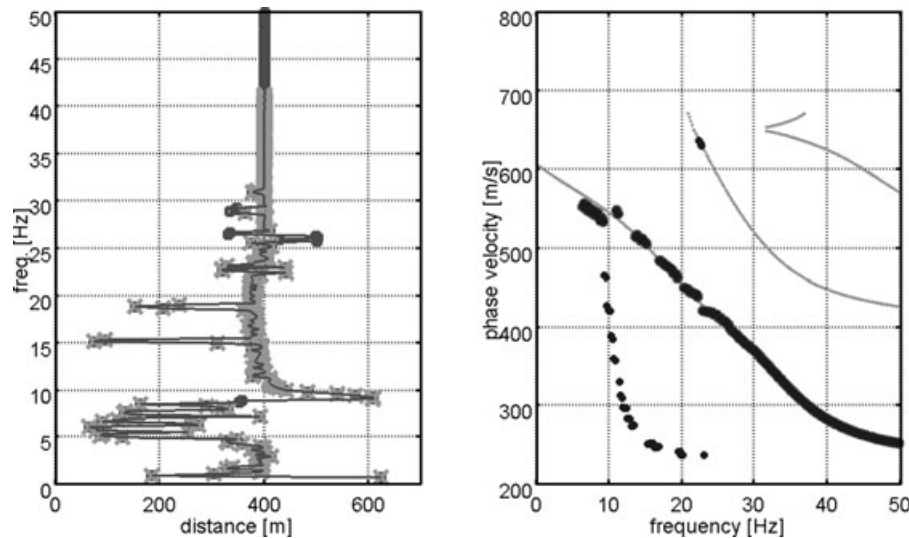
along the acquisition line (in this case the portion to the left side of the discontinuity), while little or no evidence exists of the presence of the discontinuity. The multi-offset phase analysis, on the contrary, very clearly identifies the presence of the discontinuity (Fig. 17).

We repeated the multi-offset phase analysis to the left side of the profile only: no meaningful location of discontinuities

can be identified for all frequencies (Fig. 19). The  $f$ - $k$  extracted dispersion curve lies precisely on top of the theoretical fundamental modal curve for the velocity profile of the left side and no energy leakage is observed here to modes related to the right side, as observed in Fig. 17. As a consequence, the velocity profile of the left side is identified with good approximation (Fig. 20).



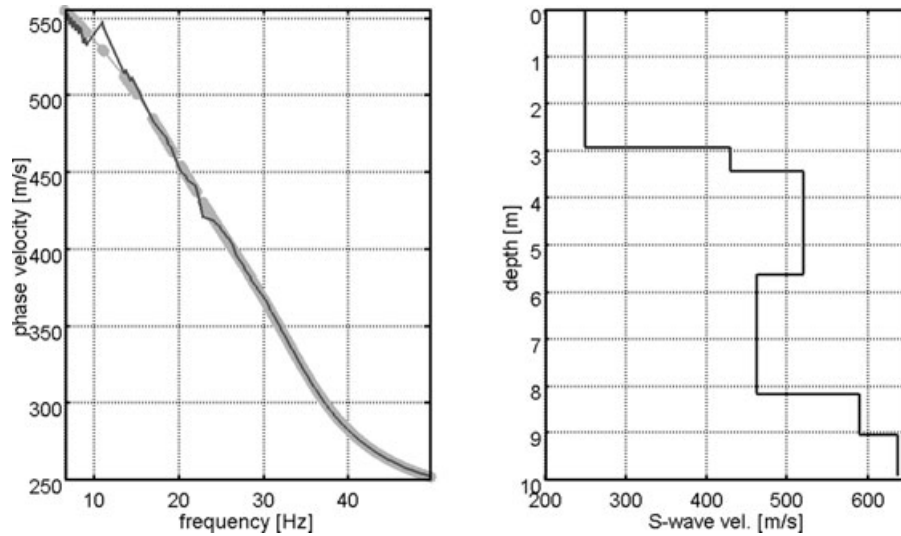
**Figure 16** Model B. The left panel shows the geometry of the system consisting of 3 horizontal layers. The top layer is 3 m thick with S-wave velocity equal to 250 m/s; the second layer is 7 m thick and has a sharp discontinuity of S-wave velocity at 400 m offset: on the left, its velocity equals 500 m/s, on the right 250 m/s; the bottom half-space has a velocity = 650 m/s. The corresponding synthetic seismogram obtained using SEM2DPAK (Ampuero 2008) is shown in the right-hand panel.



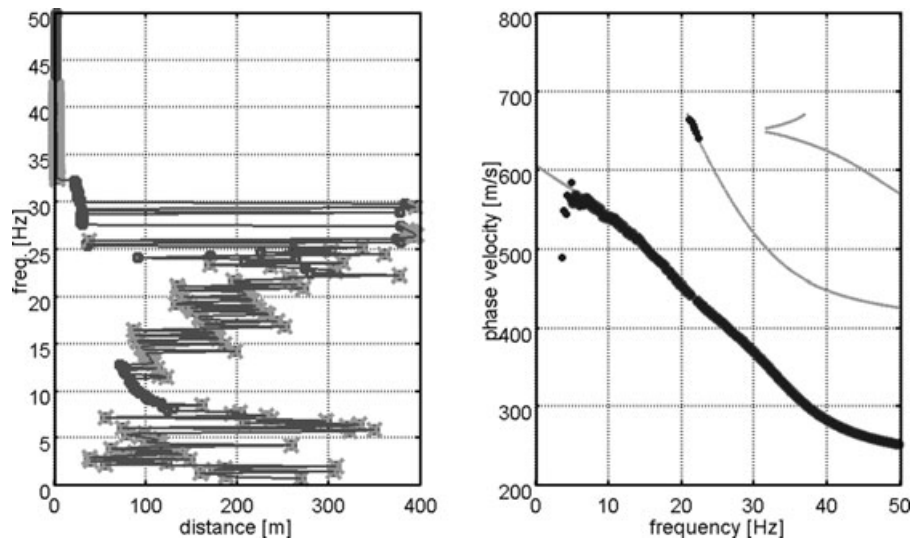
**Figure 17** Model B. In the left panel the locations are shown of the knee points  $x_i$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. In the right-hand panel the dispersion curve extracted from the synthetic seismogram (Fig. 16, right panel) is compared with the first two modal curves calculated for a velocity profile equal to the left part (0–400 m) of Model B. Even though the majority of energy peaks align on the fundamental mode, some serious discrepancy is observed at low frequencies (below  $\sim 20$  Hz); this energy is aligning on the fundamental mode of the right-hand (401–640 m) side of the model (Fig. 21).

When the multi-offset phase analysis is performed on the right side, again no meaningful knee point is identified (Fig. 21, left panel). However, the  $f$ - $k$  processing of this part of the seismogram leads to identifying an apparent dispersion curve (Fig. 21, right panel) that is in fact generated

by the jump of energy from the fundamental mode to the second mode at frequencies lower than 10 Hz. This phenomenon is due to the quasi-oscillating nature of the first two modes and has been observed and discussed elsewhere (Cercato 2009). In real (noisy) cases, mixing up of quasi-



**Figure 18** Model B. In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.6%. On the right-hand panel the inversion result. Similar to Model A, the result shows that inversion of the entire seismogram (Fig. 16, right panel) leads to a good identification of the subsurface structure between 0–400 m offset, while the velocity step at 400 m is basically invisible to this analysis.

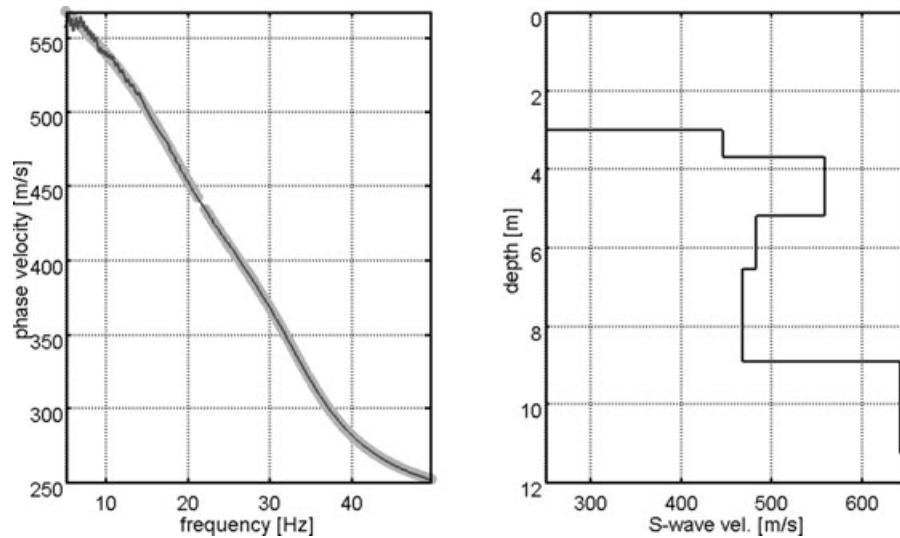


**Figure 19** Model B, analysis restricted to the left part (0–400 m). In the left-hand panel the locations are shown of the knee points  $x_0$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of slope coefficient greater than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset (1–400 m) is compared with the first three modal curves calculated for corresponding velocity profile.

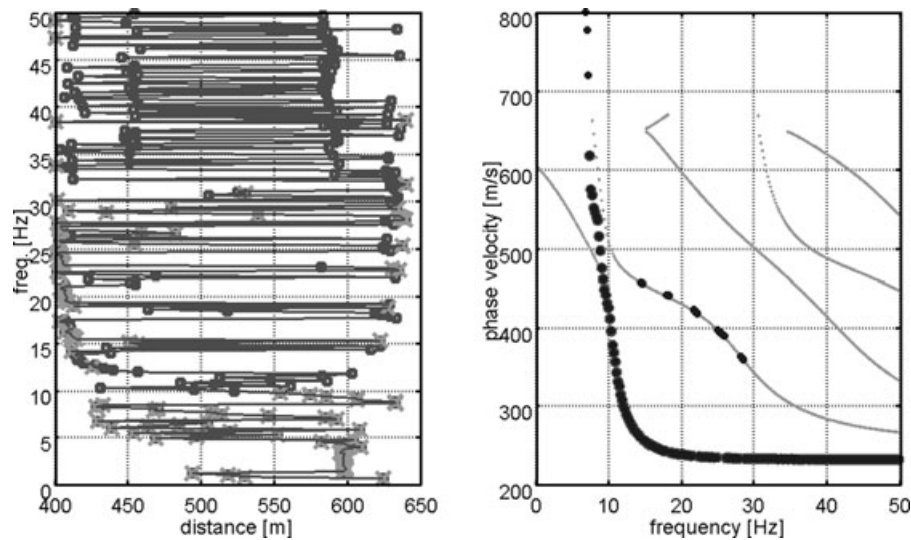
oscillating modes can be even more severe because of poor velocity resolution at low frequency; indeed inaccuracy in the extracted phase velocity  $v_r = \omega/k(\omega)$  is proportional to the inaccuracy of  $k$  and inversely proportional to the square of the  $k$  value:  $\sigma_{v_r} = |\omega/k^2| \sigma_k$ ; which implies a higher degree of uncertainty in the low-frequency range where wavenumbers are small. For this example, the data misinterpretation due to the

quasi-oscillating modes causes that the inverted model (Fig. 22) reproduces fairly accurately the true velocity profile only to about 20 m depth while, deeper than that, an exceedingly high shear velocity is recovered from the inversion.

In summary, also for Model B the multi-offset phase analysis leads to a meaningful identification of the lateral discontinuity, while the application of the traditional  $f$ - $k$  analysis and



**Figure 20** Model B, analysis restricted to the left part (0–400 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.4%. On the right-hand panel the inversion result. The result is very similar to the one obtained using the entire seismogram (Fig. 18).



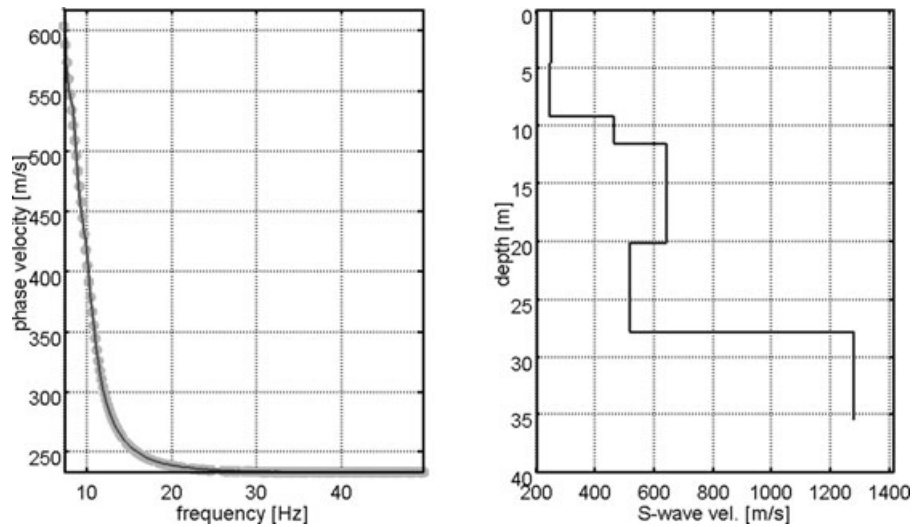
**Figure 21** Model B, analysis restricted to the right part (401–640 m). In the left-hand panel the locations are shown of the knee points  $x_0$  minimizing equation (10); the light grey crosses represent knee points characterized by a variation of slope coefficient larger than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset is compared with the first five modal curves calculated for a 1D velocity profile corresponding to the 401–640 m interval of Model B (Fig. 16). Unlike for Model A (Fig. 11), the curve extracted from the seismogram does not lie on the fundamental mode: the energy peaks jump at low frequency (about 10 Hz) from the fundamental to the second mode that is nearly oscillating the fundamental one. This phenomenon leads to an apparent dispersion curve that can be mistaken for a fundamental mode (Cercato 2009).

inversion on the entire seismogram leads to a 1D velocity profile that corresponds to the structure of the largest part of the subsurface along the acquisition line and can produce highly misleading results.

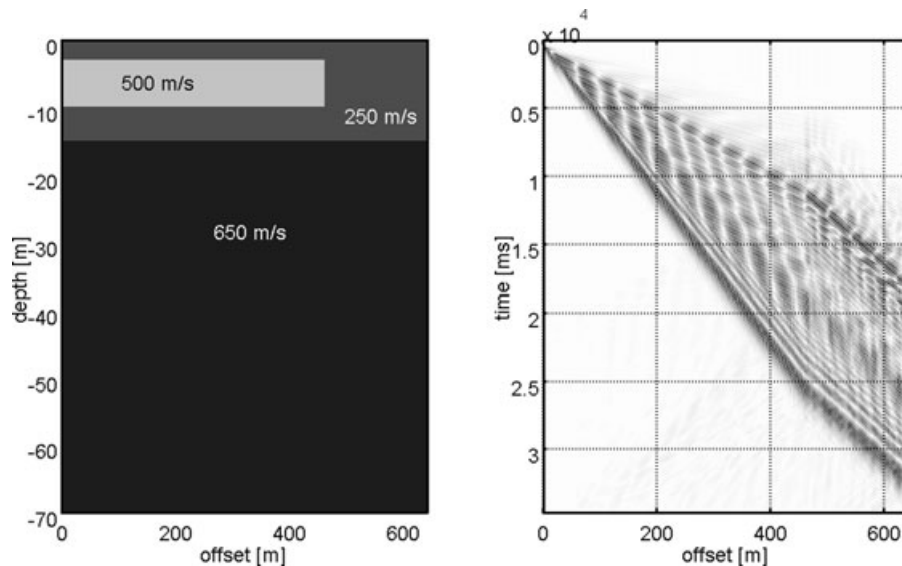
### Model C

In this example we investigate the possibility of detecting the boundaries of a high velocity inclusion (500 m/s), 7 m thick, extended for 460 m, embedded in a layer with a thickness of





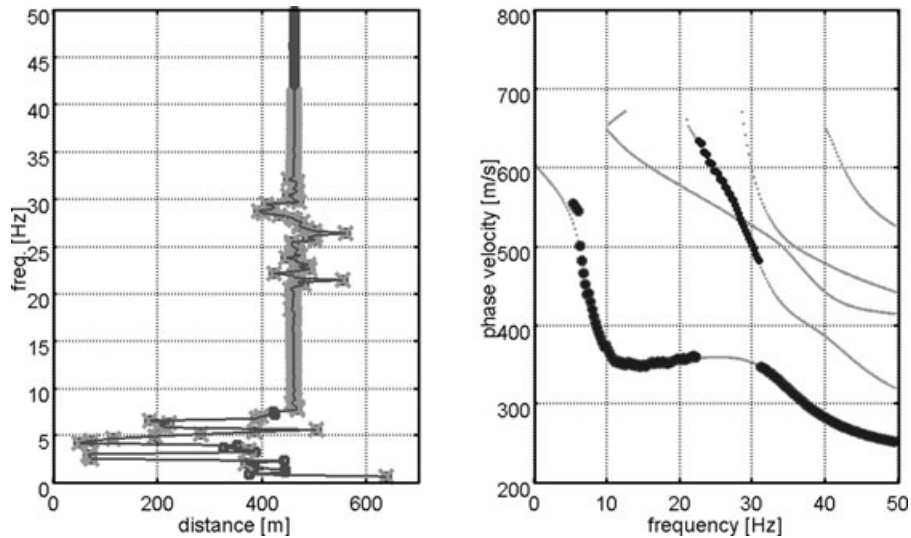
**Figure 22** Model B, analysis restricted to the right part (401–640 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.9%. On the right-hand panel the inversion result. Note that in this case the inverted model reproduces fairly accurately the true velocity profile between 401–640 m only in the upper part, to about 20 m depth. Deeper than that, the apparent dispersion curve (Fig. 21) points towards a very high velocity at low frequency, which is not present in the true model but is recovered from the inversion.



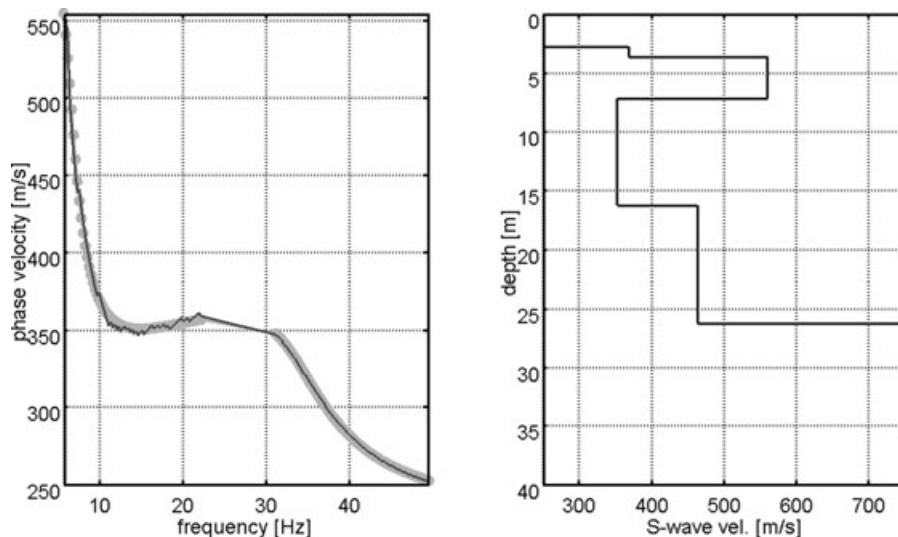
**Figure 23** Model C. The left-hand panel shows the system geometry, consisting of 4 horizontal layers. The top layer is 3 m thick with S-wave velocity equal 250 m/s; the second layer is 7 m thick and has a sharp discontinuity of S-wave velocity at 460 m: on the left the velocity equals 500 m/s, on the right 250 m/s; the third layer is 5 m thick and has a velocity of 250 m/s; the bottom half-space has a velocity = 650 m/s. This model, unlike models A and B, presents a velocity inversion. The corresponding synthetic seismogram obtained using SEM2DPAK (Ampuero 2008) is shown in the right-hand panel.

15 m and a velocity of 250 m/s over a half-space with 650 m/s velocity (Fig. 23). For this example, as for the previous ones, the multi-offset phase analysis determines the existence and the precise location of the vertical discontinuity (Fig. 24, left panel). However, the presence of the velocity inversion affects

multi-offset phase analysis. In the right panel of Fig. 24 it is apparent how the velocity inversion brings the energy peaks to align with the higher modes corresponding to the left side of the profile (where the inclusion is present) particularly in the frequency range between 22–31 Hz. In the same range, the



**Figure 24** Model C. In the left panel the locations are shown of the knee points  $x_j$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%. In the right-hand panel the dispersion curve extracted from the synthetic seismogram (Fig. 23, right panel) is compared with the first five modal curves calculated for a velocity profile equal to the left part (0–460 m) of Model C. Here the velocity inversion induces an energy jump from the fundamental to the higher modes between about 22 and 30 Hz: note that in the same frequency range the multi-offset phase analysis fails to identify the discontinuity at 460 m (left panel). In essence the dispersion curve behaviour is totally controlled by the prevalent left-hand side of the velocity model (Fig. 23).

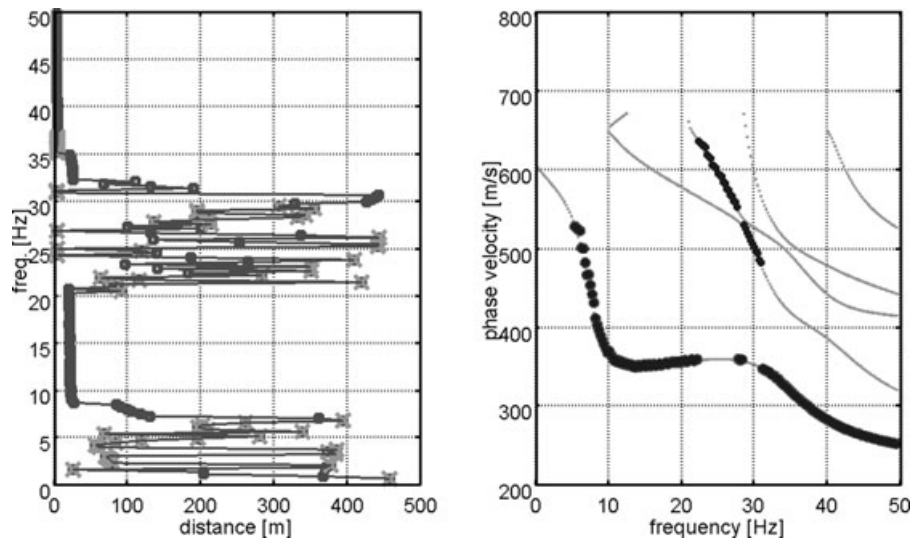


**Figure 25** Model C. In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.9%. On the right-hand panel the inversion result. Similar to Models A and B, the result shows that inversion of the entire seismogram (Fig. 23, right-hand panel) leads to a good identification of the subsurface structure, in this case including the velocity inversion, between 0–460 m offset, while the velocity step at 460 m is invisible to this analysis.

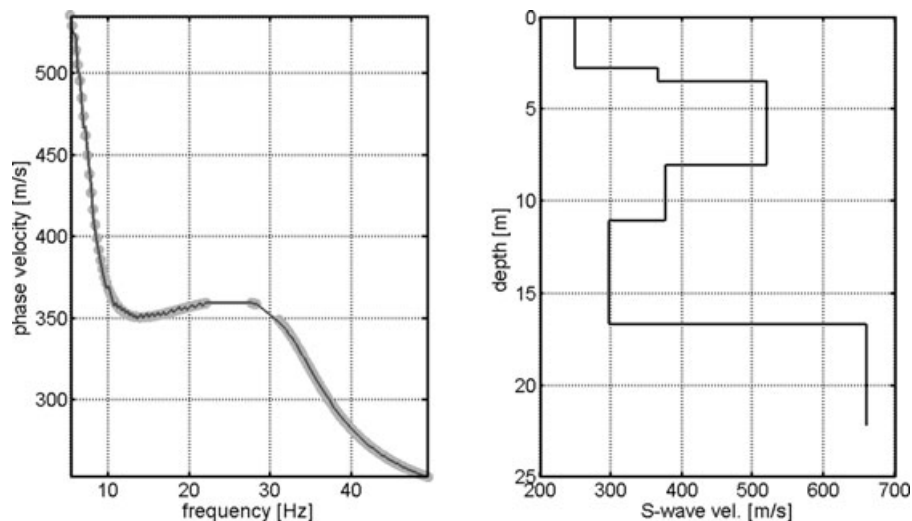
multi-offset phase analysis fails to locate the discontinuity at 460 m. This is hardly surprising, as the MOPA is inherently tuned to work on one mode at a time (equations (3)–(4)).

Inversion of the entire seismogram (Fig. 25) leads to the identification of the velocity inversion, even though both

thickness and velocity value are not estimated precisely. As in the previous examples, the classical processing and inversion procedure applied to the entire profile leads to the identification of a 1D velocity profile that corresponds (at least roughly) to the structure of the largest part of the subsurface



**Figure 26** Model C, analysis restricted to the left part (0–460 m). In the left-hand panel the locations are shown of the knee points  $x_0$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of slope coefficient greater than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset (1–460 m) is compared with the first three modal curves calculated for corresponding velocity profile.



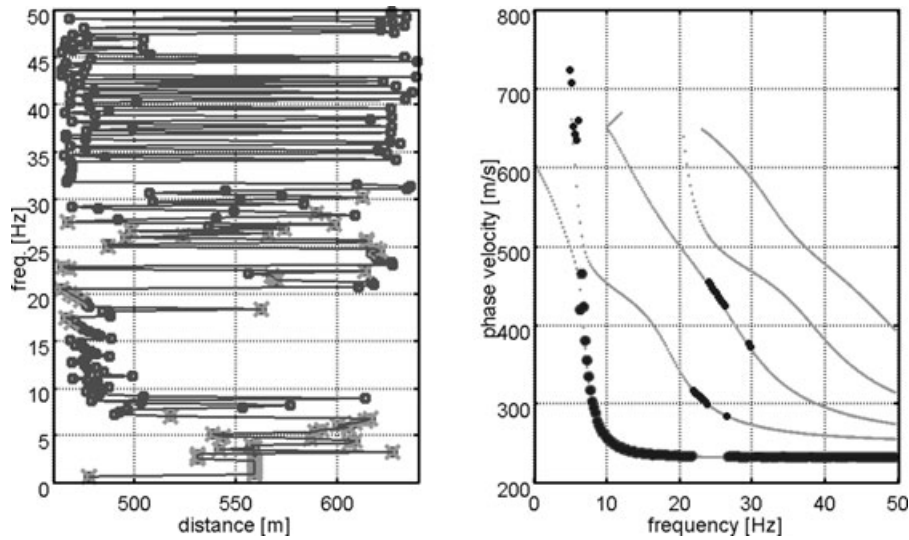
**Figure 27** Model C, analysis restricted to the left part (0–460 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 0.5%. On the right-hand panel the inversion result. The result is more accurate but similar to the one obtained using the entire seismogram (Fig. 25).

along the acquisition line (in this case the portion to the left side of the discontinuity), while no evidence exists of the presence of the discontinuity. This is potentially a major pitfall of the classical approach, which cannot be corrected by forcing lateral continuity (e.g., Socco *et al.* 2008) or by other ad-hoc adjustments.

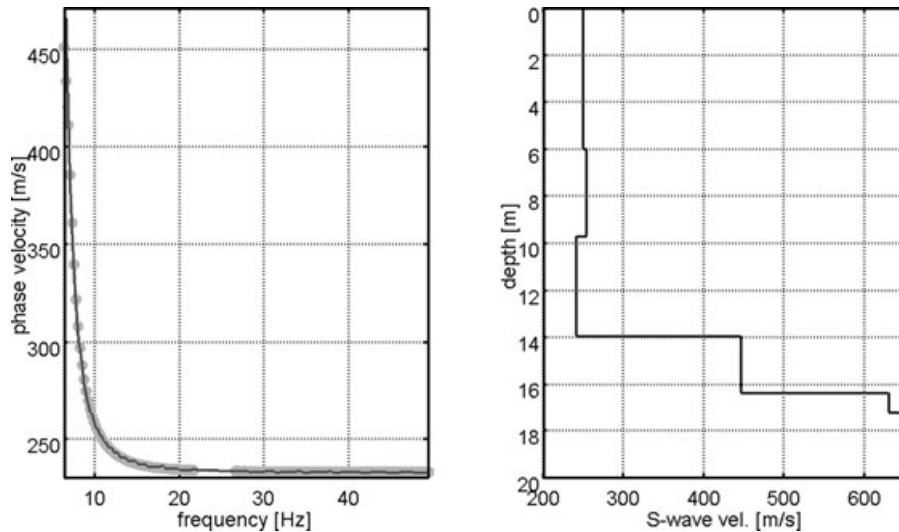
We repeated the multi-offset phase analysis to the left side of the profile only: no meaningful location of discontinuities can be identified for all frequencies (Fig. 26) except for some high-

frequency alignment at small offset that can be interpreted as a near-field effect. The  $f$ - $k$  extracted dispersion curve is very similar to the one extracted for the entire profile (Fig. 24) but differs as much as necessary to lead to a more accurate inversion of the 1D velocity profile (Fig. 27) that now is a good approximation of the true velocity profile with the high-velocity inclusion.

When the multi-offset phase analysis is performed on the right side of the profile, no meaningful knee point is identified



**Figure 28** Model C, analysis restricted to the right part (461–640 m). In the left-hand panel the locations are shown of the position of the knee points  $x_0$  minimizing equation (10) for all frequencies; the light grey crosses represent knee points characterized by a variation of slope coefficient larger than 20%. In the right-hand panel the extracted apparent curve for the same receiver subset is compared with the first three modal curves calculated for a 1D velocity profile corresponding to the 461–640 m interval of Model C (Fig. 23). Similar to Model B (Fig. 21) the curve extracted from the seismogram lies only partly on the fundamental mode: the energy peaks jump at low frequency (about 7 Hz) from the fundamental to the second mode that is nearly osculating the fundamental one. This phenomenon leads to an apparent dispersion curve that can be mistaken for a fundamental mode (Cercato 2009).

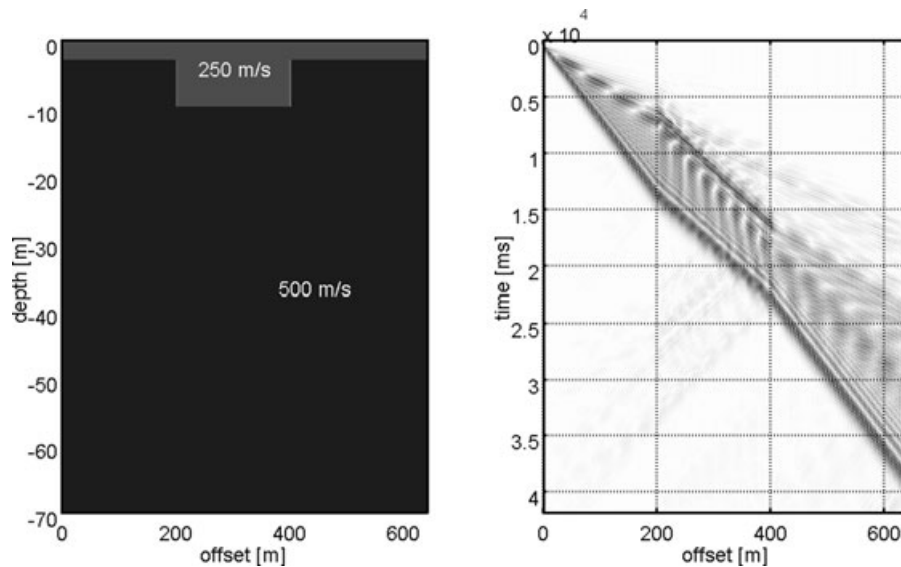


**Figure 29** Model C, analysis restricted to the right part (461–640 m). In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 1.5%. On the right-hand panel the inversion result. Note that the inverted model reproduces accurately the true velocity profile between 401–640 m.

(Fig. 28, left panel). The  $f$ - $k$  derived dispersion curve (right panel), similar to Model B, is including a jump of energy from the fundamental mode to the second mode at low frequencies. Here too the phenomenon is due to the quasi-osculating nature of the first two modes and has been observed and discussed elsewhere (Cercato 2009). Here, unlike in Model B,

the phenomenon is less pronounced and does not seriously affect the recovered velocity profile (Fig. 29) that is in fact an excellent approximation of the true velocity profile to the right of the lateral discontinuity.

In summary, Model C confirms the strength of multi-offset phase analysis in identifying lateral discontinuities, even in



**Figure 30** Model D. The left-hand panel shows the geometry of the system consisting of 3 horizontal layers. The top layer is 3 m thick with S-wave velocity equal to 250 m/s; the second layer (500 m/s) is 7 m thick and has two sharp discontinuity of S-wave velocity at 200 m and 400 m offset bounding an inclusion with velocity equal to 250 m/s; the bottom half-space has a velocity = 500 m/s. The corresponding synthetic seismogram obtained using SEM2DPACK (Ampuero 2008) is shown in the right-hand panel.

the presence of a velocity inversion due to a local inclusion. The only caveat is that in the presence of a velocity inversion, at frequencies where the energy jumps to higher modes, correspondingly the multi-offset phase analysis fails to identify knees in the phase-versus-offset plot.

For Model C also, the application of the traditional  $f$ - $k$  analysis and inversion on the entire seismogram leads to a 1D velocity profile that corresponds to the structure of the largest part of the subsurface along the acquisition line.

#### Model D

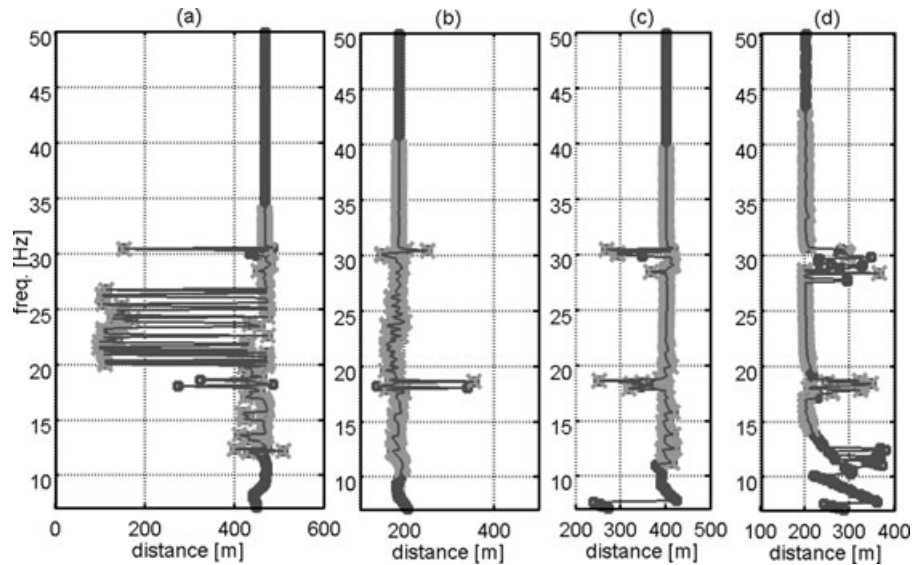
The last example is the most complex: here we investigate the possibility of detecting more than one lateral discontinuity. In fact Model D consists of a top layer (250 m/s and 3 m thick) having a thicker central part that sinks to a depth of 10 m into a half-space with higher velocity (500 m/s); this lower velocity inclusion is bounded by two sharp discontinuities at 200 m and 400 m offset (Fig. 30, left panel).

Multi-offset phase analysis applied to the entire profile detects a potential discontinuity at 470 m (Fig. 31, panel a), as a result of fitting a two-piece linear function to what should be in fact a three-piece linear trend. Then, as usual, to check consistency of this guess, we consider the receivers from the beginning to the found discontinuity at 470 m (Fig. 31, panel b); multi-offset phase analysis reveals a new possible discontinuity at 190 m. After that, in an iterative process, we perform the enhanced multi-offset phase analysis on the traces between 190–470 m offset (Fig. 31, panel c); the result is a correct identification of a discontinuity at 400 m. Then, processing traces from 190–400 m, (Fig. 31, panel d) a discontinuity at 200 m is evidenced.

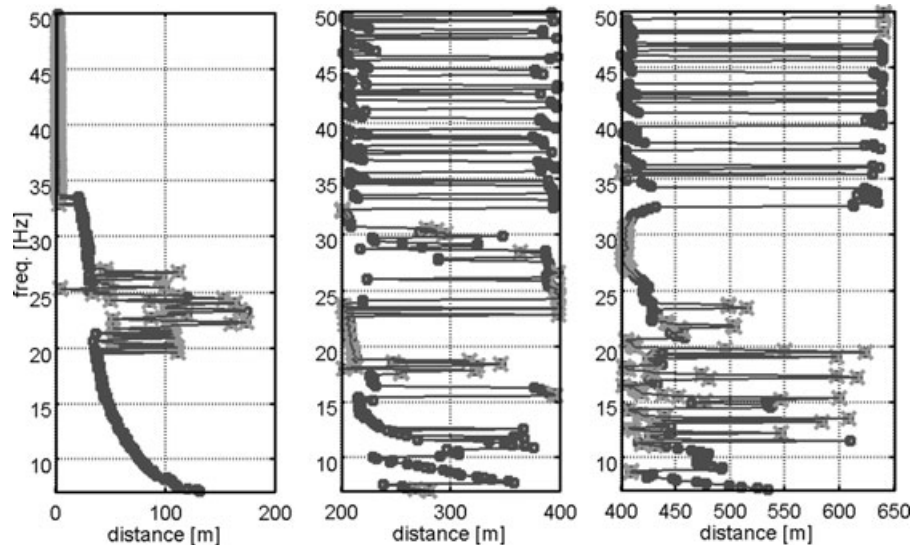
In the end (Fig. 32, left panel) we study the interval from 0–200 m and we realize that evidence at 190 m disappears (it was an effect simply due to the presence of two discontinuities in the considered data set). The same for the interval 400–640 m and the discontinuity at 470 m (Fig. 32, right panel). In conclusion, Fig. 32, left, central, right panels show the three largest data subset (0–200 m; 201–400 m; 401–640 m) with no presence of lateral discontinuities clues; for all of them the knee points indicating potential discontinuity presence are scattered without any significant correlation.

Note that the rather involved iterative procedure above is due to using a two-piece linear function at each step. We guess that fitting phase-versus-offset curves with a piecewise linear function made of more than two straight lines would speed up detection of multiple discontinuities. However, more complex piecewise linear functions would probably lead to finding artificial secondary discontinuities even when they do not exist.

Inversion of the entire seismogram (Fig. 33) leads to a misleading model with a gradual velocity increase from



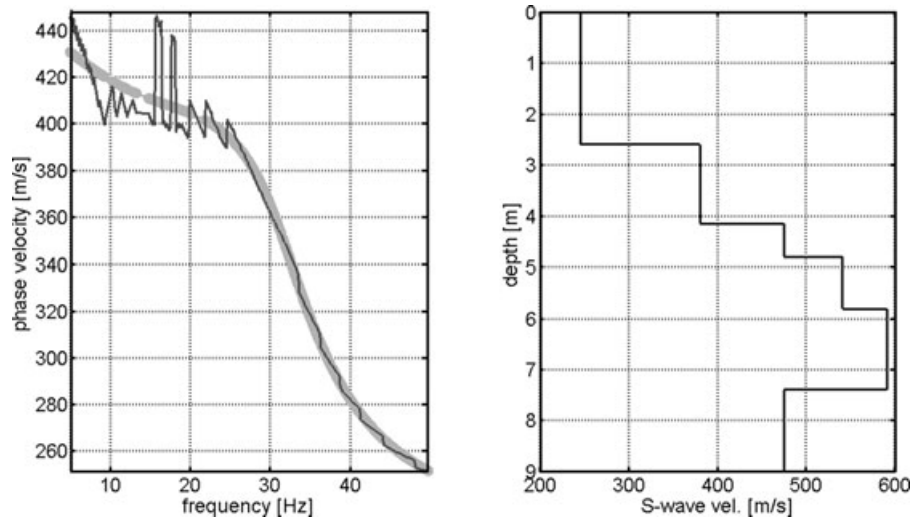
**Figure 31** Model D. The locations are shown of the knee points  $x_0$  minimizing equation (10) for all frequencies. Panels a, b, c and d show the results of multi-offset phase analysis applied to different receiver subsets. Panel a shows the result of multi-offset phase analysis of all available receivers and suggests the presence of a discontinuity at about 470 m; a further investigation considering receivers from the beginning of the profile to 470 m points out a new discontinuity at 190 m (panel b); by studying traces from 190–470 m (panel c) the discontinuity assumed at first (panel a) is now relocated at 400 m (panel c). The discontinuity guessed at 190 m (panel b) is moved to 200 m by processing data from 190–400 m. As usual, the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%.



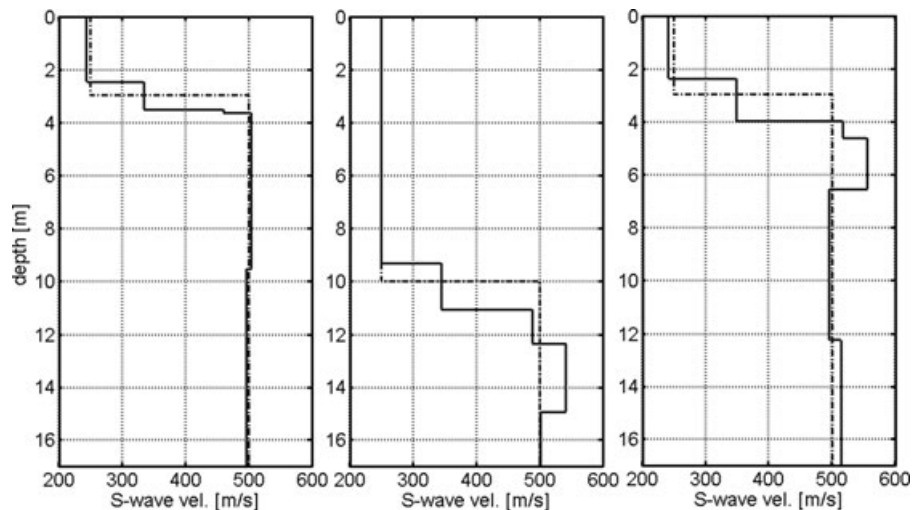
**Figure 32** Model D. The locations are shown of the knee points  $x_0$  minimizing equation (10) for all frequencies but for three different trace subsets selected as shown in Fig. 31. Left-hand, central and right-hand panels respectively show that there are not significant occurrences of knee points (hence of discontinuities) if we consider separately the profile portions going: from the beginning to 200 m, from 200–400 m and from 400 m to the end. Again the light grey crosses represent knee points characterized by a variation of angular coefficient greater than 20%.

250 m/s to 500 m/s occurring between 3–10 m of depth. The three inversion results corresponding to the three different trace subsets selected as shown in Figs 31 and 32 are plotted in Fig. 34; they reproduce very well the true velocity profiles.

Hence Model D shows the applicability and efficacy of multi-offset phase analysis even when multiple discontinuities exist. For this example the traditional  $f$ - $k$  analysis and inversion on the entire data set provide a 1D velocity profile that is not even representative of structure of the largest part of the



**Figure 33** Model D. In the left-hand panel calculated (light line) and extracted (dark line) dispersion curves are compared: the percentage cumulative data misfit is 2.3%. On the right-hand panel the inversion result. The result shows that inversion of the entire seismogram (Fig. 30, right-hand panel) leads to an averaging of the true model and the abrupt velocity jump from 250 m/s and 500 m/s (at 3 m or 10 m depth depending on the considered offset) is turned into a smooth velocity increase.



**Figure 34** Model D. The inversion results corresponding to the seismogram portions selected by multi-offset phase analysis. Left-hand, central and right-hand panels respectively correspond to portion: from 0–200 m, from 201–400 m, from 401–640 m. Note that the inverted model reproduces accurately the true velocity profile (dot-dash lines).

subsurface under the acquisition profile; indeed, in this case, the reconstructed model is rather an average of the different parts of the true model.

## CONCLUSIONS

In this paper a novel approach for detecting and locating lateral heterogeneities in horizontal stratified soil is presented. The approach is based on the multi-offset phase analysis of

surface wave data originally developed by Strobbia and Foti (2006). The advantages are that, at least in principle, the approach does not need any noise evaluation and more than one shot in order to detect the discontinuities. This is possible because our approach considers all frequencies at once. Slope discontinuities in the phase-versus-offset curves are automatically identified and, in presence of a consistent identification at multiple frequencies, indicate the location of sharp lateral S-wave discontinuities.

We tested our enhanced multi-offset phase analysis procedure on four synthetic cases, all including lateral discontinuities. The general conclusion is that the procedure works very effectively at identifying the location of the discontinuity, even in: 1) the important situation of a velocity inversion, 2) the realistic cases of relatively high level of noise and 3) presence of more than a single discontinuity. Once discontinuities are identified, a classical  $f$ - $k$  dispersion curve extraction and fundamental mode inversion is capable of reconstructing correctly the velocity profiles on every part in which the entire seismogram is subdivided: the final outcome is a correct 2D reconstruction of the discontinuous velocity profile, with no need for ad hoc hypotheses. The possible weaknesses in the inversion, e.g., in presence of osculating modes, are the same commonly found in 1D inversion and are not to be ascribed to the multi-offset phase analysis.

When the multi-offset phase analysis is not performed and the entire seismogram is  $f$ - $k$  processed and inverted, if a prevailing 1D part exists, we consistently find that the recovered 1D velocity profile corresponds to the 'true' velocity profile in the largest part of the subsurface along the acquisition line: i.e., it corresponds to the part of the seismogram carrying the most overall energy. If no prevailing 1D structure is present, the reconstructed 1D profile is an average of the different parts. This result is really disturbing as it demonstrates that the classical analysis is perfectly capable of overriding a sharp velocity contrast that shows itself even in the raw data.

In view of the above results, we strongly advocate the use of a multi-offset phase analysis prior to any attempt to use 1D Rayleigh wave inversion, in order to identify lateral heterogeneities from the data themselves. This analysis can be particularly useful in regions where strong lateral heterogeneities in the shallow low-velocity-layer are expected (Vesnaver 2004).

## ACKNOWLEDGEMENTS

We would like to thank Dr Giuseppe Pagliara from the Institut für Theoretische Physik, Heidelberg, Germany for his extremely valuable contribution to this research, for his help and fruitful comments and, no less important, for his friendship.

## REFERENCES

Abraham O., Chammas R., Cote P., Pedersen H. and Semblat J.F. 2004. Mechanical characterization of heterogeneous soils with surface waves. *Near Surface Geophysics* **2**, 249–258.

- Aki K. and Richards P.G. 2002. *Quantitative Seismology*, 2<sup>nd</sup> edn. University Science Book. ISBN 0935702962.
- Al-Hunaidi M.O. and Rainer J.H. 1995. Analysis of multi-mode signals of the SASW method. *Proceedings of the 7<sup>th</sup> International Conference on Soil Dynamics and Earthquake Engineering*, 259–266.
- Ampuero J.-P. 2008. SEM2DPACK: A spectral element method tool for 2D wave propagation and earthquake source dynamics – User's Guide Version 2.3.4. Available at <http://sourceforge.net/projects/sem2d/>
- Bergström S.G. and Linderholm S. 1946. A dynamic method for determining average elastic properties of surface soil layers. Handlarar Svenska Forskningsinstitutet för Cement och Betong vid Kungl., Stockholm, N:R7.
- Cercato M. 2009. Addressing non-uniqueness in linearized multichannel surface wave inversion. *Geophysical Prospecting* **57**, 27–47.
- Coleman T.F. and Li Y. 1996. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization* **6**, 418–445.
- Dunkin J.W. 1975. Computation of modal solution in layered, elastic media at high frequencies. *Bulletin of the Seismological Society of America* **12**, 335–358.
- Ewing W.M., Jardetzky W.S. and Press F. 1957. *Elastic Waves in Layered Media*. McGraw-Hill.
- Foti S. 2000. *Multistation methods for geotechnical characterization using surface waves*. PhD thesis, Politecnico di Torino, Italy.
- Fowler C.M.R. 1990. *The Solid Earth*. Cambridge University Press.
- Gabriels P., Snieder R. and Nolet G. 1987. *In situ* measurements of shear-wave velocity in sediments with higher-mode Rayleigh waves. *Geophysical Prospecting* **35**, 187–196.
- Heisey J.S., Stokoe II K.H. and Meyer A.H. 1982. Moduli of pavement systems from Spectral Analysis of Surface Waves. *Transport Research Record* **852**, 22–31.
- Herrmann R.B. 1987. *Computer Programs in Seismology*. St Louis University.
- Horn R.A. and Johnson C.R. 1991. *Topics in Matrix Analysis*. Cambridge University Press.
- Jones R. 1955. A vibration method for measuring the thickness of concrete road slabs *in situ*. *Magazine of Concrete Research* **20**, 97–102.
- Keilis-Borok V.I., Levshin A.L., Yanovskaya T.B., Lander A.V., Buckcin B.G., Barmin M.P. *et al.* 1989. *Seismic Surface Waves in Laterally Inhomogeneous Earth*. Kluwer Academic Publishers.
- Lin C.-P. and Chang T.-S. 2004. Multi-station analysis of surface wave dispersion. *Soil Dynamics and Earthquake Engineering* **24**, 877–886.
- Lu L. and Zhang B. 2004. The analysis of dispersion curves of Rayleigh waves in frequency-wavenumber domain. *Canadian Geotechnical Journal* **41**, 583–598.
- Luo Y., Xia J., Liu J., Xu Y. and Liu Q. 2008. Generation of a pseudo-2D shear-wave velocity section by inversion of a series of 1D dispersion curves. *Journal of Applied Geophysics* **64**, 115–124.
- McMechan G.A. and Yedlin M.J. 1981. Analysis of dispersive waves by wave field transformation. *Geophysics* **46**, 869–874.



- Miller R.D., Xia J., Park C.B. and Ivanov J.M. 1999. Multichannel analysis of surface wave to map bedrock. *The Leading Edge* 18, 1392–1396.
- 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.
- Nazarian S., Stokoe II K.H. and Hudson W.R. 1983. Use of spectral analysis of surface waves method for determination of moduli and thicknesses of pavement systems. *Transport Research Record* 930, 38–45.
- Nolet G. 1981. Linearized inversion in (teleaseismic) data. In: *The Solution of the Inverse Problem in Geophysical Interpretation* (ed. R. Cassinis). Plenum Press.
- Macon N. and Spitzbart A. 1958. Inverses of Vandermonde matrices. *The American Mathematical Monthly* 65, 95–100.
- Matthews M.C., Hope V.S. and Clayton C.R.I. 1996. The use of surface waves in the determination of ground stiffness profiles. *Proceedings of the Institution of Civil Engineers, Geotechnical Engineering* 119, 84–95.
- Mordecia A. 2003. *Nonlinear Programming: Analysis and Methods*. Dover Publishing.
- Park C.B., Miller R.D. and Xia J. 1998. Imaging dispersion curves of surface waves on multi-channel record. 68<sup>th</sup> SEG meeting, New Orleans, Louisiana, USA, Expanded Abstract.
- Park C.B., Miller R.D. and Xia J. 1999. Multichannel analysis of surface waves. *Geophysics* 64, 800–808.
- Socco L.V., Boiero D., Wisén R. and Foti S. 2006. Laterally constrained inversion of ground roll of seismic reflection records. 76<sup>th</sup> SEG meeting, New Orleans, Louisiana, USA, Expanded Abstract.
- Socco L.V., Boiero D., Comina C., Foti S. and Wisén R. 2008. Seismic characterization of an Alpine site. *Near Surface Geophysics* 6, 255–267.
- Socco L.V. and Strobba C. 2004. Surface wave methods for near-surface characterisation: A tutorial. *Near Surface Geophysics* 2, 165–185.
- Song X. and Gu H. 2007. Utilization of multimode surface wave dispersion for characterizing roadbed structure. *Journal of Applied Geophysics* 63, 59–67.
- Strobba C. 2003. *Surface wave methods: acquisition, processing and inversion*. PhD thesis, Politecnico di Torino, Italy.
- Strobba C. and Foti S. 2006. Multi-offset phase analysis of surface wave data (MOPA). *Journal of Applied Geophysics* 59, 300–313.
- Tarantola A. 1987. *Inverse Problem Theory*. Elsevier Science.
- Van Der Pol C. 1951. Dynamic testing of road constructions. *Journal of Applied Chemistry* 1, 281–290.
- Vesnaver A. 2004. The near-surface information gap for time and depth imaging. *Geophysical Prospecting* 52, 653–661.
- Wathelet M. 2005. *Array recordings of ambient vibrations: surface-wave inversion*. PhD thesis, University of Liège, Belgium.
- Xia J., Miller R.D. and Park C.B. 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics* 64, 691–700.
- Xia J., Miller R.D., Park C.B. and Tian G. 2003. Inversion of high frequency surface waves with fundamental and higher modes. *Journal of Applied Geophysics* 52, 45–57.
- Xianhai S. and Hanming G. 2007. Utilization of multimode surface wave dispersion for characterizing roadbed structure. *Journal of Applied Geophysics* 63, 59–67.
- Yoon S. 2005. *Array-based measurements of surface wave dispersion and attenuation using frequency-wavenumber analysis*. PhD thesis, Georgia Institute of Technology.
- Zhdanov M.S. 2002. *Geophysical Inverse Theory and Regularization Problems*. Elsevier Science.
- Zywicki D.J. and Rix G. J. 1999. Frequency-wavenumber analysis of passive surface waves. *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems*, 75–84.