Center for Wave Phenomena

Expanded abstracts submitted by CWP to the

EAGE 74th Conference and Exhibition
Copenhagen, Denmark
June 4-7, 2012

and

SEG International Exhibition and
82nd Annual Meeting
Las Vegas, Nevada
November 4-9, 2012

www.cwp.mines.edu
## Contents - Expanded abstract submissions

### EAGE 74th Conference and Exhibition
Copenhagen, Denmark - June 4-7, 2012

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation of interval velocity and attenuation anisotropy from reflection data at the Coronation Field</td>
<td>Jyoti Behura and Ilya Tsvankin, CWP, Edward Jenner, ION Geophysical and Alex Calvert, ION Geophysical and Maersk Oil</td>
<td>1</td>
</tr>
<tr>
<td>Adjoint-state reverse-time migration for 4C seismic imaging</td>
<td>Clément Fleury, CWP and Ivan Vasconcelos, Schlumberger Cambridge Research</td>
<td>7</td>
</tr>
<tr>
<td>Increasing illumination and sensitivity of reverse-time migration with internal multiples</td>
<td>Clément Fleury and Roel Snieder, CWP</td>
<td>11</td>
</tr>
<tr>
<td>Wavefield tomography based on local image correlations</td>
<td>Francesco Perrone and Paul Sava, CWP, Clara Andreoletti and Nicola Bienati, Eni E&amp;P S.p.A.</td>
<td>15</td>
</tr>
<tr>
<td>Migration velocity analysis in the presence of quadratic lateral velocity variation</td>
<td>Mamoru Takanashi, JOGMEC/CWP and Ilya Tsvankin, CWP</td>
<td>19</td>
</tr>
<tr>
<td>Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination</td>
<td>Kees Wapenaar, Jan Thorbecke and Joost van der Neut, Delft University of Technology, Filippo Broggini and Roel Snieder, CWP</td>
<td>23</td>
</tr>
<tr>
<td>Illumination compensation for image-domain wavefield tomography</td>
<td>Tongning Yang, CWP, Jeffrey Shragge, University of Western Australia, and Paul Sava, CWP</td>
<td>29</td>
</tr>
</tbody>
</table>

### SEG International Exhibition and 82nd Annual Meeting
Las Vegas, Nevada - November 4-9, 2012

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A practical approach to internal multiples and ghosts</td>
<td>Jyoti Behura, CWP, Farnoush Forghani-Arani, Colorado School of Mines</td>
<td>33</td>
</tr>
<tr>
<td>Newton-Marchenko-Rose Imaging: Image reconstruction based on inverse scattering</td>
<td>Jyoti Behura, CWP, Kees Wapenaar, Delft University of Technology and Roel Snieder, CWP</td>
<td>39</td>
</tr>
</tbody>
</table>
Creating a virtual source inside a medium from reflection data with internal multiples: a stationary-phase analysis
Filippo Broggini and Roel Snieder, CWP and Kees Wapenaar, Delft University of Technology

Focusing inside an unknown medium using reflection data with internal multiples: numerical examples for a laterally-varying velocity model, spatially-extended virtual source, and inaccurate direct arrivals
Filippo Broggini and Roel Snieder, CWP and Kees Wapenaar, Delft University of Technology

Joint migration velocity analysis of PP- and PS-waves for VTI media
Pengfei Cai and Ilya Tsvankin, CWP

Understanding the reverse time migration backscattering: Noise or signal?
Esteban Díaz and Paul Sava, CWP

Migration velocity analysis for nonlinear reverse-time migration
Clément Fleury, CWP

Bi-objective optimization for the inversion of seismic reflection data: Combined FWI and MVA
Clément Fleury and Francesco Perrone, CWP

Fault surfaces and fault throws from 3D seismic images
Dave Hale, CWP

Full-waveform inversion of multicomponent data for layered VTI media
Nishant Kamath and Ilya Tsvankin, CWP

Estimation of time-lapse velocity change using repeating earthquakes with different locations and focal mechanisms
Chinaemerem Kanu and Roel Snieder, CWP and Dan O’Connell, Fugro Consultants, Inc.

Unfaulting and unfolding 3D seismic images
Simon Luo and Dave Hale, CWP

Application of image-guided full waveform inversion to a 2D ocean-bottom cable data set
Yong Ma, CWP, Jianxin (Jerry) Yuan and Yunqing Shen, Seismic Technology, ConocoPhillips Company and Bin Gong, In-Depth Geophysical

Time-lapse change in near-surface shear-wave velocities caused by rainfall and large earthquakes detected by applying seismic interferometry to earthquake data
Nori Nakata and Roel Snieder, CWP

Wavefield tomography based on local image correlations
Francesco Perrone and Paul Sava, CWP, Clara Andreoletti and Nicola Bienati, Eni E&P S.p.A.

Anisotropy signature in extended images from reverse-time migration
Paul Sava, CWP and Tariq Alkhalifah, King Abdullah University of Science and Technology
Attenuation analysis for heterogeneous transversely isotropic media
Bharath Shekar and Ilya Tsvankin, CWP

Time-lapse image-domain velocity analysis using adjoint-state methods
Jeffrey Shragge, University of Western Australia, Tongning Yang and Paul Sava, CWP

Feasibility of inverting compaction-induced travelt ime shifts for reservoir pressure
Steven Smith and Ilya Tsvankin, CWP

Multiparameter TTI tomography of P-wave reflection and VSP data
Xiaoxiang Wang and Ilya Tsvankin, CWP

Integrated migration and internal multiple elimination
Kees Wapenaar, Jan Thorbecke, Joost van der Neut and Evert Slob, Delft University of Technology, Filippo Broggini, Jyoti Behura and Roel Snieder, CWP

Pitfalls in constraining attenuation from ambient seismic noise
Cornelis Weemstra, Spectraseis AG and ETH Zürich, Lapo Boschi, ETH Zürich, Roel Snieder, CWP and Alex Goertz, ETH Zürich

Illumination compensation for subsalt image-domain wavefield tomography
Tongning Yang, CWP, Jeff Shragge, University of Western Australia and Paul Sava, CWP
Estimation of interval velocity and attenuation anisotropy from reflection data at the Coronation Field
J. Behura, I. Tsvankin, E. Jenner & A. Calvert

Abstract
Attenuation can be extremely valuable in characterizing gas accumulations in shales and sands. In fractured reservoirs, anisotropy can provide additional information about the distribution of fractures. Here, we apply a layer-stripping approach to wide-azimuth P-wave data acquired over a gas reservoir in the Coronation Field, Alberta. The main processing steps involve estimation of traveltimes from the top and bottom of the target layer followed by computation of the interval NMO velocity and attenuation coefficients using velocity-independent layer stripping and the spectral-ratio method. The vertical attenuation coefficient shows a reasonable correlation with existing gas-producing well locations, and, therefore, serves as an indicator of gas accumulation in pore space. Based on the attenuation information, we conclude that the lower half of the survey area has significant gas reserves. In these areas, the estimated azimuthal velocity anisotropy can be used to plan horizontal wells oriented orthogonal to the fracture direction.

Introduction
Laboratory studies clearly indicate that attenuation is closely related to fluid saturation and mobility (Gautam et al., 2003). Well-log analysis shows that P-wave attenuation is higher in gas-bearing rocks than in those saturated with oil or water. Extremely low P-wave quality factors (Q), ranging between 5 and 10, have been observed in some gas reservoirs. Attenuation anisotropy might carry valuable information about fluid- or gas-filled fractures. In fact, preferential flow of fluids in rocks is believed to be the primary cause of attenuation anisotropy.

Behura and Tsvankin (2009a) introduce a layer-stripping technique to extract interval attenuation from reflection data using a variation of the spectral-ratio method. While no information about velocity and attenuation anywhere in the medium is required, the overburden has to be laterally homogeneous with a horizontal symmetry plane. Using synthetic examples for vertical transverse isotropy (VTI) and orthorhombic models, Behura and Tsvankin (2009a) demonstrate that this algorithm can successfully estimate interval anisotropic attenuation in 2D and 3D. Here, we apply this technique to wide-azimuth data acquired at Coronation Field, Alberta and investigate the spatial distribution of gas accumulation and fracturing using the obtained velocity and attenuation fields.

Data Acquisition and Processing
The Coronation Field is a gas reservoir located in East Central Alberta. Light oil and gas are usually trapped in numerous fluvial and valley-fill reservoir sandstones, where the sand channels vary in thickness from 4 m to 10 m and in width between 200 m and 300 m. Because of the high cost of developing this field, drilling success is critical. To understand the lithology of the channel sands and help optimize well placement, Apache Corporation acquired a 3D multicomponent seismic survey (now owned by VGS Seismic Canada Inc.). In this study, we use only the vertical-component data to estimate the P-wave velocity and attenuation fields. The entire survey was shot using single-hole dynamite on a “shoot and roll” template of 12 × 94 receivers (Table 1) with the source at the center of the patch. This shooting template is roughly square yielding an excellent azimuthal and offset distribution. The shot lines are oblique to the receiver lines. The subsurface structure is fairly close to layer-cake (no structural dip), which facilitates the layer-stripping technique described in Behura and Tsvankin (2009a).

Prior to attenuation analysis, refraction statics corrections were applied to the data, which were shifted to a smooth floating datum. Some traces were subsequently edited to remove spikes and random noise. Denoising was followed by three passes of residual statics corrections. The ground roll was suppressed using f-k filtering followed by surface-consistent median gain applied to account for the variation in the dynamite source strength.
The most critical step in estimating interval parameters is to pick traveltimes of reflections from the top and bottom of the target layer which can be done using either an auto-picker or semblance analysis. Here, we use a 3D nonhyperbolic semblance algorithm for orthorhombic media based on the generalized Alkhalifah-Tsvankin (Alkhalifah and Tsvankin, 1995) equation:

$$t^2(x, \alpha) = t_0^2 + \frac{x^2}{V_{nmo}^2(\alpha)} - \frac{2\eta(\alpha)x^4}{V_{nmo}^2(\alpha)[t_0^2 V_{nmo}^2(\alpha) + (1 + 2\eta(\alpha))x^2]},$$

where $t$ is the reflected traveltime, $x$ is the offset, $t_0$ is the two-way zero-offset traveltime, $\alpha$ is the source-to-receiver azimuth, $V_{nmo}(\alpha)$ is the azimuthally-varying normal-moveout velocity, and $\eta(\alpha)$ is the “anellipticity” coefficient responsible for the deviation from hyperbolic moveout at long offsets. The velocity $V_{nmo}(\alpha)$ is obtained from the equation of the NMO ellipse:

$$V_{nmo}^{-2}(\alpha) = \frac{\sin^2(\alpha - \varphi)}{V_{nmo}^{(1)}(\alpha)^2} + \frac{\cos^2(\alpha - \varphi)}{V_{nmo}^{(2)}(\alpha)^2},$$

$\varphi$ is the azimuth of the $[x_1, x_3]$ symmetry plane, and $V_{nmo}^{(1)}$ and $V_{nmo}^{(2)}$ are the NMO velocities in the vertical symmetry planes $[x_2, x_3]$ and $[x_1, x_3]$, respectively. The parameter $\eta$ is approximately given by (Xu and Tsvankin, 2006):

$$\eta(\alpha) = \eta^{(1)}\sin^2(\alpha - \varphi) + \eta^{(2)}\cos^2(\alpha - \varphi) - \eta^{(3)}\sin^2(\alpha - \varphi)\cos^2(\alpha - \varphi),$$

where $\eta^{(1)}$, $\eta^{(2)}$, and $\eta^{(3)}$ are the anellipticity coefficients defined in the $[x_2, x_3]$, $[x_1, x_3]$, and $[x_1, x_2]$ symmetry planes, respectively. The semblance algorithm estimates the parameters of the NMO ellipse ($V_{nmo}^{(1,2)}$ and $\varphi$), and the anellipticity parameters $\eta^{(1,2,3)}$. By assuming not just the overburden, but also the target layer to be laterally homogeneous, we compute the best-fit effective moveout parameters for the source as well as receiver gathers. The wide-azimuth reflection traveltimes for the top and bottom of the target layer are then used for estimating the interval traveltime, as described by Dewangan and Tsvankin (2006) in 2D and Wang and Tsvankin (2009) for 3D wide-azimuth data.

The interval traveltime and windowed events along the moveout curves serve as the input data to estimate the interval attenuation coefficient using the technique of Behura and Tsvankin (2009a). Assuming velocity anisotropy to have orthorhombic symmetry, we invert for the moveout parameters by fitting equation 1 to the interval traveltimes. The attenuation-anisotropy parameters are estimated by fitting the P-wave phase attenuation function $A_P$ in orthorhombic media (Zhu and Tsvankin, 2007) to the estimated interval attenuation coefficients:

$$A_P(\theta, \phi) = A_{P0} \left[1 + \delta_Q(\phi)\sin^2 \theta \cos^2 \theta + \epsilon_Q(\phi) \sin^4 \theta\right],$$

where $A_{P0}$ is the vertical attenuation coefficient, $\theta$ and $\phi$ are the polar and azimuthal phase angles, respec-
Estimation of interval velocity and attenuation anisotropy from reflection data at the Coronation Field

J. Behura, I. Tsvankin, E. Jenner & A. Calvert

3

Area, where attenuation anomalies should correspond to gas accumulations.

Optimal strategy would be to drill horizontal wells orthogonal to the fractures in the lower half of the survey area. This would also imply that in the areas showing strong azimuthal anisotropy, the target layer is more intensively fractured. For increased production from fractured formations, horizontal wells should be drilled orthogonal to the fracture strike. In the presence of more than one fracture set, however, interpretation of NMO ellipses becomes more complicated. Additional data such as well-bore resistivity images could help govern the angular variation of the P-wave attenuation coefficient (Zhu and Tsvankin, 2007). We do not assume the same orientation for attenuation as estimated for velocity anisotropy.

Although equation 4 describes the angular variation of the phase attenuation coefficient $A_0$ estimated from seismic data because $A_0 = \frac{\rho_0}{\epsilon_0}$ (Behura and Tsvankin, 2009b). As discussed by Behura and Tsvankin (2009a), inversion for the attenuation-anisotropy parameters requires knowledge of the anisotropic velocity field which can be used to compute the phase angle from the measured group angle. We use a well-log derived interval vertical P-wave velocity $V_{1P}$ of 5000 m/s (Monk et al., 2006) to estimate the velocity anisotropy parameters $\delta_{1,2,3}$ and $\epsilon$ from the moveout parameters. Using the obtained velocity model, we compute the phase angles from the group angles. Note that it is possible to perform velocity analysis prior to attenuation processing using the interval traveltime because the influence of attenuation on velocity for a fixed frequency typically is of the second order (Behura and Tsvankin, 2009b).

Results

Here, we discuss the results of azimuthal velocity and attenuation analysis for an interval starting at about 0.8 s which contains the reservoir. The interval NMO ellipse of the target layer is represented by its eccentricity and the azimuth $\phi_{nmo}$ of the major axis (Figure 1a) for each common midpoint. The NMO eccentricity is defined as $|V_{nmo}^{(1)} - V_{nmo}^{(2)}|/V_{nmo}^{fast}$ where $V_{nmo}^{fast} = \max\{V_{nmo}^{(1)}, V_{nmo}^{(2)}\}$. Azimuthal velocity anisotropy within the target layer varies smoothly throughout the field and is fairly large over most of it. If the field had only one dominant fracture set, the major axis of the NMO ellipse would be parallel to the fractures. This would also imply that in the areas showing strong azimuthal anisotropy, the target layer is more intensely fractured. For increased production from fractured formations, horizontal wells should be drilled orthogonal to the fracture strike. In the presence of more than one fracture set, however, interpretation of NMO ellipses becomes more complicated. Additional data such as well-bore resistivity images could help resolve this ambiguity.

The vertical attenuation coefficient $A_0$ is displayed in Figure 1b. Zones of higher attenuation (hot colors) in the lower half of the survey area show good correlation with locations of existing gas-producing wells (Monk et al., 2006). All the producing wells overlie areas showing high attenuation; on the other hand, two of the dry wells lie on the edge of the high attenuation zone and one of them lies in an area showing extremely low attenuation (Figure 1b). Therefore, $A_0$ could possibly be used as a reliable indicator of gas distribution in the field. In particular, the area in the vicinity of $x = 9$ km, $y = 3$ km possibly has significant gas reserves. The coefficient $A_0$ should be sensitive to gas present in pores but not in vertical fractures; it could, therefore, be used as an indicator of gas accumulation within the pore space. It is interesting to note that the quality factor of these gas sands can be as low as 5 (close to the values reported by Mavko and Dvorkin (2005)). Such small $Q$-values are especially remarkable because $A_0$ corresponds to the effective attenuation of the whole target layer that includes lithologies other than gas sands.

The inversion for the attenuation-anisotropy parameters $\delta_{1,2,3}$ and $\epsilon$ appears to be ill-constrained because of the sensitivity of amplitude analysis to noise. This is evident from the lack of large-scale spatially coherent patterns of these parameters within the survey area.

Well planning should take both interval velocity anisotropy and the attenuation coefficient into account. The optimal strategy would be to drill horizontal wells orthogonal to the fractures in the lower half of the survey area, where attenuation anomalies should correspond to gas accumulations.
Figure 1: (a) Eccentricity (tick lengths) and azimuth $\varphi^{fast}$ (tick orientations) of the interval NMO ellipses in the target layer. The maximum value of the NMO eccentricity for the target layer is 0.2. (b) Interval vertical attenuation coefficient $A_0 = 1/(2Q_0)$ for the target layer. Location of gas-producing (circles with eight lines) and dry wells (circles with four lines) in the Coronation Field (modified from Monk et al., 2006) are also marked.
Discussion and conclusion

P-wave anisotropic velocity and attenuation analysis provides valuable information for reservoir characterization at the Coronation Field. The interval velocity and attenuation within a subsurface zone containing the reservoir sands were estimated using the layer-stripping approach introduced by Dewangan and Tsvankin (2006) and Behura and Tsvankin (2009a). Nonhyperbolic semblance analysis yielded the wide-azimuth traveltimes for reflections from the top and bottom of the reservoir. Then velocity-independent layer-stripping was employed to generate interval traveltime surfaces, which were inverted for the interval velocity-anisotropy parameters of orthorhombic media. The windowed events along the moveout surfaces were used for computing the interval anisotropic attenuation.

The lower half of the survey area shows strong azimuthal velocity anisotropy as well as high attenuation. Areas of large interval vertical attenuation coefficient $A_0$ correlate with locations of gas-producing wells, which makes $A_0$ a reliable indicator of gas sands in the field. In such high-attenuation areas, the orientation of the estimated NMO ellipses can be used to plan horizontal wells, which should be orthogonal to the predominant fracture strike.

Acknowledgments

We thank ION Geophysical and VGS Seismic for providing the field data and for permission to publish this work.

References


Introduction

In recent years, marine seismic acquisition has seen the introduction of dual-field acoustic systems. Seismic streamer technology is now able to combine pressure and three-component displacement measurements (Robertsson et al., 2008). These new acquisition developments lead to novel data-domain processing techniques for noise attenuation (Cambois et al., 2009), signal interpolation (Vassallo et al., 2010), 3D deghosting (Özbek et al., 2010), or multiple attenuation/separation (Frijlink et al., 2011). In parallel, novel dual-source technology combines both monopole and dipole marine sources (Robertsson et al., 2012). These multisource multicomponent acquisition systems directly impact the development of depth imaging methods. Vasconcelos (2011) and Vasconcelos et al. (2012) utilize the directional and dynamic information of dual-source dual-field acoustic data in a vector-acoustic reverse-time imaging method (VARTM) that handles ghost and multiple energy, but does not easily utilize directional information contained in horizontal gradient data. Horizontal gradient data contain useful additional information for imaging, as used in, e.g., map migration methods (Kleyn, 1977; Douma and Hoop, 2006). To make use of 3C gradient data in arbitrary acquisition configurations, we propose an alternative VARTM method for dual-sensor 4C data based on the adjoint-state formulation (e.g., Plessix, 2006).

Theory

Dual-field VARTM within the adjoint-state method framework

Novel acquisition systems provide 4C dual-field acoustic data (pressure and displacement fields) at receivers $\mathbf{x}_r$, i.e., $\mathbf{d} = (p^{rec}, \mathbf{u}^{rec})$. Novel source technology allows for the use of pressure and point-force sources, i.e., $\mathbf{s} = (q, \mathbf{f})$. State variable $\mathbf{w} = (p, \mathbf{u})$ and adjoint variable $\mathbf{w}^\dagger = (p^\dagger, \mathbf{u}^\dagger)$ are solutions of equations (1) and (2), respectively. For brevity, we do not expose the details of how to handle initial and boundary conditions and refer the reader to Liu and Tromp (2006). The state equations are

$$
\begin{bmatrix}
\mathbf{I} & \mathbf{k}^{-1} \nabla \\
\nabla & \rho \frac{\partial^2}{\partial t^2}
\end{bmatrix}
\begin{bmatrix}
p \\
\mathbf{u}
\end{bmatrix} =
\begin{bmatrix}
q \\
\mathbf{f}
\end{bmatrix},
$$

(1)

and

$$
\begin{bmatrix}
\mathbf{I} & -\nabla (\mathbf{k}^{-1}) \\
-\nabla (\mathbf{k}^{-1}) & \rho \frac{\partial^2}{\partial t^2}
\end{bmatrix}
\begin{bmatrix}
p^\dagger \\
\mathbf{u}^\dagger
\end{bmatrix} =
\begin{bmatrix}
q^\dagger \\
\mathbf{f}^\dagger
\end{bmatrix},
$$

(2)

where adjoint source $s^\dagger = (q^\dagger, \mathbf{f}^\dagger)$ is key to achieve successful adjoint VARTM and is function of misfit function $J$. The choice of $J$ is goal-dependent. For example, one may make use of the directional dynamic information contained in 4C data. We introduce weighting operators $\mathbf{W}_s$ and $\mathbf{W}_r$ to define

$$
J = \frac{1}{2} \sum_{s,r} \mathbf{W}_s \int_0^T \| \mathbf{W}_r [\mathbf{w}(\mathbf{x}_r, \mathbf{x}_s, t) - \mathbf{d}(\mathbf{x}_r, \mathbf{x}_s)] \|_2^2 dt
$$

(3)

and derive the corresponding adjoint source for shot-profile migration:

$$
s^\dagger(\mathbf{x}, \mathbf{x}_s, t) = \mathbf{W}_s \sum_r \mathbf{W}_r^\dagger \mathbf{W}_r [\mathbf{w}(\mathbf{x}_r, \mathbf{x}_s, T-t) - \mathbf{d}(\mathbf{x}_r, \mathbf{x}_s, T-t)] \delta(\mathbf{x} - \mathbf{x}_r).
$$

(4)

We define image $I$ as the gradient of objective function $J$ with respect to normalized change in impedance for constant velocity (Douma et al., 2010) and obtain

$$
I = \frac{1}{2} \sum_s \int_0^T p^\dagger(\mathbf{x}, \mathbf{x}_s, t) p(\mathbf{x}, \mathbf{x}_s, t) dt + \frac{1}{2} \sum_s \int_0^T \rho(\mathbf{x}) \mathbf{u}^\dagger(\mathbf{x}, \mathbf{x}_s, t) \cdot \frac{\partial^2}{\partial t^2} \mathbf{u}(\mathbf{x}, \mathbf{x}_s, t) dt,
$$

(5)

where the state and adjoint fields are solutions of the state equations (1) and (2).

Directionally balanced adjoint VARTM

4C dual-field acoustic data provide an exact description of the wave state at a particular space and time locus and contain directional dynamic information that is absent from conventional pressure or 2C seismic data. Map migration techniques have been investigated to incorporate such directional information.
(as slope estimate) into a ray-based migration process. The adjoint VARTM method appears as a natural finite-frequency extension of map migration. Our formulation allows for extrapolating finite-frequency seismic data with directionally balanced energy (see imaging impulse responses in Figure 1) by selecting the appropriate weighting operators. For receiver-directional balancing, we utilize 4C dual-field acoustic data and choose $W_r$ (equation (6)) so that $J$ has both consistent physical dimensions and balanced magnitudes. For source-directional balancing, we utilize dual-source configurations and choose $W_s$ (equation (7)) so that the different sources have normalized energy contribution. These choices yield:

$$W_r = \begin{bmatrix} \sqrt{\kappa} & 0 \\ 0 & \sqrt{\rho} \frac{\partial}{\partial t} \end{bmatrix}$$

$$W_s = E_{q,t}^{-1} I,$$

where $E_{q,t}$ are energies radiated by pressure source $s_q$ and point-force source $s_f$.

Examples

Advantages of directional balancing for marine seismic imaging

A direct application of directional balancing is deghosting and dealiasing. Marine data typically suffer from what is commonly referred to as “ghosts” generated at the water free surface. Marine data usually subsample wavefield recordings in the shot direction leading to aliasing effects. Our reverse-time map migration method reduces both aliasing artifacts and receiver-side ghosts in the wavefield extrapolation process which leads to reduced artifacts in the image space (first row in Figure 2). Without additional data processing, receiver-directional balancing allows for focusing the ghost energy at the point scatterer that reduces the ambiguity on its location. We push the experiment further by assuming that data have been collected from both pressure and vertical point-force sources (second row in Figure 2). With this dual-source configuration, source-directional balancing leads to better resolution of the scatterer by attenuating source-side ghost energy. We note that these improvements are further enhanced while using 3C displacement fields instead of only 1C vertical displacement fields.

2D SEAM model synthetic example

We apply our weighted adjoint VARTM method to a 2D section (inline 1199) of the SEG Advanced Modeling (SEAM) Project model. We generate the synthetic data using fixed receivers at 50 m depth with 10 m spacing, sources at 30 m depth with 130 m spacing and a Ricker wavelet source pulse with 10 Hz peak frequency. We focus our analysis on the salt-flank area in Figure 3 for its stratigraphic complexity. The stratigraphic layers and salt boundaries are better resolved when simultaneously migrating direct and receiver-side ghost energy using 2C or 4C dual-field data. We observe gain in resolution and bandwidth for the salt flank and main reflectors with receiver-directional balancing (green arrows). The lateral directionality gained from using horizontal displacements of 4C seismic data attenuates migration artifacts (blue arrows), and enhances reflector continuity particularly close to the salt flank. The use of 4C rather than 2C dual-field data provides better performances of the adjoint VARTM method because it takes advantage of all directions of propagation. This also makes the method more robust for any acquisition geometry. For dual-source acquisition, source-directional balancing further improves the gain in resolution and bandwidth (magenta arrows) when source-side ghosts get mitigated.

Conclusions

Our strategy for adjoint VARTM imaging uses the full directional information contained in dual-source 4C acoustic data in seismic imaging. The adjoint VARTM formulation allows for great flexibility by means of the use of source and receiver weighting operators that control how directivity information is incorporated into the image. Because the adjoint-state VA method jointly images pressure wavefields and its 3C gradients, it is in effect an extension of conventional map migration to a two-way finite-frequency map migration. In addition to being a formulation for migration, the VA adjoint-state method also allows for the calculation of model gradients for full-waveform inversion (FWI) using dual-source...
Figure 1 Single-scatterer adjoint VARTM impulse responses for single-source, single-receiver geometry:
1C pressure data (top line) vs. 4C pressure-displacement data (bottom line) with receiver-directional balancing for (a) and (e) pressure source, (b) and (f) vertical point-force source, (c) and (g) horizontal point-force source, and (d) and (h) dual-source configuration. The green circle and blue square denote the source and receiver positions, respectively. The cyan circle marks the location of the scatterer.

4C seismic data.

Acknowledgements

We thank James Rickett and Colin Thomson for some valuable comments and stimulating discussions. We thank Schlumberger for supporting this research and granting us permission to publish it.

References


74th EAGE Conference & Exhibition incorporating SPE EUROPEC 2012
Copenhagen, Denmark, 4–7 June 2012
Figure 2 Single-scatterer adjoint VARTM impulse responses for marine seismic imaging: imaging source- and receiver-side ghosts and removing their imprints on the image. Given a pressure source (green circle) and fix-spread receivers (blue line), images of the point scatterer (cyan circle) for (a) 1C pressure data, (b) 2C pressure-vertical displacement data, and (c) 4C pressure-displacement data. For 4C pressure-displacement data, changing the source mechanism gives the images of the scatterer for (d) vertical point-force source and (e) dual-source configuration.

Figure 3 2D SEAM model example: imaging source- and receiver-side ghosts and removing their imprints on the image. Given pressure sources, salt-flank images for (a) 1C pressure data, (b) 2C pressure-vertical displacement data, and (c) 4C pressure-displacement data. Given 4C pressure-displacement data, salt-flank images for (d) vertical point-force sources and (e) dual-source configuration. For reference, (f) band-pass filtered impedance contrast model. Arrows mark features of interest in the images.
Introduction

Reverse-time migration (RTM) is now a standard imaging technique in the industry. The method provides high-quality images for oil and gas exploration and is of special interest for complex subsurface geologies. Yet, conventional RTM migration does not resolve the problem of imaging waves multiply scattered by complex subsurface structures. This limitation comes in part from the underlying single-scattering assumption of most migration techniques (Fleury and Vasconcelos, 2012). To utilize the energy, illumination and sensitivity contained in multiply scattered data, we propose a nonlinear RTM migration algorithm (Fleury and Snieder, 2011). In this paper, we improve this technique and propose a new strategy for using multiply scattered waves in interpretation and migration velocity analysis.

Theory of nonlinear reverse-time migration

The division of Earth properties into subdomains of short and long wavelengths justifies the use of smooth model \( m_0 \) perturbed by rough model \( \Delta m \) in seismic migration (\( m = (\rho, c) \) in the acoustic assumption). Model \( \Delta m \) acts as a scattering source for scattered wavefield \( w_s \) superimposed on reference wavefield \( w_0 \). The relation between wavefield \( w_s \) (equal to seismic data \( d \), at receivers) and perturbation \( \mathbf{P} \) of the Helmholtz operator \( \mathbf{H}_0 \) (function of models \( m_0 \) and \( \Delta m \)) is nonlinear:

\[
  w_s = w_0 \cdot \mathbf{P} \cdot w_0 + w_0 \cdot \mathbf{P} \cdot w_0 \cdot \mathbf{P} \cdot w_0 + \ldots
\]  

To preserve this nonlinear relation between wavefield and perturbation, we modify the conventional RTM extrapolation technique. We introduce both source- and receiver-extrapolated scattered wavefields \( (w_{s,sou} \text{ and } w_{s,rec}) \) and assume to know an estimate of source signature \( s \). The modified right-hand side equations (5-7) replace the conventional left-hand side equations (2-4).

\[
  H_0 \cdot w_0 = s \tag{2}
\]

\[
  H_0^\dagger \cdot w_{s,rec} = d \tag{3}
\]

\[
  w_{s,sou} = 0 \tag{4}
\]

\[
  (H_0^\dagger - \mathbf{P}^\dagger) \cdot w_{s,rec} = d_s + \mathbf{P}^\dagger \cdot w_0 \tag{6}
\]

\[
  (H_0^\dagger - \mathbf{P}^\dagger) \cdot w_{s,sou} = \mathbf{P}^\dagger \cdot w_0 \tag{7}
\]

Symbol \( ^\dagger \) denotes the adjoint operation. Operator \( \mathbf{P} \) being an unknown of the imaging problem, we use an estimate referred to as \( \mathbf{P}_{est} \) and obtained from a conventional RTM image by either picking reflectors or directly scaling the image in order to build a scattering model \( (\Delta m = (\Delta \rho, \Delta \mathbf{c})) \). Also, \( \mathbf{P}_{est} \) can be derived when integrating our nonlinear RTM method with full waveform inversion technology.

With the newly extrapolated reference and scattered wavefields, we define image \( i \) as

\[
  i = \sum_{\text{sources}} \text{sources} \cdot w_0 \ast \text{sources} \cdot w_{s,rec} + \frac{1}{2} \sum_{\text{sources}} \text{sources} \cdot w_{s,sou} \ast \text{sources} \cdot w_{s,rec},
\]

where \( \ast \) denotes zero-time crosscorrelation or deconvolution. The imaging condition maps into the subsurface the interaction of source-extrapolated wavefields \( w_0 \) and \( w_{s,sou} \) with receiver-extrapolated wavefield \( w_{s,rec} \). This imaging condition is nonlinear with respect to the perturbation (Fleury and Vasconcelos, 2012) while preserving a linear relation between image and data. We can either choose a crosscorrelation or deconvolution imaging condition. The former relates to an energy-based representation of the subsurface while the latter is reflectivity-based. We use a deconvolution imaging condition following the recommendations of Schleicher et al. (2008). Image \( i_0 \) maps the interaction of reference and scattered wavefields and is similar to conventional imaging conditions. Image \( i_s \) maps the interaction of only scattered wavefields which corresponds to higher orders of scattering. It has no linear contribution to the image and is therefore negligible under the linear single-scattering assumption. To highlight structural contrasts in the images, we modify \( i_0 \) and \( i_s \) to only consider the interactions of up- and down-going wavefields, which we denote with superscripts \( (u) \) and \( (d) \), respectively. This results in four sub-images \( i_{0,(u,d)} \), \( i_{s,(u,d)} \), \( i_{s,(d,u)} \), and \( i_{s,(u,d)} \) in equations 9 and 10 which are key for our analysis.
\[ i_0 = \sum_{\text{sources}} w_{0}^{(d)} \ast w_{\text{s,rec}}^{(u)} + \sum_{\text{sources}} w_{0}^{(u)} \ast w_{\text{s,rec}}^{(d)} \] (9) \[ i_k = \sum_{\text{sources}} w_{s,\text{sou}}^{(d)} \ast w_{s,\text{rec}}^{(u)} + \sum_{\text{sources}} w_{s,\text{sou}}^{(u)} \ast w_{s,\text{rec}}^{(d)} \] (10)

**Strategies for nonlinear reverse-time migration**

Multiply scattered waves are waves that have been reflected, diffracted, and more generally scattered more than once. They transport energy that contains information about the subsurface but that is not traditionally mapped into conventional RTM images. Multiply scattered waves both illuminate the subsurface with better coverage than singly scattered waves and are generally more sensitive to the Earth model. Sub-images \( i_0^{(d,u)} \) and \( i_0^{(u,d)} \) emphasize single scattering events while sub-images \( i_k^{(d,u)} \) and \( i_k^{(u,d)} \) emphasize multiple scattering events. Interestingly, all of these sub-images are representations of the same subsurface structure and must consequently share common structural features. A comparative study of these sub-images therefore provides valuable information about the subsurface.

![Figure 1](image)

**Figure 1** Modeling data for experiments (a) one-scatterer velocity model, (b) lense and two-scatterer velocity model, and (c) low-anomaly and one-scatterer velocity model. Density model (d) is the same for all experiments. The red dots and green line indicate the fix-spread source/receiver geometry.

The three synthetic experiments in Figure 1 illustrate this concept. We migrate these data with the same reference model corresponding to background velocity gradient \( \epsilon_0 \) in Figure 1(a) and smooth density \( \rho_0 \) in Figure 1(d). Despite small velocity anomalies, conventional RTM images provide a clear image of the density reflector in the three cases. Interpreting this reflector and introducing it into a density contrast model leads to perturbation operator \( \Delta \rho \) used in the modified extrapolation procedure. Sub-images \( i_0^{(d,u)} \) and \( i_0^{(u,d)} \) do not contribute to the image of the scatterers because wavefields \( w_{0}^{(u)} \) and \( w_{s,\text{sou}}^{(d)} \) do not illuminate the scatterer. We ignore them in the analysis of the results. Figure 2 shows that with the correct velocity model (experiment (a)), sub-images \( i_0^{(d,u)} \) and \( i_0^{(u,d)} \) focus scattered energy at the correct location (blue dot). In sub-image \( i_0^{(d,u)} \), reference waves illuminate the scatterer from above. In sub-image \( i_k^{(u,d)} \), scattered waves illuminate the scatterer from below. Adding the two contributions improve the spatial resolution of the scatterer. In experiment (b) (Figure 3), the high-velocity lens causes the two scatterers to be focused at the wrong locations in depth in sub-image \( i_0^{(d,u)} \). Sub-image \( i_k^{(u,d)} \) utilizes different illumination and maps the scatterers at their correct locations (blue dots). Sub-image \( i_k^{(u,d)} \) does not
show sensitivity to the velocity lens because the lens does not principally affect waves multiply scattered by the object "reflector-scatterer-reflector" in that order. Experiment (c) shows similar conclusions in Figure 4. Sub-image $i_{0}^{(d,a)}$ maps the scatterer at the correct location (blue dot) while sub-image $i_{0}^{(a,d)}$ shows defocus of the multiply scattered energy. Scattered wavefields $w_{s,sou}$ and $w_{s,rec}$ that reflect at the density contrast before or after interacting with the scatterer are sensitive to the low-velocity anomaly. Wavefields $w_{0}$ and $w_{s,rec}$ that directly interact at the scatterer are not. Experiment (a) illustrates how to utilize illumination and energy of multiply scattered waves to provide extra information about the subsurface. Experiments (b) and (c) demonstrate how to utilize sensitivity of multiply scattered waves to cross-validate the accuracy of a given migration velocity model.

**Figure 2** Point scatterer sub-images for experiment (a): adding sub-images (a) and (b) provides extra-illumination with multiply scattered energy and results in increased spatial resolution in sub-image (c).

**Figure 3** Point scatterer sub-images for experiment (b): contrary to image (a), image (b) focuses the scatterers at their correct locations and shows no sensitivity to the velocity lense.

**Figure 4** Point scatterer sub-images for experiment (c): contrary to image (a), image (b) is sensitive to the low-velocity anomaly and shows defocus of the scatterer.

A practical application for our nonlinear RTM method is target-oriented sub-salt imaging. Sub-salt imaging is challenging because of the structural complexity of salt bodies and because of the general lack of illumination below salt. We consider the Sigsbee 2A data (Paffenholz et al., 2002) and migrate with the velocity models in Figure 5. After conventional RTM, we select a poorly illuminated area (green box in Figure 5) and apply our nonlinear RTM algorithm. Our goal is to improve the image quality and resolve ambiguities in this targeted poorly illuminated sub-salt area. Figure 6 shows some of the resulting images. Despite different wavelets (no Laplacian filtering), sub-image $i_{0}^{(d,a)}$ exhibits similar structure and amplitude to the conventional image. The illumination is relatively poor apart from the left...
Figure 5 Migrating Sigsbee data: (a) reference velocity model and (b) velocity model associated to operator $P_{est}$. Scaling a conventional RTM image provides the contrast $\Delta c_{est}$ used to define the estimate of operator $P$. The red and purple lines indicate the constant-offset source/receiver geometry.

Figure 6 Sigsbee images in poorly illuminated sub-salt area: (a) conventional RTM image after Laplacian filtering compared to (b) sub-image $i_{0}^{(d,a)}$ and (c) sub-image $i_{s}^{(u,d)}$. Side of the image (between 10 and 11 km offset). Sub-image $i_{s}^{(u,d)}$ maps the interaction of multiply scattered waves. This results in better sub-salt illumination. The coherent structures in sub-image $i_{s}^{(u,d)}$ are similar to the ones present in sub-image $i_{0}^{(d,a)}$. Migration velocity model $c_{0}$ (Figure 5(a)) is therefore quite accurate. Sub-image $i_{s}^{(u,d)}$ also shows better coherence and reveals additional features in the image. For example, sediment layering is more visible especially across the two faults. The nonlinear RTM method helps for interpreting the sub-salt image.

Conclusions

Our strategy for nonlinear RTM imaging outlines the potentials of utilizing energy, illumination, and sensitivity of multiply scattered waves in seismic imaging. The comparative study of nonlinear sub-images shows as a tool for both interpretation and possibly model sensitivity analysis.

Acknowledgements

This work was supported by the Consortium Project on Seismic Methods for Complex Structures at the Center for Wave Phenomena.

References


74th EAGE Conference & Exhibition incorporating SPE EUROPEC 2012
Copenhagen, Denmark, 4-7 June 2012
Figure 1: Cartoon describing the behavior of local correlation for the correct and for an incorrect model. The correlation of two images peaks in the direction aligned with the slope of the reflector (a); when the velocity is incorrect the correlation peaks away from it (b).

Introduction

Seismic imaging includes the estimation of both the position of the structures that generate the data recorded at the surface and a model that describes the propagation in the subsurface. The two problems are related since a model is necessary to infer the position of the reflectors. The waves recorded at the surface are extrapolated in the model by solving a wave equation, and they are crosscorrelated with a synthetic source wavefield simulated in the same model (Claerbout, 1985). Under a single scattering approximation, reflectors are located where the source and receiver wavefields match in time and space. If the velocity model is inaccurate, the reflectors are positioned at incorrect locations.

Wave-equation tomography is a family of techniques that estimate the velocity model parameters from finite bandwidth signals recorded at the surface. The inversion is usually formulated as an optimization problem, where the correct velocity minimizes a certain objective function that measures the inconsistency between the data simulated in a trial velocity model and the observed data. The objective function can be defined either in the data space (full-waveform inversion (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004)) or in the image space (migration velocity analysis).

Migration velocity analysis (Fowler, 1985; Faye and Jeannot, 1986; Al-Yahya, 1989; Chavent and Jacewitz, 1995; Biondi and Sava, 1999; Albertin et al., 2006; Yang and Sava, 2010) defines the objective function in the image space and is based on the semblance principle (Al-Yahya, 1989). Groups of experiments and the associated images are analyzed. If the velocity model is correct, the images from different experiments must be consistent since a single earth model generates the recorded data. Migration velocity analysis leads to smooth objective functions and well-behaved optimization problems (Symes, 1991; Symes and Carazzone, 1991), and it is less sensitive to the initial model than full-waveform inversion. Migration velocity analysis measures either the invariance of the migrated images in an auxiliary dimension (reflection angle, shot, etc.) (Al-Yahya, 1989; Rickett and Sava, 2002; Xie and Yang, 2008) or focusing in an extended space (Rickett and Sava, 2002; Symes, 2008; Sava and Vasconcelos, 2009).

We propose a new objective function that operates in the image space and does not need common-image gathers (CIGs). We consider pairs of images from adjacent experiments and reformulate the semblance principle in the physical space, instead of the extended space at selected CIGs. We use the local correlations of two images to define an objective function based on multi-dimensional shifts in the image space. This approach allows us to include all image points in the velocity analysis step.

Theory

Migration velocity analysis is based on the semblance principle (Al-Yahya, 1989): a single earth model generates the recorded data and if the velocity model is correct, different experiments must produce consistent images of the reflectors.

Two possible measures of similarity between two images are their difference and their correlation.
We evaluate the semblance of two images by computing local correlations at each image point and define the objective function

$$J(m) = \frac{1}{2} \sum_i \| K(x) \sum_\lambda P(x, \lambda) c_i(x, \lambda) \|^2_2,$$  \hspace{1cm} (1)

where $c_i(x, \lambda) = \int_{w(x)} R_{i+1}(\xi - \frac{\lambda}{2}) R_i(\xi + \frac{\lambda}{2}) \, d\xi$ is the local correlation of the images $R_i$ and $R_{i+1}$ corresponding to shots with index $i$ and $i+1$, and $P$ is a penalty operator that highlights features that are related to velocity errors. The correlations are computed over the windows $w(x)$. Local correlations measure phase shifts between two images, thus our approach is robust against amplitude mismatches between shots (unlike an objective function based on the difference of two images).

If the velocity model is correct, two images from neighboring experiments build up the final image along the direction of the reflectors. This is equivalent to having no moveout in shot-domain common-image gathers (Xie and Yang, 2008). The similarity between the two images is evaluated at every point through local correlations. If the velocity model is correct, the maximum of the correlation lies along the reflector slope (Figure 1(a)); otherwise the maximum of the correlation is displaced from the reflector slope in a direction that depends on the sign and extent of the velocity error (Figure 1(b)). A measure of velocity error is obtained by penalizing the local correlations perpendicular to the reflector slope (Figure 1). If the velocity model is correct, the penalty operator annihilates the signal in the local correlation panels, otherwise a residual is obtained and can be exploited for wavefield tomography.

We compute the gradient of the objective function with the adjoint-state method (Plessix, 2006). In equation 1, we have two images, $R_i$ and $R_{i+1}$, and 4 wavefields $u = [u_{r,i} \; u_{s,i} \; u_{s,i+1} \; u_{r,i+1}]$, which are the state variables of the system. The main step of the adjoint-state method is the computation of the adjoint-state variables, which are the solution of the state-equation for a new source term $g$. The new adjoint sources are defined as the partial derivative of the objective function with respect to the state variables $u$. We consider the dip field a slowly varying function of state variables and neglect its derivative with respect to the state variables, although these can also be included in the adjoint-source calculation. The source-side adjoint source has the following expressions:

$$g_{s,x_s}(x, \eta, t) = \left[ r_{x_{s-1}}(x) \int P(x, \lambda) w \left( x - \eta + \frac{\lambda}{2} \right) R_{x_{s-1}}(\eta - \lambda) \, d\lambda \right. + \left. r_{x_s}(x) \int P(x, \lambda) w \left( x - \eta - \frac{\lambda}{2} \right) R_{x_{s+1}}(\eta + \lambda) \, d\lambda \right] u_{r,x_s}(\eta, t).$$  \hspace{1cm} (2)

Similarly, the receiver-side adjoint source is obtained by replacing the background receiver wavefield $u_{r,x_s}$ with the background source wavefield $u_{s,x_s}$ in equation 2. The gradient of the objective function is computed then as the zero-time, zero-space lag crosscorrelation in time (or frequency) of state and adjoint-state variables, analogously to the procedure employed in conventional FWI.

**Synthetic Examples**

For illustration, we consider the simple synthetic heterogeneous model shown in Figure 2(a). We generate full-acoustic data with absorbing boundary conditions. The optimal migrated image is shown in Figure 2(b). The initial model for migration is a severely smoothed version of the correct model and is shown in Figure 2(c). The initial image (Figure 2(d)) is misfocused, and the reflectors are shifted from their correct location. After 4 inversion steps we recover the model in Figure 2(e) and the associated migrated image in Figure 2(f); the reflectors are better focused and closer to the actual position of the interfaces in the correct model. The edges of the reflector are frowning because the sides of the model are not well illuminated and the gradient of the objective function does not affect them. Figure 3 shows the evolution of the objective function with the number of iterations. After 4 iterations the objective function converges and the image is not updated further.

**Conclusions**

We present a new approach to wavefield tomography based on a restatement of the semblance principle in the image space. We define an objective function using appropriately penalized local image correlations.
Figure 2: Velocity model used for generating the full-acoustic data (a) and the image obtained with the
correct velocity model (b). The initial migration velocity model (b) and associated image (d). Migration
velocity model after 4 inversion steps (e) and associated image (f). Observe the better focusing and
positioning of the reflectors.

Figure 3: Evolution of the objective function with iterations. The inversion decreases the residual.
This formulation avoids both the construction of common-image gathers and the picking of moveout curves. The gradient with respect to the model parameters is computed with the adjoint-state method. A synthetic example with a laterally heterogeneous velocity model shows the ability of our strategy to measure the velocity error and focus the reflectors in the image space.

Acknowledgments
We would like to thank Eni E&P for the permission to publish this work. The reproducible numeric examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org.

REFERENCES
Migration velocity analysis (MVA) has been extended to heterogeneous transversely isotropic media with a vertical (VTI) and tilted (TTI) symmetry axis. However, MVA suffers from trade-offs between anisotropy parameters, lateral velocity variation, and the shapes of reflecting interfaces. Sarkar and Tsvankin (2003, 2004) present a 2D MVA algorithm designed to estimate both the spatially varying vertical velocity $V_0$ and parameters $\varepsilon$ and $\delta$ of VTI media. They divide the model into factorized blocks, in which the ratios of the stiffness coefficients $c_{ij}$ (and, therefore, the anisotropy parameters) are constant. If $V_0$ is known at a single point in each factorized VTI block, the MVA algorithm can estimate the parameters $\varepsilon$ and $\delta$ along with the velocity gradients. Behera and Tsvankin (2009) extend this technique to “quasi-factorized” TTI media under the assumption that the symmetry axis is orthogonal to the reflector beneath each layer.

Most existing methods, however, are designed for relatively large-scale lateral heterogeneity. Lateral velocity variation on a scale comparable to spreadlength (e.g., caused by velocity lenses) can distort the estimated parameters and reduce the quality of stack, even if velocity analysis is based on reflection tomography (Takanashi et al., 2009). To properly account for small-scale lateral heterogeneity, here we extend the MVA algorithms of Sarkar and Tsvankin (2004) and Behera and Tsvankin (2009) to TI media with quadratic lateral variation of the symmetry-direction velocity $V_0$.

Influence of quadratic lateral velocity variation on image gathers

First, we consider a piecewise-factorized VTI model with the vertical velocity $V_0$ in each block described by:

$$V_0(x, z) = V_0(0, 0) + k_{x1} x + k_{z1} z + k_{x2} x^2,$$  \hspace{1cm} (1)

where $k_{x1}$ and $k_{z1}$ are vertical and lateral velocity gradients, and $k_{x2}$ is the quadratic coefficient. For instance, a smooth (e.g., parabola-shaped) low-velocity lens centered at $x = 0$ can be approximated by the velocity function $V_0(x, z)$ with a positive $k_{x2}$. Takanashi and Tsvankin (2011) discuss the influence of thin laterally heterogeneous (LH) layers on the reflection moveout from deeper interfaces. They show that the distortion of the NMO velocity or ellipse depends on the curvature of the vertical interval traveltime, and the magnitude of the distortion increases with the distance between the LH layer and the target. For instance, the NMO velocity for the reflection from the bottom of a model that includes an LH layer sandwiched between two laterally homogeneous layers is given by (Takanashi and Tsvankin, 2011):

$$
V_{nmo,het}^{-2} = V_{nmo,hom}^{-2} + \frac{\tau_0}{3} \frac{\partial^2 \tau_{02}}{\partial x^2}, \quad D = k^2 + 3kl + 3l^2,
$$  \hspace{1cm} (2)

where $V_{nmo,hom}$ is the NMO velocity for the reference laterally homogeneous medium with the parameters corresponding to the CMP location, $\tau_0$ is the total zero-offset traveltime, and $\tau_{02}$ is the interval traveltime for the second (LH) layer. The parameters $k$ and $l$ are close to the thicknesses of the second and third layer, respectively, divided by the total thickness of the model.

Under the assumption that the model is horizontally layered and lateral heterogeneity is weak, the second derivative of the vertical traveltime can be replaced with that of the vertical velocity (Grechka and Tsvankin, 1999), and equation 2 can be rewritten as

$$
V_{nmo,het}^{-2} = V_{nmo,hom}^{-2} - \frac{2 \tau_0 \tau_{02} D k_{x2}^{(2)}}{3 V_0^2},
$$  \hspace{1cm} (3)

where $k_{x2}^{(2)}$ is the quadratic coefficient from equation 1 for the second layer. Since $V_{nmo,het}$ is responsi-
The parameters of the isotropic LH layer are $V_0(0) = 2280\,\text{m/s}$, $k_{x1} = 0.24\,\text{s}^{-1}$, and $k_{x2} = 2.4 \times 10^{-4}\,\text{s}^{-1}\text{m}^{-1}$. The parameters of the VTI medium beneath the LH layer are $V_0 = 3000\,\text{m/s}$, $k_{x1} = 0.1\,\text{s}^{-1}$, $k_{z1} = 0$, $\varepsilon = 0.2$, and $\delta = 0.1$. Image gathers produced with an inaccurate parameter of the LH layer: (b) $k_{x1} = 0$; (c) $k_{x2} = 0$ (the other parameters are correct). The maximum offset is 4 km.

The influence of errors in $k_{x1}$ and $k_{x2}$ in a thin layer on image gathers obtained after Kirchhoff prestack depth migration is illustrated by Figure 1. Prestack synthetic data are produced by a finite-difference algorithm. The contribution of either $k_{x1}$ or $k_{x2}$ leads to a velocity variation of 960 m/s between $x = -2$ km and $x = 2$ km. Although the error in $k_{x1}$ does distort $V_0$ at $x \neq 0$ and reflector positions, the corresponding residual moveout is relatively small at all depths (Figure 1b). In contrast, ignoring $k_{x2}$ leads to a substantial overcorrection (i.e., the imaged depth decreases with offset) for the reflectors from interfaces far below the thin layer (Figure 1c). Consequently, failure to correct for $k_{x2}$ in the overburden
causes errors in the medium parameters at depth (e.g., Takanashi et al., 2009).

Similar conclusions can be drawn for TTI media with quadratic lateral velocity variation. We assume that the symmetry axis is orthogonal to the bottom reflector in each block and that the symmetry-direction velocity \( V_0 \) is represented as

\[
V_0(x, z) = V_0(0, 0) + k_{x1}x + k_{z1}z + k_{x2}x^2,
\]

where \( x \) and \( z \) are the rotated coordinate axes parallel and perpendicular to the layer boundaries (Figure 2a). For models with close dips similar to the one in Figure 2, the moveout distortion in image gathers is caused primarily by \( k_{x2} \).

**MVA for models with quadratic velocity variation**

Our algorithm includes iterative application of Kirchhoff PSDM and velocity updating until the residual moveout becomes sufficiently small. To estimate the residual moveout for long-offset data, Sarkar and Tsvankin (2004) employ 2D semblance analysis using the following nonhyperbolic equation in the migrated domain:

\[
z_M^2(h) \approx z_M^2(0) + Ah^2 + B \frac{h^4}{h^2 + z_M^2(0)},
\]

where \( z_M \) is the migrated depth, \( h \) is the half-offset, and \( A \) and \( B \) are dimensionless coefficients responsible for the residual moveout. In our model, each block is described by the parameters \( V_0, \delta, \epsilon, k_{x1}, k_{z1}, \) and \( k_{x2} \) (or \( k_{x2} \) for TTI models). Inversion for the parameters of a thin layer in the layer-stripping mode is generally unstable. Also, the moveout at the bottom of the thin layer in Figure 1 is distorted by NMO stretch, which can lead to further instability in the parameter updates. In contrast, the residual moveout of deep events is sensitive to errors in the coefficient \( k_{x2} \) in the overburden. Therefore, in the presence of velocity lenses it is highly beneficial to invert the residual moveout for all available reflectors simultaneously. To make the algorithm of Sarkar and Tsvankin (2004) suitable for such simultaneous update,
the perturbations of the migrated depths are expressed as linear functions of the parameter perturbations in all blocks.

**Synthetic test**

The test results in Figure 3 are obtained for a model that includes two thin isotropic layers with quadratic velocity variation embedded in a TTI medium (Figure 2). The depth profile of the symmetry-direction velocity is assumed to be known at one location \((x = 0)\). The initial model for MVA is composed of homogeneous, isotropic blocks. Using the residual moveout for all reflectors, we invert for \(V_0\), \(k_{x1}\), and \(k_{x2}\) in the thin isotropic layers and \(k_{x1}\), \(k_{x2}\), \(\epsilon\), and \(\delta\) in the TTI layers. After 30 iterations, MVA practically removes the residual moveout for all events (Figure 3b) and accurately recovers the interval medium parameters. In particular, the errors in \(\epsilon\), \(\delta\), and \(k_{x1}\) in the TTI layers are less than 0.01, 0.03, and 0.01, respectively. Next, we apply the algorithm with the value of \(k_{x2}\) set to zero and the thin layers subdivided into smaller blocks (1.5 km wide, which is close to half the effective spreadlength; Figure 3c). Despite ignoring the contribution of \(k_{x2}\), the velocity variations in the thin layers are well-resolved and the errors in the parameters of the TTI layers are just slightly higher than those in the previous test. Note, however, that when \(k_{x2}\) is not taken into account, a block boundary has to be close to the center of the lens. In contrast, running MVA in the layer-stripping mode leads to relatively large residual moveout (Figure 3d) and significant errors in the TTI parameters. The instability of parameter estimation in the layer-stripping mode is partially caused by the NMO stretch at the bottom of the thin LH layers (Figure 3).

**Conclusions**

Our analytic and numerical results show that small-scale lateral heterogeneity in thin shallow layers (e.g., due to velocity lenses) can seriously distort P-wave image gathers for VTI and TTI media. In particular, the residual moveout for reflectors at depth is sensitive to the quadratic lateral variation of the symmetry-direction velocity \(V_0\) in the overburden. To account for such heterogeneity, we devised migration velocity analysis for TI models composed of “quasi-factorized” blocks, in which the anisotropy parameters are constant, while \(V_0\) is described by a quadratic function of the lateral coordinate. It is essential to carry out model updating for all blocks or layers simultaneously because the parameters of thin LH layers cannot be constrained in the layer-stripping mode. Under the assumption that the vertical profile of \(V_0\) is known at one location in each block, the algorithm accurately reconstructs the laterally-varying velocity fields and anisotropy parameters throughout the medium. Therefore, our model representation helps avoid instability of parameter estimation typical for reflection tomography.

**Acknowledgments**

This work was supported by Japan Oil, Gas and Metals National Corporation (JOGMEC) and the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena.

**References**


Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination

Kees Wapenaar∗ (Delft University of Technology), Jan Thorbecke (Delft University of Technology), Joost van der Neut (Delft University of Technology), Filippo Broggini (Colorado School of Mines) and Roel Snieder (Colorado School of Mines).

It has recently been shown that the response to a virtual source in the subsurface can be derived from reflection data at the surface and an estimate of the direct arrivals between the virtual source and the surface. Hence, unlike for seismic interferometry, no receivers are needed inside the medium. This new method recovers the complete wavefield of a virtual source, including all internal multiple scattering. Because no actual receivers are needed in the medium, the virtual source can be placed anywhere in the subsurface. With some additional processing steps (decomposition and multidimensional deconvolution) it is possible to obtain a redatumed reflection response at any depth level in the subsurface, from which all the overburden effects are eliminated. By applying standard migration between these depth levels, a true amplitude image of the subsurface can be obtained, free from ghosts due to internal multiples. The method is non-recursive and therefore does not suffer from error propagation. Moreover, the internal multiples are eliminated by deconvolution, hence no adaptive prediction and subtraction is required.
Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination
K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini & R. Snieder

Introduction
We discuss a new approach to creating the response to a virtual source inside an unknown medium that goes beyond seismic interferometry and that can be used as a basis for internal multiple elimination. Broggini et al. (2011, 2012) show that, given the reflection response of a 1D layered medium, it is possible to obtain the response to a virtual source inside the medium, without the need to know the medium parameters. The method consists of an iterative scheme, akin to earlier work of Rose (2002). Interestingly, the response retrieved by this new method contains all scattering effects of the layered medium. Note that, in order to obtain the same virtual-source response by seismic interferometry, one would need a receiver at the position of the virtual source inside the medium, and real sources at the top and bottom of the medium. Hence, the advantage of the new 1D approach over 1D seismic interferometry is that no receivers are needed inside the medium and that the medium needs to be illuminated from one side only.

Last year we made a first step towards generalizing this idea to the 3D situation (Wapenaar et al., 2011). Using physical arguments we proposed an iterative scheme that transforms the reflection response of a 3D medium into the response to a virtual source inside the 3D medium. Whereas controlled-source seismic interferometry (Schuster et al., 2004; Bakulin and Calvert, 2006) requires receivers in the subsurface (Figure 1a), the proposed scheme requires reflection data at the surface (Figure 1b) and an estimate of the direct arrivals between the virtual source and the acquisition surface. The method recovers the complete wavefield of a virtual source, including internal multiple scattering, without needing any information about the location and strength of the interfaces. Similar as in seismic interferometry, the virtual source is closer to the reservoir. However, unlike for interferometry, no actual receivers are needed inside the medium. Moreover, by repeated application, virtual sources and virtual receivers can be obtained anywhere inside the medium (Figure 1c). We show in this paper that from the reconstructed wavefields inside the medium it is possible to obtain improved images, free of ghosts caused by internal multiple scattering. The proposed method requires a background model for the primary arrivals, but the internal multiple prediction and elimination is data-driven. The method is non-recursive and therefore does not suffer from error propagation. The internal multiples are eliminated by deconvolution, hence no adaptive prediction and subtraction is required.

Figure 1: (a) Controlled-source seismic interferometry creates virtual sources in the subsurface at the position of receivers in a borehole (Schuster et al., 2004; Bakulin and Calvert, 2006). (b) Data-driven wavefield reconstruction (Broggini et al., 2011; Wapenaar et al., 2011) also creates virtual sources in the subsurface, but uses sources and receivers at the surface. (c) Repeated application gives virtual sources and virtual receivers throughout the medium, leading to improved images, free of ghosts caused by internal multiple scattering.

Initiating the iterative process
We discuss an iterative scheme that creates the response to a virtual source in the subsurface from reflection data at the surface (for details see Wapenaar et al., 2011, 2012). The scheme is initiated with an estimate of the direct arrivals between the virtual source at $x_{VS} = (x_{VS}, z_{VS})$ and the surface. These direct arrivals are reversed in time, which turns the upgoing field at the surface $z = 0$ into a downgoing field $p_0^-(x, t)$, see Figure 2a (the red curves will be discussed in the next section). The field $p_0^-(x, t)$ is used as the initial downgoing field at $z = 0$. It is convolved with the reflection response of the medium, according to

$$p_0^-(x_R, t) = \int_{-\infty}^{\infty} [R(x_R, x, t) * p_0^+(x, t)]_{z=0} \, dx,$$

for $z_R = 0$. Here $R(x_R, x, t)$ is the measured reflection response, after surface-related multiple elimination and deconvolution for the source wavelet. The result of this convolution is the upgoing wave field $p_0^-(x, t)$.
Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination

K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini & R. Snieder

Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination

K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini & R. Snieder

Creating the virtual-source response

After finalizing the iterative process, we define \( p(x,t) \) as the superposition of the final incident and reflected wave fields at \( z = 0 \). Because, for this example, iteration \( k = 1 \) was the final iteration, we have \( p(x,t) = p^\text{sym}(x,t) + p^\text{as}(x,t) \). This total field is shown in Figure 3b (left frame). Within the time window this field is antisymmetric in time (this was the design criterion for the iterative scheme). Hence, if we superpose the total field and its time-reversed version, i.e., \( p^\text{sym}(x,t) = p(x,t) + p(x,-t) \), all events within the time window will cancel each other, whereas outside the time window this superposition will be symmetric, see Figure 3b (right frame). Note that, because we consider a lossless medium, \( p^\text{sym}(x,t) \) obeys the wave equation. Since time-reversal changes the propagation direction, it follows that the causal part is upward propagating at \( z = 0 \) and the acausal part is downward propagating at \( z = 0 \). The first arrival of the causal part of \( p^\text{sym}(x,t) \) in Figure 3b corresponds with the direct arrival of the response to the virtual source at \( x_{VS} \) (Figure 2a). Given this last observation, combined with the fact that the causal part is upward propagating at \( z = 0 \) and that the total field obeys the wave equation and is

The iterative process

We define two travelt ime curves, indicated by the red solid lines. The lower curve is taken just before the direct arrival of Figure 2a and the upper curve is defined as the time-reversal of the lower curve, hence, it comes directly after the initial incident field \( p^+_0(x,t) \). We define a window function \( w(x,t) \) that equals 1 between these red curves and 0 elsewhere. We discuss an iterative scheme, which has the aim to modify the incident field in such a way that, within the time window, the incident field is equal to minus the time-reversed reflection response. We will motivate this peculiar condition in the next section. To achieve this goal, the \( k \)th iteration is defined as follows

\[
p^+_k(x,t) = p^+_0(x,t) - w(x,t)p^+_{k-1}(x,-t),
\]

\[
p^-_k(x_R,t) = \int_{-\infty}^{\infty} [R(x_R,x,t) * p^+_k(x,t)]_{x=0} \, dx,
\]

for \( x \) and \( x_R \) at \( z = 0 \). The reflection response \( p_R(x,t) \) is shown in Figure 2b. Only the response of the first reflector (the first event) falls within the time window. We subtract the time-reversal of this event from \( p^+_0(x,t) \), which gives, according to equation (2) for \( k = 1 \), the modified incident wave field \( p^+_1(x,t) \), see Figure 3a. Using equation (3) we evaluate the reflection response \( p^-_1(x,t) \) to this modified incident wave field, which is also shown in Figure 3a. Note that, within the time window, \( p^-_1(x,t) \) is identical to \( p^-_0(x,t) \), hence, further iterations will not cause any changes. Already after one iteration we have achieved the desired situation mentioned at the beginning of this section. This is a consequence of analyzing the simple configuration with two interfaces only. For more complex configurations more iterations will be required.

Figure 2: Initiating the iterative process. (a) An estimate of the direct field between the virtual source and the surface is reversed in time, giving the initial downgoing field \( p^-_0(x,t) \). (b) The initial downgoing field \( p^-_0(x,t) \) is convolved with the reflection response of the medium, giving the upgoing field \( p^+_0(x,t) \).

Creating the virtual-source response

After finalizing the iterative process, we define \( p(x,t) \) as the superposition of the final incident and reflected wave fields at \( z = 0 \). Because, for this example, iteration \( k = 1 \) was the final iteration, we have \( p(x,t) = p^\text{sym}(x,t) + p^\text{as}(x,t) \). This total field is shown in Figure 3b (left frame). Within the time window this field is antisymmetric in time (this was the design criterion for the iterative scheme). Hence, if we superpose the total field and its time-reversed version, i.e., \( p^\text{sym}(x,t) = p(x,t) + p(x,-t) \), all events within the time window will cancel each other, whereas outside the time window this superposition will be symmetric, see Figure 3b (right frame). Note that, because we consider a lossless medium, \( p^\text{sym}(x,t) \) obeys the wave equation. Since time-reversal changes the propagation direction, it follows that the causal part is upward propagating at \( z = 0 \) and the acausal part is downward propagating at \( z = 0 \). The first arrival of the causal part of \( p^\text{sym}(x,t) \) in Figure 3b corresponds with the direct arrival of the response to the virtual source at \( x_{VS} \) (Figure 2a). Given this last observation, combined with the fact that the causal part is upward propagating at \( z = 0 \) and that the total field obeys the wave equation and is
Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination

K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini & R. Snieder

Figure 3: (a) The iterative process. The event(s) of the upgoing field between the red curves is (are) reversed in time and subtracted from the initial downgoing field. The modified downgoing field \( p_k^- (x, t) \) (here \( k = 1 \)) is convolved with the reflection response of the medium, giving the modified upgoing field \( p_k^+ (x, t) \). (b) Left frame: Superposition of downgoing and upgoing field. Right frame: superposition of total field and its time-reversal. This is interpreted as \( G(x, x_{VS}, t) \) for the Green’s function in the real medium. This argumentation holds for more general situations, but it has been checked explicitly with the stationary-phase method for the response in Figure 3b (right frame) (Wapenaar et al., 2012). Figure 4a shows once more the direct arrival of Figure 2a, with which we initiated the iterative process, and the causal part of \( p_{\text{sym}}(x, t) \) (Figure 3b, right frame), which is interpreted as the Green’s function \( G(x, x_{VS}, t) \). Note that the direct arrival in this retrieved Green’s function comes from our initial estimate, whereas the internal multiples come entirely from the measured reflection response.

Decomposition

The Green’s function \( G(x, x_{VS}, t) \) is the response to the virtual source at \( x_{VS} \). In Figure 4a it is clearly seen that this source radiates upward as well as downward propagating fields. The latter arrive at the surface after having been reflected at the interface below the source. We now discuss how the source can be decomposed into an upward and a downward radiating source. To this end we first repeat the iterative process, but we replace the minus sign in equation (2) and Figure 3a by a plus sign. When this iterative process is finished, we define again the total field as \( p(x, t) = p_1^+ (x, t) + p_1^- (x, t) \), of which the part within the time window is now symmetric in time (instead of antisymmetric). Hence, by evaluating \( p_{\text{asym}}(x, t) = p(x, t) - p(x, -t) \), all events within the time window will again cancel each other (like in Figure 3b), whereas outside the time window this field is antisymmetric. The new created virtual source is therefore also antisymmetric (like a dipole). If we subtract this new response from our retrieved monopole response, i.e., if we evaluate \( \frac{1}{2} \{ p_{\text{sym}}(x, t) - p_{\text{asym}}(x, t) \} \) and take the causal part, we obtain the response to an upward radiating virtual source, whereas the causal part of the superposition \( \frac{1}{2} \{ p_{\text{sym}}(x, t) + p_{\text{asym}}(x, t) \} \) yields the response to a downward radiating virtual source, see Figure 4b.

Figure 4: (a) Comparison of the direct arrivals (Figure 2a) and the retrieved Green’s function \( G(x, x_{VS}, t) \) (the causal part of Figure 3b). (b) Decomposition of the virtual source into an upward and downward radiating virtual source.
Creating virtual sources inside an unknown medium from reflection data: a new approach to internal multiple elimination
K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini & R. Snieder

Figure 5: (a) As in Figure 4b, but after applying source-receiver reciprocity. (b) Result of applying multidimensional deconvolution to the responses of (a). This is the redatumed reflection response, free of overburden effects. Repeating this for all depth levels leads to a multiple-free image of the subsurface.

**Imaging without internal multiple ghosts**

Using source-receiver reciprocity, a virtual-source response observed by receivers at the surface is equal to the response to sources at the surface, observed by a virtual receiver in the subsurface. Hence, the decomposed responses of Figure 4b are, after applying source-receiver reciprocity, interpreted as the downward and upward propagating fields at a virtual receiver in the subsurface, due to sources at the surface, i.e., as $G^+(x, x_S, t)$ and $G^-(x, x_S, t)$, respectively, with $x$ at the virtual-source depth $z_{VS}$, see Figure 5a. For a range of virtual receivers at $z_{VS}$, these responses are related via

$$G^-(x_R, x_S, t) = \int_{-\infty}^{\infty} R(x_R, x, t) * G^+(x, x_S, t) \bigg|_{z=z_{VS}} \, dx,$$

for $x_R$ at $z_{VS}$. Here $R(x_R, x, t)$ is the reflection response at the virtual-source depth $z_{VS}$, for the situation of a homogeneous overburden. It can be resolved from $G^+(x, x_S, t)$ and $G^-(x, x_S, t)$ by multidimensional deconvolution (MDD), similar as in seismic interferometry by MDD. Figure 5b shows the MDD result (again without artifacts because for this example we evaluated it by the method of stationary phase). It can be seen as the redatumed reflection response from which all the overburden effects have been eliminated. By repeating the entire process for many depth levels, as indicated in Figure 1c, and applying standard migration in between these depth levels, a true amplitude image of the subsurface can be obtained, free from ghosts due to internal multiples. The advantages and disadvantages of the proposed approach with respect to other internal multiple elimination schemes (Weglein et al., 1997; Berkhourt and Verschuur, 1997; Jakubowicz, 1998; etc.) need further investigation.

**Conclusions**

For a simple configuration we have shown that it is possible to create the response to a virtual source in the subsurface from the reflection response at the surface and an estimate of the direct arrivals. Following the physical arguments in the section “Creating the virtual-source response”, it is plausible that this procedure will also work in more complex environments. Nevertheless, the method will also have its limitations. The effects of a finite acquisition aperture, errors in the direct arrivals, triplications, head waves, diving waves, fine-layering etc. need further investigation. A first numerical test by one of us (FB) with a variable-velocity syncline model shows promising results. For those configurations for which it is possible to create virtual sources from reflection data, the method can be used as a basis to image the subsurface without internal multiple ghosts. Because the proposed method is non-recursive, the data-driven prediction of internal multiples will not suffer from error propagation. Because the internal multiples are eliminated by deconvolution, no adaptive prediction and subtraction is required. Last, but not least, the internal multiples contribute to the restoration of the amplitudes of the primary reflections.

**References**

Introduction

Waveform inversion (WI) builds velocity models using seismic wavefields (Tarantola, 1984; Woodward, 1992; Pratt, 1999; Sirgue and Pratt, 2004; Plessix, 2006; Vigh and Starr, 2008; Plessix, 2009; Symes, 2009). Although costly, this methodology is gaining momentum due to its accuracy and due to its potential for robust model updates in complex geology. This method requires accurate wavefields simulated using a wave-equation, usually constant-density acoustic. The technique also employs an objective function (OF) which captures the similarity between simulated and observed seismic wavefields. Though commonly overlooked, WI can also be implemented in the image-domain, which requires a modest alteration of the OF definition when considering multiple seismic experiments (Symes, 2009; Yang and Sava, 2011).

Both data- and image-domain implementations have weaknesses which degrade their effectiveness. For the data-domain implementation, we assume that we can accurately simulate the physics of wave propagation, which is difficult when the earth is characterized by strong (poro)elastic effects. For the image-domain implementation, we assume that the migrated images are perfectly focused when the model is correct, which is not the case in areas of poor illumination, e.g. sub-salt. Here we concentrate on the image-domain WI, sometimes known as wavefield tomography (WT). We describe a method for illumination compensation which increases the robustness of WT in areas of poor illumination.

Theory

An effective gradient calculation uses the adjoint state method (Plessix, 2006; Symes, 2009). This method consists of the following steps: (1) compute the state variables, i.e. the seismic wavefields obtained from the source by forward modeling; (2) compute the adjoint source, i.e. a calculation based on the OF and on the state variables; (3) compute the adjoint state variables, i.e. the seismic wavefields obtained from the adjoint source by backward modeling; (4) compute the gradient using the state and adjoint state variables.

The state variables relate the OF to the model parameters as a solution of the acoustic wave equation

\[
\begin{bmatrix}
L(x, \omega, m) & 0 \\
\mathcal{L}^*(x, \omega, m)
\end{bmatrix}
\begin{bmatrix}
\mathbf{u}_s(x, \lambda) \\
\mathbf{u}_r(x, \lambda)
\end{bmatrix} = \begin{bmatrix}
\mathbf{f}_s(x, \lambda) \\
\mathbf{f}_r(x, \lambda)
\end{bmatrix},
\]

(1)

\(\mathcal{L}\) and \(\mathcal{L}^*\) are forward and adjoint frequency-domain wave operators, \(\mathbf{u}_s\) and \(\mathbf{u}_r\) are the simulated source and receiver wavefields, \(\mathbf{f}_s\) and \(\mathbf{f}_r\) are the source and recorded data, \(m\) are the model parameters, \(\lambda\) is the experiment index, \(\omega\) is the frequency, and \(x\) are the space coordinates \(\{x, y, z\}\).

The OF is formulated to minimize image inconsistencies caused by the model errors

\[
\mathcal{H} = \frac{1}{2} \| K_I(x) P(\lambda) r(x, \lambda) \|^2_{x,\lambda},
\]

(2)

and it is based on extended image

\[
r(x, \lambda) = \sum_{\mathbf{e}} \sum_{\omega} \mathbf{u}_s(\mathbf{e}, x - \lambda, \omega) \mathbf{u}_r(\mathbf{e}, x + \lambda, \omega)
\]

(3)

The mask operator \(K_I(x)\) restricts the evaluation of the OF to some locations in the image, and \(P(\lambda)\) is a penalty operator acting on the extended image to highlight defocusing, i.e. image inaccuracy. It is typically assumed that defocusing is only due to velocity error, such that for correct models images focus completely at \(\lambda = 0\). In this case, we can define \(P(\lambda) = |\lambda|\), which is the differential semblance optimization (DSO) operator of Symes (2009). However, this penalty operator is not effective when poor illumination affects the image accuracy and leads to additional defocusing. The DSO operator simply translates all image defocusing into model updates, which is incorrect.

As an alternative, this paper advocates a method to capture the illumination effects in the penalty operator by making it image-dependent, as opposed to a prescribed penalty based on an assumption of perfect focusing. Then, we can compare the current extended images with this partially defocused image due to illumination, as opposed to an ideal image assumed by DSO. The challenge is to construct a template image that does not suffer from velocity errors, but merely from illumination distortions. One way
Figure 1: A shot gather for a source at $x = 2.0$ km showing the missing traces which simulate partial illumination due to an acquisition obstacle.

Figure 2: Velocity models (a and b), conventional images (c and d), and angle gathers (e and f) for the correct and starting models, respectively.

to achieve this goal is by modeling and migration in the current model starting from an ideal image based on the current image and velocity model. The resulting template image captures all illumination effects, but it is not subject to velocity error. We define the penalty operator based on the inverse energy of this template image, thus generalizing the DSO technique which assumes that all image energy concentrates at $\lambda = 0$.

Using the penalty function, we can derive the adjoint sources used to model the adjoint state variables as the derivatives of $\mathcal{H}$ (equation 2) with respect to the state variables $u_s$ and $u_r$, respectively:

$$
\begin{bmatrix}
g_s(\mathbf{e}, \mathbf{x}, \omega) \\
g_r(\mathbf{e}, \mathbf{x}, \omega)
\end{bmatrix} = 
\begin{bmatrix}
K f(\mathbf{x}) \sum_{\lambda} P(\lambda) \tilde{P}(\lambda) r(\lambda) u_r(\mathbf{e}, \mathbf{x} + \lambda, \omega) \\
K f(\mathbf{x}) \sum_{\lambda} P(\lambda) \tilde{P}(\lambda) r(\lambda) u_s(\mathbf{e}, \mathbf{x} - \lambda, \omega)
\end{bmatrix}
$$

(4)
The adjoint state variable $a_s$ and $a_r$ are the wavefields obtained by backward and forward modeling using the corresponding adjoint sources:

$$
\begin{bmatrix}
L^* (x, \omega, m) & 0 \\
0 & L(x, \omega, m)
\end{bmatrix}
\begin{bmatrix}
a_s(e, x, \omega) \\
a_r(e, x, \omega)
\end{bmatrix} =
\begin{bmatrix}
g_s(e, x, \omega) \\
g_r(e, x, \omega)
\end{bmatrix},
$$

(5)

and the gradient is the correlation of the state and adjoint-state variables:

$$
\frac{\partial H}{\partial m} = \sum_{e} \sum_{\omega} \frac{\partial L}{\partial m} (u_s(e, x, \omega) a_s(e, x, \omega) + u_r(e, x, \omega) a_r(e, x, \omega)).
$$

(6)

The model is then updated using gradient line search aimed at minimizing the OF given by equation 2.

**Example**

We illustrate our new penalty operator with the model shown in Figures 2(a) and 2(c). The model includes high and low Gaussian anomalies at $x = 1.5$ km and at $x = 2.5$ km. We consider 5 horizontal interfaces and use finite-difference modeling to generate 73 shots like the one shown in Figure 1. We use the data gap between $x = 2.25 - 2.75$ km to simulate imaging illumination gaps. The migrated images and angle gathers for the true and starting models are plotted in Figures 2(c)-2(f). We invert using the DSO and illumination penalty operators, shown in Figures 3(a)-3(b), respectively. The updated model after 30 nonlinear iteration are shown in Figures 3(c) and 3(d). The migrated images, space-lag and angle gathers for the two updated models are shown in Figures 3(e)-3(j). The result obtained from the illumination-based penalty is significantly better than the DSO result. Although the space-lag gathers obtained from the DSO penalty are more focused at $\lambda = 0$, the inversion fails to produce the correct answer since the penalty does not take the uneven illumination into account, thus translating illumination defocusing into velocity updates. In contrast, the space-lag gathers for the illumination-based penalty have residual moveout, which is consistent with the subsurface illumination, and the image is closer to the correct answer. The angle gathers also show the improved image quality of the model obtained with the illumination-based penalty.

**Conclusions**

Wavefield tomography using differential semblance optimization make an implicit assumption of perfect subsurface illumination and infinite bandwidth data. The penalty operator used in the definition of the objective function corresponds to this ideal case, thus this technique fails to produce robust results in areas of poor illumination. This problem can be addressed with the definition of an improved penalty function which uses the actual migrated image based on the current acquisition and velocity model. Replacing the conventional DSO penalty operator with one taking the actual subsurface illumination into account increases the robustness and accuracy of WT, especially in poorly illuminated areas, e.g. sub-salt.

**Acknowledgments**

We acknowledge the sponsors of the Center for Wave Phenomena at Colorado School of Mines. The reproducible numeric examples use the Madagascar open-source package (http://www.reproducibility.org).

**REFERENCES**


Figure 3: Penalty operators (a and b), velocity models (c and d), conventional images (e and f), space-lag gathers (g and h) and angle gathers (i and j) for inversion using the DSO and the illumination-based penalties, respectively. In panels (a) and (b), the penalty values increase from light to dark colors.

A practical approach to prediction of internal multiples and ghosts

Jyoti Behura† & Farnoush Forghani‡
†Center for Wave Phenomena, ‡Center for Rock Abuse, Colorado School of Mines

SUMMARY

Existing methods of internal multiple prediction are either computationally expensive or not automated. Here, based on stationary phase arguments, we introduce a method for predicting internal multiples that is not only fully automated but also computationally inexpensive. The procedure is completely data driven and requires no velocity information or reflector identification. An additional advantage of the proposed method is that it can also be used to predict source- and receiver-ghosts. The method, however, is limited to gently-dipping reflectors. Through synthetic examples and field data, we demonstrate the effectiveness of our methodology.

INTRODUCTION

In most stages of seismic data processing, it is assumed that the data is devoid of multiply scattered energy and contains only primaries. For example, the contribution of multiples to semblance panels (for velocity analysis) might result in erroneous picking of velocities. More importantly, most imaging algorithms are based on the Born approximation (single scattering). If not suppressed, the multiples might show up as spurious events in the image or even interfere with the primaries resulting in a degraded image.

Suppressing surface-related multiples using SRME (Verschuur et al., 1992) is a standard practice now. Internal-multiple suppression, on the other hand, is not prevalent either because it is not fully automated or is computationally expensive. Berkhourt & Verschuur (1997) proposed an iterative algorithm that includes extrapolating the shot records to a reflecting boundary responsible for the generation of the internal multiples. On similar lines, Jakubowicz (1998) proposed isolating the primary events on data followed by a convolution and deconvolution process to predict the internal multiples. The above two methods are not computationally intensive but are not fully automated. The inverse scattering approach of Weglein et al. (1997) overcomes the above drawback and can also predict all possible internal multiples. The method, however, is computationally expensive.

In marine acquisition, since sources and receivers lie under the water surface, the air-water interface acts as a mirror. So the source would have a mirror source (ghost) above the surface; the same argument applies for the receiver as well. The source- and receiver-ghosts act as filters on the data with zeros at many frequencies. Such notch filters result in poor data quality by limiting the usable bandwidth of the data.

Existing methods of source- and receiver-ghosts suppression are primarily confined to the acquisition stage. Ziolkowski (1971) suggested recording at two different depths so as to fill in the zeros in the spectra. Similar recommendations have been made by others including Ghosh (2000). Such an acquisition is termed as over/under streamer acquisition (Moldoveanu et al., 2007). The extreme case of all hydrophones having different depths was recently proposed by Soubaras & Whiting (2011). Ghosts could also be attenuated using dual-sensors (pressure and velocity) in a streamer (Carlson et al., 2007).

Here, we introduce a practical approach to predict internal multiples with the primary aim of making the process computationally tractable as well as automatic. We also demonstrate how the proposed method could also be used to predict source- and receiver-ghosts.

INTERNAL MULTIPLES

As proposed by Jakubowicz (1998), any first-order internal multiple can be represented as a combination of three primaries. For example, in Figure 1a, the internal multiple $M_{ABCDE}$ is given by a convolution of primaries $P_{ABG}$ and $P_{FDE}$ followed by a deconvolution with the third primary $P_{FCG}$. To automate the process, we replace the windowed primaries (Jakubowicz, 1998) with the full Green’s function between the source and receiver.

Fig. 1: (a) Illustration of the computation of the internal multiple for a 2D acquisition. The summation zone spans between the source at $x_A$ and the receiver at $x_E$. The dashed red line denotes the primary that is common to the two other primaries and needs to be deconvolved in equation 1. (b) Plan view of a 3D surface-seismic acquisition. The grey area represents the proposed summation zone for a given source-receiver pair.
Internal multiples and ghosts prediction

In practice, the internal multiple is obtained by following dual-summation over the locations of $F$ and $G$:

$$
\mathcal{M}_{ABCDE}(\omega) = \sum_{x_F} \sum_{x_G} \frac{\mathcal{G}(x_A, x_G, \omega) \mathcal{G}(x_E, x_E, \omega)}{\mathcal{G}(x_F, x_G, \omega)},
$$

(1)

where $x$ represents the coordinate location, $\mathcal{G}$ is the Green’s function between any source and receiver and $\omega$ is the frequency. Equation 1 predicts all orders of internal multiples at the same time. The above procedure is akin to seismic interferometry (Lobkis & Weaver, 2001; Wapenaar & Fokkema, 2006) that retrieves the Green’s function between a virtual source and a receiver. The maximum contribution to the summation 1 comes from the stationary points at $x_F$ and $x_G$.

To predict only multiples, we limit the summation over locations $x_F$ and $x_G$ to a specific spatial location and aperture. It is crucial that points $x_F$ and $x_G$ lie in between the source at $x_A$ and the receiver at $x_E$. If $F$ and $G$ lie outside this region or are close the points $A$ and $E$, respectively, then $\mathcal{M}_{ABCDE}$ would contain primaries in addition to multiples. For example, in the special case of $F$ coinciding with $A$ and $G$ coinciding with $E$, equation 1 would reduce to $\mathcal{M}_{ABCDE}(\omega) = \mathcal{G}(x_A, x_G, \omega)$, i.e. the ‘predicted multiple’ would be the Green’s function between source at $x_A$ and receiver at $x_E$ and contain all primaries and multiples. A significant advantage of limiting the summation to a small aperture (determined by the source and receiver locations) is the reduction in computational cost. Summing over all possible locations of $F$ and $G$ would not only be prohibitively expensive, but also would not yield the internal multiples. In 3D, the summation can be carried over all surface points in a region (grey area in Figure 1b) that would likely contain the stationary points $F$ and $G$. In the inverse scattering approach (Weglein et al., 1997), on the other hand, the summation is over all possible locations of $F$ and $G$ which adds to its computational cost.

The above modifications (using the Green’s function in equation 1 and limiting the region of summation), however, result in other drawbacks and limitations. Usage of the full Green’s function in equation 1 results in spurious events that have been extensively studied in seismic interferometry (Snieder et al., 2006; Forghani & Snieder, 2010). However, since there is no stationary phase contribution to the spurious events, including more traces into the summation can suppress these events. Another drawback is that the near-offset prediction would not be accurate because of the proximity of $F$ and $G$ to $A$ and $E$, respectively and also because of the limited number of traces going into the summation in equation 1. The limitation imposed on the region of summation assumes that the stationary points $x_F$ and $x_G$ lie in between $x_A$ and $x_E$. This might not be the case for complicated geology and steeply dipping reflectors. For example, in Figure 2, since $G$ lies outside the zone of integration (between $x_A$ and $x_E$), the internal multiple would not be predicted accurately. So the method proposed here would work well for gently dipping layers and non-complicated subsurfaces. This is the primary limitation of our methodology compared to the methods of Berkhout & Verschuur (1997); Jakubowicz (1998); Weglein et al. (1997).

**Synthetic examples:** Two synthetic examples of internal multiple prediction are presented here. The first synthetic test is for a 1D velocity model comprising of 3 layers. A shot gather for this test is shown in Figure 3a. The predicted internal multiples are shown in Figure 3b. As explained above, the near-offset prediction is not accurate; however, the mid- and far-offsets clearly show the first- and second-order internal multiples.

Internal multiple prediction for a common-offset section for the Sigsbee model is shown in Figure 4. Most of the internal multiples have been predicted accurately; multiples corresponding to the steeply dipping reflectors, however, are not present in the prediction.

**GHOSTS**

The geometry of a source+receiver ghost (Figure 5a) is
similar to that of an internal multiple (Figure 1a). So the same algorithm, that is used for predicting internal multiples, can be exploited for predicting source+receiver ghosts. The source+receiver ghost \( g(x_A, x_G, \omega) \) between the source at \( x_A \) and the receiver at \( x_G \) is given by the dual-summation:

\[
g(x_A, x_G, \omega) = \sum_{x_C} \sum_{x_E} \frac{G(x_A, x_E, \omega) G(x_C, x_G, \omega)}{G(x_C, x_E, \omega)}. \tag{2}
\]

**Viking Graben data:** The above ghost-prediction algorithm is tested on a 2D seismic line from North Sea, acquired over the Viking Graben (Keys & Foster, 1998). The shallow seafloor results in strong source and receiver ghosts. A set of common-offset gathers from the field data and the predicted ghosts is shown in Figure 6. As expected, the ghost sections (Figure 6b) are dominated by a few mono-frequency signals. In addition, note that the dominant frequency of the ghost decreases with larger offsets which results from the increasing length of the ghost raypath. These ghost events can also be seen on the recorded data (Figure 6a). In addition to the presence of ghosts, note the presence of surface-related and internal multiples in Figure 6b.

Prediction of only source-side (Figure 5b) or receiver-side (Figure 5c) ghosts can be done using a variation of equation 2 that is similar to the SRME integral. The single summation is over the location \( x_C \) in Figure 5b (to predict the source ghost) or over \( x_E \) in Figure 5c (to predict the source ghost) and involves only one convolution. For example, in Figure 5b, the source ghost is given by

\[
g(x_A, x_E, \omega) = \sum_{x_C} G(x_A, x_C, \omega) G(x_C, x_E, \omega). \tag{3}
\]

**CONCLUDING REMARKS**

Although not presented here, prediction of internal multiples should be followed by an adaptive subtraction method that eliminates these multiples from the data. The predicted ghosts could be suppressed in a similar fashion or using deconvolution techniques. Also, some interpretation of the predicted multiples might be necessary because of the generation of spurious events during the interferometry process. Additional processing steps might also be adopted to suppress these artifacts (Mehta, 2007). The zone of summation (aperture) is an important variable in the algorithm. Too small an aperture would not predict any multiples and too large an aperture might lead to the presence of primaries in the prediction. Depending on the complexity of the subsurface, a few tests might be necessary to find this parameter.

The ability to predict internal multiples and ghosts using one single algorithm makes our methodology versatile and powerful. Although our method is limited to gently dipping interfaces, it should be applicable in numerous fields, especially in many onshore unconventional plays.

**Acknowledgments**

We thank the developers/maintainers of freeDDS. Discussions with Roel Snieder and John Etgen were very useful. Support for this work was provided by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP and by the Center for Rock Abuse.
Internal multiples and ghosts prediction

Fig. 5: Illustration of the geometry behind (a) source+receiver, (b) source, and (c) receiver ghost computation.

Fig. 6: Near, mid, and far common-offset sections of the Viking Graben data (a) and the predicted source+receiver ghost (b). The arrow points to an interval containing internal multiples.
Internal multiples and ghosts prediction

References


Newton-Marchenko-Rose Imaging

Jyoti Behura†, Kees Wapenaar‡ and Roel Snieder‡
† Center for Wave Phenomena, Colorado School of Mines; ‡ Delft University of Technology

SUMMARY

Using only surface reflection data and first-arrival information, we generate up- and down-going wavefields at every image point using the algorithm of Rose (2002b,a) and Wapenaar et al. (2011, 2012a). An imaging condition is applied to these up- and down-going wavefields directly to generate the image. Since the above algorithm is based on exact inverse scattering theory, the reconstructed wavefields are accurate and contain all multiply scattered energy in addition to the primary event. As corroborated by our synthetic examples, imaging of these multiply scattered energy helps illuminate the subsurface better than reverse-time migration. We also demonstrate that it is possible to perform illumination compensation using our imaging algorithm that results in improved imaging at large depths.

INTRODUCTION

Wapenaar et al. (2011) propose a methodology for reconstructing the 3D impulse response for any “virtual source” in the subsurface using surface reflection data and the direct arrivals from the “virtual source” to the receivers on the surface. Their proposal is the 3D extension of the 1D iterative algorithm of Rose (2002b,a) who shows that in layered media, it is possible to focus all the energy at a particular time (or depth if the velocity is known) by using a complicated source signature. It is imperative to briefly discuss the pioneering work of V. Marchenko, R.G. Newton, and J.H. Rose on inverse scattering theory (Prosser, 1969; Gopinath & Sondhi, 1971; Burridge, 1980; Bojarski, 1981; Newton, 1989) that is instrumental in the development of the methodology of Rose (2002b,a) and Wapenaar et al. (2011, 2012a) and the imaging algorithm presented here. In 1D scattering theory, Marchenko’s integral equation (Marchenko, 2011) determines the relation between the wavefield in the interior of a medium and the reflected impulse response. Newton (1980, 1981, 1982) derived a similar relation, called as the Newton-Marchenko integral equation, that uses all scattered waves (reflected and transmitted) in 2D and 3D media. In 1D, this relation is given by

\[ u^{sc}(t, e, x) = \sum_{e' = -1, 1} R(t + e' x, -e', e) + \sum_{e' = -1, 1} \int_{-\infty}^{\infty} R(\tau + e' x, -e', e) u^{sc}(\tau, e', x) d\tau, \] (1)

where \( t \) is time, \( e \) is the direction of wave propagation, \( x \) is the 1D space, \( u^{sc} \) represents the scattered wavefield, and \( R \) is the impulse response function. A physical explanation of the above inverse scattering theory was provided by Rose (2002b,a) who showed that the ideas of focusing and time-reversal in fact result in the Newton-Marchenko equation.

The first step involved in solving for the scattering potential is to solve the Newton-Marchenko integral equation and find the wavefield everywhere inside the medium. Newton (1980, 1981, 1982) solved the inverse scattering problem for the Schrödinger wave equation by combining the Newton-Marchenko integral equation with high-frequency asymptotics. A significant breakthrough was made by Rose (2002b,a) who proposed an iterative approach, which he named ‘single-sided’ autofocusing, that determines the wavefield in 1D media by focusing the incident wave at a specified time. Rose proved that the incident wave that focuses the wavefield in the interior comprises of a delta function (band-limited in practice) followed by the time-reversed solution of the Marchenko equation. Rose’s algorithm was recently implemented on 1D seismic data by Broginni et al. (2011) who again show that a ‘virtual source’ response can be generated from surface reflection data alone. Besides extending Rose’s iterative algorithm to higher dimensions, Wapenaar et al. (2011, 2012a) also showed that the wavefield at any interior location can also be decomposed into the up- and down-going wavefields. Here, we show how the up- and down-going wavefields can be used directly for imaging the subsurface. In honor of the contribution of Newton, Marchenko, and Rose, we call the imaging algorithm introduced here as Newton-Marchenko-Rose Imaging, NMRI† in short. Besides demonstrating our imaging technique on synthetic examples, we discuss its advantages over existing imaging methods, in particular reverse-time migration (RTM).

ALGORITHM

Any seismic imaging algorithm consists of two steps - wavefield reconstruction and imaging condition. For example, RTM is a two-way imaging technique that utilizes wavefields reconstructed in time by accurately implementing the wave equation (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983) in a smooth velocity model. Wavefield reconstruction in RTM is followed by the application of an imaging condition (commonly cross-correlation) to image the reflectors. In NMRI, the up- and down-going wavefields are constructed at every location in space using the recipe of Wapenaar et al. (2011, 2012a). An imaging condition is applied to these two wavefields to obtain the image. The pseudo-code for NMRI is given in Algorithm 1.

†NMRI also stands for Nuclear Magnetic Resonance Imaging which is a medical imaging technique.
Newton-Marchenko-Rose Imaging

the incident wave; while in the absence of a reflector, no reflected waves are generated at the image point. Therefore, only at a reflector location the incident-wavefield coincides with the reflected-wavefield which gives rise to a non-zero zero-lag cross-correlation (Figure 1a).

Since the Newton-Marchenko equation is based on exact inverse scattering, the reconstructed wavefields contain all multiply-scattered energy. Note the presence of these multiples in Figures 1a and 1b. Any multiply-scattered wave that is incident on a scatterer will also have a corresponding scattered wave occurring at the same time. In Figure 1a, the multiply-scattered incident- and reflected-waves occur at the same time, while in Figures 1b, they do not coincide in time. Hence, in addition to the primary wavefield, all multiply-scattered energy will also be imaged accurately using NMRI. Other advantages of NMRI are discussed later.

NMRI IN ACTION

Here, we present three synthetic data results to demonstrate the performance and effectiveness of NMRI. The layer-cake (Figure 1) and Lena (Figure 3) models are constant velocity, variable density models while the fold model (Figure 2) has a variable velocity and constant density. The data acquisition is a fixed surface spread in each case where the sources and receivers are at \( z = 0 \text{m} \). The source and receiver spacing is 10 m in all the three acquisitions. Time sampling is also the same (0.004 s) in each case. The direct arrivals were muted from the shot gather. Besides this, no other processing was performed on the data; the data contain all orders of internal multiples.

Ray-tracing was used in computing the first breaks for the layercake model and for Lena; for the fold model, the first breaks were computed using a finite-difference wave-propagation code in a smoothed version (Figure 2b) of the true velocity model.

Note that NMRI produced a satisfactory image (Figures 1d, 2c, and 3d) in each case. Some steeply-dipping events were not imaged accurately for the fold model and for Lena because of the lack of illumination of these features. Application of a low-cut filter to Lena (Figure 3e) shows that many small details have been imaged in detail. Even though the reflection data is complicated (Figure 3b) and contains all orders of internal multiples, NMRI imaged the primary as well as all the scattered events appropriately.

ADVANTAGES OVER RTM

True amplitude/AVA: NMRI is based on exact inverse scattering theory and therefore the reconstructed wavefield in the interior of the medium is accurate irrespective of the velocity and density distributions in the subsurface. Hence, the NMRI image should be closer to the true reflectivity of the subsurface. Angle gathers for NMRI can be generated in the same way as in RTM. AVA analysis (on angle gathers) for NRMI, however, should be more reliable because the wavefields are accurate.

Multiples are imaged: As mentioned above, since the wavefields in NMRI are reconstructed accurately, the image should be better than existing imaging algorithms. Also, all orders of multiples are reconstructed and imaged accurately. Reflectors not illuminated by the direct arrival, might be illuminated by internal multiples (Fleury, 2012); these reflectors would be visible on the NMRI image but not on the RTM image. This is also corroborated by the virtual-source imaging of internal multiples of Wapenaar et al. (2011) and Wapenaar et al. (2012b). Imaging of multiples also renders multiple (both surface-related and inter-bed) prediction and suppression unnecessary. Although, none of the examples presented
Newton-Marchenko-Rose Imaging

![Fig. 2](image)

Fig. 2: (a) The original velocity model of the fold system and (b) the smoothed version used for imaging with the corresponding NMRI image (c). A constant density of 1 gm/cm³ was used in generating the reflection data.

Here contain surface-related multiples, the theory underlying NMRI imposes no limitations on the type of multiples.

Illumination compensation: Illumination compensation can be performed by manipulating the amplitude of the leading delta function in the incident wavefield such that each primary arrival at the image point has the same amplitude. By using the same unit delta function for all first arrivals, we ensure that the reflector is equally illuminated from all incidence angles and has the same illumination at all depths. However, if the Green’s function is used instead of a uniform delta function, the image at larger depths is poor because of insufficient illumination. For example, Figure 3c generated using the correct Green’s function (as the first arrival) can be interpreted as the best possible RTM image (generated from data containing only primary reflections). The above argument explains why the NMRI image (Figure 3d) is significantly better than the best possible RTM image (Figure 3c). Moreover, AVA would become more reliable as all the incident waves have the same amplitude.

Targeted imaging: Since the computation of the first breaks (using ray-tracing or finite-differences) can be done independently for each image point, it is possible to perform targeted imaging using NMRI. The targeted imaging of Lena’s left eye is shown in Figure 3f.

Computationally cheap: In NMRI, wavefield computation is done by convolving the incident wavefield with the reflected impulse response for a few iterations (in our tests two iterations were enough in most cases). If ray-tracing is used to compute the incident wavefield, then NMRI would be significantly faster than RTM. A thorough quantitative analysis is necessary to ascertain this.

High frequencies: The cost of RTM increases significantly with increasing frequency content because the extrapolation grid has to be more finely sampled. Wavefield computation using NMRI, on the other hand, has no such limitation because the frequency content of the incident wave and the impulse response are only limited by the temporal Nyquist limit.

Anisotropy: Wavefield extrapolation in anisotropic media using numerical methods is expensive. In addition, depending on the dispersion relation used, the wavefield can contain shear-wave artifacts and incorrect P-wave amplitudes. In NMRI, however, if the first breaks are computed using ray-tracing, then imaging in anisotropic media becomes extremely cheap compared to RTM. Moreover, the wavefields in NRMI are accurate in amplitude even if the medium exhibits velocity anisotropy.

DISCUSSION AND CONCLUSIONS

The first arrivals at the surface from an impulse at any image point can be computed either using ray-tracing or solving the wave equation numerically using finite-differences. In the case of ray-tracing, the incident wavefield can then be designed by convolving the traveltime with a delta function (Rose, 2002b,a). If the background velocity field results in multipathing, one must make sure that the incident wavefield contains all multiple arrivals; if not, the incident wave would not focus at the image point. However, if the first arrivals are computed by numerically solving the wave-equation, one must normalize each incident wave with its energy content to make sure that all incident waves have similar energy. In the absence of normalization, deeper reflectors will not be imaged properly. First arrivals computed using kinematic ray-tracing might not yield the correct phase; instead, dynamic ray-tracing or gaussian-beam modeling could be used instead. Ray-tracing, however, has one significant advantage: it is substantially cheaper than solving the wave equation numerically (especially in anisotropic media).

Newton-Marchenko-Rose Imaging, which is based on exact inverse scattering theory, shows promise in imaging complicated subsurfaces. Besides primaries, it can be used for illumination compensation and can image both surface-related and internal multiples. This should make NMRI useful for imaging poorly illuminated areas, especially underneath salt bodies. In comparison to RTM, NMRI has other important advantages, such as, it is potentially computationally cheaper, can image arbitrarily anisotropic media, can be used for targeted imaging, and should generate accurate AVA response.

Acknowledgments

The implementation of NMRI was done using freeDDS. Discussions with Farnoush Forghani and Filippo Broggini were very useful. Support for this work was provided by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.
Fig. 3: (a) The density model used in imaging Lena. (b) A sample shot gather. A constant velocity of 2000 m/s was used for modeling and imaging. The NMRI image after one (c) and two (d) iterations. In (c), the Green’s function is used for the first arrival while in (d), all the first arrivals are the same delta function. (e) Low-cut filtered version of (d). (f) Targeted imaging of Lena’s left eye.
Newton-Marchenko-Rose Imaging

References


Creating a virtual source inside a medium from reflection data with internal multiples: a stationary-phase analysis
Filippo Broggini1 and Roel Snieder, Center for Wave Phenomena - Colorado School of Mines; Kees Wapenaar, Delft Univ. of Technology

SUMMARY

Seismic interferometry allows one to create a virtual source inside a medium, assuming a receiver is present at the position of the virtual source. We discuss a method that creates a virtual source inside a medium from reflection data measured at the surface, without needing a receiver inside the medium and, hence, going beyond seismic interferometry. In addition to the reflection data, an estimate of the direct arrivals is required. However, no information about the medium is needed. We analyze the proposed method for a simple configuration using physical arguments based on the stationary-phase method and show that the retrieved virtual-source response correctly contains the multiples due to the inhomogeneous medium. The proposed method can serve as a basis for data-driven suppression of internal multiples in seismic imaging.

INTRODUCTION

We propose and discuss a new approach to reconstruct the response to a virtual source inside a medium, going beyond seismic interferometry. Controlled-source interferometric methods (Curtis et al., 2006; Bakulin and Calvert, 2006; Schuster, 2009) allow us to retrieve such a response without the need to know the medium parameters, but these methods require a receiver at the location of the virtual source in the subsurface and assume the medium is surrounded by sources. Our new approach removes the constraint of having a receiver at the virtual source location and is based on an extension of the 1D theory proposed earlier (Broggini et al., 2011, 2012b; Broggini and Snieder, 2012). They show that, given the reflection response of a 1D layered medium, it is possible to reconstruct the response to a virtual source inside the medium, without having a receiver at the virtual source location and without knowing the medium.

Wapenaar et al. (2011) made a first attempt to generalize the 1D method to 3D media. They used physical arguments to propose an iterative scheme that transforms the reflection response of a 3D medium into the response to a virtual source inside the unknown medium. Apart from the reflection data measured at the surface, our proposed method also requires an estimate of the direct arrivals between the virtual source location and the acquisition surface. These arrivals are a key element of the method, because they specify the location of the virtual source in the subsurface. For this reason, the proposed method is not fully model-independent. However, a model that relates the direct arrival to a virtual source position is simpler than a model that correctly handles the multiples. In the proposed method, the reflection data contributes to the multiple-scattering part of the virtual-source response.

As in seismic interferometry, our goal is to retrieve the response to a virtual source inside an unknown medium, removing the impact of a complex subsurface. This is helpful in situations where waves have traversed a strongly inhomogeneous overburden (e.g., in subsalt imaging, Sava and Biondi, 2004). Following the work of Wapenaar et al. (2012a), we analyze the iterative scheme for a simple 2D configuration. We use physical arguments based on the stationary-phase method to show that the method converges and allows for the reconstruction of the wavefield originating from the virtual source location.

STATIONARY-PHASE ANALYSIS

We use a geometrical approach to the method of stationary phase to solve the Rayleigh-like integrals, which yield the reflected response to an arbitrary incident field.

Configuration

We consider a configuration of three parallel dipping reflectors in a lossless, constant velocity, variable density medium (Figure 1). We choose a constant velocity medium because the response obeys simple analytical expressions. The proposed iterative scheme is, however, not restricted to constant velocity media. We denote spatial coordinates as $x = (x, z)$. The acquisition surface is located at $z = 0$ m and is transparent (i.e., the upper half-space has the same medium parameters as the first layer). The first dipping reflector obeys the relation $z = z_1 - ax$ with $z_1 = 2000$ m and $a = 1/3$. The black dot denotes the position of the virtual source, with coordinates $x_{VS} = (x_{VS}, z_{VS}) = (475, 3425)$ m. The second and third reflector are parallel to the first one, so that all mirror images of the virtual source lie on a line perpendicular to the reflectors. This line obeys the relation $z = z_1 + x/a$. The first, second, and third reflectors cross this line at $x_1 = (x_1, z_1)$.

---

1 Broggini et al., 2012b, submitted to Geophysics.

2 Wapenaar et al., 2012a, submitted to Geophysical Journal International.
Creating a virtual source from reflection data

Figure 2: Initial incident wavefield ($t < 0$) and its reflection response ($t > 0$), both measured at $z = 0$. Initial incident wavefield is the time-reversal of the direct arrivals. We show the reflection response only until 2.5 s. The solid black lines denote the time of the direct arrivals and its time-reversed counterpart. These lines are repeated in the subsequent figures.

$x_2 = (x_2, z_2)$, and $x_3 = (x_3, z_3)$, respectively. The velocity of the medium is constant and given by $c = 2000$ m/s. The densities in the four layers are $\rho_1 = \rho_2 = 1000$ kg/m$^3$, $\rho_3 = 5000$ kg/m$^3$, and $\rho_4 = 4000$ kg/m$^3$, respectively. The reflection coefficients for downgoing waves at the three interfaces are $r_n = (p_{n+1} - p_n)/(p_{n+1} + p_n)$, where $n = 1, 2, 3$ denotes the layer. The reflection coefficients for upgoing waves are $-r_n$. The transmission coefficients for downgoing ($\times$) and upgoing ($\ast$) waves are $t_n = 1 \pm r_n$. Since the velocity is constant in this configuration, the reflection and transmission coefficients hold for all the angles of incidence, not only for normal incidence. The large contrast between the density of the layers causes strong multiple reflections.

**Primary arrivals**

We introduce the Green’s function $G(x, x_0, t)$ as a solution of the wave equation $\nabla^2 G + \frac{\rho}{\mu} \frac{\partial G}{\partial t} = \delta(x - x_0) \delta(t)$, with $L = \rho \nabla \cdot (\rho^{-1} \nabla) - c^{-2} \frac{\partial^2}{\partial t^2}$. Defined in this way, the Green’s function is the response to an impulsive point source of volume injection rate at $x_0$ (de Hoop, 1995). Using the Fourier convention $\hat{F}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-j\omega t) dt$, the frequency domain Green’s function $\hat{G}(x, x_0, \omega)$ obeys the equation $\hat{G} = -j \omega \rho \delta(x - x_0)$, with $L = \rho \nabla \cdot (\rho^{-1} \nabla) + \omega^2/c^2$. Here $j$ is the imaginary unit, $\omega$ denotes the frequency domain. We write $\hat{G} = \hat{G}_d + \hat{G}_s$, where superscripts $d$ and $s$ stand for direct and scattered waves, respectively. As mentioned in the introduction, we need an estimate of the direct arrivals. For the constant velocity model of Figure 1, the high-frequency approximation of the Fourier transform of the direct Green’s function $\hat{G}_d(x, x_0, \omega)$ is given by

$$\hat{G}_d(x, x_0, \omega) = j\omega \frac{\exp(-j|\omega|/c)}{\sqrt{8\pi |\omega||x - x_0|/c}},$$

with $\mu = \text{sign}(\omega)$ (Snieder, 2004). The event with label 1 in Figure 2 shows the time-reversed version of the direct arrivals $\hat{G}_d(x, x_0, \omega) \ast s(t)$, where $s(t)$ is a Ricker wavelet with a central frequency of 20 Hz. It is essential that $s(t)$ is zero phase.

**Reflection response**

To retrieve the virtual-source response $\hat{G}(x, x_0, t)$, we need the reflection response at the surface $R(x, x_0, t) \ast s(t)$, in addition to an estimate of the direct arrivals. We assumed that the acquisition surface is transparent, hence the reflection response does not include any surface-related multiples. For this purpose, $R(x, x_0, t) \ast s(t)$ can be obtained from reflection data measured at $z = 0$ after surface-related multiple elimination (Veschuur et al., 1992). Following Bleistein (1984), the reflection response can be derived from a Rayleigh-type integral:

$$\hat{R}(x, x_0, \omega) \delta(x) = \int_{-\infty}^{\infty} \frac{\partial \hat{G}_d(x, x_0, \omega)}{\partial z} \hat{R}(\omega) \delta(x, \omega) \, dx,$$

for $z_R = z_0 = 0$ and after multiplying both sides by the spectrum of the source wavelet $s(t)$.

**Initiating the iterative process**

We define the initial incident wavefield at $z = 0$ as the time-reversed version of the direct arrivals at the recording surface excited by the virtual source $x_{VS}$. Hence, the initial wavefield is $\hat{p}_0^\times(x, t) = G_0^\times(x, x_{VS}, t) \ast s(t)$ and is shown in Figure 2 with the label 1. The subscript 0 of $\hat{p}_0^\times(x, t)$ indicates the initial wavefield (or the 0th iteration). In Figure 2, we also define two traveltime curves, indicated by the solid black lines. The upper curve is taken directly after the initial incident wavefield $\hat{p}_0^\times(x, t)$ and the lower curve is defined as the time-reversal of the upper curve. These two curves allow us to define a window function

$$w(x, t) = \begin{cases} 1 & \text{between the solid black lines of Figure 2} \\ 0 & \text{elsewhere}. \end{cases}$$

This window function is a key component of the iterative scheme. The reflected upgoing wavefield $\hat{p}_0^\times(x, t)$ is obtained by convolving the downgoing incident wavefield $\hat{p}_0^{-\times}(x, t)$ with the deconvolved reflection response and integrating over the source positions:

$$\hat{p}_0^\times(x, t) = \int_{-\infty}^{\infty} \hat{R}(x, x_0, t) \ast \hat{p}_0^{-\times}(x, t) \, dx,$$

for $z_R = z_0 = 0$. Equation (5) is the time-domain version of the Rayleigh integral described by equation (2). We discuss and solve this integral with geometrical arguments based on the method of stationary phase (a detailed mathematical proof is given by Wapenaar et al., 2012a).
Creating a virtual source from reflection data

Figure 4: Panel a and b show the first and second iteration of the incident wavefield (t < 0) and its reflection response (t > 0), respectively (both measured at z = 0). The numbers identify events in the total field. We show the reflection response only until 2.5 s.

The iterative process

We now discuss an iterative scheme, which uses the (k-1)th iteration of the reflection response \( p_{k-1}^r(x,t) \) to create the kth iteration of the incident field \( p_k^i(x,t) \). The objective is to iteratively update the incident field in such a way that, within the upper and lower solid black lines shown in Figure 2, the field is anti-symmetric in time. The meaning of this criterion will be evident in the next section, where we show how to reconstruct \( G(x, y, z, t) \). The method requires a combination of time reversal and windowing and the kth iteration of the incident field is defined by

\[
p_k^i(x,t) = p_{k-1}^i(x,t) - w(x,t)p_{k-1}^r(x,-t), \quad \text{for } x \text{ at } z = 0, \tag{6}
\]

where the time window \( w(x,t) \) is defined by equation (4). The reflection response is then obtained using equation (5), that we rewrite here as

\[
p_k^r(x_R,t) = \int_{-\infty}^{\infty} [R(x_R,x,t) * p_k^i(x,t)]_{z=0} \, dx, \tag{7}
\]

for \( x \) and \( x_R \) at \( z = 0 \). The first and second iteration of the incident and reflected fields are shown in Figures 4a and 4b, respectively. The events of \( p_2^i(x,t) \) labeled 2, 3, 4 in Figure 4b correspond to the events of \( p_1^i(x,t) \) labeled 6, 5, 4 in Figure 4a (time-reversed and multiplied by \(-1\)). For this particular configuration, the 8th iteration of the incident field (for \( k > 2 \)) is similar to \( p_2^i(x,t) \) and it is composed by four events, as shown in Figure 4b for \( t < 0 \). The events labeled 1 and 4 remain unchanged in the iterative process. The other two events (labeled 2 and 3) correspond to \(-A_k^2G^2(x_R, x_{VS}^{(2)}) \ast s(t)\) and \(-A_k^3G^3(x_R, x_{VS}^{(3)}) \ast s(t)\), respectively. The coefficient \( A_k^2 \) varies at each iteration and is equal to the partial sum of the geometric series \( c + cx + cx^2 + cx^3 + \cdots + cx^k \) where \( c = \frac{t_2^r \cdot t_3^r}{t_1^r}, t_2^r, t_3^r \) and \( x = \frac{t_n^r}{r_n^r} \). The sum of the series converges, because \( x < 1 \), and it yields

\[
\sum_{k=0}^{\infty} cx^k = \frac{c}{1-x} = r_2^r t_1^r, \tag{8}
\]

where we used that \( 1 - r_2^r = r_3^r t_4^r \). The coefficient of the third event, \( A_k^3 \) (labeled 3 in Figure 4b), is equal to \(-r_1A_k^2 \) and it converges to \(-r_1r_2^r t_1^r \). Figure 5a shows the thirtieth iteration and, within the time window \( w(x,t) \), the wavefield is antisymmetric in time. This is the result we predicted when we described the iterative method. The antisymmetry was actually the design criterion for the iterative scheme. This procedure is expected to converge because in each iteration the reflected energy is smaller than the incident energy. We consider the proposed method as a correction scheme that minimizes the energy inside the time window \( w(x,t) \).

Wavefield reconstruction from the virtual source

After showing that the method converged to the desired result, we define \( p_k^i(x,t) \) as the superposition of the \( k \)th version of the incident and reflected wavefields: \( p_k^i(x,t) = p_k^i(x,t) + p_k^r(x,t) \). Figures 2, 4a, 4b and 5a show \( p_k^i(x,t) \) for \( k = 0, 1, 2, \) and 30, respectively. For brevity, we define \( p(x,t) = p_{30}^i(x,t) \). We remind the reader that, within the solid black lines, the total field at \( z = 0 \) is antisymmetric in time and this particular feature was the design criterion for the iterative scheme. Consequently, if we stack the total field and its time-reversed version, i.e., \( p(x,t) + p(x,-t) \), all events inside the time window cancel each other, as shown in Figure 5b. Note that \( p(x,t) + p(x,-t) \) also obeys the wave equation because we consider a lossless medium. The causal part of this superposition corresponds to \( p^c(x,t) + p^c(x,-t) \) and the anti-causal part is equal to \( p^c(x,t) + p^c(x,-t) \), as shown in Figure 5b for \( t < 0 \) and \( t > 0 \), respectively. From a physical point of view, time-reversal changes the propagation direction, hence it follows that the causal part propagates upward at \( z = 0 \) and the anti-causal part propagates downward at \( z = 0 \). The first event of the causal part of Figure 5b has the same arrival time of the direct arrival of the response to the virtual source at \( x_{VS} \). If we combine this last observation with the fact that the causal part is upward propagating at \( z = 0 \), and that the total field obeys the wave equation in the inhomogeneous medium and is symmetric, it is plausible that the total field in Figure 5b is proportional to \( G(x,x_{VS},t) + G(x,x_{VS},-t) \). More precisely, we speculate that \( G(x,x_{VS},-t) \) and \( G(x,x_{VS},t) \) are proportional to the anti-causal and causal parts of Figure 5b, respectively. This deduction does not take into account any particular feature of the configuration used in this analysis, hence it should hold for more general situ-
Creating a virtual source from reflection data

Figure 5: (a) Thirtieth iteration of the incident wavefield ($t < 0$) and its reflection response ($t > 0$), both measured at $z = 0$. Within the solid black lines, the total field is antisymmetric in time and this particular feature was the design criterion for the iterative scheme. (b) Superposition of the total field and its time-reversed version after the method has converged. Here, unlike the previous figures, we show the wavefield for the time interval $-4 < t < 4$ s.

Let us check its validity for the response in Figure 5b. Following the steps that led to the field shown in Figure 5b, we find for the causal part

$$p_{30}^2(x, t) + p_{30}^3(x, -t) =$$

$$t_1^3t_2^3 \left\{ G^d(x, x_{V S}, t) + r_3G^d(x, x_{V S}, t) - r_2(r_1 + r_3)G^d(x, x_{V S}, t) + (r_1r_3 - r_2^2(r_1 + r_3))G^d(x, x_{V S}, t) \right\} * s(t),$$

with the virtual source position and its mirror images shown in Figure 1. For the configuration of Figure 1, this is proportional to the wavefield $G(x, x_{V S}, t) * s(t)$ originated from the virtual source and recorded at the surface (with $t_1^3t_2^3$ as the coefficient of proportionality). The directly modeled response to the virtual source is shown in Figure 6 and it matches the causal part of the field shown in Figure 5b.

for the total wavefield we obtain

$$p(x, t) + p(x, -t) = t_1^3t_2^3G_0(x, x_{V S}, t) * s(t),$$

where $G_0(x, x_{V S}, t) = G(x, x_{V S}, t) + G(x, x_{V S}, -t)$. We note that $G(x, x_{V S}, t)$ obeys the same wave equation as $G(x, x_{V S}, t)$, i.e.,

$$LG(x, x_{V S}, t) = -\rho_3 \delta(x - x_{V S}) \frac{\partial^2 \delta(t)}{\partial t^2} = \rho_3 \delta(x - x_{V S}) \frac{\partial^2 \delta(t)}{\partial t^2}.$$

So, $G_0$ obeys the homogeneous equation $LG_0 = -\rho_3 \delta(x - x_{V S}) \frac{\partial^2 \delta(t)}{\partial t^2} + \rho_3 \delta(x - x_{V S}) \frac{\partial^2 \delta(t)}{\partial t^2}$. This is in agreement with the fact that $p(x, t) + p(x, -t)$ has been constructed without introducing a singularity (i.e., a real source) at $x_{V S}$. $G_0$ is called the homogeneous Green’s function, after Porter (1970) and Oristaglio (1989) (but note that these authors take the difference instead of the sum of the causal and acasual Green’s functions because of a different definition of the source in the wave equation).

CONCLUSIONS

We proposed a generalization to 2D of the model-independent wavefield reconstruction method of Broggini et al. (2011, 2012b). The proposed data-driven procedure yields the response to a virtual source and reconstructs correct internal multiples, without needing a receiver at the virtual source location and without needing detailed knowledge of the medium. The method requires (1) the direct arriving wave front at the surface originated from a virtual source in the subsurface, and (2) the reflection impulse responses for all source and receiver positions at the surface. For a simple configuration, the stationary-phase analysis gives insight into the mechanism of the 2D iterative scheme and confirms that the method converges to the virtual-source response. Following the physical arguments in the previous section, it is plausible that the proposed methodology will also apply to more complex environments. The method will also have its limitations. The effects of a finite acquisition aperture, triplications, head waves, fine-layering, errors in the direct arrivals, etc. need further investigation. A numerical test by one of us with a variable-velocity syncline model shows promising results with respect to the handling of triplications (Broggini et al., 2012a). Errors in the estimated direct arrivals will cause defocusing and mispositioning of the virtual source (as in standard imaging algorithms). Such errors, however, do not affect the handling of the internal multiples and do not deteriorate their reconstruction, which is handled by the actual medium through the reflection data (that includes all the information about the medium itself). The errors may be serious when the virtual source is created anywhere. The virtual-source responses contain all internal multiples, hence the method could be used as a basis for imaging without internal multiples (Wapenaar et al., 2012b).

ACKNOWLEDGMENTS

This work was supported by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena and the Netherlands Research Centre for Integrated Solid Earth Science (ISES).

\*Broggini et al., 2012a, SEG Technical Program Expanded Abstracts, this issue.
Creating a virtual source from reflection data

REFERENCES


——–, 2012b, Focusing the wavefield inside an unknown 1D medium - Beyond seismic interferometry: submitted to Geophysics.


Oristaglio, M., 1989, An inverse scattering formula that uses all the data: Inverse Problems, 5, 1097–1105.


SUMMARY

Seismic interferometry is a technique that allows one to reconstruct the full response from a virtual source inside a medium, assuming a receiver is present at the virtual source location. We describe a method that creates a virtual source inside a medium from reflection data measured at the surface, without needing a receiver inside the medium and, hence, presenting an advantage over seismic interferometry. An estimate of the direct arriving wavefront is required in addition to the reflection data. However, no information about the medium is needed. We illustrate the method with numerical examples in a lossless acoustic medium with laterally-varying velocity and density and take into consideration finite acquisition aperture and a spatially-extended virtual source. We examine the reconstructed wavefield when a macro model is used to estimate the direct arrivals. The proposed method can serve as a basis for data-driven suppression of internal multiples in seismic imaging.

INTRODUCTION

We present and discuss a new approach to retrieve the full response from a virtual source $x_{VS}$ inside a medium and, consequently, to focus the wavefield at the virtual source location. Conventional methods for seismic interferometry (Curtis et al., 2006; Bakulin and Calvert, 2006; Schuster, 2009) allow one to reconstruct such a response without knowing the medium parameters, but these methods necessitate a receiver in the subsurface at the location of the virtual source and assume that sources surround the medium. The approach that we propose removes the constraint of having a receiver at the virtual source location and is based on a development of the 1D theory previously proposed by Broggini et al. (2011, 2012) and Broggini and Snieder (2012)

![Image](image.png)

Figure 1: (a) True velocity model with a syncline reflector. (b) Smooth velocity model used to model the first arrivals between the virtual source location $x_{VS}$ and the acquisition surface at $z=0$. In both panels, the black dot represents the location of the extended virtual source used in the first numerical example. The square inset (bottom right corner) shows the shape of the $\Lambda$ shaped source that replaces the black dot in the second example.

In our proposed approach, the reflection data contributes to the reconstruction of the multiple-scattering part of the virtual-source response.

Our objective is to retrieve the response originating from a virtual source inside an unknown medium, removing the imprint of a complex subsurface, as in seismic interferometry (Wapenaar et al., 2005; Curtis et al., 2006; Schuster, 2009). This is valuable in situations where waves have traveled inside a strongly inhomogeneous overburden, like a salt body (e.g., in subsalt imaging, Sava and Biondi, 2004). In this paper, we demonstrate that the requirement of having an actual receiver inside the medium can be circumvented, going beyond seismic interferometry.

We present numerical examples in a lossless acoustic medium with a syncline reflector, where the direct arrivals contain triplications. We discuss the influence of errors in the estimate of the first arrivals on the reconstructed wavefield. Such errors arise when a macro model (a routine product of velocity analysis) is used to compute the first arriving waveforms when such data are not available with other approaches, e.g., check shots or microseismic events. We then examine how the finite acquisition aperture and the limitations of the modeling code affect the results. While seismic interferometry usually deals with virtual point sources, this method allows us to reconstruct the response to a spatially-extended virtual source. This feature has a potential application in beam migration (Gray et al., 2009). Finally, we show that the proposed scheme implicitly reconstructs the incident field that fo-

---

1Broggini et al., 2012, submitted to Geophysics.
ITERATIVE PROCESS AND NUMERICAL EXAMPLES

We require that the wavefield focuses at a specific location, hence the proposed method is not totally independent of knowledge about the medium. The iterative scheme requires the reflection response of the medium measured at the surface, complemented with independent information about the primary arrivals originated from the focusing location, to focus the acoustic wavefield inside the medium. The primary arrival wavefront can be estimated or measured in various ways: by forward modeling using a macro model, directly from the data by the Common Focusing Point method (CFP) (Thorbecke, 1997) when the virtual source is located at an interface, from microseismic events (Artman et al., 2010), or from borehole check shots. We denote the 2D spatial coordinates as $x = (x, z)$. We assume that the reflection response does not include any multiples due to the free surface. Hence, $R(x, t; s(t))$ can be obtained from reflection data measured at the recording surface $z = 0$ after a surface-related multiple elimination processing (Verschuur et al., 1992), where $s(t)$ is a zero-phase wavelet.

We examine a configuration with a syncline reflector, whose velocity is shown in Figure 1a. The density profile (not shown) has a similar behavior and the densities of the three layers are $\rho_1 = 6500 \text{ kg/m}^3$, $\rho_2 = 1000 \text{ kg/m}^3$, and $\rho_3 = 7000 \text{ kg/m}^3$, from top to bottom layer. The direct arriving wavefront associated with this model contains a triplication. To start the iterative scheme, we compute the direct arrivals originating from the virtual source using the macro model of Figure 1b. This is a smooth version of the velocity model of Figure 1a. We define the initial incident downgoing wavefield $p_0^d(x, t)$ at $z = 0$ as the time-reversed version of the direct arrivals at the recording surface excited by the virtual source $x_{VS}$. The initial incident wavefield is shown in Figure 2a. The subscript 0 in $p_0^d(x, t)$ denotes the 0th iteration (initial) of the incident wavefield. Note that, due to the smoothing, the triplications are not present in this field (i.e., the time-reversed version of the direct arrivals). In Figure 2a, we also define two traveltimes the wavefield in time and space at the virtual source location.

![Figure 2: (a) Initial incident wavefield $p_0^d(x, t)$. (b) First iteration of the incident wavefield $p_1^d(x, t)$; this incident field will focus the wavefield at the virtual source location $x_{VS}$ (black dot) in Figure 1a at $t = 0$. We show the wavefield for $-2 < t < 2$ s only.](image)

![Figure 3: (a) Causal part of the superposition of the total field and its time-reversed version, $p_1(x, t) + p_1(x, -t)$. Label (1) indicates an incomplete cancellation of the events inside the window $w(x, t)$. Label (2) shows the effect of edge diffraction due to finite aperture. (b) Directly-modeled full response to the virtual source (black dot) shown in Figure 1a.](image)

We require that the wavefield focuses at a specific location, hence the proposed method is not totally independent of knowledge about the medium. The iterative scheme requires the reflection response of the medium measured at the surface, complemented with independent information about the primary arrivals originated from the focusing location, to focus the acoustic wavefield inside the medium. The primary arrival wavefront can be estimated or measured in various ways: by forward modeling using a macro model, directly from the data by the Common Focusing Point method (CFP) (Thorbecke, 1997) when the virtual source is located at an interface, from microseismic events (Artman et al., 2010), or from borehole check shots. We denote the 2D spatial coordinates as $x = (x, z)$. We assume that the reflection response does not include any multiples due to the free surface. Hence, $R(x, t; s(t))$ can be obtained from reflection data measured at the recording surface $z = 0$ after a surface-related multiple elimination processing (Verschuur et al., 1992), where $s(t)$ is a zero-phase wavelet.

We examine a configuration with a syncline reflector, whose velocity is shown in Figure 1a. The density profile (not shown) has a similar behavior and the densities of the three layers are $\rho_1 = 6500 \text{ kg/m}^3$, $\rho_2 = 1000 \text{ kg/m}^3$, and $\rho_3 = 7000 \text{ kg/m}^3$, from top to bottom layer. The direct arriving wavefront associated with this model contains a triplication. To start the iterative scheme, we compute the direct arrivals originating from the virtual source using the macro model of Figure 1b. This is a smooth version of the velocity model of Figure 1a. We define the initial incident downgoing wavefield $p_0^d(x, t)$ at $z = 0$ as the time-reversed version of the direct arrivals at the recording surface excited by the virtual source $x_{VS}$. The initial incident wavefield is shown in Figure 2a. The subscript 0 in $p_0^d(x, t)$ denotes the 0th iteration (initial) of the incident wavefield. Note that, due to the smoothing, the triplications are not present in this field (i.e., the time-reversed version of the direct arrivals). In Figure 2a, we also define two traveltimes the wavefield in time and space at the virtual source location.

![Figure 2: (a) Initial incident wavefield $p_0^d(x, t)$. (b) First iteration of the incident wavefield $p_1^d(x, t)$; this incident field will focus the wavefield at the virtual source location $x_{VS}$ (black dot) in Figure 1a at $t = 0$. We show the wavefield for $-2 < t < 2$ s only.](image)

![Figure 3: (a) Causal part of the superposition of the total field and its time-reversed version, $p_1(x, t) + p_1(x, -t)$. Label (1) indicates an incomplete cancellation of the events inside the window $w(x, t)$. Label (2) shows the effect of edge diffraction due to finite aperture. (b) Directly-modeled full response to the virtual source (black dot) shown in Figure 1a.](image)

the upcoming reflection response $p_0^r(x, t)$ is obtained either by injecting the downgoing incident wavefield $p_0^d(x, t)$ into the actual medium or by convolving the downgoing incident wavefield $p_0^d(x, t)$ with the deconvolved reflection response and integrating over the source positions:

$$p_k^d(x, t) = \int_{-\infty}^{\infty} R(x_R, x, t) \ast p_k^d(x, t) \, dx,$$

for $x$ and $x_R$ at $z = 0$, and $k = 0$. Equation (2) is a Rayleigh-type integral (Wapenaar and Berkhout, 1993, equation 15a).

We discuss an iterative scheme that uses the $(k-1)$th iteration of the reflected wavefield $p_{k-1}^r(x, t)$ to construct the $k$th iteration of the downgoing incident field $p_k^d(x, t)$. The aim is to iteratively update the incident field $p_k^d(x, t)$ in such a way that, within the upper and lower dashed black lines shown in Figure 2a (i.e., inside the time window), the field is anti-symmetric in time. The meaning of this criterion will be clear in the remainder of this work, where we show the reconstruction of the response to a virtual source inside $x_{VS}$. The proposed method uses a combination of time reversal and time windowing to construct the next iteration of the incident field. The $k$th iteration of the incident field $p_k^d(x, t)$ is specified by

$$p_k^d(x, t) = p_0^d(x, t) - w(x, t) \int_{0}^{t} p_{k-1}^r(x, t') \, dt', \quad \text{for } x \text{ at } z = 0,$$

where the time window $w(x, t)$ is defined by equation (1). We compute the initial response $p_0^d(x, t)$ injecting $p_0^d(x, t)$ into the actual medium, apply the window function $w(x, t)$, and construct the next iteration of the incident field: $p_1^d(x, t) = p_0^d(x, t) - w(x, t) \int_{0}^{t} p_0^r(x, t') \, dt'$.
the first iteration of the updated incident wavefield \( p_1^+ (x, t) \) (see Figure 2b).

We define the superposition of the \( k \)th version of the incident and reflected wavefields as \( p_k (x, t) = p_k^i (x, t) + p_k^- (x, t) \). Also, we define \( p(x, t) \) as the final result of the iterative process. Note that, within the dashed black lines, the total field at \( z = 0 \) is antisymmetric in time and this particular feature was the design criterion for the iterative scheme. Hence, if we stack the total field and its time-reversed version, i.e., \( p(x, t) + p(x, -t) \), all events inside the time window cancel each other. The causal part of \( p(x, t) \) is equal to \( p^- (x, t) + p^+ (x, -t) \) and the anti-causal part corresponds to \( p^+ (x, t) + p^- (x, -t) \). From a physical point of view, time-reversal switches the direction of propagation, hence it follows that the causal part is upward propagating at \( z = 0 \) and the anti-causal part is downward propagating at \( z = 0 \).

For this particular configuration, we achieve the final result after only one iteration \( (k = 1) \). We form the field \( p(x, t) + p(x, -t) = p_1^+ (x, t) + p_1^- (x, -t) \) to reconstruct the response originating from the virtual source location. The causal part of this field is shown in Figure 3a. Labels (1) and (2) indicate an incomplete cancellation of events. This is possibly due to numerical limitations of the modeling code (e.g., numerical dispersion and influence of spatial-tapering on the incident field). The amplitude of the data shown in Figure 3 are clipped to 70% of the maximum amplitude and this enhances the features indicated by labels (1) and (2). The first event below the dashed curve of Figure 3a has the same arrival time of the direct arrival of the response to the virtual source at \( x_{VS} \). If we combine this last reasoning with the fact that the causal part propagates upward at \( z = 0 \), and that the total field is symmetric and obeys the wave equation in the inhomogeneous medium, it is reasonable that the total field in Figure 3a is proportional to the response due to a real source placed at the virtual source location \( x_{VS} \), as shown in Figure 3b. Work on a derivation of this principle is in progress.

The comparison between the two panels of Figure 3 shows that it is possible to reconstruct the full response to a virtual source inside the medium, including all multiples, using the direct arrivals computed using a smooth model. Note that this procedure is expected to converge because in each iteration the reflected energy is smaller than the incident energy. We interpret the proposed iterative method as a correction scheme that minimizes the energy of the wavefield inside the time window \( w(x, t) \). Moreover, the proposed method does not take into account any particular feature of the model used in this analysis, hence it should hold for more general situations.

We simulate the propagation of the updated incident field \( p_1^+ (x, t) \) inside the medium of Figure 1a. Figure 4 displays six snapshots extracted from this simulation. Panel d shows that the wavefield collapsed to a focus at the location indicated by the black dot in Figure 1a. We note that the approximate direct arrivals (computed with the smooth model) used to start the iterative process caused an imperfect virtual source as indicated by the artifacts around the virtual source. Note that the complex wavefield propagating inside the syncline (panels a,b,c,e, and f) annihilates at the focusing time as shown in panel d.

Finally, we repeat similar steps for the same velocity model but Figures 5a and 5b shows that the wavefield evolves in time and is reflected at the virtual source location in Figure 1a at \( t = 0 \), when the \( \wedge \)-shaped virtual source replaces the black dot. We show the wavefield for the time interval \( -2 < t < 2 \) s.

![Figure 4](image4.png)

**Figure 4**: Time snapshots extracted from the propagating wavefield when the field of Figure 2b is used as a source. Panel d shows the wavefield focused at the location indicated by the black dot in Figure 1a. Time is increasing from panel a to f.

![Figure 5](image5.png)

**Figure 5**: (a) Initial incident wavefield \( p_0^+ (x, t) \); (b) First iteration of the incident wavefield \( p_1^+ (x, t) \); this incident field will focus the wavefield at the virtual source location in Figure 1a at \( t = 0 \), when the \( \wedge \)-shaped virtual source replaces the black dot. We show the wavefield for the time interval \( -2 < t < 2 \) s.
now use a spatially-extended virtual source, i.e., a \( \wedge \)-shaped source as shown in the square inset of Figure 1a. The \( \wedge \)-shaped source replaces the black dot used previously. In contrast to the first example, we compute the first arrivals using the true model instead of the smooth macro model. We do this to show that the quality and degree of focusing at the virtual source location depends on the first arrivals. Since the direct arrivals are now exact, Figure 7d shows that the wavefield focuses well on the target source. The proposed scheme produces a better result with less artifacts in this situation (compared to Figure 4d), because we have a correct estimate of the first arriving wavefront. Furthermore, the reconstructed response shown below the dashed curve in Figure 6a is virtually identical to the directly-modeled response shown in Figure 6b.

CONCLUSIONS

We discussed a generalization to two dimensions of the model-independent wavefield focusing and reconstruction method of Broggini et al. (2011, 2012) and Broggini and Snieder (2012). Unlike the 1D method, which uses the reflection response only, the proposed multi-dimensional extension requires, in addition to the reflection response, independent information about the first arrivals.

The proposed data-driven procedure yields the response to a virtual source (Figures 3a, 6a), removes the imprint of the subsurface (as the virtual source method), and reconstructs internal multiples, without needing a receiver at the virtual source location and without needing detailed knowledge of the medium. The method requires (1) the direct arriving wave front at the surface originated from a virtual source in the subsurface, and (2) the reflection impulse responses for all source and receiver positions at the surface. The direct arriving wave front can be obtained by modeling in a macro model, directly from the data by the CFP method (Berkhout, 1997) when the virtual source is located at an interface, from microseismic events (Artman et al., 2010), or from borehole

Figure 6: (a) Causal part of the superposition of the total field and its time-reversed version, \( p_1(x,t) + p_1(x,-t) \). Label (1) indicates an incomplete cancellation of the events inside the window \( w(x,t) \). (b) Directly-modeled full response to the \( \wedge \)-shaped virtual source shown in Figure 1a.

Figure 7: Time snapshots extracted from the propagating wavefield when the field of Figure 5b is used as a source. Panel d shows the wavefield focused at the location indicated by the \( \wedge \)-shaped symbol in Figure 1a. Time increases from panel a to f.

check shots. In the first numerical example, we used a smooth version of the true model to compute the direct arrivals. The required reflection impulse responses are obtained from seismic reflection data after surface-related multiple elimination (Verschuur et al., 1992) and deconvolution for the source wavelet.

We showed that the method is not limited to point sources, but also handles spatially-extended virtual sources. This feature permits to create and steer oriented beams originating at depth under a complex overburden that generates strong multiples. This beamforming process could have a potential use in beam migration (Gray et al., 2009). Errors in the estimated first arrivals (due to a smooth macro model) cause defocusing and a mislocalization of the virtual source (similar as in standard imaging algorithms). Such errors, however, do not affect the handling of the internal multiples and do not deteriorate their reconstruction, which is handled by the actual medium through the reflection data measured at the surface (that includes all the information about the medium itself). Note that also the virtual source method (Bakulin and Calvert, 2006) does not optimally focus the wavefield at the virtual source location, but Wapenaar et al. (2011b) show that the focusing can be improved applying multi-dimensional deconvolution. Furthermore, because the proposed method is non-recursive, the reconstruction of internal multiples will not suffer from error propagation, unlike other internal multiple suppression techniques used in seismic imaging.

ACKNOWLEDGMENTS

This work was supported by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena and the Netherlands Research Centre for Integrated Solid Earth Science (ISES).
REFERENCES


——–, 2012, Focusing the wavefield inside an unknown 1D medium - Beyond seismic interferometry: submitted to Geophysics.


Joint migration velocity analysis of PP- and PS-waves for VTI media
Pengfei Cai* and Ilya Tsvankin, Center for Wave Phenomena, Colorado School of Mines

SUMMARY
Combining PP-waves with mode-converted PS-waves in migration velocity analysis (MVA) can help build more accurate VTI (transversely isotropic with a vertical symmetry axis) velocity models. To take advantage of efficient MVA algorithms designed for pure modes, here we generate pure SS-reflections from PP and PS data using the PP+PS=SS method. Then the residual moveout in both PP and SS common-image gathers is minimized during iterative velocity updates. The model is divided into square cells, with the P- and S-wave vertical velocities ($V_{P0}$ and $V_{S0}$) and the anisotropy parameters $\epsilon$ and $\delta$ defined at each grid point. The objective function also includes the differences between the migrated depths of the same reflectors on the PP and SS sections. The replacement of PS-waves with pure SS reflections in MVA allows us to avoid problems caused by the moveout asymmetry and other undesirable features of mode conversions. Synthetic examples confirm that 2D MVA of PP- and PS-waves can resolve all four relevant parameters of VTI media if reflectors with at least two distinct dips are available. After the velocity model has been reconstructed, accurate depth images can be obtained by migrating the recorded PP and PS data.

INTRODUCTION
To resolve the velocity $V_{P0}$ and anisotropy parameters $\epsilon$ and $\delta$ required for P-wave depth imaging, it is necessary to combine P-wave traveltimes with additional information. Tsvankin and Thomsen (1995) demonstrate that long-spread (nonhyperbolic) P- and SV-wave moveouts are sufficient for estimating the parameters $V_{P0}$, $V_{S0}$, $\epsilon$, and $\delta$. However, it is more practical to supplement P-waves with converted PS (PSV) data. Several authors discuss joint tomographic inversion of PP- and PS-waves (Stopin and Ehinger, 2001; Audebert et al., 1999; Broto et al., 2003; Foss et al., 2005). However, velocity analysis of mode conversions is hampered by the asymmetry of PS moveout. Therefore, MVA for PS-waves (Du et al., 2012; Foss et al., 2005; Audebert et al., 1999) has to account for the “diódic” nature of PS reflections (Thomsen, 1999). For example, common-image gathers (CIGs) of PS-waves can be computed separately for positive and negative offsets (Foss et al., 2005). To replace mode conversions in velocity analysis with pure SS reflections, Grechka and Tsvankin (2002) suggest the so-called PP+PS=SS method. By combining PP and PS events that share P-legs, that method generates SS reflection data with the correct kinematics. Grechka et al. (2002b) perform inversion of multicomponent (PP and PS) data using stacking-velocity tomography, which operates with NMO velocities for 2D lines and NMO ellipses for wide-azimuth 3D surveys (Grechka et al., 2002a). They compute SS traveltimes with the PP+PS=SS method and then apply stacking-velocity tomography to the PP and SS moveouts.

Figure 1: Matching the horizontal slownesses on common-receiver PP and PS gathers at locations $x_1$ and $x_2$ helps find the source-receiver coordinates $x_3$ and $x_4$ of the pure SS ray $x_3Rx_4$. This constructed SS ray has the same reflection point R as the PP ray $x_1Rx_2$ and PS rays $x_1Rx_3$ and $x_2Rx_4$ (Grechka and Tsvankin, 2002).

To make use of efficient MVA techniques developed for pure modes (Sarkar and Tsvankin, 2004; Wang and Tsvankin, 2011), here we apply the PP+PS=SS method to construct pure SS-wave reflections from PP and PS data. The MVA is performed by minimizing residual moveout of reflection events in both PP- and SS-wave CIGs. PP and SS images of the same reflector do not match in depth if the velocity model is incorrect. Therefore, in addition to flattening image gatherings, we penalize depth misties between PP and SS sections.

METHODOLOGY
To build a VTI model for depth migration using P-wave reflection moveout, at least one medium parameter (e.g., $V_{P0}$) must be known a priori (Sarkar and Tsvankin, 2004). As discussed in Tsvankin and Grechka (2000), combining P-wave traveltimes with the moveout of PS-waves converted at a horizontal and dipping interface can help constrain the velocities $V_{P0}$ and $V_{S0}$ and the parameters $\epsilon$ and $\delta$. The most significant problem in PS-wave velocity analysis is the moveout asymmetry with respect to zero offset in common-midpoint (CMP) geometry (Tsvankin and Grechka, 2011). Therefore, MVA designed for pure modes cannot be directly applied to converted waves. Here, we employ the PP+PS=SS method (Grechka and Tsvankin, 2002) to produce pure SS reflection events from PP and PS reflections.

Implementation of the PP+PS=SS method requires event registration, or identification of PP and PS events from the same interfaces. The main idea of the method is to combine PP and PS reflections that share the same P-legs. This is done by matching time slopes (horizontal slownesses) on common-receiver gathers of PP- and PS-waves (Figure 1). Then the traveltime of the constructed SS wave (sometimes called the “pseudo-S”
Figure 2: Critical (maximum) offset of the constructed SS-waves in a horizontal layer (Tsvankin and Grechka, 2011).

arrival is found from:

$$t_{SS}(x_3,x_4) = t_{PS}(x_1,x_3) + t_{PS}(x_2,x_4) - t_{PP}(x_1,x_2).$$  \hspace{1cm} (1)

While the generated SS-wave traveltimes are exact, the method cannot produce correct reflection amplitudes. Hence, we convolve the SS traveltimes with a Ricker wavelet to generate “pseudo” SS reflection data to be used for MVA. The maximum reflection angle of the shear wave generated by mode conversion in a horizontal isotropic layer with the P- and S-wave velocities $V_P$ and $V_S$ is $\theta_S^{crit} = \sin^{-1}(V_S/V_P)$, and the half-offset $h_S$ cannot exceed the critical value,

$$h_S^{crit} = D \tan \left[ \sin^{-1} \left( \frac{V_S}{V_P} \right) \right];$$  \hspace{1cm} (2)

D is the layer’s thickness. For example, for a typical velocity ratio $V_S/V_P = 1/2$, the maximum half-offset is less than 0.6D. Therefore, it is necessary to include long-offset PP and PS data to generate SS-waves suitable for robust velocity analysis. The offset range for the computed SS-waves is more narrow than that for the acquired PP and PS data but may be sufficient for MVA if the survey includes offsets reaching two target depths.

To perform velocity analysis of PP- and SS-waves, we extend the MVA algorithm of Wang and Tsvankin (2011) to multi-component data. The model is divided into square cells, and the parameters $V_{P0}$, $V_{S0}$, $\epsilon$, and $\delta$ are defined at each grid point. We apply prestack Kirchhoff depth migration to both PP and SS waves, typically starting with an initial isotropic velocity model. The moveouts of migrated PP and SS events in common-image gathers serve as input to the joint MVA. To constrain the anellipticity coefficient $\eta$ (and, therefore, $\epsilon$), the moveout in PP-wave CIGs is described by the nonhyperbolic equation (Sarkar and Tsvankin, 2004):

$$z^2(h) = z^2(0) + Ah^2 + B \frac{h^4}{h^2 + z^2(0)},$$  \hspace{1cm} (3)

where $z$ is the migrated depth as a function of the half-offset $h$, and the coefficients $A$ and $B$ are found by a 2D semblance scan. There is no need to apply equation 3 to SS-wave CIGs because the offset-to-depth ratio for the constructed SS events seldom exceeds 1-1.2. For the joint MVA, we not only minimize the residual moveout in PP and SS CIGs, but also perform coding, which involves tying PP and SS images of the same reflectors. The objective function includes a term that penalizes the mismatch in depth of migrated PP and SS images using a selection of key reflection points. Those points are chosen on the basis of coherency and focusing (Foss et al., 2005).

To carry out velocity update, it is necessary to compute traveltyme derivatives with respect to the model parameters (Wang and Tsvankin, 2011). The exact P- and SV-wave phase velocities in VTI media can be expressed as (Tsvankin, 2005):

$$\frac{V^2}{V_P^2} = 1 + \epsilon \sin^2 \theta - \frac{f}{2} \pm \frac{f}{2} \sqrt{\left( 1 + 2\epsilon \sin^2 \theta \frac{f}{2} \right) - \frac{2(\epsilon - \delta) \sin^2 2\theta}{f}},$$  \hspace{1cm} (4)

where $f = 1 - V_{SO}^2/V_P^2$ and $\theta$ is the phase angle with the symmetry axis. The plus in front of the radial corresponds to P-waves and minus to SV-waves. For purposes of MVA, it is convenient to replace $\epsilon$ and $\delta$ with the P-wave horizontal ($V_{hor,p}$) and NMO ($V_{nmo,p}$) velocities given by $V_{hor,p} = V_{P0} \sqrt{T + 2\epsilon}$ and $V_{nmo,p} = V_{P0} \sqrt{T + 2\epsilon}$. Therefore, we compute the traveltyme derivatives with respect to $V_{hor,p}$ and $V_{nmo,p}$ instead of $\epsilon$ and $\delta$. The MVA algorithm updates $V_{P0}$, $V_{S0}$, $V_{hor,p}$, and $V_{nmo,p}$ and then converts $V_{hor,p}$ and $V_{nmo,p}$ into $\epsilon$ and $\delta$.

Following Sarkar and Tsvankin (2004), the variance $\text{Var}$ of the migrated depths can be written as

$$\text{Var} = \sum_{j=1}^{U} \sum_{k=1}^{M} \left[ z(x_j,h_k) - \hat{z}(x_j) \right]^2;$$  \hspace{1cm} (5)

where $U$ is the number of image gathers used in the velocity update, $M$ is the number of offsets in the image gathers, $x_j$ is the midpoint, $h_k$ is the half-offset, and the average migrated depth at $x_j$ is $\hat{z}(x_j) = (1/M) \sum_{k=1}^{M} z(x_j,h_k)$. By minimizing the variances $\text{Var}_P$ (for P-waves) and $\text{Var}_S$ (for S-waves), we flatten CIGs for both modes. The difference between the migrated depths of PP- and SS-waves from the same reflector can be estimated as:

$$\text{Var}_{P-S} = \sum_{j=1}^{U} \sum_{k=1}^{M} \left[ z_P(x_j) - z_S(x_j) \right]^2.$$  \hspace{1cm} (6)

Minimizing the variance $\text{Var}_{P-S}$ makes it possible to tie the PP and SS sections in depth.

The objective function is defined as follows:

$$F(\Delta \lambda) = \mu_1 \left[ |A_P \Delta \lambda + b_P|^2 + |A_S \Delta \lambda + b_S|^2 \right] + \mu_2 \left[ |D \Delta \lambda + y|^2 + \xi |L \Delta \lambda|^2 \right],$$  \hspace{1cm} (7)

where $A_P$ and $A_S$ depend on the derivatives of the PP and SS migrated depths with respect to the medium parameters, the vectors $b_P$ and $b_S$ contain elements that characterize the residual moveout in PP- and SS-wave CIGs, the matrix $D$ describes the differences between the derivatives of the PP and SS migrated depths with respect to the medium parameters, and the vector $y$ contains the differences between the migrated depths.
Figure 3: VTI model with dipping interfaces. The parameters of the top isotropic layer are $V_P = 2000$ m/s and $V_S = 1000$ m/s; for the second layer, $V_{P0} = 3000$ m/s, $V_{S0} = 1500$ m/s, $\varepsilon = 0.2$ and $\delta = 0.1$. The maximum dip of both reflectors is $27^\circ$.

of the same reflectors on the PP and SS sections. The coefficients $\mu_1$, $\mu_2$, $\mu_3$ and $\zeta$ govern the weights of the terms responsible for flattening PP- and SS-wave CIGs ($||A_P \Delta \lambda + b_P||^2$ and $||A_S \Delta \lambda + b_S||^2$), codephasing ($||D \Delta \lambda + y||^2$), and regularization ($||L \Delta \lambda||^2$). The objective function is minimized by a least-squares algorithm.

SYNTHETIC TEST

We use anisotropic ray-tracing package ANRAY to compute PP- and PS-wave reflection traveltimes. PP and SS images are generated with Kirchhoff prestack depth migration (Seismic Unix program ‘sukdmig2d’). To create traveltime tables for migration of PP- and SS-waves, we perform ray tracing using SU code ‘rayt2dan’.

The algorithm is tested on a model (Figure 3) that includes a reflector (e.g., a fault plane) dipping at an angle approaching $30^\circ$ near the surface. To avoid instability in ray tracing, we smooth the corner of the dipping interface using bicubic spline interpolation. The synthetic data include PP and PS reflections from the top and bottom of the VTI layer (Figure 4). The maximum offset-to-depth ratio for the bottom reflector is close to two, which is sufficient for applying the PP+PS=SS method. Indeed, the maximum offset for recorded PP data is 4 km and the maximum offset for constructed SS data is 2 km. The whole PP data set is used along with the SS-waves to estimate the residual moveout in CIGs. For codephasing, however, we only use conventional-spread PP data with offsets not exceeding those for SS-waves. Tsvankin and Grechka (2000, 2011) demonstrate that the traveltimes of the PP- and PS-waves reflected from a horizontal and a moderately dipping interface are sufficient to constrain the parameters $V_{P0}$, $V_{S0}$, $\varepsilon$, and $\delta$.

Here we invert only for the parameters of the middle layer. We assume that layer to be weakly heterogeneous and apply strong regularization (the operator $L$ in equation 7). The ini-

Figure 4: CMP gathers of the recorded (a) PP-waves and (b) PS-waves at location 3000 m. (c) The SS data constructed by the PP+PS=SS method at the same location.

Figure 5: Common-image gathers of (a) PP-waves and (b) SS-waves (displayed every 100m) after migration with the initial model.
The initial model is isotropic with the velocities $V_{P0}$ and $V_{S0}$ distorted by 15%. Figures 5(a) and 5(b) show that the PP- and SS-wave image gathers obtained with the initial model exhibit substantial residual moveout. Also, the bottom of the VTI layer is not imaged at the correct depth (Figures 6(a) and 6(b)).

A set of CIGs of PP- and SS-waves from both the horizontal and dipping interfaces (for locations from 1 to 4 km) serve as the input to the joint MVA. Since we use gridded tomography, the derivatives of migrated depths with respect to the model parameters are calculated at the vertices of relatively fine grids. Therefore, we follow Wang and Tsvankin (2011) in employing a mapping matrix to convert the model updates into the parameter values at each grid point. After 11 iterations, the image gathers are flat (Figures 7(a) and 7(b)) and the reflectors on the PP and SS sections are correctly positioned (Figures 8(a) and 8(b)). The inversion produces accurate estimates of the interval VTI parameters: $V_{P0} = 3031$ m/s, $V_{S0} = 1515$ m/s, $\epsilon = 0.20$, and $\delta = 0.08$. These results confirm the feasibility of building the VTI depth model using 2D PP and PS reflection data if both horizontal and dipping events are available. We will also discuss application of the algorithm to more complicated VTI models with pronounced lateral heterogeneity.

CONCLUSIONS

In the presence of moderate dips, converted PS data may provide essential information for estimating VTI velocity models in depth. Here, we presented an efficient algorithm for joint velocity analysis of PP- and PS-waves from heterogeneous VTI media. After constructing pure SS reflections with the PP+PS=SS method, we update the velocity model by flattening PP- and SS-wave image gathers and tying the PP and SS migrated images in depth. It is essential to acquire long-spread (with the maximum offset-to-depth ratio reaching two) PP and PS data, which ensures that the offset range of the computed SS-waves is sufficient for robust velocity analysis. To better constrain the parameters $\eta$ and $\epsilon$, the residual move-out in PP-wave image gathers is evaluated using nonhyperbolic semblance scan.

Synthetic testing for layered VTI media confirmed that if both horizontal and moderately dipping PP and PS events are available, the joint MVA converges toward the correct depth model. Therefore, PS-waves can play an important role in velocity model-building for prestack depth migration. Multicomponent data also provide accurate estimates of the shear-wave vertical velocity, which can be used in lithology prediction and reservoir characterization.

ACKNOWLEDGMENTS

We are grateful to James Gaiser (Geokinetics) for helpful discussions and to the members of the A(nisotropy)-Team of the Center for Wave Phenomena (CWP) for valuable technical assistance. Support for this work was provided by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.
If a sharp boundary is present in the model, we can decompose the source wavefield into transmitted and reflected energy that originates at the sharp boundary:

\[ W_s(x, t) = W_t^s(x, t) + W_r^s(x, t), \]  

(2)

where the superscripts \( t \) and \( r \) stand for transmitted and reflected energy, respectively. The same idea can be applied to the receiver wavefield:

\[ W_r(x, t) = W_t^r(x, t) + W_r^r(x, t). \]  

(3)

By taking advantage of the linearity of equation 1, we can split the conventional imaging condition as follows:

\[ R(x) = R^{tt}(x) + R^{rr}(x) + R^{rt}(x) + R^{tr}(x). \]  

(4)

Here, the first superscript is associated with the source wavefield and the second is associated with the receiver wavefield. For example, \( R^{tr}(x) \) is an image constructed with the transmitted source wavefield and the reflected receiver wavefield.

This analysis can be used as well with the extended imaging condition (EIC) (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011). The EIC is similar to the CIC except that the cross-correlation lags between source and receiver wavefields are preserved in the output as follows:

\[ R(x, \lambda, \tau) = \sum_{s} \sum_t W_s(x - \lambda, t - \tau) W_r(x + \lambda, t + \tau). \]  

(5)

Here \( \lambda \) and \( \tau \) represent the cross-correlation space-lags and time-lags, respectively. The conventional image is a special case of the extended image \( R(x) = R(x, 0, 0) \).

By using extended images, we can measure the accuracy of the velocity model by analyzing the moveout of the events (Yang and Sava, 2010), and we can perform transformations from the extended to the angle domain (Sava and Fomel, 2003, 2006; Sava and Vlad, 2011). The extended images provide a measurement of the similarity between the source and receiver wavefields along space and time, so we can exploit these images to analyze the RTM backscattered events.

In the presence of sharp boundaries, we can also construct four partial extended images:

\[ R(x, \lambda, \tau) = R^{tt}(x, \lambda, \tau) + R^{rr}(x, \lambda, \tau) \]
\[ + R^{rt}(x, \lambda, \tau) + R^{tr}(x, \lambda, \tau). \]  

(6)

By analyzing the individual contributions to the image and extended image, we can better understand the behavior of the backscattered events. This analysis is similar to the one of Fei et al. (2010) and Liu et al. (2011) whose objective is to filter out the non-geological portions of the image. Here, we approach the problem in a broader sense by attempting to understand the physical meaning of the backscattered energy and its uses for velocity analysis.

In order to gain an understanding of the RTM backscattered events, we use a simple model with two-layers and a strong velocity contrast. Figure 2(b) shows the image obtained with the conventional imaging condition for one shot at \( x = 5 \text{km} \). This image has strong backscattered energy (indicated with letter “a”) above the reflector located at \( z = 1.5 \text{km} \). Figures 1(a), 1(b) and 2(a) show a time-lag gather, a space-lag gather and a common image (CIP) point gather. The backscattered energy \( R^{tr}(x, \lambda, \tau) + R^{rt}(x, \lambda, \tau) \) (denoted with letter “a”) maps toward zero lag for time-lag and space-lag gathers, as shown in Figures 1(a) and 1(b), respectively. This mapping to zero lag means that the reflected wavefields map in perfect synchronization with the transmitted wavefields; therefore they cannot be dissociated in the imaging condition. This synchronization is achieved only because we use the correct velocity model to obtain these images. We can also identify the cross-correlation between the reflected wavefields (denoted by letter “b”) in Figures 1(a) and 1(b) for time-lag and space-lag gathers, respectively. In the time-lag gathers, this event has an opposite slope compared with the one of the primaries, as if we changed the cross-correlation order in equation 5.

The backscattered events (identified with letter “a”) also appear in CIP gathers, as shown in Figure 2(a). Both backscattered contributions map to \( \tau > 0 \) in the \( \tau - \lambda_z \) plane. However, they map as two different events, whereas for time-lag and space-lag gathers both map to zero lag. In the \( R^{tt}(x, \lambda, \tau) \) image, they map to \( \tau > 0 \) and \( \lambda_z > 0 \), and in the \( R^{tr}(x, \lambda, \tau) \) image they map to \( \tau > 0 \) and \( \lambda_z < 0 \).

**Sensitivity to Velocity Errors**

In this section, we test the dependency of the backscattered energy on velocity errors using extended images.

---

Figure 2: Synthetic example: (a) common image point at \((x, z) = (5.0, 1.5)\text{km}\) and (b) migrated image of one shot.
REFERENCES


Foss, S., B. Ursin, and M. V. de Hoop, 2005, Depth-consistent reflection tomography using PP and PS seismic data: Geophysics, 70, U51–U65.


Thomsen, L., 1999, Converted-wave reflection seismology over inhomogeneous, anisotropic media: Geophysics, 64, 678–690.


Understanding the reverse time migration backscattering: noise or signal?
Esteban D´ıaz and Paul Sava, Center for Wave Phenomena, Colorado School of Mines

SUMMARY
Reverse time migration (RTM) backscattered events are produced by the cross-correlation between waves reflected from sharp interfaces (e.g. the top of salt bodies). Commonly, these events are seen as a drawback for the RTM method because they obstruct the image of the geologic structure. Many strategies have been developed to filter out the artifacts from the conventional image. However, these events contain information that can be used to analyze kinematic synchronization between source and receiver wavefields reconstructed in the subsurface. Numeric and theoretical analysis indicate the sensitivity of the backscattered energy to velocity accuracy: an accurate velocity model maximizes the backscattered artifacts. The analysis of RTM extended images can be used as a quality control tool and as input to velocity analysis designed to constrain salt models and sediment velocity.

INTRODUCTION
Reverse time migration (RTM) is not a new imaging technique (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983). However, it was not until the late 1990s that computational advances allowed the geophysical community to use this technology for exploratory 3D surveys. In complex geological settings, RTM produces better images than other methods.

A striking characteristic of RTM is the presence of low wavenumber events, uncorrelated with the geology, in models with sharp interfaces (e.g. salt intrusions). The seismic industry has dedicated effort and time developing algorithms and strategies to filter out the backscattered energy from the image. We can classify the filtering approaches in two general families: pre-imaging condition and post-imaging condition.

The pre-imaging condition family modifies the wavefields (either by modeling or by wavefield decomposition) such that the backscattered events do not form during the imaging process. One strategy is wavefield decomposition (Liu et al., 2011; Fei et al., 2010). In this method only wavefields propagating in opposite directions are cross-correlated. The modeling approach modifies the wavefields by not allowing waves to reflect during propagation. One way to achieve this is by introducing an absorbing boundary condition in the wave equation at the top of salt (Fletcher et al., 2005), or by impedance matching (Baysal et al., 1984).

In the post-imaging family, the artifacts are attenuated by filtering, which is a considerably cheaper process because operates in the image space and not on the much larger and more complex wavefields. A common strategy is to apply a Laplacian filter to the image (Youn and Zhou, 2001). Another option is a signal/noise separation by least-squares filtering (Guitton et al., 2007). One can also consider using extended images (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011) for filtering. This method (Kaelin and Carvajal, 2011) takes advantage of the slope difference between primary and backscattered events in the time-lag gathers.

In this paper, we analyze the information carried by the backscattered energy in the extended images. We show that the backscattered waves provide important information about the synchronicity between the reconstructed wavefields in the subsurface. The presence of backscattered energy in the image not only depends on the interpretation of the sharp interface but also on the velocity above it. We analyze the mapping patterns of the backscattered events in the extended images and conclude that backscattered energy is sensitive to the velocity model accuracy and therefore should be used as a source of information for migration velocity analysis (MVA). Counter to common practice, we assert that backscattering artifacts should be enhanced during RTM to constrain the velocity models, and they should only be removed in the last stage of imaging.

THEORY
The conventional imaging condition (CIC) (Claerbout, 1985) is a zero time-lag cross-correlation between the source wavefield and the receiver wavefields:

\[ R(x) = \sum_{s} \sum_{t} W_s(x, t) W_r(x, t). \]  

A wavefield extrapolated with RTM could show, depending on the complexity of the geology, waves traveling in both upward and downward directions, such as diving waves, head waves and backscattered waves. The correlation between forward and backscattered waves is particularly strong when sharp boundaries are present in the velocity model (e.g. for salt bodies).

Figure 1: Synthetic example: (a) time-lag and (b) space-lag gathers at x=5km.
Understanding RTM backscattering

Figure 3: Sample gathers with -12% and +12% velocity error: (a)-(b) time-lag gathers, (c)-(d) space-lag gathers, (e)-(f) CIP gathers.

We analyze the behavior of backscattered events in the presence of velocity errors. We test the sensitivity of the backscattered events with the same synthetic data discussed previously. In this case, we construct the images with different models characterized by a constant error varying from -12% to +12% in layer 1. We keep the interface consistent with the velocity used for imaging, i.e. the interface producing backscattered energy is placed in the model according to the velocity in layer 1. Figures 3(a) and 3(b) show time-lag gathers for a -12% and +12% velocity error, respectively. One can see that the backscattered energy does not map to $\tau = 0$ because the wavefields are no longer synchronized. For negative errors the artifacts map to $\tau > 0$, whereas for positive errors they map to $\tau < 0$. Figures 3(c) and 3(d) show space-lag gathers for the same velocity errors. In this case we see that the backscattered energy maps symmetrically away with respect to $\lambda_x = 0$. Figures 3(e) and 3(f) show CIP gathers for the same velocity errors. One can see that the velocity errors split the backscattered energy in the $\lambda_z - \tau$ and $\lambda_x$ planes; some of the energy goes through zero space-lag, while other part of the energy does not.

We can use the kinematic information from backscattered events contained in the extended images to design objective functions (OF) that exploit the presence of backscattered events. To isolate the backscattered events, we use wavefield decomposition to obtain the individual contributions shown in equation 6. We use the images $R^r(x, \lambda, \tau)$ and $R^i(x, \lambda, \tau)$ which contain backscattered energy. Minimizing such OF, e.g., by wavefield tomography, optimizes the sharp interface positioning (e.g. the top of salt) and the sediment velocity above it. A straightforward approach based on differential semblance optimization (Shen et al., 2003) can be adapted to use the backscattered energy seen away from zero lags by defining objective functions for time-lag gathers

$$J_r = \frac{1}{2} \| P(\tau) [R^r(z, \tau) + R^r(z, \tau)] \|^2, \quad (7)$$

and for space-lag gathers,

$$J_{\lambda_x} = \frac{1}{2} \| P(\lambda_x) [R^r(z, \lambda_x) + R^r(z, \lambda)] \|^2. \quad (8)$$

The functions $P(\tau) = |\tau|$ and $P(\lambda_x) = |\lambda_x|$ penalize the backscattered energy away from zero lags, thus defining the residual that we need to minimize through inversion. For common image points we can use

$$J_c = \frac{1}{2} \| P(\lambda, \tau) [R^r(\lambda, \tau) + R^r(\lambda, \tau)] \|^2. \quad (9)$$

The penalty functions are designed to measure the deviation error between actual and ideally focused extended images. For CIGs we have a definite criterion: we know that the backscattered energy has to map to zero lag. However, for CIPs the penalty operator is more complex. We use a correct CIP as reference for constructing the penalty function $P(\lambda, \tau)$.

For common image points we can use

$$J_c = \frac{1}{2} \| P(\lambda, \tau) [R^r(\lambda, \tau) + R^r(\lambda, \tau)] \|^2. \quad (9)$$

The objective functions for our synthetic example are shown in Figure 4 for time-lag CIGs, space-lag CIGs, and common image point gathers, respectively. In all three cases the OF minimizes at the correct model.

Figure 4: Normalized objective functions $J_r$ (red), $J_{\lambda_x}$ (blue) and $J_c$ (green).

EXAMPLE

In this section, we illustrate the backscattered events visible on extended images constructed with the Sigsbee 2A model (Paffenholz et al., 2002). We modify the
Understanding RTM backscattering

![Figure 5](image1.png)

Figure 5: Sigsbee analysis: time-shift CIG (a) and space-lag CIG (b) at \( x = 19.1 \text{km} \), and a common image point gather at \((x,z) = (19.1,3.45)\text{km}\) (c).

Figure 6: Sigsbee RTM image (d). The vertical line and blue dot show the CIG and CIP locations, respectively.

model by salt flooding (extending the salt to the bottom of the model) to avoid backscattering from the base of salt; therefore we focus on the reflections from the top of salt only. For this example we fix the receiver array on the surface, and we use 100 shots evenly distributed on the surface to build the image. For the migration model, we use the stratigraphic velocity which has sharp interfaces in the sediment section, in addition to the interface corresponding with the top of salt. Figure 6 shows the conventional image for our modified Sigsbee model; note the strong backscattered energy above the salt.

Figure 5(a) shows a time-lag gather calculated at \( x = 19 \text{km} \). We can see that the gather is very complex, but we can easily identify the backscattered energy (indicated with letter “a”) in Figure 5(a). In this case, the backscattered energy maps directly to \( \tau = 0 \) because we use the correct velocity model. We can also identify the events corresponding to the cross-correlation between reflected waves from the source and receiver side \( R^{rr}(z,\tau) \), (indicated with letter “b”). The \( R^{rr}(z,\tau) \) events have positive slope (given by the sediment velocity at the interface) and are visible for \( \tau > 0 \). We can also observe an abrupt change in the slope of the primary reflection corresponding to the sediment-salt interfaces at the top of salt, (indicated with letter “c”).

Figure 5(b) shows a space-lag common image gather extracted at the same location. The backscattered energy maps toward \( \lambda_x = 0 \) (indicated with letter “a”). We see again the \( R^{rr}(z,\lambda_x) \) case (indicated with letter “b”) where the energy is mapped away from zero lag. Even though we are using the correct model, we still see energy away from \( \lambda_x = 0 \). This indicates that additional processing is needed before we can use space-lag gathers for model update with wave equation tomography.

Figure 5(c) shows a common image point extracted at the top of the salt interface at \((x,z) = (19.05,3.4)\text{km}\). Despite the complexity of this image, we can still identify similar patterns to the ones seen in the synthetic example, shown in Figure 2(a). The backscattered events are mapped to \( \tau > 0 \) in the \( \tau - \lambda_z \) plane (indicated with letter “a”). In this plane, we can separate the individual contributions from \( R^{rt}(x,\lambda_z,\tau) \) (which maps to \( \lambda_z < 0 \) and \( \tau > 0 \)) and \( R^{tr}(x,\lambda_z,\tau) \) (which maps to \( \lambda_z > 0 \) and \( \tau > 0 \)) because they are imaged into two different events. In the common image gathers discussed before, we cannot differentiate the individual contributions because both events map to zero lag. The reflection maps as a point to zero lag in the \( \tau - \lambda_x \) plane (indicated with letter “c”).

CONCLUSIONS

RTM backscattered events map to zero lag in the extended images when the velocity is correct. This means that the reflected receiver wavefield travels in perfect synchronization with the source wavefield and vice versa. We demonstrate that the RTM backscattered energy is sensitive to kinematics errors in the velocity model. The backscattered energy should be maximized in the image in order to ensure an optimum velocity model. The analysis of backscattered energy on extended images provides a definite criterion to update velocity models with salt interfaces (i.e in the Gulf of Mexico). Further tests are needed to develop automated velocity model update based on the maximization of the backscattered energy.

ACKNOWLEDGMENTS

This work was supported by the sponsors of the Center for Wave Phenomena at Colorado School of Mines. The reproducible numeric examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org.
REFERENCES


Migration velocity analysis for nonlinear reverse-time migration  

Clement Fleury*, Center for Wave Phenomena, Colorado School of Mines

SUMMARY
Nonlinear reverse-time migration is a modified reverse-time migration that accounts for the nonlinear relation between seismic data and model in order to image multiply scattered waves including multiples. The illumination of multiply scattered waves yields a representation of the Earth’s subsurface that is more sensitive to model parameters, which allows for advanced seismic interpretation. The gain in sensitivity that multiply scattered waves provide also applies to velocity model building. I present a strategy for using the illumination and sensitivity of internal multiples for migration velocity analysis and illustrate this method with numerical examples.

INTRODUCTION
Reverse-time migration (RTM) is a two-way time-domain finite-frequency technique that accurately handles the propagation of complex scattered waves and produces a band-limited representation of the subsurface structure that is conventionally assumed to be linear in the contrasts of the model parameters (e.g., Farmer et al., 2006). Because of the underlying linear single-scattering assumption, most RTM implementations do not optimally use the illumination and sensitivity provided by multiply scattered waves (e.g., Fleury and Vasconcelos, 2012). Migrating multiply scattered waves requires preserving the nonlinear relation between the image and the contrasts of the model parameters. To utilize the energy, illumination, and sensitivity contained in multiply scattered data, Fleury and Snieder (2012) propose a nonlinear reverse-time migration (NLRTM) algorithm. NLRTM modifies the extrapolation of source and receiver wavefields to more accurately reconstruct multiply scattered waves and extends the concept of the imaging condition in order to map in the subsurface structurally coherent seismic events which correspond to the interaction of both singly and multiply scattered waves.

The NLRTM strategy reveals potential for migration velocity analysis (MVA). MVA traditionally applies to images of primary reflections only. Most MVA algorithms estimate and correct the focus of primary reflection events in the image domain by using the semblance principle and move-out corrections for image gathers (e.g., Symes, 2008b). The use of multiply scattered seismic waves, which include multiples, improves illumination, sensitivity, and redundancy in the MVA procedure. To take advantages of these properties, attempts have been made to apply MVA to surface-related multiples (e.g., Nasyrov et al., 2009). Surface-related multiples lead to more sensitive MVA methods and increase the effective surface illumination but do not solve the fundamental problem of limited illumination under complex subsurface structures, such as sub-salt or sub-basalt structures. MVA for internal multiples contributes to the solution of this problem. The same semblance principle that applies to primary events is applicable to multiple events when forming additional image gathers for internal multiples. NLRTM provides sub-images of the interactions of primaries/multiples and multiples/multiples from which one can build such gathers. Further information is extractable from these NLRTM extended sub-images. Despite the fact that each NLRTM sub-image uses different subsurface illumination (from primaries or multiples), these sub-images are still representations of the same subsurface. This consistency between NLRTM sub-images is usable for MVA when applying the semblance principle across sub-images.

In this paper, I review the principles behind NLRTM and develop an MVA strategy for applying the semblance principle to NLRTM. The results of synthetic examples prove the method to be effective for interpretation and velocity model building.

NONLINEAR REVERSE-TIME MIGRATION
A scattering-based approach for seismic imaging yields the description of the Earth’s model \( \mathbf{m} \) in terms of a migration model \( \mathbf{m}_0 \) (which generates the reference wavefield \( w_0 \) and data \( d_0 \)) and a contrast model \( \Delta \mathbf{m} \) (which generates the scattered wavefield \( w_s \) and data \( d_s \)). NLRTM yields a set \( i \) of sub-images that are nonlinear representations of model \( \Delta \mathbf{m} \) (Fleury and Snieder, 2012):

\[
i = (i_0^{(d,a)}, i_0^{(a,d)}, i_s^{(d,a)}, i_s^{(a,d)}),
\]

where the notation refers to wavefield components and their illumination (details follow). The nonlinearity of the NLRTM sub-images results from preserving the nonlinear relation between wavefield \( w_s \) (equal to data \( d_s \) at receivers) and the perturbation \( \mathbf{P} \) (a function of models \( \mathbf{m}_0 \) and \( \Delta \mathbf{m} \)) of the reference propagation operator \( \mathbf{H}_0 \) (e.g., Weglein et al., 2003). Assuming that the estimates of the source signature \( s_{est} \) and the perturbation operator \( \mathbf{P}_{est} \) are known, NLRTM preserves this nonlinear relation in the extrapolation of source- and receiver-side reference and scattered wavefields (Fleury and Snieder, 2012):

\[
\mathbf{H}_0 w_{0,sou}=s_{est}
\]

(2)

\[
(\mathbf{H}_0 - \mathbf{P}_{est}) w_{s,sou} = \mathbf{P}_{est} w_{0,sou}
\]

(3)

\[
\mathbf{H}_0 w_{0,rec} = d_0
\]

(4)

\[
(\mathbf{H}_0 - \mathbf{P}_{est}) w_{s,rec} = d_s + \mathbf{P}_{est} w_{0,rec}
\]

(5)

where symbol \( ^\dagger \) denotes the adjoint operation. The introduction of operator \( \mathbf{P}_{est} \) in the extrapolation accounts for scattering contrast model \( \Delta \mathbf{m}_{est} \) in reconstructing scattered wavefields. There are two main approaches for retrieving model \( \Delta \mathbf{m}_{est} \): an interpretation-based approach using velocity model building tools and an automatic algorithm-based approach using least-squares migration (Nemeth et al., 1999), optimal image scaling (Symes, 2008a), or gradient-based full-waveform inversion updates (e.g., Pratt et al., 1998). The set \( \mathbf{w} \) of wavefields,

\[
\mathbf{w} = (w_{0,sou}, w_{s,sou}, w_{0,rec}, w_{s,rec})
\]

(6)

decomposes into

\[
\mathbf{w} = \mathbf{w}^{(a)} + \mathbf{w}^{(d)}
\]

(7)
Illumination and sensitivity of internal multiples for migration velocity analysis

Figure 1: Sigsbee model for NLRTM. The red and purples lines indicate the constant-offset source/receiver geometry. The green box defines the poorly illuminated region of interest.

where wavefield $w^{(u/d)}$ is up-going (or down-going) and results from the application of the decomposition operator $D^{(u/d)}$:

$$D^{(u/d)}w = w^{(u/d)}.$$ (8)

In the extended image space (e.g., Sava and Vasconcelos, 2011), the NLRTM imaging condition yields the sub-images of set $i$:

$$i_0^{(u/d)} = \sum_{\text{source}} T_{\lambda, \tau} w^{(d/0, \text{sou})} \ast T_{\lambda, \tau} w^{(u)}$$ (9)

$$i^{(u/d)} = \sum_{\text{source}} T_{\lambda, \tau} w^{(d/0, \text{sou})} \ast T_{\lambda, \tau} w^{(u)}$$ (10)

$$\hat{i}^{(u/d)} = \sum_{\text{source}} T_{\lambda, \tau} w^{(d/0, \hat{\text{sou}})} \ast T_{\lambda, \tau} \hat{w}_{\text{rec}}^{(d)}$$ (11)

$$\hat{i}^{(u/d)} = \sum_{\text{source}} T_{\lambda, \tau} w^{(d/\hat{\text{sou}})} \ast T_{\lambda, \tau} \hat{w}_{\text{rec}}^{(d)}$$ (12)

where symbol $\ast$ denotes zero-time crosscorrelation and operator $T(\lambda, \tau)$ shifts wavefields by space-lag $\lambda$ and time-lag $\tau$:

$$T_{\lambda, \tau} w(x, t) = w(x+\lambda, t+\tau).$$ (13)

The wavefield decomposition in the imaging condition isolates back-scattered from forward-scattered energies to reveal sub-surface scattering contrasts. Images $i_0^{(u/d)}$ and $i^{(u/d)}$ map the interaction of scattered wavefields $w^{(d/0, \text{sou})}$ and $w^{(u)}$ and taken into account higher orders of scattering, which offers the possibility for correctly imaging internal multiples and for subsequently using these multiples for MVA.

**Figure 2:** Sigsbee NLRTM sub-images in the region of interest.

**MIGRATION VELOCITY ANALYSIS FOR NLRTM**

The semblance principle applies to each element of the set $i$ of NLRTM extended sub-images. Following the method proposed by Yang and Sava (2012), the defocused energies observed in the NLRTM extended sub-images are penalized by annihilator $A$ and combined into the objective function $J$,

$$J = \sum_{\text{sub-image } j} \frac{1}{2} \left| \left| A_i^{(u/d)} \right| \right|^2,$$ (14)

which an optimal migration model $m_{\text{opt}}$ minimizes. I solve this optimization problem using the adjoint-state method (e.g., Plessix, 2006). The set $w^{(u/d)}$ of adjoint wavefields,

$$w^{(u/d)} = (w^{(u/d)}_{0, \text{sou}}, w^{(u/d)}_{0, \text{rec}}),$$ (15)

extrapolates the residual defocused energies mapped into the set $g^{(u/d)}$ of NLRTM adjoint sources defined as

$$g^{(u/d)}_{0, \text{sou}} = (g^{(u/d)}_{0, \text{sou}}, g^{(u/d)}_{0, \text{rec}}),$$ (16)

$$g^{(u/d)}_{0, \text{rec}} = D^{(u/d)} \int T_{\lambda, \tau} \left( A_{\text{rec}}^{\lambda} \right) \ast T_{\lambda, \tau} w^{(d/n)} \, d\lambda d\tau,$$ (17)

$$g^{(u/d)}_{0, \text{sou}} = D^{(u/d)} \int T_{\lambda, \tau} \left( A_{\text{sou}}^{\lambda} \right) \ast T_{\lambda, \tau} w^{(d/n)} \, d\lambda d\tau,$$ (18)

where

$$g^{(u/d)}_{0, \text{rec}} = 0$$ (19)

$$g^{(u/d)}_{0, \text{sou}} = D^{(u/d)} \int T_{\lambda, \tau} \left( T_{\lambda, \tau} \ast \hat{A}_{\text{sou}}^{(d/u, u/d)} \right) \hat{A}_{\text{sou}}^{(d/u, u/d)} \, d\lambda d\tau.$$ (20)
Illumination and sensitivity of internal multiples for migration velocity analysis

Figure 3: Point-scatterer models: (a) and (b) true model, (c) and (d) starting model, and (e) and (f) contrast model estimate. The red dots and green line define the fix-spread source/receiver geometry. The pink dot indicates the scatterer.

The set of equations that extrapolate the adjoint wavefields is

\[ \mathbf{H}_0 w_{j, \text{sou}}^{(u/d)} = s_{j, \text{sou}}^{(u/d)} + \mathbf{P}^{\text{est}} w_{j, \text{rec}}^{(u/d)} \]  
\[ \mathbf{H}_0 w_{j, \text{rec}}^{(u/d)} = s_{j, \text{rec}}^{(u/d)} + \mathbf{P}^{\text{est}} w_{j, \text{rec}}^{(u/d)} \]  
\[ \mathbf{H}_0 - \mathbf{P}^{\text{est}} \]  (21)  
\[ \mathbf{H}_0 - \mathbf{P}^{\text{est}} \]  (22)  
\[ \mathbf{H}_0 - \mathbf{P}^{\text{est}} \]  (23)

The parametrization of the model space in terms of squared slowness yields the gradient \( \nabla \mathbf{m} J \) of functional \( J \):

\[ \nabla \mathbf{m} J = \sum_{\text{wavefield}} \frac{\partial^2}{\partial t^2} w_j^{(u/d)} \ast w_j^{(u/d)} \]  (24)

With this gradient, a Newton-based method allows one to iterate for the solution model \( \mathbf{m}_{\text{opt}} \) that best focuses the energies of the NLRTM sub-images.

The semblance principle also applies across the elements of the set of NLRTM sub-images. Because of the uniqueness of true model \( \mathbf{m} \), all elements \( i_j \) of \( \mathbf{i} \) must show consistent positioning of the subsurface structures of the model, despite their intrinsic differences in illumination. For MVA purposes, this consistency between NLRTM sub-images is assessable using local-window crosscorrelation function \( C \):

\[ C(i_j, i_k) = \int_{W(x)} i_j(y - \eta) i_k(y + \eta) \, dy \]  (25)

with the local-window \( W \) and space-lag \( \eta \). The annihilator \( \mathbf{A} \) penalizes the poor correlations between the sub-images. This

new MVA criterion complements the previous one and leads to the minimization of the modified objective function \( J' \):

\[ J' = \sum_{\text{sub-image}} \frac{1}{2} || A_i ||^2 + \sum_{\text{sub-images} \, j \neq i} \frac{1}{2} || \mathbf{A} C(i_j, i_k) ||^2 \]  (26)

EXEMPLARY

Multiply scattered waves, including internal multiples, are waves that have been reflected, diffracted, and more generally scattered more than once. These waves provide energy, illumination, and sensitivity that improve both imaging and MVA when their nonlinear relation to the model is accounted for. NLRTM applies to target-oriented sub-salt imaging. Sub-salt imaging is challenging because of the structural complexity of salt bodies and the lack of sub-salt illumination (e.g., Leveille et al., 2011). For the Sigsbee 2A model in Figure 1 (Paffenholz et al., 2002), NLRTM provides the image set \( \mathbf{i} \) (Figure 2) for refined interpretation in the poorly illuminated sub-salt area. The perturbation \( \mathbf{P}_{\text{est}} \) is obtained from scaling a conventional RTM image of the same subsurface to create model \( \mathbf{m}_{\text{est}} \) (Figure 1b). The NLRTM extrapolation reconstructs internal multiples that illuminate the sub-salt structures and map nonlinearly scattered energy in the NLRTM sub-images. This new and redundant structural information facilitates detailed seismic interpretation of the subsalt target. In Figure 2, the coherent structures are consistent across sub-images. This consistency is evidence that migration velocity model \( \mathbf{m}_0 \) (Figure 1b) is accurate and allows for more confident interpretation. The sub-images of internal multiples also more clearly reveal some structural features. For example, the sediment layering across the two faults (shown by the green solid lines) is more
Illumination and sensitivity of internal multiples for migration velocity analysis

Figure 5: Extended NLRTM sub-images of the point scatterer.

visible in sub-image $i_{0}^{(d,u)}$ (Figure 2d). This helps to resolve ambiguities in poorly illuminated sub-salt areas.

NLRTM also applies to MVA with internal multiples. In the point-scatterer example (Figure 3), multiples from the reflector interface and primaries illuminate the point scatterer and yield the NLRTM sub-images in Figure 4. The example implements the first MVA criterion (optimization of objective $J$). The starting model (Figure 3c) contains a low-slowness anomaly. The interpretation of a RTM image allows one to pick the density reflector to create the contrast model (Figure 3f). In Figure 4, sub-image $i_{0}^{(d,u)}$ seems focused while sub-image $i_{s}^{(u,d)}$ shows defocused energy at the scatterer location. After automatically picking locations for common image point gathers (Cullison and Sava, 2011), NLRTM extended sub-images estimate the energy focus in these sub-images. At the scatterer location, Figure 5 shows energy focus in sub-image $i_{0}^{(d,u)}$ and energy defocus in sub-image $i_{s}^{(u,d)}$ which the annihilator in Figure 6 penalizes. The image of the scatterer, illuminated from below by internal multiples, is sensitive to the slowness error in the starting model. The gradient $\nabla m J$ at the first iteration (Figure 7) supports this observation. Sub-image $i_{0}^{(d,u)}$ contributes to the slowness gradient (Figure 7a) because the primary reflection from the density contrast is defocused in this extended sub-image. In addition, sub-image $i_{s}^{(u,d)}$ contributes to the slowness gradient (Figure 7b) because of the defocused energy of the multiple reflection at the scatterer. The total gradient (Figure 7c) sums the two contributions and has stronger amplitude at the anomaly location. The sensitivity of internal multiples contributes to the MVA update and better constrains the spatial location of the high-slowness anomaly.

CONCLUSIONS

The NLRTM strategy uses the energy, illumination, and sensitivity of internal multiples for seismic imaging and MVA. The set of NLRTM sub-images is a new tool for seismic interpretation, as illustrated by the target-oriented subsalt imaging example. NLRTM sub-images emphasize the difference in illumination of primaries and multiples which results in additional information for the description of geological subsurface features. The consistency between this redundant information leads to more confident interpretation. This same consistency criterion reveals the gain in sensitivity to model parameters coming from the use of internal multiples. This extra-sensitivity benefits the development of MVA methods. As shown in the synthetic example, the illumination of internal multiples in NLRTM provides new constraints for velocity model building. These constraints lead to localized and sensitive updates of the velocity errors and potentially reduce the number of iterations needed for the convergence of gradient-based MVA methods.

ACKNOWLEDGMENTS

This work was supported by the Consortium Project on Seismic Methods for Complex Structures at the Center for Wave Phenomena.
Illumination and sensitivity of internal multiples for migration velocity analysis

REFERENCES


Symes, W., 2008a, Approximate linearized inversion by optimal scaling of prestack depth migration: Geophysics, 73, R23–R35.


Bi-objective optimization for the inversion of seismic reflection data: Combined FWI and MVA
Clement Fleury* and Francesco Perrone, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

The inversion of seismic reflection data is challenging due to the oscillatory nature of the seismic data, the nonlinear relation between the data and the Earth’s model parameters, and the relative sensitivity of the different subsets of the data to the model parameters. Full-waveform inversion solves this inverse problem in the data space but is known to have shortcomings, especially because of the multi-modality and the numerous locally optimal solutions that its objective function allows. Migration velocity analysis indirectly relates to the inversion of seismic reflection data. Used in velocity model building, the method yields a macro model of the Earth’s subsurface but constrains this model only to be kinematically correct. Migration velocity analysis has also its share of well-known pitfalls. Interestingly, the two methods have complementary characteristics which potentially mitigate the drawbacks of each technique. To take advantage of the capability of full-waveform inversion and migration velocity analysis, a bi-objective optimization strategy combines the two methods to solve the inversion of seismic reflection data. A numerical example illustrates the benefits of this new strategy.

INTRODUCTION

Seismic imaging aims for a representation of the contrasts in the Earth’s physical properties and for an estimate of a kinematically correct model for wave propagation in the subsurface. One general assumption is that one can distinguish two distinct scales in the velocity model that describes the Earth’s subsurface (Jannane et al., 1989; Bleistein et al., 2001). First, a smooth long-wavelength component describes the kinematic character of the waves, and second, a sharp short-wavelength component generates the reflections we record at the surface during the seismic experiment.

Sets of complementary techniques have been developed to address the estimation of the smooth and sharp components of the velocity model. Migration velocity analysis (MVA) (Fowler, 1985; Faye and Jeannot, 1986; Al-Yahya, 1989; Chavent and Jacewitz, 1995; Biondi and Sava, 1999; Sava et al., 2005; Albertin et al., 2006) reconstructs the long-wavelength component of the velocity model by application of the semblance principle, i.e., the assumption that the images of reflectors in the subsurface must be invariant with respect to seismic experiments (Al-Yahya, 1989). Seismic migrated images supply structural information (the short-wavelength component of the velocity model) by mapping in the subsurface the contrasts in physical properties that generated the echoes recorded at the surface. Full-waveform inversion (FWI) (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004) is an adjoint-state technique that reconstructs the sharp transitions of the model parameters. This technique involves matching the observed data with synthetic wavefields directly modeled in a trial model. By matching the kinematics and dynamics of the wavefields, FWI obtains high-resolution images of the changes in subsurface model parameters. FWI and MVA have very different behavior and degrees of robustness. Since FWI tries to exactly match the recorded data (i.e., bandlimited oscillatory signals), this method is very sensitive to the initial model. Unless the initial model accurately describes the kinematic of the waves, FWI may easily converge to spurious solutions (local minima of its objective function). MVA is essentially an estimator of the kinematic of wave propagation and is characterized by well-behaved, smooth, and mono-modal objective functions (e.g., Santosa and Symes, 1986; Symes and Carazzone, 1991; Stolk and Symes, 2004). Since MVA relies on the single-scattering assumption (i.e., the first-order Born approximation) underlying most migration schemes, this method is unable to handle nonlinear wave phenomena (such as multiple reflections), and these nonlinear events must be removed from the data before starting the estimation procedure. Because MVA focuses on kinematic features of wavefields and images, this technique also intrinsically limits the model resolution one can achieve.

Standard practice considers MVA and FWI as two distinct steps because the two methods estimate different scale components of the model: first, one estimates a kinematically reliable model through MVA; second, using the MVA solution, one applies FWI to obtain a high-resolution image of the contrasts in model parameters that generated the reflected data. Here, we show that it is indeed possible to integrate the two steps of MVA and FWI into a single bi-objective optimization problem. Our goal is two-pronged: to place a kinematic constraint on FWI to ensure the FWI gradient focuses seismic energy, as well as a data constraint on MVA to match the data which are the only seismic experimental measurements. The images used by the MVA part of our bi-objective optimization strategy are in fact the gradients of the FWI cost function (also known as sensitivity kernels for FWI). The advantages of combining the two methodologies are multifold: MVA constrains the smooth part of the model and makes the FWI contribution less sensitive to cycle skipping and local minima; FWI introduces sharp variations in the optimal solution model that are useful to further constrain the shape of smooth anomalies and to reduce the sensitivity of MVA to the nonlinearity of the data with respect to model parameters (e.g., multiply scattered waves).

THEORY

The physical model $m$ describes the Earth’s medium in terms of squared slowness. The inversion of the reflection seismic data $d$, acquired in a series of shot-profile experiments (sources $x_S$ and receivers $x_R$), is solved in the extended model space $m_{ext}$ (Symes, 2008): 

$$m_{ext} = T_{2\hbar \to 2\hbar} m.$$  

(1)
Bi-objective optimization for the inversion of seismic reflection data

where operator $\mathbf{T}(\mathbf{A}, \tau)$ shifts a given space- and/or time-dependent variable $\mathbf{u}$ by space-lag $\mathbf{A}$ and time-lag $\tau$:

$$\mathbf{T}(\mathbf{A}, \tau)_\mathbf{u}(x, t) = \mathbf{u}(x + \mathbf{A}, t + \tau).$$  

(2)

We consider the bi-objective vector $\mathbf{J} = (J_{\text{FWI}}, J_{\text{MVA}})$ that collects the FWI and MVA cost functions and use the weighting method to scalarize the optimization problem by defining the cost function $J$ (e.g., Marler and Arora, 2004):

$$J = (1 - \alpha) w_{\text{FWI}} J_{\text{FWI}} + \alpha w_{\text{MVA}} J_{\text{MVA}},$$  

(3)

where $w_{\text{FWI}}$ and $w_{\text{MVA}}$ are normalization factors, and the scalar $\alpha$ controls the weight given to MVA with respect to FWI. To solve for an optimal model $\mathbf{m}_{\text{opt}}$, a gradient-based method minimizes cost function $J$ using the adjoint-state method to compute the gradient $\nabla_{\mathbf{m}_{\text{opt}}} J$ at each iteration (e.g., Plessix, 2004):

$$\nabla_{\mathbf{m}_{\text{opt}}} J = (1 - \alpha) w_{\text{FWI}} \nabla_{\mathbf{m}_{\text{opt}}} J_{\text{FWI}} + \alpha w_{\text{MVA}} \nabla_{\mathbf{m}_{\text{opt}}} J_{\text{MVA}}.$$  

(4)

The bi-objective optimization strategy integrates the two complementary approaches of FWI and MVA. First, FWI matches the data $\mathbf{d}$ with the synthetic state wavefields $\mathbf{w}$ modeled with the trial model $\mathbf{m}$ (e.g., Tarantola, 1984) so that the misfit between the two signals is measured by cost function $J_{\text{FWI}}$:

$$J_{\text{FWI}} = \frac{1}{2} \sum_{x_i} ||\mathbf{d} - \mathbf{\delta}_{x_i} \mathbf{w}||^2,$$  

(5)

where operator $\mathbf{\delta}_{x_i}$ restricts wavefield $\mathbf{w}$ to the spatial locations of the receivers $x_i$. Given a source estimate $\mathbf{s}$, the state wavefield $\mathbf{w}$ and corresponding adjoint wavefield $\mathbf{w}^\dagger$ propagate according to

$$\begin{cases} \mathbf{H}(\mathbf{m}) \cdot \mathbf{w} = \mathbf{s} \\ \mathbf{H}^\dagger(\mathbf{m}) \cdot \mathbf{w}^\dagger = -(\mathbf{d} - \mathbf{\delta}_{x_i} \mathbf{w}), \end{cases}$$  

(6)

where $\mathbf{H}$ is the wave operator and symbol $\dagger$ denotes the adjoint operation associated with FWI. The gradient of functional $J_{\text{FWI}}$ in the extended model space is

$$\nabla_{\mathbf{m}_{\text{opt}}} J_{\text{FWI}} = \sum_{x_i} \mathbf{T}_{-\mathbf{A}_{x_i}} \partial^2 \nabla \mathbf{s} \mathbf{w}^\dagger.$$  

(7)

where symbol $\star$ denotes zero-time crosscorrelation.

Second, MVA measures the semblance between seismic events in image gathers, penalizes unfocused events, and minimizes the residual energy of these events via an optimization scheme (e.g., Symes, 2008; Yilmaz, 2001). In our formulation, MVA applies directly to the extended FWI gradient, instead of image gathers, by extension of the semblance principle to sensitivity kernels (e.g., the FWI sensitivity kernel for squared slowness in this paper). We adapt the method proposed by Yang and Sava (2012) to our bi-objective optimization framework. The annihilator $\mathbf{A}$ enhances the defocused energy observed in gradient $\nabla_{\mathbf{m}_{\text{opt}}} J_{\text{MVA}}$, which is then measured by cost function $J_{\text{MVA}}$:

$$J_{\text{MVA}} = \frac{1}{2} ||\mathbf{A} \nabla_{\mathbf{m}_{\text{opt}}} J_{\text{FWI}}||^2.$$  

(8)

Considering the MVA state vector $(\mathbf{w}, \mathbf{w}^\dagger)$, the corresponding adjoint vector $(\mathbf{w}^\ddagger, \mathbf{w}^{\ddagger\dagger})$ is a solution to the following system of equations:

$$\begin{cases} \mathbf{H}^\ddagger(\mathbf{m}) \cdot \mathbf{w}^\ddagger = \mathbf{g}^\ddagger \\ \mathbf{H}^{\ddagger\dagger}(\mathbf{m}) \cdot \mathbf{w}^{\ddagger\dagger} = \mathbf{g}^{\ddagger\dagger}, \end{cases}$$  

(9)

where symbol $\ddagger$ denotes the adjoint operation associated with
Bi-objective optimization for the inversion of seismic reflection data

![Graphical representation of bi-objective cost functions](image)

Figure 2: Bi-objective cost functions as functions of steepest-descent iterations ($\alpha = 0.7$): weighted contributions of the FWI and MVA components (blue and red curves, respectively) to the global cost function $J$ (magenta curve).

MVA, and the adjoint sources $g^\dagger$ and $g^\ddagger$ are equal to

$$
\begin{align*}
\{g^\dagger = & \int_{\lambda, \tau} T_{\lambda, \tau} \{A^\dagger A \nabla_{m,0} J_{\text{FWI}} \times T_{\lambda, \tau} \frac{\partial^2}{\partial t^2} w^\ddagger \} d\lambda d\tau \\
\{g^\ddagger = & \int_{\lambda, \tau} T_{\lambda, \tau} \{T_{\lambda, \tau} \frac{\partial^2}{\partial t^2} w^\ddagger \times A^\dagger A \nabla_{m,0} J_{\text{FWI}} \} d\lambda d\tau.
\end{align*}
$$

(10)

The gradient of objective $J_{\text{MVA}}$ in the extended model space is

$$
\nabla_{m,0} J_{\text{MVA}} = \sum_{x_i} T_{-\lambda, \tau} \frac{\partial^2}{\partial t^2} w \times T_{\lambda, \tau} w^\ddagger.
$$

(11)

The computation of the gradient of cost function $J$ requires the definition of normalization factors $w_{\text{FWI}}$ and $w_{\text{MVA}}$; here,

$$
w_{\text{FWI}} = \sum_{x_i} ||d||_2^2 \text{ and } w_{\text{MVA}} = \sum_{x_i} ||\nabla_{m,0} J_{\text{FWI}}||_2^2.
$$

(12)

where model $m_{\text{ext},0}$ is the starting model for the optimization. We refer to the review article by Marler and Arora (2004) for references to this normalization approach. The combined FWI and MVA gradient $\nabla_{m,0} J$ leads to an optimal solution $m_{\text{opt}}$ by means of a Newton-type algorithm that iteratively updates the current model in a gradient-related direction restricted to the physical space (i.e., for $\lambda = 0$ and $\tau = 0$). In the following, notation $V_m$ refers to the restriction of the extended gradient (indicated by $\nabla_{m,0}$) to the physical space.

EXAMPLES

The bi-objective optimization strategy applies to the acoustic reflection data modeled with the synthetic models in Figure 1 after plane-wave encoding (Whitmore, 1995; Zhang et al., 2005). Our implementation of the optimization procedure uses a finite-difference time-domain modeling engine. Because of the simplicity of the model, plane-wave migration allows us to completely image the data with fewer synthetic shot-gathers compared to conventional shot-record migration. True model $m_{\text{true}}$ (Figure 1a) contains a wide-spread low-slowness rectangular anomaly and a horizontal reflector that are absent from starting model $m_0$ (Figure 1b). After 53 iterations of a steepest-descent algorithm, optimal model $m_{\text{opt}}$ (Figure 1c) for the combined FWI and MVA optimization ($\alpha = 0.7$) resolves the rectangular anomaly and the reflector at their correct locations in the well-illuminated region of the model. The limited acquisition geometry and the poor structural model constraints (a single reflector below the rectangular anomaly) result in an ambiguity for the resolution of the slowness model below the rectangular anomaly which slightly distorts the shape of the deeper reflector on its sides. It is important to note that the amplitude of the rectangular slowness anomaly, in addition to its shape, is correctly estimated despite the lack of very low-frequency and diving waves in the data. This result presents a positive contrast to the results obtained when applying only MVA or FWI. For FWI ($\alpha = 0.0$), optimal model $m_{\text{opt}}$ (Figure 1d) resolves the contours of the rectangular anomaly and the reflector below, but FWI mispositions these interfaces with the exception of the top reflector of the rectangular anomaly. The slowness anomaly inside the rectangular shape is also not re-
Bi-objective optimization for the inversion of seismic reflection data

For the geometry of this inverse problem, the resolution power of MVA alone is too low to retrieve the true rectangular anomaly because of its limited size and its short-wavelength content. MVA, however, has the potential for constraining the inversion of the long-wavelength features of true model \( \mathbf{m}_{\text{true}} \). FWI, contrary to MVA, is able to constrain the short-wavelength component of optimal model \( \mathbf{m}_{\text{opt}} \). Unfortunately, FWI alone does not converge to the correct solution because the data is lacking information about the long wavelengths in model \( \mathbf{m}_{\text{true}} \). The bi-objective optimization strategy combines FWI and MVA to take advantage of the properties of each method. The parameter \( \alpha \) controls the weight given to MVA in the inversion procedure. For a high \( \alpha \), MVA is dominant and FWI acts as a regularizer by data constraint. For a low \( \alpha \), FWI is the main component of the optimization and the semblance principle of MVA becomes a constraint for the FWI gradient. Figure 2 shows the convergence of our steepest-descent algorithm for the combined FWI and MVA inversion with \( \alpha = 0.7 \), which is an intermediate value with higher weight given to MVA. The optimization problem is at first mostly driven by the MVA contribution. We adopt a multi-scale approach for FWI so that the frequency content in the data is at first band-limited to the low frequencies and slowly increases with the number of iterations. This strategy mitigates the sensitivity of FWI to cycle skipping. As a result, the FWI contribution becomes increasingly significant for later iterations when the long-wavelength features are already resolved and short-wavelength components are introduced in optimal model \( \mathbf{m}_{\text{opt}} \). Interestingly, the FWI and MVA cost functions, \( J_{\text{FWI}} \) and \( J_{\text{MVA}} \), simultaneously decrease during the joint-inversion process so that despite the different weight given to FWI and MVA, the two contributions are jointly optimized. The choice of parameter \( \alpha \) is key to regulate this convergence behavior of the combined FWI and MVA method.

To give the reader a better sense of the underlying components of our method, we describe in more detail (Figure 3) one step of the gradient-based algorithm used in the optimization procedure (\( \alpha = 0.7 \)). At iteration 6, we calculate the FWI gradients \( \nabla_{\mathbf{m}} J_{\text{FWI}} \) for the full-band data (Figure 3a) and the band-limited data used in the multi-scale FWI approach (Figure 3b). The latter FWI gradient \( \nabla_{\mathbf{m}} J_{\text{FWI}} \) is updated at each iteration. The former FWI gradient \( \nabla_{\mathbf{m}} J_{\text{FWI}} \) aids the computation of the full-band extended FWI gradient \( \nabla_{\mathbf{m}} J_{\text{FWI}} \) which we restrict to the locations indicated by the dots in Figure 3a. For example, the extended FWI gradient \( \nabla_{\mathbf{m}} J_{\text{FWI}} \) in Figure 3c is calculated at the location represented by the magenta dot. At this location, the annihilator \( \mathbf{A} \) in Figure 3d penalizes the defocused energy. This MVA procedure in the extended space applies to all locations indicated by the dots in Figure 3a. Conserving the full-band data yields better resolution for the focus estimation in the extended FWI gradients \( \nabla_{\mathbf{m}} J_{\text{FWI}} \). Figure 3e shows the resulting MVA gradient \( \nabla_{\mathbf{m}} J_{\text{MVA}} \). Weighting and combining the FWI and MVA gradients (Figure 3b and 3e, respectively) yields the gradient \( \nabla_{\mathbf{m}} J \) for the combined FWI and MVA method, shown in Figure 3f. A line-search algorithm determines the scale to apply to the combined gradient \( \nabla_{\mathbf{m}} J \) which is necessary to update model \( \mathbf{m} \) at each iteration. The combined gradient \( \nabla_{\mathbf{m}} J \) exhibits the joint properties of the FWI and MVA gradients: MVA contributes to the long-wavelength component of the gradient while FWI contributes to its short-wavelength component. In other words, MVA essentially estimates the component of the model responsible for the kinematic of seismic wave propagation while FWI simultaneously updates model features with higher resolution. The joint FWI and MVA update yields the convergence to a high-resolution optimal model \( \mathbf{m}_{\text{opt}} \), which exhibits the blended behavior and robustness of FWI and MVA. The method is not as sensitive to the initial model as FWI and is less likely to converge to a local minima solution (because of cycle skipping). The method also achieves higher resolution than MVA and is more likely to handle nonlinear wave phenomena since the MVA procedure does not directly apply to a seismic image but to the FWI gradient, which results from only the backprojection of the data misfit into the Earth’s model. It is also worth mentioning that in terms of computational costs, our method costs the same as the selected MVA method because the apparent extra calculation for the FWI gradient is equivalent to the calculation of the seismic image for MVA.

CONCLUSIONS

FWI and MVA are two complementary techniques for retrieving a model of the subsurface in seismic imaging. The combined FWI and MVA method proposed in this abstract facilitates the inversion of seismic reflection data and contributes to solving this inverse problem without the need for separating the long- and short-wavelength components of the inverted model. Instead, our bi-objective optimization strategy solves for a multi-scaled optimal model that jointly minimizes the FWI and MVA cost functions. A global criterion method combines these two cost functions. Our strategy extends the semblance principle to sensitivity kernels instead of just seismic images. Conventional MVA techniques are then applicable to the FWI gradients, and the approach we propose is not limited to the MVA strategy that we selected for this work. Our numerical results motivate further exploration of the application of a bi-objective optimization to the inversion of seismic reflection data. We are currently working on the expansion of our method to geological models and real data.

ACKNOWLEDGMENTS

This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena. This research was supported in part by the Golden Energy Computing Organization at the Colorado School of Mines using resources acquired with financial assistance from the National Science Foundation and the National Renewable Energy Laboratory.
Bi-objective optimization for the inversion of seismic reflection data

REFERENCES

Yang, T., and P. Sava, 2012, Image-domain wavefield tomography for complex geologic structures: Presented at the EAGE/ACGGP First Latin American Geophysics Work-
Fault surfaces and fault throws from 3D seismic images
Dave Hale, Center for Wave Phenomena, Colorado School of Mines

SUMMARY
A new method for processing 3D seismic images yields images of fault likelihoods and corresponding fault strikes and dips. A second process automatically extracts from those images fault surfaces represented by meshes of quadrilaterals. A third process uses differences between seismic image sample values alongside those fault surfaces to automatically estimate fault throw vectors. While some of the faults found in one 3D seismic image have an unusual conical shape, displays of unfaulted images illustrate the fidelity of the estimated fault surfaces and fault throw vectors.

INTRODUCTION
Fault surfaces like those shown in the close-up views of Figure 1 are an important aspect of subsurface geology that can be derived from seismic images. Therefore, various fault tracking methods, including those proposed by Pedersen et al. (2002, 2003), Admasu et al. (2006), Kadlec et al. (2008) and Kadlec (2011), have been developed to extract such surfaces.

The fault throws shown in Figure 1 are important as well, as they enable correlation of subsurface properties across faults. Among methods developed to estimate fault throws are those described by Borgos et al. (2003), Aurnhammer and Tönnes (2005) and Admasu (2008).

Figure 1: Close-up views of roughly conical (a) and planar (b) fault surfaces and fault throws computed automatically from a 3D seismic image. Vertical and horizontal image slices are shown in the background. Vertical fault throws are measured in ms because the vertical axis of the image is time. Each quadrilateral intersects exactly one edge in the 4 ms by 25 m by 25 m image-sampling grid.

This paper describes a sequence of three new methods to (1) compute 3D fault images, (2) extract fault surfaces, and (3) estimate fault throws. I used this three-step sequence to compute the fault surfaces and throws displayed in Figures 1 and 2. Although each of the three steps was designed in conjunction with the others in this sequence, aspects of any one of them could be used to enhance other methods cited above.

FAULT IMAGES
Before extracting fault surfaces like those shown in Figures 1 and 2, I first compute images of faults. The method I use for this first step is based on semblance (Taner and Koehler, 1969), and is therefore similar to methods proposed by Marfurt et al. (1998). Like Marfurt et al. (1999), I compute semblances from small numbers (3 in 2D, 9 in 3D) of adjacent seismic traces, after aligning those traces so that any coherent events are horizontal.

Semblance is a measure of coherence in the range $[0, 1]$; it is a normalized ratio, the square of an average value divided by an average of squared values. Because faults are most likely to exist where semblance $s$ is low, I (somewhat arbitrarily) define and compute a measure of fault likelihood $f = 1 - s^8$.

When used to highlight faults, some sort of averaging or smoothing of semblance (or some other attribute) is required, as emphasized by Gersztenkorn and Marfurt (1999) and Aqrawi and Boe (2011). These authors describe vertical smoothing of fault attributes.

However, faults need not be vertical. Therefore, when averaging the numerators and denominators of normalized semblance ratios, I vary the orientation of the smoothing filter in a scan over possible fault orientations.
Fault surfaces and fault throws

Figure 2: Fault surfaces and fault throws computed for two different subsets of a 3D seismic image. Faults extracted from the shallower subset (a) have conical shapes, while those extracted from the deeper subset (b) have more typical planar shapes. Seismic reflectors are more continuous in the corresponding unfaulted images (c and d).

Figure 3 illustrates for a 2D image the results of non-vertical smoothing for two different fault dips $\theta$ in this scan. This example shows that fault likelihoods tend to be largest when the smoothing of semblance numerators and denominators is aligned with the faults, which are not vertical.

Much like Cohen et al. (2006), I scan over multiple fault strikes and dips to determine the orientation that maximizes fault likelihood. In the 3D examples shown in this paper, this scan included $N_\phi = 26$ fault strikes $\phi$ and $N_\theta = 22$ fault dips $\theta$, for a total of $N_\phi N_\theta = 572$ possible fault orientations.

The computational cost of the scan is proportional to this number of orientations. To reduce this cost, for each orientation I use efficient recursive smoothing filters to perform the averaging of semblance numerators and denominators. The computational cost of these recursive filters is independent of the spatial extent of their impulse responses, which may include well over 1000 samples in 3D images. This number represents the number of samples that contribute to the computation of fault likelihood for one orientation at one sample location in a 3D image (Cohen et al., 2006). With recursive smoothing filters I avoid this large factor in computational cost.

Figure 4a shows fault likelihoods computed with a scan over $N_\theta = 22$ fault dips for the 2D seismic image in Figure 3a. Ridges of fault likelihood in this fault image generally coincide with faults apparent in the seismic image. These ridges can be found by simply scanning each row of the fault image, preserving only local maxima, and setting fault likelihoods elsewhere to zero, as shown in Figure 4b. In effect, this process thins the fault image, significantly reducing the number of image samples at which a fault may be considered to exist.

It is important to remember that for any images of fault likelihood, such as those shown in Figure 4, we have corresponding images of fault orientation, the fault strikes and dips for which fault likelihood is maximized. These images of fault orientation are especially useful when extracting fault surfaces from 3D images.

**FAULT SURFACES**

For 3D seismic images, ridges of fault likelihood correspond to potential fault locations. However, it is more difficult to extract ridge surfaces from 3D images than to extract ridge curves from 2D images as illustrated in Figure 4.

I extract fault surfaces from 3D images of fault likelihoods $f$,
Fault surfaces and fault throws

Figure 3: Fault likelihoods computed for a 2D seismic image (a) and for two different fault dips $\theta$, one positive (b) and the other negative (c), in the scan used to estimate fault likelihoods and orientations.

strikes $\phi$, and dips $\theta$ using a method for extracting ridge surfaces similar to that proposed by Schultz et al. (2010), which they demonstrate for 3D medical images of the human brain. The fault surfaces shown in Figures 1 and 2 are ridges in 3D images of fault likelihood, and are represented by meshes of quadrilaterals (hereafter referred to as quads).

As shown in Figure 5, each quad in a fault surface intersects exactly one edge of the 3D sampling grid for the fault image. Each of the four nodes of a quad lies within exactly one cell of that grid. The coordinates of a quad node within any such cell are averages of the coordinates of all quad-edge intersections for that cell. This averaging enables representation of a fault surface with sub-voxel precision. Therefore, to find the locations of the quad nodes, we must first find the intersections of the fault surface and edges of the 3D sampling grid.

I find those surface-edge intersections using an adaptation of the method proposed by Schultz et al. (2010), in which I ensure that the orientations of ridge surfaces extracted from 3D images of fault likelihoods are consistent with the correspond-

Figure 4: Fault likelihoods computed by scanning over fault dips $\theta$, before (a) and after (b) ridge extraction.

I have made no attempt to fill any of the small holes apparent

Figure 5: Four adjacent quads in a fault surface share a node that lies within one cell of the 3D fault image sampling grid. Spatial coordinates of the quad node are averages of the coordinates of intersections of the fault surface and edges of the image sampling grid.
Fault surfaces and fault throws

in these surfaces, although such a filling process would be easy to implement because each quadrilateral is linked to its neighbors. The fact that holes are small is due to the continuity of ridges in the 3D images of fault likelihood.

**FAULT THROWS**

Because each quad in fault surfaces like those shown in Figure 1 corresponds to exactly one edge in the 3D image sampling grid, it is straightforward to walk up and down fault curves and gather samples of the 3D seismic image on opposite sides of a fault. I then compute fault throws that minimize sums of squared differences of those sample values.

This new method for computing fault throws is an adaptation of a classic dynamic programming solution (Sakoe and Chiba, 1978) to the problem of automatic speech recognition. That solution today is often called dynamic time warping and is here adapted to find a spatial warping that best aligns samples of 3D seismic images alongside faults, as illustrated in Figure 2.

One of the most attractive features of the dynamic time warping algorithm is that it optimally aligns two time series while constraining the amount of stretching or squeezing of sequences that is permitted during alignment. The relative shift (here, fault throw) between two sequences may vary with time (or depth), but dynamic time warping constrains the rate at which the shift changes with time.

My adaptation of this algorithm is to constrain the rate at which fault throw varies in both strike and dip directions along a fault. This constraint is much like that imposed in dynamic image warping (e.g., Pishchulin, 2010), in which shifts are constrained to vary slowly in both horizontal and vertical directions.

However, we cannot simply estimate fault throws by aligning a 2D image extracted from the footwall side of a fault surface with another 2D image extracted from the hanging-wall side. Consider for example the fault surface shown in Figure 6, where part of the surface lies in front of another part of that same surface. This situation occurs often in the fault surfaces shown in Figure 2. Nevertheless, quad meshes provide the left-right and up-down connectivity required to constrain changes in fault throws in both the strike and dip directions within such surfaces.

**CONCLUSION**

The methods proposed in this paper were designed as parts of a three-step process to (1) compute images of fault likelihood, strike and dip, (2) extract fault surfaces, and (3) estimate fault throws.

It is significant that the scan in the first step yields images of fault strikes and dips for which fault likelihood is maximized. These estimates of fault orientations are useful in several consistency tests performed in the second step used to extract fault surfaces.

![Figure 6: Close-up view of fault throws computed for a fault surface in which one part of the surface lies in front of another part. For such surfaces we cannot simply compute throws from footwall and hanging-wall images extracted alongside faults.](image)

The quad-mesh representation for those fault surfaces facilitates the third step of estimating fault throws. Because throw vectors connect image samples on one side of a fault to those on the other side, it is especially convenient that a quad in the fault surface lies between two adjacent samples of the seismic image. In addition, the quad mesh provides left-right and up-down connectivity needed to implement the dynamic warping algorithm used to estimate fault throws.

Most of the computation time in this three-step process lies in the first step, which currently requires a scan over all possible fault orientations. I improve the computational efficiency of this scan by using fast recursive smoothing filters for each orientation, but further improvements may be worthwhile. My current implementation of this scan for 500 fault orientations requires about two hours to process a 3D image of $1000^3$ samples on a 12-core workstation.

I did not expect to find the conical shapes of faults apparent in Figures 1 and 2, in part because I had not recognized their hyperbolic appearance in horizontal and vertical slices of the 3D seismic image. An important benefit in using an automated process to extract faults from 3D seismic images is that the process cannot exclude such shapes simply because they are unexpected.

**ACKNOWLEDGMENTS**

In this work I benefited greatly from discussions with many others, including Luming Liang, Marko Maučec, Bob Howard, Dean Witte, and Anastasia Mironova. The 3D seismic image used in this study was graciously provided via OpendTect by dGB Earth Sciences B.V.
REFERENCES


Full-waveform inversion of multicomponent data for layered VTI media
Nishant Kamath & Ilya Tsvankin, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

Although full-waveform inversion (FWI) has shown significant promise in reconstructing heterogeneous velocity fields, most existing methodologies are limited to acoustic models. We extend FWI to multicomponent (PP and PS) data from anisotropic media, with the current implementation limited to a stack of horizontal, homogeneous VTI (transversely isotropic with a vertical symmetry axis) layers. The algorithm is designed to estimate the interval vertical P- and S-wave velocities ($V_{P0}$ and $V_{S0}$) and Thomsen parameters $\varepsilon$ and $\delta$ from long-spread PP and PSV reflections. The forward-modeling operator is based on the anisotropic reflectivity technique, and the inversion is performed in the time domain using the gradient (Gauss-Newton) method. To build the initial model, we perform nonhyperbolic semblance analysis of PP and PS data. Analysis of the eigenvectors of the approximate Hessian matrix shows that the objective function is weakly sensitive to density, but the parameters $V_{P0}$, $V_{S0}$, $\varepsilon$, and $\delta$ can be resolved not only by joint inversion of PP and PS data, but also with PP reflections alone. The inversion becomes more stable with increasing spreadlength-to-depth ($X/Z$) ratio, especially for layers in the deeper part of the section. The insights gained by examining simple layered models should help guide the inversion for more realistic heterogeneous TI media.

INTRODUCTION

Transversely isotropic media with a vertical axis of symmetry (VTI) are described by the vertical velocities $V_P$ and $V_S$ and the anisotropy parameters $\varepsilon$, $\delta$, and $\gamma$. However, traveltime analysis of PP-wave reflection data typically yields just the P-wave normal-moveout velocity $V_{nmo,P}$ and anellipticity coefficient $\eta$ (Alkhalifah and Tsvankin, 1995):

$$V_{nmo,P} = V_P \sqrt{1 + 2\delta},$$

$$\eta = \frac{\varepsilon - \delta}{1 + 2\delta}.\tag{1}$$

In moveout inversion, $\eta$ is often replaced with the P-wave horizontal velocity:

$$V_{hor,P} = V_P \sqrt{1 + 2\varepsilon}.$$

Even the combination of long-spread reflection traveltimes of P-waves and mode-converted PSV-waves is insufficient for reconstructing layer-cake VTI models in depth (Grechka and Tsvankin, 2002).

Most existing FWI algorithms (Plessix and Rynja, 2010; Plessix and Cao, 2011; Gholami et al., 2011) for anisotropic media use the acoustic approximation and operate primarily with diving P-waves. Chang and McMechan (2009) perform a feasibility study of FWI applied to PP and PS data from a horizontal anisotropic layer sandwiched between isotropic media. Here, we examine the reconstruction of more realistic, multilayered VTI models in depth using full-waveform inversion of PP and PS data. By taking both anisotropy and elasticity into account, our method properly models reflection amplitudes and handles multicomponent data.

METHODOLOGY

We generate 2D synthetic PP and PSV data from a point explosive source with the anisotropic reflectivity method (Mallick and Frazer, 1990). The parameters of the top layer are assumed to be known and fixed at the correct values during the inversion.

Building the initial model

The initial model is obtained by commonly used moveout-inversion techniques. Time processing of PP reflection data in layer-cake VTI media is fully controlled by the parameters $V_{nmo,P}$ and $\eta$, which can be estimated from PP-wave traveltimes. The PP-wave long-spread reflection moveout in a horizontal VTI layer is described by the nonhyperbolic equation of Alkhalifah and Tsvankin (1995):

$$t^2 = t_{P0}^2 + \frac{x^2}{V_{nmo,P}^2} - \frac{2 \eta x^4}{V_{nmo,P}^2 V_{nmo,P}^2 + (1 + 2\eta) x^2},\tag{4}$$

where $x$ is the offset and $t_{P0}$ is the PP-wave two-way zero-offset time. Equation 4 remains valid for layered VTI media, with $V_{nmo,P}$ and $\eta$ becoming effective quantities for the stack of layers above the reflector. If the spreadlength-to-depth ratio $X/Z$ is less than 1.5, the magnitude of nonhyperbolic moveout is insufficient for constraining $\eta$. For $X/Z$ reaching 1.5-2, equation 4 is used to perform 2D semblance scanning for the effective parameters $V_{nmo,P}$ and $\eta$. Then the interval velocity $V_{nmo,P}$ is found from the conventional Dix equation and the interval $\eta$ from the Dix-type equation given in Tsvankin (2005). The initial value of $\delta$ is set to zero, which allows us to find the parameters $V_{P0}$ and $\varepsilon$ from $V_{nmo,P}$ and $\eta$. The density $\rho$ and shear-wave vertical velocity $V_S$ (if only PP data are available) for the initial model are supposed to be found from well logs.

For multicomponent data, it is necessary to identify the PP and PS (PSV) reflections from the same interfaces (i.e., perform event registration). The interval values of $V_{nmo,P}$ and $\eta$ can be calculated from P-wave data as described above. To estimate the effective PS-wave NMO velocity ($V_{nmo,PS}$), we apply a 2D semblance scan based on equation 4 to long-spread PS data. Then the effective NMO velocity $V_{nmo,SV}$ of the pure SS reflection can be found from (Seriff and Sriram, 1991):

$$2 t_{P0} V_{nmo,PS}^2 = t_{P0} V_{nmo,P}^2 + t_{S0} V_{nmo,SV}^2,$$\tag{5}

where $t_{P0}$ and $t_{S0}$ are the zero-offset traveltimes of PS- and SS-waves respectively, so $t_{S0} = 2t_{P0} - t_{P0}$. The vertical traveltimes are used to obtain the ratio $V_{P0}/V_{S0} = t_{S0}/t_{P0}$. The
Full-waveform inversion of multicomponent data for layered VTI media

interval SV-wave NMO velocity, which can be obtained from the Dix equation, is given by (Tsvankin, 2005):

\[ V_{nmo,SV} = V_{S0} \sqrt{1 + 2\sigma}, \quad \sigma = \left( \frac{V_{S0}}{V_{P0}} \right)^2 (\varepsilon - \delta). \]  

(6)

Grechka and Tsvankin (2002) show that in principle it is possible to calculate all four parameters \((V_{P0}, V_{S0}, \varepsilon, \text{ and } \delta)\) from \(V_{nmo,P}, V_{nmo,SV}, V_{P0}/V_{S0},\) and \(\eta.\) However, small errors in the NMO velocities and \(\eta\) propagate into the other VTI parameters with amplification and make the results too unstable in practice. Still, this approach provides us with an initial model to be updated by full-waveform inversion.

**Inversion algorithm**

We perform time-domain inversion of either PP data alone or the combination of PP and PS reflections. The least-squares objective function is defined as:

\[ \mathcal{F}(m) = \frac{1}{2} ||d_{obs} - d_{cal}(m)||^2, \]

(7)

where \(d_{obs}\) is the observed data and \(d_{cal}(m)\) is the data calculated for a certain model \(m.\) The model updating is carried out via the Gauss-Newton method,

\[ (\Delta m) = [J^T J]^{-1} J^T \Delta d, \]  

(8)

where \(J\) is the Fréchet derivative matrix obtained by perturbing each model parameter, \(J^T J\) is the approximate Hessian matrix, and \(\Delta d\) is the difference between the observed data and those computed for a trial model. Forward modeling is carried out with the anisotropic reflectivity algorithm of Mallick and Frazer (1990).

Since the vertical velocities and anisotropy parameters do not have the same units, it is more convenient to invert for the vertical and NMO velocities. If only PP data are used, each layer is described by the parameters \(V_{P0}, V_{nmo,P}, V_{hor,P}, V_{P0}/V_{S0},\) and density \(\rho.\) In the case of joint inversion of PP and PS data, we estimate the interval values of \(V_{P0}, V_{S0}, V_{nmo,P}, V_{nmo,SV},\) and \(\rho.\)

**INVERSION RESULTS**

First, the FWI algorithm is applied to the simple three-layer model in Figure 1. The top layer is isotropic, and its velocities and density are assumed to be known. The bottom halfspace is also known to be isotropic, but its parameters are estimated by FWI. We perform tests for data with the spreadlength-to-depth ratio \(X/Z\) ranging from one to three. For \(X/Z=1,\) \(\eta\) cannot be constrained by PP reflection traveltimes, so the initial values of \(\varepsilon\) and \(\delta\) are set to zero.

The testing shows that the interval parameters \(V_{P0}, V_{S0}, \varepsilon, \text{ and } \delta\) can be resolved by FWI, but the inversion is extremely sensitive to the starting model when the data include both PP and PS reflections. If PP and PS data are inverted with the initial \(\delta = 0,\) the algorithm converges to the correct values only for \(X/Z=1.\) For longer spreads, accurate parameter estimation requires calculating \(\delta\) from moveout inversion of PP and PS data. This is likely due to the shape of the objective function, which causes the inversion for the initial \(\delta = 0\) to get trapped in local minima.

Figure 1: Three-layer model used to test the algorithm. The parameters of the top isotropic layer are \(V_P = 2800\) m/s, \(V_S = 1400\) m/s, and \(\rho = 1.8\) g/cm\(^3\). For the VTI layer, \(V_{P0} = 3000\) m/s, \(V_{S0} = 1632\) m/s, \(\varepsilon = 0.25,\) \(\delta = 0.1,\) and \(\rho = 2.4\) g/cm\(^3\). For the bottom halfspace, \(V_P = 3400\) m/s, \(V_S = 1800\) m/s, and \(\rho = 3.2\) g/cm\(^3\).

Figure 2: Components of the eigenvectors (numbered 1 to 4) associated with the four largest eigenvalues of the Hessian. The input data include PP and PS reflections for the model in Figure 1 for \(X/Z=1.5.\) The superscript (2) denotes the VTI layer and (3) the bottom isotropic halfspace.

To evaluate the sensitivity of the objective function to the model parameters, we perform the eigenvector/eigenvalue decomposition of the Hessian matrix (Plessix and Cao, 2011) for joint inversion of PP and PS data. Each component of an eigenvector (called the “direction cosine”) indicates the relative sensitivity of the objective function to the model parameters. Figure 2 shows that the objective function is most sensitive to the layer thickness \(D\) (and hence to \(V_{P0,}\) since the vertical traveltimes are well-constrained), followed by \(V_{S0}, V_{nmo,P,}\) and \(V_{nmo,SV}.\) In contrast, all our tests demonstrate that the objective function is weakly sensitive to density. Hence, in the subsequent
Full-waveform inversion of multicomponent data for layered VTI media

tests the interval densities are fixed at the correct values.

Next, we generate only PP data for the same model and invert for the parameters \(V_{P0}, V_{nmo,P}, V_{hor,P}, \) and \(V_{P0}/V_{S0}\) using the same range of spreads. For all values of \(X/Z,\) the algorithm converges to the correct parameters, even when the initial value of \(\delta\) is set to zero. Evidently, the objective function has a simpler shape with fewer local minima, if only PP data are included.

Interestingly, for \(X/Z=1\) the inversion yields accurate VTI parameters despite the absence of PS data. This is an unexpected result since traveltimes on short spreads are not sufficient for constraining the horizontal velocity (or the parameter \(\eta\) and, therefore, \(\epsilon\)). However, the geometrical-spreading factor near the symmetry axis is sensitive to \(\eta\) (Tsvankin, 1995, 2005; Xu et al., 2005), which helps estimate all relevant parameters using full-waveform data.

For PP-waves, the eigenvectors associated with the two largest eigenvalues of the Hessian (Figure 3) point equally in the direction of two model parameters \((V_{P0} \text{ and } D),\) so the objective function is sensitive to the combination of \(V_{P0}\) and \(D.\) The third eigenvector, on the other hand, points almost entirely in the direction of \(V_{hor,P} \text{.}\) As mentioned above, the near-offset P-wave amplitude is influenced by \(\eta,\) which may help resolve \(V_{hor,P}.\)

The objective function for PP-wave inversion is not as sensitive to the \(V_{P0}/V_{S0}\) ratio as it is to \(V_{P0}\) and \(D.\) However, the P-wave amplitude-variation-with-offset (AVO) gradient (and the P-wave reflection coefficient as a whole) includes the jump in the shear-wave vertical rigidity modulus \(G = \rho V_{S0}^2\) (Rüger, 1997, 2002), which creates a dependence of the FWI objective function on \(V_{S0}.\)

When larger offsets are included, the velocity \(V_{hor,P}\) (or \(\eta\)) is well-resolved even in the presence of random noise because it governs the magnitude of nonhyperbolic moveout (Figure 4). The small errors (up to 0.02) in the inverted parameters \(\epsilon\) and \(\delta\) are mostly related to the slight distortion in the vertical velocity \(V_{P0}.\)

![Figure 3: Components of the eigenvectors (numbered 1 to 4) associated with the four largest eigenvalues of the Hessian. The input data include PP reflections for the model in Figure 1 for \(X/Z=1\).](image)

![Figure 4: Parameters of the VTI layer (circles) after each iteration of FWI. The actual values are marked by the horizontal solid lines. The input data include PP reflections for the model in Figure 1 for \(X/Z=2\). The data are contaminated with band-limited (10-25 Hz) random noise; the signal-to-noise ratio is five.](image)

Next, we test the algorithm for a more complicated multilayered VTI model (Figure 5). Again, the parameters of the top layer are fixed at the correct values, and the bottom half-space is known to be isotropic. We contaminate PP data with band-limited (10-25 Hz) random noise. The eigenvector/eigenvalue decomposition of the Hessian matrix indicates that the objective function is most sensitive to the parameters \(V_{P0}, V_{hor,P},\) and \(D\) of the shallow VTI layer (layer 2) and to the P-wave velocity in the isotropic layer immediately below it. The influence of the parameters of the deeper layers on the objective function is much weaker. When the spreadlength is equal to the depth of the bottom of the model, the spreadlength-to-depth ratio for the bottom of layer 2 (\(X/Z\)) is close to 2.2. Then the parameters of that layer are well-constrained, but there are significant errors for the deeper VTI layer (layer 4, Figure 6).

However, for longer spreads (the ratio \(X/Z\) for the bottom of the model is equal to two), the parameters of layer 4 are accurately resolved (Figure 7). Therefore, when data include sufficiently long offsets for the target horizon (which may not be typical in practice), it is possible to invert for \(V_{P0}, V_{S0}, \epsilon,\) and \(\delta\) with only PP-waves. Even in the presence of band-limited random noise, the error in \(V_{P0}\) for layer 4 is less than 2.2%, and the errors in \(\epsilon\) and \(\delta\) do not exceed 0.03 (Figure 7). Performing isotropic FWI for VTI media, on the other hand, can lead to depth stretching or overestimation of velocity in the deeper layers (Gholami et al., 2011).

**CONCLUSIONS**

It is well known that the depth scale of horizontally layered VTI models is not constrained by reflection traveltimes of PP- and PS-waves, even if long-spread data are acquired. We show that the interval vertical P- and S-wave velocities and anisotropy parameters \(\epsilon\) and \(\delta\) of layer-cake VTI media can be estimated by full-waveform inversion of PP and PS reflection data.
If the densities are fixed at the correct values, the parameters $V_{P0}$, $V_{S0}$, $\varepsilon$, and $\delta$ may be constrained by PP-waves alone. The sensitivity of the objective function to the interval parameters decreases for the deeper layers. However, if the spreadlength-to-depth ratio $X/Z$ for the bottom of a VTI layer reaches two, its parameters can be obtained from the inversion of PP data. Still, it might be beneficial to use multicomponent (PP and PS) data if the level of noise is high. The results of 1D inversion provide useful insights for designing the inversion operator capable of handling more complicated heterogeneous structures.

The analysis performed for stratified VTI media can be generalized for vertical symmetry planes of azimuthally anisotropic models (e.g., orthorhombic). However, geometrical spreading in the symmetry planes of orthorhombic media is influenced by azimuthal velocity variations and has to be modeled with a 3D algorithm.

ACKNOWLEDGMENTS

We are grateful to the members of the A(nisotropy)-Team of the Center for Wave Phenomena (CWP), Colorado School of Mines, for fruitful discussions. We would also like to thank Jyoti Behura (CWP) for participating in our numerous brainstorming sessions. This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.
REFERENCES

Estimation of time-lapse velocity change using repeating earthquakes with different locations and focal mechanisms

Chinaemerem O. Kanu1, Roel Snieder1 and Dan O’Connell2
Center for Wave Phenomena, Colorado School of Mines, Golden CO 804011
Fugro Consultants, Inc., William Lettis Associates Division, Lakewood, CO 804012

SUMMARY

Codas of repeating earthquakes carry information about the time-varying properties of the subsurface or reservoirs. Some of the changes within a reservoir change the seismic velocity and thereby the seismic signals that travel through the reservoir. In characterizing these velocity changes, seismic signals are often influenced by changes in other properties of the reservoir such as fluid migration or the properties of the seismic sources of the signals. We investigate the impact of the perturbations in seismic source properties on time-lapse velocity estimation. We suggest a criterion for selecting seismic events that can be used in velocity analysis. This criterion depends on the dominant frequency of the signals, the center-time of the used time window in a signal, and the estimated relative velocity change. The criterion provides a consistent framework for monitoring changes in subsurface velocities using microseismic events.

INTRODUCTION

Monitoring temporal changes within the Earth’s subsurface is a topic of interest in many areas of geophysics. These changes can result from an earthquake and its associated change in stress (Cheng et al., 2010), fluid injection or hydrofracturing (Davis et al., 2003), and oil and gas production (Zoback and Zinke, 2002). Some of the subsurface perturbations induced by these processes include temporal and spatial velocity changes, stress perturbations, changes in anisotropic properties of the subsurface, and fluid migration. Many of these changes span over a broad period of time and might even influence tectonic processes, such as induced seismicity (Zoback and Harjes, 1997). A seismic velocity perturbation of the subsurface leads to progressive time shifts across the recorded seismic signals. Various methods and data have been used to resolve the velocity perturbations. These methods include seismic coda wave interferometry (Snieder et al., 2002), doublet analysis of repeating microseismic and earthquake coda signals (Poupinet et al., 1984), time-lapse tomography (Vesnaver et al., 2003), and ambient seismic noise analysis (Sens-Schönfelder and Wegler, 2006; Cheng et al., 2010; Meier et al., 2010). Earthquake coda signals have higher sensitivity to the changes in the subsurface because multiple scattering allows these signals to sample the area of interest multiple times. However, doublet analysis of the earthquake (microseismic) coda requires repeating events. Failure to satisfy this requirement can compromise the accuracy of the estimated velocity changes.

Fluid-triggered microseismic events often are repeatable, but then events occur at slightly different positions with somewhat different source mechanisms (Sasaki and Kaieda, 2002; Go-dano et al., 2012). Imprints of the source perturbation and the velocity change on the seismic waveforms can be subtle, thereby retaining the similarity between seismic signals. Therefore, we will need to ask, how do the source location, source mechanism, and subsurface perturbations affect the estimated velocity changes?

In this study, we investigate the impact of changes in source properties on the estimation of relative velocity changes. Knowledge of the impact of these perturbations on the estimated velocity change allows for a consistent framework for selecting pairs of earthquakes or microearthquakes used in velocity change analysis. This results in a more robust estimation of velocity change.

MATHEMATICAL CONSIDERATION

In this section, we use the time-shifted cross-correlation (Snieder, 2002, 2006) to develop an expression for the average value of the time perturbation of scattered waves that are excited by sources with varying source properties. This perturbation is due to changes in the velocity of the subsurface and to changes in the source properties. Figure 1 is a schematic figure showing the general setup of the problem we are investigating. Two sources ($S_1$ and $S_2$) represent a microseismic doublet (repeating microseismic events). These events occur at different locations and can have different rupture patterns. We assume that these events can be described by a double couple. We investigate the ability of using the signals of these sources for time-lapse monitoring of velocity changes, assuming that these sources occur at different times. We express the signals of the two sources as unperturbed and perturbed signals.

The unperturbed seismic signal $U(t)$ is given as

$$U(t) = A \sum_T U^{(T)}(t)$$

and the perturbed seismic signal $\hat{U}(t)$ is given as

$$\hat{U}(t) = \hat{A} \sum_T (1 + r^{(T)}) U^{(T)}(t - t_p^{(T)})$$

where $A$ and $\hat{A}$ are the amplitudes of the unperturbed and perturbed source signals, respectively. These amplitudes represent the strengths of the sources. The recorded waves are a superposition of wave propagation along all travel paths as denoted by the summation over paths $T$. The change in the source focal mechanism only affects the amplitude of the wave traveling along each trajectory $T$ because the excitation of waves by a double couple is real (Aki and Richards, 2002). The change in the signal amplitudes - due to changes in the source mechanism angles - is defined by $r^{(T)}$ for path $T$, and $t_p^{(T)}$ is the
time shift on the unperturbed signal due to the medium perturbation for path $T$. The time-shifted cross-correlation of the two signals is given as

$$C(t_c) = \int_{t-c}^{t+c} U(t')\hat{U}(t' + t_c) \, dt',$$

where $t$ is the centertime of the employed time window and $2\tau_w$ is the window length. The normalized time-shifted cross-correlation has a maximum at a time lag equal to average time perturbation ($t_s = \langle t_p \rangle$) of all waves that arrive in the used time window:

$$\frac{1}{C(0)} \left. \frac{\partial C(t_s)}{\partial t} \right|_{t_2-t_1} = 0.$$  \hfill (4)

Equation (4) allows for the extraction of the average traveltime perturbation from the cross-correlation. In this study, the average of a quantity $F$ is a normalized intensity weighted sum of the quantity (Snieder, 2006):

$$\langle F \rangle = \frac{\sum A_i^2 F_i}{\sum A_i^2}.$$ \hfill (5)

where $A_i^2 = \int U(2t')^2 \, dt'$ is the intensity of the wave that has propagated along path $T$. The average value of the time perturbation and its variances are given by

$$\langle t_p \rangle = -\left( \frac{\delta V}{V_0} \right) t_c$$ \hfill (6)

and

$$\sigma^2 = \langle t_p^2 \rangle - \langle t_p \rangle^2 \simeq \frac{D^2}{3V_0^2}.$$ \hfill (7)

In the above equations, $(\delta V / V_0)$ is the average relative velocity change, $D$ is the shift in the source location, $t_c$ is the centertime of the processed time window and $V_0$ is the unperturbed velocity. Equation 6 suggests that the average time shift in the multiple scattered signals results only from the velocity changes within the subsurface. The variance of the time shifts depends, however, on the perturbations of the source location.

Equation (3)

\begin{equation}
C(t_c) = \int_{t-c}^{t+c} U(t')\hat{U}(t' + t_c) \, dt',
\end{equation}

where $t$ is the centertime of the employed time window and $2\tau_w$ is the window length. The normalized time-shifted cross-correlation has a maximum at a time lag equal to average time perturbation ($t_s = \langle t_p \rangle$) of all waves that arrive in the used time window:

\begin{equation}
\frac{1}{C(0)} \left. \frac{\partial C(t_s)}{\partial t} \right|_{t_2-t_1} = 0.
\end{equation}

Equation (4) allows for the extraction of the average traveltime perturbation from the cross-correlation. In this study, the average of a quantity $F$ is a normalized intensity weighted sum of the quantity (Snieder, 2006):

\begin{equation}
\langle F \rangle = \frac{\sum A_i^2 F_i}{\sum A_i^2}.
\end{equation}

where $A_i^2 = \int U(2t')^2 \, dt'$ is the intensity of the wave that has propagated along path $T$. The average value of the time perturbation and its variances are given by

\begin{equation}
\langle t_p \rangle = -\left( \frac{\delta V}{V_0} \right) t_c
\end{equation}

and

\begin{equation}
\sigma^2 = \langle t_p^2 \rangle - \langle t_p \rangle^2 \simeq \frac{D^2}{3V_0^2}.
\end{equation}

In the above equations, $(\delta V / V_0)$ is the average relative velocity change, $D$ is the shift in the source location, $t_c$ is the centertime of the processed time window and $V_0$ is the unperturbed velocity. Equation 6 suggests that the average time shift in the multiple scattered signals results only from the velocity changes within the subsurface. The variance of the time shifts depends, however, on the perturbations of the source location.

Figure 1: Geometry of wave paths. The sources are $S_1$ and $S_2$. Path $T$ shows the scattering path for the unperturbed signal and the scattering path for the perturbed signal is defined by Path $T'$. The sources are separated by a distance $D$ and the source distances are $L_T$ and $L_{T'}$, distance away from the first scatterer along path $T$ and $T'$, respectively.

NUMERICAL VALIDATION

Figure 2: The experiment geometry of the numerical simulation. The receivers (squares) are surrounding the point scatterers (black dots). The source is positioned in the origin (cross). All perturbations of the source location is done from this position. The stations marked (NW, NE, E, SE, and SW) are used in the presentation of results in Figures 4 and 5.

Description of numerical experiments

We test the equations in section 2 with the numerical simulation using Foldy’s multiply scattering theory (Foldy, 1945) described by (Groenenboom and Snieder, 1995). The theory models multiple scattering of waves by isotropic point scatterers. We conduct our numerical experiments using a circular 2D geometry (Figure 2) with point scatterers surrounded by 96 receiver stations. We uniformly assign the imaginary component of the scattering amplitude $|ImA|$ = 4 to all the scatterers. In 2D, this is the maximum scattering strength that is consistent with the optical theorem that allows for conservation of energy (Groenenboom and Snieder, 1995). We assume an acoustic propagating wavefield which is a sum of the direct wavefield $U_0(r)$ and the scattered wavefield from all the scatterers:

\begin{equation}
U(r) = U_0(r) + \sum_{i=1}^{n} G^{(0)}(r, r_i) A_i U(r_i),
\end{equation}

where the scattered wavefield is described by the second term. The Green’s function at $r$ due to scatterer $i$ at position $r_i$ is $G^{(0)}(r, r_i)$ while $U(r_i)$ is the incoming wavefield at scatterer $i$. The scattering amplitude of each scatterer is given by $A_i$. The direct wave is modulated by the far-field P-wave radiation pattern $F^p(\lambda, \delta, \phi)$ (Aki and Richards, 2002):

\begin{equation}
U_0(r) = F^p(\lambda, \delta, \phi) G^{(0)}(r, r_s),
\end{equation}

with $r_s$ the source location, where $\lambda$, $\delta$, and $\phi$ are the source parameters (rake, dip and strike, respectively). Sources are located at the center of the scattering area. The source spectrum
has a dominant frequency \( f_d \) of approximately 30 Hz and a frequency range of 10-50 Hz. We assume a reference velocity \( V_0 = 3500 \text{ m/s} \). Because the model we are using is an isotropic multiple scattering model, the transport mean free path is the same as the scattering mean free path: \( l' = l \) which is approximately 12.2 km. There is no intrinsic attenuation in the numerical model. We generate an unperturbed multiple scattered signals with a reference model defined by the following parameter values: the source angles \( \phi = 0^\circ \), \( \lambda = 0^\circ \), \( \delta = 90^\circ \); change in medium velocity \( \Delta V = 0 \text{ m/s} \); and shift in the source location \( D = 0 \text{ m} \). In order to understand the effect of the perturbation of these parameters on velocity change estimation, we generate signals from the perturbed version of the model.

**Data processing**

To estimate velocity change imprints on the synthetic signals due to the perturbation of the source location or its radiation properties, we use the stretching algorithm of (Hadziioannou et al., 2009). In the stretching algorithm, we multiply the time of the perturbed signal with a stretching factor \((1 - \epsilon)\) and interpolate the perturbed signal at this stretched time. We then stretch the perturbed signal at a regular interval of \( \epsilon \) values. To resolve the value of \( \epsilon \), we use an \( L_2 \) objective function rather than the cross-correlation algorithm as suggested by (Hadziioannou et al., 2009). For events of equal magnitude \((A = \hat{A})\), the objective function is

\[
R(\epsilon) = \| \hat{U}(t(1 - \epsilon)) - U(t) \|_2,
\]

where \( \| \cdots \|_2 \) is the \( L_2 \) norm. The minimum of the objective function depends on the amplitude changes between the two signals and on the travel time perturbations due to velocity changes and shifts in the source location.

**Effect of perturbation of source properties on the estimated velocity change**

To understand the effect of the changes in the source properties on the estimation of the relative velocity change, we conduct our numerical experiment over a range of parameter changes. We perturb the source location and the orientation of the source angles. The perturbation of the source radiation parameters is characterized by the weighted root mean square change in source parameters \( \langle r \rangle \) (Robinson et al., 2007):

\[
\langle r \rangle = -\frac{1}{2}(4\Delta \phi^2 + \Delta \lambda^2 + \Delta \delta^2),
\]

where \( \Delta \phi \) is the change in strike, \( \Delta \lambda \) is the change in rake, and \( \Delta \delta \) is the change in dip. These changes represent the angle differences between the two sources (doublet).

Figure 3 shows that the estimated relative velocity changes \( \langle \delta V/V_0 \rangle \) are near the true value \( \langle \delta V = 0 \rangle \) for models with perturbations of either the source location or the source radiation parameters. The velocity change estimated from individual stations varies around zero, but with a shift in the source location of \( D = 0.143 \lambda_d \), with \( \lambda_d \) the dominant wavelength and source angle perturbations as large as \( \Delta \phi = 20^\circ \), \( \Delta \lambda = 20^\circ \), and \( \Delta \delta = 20^\circ \), \( \langle r \rangle \approx -0.366 \), the magnitude of these variations is smaller than 1/20th of the typical velocity changes inferred from seismic signals (Figure 3). These results agree with equation (6) which predicts that the average value of time shifts in the perturbed signal results only from changes in the medium velocity and is not affected by changes in source properties.

**Limiting regimes of the estimations**

To investigate the extent of the perturbation in the source location and source radiation perturbations that can be allowed in the estimation of relative velocity changes, we generate synthetic signals with models having a 0.1% relative velocity change and various perturbations of the source parameters. Comparing the phase changes due to shift in source location to those due to velocity changes, the shift in the source location has to satisfy

\[
\frac{D}{\lambda_d} < \sqrt{2} \left( \frac{|\delta V|}{V_0} \right) f_d t_c,
\]

for an accurate estimation of the velocity changes. Figure 4 shows the estimated relative velocity changes from signals generated from sources at different locations. Using the model parameters \( |\delta V/V_0| = 0.001 \), \( t = 10 \text{ s} \), and \( f_d = 25 \text{ Hz} \), the constraint on the source location shift for Figure 4 is \( D/\lambda_d < 0.35 \). Figure 4 shows that for \( D/\lambda_d \geq 0.3 \), the estimated velocity change deviates significantly from the real velocity change; this is in agreement with equation (12). The criterion in equation (12) imposes a constraint on the spacing requirements for the source locations of the doublets used for time-lapse velocity change monitoring with microearthquakes. Alternatively, equation (12) gives the magnitude of a velocity change that is resolvable with a given shift in the source location. According to equation (12), the allowable source separation increases with the centertime \( t_c \) of the employed time window. This is
due to the fact that the imprint of the velocity change is more pronounced as the waves have propagated over a greater distance through the perturbed medium.

We also investigate the effect of the source radiation properties on the estimated velocity change of the medium of interest. Figure 5 shows the estimated velocity changes from a model with 0.1% velocity change using sources with perturbed radiation angles (measured by $r$). In Figure 5A, the estimated velocity change at the individual stations progressively deviates from the true velocity change of 0.1% with increasing change in the orientations of the source angles. This deviation is due to the decorrelation between the perturbed and the unperturbed signals as shown in Figure 5B, which shows the maximum normalized cross-correlation of the codas within the processed time window. With an increasing change in the orientation of the sources, the maximum cross-correlation value of the waves excited by the doublets decreases. However, for source angle perturbations as large as $\Delta \phi = 28^\circ$, $\Delta \lambda = 28^\circ$, and $\Delta \delta = 28^\circ$ corresponding to $|\langle r \rangle| = 0.72$ (Figure 5A), the maximum deviation from the 0.1% model velocity change is approximately 0.01%. This is a small change compared to velocity changes resolved from seismic signals in practice. The maximum cross-correlation (Figure 5B) can be used as a diagnostic of the accuracy of the estimated velocity change. In this example, a maximum cross-correlation of 0.7 indicates an error of about 10% in the estimated velocity change.

CONCLUSION

In this study, we investigate the influence of perturbation in source properties (location and radiation) on the estimation of velocity changes. These velocity changes are extracted from multiply scattered signals (codas) of repeating events. We show that we can resolve accurate values of relative velocity changes to source angle perturbations, measured by maximum cross-correlation values. Stations SW, SE, NE, and NW positions are given in Figure 2.

Figure 5: Effect of source angle perturbation on estimated velocity change. A. The estimated relative velocity changes are from a 0.1% model velocity change and various source radiation perturbations. B. The decorrelation of the doublets, due to source angle perturbations, measured by maximum cross-correlation values. Stations SW, SE, NE, and NW positions are given in Figure 2.

if the shift in the source location satisfies the constraint (equation 12). This constraint depends on the dominant frequency of the signal, the estimated relative velocity change, and the center time of the employed time window. This places a restriction on the relative event locations that can be used to estimate the relative velocity change of the subsurface. Using doublets that do not satisfy the constraint result to an inaccurate estimate of the velocity change.

A significant change in the source mechanism of double couple sources can introduce a bias in the estimation of relative velocity change. This bias is due to the decorrelation of the perturbed and unperturbed signals which lowers the accuracy of the estimated velocity change. However, this bias is negligible for the typical velocity changes resolved from seismic signals in practice. This result permits the use of sources of different orientations for the estimation of velocity changes, provided that the maximum cross-correlation of the source signals is greater than 0.7 as shown in Figure 5B. For a consistent estimate of the velocity change, using multiple stations is useful to ascertain the accuracy of the estimated velocity change in an isotropic subsurface.

ACKNOWLEDGMENTS

We are grateful for the financial support of the Department of Energy (DOE) through grant DE-EE0002758.
REFERENCES


Unfaulting and unfolding 3D seismic images
Simon Luo & Dave Hale
Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401, USA

SUMMARY

One limitation of automatic interpretation methods such as seismic image flattening is their inability to handle geologic faults. To address this limitation, we propose to combine a method for automatic image unfaulting with seismic image flattening. First, using fault surfaces and fault throw vectors estimated from an image, we interpolate throw vectors to produce a throw vector field, which we use to unfault the image. Then, using seismic image flattening, we unfold the unfaul	ted image to obtain a new image in which sedimentary layers are horizontal and also aligned across faults. From this flattened unfaul	ted image, we can automatically extract geologic horizons.

INTRODUCTION

Extracting isochronal geologic surfaces—geologic horizons of the same age—is a common problem in geophysics and geology. Such horizons are useful for interpretation of stratigraphic features and analysis of structural deformation, as well as interpolation and correlation of subsurface properties. A geologic horizon is assumed to have been initially deposited as a horizontal layer and subsequently subjected to faulting and folding. So in order to extract such a horizon, it is necessary to quantify faulting and folding apparent in a seismic image.

Perhaps the most straightforward way to extract a geologic horizon is by manual picking. Manual picking is often used in conjunction with autotracking methods (e.g., Howard, 1990), which track seismic events by following local extrema or zero crossings in amplitude in a seismic image. Alternatives to horizon autotracking methods include volume interpretation methods (e.g., Stark, 1996), which, rather than tracking single events, process simultaneously an entire seismic volume. Automatic seismic image flattening (Stark, 2004; Lomask et al., 2006; Parks, 2010) is an example of a volume interpretation method. Automatic seismic image flattening could potentially identify all horizons in an image, but the method is unable to match horizons across faults unless additional information (e.g., fault throw) is provided.

A fully automatic method for extracting geologic horizons is ideal. Toward this end, we propose an automatic method that can be used to extract all geologic horizons in an image, consisting of two steps: image unfaulting followed by image unfolding (i.e., image flattening). To unfault an image, we first use the method described by Hale (2012) to estimate fault locations and fault throw vectors, displacement vectors along the dip direction of a fault surface. To unfold an unfaulted image, we use non-vertical image flattening (e.g., Luo and Hale, 2011). By unfaulting and then unfolding an image, we obtain an image in which a surface of constant relative geologic time (i.e., a horizontal slice) maps to a geologic horizon.

IMAGE UNFAULTING

To unfault an image, we must first estimate fault locations and fault slip. For the examples shown in this paper, we use the method described by Hale (2012) to automatically compute fault surfaces and fault throws from a 3D seismic image. Although we use Hale’s (2012) method, other methods (e.g., Borgos et al., 2003; Carrillat et al., 2004; Skov et al., 2004; Aurnhammer and Tönnes, 2005; Admasu, 2008; Liang et al., 2010) could also be used to estimate fault locations and fault throw.

For an image \( f(x) \), where \( x = (x_1, x_2, x_3) \) are coordinates in the present-day space, the estimated fault throw vectors \( \tilde{t}(w) \), where \( w = (w_1, w_2, w_3) \) are coordinates in the unfaul	ted space, can be used to compute an image

\[
\tilde{f}(w) = f(w + \tilde{t}(w)) ,
\]

in which seismic events are aligned across faults where \( \tilde{f}(w) \) is specified. An example of a fault throw vector for a synthetic 2D seismic image is shown in Figure 1. In the figure, \( w_f \) indicates the location of an image sample on the footwall side of the fault, while \( w_h \) indicates the location of the corresponding sample on the hanging wall side of the fault. The fault throw vector \( \tilde{t}(w_h) \) specifies the location of the image sample that, once shifted to \( w_h \), on the hangi
g wall side, aligns with the image sample at \( w_f \) on the footwall side. Note from equation 1 that events are shifted only at locations where the fault throw \( \tilde{f}(w) \) is specified. Because we estimate fault throw only at locations where we have identified a fault surface, we must interpolyate fault throw vectors at locations between faults to avoid creating new discontinuities in an image when unfaulting. To interpolate fault throw vectors at locations between faults, we use blended neighbor interpolation (Hale, 2009).

Figure 1: A fault throw vector. At \( w_f \) on the footwall side of the fault, the fault throw is zero. At \( w_h \) on the hanging wall side of the fault, the fault throw vector \( \tilde{t}(w_h) \) specifies the location of the image sample that, once shifted, aligns with the sample on the footwall side.
Unfaulting and unfolding

Figure 2: A seismic image (a) overlaid with the vertical component of fault throw vectors, the unfaulted image (b), and the flattened unfaulted image (c) computed using the composite shift vectors (d).

Our convention is that fault throw vectors \( \mathbf{t}(w) \) specify throws on the hanging wall side of a fault. Because we will interpolate these throw vectors (e.g., Figure 2a) between faults, we must also specify fault throw vectors on the footwall side of a fault so that the relative throws on opposing sides of a fault do not change after interpolation. Because fault throw vectors specify throws on only the hanging wall side of a fault, the fault throws on the footwall side must be zero (see Figure 1).

Figures 2a and 3a show subsections of a 3D seismic image from offshore Netherlands with the vertical component of the estimated fault throw vectors overlaid. Using the interpolated throw vectors \( \mathbf{t}(w) \) estimated from an input image \( f(x) \), the unfaulted image \( h(w) \) is computed as

\[
h(w) = f(w + \mathbf{t}(w)).
\]

(2)

Figure 2b shows the unfaulted image computed according to equation 2 from the input seismic image shown in Figure 2a and the blended neighbor interpolation of the fault throw vectors also shown in Figure 3a.

### IMAGE FLATTENING

To flatten an (unfaulted) image \( h(w) \), we must find a mapping \( w(u) \), where \( u = (u_1, u_2, u_3) \) are coordinates in the flattened space, such that the image \( g(u) = h(w(u)) \) is flat. We write the mapping \( w(u) \) in terms of a shift vector field \( \mathbf{r}(u) \):

\[
w(u) = u - \mathbf{r}(u),
\]

(3)

which has a corresponding Jacobian matrix \( \mathbf{J} = \partial w / \partial u \):

\[
\mathbf{J} = \begin{bmatrix}
1 - \partial r_1 / \partial u_1 & -\partial r_1 / \partial u_2 & -\partial r_1 / \partial u_3 \\
-\partial r_2 / \partial u_1 & 1 - \partial r_2 / \partial u_2 & -\partial r_2 / \partial u_3 \\
-\partial r_3 / \partial u_1 & -\partial r_3 / \partial u_2 & 1 - \partial r_3 / \partial u_3
\end{bmatrix}.
\]

(4)

Next, given normal vectors \( \mathbf{n} = (n_1, n_2, n_3) \), which we compute from an image using structure tensors (van Vliet and Verbeek, 1995; Fehmers and Höcker, 2003), we can write the Jacobian matrix for rotation:

\[
\mathbf{J}_r = \begin{bmatrix}
1 & n_3 & -n_2 \\
-n_3 & 1 & n_1 \\
-n_2 & -n_1 & 1
\end{bmatrix}
\]

(5)

To flatten an image, we solve for an approximately isometric mapping \( w(u) \) with Jacobian matrix \( \mathbf{J} \) that satisfies

\[
\mathbf{J}^\top \mathbf{J}_r = \mathbf{I}.
\]

(6)
The columns of the matrix \( \mathbf{J} \) contain vectors \( \mathbf{w}_1(\mathbf{u}), \mathbf{w}_2(\mathbf{u}), \) and \( \mathbf{w}_3(\mathbf{u}) \). That is,

\[
\mathbf{J} = \begin{bmatrix} \mathbf{w}_1(\mathbf{u}) & \mathbf{w}_2(\mathbf{u}) & \mathbf{w}_3(\mathbf{u}) \end{bmatrix},
\]

where

\[
\mathbf{w}_1(\mathbf{u}) = \frac{\partial \mathbf{w}(\mathbf{u})}{\partial u_1} = \begin{bmatrix} 1 - \frac{\partial r_1}{\partial u_1} & \frac{\partial r_2}{\partial u_1} & \frac{\partial r_3}{\partial u_1} \end{bmatrix}^\top, \tag{8}
\]

\[
\mathbf{w}_2(\mathbf{u}) = \frac{\partial \mathbf{w}(\mathbf{u})}{\partial u_2} = \begin{bmatrix} \frac{\partial r_1}{\partial u_2} & 1 - \frac{\partial r_2}{\partial u_2} & \frac{\partial r_3}{\partial u_2} \end{bmatrix}^\top, \tag{9}
\]

\[
\mathbf{w}_3(\mathbf{u}) = \frac{\partial \mathbf{w}(\mathbf{u})}{\partial u_3} = \begin{bmatrix} -\frac{\partial r_1}{\partial u_3} & -\frac{\partial r_2}{\partial u_3} & 1 - \frac{\partial r_3}{\partial u_3} \end{bmatrix}^\top. \tag{10}
\]

Vectors \( \mathbf{w}_1(\mathbf{u}) \) and \( \mathbf{w}_2(\mathbf{u}) \) are tangent to a surface (e.g., a horizon) at \( \mathbf{u} \) and thus are orthogonal to a vector \( \mathbf{n}(\mathbf{u}) \) normal to the surface at \( \mathbf{u} \). The vector \( \mathbf{w}_3(\mathbf{u}) \) is tangent to the line for which the horizontal coordinates \( u_1 \) and \( u_2 \) in the flattened space are constant, i.e., the line in coordinates \( \mathbf{w} \) that maps to a vertical line in coordinates \( \mathbf{u} \) (Mallet, 2004). For an exactly isometric mapping \( \mathbf{w}(\mathbf{u}) \), tangent vectors \( \mathbf{w}_1(\mathbf{u}), \mathbf{w}_2(\mathbf{u}), \) and \( \mathbf{w}_3(\mathbf{u}) \) are orthonormal vectors, and the corresponding Jacobian matrix is orthogonal.

Next, if we denote the columns of \( \mathbf{J} \) as \( \mathbf{w}_1(\mathbf{u}), \mathbf{w}_2(\mathbf{u}), \) and \( \mathbf{w}_3(\mathbf{u}) \), then

\[
\mathbf{J}_r = \begin{bmatrix} \mathbf{w}_1(\mathbf{u}) & \mathbf{w}_2(\mathbf{u}) & \mathbf{w}_3(\mathbf{u}) \end{bmatrix}, \tag{11}
\]

and equation 6 states

\[
\begin{bmatrix} \mathbf{w}_1 \hat{\mathbf{w}}_1 & \mathbf{w}_2 \hat{\mathbf{w}}_2 & \mathbf{w}_3 \hat{\mathbf{w}}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}
\]

The matrix \( \mathbf{J}^\top \mathbf{J}_r \) is a metric tensor characterizing local metric properties such as length, angle, area, and volume (Mallet, 2002, 2004), so by setting this matrix equal to the identity, we constrain the type of deformation parameterized by the mapping \( \mathbf{w}(\mathbf{u}) \). Equation 12 gives nine equations for the partial derivatives of the shift vector field \( \mathbf{r}(\mathbf{u}) \):

\[
\begin{align*}
n_1 \left( 1 - \frac{\partial r_1}{\partial u_1} \right) - n_2 \frac{\partial r_2}{\partial u_1} - n_3 \frac{\partial r_3}{\partial u_1} &= 0, \\
-n_1 \frac{\partial r_1}{\partial u_2} + n_2 \left( 1 - \frac{\partial r_2}{\partial u_2} \right) - n_3 \frac{\partial r_3}{\partial u_2} &= 0,
\end{align*} \tag{13}
\]
Unfaulting and unfolding

Figure 4: Geologic horizons extracted using the composite shift vector fields shown in Figure 2d (a) and Figure 3d (b).

\[ \alpha \left(1 - \frac{\partial r_1}{\partial u_1}\right) - \gamma \frac{\partial r_2}{\partial u_1} + n_1 \frac{\partial r_3}{\partial u_1} = 1, \]
\[ \gamma \left(1 - \frac{\partial r_1}{\partial u_1}\right) - \beta \frac{\partial r_2}{\partial u_1} + n_2 \frac{\partial r_3}{\partial u_1} = 0, \]
\[ -\alpha \frac{\partial r_1}{\partial u_2} + \gamma \left(1 - \frac{\partial r_2}{\partial u_2}\right) + n_1 \frac{\partial r_3}{\partial u_2} = 0, \]
\[ -\gamma \frac{\partial r_1}{\partial u_2} + \beta \left(1 - \frac{\partial r_2}{\partial u_2}\right) + n_2 \frac{\partial r_3}{\partial u_2} = 1, \]
\[ -\frac{\partial r_1}{\partial u_3} - \gamma \frac{\partial r_2}{\partial u_3} - n_1 \left(1 - \frac{\partial r_3}{\partial u_3}\right) = 0, \]
\[ -\gamma \frac{\partial r_1}{\partial u_3} - \beta \frac{\partial r_2}{\partial u_3} - n_2 \left(1 - \frac{\partial r_3}{\partial u_3}\right) = 0, \]
\[ -n_1 \frac{\partial r_1}{\partial u_3} - n_2 \frac{\partial r_2}{\partial u_3} + n_3 \left(1 - \frac{\partial r_3}{\partial u_3}\right) = 1, \]

where \( \alpha = n_3 + n_2^2/(1 + n_3), \) \( \beta = n_2 + n_1^2/(1 + n_3), \) and \( \gamma = -n_1 n_2/(1 + n_3). \) We solve equations 13, 14, and 15 for the components of the shift vector field \( \mathbf{r}(\mathbf{u}) \) by weighted least-squares using conjugate gradient iterations.

Equation 6 describes an isometric mapping of an image to a flattened image, but in general, we cannot expect to find an exactly isometric mapping for all images. In practice, this means equations 13, 14, and 15 cannot be satisfied exactly, and we must decide which equations to emphasize. For image flattening, we give most weight to equations 13, which determine the angle between the surface tangent vectors \( \mathbf{w}_1(\mathbf{u}) \) and \( \mathbf{w}_2(\mathbf{u}) \) and the normal vector \( \mathbf{n}(\mathbf{u}) \). If these equations are satisfied, then the image \( g(\mathbf{u}) \) obtained by applying the shifts \( \mathbf{r}(\mathbf{u}) \) will be flat. We give less weight to equations 14. The four corresponding entries in the metric tensor on the left side of equation 12 form what is referred to as the first fundamental form (Floater and Hormann, 2005), which characterizes lengths, areas, and angles measured on a surface. Finally, we give least weight to equations 15, which determine the length of the tangent vector \( \mathbf{w}_3(\mathbf{u}) \) and the angles it forms with the surface tangent vectors. If \( \mathbf{w}_3(\mathbf{u}) \) is a unit vector parallel to the normal vector \( \mathbf{n} \), then the thickness of sedimentary layers will be preserved in the flattening process. For most images, however, we cannot preserve thickness while flattening, so we give the corresponding equations least weight.

**HORIZON EXTRACTION**

We extract a horizon by first selecting a horizontal slice of constant \( u_3 \) in a flattened unfaulted image \( g(\mathbf{u}) \). Then, we form a composite shift vector field \( s(\mathbf{u}) \) by combining the interpolated throw vectors \( t(\mathbf{w}) \) and flattening shift vectors \( \mathbf{r}(\mathbf{u}) \) as

\[ s(\mathbf{u}) = \mathbf{r}(\mathbf{u}) - t(\mathbf{u} - \mathbf{r}(\mathbf{u})), \]

The composite shift vector field allows for a direct mapping from an image \( f(x) \) to a flattened unfaulted image \( g(\mathbf{u}) \) with

\[ g(\mathbf{u}) = f(\tilde{x}(\mathbf{u})) = \tilde{x}(\mathbf{u}) = \tilde{x}(\mathbf{u}). \]

Using the composite shift vector field \( s(\mathbf{u}) \), we map a surface of constant \( u_3 \), which corresponds to constant geologic time or constant depositional time, to a geologic horizon in present-day coordinates. For example, for \( \tilde{\mathbf{u}} = (u_1, u_2, k_3) \) where \( k_3 \) is constant, the coordinates \( \tilde{x} \) of the horizon in present-day space are simply \( \tilde{x} = x(\tilde{\mathbf{u}}) = \tilde{x}(\mathbf{u}). \)

Figures 4a and 4b show geologic horizon surfaces extracted from the composite shift vector fields shown in Figures 2d and 3d, respectively. In the horizon in Figure 4a, notice the en echelon faults that can be clearly seen in the seismic image in Figure 2a. In the horizon in Figure 4b, notice the roughly circular fault polygons, which correspond to the conical fault surfaces described by Hale (2012).

**ACKNOWLEDGEMENTS**

Thanks to dGB Earth Sciences for providing the 3D seismic images shown in this report. This research was supported by the sponsors of the Center for Wave Phenomena at the Colorado School of Mines.
REFERENCES


——, 2012, Fault surfaces and fault throws from 3D seismic images: Submitted for presentation at the 82nd Annual International Meeting, SEG, Expanded Abstracts.


——, 2004, Space-time mathematical framework for sedimentary geology: Mathematical Geology, 36, 1–32.


——, 2004, Relative geologic time (age) volumes — Relating every seismic sample to a geologically reasonable horizon: The Leading Edge, 23, 928–932.

Application of image-guided full waveform inversion to a 2D ocean-bottom cable data set

Yong Ma*, Center for Wave Phenomena, Colorado School of Mines; Jianxin (Jerry) Yuan and Yunqing Shen, ConocoPhillips Company; Bin Gong, Formerly ConocoPhillips Company, now In-Depth Geophysical

SUMMARY

Image-guided full waveform inversion (IGFWI) was proposed and synthetically tested in Ma et al. (2011). Compared with conventional FWI, IGFWI yields models (e.g., velocities) that make better geologic sense because it takes into consideration the subsurface structures, which can be apparent in migrated seismic images. Moreover, IGFWI converges faster in fewer iterations, especially when reflection energy is used to invert for high-wavenumber details in the model. We test the IGFWI method on a 2D ocean-bottom cable (OBC) data set. In this test, we first use refraction data to update the low-wavenumber component of the model and then proceed to invert for high-wavenumber details with reflection data. In the reflection stage, we incorporate structural constraints into inversion; as a result, the estimated model make plausible geologic sense.

INTRODUCTION

Full waveform inversion (FWI) (Tarantola, 1984) uses recorded seismic data \( \mathbf{d} \) to estimate parameters of a subsurface model \( \mathbf{m} \) by minimizing the data misfit function. In FWI, the objective function often takes an L2 norm: 
\[
E(\mathbf{m}) = \frac{1}{2} \| \mathbf{d} - \mathbf{F}(\mathbf{m}) \|^2. 
\]
FWI is a computationally intensive tool. It requires multiple iterations to minimize the data misfit; in each iteration, both the gradient calculation and the line search are equivalent to the cost of several seismic waveform simulations and reconstructions, which are especially expensive in 3D.

In addition to its computational cost, FWI also suffers from the nonuniqueness problem, which is caused mainly by local minima in the data misfit function. The presence of local minima is due to cycle skipping and the nonlinearity in the forward operation \( \mathbf{F}(\mathbf{m}) \) (Snieder et al., 1989). In practice, the cycle-skipping problem typically appears because it can be difficult to obtain an adequate initial model that is consistent with unrecorded low frequencies.

Both local-minima and cycle-skipping problems lead to models that poorly approximate the subsurface. To mitigate such problems, multiscale approaches (Bunks, 1995; Sirgue and Pratt, 2004; Boonyasiriwat et al., 2009) have been proposed. The fidelity of multiscale techniques depends fundamentally on the fidelity of low-frequency content in recorded data. In practice, however, the low frequencies required to bootstrap the multiscale approach may be unavailable.

Ma et al. (2011) propose image-guided FWI (IGFWI) to overcome these problems in a sparse model space. The model \( \mathbf{m} \) estimated in conventional FWI is densely sampled for simulating wave propagation. The number of dense samples is far beyond the number of necessary samples needed to geologically explain the model. Therefore, in inversion, we should only invert for a sparse model \( \mathbf{s} \), which contains as few samples as possible while still maintaining as many geological features as possible. Between the dense model \( \mathbf{m} \) and the sparse model \( \mathbf{s} \), we define a linear relationship \( \mathbf{m} = \mathbf{R}s \), where \( \mathbf{R} \) can take different forms. In IGFWI, we implement \( \mathbf{R} \) and \( \mathbf{R}^T \) with the image-guided interpolation (Hale, 2009) and its adjoint operator (Ma et al., 2010), respectively, in order to apply structural constrains derived from migrated images.

In the sparse model space \( \mathbf{s} \), we represent FWI as a sparse inverse problem, in which we minimize a new objective function: 
\[
E(\mathbf{s}) = \frac{1}{2} \| \mathbf{d} - \mathbf{F}(\mathbf{Rs}) \|^2. 
\]
The nonlinear conjugate-gradient method is still valid for solving the sparse model \( \mathbf{s} \) iteratively (Ma et al., 2010). In reality, we need a dense model \( \mathbf{m} \) to compute synthetic data \( \mathbf{F}(\mathbf{m}) \) and to fit recorded data \( \mathbf{d} \). For this reason, an image-guided conjugate-gradient method is used to solve the sparse inversion:

\[
\begin{align*}
\mathbf{h}_0 &= \mathbf{R}^T \mathbf{g}_0, \\
\beta_i &= \frac{(\mathbf{R}^T \mathbf{g}_i)^T (\mathbf{R}^T \mathbf{g}_i - \mathbf{R}^T \mathbf{g}_{i-1})}{(\mathbf{R}^T \mathbf{g}_{i-1})^T \mathbf{R}^T \mathbf{g}_{i-1}}, \\
\mathbf{h}_i &= \mathbf{R}^T \mathbf{g}_i + \beta_i \mathbf{h}_{i-1},
\end{align*}
\]

where \( \mathbf{g}_i \) and \( \mathbf{h}_i \) are the gradient of the objective function and the conjugate-gradient update direction, respectively.

An implementation of IGFWI based on conjugate gradients consists of four steps performed iteratively, beginning with an initial model \( \mathbf{m}_0 \) and a sparse model \( \mathbf{s}_0 \):

1. compute the gradient \( \mathbf{g}_0 = \mathbf{R}^T \mathbf{q}_0 \); 
2. search for a step length \( \alpha_i \), in the direction \( \mathbf{h}_i \); 
3. update the model with \( \mathbf{m}_{i+1} = \mathbf{m}_i - \alpha_i \mathbf{h}_i \); 
4. remigrate and reselect the sparse model (optional).

Although the four-step procedure finally gives a dense model solution, it in fact maintains the advantages of sparse inversion. In the sparse inversion, the sparse representation causes more blocky updates in the model and the blockiness can mitigate the absence of low frequencies in field data. Because the structure features of the subsurface are taken into consideration as constraints in the sparse inversion, IGFWI generates models that make better geological sense than conventional FWI does. Ma et al. (2011) test IGFWI in the synthetic Marmousi model. In this paper, we test IGFWI on a real 2D OBC data set.

DATA

The entire 2D data set consists of 24 shots and 235 receivers. Shot and receiver spacings are about 500 m and 50 m, respectively. Sources are 5 m below the water surface; receivers are on the sea floor (70 m beneath the water surface). Figure 1a shows one F-wave shot gather for the 2D data set; the largest
Application of IGFWI to OBC data

offset is about 12 km. Due to the large offset of the survey, this data set contains significant refraction energy, especially in the far offset. As illustrated in Figure 1a, the red line indicates the traveltime of the direct wave, and refraction energy arrives earlier than the direct wave. Because of the shallow water depth (70 m), strong surface-related multiples are present in the original data set. However, no processing is done to remove these surface-related multiples. The original data set contains a very wide frequency band. A bandpass filter is applied to the original data set to obtain the data between 4 Hz and 20 Hz, shown in Figure 1a.

METHODOLOGY

Synthetic data

The survey area is known to have strong vertical transparent isotropy (VTI). Figure 2 shows the initial P-wave velocity model that is provided by traveltime tomography and used in FWI. Figure 3a and 3b show 1D profiles of the Thomsen parameters \(\varepsilon\) and \(\delta\), respectively. To honor the strong VTI effect in wave propagation, we solve an acoustic VTI wave equation (Chu et al., 2011) that uses the pseudo-acoustic approximation (Alkhalifah, 2000) to simulate the synthetic data. One example of the 2D acoustic VTI wave equation contains two coupled partial differential equations, which can be found in Du et al. (2008). Figure 1b shows the synthetic P-wave data simulated with the initial velocity and the Thomsen parameters. If we ignore the VTI anisotropy by using an isotropic acoustic wave equation, the synthetic data has more significant traveltime error in the far offset.

Refraction FWI + reflection FWI

Because significant refraction energy is available in the far offset, we can first use the refraction data to update the velocity model. In this stage of FWI, we only match seismograms that arrive earlier than the direct wave, indicated by the red line in Figure 1a. After updating low-wavenumber components of the velocity with refractions, we then use the reflection information in the data set to update high-wavenumber details of the model. The reflection data that we use for inversion is within the red triangle region, where the maximum zero-offset time is approximately 3 s.

Conventional FWI is performed to implement the inversion with refraction data. When proceeding to the inversion with reflection data, we use IGFWI and compare IGFWI results with conventional FWI results. Both conventional FWI and IGFWI are implemented with a time-domain approach; it is equivalent to inverting all the frequencies (4 – 20 Hz) simultaneously in the frequency domain.
Application of IGFWI to OBC data

INVERSION RESULTS

Inversion with refraction data

Figure 4a displays the estimated velocity model after 5 iterations of FWI with refraction data. Compared to the initial velocity in Figure 2, the estimated velocity in Figure 4a mainly changes the initial model with low-wavenumber components because FWI only uses the refraction energy.

Inversion with reflection data

In order to allow FWI to update high-wavenumber details, we need to use reflection data. In this stage of reflection FWI, we employ the refraction-FWI-updated velocity model (Figure 4a) as the initial velocity model. Both conventional FWI and IGFWI are tested using the reflection data.

Conventional FWI

We first compute the gradient of the objective function using the adjoint-state method (Tromp et al., 2005). In other words, the gradient is achieved by performing a reverse-time migration (RTM) of the data residual $d - F(m_i)$ with the current velocity model $m_i$. We then use a quadratic line search method to update the velocity model. Figure 4b shows the velocity updated by the conventional FWI, which employs the reflection data in 5 iterations. Unlike the refraction inversion (Figure 4a), the inversion with reflection data updates the velocity model with more details by generating high-wavenumber components, as illustrated in Figure 4b. However, the high-wavenumber details are contaminated by significant artifacts that are not geologically sensible.

Image-guided FWI

Instead of using the gradient $g_i$, IGFWI computes the projected gradient $R^T g_i$ and the image-guided gradient $RR^T g_i$. In order to do this, we must first know the structural information of
Application of IGFWI to OBC data

Figure 6: RTM with the initial model (a), refraction FWI model (b), conventional reflection FWI model (c), and image-guided reflection FWI model (d).

The nonlinear conjugate-gradient method (equation 1) takes $\mathbf{R}\mathbf{r}_i$ and $\mathbf{R}\mathbf{r}_i^T\mathbf{g}_i$ to compute the update direction $\mathbf{h}_i$ that is employed in the subsequent quadratic line search. Figure 4c shows the velocity updated by the image-guided FWI, which uses the reflection data in 5 iterations. Compared to the conventional reflection FWI (Figure 4b), image-guided reflection FWI generates less artifacts and better honors the structural features. Moreover, the gather-scatter process $\mathbf{R}\mathbf{r}$ (Ma et al., 2011), which is implied by a joint operation of the image-guided interpolation and its adjoint operator, produces additional low wavenumber components.

DISCUSSION

In this test, the synthetic data takes into account VTI anisotropy, but the anisotropic parameters may not be exact because they are estimated for another nearby field. Therefore, the estimated velocity model may have errors to trade off the ambiguity between the velocity itself and the anisotropy parameters.

RTM is used to test the inversion results. Figure 6 displays the RTM images of the 2D OBC data set, which is migrated with the initial velocity, refraction-FWI velocity, conventional reflection-FWI velocity, and image-guided reflection-FWI velocity, respectively. In the oval-highlighted area of Figure 6 have gas clouds been discovered. Despite the existence of gas clouds, rock layers should maintain the structural continuity. However, broken structures are observed in the highlighted area of Figures 6a and 6b. In contrast, Figures 6c and 6d show more interpretable coherent structures, especially in the highlighted gas cloud area. Moreover, the migrated image in Figure 6d, which is done with the image-guided reflection-FWI velocity, contains less artifacts than Figure 6c.

CONCLUSION

We have demonstrated how IGFWI can improve estimating the velocity model with a 2D OBC data set. Compared to conventional reflection FWI, IGFWI with reflection data, which essentially solves a sparse inverse problem, generates velocity models that make better geological sense. This improvement is due to the fact that IGFWI uses the subsurface structures, extracted from the migrated seismic image, to constrain the inversion in the sparse model space.

ACKNOWLEDGMENTS

Authors thank ConocoPhillips for the permission to publish this work. Yong wants to thank Yanqing Zeng for his generous help to make RTM images. Yong is grateful to Chunlei Chu and Yi Wang for the reference to the acoustic VTI modeling. Yong also wants to thank Kirk Wallace for his help to generate common-image gathers, which are used to QC the inversion.
REFERENCES

SUMMARY

Shear-wave velocity is an indication of the shear strength of the ground since the velocity is related to the shear modulus. Therefore, monitoring the velocities is useful for site characterization and disaster prevention. We estimate the time-lapse change in shear-wave velocity as well as shear-wave splitting in the shallow subsurface throughout Japan by applying seismic interferometry to the data recorded with KiK-net, a strong-motion network in Japan. Each KiK-net station has two receivers; one is on the surface and the other is in a borehole. Using seismic interferometry, we extract the shear wave that propagates between these two receivers. Because KiK-net has continuously recorded strong motion seismograms since the end of 1990s, the data are available for time lapse measurements. After the Tohoku-Oki earthquake, the shear-wave velocity decreases about 5% in a region 1200 km wide and anisotropy increases more than 10%. From seasonal averages, we find the velocity and precipitation have a negative correlation.

INTRODUCTION

Seismic interferometry is a technique to obtain the wavefields that propagate between receivers [e.g., Claerbout, 1968; Lobkis and Weaver, 2001; Veenstra et al., 2004]. Seismic interferometry has been applied to ambient noise [e.g., Dragovan et al., 2009], traffic noise [e.g., Nakata et al., 2011], production noise [e.g., Vasconcelos and Snieder, 2008], and earthquake data [e.g., Sens-Schonfelder and Wegler, 2006].

In Japan, about 700 KiK-net stations are distributed to record strong motion. KiK-net is operated by the National Research Institute for Earth Science and Disaster Prevention (NIED). Each station has a borehole, a few hundred meters deep, and two seismometers at the bottom and top of the borehole. Each seismometer has three components: one vertical component and north-south and east-west horizontal components.

By applying seismic interferometry to KiK-net data, we analyze near-surface shear-wave velocities throughout Japan. Measuring time-lapse changes of the shallow subsurface is important for civil engineering and for estimating the site response to earthquakes. KiK-net has recorded seismograms continuously since the end of 1990s, and the data are available for time-lapse measurements (Sawazaki et al., 2009; Yamada et al., 2010; Nakata and Snieder, 2012a). Many foreshocks and aftershocks of the Mw 9.0 Tohoku-Oki earthquake on March 11, 2011 (see epicenters in Figure 1), which are recorded by KiK-net, allow us to analyze the changes in velocity and anisotropy caused by the main shock.

Here we present the change in shear-wave velocity and shear-wave splitting because of the Tohoku-Oki earthquake and heavy rainfall. We first introduce the data analysis method. Then we present the changes in velocity and anisotropy after the Tohoku-Oki event. Finally, we show the correlation between the velocity and precipitation.

DATA ANALYSIS

Angle of incidence of earthquake data

All the earthquakes used in this study are at a depth greater than 7 km, which is a large depth compared to the depth of the boreholes. The velocity in the near-surface is much slower than it is at greater depths. Since we consider events much deeper than the borehole, and because of the slow velocities
at the near surface, we assume the waves propagate from the borehole to the surface receivers as plane waves in the vertical direction at each station.

To confirm this assumption, we compute the angle of incidence $\theta$ by employing the procedure proposed by Nakata and Snieder (2012a) using ray tracing. All earthquake data used have $\cos \theta > 0.975$, which means the maximum of the estimated velocity bias is 2.5%. The bias is, in practice, much smaller because of averaging over many earthquakes. This inaccuracy is even less sensitive to the analysis of shear-wave splitting because $\cos \theta$ is the same for the waves in all polarization directions.

**Estimating shear-wave velocities**

We first apply deconvolution interferometry (deconvolving the seismogram at a surface receiver with that at a borehole receiver) to each KiK-net data in the north-south component and retrieve the shear wave propagating from the borehole seismometer to the surface seismometer at each station for each earthquake (Nakata and Snieder, 2012a). After applying a 1-13 Hz bandpass filter for all deconvolved waves, we estimate the shear-wave velocity by seeking the arrival time of the retrieved waveform and using the known depth of the borehole. To find the travel times, we pick three adjacent samples which have the largest amplitude and interpolate using a quadratic function because the sampling interval is not small enough to estimate the changes in the travel time caused by an earthquake (Nakata and Snieder, 2012a).

We estimate the fast and slow shear velocities caused by shear-wave splitting through seismic interferometry following the method proposed by Miyazawa et al. (2008). First, we rotate the seismograms recorded at the surface and borehole sensors in 10-degree increments in polarization. Then we deconvolve the waveforms in each polarization to obtain the wave that propagates from the borehole to the surface for all polarizations. We determine the travel times from the propagating waves and calculate the shear-wave velocities in each polarization using the depth of the borehole. Comparing the shear-wave velocities or the travel times as a function of the polarization, we find the fast and slow shear-wave polarization directions.

**CHANGES IN VELOCITY AND ANISOTROPY CAUSED BY THE TOHOKU-OKI EARTHQUAKE**

Figure 2a shows deconvolved waveforms of earthquakes between January 1, 2011 and May 26, 2011 for KiK-net station FKSH18 in the Fukushima prefecture (200 km west-southwest from the main-shock epicenter); Figure 2b shows the epicenters of the earthquakes that occurred during the periods before and after the event. The arrival times are shown with circles in Figure 2a.

The travel time measured during the main shock of the Tohoku-Oki earthquake (the large magenta circle in Figure 2a) is significantly longer than that from the other earthquakes. This indicates a reduction in the shear-wave velocity of about 22% during the shaking caused by the Tohoku-Oki event. Note also the delay of the waves in the early aftershocks indicated in red in Figure 2a. The delay of the waveforms after the Tohoku-Oki earthquake compared with the waveforms recorded before the event indicates that the shear waves propagate with a reduced shear-wave velocity after the Tohoku-Oki earthquake (Figure 2a). Nakata and Snieder (2011) discover the travel-time change during the shaking of the main event by applying short-time moving-window seismic interferometry. The reduction in velocity starts at 30 s after the origin time of the main shock and the velocity reaches the greatest reduction at 130 s after the origin time.

We average the deconvolved waves at each KiK-net station
Figure 3 shows the changes in velocity and anisotropy at KiKnet station FKSH12, which is in the Fukushima prefecture (220 km west-southwest from the main-shock epicenter). Earthquakes used in Figure 3 were recorded from May 1, 2010 to December 31, 2011. We compute the anisotropy coefficients, $(v_{\text{fast}} - v_{\text{slow}})/v_{\text{fast}}$ (where $v_{\text{fast}}$ and $v_{\text{slow}}$ are the fast and slow velocities, respectively), of each earthquake (the bottom panel in Figure 3). The top panel in Figure 3 illustrates the change in velocity. We estimate mean velocities and mean anisotropy coefficients in time periods: May 1, 2010–March 10, 2011, March 12, 2011–May 26, 2011, and May 27, 2011–December 31, 2011.

Based on Student’s $t$-test [e.g., Bulmer, 1979], the mean velocities and mean anisotropy coefficients are significantly different between each pair of periods (greater than 99.7% confidence). After the Tohoku-Oki event, the shear-wave velocity decreases and the anisotropy coefficient increases, and these changes recover with time (velocity: 770→723→743 m/s and anisotropy coefficient: 7.8→12.5→10.8%). Because the fluid condition in cracks is one major cause of anisotropy (Crampin et al., 1984), large and intermediate earthquakes, which induce the stress change to open or close cracks (Nur and Simons, 1969) and extend cracks (Atkinson, 1984), can change the anisotropy coefficient.

As shown by the moving average of the anisotropy coefficient (the blue line in the bottom panel of Figure 3), not only the main shock but also several large aftershocks might contribute to the change in polarization anisotropy since the anisotropy coefficient continues to increase for more than one month after the main shock. In contrast, the velocity suddenly decreases after the main shock (see the blue line in the top panel of Figure 3).

We compute the anisotropy coefficient for all available stations throughout Japan for January 1, 2011–March 10, 2011 and March 12, 2011–May 26, 2011. To reduce uncertainty, we use only stations which have 1) more than 3 earthquake records during both time intervals, 2) travel times of interferometric waves greater than 0.1 s, 3) anisotropy coefficients greater than 1%, and 4) a standard deviation of velocity measurements smaller than 5%.

The anisotropy coefficient in most parts of northeastern Japan increases after the earthquake. To evaluate the change in the anisotropy coefficient due to the event, we define the change in anisotropy as $(AC_{\text{after}} - AC_{\text{before}})/AC_{\text{before}}$, where $AC_{\text{before}}$ and $AC_{\text{after}}$ are the anisotropy coefficients before and after the main shock, respectively. The change in the anisotropy coefficient is shown in the left map in Figure 1. Nakata and Snieder (2012b) compare the change in the anisotropy coefficient with the change in shear-wave velocity and the change in the largest principle stress direction.
On the west side of the Median Tectonic Line and the Itoigawa-Shizuoka Tectonic Line (the black dashed lines in Figure 1), the shear-wave velocity increases and the anisotropy coefficient decreases after the main shock. We speculate that the changes could be explained by increase in the compressional stress because Kern (1978) showed with rock-physics experiments that as the confining pressure increases, velocity increases and anisotropy decreases. We cannot, however, directly measure the compressional stress in this study, so that we cannot validate this hypothesis.

CHANGES IN VELOCITY CAUSED BY HEAVY RAINFALL

We compute the monthly-averaged shear-wave velocities of the north-south polarization to investigate a possible seasonal velocity variation related to precipitation. We use only the data in the southern half of Japan because that region has a more pronounced seasonal precipitation cycle than the northern half of Japan. We calculate the mean velocities over the stations with the 15% slowest shear-wave velocities in the area since these stations are located at soft-rock sites and are therefore influenced more by precipitation than the station at hard-rock sites. In Figure 4a illustrates a significant velocity difference between spring/summer and fall/winter.

We compare the monthly-averaged velocities with the monthly average of precipitation (observed by the Japan Meteorological Agency (JMA)) computed from precipitation records over 30 years (Figure 4b). Note that the shear-wave velocity and precipitation have a negative correlation (i.e., when it rains, the velocity decreases), which is consistent with the findings of Sens-Schönfelder and Wegler (2006).

Nakata and Snieder (2012a) shows the variation in velocities by using the stations with the 85% fastest shear-wave velocities in the area, and the shear-wave velocity does not vary with precipitation. The cause of the velocity reduction is the decreased effective stress of the soil due to the infiltration of water that increases the pore pressure [Das, 1993, Section 4.19; Chapman and Godin, 2001; Snieder and van den Beukel, 2004]. We assume that for soft-rock sites most of the velocity change is caused by the effective stress change because Snieder and van den Beukel (2004) show that the relative density change with pore pressure is much smaller than the relative change in the shear modulus.

CONCLUSION

We obtain near-surface shear-wave velocities throughout Japan and time-lapse changes caused by the Tohoku-Oki earthquake and heavy rainfall by applying seismic interferometry to KiK-net data. The Tohoku-Oki event decreased the shear-wave velocity about 5% throughout northeastern Japan. The precipitation and the velocity have negative correlation in soft-rock sites. We also analyze the shear-wave splitting while using two horizontal records and the Tohoku-Oki earthquake increased the polarization anisotropy in northeastern Japan. The changes in velocity and anisotropy caused by the earthquake recovers with time.

ACKNOWLEDGEMENTS

We are grateful to NIED for providing us with the KiK-net data and to JMA for precipitation data.
REFERENCES


Miyazawa, M., R. Snieder, and A. Venkataraman, 2008, Application of seismic interferometry to extract P- and S-wave propagation and observation of shear-wave splitting from noise data at Cold Lake, Alberta, Canada: Geophysics, 73, D35–D40.


Waveform tomography based on local image correlations
Francesco Perrone∗ and Paul Sava, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

Common velocity analysis is based on the invariance of migrated images with respect to the experiment index (shot number, plane-wave take-off angle, etc.) or extension parameters (reflection angle, correlation lags in extended images, etc.). All the information available, i.e. the entire survey, is used for assessing the quality of the model used for imaging. This approach is effective but forces a clear separation between imaging and velocity model building. Here, we ask a complementary question: how much information about the velocity model is contained in a minimum number of images? Starting from an alternative statement of the semblance principle, we propose a measure of velocity error based on local correlations of pairs of migrated images. We design an objective function and implement a local, gradient-based optimization scheme to reconstruct the velocity model. Our methodology is “full-wave” because it is not based on a linearization of the imaging operator (in contrast with linearized wave-equation migration velocity analysis techniques). The gradient of the objective function is computed using the adjoint-state method.

INTRODUCTION

Seismic imaging includes estimation of both the position of the structures responsible for the recorded data and a model that describes the propagation in the subsurface. The two problems are related since a model is necessary to infer the position of the reflectors. The recorded wavefields are extrapolated in the model by solving a wave equation and crosscorrelated with a synthetic source wavefield simulated in the same model (Claerbout, 1985). Under a single scattering approximation, if the source and receiver wavefields match in time and space. If the velocity model is inaccurate, the reflectors are positioned at incorrect locations.

Wave-equation tomography is a family of techniques that estimate the velocity model from finite bandwidth signals. The inversion is usually formulated as an optimization problem, where the correct velocity minimizes an objective function that measures the inconsistency between simulated and observed data. Full-waveform inversion (FWI) (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004) addresses the estimation problem in the data space and measures the mismatch between observations and simulated data. FWI aims to reconstruct the exact model that generates the recorded data. By matching both traveltime and amplitude information, FWI allows one to achieve high-resolution results (Sirgue et al., 2010). Nonetheless, it needs a source estimate, the physics of wave propagation must be correctly modelled and a good parametrization (for example, impedance vs. velocity contrasts) is crucial (Kelly et al., 2010). Moreover, an accurate initial model is key to avoid cycle skipping and convergence to the global minimum of the objective function. Migration velocity analysis (MVA) (Fowler, 1985; Faye and Jeannot, 1986; Al-Yahya, 1989; Chavent and Jacewitz, 1995; Biondi and Sava, 1999; Sava et al., 2005; Albertin et al., 2006) defines the objective function in the image space and is based on the semblance principle (Al-Yahya, 1989). Groups of experiments and the associated images are analyzed. If the velocity model is correct, the images from different experiments must be consistent since the Earth is assumed stationary on the time scale of the seismic experiments. MVA leads to smooth objective functions and well-behaved optimization problems (Symes, 1991; Symes and Carazzone, 1991), and it is less sensitive to the initial model than FWI. On the other hand, because we lose the information about the data amplitudes, the estimated model shows lower resolution than the FWI result. MVA measures either the invariance of the migrated images in an auxiliary dimension (reflection angle, shot, etc.) (Al-Yahya, 1989; Sava and Fomel, 2003; Xie and Yang, 2008) or focusing in an extended space (Rickett and Sava, 2002; Symes, 2008; Sava and Vasconcelos, 2009). A large number of shots and good illumination are necessary to pick moveout curves or assess focusing. Moreover, because of the high memory requirement for storing the information from each experiment, only a subset of the image is considered in the evaluation of the objective function. For example, common-image gathers (CIGs) (Rickett and Sava, 2002; Yang and Sava, 2011) or common-image point (CIPs) (Sava and Vasconcelos, 2009) are computed at fixed lateral positions or picked image points along reflectors, respectively, and both methods require migration of all the shots that illuminate the location where the CIGs/CIPs are constructed.

We propose an objective function that operates in the image space and does not need CIGs. We consider small groups of images from neighboring experiments and formulate the semblance principle in the shot-image domain, instead of the extended space at selected CIGs. We use the morphologic relationship between images from neighboring experiments to define a measure of relative shifts in the image space. This approach reduces the memory requirements and allows us to include all image points in the velocity analysis step.
Waveform tomography based on local image correlations

Figure 1: Images of a flat reflector obtained using a velocity model that is (a) too high, (b) correct, and (c) too low. For each model we compute the sensitivity kernel associated with the highlighted box. Panels (d), (e), and (f) show the sensitivity kernels for the corresponding models.

Figure 2: Local image correlations ((a), (b), and (c)), penalty operators ((d), (e), and (f)), and penalized local correlations ((g), (h), and (i)) for the box highlighted in Figure 1 and the three models (velocity too high, correct, and too low, respectively).

THEORY

MVA is based on the semblance principle (Al-Yahya, 1989): a single Earth model generates the recorded data; then, if the velocity model is correct, different experiments must produce consistent images of the reflectors. The similarity between migrated images is usually assessed through the analysis of common-image gathers (CIGs), which can be constructed in different domains (reflection angle, ray parameter, sub-surface offset, etc.) (Sava and Fomel, 2003; Shen and Symes, 2008; Xie and Yang, 2008), or extended images (EIs) (Rickett and Sava, 2002; Symes, 2008; Sava and Vasconcelos, 2009). These approaches require migration of a large number of shots in order to measure moveout in the CIGs or focusing in EIs. Alternatively, we can measure the similarity of pairs of images by computing local correlations at each image point. We define the objective function

$$J(m) = \frac{1}{2} \sum_i \| \sum_{\lambda} P(x, \lambda) c_i(x, \lambda) \|^2_{x, \lambda},$$

where $c_i(x, \lambda) = \int_{w(x)} R_{i+1}(\xi - \frac{1}{2}) R_i(\xi + \frac{1}{2}) d\xi$ is the local correlation of $R_i$ and $R_{i+1}$ and $P$ is a penalty operator that highlights features that are related to velocity errors. The correlations are computed in local windows $w(x)$, which allow us to consider small subsets of the data. The index $i$ spans the shots and $m$ denotes the model. In this work, we define the objective function using pairs of shots, but we use the entire images obtained from them, instead of analyzing only sparsely selected CIGs or CIPs. The reduced input data leads to lower memory requirements with respect to other techniques based on common image-gathers or extended images. A possible alternative to measure inconsistencies between migrated images is the energy of the difference between them. Plessix (2006) uses the image difference as a regularization term for full-waveform inversion and not as a stand-alone objective function. The difference operator is a high-pass filter, enhances noise, and is sensitive to amplitude patterns. Images from different shots do not have identical amplitudes at a given spatial location because the wavefields experience different propagation paths. The bigger the separation between the shots, the higher the amplitude mismatch. The difference objective function considers this discrepancy a velocity error indicator and thus this objective function is intrinsically biased. Local correlations measure phase shifts between two images. Correlation is more robust because amplitudes are not directly used for measuring velocity errors. Nonetheless, the size of the correlation window must be such that it captures the relative shift between the two images. Local correlation in the image space is analogous to data correlation for FWI, as proposed by van Leeuwen and Mulder (2008, 2010).
Waveform tomography based on local image correlations

We restate the semblance principle as follows: if the velocity model is correct, two images from nearby experiments constructively interfere along the direction of the reflectors (Perrone and Sava, 2012). This is equivalent to having no moveout in shot-domain CIGs (Xie and Yang, 2008). Figure 1 shows the migrated images of a horizontal reflector with three velocity models (too high (a), correct (b), and too low (c)) and the sensitivity kernels ((d), (e), and (f)) for the local window highlighted by the box. The sign of the sensitivity kernel is consistent with the velocity error. Figure 2 shows the local correlations, penalty operators, and penalized correlations for the box in Figure 1. The similarity between two images is measured by the inconsistency between the local correlation and the dip estimated from one image. If the velocity model is correct, the maximum of the correlation lies along the reflector slope (the central column of Figure 2); the penalized correlation is an odd function, and the stack over correlation lags is practically zero, i.e., the two image are aligned and the model is correct. If the velocity is incorrect (the first and third column of Figure 2), we observe a deviation that depends on the sign and extent of the velocity error. This deviation is measured as a mismatch between the orientation of the penalty operator (constructed from the estimated dip in the image) and of the local correlation. The penalized correlation are not odd functions and their mean value measures the relative shift between the two images.

The penalty operator $P$ is constructed from the estimated dip field (which depends on the model parameters as well). The dependence of $P$ on the wavefield $u$ can be accounted for by expressing the dip as a function of the image through gradient square tensors (van Vliet and Verbeek, 1995). This relationship is highly nonlinear and a thorough study of $\partial P / \partial u$ and its impact on the adjoint-state calculations is subject of further research.

SYNTHETIC EXAMPLES

We test our methodology on a synthetic example (Figure 3). Figures 3(a) and 3(c) show the initial model and the migrated image, respectively. The inaccuracy of the velocity model can be assessed from the shot-domain CIGs Figure 3(e): the depth of the reflectors changes as a function of experiment, i.e., the semblance principle is not satisfied. After 14 steepest-descent iterations, we recover the model in Figure 3(b). The layers can now be identified, and the image shows better positioning of the interfaces (Figure 3(d)). Figure 3(f) shows alignment of the partial images, which indicates an accurate model. The simplicity of the structures in the model helps the inversion because the definition of the penalty operators at each image point is unequivocal. We can define a dip vector at each image point; the pinch-outs of the sincline ($x = 2$ km and $x = 8$ km) do not influence the computed gradient. For more complicated models with conflicting dips and complicated geologic features (where there is no clear definition of the dip field), a more sophisticated design of the penalty operator is necessary. The eigenvalues and eigenvectors of the gradient square tensors (van Vliet and Verbeek, 1995) can be used to define ellipses that, in turn, may offer a structure oriented criterion for the definition of the penalty operators.

We run a second test using the Marmousi model. We simulate 78 shots with a finite-difference single-scattering modeling code and absorbing boundary conditions on the 4 sides of the model. The spacing between the shots is 0.08 km, and the first source is at $x = 0.96$ km; receivers are placed 0.008 km apart on the surface. The initial model is a heavily smoothed and scaled version of the original Marmousi slowness model. Figures 4(a), 4(c), and 4(e) show the initial model, migrated image, and shot-domain CIGs extracted every 1 km from $x = 0$ km, respectively. Conflicting dips are not yet handled correctly by our methodology because they make the definition of the penalty operators ambiguous. For this reason, we restrict the inversion to the first 1.5 km in depth, where the reflectors are more coherent.

The migrated image (Figure 4(c)) presents crossing interfaces and defocused fault planes; the shot-domain CIGs (Figure 4(e)) show variation of the depth of reflectors as a function of experiment, which indicates model inaccuracy. After 14 steepest-descent iterations, we obtain the results in Figures 4(b), 4(d), and 4(f). The interfaces are moved toward the correct position, and we observe better focusing of the fault planes. The shot-domain CIGs in Figure 4(f) show flatter events, which indicate a more kinematically accurate model.

CONCLUSIONS

We present an approach to waveform tomography based on the semblance principle in the shot-image domain. We define an objective function using appropriately penalized local image correlations that measure the relative shift between two images obtained from nearby experiments; the penalty operator enhances the mismatch between the structural information (the orientation of the reflector) and the shift between two images. The gradient of the objective function with respect to the model is computed with the adjoint-state method, which makes our technique a close relative to the more conventional full waveform inversion. Inversion on a simple synthetic and the Marmousi model shows the ability of our strategy to correct model errors.

ACKNOWLEDGMENTS

We would like to thank eni E&P for the permission to publish this work. The reproducible numerical examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org.
Waveform tomography based on local image correlations

Figure 3: (a) Initial slowness model and (b) slowness model after 14 iterations of waveform tomography; (c) initial image and (d) migrated image with the updated model; (e) initial shot-domain CIGs and (f) shot-domain CIGs with the updated model. Observe the improved positioning of the reflectors after the inversion. The curvature of the CIGs is corrected by the inversion algorithm.

Figure 4: (a) Initial slowness model and (b) slowness model after 14 iterations of waveform tomography; (c) initial image and (d) migrated image with the updated model; (e) initial shot-domain CIGs and (f) shot-domain CIGs with the updated model.
Waveform tomography based on local image correlations

REFERENCES


Anisotropy signature in extended images from reverse-time migration
Paul Sava, Center for Wave Phenomena, Colorado School of Mines
Tariq Alkhalifah, King Abdullah University of Science and Technology

SUMMARY

Reverse-time migration can accurately image complex geologic structures in anisotropic media. Extended images at selected locations in the earth, i.e. at common-image-point gathers (CIPs), carry enough information to characterize the angle-dependent illumination and to provide measurements for migration velocity analysis. Furthermore, inaccurate anisotropy leaves a distinctive signature in CIPs, which can be used to evaluate anisotropy through techniques similar to the ones used in conventional wavefield tomography.

INTRODUCTION

Wave-equation depth migration is powerful and accurate for imaging complex geology, but its potential can only be achieved with high quality models of the earth (Gray et al., 2001). Imaging using reverse-time migration (Baysal et al., 1983; McMechan, 1983), addresses this challenge, although this technique is still fairly computationally intensive, especially in anisotropic media.

Wave-equation imaging of wide-azimuth data (Regone, 2006; Michell et al., 2006; Clarke et al., 2006) is particularly challenging due to the large data volumes and due to the difficulty of interpreting multi-dimensional data. Imaging using extended common-image-point-gathers provides computational savings over alternative methods based on common-image-gathers (Sava and Vasconcelos, 2011). Such efficiency is especially needed in anisotropic media where the cost of imaging is high.

Here we analyze the extended image in anisotropic media, and specifically transversely isotropic (TI) media. These extended images provide image point extensions as a function of space- and time-lags. Embedded in these lag sections is valuable image focusing information, which can be used to analyze the velocity accuracy used in the wavefield extrapolation. As a result, we use such lag sections to characterize the response of the extended images to ignoring anisotropy in RTM. This info, partially decoupling the effects of velocity error from anisotropy error can be used for model building.

THEORY

Conventional wavefield-based imaging consists of two major steps: the wavefield reconstruction and the imaging condition (Berkhout, 1982; Clærbout, 1985). The major driver for the accuracy of this technique are the source and receiver wavefields which depend on space

\[ x = \{x, y, z\} \] and time \( t \). Wavefield reconstruction in tilted TI anisotropic medium, requires numeric solutions to a pseudo-acoustic wave-equation, e.g. the time-domain method of Fletcher et al. (2009) and Fowler et al. (2010), which consists of solving a system of second-order coupled equations:

\[
\frac{\partial^2 p}{\partial t^2} = v_{pz}^2 H_2 [p] + v_{pz}^2 H_1 [q] ,
\]

\[
\frac{\partial^2 q}{\partial t^2} = v_{pz}^2 H_2 [q] + v_{pz}^2 H_1 [q] .
\]

Here, \( p \) and \( q \) are two wavefields depending on space \( x \) and time \( t \), \( H_1 \) and \( H_2 \) are differential operators applied to the quantity in the square brackets Alternative formulations, e.g. Duveneck and Bakker (2011) and Zhang et al. (2011) provide expressions for stable extrapolation in more general cases, including TTI. The velocities \( v_{px}, v_{pz}, \) and \( v_{pq} \) are the vertical, horizontal and “NMO” velocities used to parametrize a generic TTI medium. If we describe the medium using the parameters introduced by Thomsen (2001), the velocities are related to the anisotropy parameters \( \epsilon \) and \( \delta \) by the relations \( v_{pn} = v_{px} \sqrt{1+2\delta} \) and \( v_{pz} = v_{px} \sqrt{1+2\epsilon} \). Thus, the quantities \( H_1 \) and \( H_2 \) depend on the medium parameters, as well as \( \theta \), which describes the angle made by the TI symmetry axis with the vertical and \( \phi \), the azimuth angle of the plane that contains the tilt.

A conventional imaging condition based on the reconstructed wavefields defines the image as the zero lag of the cross-correlation between the source and receiver wavefields (Clærbout, 1985). An extended imaging condition is a generalization of the conventional imaging condition in that it retains in the output image acquisition or illumination parameters. For example, we can generate image extensions by correlation of the wavefields shifted symmetrically in space (Rickett and Sava, 2002; Sava and Fomel, 2005) or in time (Sava and Fomel, 2006). This separation is simply the lag of the cross-correlation between the source and receiver wavefields (Sava and Vasconcelos, 2011):

\[
R(x, \lambda, \tau) = \sum_{\text{shots}} \sum_{t} W_s(x - \lambda, t - \tau) W_r(x + \lambda, t + \tau) .
\]

Here \( R \) represents the migrated image which depends on position \( x \), and the quantities \( \lambda \) and \( \tau \) are cross-correlation lags in space and time, respectively. The source and receiver wavefields used for imaging, \( W_s \) and \( W_r \) are the main wavefields indicating by \( p \) in equations 1-2. The conventional imaging condition represents a special case of equation 3 for \( \lambda = 0 \) and \( \tau = 0 \). Various techniques have been proposed to convert the space- and time-lag extensions into reflection angles (Sava and Fomel, 2003, 2006; Sava and Vlad, 2011;
Anisotropic extended images

Sava and Alkhalifah, 2012), thus facilitating amplitude and velocity analysis in complex geologic structures.

Figure 1: Image point defocusing corresponding to isotropic poststack migration with the correct velocity of 2 km/s in an anisotropic medium with (a) \( \eta = 0.1 \) (grey curve), (b) \( \eta = 0.2 \) (black curve), (c) \( \eta = -0.1 \) (grey curve), and (d) \( \eta = -0.2 \) (black curve).

Alkhalifah and Fomel (2011) show using anisotropic parameter continuation that the residual anisotropic response for post stack imaging in homogeneous media, and specifically residuals corresponding to the anisotropy parameter \( \eta \), has a “V”-like shape. This response stems from the fourth-order nature of the \( \eta \) influence on travel-time as a function of offset. A residual reconstruction of such travel-time maps the fourth order influence to a “V” shaped (rather than a residual hyperbolic) moveout as shown in Figure 1. The slope of the “V” depends on the value of \( \eta \), where negative \( \eta \) provides a flipped “V” signature. We will later see a similar response for the prestack case even in complex media.

In this paper, we show that such a behavior also characterizes the extended images for pre-stack data in complex anisotropic media. This is because the major wavefield complexities are compensated during the wavefield reconstruction in anisotropic media. All that remains are the effects of anisotropy which, for smooth models, resembles the established behavior of poststack depth images. This property could be used to separate the effects of anisotropy from those of velocity, with applications to anisotropic earth model building.

**EXAMPLES**

The model for our first example, Figure 2(c), consists of a horizontal reflector in a 3D constant velocity model. Figures 2(a)-2(b) depict one snapshot of the wavefields for a source located on the surface in VTI and TTI models, respectively. The models are characterized by the parameters \( \epsilon = 0.30, \delta = -0.10 \) and, for the TTI model, \( \theta_a = 35^\circ \) and \( \phi_a = 45^\circ \). Figures 3(a)-3(b) depict extended CIPs at \( \{x, y, z\} = \{4, 4, 1\} \) km for many sources distributed uniformly on the surface and correct earth models. All CIPs show focusing at zero lags in space and time, thus indicating wavefield reconstruction with a correct earth model. However, the CIPs constructed with isotropic migration, Figures 3(c) and 3(d), are not focused, thus indicating that the CIPs are sensitive to anisotropy inaccuracy.

Figure 3: Extended CIPs at \( \{x, y, z\} = \{4, 4, 1\} \) km for many sources distributed uniformly on the surface. The images are obtained by (a)-(b) isotropic migration, and (c)-(d) anisotropic migration of the VTI data (left) and TTI data (right).

The next examples illustrate the behavior of the extended images as a function of anisotropy inaccuracy in constant and laterally heterogeneous models, respectively. The models are shown in Figures 4(a) and 6(a). The data, exemplified in Figures 4(b) and 6(b), are constructed with anisotropy characterized by parameter \( \eta = 0.25 \). Figures 7(a)-7(c) show the extended images for different shot locations obtained by reverse-time migration with the correct salt model. The figures indicate a dependency of the CIPs with the angle of illumination.
Anisotropic extended images

Figure 2: 3D model used to illustrate wide-azimuth extended images in anisotropic media. Panels (a) and (b) show the wavefields corresponding to a single source at \( \{x, y\} = \{4, 4\} \) km for VTI and TTI media, respectively. The model (c) consists of a horizontal reflector in constant velocity.

which can be exploited for angle decomposition. The summation for all shots leads to the CIP shown in Figure 7(e) which shows focusing of the image at zero space and time lags. This is due to the fact that all shots leave an imprint on the CIPs at zero space and time lags, but at different slopes which depend on the illumination angle, Figures 7(a)-7(c).

For incorrect \( \eta \), the resulting response has an overall “V” shape similar to that seen for residual poststack migration. The slope of the “V” flanks, as illustrated in Figures 7(d)-7(f), depend on \( \eta \). For \( \eta \) close to the correct value, the “V” energy tends to approach the apex, and for correct \( \eta \) it focuses at its apex, Figure 7(e). This general behavior is consistent regardless of the complexity of the medium, especially with respect to the portion of the signature that is near the apex. For comparison, Figures 5(d)-5(e) show extended gathers corresponding to the same values of \( \eta \) as in Figures 7(d)-7(e). Despite the fact that the background velocity model is significantly different, constant vs. salt, the CIPs are comparable, thus indicating that they are mainly influenced by the anisotropy parameters. Of course, for a varying \( \eta \) such signatures reflect an effective value under the effective medium theory. This property can be exploited for anisotropic migration velocity analysis.

CONCLUSIONS

The distinct anisotropic signatures, needed for parameter estimation, tend to get lost in the mist of the complexity of the velocity in the medium. Extended images preserve such signatures, which allows for direct anisotropy analysis. The slope of the residual moveout in extended CIPs mainly depends on \( \eta \), which simplifies parameter estimation in anisotropic media. A comparison of extended images between simple and complex models supports our assertion.

ACKNOWLEDGMENTS

This work was supported by the sponsors of the Center for Wave Phenomena at Colorado School of Mines. The authors also thank KAUST for financial support. The reproducible numeric examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org. This research was supported in part by the Golden Energy Computing Organization at the Colorado School of Mines using resources acquired with financial assistance from the National Science Foundation and the National Renewable Energy Laboratory.
Anisotropic extended images

Figure 4: (a) 2D constant model used to study the dependence of the extended CIPs with anisotropy, and (b) one shot gather for a source located at \( \{x, z\} = \{3.4, 0\} \) km. The earth model is anisotropic and characterized by parameter \( \eta = 0.25 \).

Figure 5: 2D extended CIPs at \( \{x, z\} = \{4, 0\} \) km for the constant model, Figure 6(a). The extended images (a)-(c) correspond to individual shots at surface coordinates \( x = \{3.2, 4.0, 4.8\} \) km and are obtained from wavefields reconstructed with the correct \( \eta = 0.25 \). The extended images (d)-(f) correspond to anisotropy characterized by parameter \( \eta = \{0.15, 0.25, 0.35\} \).

Figure 6: (a) 2D salt model used to study the dependence of the extended CIPs with anisotropy, and (b) one shot gather for a source located at \( \{x, z\} = \{3.4, 0\} \) km. The earth model is anisotropic and characterized by parameter \( \eta = 0.25 \).

Figure 7: 2D extended CIPs at \( \{x, z\} = \{4, 0\} \) km for the salt model, Figure 6(a). The extended images (a)-(c) correspond to individual shots at surface coordinates \( x = \{3.2, 4.0, 4.8\} \) km and are obtained from wavefields reconstructed with the correct \( \eta = 0.25 \). The extended images (d)-(f) correspond to anisotropy characterized by parameter \( \eta = \{0.15, 0.25, 0.35\} \).
REFERENCES


Attenuation analysis for heterogeneous transversely isotropic media

Bharath Shekar & Ilya Tsvankin, Center for Wave Phenomena, Colorado School of Mines, Golden CO 80401

SUMMARY

Attenuation coefficients obtained from seismic data may provide sensitive attributes for reservoir characterization and increase the robustness of AVO (amplitude variation with offset) analysis. Here, we present an algorithm for ray tracing in attenuative anisotropic media based on the methodology of Červený and Pšenčík. Both kinematic and dynamic ray tracing are carried out in an elastic reference medium, with the attenuation terms incorporated as perturbations along the ray. Numerical examples for smoothly varying transversely isotropic (TI) media demonstrate the accuracy of the method. The dynamic ray-tracing technique provides an efficient forward-modeling operator that can be used in the inversion for anisotropic attenuation. We outline a methodology for estimating attenuation coefficients in 2D heterogeneous transversely isotropic (TI) media. A necessary prerequisite for accurate attenuation analysis is reconstruction of the heterogeneous velocity field. If the subsurface can be approximated by a piecewise-factorized VTI (TI with a vertical axis of symmetry) medium, the velocity model can be built using the migration velocity analysis algorithm proposed by Sarkar and Tsvankin. Attenuation coefficients in each factorized block can then be estimated using gradient-based inversion that employs dynamic ray tracing.

INTRODUCTION

Seismic attenuation is sensitive to lithology and physical properties of subsurface rocks (Johnston and Toksöz, 1981), so it can provide valuable attributes for reservoir characterization (Lynn, 2004). Compensation for attenuation also helps improve the quality of images obtained by depth migration (e.g., Xin et al., 2008). However, estimation of attenuation from surface seismic data is a challenging problem, and stable attenuation analysis typically requires the subsurface structure to be laterally homogeneous. For example, Brzostowski and McMechan (1992) propose a tomographic algorithm that relies on the spectral-ratio method to compute isotropic attenuation coefficients from a field data set. They first estimate the 3D velocity structure and then use the simultaneous iterative reconstruction technique (SIRT) (Menke, 1984) to invert for attenuation.

Zhu et al. (2007) extend the spectral-ratio method to anisotropic media and estimate angle-dependent attenuation coefficients from physical-modeling data. Behura and Tsvankin (2009a) introduce a layer-stripping algorithm that combines the velocity-independent layer stripping (VILS) method of Dewangan and Tsvankin (2006) with the spectral-ratio technique to estimate interval attenuation coefficients of pure (PP or SS) reflected waves. The overburden is assumed to be laterally homogeneous (but possibly vertically heterogeneous) with a horizontal symmetry plane, while the target layer may be arbitrarily anisotropic and heterogeneous. Shekar and Tsvankin (2011) extend the attenuation layer-stripping method of Behura and Tsvankin (2009a) to mode-converted PS-waves and estimate the interval S-wave attenuation coefficient. The attenuation layer-stripping method, however, is restricted to models with a laterally homogeneous overburden. Here, we build the foundation for extending attenuation analysis to transversely isotropic models with spatially varying velocity and attenuation functions.

MODELING OF SEISMIC WAVE PROPAGATION IN ATTENUATIVE MEDIA

We compute asymptotic Green’s functions in attenuative anisotropic media by dynamic ray tracing (Červený, 2001). So-called “complex” ray theory treats ray trajectories and parameters computed along the ray as complex quantities (Thomson, 1997; Hanyga and Seredyńśka, 2000), but its numerical implementation is not straightforward. Alternatively, ray tracing in attenuative media can be performed using perturbation methods (Gajewski and Pšenčík, 1992; Červený and Pšenčík, 2009), which involve computation of rays in a reference elastic medium with the influence of attenuation modeled as a perturbation along the ray. Next, we briefly review the perturbation-based calculation of the asymptotic Green’s function in attenuative, anisotropic, heterogeneous media with smooth spatial variations of the stiffness tensor.

Dynamic ray tracing in viscoelastic media

The density-normalized stiffness tensor for attenuative media is complex-valued:

\[ \bar{a}_{ijkl} = a_{ijkl}^R + ia_{ijkl}^I, \]

where \( a_{ijkl}^R \) and \( a_{ijkl}^I \) are the real and imaginary parts of \( \bar{a} \). The ray-theoretical frequency-domain displacement of P-waves propagating in heterogeneous, attenuative, anisotropic media has the form:

\[ \mathbf{u}(x_0, \omega) = S(\omega) \mathbf{A}(x_0) \exp[-i \omega (t - \text{Re} \tau(x_0))] \times \exp[-\omega \text{Im} \tau(x_0)], \]

where \( \mathbf{u}(x_0, \omega) \) is the displacement vector, \( S(\omega) \) is the source spectrum, \( \mathbf{A}(x_0) \) (assumed to be frequency-independent) incorporates the geometrical spreading and transmission coefficients along the raypath, the polarization vector, and the source/receiver directivity. The real and imaginary parts of the traveltime \( \tau(x_0) \) contribute to the phase and amplitude functions along the ray, respectively. The imaginary part of the traveltime is often called the “dissipation factor” and denoted by \( t^* \) (Gajewski and Pšenčík, 1992):

\[ t^*(x_0) = \text{Im} \tau(x_0). \]

The factor responsible for the exponential amplitude decay is called the dissipation filter \( D(\omega) \):

\[ D(\omega) = \exp[-\omega t^*(x_0)]. \]

Červený and Pšenčík (2009) show that for weakly attenuative media it is sufficient to perform dynamic ray tracing in the
Attenuation analysis for TI media

reference elastic medium (Červený, 2001). Then, the complex traveltime can be computed as a perturbation of the real-valued traveltime along the ray. Hence, the density-normalized stiffnesses $\alpha_{ijkl}$ in equation 1 correspond to the reference medium, and the term $\alpha_{ijkl}$ is the perturbation that makes the medium viscoelastic. Červený and Pšenčík (2009) define the “perturbation Hamiltonian” $\mathcal{H}(x_a, p_b, \alpha)$ as follows:

$$\mathcal{H}(x_a, p_b, \alpha) = \mathcal{H}^0(x_a, p_b) + \alpha \cdot \triangle \mathcal{H}(x_a, p_b), \quad (5)$$

with

$$\triangle \mathcal{H}(x_a, p_b) = \mathcal{H}(x_a, p_b) - \mathcal{H}^0(x_a, p_b), \quad (6)$$

where $\mathcal{H}^0$ and $\mathcal{H}$ correspond to the (elastic) reference and (viscoelastic) perturbed medium, respectively, and $\alpha$ is the perturbation parameter.

The traveltime $\tau(x_a, \alpha)$ and its spatial derivatives $\partial \tau(x_a, \alpha)/\partial x_i$, which correspond to the perturbation Hamiltonian defined in equation 5, can be expanded in the perturbation parameter $\alpha$:

$$\tau(x_a, \alpha) \approx \tau(x_a) + \alpha \cdot \frac{\partial \tau(x_a, \alpha)}{\partial \alpha} + \frac{1}{2} \alpha^2 \cdot \frac{\partial^2 \tau(x_a, \alpha)}{\partial \alpha^2}, \quad (7)$$

$$\frac{\partial \tau(x_a, \alpha)}{\partial x_i} \approx \frac{\partial \tau(x_a)}{\partial x_i} + \alpha \cdot \frac{\partial^2 \tau(x_a, \alpha)}{\partial x_i \partial \alpha}, \quad (8)$$

where $\tau(x_a)$ and $\partial \tau(x_a)/\partial x_i$ are computed in the reference medium. The partial derivatives in equations 7 and 8 are computed as quadratures along the ray using the Hamiltonians defined in equations 5–6, and are evaluated for $\alpha = 0$.

Equations 7 and 8 can be separated into the real and imaginary part:

$$\tau(x_a, \alpha) = \text{Re} \{ \tau(x_a, \alpha) \} + i \cdot \text{Im} \{ \tau(x_a, \alpha) \}, \quad (9)$$

$$\frac{\partial \tau(x_a, \alpha)}{\partial x_i} = \text{Re} \left\{ \frac{\partial \tau(x_a, \alpha)}{\partial x_i} \right\} + i \cdot \text{Im} \left\{ \frac{\partial \tau(x_a, \alpha)}{\partial x_i} \right\}. \quad (10)$$

The dissipation factor and its spatial gradient in the perturbed attenuative medium can be found by substituting $\alpha = 1$ into equations 7 and 8 and using equations 3, 9, and 10:

$$t^*(x_a) = \text{Im} \{ \tau(x_a, \alpha) \} \approx \text{Im} \{ \tau(x_a, 1) \} + \frac{1}{2} \cdot \tau_{aa}(x_a), \quad (11)$$

$$t^*_a \left( \frac{\partial^*}{\partial x_i} \right) = \text{Im} \{ \tau_{a, i}(x_a) \} \approx \text{Im} \{ \tau_{a, i}(1) \}, \quad (12)$$

where $\tau_{a, i} = \partial \tau_a / \partial x_i$, $\tau_{aa} = \partial^2 \tau_a / \partial \alpha^2$, and $\tau_{a, ai} = \partial^2 \tau_a / \partial x_i \partial \alpha$.

The phase attenuation coefficient $\Delta$ is defined as

$$\Delta = k^p / k^R, \quad (13)$$

where $k^R$ and $k^I$ are the real and imaginary parts of the wave vector, respectively. The local group attenuation coefficient at any spatial location $x_a$ is given by

$$\Delta(x_a) = -\text{Im} \{ k^R(x_a) \} \approx -\frac{1}{2Q(x_a)}, \quad (14)$$

where $Q(x_a)$ is defined as the local quality factor. Červený and Pšenčík (2009) show that equation 14 produces the phase attenuation coefficient computed in the phase direction corresponding to the ray (or to the group angle), and that it is not influenced by the inhomogeneity angle (the angle between the real and imaginary parts of the wave vector). This result agrees with the general perturbation analysis of the influence of the inhomogeneity angle presented by Behura and Tsvankin (2009b).

The exact phase attenuation coefficient in TI media can be found by solving the complex Christoffel equation (Červený and Pšenčík, 2005; Zhu and Tsvankin, 2006). Zhu and Tsvankin (2006) also present the following linearized approximation for the P-wave phase attenuation coefficient obtained under the assumptions of weak attenuation and weak velocity and attenuation anisotropy:

$$\omega_p(\theta) = \frac{1}{2Q_{p0}} = \omega_{p0}(1 + \delta_{0} \sin^2 \theta \cos^2 \psi + \epsilon_{0} \sin^4 \theta), \quad (15)$$

where $Q_{p0}$ is the P-wave vertical quality factor $Q_{p0}$ and vertical quality factor $Q_{p0}$ (equations 16 and 17).

<table>
<thead>
<tr>
<th>Model</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{p0}^{(0)}$ (km/s)</td>
<td>3.00</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>$V_{s0}^{(0)}$ (km/s)</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td>$\delta_2$ (km)</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>$k_x$ (km$^{-1}$)</td>
<td>0</td>
<td>0</td>
<td>0.20</td>
</tr>
<tr>
<td>$k_z$ (km$^{-1}$)</td>
<td>0</td>
<td>0</td>
<td>0.80</td>
</tr>
<tr>
<td>$Q_{p0}^{(0)}$</td>
<td>20</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>$Q_{s0}^{(0)}$</td>
<td>20</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>$\epsilon_2$</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>$\delta_x$</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>$f_x$ (km$^{-1}$)</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>$f_z$ (km$^{-1}$)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: TI models used to test the algorithm. Models 1 and 2 are homogeneous, while model 3 has linearly varying P-wave vertical velocity $V_{p0}$ and vertical quality factor $Q_{p0}$ (equations 16 and 17).

Numerical examples

To test the algorithm described above, we consider an anisotropic halfspace with smoothly varying velocity and attenuation parameters. The velocity parameters $\epsilon$ and $\delta$ and attenuation-anisotropy parameters $\epsilon_2$ and $\delta_2$ are constant, while the P-wave vertical velocity $V_{p0}$ and vertical quality factor $Q_{p0}$ are either constant or vary as linear functions of the spatial coordinates:

$$V_{p0}(x,z) = V_{p0}^{(0)}(1 + k_x x + k_z z), \quad (16)$$

$$Q_{p0}(x,z) = Q_{p0}^{(0)}(1 + j_x x + j_z z), \quad (17)$$

...
Attenuation analysis for TI media

Three sets of the velocity and attenuation parameters used in the tests are listed in Table 1. Models 1 and 2 are homogeneous, while model 3 has linearly varying parameters $V_{P0}$ and $Q_{P0}$ defined by equations 16 and 17. There is a good agreement between the ray-traced and exact quality factors, even for the strongly attenuative model 2 (Figures 1a and 1b). While the linearized quality factor computed from equation 15 does not differ significantly from the exact quality factor for model 1, the deviation is more pronounced for model 2. Hence, the linearized approximation can be used in solving the inverse problem for moderately attenuative models. The attenuation-induced perturbation of the real-valued traveltime (equation 9) is significant (on the order of milliseconds) only for strongly attenuative media (model 2, Figure 1c).

Figure 2a displays ray trajectories for model 3, in which $V_{P0}$ and $Q_{P0}$ vary linearly with $x$ and $z$ (Table 1). The perturbation of the real-valued traveltime due to attenuation is almost negligible. Figures 2b displays the P-wave quality factor (equation 18) as a function of the traveltime along the ray. The values of $Q$ computed by ray tracing practically coincide with those obtained from the exact and linearized equations. The spatial derivatives of the dissipation factor (Figure 2c) do not vanish, as expected for this heterogeneous model.

**Figure 1**: Ray-tracing results for models 1 and 2 from Table 1. The quality factor $Q_P$ as a function of phase angle for (a) model 1 and (b) model 2. The exact coefficient is plotted in gray, the red stars denote the value obtained from equation 14, and the black line is the linearized quality factor from equation 15. (c) Perturbation of the real-valued traveltine as a function of the unperturbed traveltine computed along the ray with a take-off phase angle of 18° in model 2.

**Figure 2a** displays ray trajectories for model 3, in which $V_{P0}$ and $Q_{P0}$ vary linearly with $x$ and $z$ (Table 1). The perturbation of the real-valued traveltime due to attenuation is almost negligible. Figures 2b displays the P-wave quality factor (equation 18) as a function of the traveltime along the ray. The values of $Q$ computed by ray tracing practically coincide with those obtained from the exact and linearized equations. The spatial derivatives of the dissipation factor (Figure 2c) do not vanish, as expected for this heterogeneous model.

**Figure 2**: Ray trajectories for model 3, in which $V_{P0}$ and $Q_{P0}$ vary linearly with $x$ and $z$ (Table 1). The perturbation of the real-valued traveltime due to attenuation is almost negligible. Figures 2b displays the P-wave quality factor (equation 18) as a function of the traveltime along the ray. The values of $Q$ computed by ray tracing practically coincide with those obtained from the exact and linearized equations. The spatial derivatives of the dissipation factor (Figure 2c) do not vanish, as expected for this heterogeneous model.

**Figure 2a** displays ray trajectories for model 3, in which $V_{P0}$ and $Q_{P0}$ vary linearly with $x$ and $z$ (Table 1). The perturbation of the real-valued traveltime due to attenuation is almost negligible. Figures 2b displays the P-wave quality factor (equation 18) as a function of the traveltime along the ray. The values of $Q$ computed by ray tracing practically coincide with those obtained from the exact and linearized equations. The spatial derivatives of the dissipation factor (Figure 2c) do not vanish, as expected for this heterogeneous model.

**BASIC ELEMENTS OF INVERSION METHODOLOGY**

Due to numerous complications involved in attenuation estimation, robust attenuation analysis typically requires the subsurface structure to be relatively simple. If the overburden is laterally homogeneous and has a horizontal symmetry plane, anisotropic attenuation coefficients of P- and S-waves can be estimated without knowledge of the velocity field using the layer-stripping method of Behura and Tsvankin (2009a) and Shekar and Tsvankin (2011). That method yields attenuation coefficients corresponding to a given source-receiver offset or group angle. To invert for attenuation-anisotropy parameters, it is necessary to estimate the corresponding phase angle, which requires knowledge of an approximate velocity model. However, for more complicated, heterogeneous subsurface models, attenuation analysis has to be preceded by accurate velocity estimation.

Sarkar and Tsvankin (2004) introduce a 2D method to reconstruct the P-wave velocity field in VTI media via migration velocity analysis (MVA). They divide the subsurface into factorized VTI blocks, in which the parameters $\varepsilon$ and $\delta$ are constant, while the P-wave vertical velocity varies linearly with depth and lateral position (equation 16). The methodology of Sarkar and Tsvankin (2004) helps resolve the vertical and lateral velocity gradients along with $\varepsilon$ and $\delta$, provided that the velocity $V_{P0}$ is known at a single point in each block. To estimate the block-based spatially varying attenuation coefficient, we consider factorized VTI models for both velocity and attenuation (Figure 3). Within each factorized block, the velocity-anisotropy parameters as well as the attenuation-anisotropy parameters are constant, while the P-wave vertical velocity $V_{P0}^{(0)}$ and vertical quality factor $Q_{P0}^{(0)}$ vary linearly with...
Figure 2: Ray-tracing results for model 3 from Table 1. (a) Ray trajectories. (b) The quality factor $Q_P$ (same display as in Figure 1). (c) The spatial derivative of the dissipation factor with respect to $x$ (black line) and $z$ (gray line) computed along the ray with a take-off phase angle of 18°.

Figure 3: 2D model for velocity and attenuation estimation. The subsurface is divided into factorized blocks with VTI symmetry for both velocity and attenuation.

depth and lateral position (equations 20 and 21). The velocity field can then be built using the algorithm of Sarkar and Tsvankin (2004). The layer-stripping method of Behura and Tsvankin (2009a) is used to estimate the initial attenuation coefficients in each layer or block. The attenuation estimates are refined by employing a gradient-based inversion algorithm. Since the ray-tracing methodology described above produces the Fréchet derivatives of the complex-valued traveltime along the ray, it can be efficiently used for gradient computation. We will present synthetic examples illustrating the accuracy and limitations of the method.

CONCLUSIONS

We implemented a dynamic ray-tracing algorithm that relies on ray perturbation theory to model P-wave amplitudes for 2D heterogeneous attenuative TI media. The attenuation-related terms are computed along rays traced in an elastic background medium by perturbing the corresponding reference quantities. The algorithm produces accurate P-wave traveltimes and attenuation coefficients for smoothly varying TI media, even in the presence of extremely strong attenuation. The linearized expressions for the attenuation coefficient provide acceptable accuracy for moderately attenuative media. The examples also show that the influence of attenuation on traveltime is significant only if the quality factor is uncommonly small. We also proposed a methodology of attenuation analysis based on dividing the model into factorized VTI blocks. The attenuation coefficient in each factorized block is represented as a linear function of the spatial coordinates and obtained by gradient-based inversion. The perturbation ray-tracing method substantially reduces the cost of modeling and provides essential quantities for the inversion operator.

ACKNOWLEDGMENTS

We are grateful to the members of the A(nisotropy)-Team of the Center for Wave Phenomena (CWP), Colorado School of Mines, for fruitful discussions. Bharath Shekar wishes to acknowledge John Stockwell (CWP) and Sribharath Kainkaryam (Purdue) for their help with ray-theory concepts. Support for this work was provided by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP and by the Research Partnership to Secure Energy for America (RPSEA).
Attenuation analysis for TI media

REFERENCES

Behura, J., and I. Tsvankin, 2009a, Estimation of interval anisotropic attenuation from reflection data: Geophysics, 74, no. 6, A69–A74.
Tsvankin, I., and V. Grechka, 2011, Seismology of azimuthally anisotropic media and seismic fracture characterization: SEG.
Time-lapse image-domain velocity analysis using adjoint-state methods
Jeffrey Shragge*, CPGCO2, The University of Western Australia,
Tongning Yang and Paul Sava, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

Adjoint-state methods (ASMs) have proven successful for calculating the gradients of the functionals commonly found in geophysical inverse problems. The 3D image-domain formulation of the seismic velocity estimation problem uses imperfections in 3D migrated images to form an objective function, which is minimized using a combined ASM plus line-search approach. While image-domain methods are less sensitive than their data-domain counterparts because they are based largely on wavefield kinematics and not directly matching amplitudes, they are more robust to poorer starting models, which makes them attractive for the early stages of seismic velocity estimation. For time-lapse (4D) seismic scenarios, we show that the 3D ASM approach can be be extended to multiple datasets to offer high-quality estimates of production- and/or injection-induced subsurface change. We discuss two different penalty operators that lead to what we term absolute and relative inversion strategies. The absolute approach straightforwardly uses the difference of two independent 3D inversions to estimate a 4D slowness perturbation. The relative approach directly incorporates the baseline image into the penalty function to highlight where the baseline and monitor images are different and to mask where they are similar - even if reflectors are imperfectly focused. Both these techniques yield good 4D slowness estimates for synthetic data; however, we assert that the relative approach is more robust and preferable to the absolute strategy in the presence of 4D field noise because it represents a less-demanding inversion goal.

INTRODUCTION

Adjoint-state methods (ASMs) have been used with success for a number of years in seismic exploration as an effective approach for calculating the gradient of a functional [see the overviews of Plessix (2006) and Symes (2009)]. While the majority of studies have focused on data-domain applications - in particular full waveform inversion (FWI) (Tarantola, 1984; Pratt, 1999) - a number of authors have explored the complementary image-domain tomography strategy (Girard and Vasconcelos, 2010; Yang and Sava, 2011). This inversion approach, based largely on wavefield kinematics, uses an objective function (OF) derived from observed imperfections in migrated images (i.e., poorly focused subsurface-offset panels). Because image-domain approaches do not formally match field data amplitudes, they afford lower resolution than data-domain methods; however, they are less sensitive to the manifold factors that can affect amplitude (e.g., illumination, anisotropy, etc). While this is usually considered a negative trait one important corollary is that kinematically oriented image-domain ASMs are usually more robust than data-domain approaches because they satisfy less demanding inversion criteria. This leads to an increased likelihood of converging toward the correct - though more bandlimited - inversion result.

The goal of image-domain ASM tomography is to invert for model slowness perturbations, \( \Delta s_1 \equiv s_1 - s_0 \), that represent the difference between the true and background models, \( s_1 \) and \( s_0 \), respectively. The key steps of this non-linear inversion approach are similar to data-domain implementations: i) compute the state variables represented by the forward modeled seismic wavefields; ii) calculate the adjoint sources based on the OF and state variables; iii) compute the adjoint state variables represented by the backpropagated adjoint sources; and iv) calculate the gradient estimate by combining the forward and adjoint state variables through an imaging condition. In order to speed up convergence in ASM inversion problems, one normally incorporates a judicious penalty operator that upweights energy located away from zero subsurface offset while excluding that already focused about zero offset. This is often accomplished through use of a differential semblance operator (DSO) (Shen et al., 2005) that, by definition, cancels out a perfectly focused image. ASM-derived model perturbations are useful when used outright for generating a final migration image, but they could also serve as input to a further higher-resolution data-domain velocity inversion analysis (e.g., FWI).

The 4D ASM velocity estimation problem shares many similarities with - and can be viewed as an extension of - the corresponding 3D ASM inversion problem. One key difference is that there are now multiple data sets to work with (i.e., baseline and monitor) as well as multiple slowness perturbations to recover (i.e., the baseline \( \Delta s_1 \), monitor \( \Delta s_2 \), and time-lapse \( \Delta s_{TL} \) differences). Unlike the 3D problem where, by definition, one seeks the absolute slowness perturbation that optimally focuses subsurface offset gatherers at zero offset, we discuss how the 4D scenario can be solved by implementing one of two strategies. First, one can set up two separate tomographic inversions to independently estimate \( \Delta s_1 \) and \( \Delta s_2 \), and then take their difference to form an absolute time-lapse slowness perturbation. Second, one may directly compute a relative estimate of the time-lapse slowness difference by appropriately coupling the baseline and monitor datasets in the inversion. We accomplish this through a strategy where baseline image information is introduced into a new penalty function that we apply when inverting the monitor dataset for the \( \Delta s_2 \) estimate.

In this abstract we examine the relative 4D approach that uses an image-derived penalty function to upweight energy in the monitor image not spatially coincident with that in the baseline image while downweighting that which is co-located - even if not optimally focused. Thus, this represents a relative change between the monitor and baseline image that is analogous to the illumination compensation discussed in Yang et al. (2012). We begin by briefly reviewing the 3D ASM theory and detailing the two 4D extensions. We then present the results of two
inversion experiments that examine the similarities and differences between the absolute and relative 4D ASM approaches.

**ADJOINT-STATE METHOD THEORY**

Given a one-way frequency-domain wave operator, \( \mathcal{L} \), and its adjoint, \( \mathcal{L}^* \), we write the 3D forward modeling problem as:

\[
\begin{bmatrix}
\mathcal{L}(x, \omega, s_0) & 0 \\
0 & \mathcal{L}^*(x, \omega, s_0)
\end{bmatrix}
\begin{bmatrix}
u_1(x, \omega, s_0) \\
u_2(x, \omega, s_0)
\end{bmatrix}
= \begin{bmatrix}f_1(x, \omega, s_0) \\
f_2(x, \omega, s_0)
\end{bmatrix}
\]

where \( x \) are the spatial coordinates; \( \omega \) is frequency; \( s_0 \) is the background slowness field (formally the square root of model parameters \( m \)); \( \upsilon_1 \) and \( \upsilon_2 \) are the computed source and receiver wavefields corresponding to source index \( e \); and \( f_1 \) and \( f_2 \) are the source and recorded data. State variables, \( \upsilon_1 \) and \( \upsilon_2 \), both solutions to the (acoustic) wave equation, are used to formulate an objective function (OF) \( H \) that, for the imaging-domain ASM tomography problem, is based on minimizing image inconsistencies (i.e., poorly focused subsurface-offset gather) caused by an unknown slowness perturbation \( \Delta s_1 \)

\[
H = \frac{1}{2} \|K_\ell(x)P(x, \lambda)r(x, \lambda)\|_{\lambda}^2,
\]

where \( r(x, \lambda) \) is an extended image volume (Sava and Vasconcelos, 2010) formed by cross-correlating the state variables in a direction denoted by \( \lambda \) (herein taken to be horizontal)

\[
r(x, \lambda) = \sum_{e, \omega} \upsilon_1(x, \omega - \lambda, \omega) \upsilon_2(x, \omega + \lambda, \omega).
\]

Mask operator \( K_\ell \) is used to restrict the OF evaluation to certain image locations, while penalty operator \( P(x, \lambda) \) highlights the defocusing in \( [x, \lambda] \) hypercube within extended image \( r(x, \lambda) \). In the absence of any other information the DSO penalty function, defined by \( P(\lambda) = |\lambda| \) and shown in Figure 1(a), has been shown to produce impressive ASM inversion results at a reasonable rate of convergence.

\[
\begin{pmatrix}
\mathcal{L}^*(x, \omega, s_0) & 0 \\
0 & \mathcal{L}(x, \omega, s_0)
\end{pmatrix}
\begin{bmatrix}
u_1(x, \omega, s_0) \\
u_2(x, \omega, s_0)
\end{bmatrix}
= \begin{bmatrix}f_1(x, \omega, s_0) \\
f_2(x, \omega, s_0)
\end{bmatrix}
\]

Having defined a penalty function we can specify the adjoint sources, \( g_r \) and \( g_s \), used to form the adjoint state variables as the derivatives of \( H \) (equation 2) with respect to state variables \( \upsilon_1 \) and \( \upsilon_2 \):

\[
\begin{aligned}
g_r(\mathbf{e}, \mathbf{x}, \omega) & = \frac{\partial H}{\partial \upsilon_1(\mathbf{e}, \mathbf{x}, \omega)} \\
g_s(\mathbf{e}, \mathbf{x}, \omega) & = \frac{\partial H}{\partial \upsilon_2(\mathbf{e}, \mathbf{x}, \omega)}
\end{aligned}
\]

\[
\begin{aligned}
\left[ g_r(\mathbf{e}, \mathbf{x}, \omega) \right] & = \left[ \|K^\dagger(\mathbf{x})\|_2^2 \frac{\sum_{\lambda} |P(x, \lambda)|^2 r(x, \lambda) u_1(\mathbf{e}, \mathbf{x} + \lambda, \omega)}{\sum_{\lambda} |P(x, \lambda)|^2 r(x, \lambda) u_2(\mathbf{e}, \mathbf{x} - \lambda, \omega)} \\
\left[ g_s(\mathbf{e}, \mathbf{x}, \omega) \right] & = \left[ \|K^\dagger(\mathbf{x})\|_2^2 \frac{\sum_{\lambda} |P(x, \lambda)|^2 r(x, \lambda) u_2(\mathbf{e}, \mathbf{x} + \lambda, \omega)}{\sum_{\lambda} |P(x, \lambda)|^2 r(x, \lambda) u_1(\mathbf{e}, \mathbf{x} - \lambda, \omega)} \right].
\end{aligned}
\]

The adjoint-state variables, \( a_r \) and \( a_s \), are the wavefields obtained by backward and forward modeling using the corresponding adjoint sources

\[
\begin{bmatrix}
\mathcal{L}^*(x, \omega, s_0) & 0 \\
0 & \mathcal{L}(x, \omega, s_0)
\end{bmatrix}
\begin{bmatrix}
a_1(x, \omega, s_0) \\
a_2(x, \omega, s_0)
\end{bmatrix}
= \begin{bmatrix}g_r(x, \omega, s_0) \\
g_s(x, \omega, s_0)
\end{bmatrix}
\]

The gradient estimate is then formed by correlating the state and adjoint-state variables,

\[
\frac{\partial H}{\partial m} = \sum_{e, \omega, s_0} \frac{\partial \mathcal{L}^*}{\partial m} \begin{bmatrix}a_1(x, \omega, s_0) a_2(x, \omega, s_0) + u_1(x, \omega, s_0) a_2(x, \omega, s_0) \end{bmatrix}
\]

\[
\Delta s_1^{(i)} = \frac{\partial H}{\partial m}.
\]

To find the slowness perturbation estimate, \( \Delta s_1^{(i)} \), we use a gradient line-search that minimizes \( H \) at the current \( n \)th iteration, which e then use to update model parameters for the \( n+1 \)th iteration (i.e., \( s_1^{(n+1)} = s_0 + \sum_{i=1}^{n} \Delta s_1^{(i)} \)). The inversion procedure continues until the convergence criterion is met at the \( N \)th iteration and the optimal slowness perturbation estimate \( \Delta s_1 = \sum_{n=1}^{N} \Delta s_1^{(n)} \) found.

As discussed above there are two different strategies to the 4D ASM velocity inversion problem: **independent** and **relative**. In the independent approach one performs separate inversions to obtain independent 3D ASM tomography estimates of the baseline and monitor slowness perturbations from the common \( s_0 \) model (i.e., \( \Delta s_1 \) and \( \Delta s_2 \)) . The time-lapse estimate is assumed to be their difference: \( \Delta s_{TL} = \Delta s_2 - \Delta s_1 \). Judicious 4D practice would suggest that one should use the same model weighting function \( K_\ell \) and penalty operator \( P \) in the two independent inversion problems; to do otherwise would lead to differently weighted back-projections that could cause an erroneous \( \Delta s_{TL} \) estimate.

The relative 4D approach recognizes that, in fact, there is prior information in the baseline image that can be incorporated directly into a penalty function for monitor inversions. Figure 1(b) presents an example of an image-derived weight that we apply as a penalty function. We use a smoothing operator on the baseline image envelope and then apply a 2D AGC filter to upweight weaker reflectors toward the amplitudes of the stronger ones. While most energy in the resulting panel is focused around \( \lambda = 0 \), some residual exists away from zero subsurface offset indicating that we used an imperfect migration slowness model. (Figure 1(b) also includes a multiplicative DSO penalty function that provides the observed increased weighting at farther absolute offsets.) Applying this weighting function in the inversion scheme will penalize and largely cancel out energy common to both images - even where it is.

Figure 1: Penalty functions \( P \) where the blue-to-red color scheme goes from zero to unity. (a) DSO operator. (b) Weight operator function derived from an existing (baseline) image.
4D Adjoint-source velocity inversion
equally poorly focused. Accordingly, the only contributions to the time-lapse estimate will be from the relative changes between the images. Importantly for 4D practice, this approach does not rely on baseline/monitor data or image difference volumes that are notoriously noisy due to non-repeatable 4D acquisition. As such, we assert that this approach (where required) is likely to be more robust than other 4D velocity inversion approaches that involve computing data differences.

EXPERIMENT

Our numerical experiment tests the validity of the two 4D inversion approaches in a relatively noise-free environment. Figure 2(a) presents the baseline slowness perturbation $\Delta s_1$ from a constant background model, while Figure 2(b) presents the baseline and monitor perturbations together, $\Delta s_1 + \Delta s_2$. (We plot these slowness panels and those that follow at the same color scale and will henceforth omit the scale bar.) Using an elastic finite difference modeling operator we generated baseline and monitor datasets, $f_{r_1}$ and $f_{r_2}$, using the slowness models shown in Figure 2, a shear-wave slowness profile $\sqrt{3}$, and a density profile comprised of six equally space horizontal reflectors running the full model width.

Figure 2: Slowness perturbations from a constant $s_0=0.5$ s/km background slowness model used in the synthetic tests. (a) Baseline perturbation $\Delta s_1$. (b) Baseline plus monitor perturbations $\Delta s_1 + \Delta s_2$.

Figure 3(a) presents the one-way wave-equation migration image of the baseline dataset using $s_0 = 0.5$ s/km as the migration slowness model. The reflectors are imaged horizontal except near the location of the baseline perturbation at $x = 1.5$ km. We applied 15 iterations of image-domain ASM tomography inversion scheme discussed above to generate the $\Delta s_1$ slowness perturbation estimate shown in Figure 3(b). Figure 3(c) presents the baseline image remigrated with slowness model $s_0 + \Delta s_1$. We note that the baseline slowness estimate is a good approximation of the true perturbation in Figure 2(a) and has done a decent job of flattening the six reflectors.

![Figure 3](image-url)

Figure 3: Baseline inversion experiment. (a) Baseline image generated using $s_0 = 0.5$ s/km. (b) Inverted baseline perturbation $\Delta s_1$. (c) Baseline image using $s_0 + \Delta s_1$.

Figure 4(a) shows the migrated monitor image constructed from the estimated baseline model $s_0 + \Delta s_1$. The imaged reflectors are again nearly horizontal save for the pull down centered about $x = 2.5$ km as expected from the introduced monitor slowness perturbation in Figure 2(b). Figure 4(b) and 4(c) presents a horizontal concatenation of a subset of the penalized image offset gathers corresponding to the image in Figure 4(a). Figure 4(b) shows the effect of applying the DSO-only penalty function from Figure 1(a). We observe that while the majority of residual energy is located in the vicinity of where the monitor slowness perturbation occurs, some unflattened energy remains between $x = 0.5$ km and $x = 2.0$ km due to the imperfect baseline velocity analysis. In general, this suggests that using the DSO penalty alone for 4D velocity analysis can lead to contaminated estimates of $\Delta s_{TL}$ because residual energy from the baseline analysis can leak into the monitor inversion.

Figure 4(c) presents the monitor image residuals after applying the combined DSO+4D penalty, an example panel of which is shown in Figure 1(b). Here, the consistency in the migrated baseline and monitor images - as encapsulated in the 4D penalty operator - effectively masks most-to-all of the residual energy occurring between $x = 0.5$ km and $x = 2.0$ km. The remaining energy residuals are now more closely associated with those directly stemming from the monitor perturbation, and are likely to give an equal or better result than when following the
4D Adjoint-source velocity inversion

Figure 5(a) presents the $\Delta s_{TL} \equiv \Delta s_2 - \Delta s_1$ time-lapse slowness estimate from the independent strategy, while Figure 5(b) shows that of the relative approach. We observe that having the more restricted energy residuals using the 4D+DSO penalty function helps to spatially localize the imaged perturbation. We also observe in Figure 5(a) that the ASM inversion is still trying to image the baseline perturbation that was not completely accounted for in the first stage of analysis. As discussed above this represents “inversion leakage” between the baseline and monitor surveys and could possible introduce erroneous interpretation in more realistic tests.

CONCLUSIONS

We present an extension of 3D adjoint-state methods in the image domain to the 4D seismic velocity inversion problem. We discuss different absolute and relative inversion strategies that use different penalty functions to down- and upweight different components of imaged wavefield energy. The independent 4D inversion approach uses the difference between two separate 3D inversion estimates to compute the 4D perturbation. The relative inversion strategy couples the baseline and monitor datasets together by incorporating the baseline image directly into the penalty operator. This allows monitor image energy matching that in the baseline image - but itself not necessarily optimally focused - to be masked and precluded from inversion. We assert that the masking strategy should make ASM velocity inversion more robust in the present of 4D field data noise such as non-repeatable acquisition.

ACKNOWLEDGMENTS

We thank David Lumley for helpful discussions. Shragge acknowledges WAERA support through a Research Fellowship. This research was partly funded by the sponsors of the UWA:RM consortium. Simulation results were computed using IVEC HPC facilities. The reproducible numerical examples use the Madagascar open-source package (http://www.reproducibility.org).

Figure 4: Monitor inversion experiment. (a) Image generated using estimated slowness field $s_0 + \Delta s_1$. (b) Horizontal concatenation of penalized subsurface offset gathers for th monitor survey data for DSO penalty operator. (c) As in (b) but for 4D+DSO penalty operator.

Figure 5: Monitor inversion results. (a) Absolute 4D strategy perturbation estimate $\Delta s_2$. (b) Relative 4D strategy perturbation estimate $\Delta s_2$. 
REFERENCES

SUMMARY

Pore-pressure variations inside producing reservoirs result in excess stress and strain that cause time shifts of reflected waves. Inversion of seismic data for pressure changes requires better understanding of the dependence of compaction-induced time shifts on reservoir pressure reduction. Here, we investigate pressure-dependent behavior of P-, S-, and PS-wave time shifts from reflectors located above and below a rectangular reservoir embedded in a homogeneous half-space. Our geomechanical modeling algorithm generates the excess stress/strain field and the stress-induced stiffness tensor as linear functions of reservoir pressure. Analysis of time shifts obtained from full-waveform synthetic data shows that they vary linearly with pressure for reflectors above the reservoir, but become nonlinear for reflections from the reservoir or deeper interfaces. Time-shift misfit curves computed with respect to noise-contaminated data from a reference reservoir for a wide range of pressure reductions display well-defined minima. Our results indicate that stable pressure estimation requires including multicomponent time shifts from reflectors below the reservoir and applying a nonlinear global inversion algorithm.

INTRODUCTION

Compaction-induced seismic traveltime shifts can potentially be inverted for pressure and fluid distribution inside a producing reservoir to monitor hydrocarbon movement and intercompartment pressure communication (Landrø, 2001; Lumley, 2001; Calvert, 2005; Hodgson et al., 2007). Knowledge of reservoir pressure can also be used to estimate stress and strain variations outside the reservoir (Hewanger and Horne, 2005; Dusseault et al., 2007; Scott, 2007), which helps guide drilling decisions and reduce the cost of repairing or replacing wells snapped or sheared by high stresses. Conventional methodologies employ poststack data and compaction-induced vertical stress-strain to estimate time-lapse volume and velocity changes (Hatchell and Bourne, 2005; Janssen et al., 2006; Roste, 2007). However, migration and stacking of data represents a complex filtering process that can corrupt phase relationships and arrival times. Additionally, it has been shown that shear (deviatoric) strains generate significant time shifts, requiring the use of triaxial strain in geomechanical interpretation of time-lapse data (Herwanger, 2008; Sayers, 2010; Smith and Tsvankin, 2011).

Fuck et al. (2009, 2011) develop a modeling methodology using a triaxial strain formulation and the nonlinear theory of elasticity, and estimate P-wave time shifts using anisotropic ray tracing. Smith and Tsvankin (2011) confirm the P-wave results of Fuck et al. (2009) and analyze time shifts for S- and PS-waves using finite-difference elastic modeling. These studies demonstrate that both volumetric (hydrostatic) and deviatoric (shear) strains generate significant time shifts for all three (P, S, and PS) modes. Sensitivity of time shifts to reservoir pressure, analyzed by Smith and Tsvankin (2011), varies with wave type and reflector location. Here, we use the geomechanical and seismic modeling methodology of Smith and Tsvankin (2011) to study the dependence of P-, S-, and PS-wave time shifts on reservoir pressure. For a set of reflectors located above and below the reservoir, we examine the linearity of time shifts expressed as a function of pressure for different noise levels in the data. We also examine time-shift misfits with respect to a reference reservoir for a realistic range of pressure reductions.

THEORETICAL BACKGROUND

Modeling traveltimes shifts caused by hydrocarbon production from a reservoir is a three-step process. First, changes in reservoir parameters (here, pressure reduction) result in excess stress and strain in and around the reservoir. Second, the excess stress/strain perturbs the stiffness coefficients that govern the velocities and traveltimes of seismic waves. Third, the stress-induced stiffnesses are used to model seismic data and compute the time shifts between the baseline and monitor surveys. Note that time shifts are generally nonlinear in the relevant stiffness coefficients even when the model is homogeneous and isotropic.

We employ a simplified, 2D rectangular reservoir model after Fuck et al. (2009, 2011) (Figure 1), comprised of isotropic Berea sandstone that follows standard Biot-Willis compaction theory (Hofmann et al., 2005; Zoback, 2007). The effective pressure in the reservoir \( P_{\text{eff}} \) changes according to a reduction of the pore fluid pressure \( P_{\text{fluid}} \):

\[
P_{\text{eff}} = P_c - \alpha P_{\text{fluid}} ,
\]

where \( P_c \) is the confining pressure of the overburden, and \( \alpha \) is known as the effective stress coefficient (Biot-Willis coefficient for “dry” rock, where air is the only pore infill). Pressure changes occur only in the reservoir block. The resulting displacement, stress, and strain changes throughout the section are computed using a finite element, plane-strain solver (COMSOL AB, 2008). The modeled strains are linear in the pressure drop, and 1-2 orders of magnitude higher inside the reservoir than in the overburden.

The strain-induced variations of the stiffness tensor can be expressed using the so-called nonlinear theory of elasticity (Hearmon, 1953; Thurston and Brugger, 1964; Fuck et al., 2009):

\[
c_{ijkl} = c_{ijkl}^{0} + \frac{\partial c_{ijkl}^{0}}{\partial e_{mn}} \Delta e_{mn} = c_{ijkl} + c_{ijklmn} \Delta e_{mn} .
\]
Compaction-induced traveltimes: Methodology and study

Figure 1: Reservoir geometry after Fuck et al. (2009). Pore-pressure ($P_p = P_{\text{fluid}}$) reduction occurs only within the reservoir, resulting in an anisotropic velocity field due to the excess stress and strain. For geomechanical modeling, the reservoir is located in a model space measuring 20 km x 10 km, which is sufficient for obtaining stress, strain, and displacement close to those for a half-space. The reservoir is comprised of and embedded in homogeneous Berea sandstone ($V_p = 2300$ m/s, $V_S = 1456$ m/s, $\rho = 2140$ kg/m$^3$) with the following third-order stiffness coefficients: $C_{111} = -13,904$ GPa and $C_{112} = 533$ GPa (Prioul et al., 2004). The effective stress (Biot) coefficient ($\alpha$) for the reservoir is 0.85. Velocities in the model are reduced by 10% from the laboratory values to account for the difference between static and dynamic stiffnesses in low-porosity rocks (Yale and Jamieson, 1994). Results for reflectors labeled A and C are shown in Figures 3 and 4 and discussed in the text.

![Figure 1: Reservoir geometry after Fuck et al. (2009).](image)

Figure 2: Typical two-way traveltime shift surfaces for (a) P-waves, (b) S-waves, and (c) PS-waves, measured using 22 reflectors around the reservoir (white box) of Figure 1 (Smith and Tsvankin, 2011). The time shifts correspond to hypothetical specular reflection points at each (X,Z) location in the subsurface. Positive shifts indicate lags where monitor survey reflections arrive later than the baseline events; negative shifts are leads. Source location is indicated by the white asterisk on top; the pressure drop $\Delta P = 20\%$.

![Figure 2: Typical two-way traveltime shift surfaces.](image)

where $c_{ijkl}^{\alpha\beta\gamma}$ is the second-rank stiffness tensor of the background (unperturbed) medium, $c_{ijklmn}$ is a sixth-order strain-sensitivity tensor, and $\Delta\kappa_{mn}$ is the excess strain. The tensors $c_{ijkl}$ and $c_{ijklmn}$ can be represented in matrix notation ($C_{\alpha\beta\gamma}$ and $C_{ijklmn}$, respectively) using the Voigt convention (Tsvankin, 2005; Fuck and Tsvankin, 2009). Despite the term “nonlinear,” which refers to Hooke’s law, equation 2 expresses the stiffnesses $c_{ijkl}$ as linear functions of the strain $\Delta\kappa_{mn}$. Wave propagation through the stressed medium may then be modeled using Hooke’s law with the stiffness tensor $c_{ijkl}$. For our 2D models, we need only two elements of the strain-sensitivity tensor, and employ empirical values of the coefficients $C_{\alpha\beta\gamma}$ measured by Sarkar et al. (2003).

Fuck et al. (2009, 2011) present a Fermat integral expression for traveltimes along a raypath through a stressed medium. The integrand is linearized in the volumetric (hydrostatic) and deviatoric (shearing) strains under the assumption that pressure-induced strain is small. In general, however, time shifts are nonlinear in the stiffness coefficients, and therefore, strains. Indeed, even the P-wave velocity in homogeneous, isotropic media is given by $\sqrt{\kappa / \rho}$, and traveltimes over distance R is $T = R \sqrt{\kappa / \rho}$. Therefore, the issue is the range of $\Delta\kappa_{mn}$ and $\Delta\kappa_{ij}$ for which traveltime shifts can be accurately described as linear functions of the stiffnesses and therefore, of reservoir pressure reduction.

Smith and Tsvankin (2011) employ an elastic finite-difference algorithm to compute time shifts of P-, S-, and PS-reflections for the model in Figure 1. Figure 2 shows typical two-way time shifts for a reservoir at 1.5 km depth with a pressure drop of 20%. For P-waves (Figure 2a) the results are close to those obtained by ray tracing (Fuck et al., 2009). Strain-induced P-wave velocity anisotropy around the reservoir causes offset-dependent traveltime shifts. Laterally varying P-wave time-shift patterns below the reservoir are due to elevated shearing (deviatoric) strains at the reservoir endcaps. The combination of higher volumetric and deviatoric strains inside the reservoir generates large S-wave and PS-wave time shifts from reflectors beneath it (Figures 2b,c).

The time-shift distributions in Figure 2 show that specific wave types/reflector locations exhibit different sensitivities to reservoir pressure changes, which provides useful guidance for designing a pressure-inversion procedure. If time shifts change linearly with pressure, they can be represented by a system of linear equations for a multicompartment reservoir: $A \Delta P = \Delta t$, where $A \Delta P$ is a vector containing pressure drops in all reservoir compartments, and $\Delta t$ is a vector of P-, S-, and/or PS-wave time shifts for a range of receiver offsets. Elements of matrix A are the pressure-dependent time-shift gradients ($\partial t_i (x,z) / \partial P_j$) (i denotes offset and j reservoir compartment) of reflections for specific source-receiver pairs. Therefore, for small stiffness perturbations and specific combinations of reservoir geometry, source location, and reflector depths, the coefficients of A can be computed using standard linear inversion techniques.
Compaction-induced travelt ime shifts: Methodology and study

METHODOLOGY

We investigate the behavior of time shifts for reflectors A & C shown in Figure 1 for a range of pressure drops between 0.01% and 30%. The initial reservoir pressure of modeled and reference reservoirs corresponds to a state of stress/strain equilibrium. Thus, the initially homogeneous stiffness/velocity field across the section becomes heterogeneous as reservoir pressure is reduced. Following geomechanical finite-element modeling of the pressure-induced strain $\Delta e_{mn}(x, z)$, changes in the stiffnesses are computed from equation 2. These stiffnesses are used by an elastic finite-difference code (Sava et al., 2010) to generate shot records from the source located at X=0 km. Reflectors A and C (Figure 1) are inserted in the model to sample travel times and estimate time shifts. Then the multicomponent synthetic data are processed to isolate arrivals from the specific reflector, and time shifts between the initial/reference (baseline) and monitor reservoir models are computed by cross-correlation. P-wave time shifts are measured from the vertical displacement, while S- and PS-wave shifts are measured on the horizontal component. Additional smoothing is applied to reduce interference-related distortions.

The first step in computing time-shift misfits is to generate time shifts for a wide range of pressure reductions. Then the model corresponding to $\Delta P = 15\%$ is designated as the reference one (i.e., $\Delta P = 15\%$ is the actual pressure drop we would like to invert for). The misfit for each value of $\Delta P$ for a specific wave type and reflector is computed as the L2-norm of the difference between the time shifts for that $\Delta P$ and those for the reference model. Joint misfit is the sum of the L2-norms for all wave types at a specific reflector. Misfits shown here have not been normalized in order to facilitate comparison of results for different wave types at specific reflectors.

ANALYSIS

Figure 3 shows measured time shifts of P- and S-waves reflected from interfaces A and C of Figure 1 for 20 reservoir pressure drops ranging between 0.01% and 30% of the initial reservoir fluid pressure. Data for S- and PS-waves (excited by a vertical force) do not include time-shift estimates at X=0 due to the low amplitudes of the horizontal displacement at small offsets.

In general, P-wave time shifts at reflector A are lags linear in pressure drop, caused by a P-wave velocity reduction above the reservoir. S-waves at reflector A experience small velocity increases and time-shift leads due to changes in the stiffness $C_{55}$ in the overburden (Smith and Tsvankin, 2011). PS-wave shifts above the reservoir are close to zero, because P-lags are almost canceled by S-leads. Time shifts for all wave types from reflector C are clearly nonlinear as a function of pressure because of large stiffness perturbations inside the reservoir. Therefore, we can apply linear inversion to time shifts only at reflectors above the reservoir. While time shifts for reflectors beneath the reservoir may be approximated as linear for small pressure drops, they become nonlinear after 10-20% pressure reduction, which necessitates application of a nonlinear inversion algorithm.

As another aspect of our 4D inversion feasibility study, we evaluate the sensitivity of “compound” time shifts for a certain mode and a given reflector to a range of reservoir depressurizations. The term “higher sensitivity” indicates a combination of high time-shift values with respect to reservoir pressure variation, and a steeply-sloped misfit curve with a distinct minimum at the reference (true) pressure value. L2-norm time-shift misfits were computed against a reference reservoir with a pressure drop of 15%. The results for P-, S-, PS-wave and joint misfits at reflectors A and C are shown in Figure 4. Time-shift misfits for 2-ms noise (top row) do not differ significantly from the corresponding noise-free estimates. Misfit slopes correlate well with the time-shift magnitudes for each wave type and reflector depth in Figures 2 and 3. At reflector A, PS-wave time shifts at larger offsets are more sensitive to lower pressure drops than time shifts of P- and S-waves. S-wave shifts clearly provide the highest sensitivity for all pressure drops beneath the reservoir at reflector C. However, in all cases, the joint misfit is more sensitive to pressure than the misfit for any single wave type.

The influence of increasing Gaussian noise with standard deviations of 2 ms, 5 ms, and 10 ms added to the reference reservoir time shifts is shown in Figure 4. Substantial degradation in sensitivity begins for reflectors above the reservoir for approximately 5-ms noise (Figure 4c). The misfit curves develop local minima, indicating that a linear inversion algorithm may fail at moderate noise levels. For strong noise reaching 10 ms (approximately 2/3 of the maximum P-wave time shifts for noise-free data) (Figures 4e,f), the misfits for reflectors above the reservoir are significantly distorted. However, time shifts of all wave types for reflector C are sufficiently large to provide smooth sensitivity curves with a clear global minimum. In general, as noise levels increase, the sensitivity of time shifts for individual wave types/reflectors to pressure reduction declines, but S-wave and joint misfits at or beneath the reservoir remain reasonably high. These results suggest that in the presence of substantial noise, joint and S-wave time shifts for reflectors below the reservoir provide the most reliable input data for pressure inversion.

INVERSION ALGORITHM

Reservoir pressure is estimated with a global inversion technique that follows the nearest-neighborhood algorithm of Sambridge (1999). The inversion process concentrates model updates in the region around a number of minimum-error points in parameter space. At these points, the algorithm makes a simple estimate of the local gradient by comparing misfits at a limited number of nearby points. The next step is addition of models within the local region along and perpendicular to the gradient direction. As a result, successive minimum misfits progress toward the local and/or global minima. The algorithm also inserts additional models across the parameter space to find other local minima and reconstruct the shape of the error surface (Figure 5).
Figure 3: Time shifts for reservoirs with pressure drops ranging from 0.01% and 30%. The source is located above the center of the reservoir at X=0. (a,b) P-waves, (c,d) S-waves, and (e,f) PS-waves for (a,c,e) reflector A and (b,d,f) reflector C. Plot legends indicate receiver (surface) coordinates between X=0 km and X=2.0 km.

Figure 4: P-, S-, PS-wave, and joint L2-norm misfit curves computed using noise-contaminated reference time shifts. Reference reservoir pressure drop is 15%. (a,c,e) reflector A, (b,d,f) reflector C. Standard deviation of Gaussian noise in the reference time shifts increases by row: (a,b) ±2 ms, (c,d) ±5 ms, and (e,f) ±10 ms.

Figure 5: Error surface generated by the inversion algorithm for a single-compartment reservoir (pressure drop ΔP = 15.28%, reservoir width = 1800 m) using joint misfits from reflector C with 5 ms noise. Six iterations with 185 total models (black circles) were completed. The global minimum of the error surface practically corresponds to the reference values. Here, reservoir width was used as a second parameter for a single-compartment inversion.

CONCLUSIONS

We have used a 2D rectangular reservoir model to study the dependence of P-, S- and PS-wave time shifts on reservoir pressure with the goal of assessing the feasibility of pressure estimation. While the stress-induced stiffness tensor is linear in pressure and excess strain, traveltime shifts generally depend on the stiffness coefficients in a nonlinear fashion. In the regions with relatively small strain, pressure-related perturbations in the stiffnesses are not sufficiently large to cause nonlinearity of time shifts. For example, time shifts are linear in pressure reduction for reflectors above the reservoir. However, strains inside the reservoir are much larger than those in the surrounding medium, creating large stiffness changes. Thus, waves reflected from points at and below the reservoir exhibit nonlinear time-shift dependence on pressure.

L2-norm misfits of time shifts computed with respect to a reference reservoir show that S-wave time shifts from reflectors beneath the reservoir are generally most sensitive to pressure. Misfit curves for S-wave reflections from deep interfaces provide robust indicators of pressure change even for reference time shifts contaminated with 10-ms noise. Therefore, inversion of noisy data for reservoir pressure needs to operate with reflections from beneath the reservoir. We have also designed a nonlinear inversion algorithm that can handle significant time shifts caused by pressure reductions exceeding 10%.
Compaction-induced traveltime shifts: Methodology and study

REFERENCES


Hatchell, P. and S. Bourne, 2005, Rocks under strain: Strain-induced time-lapse time shifts are observed for depleting reservoirs: The Leading Edge, 24, no. 12, 1222–1225.


Herwanger, J., 2008, R we there yet?: 70th Annual International Meeting, EAGE, Expanded Abstracts.


Hofmann, R., X. Xu, M. Batzle, M. Prasad, A.-K. Furre, and A. Pillitteri, 2005, Effective pressure or what is the effect of pressure?: The Leading Edge, 24, no. 12, 1256–1260.


Multiparameter TTI tomography of P-wave reflection and VSP data
Xiaoxiang Wang and Ilya Tsvankin, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

Transversely isotropic models with a tilted symmetry axis (TTI media) are widely used in depth imaging for complex geologic structures. Here, we present a modification of a previously developed 2D P-wave tomographic algorithm for building heterogeneous TTI models and apply it to synthetic data. For stable reconstruction of the symmetry-direction velocity \( V_{P0} \) and the anisotropy parameters \( \epsilon \) and \( \delta \), reflection data are combined with walkaway VSP (vertical seismic profiling) travel times. To improve the convergence of the inversion algorithm, we propose a three-stage model-updating procedure that gradually relaxes the constraints on the spatial variations of the anisotropy parameters \( \epsilon \) and \( \delta \), while the symmetry axis is kept orthogonal to the reflectors. We also incorporate geologic constraints into tomography by designing regularization terms that penalize parameter variations in the direction parallel to the interfaces. The performance of the regularized joint tomography of reflection and VSP data is examined for two sections of the BP TTI model that contain an anticline and a salt dome. The tests provide useful guidance for building accurate anisotropic models for prestack depth migration.

INTRODUCTION

Depth imaging for complex geologic environments (including fold-and-thrust belts and subsalt plays) requires heterogeneous anisotropic velocity models such as transverse isotropy with a vertical (VTI) or tilted (TTI) axis of symmetry. Because P-wave reflection traveltimes do not provide sufficient constraints for resolving all relevant TTI parameters, it is necessary to use additional information, such as borehole data (Moric et al., 2004; Tsvankin, 2005; Bakulin et al., 2010). Also, nonuniqueness of the inversion can be mitigated by regularization which imposes a priori constraints on the estimated model (Engl et al., 1996). Zhou et al. (2011) develop multiparameter reflection tomography for TTI media and find that simultaneous estimation of all three relevant parameters \( V_{P0}, \epsilon, \) and \( \delta \) helps flatten common-image gathers (CIGs) better than single-parameter (only \( V_{P0} \)) inversion.

In a previous publication (Wang and Tsvankin, 2011; hereafter, referred to as Paper I), we develop a 2D ray-based tomographic algorithm for iteratively updating the parameters \( V_{P0}, \epsilon, \) and \( \delta \) of TTI media defined on rectangular grids (the symmetry axis is set orthogonal to the imaged reflectors). Here, we present a three-stage inversion methodology in which gridded tomography is preceded by two partial parameter-updating steps designed to stabilize the inversion. Then the tomographic algorithm is tested on two sections of the TTI model devised by BP.

METHODOLOGY

The tomographic algorithm employed here is described in detail in Paper I. The residual moveout in common-image gathers produced by prestack Kirchhoff depth migration is minimized by iterative parameter updates. If walkaway VSP surveys are available, VSP traveltimes are computed for each trial model and included in the objective function:

\[
F(\Delta \lambda) = \|A \Delta \lambda - b\|^2 + \xi_{\text{vhp}}^2 \|E \Delta \lambda - d\|^2 + R(\Delta \lambda),
\]

where the vector \( \Delta \lambda \) represents the parameter updates, the elements of the matrix \( A \) are the traveltimes derivatives with respect to the medium parameters at each grid point (\( A \) is obtained analytically along the raypaths), \( b \) is a vector containing the residual moveout in CIGs, the matrix \( E \) is composed of the VSP traveltime derivatives, and the vector \( d \) is the difference between the observed and calculated VSP traveltimes for each source-receiver pair. The regularization term in equation 1 has the form:

\[
R(\Delta \lambda) = \xi_1^2 \|\Delta \lambda\|^2 + \xi_2^2 \|L_1(\Delta \lambda + \lambda^0)\|^2 + \xi_3^2 \|L_2(\Delta \lambda + \lambda^0)\|^2,
\]

where \( \xi_2 \|\Delta \lambda\|^2 \) restricts the magnitudes of parameter updates that have small derivatives in the matrices \( A \) and \( E \), the operators \( L_1 \) and \( L_2 \) are designed to make parameter variations more pronounced in the direction normal to the interfaces, and \( \xi_1, \xi_2, \) and \( \xi_3 \) are the regularization coefficients/weights. In 2D, the reflector normal is defined by the dip angle with the vertical, which is computed from the depth image using Madagascar program “sfdip.” Then the symmetry-axis tilt \( \nu \) at each grid point is set equal to the corresponding dip. To construct the matrix \( L_1 \), we first compute two components of the gradient vector \( \nabla \lambda \) from the following finite-difference approximation:

\[
\lambda' (x) = \frac{\lambda(x + dx) - \lambda(x - dx)}{2dx} + O[(dx)^2],
\]

\[
\lambda' (z) = \frac{\lambda(z + dz) - \lambda(z - dz)}{2dz} + O[(dz)^2],
\]

where \( \lambda \) is the parameter (\( V_{P0}, \epsilon, \) or \( \delta \)) at the grid point with the coordinates \( x \) and \( z \), and \( dx \) and \( dz \) are the cell dimensions. Since the dip field yields the vector \( \mathbf{n} \) orthogonal to reflectors, we minimize the norm of the cross-product \( |\mathbf{n} \times \nabla \lambda| \) at all grid points, which is equivalent to aligning the direction of the largest parameter variation with \( \mathbf{n} \) and restricting the variations along interfaces. A higher-order finite-difference approximation can be used to include more cells. Similarly, two components of the Laplacian operator \( \nabla^2 \lambda \) can be approximated as follows:

\[
\lambda'' (x) = \frac{\lambda(x + dx) - 2\lambda(x) + \lambda(x - dx)}{(dx)^2},
\]

\[
\lambda'' (z) = \frac{\lambda(z + dz) - 2\lambda(z) + \lambda(z - dz)}{(dz)^2}.
\]
Then the cross-product $\|n \times \nabla^2 \lambda\|$ is minimized to mitigate the parameter variations along the interfaces.

As described in Paper I, the anisotropic velocity field is updated using a linearized iterative algorithm. In the first few iterations, the symmetry-direction velocity $V_{0z}$ is typically inaccurate and simultaneous inversion for all TTI parameters may result in unacceptably large updates for $\varepsilon$ and $\delta$. If the anisotropy parameters are moderate ($\varepsilon < 0.25$ and $\delta < 0.15$ in our tests), it is convenient to fix them temporarily at the initial values (which are typically small) and limit the updates to the velocity $V_{0z}$. At the second stage of parameter updating, the model is divided into several layers based on the picked reflectors, and the anisotropy parameters are assumed to be constant within each layer. Then $\varepsilon$ and $\delta$ are updated in each layer, while the velocity $V_{0z}$ is still inverted on a grid. Such a “quasi-factorized” assumption, equivalent to strong smoothing of $\varepsilon$ and $\delta$, helps resolve all TTI parameters if $V_{0z}$ is a linear function of the spatial coordinates and at least two distinct dips are available (Behera and Tsvankin, 2009). At the third and last stage of velocity analysis, the parameters $\varepsilon$ and $\delta$ are updated on the same grid as that for $V_{0z}$ to allow for more realistic treatment of heterogeneity. Still, because P-wave kinematics are less sensitive to the anisotropy parameters than to the symmetry-direction velocity, the $\varepsilon$- and $\delta$-fields in function 2 should be regularized with larger weights.

SYNTHETIC EXAMPLES

We test the joint tomography of P-wave long-spread reflection data and walkaway VSP traveltimes on two sections of the BP TTI model that contains an anticline structure and a salt dome (http://www.freeusp.org/2007_BP_Ani_Vel_Benchmark/). Here, we show only the results for the salt section (Figures 1, 2).

Migration velocity analysis (MVA) is applied to CIGs from $x = 16$ km to 46 km with an interval of 150 m (the maximum offset is 10 km). The dataset contains a vertical “well” at location $x_{\text{well}} = 29.9$ km to the left of the salt body (Figure 1). Two sets of 24 receivers (one between $z = 5275$ m and 5562.5 m and the other between $z = 8400$ m and 8687.5 m) were placed at even intervals in the well to record a walkaway VSP survey. The maximum offset for the VSP data is 10 km with a source interval of 50 m. The input data also include check-shot traveltimes in the same well obtained every 50 m from $z = 1743.75$ m to 9093.75 m.

During the inversion, the water layer and salt body are kept isotropic with the velocities fixed at the correct values. Also, the actual positions of the top and flanks of the salt dome are assumed to be known, and the update is performed only for the sedimentary formations around the salt body. Since ray tracing becomes unstable in the presence of sharp velocity contrasts, we apply 2D smoothing to the velocity model to find the ray-paths crossing the salt. The traveltimes and their derivatives are then calculated in the original (unsmoothed) model. To build an initial model, we compute a 1D profile of $V_{0z}$ from the check-shot traveltimes and then obtain the 2D velocity field by extrapolation that conforms to the picked interfaces. The CIGs with the isotropic initial model display substantial residual moveout, and most reflectors are misplaced because of the large velocity errors.

After two iterations of velocity ($V_{0z}$) updating with fixed $\varepsilon$ and $\delta$ and two more iterations with the “quasi-factorized” TTI model assumption (see above), the estimated parameter fields produce relatively flat CIGs and an improved image. To further reduce the residual moveout in CIGs and the VSP traveltimes misfit, the anisotropy parameters are estimated on the same grid as that for $V_{0z}$. After three more iterations, the velocity $V_{0z}$ (Figure 3(a)) above $z = 7$ km on the left side of the salt body is relatively well-resolved (errors in most areas do not exceed 3%); however, the errors in $V_{0z}$ on the right side are higher because of the limited constraints from VSP data. (The VSP rays originated to the right of the well cross the high-velocity salt body, and we assign smaller weights in the objective function to those traveltimes.) The spatial variations of $\varepsilon$ and $\delta$ are partially recovered from the water bottom down to $z = 5$ km (Figure 3(c) and 3(d)). Because of the limited offset-to-depth ratio and poor coverage of VSP rays (especially for the right part of the model) at depth, the anisotropy parameters for grid points below 5 km could not be updated after the first iteration. The final inverted model practically removes the residual moveout in CIGs and the VSP traveltimes misfit, which results in an accurate image (Figure 4).

CONCLUSIONS

We presented a three-stage tomographic algorithm for TTI media that operates with P-wave reflection and VSP data. The parameters $V_{0z}$, $\varepsilon$, and $\delta$ are updated iteratively, while the tilt field is computed from the depth image by setting the symmetry axis perpendicular to the reflectors. The regularization terms in the objective function suppress parameter variations in the direction parallel to boundaries. Such structure-guided regularization also helps propagate along interfaces the most reliable updates corresponding to large derivatives in the Fréchet matrix (e.g., those in the cells crossed by dense VSP rays).

The joint tomography regularized by the structure-guided terms was successfully applied to two sections of the BP TTI model that include an anticline and a salt dome. A purely isotropic velocity field, which was obtained from check-shot traveltimes and extrapolated along the horizons, served as the initial model. With constraints from P-wave reflection and VSP data, the TTI parameters in the shallow part (above 5 km) of the section are well-resolved. However, the errors in the anisotropy parameters $\varepsilon$ and $\delta$ increase with depth due to the small offset-to-depth ratio and poor coverage of VSP rays.

ACKNOWLEDGMENTS

We are grateful to BP for providing the synthetic data set. This work was supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena, Colorado School of Mines.
Figure 1: Section of the BP TTI model with a salt dome (the grid size is 6.25 m × 6.25 m). The top water layer and the salt body are isotropic with the P-wave velocity equal to 1492 m/s and 4350 m/s, respectively. (a) The symmetry-direction velocity $V_{P0}$. The vertical “well” at $x = 29.9$ km is marked by a black line. (b) The tilt of the symmetry axis, which is set orthogonal to the interfaces. The anisotropy parameters (c) $\epsilon$ and (d) $\delta$.

Figure 2: Depth image produced with the actual parameters from Figure 1. The strong reflections from the top of the salt dome and the flanks right beneath it are well-imaged by Kirchhoff depth migration, but the deeper segments of the flanks are blurred.
Figure 3: Inverted TTI parameters (a) $V_{P0}$, (c) $\epsilon$, and (d) $\delta$ after the final iteration of joint tomography. All three parameters are estimated on a 200 m $\times$ 100 m grid. (b) The symmetry-axis tilt $\nu$ computed from the depth image obtained before the final iteration.

Figure 4: The migrated section computed with the final inverted TTI model from Figure 3.
REFERENCES


Integrated migration and internal multiple elimination
Kees Wapenaar*, Jan Thorbecke, Joost van der Neut and Evert Slob, Delft University of Technology, Filippo Broggini, Jyoti Behura and Roel Snieder, Colorado School of Mines

SUMMARY
Standard seismic migration, applied to data with internal multiples, leads to images with ghosts. The reason is that one-way wave field extrapolation operators give erroneous downgoing and upgoing fields in the subsurface, which in the correlation process lead to ghosts. By using “data-driven wave field reconstruction” it is possible to obtain the correct downgoing and upgoing wave fields in the subsurface. Using these in the correlation process leads to ghost-free images.

INTRODUCTION
Standard seismic depth migration involves one-way downward extrapolation of downgoing and upgoing wave fields from the acquisition surface to a depth level in the subsurface, correlating the downward extrapolated downgoing and upgoing fields and selecting the t = 0 component to get an image of the reflectivity at the chosen depth level (Cheroutb, 1971; Berkhout, 1982). By repeating this procedure for all depth levels of interest, a reflectivity image of the subsurface is obtained. Migrated data usually suffer from ghost images due to multiple reflections. Assuming the surface-related multiples are suppressed prior to migration (Verschuur et al., 1992), these ghost images are caused by the internal multiples (e.g. event 3 in Figures 1b and c).

Fig. 1: Example of standard depth migration. (a) Medium with two dipping interfaces. (b) Reflection response at the surface (one shot record). (c) Migration result with ghosts. (d) Downward extrapolated upgoing wave field at virtual receivers (light-blue triangles in Fig. a) below the deepest interface.

The obvious explanation is that a standard depth migration scheme cannot distinguish primaries from multiply reflected waves, hence, a multiply reflected event from a relatively shallow interface (as indicated by the ray-path in Figure 1a) is interpreted as a primary reflection from a deeper interface and thus erroneously imaged at a too large depth. Another way of explaining the same phenomenon is that one-way downward extrapolation operators yield erroneous downgoing and upgoing wave fields at depth. For example, Figure 1d shows the downward extrapolated upgoing wave field at the virtual receivers (the light-blue triangles in Figure 1a) below the deepest interface. Because the source is at the surface there can only be downgoing waves below the deepest interface, hence, the upgoing field in Figure 1d should be zero. The erroneous non-zero field correlates with the downgoing field below the deepest reflector and gives rise to the ghost images in Figure 1c.

Here we introduce a new approach to deal with internal multiple reflections in depth migration. The main idea is that we replace the independent one-way extrapolation operators for downgoing and upgoing waves by a new downward extrapolation scheme that gives the correct downward and upward propagating fields at depth, due to sources at the surface (hence, for the example in Figure 1 this scheme gives a zero upgoing wave field below the deepest reflector, see Figure 2b). A ghost-free image is obtained by correlating the correct downgoing and upgoing fields at depth, followed by selecting the t = 0 component (and repeating the whole procedure for all depth levels of interest), see Figure 2a. Following this approach, not only the ghost images are suppressed, but in addition the multiples in the extrapolated downgoing and upgoing fields at depth contribute to the primary image. This approach is stable with respect to errors in the macro velocity model (Figure 3). Like in standard depth migration, velocity errors cause mispositioning of reflectors and defocusing of diffractions, but the ghost suppression is not affected by these errors.
Migration and multiple elimination

![Fig. 4: Data-driven “wave field reconstruction” (Broglini et al., 2011; Wapenaar et al., 2011a) uses the reflection response at the surface and an estimate of the first arrivals (a) to obtain the full response to a virtual source in the subsurface (b). Using reciprocity and decomposition, this response is turned into the downgoing (c) and upgoing (d) fields in the subsurface, due to sources at the surface. These are the fields that are needed to form a ghost-free image of the subsurface.](image)

DOWNWARD WAVE FIELD EXTRAPOLATION

A possible method to obtain the correct downward and upward propagating fields at depth is two-way wave field extrapolation, followed by up/down decomposition. We showed previously that this leads in principle to ghost-free images (Wapenaar et al., 1987). However, we did not pursue this approach because of the high sensitivity of the two-way extrapolation operator with respect to errors in the macro velocity model. Ideally, we would like to use a downward extrapolation scheme that is not more sensitive to errors in the macro velocity model than the one-way operators used in standard depth migration. The data-driven “wave field reconstruction method” that we introduced last year (Broglini et al., 2011; Wapenaar et al., 2011a) provides the required operator. Recall that this method uses the reflection data at the surface (Figure 1b) and an estimate of the first arrivals from a virtual source in the subsurface (Figure 4a) to obtain the full response at the surface to a virtual source in the subsurface (Figure 4b). In the terminology of depth migration, an “estimate of the first arrivals from a virtual source in the subsurface” is nothing but the “one-way extrapolation operator in the macro velocity model”. Similarly, the “full response at the surface to a virtual source in the subsurface” is, via reciprocity, nothing but the “total field at depth, due to sources at the surface”. Hence, the wave field reconstruction method actually applies a one-way operator defined in a macro model (Figure 4a) to the reflection data at the surface (Figure 1b) and gives the total field at depth (Figure 4b, with reciprocity applied). Two different versions of the scheme exist, one giving the total field observed by a monopole receiver and the other by a dipole-type receiver. By subtracting and adding these responses, respectively, the total field is decomposed into the required downgoing and upgoing wave fields at depth (Figure 4c,d). In the following we show step-by-step how the downgoing and upgoing wave fields in Figure 4c,d are obtained from the reflection data at the surface and how these fields (for all depth levels of interest) are used to form the ghost-free image of the subsurface (Figure 2a).

CREATING THE VIRTUAL-SOURCE RESPONSE

We discuss the procedure by which the virtual-source response of Figure 4b is obtained from the reflection data at the surface (Figure 1b).

The medium configuration

The medium configuration in Figures 1a and 4 is defined as follows. The reflection-free acquisition surface is located at z = 0. The first reflector is defined as z = z1 = αz, with z1 = 1000 m and α = 1/4. The second reflector is parallel to the first reflector and the layer thickness is 618 m. The virtual source in Figure 4a,b is defined at xVS = (xVS, zVS) = (0, 1320) m. The velocity is constant throughout the medium (2000 m/s), and the densities of the different layers are 1000, 5000 and 1000 kg/m³. Furthermore, the medium is lossless. For this medium we previously evaluated the creation of the virtual source response by the method of stationary phase (Wapenaar et al., 2012). The velocity was chosen constant to facilitate this analysis, but the method we discuss here is also valid for variable velocity media (for a numerical experiment in a variable-velocity syncline model, see Broglini et al., 2012). In the following numerical experiment we use the constant velocity model, so that the results can easily be compared with our earlier analytical result. The sources at the surface are located between -1100 and 1100 m and the receivers between -2250 and 2250 m. The source and receiver spacing is 10 m.

Initiating the iterative process

The iterative procedure is a 2D extension of the iterative 1D scheme discussed by Roe (2002) and Broglini et al. (2011). Figure 5b shows again the direct arrivals of the virtual source at xVS in the subsurface. These direct arrivals are reversed in time, which turns the upgoing field at the surface z = 0 into a downgoing field \( p^+_0(x, t) \), see Figure 5a. The red curves along these events define a window function \( u(x, t) \), which equals 1 between these curves and 0 elsewhere. This window will be used later. The field \( p^+_0(x, t) \) is used as the initial downgoing field at the surface z = 0, which illuminates the medium from above (Figure 5c). The response to this field is obtained by convolving it with the reflection response of the medium, according to

\[
p^+_0(x_R, t) = \int_{-\infty}^{\infty} R(x_R, x, t) * p^+_0(x, t) \big|_{z = 0} \, dx,
\]

for \( x_R = 0 \). In practical situations, \( R(x_R, x, t) \) is the measured reflection response, after surface-related multiple elimination and deconvolution for the source wavelet. For this example we use a numerically modeled reflection response (Figure 1b). Equation 1 is evaluated numerically as well. The result of this convolution is the upgoing wave field \( p^+_0(x, t) \) at the surface, see Figure 5d.

The iterative process

Next, we apply an iterative scheme, which has the aim to modify the incident field in such a way that, with the time window, the incident field is equal to minus the time-reversed up-
Migration and multiple elimination

![Diagram](image)

**Fig. 5:** Initiating the iterative process. An estimate of the direct field between the virtual source and the surface (b) is reversed in time, giving the initial downgoing field \( p_0^-(x, t) \) (a, c). This initial downgoing field \( p_0^-(x, t) \) is convolved with the reflection response of the medium, giving the upgoing field \( p_0^+(x, t) \) (d).

The upgoing field \( p_0^+(x, t) \) is needed in equation 2 is shown in Figure 5d. Only the first event falls within the time window, defined by the red curves. Following equation 2 for \( k = 1 \), we subtract the time-reversal of this event from \( p_0^+(x, t) \), which gives the modified incident wave field \( p_1^{-}(x, t) \), see Figure 6a. Using equation 3 we evaluate the reflection response \( p_1^{-}(x, t) \) to this modified incident wave field [Figure 6b]. Note that, within the time window, \( p_1^{-}(x, t) \) is identical to \( p_0^-(x, t) \) in Figure 5d, hence, further iterations will not cause any changes. Already after one iteration we have achieved the italicized condition mentioned above. This is a consequence of analyzing the simple configuration with two interfaces only. For more complex configurations more iterations will be required.

**Creating the virtual-source response**

After finalizing the iterative process, we define \( p(x, t) \) as the superposition of the final incident and reflected wave fields at \( z = 0 \). Because, for this example, iteration \( k = 1 \) was the final iteration, we have \( p(x, t) = p_1^-(x, t) + p_1^+(x, t) \). This total field is shown in Figure 7a. Within the time window this field is an asymmetric in time. Hence, if we superpose the total field and its time-reversed version, i.e., \( p_{sym}(x, t) = p(x, t) + p(x, -t) \), all events within the time window cancel each other, whereas outside the time window this superposition is symmetric, see Figure 7b. Note that, because we consider a lossless medium, \( p_{sym}(x, t) \) obeys the wave equation. Since time-reversal changes the propagation direction, it follows that the causal part is upward propagating at \( z = 0 \) and the acausal part is downward propagating at \( z = 0 \). The first arrival of the causal part of \( p_{sym}(x, t) \) in Figure 7b corresponds with the direct arrival of the response to the virtual source at \( x_{VS} \) [Figure 5b]. Given this last observation, combined with the fact that the causal part is upward propagating at \( z = 0 \) and that the total field obeys the wave equation and is symmetric, it is plausible that \( p_{sym}(x, t) \) is proportional to \( G(x, x_{VS}, t) + G(x, x_{VS}, -t) \) (where \( G \) stands for the Green’s function in the real medium). This argumentation holds for more general situations, but it has been checked explicitly.
with the stationary-phase method for the response in Figure 7b (Wapenaar et al., 2012). The causal part of $p_{\text{sym}}(x, t)$ in Figure 7b is the virtual-source response that we showed earlier in Figure 4b. It is interpreted as the Green’s function $G(x, x_{VS}, t)$. Note that the direct arrival in this retrieved Green’s function comes from the virtual source of the first arrivals (Figure 4a), whereas the internal multiples come entirely from the reflection response (Figure 1b).

**DECOMPOSITION**

The Green’s function $G(x, x_{VS}, t)$ is the response to the virtual source at $x_{VS}$. In Figure 4b it is clearly seen that this source radiates upward as well as downward propagating fields. The latter arrive at the surface after having been reflected at the interface below the source. We now discuss how the source can be decomposed into an upward and a downward radiating source. To this end we repeat the iterative process, but we replace the minus sign in equation 2 and Figure 6a by a plus sign, hence

$$q_{\text{up}}^+(x, t) = p_{\text{up}}^+(x, t) + w(x, t)q_{\text{sym}}^+(x, t),$$

$$q_{\text{down}}^+(x, t) = \int_{-\infty}^{+\infty} [R(x_{\text{up}}, x, t) + q_{\text{sym}}^+(x, t)]_{\tau=0}dx,$$

with $q_{\text{sym}}^+(x, -t)$ following again from equation 1. When this iterative process is finished (here again for $k = 1$), we define the total field as $q(x, t) = q_{\text{up}}^+(x, t) + q_{\text{sym}}^+(x, t)$, of which the part within the time window is now symmetric in time (instead of antisymmetric). Hence, by evaluating $p_{\text{sym}}(x, t) = q(x, t) - q(x, -t)$, all events within the time window again cancel each other (like in Figure 7b), whereas outside the time window this field is antisymmetric. The new created virtual source is therefore also antisymmetric (like a dipole). If we subtract this new response from our retrieved monopole response, i.e., if we evaluate $\frac{1}{2}[p_{\text{sym}}(x, t) - p_{\text{sym}}(x, t)]$ and take the causal part, we obtain the response to an upward radiating virtual source (Figure 8a), whereas the causal part of the superposition $\frac{1}{2}[p_{\text{sym}}(x, t) + p_{\text{sym}}(x, t)]$ yields the response to a downward radiating virtual source (Figure 8b).

**IMAGING WITHOUT GHOSTS**

Having obtained the correct downward and upward propagating fields at depth, we are now ready to use these fields for imaging. For a range of virtual receivers at $z_{VS}$, these fields are related via

$$G^+(x_{R}, x_{S}, t) = \int_{-\infty}^{+\infty} [R(x_{R}, x, t) * G^+(x, x_{S}, t)]_{\tau=0}dx,$$

for $x_{R}$ at $z_{VS}$. Here $R(x_{R}, x, t)$ is the reflection response at the virtual-source depth $z_{VS}$, for the situation of a homogeneous overburden. It can be resolved from $G^+(x, x_{S}, t)$ and $G^-(x, x_{S}, t)$ by multidimensional deconvolution (MDD), similar as in seismic interferometry by MDD. To this end, we correlate both sides of equation 6 with the downgoing field and integrate over the sources at the surface. This yields

$$C(x_{R}, x', t') = \int_{-\infty}^{+\infty} [R(x_{R}, x, t) * G^-(x, x_{S}, t)] dx,$$

for $x_{R}$ and $x'$ at $z_{VS}$, with the correlation function and point-spread function defined as

$$C(x_{R}, x', t) = \int_{-\infty}^{+\infty} G^+(x, x_{S}, t) * G^+(x', x_{S}, t) dx,$$

$$G^+(x, x_{S}, t) = \int_{-\infty}^{+\infty} G^+(x, x_{S}, t) * G^+(x', x_{S}, t) dx,$$

respectively (van der Neut et al., 2010; Wapenaar et al., 2011). Inverting equation 7 gives the reflection response $R(x_{R}, x, t)$ at depth $z_{VS}$. This is the ideal approach, which leads to true-amplitude angle-dependent reflection information. A much more simple approach is to evaluate the correlation function via equation 8 for zero-offset and zero time, i.e., $C(x', x', 0)$, and treat this as an estimate of the reflectivity for all image points $x'$ of interest. We used an intermediate approach to obtain the image of Figure 2a. We evaluated the correlation function $C(x_{R}, x', t)$ at three depth levels $z_{VS} = 600, 1320, 2150$ m, which are situated above, between and below the two reflectors, respectively, and we applied standard poststack migration to these responses to image the regions between these depth levels. The resulting image of Figure 2a is free of ghosts because we used the correct downgoing and upgoing fields (Figure 4c,d; Figure 2b) in equation 8.

**CONCLUSIONS**

We have discussed a new approach to integrated migration and internal multiple elimination. The method uses a downward extrapolation scheme that gives the correct downgoing and upgoing wave fields in the subsurface. Like standard migration, the method requires an estimate of the one-way operator, but no knowledge of the reflector is required; this information comes from the data itself. The method is stable with respect to errors in the macro model (Figure 3). Several other approaches exist that deal with the internal multiple problem (Weglein et al., 1997; Berkhout and Verschuur, 1997; Jakubowicz, 1998; Ten Kroode, 2002). It is beyond the scope of this paper to compare these with our method. We conclude by stating some properties of our proposed scheme. An essential aspect is that we use non-recursive one-way operators, hence, the method does not suffer from error propagation. As a matter of fact, it can start at any desired depth, or be used in a target-oriented approach. Another aspect is that no adaptive prediction and subtraction is required. Last, but not least, the internal multiples contribute to the restoration of the amplitudes of the primary reflections. The effects of triplications, head waves, diving waves, fine-layering, etc., need further investigation. A first numerical test with a variable-velocity syncline model shows promising results (Brockini et al., 2012).
Migration and multiple elimination

References


Pitfalls in constraining attenuation from ambient seismic noise
Cornelis Weemstra*, Spectraseis AG, ETH Zürich, Lapo Boschi, ETH Zürich, Roel Snelder, Colorado School of Mines, Alex Goertz, ETH Zürich

SUMMARY

We numerically investigate the ability of ambient noise surface wave interferometry to invert for anelastic attenuation. Shorter correlation time windows leads to higher signal-to-noise ratio of empirical Green’s functions (EGF’s). We show, however, that it is necessary to correct for the window length and that this is especially important if short time windows are employed. Furthermore, we show how source distribution varying with azimuth can severely affect the estimation of attenuation from ambient-noise surface waves.

INTRODUCTION

Claerbout (1968) derived a relation between the transmission response and the reflection response of a horizontally layered medium. He showed that the autocorrelation of the transmission response associated with a seismic noise source located vertically below the reflectors equals the reflection response plus its time-reversed version. He later generalized his theory to the crosscorrelation of noise recorded at two locations over a 3-D inhomogeneous subsurface. This is now known as passive seismic interferometry and its first successful application in seismology is due to Campillo and Paul (2003).

Since then, the field of passive seismic interferometry has taken a flight. In recent years, several researchers have focused on estimating attenuation based on interferometric measurements of regional scale surface waves. (Prieto et al., 2009; Lin et al., 2011; Lawrence and Prieto, 2011). Weemstra et al. (2012) show that the same methodology can be applied to local scales, i.e. ∼10-100 km. Robust estimates of seismic attenuation at this scale are needed for several reasons. Hydrocarbon reservoirs e.g., exhibit an abnormally strong attenuation contrast in oil and gas exploration. Furthermore, a good S-wave velocity and attenuation model of the shallow subsurface would help to improve S-wave static corrections needed in oil and gas exploration.

The stochastic nature of the ambient seismic wavefield supports time-averaging of crosscorrelations by cutting data into shorter time windows (Bensen et al., 2007; Seats et al., 2012). The crosscorrelations associated with these individual time windows are subsequently stacked. Seats et al. (2012) show that shorter time windows significantly improve the quality of the EGF which is beneficial for traveltime tomography studies and decreases the total required amount of data. However, care should be taken while applying this procedure in noise attenuation studies. We show that one should correct for the window length in order to prevent overestimation of attenuation values.

We model the wavefield with a simple two-dimensional ray-theory approach, which allows for intrinsic attenuation of energy and a dispersive wavefield. In this way, we are able to test the effect of source distribution, window length and subsurface attenuation on the retrieved ambient seismic noise attenuation. In general, attenuation of energy is due to the joint contribution of dissipation and scattering. We extend our model by adding isotropic point scatterers (Foldy, 1945; Groenenboom and Snieder, 1995). This approach enables us to selectively model the effect of scattering and intrinsic attenuation, which may lead to a better understanding of the relative importance of the respective phenomena. This will aid in the interpretation of attenuation values constrained from ambient seismic noise processing of real data.

THEORY

Given the recordings \( u(x_1) \) and \( u(x_2) \), captured at surface locations \( x_1 \) and \( x_2 \), the normalized time-domain cross-correlation \( C_{x_1x_2} \) is defined

\[
C_{x_1x_2}(t) \equiv \frac{1}{2T} \int_{-T}^{T} u(x_1, \tau) u(x_2, \tau + t) d\tau,
\]

where \( t \) is time, \( \tau \) is integration time and correlation time is given by \( 2T \).

Considering a single deterministic wave of angular frequency \( \omega \) and assuming the correlation time \( 2T \) to be sufficiently long with respect to the period of the waves, Tsai and Moschetti (2010) combine (through a simple integral) multiple sources acting at the same time. A uniform velocity and non-attenuating 2-D medium is assumed; velocity may vary with \( \omega \) though. They assume the sources to be uncorrelated in the sense that peaks due to crosscorrelation of signals generated by different sources, the so called “cross-terms”, cancel out through ensemble averaging. This assumption is equivalent to the assumption of “stationarity” by Aki (1957), because the stochastic behavior of the sources makes that these cross-terms do not interfere constructively. Similar assumptions have been made by other authors (Lobkis and Weaver, 2001; Snieder, 2004; Wapenaar, 2004). Through this procedure, Tsai and Moschetti (2010) show analytically that the real part of the azimuthally averaged cross-spectrum equals Aki’s normalized azimuthally averaged Spatial autocorrelation function (Aki, 1957),

\[
\Re[F[C_{x_1x_2}(\tau)]](\omega) = \rho(\omega, r) = J_0\left(\frac{\omega r}{c(\omega)}\right),
\]

where the operator \( \Re[...] \) maps its complex argument into its real part, \( F \) is the Fourier-transform operator, \( c(\omega) \) is the wave velocity as function of angular frequency and \( J_0 \) denotes the 0th order Bessel function of the first kind. The overbar over the left terms denotes azimuthal averaging and the left term in equation (2) is hence averaged over all couples for which
\[ |\mathbf{x}_1 - \mathbf{x}_2| = r. \] We dub \( \hat{\rho}(r, \omega) \) the “averaged complex coherency” and the (non-averaged) complex coherency is therefore given by \( \rho(r, \omega) \).

In this form, however, equation (2) does not account for the effects of attenuation. Prieto et al. (2009) propose to do so “empirically” through multiplication by an exponential factor \( e^{-\alpha(\omega)r} \),

\[ \rho(\omega, r) = J_0 \left( \frac{\omega r}{c(\omega)} \right) e^{-\alpha(\omega)r}, \quad (3) \]

The attenuation coefficient \( \alpha(\omega) \) accommodates a more rapid decrease of the Bessel function with interstation distance \( r \), representing the effect of energy dissipation and scattering. Expression (3) will be referred to as the “damped Bessel function” henceforth.

A ray-theoretical framework that describes how amplitudes of the complex coherency depend on source distribution has been established by Tsai (2011). He illustrates that the source distribution significantly impacts the behavior of \( \rho(r) \). His results indicate that equation (3) is only valid for a set of sources that is uniformly distributed with respect to \( r \). Throughout this paper we will assume a uniform radial distribution of sources and only vary the azimuthal distribution of sources. Tsai (2011) also shows that a non-uniform distribution of sources with azimuth may result in decay of amplitude with distance which could be misinterpreted as attenuation associated with subsurface characteristics.

The wavefield at \( \mathbf{r} \) due to a source at \( \mathbf{r}_j \) is described by the impulse response, which for the two-dimensional scalar wave equation is given by:

\[ G^{(0)}(\mathbf{r}, \mathbf{r}_j) = \frac{-i}{4} H_0^{(1)}(k|\mathbf{r} - \mathbf{r}_j|) \quad (4) \]

where \( H_0^{(1)} \) is a Hankel function of the first kind and order \( l \) and \( k = |\mathbf{k}| \), i.e. the absolute value of the wave vector \( \mathbf{k} \), representing the dispersion relation of the medium with \( k = \omega/c(\omega) \). Intrinsic attenuation can be included in the model through multiplication of expression (4) with \( e^{-\alpha(\omega)r} \).

**NUMERICAL SETUP**

A total number of 60000 sources are randomly distributed throughout a circle with a radius of 70000 m (Figure 1a). We place 600 receivers along three lines (200 receivers per line) and receivers positions are incremented by 100 m in both directions away from the center (Figure 1b). \( \rho(\omega, r) \) is calculated for receivers equidistant from the center for each line individually and hence the wavefield is sampled at interstation distances of 200, 400, ..., 20000 m (Receivers with similar color coding in Figure 1b). We neglect the cross-terms through crosscorrelation of signal emitted by each source separately and subsequent stacking; hence \( \rho(\omega, r) \) is a sum of 60000 crosscorrelations. Averaging over azimuth gives us \( \hat{\rho}(\omega, r) \), which is a sum of 180,000 crosscorrelations: each due to a different randomly placed source. This azimuthal averaging is only employed to triple the number of realizations as each relocation of a receiver couple is essentially similar to placing another 60000 sources randomly on the circle.

![Figure 1: The two-dimensional experiment setup is given by a, with a radius of the circle of 70 km. Black lines denote receiver lines. A zoom in on the very center of a is shown in b, where colored diamonds represent receivers. The blue stars in both a and b denote the randomly placed sources.](image)

We prescribe a constant velocity \( c \) of 750 m/s and the intrinsic attenuation is set to zero. The source spectrum is defined between 0.15 and 0.45 Hz, but we shall only analyze one single frequency. We crosscorrelate windows of 30, 60 and 90 seconds and evaluate \( \hat{\rho}(\omega, r) \) for a frequency of 0.20 Hz. Evaluation involves fitting a damped Bessel function to \( \hat{\rho}(\omega, r) \) by minimizing the L1-norm on the differences between the two. This misfit function therefore is a function of \( c \) and \( \alpha \). The results are shown in Figure 2 for the three respective window lengths.

**EFFECTS OF WINDOW LENGTH**

Although an attenuation coefficient of zero is prescribed, it is clear for all three cases that the decay of \( \hat{\rho}(r) \) with distance is stronger than a non-damped \( J_0 \). For windows of 30 s, its amplitude is even close to zero for large interstation distances (light green dots in Figure 2a). The window length appears to have a profound effect on the amplitude of \( \hat{\rho}(r) \) for longer...
distances. We define energy traveling between a receiver pair that is captured by both receivers as “Coherent energy”. We argue here that the loss of coherency with distance of \( \rho(r) \) is a function of the ratio of coherent energy and the total energy of signal traveling in a similar direction. In order to test this, we define empirically a correction factor, dubbed \( C(r) \), which would correct for such an effect. \( C(r) \) is a function of the window length \( 2T \) and the traveltimes \( \Delta t \) and \( \Delta t' \) according to:

\[
C(r) = \frac{1}{1 - (\Delta t/2T)}
\]  

(5)

Correcting the values of \( \rho(r) \) for the traveltimes associated with the distance between a receiver pair yields values denoted by the light blue dots in Figure 2. The pure Bessel function and the corrected values of \( \rho(r) \) clearly coincide in Figure 2b and 2c. The coincidence in Figure 2a is less exact, which might be due to the very low portion of coherent energy compared by the crosscorrelation process. For receiver pairs as far as 20000 meters apart, the fraction of coherent energy recorded by both receivers is only \([30 - (20000/750)]/30 \approx 0.11\) while the noise introduced by the finite windowing length is notably high as the requirement \( T \gg 1/\omega \) is less valid for shorter time windows (Tsai and Moschetti, 2010).

The apparently simple correction factor can have serious implications if not accounted for. In case of short window lengths (i.e. \( \Delta t/2T \) has a significant non-zero value) the decay of \( \tilde{\rho}(r) \) could be mistakenly interpreted as loss of energy due to the medium. Such misinterpretation is demonstrated by our example in Figure 2: significant estimates for \( \alpha \) are obtained fitting a damped Bessel function to \( \tilde{\rho}(r) \) while no attenuation is prescribed. Furthermore, the resolved phase velocities are also slightly modified due to the effect of short window lengths, i.e. higher estimates than the prescribed 750 m/s are obtained.

Figure 3 shows that the model doesn’t accurately fit \( \rho(r) \). Furthermore, it should be understood that if one corrects for the window length by multiplying the estimates of \( \tilde{\rho}(r) \) with \( C(r) \), the error on \( \tilde{\rho}(r) \) will also be increased by a factor \( C \). This can have serious implications in real-data applications, where, contrary to our investigation here, the cross-terms cannot be neglected and therefore significant error on the obtained \( \rho(r) \) may be found (see e.g. Lawrence and Prieto, 2011; Weemstra et al., 2012). We conclude that either sufficiently long time windows or a correction factor as in equation (5) should be employed to overcome erroneous interpretation of attenuation values from damped Bessel function fitting.

AZIMUTHALLY VARYING SOURCE DISTRIBUTIONS

Tsai (2011) derives the effect of an arbitrary azimuthal distribution of far-field surface waves on \( \rho(r) \). He does so, writing the square of the azimuthal source distribution as a Fourier series and calculating the real part of the complex coherency by integrating over the source distribution. He finds that only even numbered coefficients of the Fourier series contribute to \( \rho(r) \).

We show how different, cosine distributed, azimuthal source distributions affect the behavior of \( \rho(r) \). This means that we, similarly to Tsai (2011), don’t average station pairs over azimuth, but evaluate the complex coherency along one individual line. The source distribution is varied with azimuth according to the probability density functions shown in Figures 4a and 4c. The two configurations shown in Figures 4b and 4d correspond to the blue graphs in 4a and 4c, respectively. These two distribution are symmetric with respect to an azimuth of 0 degrees, while the red, green and yellow distributions are rotated 30, 90 and 120 degrees clockwise, respectively. We calculate crosscorrelations along one line of receivers and correct

Figure 2: Variation of \( \bar{\rho}(r) \) with distance for a frequency of 0.20 Hz employing window lengths of 30 s (a), 60 s (b) and 90 s (c). The light green dots are the real parts of \( \bar{\rho}(r) \), while the imaginary parts are denoted by the red dots. The best fitting damped Bessel function is given by the solid black line, while the corresponding pure Bessel function is denoted by the dashed black line. Corrected values of \( \bar{\rho}(r) \) are given by the light blue dots. Attenuation coefficients (a) and phase velocities (c) associated with best fitting damped Bessel function are given in the top right corner. The two black boxes in a are shown enlarged in Figure 3.
Low frequency passive seismic interferometry

Figure 3: Blow-ups of the two boxed regions in Figure 2a. For explanation of the symbols and lines see caption of Figure 2.

Figure 4: Probability density functions of the single (a) and double (c) cosine distributed azimuthal source distributions. Colors represent different (clockwise) rotations of the distributions. Examples of the blue single and double cosine distributed azimuthal source distributions are shown in b and d, respectively. The radius of the circle is 70 km and a single, fixed line of receivers is shown in black.

The resulting $\rho(r)$ are shown in Figure 5a and 5b for the distributions in Figure 4a and 4c, respectively. We see that in case of a single cosine distribution the observed $\rho(r)$ is coinciding with a pure Bessel function associated with a velocity of 750 m/s and a frequency of 0.20 Hz (Figure 5a), i.e. our prescribed values. This is true for each of the four single cosine distributions and therefore a rotation of the source distribution doesn’t have any effect on obtained complex coherency values. The behavior of the imaginary parts of the cross-spectra does change as a function of rotation of the source distribution.

However, when we prescribe a double cosine distribution with azimuth, the behavior of $\rho(r)$ is significantly affected (Figure 5b). None of the four source distributions yields a complex coherency that fits the non-damped Bessel function. Source distributions with a higher density of sources on either ends of the receiver line than perpendicular to it, give complex coherency amplitudes that are even higher than the amplitude of the pure Bessel function (green and yellow dots in Figure 5b).

The same phenomenon is observed by Tsai (2011) after derivation of $\rho(r)$ for similar azimuthal distributions of sources, but a distribution of sources with $r$ that places all sources in the far-field. Sufficient averaging over azimuth would however make the different positive and negative contributions to $\rho(r)$ cancel each other and hence converges towards a pure Bessel function.

CONCLUSIONS

The crosscorrelation window length affects the decay of the averaged complex coherency with distance. An empirically defined function allows for correction of this phenomenon and therefore ensures that amplitude decay due to short window lengths is not erroneously attributed to the medium. Given a uniform distribution of sources with radius, a varying distribution of sources with azimuth affects the complex coherency.

ACKNOWLEDGEMENTS

We greatly acknowledge support from the QUEST Initial Training Network funded within the EU Marie Curie actions programme.
Low frequency passive seismic interferometry

REFERENCES


Illumination compensation for subsalt image-domain wavefield tomography
Tongning Yang*, Center for Wave Phenomena, Colorado School of Mines,
Jeffrey Shragge, School of Earth and Environment, University of Western Australia,
and Paul Sava, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

Wavefield tomography represents a family of velocity model building techniques based on seismic waveforms as the input and seismic wavefields as the information carrier. For wavefield tomography implemented in the image domain, the objective function is designed to optimize the coherency of reflections in extended common-image gathers. This function applies a penalty operator to the gathers, thus highlighting image inaccuracies due to the velocity model error. Uneven illumination is a common problem for complex geological regions, such as subsalt. Imbalanced illumination results in defocusing in common-image gathers regardless of the velocity model accuracy. This additional defocusing violates the wavefield tomography assumption stating that the migrated images are perfectly focused in the case of the correct model and degrades the model reconstruction. We address this problem by incorporating the illumination effects into the penalty operator such that only the defocusing due to model errors is used for model construction. This method improves the robustness and effectiveness of wavefield tomography applied in areas characterized by poor illumination. The Sigsbee synthetic example demonstrates that velocity models are more accurately reconstructed by our method using the illumination compensation, leading to more coherent and better focused subsurface images than those obtained by the conventional approach without illumination compensation.

INTRODUCTION

Building an accurate and reliable velocity model remains one of the biggest challenges in current seismic imaging practice. In regions characterized by complex subsurface structure, prestack wave-equation depth migration, (e.g., one-way wave-equation migration or reverse-time migration), is a powerful tool for accurately imaging the earth’s interior (Gray et al., 2001; Etgen et al., 2009). The widespread use of these advanced imaging techniques drives the need for high-quality velocity models because these methods are very sensitive to model errors (Symes, 2008; Woodward et al., 2008; Virieux and Operto, 2009).

Wavefield tomography represents a family of techniques for velocity model building using seismic wavefields (Tarantola, 1984; Woodward, 1992; Pratt, 1999; Sirgue and Pratt, 2004; Vigh and Starr, 2008; Plessix, 2009). The core of wavefield tomography is using a wave equation (typically constant density acoustic) to simulate wavefields as the information carrier. Wavefield tomography is usually implemented in the data domain by adjusting the velocity model such that simulated and recorded data match (Tarantola, 1984; Pratt, 1999). This match is based on the strong assumption that the wave equation used for simulation is consistent with the physics of the earth. However, this is unlikely to be the case when the earth is characterized by strong (poro)elasticity. Significant effort is often directed toward removing the components of the recorded data that are inconsistent with the assumptions used.

Wavefield tomography can also be implemented in the image domain rather than in the data domain. Instead of minimizing the data misfit, the techniques in this category update the velocity model by optimizing the image quality (Yilmaz, 2001). Differential semblance optimization (DSO) is one realization of image-domain wavefield tomography (Symes and Carazzo, 1991). The essence of the method is to minimize the difference between same reflection observed at neighboring offsets or angles (Shen and Calandra, 2005; Shen and Symes, 2008). DSO implemented using space-lag gathers constructs a penalty operator which annihilates the energy at zero lag and enhances the energy at nonzero lags (Shen et al., 2003). This construction assumes that migrated images are perfectly focused at zero lag when the model is correct. This assumption, however, is violated in practice when the subsurface illumination is uneven. In complex subsurface regions, such as subsalt, uneven illumination is a general problem and it deteriorates the quality of imaging and velocity model building (Leveille et al., 2011). Several approaches have been proposed for illumination compensation of imaging (Gherasim et al., 2010; Shen et al., 2011), but not for velocity model building. In this paper, we address the problem of uneven illumination associated with image-domain wavefield tomography. The main idea is to include the illumination information in the penalty operator used by the objective function such that the defocusing due the illumination is excluded from the model updating process. We illustrate our technique with a subsalt velocity model.

THEORY

The core element for image-domain wavefield tomography using space-lag extended images (subsurface-offset CIGs) is an objective function and its gradient computed using the adjoint-state method (Plessix, 2006; Symes, 2009). The state variables relate the objective function to the model parameter and are defined as source and receiver wavefields \( u_s \) and \( u_r \) obtained by solving the following acoustic wave equation:

\[
\begin{bmatrix}
\mathcal{L} (x, \omega, m) \\
0
\end{bmatrix}
\begin{bmatrix}
0 \\
\mathcal{L}^* (x, \omega, m)
\end{bmatrix}
= \begin{bmatrix}
u_s (j, x, \omega) \\
u_r (j, x, \omega)
\end{bmatrix} = \begin{bmatrix} f_s (j, x, \omega) \\
fr (j, x, \omega)
\end{bmatrix},
\]

where \( \mathcal{L} \) and \( \mathcal{L}^* \) are forward and adjoint frequency-domain wave operators, \( f_s \) and \( fr \) are the source and record data, \( j = 1, 2, \ldots, N_s \) where \( N_s \) is the number of shots, \( \omega \) is the angular frequency, and \( x \) are the space coordinates \{x, y, z\}. The wave operator \( \mathcal{L} \) and its adjoint \( \mathcal{L}^* \) propagate the wavefields
Illumination compensation tomography

forward and backward in time respectively using a two-way wave equation, e.g., \( L = -\omega^2 m - \Delta \), where \( m \) represent slowness squared. The objective function for image-domain wavefield tomography measures the image incoherence caused by the model errors

\[
\mathcal{H}_d = \frac{1}{2} ||K_f(x) P(M) r(x, \lambda)||_{X, \lambda}^2 ,
\]

where \( r(x, \lambda) \) are extended images:

\[
r(x, \lambda) = \sum_j \sum_\omega u_s(j, x - \lambda, \omega) u_r(j, x + \lambda, \omega),
\]

The online represents complex conjugate, and the lag vector has, in this case, only horizontal components: \( \lambda = \{ \lambda_x, \lambda_y, 0 \} \).

The mask operator \( K_f(x) \) limits the construction of the gatherers to select locations in the subsurface. \( P(M) \) is a penalty operator acting on the extended image to highlight defocusing, i.e. image inaccuracy. It is typically assumed that defocusing is only due to velocity error, an assumption which leads to the differential semblance optimization (DSO) operator of Symes (2008). However, this penalty operator is not effective when poor illumination affects the image accuracy and leads to additional defocusing.

To alleviate the negative influence of poor illumination, we need to include the illumination distribution in the tomographic procedure. Illumination can be assessed by applying illumination analysis, which is formulated based on the migration deconvolution, given by the expression

\[
\tilde{r}(x) = (M^* M)^{-1} r(x),
\]

where \( \tilde{r} \) is a reflectivity distribution, \( r \) is a migrated image, \( M \) is a demigration operator which is linear with respect to the reflectivity. This operator is different from the modeling operator \( L \). The adjoint \( M^* \) represents the migration operator. \( (M^* M)^{-1} \) is a blurring operator, and represents the Hessian (second-order derivative of the operator with respect to the model) for the operator \( M \). This term includes the subsurface illumination information associated with the velocity structure and the acquisition geometry. In practice, the full \((M^* M)^{-1}\) matrix is too costly to construct, but we can evaluate its impact by applying a cascade of demigration and migration \((M^* M)\) to a reference image. For example, using extended images, we can write:

\[
r_e(x, \lambda) = M^* M r(x, \lambda).
\]

The resulting image \( r_e \) approximates the diagonal elements of the Hessian and captures defocusing associated with illumination effects. Such defocusing is the consequence of uneven illumination and should not be used in the velocity update. Therefore, an illumination-based penalty operator can be constructed as

\[
P(x, \lambda) = \frac{1}{E[r_e(x, \lambda)] + \epsilon},
\]

where \( E \) represents envelope and \( \epsilon \) is a damping factor used to stabilize the division. Replacing the conventional penalty \( P = |\lambda| \) with the one in equation 6 is the basis for our illumination compensated image-domain wavefield tomography. Note that the DSO penalty operator is a special case of our new penalty operator and corresponds to the case of perfect subsurface illumination and wide-band data.

The adjoint sources are computed as the derivatives of the objective function \( \mathcal{H}_d \) shown in equation 2 with respect to the state variables \( u_s \) and \( u_r \):

\[
\begin{align*}
\mathcal{L}^* (x, \omega, m) & \quad 0 \\
\mathcal{L} (x, \omega, m) & \quad \alpha_s(x, \omega) \\
\end{align*}
\]

where

\[
\mathcal{L}^* (x, \omega, m) = \sum_\lambda P(M) K_f(x) \tilde{K}_f(x) P(M) r(x, \lambda) u_r(j, x + \lambda, \omega)
\]

and the gradient is the correlation between state variables and adjoint state variables:

\[
\frac{\partial \mathcal{H}_d}{\partial m} = \sum_j \sum_\omega \frac{\partial \mathcal{L}}{\partial m} \left( u_s(j, x, \omega) \alpha_s(j, x, \omega) + u_r(j, x, \omega) \alpha_r(j, x, \omega) \right)
\]

The model is then updated using gradient line search aimed at minimizing the objective function given by equation 2.

EXAMPLES

In this section, we use the Sigsbee 2A model (Paffenholz et al., 2002) and concentrate on the subsalt area to test our method in regions of complex geology with poor illumination. The target area ranges from \( x = 6.5 \) to \( 20 \) km, and from \( z = 4.5 \) to \( 9 \) km. The model, migrated image, and angle-domain gatherers for correct and initial models are shown in Figures 1(a)-1(c) and Figures 2(a)-2(c), respectively. The angle gatherers are displayed at selected locations corresponding to the vertical bars overlain in Figure 1(b). The actual spacing of the gatherers and penalty operators are 0.45 km. Note that the reflections in the angle gatherers appear only at positive angles, as the data are simulated for towed streamers and the subsurface is illuminated from one side only. We run the inversion using both the conventional DSO penalty and the illumination-based penalty operators. The DSO penalty operator is shown in Figure 3(a). For the illumination-based operator, we first generate gathers containing defocusing due to illumination effects (Figure 3(b)), and then we construct the penalty operator using equation 6, as shown in Figure 3(c). For the gatherers characterizing the illumination effects, we can observe significant defocusing in the subsalt area, as the salt distorts the wavefields used for imaging and causes the poor illumination.

We run both inversions for 10 iterations, and obtain the reconstructed model, migrated image, and angle-domain gath-
Illumination compensation tomography

Figure 1: (a) The true model in the target area of the Sigsbee model. (b) The corresponding migrated image, and (c) the angle-domain gathers.

Figure 2: (a) The initial model obtained by scaling the subsalt sediments of the true model. (b) The corresponding migrated image, and (c) the angle-domain gathers.

Figure 3: (a) The DSO penalty operator. (b) The gathers characterizing the illumination effects. (c) The illumination-based penalty operator constructed from the gathers in Figure 9(b).

The figures show that we update the models in the correct direction in both cases and that the reconstructed models are closer to the true model than the starting model. We find, however, that the model obtained using the illumination-based penalty is closer to the true model than the model obtained using DSO penalty. The model obtained using DSO penalty is not sufficiently updated and is still too slow. This is because the severe defocusing due to the salt biases the inversion when we do not take into account the uneven illumination. The comparison of the images also suggests that the inversion using the illumination-based penalty is superior to the inversion using DSO penalty. Both images are improved due to the updated model, as illustrated by the better focused diffractions distributed at $z = 7.6$ km, and by the faults located between $x = 14.0$ km, $z = 6.0$ km and $x = 16.0$ km, $z = 9.0$ km which are more visible in the images. If we concentrate on the bottom reflector (around 9 km), we can distinguish the extent of the improvements on the image quality for both inversions. The bottom reflector is corrected to the right depth for inversion using the illumination-based penalty, while for inversion using the DSO penalty, this reflector is still misplaced from the correct depth and it is not as flat as the reflector in Figure 5(b). Figures 6(a)-6(d) compare the angle gathers at $x = 10.2$ km for the correct, initial, and reconstructed models using DSO and illumination-based penalties. The gathers for both reconstructed models show flatter reflections, indicating that the reconstructed models are more accurate than the
Illumination compensation tomography

Figure 4: (a) The reconstructed model from inversion using the DSO penalty. (b) The corresponding migrated image, and (c) the angle-domain gathers.

Figure 5: (a) The reconstructed model from inversion using the illumination-based penalty. (b) The corresponding migrated image, and (c) the angle-domain gathers.

Figure 6: Angle-domain gathers at $x = 10.2$ km for (a) the correct model, (b) the initial model, the reconstructed models using (c) the DSO penalty, and (d) the illumination-based penalty.

CONCLUSIONS

We demonstrate an illumination compensation strategy for wavefield tomography in the image domain. The idea is to measure the illumination effects on space-lag extended images, and replace the conventional DSO penalty operator with another one that compensates for illumination. This workflow isolates the defocusing caused by the illumination such that image-domain wavefield tomography minimizes only the defocusing related to velocity error. The synthetic examples show the improvements of the inversion result and of the migrated image after the illumination information is included in the penalty operator. Our approach enhances the robustness and effectiveness of wavefield tomography in the model building process when the subsurface illumination is uneven due to complex geologic structures such as salt. The cost of this technique is higher than that of the conventional approach since we periodically need to re-evaluate the subsurface illumination.

ACKNOWLEDGMENTS

We acknowledge the support of the sponsors of the Center for Wave Phenomena at Colorado School of Mines. The reproducible numeric examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org.
REFERENCES


