19200

Toward a Better Full Waveform Inversion of Surface Waves

I. Masoni (Total E&P / Joseph Fournier University), R. Valensi* (IFFSTAR), R. Brossier (Joseph Fournier University), J. Virieux (Joseph Fournier University), J.L. Boelle (Total E&P), D. Leparoux (IFFSTAR) & P. Cote (IFFSTAR)

SUMMARY

This study presents advances in methodology for using seismic Full Waveform Inversion (FWI) for near surface characterization. By exploiting the full (visco-)elastic waveform and including surface waves in an FWI scheme, the classical difference-based misfit in the (t,x) domain is no longer sufficiently robust to handle the non-linearities due to the initial model definition and associated cycle-skipping problem. Alternative misfits based on the difference or cross-correlation of the data in alternative domains are therefore investigated using a grid analysis, to obtain a more convex valley of attraction and avoid the presence of local minima. Furthermore, in the case of available multi-component data, the use of new observables for the misfit function, such as the vertical over horizontal component spectral ratio (V/H ratio), may provide different information that can be significant for robust amplitude independent waveform inversion. The grid analysis results, obtained using simple 1D layered models, give promising perspectives for using robust FWI to image near-surface targets.

Introduction

Near-surface characterization by seismic methods is a challenging problem when complex targets are tackled. Most of the classical seismic methods are based on a simple velocity distribution assumption, such as a (locally) layered Earth for dispersion analysis of surface waves (MASW) or an increasing velocity with depth for refraction seismic. In such cases, lateral variations, velocity inversion or strong heterogeneities push classical seismic methods toward their limits of applicability and reliability. To overcome these limitations, the Full Waveform Inversion (FWI) method has recently been investigated to avoid any assumptions on subsurface complexities. Using FWI for near-surface characterization can be tackled by: (1) early waveform tomography using either P or SH waves to improve resolution compared to classical ray-based tomography (Ellefsen, 2009; Smithyman et al., 2009); (2) exploiting the full (visco-)elastic waveform including surface-waves (Gélis et al., 2007; Romdhane et al., 2011; Bretaudeau et al., 2013). Both strategies show promising results although FWI remains a difficult tool to manage, and in particular at the near-surface scale there may be difficulties due to the strong non-linearities of the inverse problem. The non-linearity of FWI is mainly due to the initial model definition and associated cycle-skipping issue, and to the amplitude modeling which is biased with 2D modeling or when an inaccurate attenuation model is used. These two effects are enhanced when using the classical least-square minimization of the data difference.

This study focused on the methodology, presents how alternative misfit functions may improve the convexity of the inverse problem, and therefore allow the use of less accurate initial models or amplitude modeling. These misfit functions are based on difference or cross-correlation of the data in alternative domains for vertical geophone data, and also take benefit of the vectorial property of the wavefield for multi-component data.

Classical FWI

FWI is a high-resolution technique used to derive quantitative models of the subsurface by matching the full observed seismogram with a corresponding synthetic seismogram calculated from an estimated velocity model, and solving a local optimization problem. The L_2 norm of the difference is conventionally used to calculate the misfit (Tarantola, 1984), fitting both the amplitude and phase of the waveforms:

$$C_{diff} = \sum_t \sum_x \frac{1}{2} \left(\mathbf{d_{obs}}(t,x) - \mathbf{d_{cal}}(t,x) \right)^2, \tag{1}$$

where $\mathbf{d_{obs}}(t,x)$ is the measured data (vertical or horizontal particle velocity, or both) and $\mathbf{d_{cal}}(t,x)$ is the calculated data, both recorded at time t and offset x. As the misfit function is minimized in a least-squares sense, the model is iteratively updated with a gradient-based descent method until a minimum is reached (Virieux and Operto, 2009). By using a strict data-matching approach, this method may not be sufficiently robust. Non-linearities, such as cycle skipping, can reduce the convexity of the misfit function (Bunks et al., 1995; Mulder and Plessix, 2008) and the minimization may get stuck in a local minimum. In the absence of very low-frequency data, the initial velocity model needs to explain the data to within half a wavelength, so that it lies within the small basin of attraction of the global minimum and can converge. When exploiting slow surface waves propagating in the low velocity near surface, the problem of cycle-skipping is even greater due to their small wavelengths.

Alternative misfit functions

Current solutions to calculating the misfit more robustly are based on other norms such as the hybrid L_1/L_2 or Huber norm (Brossier et al., 2010; Guitton and Symes, 2003) or on zero-lag cross-correlation (Routh et al., 2011), but these also suffer from cycle-skipping.

One way to avoid the strong dependency to amplitude and cycle-skipping would be to use a weighted cross-correlation as proposed by van Leeuwen and Mulder (2008). The misfit is calculated by a cross correlation on the time axis of the observed and calculated data.

$$C_{xcorr} = \sum_{\Delta t} \sum_x \frac{1}{2} \left(W(\Delta t) \sum_t \mathbf{d_{obs}}(t+\Delta t,x) \mathbf{d_{cal}}(t,x) \right)^2. \tag{2}$$

The weighting $W(\Delta t) = e^{-\alpha \Delta t^2}$ is applied to each time sample to maximize zero-lag energy, giving a misfit function whose negative is minimized. An appropriate width controlled by the α parameter needs to be chosen to be within the order of the length of the wavelet, since it can greatly influence the convexity of the misfit function (van Leeuwen and Mulder, 2008).

The data domain in which observed and calculated datasets are compared also affects the sensitivity of the misfit function. The (τ, p) domain or "slant-stack" may provide a better behavior of the misfit, projecting the data to plane waves. Both the difference-based L_2 norm (Eq 3) and a weighted cross-correlation applied on the slowness p axis (Eq 4) can be considered:

$$C_{diff} = \sum_{\tau} \sum_{p} \frac{1}{2} \left(\mathbf{d_{obs}}(\tau, p) - \mathbf{d_{cal}}(\tau, p) \right)^2, \tag{3}$$

$$C_{xcorr} = \sum_{\tau} \sum_{\Delta p} \frac{1}{2} \left(W(\Delta p) \sum_{p} \mathbf{d_{obs}}(\tau, p + \Delta p) \mathbf{d_{cal}}(\tau, p) \right)^2. \tag{4}$$

Other domains such as the (ω, p) or (ω, k) domain can also be investigated, applying a difference-based or correlation-based misfit (Perez Solano et al., 2012). A parallel may be drawn to MASW methods, which make use of data in the frequency domain to invert dispersion curves (Forbriger, 2003).

A grid analysis is carried out on these misfit functions, all tested on a canonical 1D case typical of experimental setups used for onshore depth exploration. Synthetic datasets were created using a discrete wavenumber summation method (Dietrich, 1988), for horizontally layered media with a free surface, and simulating 3D elastic wave propagation with a Ricker wavelet source of 10 Hz peak frequency. Figure 1a shows the two-layer model used to create the reference dataset to which random Gaussian noise is added. A grid analysis is performed on the shear velocity (V_s) and the depth of the first layer to investigate the accuracy required for the initial model to converge to the global minimum.

The results in Figure 2 show that even for this simple framework and only small shifts in the model parameters the classical difference-based L_2 norm misfit (Figure 2a) contains many local minima due to the high amplitude of the surface waves that dominate the misfit. However, the difference-based L_2 norm becomes more convex in all alternative data domains tested as shown by Figures 2b-2d. Results are especially successful in the (ω, k) and (ω, p) domains. The frequency-domain combined with slope information appears to efficiently mitigate non-linearities of the problem. Weighted cross-correlation also appears to improve the convexity of the inverse problem, even in the time domain (Figure 2e-2h).

This test shows that the classical difference-based misfit in the (t, x) domain is clearly not sufficiently robust to handle surface wave inversion for near-surface targets, and alternative misfits in other domains should significantly improve the behavior of FWI scheme.

Handling multi-component data with the V/H ratio

The classical formalism of multicomponent data in FWI only considers a summation over each component in Eq 1. However, several observables can be extracted from multicomponent data sets. In this study, we will illustrate the advantages of using the vertical over horizontal component spectral ratio (V/H ratio) in an FWI framework. Boore and Nafi Toksoz (1969) first showed that the ellipticity and the phase velocity observables are complimentary. The H/V ratio has now become a widely used tool to study site effects with microtremor data, and more recently, there have been several attempts to combine the inversion of the surface wave dispersion curves with the inversion of the H/V ratio (Arai and Tokimatsu, 2005). In the context of the FWI method, the polarization observables may be of interest for several reasons: unknown source excitation effects may be removed, multicomponent data can be introduced in an established framework and these observables are sensitive to velocity contrasts in the shallow part of the media (Tanimoto and Rivera, 2008). In an inversion framework, it also has the advantage to be less sensitive to the biases generated by 2D/3D modeling effects or seismic attenuation. In this case study the V/H ratio is preferred to the H/V ratio to avoid potential peaks (Tanimoto and Rivera, 2008).

To illustrate the properties of this observable, another grid analysis test on a 1D layered configuration is performed by varying the parameters V_s, V_p and depth of the interface. Two parameters are varied independently: the V_s parameter, related to V_p with a constant V_p/V_s ratio equal to 2.19, and the depth of the layer interface. The reference model is composed of two layers with a V_s in the top layer of 520 m/s, and an interface at 5 m depth (Figure 1b). Compared to the previous grid analysis, the experimental setup is more similar to engineering scale experiments where typically MASW methods are used. The far offset is 60 m, the source signal is a Ricker wavelet with a central frequency of 60 Hz and data are analyzed in the $[40, 120]$ Hz frequency band. Moreover, noise has been added to the reference data, such that the signal to noise ratio is 15 dB for the horizontal component and 20 dB for the vertical component.

Phase-difference based misfit functions are less sensitive to amplitude errors in modeling and have already been used by Ellefsen (2009) for near surface imaging when considering only early arrival waveforms. The grid analysis is carried out with a phase-difference misfit function using different components (Figures 3a-3c), all showing a similar behavior: the occurrence of local minima and a very small basin of attraction. However, the V/H-based misfit function (Figure 3d) appears to contain less local minima and to be more convex, being sensitive to the deeper parts of the model. This result shows that the V/H observable seems to provide different information than phase observables (Boore and Nafi Toksoz, 1969) and might be significant for robust amplitude independent waveform inversion.

Conclusions

This study on methodology shows with tests on simple layered configurations, that the classical usage of the difference-based misfit function for single or multi-component data appears to be highly sensitive to the starting model accuracy to ensure convergence to the global minimum for highly non-convex problems. The use of local slope, through alternative domains to measure the misfit function, gives a better well-posed problem and should ensure a more robust behavior of FWI when applied to near surface single component data containing surface waves. Moreover, when multi-component data are available, it appears that a new observable derived from the vectorial nature of the wavefield also allows to improve the well-posedness of the problem, rather than a simple summation over components. This study gives promising perspectives for using robust FWI to image near-surface targets.

Acknowledgements

The authors would like to thank the SEISCOPE consortium sponsors and TOTAL E&P for funding and permission to show these results. This work was produced using the CIMENT high-performance computing facilities (Université Joseph Fourier, Grenoble).

References

Arai, H. and Tokimatsu, K. [2005] S-wave velocity profiling by joint inversion of microtremor dispersion curve and horizontal-to-vertical (h/v) spectrum. *Bulletin of the Seismological Society of America*, **95**(5), 1766–1778.

Boore, D.M. and Nafi Toksoz, M. [1969] Rayleigh wave particle motion and crustal structure. *Bulletin of the Seismological Society of America*, **59**(1), 331–346.

Bretaudeau, F., Brossier, R., Leparoux, D., Abraham, O. and Virieux, J. [2013] 2D elastic full waveform imaging of the near surface: Application to a physical scale model. *Near Surface Geophysics*, **11**(3), 307–316.

Brossier, R., Operto, S. and Virieux, J. [2010] Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics*, **75**(3), R37–R46.

Bunks, C., Salek, F.M., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473.

Dietrich, M. [1988] Modeling of marine seismic profiles in the t-x and tau-p domains. *GEOPHYSICS*, **53**(4), 453–465.

Ellefsen, K.J. [2009] A comparison of phase inversion and traveltime tomography for processing near-surface refraction traveltimes. *Geophysics*, **74**(6), WCB11–WCB24.

Forbriger, T. [2003] Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms. *Geophysical Journal International*, **153**(3), 735–752.

Gélis, C., Virieux, J. and Grandjean, G. [2007] 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain. *Geophysical Journal International*, **168**, 605–633.

Guitton, A. and Symes, W.W. [2003] Robust inversion of seismic data using the Huber norm. *Geophysics*, **68**(4), 1310–1319.

Mulder, W. and Plessix, R.E. [2008] Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting*, **56**(6), 827–841.

Perez Solano, C., Donno, D. and Chauris, H. [2012] Alternative objective function for inversion of surface waves in 2D media. *Expanded Abstracts, 18th European Meeting of Environmental and Engineering Geophysics*, Near Surface Geoscience, A21.

Romdhane, A., Grandjean, G., Brossier, R., Réjiba, F., Operto, S. and Virieux, J. [2011] Shallow structures characterization by 2d elastic waveform inversion. *Geophysics*, **76**(3), R81.

Routh, P. et al. [2011] Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data. *SEG Technical Program Expanded Abstracts*, **30**(1), 2433–2438.

Smithyman, B., Pratt, R.G., Hayles, J. and Wittebolle, R. [2009] Detecting near-surface objects with seismic waveform tomography. *Geophysics*, **74**(6), WCC119–WCC127.

Tanimoto, T. and Rivera, L. [2008] The zh ratio method for long-period seismic data: sensitivity kernels and observational techniques. *Geophysical Journal International*, **172**(1), 187–198.

Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266.

van Leeuwen, T. and Mulder, W. [2008] Velocity analysis based on data correlation. *Geophysical Prospecting*, **56**(6), 791–803, ISSN 1365-2478.

Virieux, J. and Operto, S. [2009] An overview of full waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC127–WCC152.

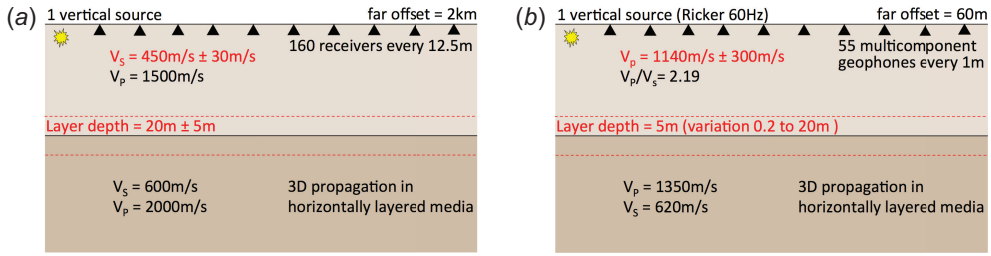

Figure 1 *Schematic of model used to test alternative misfit functions based on cross-correlation or different data domains (a); and model used to test multicomponent data misfit functions (b).*

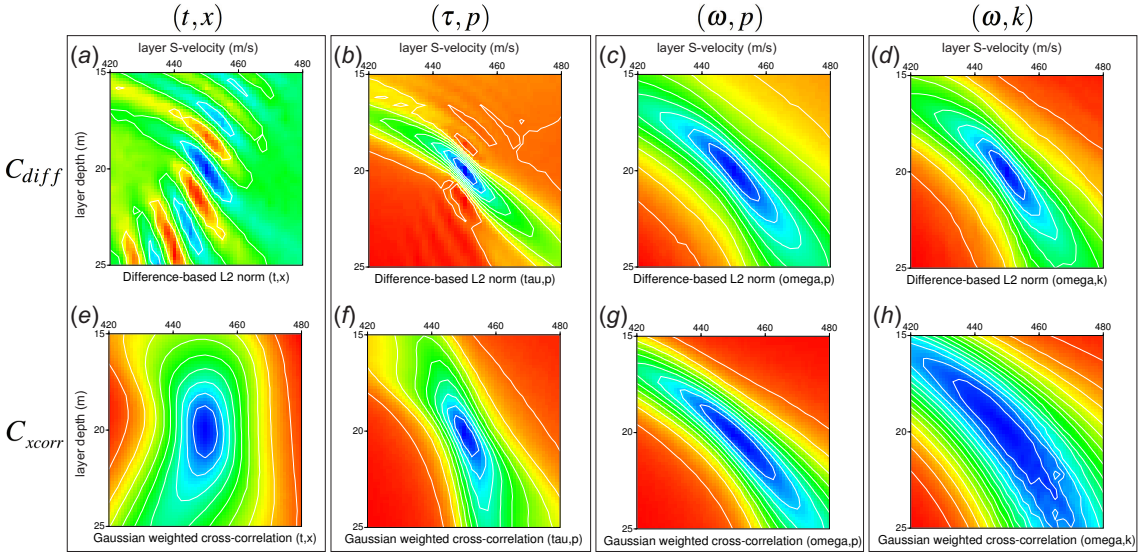

Figure 2 *Two-parameter grid search results using classical and alternative misfit functions, for a reference dataset with the true global minimum at* 450 *m/s layer S-velocity and* 20 *m layer depth.*

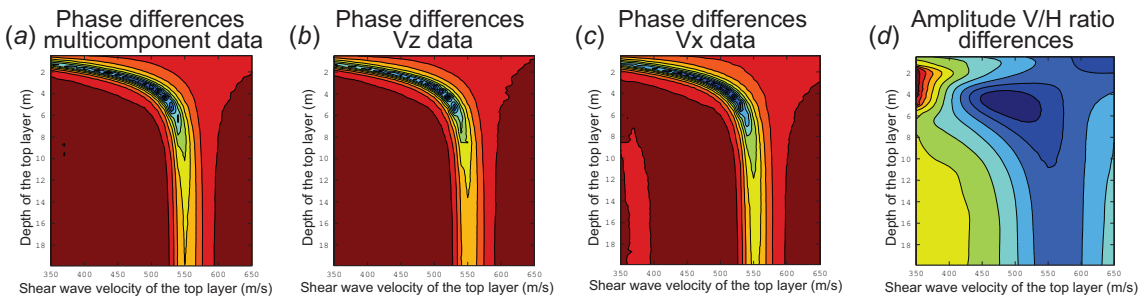

Figure 3 *Two-parameter grid search results for multicomponent data misfit functions, for a reference dataset with the true global minimum at* 520 *m/s layer S-velocity and* 5 *m layer depth.*