Laterally constrained inversion of ground roll from seismic reflection records

GEOPHYSICS, VOL. 74, NO. 6 (NOVEMBER-DECEMBER 2009); P. G35–G45, 19 FIGS., 2 TABLES.
10.1190/1.3223636

Laterally constrained inversion of ground roll from seismic reflection records

Laura Valentina Socco[1], Daniele Boiero[1], Sebastiano Foti[2], and Roger Wisén[3]

ABSTRACT

Seismic reflection data contain surface waves that can be processed and interpreted to supply shear-wave velocity models along seismic reflection lines. The coverage of seismic reflection data allows the use of automated multifold processing to extract high-quality dispersion curves and experimental uncertainties in a moving spatial window. The dispersion curves are then inverted using a deterministic, laterally constrained inversion to obtain a pseudo-2D model of the shear-wave velocity. A Monte Carlo global search inversion algorithm optimizes the parameterization. When the strategy is used with synthetic and field data, consistent final models with smooth lateral variations are successfully retrieved. This method constitutes an improvement over the individual inversion of single dispersion curves.

INTRODUCTION

Seismic reflection records contain surface waves (ground roll) that can be exploited to provide near-surface S-wave velocity models by inverting dispersion curves estimated from seismic data (e.g., Mari, 1984; Yilmaz et al., 2006; Socco et al., 2008). Dispersion-curve inversion usually assumes a 1D model that contains a stack of homogeneous linear elastic layers (e.g., Socco and Strobbia, 2004), but the role of lateral variations along a profile deserves special attention. Attempts to assess lateral variations from surface waves are usually based on a moving spatial window to extract the dispersion curves, which are then inverted separately. The inversions supply independent S-wave velocity profiles, which are interpolated and presented as a 2D shear velocity section (Tian et al., 2003; Bohlen et al., 2004; Hayashi and Suzuki, 2004; Grandjean and Bitri, 2006).

In linearized surface-wave inversion, one of the most critical aspects is the sensitivity to the initial model because it can drive the inversion process into local minima that could be significantly far from the true model (Calderón-Macías and Luke, 2007). Therefore, it is important to parameterize the model correctly. In layered systems, the number of layers should be selected according to a minimum parameterization criterion. This number of layers must be large enough to describe the velocity profile properly but should not lead to overparameterization. If a priori information is not available, the initial model should be chosen according to the data quality and the information content. A global search method can be used to optimize model parameterization (Socco and Boiero, 2008).

Constraints and a priori information can be introduced into inversion to mitigate solution nonuniqueness. If several dispersion curves are available along a seismic line, a laterally constrained inversion (LCI) scheme can be adopted. LCI, first presented by Auken and Christiansen (2004) for interpreting resistivity data, is a deterministic inversion in which each 1D model is linked to its neighbors with a mutual constraint to provide a single pseudo-2D model. Lateral constraints can be considered as a priori information; the smaller the expected variation of a model parameter, the more rigid the constraint. It is also possible to use any available a priori information, e.g., from drilling, to constrain the inversion.

The LCI approach has been validated through several applications to resistivity and EM data. Wisén et al. (2005) compare the LCI of 1D resistivity soundings and a 2D smoothed inversion, showing that, in layered media, LCI has a better vertical resolution. They also introduce constraints from boreholes into LCI. A similar comparison is made by Auken et al. (2005) for synthetic and field resistivity data with lateral variations. Mansoor et al. (2006) successfully apply LCI to a frequency-domain EM data set acquired over a shallow-water wetlands site and show significant improvement in the retrieved information compared to traditional mapping. A laterally and mutually constrained inversion of continuous vertical electric soundings (CVES) and time-domain electromagnetic (TEM) data is presented by Christiansen et al. (2007). The concept of LCI applied to surface-

Manuscript received by the Editor 23 May 2008; revised manuscript received 25 March 2009; published online 6 November 2009.
[1]Politecnico di Torino, Department of Land, Environment, and Geoengineering (DITAG), Torino, Italy. E-mail: valentina.socco@polito.it; daniele.boiero@polito.it.
[2]Politecnico di Torino, Department of Structural and Geotechnical Engineering (DISTR), Torino, Italy. E-mail: sebastiano.foti@polito.it.
[3]Formerly Lund University, Lund, Sweden; presently Rambøll Danmark A/S, Virum, Denmark. E-mail: RGW@ramboll.dk.

wave data using a walk-away data set was originally presented by Wisén and Christiansen (2005). Here, we apply LCI as the final step of analyzing surface waves in seismic reflection surveys.

We present an alternative approach to retrieve a pseudo-2D S-wave velocity model from land seismic reflection data. The key elements are a moving spatial window along the seismic line to retrieve a number of dispersion curves with experimental uncertainties; a Monte Carlo inversion (MCI) to obtain a rigorous model parameterization; and, finally, the inversion of a series of dispersion curves along a seismic line, using linearized, laterally constrained inversion. In the following, the processing and inversion methods are described and applied to a synthetic data set and to field data from Italian alpine valleys.

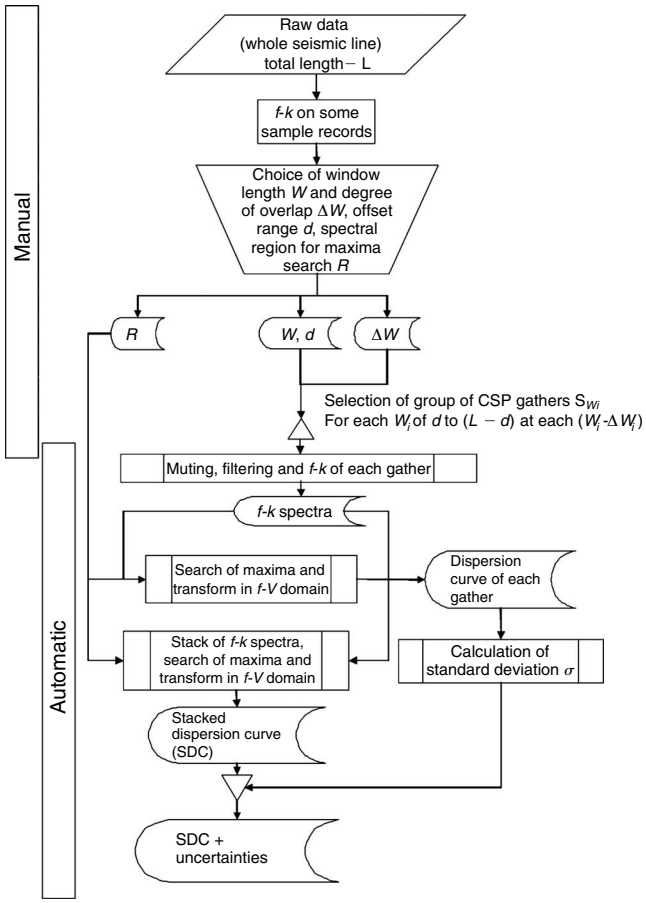

Figure 1. Flowchart of the processing procedure; W = processing window length, ΔW = overlap between consecutive positions of the processing window, d = offset range for CSP gather selection, R = spectral region for dispersion curve estimation, DC = dispersion curve for the single CSP gather, SDC = stacked dispersion curve estimated on the stacked f-k spectra of all the gathers selected for each window position, σ_e = standard deviation of the dispersion curve.

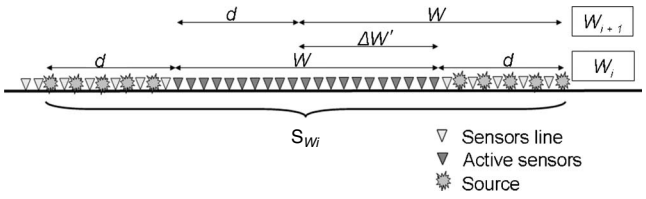

Figure 2. Gather selection. For a defined position of the moving window W_i, the sensors within the length W and the shots within the offset range d are selected to obtain S_{Wi}. Then, W is shifted by ΔW to the position W_{i+1} and the procedure is repeated.

METHOD

Preliminary data assessment

Acquisition of reflection data is not optimized for surface-wave data analysis, so we first must assess whether the data fulfill the quality requirements for surface-wave analysis. The data must have a good signal-to-noise ratio (S/N) in a wide-frequency band and sufficient spectral resolution to allow modal separation (Socco and Strobbia, 2004). The choice of the source type, sensor frequency, and sampling in time and space may not be adequate in traditional seismic reflection data.

Seismic reflection sources often are powerful enough to supply very high S/N surface-wave signals at far offsets. The time sampling and trace length are usually adequate and should contain only the traces in which the surface waves are not truncated. The offset often is sufficient to guarantee high wavenumber resolution and, hence, good modal separation; on the other hand, the spatial sampling often is too coarse to retrieve dispersion curves without spatial aliasing.

In seismic reflection surveys, low-frequency surface waves are considered as coherent noise, so they can be filtered using high-frequency sensors, sensor groups, source arrays, and/or analog low-cut filters during acquisition. The ideal recording for surface-wave analysis would be single, low-frequency sensors (no sensor groups) with no filter. If these requirements are not obtained, the data should be evaluated carefully before using them for surface-wave analysis. If the acquisition parameters are favorable, signal processing can be performed to retrieve dispersion curves (Dziewonski and Hales, 1972; Nolet and Panza, 1976; McMechan and Yedlin, 1981; Park et al., 1999).

Processing

One of the most attractive aspects of surface waves in seismic reflection records is the large amount of data that can be considered a resource but that also requires an automatic processing approach to handle the full data set efficiently. This approach also allows uncertainties of the experimental dispersion curve to be estimated. The procedure we propose, summarized in the flowchart in Figure 1 and explained in Figure 2, is based on frequency-wavenumber (f-k) analysis. The first (manual) part of the procedure concerns tests on representative sample records to select the processing parameters: optional muting and filtering, the length of the moving window W, the overlap of the spatial windows ΔW, the maximum offset range d for shot selection, and the spectral region R for dispersion curve search.

The criterion used to choose the optimum processing window W is based on the fact that widening W increases the spectral resolution but also introduces noise because far-offset traces with low S/N are included; a compromise must be reached. The choice of the maximum offset range d is based on the S/N. To optimize the search for the dispersion curves and to avoid mixing different modes, it is useful to limit the spectral region R in which to search for the spectral maxima. Using the chosen W, we find the optimum R on some representative records and apply it to the whole data set.

The second section of the processing procedure (Figure 1) is fully automatic and is driven by the parameters defined in the first section. A group of common-shot-point (CSP) gathers (S_{Wi}) is selected for each *i*th position of *W* along the seismic line, considering the receivers that fall in W_i and the shots in the nearest *d* interval (Figure 2). Then *W* is moved by an increment of (*W*-Δ*W*) to the position W_{i+1}, and the procedure is repeated for each window along the survey line (Figure 2).

The *f*-*k* spectra are computed for each group of gathers S_{Wi}, and the spectral maxima inside *R* are found. The coordinates of the maxima are stored and transformed to the surface-wave phase velocity versus frequency curve of each gather by applying the relation $V = 2\pi f/k$. This leads to the extraction of *n* dispersion curves for each W_i, where *n* is the number of shots that fall into *d*. These dispersion curves are then used to estimate the experimental uncertainty σ_e as the standard deviation of the velocity values at each frequency. All amplitude spectra related to S_{Wi} are summed, and the spectral maxima of the sum are found inside *R* to obtain a smooth dispersion curve with a higher S/N (Grandjean and Bitri, 2006; Neducza, 2007); we call the latter the stacked dispersion curve (SDC). The SDC and its uncertainties are then associated with the spatial coordinate of the center of W_i

The procedure is repeated for each W_i; the result is several dispersion curves (with experimental uncertainties) that are regularly spaced with (*W*-Δ*W*) steps along the seismic line.

Inversion

The inversion has two steps. First, we apply a Monte Carlo inversion (MCI) to obtain a rigorous model parameterization. Then we use the LCI to invert the entire data set (i.e., all stacked dispersion curves with uncertainties and the lateral constraints) as one system, starting from the parameterization obtained from the MCI. LCI and MCI use 1D forward modeling based on the Haskell and Thomson approach (Thomson, 1950; Haskell, 1953). The model is a layered linear elastic medium defined by the velocities, densities, and boundary depths of each layer. The unknown model parameters in the inversion are the layer S-wave velocities and thicknesses. Poisson's ratio and density are assumed a priori according to the expected materials and water-table level.

Monte Carlo inversion (MCI)

We use an MCI algorithm (Socco and Boiero, 2008), which is efficient because of the use of the scale properties of the dispersion curves. These properties are linked to the scaling of the modal solution with the wavelength. If the model parameters are scaled, the modal dispersion curve scales correspondingly; phase velocities and frequencies scale if all the layer velocities are scaled, but only the frequencies scale if all the layer thicknesses are scaled (Socco and Strobbia, 2004).

We generate a set of random models after defining the number of layers and the upper and lower starting values for each model parameter (shear-wave velocity and thickness of each layer). We then use this set of models to calculate the associated synthetic dispersion curves. Each of the obtained synthetic dispersion curves is shifted as close as possible to the experimental one by equalizing the curve barycenters (the barycenter coordinates are computed as the average values of the phase velocity and frequency, respectively). This operation supplies the scale factors by which the models are updated relative to the shifted curves. Thus, the global distance between the experimental and theoretical curves is artificially reduced, and each randomly generated model moves closer to the true model (Socco and Boiero, 2008).

Next, the obtained scaled models are used in the inversion by calculating the misfit between the experimental dispersion curve and the shifted ones and selecting the final models according to a statistical test. This procedure concentrates the sampling in the high-probability density zones (the low-misfit regions) of the model parameter space and allows significant optimization of the process, reducing the number of needed simulations (usually 10^4–10^5).

Assuming the experimental dispersion curve has a Gaussian error (Lai et al., 2005) with a known standard deviation, we can use a misfit function with a chi-square χ^2 structure:

$$S = \chi^2 = \frac{\sum_{i=0}^{l} (V_{ti} - V_{ei})^2 \sigma_{ei}^{-2}}{\ell - (2n-1)}, \quad (1)$$

where V_{ti} and V_{ei} are the calculated and experimental phase-velocity vector elements, respectively; σ_{ei} are the data-uncertainty vector elements; ℓ is the number of data points in the dispersion curve; and *n* is the number of layers in the model. In this way, the denominator of equation 1 is the number of degrees of freedom of the problem, which is the difference between the amount of experimental information (the number of data points) and the number of unknowns.

We then apply a one-tailed test, considering that the ratio between the χ^2 variables follows a Fisher distribution (Ostle, 1963). This procedure selects the set of acceptable velocity models that represents the final result of the inversion for a chosen level of confidence α. Thus, the data quality and the dimensionality of the problem are accounted for. The experimental uncertainties are included in χ^2, and the statistical test is ruled by the number of degrees of freedom of the Fisher distribution. The lower the level of confidence α, the wider the set of selected profiles. This stochastic approach improves our knowledge about the high probability density regions of the model parameter space (the V_S and thickness of each layer) and allows the nonuniqueness of the solution to be analyzed and the general pattern of the expected result to be identified.

The MCI uses overparameterized models as a preliminary step. The task in this step is not to estimate the model parameters but to identify a model trend to be used as a guide to parameterize the initial model for the LCI correctly. Therefore, the fact that the model parameters are poorly resolved is not critical at this stage. The initial model is chosen by the operator according to the retrieved model trend.

Laterally constrained inversion (LCI)

LCI was developed (Auken and Christiansen, 2004) to invert sounding data along a profile, using a pseudo-2D layered parameterization; it is restricted to quasilayered geological environments.

The inversion result contains a set of 1D models in which each separate model corresponds to a dispersion curve. All dispersion curves are inverted simultaneously by minimizing a common objective function, which contains the data, the a priori information, and the constraints (Auken and Christiansen, 2004). The lateral and a

priori constraints are scaled according to the model separation and are thereby weakened with increasing distance between neighboring soundings.

The lateral constraint is defined by the variance of the difference allowed for the same parameter between neighboring models: the greater the variance, the weaker the constraints. The constraints, the a priori information, and the dispersion data are all inversion input. Consequently, the output models form a balance between the constraints, the physics, and the data. Model parameters with little influence on the data are controlled by the constraints. Information from one model spreads to the neighboring models through the lateral constraints. The result is a smoothly varying pseudo-2D model.

Because LCI solves an overdetermined problem (more data than model parameters), a sensitivity analysis of the estimated model parameters can be performed. The parameter sensitivity analysis of the final model is the linearized approximation of the covariance of the estimation error $\mathbf{C}_{\text{est}}$ (Tarantola and Valette, 1982):

$$\mathbf{C}_{\text{est}} = (\mathbf{G}'^{T}\mathbf{C}'\mathbf{G}')^{-1}, \tag{2}$$

where $\mathbf{G}'$ contains the Jacobian, the a priori information, and the regularization, and where $\mathbf{C}'$ contains the experimental uncertainties, the uncertainties on the a priori information, and the constraints (Auken and Christiansen, 2004).

The standard deviations of the model parameters are calculated as the square root of the diagonal elements in $\mathbf{C}_{\text{est}}$. The model parameters are represented as logarithms, so the analysis gives a standard deviation factor (STDF) on the parameter m_s, defined as

$$\text{STDF}(m_s) = \exp(\sqrt{C_{\text{est},ss}}), \tag{3}$$

The theoretical case of a perfect model parameter resolution has an STDF of one. We define well-resolved parameters as having an STDF less than 1.2, which is approximately equivalent to an error of 20%; moderately resolved parameters fall within $1.2 < \text{STDF} < 1.5$, poorly resolved parameters have $1.5 < \text{STDF} < 2$, and unresolved parameters have $\text{STDF} > 2$ (Auken and Christiansen, 2004).

The normalized residual error describes the success of the inversion of each sounding. This error is normalized to the experimental uncertainty σ_e of the corresponding dispersion curve so that a value of one for the normalized residual equals the experimental uncertainty and a value of less than one corresponds to a residual error smaller than the experimental uncertainty. The normalized residual per point for a specific sounding is defined as

$$\text{NR} = \sqrt{\frac{\sum_{i=1}^{l}\left(\frac{V_{ei} - V_{ti}}{\sigma_{ei}}\right)^2}{\ell}}, \tag{4}$$

where V_{ei} and V_{ti} are the phase velocity vector elements of the experimental and theoretical dispersion curves, respectively; σ_{ei} are the respective experimental uncertainties; and ℓ is the number of data points in the dispersion curve. This differs from the objective function of the inversion process, which is defined for the whole data set and includes the constraints and any a priori information.

The LCI results differ from those obtained from a simple smoothing of the 1D soundings inverted separately. In LCI, we regularize the solution by minimizing a common objective function for the whole data set. The a posteriori regularization obtained from the interpolation of independent models does not account for the misfit between the experimental and calculated dispersion curves; hence, it does not guarantee compliance with the data.

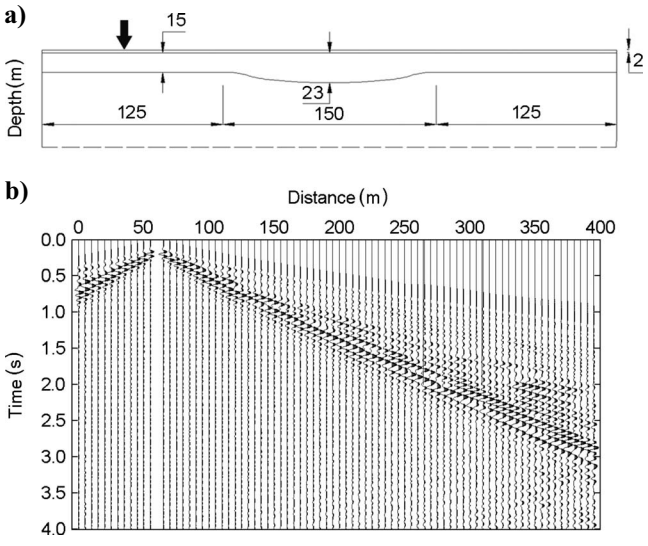

Figure 3. (a) Synthetic model geometry. (b) Synthetic seismogram for the shot position at 60 m—black arrow in (a).

SYNTHETIC EXAMPLE

The synthetic model shown in Figure 3a and described in Table 1 is a linear elastic system with a lateral variation represented by a 150-m-wide and 8-m-deep valley in the second interface. We obtained the synthetic data set using a finite-element-method code created by Comsol Multiphysics (http://www.comsol.com) with an axially symmetric approach. The source was a Ricker wavelet with a peak frequency of 10 Hz, discretized with a 1-ms time interval and applied as a vertical force to the free surface. The full waveform in the vertical section was simulated. Vertical particle velocity time histories were saved every 5 m along the free surface (Figure 3b) for each of 21 simulated shots, with a spacing of 20 m, to obtain a synthetic seismic reflection data set.

Dispersion curves were extracted as described in the processing section for several processing window lengths W; the effect of different W on the final result of the inversion was assessed. We finally chose the optimum W of 24 channels and an overlap ΔW of 75% of W ($\Delta W = 18$ channels). This allowed 12 evenly spaced dispersion curves to be retrieved along the model. Each dispersion curve was retrieved from the coordinates of the spectral maxima of the spectrum obtained by stacking 10 f-k gathers (S_{Wi}); the uncertainties were also calculated.

Table 1. Synthetic model parameters.

Layer	V_P (m/s)	V_S (m/s)	ρ (kg/m^3)
1	180	90	1800
2	240	180	2100
3	350	200	2400

The MCI applied to one of the retrieved SDC supplied the final model set shown in Figure 4. On the basis of the MCI results, we selected a three-layer initial model for the LCI (Figure 4). The LCI inversion result obtained with a medium constraint is shown and compared with the true model in Figure 5. The lateral variation in the model is well retrieved using the LCI (Figure 5). A slight smoothing effect on the layer thickness can be seen.

Figures 6 and 7 compare the true model and the inversion results for each parameter (V_S, thickness, and depth) in terms of percent error with respect to the true value, according to

$$e = \frac{m_{\text{est}} - m_{\text{true}}}{m_{\text{true}}} 100, \tag{5}$$

where m_{true} and m_{est} are the true value and the value estimated by the inversion, respectively, for the model parameter m. Figure 6a shows the errors relative to the true model for individual (unconstrained) inversions and for the LCI with different levels of constraint. The values of the lateral constraints are 100, 1, and 0.1 (units are meters/second for the velocity and meters for the thickness) for weak, medium, and strong constraints, respectively.

The results of the constrained inversion are almost identical for the weak and medium constraints and show improvement over the unconstrained ones. The velocities of the first two layers and the thickness of the first layer are retrieved very accurately (with errors $<3\%$) by the individual inversions and the LCIs. The velocity of the half-space is underestimated slightly (with errors $<10\%$) by the unconstrained inversions and by the LCI with weak and medium constraints, yet it presents an underestimation of 20% for the strongly constrained LCI. The most significant parameter, in terms of correctly retrieving the lateral variations, is the thickness of the second layer. In the LCI results obtained with weak and medium constraints (Figures 5 and 6a), the thickness is underestimated where lateral variation occurs (with errors $<20\%$) but is well retrieved in the 1D portion of the model (with errors $\sim 5\%$). The individual inversions and the strongly constrained LCI produce errors ranging from $+10\%$ to -30%.

The LCI results obtained with weak and medium constraints are improved compared to the individual inversions. The results obtained applying strong constraints (Figure 6a) show that if the constraints are not consistent with the geologic variability of the site, an oversmoothed final model is obtained. This can also be recognized by analyzing the normalized residuals of the last iteration (Figure 6b). Although the normalized residuals remain unchanged for the individual, weak, and medium constrained inversions, they increase for the strongly constrained one.

Finally, we show the effect of different processing window widths W in Figure 7. The errors, with respect to the true model, are plotted for the inversion performed adopting medium constraints for W equal to 12, 24, and 36 channels, respectively. The narrowest window (Figure 7, top) supplies relatively good results for the uppermost portion of the model but produces larger errors for the deeper part of the model and for the zone related to the lateral variation with respect to the optimum window (Figure 7, middle). The longest window (Figure 7, bottom) accurately estimates the velocity of the deeper layers (because it allows a better estimation of the low-frequency band of the dispersion curves) but produces greater errors for the uppermost part of the model. Moreover, the choice of a wider window reduces the horizontal extent of the final model (Figure 7).

FIELD DATA EXAMPLES

The first example (from La Salle) contains a very good data set from which pseudo-2D sections of shear properties are obtained down to 60 m. The other two examples (from Torre Pellice) are in an urban environment. They show the improvement that can be obtained with LCI compared to individual inversions for noisy data sets. They also show that constraints from other geophysical tests can be introduced into the LCI process to obtain more consistent models. From previous knowledge about the geologic environments, we do not expect significant lateral variations along the seismic lines for either site.

We extracted the experimental dispersion curves using the procedure described in the processing section for all of the examples, and we show how the choice of different processing parameters (W,ΔW,d) can lead to different results (see Table 2).

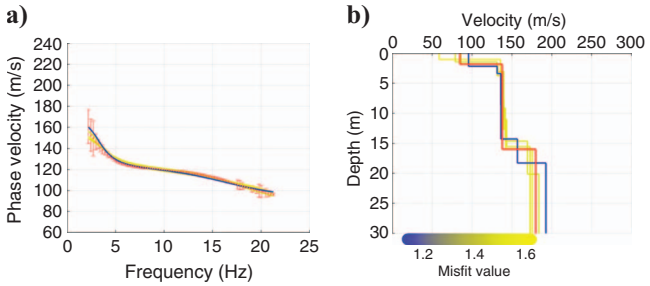

Figure 4. Monte Carlo inversion of the synthetic data. (a) Dispersion curve (SDC) with uncertainties extracted from the synthetic data (red dots and red bars) with the theoretical curves relative to the MCI final models. (b) The MCI results plotted with a color scale that represents the misfit together with the chosen initial model for the LCI (red).

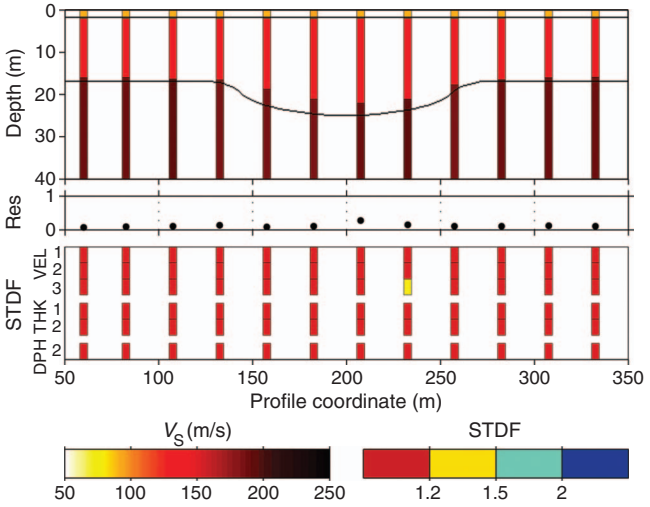

Figure 5. LCI of the synthetic data compared with the boundary position of the true model. Starting from the top, V_S profiles for each dispersion curve along the seismic line (the V_S color scale is depicted at the bottom left); normalized residuals for each V_S model; and standard deviation factor for each model parameter (from top to bottom V_S, thickness, and depth). The STDF color scale is depicted at the bottom right; red is well resolved, and blue is poorly resolved.

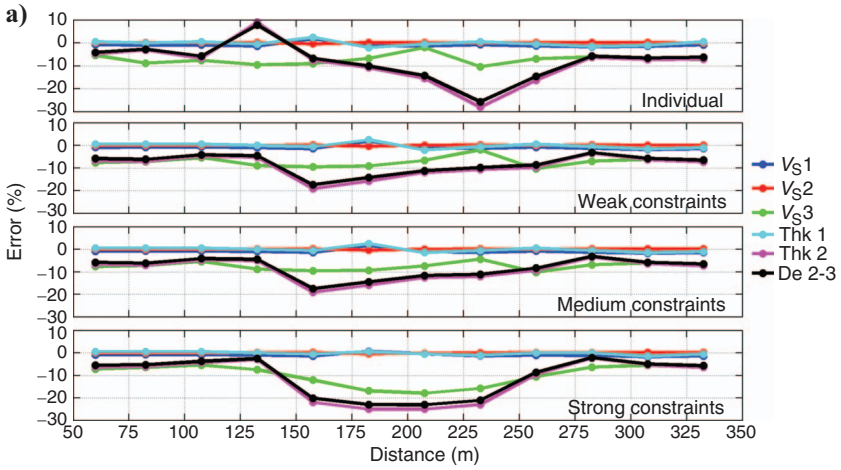

Figure 6. Errors with respect to the true model for each model parameter: (a) from the top: individual inversion, LCI with weak constraints, LCI with medium constraints, and LCI with strong constraints; (b) normalized residuals for the results shown above.

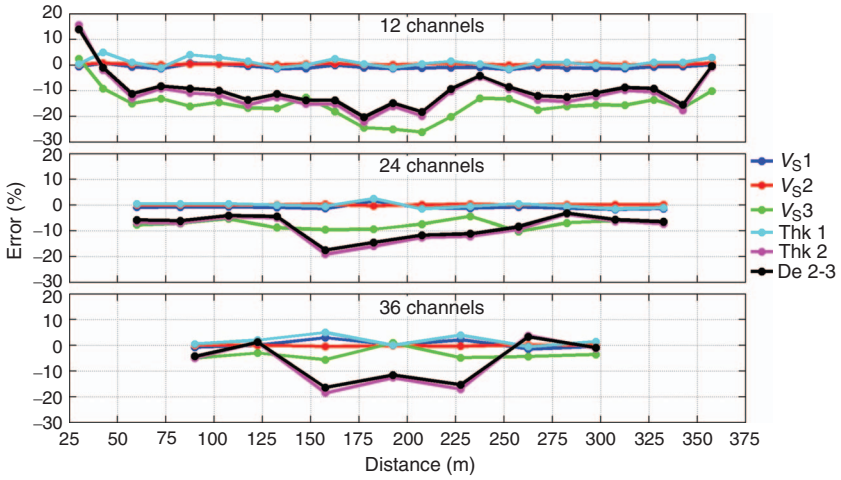

Figure 7. Effect of the processing window length W on the synthetic data inversion results for 12, 24 (optimum window), and 36 channels.

Field example 1: La Salle site

This example used data acquired over a wide alluvial fan located in La Salle, Aosta Valley (Figure 8), Italy. The maximum thickness of the Quaternary deposits is 200 m; the fan is composed mainly of alluvial deposits (sands and gravels), polygenic slivers, pebbles, and blocks. The deposit initially was investigated through two boreholes drilled down to a depth of 50 m, which were later used for the downhole tests (DH1 and DH2 in Figure 8). The stratigraphy shows the typical chaotic sequences of gravelly soils in alpine alluvial fans, with no marked layering down to a depth of about 50 m.

The survey performed at the La Salle site is described in detail in Socco et al. (2008). Here we present two seismic reflection lines (L1 and L2 in Figure 8). Each profile is about 1000 m long, acquired with 240 active channels with 10-Hz vertical geophones, 2-m geophone spacing, 6-m shot spacing, 1-ms sampling rate, 2-s recording time, antialias filter, and no low-cut filters.

The quality of the ground roll data was very good. Preliminary tests were performed to select the optimum processing window length (Figure 9). The shorter windows (24 and 36 channels in Figure 9a and b, respectively) do not allow the low-frequency band of the dispersion curve to be retrieved. The longer windows (48 and 60 channels in Figure 9c and d, respectively) supply very similar results, showing that an extension of the window length beyond 48 channels does not significantly increase the information in the dispersion curve. We therefore chose a window length equal to 48 channels to retrieve a closely spaced series of dispersion curves along each line. The other processing parameters are in Table 2. An example of processing is shown in Figure 10.

The SDCs of the two lines are in Figure 11. The frequency band of the dispersion curves spans from 3–4 Hz to more than 70 Hz, which corresponds to wavelengths from 3 m to more than 300 m, allowing both large investigation depths and great detail close to the ground surface. We ran the preliminary MCI with a 10-layer model. The result (Figure 12) shows a sharp interface at a

Table 2. Processing parameters for the field data sets.

Data set	W (m)	W (number of channels)	ΔW (%)	ΔW (number of channels)	d (m)	Number of shots	Number of SDC
La Salle L1	94	48	80	38	42	14	58
La Salle L2	94	48	80	38	42	14	38
Torre Pellice L1	46	24	50	12	60	20	4
Torre Pellice L2	46	24	75	18	6	2	20

depth of about 5 m and then a smoothly increasing velocity with depth. This trend suggests the adoption of a vertically smoothed approach for the LCI. In this case, we chose a model with a large number of thin layers. The layer thicknesses are fixed; only the velocity is estimated.

In the LCI results along L1 and L2 (Figure 13), each slice of the section is a 1D velocity model located at the same position of the respective dispersion curve. The final result is a pseudo-2D section of V_S. The normalized residuals of each sounding are also printed below each of the velocity profiles.

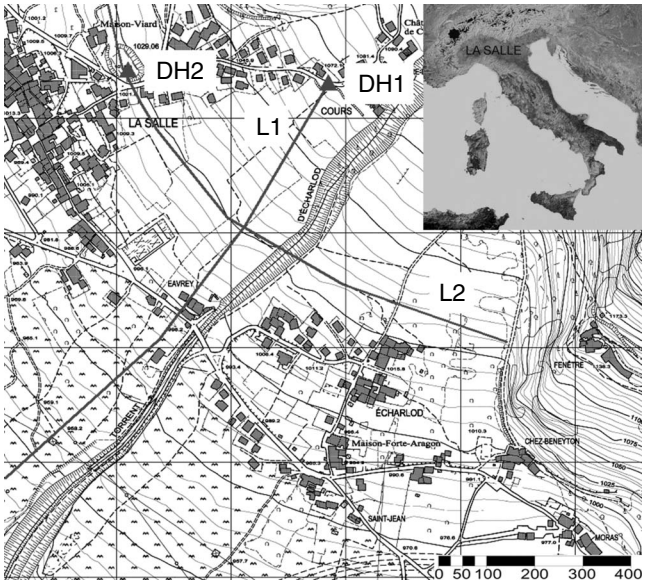

Figure 8. Survey locations at the La Salle site. L1 and L2 are the seismic reflection lines analyzed, and DH1 and DH2 are downhole test locations.

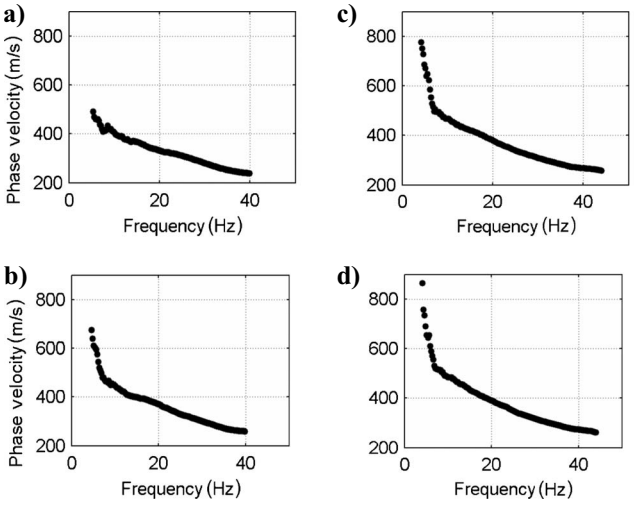

Figure 9. An example of the effect of the processing window length W on field data (La Salle site, line L1, channel spacing 2 m). Different processing windows supply different dispersion curves: (a) 24 channels, (b) 36 channels, (c) 48 channels, and (d) 60 channels.

Field example 2: Torre Pellice site

This data set was acquired in a populated alpine valley in Italy for a seismic response study. The site geology is characterized by shallow fluvial sediments with an expected thickness of 10–50 m, overlying lacustrine sediments. The bedrock depth is expected to be more than 100 m in the central part of the valley. High-resolution reflection surveys were performed along two lines across the valley (L1 and L2 in Figure 14), with the main task of identifying the bedrock position. Local stratigraphic information down to 30 m close to L1 and down to 50 m close to L2 was available from the boreholes drilled for the downhole tests (DH1 and DH2 in Figure 14).

Each seismic line is about 800 m long, and the acquisition parameters are the same as those of the La Salle site. Significant ground roll is present in the seismic records, and we have analyzed it to retrieve information about the V_S distribution in the overburden. The processing parameters used for surface-wave analysis are summarized in Table 2. Thanks to the dense spatial sampling and long offsets, the data provide a good spectral resolution and recovery of dispersion curves over a wide frequency band, but they also have a variable S/N, caused by human activities.

The first example is a subsection of L1 (the shaded area in Figure 14). Because of the low S/N, many shots are stacked for each processing window to improve the quality of the extracted dispersion curves (SDCs). The whole data set is processed in just four processing windows (Table 2). The experimental SDCs are shown in Figure 15. The curves present quite smooth and regular patterns but high experimental uncertainties, particularly at low frequencies.

The MCI clearly suggests a layered model (Figure 16), with a sharp shallow interface overlying a quite stiff layer. A velocity decrease is seen at about 10-m depth, and an increase is seen at ~20-m depth. The obtained V_S profiles do not identify a clear or consistent trend below this interface because of the lack of information; therefore, a single layer is used below 20 m in the initial model for the LCI.

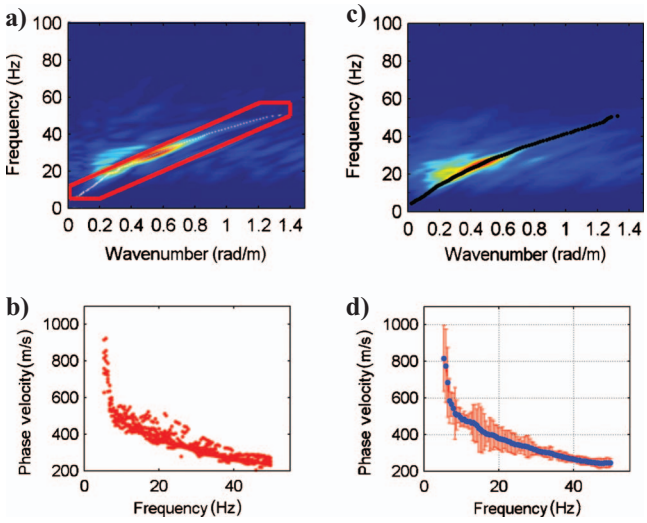

Figure 10. Example processing results from the La Salle site for a given W_i. (a) Example f-k spectrum for a single gather from group S_{Wi}. The red line delimits the region R; the white dots are the searched for maxima. (b) Dispersion curves obtained for the whole group of CSP, S_{Wi}. (c) Stacked spectrum. The black dots are the maxima corresponding to the stacked dispersion curve, SDC. (d) SDC (blue dots) and uncertainties (red bars) retrieved from the group of dispersion curves.

The comparison between the inversions of the four stacked dispersion curves and the corresponding LCIs is presented in Figure 17. The velocity decrease in the initial model is confirmed in three of the four inverted profiles in the individual and constrained inversions and is confirmed by the DH1. The V_S profile that does not show the velocity decrease is relative to the dispersion curve at the top of Figure 15, which contains higher uncertainties, particularly in the 7–10-Hz frequency band. The result of the LCI is smoother and more regular than the individual inversions, and the STDF of the model parameters is significantly reduced, indicating the resolution of the model parameters has improved.

The second example is for a subsection of line L2 (the shaded area in Figure 14). An example of raw data is presented in Figure 18a. Unlike the L1 line example, here we selected several W_i (Table 2) with the aim of obtaining a pseudo-2D section of V_S. We used an S_{Wi} made of just two gathers for each W_i. This approach simulates the classical surface-wave data acquisition made by single-sensor land cables (Yilmaz et al., 2006). The S/N of the data set is not very good, and the frequency band of the dispersion curves is quite variable along the line (Figure 18b). On the basis of visual assessment, we discarded the experimental dispersion curves that were characterized by a very poor S/N or a very narrow frequency band.

After running a preliminary MCI, we selected a five-layer initial model for the LCI. The results are presented in Figure 19; the individual inversions are presented on the left and are compared with the LCI on the right. We also introduced a priori information from DH2 in the vicinity of L2 as a fixed velocity model that influences the surface-wave models through lateral constraints. The DH2 is located away from the seismic line, so its influence is reduced because the lateral constraints are scaled with distance. The LCI gives a smoother, more consistent model than the individual inversions. A velocity decrease was retrieved, particularly in the left part of the section. The normalized residual values at the last iteration are shown below the velocity models.

The STDF for each model parameter shows that the individual inversions supply reliable results only for the uppermost layer. The LCI shows a low STDF, implying very good resolution of the model parameters for the first three layers and acceptable values for the other layers, except for the thickness of the fourth layer, which appears poorly resolved in all of the velocity profiles.

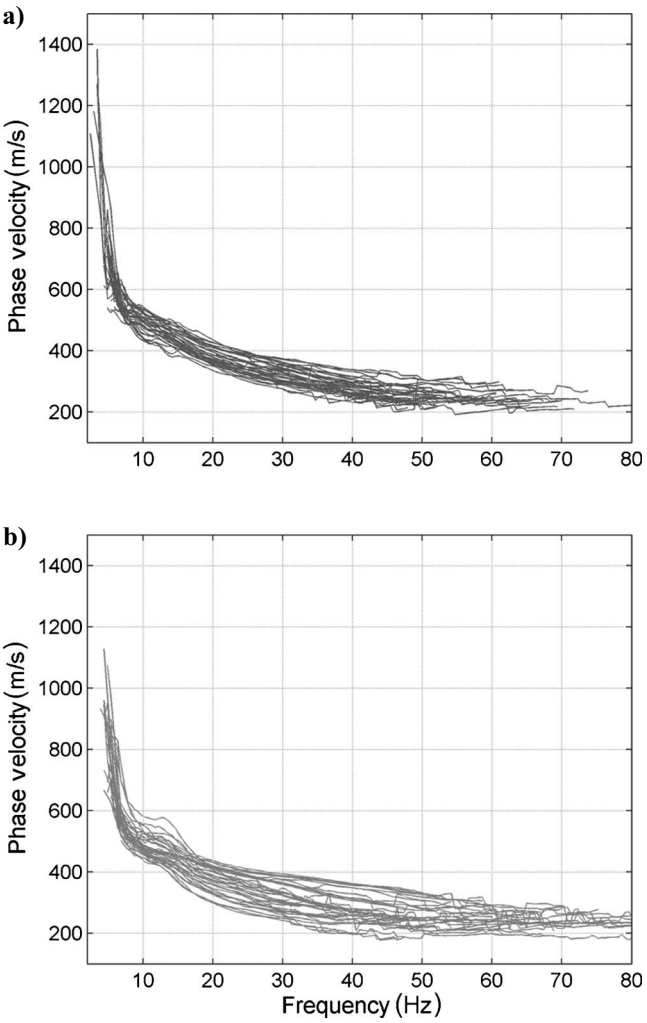

Figure 11. La Salle site, experimental dispersion curves (SDCs): (a) L1, (b) L2. The curves are superimposed, but each curve corresponds to a different spatial position along the line.

DISCUSSION

We have presented a processing and inversion strategy to analyze surface waves present in seismic land reflection records. We have tested several data sets to show the influence of the choices of processing parameters. The optimal width of the moving window for the *f-k* search of the dispersion curves is data dependent and varies according to the S/N, the frequency bandwidth, and the effect of lateral variations. The La Salle real data example shows how the quality of the retrieved dispersion curve is influenced by the processing window width (Figure 9). The frequency band, in which the dispersion curve can be successfully retrieved, increases with the width of the processing window. A trade-off occurs between dispersion curve bandwidth (related to window length) and lateral resolution. We have shown this on synthetic data, comparing the results of the LCI of the dispersion curves obtained with different processing windows. The optimum window also produces better results in terms of the final model identification (Figure 7).

If lateral variations are expected, the width of the processing window should be reduced to be comparable to the scale of the lateral variations. The presence of lateral variations within the processing

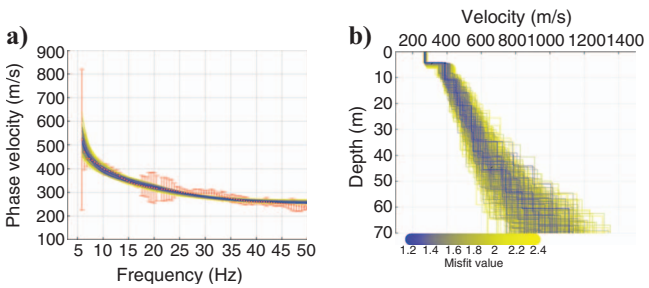

Figure 12. Monte Carlo inversion for the choice of the LCI initial models. (a) Experimental dispersion curve (SDC) with uncertainties for the La Salle site (red dots and red bars) together with the theoretical curves relative to the MCI final models. (b) The MCI results plotted with a color scale that represents the misfit.

window can increase the experimental uncertainties that affect the inversion result.

The overlap between successive positions of the moving window influences the smoothness of the results. If the dispersion-curve quality is privileged with respect to quantity (as for the first field data example at the Torre Pellice site), a limited number of smooth dispersion curves over a wide frequency band can be obtained. If the dispersion-curve quantity is privileged with respect to quality (as for the second field data example at the Torre Pellice site), more single-fold dispersion curves can be retrieved. The latter choice can lead to a more heterogeneous quality of the dispersion curves along the seismic line (Figures 15 and 18b).

The processing and the constraints applied in the inversion produce a smoothing effect on the final result. It is not straightforward to separate or quantify the smoothing associated with the processing and with the constraints. The *f-k* transform applied to retrieve the dispersion curves averages the information contained in the data and therefore introduces smoothing. The choice of the processing window is driven by achieving a good-quality dispersion curve for the specific data set rather than by mitigating this smoothing effect. Moreover, the choice of processing parameters (window width and overlap between the neighboring windows) influences the strength of the constraints applied in the inversion because the constraints are scaled with the distance between the neighboring soundings (which comes from the window width and the overlap); the larger the window and the smaller the overlap, the larger the distance and the weaker the constraints. There is no general rule for the choice of these parameters, and they should be consistent to the data quality and the knowledge of the site.

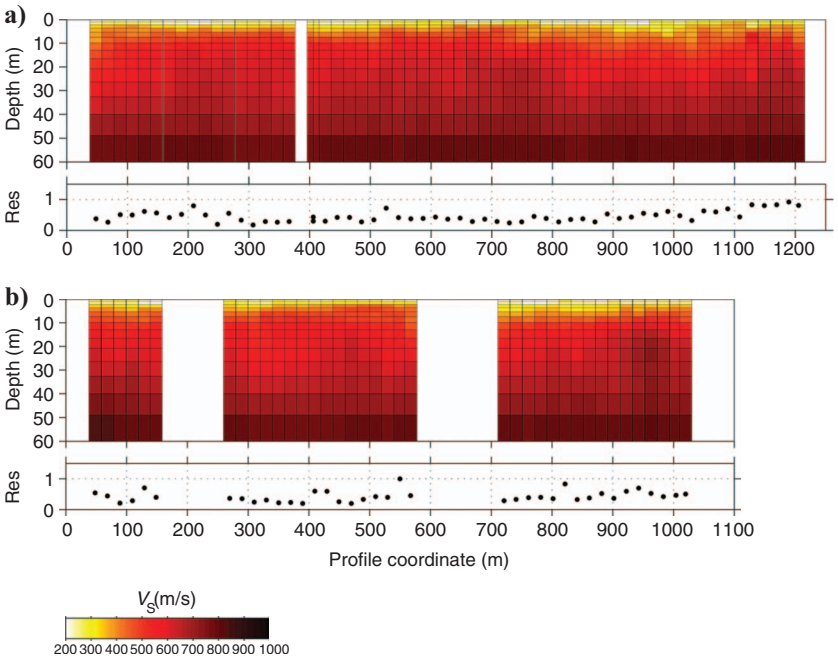

Figure 13. La Salle site, laterally and vertically constrained inversion results: (a) L1, (b) L2. The blank sections are related to topographic anomalies where portions of the data set had to be discarded (Socco et al., 2008).

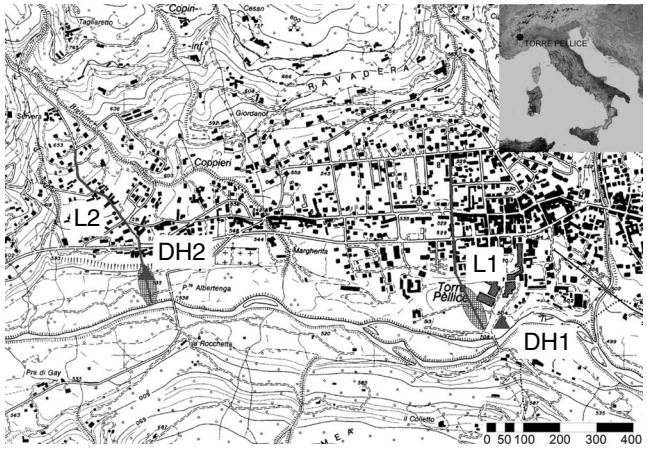

Figure 14. Map of the Torre Pellice site with the survey location. Ground roll LCI was performed along parts of lines L1 and L2 (shown by the shaded areas). DH1 and DH2 represent the downhole test locations.

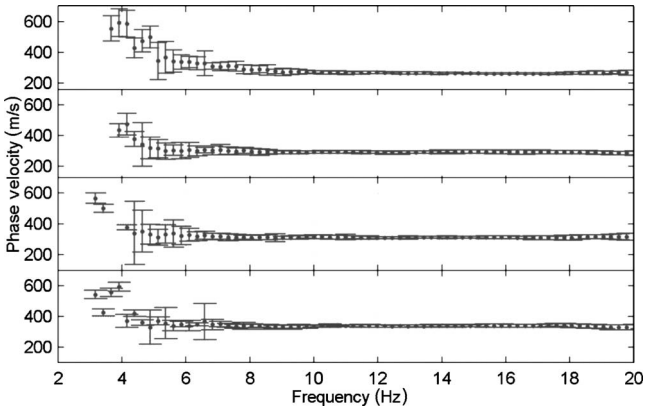

Figure 15. Torre Pellice site, L1 — experimental stacked dispersion curves.

The strength of the constraints introduces different levels of smoothing into the inversion result and should therefore be chosen according to the expected geologic variability. To assess the correctness of the choice of constraint strength, we consider two parameters: the STDF (related to model parameter resolution) and the normalized residuals. The introduction of constraints improves model-parameter resolution: the stronger the constraints, the better the model parameters are resolved. On the other hand, if the constraints are too strong, they produce an unrealistic smoothing of the model parameters, shown by the increase in the normalized residuals. If the normalized residuals obtained from the unconstrained inversions and the ones obtained from the LCI are of the same magnitude, the improvement in model resolution does not correspond to an artificial smoothing. We have shown this for the synthetic data set (Figure 6) and for the field data (Figures 17 and 19).

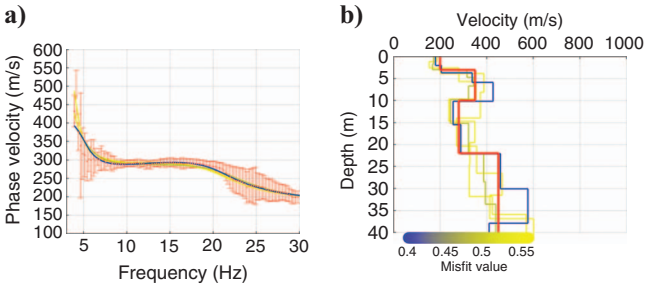

Figure 16. Monte Carlo inversion for the choice of the LCI initial models. (a) Experimental dispersion curve (SDC) with uncertainties for the Torre Pellice site (red dots and red bars) together with the theoretical curves relative to the MCI final models. (b) MCI results plotted with a color scale that represents the misfit, together with the chosen initial model for the LCI (red).

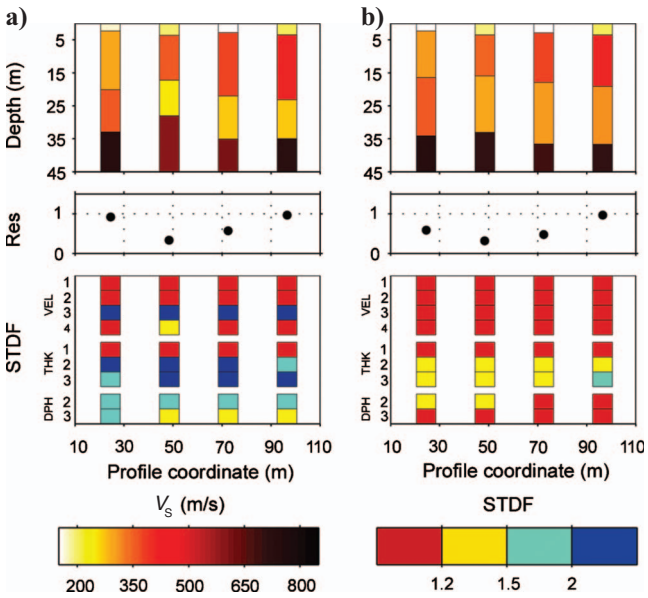

Figure 17. Torre Pellice site, line L1. (a) The individual inversion and (b) the LCI. Starting from the top are the V_S profiles for each dispersion curve along the seismic line (the V_S color scale is at the bottom left), the normalized residual for each V_S model, and the STDF for each model parameter. The STDF color scale is at the bottom right; red depicts well resolved, and blue depicts poorly resolved.

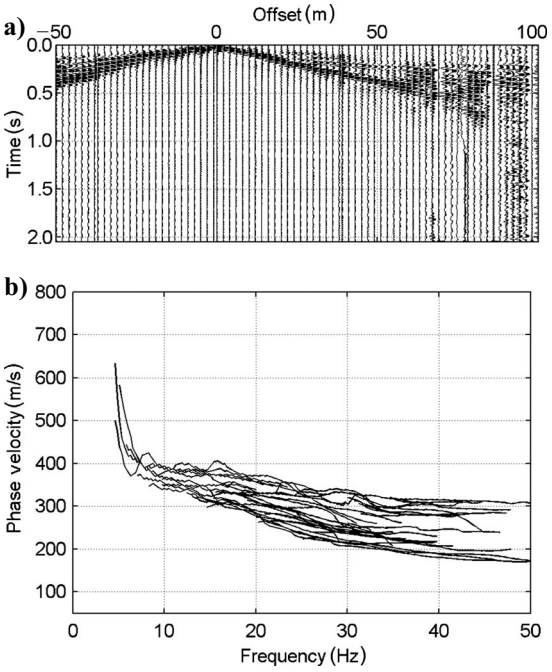

Figure 18. Torre Pellice site, line L2. (a) Example of the raw seismic data. (b) Experimental dispersion curves. The curves are superimposed, but each curve corresponds to a different position along the line.

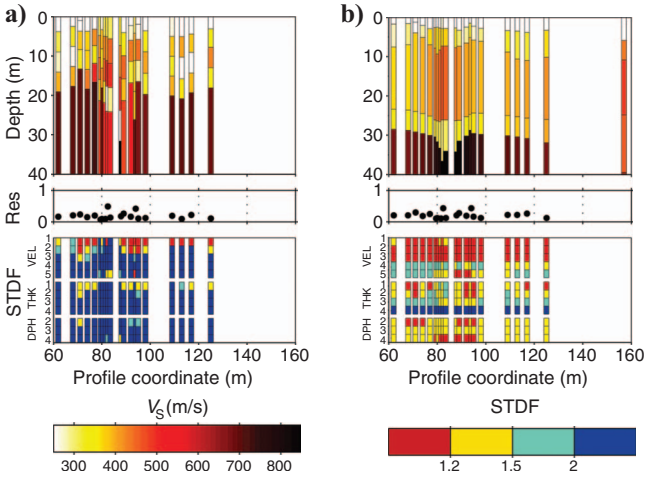

Figure 19. Torre Pellice site, line L2. (a) The individual inversion and (b) the LCI. Starting from the top are the V_S profiles for each dispersion curve along the seismic line (the V_S color scale is at the bottom left), the normalized residual for each V_S model, and the STDF for each model parameter. The STDF color scale is at the bottom right; red depicts well resolved, and blue depicts poorly resolved.

CONCLUSION

Using synthetic and field data, we have demonstrated a procedure to obtain a pseudo-2D shear wave velocity model by analyzing surface waves in seismic reflection records. A semiautomatic processing approach extracts a set of dispersion curves with experimental uncertainties from multifold data. Several gathers are extracted in a moving window and stacked in the f-k domain to improve the S/N. A dispersion curve and its experimental uncertainties are extracted for each position of the moving window along the line. The dispersion-curve data set is inverted through least-squares laterally constrained inversion. The initial model for the least-squares inversion is obtained through a Monte Carlo inversion.

The test on the synthetic data shows that LCI is a powerful tool for consistent and reliable estimation of a V_S model with smooth lateral variations. The results obtained from the synthetic and field data sets show good resolution of the model parameters (expressed in terms of the standard deviation factor). This also applies for challenging targets, such as velocity decreases.

This method is not applicable when abrupt lateral changes in the layer parameters are expected. Only smooth lateral variations occur in the synthetic and the field examples we have shown.

Our method offers several improvements over traditional surface-wave analysis. The processing implemented to retrieve the dispersion-curve data set is based on preliminary tests on the data quality and then becomes fully automatic. It applies stacking to improve dispersion-curve quality and also allows the experimental uncertainties to be calculated and later included in the inversion. The initial model for the inversion is selected using a preliminary global search inversion and is not based on subjective choices. Finally, a laterally constrained inversion is applied to exploit the global information of the data set and to retrieve a pseudo-2D final model. The whole procedure allows the surface wave contained in a seismic reflection data set to be exploited.

Considering future developments of this work, tuning the constraint strength according to data information and variability along the seismic line may improve the approach. It would also be interesting to weight the constraints for different parameters (velocity and thickness), based on a priori information. Other studies are in progress on synthetic data to better tune the level of the lateral constraints in the LCI process. A foreseeable upgrade of the procedure concerns the analysis of 3D seismic and an inversion process in which constraints from neighboring zones are applied in several directions.

ACKNOWLEDGMENTS

The Seismic Service of the Environment Regional Agency of the Piemonte Region (ARPA Piemonte) and the Natural Risk Management and Observational Seismology Department of the Aosta Valley Region supported the data acquisition and allowed publication of the data. The European Union financed the Interreg III B-Alpinespace-Sismovalp Seismic Hazard and Alpine Valley Response Analysis project through which the field data were acquired. Esben Auken and Anders Christiansen, Aarhus University, kindly shared their LCI code for research purposes. We would also like to thank Margherita Maraschini for her valuable contribution to the LCI forward-modeling code. Thanks are also extended to Luigi Sambuelli for acting as a peer reviewer of this paper and to two anonymous reviewers for their valuable suggestions.

REFERENCES

Auken, E., and A. V. Christiansen, 2004, Layered and laterally constrained 2D inversion of resistivity data: Geophysics, **69**, 752–761.

Auken, E., A. V. Christiansen, B. H. Jacobsen, N. Foged, and K. I. Sørensen, 2005, Piecewise 1D laterally constrained inversion of resistivity data: Geophysical Prospecting, **53**, 497–506.

Bohlen, T., S. Kugler, G. Klein, and F. Theilen, 2004, 1.5-D inversion of lateral variation of Scholte-wave dispersion: Geophysics, **69**, 330–344.

Calderón-Macías, C., and B. Luke, 2007, Improved parameterization to invert Rayleigh-wave data for shallow profiles containing stiff inclusions: Geophysics, **72**, no. 1, U1–U10.

Christiansen, A. V., E. Auken, N. Foged, and K. I. Sørensen, 2007, Mutually and laterally constrained inversion of CVES and TEM data: A case study: Near Surface Geophysics, **5**, 115–123.

Dziewonski, A. M., and A. L. Hales, 1972, Numerical analysis of dispersive seismic waves, *in* B. A. Bolt, ed., Methods in computational physics: Academic Press, 11, 271–295.

Grandjean, G., and A. Bitri, 2006, 2M-SASW: Multifold multichannel seismic inversion of local dispersion of Rayleigh waves in laterally heterogeneous subsurfaces: Application to the Super-Sauze earthflow, France: Near Surface Geophysics, **4**, 367–375.

Haskell, N. A., 1953, The dispersion of surface waves on multilayered media: Bulletin of the Seismological Society of America, **43**, 17–34.

Hayashi, K., and H. Suzuki, 2004, CMP cross-correlation analysis of multichannel surface-wave data: Exploration Geophysics, **35**, 7–13.

Lai, C. G., S. Foti, and G. J. Rix, 2005, Propagation of data uncertainty in surface wave inversion: Journal of Engineering and Environmental Geophysics, **10**, 219–28.

Mansoor, N., L. Slater, F. Artigas, and E. Auken, 2006, High-resolution geophysical characterization of shallow-water wetlands: Geophysics, **71**, no. 4, B101–B109.

Mari, J. L., 1984, Estimation of static corrections for shear-wave profiling using the dispersion properties of Love waves: Geophysics, **49**, 1169–1179.

McMechan, G. A., and M. J. Yedlin, 1981, Analysis of dispersive wave by wave field transformation: Geophysics, **46**, 869–874.

Neducza, B., 2007, Stacking of surface waves: Geophysics, **72**, no. 2, V51–V58.

Nolet, G., and G. F. Panza, 1976, Array analysis of seismic surface waves: Limits and possibilities: Pure and Applied Geophysics, **114**, 776–790.

Ostle, B., 1963, Statistics in research: Basic concepts and techniques for research workers: Iowa State University Press.

Park, C. B., R. D. Miller, and J. Xia, 1999, Multichannel analysis of surface waves: Geophysics, **64**, 800–808.

Socco, L. V., and D. Boiero, 2008, Improved Monte Carlo inversion of surface wave data: Geophysical Prospecting, **56**, 357–371.

Socco, L. V., D. Boiero, C. Comina, S. Foti, and R. Wisén, 2008, Seismic characterisation of an alpine site: Near Surface Geophysics, **6**, 253–265.

Socco, L. V., and C. Strobbia, 2004, Surface wave methods for near-surface characterisation: A tutorial: Near Surface Geophysics, **2**, 165–185.

Tarantola, A., and B. Valette, 1982, Generalized nonlinear inverse problems solved using a least squares criterion: Reviews of Geophysics and Space Physics, **20**, 219–232.

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

Tian, G., D. W. Steeples, J. Xia, and K. T. Spikes, 2003, Useful resorting in surface-wave method with the Autojuggie: Geophysics, **68**, 1906–1908.

Wisén, R., and A. V. Christiansen, 2005, Laterally and mutually constrained inversion of surface wave seismic data and resistivity data: Journal of Environmental & Engineering Geophysics, **10**, 251–262.

Wisén, R., E. Auken, and T. Daklin, 2005, Combination of 1D laterally contrained inversion and 2D smooth inversion of resistivity data with a priori data from boreholes: Near Surface Geophysics, **3**, 71–79.

Yilmaz, Ö, M. Eser, and M. Berilgen, 2006, A case study of seismic zonation in municipal areas: The Leading Edge, **25**, 319–330.