Rock physics constrained anisotropic WEMVA: Part I - Theory and synthetic test

Yunyue (Elita) Li, Robert Clapp, Biondo Biondi, and Dave Nichols

ABSTRACT

We present a regularization scheme utilizing available rock physics data to better constrain the anisotropic wave-equation migration velocity analysis (WEMVA) and to better resolve the ambiguity among the anisotropic parameters. In addition to the spatial covariance to constrain the spatial correlation of each VTI parameter individually, we propose a cross-parameter covariance at each subsurface location to link the VTI parameters. There are two significant effects that this regularization scheme brings to the updates for the VTI parameters. First, instead of spreading the updates evenly along the wavepath, the regularization term allows more updates in the regions where the models are highly uncertain. Second, the regularization term brings extra information for parameter updates from the correlation with the other parameters. These improvements help the inversion converge faster and yield VTI models that are more consistent with the underlying geological and lithological assumptions. We demonstrate these improvements on a synthetic dataset.

INTRODUCTION

Anisotropic model building tries to resolve more than one parameter at each model location. This number is three for a vertical transverse isotropic (VTI) medium, and five for a tilted transverse isotropic (TTI) medium. Traditional surface seismic tomography may be able to produce an accurate isotropic earth model efficiently for a large area when the acquisition is dense and the earth is well-illuminated by rays at a wide range of angles. However, surface seismic data inversion becomes ill-posed and highly underdeterimined due to the rapidly increasing model space with the increasing complexity of the subsurface.

One important disadvantage of surface seismic tomography is its lack of depth information. During tomography, neither the low wavenumber velocity nor the high wavenumber reflectivity are known. This issue is more severe when we consider anisotropy. Several localized tomography around the wells were studied to add the depth dimension into the inversion (Bear et al., 2005; Bakulin et al., 2010b,a). Joint inversion of surface seismic data and borehole data (check-shots and walkaway VSPs) in these studies helps to yield better defined earth models. Due to the ambigu-

ity among the parameters, it is difficult to resolve a reliable and unique anisotropic model in 3D even with the borehole aided localized tomography (Bakulin et al., 2009).

In the anisotropic case, the uncertainties are further increased because the sensitivities of the kinematics of the surface seismic data to the anisotropic parameters are much lower than to velocity. Large offsets and a wide range of illumination angles are required to constrain the anisotropic parameters. Consequently, the recoverable depth range for the anisotropic parameters is much shallower than when a simple isotropic velocity is estimated. Even in the shallow region, where the seismic waves travel with wide angles and large offsets, the kinematic effects of the velocity can still overwhelm the inversion.

To help with the inversion for anisotropy, we can use our prior knowledge of the subsurface. In addition to the two-point (spatial) covariance describing the smoothness in the subsurface (Clapp, 2000; Woodward et al., 2008; Li and Biondi, 2011), a single point (local) cross-parameter covariance can be used to better describe the subsurface. One way to estimate the cross-parameter covariance is from rock physics studies (Hornby et al., 1995; Sayers, 2004, 2010; Bachrach, 2010b). Many authors (Dræge et al., 2006; Bandyopadhyay, 2009; Bachrach, 2010a) have built depth trends for seismic purposes. In particular, Bachrach (2010a) developed both deterministic and stochastic modeling schemes based on the rock physics effective-media models for compacting shale and sandy shale. The stochastic modeling enables us to explore the range of possible anisotropic parameters based on the rock physics modeling parameters. Further corroborated by core measurements, the parameters allowed by the rock physics model are limited to a certain range, which greatly reduces the possible range of the VTI parameters. These rock physics modeling results can be used to construct initial VTI models and the covariance relationships among the VTI parameters. Li et al. (2011) and Yang et al. (2012) have demonstrated that the rock physics prior models are helpful in constraining ray-based tomography.

In this paper, we first analyze the sensitivities of the WEMVA objective function with respect to different VTI parameters in a homogeneous VTI model. Using an interpolation example, we demonstrate the additional information that the cross-parameter covariance brings to the inversion. We then test three different regularization schemes when a synthetic dataset is inverted using anisotropic WEMVA. The inversion results show that when the accurate full covariance matrix is applied, the convergence of the WEMVA inversion as well as the lithological definition of the inverted models are improved.

WEMVA AND ROCK-PHYSICS REGULARIZATION

Anisotropic WEMVA aims at building an anisotropic earth model that minimizes the residual image from the surface seismic data (Li and Biondi, 2011). This optimization problem is highly nonlinear and underdetermined. We commonly add a model regularization term to the anisotropic WEMVA objective function defined in the image

space to constrain the null space and stabilize the inversion. Our resulting objective function is:

$$S(\mathbf{m}) = \frac{1}{2}||\mathbf{D}_\theta \mathbf{I}(\mathbf{x},\theta)|| - \alpha\frac{1}{2}||\sum_\theta \mathbf{I}(\mathbf{x},\theta)|| + \beta\frac{1}{2}(\mathbf{m}-\mathbf{m}_{\text{prior}})^T\mathbf{C}_M^{-1}(\mathbf{m}-\mathbf{m}_{\text{prior}}), \quad (1)$$

where the first two terms define the "data fitting" objective, and the third defines the "model regularization" objective. The first term is to minimize the differential semblance in the data fitting objective, and the second term is to maximize the stacking power. Model $\mathbf{m}$ is the VTI subsurface model, $\mathbf{I}(\mathbf{x},\theta)$ is the migration image in the angle domain with θ the aperture angle and $\mathbf{D}_\theta$ a derivative operator along the angle axis. In the model regularization objective, $\mathbf{m}_{\text{prior}}$ and $\mathbf{C}_M$ define a Gaussian distribution of a prior model that is ideally independent of the seismic data. This regularization will bring more information into the optimization. Parameters α and β balance the relative weights among different objectives.

The data fitting objective relates the incoherence in the angle domain common image gathers to the inaccuracy in the subsurface models. To test the data objectives, we model a simple synthetic dataset using a homogeneous VTI model ($v_v = 2$ km/s, $\epsilon = 0.2$, and $\delta = 0.1$) with one flat reflector. The maximum offset is 6 km, and the depth of the reflector is 1.5 km. The data are then migrated using all possible combinations of v_v, ϵ, and δ when v_v varies in $[1.5, 2.5]$ km/s, $\epsilon \in [0.1, 0.3]$, and $\delta \in [0, 0.2]$. Based on the migrated images in the angle domain, the data fitting objective at each model point in this subspace is then computed according to the first two terms in Equation 1.

Figure 1 shows the data fitting objective function assuming one of the three VTI parameters is accurate. Panels (a), (b), and (c) are extracted from the $v_v = 2$km/s plane, $\delta = 0.1$ plane, and $\epsilon = 0.2$ plane, respectively. When the third parameter is accurate, the data fitting objective function is convex and comes to a minimum at the correct solution for the other two parameters. However, the resolution of the objective function with respect to each parameter is dramatically different. A much higher resolution for velocity than for ϵ and δ can be seen from Figures 1(b) and (c). The resolution of ϵ and δ is at least an order of magnitude lower than that of velocity (notice the smaller range of values on the colorbar in Figure 1(a)).

Moreover, severe tradeoffs among the VTI parameters can be observed from Figure 1. A strong tradeoff can be seen in Figure 1(a). The WEMVA objective function cannot resolve ϵ or δ independently as long as the summation of the two remains the same (unless events propagating at more than 60° angles are recorded). Tradeoffs between the anisotropic parameters and the vertical velocity can also be seen from Figures 1(b) and 1(c), where positive vertical velocity errors are compensated by negative errors in the anisotropic parameters and vice versa. Nevertheless, these tradeoff effects are much less obvious, further demonstrating the low sensitivity of seismic data to the anisotropic parameters.

For pressure waves, velocity has a dominant effect on the kinematics due to its first-order influence. As a result, the objective function is biased towards the velocity

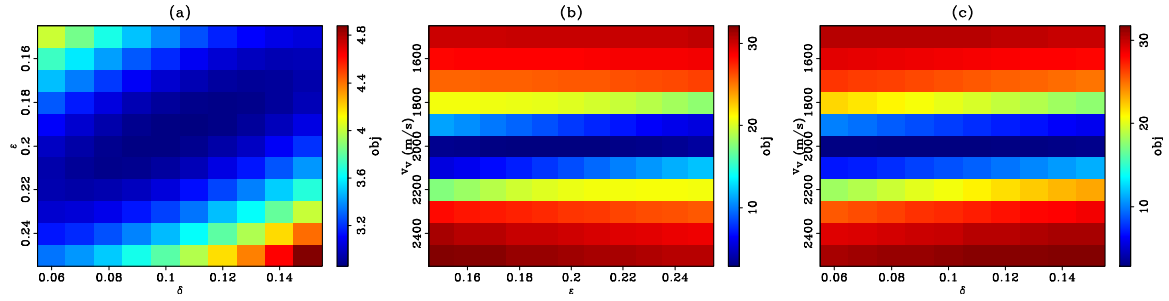

Figure 1: The value of data fitting objective function extracted from (a) $v_v = 2$ km/s plane, (b) $\delta = 0.1$ plane, and (c) $\epsilon = 0.2$ plane. Notice the different color scale in panel (a). [**CR**]

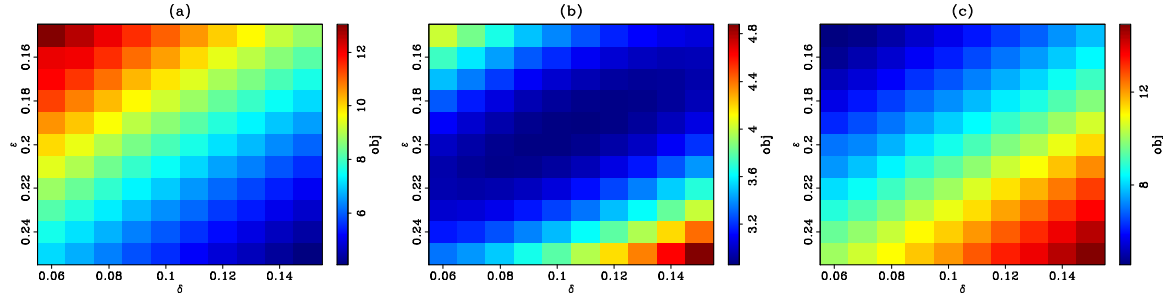

Figure 2: The value of data fitting objective function extracted from (a) $v_v = 1.9$ km/s plane, (b) $v_v = 2$ km/s plane, and (c) $v_v = 2.1$ km/s plane. [**CR**]

error despite the error in the other VTI parameters. Figure 2 shows the topography of the data fitting objective function extracted from the (a) $v_v = 1.9$ km/s plane, (b) $v_v = 2$ km/s plane, and (c) $v_v = 2.1$ km/s plane, respectively. When the velocity is inaccurate, the objective function loses its convexity in the ϵ–δ plane. The objective function always dips towards higher ϵ and higher δ when velocity is slower, and towards lower ϵ and lower δ when velocity is faster. In these cases, the gradient of the objective function would not be able to guide the inversion to the correct solution for the anisotropic parameters if the errors in ϵ and δ are in the opposite direction than the velocity error.

For field data, the "data fitting" objective function has worse behavior because the structure of the earth subsurface is highly complex with heterogeneities at all scales and densely distributed dipping reflectors. These complexities weaken the convexity of the objective function and create local minima.

To better constrain the model and to mitigate severe ambiguities among the VTI parameters, we include additional information in the inversion. We utilize the regularization term in Equation 1 by assuming a multivariate Gaussian distribution for the VTI model parameters. To speed up the convergence, we use a preconditioning scheme instead of the original regularization scheme.

WEMVA with the rock physics regularization

Assuming a Gaussian distribution, Tarantola (1984) characterizes the prior information using the mean and the covariance of the model and includes it as a regularization term. We separate the covariance into two parts: a spatial covariance between the same parameter at different locations, and a cross-parameter covariance between different parameters at the same location.

The spatial covariance is mainly defined by the structure of the area, and it can be estimated using a set of steering filters (Clapp, 2000). The cross-parameter covariance can be inferred from the lithological information at the model location. In a previous work (Li et al., 2013), we introduced the cross-parameter matrix with only diagonal terms due to the lack of the lithological information. This diagonal preconditioning balances the relative scales among the VTI parameters, but it ignores the correlation among them.

When we have a rough estimate of the lithological environment, we can build a more complete cross-parameter covariance using rock physics modeling. As an example, the Gulf of Mexico is populated with large shale deposits. Therefore, rock physics principles can be used to estimate the range of anisotropic parameters considering a compacting shale model (Bachrach et al., 2011). As demonstrated in (Li et al., 2014), the rock physics stochastic modeling for shale anisotropy shows that in the shallow sediments, a high velocity zone is generally collocated with sand layers where the anisotropy is low. On the contrary, diagenesis processes in the more compacted deeper sediments alter the clay mineral from smectite to illite, which increases

both velocity and anisotropy simultaneously. Valuable prior knowledge about the subsurface can be included in the inversion via both the diagonal and the off-diagonal terms in the cross-parameter covariance matrix.

We assume the spatial covariance and local cross-parameter covariance components are independent of each other (Li et al., 2011). To speed up the convergence, we use a preconditioning scheme where we use steering filters to approximate the square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the preconditioning variable $\mathbf{n}$ is related to the original model $\mathbf{m}$ as follows:

$$\mathbf{m} = \mathbf{\Sigma S n}. \tag{2}$$

In Equation 2, the smoothing operator $\mathbf{S}$ is a band-limited diagonal matrix:

$$\mathbf{S} = \begin{vmatrix} \mathbf{S}_v & 0 & 0 \\ 0 & \mathbf{S}_\epsilon & 0 \\ 0 & 0 & \mathbf{S}_\delta \end{vmatrix}, \tag{3}$$

with potentially different smoothing operators for velocity, ϵ, and δ, according to the geological information in the study area. The standard deviation matrix $\mathbf{\Sigma}$ is the square-root of the covariance matrix:

$$\mathbf{\Sigma} = \begin{vmatrix} C_{vv}\,\mathbf{I} & C_{v\epsilon}\,\mathbf{I} & C_{v\delta}\,\mathbf{I} \\ C_{\epsilon v}\,\mathbf{I} & C_{\epsilon\epsilon}\,\mathbf{I} & C_{\epsilon\delta}\,\mathbf{I} \\ C_{\delta v}\,\mathbf{I} & C_{\delta\epsilon}\,\mathbf{I} & C_{\delta\delta}\,\mathbf{I} \end{vmatrix}^{1/2}. \tag{4}$$

The diagonal elements C_{vv}, $C_{\epsilon\epsilon}$, and $C_{\delta\delta}$ denote the variance of the velocity, ϵ, and δ, respectively. The off-diagonal elements $C_{v\epsilon}$, $C_{v\delta}$, and $C_{\epsilon\delta}$ denote the cross-variance between the velocity and ϵ, between velocity and δ, and between ϵ and δ, respectively. These elements can be obtained by rock-physics modeling and/or lab measurements (Bachrach et al., 2011; Li et al., 2011). The covariance matrix $\mathbf{\Sigma}$ is symmetric. As a result, there are only six independent components in the covariance matrix. In an ideal case, matrix $\mathbf{\Sigma}$ should be estimated at each subsurface location to reflect the local lithological information.

Interpolation test with the preconditioning operators

To demonstrate the extra information that the preconditioning operators introduce to the inversion, we test these preconditioners on a classic missing data problem (Claerbout, 2009) on the BP2007 model. Consider the synthetic example in Figure 3, where the left column shows the current best spatial estimation of the VTI parameters (initial model) and the right column shows the true VTI parameters (true model). To build the initial velocity model, we assume the water bottom topography is known and the water velocity is 1.5 km/s. The basement interface is assumed to be flat

and the velocity of the basement rock is 3.5 km/s. Then, we linearly interpolate the velocity between the water bottom and the basement interface. We follow a similar approach to build the initial ϵ and δ model. The resulting initial VTI model shows smooth, linearly increasing trends for velocity, ϵ, and δ.

The true velocity, ϵ, and δ updates are plotted in Figures 4(a), (b) and (c), respectively. The true velocity updates show two anticline structures up shallow and a major anticline structure down deep. The updates in the anisotropic parameters contain more layering information than the velocity update. These are the solutions to the model building problem.

Suppose that, among the three unknown parameters, only velocity has been measured at random well locations in this section (Figure 4(d)). The goal of the inversion is to interpolate the well information onto a regular 2D grid and to provide reasonable estimates for the unmeasured anisotropic parameters.

The interpolation problem can be formulated as a data fitting objective

$$\mathbf{WB}d \quad \approx \quad \mathbf{W}m; \tag{5}$$

to match the interpolated model m with the binned data $\mathbf{B}d$ at the well location defined by $\mathbf{W}$. Additionally, the distribution of the model must follow a user-defined prior covariance $\mathbf{C}$:

$$0 \approx \quad \mathbf{C}^{-1/2}m. \tag{6}$$

To speed up the inversion, a preconditioning strategy is used by introducing a preconditioning variable

$$n = \mathbf{C}^{-1/2}m. \tag{7}$$

Now, the system of equations we solve is as follows:

$$\mathbf{WB}\ d \quad \approx \quad \mathbf{WC}^{1/2}n; \tag{8}$$

$$0 \quad \approx \quad n. \tag{9}$$

Different choices can be made for the preconditioner $\mathbf{C}^{1/2}$. As discussed in the previous work (Li et al., 2013), structure dip filters can be good estimates of the spatial covariance of the velocity and anisotropic model individually. Figure 5 shows a stacked image of the studied area (panel (a)) and the dip field estimated from it (panel (b)). Steering filters $\mathbf{S}_v$ (Clapp, 2000) are estimated at each point on this section according to the dip field. We use the same steering filters for ϵ and δ; that is $\mathbf{S}_\epsilon = \mathbf{S}_\delta = \mathbf{S}_v$.

In addition, we observe correlations among the VTI parameters at a single location. Point-by-point, we calculate the variance and the cross-variance using the true updates in Figure 4, and we plot the six independent components of the covariance matrix in Figure 6 and 7. The diagonal components in Figure 6 describe the spatial distribution of the variance between the initial and true model for each VTI parameter. The off-diagonal components in Figure 7 describe the cross-covariance between two given VTI parameters.

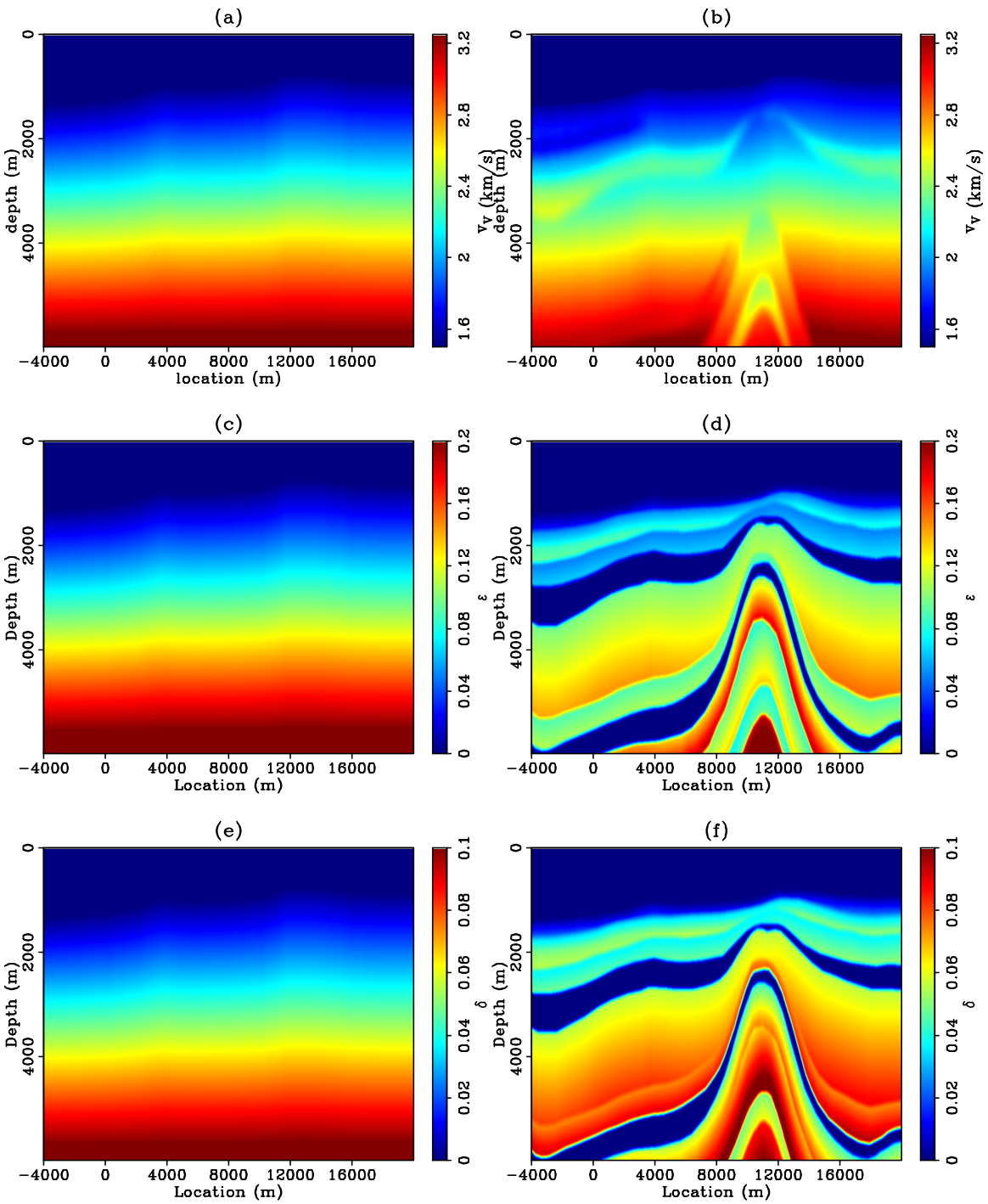

Figure 3: Comparison of the initial (left column) and the true (right column) models for BP2007 synthetic test. Top row: velocity models; middle row: ϵ models; bottom row: δ models. [**CR**]

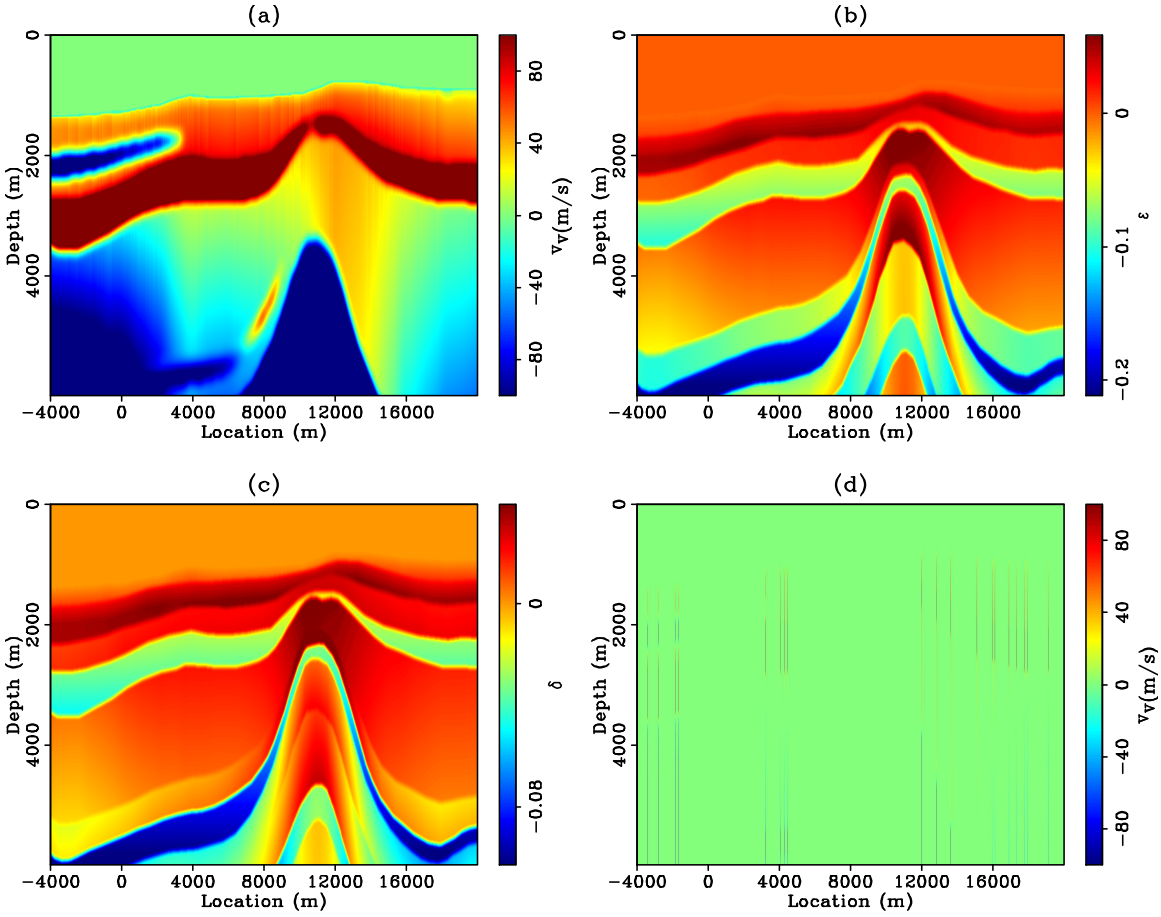

Figure 4: True updates in velocity in (a), ϵ in (b) and δ in (c). Pseudo velocity logs at random locations are plotted in (d). [**CR**]

To demonstrate the value of the cross-parameter covariance, we compare the inversion results of the following two preconditioners:

$$\mathbf{C}_1^{1/2} = \mathbf{S}, \tag{10}$$

and

$$\mathbf{C}_2^{1/2} = \mathbf{S}\boldsymbol{\Sigma}. \tag{11}$$

The interpolation results of both tests after six iterations are plotted in Figure 8.

The interpolation result for velocity using $\mathbf{C}_1^{1/2}$ is shown in Figure 8(a). We do not plot the interpolation results for ϵ and δ because both of them are zero due to the lack of measurements. The inversion results are better to the right side of the section where the velocity logs are denser. The shallow sand layer has been almost perfectly recovered. On the left side, the inversion interpolates between and extrapolates from the well logs along the structural dip. The result hints at the shallow sand layer and the deep anticline structure. However, the blank spaces between the logs are too large to be filled.

The inverted velocity using $\mathbf{C}_2^{1/2}$ (Figure 8(b)) is very similar to that using $\mathbf{C}_1^{1/2}$. However, extra information added by the off-diagonal terms in the covariance matrix shows correct updates in the inverted results for ϵ and δ. On the right side, the inversion almost perfectly recovers the ϵ and δ updates along with the proper layering in both the shallow and the deep sections. The reconstruction of the deep layering is mostly attributed to the accurate correlation between velocity and the anisotropic parameters. To the left, the correct ϵ and δ updates are obtained wherever the velocity has been correctly resolved. In the regions where velocity is missing, the off-diagonal terms do not falsely "create" updates for ϵ and δ.

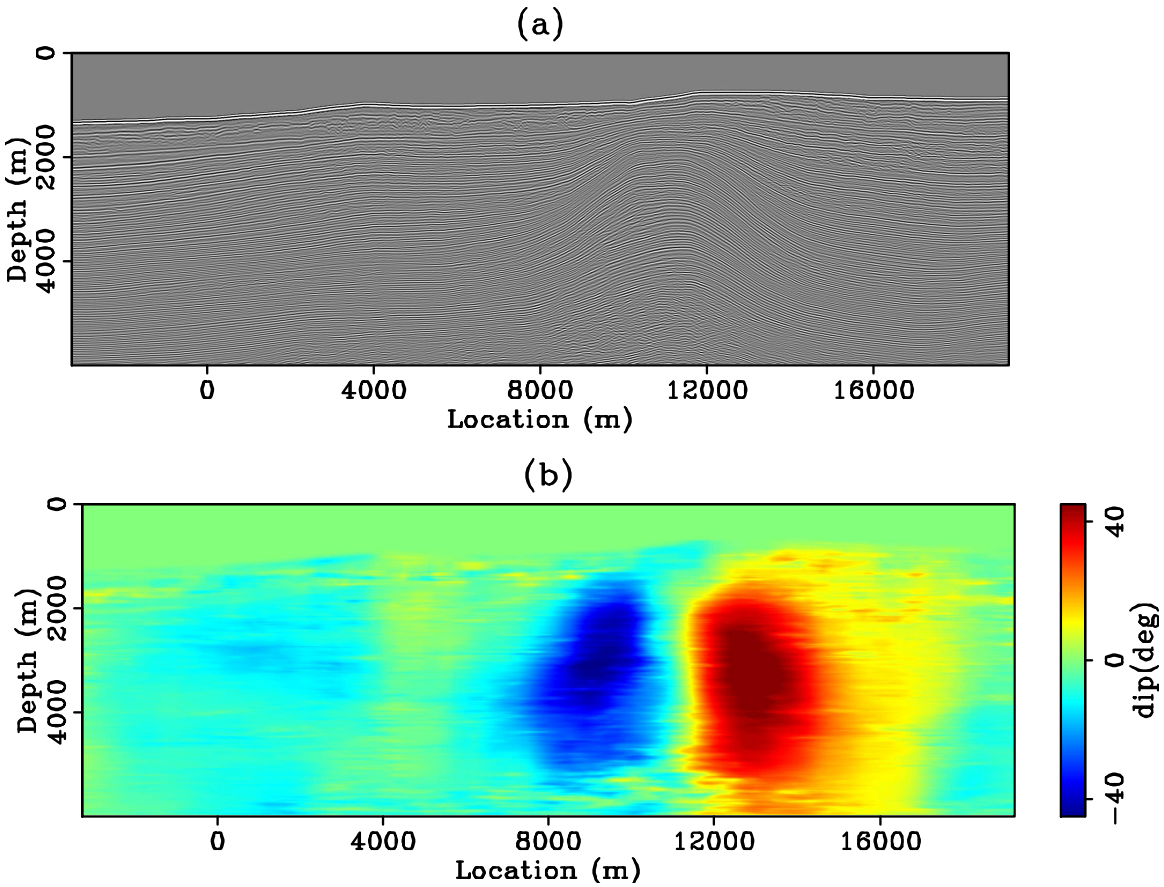

Figure 5: A stacked image of the studied area in panel (a), and the corresponding dip field in panel (b). [**CR**]

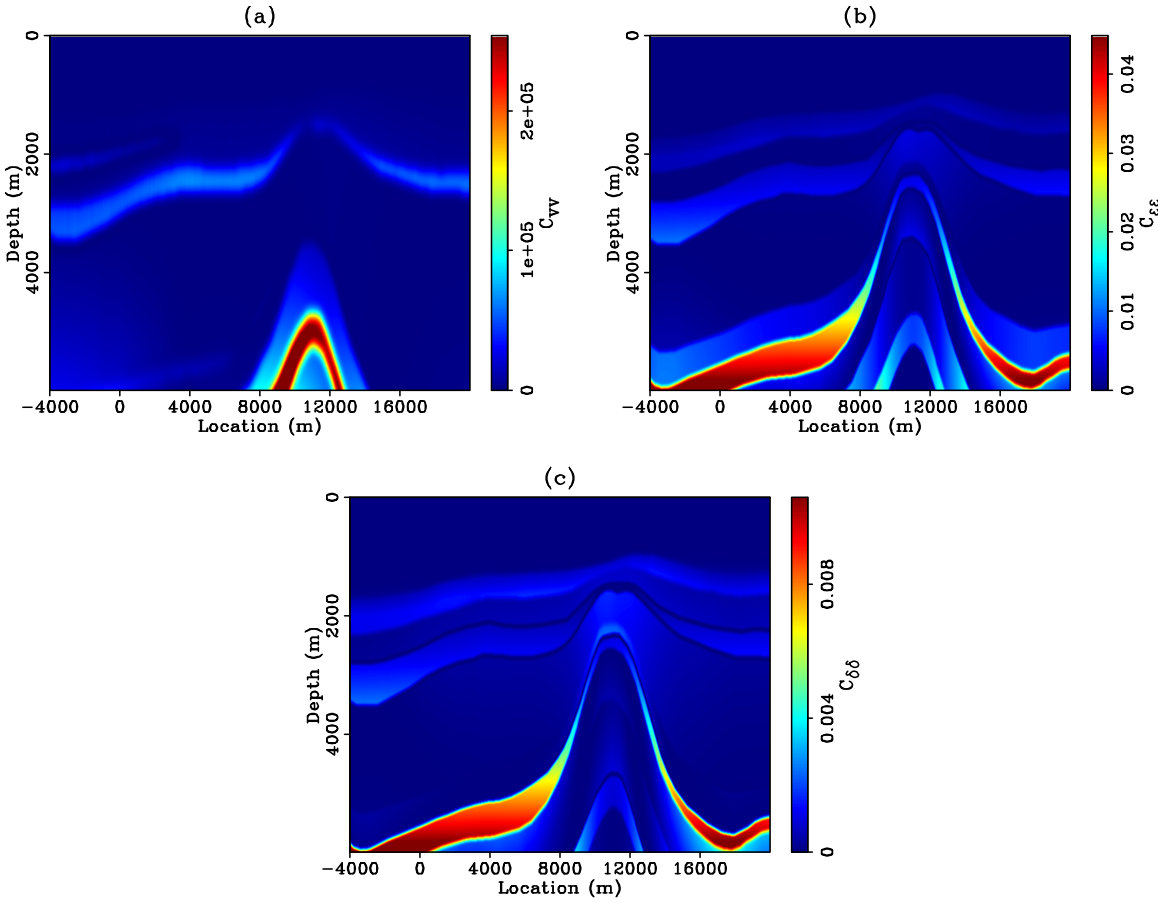

Figure 6: Diagonal elements of the covariance matrix of the VTI model. (a) C_{vv}: map of variance for velocity, (b) $C_{\epsilon\epsilon}$: map of variance for ϵ, and (c) $C_{\delta\delta}$: map of variance for δ. [**ER**]

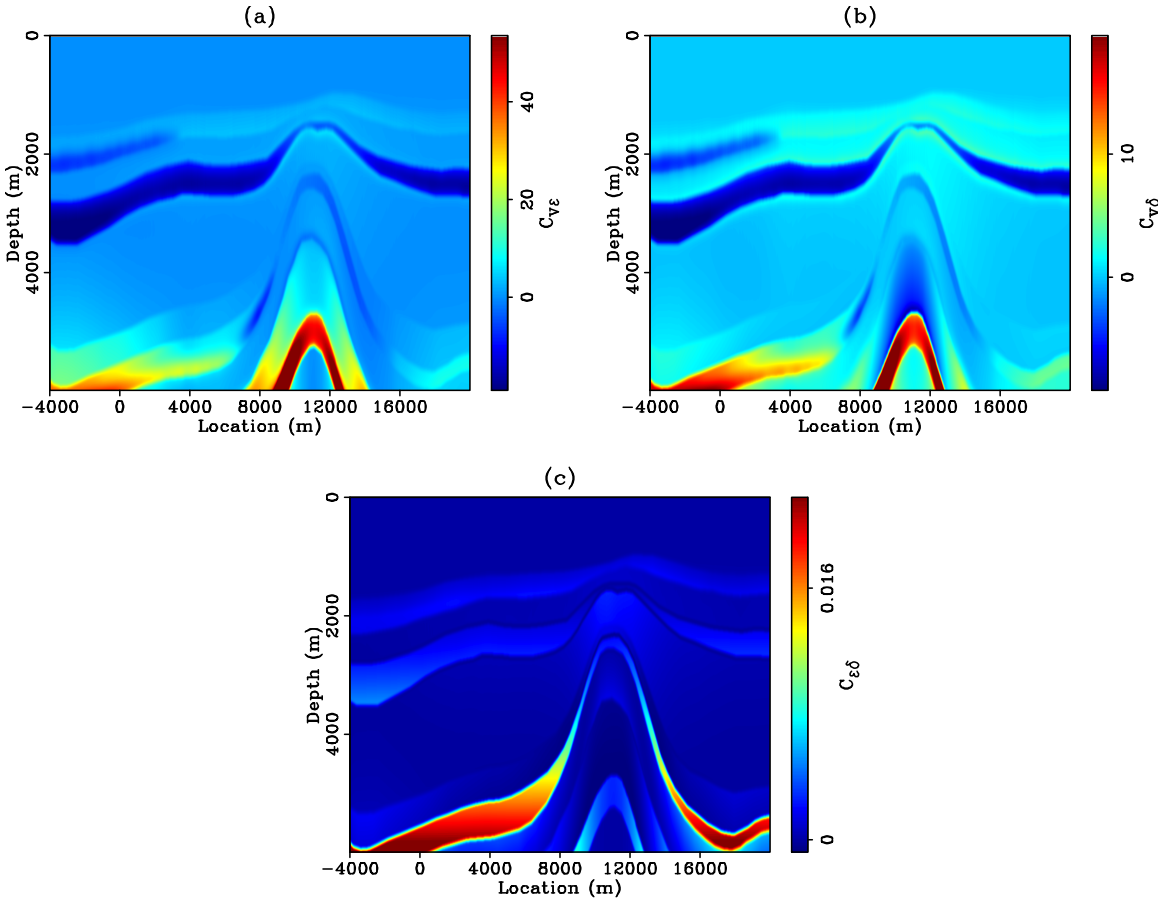

Figure 7: Off-diagonal elements of the covariance matrix of the VTI model. (a) $C_{v\epsilon}$: map of covariance between velocity and ϵ, (b) $C_{v\delta}$: map of covariance between velocity and δ, and (c) $C_{\epsilon\delta}$: map of covariance between ϵ and δ. [**ER**]

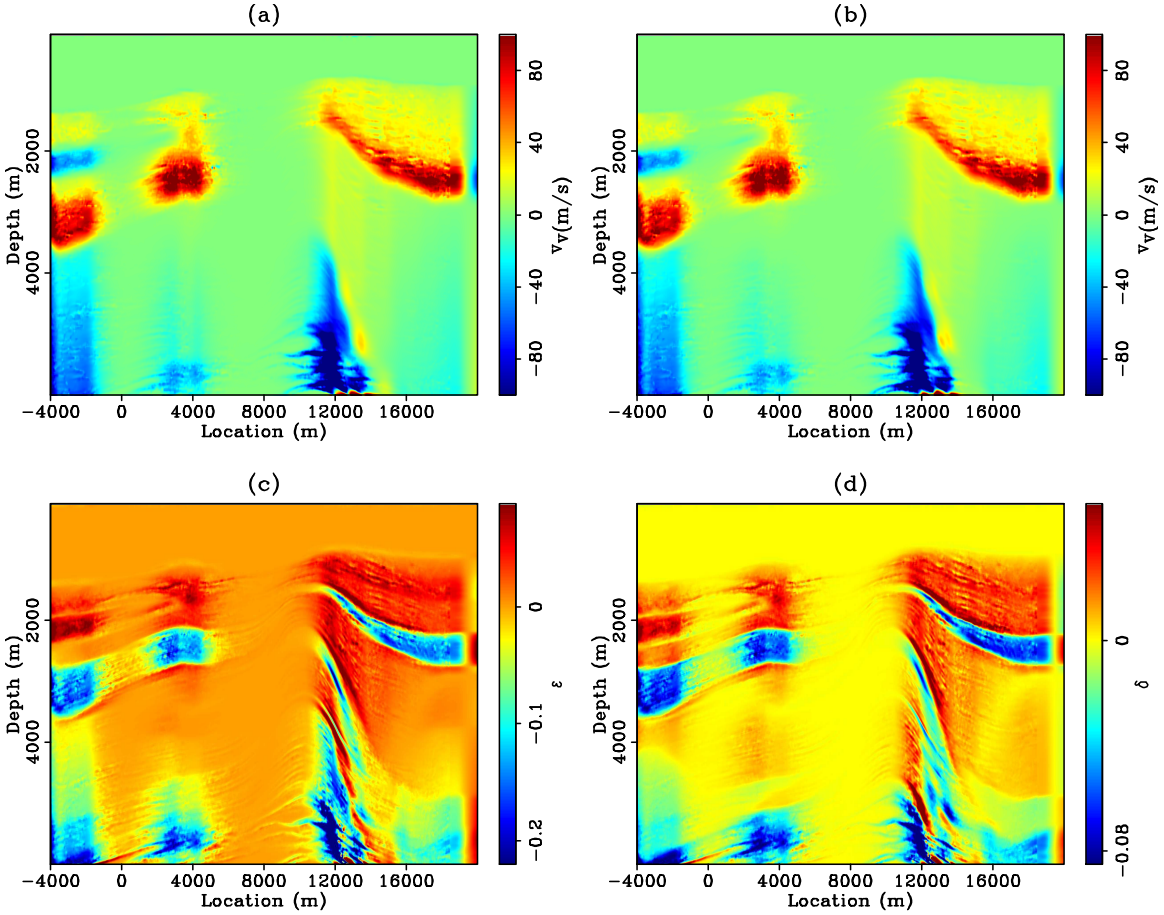

Figure 8: Comparison of the interpolation results. (a): Inverted velocity using $\mathbf{C}_1^{1/2}$. (b): Inverted velocity using $\mathbf{C}_2^{1/2}$. (c): Inverted ϵ using $\mathbf{C}_2^{1/2}$. (d): Inverted δ using $\mathbf{C}_2^{1/2}$. [**CR**]

WEMVA SYNTHETIC EXAMPLE WITH ROCK PHYSICS CONSTRAINTS

In this section, we will test the preconditioning scheme using different estimations of the covariance matrix on the modified BP2007 VTI dataset. The synthetic data are modeled with a streamer geometry where the maximum offset is 6km. Shot spacing is 100m, and receiver spacing is 12.5m. A total of 100 shots have been modeled. The synthetic data are then inverted using the preconditioned anisotropic WEMVA technique.

The initial and the true models are shown in Figure 3. Compared with the true models, the initial model captures the overall increasing trend of the VTI parameters. However, the following two important lithological layers in the shallow region (above 4km) are missing from the initial models: the first shale layer with low velocity and high anisotropy, and the third sand layer with high velocity and low anisotropy. The sand-shale inter-layering in the deep region (below 4km) is obvious along the anticline in the true model. In contrast with the shallow region, the high velocity layers in the deep section correlate with the shale layers and with high anisotropy. Because the deeper structures below 4km will be difficult to resolve due to the limited acquisition, we will focus our discussion on the shallow sediments.

The initial migration image is shown in Figure 9. There are large vertical shifts between the initial image and the true image. The reduced amplitudes in the deeper region on the initial stacked image indicate a weaker focusing effect compared with the true stacked image. Figure 10 shows the comparison of the initial ADCIGs in (a) with the true ADCIGs in (b). The upward moveouts in the initial ADCIGs indicate that the average migration velocity is lower than the true average velocity along the wavepaths. Although the true anisotropy updates in the sand layer are negative by my construction, the overwhelming effects of the slower velocity on the kinematics of the acoustic waves are predominant. In the iterative WEMVA process, these upward moveouts in the ADCIGs will translate into positive updates in both velocity and anisotropic parameters ϵ and δ.

The six independent components of the covariance matrix at each point are plotted in Figures 6 and 7. Notice the negative correlation between velocity and anisotropy in the shallow region, and the positive correlation in the deep region (Figure 7(a) and 7(b)). Figure 7(c) shows that the positive correlation trend between ϵ and δ is largely independent of the depth in this section.

To test the effect of different preconditioning schemes, we perform anisotropic

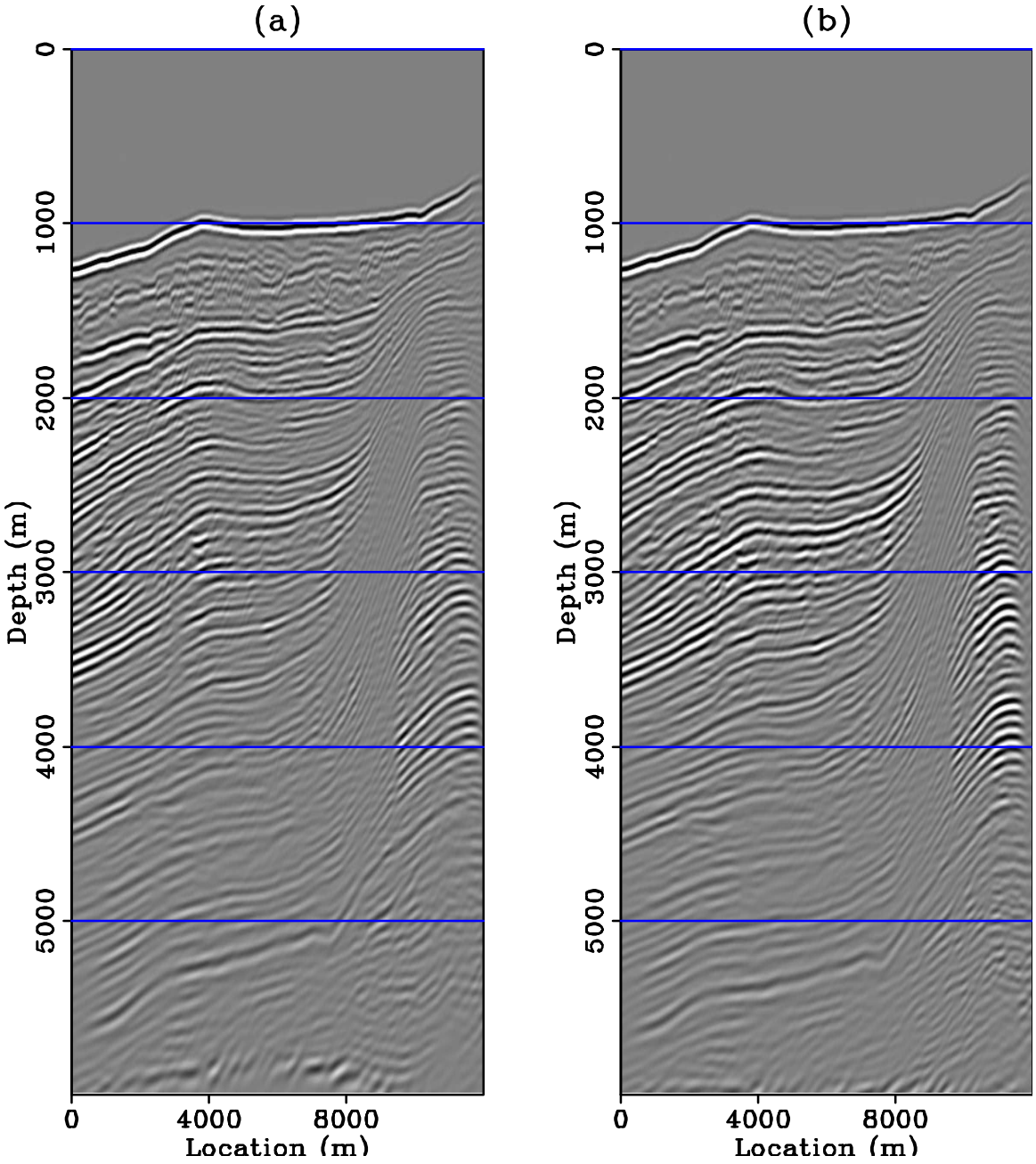

Figure 9: Comparison of the initial stacked image (a) and the true stacked image (b). Notice the depth shift between the images and the unfocused reflectors around $z = 3$ km and $x = 4$ km. [**CR**]

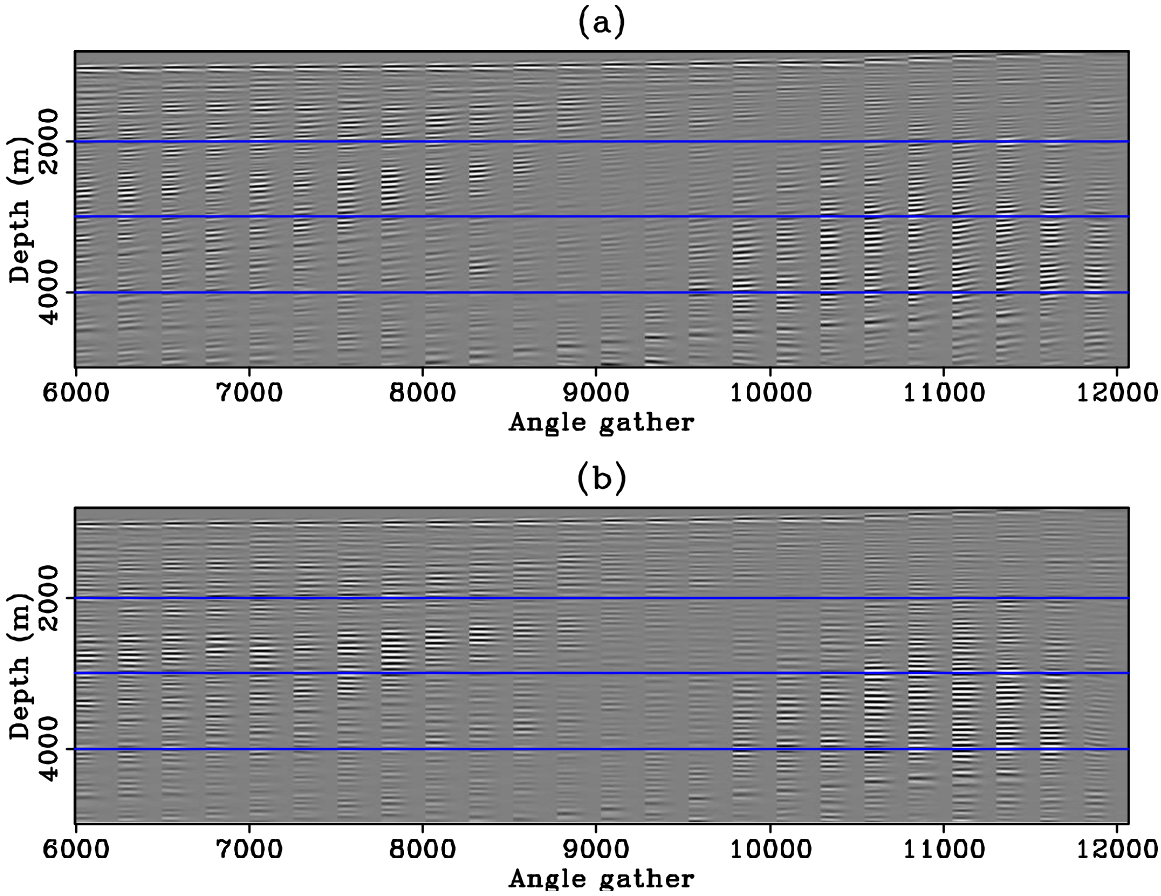

Figure 10: Comparison of the initial ADCIGs (a) and the true ADCIGs (b). Notice the upward moveout throughout the section in the initial gathers. [**CR**]

WEMVA with three different preconditioning matrices:

$$\mathbf{\Sigma}_1 = \left| \begin{array}{ccc} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{array} \right|^{1/2}, \tag{12}$$

$$\mathbf{\Sigma}_2 = \left| \begin{array}{ccc} C_{vv}\,\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{\epsilon\epsilon}\,\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_{\delta\delta}\,\mathbf{I} \end{array} \right|^{1/2}, \tag{13}$$

$$\mathbf{\Sigma}_3 = \left| \begin{array}{ccc} C_{vv}\,\mathbf{I} & C_{v\epsilon}\,\mathbf{I} & C_{v\delta}\,\mathbf{I} \\ C_{\epsilon v}\,\mathbf{I} & C_{\epsilon\epsilon}\,\mathbf{I} & C_{\epsilon\delta}\,\mathbf{I} \\ C_{\delta v}\,\mathbf{I} & C_{\delta\epsilon}\,\mathbf{I} & C_{\delta\delta}\,\mathbf{I} \end{array} \right|^{1/2}. \tag{14}$$

The preconditioning matrix Σ_1 indicates that no preconditioning is applied; Σ_2 indicates that a diagonal preconditioning is applied; and Σ_3 indicates that the full preconditioning approach is applied. The spatial covariance matrix for the three tests are the same.

The initial preconditioning model $\mathbf{n}_0$ is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \left\langle \mathbf{m}_0 - \mathbf{\Sigma B n}_0, \mathbf{m}_0 - \mathbf{\Sigma B n}_0 \right\rangle. \tag{15}$$

The gradient of the WEMVA objective function (1) with respect to this preconditioning variable $\mathbf{n}$ is

$$\begin{aligned} \nabla_{\mathbf{n}} J &= (\frac{\partial \mathbf{m}}{\partial \mathbf{n}})^* \nabla_{\mathbf{m}} J \\ &= \mathbf{B}^* \mathbf{\Sigma}^* \nabla_{\mathbf{m}} J, \end{aligned} \tag{16}$$

where $\nabla_{\mathbf{m}} J = [\nabla_v J \;\; \nabla_\epsilon J \;\; \nabla_\delta J]^T$.

To understand the preconditioning scheme, we analyze the preconditioning effect assuming a nonlinear steepest decent inversion framework. The initial preconditioned model $\mathbf{n}_0$ is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \left\langle \mathbf{m}_0 - \mathbf{\Sigma S n}_0, \mathbf{m}_0 - \mathbf{\Sigma S n}_0 \right\rangle. \tag{17}$$

For the ith iteration, the preconditioned variable is obtained by

$$\mathbf{n}_{i+1} = \mathbf{n}_i + \lambda_i \nabla_{\mathbf{n}} J, \tag{18}$$

with λ_i the step length in the ith iteration. Hence the original model variable is

$$\begin{aligned} \mathbf{m}_{i+1} &= \mathbf{S\Sigma n}_{i+1} \\ &= \mathbf{S\Sigma n}_i + \lambda_i \mathbf{S\Sigma} \nabla_{\mathbf{n}} J \\ &= \mathbf{m}_i + \lambda_i \mathbf{S\Sigma\Sigma}^* \mathbf{S}^* \nabla_{\mathbf{m}} J. \end{aligned} \tag{19}$$

Equation 19 suggests that preconditioning a non-linear inversion is equivalent to filtering the gradients so that the resulting updates have the desired spectrum. Therefore, instead of explicitly reformulating the preconditioned inversion, we can make use of the original non-linear conjugate gradient algorithm implementation with minimal changes. The preconditioning step is highlighted in red in Algorithm 1.

Algorithm 1 Optimization algorithm

initialize the model: $\mathbf{m}_0$
compute the migrated image: I_0
compute the gradient: $\mathbf{g}_0$
precondition the gradient: $\mathbf{g}_s^0 = \mathbf{S}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*\mathbf{S}^*\mathbf{g}_0$
initialize the search direction: $\mathbf{p}_0 = -\mathbf{g}_s^0$
for $k = 1 \cdots N_k$ **do**
  perform a line search: optimize λ, $\underset{\lambda}{\text{argmin}}\ J(\mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1})$
  update the velocity model: $\mathbf{m}_k = \mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1}$
  compute the migrated image: I_k
  compute the gradient: $\mathbf{g}_k$
  precondition the gradient: $\mathbf{g}_s^k = \mathbf{S}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^*\mathbf{S}^*\mathbf{g}_k$
  find the search direction: $\mathbf{p}_k = -\mathbf{g}_k + \frac{(\mathbf{g}_s^k)^T(\mathbf{g}_s^k - \mathbf{g}_s^{k-1})}{(\mathbf{g}_s^{k-1})^T\mathbf{g}_s^{k-1}}$
end for

Inversion results

Figure 11 shows the velocity updates after the first iteration. In panel (a), we display the raw update from the moveouts on the ADCIGs. The overall update direction for velocity is positive. The vertical resolution of the updates is very low after the first iteration: a bulk positive velocity shift is indicated by the raw gradient. In panel (b), the spatial distribution of the velocity update has been modified by the diagonal elements of the covariance matrix. Because of the prior knowledge of a sand layer with high velocity perturbation, the preconditioning concentrates the strong velocity update within that layer. Compared with the true velocity update in panel (d), the velocity update with the diagonal preconditioning already achieves high vertical resolution from the first iteration. In panel (c), we utilize all the elements of the covariance matrix to precondition the gradient. The resulting update is very similar to the update with diagonal preconditioning. This similarity suggests that the off-diagonal elements of the covariance matrix have limited influence on velocity.

Figure 12 shows the updates in ϵ after the first iteration. In panel (a), the raw update in ϵ appears lower resolution than the raw updates for velocity. All the updates are in the positive direction, thus compensating for the upward moveouts in the initial ADCIGs. However, the updates in ϵ are very small due to its low influence to the kinematics of the waves. Although the diagonal preconditioning has changed the spatial distribution of the update with better definition of the layers,

the positive values in the ϵ update in the sand layer is still in the opposite direction of the true update (Figure 12b). In panel (c), when the off-diagonal elements of the covariance matrix are included to account for the cross-variance between velocity and ϵ, the preconditioning provides ϵ updates in accordance with the lithology as follows: positive ϵ updates in the shale layer and negative ϵ updates in the sand layer. Considering the significant difference between panel (b) and panel (c), we conclude that the lithological information mainly comes from the off-diagonal elements of the covariance matrix. Similar analysis and conclusions can also be made for anisotropic parameter δ (Figure 13).

Figure 14 shows the updates in velocity after 20 iterations. Panel (a) shows the velocity update when no preconditioning is applied. The spatial resolution for velocity gradually improves with iterations. However, the resolution remains significantly lower than the true update at iteration 20. More iterations could further improve the resolution, with a higher weighting on the stacking power term. However, to make a fair comparison, we stop all the tests at 20 iterations. When implementing diagonal preconditioning (panel (b)) and full preconditioning (panel (c)), the overall structure and resolution of the velocity updates remain stable with increasing iterations. This shows an early convergence. Nevertheless, the later iterations improve the definition of the thin sand layer with high velocity at $z = 1.5$km and $x = 10$km, and the anticline structure with low velocity around $z = 4$km and $x = 11.5$km. The final inversion results for velocity are nearly identical between the diagonal and the full preconditioning schemes. As will be shown by the final stacked image, the effective imaging velocity from the three different preconditioning schemes are very similar.

Figure 15 shows the updates in ϵ after 20 iterations. When no preconditioning is applied (panel (a)), the resolution of the ϵ is lower than the resolution of velocity as predicted by the topography of the objective function in the previous section. Compared with Figure 14(a), although the updates in ϵ in Figure 15(a) are mostly in phase with the updates in velocity, the positive updates on the flanks and the top of the anticline are in the correct direction despite the negative velocity updates at the same locations. When preconditioned with diagonal covariance (panel (b)), the inversion has a better resolution for the shale layer near the water bottom. However, the updates in the sand layer are in the opposite direction of the true ϵ updates. The inverted updates for ϵ are very close to the true updates when the off-diagonal components of the covariance matrix are also included in the preconditioning (panel (c)). The inversion almost perfectly recovers both the shale and the sand layer due to the prior knowledge of the cross-correlation between the velocity and ϵ perturbation within each layer. Similar analysis and conclusions can also be made for the inverted updates in δ (Figure 16). Notice that on the flanks and the top of the anticline, the inversion successfully resolves the ϵ perturbation no matter which preconditioning scheme is applied. This success is due to the rich angular coverage in these regions, especially wide angle imaging rays with higher sensitivity to the anisotropic parameters.

Figure 17 shows the inverted vertical velocity models after twenty iterations.

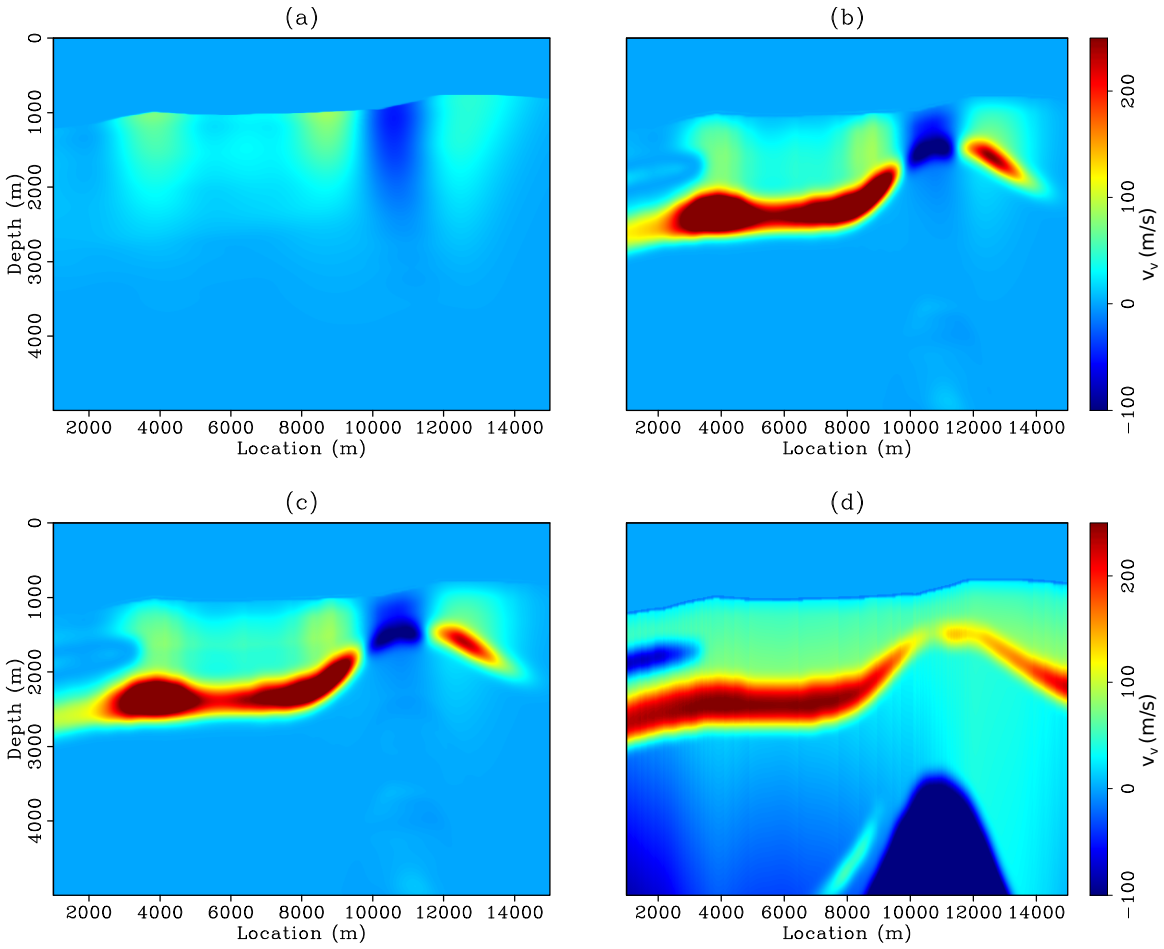

Figure 11: Velocity updates after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [**CR**]

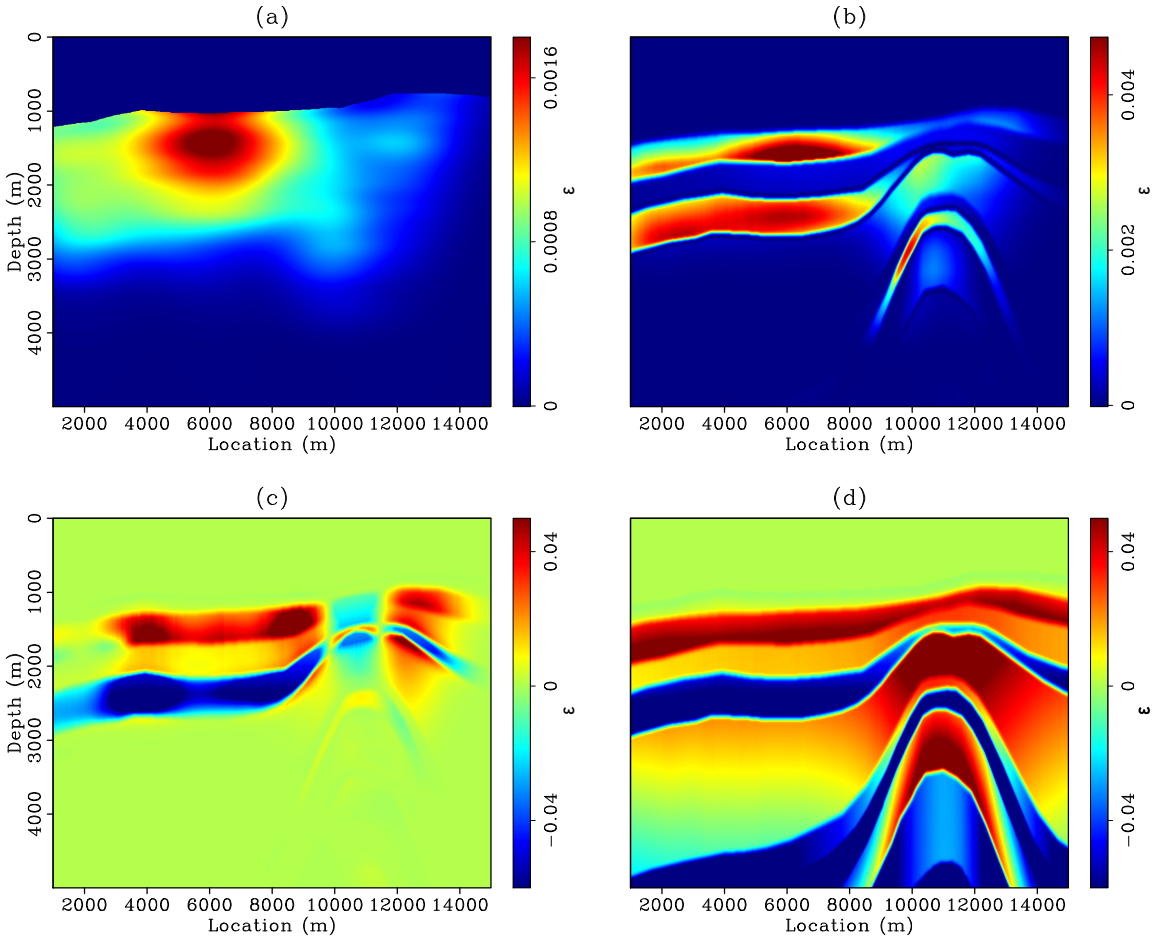

Figure 12: Updates in ϵ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. Notice the enhanced amplitude by the full preconditioning in (c). [**CR**]

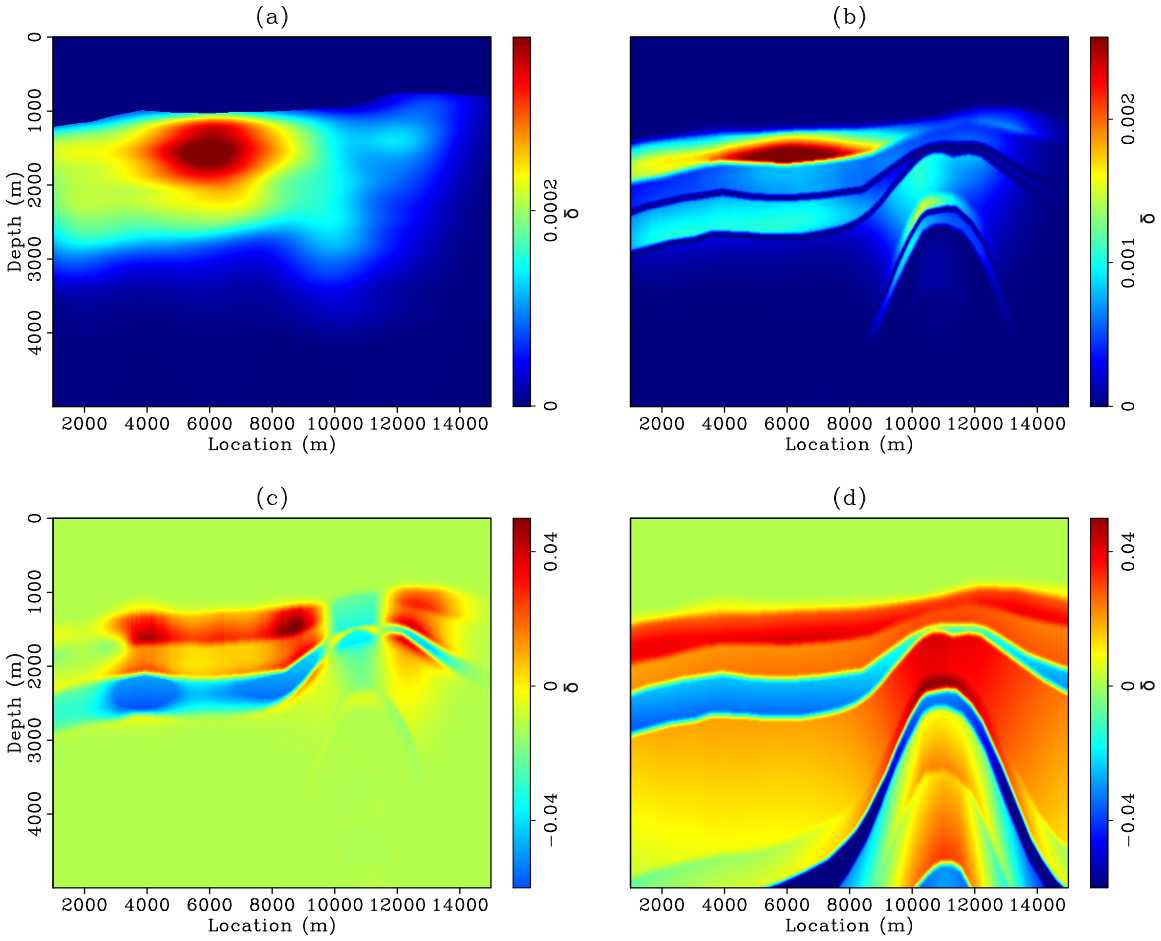

Figure 13: Updates in δ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. Notice the enhanced amplitude by the full preconditioning in (c). [**CR**]

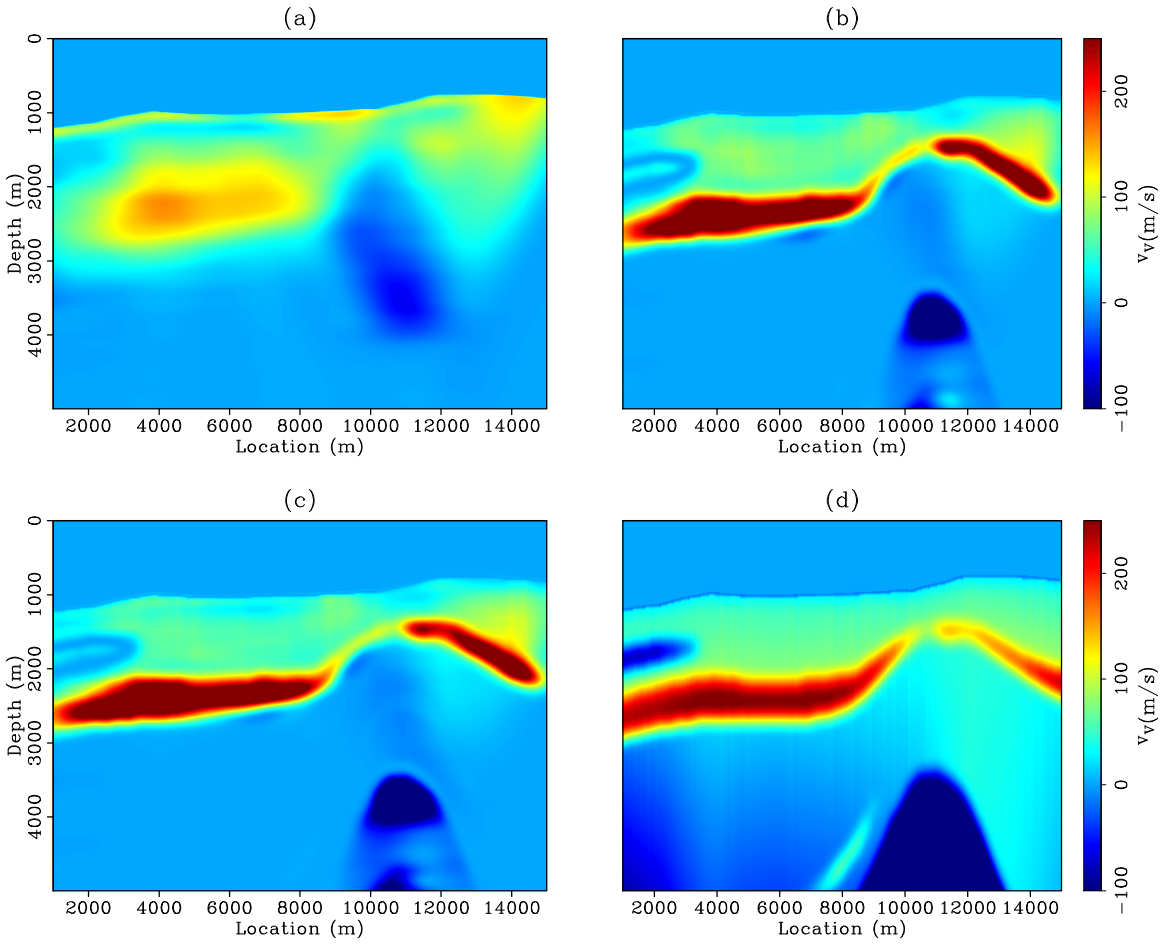

Figure 14: Velocity updates after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [**CR**]

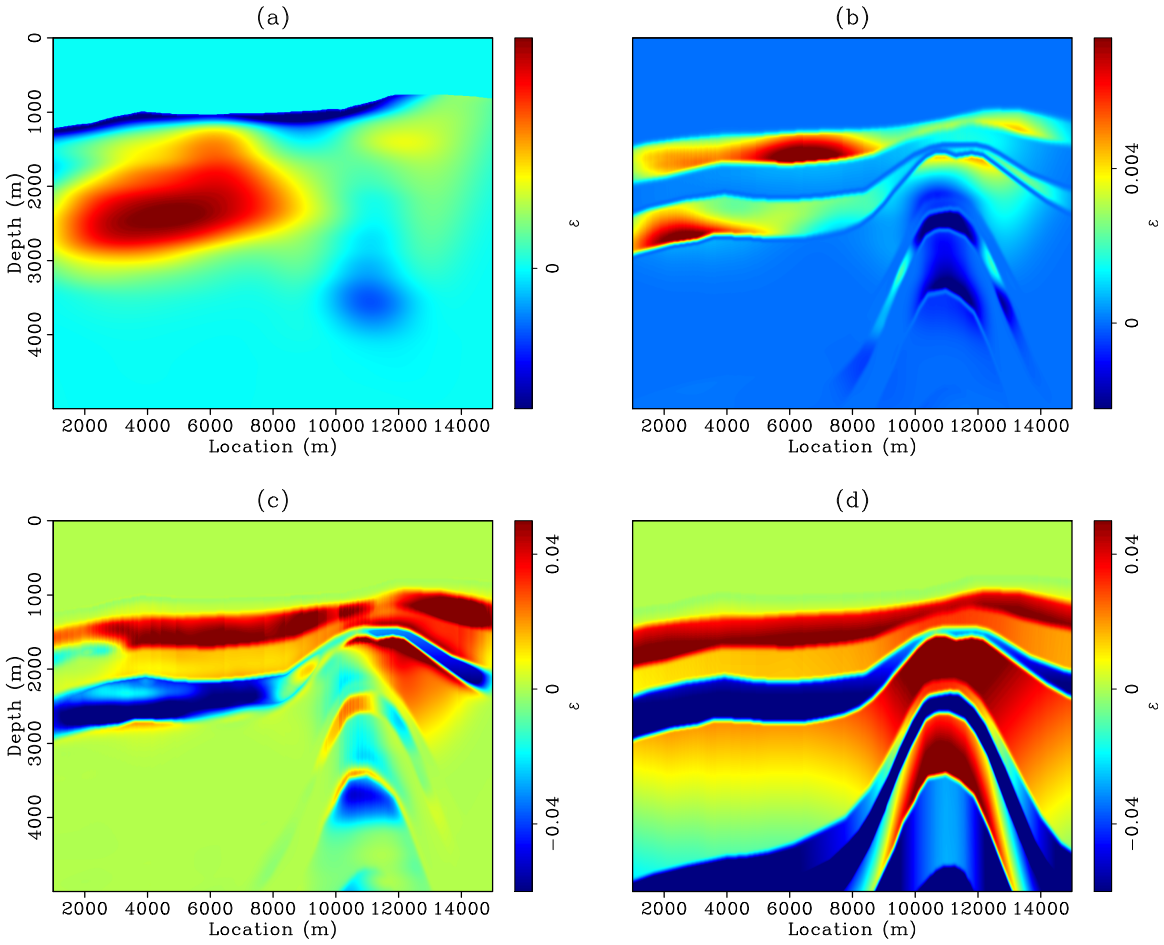

Figure 15: Updates in ϵ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. [**CR**]

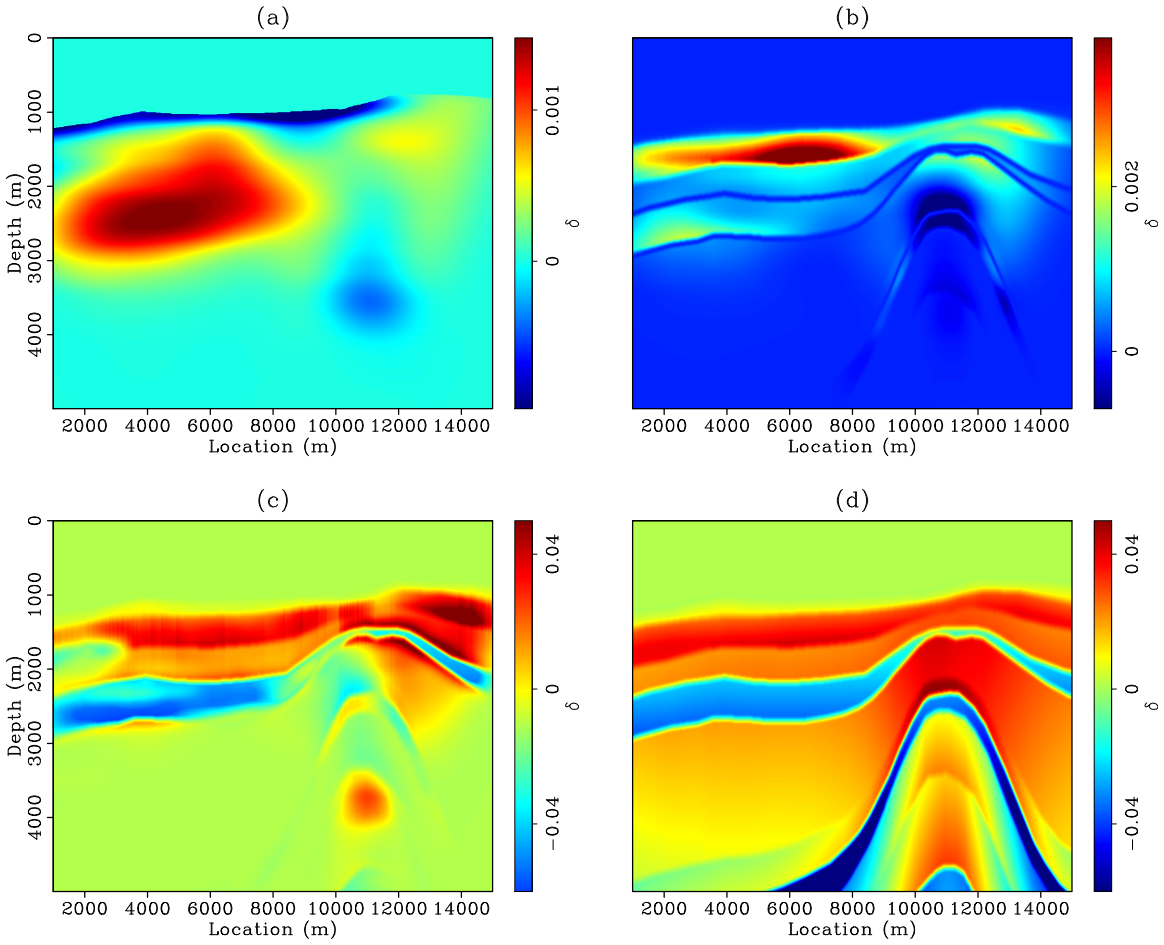

Figure 16: Updates in δ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. [**CR**]

When no preconditioning is applied (panel (a)), the inversion only recovers low wavenumber components of the velocity model. When diagonal or full preconditioning is applied (panels (b) and (c), respectively), the layering with variable depths in the shallow region is very well recovered. In the deeper section, only the strong low velocity anomaly at the top of the anticline has been retrieved due to the limited angle illumination. The lack of inversion success in the deeper region suggests that powerful as our preconditioning scheme is, it cannot "create" information where the seismic data has little information.

Figure 18 shows the inverted ϵ models after twenty iterations. Inversions without the cross-parameter correlation (panel (a) and (b)) have barely moved the solution from the initial models. The inversion can resolve a high resolution and rock physics plausible ϵ model only when the full preconditioning scheme is applied.

Figure 19 shows the inverted δ models after twenty iterations. Similar to the inversion results for ϵ, the information contributed by the off-diagonal terms in the full covariance matrix provides much better spatial and rock physics constraints. Notice that in the shallow shale layers, the stronger anisotropic thin shale layers are very well resolved on panel (c). This high resolution result is attributed to both the high resolution preconditioning and the high resolution stacking power objective function.

It is important to note that some ambiguities remain in the VTI parameters even when the full preconditioning scheme is used. One example is highlighted by the ellipses in the figures. In Figure 17(c), a low velocity wedge can be seen when compared with the true velocity model in Figure 17(d). In the corresponding region, a thinner isotropic sand layer has been observed from the inverted anisotropic parameters (Figures 18(c) and 19(c)) when compared with the true models (Figures 18(d) and 19(d)). The low velocity anomaly and the high anisotropic anomalies, although of different shapes, compensate each other and focus the seismic data associated with this area.

Figure 20 shows the comparison between the final stacked images produced by different preconditioning schemes with the true stacked image. All three inverted images are greatly improved from the initial stacked image in Figure 9(a) with better focusing and better defined depths of the reflectors. Despite the significant differences in the inverted velocity and anisotropic models from three preconditioning schemes, the effective imaging velocities along the wavepaths are very similar.

Figure 21 shows the comparison of the final ADCIGs with different preconditioning schemes. Compared with the initial ADCIGs in Figure 10(a), the inversion flattens the angle domain events no matter which preconditioning scheme is applied. The three inversion tests give very similar results in the center of the section between $x = 7.5$ km and $x = 10.5$ km where the reflectors are illuminated by the data with rich angles. When the subsurface is less well illuminated by the data, for example in the regions highlighted by the circles, the full preconditioning scheme produces flatter events, stronger amplitudes, and higher angle coverage.

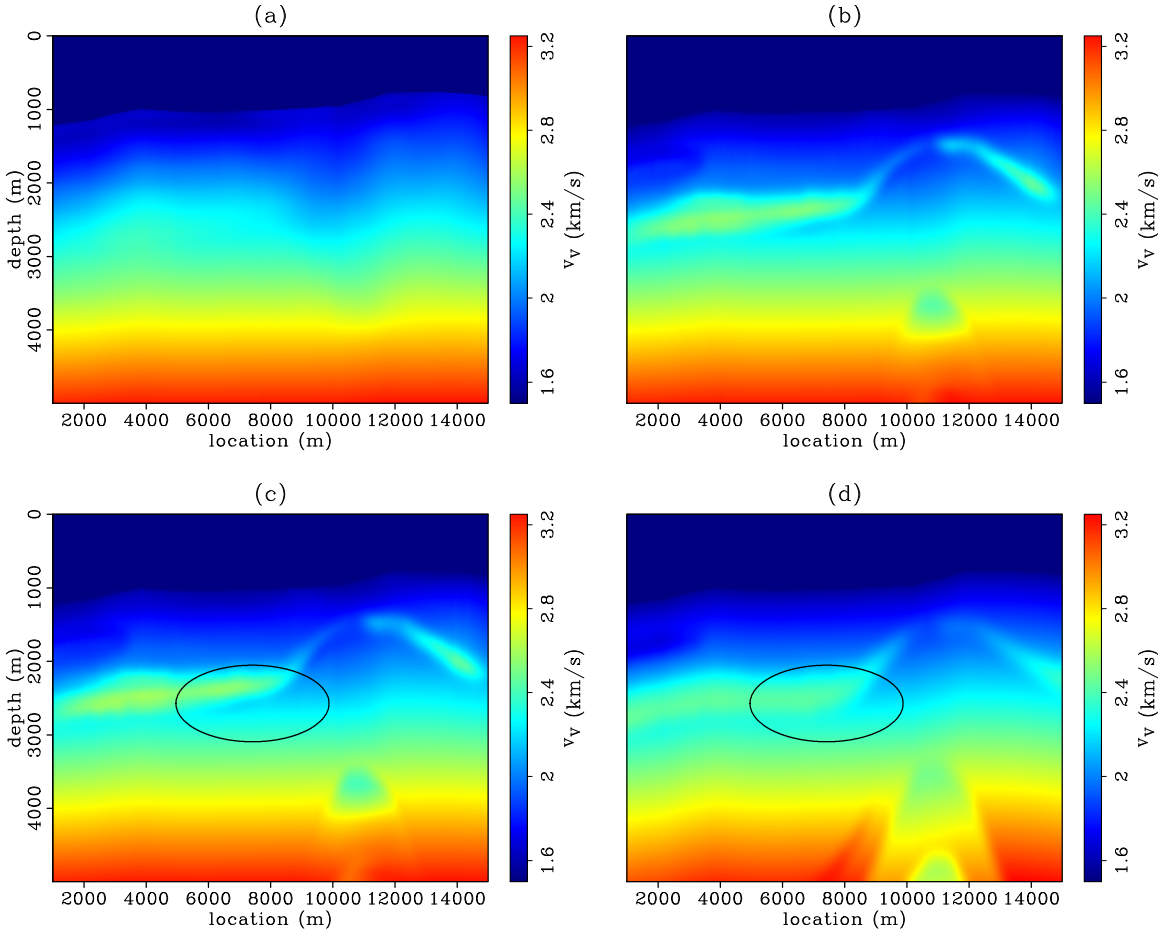

Figure 17: Inverted vertical velocity model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true vertical velocity model for comparison. [**CR**]

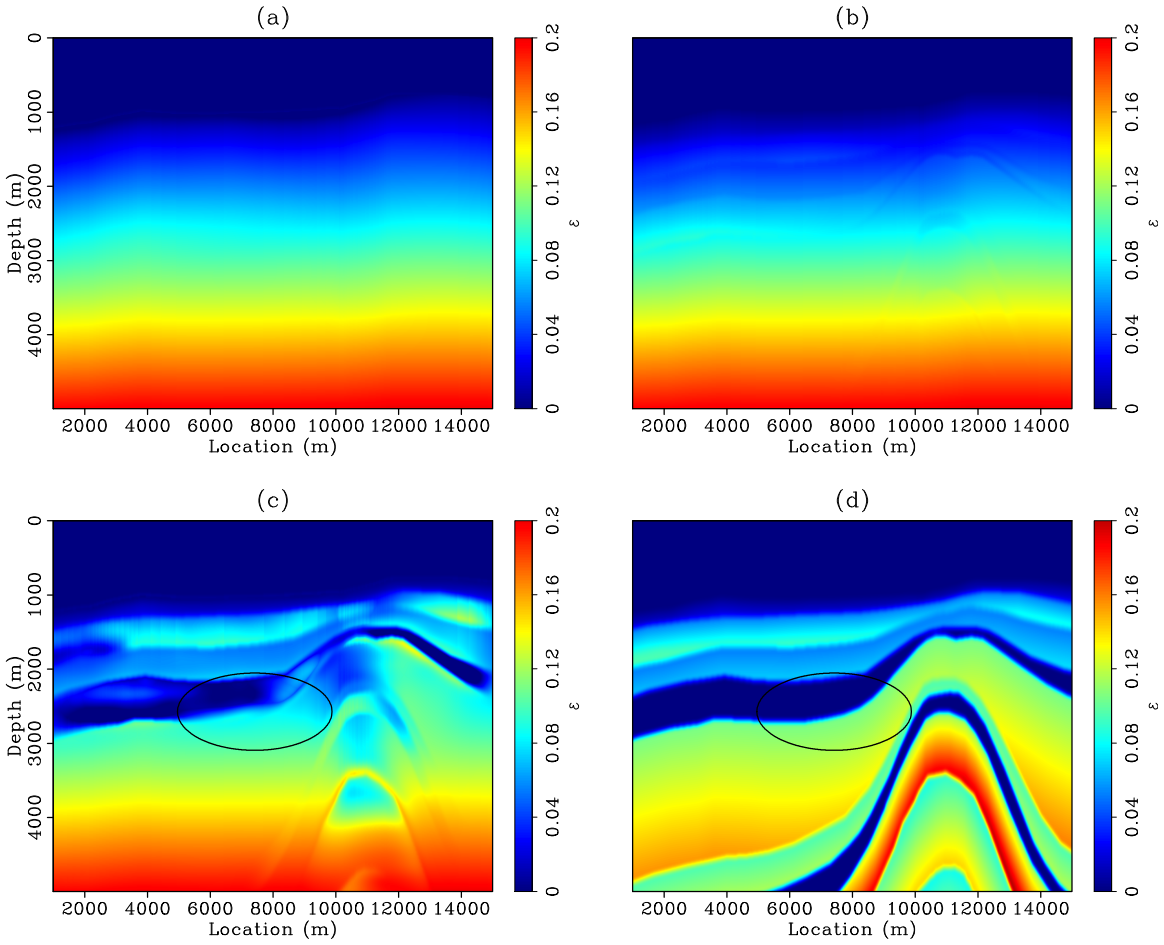

Figure 18: Inverted ϵ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ model for comparison. [**CR**]

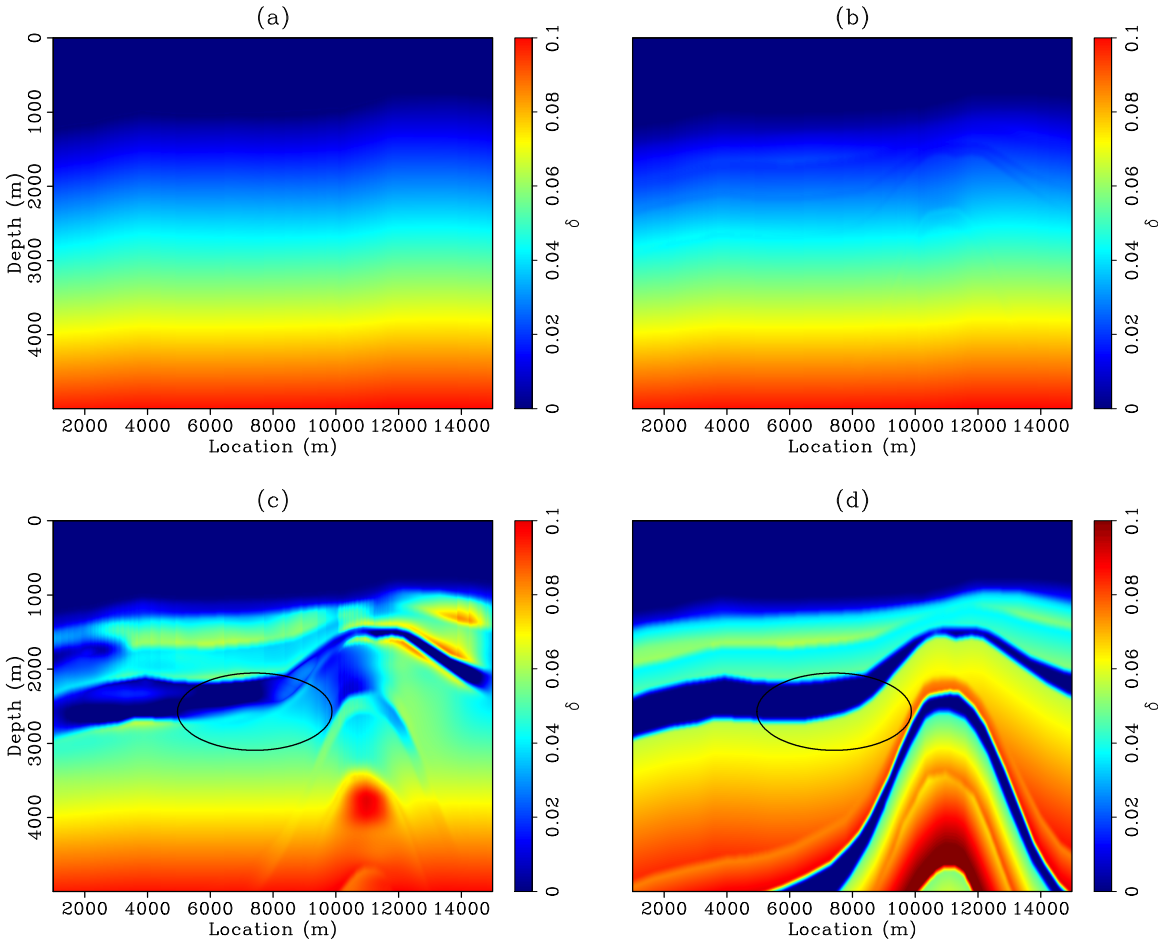

Figure 19: Inverted δ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ model for comparison. [**CR**]

The WEMVA data fitting objective function can be evaluated at each imaging point. We compare the initial objective function map with the inverted objective function map with different preconditioning schemes in Figure 22. The warmer color indicates higher data fitting error, whereas the cooler color indicates lower data fitting error. In the initial objective function map, the error is small in the shallow region (above 2.5 km) but it gets stronger with depth. As shown by the inverted objective function map, the data fitting error has been significantly reduced in all three cases, especially in the regions between $x = 1.5$ km and $x = 4$ km and between $x = 10$ km and $x = 12$ km.

To compare the objective function maps quantitatively (the visual comparison is difficult due to their similarity), we subtract the objective function map in Figure 22(d) from Figure 22(b) and plot the difference in Figure 23(a). Similarly, the difference between Figures 22(d) and 22(c) is plotted in Figure 23(b). The region in red denotes the area where the inversion with full preconditioning fits the data better than the other two tests, and the region in blue denotes the opposite. The inversion with full preconditioning does not guarantee a better fit for the data globally, but it does show better fits locally (Figure 22). Furthermore, the differences among different preconditioning schemes increase with depth as the data constraint becomes weaker with depth.

Figure 24(a) plots the value of the data fitting objective function (sum of the objective function map in Figure 22) as a function of iteration for three different preconditioning tests. The objective function value is normalized with respect to the objective function value of the image migrated with the true models. All three inversion tests reduce the objective function value to about 30% of the initial value. During the 20 iterations, the objective function values with the full covariance matrix are almost always the lowest among three tests; however, the differences among the three objective values at each iteration are small. The curvature of the objective function curves suggests that the inversion tests with diagonal or full preconditioning have converged at iteration 20, whereas the inversion test without any preconditioning may need more iterations to fully converge.

Figure 24(b) plots the normalized length of the gradient as a function of iteration for three inversion tests. Although not guaranteed to be monotonically decreasing, the length of the gradient should have a decreasing trend and should approach zero as the inversion converges, as shown by the curves with diagonal or full preconditioning. On the other hand, the gradient when no preconditioning is applied still has significant magnitudes at the last few iterations. This is consistent with the objective function curve (magenta line in Figure 24(a)) that more iterations may be needed to improve the convergence.

DISCUSSION

The nonlinear and underdetermined nature of the anisotropic model building problem leads to multiple VTI models that fit the same surface seismic data equally well. These models are referred to as equiprobable models (Yang et al., 2012; Osypov et al., 2008). The key element in this paper is the inclusion of the covariance matrix among the VTI parameters so that these equiprobable models can be differentiated from the prospective of the geological and lithological environment.

At first glance, one may have concerns about the reliability of the "extra" information from the preconditioning. However, our tests using different preconditioning on the synthetic example show that if the VTI parameters are well constrained by the data in a certain region, different preconditioning schemes will not change the solutions to those parameters when the convergence has been achieved in all the inversion tests. For those strongly constrained parameters, proper preconditioning simply accelerates the convergence and enhances the resolution. The "extra" information mostly affects the weakly constrained parameters, especially in the poorly illuminated region. In these cases, the solutions to the model building problem can be dramatically different. It is possible that the rock physics model is inaccurate, which eventually leads to erroneous VTI models in the less constrained region. However, the inverted VTI models are always consistent with the input rock physics information. Therefore, quick evaluation and modification of the models are possible when more accurate rock physics information is available.

As shown by the inversion results, even when the exact full preconditioning was applied in the tests, ambiguities among the VTI parameters cannot be completely resolved. Two potential reasons may explain the residual ambiguity. First, the preconditioning scheme proposed in this paper utilizes the multivariate Gaussian distribution assumption among the VTI parameters. When the VTI parameters do not follow the Gaussian distribution, the proposed formulation may not provide a sufficient description of the correlation among them. Moreover, the covariance matrix among the VTI parameters is estimated based on the current VTI model and/or a previous lithological interpretation. The conditional probability distribution might have been changed as the inversion updates the VTI parameters. Therefore, reevaluation of the covariance matrix might be necessary after a few nonlinear iterations to better mitigate the ambiguities.

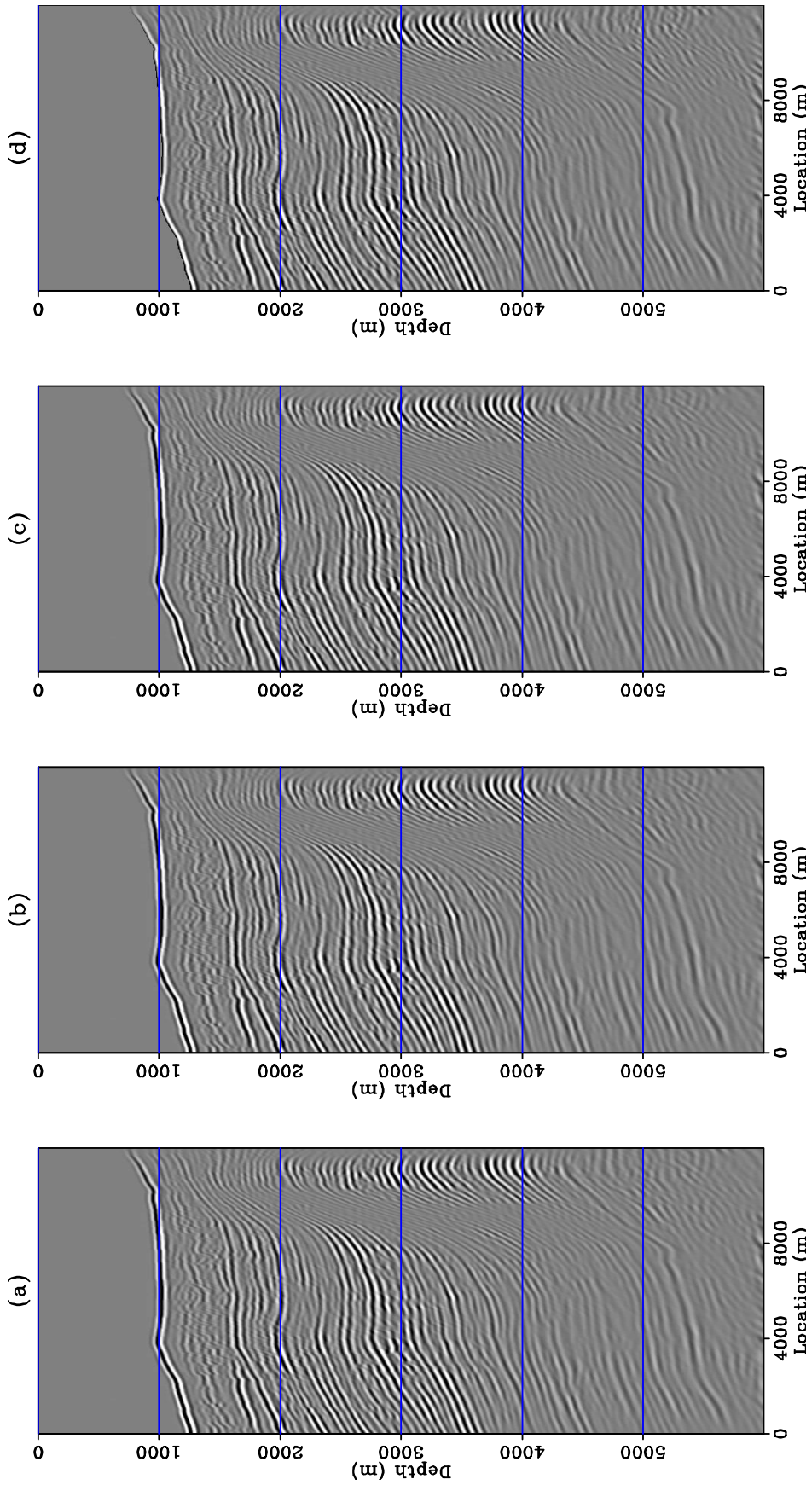

Figure 20: Final stacked images with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Panel (d) shows the true stacked image. [**CR**]

CONCLUSION

This paper presents a preconditioned anisotropic WEMVA scheme to better constrain the anisotropic model building process. The proposed preconditioning method includes lithological information in order to guide the inversion towards a plausible geological and rock physics solution. Numerical examples on a 2-D synthetic dataset show that when proper preconditioning is applied during inversion, the inversion achieves the best resolution from the first iteration. By utilizing the cross-variance among the VTI parameters, the inversion also correctly resolves the less constrained anisotropic parameter. Therefore, a properly preconditioned WEMVA inversion provides a reliable tool for anisotropic model building.

ACKNOWLEDGEMENTS

The author thanks BP for providing the synthetic anisotropic model.

REFERENCES

Bachrach, R., 2010a, Applications of deterministic and stochastic rock physics modeling to anisotropic velocity model building: SEG Expanded Abstracts, **29**, 2436–2440.

———, 2010b, Elastic and resistivity anisotropy of compacting shale: Joint effective medium modeling and field observations: SEG Expanded Abstracts, **29**, 2580–2584.

Bachrach, R., Y. K. Liu, M. Woodward, O. Zradrova, Y. Yang, and K. Osypov, 2011, Anisotropic velocity model building using rock physics: Comparison of compaction trends and check-shot-derived anisotropy in the gulf of mexico: SEG Expanded Abstract, **30**, 207–211.

Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2009, Can we distinguish TTI and VTI media?: SEG Expanded Abstracts, **28**, 226–230.

———, 2010a, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, 27–36.

———, 2010b, Localized anisotropic tomography with well information in VTI media: Geophysics, **75**, 37–45.

Bandyopadhyay, K., 2009, Seismic anisotropy: geological causes and its implications: PhD thesis, Stanford University.

Bear, L., T. Dickens, J. Krebs, J. Liu, and P. Traynin, 2005, Integrated velocity model estimation for improved positioning with anisotropic PSDM: The Leading Edge, 622–634.

Claerbout, J. F., 2009, Image estimation by example.

Clapp, R., 2000, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.

Dræge, A., M. Jakobsen, and T. A. Johansen, 2006, Rock physics modeling of shale diagenesis: Petroleum Geoscience, **12**, 49–57.

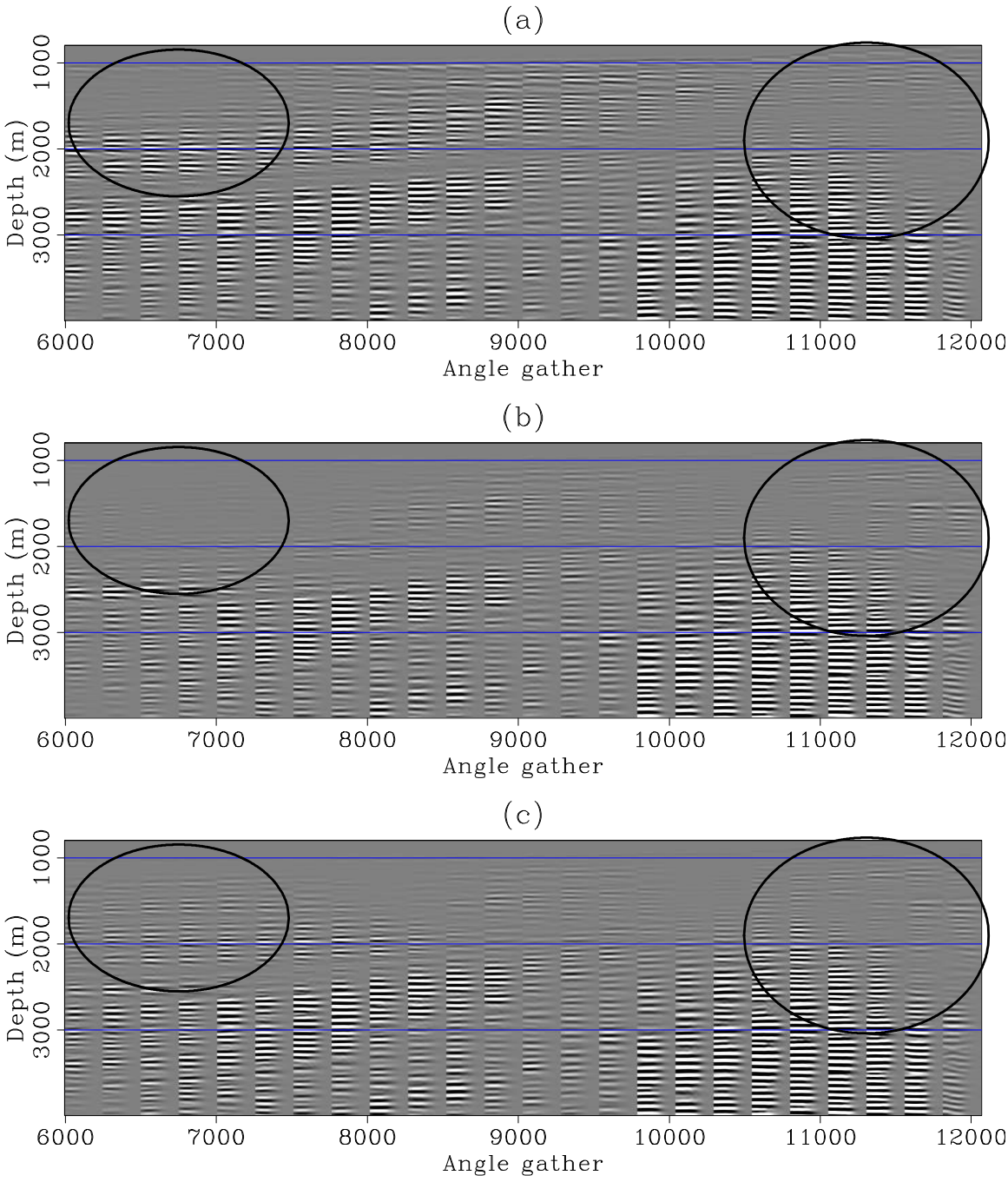

Figure 21: Final ADCIGs with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Notice the improved angle coverage in the ellipse on the left and the improved flatness on the right. [**CR**]

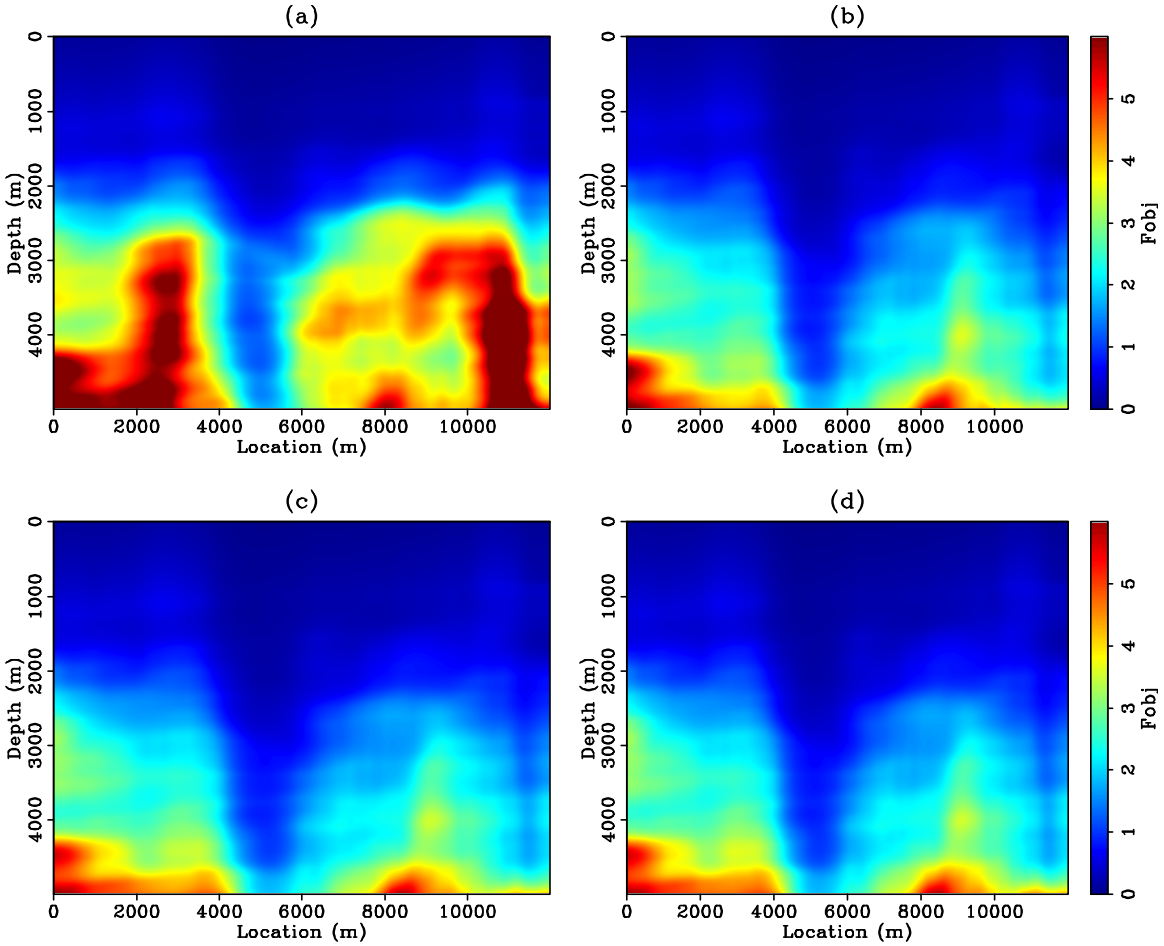

Figure 22: Comparison of the initial objective function map (a), with the final objective function map with (b) no preconditioning, (c) diagonal preconditioning, and (d) full preconditioning. [**CR**]

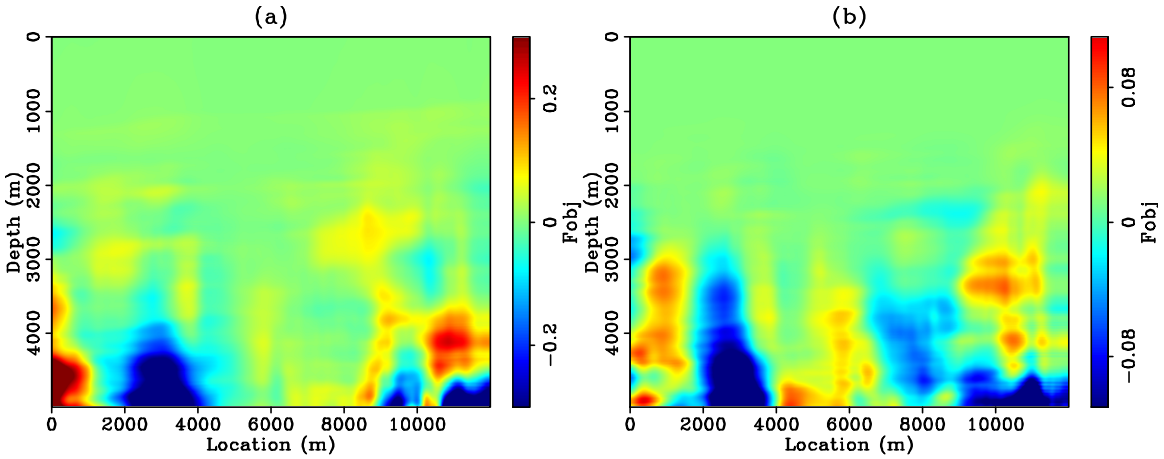

Figure 23: Differences between the objective function maps. (a) Difference between Figure 22(d) and 22(b). (b) Difference between Figure 22(d) and 22(c). [**CR**]

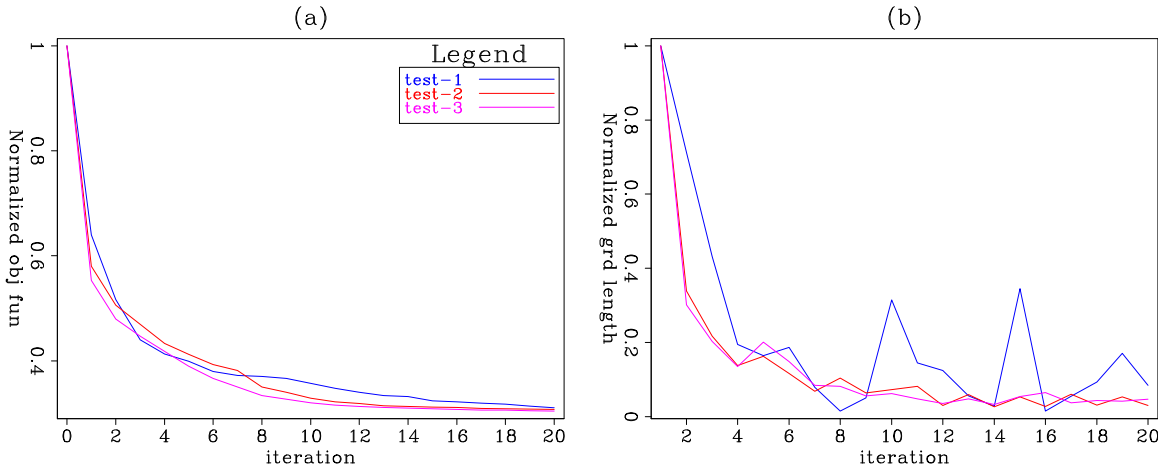

Figure 24: Objective function value of the data fitting goal in (a) and the size of the gradient in (b) as a function of the iteration. Blue line: no preconditioning; red line: diagonal preconditioning; magenta line: full preconditioning. [**CR**]

Hornby, B., D. Miller, C. Esmersoy, and P. Christie, 1995, Ultrasonic-to-seismic measurements of shale anisotropy in a North Sea well: SEG Expanded Abstracts, **14**, 17–21.

Li, Y. and B. Biondi, 2011, Migration velocity analysis for anisotropic models: SEG Expanded Abstract, **30**, 201–206.

Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011, Anisotropic tomography using rock physics contraints: 73^{rd} EAGE Conference & Exhibition.

Li, Y. E., B. Biondi, R. Clapp, and D. Nichols, 2013, Wave equation migration velocity analysis for VTI models: SEP-Report, **150**, 157–168.

Li, Y. E., R. Clapp, B. Biondi, and D. Nichols, 2014, Rock physics constrained anisotropic WEMVA: Part II - Field data test : SEP-Report, **152**, 105–128.

Osypov, K., D. Nichols, M. Woodward, O. Zdraveva, and C. E. Yarman, 2008, Uncertainty and resolution analysis for anisotropy tomography using iterative eigendecomposition: SEG Expanded Abstracts, **27**, 3244–3249.

Sayers, C., 2004, Seismic anisotropy of shales: What determines the sign of Thomsen's delta parameter?: SEG Expanded Abstracts, **23**, 103–106.

———, 2010, The effect of anisotropy on the Young's moduli and Poisson's ratios of shales: SEG Expanded Abstracts, **29**, 2606–2611.

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

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

Yang, Y., K. Osypov, R. Bachrach, M. Woodward, O. Zdraveva, Y. Liu, A. Fournier, and Y. You, 2012, Anisotropic tomography and uncertainty analysis with rock physics constraints: Green Canyon case study: SEG Expanded Abstract, 1–5.