Iterative multiparameter elastic waveform inversion using prestack Kirchhoff approximation

Hassan Khaniani, John C. Bancroft and Eric von Lunen

ABSTRACT

Elastic Full Waveform Inversion (FWI) is an iterative method that simultaneously uses traveltime and amplitude of the seismic data to recover subsurface elastic properties. To date, despite the advancements in mathematical aspects of FWI, this method has not found much application in commercial seismic data processing, mainly because of its computation cost. Conventional elastic FWI methods require a depth imaging algorithm for forward modeling (e.g., Finite Difference Time Domain (FDTD)) and a depth migration for the inversion (e.g., Reverse Time Migration (RTM)).

Our aim is to propose a revised "standard strategy" for the inversion of elastic properties from the linearized reflected elastic waves. We use the direct relationship between the scattering potential of the Born approximation with the reflectivity function of the asymptotic Kirchhoff approximation. To estimate the gradient function of algorithm, we implemented the direct inversion strategy of Beylkin and Burridge (1990) to obtain an iterative well-known standard AVO inversion. Both of the forward and inverse operators use prestack time imaging methods that map the migrated P-to-S traveltime to P-to-P traveltime. We obtain two registered volumes in a pseudo-depth eliminating the need of ray tracing for registration issues. For complex structures, one can add ray tracing to the algorithm.

INTRODUCTION

Seismic waveforms contain traveltime and amplitude information of a scattered wavefield that can be used to determine physical properties of the subsurface. Perturbation theory (e.g., Cohen and Bleistein, 1979, Beylkin, 1985, Bleistein et al. 1985) plays an important mathematical role in studying and understanding the scattered waves. In essence, it finds the relationship between the difference between recorded wavefield and a modelled wavefield, called the data residual, and the difference between their elastic properties, called perturbed model. Once the relationship is established, it performs the inversion of the perturbation from data residuals. In order to find a unique solution of the perturbed model, this method requires that the perturbed model be small. The inverted perturbed model is added to the initial background model to approximate the true solution.

In the framework of perturbation theory, several studies have linearized the nonlinear acoustic wave equation by Born approximation (e.g., Morse and Feshbach, 1953) for the direct inversion of elastic properties. For example, Cohen and Bleistein (1979) used the Born approximation for direct velocity inversion of a migrated section (poststack data). Clayton and Stolt (1981) used Born and WKBJ approximations of prestack F-K migration (Stolt, 1978) for the direct inversion of density and bulk modulus. Beylkin (1985) presented a mathematical framework for migration/inversion of acoustic waves. His Born approximation forward problem was modified by Bleistein (1986) by including

the Kirchhoff approximation the ray-theoretic reflection coefficient of Bleistein (1984). Bleistein (1986) showed that the amplitude of his inversion is the same as the angle-dependent geometric-optic reflection coefficient. Parson (1986) assumed that output of Bleistein's (1986) inversion is equivalent to the linearized reflection coefficient of Aki and Richard (1980) and showed the direct inversion of the relative change of elastic properties with a P-to-P survey. A major problem with a direct inversion method is that the inversion operators require reasonable *a priori* knowledge of the medium to produce accurate results (Tarantola 1984a). This creates a need to iteratively update the inversion operators from the results of the previous steps.

Several authors addressed the iterative approach by the technique of nonlinear least squares (mostly l^2 norm) for the inverse problem. In an acoustic Born approximation, Tarantola (1984) and later LeBras and Clayton (1988) showed that the inverse problem for linearized reflection data (e.g., multiple-free) can be solved using iterative classical Kirchhoff migration and forward modeling. Tarantola (1984b) provided an inversion strategy by iterative use of consistent migration and forward modeling. He did not specify the exact type of forward modeling and migration, but he showed that the inversion of a nonlinear acoustic (full) wavefield can be done using Reverse Time Migration (RTM) and two-way acoustic modeling. Similar strategies were later presented for inversion of the elastic two-way wave approximation (Tarantola, 1984c, Tarantola, 1986 and Mora, 1988). Gauthier el. al (1986) applied the numerical method of Tarantola (1984b) by FDTD operators on a synthetic data set. The procedure of FWI by FDTD operators has been in the subject of much research; unfortunately, for nearly three decades, FWI-based FDTD has had a computational time burden that prevented its "standard" application.

Besides FDTD, there were attempts to improve computational efficiency by implementing cheaper numerical schemes for forward modeling and migrations. Beydoun and Mendez (1989) and Jin et al., (1990) combined ray tracing and the Born approximation amplitude radiation patterns of elastic waves for the multi-parameter inversion of elastic properties. The technique was then improved by Thierry et al. (1999) and Lambare et al. (2003) in terms of computational costs and resolution. Operto et al., (2003) applied the method to a synthetic 3D survey on an SEG/EAGE overthrust acoustic model for velocity inversion.

In this study, we first describe the mathematics of the forward problem and the conventional inversion scheme which is based on the Born approximation (e.g., Beylkin and Burridge, 1990, Tarantola 1984a, Tarantola (1986) and Mora, 1988). We then propose the implementation of prestack time Kirchhoff-based methods for both the forward problem and inversion of linearized elastic reflection data. The framework of the inversion is similar to that of conventional FWI. The main difference is that we use the asymptotic Kirchhoff approximation. Hence, the perturbation of the model from data residual is found by the direct relationships between the scattering potential of the Born approximation with the linearized Zoeppritz approximation of Aki and Richards (1980). These relationships already have been derived by several authors; we apply them as suitable approximations for our problem. For more details on derivations, the reader is referred to Beylkin and Burridge (1985), Jaramillo and Bleistein (1999), Kroode (2013) and Shaw and Sen (2004).

In the Kirchhoff-based approach (Bleistein,1987), the migrated data residuals produce residual reflectivity functions when the migration operators use true amplitudes (e.g., accurate weight functions) and also remove the source signatures. In the form of steepest descent inversion, the integration of the residual reflectivity on the vectors perpendicular to the reflectors (e.g., normal rays) produces a scaled gradient of elastic properties which is then used in the frame of conventional FWI for updating the initial model. The approach is a multi-parameter elastic inversion scheme that is applicable to all types of reflected wavefields such as P-P, P-S, S-S and S-P; however, we only show the results of the inversion of P- and S- wave velocity from the P-P and P-S wavefields respectably.

The techniques required for conditioning the data are explained using a field data example. The nature of the technique is intended to be a self-updating iterative scheme. For inversion of the synthetic data (with the amplitudes under the control of modeling and migration), we can expect the algorithm to automatically converge to the true model. Field data typically have a lower SNR compared with the synthetic data. In this case, the challenge of SNR is controlled using the calibration of the amplitude and phase of the migrated true data with the modelled data obtained from well log information. So, instead of a self-updating procedure, at each step, a visual panel is created to determine the step length.

FORWARD PROBLEM FOR SCATTERED ELASTIC WAVEFORM

Consider an isotropic medium as shown in Figure (1), with true elastic properties $M_{true}(\lambda_{true}, \mu_{true}, \rho_{true})$, that can be decomposed into an initial model $M_0(\lambda_0, \mu_0, \rho_0)$ and a perturbation $\Delta M(\Delta\lambda, \Delta\mu, \Delta\rho)$. The components of $M_{true}(\lambda_{true}, \mu_{true}, \rho_{true})$ are given by

$$\begin{aligned} \lambda_{true}(x) &= \lambda_0(x) + \Delta\lambda(x), \\ \mu_{true}(x) &= \mu_0(x) + \Delta\mu(x), \\ \rho_{true}(x) &= \rho_0(x) + \Delta\rho(x). \end{aligned} \tag{1}$$

where $\rho(x)$ is density, and $\lambda(x)$ and $\mu(x)$ are Lame`s parameters as defined by Aki and Richards, (1980). The subscript `true` refers to the true model to be found, the subscript `0` refers to the initial model to be updated. Initially it has the low frequency component of the true model and is assumed to be close to the true model. The variable x is the subsurface coordinate. The terms $\Delta\rho$, $\Delta\lambda$ and $\Delta\mu$ are the perturbations of the elastic properties which contain the higher frequencies of the true model, and in our problem are assumed to generate the perturbation of the total scattered elastic wavefield $\Delta U(s,x)$ from the medium. The variable s is the source location. It is assumed that the traveltime of the wavefield traveling through the initial model is close to the true model. Using the Born approximation, Beylkin and Burridge (1990) derived the single scattered wavefield components of the P-P wave $\Delta U_{jk}^{PP}(s,r,t)$, the mode converted P-S wave $\Delta U_{jk}^{PS}(s,r,t)$, S-P wave $\Delta U_{jk}^{SP}(s,r,t)$ and the S-S wave $\Delta U_{jk}^{SS}(s,r,t)$ using a high frequency (ray based) approximation solution of wave equation. The total scattered wavefield is given by

$$\Delta U_{jk}(s,r,t) = \Delta U_{jk}^{PP}(s,r,t) + \Delta U_{jk}^{PS}(s,r,t) + \Delta U_{jk}^{SP}(s,r,t) + \Delta U_{jk}^{SS}(s,r,t), \qquad (2)$$

where the subscripts jk are the displacement in the $k-$direction due to a point force in the $j-$direction at point s.

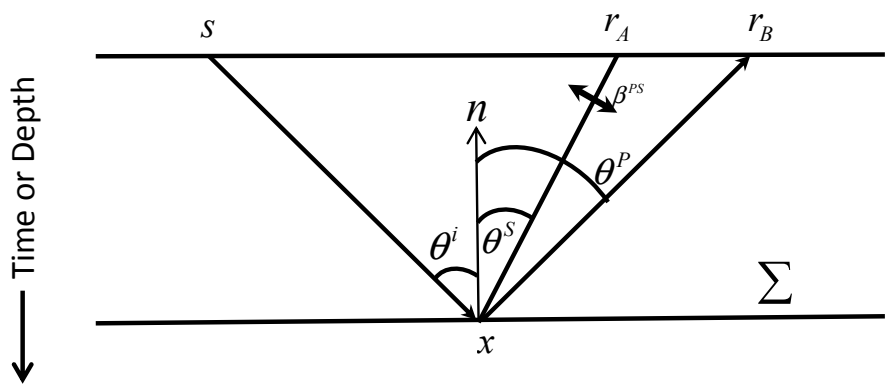


FIG. 1: Adaptation of Beylkin and Burridge's (1990) elastic Born approximation to the Kirchhoff approximation. In the Kirchhoff approximation, the elastic model parameters are thought to be independently decomposed to a homogenous model M_0 and the perturbed model ΔM across the interface.

In 3D models, ΔU_{jk} consists of 9 components which are the result of combination of a 3-component acquisition source with 3-component receivers. These 9-component are the resultant of 36-component wavefield combination on the right side of equation (2). Although the inversion approach presented here is applicable to all wavefields in equation (2), due to the low SNR of the amplitude information observed in the field data and the acquisition source type deployed, in this study the wavefields $\Delta U_{jk}^{SP}(s,r,t)$ and $\Delta U_{jk}^{SS}(s,r,t)$ have been ignored. In addition for analysing P-to-P and P-to-S data with an 18- component resultant wavefield, we consider two configuration of the vertical component source for vertical (for P-to-P wavefield) and radial components (for P-to-S wavefield). The term $\Delta U_{jk}^{PP}(s,r,t)$ is defined by

$$\Delta U_{jk}^{PP}(s,r,t) = -\partial_t^2 \int S^{PP} A_j^P A_k^P \delta\left(t - \phi^P - \phi^P\right) dx, \qquad (3)$$

and the term $\Delta U_{jk}^{PS}(s,r,t)$ is defined by

$$\Delta U_{jk}^{PS}(s,r,t) = -\partial_t^2 \int_D S^{PS} A_j^P A_{kl}^S \beta_l^{PS} \delta\left(t - \phi^P - \phi^S\right) dx, \qquad (4)$$

where the factor β is the polarization unit vector of the shear wavefield that reaches to the surface (Figure 1).

The derived angle-dependent scattering potentials S^{PP} and S^{PS} of Beylkin and Burridge (1990) are in terms of Lame's constants shown in equation (1) and are given by

$$S^{PP}(\Delta\lambda,\Delta\mu,\Delta\rho)=\rho^0\left(\frac{\Delta\lambda}{\lambda+2\mu}+\frac{\Delta\rho}{\rho}\cos\theta^{PP}+\frac{2\Delta\mu}{\lambda+2\mu}\cos^2\theta^{PP}\right), \tag{5}$$

and

$$S^{PS}(\Delta\mu,\Delta\rho)=\rho^0\left(\frac{\Delta\rho}{\rho}\sin\theta^{PS}+\frac{\Delta\mu}{\mu}\frac{v_S}{v_P}\sin 2\theta^{PS}\right). \tag{6}$$

As shown in Figure (1), the angles θ^{PP} and θ^{PS} are defined by $\theta^{PP}=\theta^i+\theta^P$ and $\theta^{ps}=\theta^i+\phi^s$, where the terms θ^i, θ^P and θ^S are the incident and reflected P- and S-waves angles, respectively. For an isotropic model, one can change the variables of the scattering potentials (5) and (6) to the perturbation of P-wave velocity v_P and S- velocity v_S by combining Snell's law with the differential operator Δ in the following relationships,

$$\Delta\lambda=\Delta(\rho v_P{}^2-2\rho v_S{}^2), \tag{7}$$

$$\Delta\mu=\Delta(\rho v_S{}^2), \tag{8}$$

to get

$$S^{PP}(\Delta v_P,\Delta v_S,\Delta\rho)=\rho\left(2\frac{\Delta v_P}{v_P}-4k^2\sin^2\theta^{PP}\frac{\Delta v_S}{v_S}+\left(2\cos^2\theta^P-4k^2\sin^2\theta^P\cos^2\theta^P\right)\frac{\Delta\rho}{\rho}\right), \tag{9}$$

and

$$S^{PS}(\Delta v_s,\Delta\rho)=\rho\sin\theta^{PS}\left(\begin{array}{l}\left(4k\cos\theta^P\cos\theta^S-4k^2\sin^2\theta^P\right)\frac{\Delta v_S}{v_S}\ldots\\ \qquad+\left(1+2k\cos\theta^P\cos\theta^S-2k^2\sin^2\theta^P\right)\frac{\Delta\rho}{\rho}\end{array}\right), \tag{10}$$

where $k=v_S/v_P$. (See e.g., Mora, 1987 and Jin et al., 1992)

The source signatures are absent here, but in practice appropriate phase corrections and deconvolution schemes can be employed to remove the effects of the source. Available well log information can assist in removing the source signatures and aid in enhancing the true amplitude terms (A^P and A^S) by calibration of the forward models with the true data. More details about these steps are provided in our numerical field data example.

THE INVERSE PROBLEM FOR A SCATTERED ELASTIC WAVEFIELD

The forward problem (3) and (4) can be written in a more compact notation as

$$\Delta U_{jk}^{pp} = f^{pp} \Delta M \tag{11}$$

and

$$\Delta U_{jk}^{ps} = f^{ps} \Delta M \tag{12}$$

where f is the corresponding forward operator. Beylkin and Burridge (1990) derived the adjoint operator f^* for the direct inversion of scattering potential S^{PP} and S^{PS} by considering the forward operator f of the integral function in the equations (11) and (12) as a generalized Radon formula (see e.g., equations 5.6 and 5.16 of Baylkin and Burridge, 1990). In principle, the result of direct inversion can be enhanced if the forward modeling and inversion of the Born approximation formulations of (11) and (12) are adapted in an iterative scheme. The iterative approach uses a generalized l^2 norm to minimize an objective function J defined by

$$J = \left\| \Delta U_{ij} \right\|^2, \tag{13}$$

where, ΔU_{ij} is the scattered field defined in equations (11) and (12). The inverse problem typically uses a steepest descent algorithm with the following steps:

$$\gamma_k = -\frac{\partial J}{\partial M} = F^T(\Delta U_k), \tag{14}$$

$$\Delta M_k = \alpha_k \gamma_k, \tag{15}$$

and

$$M_{k+1} = M_k + \Delta M_k, \tag{16}$$

with γ_k and α_k being the gradient function and step length to minimize (13). The operator F is the derivative of f at the point M (i.e., $F = \frac{\partial f}{\partial M}$) which satisfies the linearized forward problem defined by

$$f(M_0 + \Delta M) \approx f(M_0) + F\Delta M. \tag{17}$$

Note that the small perturbation of the model parameters facilitates the linearization of the forward problem using a Taylor series expansion. In seismic problems, the nonlinear inversion algorithms of the scattered wavefield include two common steps:

1. Forward modeling, to minimize the norm of data residuals or the missing diffracted field ΔU (i.e, $\Delta U = f(\Delta M)$).

2. Inversion, using migration operator on the data residuals. (i.e., $\gamma = F^*(\Delta U)$).

Equations (3) and (4) for f and F^* describes the first order Born approximation of the isotropic homogenous elastic wave equation. In the time domain, to include higher order

terms to this approximation, one may implement two wave equation operators such as FDTD and RTM for f and F^* respectively. Although Tarantola (1984) did not specify the types of f and F^*, most authors used FDTD and RTM, and a few authors used ray based approaches for designing f or F^* (or both of them). The use of ray tracing alleviated the computational costs for f and F^* compared to time-stepping methods; however, both of the above steps are still computationally expensive if the operators use depth imaging algorithms. To date, this has been one of the major obstacles for the method to be applicable commercially.

SOLUTION OF THE INVERSE PROBLEM USING THE KIRCHOFF APPROXIMATION

Our approach further reduces the computational time for the iterative scheme of (13) to (16), but with the operators for both f and F^* being conventional PSTM. We show that for both of the f and F^* operators the Born approximation is replaced by Kirchhoff approximation, as they are asymptotically similar with a smooth error. Detailed mathematical proofs of this similarity are discussed by Beylkin and Burridge (1990), Jaramillo and Bleistein (1999), Shaw and Sen (2004) and Kroode (2012); we will not repeat them here. The main difference between the Born and Kirchhoff approximations is in the shape of the scattering and reflectivity function. The scattering potentials S^{PP} and S^{PS} of the Born approximation are derived from perturbation of the elastic properties ΔM that has Heaviside-type singularity. If the singularity of ΔM is defined by a delta function in a coordinate orthogonal to the reflector $\gamma_R(x)$, we have

$$\frac{\partial}{n}\Delta M(x) = \Delta M(x)\gamma_R(x), \tag{18}$$

which after replacing in equations (3) and (4) transforms to the surface integral in the Kirchhoff approximation from the Born approximation.

Ursin and Tygel (1997) provides a more physical interpretation of the Kirchhoff surface scattering integral by applying the divergence theorem to the Born scattering volume integral. As shown in Figure (2), they define a surface Σ along which the medium M varies steeply (e.g., same as in equation (18)). So, the gradient of surface Σ is parallel to the gradient direction of M. Now, instead of equation (1) (i.e., $M_{true} = M_0 + \Delta M$) we have

$$M_{true} = \begin{cases} M_0 + \Delta M^-, & above\ \Sigma \\ M_0 + \Delta M^+, & below\ \Sigma \end{cases}, \tag{19}$$

where M_0 is a function in D, and $\Delta M^{\mp}$ are smooth functions (perturbations) above and below Σ, respectably. As shown in Figure (2), if the integration domain be divided into two upper and lower hemispherical volumes D^+ and D^- respectably, such that $D = D^- + D^+$, then we can decompose the total volume integrals of equations 3 or 4 over D to two volume integrals over D^- and D^+

$$\Delta U = \int_D S(x)e^{i\omega\phi}dx = \int_{D^-} S(x)e^{i\omega\phi}dx + \int_{D^+} S(x)e^{i\omega\phi}dx. \tag{20}$$

The asymptotic approximation of Bleistein (1984, equation 2.8.2) for the Fourier integral shows how to decompose a volume integral into a surface over ∂D and a volume integral over D as

$$\Delta U = \int_D S(x)e^{i\omega\phi}dx = \frac{1}{i\omega}\left\{\int_{\partial D} S(x)\frac{n.\nabla\phi(x)}{|\nabla\phi(x)|^2}e^{i\omega\phi}d\sigma - \int_D \nabla.\left[\frac{S(x)\nabla\phi(x)}{|\nabla\phi(x)|^2}\right]e^{i\omega\phi}dx\right\}. \tag{21}$$

For the case of volume integral, the supporting scattering is inside the volume, and because of Sommerfield radiation conditions the contribution of the integration over the surface ∂D is small compared to the total volume D. For the volumes D^- and D^+ this scenario is different; surface integrals are dominant contributors. This is because the surface Σ (which is the main contributor to both region integrals) belongs to the surface ∂D^- and ∂D^+. For the rest of ∂D^- and ∂D^+ the Sommerfield radiation condition applies and we have

$$I^- = -i\omega\int_\Sigma S^-(x)\frac{n.\nabla\phi(x)}{|\nabla\phi(x)|^2}e^{i\omega\phi}d\sigma, \tag{22}$$

and

$$I^+ = i\omega\int_\Sigma S^+(x)\frac{n.\nabla\phi(x)}{|\nabla\phi(x)|^2}e^{i\omega\phi}d\sigma. \tag{23}$$

Now from $I = I^- + I^+$ and using the definition of $S(x)$ (e.g., see Beylkin and Burridge 1990) we have

$$S(x) = [\frac{\Delta\rho^+(x) - \Delta\rho^-(x)}{\rho_0(x)}\delta_{ik} + \frac{\Delta c^+_{ijkl}(x) - \Delta c^-_{ijkl}(x)}{\Delta c_{ijkl_0}(x)}p^r_j p^s_l]h^s_k h^r_i, \tag{24}$$

where p^r and p^s are slowness vectors for the reflected (scattered) and incident wavefields, respectively. The terms h^r and h^s are polarization vectors for the reflected (scattered) and incident from source wavefield respectively. The Einstein summation convention is used here.

We emphasize that equation (24) is a general description of reflectivity (scattering potential) for various types of waves. Coupling the definition of the stiffness c_{ijkl} for the type of model (e.g., elastic, anisotropic, poroelastic, etc), the slowness as well as the polarization vectors of the incident and reflected waves have been the basis for modeling and inversion of the wavefields. We refer the reader to the work of Shaw and Sen (2004) for examples of setting up the problem for isotropic and anisotropic models.

In addition, equation (24) describes the main difference between the results of Born approximation modeling and inversion. As previously stated the radiation patterns of

scattering potential is step-like while the radiation pattern of Kirchhoff approximation is spike-like resulting from singularities over the surface Σ (i.e., steep gradient of elastic properties).

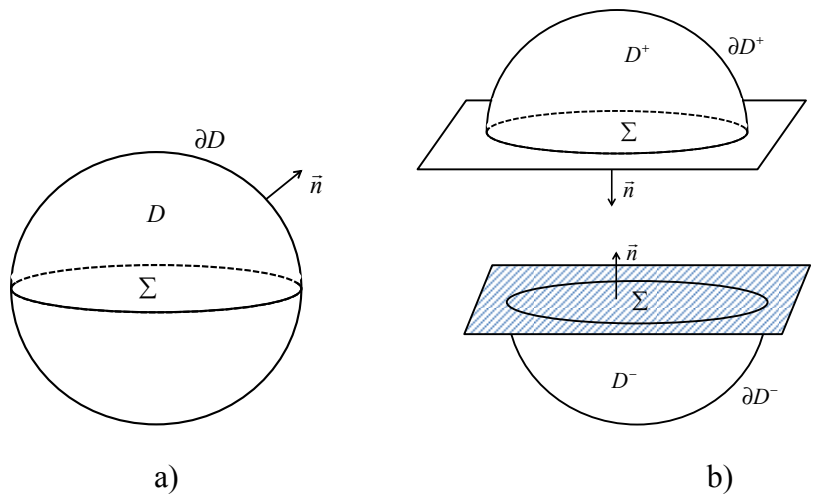


FIG. 2: The relationship between the Born approximation (volume integral) and the Kirchhoff approximation (surface integral). Adapted from Bleistein et al. (2001).

For the iterative scheme of equations (13) to (16), we use a standard Kirchhoff approximation with Zoeppritz reflectivity functions (Aki and Richards, 1980) for forward modeling and migration operators. Note that the Zoeppritz reflectivity functions of P-to-P and P-to-S waves; R^{PP} and R^{PS} are related to equations (9) and (10) by (Shaw and Sen, 2004 and Ursin and Tygel , 1997)

$$R^{PP} = \frac{1}{4\rho_0 \cos^2 \theta^P} \frac{\partial}{\partial n} S^{PP}, \tag{25}$$

and

$$R^{PS} = \frac{\tan \theta^S}{2\rho_0 \sin(\theta^{PS})} \frac{\partial}{\partial n} S^{PS}. \tag{26}$$

The missing diffracted field in equations (3) and (4) can be approximated by R^{PP} and R^{PS} and the angle-dependent weight functions of (25) and (26). Equations (25) and (26) show that the final output of the inversion scheme by PSDM Kirchhoff operators should give similar result to the use of the Born l^2 ray-based methods (Beydoun and Mendez; 1989). For smoothly varying elastic properties, we successfully reduced the computational cost of the ray tracing operators for both f and F^* by traveltime approximation techniques of PSTM.

In practical schemes of using Kirchhoff operators, the optimum weight functions K can be assigned to account for several amplitude factors such as attenuation, reflection, transmission and geometrical spreading factor.

NUMERICAL IMPLEMENTATION FOR GRADIENT ESTIMATION

Figure (3) shows a geology model where the elastic properties vary differently in depth. The source and receivers are on the surface. The migrated shot record demonstrates the different radiation patterns and reflectivity of scattered P-to-P and P-to-S wavefields. For all individual shot records, each scatter point response is migrated to its P-to-P time (See Figure 4). By gathering all migrated shot data, we are able to sort all individual scatter point responses to form a common image gather. Since the algorithm is based on the common traveltime of scatter points, the traveltime acts as a pseudo-depth and removes the need for ray tracing. For complex structures one may add the ray tracing to the algorithm.

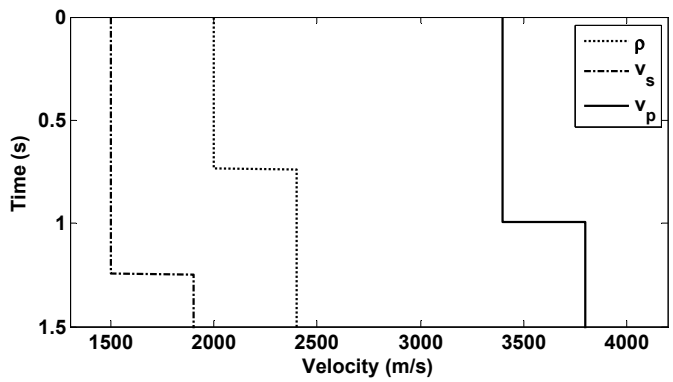


FIG. 3: Geological model having changes in elastic properties at three different depths. The unit for density is kg/m^3.

The multiparameter gradient function is obtained in a similar way to conventional AVO inversion. For each scatter point, two systems of equations are formed for P-to-P (to solve for three unknowns Δv_P, Δv_S and ρ) and for P-to-S (to solve for two unknowns Δv_S and ρ). An example of this operator at 0.75 s is shown in Figure (5). It shows that the operator for the inversion is dependent on the knowledge of the velocity field at far offsets. In addition, note the similarity of inversion operators of Δv_P and $\Delta\rho$ for P-to-P wavefield and Δv_S and $\Delta\rho$ for smaller angle of incident. This creates deficiency in rank of inversion matrix. For this problem, we need to identify the density field in the real data to enhance the gradient functions.

The inversion of P-to-P and P-to-S data can be done separately or simultaneously; however, simultaneous inversion is more stable because the results can be combined and compared with well log data. Note that the our term "simultaneous inversion" refers to the traveltime; it is different from other authors's terminology, which assumes a relationships between v_P, v_s and ρ to reduce equations (29) and (30) to two equations and two unknowns. Examples of their strategy varies from Garner's relationship (Smith and Gidlow; 1987, Stewart, 1990) or a constant v_P to v_s ratio. Without loss of generality, we can implement them in the frame of FWI for gradient calculations.

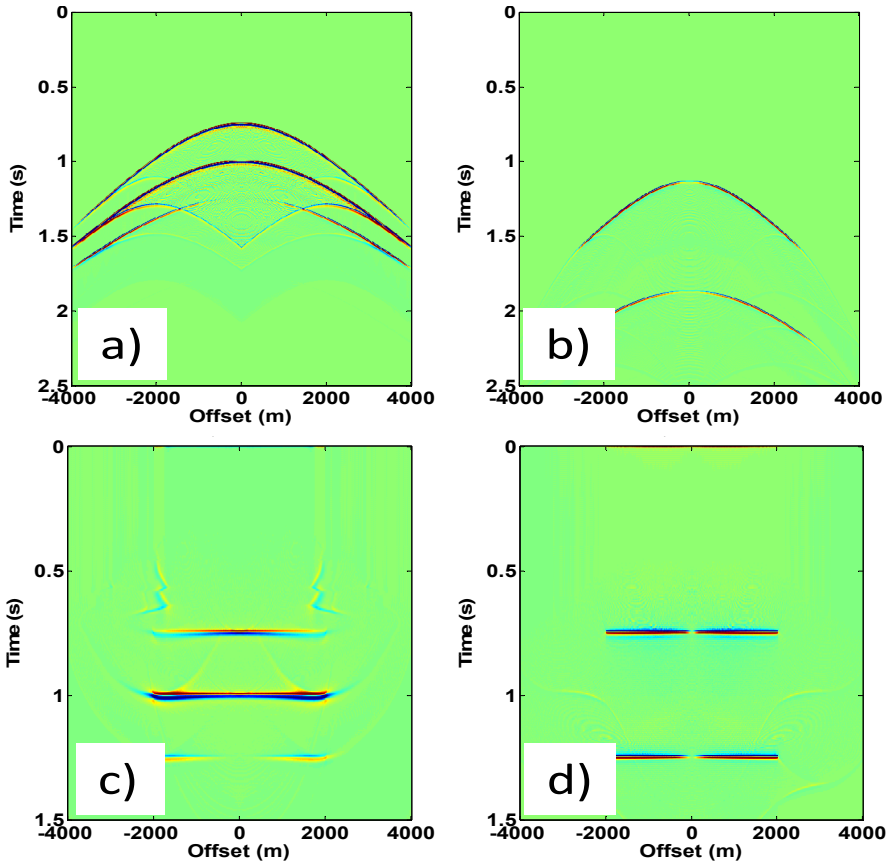


FIG. 4: Radiation patterns and reflectivity of a scattered wavefield recorded at surface. The model is shown in Figure (3) and the offset corresponds to the source position. a) The vertical component of a shot record of P-to-P data. b) The horizontal component of a shot record of P-to-S data c) The migration of (a) with true amplitude correction d) The migration of (b) with true amplitude correction.

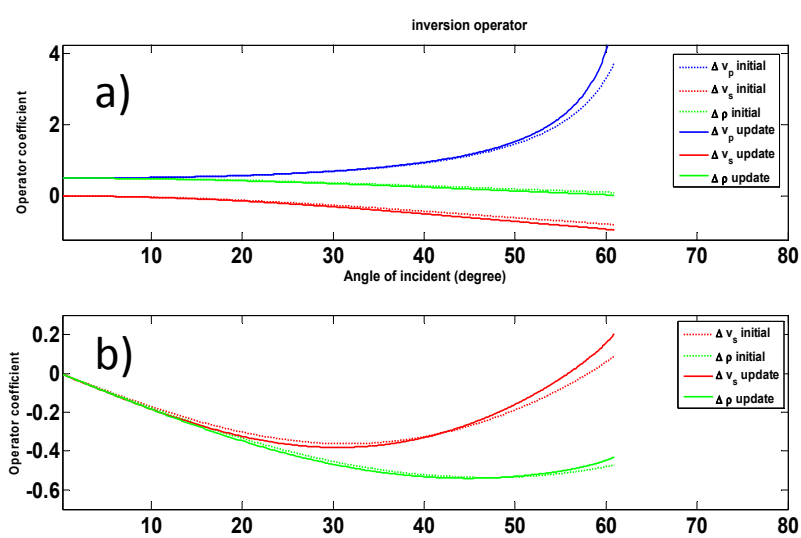


FIG. 5: A sample scatter point: the angle-dependent matrix operator coefficient (linearized AVO matrix) based on Aki and Richard (1980) for: a) P-to-P radiated waves b) P-to-S radiated waves. Note that the legend shows that dotted lines are initial models.

DESCRIPTION OF THE INVERSION PROCESS

To demonstrate a our form of FWI technique we use a simple numerical model. Figure (6) show a numerical estimation of the reflectivity of P-to-P and P-to-S data from a single shot in a 1-D geological model that has an upper horizontal layer with $v_p = 3000$ m/s, $v_s = 1500$ m/s and $\rho = 2000$ kg/m3 and a lower horizontal layer that has $v_p = 4000$ m/s, $v_s = 2000$ m/s and $\rho = 2300$ kg/m3. For demonstration purposes let us assume that the final output of the migration is a delta function with the amplitude of the reflectivity. Figures (6a & 6c) show the migrated data residual between the true data reflectivity of the P-to-P and P-to-S waves (R_{pp}^{true} and R_{ps}^{true}) located at the true depth z^{true} (or true time τ^{true}) and the modeled reflectivity R_{pp}^{m} and R_{ps}^{m} located at initial depth z^{m} (or initial time τ^{m}). The angle-dependent elastic parameters for calculation of the data residual from the Zoeppritz equations can be written as

$$\mathrm{I}_{pp}^{residual}(z) = R_{pp}^{true}\delta(z-z') - R_{pp}^{\mathrm{m}}(\Delta v_p)\delta(z-z') - R_{pp}^{\mathrm{m}}(\Delta v_s)\delta(z-z') - R_{pp}^{\mathrm{m}}(\Delta\rho)\delta(z-z'), \quad (27)$$

which is defined for P-to-P waves. In the shot record, for normal incident P-waves, the contribution of Δv_s is neglected. Similarly, for the P-S wave at an arbitrary non-normal angle of incidence $\theta = 20°$ we have

$$\mathrm{I}_{ps}^{residual}(z) = R_{ps}^{true}\delta(z-z') - R_{ps}^{\mathrm{m}}(\Delta v_s)\delta(z-z') - R_{ps}^{\mathrm{m}}(\Delta\rho)\delta(z-z'), \quad (28)$$

where for $\mathrm{R}_{\mathrm{ps}}^{\mathrm{m}}$, the reflectivity has been scaled to a non-normal angle of incidence ($\theta^{ps} = 20^{\circ}$). An integration of the data residual (Figure 6b & 6d) is equivalent to the gradient function of equation (14) for all parameters

$$\Delta M(v_p, v_s, \rho) = \Delta v_p(z) + \Delta v_s(z) + \Delta\rho(z) = \alpha^{pp}\int \mathrm{I}_{pp}^{residual}(z)dn, \quad (29)$$

and

$$\Delta M(v_s, \rho) = \Delta v_s(z) + \Delta\rho(z) = \alpha^{ps}\int \mathrm{I}_{ps}^{residual}(z)dn, \quad (30)$$

where for a flat reflector n is the z (or τ) direction. Note that z and τ are interchangeable using depth-to-time (or time-to-depth) conversion techniques (e.g., Hubral, 1997). The normal ray n for the horizontal reflectors are in the direction of the time τ or depth z. Equations (29) and (30) show that if the images inside the integration are defined in τ (e.g., PSTM outputs), then the ΔM and M_0 can be defined in τ. This means that during the inversion process with equations (21), (22), (25) and (26), there is no need to transform the model to the depth z. Kirchhoff approximation provides the opportunity to estimate the gradient function for different elastic properties.

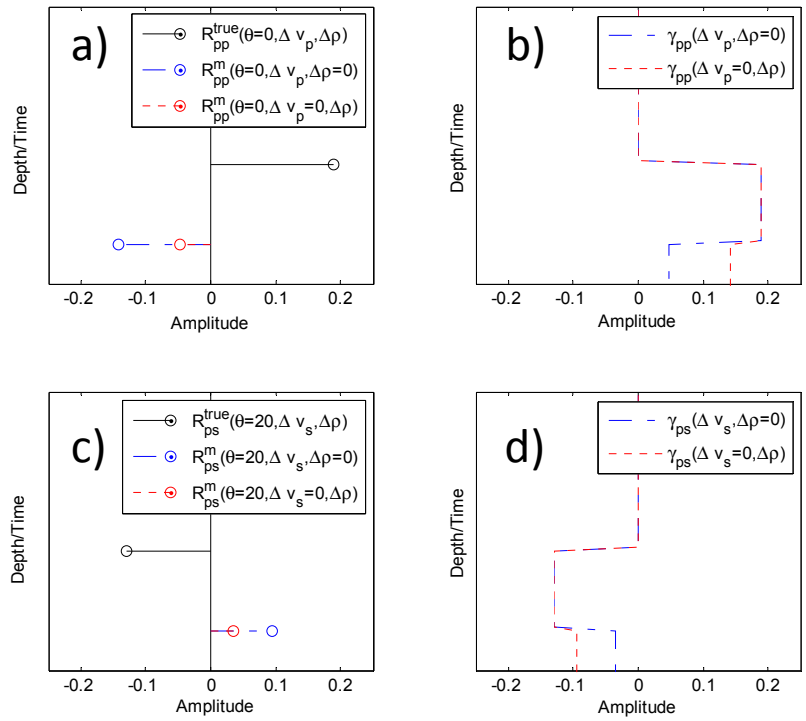


FIG. 6: A numerical example of FWI and the contribution of elastic properties to the gradient calculation. a) The true amplitude migrated P-to-P data residual after removal of the source signature (assuming full band reflectivity is a spike). b) Integration of Figure (6a) to give the gradient function of P-to-P data. c) The true amplitude migrated P-to-S data residual after removal of the source signature (assuming full band reflectivity is a spike). d) Integration of Figure (6c) to give the gradient function of P-S data.

THE EFFECTS OF DATA NOISE AND BANDLIMITATION ON FWI

Figure (7) demonstrates the effects of low SNR and limitation if frequency bands of the signal; two common problems found in real data. The blue graph in Figure (7a) is a data residual signal to be added to the noise shown by the green line. The integration of the noisy residual is shown by the dotted black line. Tarantola (1984) suggested using appropriate smoothing operators to improve the inversion result. For our available real field P-to-S data that has low SNR, we migrated all the data and used visual comparisons to determine step length. Further discussion on this issue is found in our field data example. We also note that the reflectivity functions of field data are band-limited. The band-limited reflectivity functions produce artefacts in the inversion. The lack of zero frequency causes the integral output of equations (29) and (30) not to produce an accurate gradient function, which should be a perfect step function as in Figures (7b). For more examples of artefacts of band limitation, the reader is referred to Bleistein et al. (2001) , Innanen (2011) and Kroode (2013).

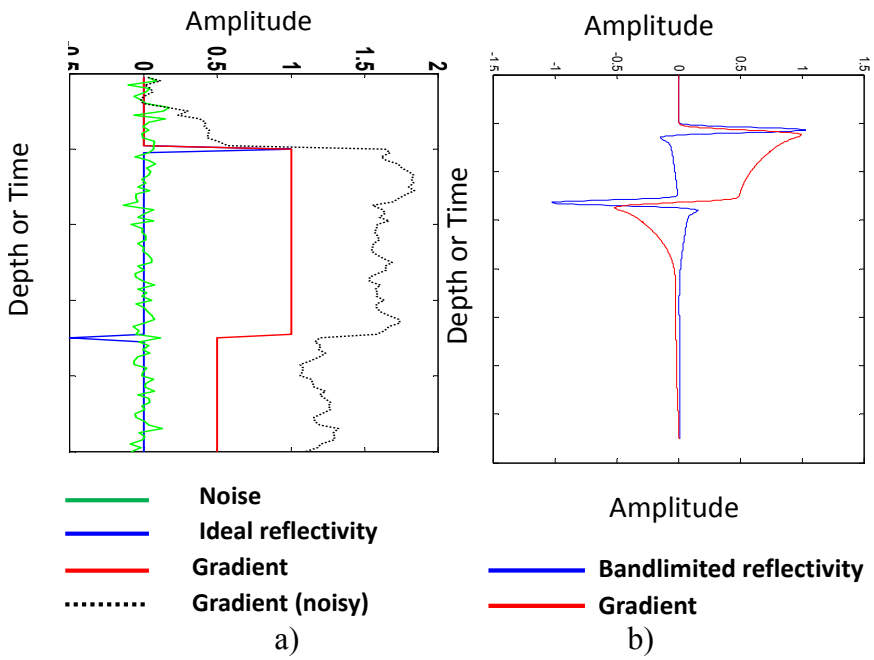


FIG. 7: Numerical considerations of FWI. a) The influence of random noise on the gradient function. b) The influence of band limitation (low cut 5Hz filtered data residual) on the gradient function. Note that the ideal gradient function is expected to be similar to the gradient in (a).

THE INITIAL MODEL AND DATA CONDITIONING

As in the inversion strategy proposed by Tarantola (1986), we can first invert for the more contributing model M parameters such as Δv_p and Δv_s, then invert for $\Delta\rho$. The starting v_p and v_s can be obtained by conventional velocity analysis such as prestack migration gathers or available well log information. The updated model is used to update the forward modeling operator f and its adjoint operators f^* to improve the traveltime in each step. This is done using the relationship between the updated interval velocity to RMS velocity (Dix, 1955). The data conditioning of the full wavefield examples were studied by several authors (e.g., Sears et al., 2008 and Warner et. al., 2013). However, in time imaging techniques, the procedure of waveform inversion is slightly different because it is based on the linearization of seismic reflection data (primary reflection data). Therefore, in real field data we will have to eliminate noise and preserve true amplitude in order to minimize the objective function defined by equation (13). Noise in this case is any signal that is not a reflected wave and is not included in the forward modeling. Examples of noise are multiples, surface waves, surface related noise, and dead traces. The effects of the source wavelet should be removed from data residuals by appropriate deconvolution methods prior to calculating the gradient function. As previously stated improvements in the bandwidth of the signal can improve gradient functions.

SYNTHETIC NUMERICAL EXAMPLE

As shown in Figure (8), a synthetic numerical model from an available well log from NEBC (Figure 8a) is created. The Kirchhoff operators for the forward modeling produced a single shot record, with a maximum of 4000m offset in split spread configuration with a receiver spacing of 12.5m. A minimum phase wavelet with a dominant frequency of 45 Hz was arbitrarily chosen. For demonstration purpose we assume that *a priori* information is available by smoothing the true model to construct G for inversion. Making use of the radiation pattern of scatter points shown in Figure (8), two least squares engines are implemented in each step for P-to-P and P-to-S data. This leads to convergence using l^2 norm convergence criterion after a few iterations. Figure (9) compares the differentiated true model (convolved with wavelet and scaled) with the inverted reflectivity. As shown in Figures (7 & 8), our control on the traveltime and amplitude of synthetic waves means there is no difficulty getting high correlation match of modeled and observed data l^2 norm as well as a high convergence rate toward true model.

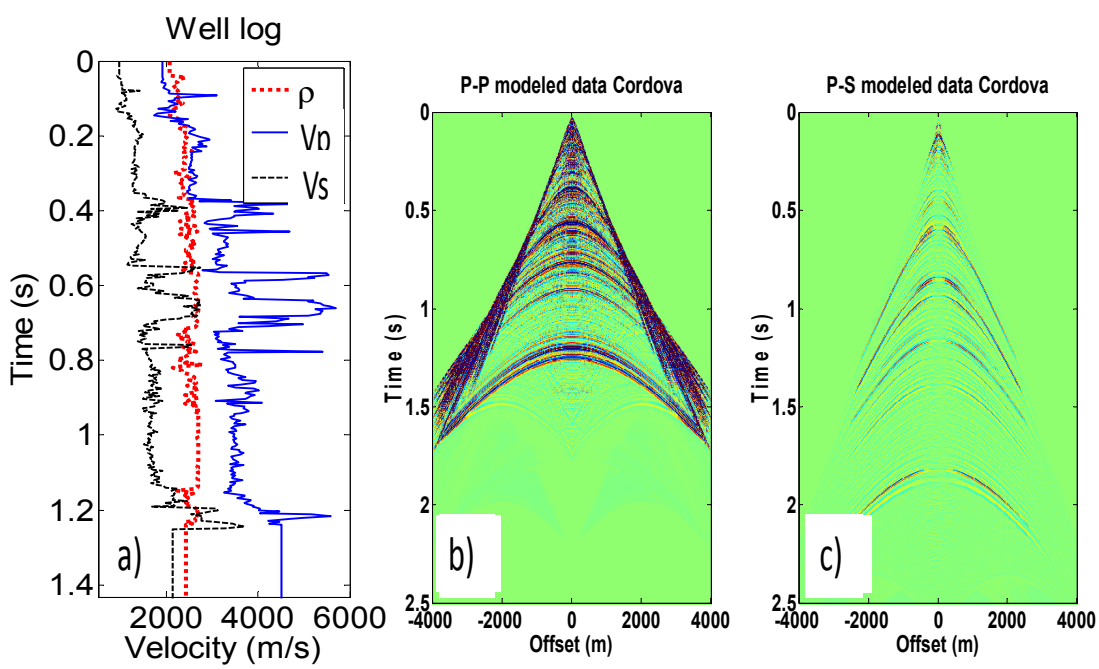


FIG. 8: Radiation patterns and reflectivity of scattered wavefield. The algorithm is same as Figure (4) but the model is obtained from a real well log data. a) Well log data. b) The vertical component of shot record of P-to-P data c) The horizontal component of shot record of P-to-S data

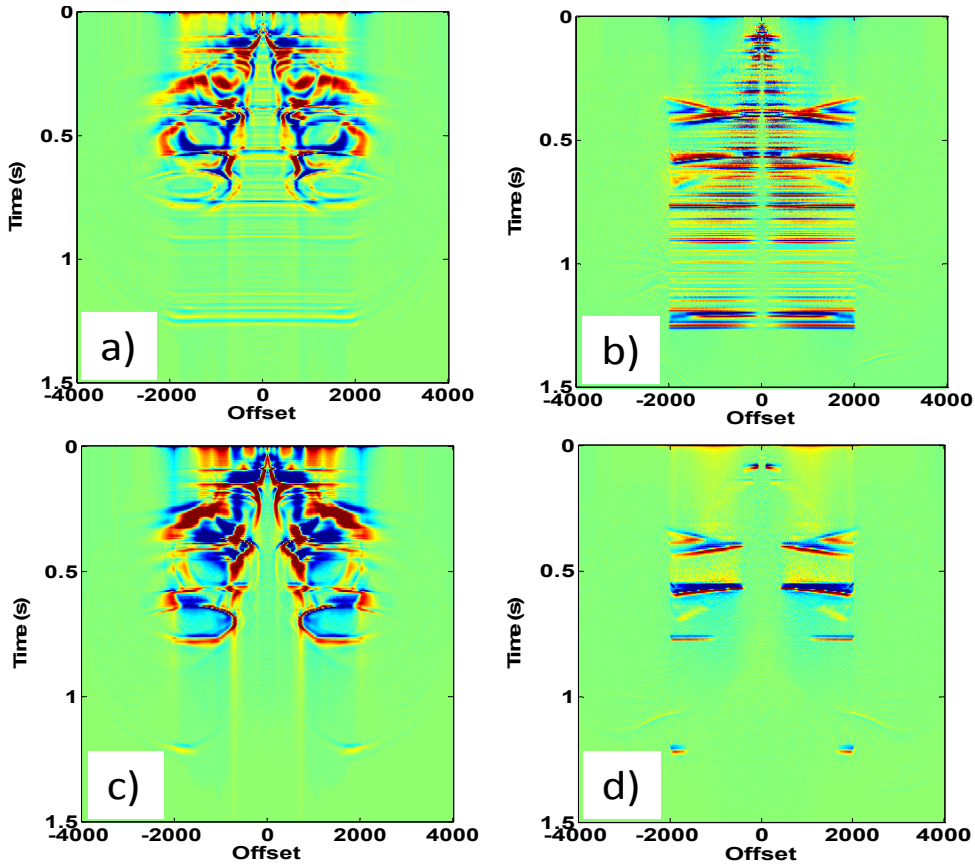


FIG. 9: Radiation patterns of scattered wavefield on prestack gathers. c) the migration of real part of Figure (8b). d) The migration of real part of Figure (8c). b) The migration of imaginary part of Figure (8b). d) The migration of imaginary part of Figure (8c).

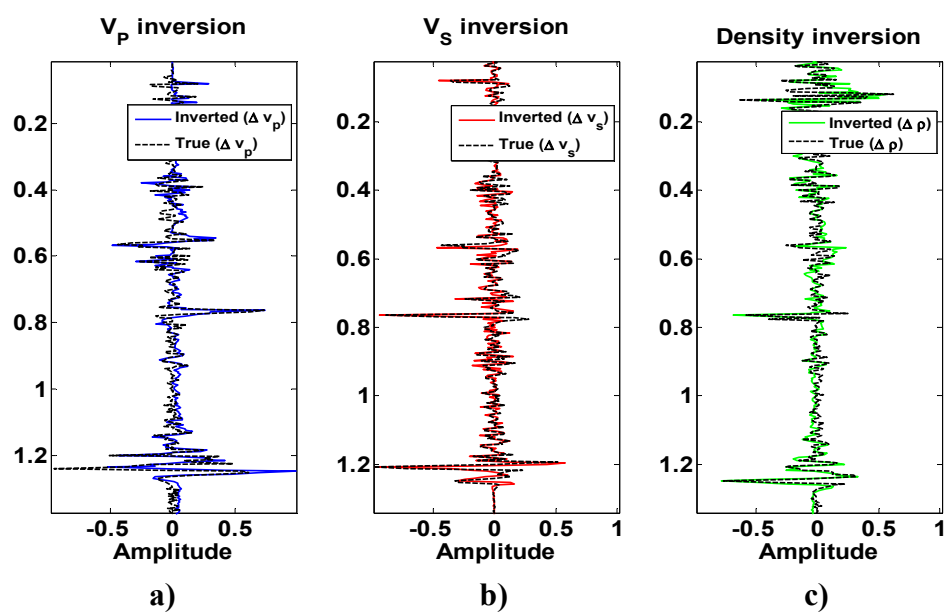


FIG. 10: Synthetic model multiparameter inversion. a) Inverted P-wave velocity. b) inverted S-wave velocity. c) Inverted density. Corresponding well log data are plotted as dotted lines. To perform a comparison the well data is differentiated and convolved with a suitable wavelet.

A FIELD DATA EXAMPLE: VELOCITY INVERSION (NEBC)

We chose 51 shot records that were acquired by Nexen Energy ULC in the Northeast British Columbia (NEBC) region of Canada. The receiver spacing is 10 m and the source spacing is 60 m. The source type was Vibroseis used vertical vibration mode. The sample rate was 4 ms and the grid lateral spacing for the inversion is 12.5 m. The maximum offset for P-to-P and P-to-S data are arbitrarily chosen to be 1000 m and 1500 m, respectively. The vertical and radial components of the 3C geophones were employed and are assumed to contain predominantly P-to-P and P-to-S waves, respectively. For the limitation of offset on our data we could not reconstruct reliable density, so we aim to invert for v_P from P-to-P and v_S from P-to-S data.

Well log information in the study area is shown in Figure 8a. To obtain the true amplitude of P-to-P data, transmission loss and spherical divergence are considered in the forward modeling. Attenuation compensation (Kjartansson, 1979) with Q=80 for shallower events (0-0.7 s) and Q=120 for deeper (0.7-1.5 s) was applied to the data (Figure 8b). For the P-S data, attenuation compensation with Q=45 for shallow (0-1 s) and Q=80 for deeper events (1-2 s) was applied. Spherical divergence (Miao et al., 2005) and transmission loss compensation have been applied on P-to-S forward modeling shown in Figure 11.

To perform quality assurance on the forward operator, the results of forward modeling of all shots are migrated and then compared with the migrated section of real data (see Figures 11-12). Higher amplitude compensation at the time $\tau = 1.0$ s in Figure 12d is due to the attenuation (Q) logarithmic factor applied to field shot records to be calibrated to the modeled data in Figure 12c.

Given the low SNR of the field data, we can visually compare the migrated residuals to find an optimum α. This means that for the purpose of quality assurance the objective function (13) is coupled with $\|\Delta M\|^2$ such that

$$\min J = \|\Delta U\|^2 + \|\Delta M\|^2 . \tag{31}$$

Equation (31) is the objective function of Tarantola (1984); however, a slight difference is that the term ΔM, essentially a step function, was obtained by the integration of a spike-like migrated reflectivity function. Here, to overcome the problem of numerical convergence of the equation (31), the band-limited spike-like reflectivity is used for visual comparison, as it is easier to identify (see e.g., Figure (7b) or Bleistein et al., 2001).

The inversion result of P-wave and S-wave velocity is shown in Figures (13). The initial velocity for inversion was obtained using a linearization of the well log as shown in Figure (14). The migration algorithm was not efficient for shallow reflectors, so we inverted data for reflectors deeper than 0.5 s on the vertical component (assumed to be P-to-P dominant) and 0.3 s for radial component (P-to-S dominant). The lack of low frequency contents of seismic data prevented the integration of reflectivity function to produce an ideal step like gradient function. Consequently, as shown by the color scale in the Figure 13, we display the inversion result only within the range of 2000 to 6000 m/s

for the P- wave and 1000 to 2200 m/s for the S- wave velocity inversion. However, the lateral variation of the resulting inversions shows good correlation with the well log data.

To improve the inversion result, the low frequency components of the well log added to the gradient function in order to update the velocity. As shown in Figure (14), the final inversion result showed good correlation with well log information.

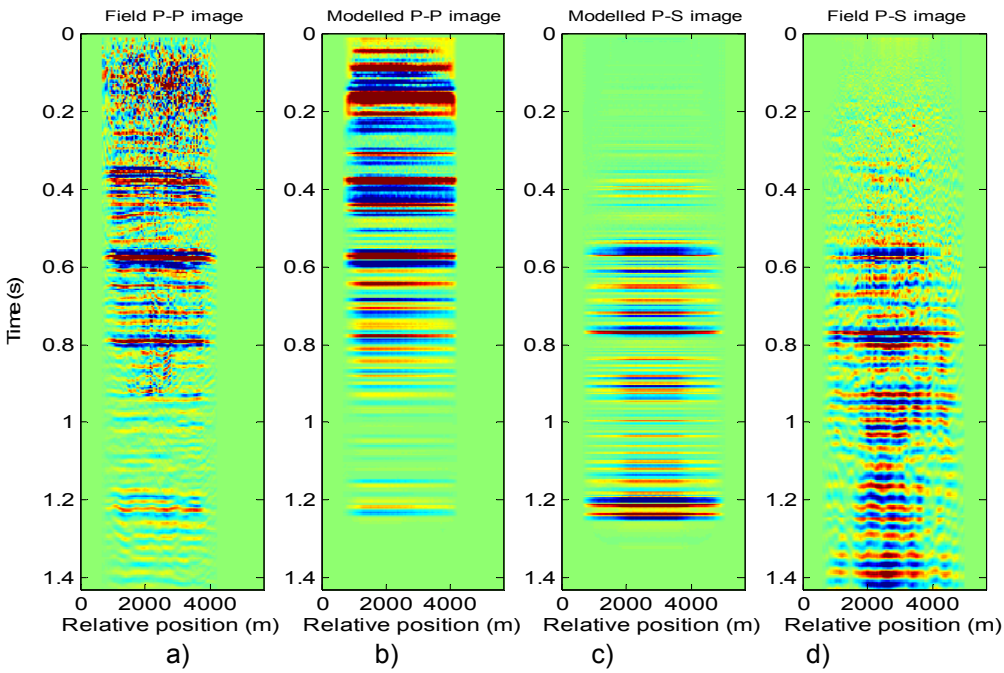


FIG. 11: Comparison of PSTM mapped in P-to-P time. a) Migrated field P-P data b) Migrated modeled P-P data c) migrated modeled P-S data d) migrated field P-S data.

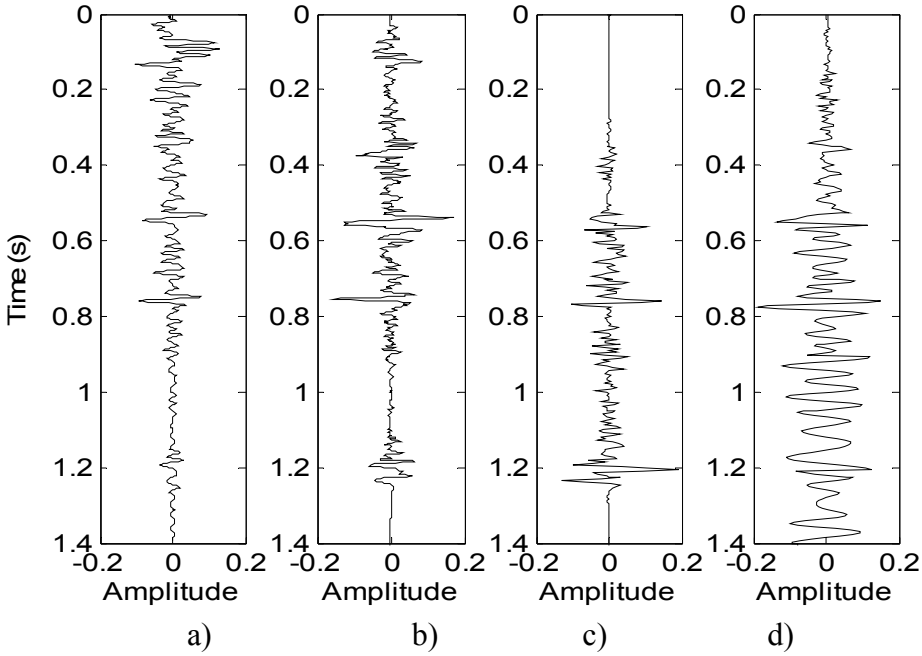


FIG. 12: Comparison of migration results at a control well. a) Migrated field P-to-P data b) Migrated modeled P-to-P data c) migrated modeled P-to-S data d) migrated field P-to-S data.

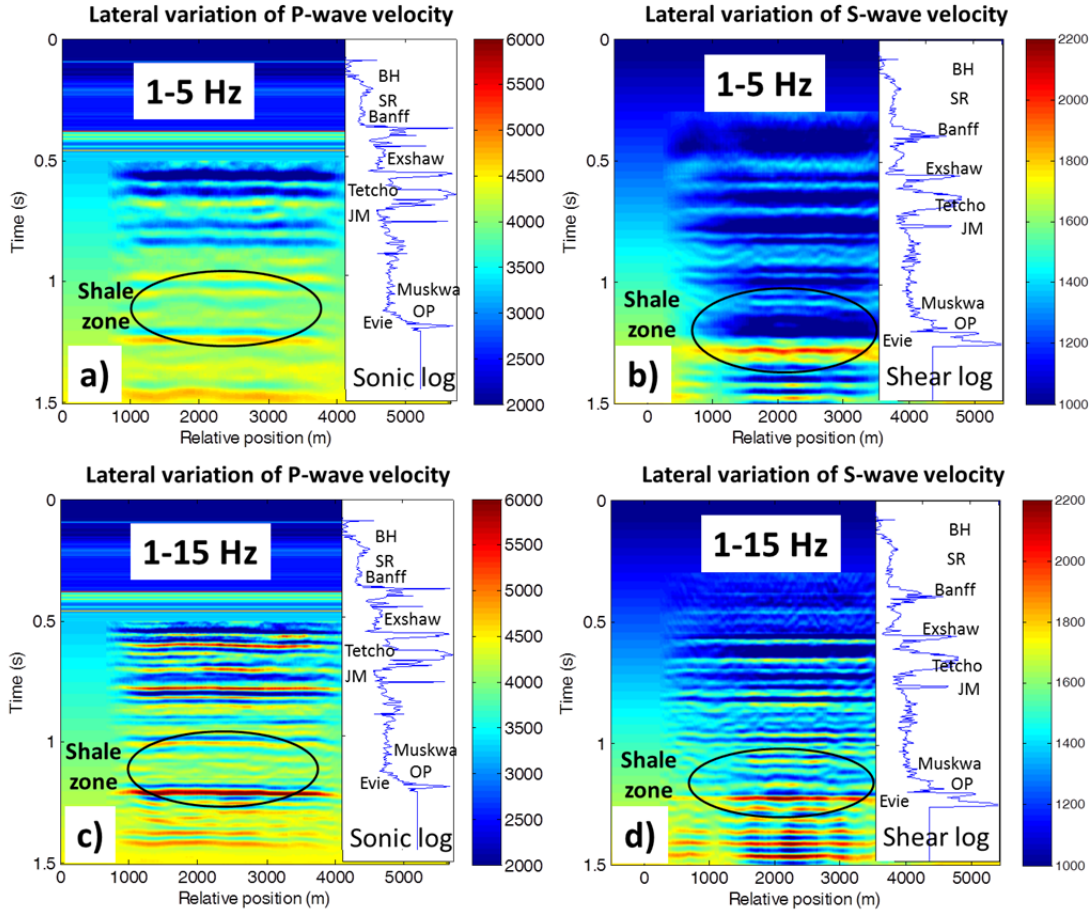


FIG. 13: P-to-P and P-to-S iterative waveform inversion a) P-wave velocity inverted from the 1-5 Hz frequency band of P-to-P data. b) S-wave velocity inverted from 1-5 Hz P-to-S data. c) Same as (a) but the frequency range is 1-15 Hz. d) Same as (b) but the frequency range is 1-15 Hz.

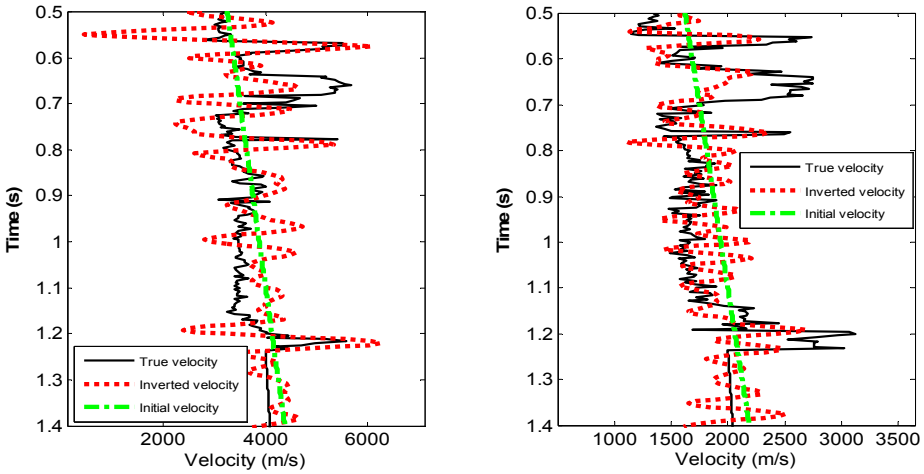


FIG. 14: Inversion results at a control well. a) P- wave velocity, b) S- wave velocity compared with sonic well log (solid blue curve). Low frequency of the well log is added to the gradient function. Initial model is shown as green dotted line.

FIELD DATA EXAMPLE: SHEAR VELOCITY INVERSION FROM SH-to-SH WAVEFIELDS

In addition to recording vertical and radial data using a vertical Vibroseis source, we also recorded horizontal data using a Vibroseis source in horizontal mode. As shown in Figure (15), a conventional velocity analysis have been done byforming S-to-S wavefeild CSP gather (Bancroft, 1998). As shown in Figure (16a), this approach can provide the low frequency component of the shear wave velocity. Low SNR prevented an accurate picking of the events.

In Figure (16), we show a Band Limited Impedance (BLIMP) inversion (Ferguson and Margrave, 1996) of poststack SH-to-SH migrated section. The BLIMP algorithm adds the low frequency of a well log to the integrated reflectivity of the poststack section. The similarity of BLIMP with FWI can be thought of as the first step of the FWI algorithm with the initial velocity being constant. The constant velocity makes the data residual to be equivalent to true data (i.e., $\Delta v = 0$). The result of inversion can be optimized by adjusting the range of well log frequencies be added to the integrated trace. Similarly, in FWI the initial model is a smoothed version of true model implying that the reflectivity of initial model is small compared to true model. Therefore, one can take advantage of the BLIMP methodology to estimate the gradient function by including low frequency of well data. In contrast the FWI has the advantage of forward modeling to scale the gradient toward the true model.

As shown in Figure (16), after 1.5 s (SH-to-SH time) the inversion gives more correlation with well log data and v_s inversion from P-to-S data. Accurate deconvolution, true amplitude and phase corrections could enhance the result of inversion.

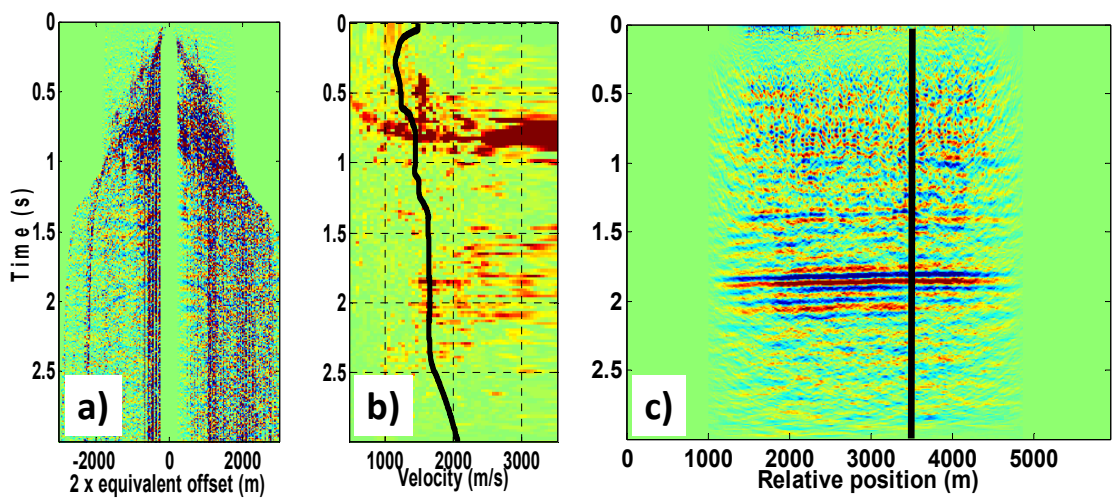


FIG. 15: Shear velocity analysis. a) Sample CSP gather for conventional velocity analysis. b) Hyperbolic semblance of (a). The black line indicates the RMS velocity obtained from well log c) The migration of the SH-to-SH data. The vertical black line shows the location of CSP gather.

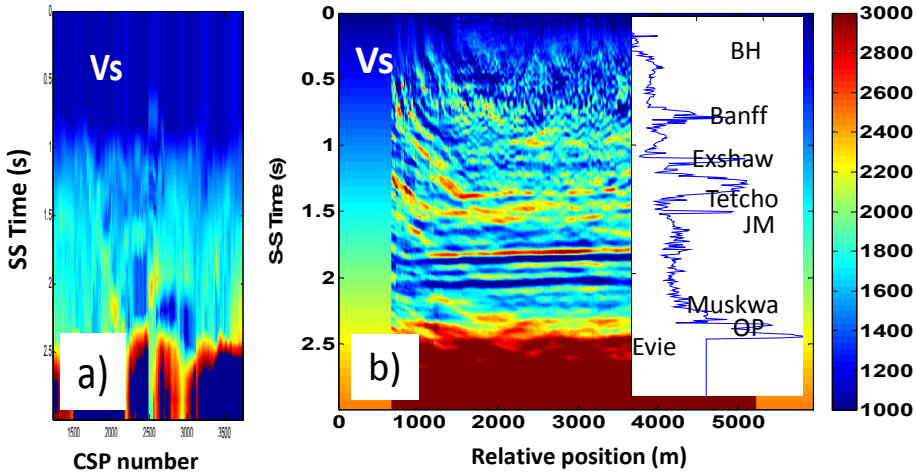
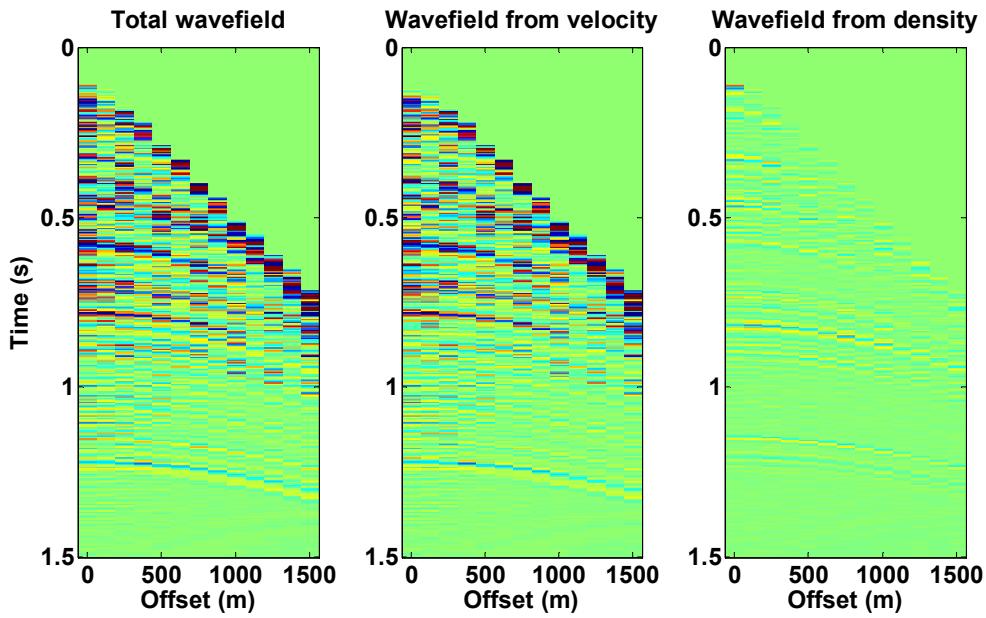


FIG. 16: Shear velocity inversion. (a) The picked velocity from semblance of CSP gathers. b) One step inversion of SH-to-SH wavefield using the poststack inversion (Fergusen and Margrave 1998). The shear velocity is assumed to be the dominant contributor of the reflectivity. As shown in Figure (15a), the level of SNR is low, so we were not able to successfully iterate the inversion for more accurate result.

FIELD DATA EXAMPLE: DENSITY FIELD DETERMINATION WITH LIMITED OFFSET DATA (NEBC)

The inversion for density is a difficult task and usually requires high angle of incident to be able to distinguish between the effects of velocity and density (see e.g., Figure 5) on the data. Another problem is that the variation of density of the models is smaller than the variation of velocity. Fortunately, the method of FWI has the capability to perform forward modeling. So we can decompose the total linearized wavefield into velocity field and density field as shown in equation (29) and (30). Figure (17a) shows a synthetic data computed from the available velocity and density. Figure (17b) obtained by the same parameters of Figure (17a) except the density is set to be constant. Figure (17c) is the difference between Figure (17a) and Figure (17b) or the same as the density field. Once the density field is obtained we can assign them to the real data. Figure (18a) shows the migrated real data. Figure (18b) and Figure (18c) are the migration of the modeled total wavefield field and velocity field respectively (as obtained in Figure 17). Figure (18d) is the estimated density field which gives the trend of density field in the real data. From Figure (18), we obtained the velocity and the density field in Figure (19a) and Figure (19b) which improve the gradient calculation in FWI procedure.

Obviously, noise is present in the real data, which prevents a prefect reconstruction of the wavesfield. It is a common sense that the implementation of human judgment (e.g., manual picking and editing) improves the recovered fields.



a) b) c)

FIG. 17: Strategy for estimation of density field in prestack data. a)The total wavefield modeled from a well control is decomposed into b) the velocity field, and c) the density field.

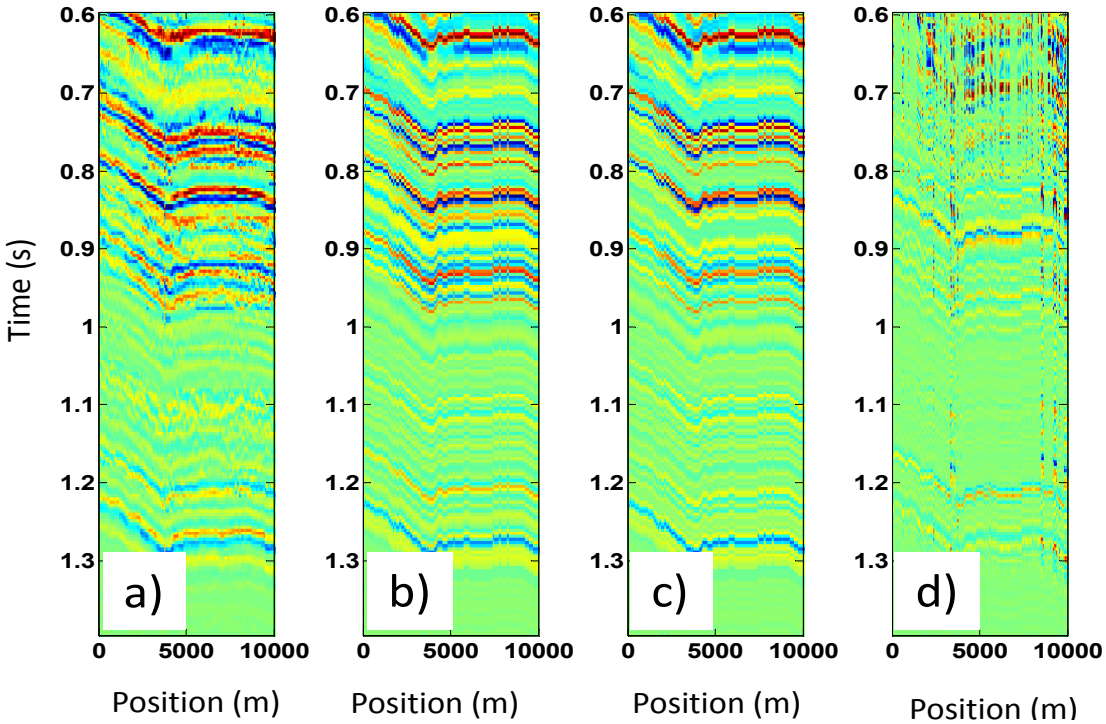


FIG. 18: Density field estimation strategy. a) An inline migrated section from a real 3D data. b) The migrated section of the modelled data obtained from (a). c) Velocity field of (b). d) Density field of (b). Note that the lateral extension of the field data is compressed.

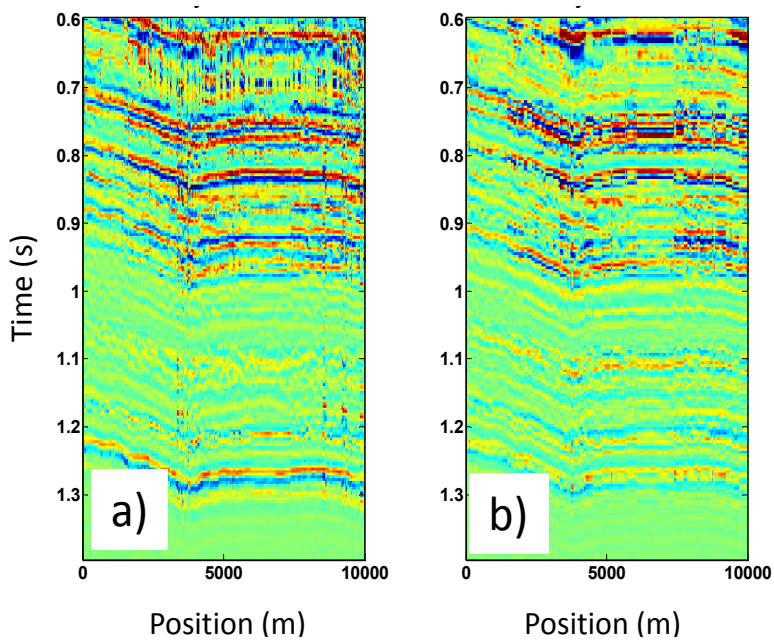


FIG. 19: Linearized FWI gradient calculation with the density in Figure (18). a) Estimated reflectivity from the velocity field. b) Estimated reflectivity from a density.

CONCLUSIONS

We proposed an efficient algorithm for multiparameter inversion of the elastic properties in the frame of Full Waveform Inversion (FWI). For reducing the computational costs, we demonstrated that the linearized solution of the seismic reflection inverse problem can be obtained using standard Prestack Time Migration (PSTM) and corresponding forward modeling in an iterative scheme. We showed that we can use the linearized Zoeppritz solvers for amplitudes radiation pattern of scatter points and Double Square Root (DSR) equation for traveltime consideration of P-to-P and P-to-S data during the waveform inversion. We showed that gradient function of objective function can be obtained by the well-known AVO inversion schemes. The method is practical for real data applications since all engines of the inversion are known to be applicable for standard methods. We showed examples of the inversion of real data using the waveform characteristics of P-to-P and P-to-S data. Without the loss of generality, the approach is adaptable for different combination of modes such as P-to-P, P-to-S, S-to-S and S-to-P, and for different mediums such as anisotropic, anelastic, poroelastic and viscoelastic. Implementing Kirchhoff based methods for higher order Born scattering data (multiples) and 3D data are being investigated and will be the subject of another paper.

ACKNOWLEDGEMENTS

We thank the faculty and sponsors of CREWES and also Nexen Energy ULC for their generous support of this work, as well for permission to the publish the results. We also would like to thank Neda Boroumand and Steve Jensen and Shahpour Moradi for their useful comments and editing of this work.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative Seismology: Theory and Methods, W.H. Freeman and Company. Vol. 1.

Bancroft, J. C., Geiger, H. D., and Margrave, G. F., 1998, The equivalent offset method of prestack time migration: Geophysics, **63**, 2042-2053.

Beylkin, C., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a casual generalized Radon transform: J.Math. Phys., **26**.99-108.

Beylkin G., and Burridge R., 1990, Linearized inverse scattering problems in acoustics and elasticity, Wave Motion, **12**, 15–52,

Bleistein, N., J. K. Cohen, and J. W. S. Jr., 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer Verlag.

Bleistein. N.. 1984, Mathematical methods for wave phenomena: Academic Press Inc.

Bleistein, N. (1987). "On the imaging of reflectors in the earth." GEOPHYSICS, 52(7), 931–942.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**,467-481.

Cohen, J, .K. and Bleistein, N., 1979, Velocity inversion procedure of acoustic waves: Geophysics, **44**, 1077-1087.

Dix, C. H., 1955, Seismic velocities from surface measurements: Geophysics, **20**, 68-86.

ten Kroode, F., Bergler, S., Corsten, C., de Maag, J., Strijbos, F., and Tijhof, H. (2013). Broadband seismic data — The importance of low frequencies: Geophysics, 78(2), WA3–WA14.

Hubral, P., 1977, Time-migration - some ray-theoretical aspects: Geophys. Prosp. 25, 738-745.

Innanen,K., and Margrave G. F., 2011, Seismic inversion and the importance of low frequencies: CREWES annual research report, U of Calgary

Jaramillo, H. and Bleistein, N. (1999). "The link of Kirchhoff migration and demigration to Kirchhoff and Born modeling." GEOPHYSICS, 64(6), 1793–1805.

Khaniani H., Bancroft J. C., and Margrave G. F., 2012, Full waveform inversion algorithm using Common Scatter Point (CSP) gathers: SEG Expanded Abstracts.

Kjartansson, E. (1979), Constant *Q*-wave propagation and attenuation, J. Geophys. Res., 84(B9),4737–4748.

ten Kroode, F. (2012). A wave-equation-based Kirchhoff operator. Inverse Problems, 28(11), 115013.

Manning, P. M., 2008, Techniques to enhance the accuracy and efficiency of finite-difference modeling for the propagation of elastic waves: PhD thesis, The University of Calgary.

Margrave, G. F., Ferguson, R. J., and Hogan, C. M, 2011, Full Waveform Inversion (FWI) using one-way migration and well calibration: SEG Expanded Abstracts **30**.

Mi Y., and Margrave G. F., 2001, Converted-wave prestack depth imaging with the nonstationary wavefield extrapolators: CREWES annual research report, U of Calgary

Ostrander, W. (1984). "Plane wave reflection coefficients for gas sands at nonnormal angles of incidence."GEOPHYSICS, 49(10), 1637–1648.

Parson, R., 1986, Estimating reservoir mechanical properties using constant offset images of reflection coefficients and incident angles: 56th Ann. Internat. Mtg., Expanded Abstracts, 617-620.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49-76.

Sears, T., S. Singh, and P. Barton, 2008, Elastic full waveform inversion of multi-component OBC seismic data: Geophysical Prospecting, **56**, 843–862

Shaw, R. K., and M. K. Sen, 2004, Born Integral, stationary phase, and linearized reflection coefficients in anisotropic media, Geophysical Journal International, 158, 225-238.

Sirgue, L., and Pratt, R. G., 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 1, 231.

Smith, G.C., and Gidlow, P.M., 1987, Weighted Stacking for Rock Property Estimation and Detection of Gas: Geophysical Prospecting, 35, 993-1014.

Stewart, R.R., 1990, Joint P and P-SV Inversion: CREWES Research Report.

Taner, M. T., and F. Koehler, 1969, Velocity spectra - Digital computer derivation and applications of velocity functions: Geophysics, **34**, 859-881.

Tarantola, A., 1984a, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998-1015.

- 1984b, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

- 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893-1903.

-1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier Science Publishing Co., Inc.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C.,Conroy, G., and Bertrand, A. (2013). ”Anisotropic 3D full-waveform inversion.” GEOPHYSICS, 78(2), R59–R80.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, WCC127-WCC152

Yilmaz, Ö., 1989, Velocity-stack processing: Geophysical Prospecting, **37**, 357–382

Zuleta L., 2012, Near-surface characterization and Vp/Vs analysis of a shale gas basin: M.Sc. Thesis, University of Calgary