Shallow-structure characterization by 2D elastic full-waveform inversion

Anouar Romdhane, Gilles Grandjean, Romain Brossier, Fayçal Réjiba, Stéphane Operto, Jean Virieux

To cite this version:


HAL Id: hal-00593115
https://hal-brgm.archives-ouvertes.fr/hal-00593115
Submitted on 13 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Shallow-structure characterization by 2D elastic full-waveform inversion

Anouar Romdhane\(^1\), Gilles Grandjean\(^2\), Romain Brossier\(^3\), Faycal Rejiba\(^4\), Stephane Operto\(^5\), and Jean Virieux\(^3\)

ABSTRACT

Assessing the effectiveness of elastic full-waveform-inversion (FWI) algorithms when applied to shallow 2D structures in the presence of a complex topography is critically important. By using FWI, we overcome inherent limitations of conventional seismic methods used for near-surface prospecting (acoustic tomography and multichannel spectral analysis of surface waves). The elastic forward problem, formulated in the frequency domain, is based on a mixed finite-element P0-P1 discontinuous Galerkin method to ensure accurate modeling of complex topography effects at a reasonable computing cost. The inversion problem uses an FWI algorithm to minimize the misfit between observed and calculated data. Based on results from a numerical experiment performed on a realistic landslide model inspired from the morphostructure of the Super-Sauze earthflow, we analyzed the effect of using a hierarchical preconditioning strategy, based on a simultaneous multifrequency inversion of damped data, to mitigate the strong nonlinearities coming from the surface waves. This strategy is a key point in alleviating the strong near-surface effects and avoiding convergence toward a local minimum. Using a limited-memory quasi-Newton method improved the convergence level. These findings are analogous to recent applications on large-scale domains, although limited source-receiver offset ranges, low-frequency content of the source, and domination of surface waves on the signal led to some difficulties. Regarding the impact of data decimation on the inversion results, we have learned that an inversion restricted to the vertical data component can be successful without significant loss in terms of parameter imagery resolution. In our investigations of the effect of increased source spacing, we found that a sampling of 4 ms (less than three times the theoretical maximum of one half-wavelength) led to severe aliasing.

INTRODUCTION

Accurate subsurface imaging based on seismic methods constitutes one of the main issues encountered in the environmental and civil engineering fields. It offers the possibility of taking advantage of a noninvasive technique to depict subsoil structures of the first 100 m as a reconstructed image from a seismic wavefield recorded at the surface. This can be achieved using several reconstruction techniques that analyze different kinds of waves associated with propagation phenomena (diffraction, reflection, dispersion, refraction, etc.). The most conventional technique is based on the inversion of body-wave arrival traveltimes, particularly P-waves, using direct or refracted waves. Efficiency of the process closely depends on the realism of the associated forward problem to account for the characteristics of the medium (heterogeneities, contrasts) in calculating traveltimes. In this context, robustness and efficiency of the ray-tracing technique, based on the asymptotic ray theory in the high-frequency approximation, are restricted to the case of smoothed media (Cerveny et al., 1977; Cerveny, 2001) and consequently are unsuitable for highly heterogeneous subsurface domains.
An alternative to ray tracing and the more robust wavefront construction technique (Vinje et al., 1993, 1996a, 1996b) consists of applying finite differences to solve the eikonal equation numerically (Vidale, 1988; Podvin and Lecomte, 1991), making it possible to deal with more heterogeneous media. Important progress has been achieved to handle the associated inverse problem efficiently, using the popular simultaneous iterative reconstruction technique (SIRT) (van der Sluijs and van der Vorst, 1987; Grandjean and Sage, 2004) or the more appealing adjoint state method (Taillardier et al., 2009). However, applications to real data in the context of shallow prospecting (Grandjean and Leparoux, 2004; Ellefsen, 2009) reveal restrictions. This is particularly the case when later arrivals must be included in the inversion scheme or when dealing with real data where surface waves, which always represent the main component (about two-thirds) of the seismic energy, can seriously mitigate signals used in the inversion.

On the other hand, with the introduction of the spectral analysis of surface waves (SASW) method (Nazarian and Stokoe, 1984, 1986; Stokoe and Nazarian, 1985; Stokoe et al., 1988), surface waves have received much attention. The good signal-to-noise ratio (S/N) of these waves associated with the relative ease of their acquisition gives rise to a variety of applications (Lai, 1998; Park et al., 1999; Rix et al., 2001). Early studies were devoted to reconstructing 1D shear-wave velocity distributions by calculating phase differences between two receivers. The SASW method was later extended to the multichannel analysis of surface waves (MASW), which is based on the phase-velocity variation with frequency from a multichannel recording system. The 1D assumption of the MASW method is imposed by the formulation used for solving the inverse problem (Herrmann, 1991). To overcome this limitation, some extensions have been made (Park et al., 1998; Xia et al., 1999; Grandjean and Bitri, 2006) to adapt the methodology to 2D contexts by narrowing offset windows or and using a summation principle to increase the S/N. The resulting 1D velocity profiles are then interpolated along a seismic line to produce a 2D view of the shear velocity.

Despite all of these developments, some limitations still affect the potential of surface-wave methods. These limitations are mainly the result of difficulties encountered when identifying and separating the first (fundamental) propagation mode from higher modes (possible propagation modes of surface waves in a layered medium), which form the basis of the inversion process. This phenomenon, in addition to the errors resulting from fitting data including 2D or 3D effects (phase-velocity changes) under a 1D assumption, drastically mitigate the efficiency of the MASW method (Bodet, 2005). Some recent results have also shown that the dispersion curve is not an intrinsic property of the medium by emphasizing the influence of acquisition parameters (Socco and Strobia, 2004).

An important point is the common feature of MASW and first-break acoustic tomography. Both use a restrictive part of the information contained in the seismic signal: the dispersion of Rayleigh waves and the first P-wave arrivals. A strategy integrating both signals should be more efficient and physically consistent to reduce the possible solutions satisfying the approaches. To overcome this issue, an alternative approach consists of taking advantage of recent advances in quantitative imaging based on full-waveform inversion (FWI) in the time (Tarantola, 1984) or frequency domains (Pratt et al., 1998). In theory, these approaches offer important possibilities because they use all information contained in seismic signals (P-waves in the acoustic case and P-SV-SH-waves in the elastic case) in the inversion strategy. The inverse problem formulation in the frequency domain has been implemented and applied to synthetic and real data concerning large-scale domains (kilometric scale) (Ravaut et al., 2004; Brenders and Pratt, 2007; Brossier et al., 2009). This context is very different from the subsurface one because low-frequency sources and long offsets can be used. Surface waves can be separated easily from body waves and traditionally are muted.

However, in the context of near-surface imaging, the greatest part of the energy emitted by a surface seismic source contributes to the generation of surface waves. To overcome this limitation, some workers have proposed applying a time window to the early arrivals and performing acoustic waveform tomography with near-surface data (Gao et al., 2006, 2007; Sheng et al., 2006; Smithyman et al., 2009). Results show that this strategy outperforms traveltime tomography and is well suited to data coming from refraction surveys where far and intermediate offsets are considered. However, when the offset range is too small (which is usually the case in near-surface prospecting) to allow separation between body waves and surface waves, the efficiency of this strategy may be severely altered. Moreover, this strategy does not take advantage of the information included in shear and surface waves that are usually considered as a source of noise in the inversion. These waves propagate with a lower velocity than compressional waves and may therefore lead to higher resolution of the images.

To our knowledge, only a few studies of waveform inversion involving body and surface waves have been performed for near-surface (0–100 m) investigations (Gélos et al., 2007; Romdhane et al., 2008). These numerical investigations were performed to image near-surface heterogeneities with various contrasts in a well-known background medium. In these cases, surface topography was considered to be flat, mainly because of the computational difficulties encountered to model surface waves accurately in the presence of a complex topography (Moczo et al., 2007). The effects of irregular topography on seismic wave motion have been the subject of some numerical investigations (Bleibinhaus and Rondenay, 2009; Shiam-Jong et al., 2009). It is well established that topography can drastically influence amplitudes and phases of the seismic signal. Consequently, correct modeling of free-surface effects is a critical requirement for any seismic-inversion process.

Recent work conducted by Brossier et al. (2009) focuses on the impact of applying several multiresolution strategies to mitigate the strong nonlinearity inherent in surface waves. Simulations performed with a section of the well-known SEG/EAGE overthrust model reveal that preconditioning provided by time damping associated with successive inversions of overlapping frequency groups is critical to converge toward acceptable velocity models.

The objective of our work is thus to evaluate the effectiveness of using an FWI algorithm to take advantage of the information contained in surface waves to image heterogeneous shallow structures in the context of a complex surface topography. Our paper is organized in two sections. In the first section, we present a brief review of the basis of the elastic FWI technique used. In the second section, we apply it to a synthetic but
realistic landslide case, derived from the structure of the Super-
Sauze earthflow. We evaluate the efficiency of using precondition-
ing strategies to reconstruct the shallow velocity structure.
We also address the effects of some practical considerations,
particularly the restriction to the vertical data component and
the impact of acquisition decimation, typically related to subsur-
face prospecting, on inversion results.

ELASTIC FULL-WAVEFORM INVERSION

We first consider the 2D P-SV-wave modeling case. The for-
ward and inverse problems are solved in the frequency domain.
The forward problem is based on a discontinuous Galerkin (DG)
approach. An FWI algorithm is used to solve the inverse prob-
lem. It is based on the preconditioned conjugate-gradient (PCG)
method or a limited-memory quasi-Newton Broyden-Fletcher-
Goldfarb-Shanno (L-BFGS) approach and is implemented on a
parallel computation architecture. We then present the precondi-
tioning strategies used for our numerical simulations.

The forward problem

In an isotropic elastic medium, the equation system governing
the wave propagation in 2D media relates velocities $\Delta V_x$ and $V_z$
to stresses $\sigma_{xx}$, $\sigma_{zz}$, and $\sigma_{xz}$. It can be written in the frequency
domain as

$$
-\iota \omega p V_x = \frac{1}{\rho(x)} \left[ \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right] + F_x,
$$

$$
-\iota \omega p V_z = \frac{1}{\rho(x)} \left[ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right] + F_z,
$$

$$
-\iota \omega \sigma_{xx} = \left( \lambda(x) + 2 \mu(x) \right) \frac{\partial V_x}{\partial x} + \lambda(x) \frac{\partial V_z}{\partial z},
$$

$$
-\iota \omega \sigma_{xz} = \lambda(x) \frac{\partial V_x}{\partial x} + \left( \lambda(x) + 2 \mu(x) \right) \frac{\partial V_z}{\partial z},
$$

$$
-\iota \omega \sigma_{xz} = \mu(x) \frac{\partial V_x}{\partial x} + \frac{\partial V_z}{\partial x}.
$$

(1)

where $\lambda$ and $\mu$ are the Lamé coefficients, $x$ is spatial position, $\rho$
is density, and $\omega$ is angular frequency. The physical properties
of the medium are supposed to be constant inside each cell, and
central numerical fluxes are used. Details of the mixed DG
interpolation orders P0-P1 formulation, used in this study, are
provided in Brossier (2009).

System 1 can be written with respect to a linear matrix for-
malism for each frequency considered:

$$
Ax = s
$$

(2)

where vector $x$ denotes the unknowns, consisting of the particle
velocities and stresses, $s$ is the source term, and $A$ is the imped-
ance matrix. To solve the linear system resulting from discretiz-
ing equation 2, the impedance matrix is first factorized with an LU
decomposition independent of the source term. Solutions for multi-
ple sources (i.e., multiple right-hand-side terms) can then be
obtained efficiently by forward and backward substitutions. Parallel
factorization of the impedance matrix is performed using the
MUMPS massively parallel direct solver package (MUMPS, 2009).

Some recent results reveal promising prospects for applying
the DG method to elastic-wave propagation. The use of high
orders of interpolation is especially appealing because they
allow unstructured meshes and thus offer the possibility
of locally adapting the mesh size to local medium parameters (h-
adaptive mesh). They also ensure high accuracy with a coarse
discretization of the medium (Dumbser and Käser, 2006). How-
ever, this coarse discretization may be inconsistent with the
expected resolution of the FWI, which necessitates a discretiza-
tion close to $\lambda/4$ (Sourbier et al., 2009).

In this study, we use a lower order of interpolation. Applying
the DG method based on the lowest interpolation order (P0)
turns out to be very efficient, in terms of computational cost, in
comparison with classical finite-difference formulations in the
context of contrasted media and smooth surface topography
(Brossier et al., 2008). The accuracy is guaranteed with only
10–15 cells per minimum wavelength compared to the 60 grid
points necessary with the rotated second-order stencil and the
vacuum formalism (Saenger et al., 2000; Bohlen and Saenger,
2006). In addition, an interesting compromise between accuracy
of wavefield estimation and computational cost consists of using
the mixed P0-P1 DG interpolation to overcome some particular
restrictions related to topography complexity. The use of
unstructured meshes (for P1) ensures precise implementation of
the source term and accurate modeling of the complex topogra-
phy, taking into account the free-surface boundary conditions
(Brossier, 2010). In addition, it offers the possibility to adapt
mesh size to the local physical parameters. This property is of
great interest, especially in the context of near-surface modeling
with weathered zones (with very low velocities).

The inverse problem

In this section, we briefly review the principles of FWI. An
extensive overview of the method can be found in Virieux
and Operto (2009).

In the case of weighted least-squares linearized inversion, the
misfit function $E$ can be expressed (Tarantola, 1987) as

$$
E(m) = \frac{1}{2} \left( d_{obs} - d_{cal} \right)^T S_d S_d \left( d_{obs} - d_{cal} \right),
$$

(3)

where the dagger $\dagger$ denotes the transpose conjugate; $d_{obs}$
and $d_{cal}$ denote observed and calculated data (particle velocities),
respectively; $d_{obs} - d_{cal}$ corresponds to the data-misfit
vector in model $m$, and $S_d$ is a weighting operator applied to
the data. A solution to equation 3 is to linearize it in the second
order around an initial model $m_0$, which corresponds to the
model of the $l$th iteration as follows:

$$
E(m^{(l-1)} + \delta m^{(l)}) = E(m^{(l-1)}) + \nabla_m E(m^{(l-1)}) \delta m^{(l)} + \frac{1}{2} \delta m^{(l)} H(m^{(l-1)}) \delta m^{(l)},
$$

(4)

where $\delta m^{(l)}$ is the model perturbation and where $\nabla_m E(m^{(l-1)})$ and $H(m^{(l-1)})$ are the gradient and Hessian of the misfit function,
respectively. Minimizing $E$ leads to the Newton equation, which
relates the model perturbation to the gradient and Hessian as

$$
\delta m^{(l)} = - \left( H(m^{(l-1)}) \right)^{-1} \nabla_m E(m^{(l-1)}).
$$

(5)

The gradient direction is computed efficiently following the
adjoint-state formulation (Plessix, 2006). For one model parame-
ter $k$, the system can be recast in matrix form:
\[ \nabla_m E(m_k) = R \left[ x^t \left( \frac{\delta A}{\delta m_k} \right)^t A^{-1} P S_D^t S_d d^t \right], \]  
(6) 

where \( P \) is an operator that projects the data residual vector in the data space to the model space, \( R \) is the real part of a complex number, \( \delta d \) corresponds to the data misfit vector, and \( t \) and * are the transpose and conjugate operators. Equation (6) shows that the gradient is formed by a weighted product of the incident wavefield \( x \) and the adjoint wavefield \( A^{-1} P S_D^t S_d d^t \). The gradient of the misfit function therefore requires computing only two forward problems per shot.

In practice, for realistically sized problems, resolving the Newton equation (equation (6)) is avoided because of the large inherent cost. One alternative used in this study consists of considering only diagonal terms of the Hessian or the pseudo-Hessian matrix (Pratt et al., 1998; Shin et al., 2001) as a preconditioner for the optimization algorithm.

To overcome the diagonal estimation of the Hessian, an L-BFGS method can be used (Nocedal and Wright, 1999). This algorithm is more efficient than the preconditioned nonlinear conjugate gradient for solving FWI problems (Brossier et al., 2009). The algorithm estimates a nondiagonal inverse Hessian from the \( m \) most recent gradient and model vectors. An initial estimate of the Hessian can be provided from the diagonal terms of an approximate Hessian. An example of the contribution of this method to improve the convergence level of the misfit function is illustrated later.

Once the right-hand side of equation (6) is estimated, the model is updated iteratively:

\[ m'(l) = m(l-1) + \alpha \delta m(l), \]  
(7) 

where \( \alpha \) denotes the step length, estimated in this study by parabola fitting.

Efficient mitigation of nonlinear effects

FWI is carried out by proceeding iteratively from low to high frequencies. This allows short wavelengths to be introduced progressively in the parameter images and thus helps to mitigate the nonlinearity of the inverse problem. The strategy has proven effective for the acoustic inverse problem (Pratt, 1999; Ravaut et al., 2004; Operto et al., 2006). In the elastic case, work conducted to evaluate the ability of the method to locate small heterogeneities in shallow subsurface structures in the presence of a flat topography (Gélis, 2005) reports many difficulties stemming from the presence of complex wave phenomena, particularly surface waves. Because the waves contain most of the seismic energy and because they interact strongly with the topographic irregularities, we speculate that they will significantly govern the optimization process and constraint the algorithm to explore a wrong solution and reach a local minimum.

To fulfill our objective, i.e., imaging shallow and highly contrasted velocity structures in the presence of a complex topography, we must take into consideration three critical points.

1) An accurate starting model is required. It must be close enough to the true velocity model to avoid the cycle-skipping phenomenon, which may occur when the error traveltime is greater than half a period.

2) The receiver antenna must be long enough to ensure good model illumination. Limited-aperture acquisition geometries can result in the algorithm being trapped in a local minimum.

3) The choice of inverted frequencies is critical to guarantee accurate coverage in terms of long and short wavelengths, especially for the S-wave velocity parameter. Low frequencies must be considered to avoid convergence toward a local minimum at an early stage. This restriction also explains the necessity of considering a starting model close enough to the real one. Moreover, selection of the inverted frequencies must ensure a continuous window illumination following, for example, the strategy proposed by Sirgue and Pratt (2004).

An alternative to mitigate the strong nonlinearities resulting from complex wave phenomena consists of defining two levels of hierarchy (Brossier et al., 2009). The first is to perform successive inversions of overlapping groups of finite frequencies to better constrain the algorithm and take into account the redundant information contained in the selected frequencies. Frequencies of each group are inverted simultaneously, and the overlapping (frequencies in common) between two successive groups is minimized. Application of this strategy to the SEG/EAGE overtrust model reveals some improvements in comparison to a sequential single-frequency approach.

The second level consists of progressively introducing later arrivals (converted waves, surface waves) in the inversion. In the frequency domain, this level of hierarchy can be implemented in a flexible way by time windowing (Pratt and Shipp, 1999). In the frequency domain, only time damping can be used (Shin et al., 2002). Time damping of a seismic signal \( x(t) \) with respect to the first-arrival traveltime \( t_0 \), for example, can be implemented in the frequency domain by introducing a complex-valued frequency following the expression

\[ X(w + i\gamma) = \int_{-\infty}^{+\infty} x(t) \exp(-\gamma(t-t_0)) \exp(-iwt) dt, \]  
(8) 

where \( \gamma \) denotes the applied damping factor.

To assess the effectiveness of these strategies in our context, we conducted a numerical study for a realistic landslide model.

For all tests presented hereafter, the inverted model parameters are P- and S-wave velocities. The source-parameter estimate is not addressed, although it is a critical issue when applying FWI to real data. The proposed numerical tests are performed to highlight two aspects. In the first section, we evaluate the performance of the defined preconditioning strategies to recover the velocity structures and to assess the contribution of the L-BFGS optimization method. In the second section, we study the effect of decimating the acquisition geometry on the inversion results, notably in term of number of sources, to be as close as possible to realistic cases.

LANDSLIDE SYNTHETIC CASE STUDY: A NUMERICAL EXPERIMENT

The landslide model was inspired from a transverse section of the Super-Sauze earthflow located in the French Alps (Plageollet et al., 2000). It consists of a 210 × 60-m section composed of several velocity layers, as proposed by Grandjean et al. (2006), after performing first-arrival tomography. The medium is
Elastic FWI for shallow structures

characterized by strong lateral velocity variations associated with highly contrasted media, with P- and S-wave velocities varying from 800 to 3200 m/s and 480 to 1600 m/s, respectively (Figure 1; Table 1) with an inconstant Poisson’s ratio. Here we used a constant density of 1600 g/cm³ for forward and inverse problems. The surface topography is highly irregular. Simulations were performed using a Ricker source wavelet with a peak frequency of 60 Hz. In a real data context, the source signature and radiation pattern are additional unknowns that can be estimated by solving a linear inverse problem (Pratt, 1999; Virieux and Operto, 2009). The chosen parameters correspond to wavelengths (at the central frequency) varying between 53.3 and 13.4 m for V_P and between 26.6 and 8 m for S-wave velocity V_S. The frequency bandwidth covers the interval [10,150] Hz. Detecting shallow structures of metric scale is thus affected by the bad resolution of the thin layers of the model. From a numerical point of view, a source with higher-frequency content should overcome this limitation. However, this assumption is meaningless in practice because high-frequency signals are strongly attenuated in the shallow, fissured layers of the medium and thus useless.

The mesh is divided into a 1-m-thick unstructured layer for P1 interpolation and a structured layer (made of equilateral triangles) for P0. This choice ensures at least 15 grid cells per minimum propagated wavelength (corresponding to the surface wave estimated from the shear wave and the Poisson’s ratio with the Viktorov formula (Viktorov, 1965)) for the highest modeled frequency.

A total of 197 explosive (Ricker) sources were considered 1 m below the surface, with a 1-m spacing along the horizontal axis; 197 receivers were located 0.5 m below the surface. Vertical and horizontal particle velocities were computed. An example of vertical and horizontal components of one shot gather (Figure 2) shows that most of the seismic energy is radiated in the form of surface waves. It also highlights the footprint of the irregular topography on the seismic signal, which is drastically warped. Strong diffractions resulting from the topography shape can be observed, particularly for the incident surface waves.

Impact of inversion conditioning

In our tests, V_P and V_S are the inverted model parameters, and density is supposed to be known. Starting models are smoothed versions of true ones, obtained after applying a 2D Gaussian smoothing function with a spatial correlation length of 6 m for V_S (Figure 3) and V_P parameters. This constitutes a good compromise between severely altering the delineation of layer interfaces and avoiding the cycle-skipping phenomenon that may occur when the starting models are too far from the real ones. For the shallow layers, it is a realistic model that can be obtained by conventional methods. These models also pose that we have a priori knowledge on the shape of the bedrock. Examples of vertical and horizontal components of shot gather are shown in Figures 4a and 4b, for comparison with those of Figures 2a and 2b.

Successive inversions of single frequencies

In a first step, sequential inversion is performed with respect to the selected frequencies of Table 2 to ensure a continuous wavenumber illumination (Sirgue and Pratt, 2004). Horizontal and vertical components are considered. A maximum of 25 iterations per frequency is performed to ensure convergence of the algorithm at reasonable computing cost.

The final models obtained are shown in Figures 5a and 6. The inversion fails to converge toward an acceptable model for V_P and V_S. Indeed, the main features of the layered structure are not recovered. In addition, strong artifacts are observed. The algorithm has converged into a local minimum because we observe strong, unrealistic anomalies near the free surface (Figure 5b and Figure 7b). This failure can be attributed to the dominant contribution of surface waves that prevents the high-frequency signals associated with body waves to be considered in the inversion. Similar effects have been observed by Gélis et al. (2007) and Romdhane et al. (2009).

Frequency group inversion of damped data

In this section, we investigate the performance of a simultaneous inversion of damped data. We consider three overlapping groups of three frequencies (see Table 2), with damping

Table 1. Maximum and minimum velocity parameters for the landslide model. Maximum and minimum wavelengths are calculated for the lowest and highest inverted frequencies, respectively.

<table>
<thead>
<tr>
<th>V_P (m/s)</th>
<th>V_S (m/s)</th>
<th>( \lambda_{P_{	ext{max}}} ) (m)</th>
<th>( \lambda_{P_{	ext{min}}} ) (m)</th>
<th>( \lambda_{S_{	ext{max}}} ) (m)</th>
<th>( \lambda_{S_{	ext{min}}} ) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>800</td>
<td>480</td>
<td>23.8</td>
<td>5.9</td>
<td>11.9</td>
<td>3.57</td>
</tr>
<tr>
<td>3200</td>
<td>1600</td>
<td>150.2</td>
<td>37.6</td>
<td>75.1</td>
<td>22.5</td>
</tr>
</tbody>
</table>

Figure 1. (a) V_S true velocity model for the (realistic) landslide model. The gray dashed line and the black star correspond to the position of the extracted vertical profiles and the source position of shot gathers depicted in the following figures, respectively. (b) Zoom of the boxed area, showing the mesh used for the landslide model with the mixed DG P0-P1 method.
coefficients varying between 20 and 1.5. Figures 4c and 4d shows the vertical and horizontal components obtained with a damping coefficient of 20. The shot position used in Figure 2 is considered. Comparison of shot gathers of Figures 2 and 4 highlights the role of data damping to progressively introduce surface waves as well as complex free-surface reflections, particularly for the far offsets. A maximum of 25 iterations was performed for each damped frequency group.

Final results, obtained after inverting the three groups (Figures 8 and 9), reveal how crucial this strategy is to converge successfully toward an acceptable solution. For $V_S$, focusing the shallower structures is defined with a high level of resolution. For the deeper layers, the model is slightly improved and the contribution is less significant as the velocity values (and thus

Figure 2. Examples of (a) horizontal and (b) vertical synthetic shot gathers of the landslide model. The shot position correspond to an abscissa of 100 m on the horizontal distance axis of Figure 1. DP, RP, and RPW correspond to direct, refracted, and reflected P-waves, respectively. RW and RRW correspond to Rayleigh waves (fundamental mode) and back-propagated Rayleigh waves.

Figure 3. Starting $V_S$ model considered for the landslide case.

Figure 4. (a) Horizontal and (b) vertical synthetic shot gathers of the starting landslide model. (c) Damped horizontal and (d) vertical synthetic shot gathers of the starting landslide model ($\gamma = 20$). The shot position correspond to an abscissa of 100 m on the horizontal distance axis of Figure 1. DP, RP, and RPW correspond to direct, refracted, and reflected P-waves, respectively. RW correspond to Rayleigh waves (fundamental mode).

Table 2. Sequential inverted frequencies, frequency groups, and damping coefficients considered for the landslide model.

<table>
<thead>
<tr>
<th>Frequency group</th>
<th>Sequential frequency (Hz)</th>
<th>Simultaneous frequency (Hz)</th>
<th>Damping coefficient (1/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21.3</td>
<td>21.3, 27.5, 42.7</td>
<td>20, 5, 1.51</td>
</tr>
<tr>
<td>2</td>
<td>27.5</td>
<td>42.7, 61.0, 82.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>42.5</td>
<td>82.4, 106.8, 134.3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>61.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>82.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>106.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>134.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the associated wavelengths) increase (Figures 8 and 10b). The weak contribution of the inversion process to reconstruct \( V_P \) (Figures 9 and 10a) was expected and can be explained by the lack of short wavelengths illuminated with respect to the \( V_P \) model velocities (see the wavelengths associated with \( V_P \) in Table 2). Few artifacts can, however, be noticed in the zones corresponding to the highest velocity contrasts (between the shallowest layer and the bedrock, in Figure 8b).

Examples of an initial differential seismogram (difference between data calculated with the true model and data calculated with the starting model) and a final differential seismogram (difference between data calculated in the final model and data calculated with the true model) are depicted in Figures 10c and 10d, respectively. The comparison shows that the unexplained energy mainly comes from the back-propagated Rayleigh waves at the highest velocity contrasts.

The penetration depth of the Rayleigh wave is approximately half of its wavelength. This means it will dominate the low-frequency part of the data spectrum, whereas body waves will dominate the high-frequency part of the spectrum. Applying strong damping coefficients to the high frequencies to favor the use of body waves is therefore unnecessary, in our opinion.

We have also investigated the effect of resampling the frequency interval in the inversion group. We have divided the frequency interval by a factor of two and considered three groups of five (instead of three) frequencies. This resampling is expected to strengthen the spectral redundancy and yield a higher definition of layers. The same damping coefficients were considered as in the previous example. Results in Figure 11 show a slight improvement of \( V_S \) parameter reconstruction at the expense of a significantly higher computing cost.

**Contribution of L-BFGS method**

We performed an inversion test using the L-BFGS optimization method with the same frequency groups and damping coefficients as in the previous section (with respect to Table 2). The initial estimate of the Hessian is provided by the diagonal elements of the pseudo-Hessian (Shin et al., 2001), and five differences of cost-function gradients and model vectors are used for the L-BFGS algorithm. Figures 12a and 12b and Figures 13a and 13b show final \( V_P \) and \( V_S \) reconstructed velocity models and vertical extracted profiles for each parameter, respectively.

![Figure 5. (a) Final \( V_S \) model and (b) relative velocity error (ratio of the velocity error to the true velocity), obtained after sequential inversion of seven frequencies varying from 21.3–134.3 Hz as indicated in Table 2. Vertical and horizontal components are inverted.](image)

![Figure 6. Final \( V_P \) model obtained after sequential inversion of seven frequencies varying from 21.3 to 134.3 Hz as indicated in Table 2. Vertical and horizontal components are inverted.](image)

![Figure 7. (a) \( V_P \) and (b) \( V_S \) parameter profiles extracted along a vertical line (located at a distance of 100 m) obtained after sequential inversion of seven frequencies varying from 21.3 to 134.3 Hz as indicated in Table 2. Profiles of the true model are plotted with solid black lines, the initial model is the dotted lines, and the final model is the dashed lines.](image)
Amplitudes of the structures are significantly better defined for the shallow and the deep layers of the model in comparison with those obtained with the PCG optimization method (see Figures 8a, 9, 10a, and 10b). Fewer artifacts can be observed even in the shallow zone corresponding to the highest velocity contrasts.

Figure 12c depicts the evolution of the logarithm of the misfit function with respect to the iteration number with the L-BFGS and PCG methods for the first frequency group. The convergence level is drastically improved with the L-BFGS algorithm when compared to PGC. The final differential seismograms (Figure 14) show that amplitude residuals are strongly attenuated, even for the longer recording times. The strong residuals (observed in Figure 10d) from the highest velocity contrasts are significantly lessened. Similar effects have been noticed by Brossier et al. (2009) and attributed to the contribution of the off-diagonal terms of the Hessian matrix estimated by the L-BFGS algorithm. This confirms the promising prospects for realistic applications.

Impact of data decimation

In this section, we analyze the impact of data decimation on the inversion results. The percentage of model degradation with respect to a reference case is estimated with the root mean square (rms) of the relative velocity error.
groups and damping terms were used as in the previous section for consistency. The final distribution of $V_S$ reconstructed by considering the vertical component has a higher resolution (Figures 15a and 15c) compared to the one obtained by considering the horizontal component. In the better case, we observe large velocity errors close to the free surface in the zones with a very high contrast between the steeply dipping bedrock and the shallowest layer (Figures 15b and 15d). Percentages for $V_S$ model degradation (with respect to the reference case of Figure 8b) are 24% with the vertical component and 102% with the horizontal component. As a result, the algorithm appears more sensitive to the information provided by the vertical data component.

In addition, final inverted $V_S$ images obtained by considering the vertical component do not differ significantly from the ones calculated by considering vertical and horizontal components. However, a comparison of the figures of associated relative velocity errors (Figures 8b and 15b) shows the lower resolution of the deeper layers of the model.

**Acquisition configuration**

We finally investigate the sensitivity of inversion results to the parameters of the recording geometry. The impact of
decimating survey geometries on the waveform tomography for lithospheric imaging is addressed by Brenders and Pratt (2007). The requirement of fully unaliased surface sampling $\Delta_{\text{samp}}$ is given by the relationship $(\Delta r, \Delta s) \leq \Delta_{\text{samp}} = \lambda/2$, where $\Delta r$ and $\Delta s$ denote receiver and source spacings and $\lambda$ is the calculated wavelength for a specific frequency with respect to the minimum velocity of the medium. Brenders and Pratt (2007) suggest that for a receiver spacing below $\lambda/2$, the image quality remains acceptable for $\Delta s' > 3\Delta_{\text{samp}}$. In the context of near-surface imaging, field and logistic limitations often prevent the use of a dense sampling of source array. We keep the number of receiver constant and evaluate the effect of decimating the number of sources with a sparser grid consisting of 99 sources (with a 2-m spacing) and then 49 sources (with a 4-m spacing), respectively.

Recalling that the minimum wavelength with respect to the S-wave velocity is computed from $V_{S_{\text{min}}} = 480$ m/s and $f_{\text{max}} = 134.2$ Hz, we have $\Delta r = 1m \leq \Delta_{\text{samp}} = 1.78$ m and $\Delta s < 3\Delta_{\text{samp}} \approx 3\lambda_{\text{min}}/2 \approx 5.34$ m for both cases. The minimum and maximum offset coverages along the model are retained. Only the vertical data component is considered for the inversion.

Although the acquisition aperture is not modified, this acquisition geometry is expected to mitigate the inversion performance slightly. Results show that a source sampling of 2 m leads to acceptable results (Figures 16a and 16b) in comparison to those obtained with a 1-m source sampling. The percentage of model degradation is 12%. However, associated vertical profiles depicted in Figures 17a and 17b demonstrate that the deep structure (with a $V_S$ velocity of 800 m/s) is defined with a lower definition. A source sampling of 4 m introduces significant aliasing effects near the free surface (Figures 16c and 16d), translated into strong artifacts. The inversion obviously fails to converge toward the true model with an acceptable level of resolution. The final $V_S$ image obtained is severely altered, with respect to Figure 15a. The percentage of model degradation is 43%.

Extracted vertical profiles (Figures 17c and 17d) confirm the aliasing effect observed near the free surface and the poor resolution of the deep structures. This effect can be even more pronounced in the presence of noise in real data, which seriously mitigates the effectiveness of the algorithm to improve layer definition.

**Computing time**

We used a constant mixed P0-P1 mesh with a total of 266,709 cells composed of 8232 P1 cells (with three degrees of freedom per field) and 258,477 P0 cells (with one degree of freedom per field), giving 1,415,865 degrees of freedom. The forward modeling required an average time of 89 s to be solved.
for 197 sources per modeled frequency. For the inversion, each iteration required an average time of 400 s and a total memory of 6.7 Gb for factorization. All simulations were performed using 24 cores on a BRGM HP DL 165 G2 cluster, which consists of 32 nodes with Myrinet interconnection. Each node comprises two quad-core 2.3-GHz AMD Opteron processors, providing 16 Gb of RAM.

DISCUSSION

We have applied elastic FWI inversion to a realistic landslide model characterized by strong lateral velocity variations and a complex surface topography. In this particular context, the seismic signal is dominated by surface waves that cannot be separated easily from body waves because of the limited aperture of the acquisition geometry. The inversion of raw data failed to yield acceptable velocity images. This failure can be attributed to the dominant contribution of surface waves. The effects of surface waves on FWI have been investigated in small-scale field experiments by Gélis et al. (2007) for the elastic case and by Bleibinhaus and Rondenay (2009) in the presence of a complex topography for the acoustic case on a larger scale.

The performed tests reveal that a combination of inversion of overlapping groups of multiple frequencies and data damping to allow a progressive introduction of the complex seismic events (surface waves, multiples) is critical to mitigate the strong nonlinearities introduced by surface waves and to reconstruct the shallow structures accurately. The use of a quasi-Newton LBFGS optimization algorithm can significantly improve the convergence level of the method and the parameters reconstruction.

Figure 16. (a) Final $V_S$ model and (b) relative velocity error (ratio of velocity error to true velocity), obtained after simultaneous inversion of three damped-frequency groups varying from 21.3 to 134.3 Hz as indicated in Table 2. Only the vertical component is inverted and a decimated acquisition ($\Delta s = 2$ m) is used. (c, d) Final $V_S$ model and relative velocity error with $\Delta s = 4$ m.

Figure 17. (a, c) $V_P$ and (b, d) $V_S$ parameter profiles extracted along a vertical line (located at a distance of 100 m), obtained after simultaneous inversion of three damped-frequency groups varying from 21.3 to 134.3 Hz, as indicated in Table 2. Only the vertical component is inverted. Two levels of decimation are considered with a source sampling of 2 m (a, b) and 4 m (c, d) along the surface topography. True model is plotted in solid black lines, initial model is plotted in dotted lines, and final model is plotted in dashed lines.
Efficiency of the process can, however, be severely altered by an insufficient source sampling interval.

An important extension to the elastic FWI, which may be critical for challenging real data applications, should incorporate the reconstructed attenuation parameters $Q_p$ and $Q_s$. Implementation of the algorithm in the frequency domain can take advantage of the complex-velocity Kolsky-Futterman model (Toverud and Ursin, 2005). Application to real data, in the context of near-surface characterization, can be based on the waveform-to-tomography workflow proposed by Smithyman et al. (2009) to produce images of $V_p$ and $Q_p$ parameters and to locate shallow targets. It also requires an estimation of the source (Pratt, 1999), which was supposed known in our work.

**CONCLUSION**

We have presented a numerical study to evaluate the potential of a 2D FWI approach; it shows promising prospects for imaging shallow structures in the presence of a complex topography. A discontinuous Galerkin method, based on a low-order mixed P0-P1 interpolation, is used for accurate waveform modeling at a reasonable computing cost. A 2D elastic-frequency-domain FWI algorithm has been applied to a realistic landslide model, characterized by highly contrasted layers and strong lateral velocity variations.

A two-level preconditioning strategy, based on simultaneous multifrequency inversion of damped data, has been applied to mitigate difficulties inherent in classical single-frequency inver-
sions. Results confirm that simultaneous inversion of damped data, which allows a progressive introduction of converted and free surface waves, significantly outperforms the successive single-frequency inversion approach. It is a useful solution to miti-
gate strong nonlinearities resulting from surface waves and to avoid convergence toward a local minimum. We have also emphasized the high potential of the L-BFGS optimization method to improve the convergence level significantly, compared to the more classical PCG algorithm.

Finally, we have addressed the impact of some recording parameters on medium reconstruction. We have shown that restricting the inversion to the vertical component data can lead to acceptable results in terms of imaging resolution and convergence level, with a percentage of model degradation of 24%.

We have also illustrated how poor model illumination is translated in terms of imaging resolution. Future work will tackle the construction of the initial model, a key issue for FWI before considering applications to real data.

**ACKNOWLEDGMENTS**

The study was funded by BRGM and by Agence Nationale de la Recherche (ANR), project ANR-05-NT05-2-42427. The MUMPS software package was used for linear system solving and Triangle software was used for mesh generation. We are grateful to M. Dietrich and four anonymous reviewers for their comments and suggestions on this paper.


Cerveny, V., I. A. Molotkov, and I. Pliščenk, 1977, Ray theory in seismo-
logic: Charles University Press.


Ellsènes, K. J., 2009, A comparison of phase inversion and traveltime to-
mography for processing near-surface reflection traveltimes: Geophysi-
s, 74, no. 6, WCB11–WCB24, doi:10.1190/1.3196857.

Ellefsen, K. J., 2009, A comparison of phase inversion and traveltime to-
mography for processing near-surface reflection traveltimes: Geophysi-
s, 74, no. 6, WCB11–WCB24, doi:10.1190/1.3196857.


Géliès, C., J. Virieux, and G. Grandjean, 2007, Two-dimensional elastic wave inversion using Born and Rytov formulations in the fre-


Grandjean, G., C. Pennetier, A. Bitri, O. Méric, and J. Malet, 2006, Carac-
térisation de la structure interne et de l’état hydrique de glissements argilo-marnieux par tomographie géophysique: l’exemple du glissement-coulée de Super-Sauce (Alpes du Sud, France): Comptes Rendus Ge-


Grandjean, G., and S. Sage, 2004, JtTS: A fully portable seismic tomogra-
phy software based on Fresnel wavepaths and a probabilistic reconstruc-
tion approach: Computers & Geosciences, 30, 925–935, doi:10.1016/
j.angeo.2004.06.009.


Elastic FWI for shallow structures


