Elastic wave-equation reflection traveltime inversion wave mode decomposition

Tengfei Wang[1], Jiubing Cheng[2], Qiang Guo[3] and Chenlong Wang[4]

[1] *School of Ocean and Earth Science, Tongji University, Shanghai, China. E-mail:*
1110701@tongji.edu.cn

[2] *State Key Laboratory of Marine Geology, Tongji University, Shanghai, China. E-mail:*
cjb1206@tongji.edu.cn

[3] *School of Ocean and Earth Science, Tongji University, Shanghai, China. E-mail:*
1110701@tongji.edu.cn

[4] *State Key Laboratory of Marine Geology, Tongji University, Shanghai, China. E-mail:*
cjb1206@tongji.edu.cn

(May 1, 2017)

Running head: **Elastic WERTI based on mode decomposition**

ABSTRACT

Elastic reflection waveform inversion (ERWI) utilize the reflections to update the low and intermediate wavenumber in the deeper part of elastic model, which can provide good initial models for elastic full waveform inversion (EFWI). Though ERWI can mitigate the nonlinearity to some extent, it is still stuck with the cycle-skipping problem due to the objective function of waveform fitting. Building initial P and S wave velocity models for EFWI by using elastic wave-equation reflections traveltime inversion (ERTI) would be effective and robust since taveltime information relates to the background model more linearly. The reflection kernels in elastic media are complicated and difficult to use, especially when cal-

culating the gradient of S-wave velocity. The investigation of reflection kernels show that mode decomposition can suppress the artifacts in gradient calculation. Besides, P/S separation of multicomponent seismograms can isolate the P or S recordings for individual data residuals. Thus, we implement the ERTI by using the L_2 norm of the traveltime residual extracted by the Dynamic image warping (DIW) as objective function. A two-step inversion strategy is adopted to effectively reduce the nonlinearty of inversion, in which PP reflections are first used to invert V_p, followed by V_s inversion with PS reflections based on the well recoverd V_p. To avoid the mismeasurement of PS traveltime residual, the well recovered PP image are used to generate the PS reflections. Numerical example of Sigsbee2A model validates the effectiveness of the algorithms and strategies for ERTI.

INTRODUCTION

With the emergence of long-offset wide-azimuth acquisitions and broad-band sources, full waveform inversion (FWI) has been recognized as an efficient tool for constructing velocity models and quantitative seismic imaging (see Virieux and Operto (2009) for a review). Though the acoustic FWI, primarily focusing on P-wave velocity inversion, has been widely studied in the past decades (Tarantola, 1984; Pratt et al., 1998; Shipp and Singh, 2002), people are paying more attention to the waveform inversion under the elastic assumption, referred to as elastic full waveform inversion (EFWI) (Tarantola, 1986). EFWI provides high-resolution model estimation of the elastic properties but will fall into local minimal easily, because it is incapable of recovering the low and intermediate wavenumber components of the model in the absence of low frequency data and/or good initial models (Sears et al., 2008; Brossier et al., 2009). Besides, the nonlinearities of multi-parameter inversion further increase the difficulties for EFWI. To deal with the cycle-skipping and parameter trade-offs problems, more preconditions, hierarchical strategies and parameterization investigation should be considered during the inversion (Sears et al., 2008; Prieux et al., 2013; Operto et al., 2013; Wang et al., 2015b; Oh and Alkhalifah, 2016).

In the classical FWI, the long-offset data corresponding to diving waves are very important to build the long-to-intermediate wavelengths of the model. However, the penetration depths of diving waves are always far from sufficient to reach the target in the deeper part even using the wide-azimuth surveys. In addition, the low signal-to-noise ratio at the far offset is also a limit for classical FWI relying on the diving waves. Recently, the reflected waves are considered to be a potential tool to help build the macro-model containing low-to-intermediate wavenumber in the deep part(Xu et al., 2012; Wu and Alkhalifah, 2015;

Zhou et al., 2015). This velocity building task can be implemented in the image domain or the data domain. Actually, the image-domain ray-based tomography (Woodward, 1992; Woodward et al., 2008; Jones, 2010) has already been widely used to obtain the background velocity by flatting the common image gathers. But when the lateral velocity variation is strong, the ray-based method would fail to present the wave propagation underground, which will probably lead to the failure of tomography. To overcome the limit of ray theory, the reflection inversion based on wave equation has been studied by many researches.

The misfit function of reflection inversion can be built in image domain in the manner of wave equation migration velocity analysis (WEMVA), which tries to maximize the energy at zero offset location in the extended image space (Symes, 2008; Almomin and Biondi, 2012; Sun and Symes, 2012; Biondi and Almomin, 2013). Raknes and Weibull (2016) developed the image-domain method in the 3D elastic media. Wang et al. (2017) proposed to employ the extended PS image in WEMVA to update the V_s model with the help of elastic wave mode decomposition. However, the extended-domain methods are limited due to its prohibitively computational cost, especially in 3D case. While in the data domain, Xu et al. (2012) suggested using a reflection waveform inversion (RWI) method to suppress the nonlinearity in FWI, which aim to invert the long-wavelength components of the model by using the reflections predicted by migration/demigration process. RWI highly relies on the accurate reflectivity to generate the reflections that can match the observed data. However, it is very challenging and also expensive to obtain a good reflectivity model through least-square migration when initial model is far away from the true value. Zhou et al. (2015) proposed a joint FWI method to combine the diving and reflected waves to utilize both RWI and the conventional FWI. Recently, the RWI method is developed by several works (Wu and Alkhalifah, 2015; Zhou et al., 2015), and recently extended to elastic case by Guo

4

and Alkhalifah (2016).

However, compared with the waveform information, traveltime is more sensitive and linearly related to the low-wavenumber components of the model. Therefore, traveltime inversion will be more robust and helpful to build good initial models for conventional FWI (Wang et al., 2014). Ma and Hale (2013) introduced a wave equation reflected traveltime inversion (WERTI) method based dynamic image warping (DIW) to build the low wavenumber of the model. Chi et al. (2015) and Wang et al. (2015a) employed correlation-based method to extract the temporal or spatial lag to implement the reflection inversion. Elastic reflections carry the background information of the P and S wave velocities, which can be in favor of EFWI with good initial P and S wave velocity models. Unfortunately, in elastic case, traveltimes of a particular wave modes are difficult to extract due to the complicated mode-conversions. Therefore, the estimated travletime differences would be inaccurate if using the original multicomponent seismic data, which makes the elastic wave-equation reflection traveltime inversion (ERTI) difficult to implement. Besides, since the multi-parameter trade-offs will increase the nonlinearity of inversion, more hierarchical strategies should be considered to deal with this problem. Wang and Cheng (2017) obtained good results by utilizing the wave mode decomposition to precondition the gradients in EFWI. The mode decomposition would be a potential tool for hierarchical strategies in ERTI.

In this paper, we will tackle the traveltime misfits of PP and PS reflections to implement the ERTI with the aid of wave mode decomposition and DIW (Hale, 2013). First, the elastic reflection wavepath and its components of different wave modes are calculated to investigate the effectiveness of mode decomposition on suppressing the artifacts in the gradient calculation. Then, P/S separation of multicomponent seismograms is applied to the observed and predicted reflection data to extract the individual traveltime residuals through DIW.

5

Based on the analysis of elastic reflection kernels and the separated traveltime residuals, we design a two-stage workflow to implement the ERTI method, in which the traveltime of PP is firstly used to recover the background V_p model followed by inverting the background V_s model through the traveltime of PS reflections. During the V_s inversion, the PS reflections are predicted by the PP image to avoid the wrongly estimated time shifts when using PS image. And also, we precondition the V_s gradient through spatial wave mode decomposition. Finally, the numerical example of Sigsee2A model proves the robustness and validity of our ERTI method.

THEORY OF ERTI

Assume that there is a perturbation δc_{ijkl} in the background elastic media c_{ijkl}, the background wavefileds u_i and perturbed wavefields δu_i satisfy:

$$\rho \frac{\partial u_i^2}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}\frac{\partial u_k}{\partial x_l}\right] = f_i, \tag{1}$$

and

$$\rho \frac{\partial \delta u_i^2}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}\frac{\partial \delta u_k}{\partial x_l}\right] = \frac{\partial}{\partial x_j}\left[\delta c_{ijkl}\frac{\partial u_k}{\partial x_l}\right], \tag{2}$$

where δu_i can be seen as the demigrated reflection data using the image perturbation δc_{ijkl} obtained from RTM or other imaging method. In ERTI, we aim to minimize the traveltime differences between observed data $\mathbf{d}^o$ and calculated data $\mathbf{d}^c$, then the objective function is:

$$\begin{cases} \tau(\mathbf{x_r}, t; \mathbf{x_s}) = \arg\min_\tau \parallel \mathbf{d}^c(\mathbf{x}_r, t; \mathbf{x}_s) - \mathbf{d}^o(\mathbf{x}_r, t + \tau; \mathbf{x}_s) \parallel^2 \\ E = \frac{1}{2}\int \tau^2(\mathbf{x_r}, t; \mathbf{x_s})dt d\mathbf{x_r} d\mathbf{x_s}, \end{cases} \tag{3}$$

where the time differences $\tau(\mathbf{x_r}, t; \mathbf{x_s})$ can be extracted through DIW. After a similar derivation as in Ma and Hale (2013) (see Appendix A for detail), the gradients of equation (3)

can be expressed as:

$$\frac{\partial E}{\partial c_{ijkl}} = -\int (\frac{\partial u_i}{\partial x_j}\frac{\partial \delta\psi_k}{\partial x_l} + \frac{\partial \delta u_i}{\partial x_j}\frac{\partial \psi_k}{\partial x_l}), \tag{4}$$

where ψ_i and $\delta\psi_i$ are the adjoint wavefields satisfying:

$$\rho\frac{\partial \psi_i^2}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}\frac{\partial \psi_k}{\partial x_l}\right] = \tau(\mathbf{x_r}, t; \mathbf{x_s})\frac{\dot{d}_i^o(\mathbf{x_r}, t+\tau; \mathbf{x_s})}{h_i(\mathbf{x_r}, t; \mathbf{x_s})}, \tag{5}$$

and

$$\rho\frac{\partial \delta\psi_i^2}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}\frac{\partial \delta\psi_k}{\partial x_l}\right] = \frac{\partial}{\partial x_j}\left[\delta c_{ijkl}\frac{\partial \psi_k}{\partial x_l}\right], \tag{6}$$

with $h_i(\mathbf{x_r}, t) = \dot{d}_i^o(\mathbf{x_r}, t+\tau)^2 - \ddot{d}_i^o(\mathbf{x_r}, t+\tau)(d_i^c(\mathbf{x_r}, t) - d_i^o(\mathbf{x_r}, t+\tau)$. The hat dot denotes the time derivative. On the right hand side (RHS) of equation (4), the first and second term indicate the source and receiver part of the reflection wavepath, respectively. Then we can get the gradients in terms of P- and S- wave velocities through the chain rule:

$$\begin{aligned}\frac{\partial E}{\partial V_p} &= 2\rho V_p\frac{\partial E}{\partial c_{ijkl}}\delta_{ij}\delta_{kl}, \\ \frac{\partial E}{\partial V_s} &= 2\rho V_s\frac{\partial E}{\partial c_{ijkl}}(-2\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}).\end{aligned} \tag{7}$$

ELASTIC BORN REFLECTION KERNELS

RWI and ERTI utilize different objective functions which only induce different types of adjoint sources, but they share similar reflection kernels. Therefore, the key point of reflection inversion is to calculate the reflection kernel. Due to the complex mode conversions in elastic wavefields, wavepath of elastic reflections will be far more complicated than that in acoustic case. Here, we decompose the origin kernel into four components which represent cross-correlation of different wave modes.

For simplicity, we rewrite (4) as follow:

$$\nabla E(\mathbf{m_0}) = -\int (\mathbf{u} \cdot \delta\boldsymbol{\psi} + \delta\mathbf{u} \cdot \boldsymbol{\psi}) \tag{8}$$

with $\mathbf{m}_0$ is the background model, $\mathbf{u}$ and $\boldsymbol{\psi}$ are the forward and adjoint background wavefields, $\delta\mathbf{u}$ and $\delta\boldsymbol{\psi}$ are the forward and adjoint perturbed wavefields. The operator $\cdot$ denotes the cross correlation between two wavefields. Note that, equation (8) just schematically shows the manner of cross correlation. The detailed formulas should be derived according to the parameterization through chain rule as in equation (7). Considering mode decomposition, the above formula can be decomposed into four types with:

$$K_{m_0}^{MN} = -\int(\mathbf{u}^M \cdot \delta\boldsymbol{\psi}^N + \delta\mathbf{u}^M \cdot \boldsymbol{\psi}^N), \tag{9}$$

where $M, N \in \{P, S\}$. $K_{m_0}^{MN}$ represents the cross correlation between the M mode forward wavefields and the N mode adjoint wavefields. Note, it does not denote the kernel of MN mode data. For example, the reflection kernel of PS data should be $(\mathbf{u}^P \cdot \delta\boldsymbol{\psi}^P + \delta\mathbf{u}^S \cdot \boldsymbol{\psi}^S)$ but not K^{PS}.

To analyze the elastic reflection kernels, we calculate them with the single-source-receiver data which are synthesized by single reflector with pure P-wave source. We use the velocity parameterization as in equation (7). In the first model, we place a single V_p reflector in the homogeneous background (Fig 1a and b). Since there is no perturbation of V_s, only PP reflection exist in the data, which is almost the same as in acoustic media. As shown in Figure 1c and d, the reflection kernel consists of two "rabbit-ear", the source and receiver parts. We can see the migration impulse below the reflector due to the down-going perturbed wavefields (Zone D mentioned by Zhou et al. (2015)). The energy in V_s kernel focus on the edge of the wavepath rather than the first Fresnel-Zone as in the V_p kernel. One plausible reason is that the V_s kernel is relatively insensitive to the PP data generated by V_p reflector.

In the second model, we use the V_s reflector (Fig 2a and b) to generate both PP and PS reflection. The V_p kernel excludes S-wavefield automatically because of the divergence

8

operator implied in the term $\delta_{ij}\delta_{kl}$ (equation (7)). However, $\delta\boldsymbol{\psi}$ contains the non-physical converted SP wavefields generated by the back-propagated $\boldsymbol{\psi}^S$ at the location of reflector. These SP wavefields make the V_p kernel slightly different from that in Figure 1c. If we only back-propagate the PP data, V_p kernel will be the same as Figure 1c.

For V_s kernel (Figure 2d), due to mode conversions, multi-wavepaths overlapping with each other make it much more complicated and difficult to find the correct PS reflection kernel. The straightforward utilization of this kernel in gradient calculation will very likely cause severe cross-talk during reflection inversion. According to equation (9), we calculate the components of K_{V_s}, as shown in Figure 3. The $K_{V_s}^{PP}$ is similar to $K_{V_p}^{PP}$ but with an opposite sign. $K_{V_s}^{PS}$ and $K_{V_s}^{SP}$ mainly consist of high-wavenumber energy, which in fact are the migration impulse of cross-mode wavefields. Most importantly, we expect to update the V_s model through the S-wavepath in PS reflection, which is exactly the component $K_{V_s}^{SS}$. Figure 3d shows the expected S-wavepath of PS reflection affected little by the P-wavefields. Therefore, we recommend using $K_{V_s}^{SS}$ to mitigate the cross-talks in V_s gradient calculation.

WORKFLOW OF ELASTIC WERTI

In elastic case, it is common to observe that different mode-conversions, mainly PP and PS events, overlap and intersect with each other. The cross points between events would be singularities for traveltime difference estimation through DIW. Therefore, $\tau(\mathbf{x_r}, t; \mathbf{x_s})$ would be inaccurate because of these singularities if using the original multicomponent seismic data, which makes equation (7) difficult to implement. Besides, according to the reflection kernel analysis, individually injecting PP or PS recordings also can mitigate the cross-talk in gradient calculation. Thus, we decompose the observed and calculated data into P- and S-wave parts through P/S separation of multi-component seismograms (Li et al., 2016). In

this way, the traveltime differences can be decoupled into P- and S-wave part by DIW, with which we can implement the ERTI through a two-stage workflow, i.e. the P-wave stage followed by the S-wave stage.

PP reflection for V_p

In this stage, we use traveltime of PP reflection to recover the low-to-intermediate wavenumber of V_p model. The perturbation of V_p (δV_p) is obtained through elastic reverse time migration (ERTM). When the migration velocity model is inaccurate, the migration/demigration of large-offset data would produce deviated zero-offset two-way travletime, which will increase the nonlinearity of inversion. Therefore, people always assume the two way travletime of zero-offset as an invariants during reflection inversion, such as reflection traveltime tomography or RWI. This assumption leads to the usage of zero or small offset data to generate the velocity perturbation (Zhou et al., 2015). Besides, since the traveltime rather than the amplitude of predicted reflection is considered in ERTI, ERTM is sufficient to obtain the image perturbation but not a least-squares manner (LSRTM) to find the correct reflectivity. Thus, the objective function is:

$$
\begin{cases}
\tau_{pp}(\mathbf{x_r}, t; \mathbf{x_s}) = \arg\min_{\tau} \parallel \mathbf{d}_{pp}^c(\mathbf{x}_r, t; \mathbf{x}_s) - \mathbf{d}_{pp}^o(\mathbf{x}_r, t + \tau; \mathbf{x}_s) \parallel^2 \\
E_{pp} = \dfrac{1}{2} \displaystyle\int \tau_{pp}^2(\mathbf{x_r}, t; \mathbf{x_s}) dt d\mathbf{x_r} d\mathbf{x_s},
\end{cases}
\tag{10}
$$

where d_{pp}^c and d_{pp}^o are calculated and observed PP reflection obtained by surface P/S separation, respectively. According to the previous derivation, we can obtain the gradient of V_p ($\frac{\partial E}{\partial V_p}$). Note, the divergence operator is implied in $\frac{\partial E}{\partial V_p}$, which makes the gradient calculation only involve PP reflection. Therefore, the PP reflection inversion for V_p in elastic media is similar to the acoustic case except that we need the P/S separation for the former one. Besides, the separated multicomponent PP recordings will include some SP conversions (if

any) in the elastic case, which may lead to the mismatch of the traveltime inversion.

PS reflection for V_s

In the S-wave stage, we utilize the PS reflection to retrieve the background component of V_s model. Similarly, the objective function becomes:

$$
\begin{cases}
\tau_{ps}(\mathbf{x_r}, t; \mathbf{x_s}) = \arg\min_{\tau} \parallel \mathbf{d}_{ps}^c(\mathbf{x}_r, t; \mathbf{x}_s) - \mathbf{d}_{ps}^o(\mathbf{x}_r, t + \tau; \mathbf{x}_s) \parallel^2 \\
E_{ps} = \dfrac{1}{2} \displaystyle\int \tau_{ps}^2(\mathbf{x_r}, t; \mathbf{x_s}) dt d\mathbf{x_r} d\mathbf{x_s}.
\end{cases}
\tag{11}
$$

However, the implementation is a little different from the previous stage. There are two ways to predict the PS reflection after the P-wave stage.

I: Migrate the observed PS data with inverted V_p model and initial V_s model to get the high wavenumber of V_s (δV_s), then predict the PS reflection;

II: After the P-wave stage inversion, both the background and the structure of V_p have been well recovered. As we know, in most cases V_p and V_s share the same structure in the subsurface. Therefore, we can use the well-located δV_p instead of δV_s to generate the PS reflections.

Unfortunately, the predicted PS reflection with method **I** are sensitive to the error of V_p model. A slightly deviated V_p model around the true value will lead to a wrongly estimated traveltime residual of PS reflection. However, method **I** can provide correct traveltime residuals. Under the high-frequency approximation, We will explain this using a single reflector example through the map migration and demigration.

Since the PS reflection is asymmetric, the ray path of PS is shown in figure 4. According

11

to the Snell's law and geometrical relationship, we have:

$$\frac{sin\theta_1}{V_p} = \frac{sin\theta_2}{V_s},$$

$$h(tan\theta_1 + tan\theta_2) = X,$$

(12)

with V_p and V_s are the correct velocity, X is the offset, h is the depth, and θ_1 and θ_2 are the incident and reflected angle, respectively. Therefore, the moveout of PS reflection is:

$$t = \frac{h}{cos\theta_1 V_p} + \frac{h}{cos\theta_2 V_s}.$$

(13)

The difference between method **I** and **II** is the migrated depth of the two processes when velocity model is inaccurate. To show the difference of the above two method, zero (near) offset map migration is implemented to obtain the migrated depth when velocity is wrong. Then, we use this depth and the wrong velocity through equation (13) to predict the traveltime of PS reflection.

In method **I**, since the zero-offset two-way traveltime keeps unchanged during map migration, the migrated depth (h_1) satisfies:

$$\frac{h_1}{V_p'} + \frac{h_1}{V_s'} = \frac{h}{V_p} + \frac{h}{V_s},$$

(14)

where V_p' and V_s' are the wrong velocity. Easily, we can get:

$$h_1 = \frac{V_s'(V_p + V_s)}{V_s(V_p' + V_s')}h.$$

(15)

As we know, it is common to find $V_p > V_s$ and $V_p' > V_s'$. Therefore, a small perturbation of V_p' will lead to a relatively big change of the term $\frac{V_p + V_s}{V_p' + V_s'}$, which means that the migration depth h_1 is very sensitive to both V_p' and V_s'. While in method II, the depth after migration is:

$$h_2 = \frac{V_p'}{V_p}h.$$

(16)

12

Then, the demigration process can be implemented with:

$$t' = \frac{h'}{cos\theta_1 V_p'} + \frac{h'}{cos\theta_2 V_s'}.$$ (17)

where $h' \in \{h_1, h_2\}$ is the migrated depth and t' is the corresponding traveltime.

To test the sensitivity of demigrated PS traveltime to V_p' and h, we compare the predicted PS traveltime curves of the wrong velocity with that of the true velocity. Here, $V_p = 2500m/s$, $V_s = 1500m/s$, $V_s' = 1300m/s$ are fixed and the value of V_p' are perturbed ($V_p' = 2450, 2500, 2550m/s$) to calculate the curves. As shown in figure 5a, the predicted PS traveltime with method I (black) oscillates around the true value (red) up and down. Theoretically, the predicted curve should behave like the second black line ($V_p' = V_p$), where the zero-offset traveltime is correct while the long offset traveltime is larger than the true one. Especially, when $V_p' = 2450m/s$ the predicted curve (above the true value) is inconsistent with the S-wave velocity error ($V_s' < V_s$). While when $V_p' = 2550m/s$ the predicted curve is below the true value. This means that the sign of the traveltime residual will be very sensitive to the error of V_p model, even if the error of V_p is as small as $\pm50m/s$. This change of the residual's sign will cause the change of the gradient's direction during the inversion. Therefore, the nonlinearity of inversion increases obviously due to this sensitivity. However, if we use method II (blue), the predicted curves are all below the true value, whose traveltime residual are consistent with the S-wave velocity error. If the depth of reflector decreases, the sensitivity in method I reduces (figure 5b). But the inconsistency of residual and velocity error still exists, especailly in the small offset. However, method II also provides correct time residuals as the previous one. According to the above analysis, we recommend method II. Nonethelss, there are two disadvantages along with this process. First, since we use the δV_p to generate PS reflection, the amplitude of PS will be inaccurate even if

the ELSRTM is used to find the corrcet reflectivity. Luckily, ERTI does not rely on the amplitude information otherwise it will be difficult problem. Second, according to figure 5, the traveltime difference is large for the PS reflection in the deeper part, DIW may suffer from the cycle-skipping in this situation. We can utilize the "layer-stripping" strategy, during which the shallow part is inverted using the early arrived reflection followed by the deep part inversion with the late arrived reflections. In this way, the incorrect traveltime residuals of the deeper part will not mislead the inversion of shallow part. After that, the inversion of shallow part can ensure the traveltime residual of deeper part correct.

Besides, in order to make sure that reflected S-wavepath is used to update V_s, wave mode decomposition is implemented to calculate $K_{V_s}^{SS}$ with:

$$\frac{\partial E_{ps}}{\partial V_s} = -2\rho V_s \int (\frac{\partial \delta u_i^S}{\partial x_j}\frac{\partial \psi_k^S}{\partial x_l})(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{18}$$

This is similar to the gradient precondtioning for EFWI proposed by Wang and Cheng (2017). The mode decomposition can mitigate parameter trade-offs and suppress artifacts for the V_s inversion.

COMPARISON WITH CORRELATION BASED OBJECTIVE FUNCTION

Besides DIW methdo, the temporal or spatial correlation is also widely used to extract the traveltime differences (van Leeuwen and Mulder, 2010; ?; Wang et al., 2015a):

$$\begin{cases} \Delta t(h) = \underset{\Delta t}{\arg\min} \int d^o(t + \Delta t, h)d^c(t, h)dt, \\ E_t = \frac{1}{2}\sum \parallel \Delta t(h) \parallel^2 \end{cases} \tag{19}$$

or

$$\begin{cases} \Delta h(t) = \arg\min_{\Delta h} \int d^o(t, h + \Delta h) d^c(t, h) dt, \\ E_h = \frac{1}{2} \sum \| \Delta h(t) \|^2 \end{cases} \quad (20)$$

The traveltime residual extracted by correlation function is the global average delay of a single trace or time. If there are multievents in the seismograms, the correlation-based method can not extract the local traveltime information, whcih may increase the difficulty of event-to-event traveltime fitting. Nevertheless, DIW can preserves the local variation of traveltime differences according to equation (3), which have a higher resolution than the correlation-based method.

To investigate the behaviors of DIW and correlation-based objective function, we use three models of different complexity as shown in figure 7. Since the estimated PS traveltime residual would be wrong if initial model is far from the true value with method II, we only investigate the behavior of PP traveltime misfit function, which is almost the same as the acoustic case. In this investigation, a constant background velocity ($V_p = 2500m/s$) model and different reflectivity model are used to implement the Born modeling. Then, we perturb the background velocity to calculate the objective function through migration/demigration process.

In figure 7b, we can see that for the simple one-reflector case, the temporal correlation objective function has a better convexity than the spatial correlation and DIW method. But when there are multi-reflectors (figure 7c and d), the temoral correlation fails to find the correct minimal due to the multi-events in the recordings. While the behavior of spatial correlation and DIW method are robust to locate the correct solution. In the third test, we choose a part of Sigsbee2A model to simulate a complicated case. Figure 7e shows that the temporal correlation still fails to find the correct solution and the DIW even has a better

convexity than the spatial correlation, especailly when the initail velocity is high. The above investigation indicates that DIW method is robust and effective to extract traveltime residual, especially for the complicated model.

NUMERICAL EXAMPLE

We select a part of the Sigsbee2A model (Figure 6a and b) to test the inversion algorithm and strategy. The V_s model is generated through V_p model with a constant scale factor of 0.66. The initial model for E-WERTI are shown in Figure 6c and d, which increase linearly with depth. The initial model of V_p is generally lower (from 1500m/s to 1996m/s) while V_s is higher (from 990m/s to 1317m/s) than the true model. 36 shots are evenly deployed on the surface and receivers are fixed with a maximum offset of 4km. The main frequency of P-wave source is 15Hz. The spatial and time interval of forward modeling is 16m and 1.2ms, respectively.

Figure 8 is the results of ERTM and EFWI with the initial model. Since the V_p and V_s initial model are far from the true value, both the PP image (δV_p) and PS image (δV_p) are wrongly located even though only the small offset data are applied to the ERTM. Especially, the refracions are not correctly migrated in PP image and faults are almost disappear in PS image. Meanwhile, we try the EFWI with the above initial model. During the EFWI, a four-stage hierarchical strategy from low to high frequency is applied through the time-domain low-pass filter, in which the frequency bands are 0-2Hz, 0-4Hz, 0-6Hz and 0-8Hz. We can see the showllow parts of inverted results is acceptable due to the exsistence of diving wave and low-frequency data. However, the inversion of deep part suffer from the severe cycle-skipping problem because of the absence of low-to-intermediate components in the initial model, which leads to the failure of inversion.

16

Starting from this initial model, we implement the proposed ERTI workflow. During the inversion, the direct waves are muted to make sure that only the reflection data are involved. Figure 9a and 9b show the inverted V_p and V_s model of ERTI. After 40 iterations for each stage, ERTI provides a good recovery of the low-to-intermediate components of V_p and V_s model. Nonetheless, on the right part, the reflection coverage of surface observation is insufficient for ERTI to rebuild the long-wavelength components.

Note, the structure-oriented smooth filter is an appropriate regularization to stabilize and even accelerate convergence to models which are a priori more satisfying than others which fit the data. In each iteration of ERTI, we constrain the gradients of ERTI with the seismic image obtained by ERTM. Figure 12b shows the gradient of V_p smoothed by an isotropic filter in the first iteration. Though there is no high wavenumber artifacts, the illumination footprints of acquisition geometry are obvious, which will cause uncertainties during inversion. After the structure-oriented smooth, the gradient is more balanced for the inversion of background velocity. And the inverted results of the two methods also illustrate advantage of structrue-oriented smooth filter.

Using the inverted results of ERTI as starting models, we also perform the conventional ERTM. As shown in Figure 10c and d, since the low-to-intermediate wavenumber components of V_p and V_s are well recovered, both of the PP and PS image are greatly improved. The location of reflectors and faults are almost correct and the refractions are migrated as well. However, the image of the right part is not so good because the recovery of low-to-intermediate components are not well in this area . In addtion, figure 11 shows the comparison between the demigrated PP and PS reflections using initial and inverted model with the true model. We can see that the main reflections of both PP and PS are well matched after the ERTI. Therefore, the cycle-skipping problem will be mitigated greatly if

using the starting model in figure 9 for EFWI. Accordingly, we implement the EFWI using the inverted results as starting models with the same hierarchical strategy as the previous example (figure 8). Figure 13 show the inverted results of conventional EFWI. Comparing with the inverted results using initial models (figure 8), EFWI improves quite a lot, especially in the deep part. The vertical profiles at 1.4km and 3km validate the effectiveness of EFWI with the new starting model as well. But the inversion of the right part is not as good as the left part because of the mentioned unreasonable recovery for background model during ERTI.

Usually, the low-frequency components in the data and the low-to-intermediate wavenumber in the initial model are two key points to mitigate the cycle-skipping problem in FWI. In real data, the low frequency components are difficult to obtain and also more easily contaminated by the noises. Thus, the success of FWI more relies on the good intial models. In this paper, we will test the dependency of EFWI on the low-frequency data when using the above ERTI results. The low-cut frequency threshold are 3Hz, 5Hz and 7Hz. The components lower than the threshold are filtered out when applying the similar hierarchical strategy from low to high frequency in time domain. As shown in figure 15, for a starting frequency of 3Hz, EFWI provides almost the same inverted results with the no-low-frequency-cut test (figure 13). On the left part, the inverted results are acceptable even when starting from 7Hz, which means the low-to-intermediate components in this area is particularly good for EFWI. However, on the right part, the inversion are more dependent on the low-frequency components in the data due to the improper recovery of the background components.

CONCLUSIONS

We extend the wave equation reflection travletime inversion towards elastic media to build the low-wavenumber component of elastic model. RTI can recover the low-wavenumber components effectively, because time misfits are more sensitive and linearly related to the low-wavenumber model perturbation. Compared with the correlation based method, DIW method provides more stable estimation of traveltime shifts especially when seismograms contain multi-reflections. The complicated mode conversions in elastic case make the reflection kernels very complex. The kernel analysis of different wave modes show that mode decomposition can suppress artifacts and recover the correct reflection wavepath in gradient calculation. With the aid of P/S separation of 3C seismograms, we can obtain the separated PP and PS seismograms and then get travel time differences of PP and PS through DIW, respectively. To build the long-wavelength component of the model, we introduce a two-stage ERTI workflow by firstly using PP then PS reflections, through which the nonlinearity of reflection inversion is reduced effectively. In the second stage, the wave mode decomposition is introduced to calculate the gradient of V_s to mitigate the trade-off between V_p and V_s. The Sigsbee2A model example shows that even starting with a bad initial model, the two-stage E-WERTI can provide reliable starting model for conventional EFWI.

ACKNOWLEDGEMENT

REFERENCES

Almomin, A., and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: Presented at the SEG Technical Program Expanded Abstracts 2012, Society of Exploration Geophysicists.

Biondi, B., and A. Almomin, 2013, Tomographic full-waveform inversion (TFWI) by combing FWI and wave-equation migration velocity analysis: The Leading Edge, **32**, 1074–1080.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2d elastic frequency-domain full-waveform inversion: Geophysics, **74**, WCC105–WCC118.

Chi, B., L. Dong, and Y. Liu, 2015, Correlation-based reflection full-waveform inversion: Geophysics, **80**, R189–R202.

Guo, Q., and T. Alkhalifah, 2016, *in* A nonlinear approach of elastic reflection waveform inversion: Society of Exploration Geophysicists, 1421–1425.

Hale, D., 2013, Dynamic warping of seismic images: GEOPHYSICS, **78**, S105–S115.

Jones, I., 2010, Tutorial: Velocity estimation via ray-based tomography: First Break, **28**.

Li, Z., X. Ma, C. Fu, B. Gu, and G. Liang, 2016, Frequency-wavenumber implementation for p- and s-wave separation from multi-component seismic data: Exploration Geophysics, **47**, 32.

Ma, Y., and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full waveform inversion: Geophysics, **78**, R223–R233.

Oh, J.-W., and T. Alkhalifah, 2016, The scattering potential of partial derivative wavefields in 3-D elastic orthorhombic media: an inversion prospective: Geophysical Journal International, **206**, 1740–1760.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss–newton and full newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Prieux, V., R. Brossier, S. Operto, and J. Virieux, 2013, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the valhall field. part 2: imaging compressive-wave and shear-wave velocities: Geophysical Journal International, ggt178.

Raknes, E. B., and W. Weibull, 2016, Combining wave-equation migration velocity analysis and full-waveform inversion for improved 3D elastic parameter estimation, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists.

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

Shipp, R. M., and S. C. Singh, 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: Geophysical Journal International, **151**, 325–344.

Sun, D., and W. W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: Presented at the SEG Technical Program Expanded Abstracts 2012, Society of Exploration Geophysicists.

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

Tarantola, A., 1984, 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.

van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Wang, C., W. Weibull, J. Cheng, and B. Arntsen, 2017, Automatic shear-wave velocity analysis with elastic reverse time migration: 79th EAGE Conference and Exhibition 2017, Expanded Abstracts.

Wang, H., S. C. Singh, F. Audebert, and H. Calandra, 2015a, Inversion of seismic refraction and reflection data for building long-wavelength velocity models: Geophysics, **80**, R81–R93.

Wang, T., and J. Cheng, 2017, Elastic full waveform inversion based on mode decomposition: the approach and mechanism: Geophysical Journal International, **209**, 606–622.

Wang, T., J. Cheng, and C. Wang, 2015b, Elastic wave mode decoupling for full waveform inversion: 77th EAGE Conference and Exhibition 2015, Expanded Abstracts.

Wang, T., W. Kang, and J. Cheng, 2014, Migration velocity model building using local angle domain nonlinear tomography: Presented at the Beijing 2014 International Geophysical Conference and Exposition, Beijing, China, 21-24 April 2014, Society of Exploration Geophysicists.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

Woodward, M. J., D. Nichols, O. Zdraveva, P. Whitfield, and T. Johns, 2008, A decade of tomography: Geophysics, **73**, VE5–VE11.

Wu, Z., and T. Alkhalifah, 2015, Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion: Geophysics, **80**, R317–R329.

Xu, S., D. Wang, F. Chen, G. Lambare, and Y. Zhang, 2012, Inversion on reflected seismic wave: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1–7.

Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2015, Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation: Geophysical Journal International, **202**, 1535–1554.
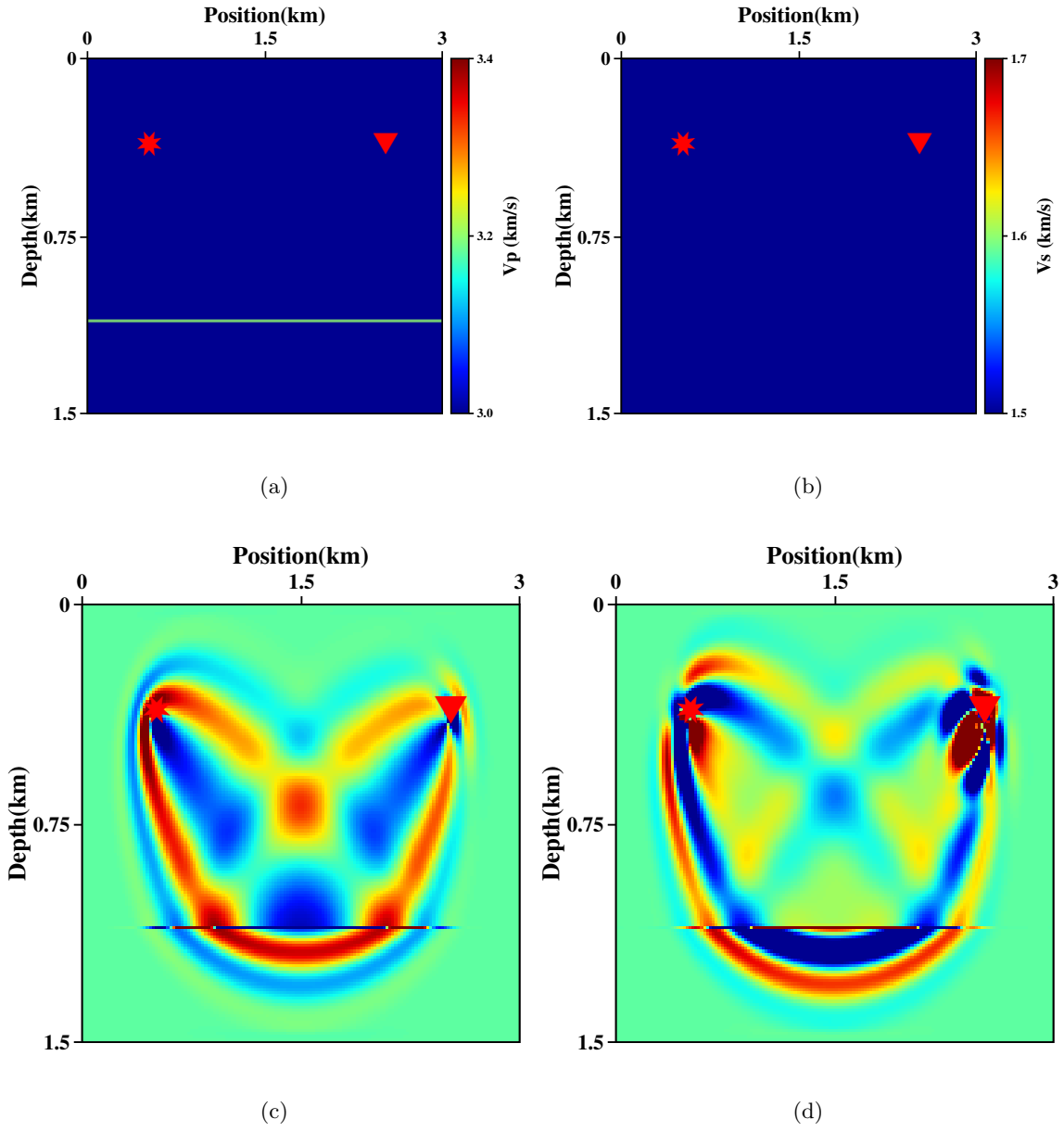
Figure 1: Kernels with single reflector in V_p model. (a) V_p model, (b) V_s model, (c) K_{V_p}, (d) K_{V_s}.
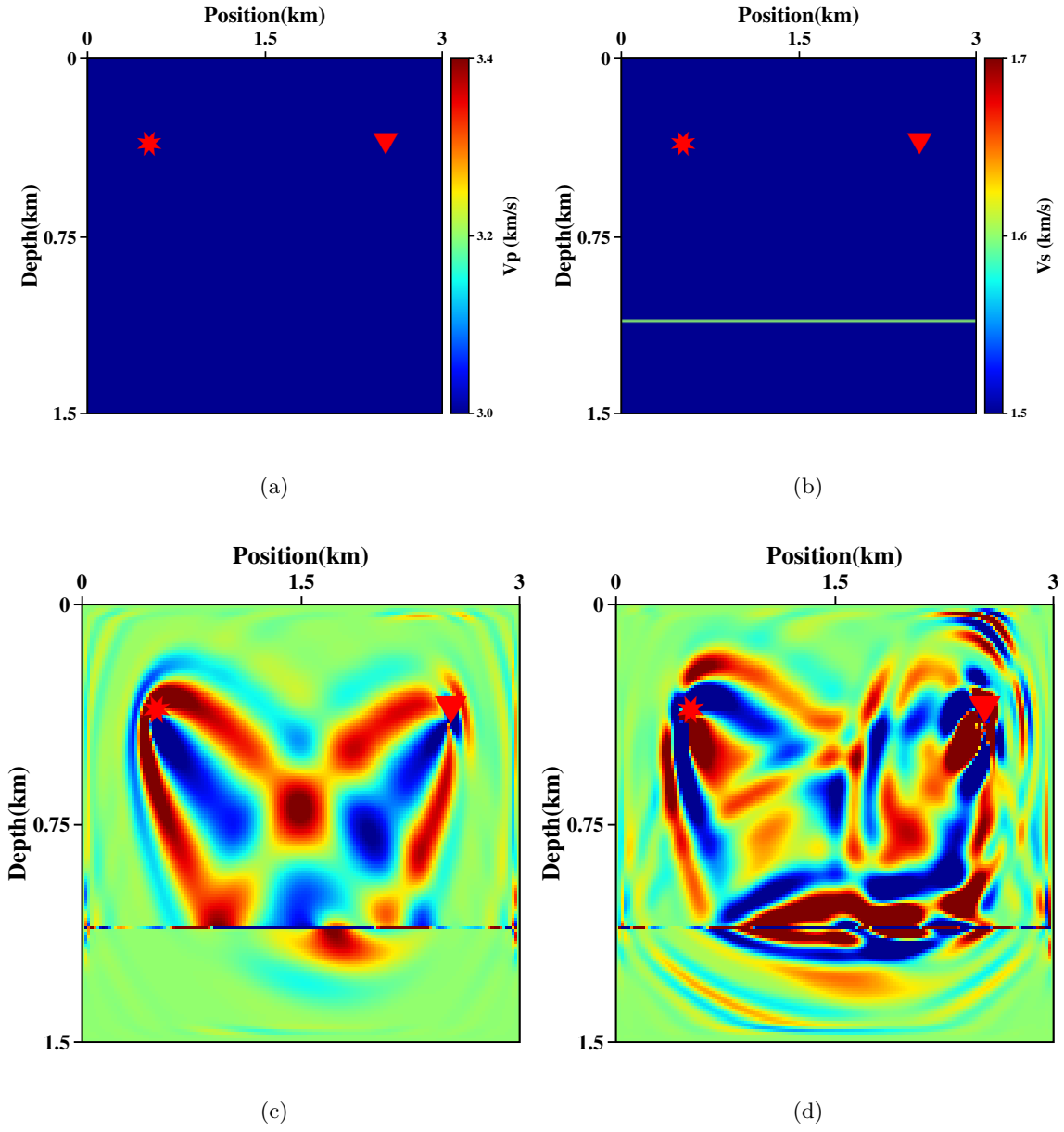
Figure 2: Kernels with single reflector in V_s model. (a) V_p model, (b) V_s model, (c) K_{V_p}, (d) K_{V_s}.

(a)

(b)

(c)

(d)
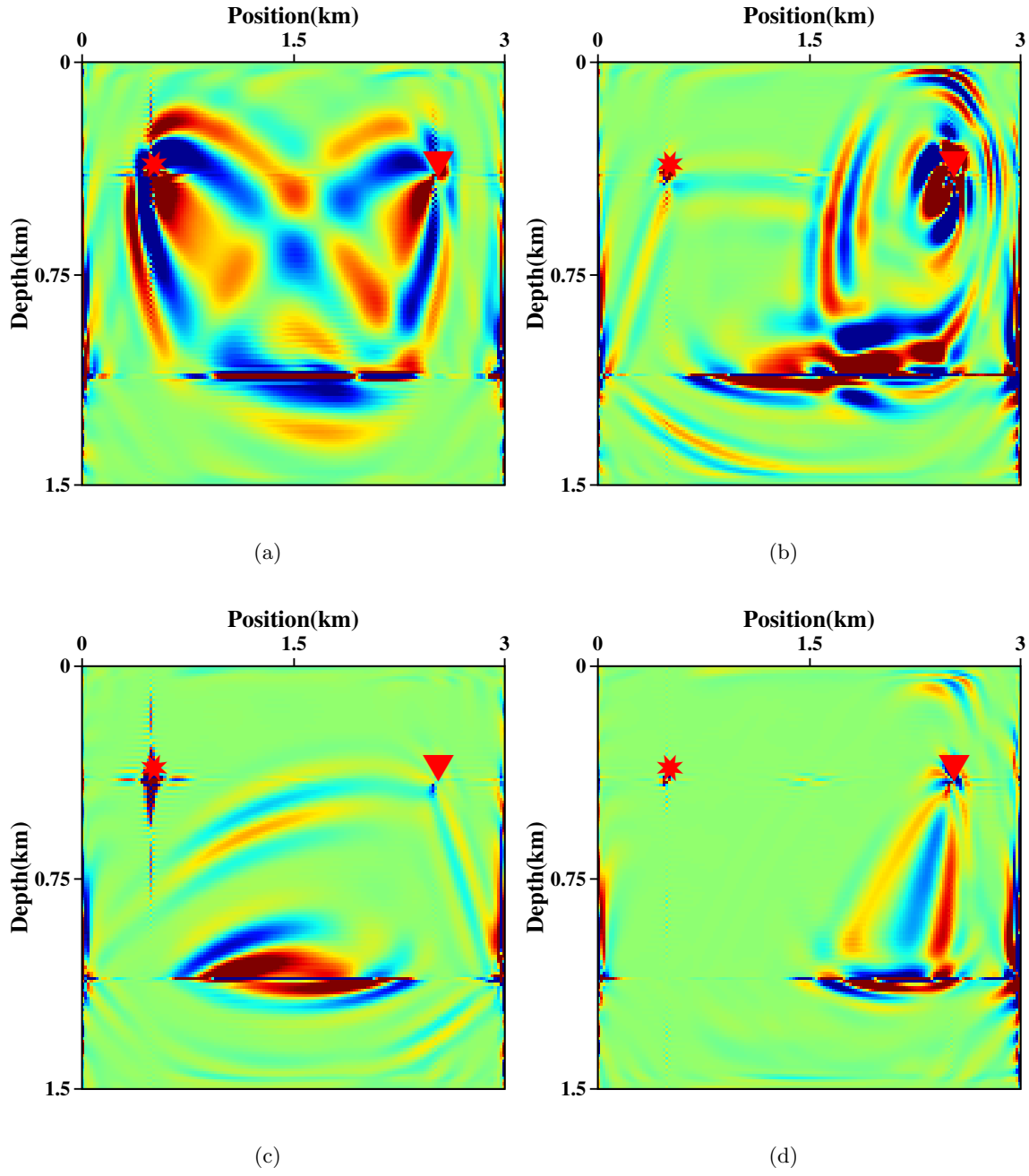
Figure 3: Four components of K_{V_s}. (a) $K_{V_s}^{PP}$, (b) $K_{V_s}^{PS}$, (c) $K_{V_s}^{SP}$, (d) $K_{V_s}^{SS}$.
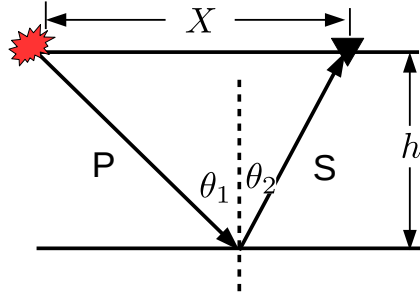
Figure 4: Ray path of PS reflection with a single reflector.

(a)

(b)
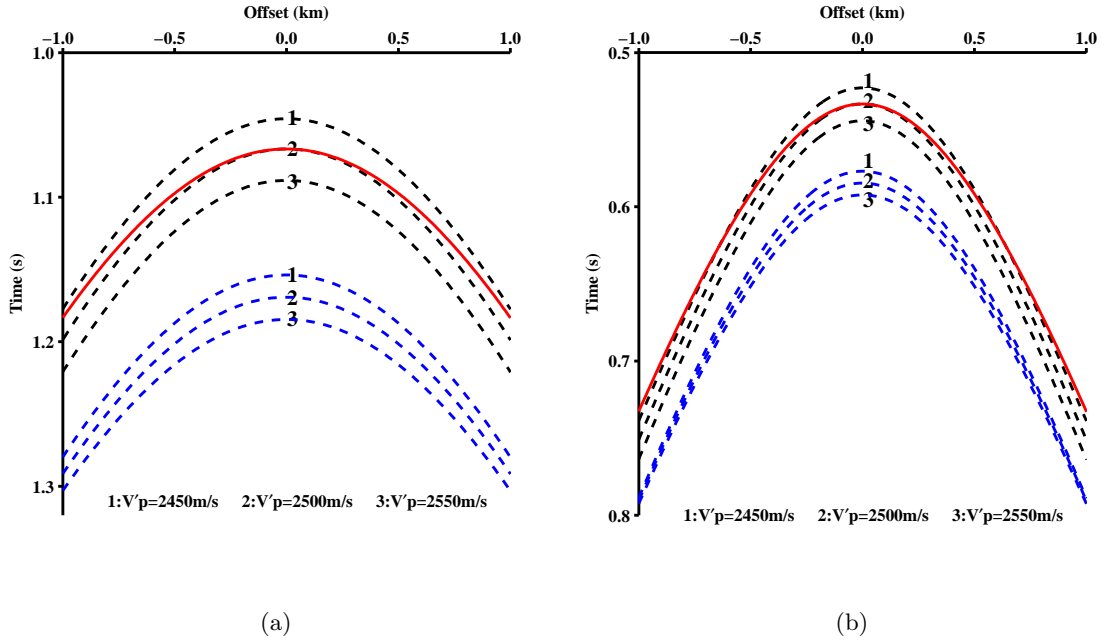
Figure 5: The comparison among demigrated PS traveltime moveout of method I (black), method II (blue) and the real (red) one. (a) $h = 1000m$, (b) $h = 500m$.
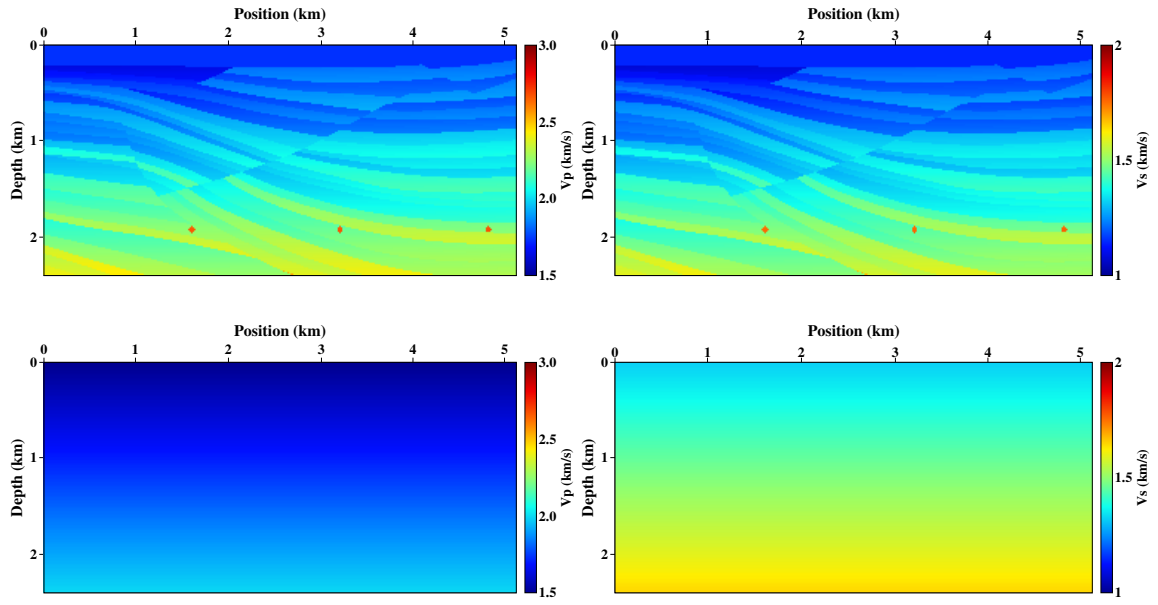
Figure 6: Sigbee2A model example. On the top are true models of V_p (a) and V_s (b). On the bottom are initial models of V_p (c) and V_s (d) linearly increasing with depth.
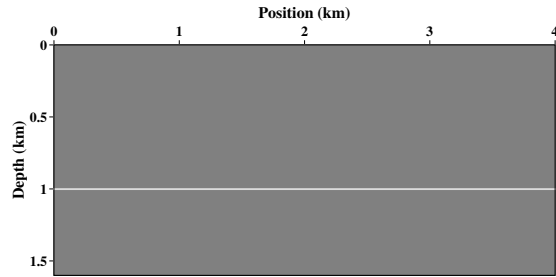
(a)
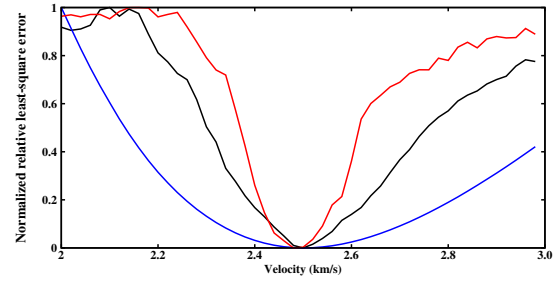
(b)
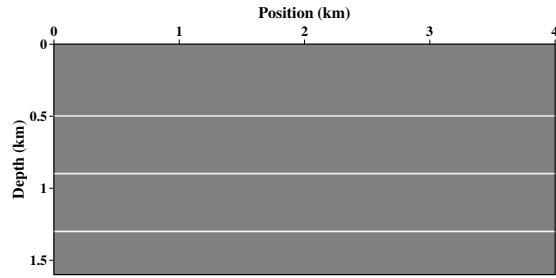
(c)
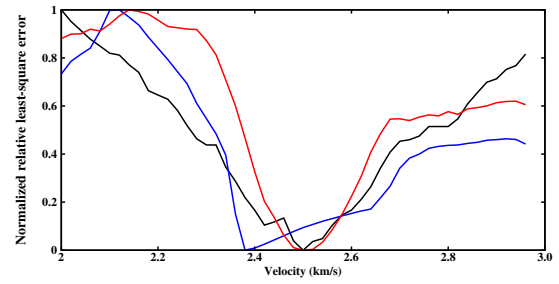
(d)
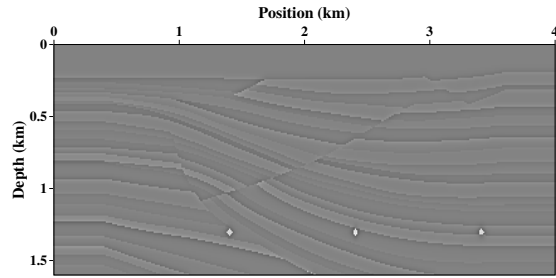
(e)
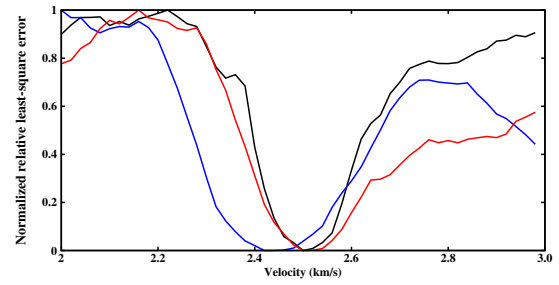
(f)

Figure 7: Normalized PP traveltime misfits for different objective functions for different model;
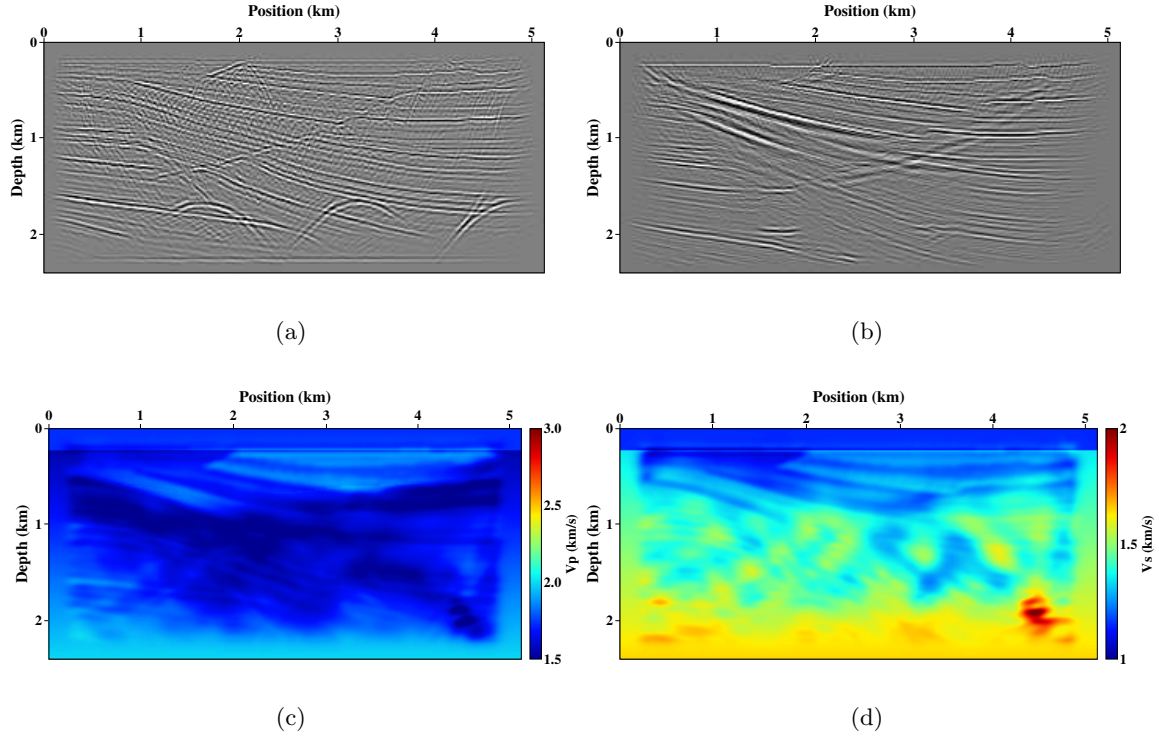
Figure 8: The results of ERTM and EFWI using initial model: (a) and (b) are PP and PS image of ERTM with near offset data, (c) and (d) are inverted V_p and V_s with EFWI.
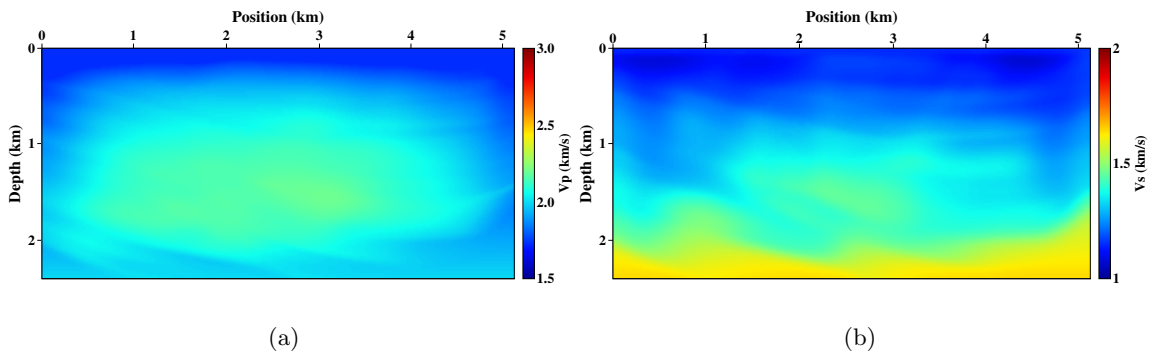
Figure 9: Inverted results of ERTI: (a) V_p, (b) V_s.
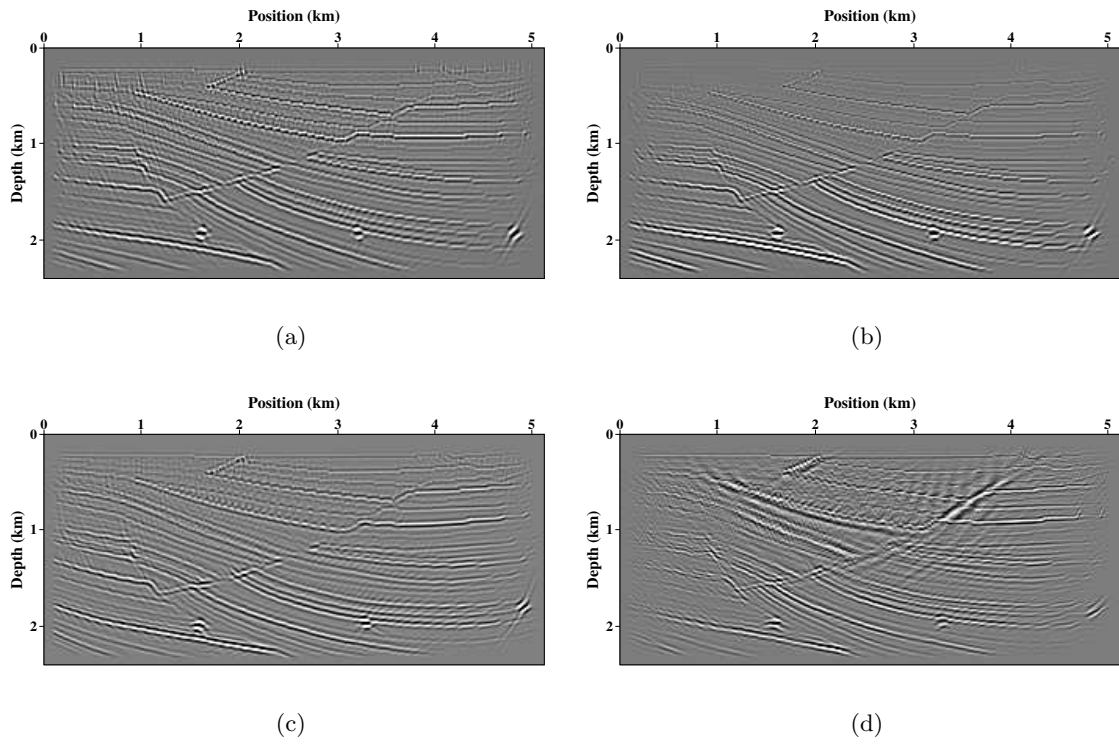
Figure 10: ERTM results using the true model (a, b) and inverted model (c, d). (a) and (c) are the PP image(b) and (d) are the PS image
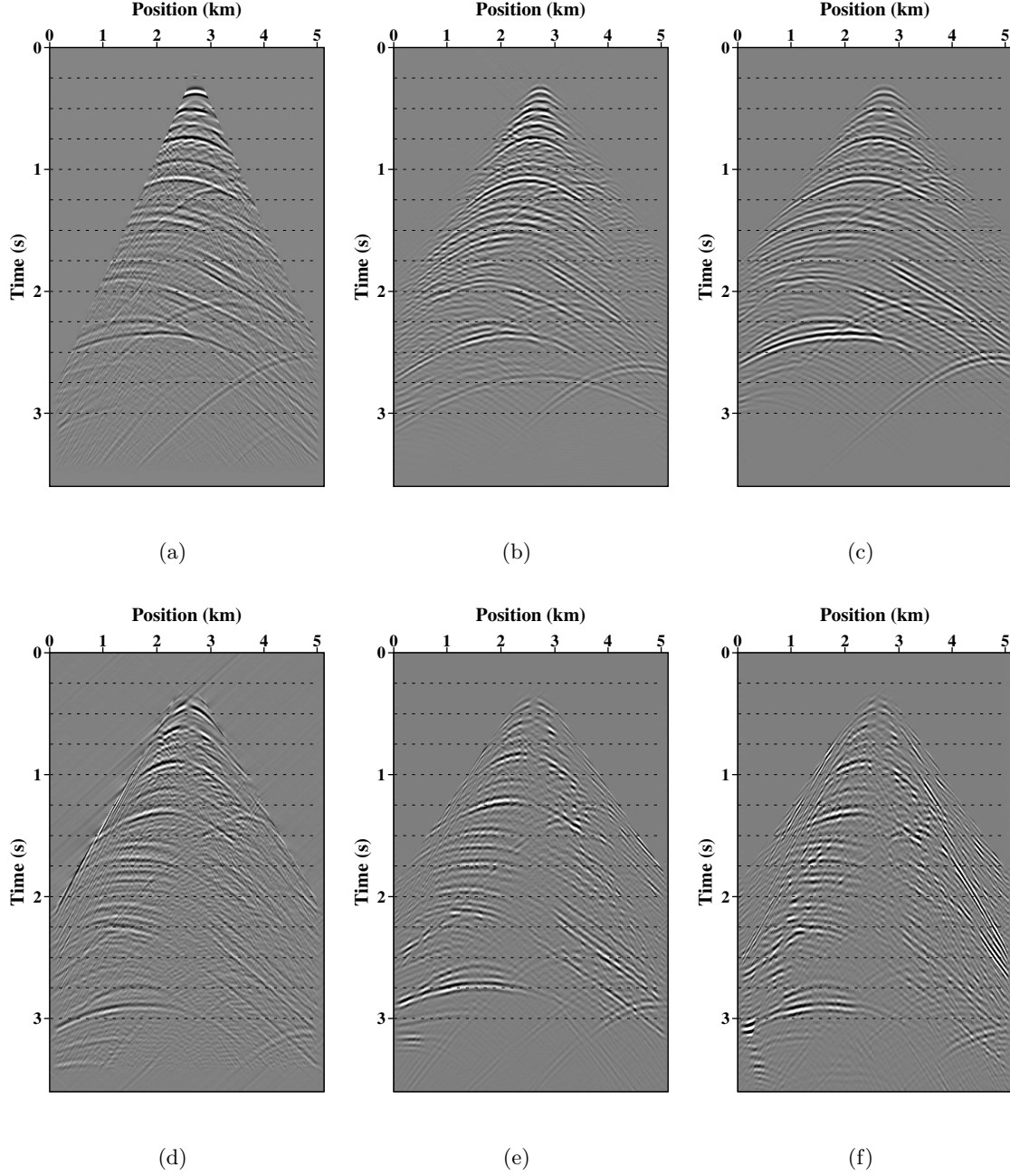
Figure 11: Comparison of the observed and the demigrated reflection data using initial model and the inverted model. The first row are the separated PP reflection, while the second row are the separated PS reflection. The left, middle and right column are the observed reflection data, the demigrated reflection data with initial model and the demigrated data with inverted model, respectively.

33
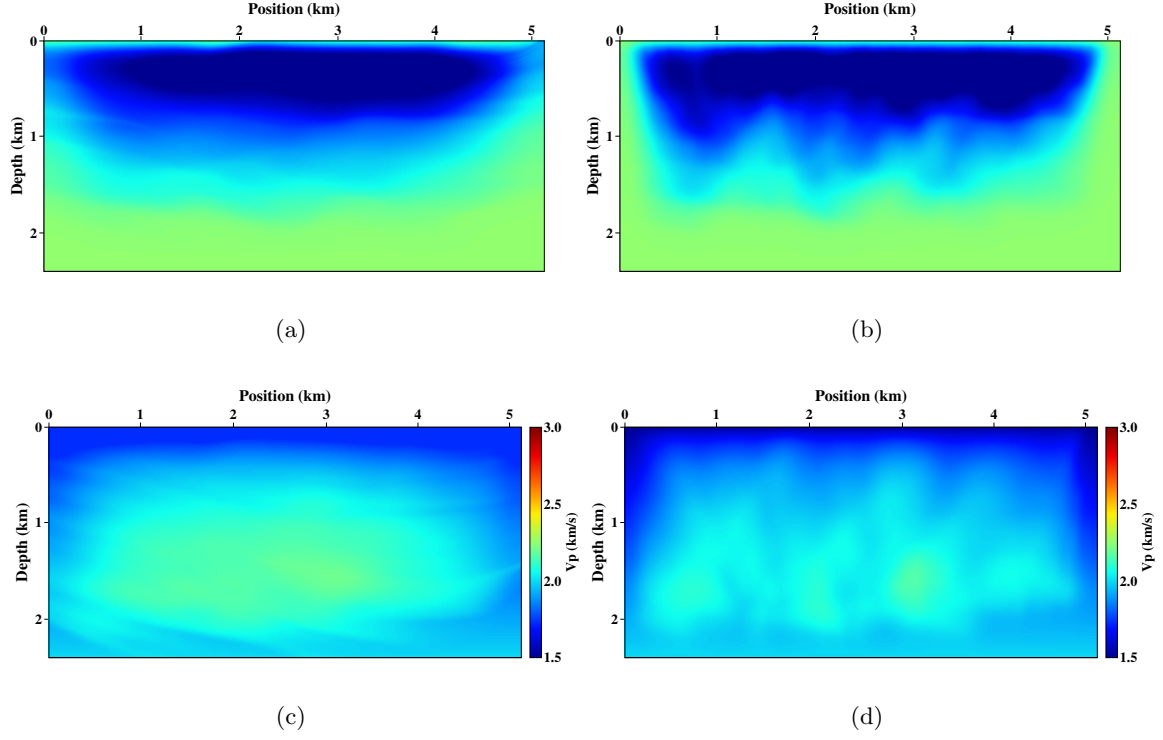
Figure 12: The comparison of gradients and final inverted results with (left) and without (right) the structure-oriented constrain: (a) and (c) are the V_p gradients in the first iteration, (b) and (d) are final inverted V_p model.
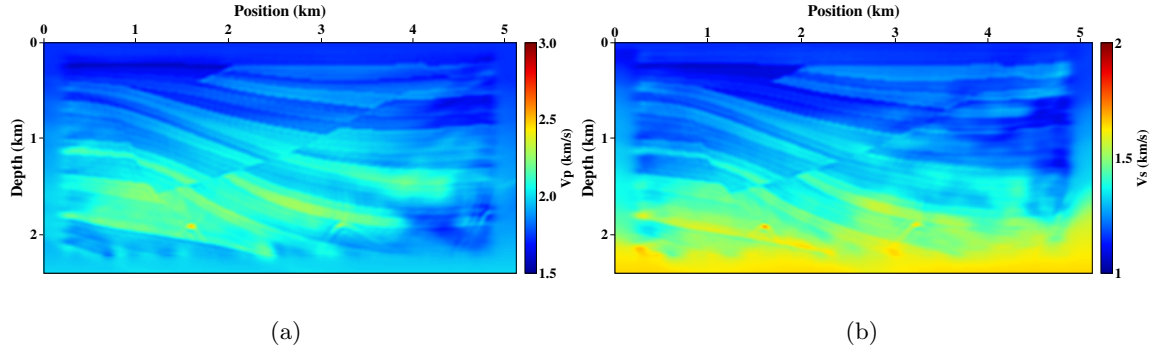
Figure 13: EFWI results using the ERTI model as starting model. (a) V_p, (b) V_s.
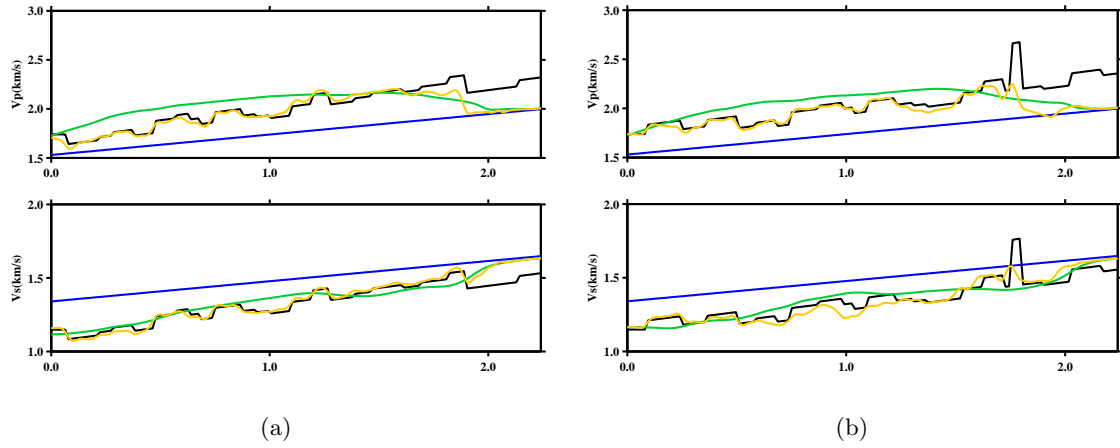

Figure 14: Vertical profile of ERTI and EFWI model at 1.4km (a) and 3km (b). The black and blue lines denote the true and initial model. The green and yellow denote the ERTI and EFWI model.
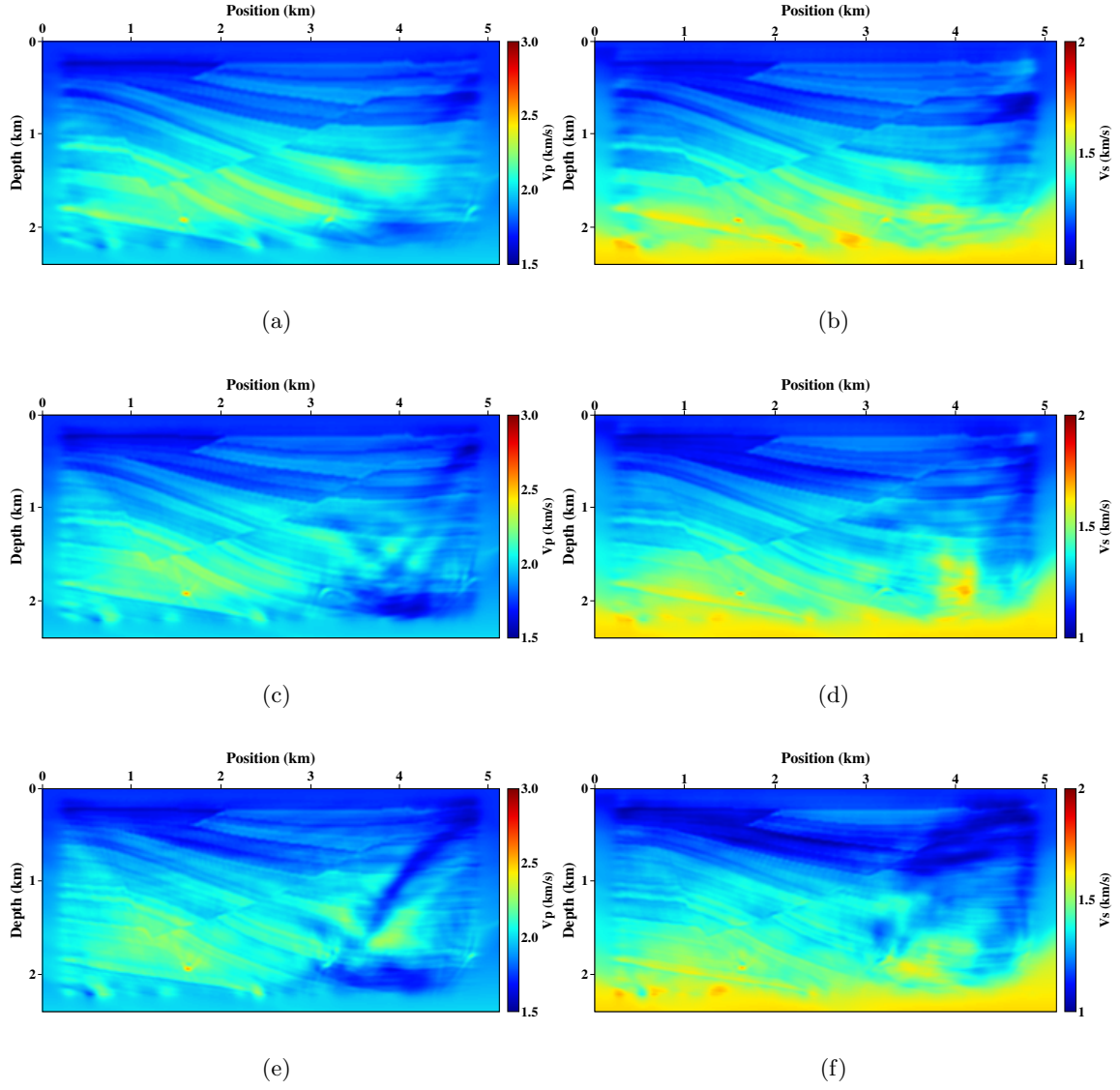
Figure 15: EFWI results with different starting frequency. (a), (c), (e) are V_p, (b), (d), (f) are V_s. The starting frequency are 3Hz, 5Hz and 7Hz from top to bottom row, respectively.