

Chapter 2

WEMVA for VTI models

Anisotropic models are needed for wave simulation and inversion where a complex geological environment exists. In this chapter, I extend the theory of wave equation migration velocity analysis (WEMVA) to vertical transverse isotropic (VTI) media. The accuracy of the VTI model is mapped to the focusing of the prestack image in the subsurface offset domain, which is then evaluated by a combined differential semblance optimization objective function and a stacking power maximization objective function. Because of the lack of constraints, I assume δ can be accurately obtained from other sources of information, and invert for the NMO slowness and the anellipticity parameter η .

To regularize the multi-parameter inversion, I build a regularization framework to adapt the geological information to guide the updates in both NMO slowness and η . This regularization step is crucial to stabilize the inversion and to produce geologically meaningful results. I test the proposed approach on both a synthetic dataset and a 2-D Gulf of Mexico dataset, in both cases started with fairly good initial VTI models. The inversion results reveal shallow anomalies collocated in NMO velocity and η and improve both the continuity and the resolution of the final stacked images.

INTRODUCTION

When compared with isotropic models, anisotropic models are recognized as more realistic representations of the subsurface in complex geological environment (McCollum and Snell, 1932; Postma, 1955; Helbig, 1956). Many authors (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008) have developed migration and processing schemes for vertical transverse isotropic (VTI) and tilted transverse isotropic (TTI) media; however, the challenge of building a reliable anisotropic Earth model remains a bottleneck for the exploration workflow.

In my thesis study, I am interested in retrieving the kinematics of the pressure waves in acoustic media. Surface seismic data inversion for anisotropy is ill-posed and highly underdetermined due to the rapidly increasing size of the model space with the increasing complexity of the subsurface (Woodward et al., 2008; Bakulin et al., 2010b,a). To reduce the size of the model space, Sarkar and Tsvankin (2003, 2004) and Takanashi and Tsvankin (2012) parametrize their model space using factorized blocks or layers, where the anisotropic parameters ϵ and δ are constant and the velocity variations are simple linear or quadratic functions of the space coordinates. These factorization methods dramatically decrease the number of unknowns to be solved by the inversion, and hence stabilize the inversion. However, these models tend to over simplify the Earth where heterogeneities may occur at all scales.

In principle, grid-based tomographic methods handle both anisotropy and heterogeneity at the same time. Most of the existing grid-based anisotropic model-building schemes are based on ray-based tomography by measuring the non-hyperbolic moveout to flatten the common image gathers (CIG). The residual moveout is then back-projected along the ray paths to update the anisotropic models (Zhou et al., 2003, 2004; Yuan et al., 2006; Cai et al., 2009; Woodward et al., 2008). However, ray-based methods are prone to inaccurate results when multi-pathing exists in areas of complex overburden. Furthermore, due to the ill-posed and underdetermined nature of the anisotropic tomography, small amounts of noise in the residual moveout on a CIG gather can lead to significant variations in the inverted model. Therefore, I propose

a wave-equation-based tomography method to build the anisotropic model.

In this chapter, I first generalize the methodology of image-space WEMVA (Biondi and Sava, 1999; Sava and Biondi, 2004a,b) from an isotropic medium to an anisotropic medium. I use grid-based representations of velocity, ϵ and δ to parametrize the VTI medium. I show that theoretically the gradient of the tomographic objective functional for an isotropic medium can be modified to describe an anisotropic medium by simply adding corresponding terms for the additional parameters. Due to the lack of constraints, I reparameterize the subsurface using NMO velocity, η and δ , but keep δ fixed during the iterative WEMVA updates. Moreover, I precondition the model using geological information. Finally, I demonstrate on a synthetic dataset and a field dataset that the proposed method can identify anomalies in the velocity and the anisotropic model, and hence produce better migrated images with better spatial continuity and higher resolution.

WAVE EQUATION MIGRATION VELOCITY ANALYSIS FOR ANISOTROPY

In this section, I extend the methodology of WEMVA to VTI media. I provide a concise derivation of the WEMVA gradients using the Lagrangian augmented functional. The derivation using perturbation theory can be found in Appendix A. The interpretations of the adjoint-state equations suggest the same implementation as perturbation theory suggests.

Governing equations

Assuming that the shear wave velocity is much smaller than the pressure wave velocity, one approximation of the up-going and down-going one-way acoustic wave equations for VTI media is shown as follows (Shan, 2009):

$$\left(\frac{\partial}{\partial z} \mp i\Lambda \right) P = 0, \quad (2.1)$$

where $P = P(x, y, z, \omega)$ is the wavefield in the space-frequency domain and Λ describes the dispersion relationship in terms of P-wave vertical velocity v_v and Thomsen parameters ϵ and δ (Thomsen, 1986):

$$\Lambda = \frac{\omega}{v_v} \sqrt{\frac{\frac{\omega^2}{v_v^2} - (1 + 2\epsilon)|\mathbf{k}|^2}{\frac{\omega^2}{v_v^2} - 2(\epsilon - \delta)|\mathbf{k}|^2}}, \quad (2.2)$$

where $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector.

When implementing the wave-equation, I adapt the optimized implicit finite difference scheme (Shan, 2009) to approximate the dispersion relation in Equation 2.2 using a rational function:

$$\Lambda = \frac{\omega}{v_v} \left(1 - \sum_{i=1}^n \frac{\alpha_i |\mathbf{k}|^2}{\frac{\omega^2}{v_v^2} - \beta_i |\mathbf{k}|^2} \right), \quad (2.3)$$

where the coefficients α_i and β_i are the optimized coefficients for particular ϵ and δ values. Prior to wave-equation simulation, I build tables of coefficients for the whole range of possible ϵ and δ pairs. During the propagation, I look up the optimized implicit finite difference coefficients according to the local ϵ and δ values. Details about the optimized implicit finite differencing scheme is described in Appendix A.

In shot-profile migration, both source wavefields \mathbf{D} and receiver wavefields \mathbf{U} are downward continued using the following one-way wave equations and boundary conditions:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda \right) D(\mathbf{x}, \mathbf{x}_s) = 0 \\ D(x, y, z = 0, \mathbf{x}_s) = f_s \boldsymbol{\delta}(\mathbf{x} - \mathbf{x}_s) \end{cases}, \quad (2.4)$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\Lambda \right) U(\mathbf{x}, \mathbf{x}_s) = 0 \\ U(x, y, z = 0, \mathbf{x}_s) = d_{\mathbf{x}_s} \boldsymbol{\delta}(\mathbf{x} - \mathbf{x}_r) \end{cases}, \quad (2.5)$$

where f_s is the source wavelet, $d_{\mathbf{x}_s}$ is the shot record by a shot at $\mathbf{x} = \mathbf{x}_s$. To

differentiate from the anisotropic parameter δ , I use $\boldsymbol{\delta}$ to denote the Dirac delta. Equations 2.4 and 2.5 can be summarized in matrix forms as follows:

$$\mathbf{L}\mathbf{D} = \mathbf{f} \quad (2.6)$$

and

$$\mathbf{L}^*\mathbf{U} = \mathbf{d}, \quad (2.7)$$

where $\mathbf{L} = \frac{\partial}{\partial z} - i\Lambda$ and $*$ denotes the adjoint.

I use an extended imaging condition (Sava and Formel, 2006) to compute the image cube with subsurface offsets:

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{D})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{U}), \quad (2.8)$$

where $\mathbf{S}_{+\mathbf{h}}$ is a shifting operator which shifts the wavefield $+\mathbf{h}$ in the \mathbf{x} direction. Notice that $(\mathbf{S}_{+\mathbf{h}})^* = \mathbf{S}_{-\mathbf{h}}$. Equations 2.6, 2.7 and 2.8 are the state equations, and \mathbf{D} , \mathbf{U} and $I_{\mathbf{h}}$ are the state variables.

WEMVA gradients for VTI parameters

To evaluate the accuracy of the subsurface model, I use a combined differential semblance optimization (DSO) (Symes and Carazzone, 1991; Shen, 2004) and stacking power objective function (Soubaras and Gratacos, 2007):

$$J = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \langle I_{\mathbf{h}}\boldsymbol{\delta}(\mathbf{h}), I_{\mathbf{h}}\boldsymbol{\delta}(\mathbf{h}) \rangle. \quad (2.9)$$

where \mathbf{h} is the subsurface offset. The first term in the objective function J is the DSO term, and the second term denotes the stacking power term. The trade-off parameter α controls the relative strength between both terms. This composite objective function shares the merit of both objectives: the global convergence of DSO and the high resolution of stacking power maximization.

To derive the gradient of the WEMVA objective function with respect to the VTI parameters, I follow the adjoint-state recipe provided by Plessix (2006). First, I form the Lagrangian augmented functional:

$$\begin{aligned}
\mathcal{L}(\mathbf{D}, \mathbf{U}, I_{\mathbf{h}}; \lambda, \mu, \gamma_{\mathbf{h}}; v_v, \epsilon, \delta) = & \\
& \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h} I_{\mathbf{h}}, \mathbf{h} I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \langle I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}), I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}) \rangle \\
& + \langle \lambda, \mathbf{f} - \mathbf{L}(v_v, \epsilon, \delta) \mathbf{D} \rangle \\
& + \langle \mu, \mathbf{d} - \mathbf{L}^*(v_v, \epsilon, \delta) \mathbf{U} \rangle \\
& + \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}}, (\mathbf{S}_{+\mathbf{h}} \mathbf{D})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) - I_{\mathbf{h}} \rangle. \tag{2.10}
\end{aligned}$$

Then the adjoint-state equations are obtained by taking the derivative of \mathcal{L} with respect to the state variables \mathbf{D} , \mathbf{U} and $I_{\mathbf{h}}$:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{D}} = -\mathbf{L}^*(v_v, \epsilon, \delta) \lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) \gamma_{\mathbf{h}} = \mathbf{0}, \tag{2.11}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = -\mathbf{L}(v_v, \epsilon, \delta) \mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^* (\mathbf{S}_{+\mathbf{h}} \mathbf{D}) \gamma_{\mathbf{h}} = \mathbf{0}, \tag{2.12}$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} - \alpha I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}) = 0, \forall \mathbf{h}. \tag{2.13}$$

Equation 2.11, 2.12, and 2.13 are the adjoint-state equations. Parameters λ , μ and $\gamma_{\mathbf{h}}$ are the adjoint-state variables, and can be calculated from the adjoint-state equations.

The physical interpretation of the adjoint-state equations offers better understanding of the physical process and provides insights for implementation. Clearly, the solution to Equation 2.13, $\gamma_{\mathbf{h}}$, is the perturbed (residual) image at a certain subsurface offset. Equations 2.11 and 2.12 define the perturbed source and receiver wavefields, respectively. Notice the perturbed source wavefield λ at location \mathbf{x} depends on the image at $(\mathbf{x} - \mathbf{h}, \mathbf{h})$ and the background receiver wavefield \mathbf{U} at $\mathbf{x} - 2\mathbf{h}$. The same rule applies to the perturbed receiver wavefield μ .

With the solutions to the equations above, I can now derive the gradients of the

objective function 2.9 by taking the derivative of the augmented functional \mathcal{L} with respect to the model variables v_v , ϵ and δ as follows:

$$\nabla_{v_v} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \mathbf{v}_v} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \mathbf{v}_v} \mathbf{U} \right\rangle, \quad (2.14)$$

$$\nabla_{\epsilon} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \epsilon} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \epsilon} \mathbf{U} \right\rangle, \quad (2.15)$$

$$\nabla_{\delta} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \delta} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \delta} \mathbf{U} \right\rangle. \quad (2.16)$$

Operators $\frac{\partial \mathbf{L}}{\partial v_v}$, $\frac{\partial \mathbf{L}}{\partial \epsilon}$, and $\frac{\partial \mathbf{L}}{\partial \delta}$ and their adjoint operators are computed based on a polynomial approximation of the dispersion relation in Equation 2.2. Details are described in Appendix A.

Note that if the forward modeling operator \mathbf{L} is isotropic, Equation 2.14 reduces to the gradient of the isotropic WEMVA objective function with respect to the isotropic velocity. Therefore, the extension of WEMVA to the anisotropic medium is simply adding related terms to the additional anisotropic parameters.

Figure 2.1 shows the gradients of the objective function when an impulsive image $I_{\mathbf{h}} = \delta(x, y, z - 1500, \mathbf{h})$ is evaluated in a homogeneous VTI medium. The background VTI parameters are: $v_v = 2\text{km/s}$, $\epsilon = 0.1$ and $\delta = 0.05$. These impulse responses are often referred as ‘‘sensitivity kernels’’ (eg. Marquering et al. (1998, 1999); Rickett (2000)). The top row displays the kernels for v_v , ϵ , and δ when a single pair of source and receiver is placed at zero offset. The zero-offset kernel for vertical velocity in panel (a) nicely covers the wavepaths from the source to the image point and from the image point to the receiver. Due to the lack of sensitivity of the vertical rays to the anisotropic parameter ϵ , the zero-offset kernel for ϵ in panel (d) is saturated with the near field artifacts. In panel (g), the zero-offset kernel for δ shows similar pattern as for v_v , except for the reduced sensitivity with depth.

The middle row in Figure 2.1 displays the kernels when a single source-receiver pair is placed at 4000m offset. All three kernels cover the wavepaths from the source to the image point and from the image point to the receiver. Reading from the scale bar in panel (b), a lower sensitivity of the slanted rays to the vertical velocity is observed compared with that of the vertical rays. To the contrary, higher sensitivities for ϵ and δ are observed from panel (e) and panel (h).

To show the relative strength of the sensitivity kernels, the summation of the kernels in the first two rows is plotted in the bottom row in Figure 2.1. It is obvious and natural that the vertical rays have higher sensitivities for vertical velocity. Meaningful sensitivity for ϵ only comes from the slanted traveling waves. Similarly, sensitivity for δ mainly comes from the slanted traveling waves.

Parameterization

Many authors (Tsvankin and Thomsen, 1994; Alkhalifah and Tsvankin, 1995) have shown that surface recorded P-wave traveltime can be characterized by the NMO velocity, v_n , and the anellipticity parameter η . Based on the definition in Equation 1.2 and 1.4, the one-way wave-equation can be formulated in terms of v_n , η and δ by a change of variables:

$$\left(\frac{1}{\sqrt{1+2\delta}} \frac{\partial}{\partial z} \mp i\Lambda' \right) P = 0 \quad (2.17)$$

where

$$\Lambda' = \frac{\omega}{v_n} \sqrt{1 - \frac{\omega^2}{s_n^2} \frac{|\mathbf{k}|^2}{2\eta|\mathbf{k}|^2}}. \quad (2.18)$$

Notice that in this formulation (Equation 2.17), Thomsen parameter δ and the derivative in depth $\frac{\partial}{\partial z}$ are coupled with each other. This is a theoretical proof of the well-observed δ and depth ambiguity (Plessix and Rynja, 2010). To constrain the δ parameter, well log or well marker information is needed to add the depth dimension into the inversion.

Considering the surface seismic data only, I assume δ can be correctly obtained

from other sources of information (such as check shots and well logs) and keep it fixed through the WEMVA iterations. When $\eta = 0$, in the so-called elliptical anisotropic medium, the dispersion relation 2.18 reduces to the isotropic dispersion relation with a vertically stretched NMO slowness.

Instead of re-deriving the gradients using the new formulation of the dispersion relation 2.17, I derive the gradients of the objective function 2.9 with respect to NMO velocity v_n and η using chain rule and the previously derived gradients (Equation 2.14 to 2.16):

$$\begin{aligned}\nabla_{v_n} J &= \left(\frac{\partial J}{\partial v_n} \right)^T = \left(\frac{\partial J}{\partial v_v} \frac{\partial v_v}{\partial v_n} \right)^T + \left(\frac{\partial J}{\partial \delta} \frac{\partial \delta}{\partial v_n} \right)^T \\ &= \left(\frac{\partial v_v}{\partial v_n} \right)^T \nabla_{v_v} J + \left(\frac{\partial \delta}{\partial v_n} \right)^T \nabla_{\delta} J,\end{aligned}\quad (2.19)$$

$$\begin{aligned}\nabla_{\eta} J &= \left(\frac{\partial J}{\partial \eta} \right)^T = \left(\frac{\partial J}{\partial \epsilon} \frac{\partial \epsilon}{\partial \eta} \right)^T + \left(\frac{\partial J}{\partial \delta} \frac{\partial \delta}{\partial \eta} \right)^T \\ &= \left(\frac{\partial \epsilon}{\partial \eta} \right)^T \nabla_{\epsilon} J + \left(\frac{\partial \delta}{\partial \eta} \right)^T \nabla_{\delta} J,\end{aligned}\quad (2.20)$$

where

$$\frac{\partial v_v}{\partial v_n} = \frac{1}{\sqrt{1+2\delta}}; \quad \frac{\partial \delta}{\partial v_n} = \frac{\sqrt{1+2\delta}}{v_v}; \quad (2.21)$$

$$\frac{\partial \epsilon}{\partial \eta} = 1+2\delta; \quad \frac{\partial \delta}{\partial \eta} = -\frac{(1+2\delta)^2}{1+2\epsilon}. \quad (2.22)$$

Figure 2.2 shows the same impulse responses as in Figure 2.1, except for a different parametrization: NMO velocity, η and δ . As shown by the first column, the sensitivity for the NMO velocity is very similar to that for the vertical velocity. Compared with Figure 2.1(b), the sensitivity for NMO velocity to the slanted rays is slightly higher; however, the increase in sensitivity is small (less than 2%) due to the small influence of δ on NMO velocity. Impulse response of η has a negative sign, reflecting the fact

that the surface reflection P-wave data have higher sensitivity for δ than for ϵ . The interferences between the ϵ gradients and δ gradients distribute the sensitivity of η more evenly across different propagation angles.

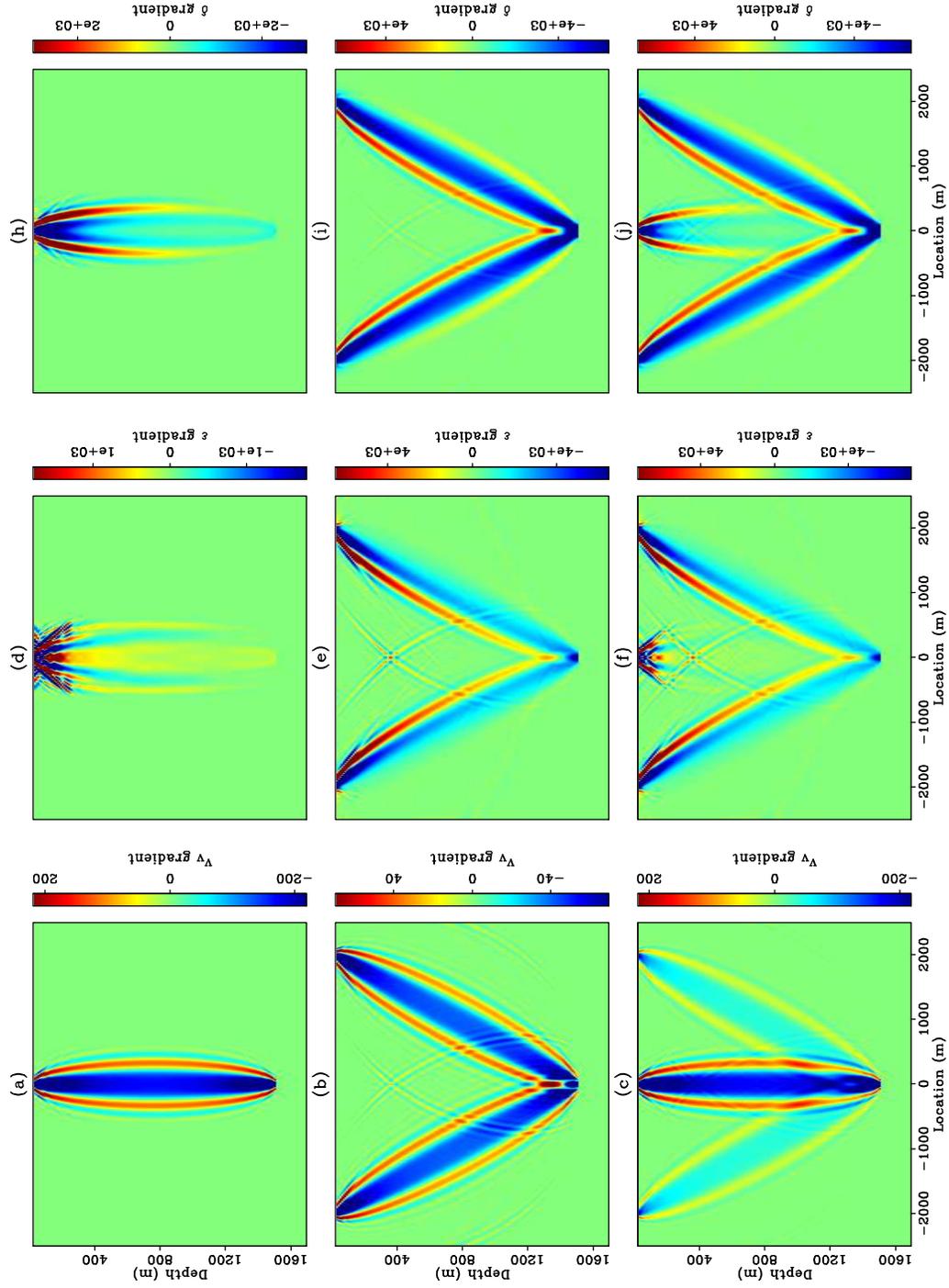


Figure 2.1: 2D impulse responses for vertical velocity (left column), ϵ (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the two rows above. [ER] chap2/. 2dkernelchap2

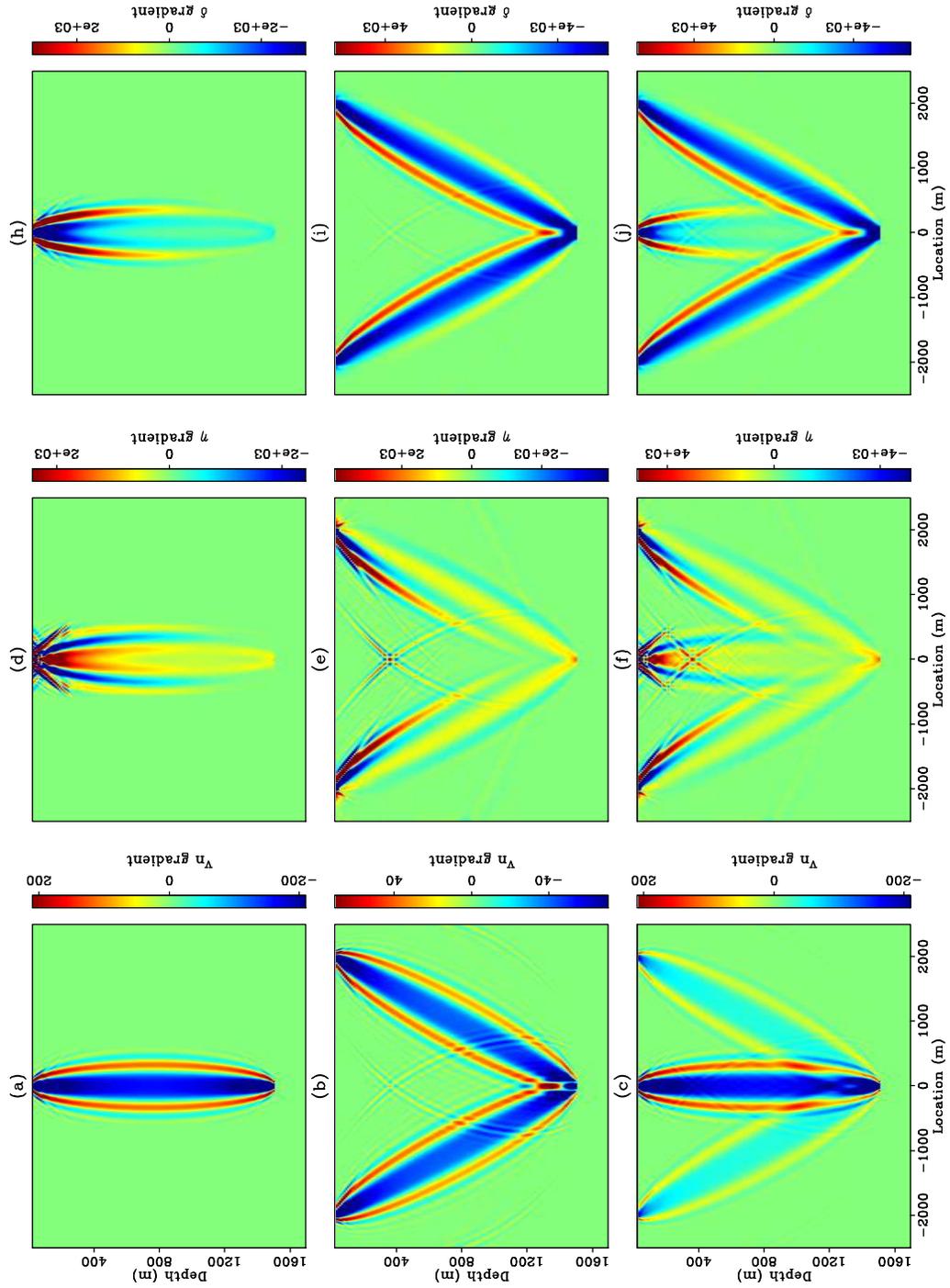


Figure 2.2: 2D impulse responses for NMO velocity (left column), η (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the top two rows. [ER] chap2/. 2dkervn

Preconditioning the WEMVA gradient

Subsurface model building is a highly underdetermined and nonlinear inverse problem. Therefore, prior knowledge of the subsurface is needed to stabilize the inversion and to define a plausible subsurface model. Assuming Gaussian distribution, Tarantola (1984) characterizes the prior information using the mean $\mathbf{m}_{\text{prior}}$ and the covariance \mathbf{C}_M of the model and includes it as a regularization term. The resulting regularized objective function is as follows:

$$J(\mathbf{m}) = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \left\| \sum_{\theta} \mathbf{I}(\mathbf{x}, \theta) \right\|^2 + \beta \frac{1}{2} (\mathbf{m} - \mathbf{m}_{\text{prior}})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{\text{prior}}). \quad (2.23)$$

where the first two terms define the “data fitting” objective, and the third “model regularization” term. The scalar β is to balance the relative strength of the objectives.

A classic regularizer is a roughening operator (Claerbout, 2009), so that the roughness of the model is penalized. However, this regularization formulation suffers from low convergence due to the small length of the roughening operator (Clapp, 2000). To speed up the convergence, a regularization problem is often reformulated to a preconditioning problem by introducing a preconditioning variable (Polak, 1997). After preconditioning, the short roughening operator is replaced by a long smoothing operator with. The convergence is hence improved both by the longer spread of the information, and by the improved condition number of the matrix we are to invert.

If we define the preconditioned variable $\mathbf{n} = \mathbf{C}_M^{-\frac{1}{2}} (\mathbf{m} - \mathbf{m}_{\text{prior}})$, the regularized objective function 2.23 is reformulated as follows:

$$J(\mathbf{n}) = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \left\| \sum_{\theta} \mathbf{I}(\mathbf{x}, \theta) \right\|^2 + \beta \frac{1}{2} \|\mathbf{n}\|^2. \quad (2.24)$$

Setting $\beta = 0$, the preconditioned objective function 2.24 is essentially the same as the original data fitting objective function 2.9, except for the different model parameters. Hence, the gradient of the preconditioned WEMVA objective function (2.24) with

respect to this preconditioned variable \mathbf{n} can be obtained by the chain rule:

$$\begin{aligned}\nabla_{\mathbf{n}}J &= \left(\frac{\partial\mathbf{m}}{\partial\mathbf{n}}\right)^*\nabla_{\mathbf{m}}J \\ &= \mathbf{C}_M^{\frac{1}{2}*}\nabla_{\mathbf{m}}J,\end{aligned}\tag{2.25}$$

where $\nabla_{\mathbf{m}}J = [\nabla_{s_n}J \ \nabla_{\eta}J \ \nabla_{\delta}J]^T$ is the gradient of the objective function 2.9 with respect to the original model parameters.

In the single-parameter (velocity) tomography, the geological dip information can be included into the inversion using the steering filters to spread the velocity updates along the structure (Clapp, 2000). The generalization of the same idea to the multi-parameter anisotropic tomography is fairly straightforward: similar steering filters can also be applied to smooth the anisotropic parameters (Woodward et al., 2008).

The additional character about the multi-parameter anisotropic tomography is that the VTI parameters are correlated with each other. The correlations are mainly determined by the local lithology. Therefore, a cross-parameter covariance is also needed to constrain the anisotropic tomography. In many exploration areas, the lithological information is not available. In this situation, the only information we know is the relative ratio between velocity and the anisotropic parameters. Assuming proportional errors in each parameter, we have a rough estimate of the diagonal elements of the cross-parameter covariance matrix.

In this chapter, I assume there are two component of the covariance matrix: spatial covariance and local cross-parameter covariance. I assume they can be separated independently from each other (Li et al., 2011). A set of precondition operators are used: smoothing filters to approximate square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the model covariance matrix is approximated by the following formulation:

$$\mathbf{C}_M^{\frac{1}{2}} = \mathbf{\Sigma}\mathbf{S}.\tag{2.26}$$

In Equation 2.26, the smoothing operator \mathbf{S} is a band-limited diagonal matrix:

$$\mathbf{S} = \begin{vmatrix} \mathbf{S}_v & 0 & 0 \\ 0 & \mathbf{S}_\eta & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad (2.27)$$

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

$$\mathbf{\Sigma} = \begin{vmatrix} \mathbf{C}_{vv} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{C}_{\eta\eta} \mathbf{I} & 0 \\ 0 & 0 & 0 \end{vmatrix}^{\frac{1}{2}}. \quad (2.28)$$

The diagonal elements \mathbf{C}_{vv} and $\mathbf{C}_{\eta\eta}$ denote the variance of the NMO velocity and η , respectively. They are used to balance the relative strength of the velocity and η updates. Notice that both the bottom rows of \mathbf{S} and $\mathbf{\Sigma}$ contain all zeros, which remove updates in the δ model.

To explain the preconditioning scheme, I analyze the preconditioning effect assuming a non-linear steepest decent inversion framework. The initial preconditioned model \mathbf{n}_0 is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \langle \mathbf{m}_0 - \mathbf{\Sigma} \mathbf{S} \mathbf{n}_0, \mathbf{m}_0 - \mathbf{\Sigma} \mathbf{S} \mathbf{n}_0 \rangle. \quad (2.29)$$

For the i_{th} iteration, the preconditioned variable is obtained by

$$\mathbf{n}_{i+1} = \mathbf{n}_i + \lambda_i \nabla_{\mathbf{n}} J, \quad (2.30)$$

and hence the original model variable is

$$\begin{aligned}
\mathbf{m}_{i+1} &= \mathbf{S}\Sigma\mathbf{n}_{i+1} \\
&= \mathbf{S}\Sigma\mathbf{n}_i + \alpha_i\mathbf{S}\Sigma\nabla_{\mathbf{n}}J \\
&= \mathbf{m}_i + \alpha_i\mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\nabla_{\mathbf{m}}J.
\end{aligned} \tag{2.31}$$

Equation C.44 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:  $\mathbf{g}_s^0 = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_0$ 
initialize the search direction:  $\mathbf{p}_0 = -\mathbf{g}_s^0$ 
for  $k = 1 \cdots N_k$  do
  perform a line search: optimize  $\lambda$ ,  $\operatorname{argmin}_{\lambda} J(\mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1})$ 
  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:  $\mathbf{g}_s^k = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_k$ 
  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

```

SYNTHETIC DATA TEST

I test the proposed anisotropic WEMVA method on a modified BP2007 model.¹ Instead of using the original TTI synthetic data, I model my synthetic data assuming

¹The dataset was created by Hemang Shah and is provided at <http://software.seg.org/> courtesy of BP Exploration Operation Company Limited.

VTI. I also add a high velocity and high anisotropy anomaly in the shallow region of the model. This simulates the geological setting that local diagenetic recrystallization of clay minerals significantly increases both the velocity and the anisotropy.

The initial and the true velocity models are shown in Figure 2.3(a) and 2.3(c). The initial and the true η models are shown in Figure 2.4(a) and 2.4(c). I assume parameter δ is accurately obtained from other studies and use the true δ model in the inversion without updating it. I model the synthetic data using one-way VTI born modeling and perform anisotropic WEMVA for 60 iterations. I construct the preconditioner using the same triangular smoother for \mathbf{S}_v and \mathbf{S}_η , as well as a stationary matrix Σ with zero off-diagonal components to scale the gradients.

The inverted NMO velocity and η model are shown in Figure 2.3(b) and 2.4(b), respectively. Compared with the initial models, the inverted models are closer to the true models with improved definition of the shallow geology. I compare the true anomaly and the inverted anomaly in Figure 2.5. The anisotropic WEMVA inversion successfully identified the location of the anomaly in both NMO velocity and the η model. I increase the trade-off parameter α from 0.5 to 0.8 after 30 iterations, because the stacking power objective function can provide higher resolution gradients once the models are close to the true ones. However, the resolution of the η model is still limited by the illumination angles. Nevertheless, the cumulative contribution of this lower resolution anomaly in η explains the kinematic error in the P-wave propagation.

Figure 2.6 shows the data fitting objective function value (Equation 2.9) as a function of the iteration number. To make the objective function value comparable, I keep $\alpha = 0.5$ for this plot even though α is increased to 0.8 after 30 iterations. The objective function curve flattens out between iteration 20 and 30, suggesting the inversion has converged at the current resolution. After increasing the trade-off parameter α , the objective function further decreases until the inversion converges around 60 iterations.

Figure 2.7 compares the initial migrated image in panel (a) with the inverted

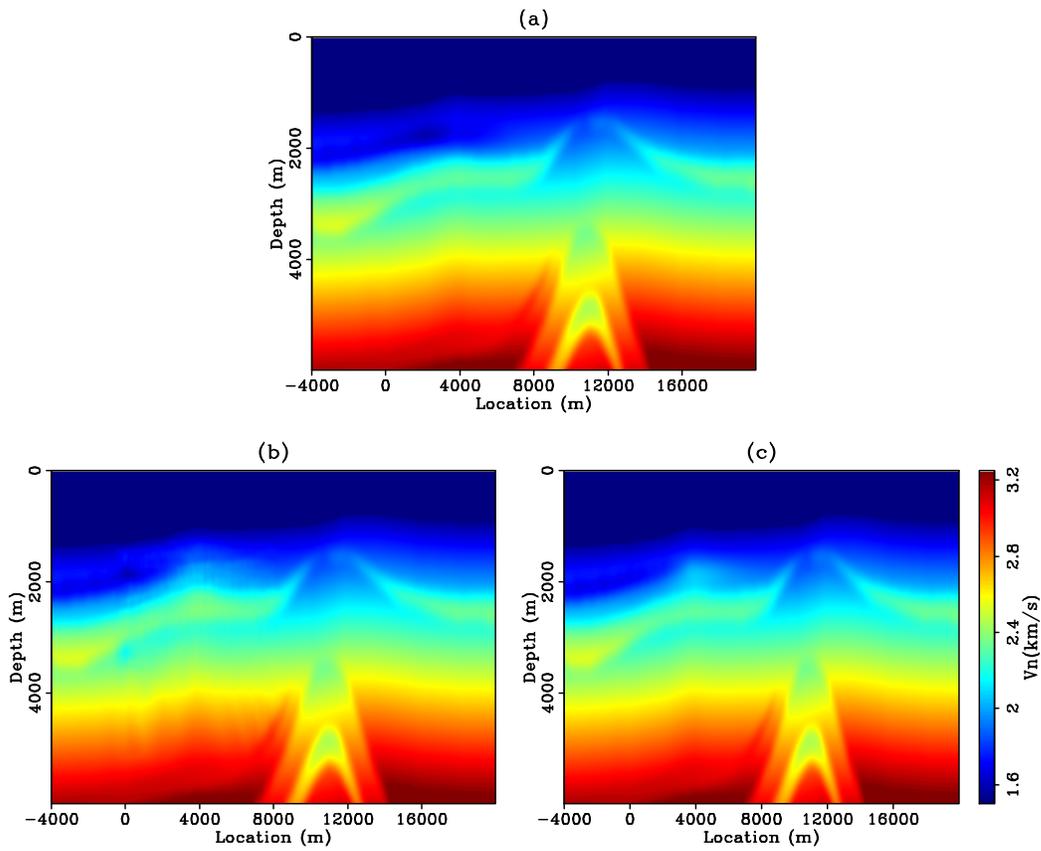


Figure 2.3: Modified BP2007 model. (a): initial NMO velocity; (b): inverted NMO velocity; and (c) true NMO velocity. [CR] chap2/. bp1-vp

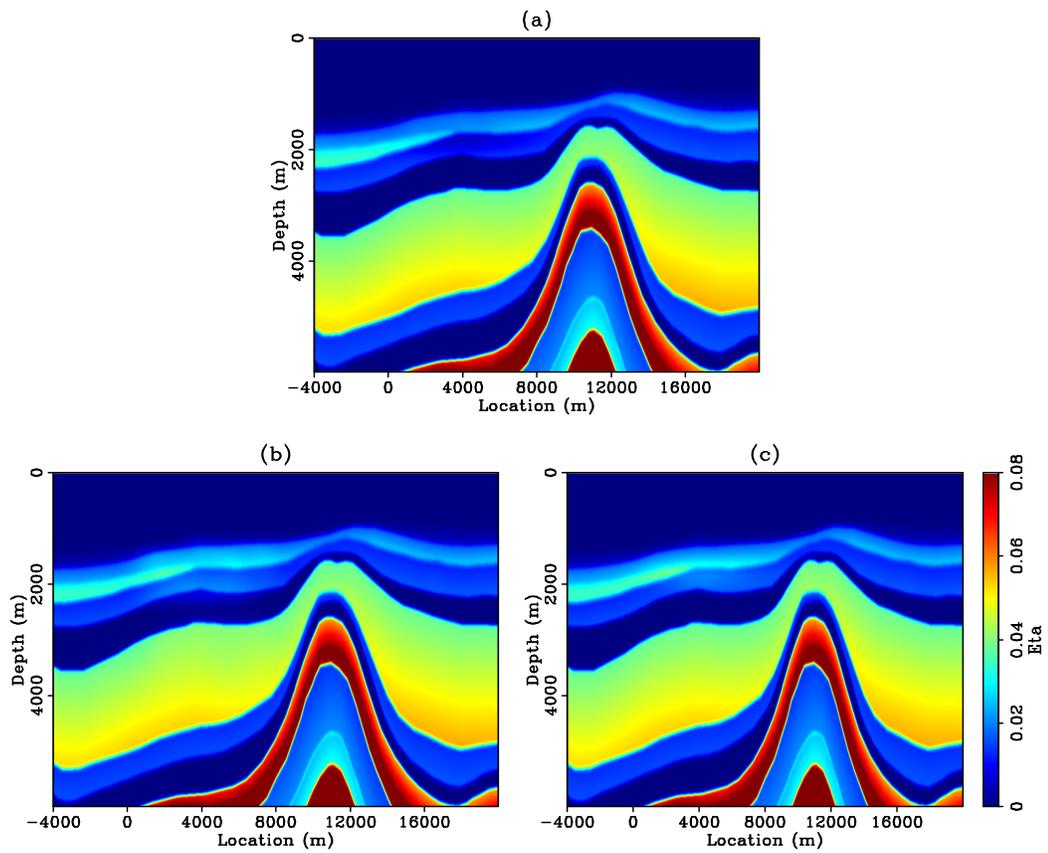


Figure 2.4: Modified BP2007 model. (a): initial η model; (b): inverted η model; and (c) true η model. [CR] chap2/. bp1-eta

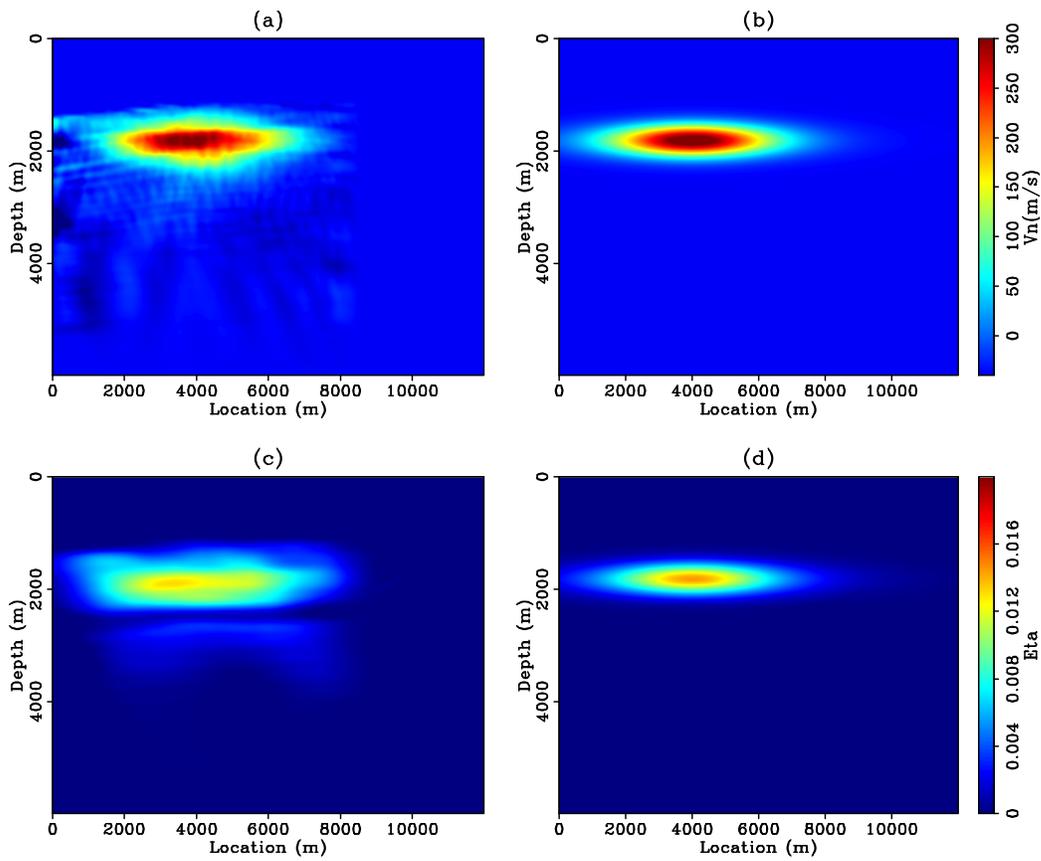
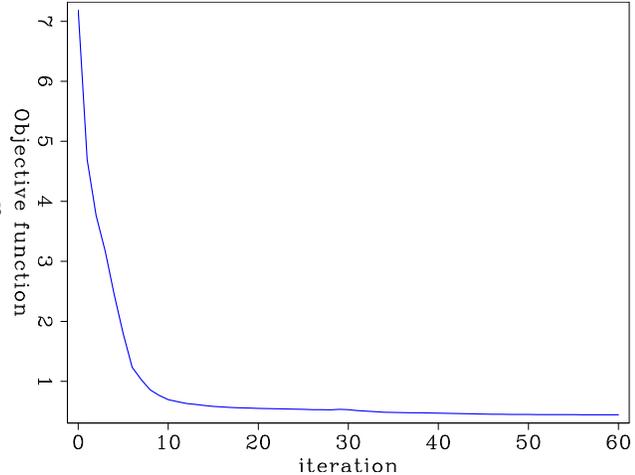


Figure 2.5: Comparison of the anomaly. (a): inverted velocity anomaly; (b): true velocity anomaly; (c) inverted η anomaly and (d) true η anomaly. [CR]

chap2/. bp1-updt

Figure 2.6: Objective function as a function of iteration. [CR]
 chap2/. bp1-obj



migrated image in panel (b) and the image obtained with true models in panel (c). The improved velocity and η model yield a better image with correct depth, better focusing, and different interpretations of the subsurface. Compared with the initial image, the reflectors in the inverted image have been pushed down to the correct depths. Flag 1 and 2 in Figure 2.7 point out two representative reflectors whose positions have been corrected by the inversion. The vertical shifts of the deeper reflectors at 5 km depth are larger than 200m. Highlighted by the ellipses, the image shows more continuous reflectors and better focusing after the anisotropic WEMVA inversion. Moreover, flag 3 marks a region where an anticline trap could have been misinterpreted with the leaking point at 4 km in the initial image. It is no longer misleading in the inverted image.

Figure 2.8 compares the initial ADCIGs in panel (a) with the inverted ADCIGs in panel (b). The ADCIGs are extracted every 250m from $x = 2000\text{m}$ to $x = 8000\text{m}$. The initial gathers show strong upward curvatures in the shallow region, indicating undermigration. Interestingly, although both the velocity and η perturbations are positive, we observe downward curvatures in the deeper sections below 4km. This is because the wavepaths bend significantly due to the velocity and η anomaly so that the effective average velocity along the initial wavepaths are higher than that along

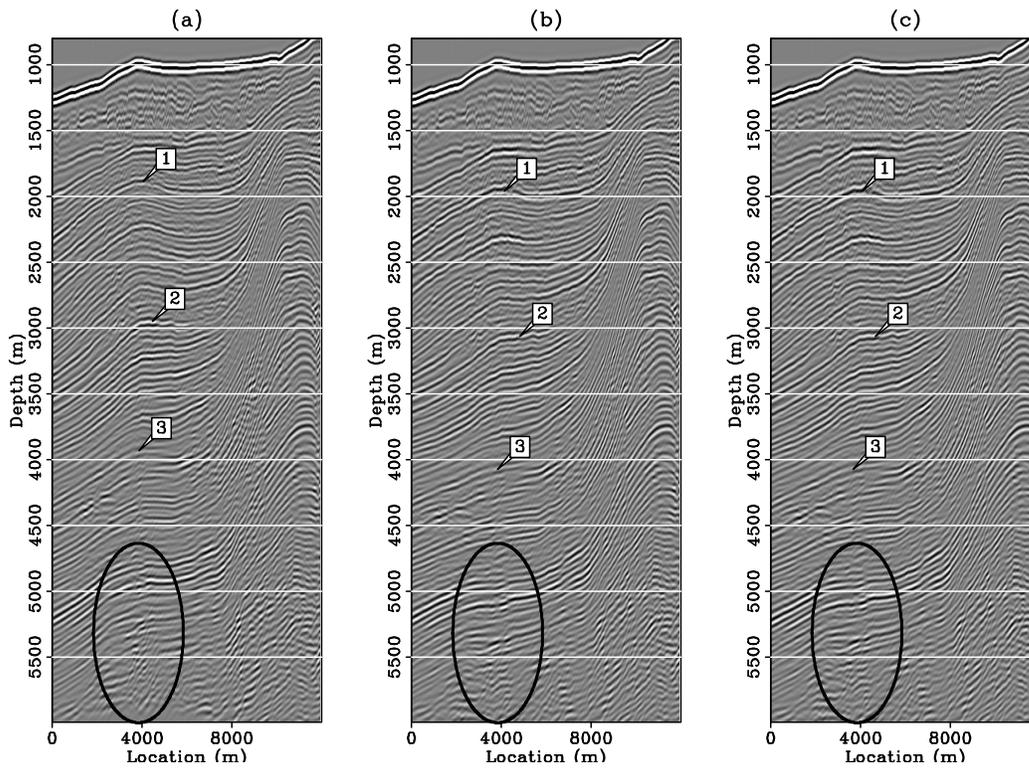


Figure 2.7: Stack image using (a): initial anisotropic model; (b): inverted anisotropic model; and (c) true anisotropic model. Flags 1 and 2 mark the depth of two representative reflectors in this section. Notice the depths of both reflectors have been corrected in the inverted image. The ellipses highlight the region with better focusing and more continuous reflectors. [CR] chap2/. bp1-imag

the correct wavepaths.

After 60 iterations of the anisotropic WEMVA inversion, the events in the ADCIGs are almost completely flattened. Notice both the upward and the downward curving residual moveouts are corrected, confirming that the inverted velocity and η model explains the kinematics of the data.

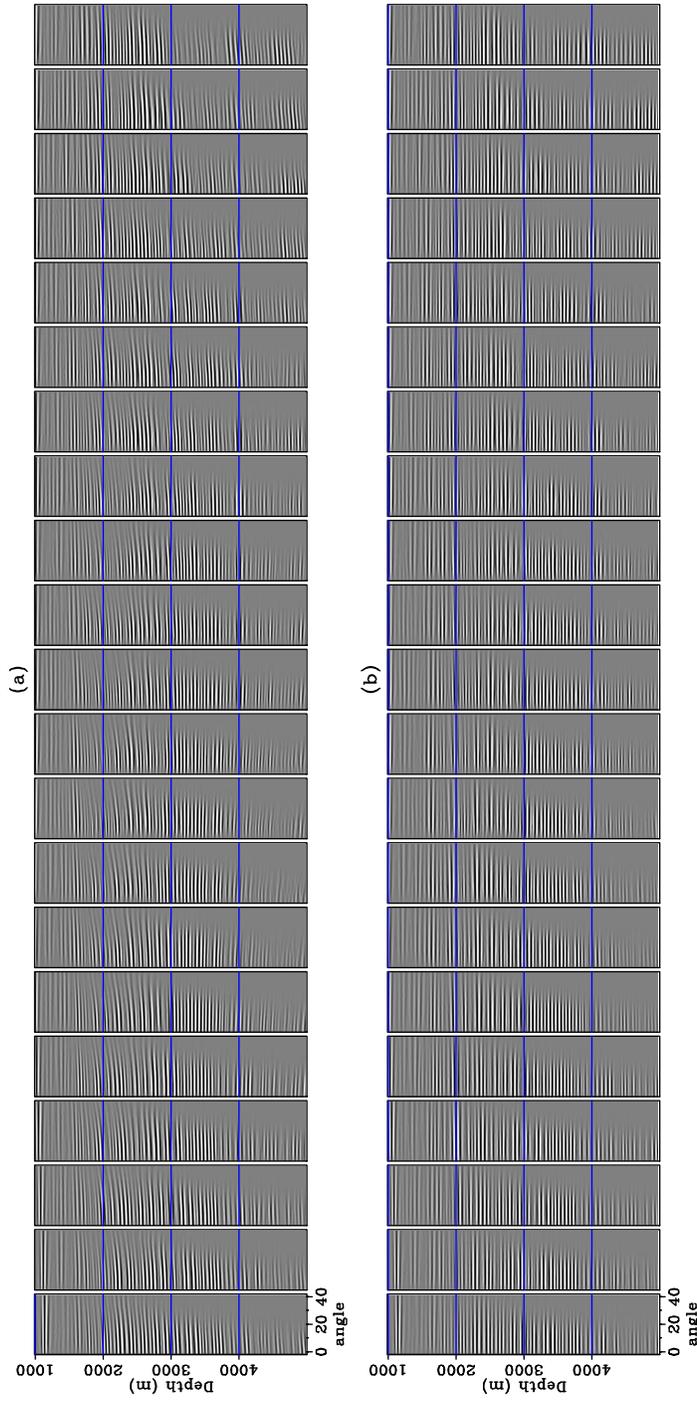


Figure 2.8: Angle-domain common image gathers extracted at 250m interval between $x = 2000\text{m}$ and $x = 8000\text{m}$. (a): ADCIGs using initial anisotropic model, (b): ADCIGs using inverted anisotropic model. [CR]

chap2/. bp1-aimg

FIELD DATA TEST

Anisotropic models are better defined in well-developed and production areas, where various well log, check shots and VSP data are available. The Gulf of Mexico dataset used in this field data test is from a producing area in an offshore sedimentary basin close to a salt body. The current VTI model was built using an interactive visualization method (Bear et al., 2005) that integrates surface seismic, sonic logs, vertical check shots, and offset check shots. Figure 2.9 illustrates the available datasets in this area.

From the bottom row in Figure 2.10, we can see that the PSDM gathers are almost flat, inverted velocities have a reasonable match with the sonic logs, and the modeled traveltimes agrees with the offset check-shot measurements. Therefore, I assume the current anisotropic model for δ is accurate enough and keep it fixed through the WEMVA iterations.

However, according to the color code in Figure 2.9, although most of the traveltimes are fitted very well for near- to mid-offset check shots (indicated by rays in white), the travel times modeled between long-offset shots and the downhole receivers are still greater than the measured travel times (indicated by rays in green). Therefore, a fully automated anisotropic model building method utilizing all types of data would be highly valuable to improve the current model.

I extract a source line from the Gulf of Mexico 3D dataset away from the salt body. The source and receiver positions are plotted in Figure 2.11. Ignoring the cross line offsets, sources are grouped every 100m, and receivers are grouped every 50m. The maximum offset is about 4 km. Using the initial models in Figure 2.12, I obtain the initial migrated image which is shown on the top panel in Figure 2.13.

Many small-scale faults are visible in this area, as shown in the initial migrated image. Migration artifacts at $x = 13$ km and $z = 2500$ m are caused by the sharp discontinuity in the velocity defined by a major fault running from $x = 14$ km on the top to the bottom of the section. The initial angle gathers are shown in the bottom row in Figure 2.13. Because of the streamer geometry, the subsurface reflectors are

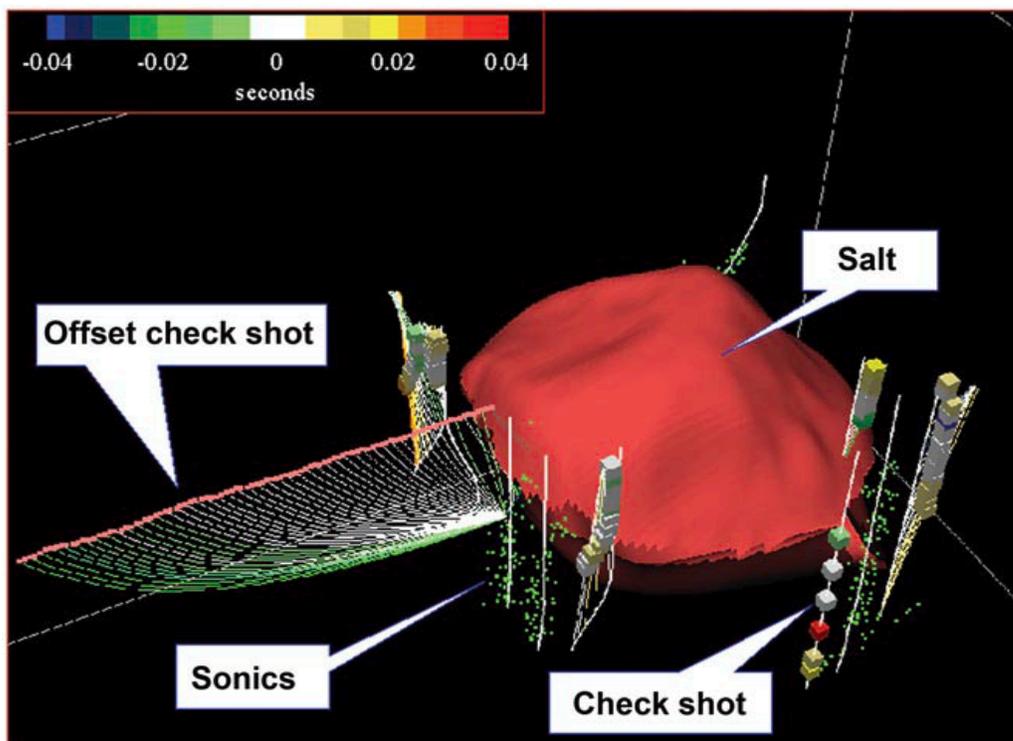


Figure 2.9: Available datasets for anisotropic model building. From Bear et al. (2005).

[NR] chap2/. alldata

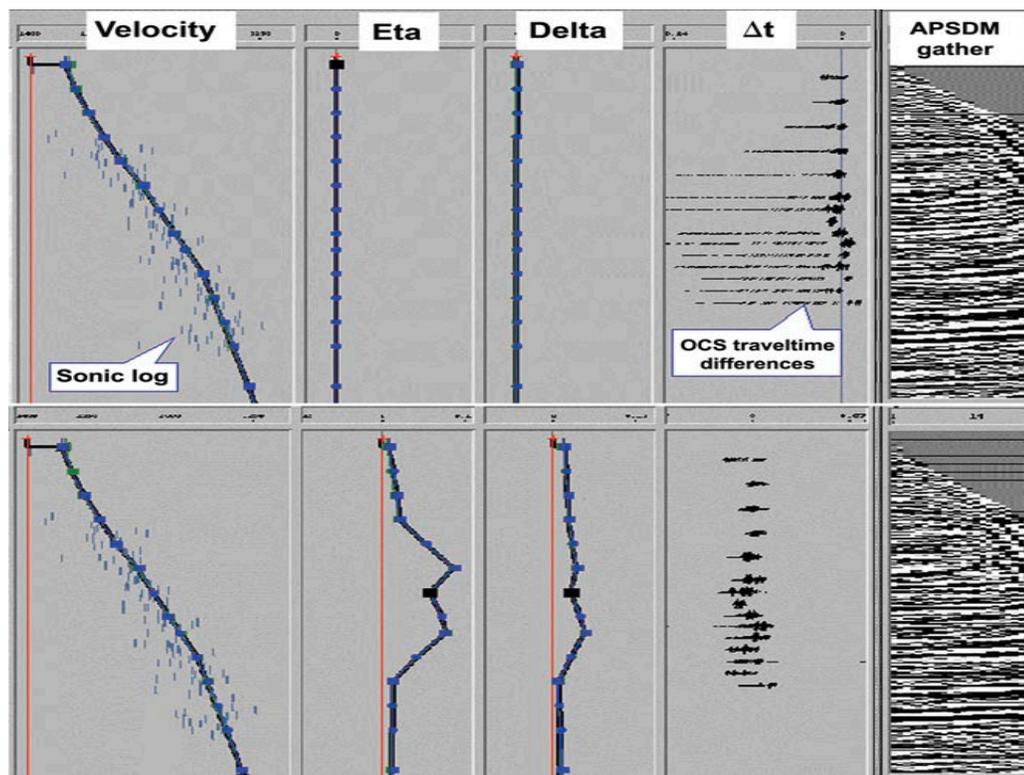


Figure 2.10: Visualization of interactive modeling for anisotropy. The best anisotropic model honors sonic logs, OCS traveltimes and flattens the prestack surface seismic gathers at the same time. From Bear et al. (2005). [NR] chap2/. interact

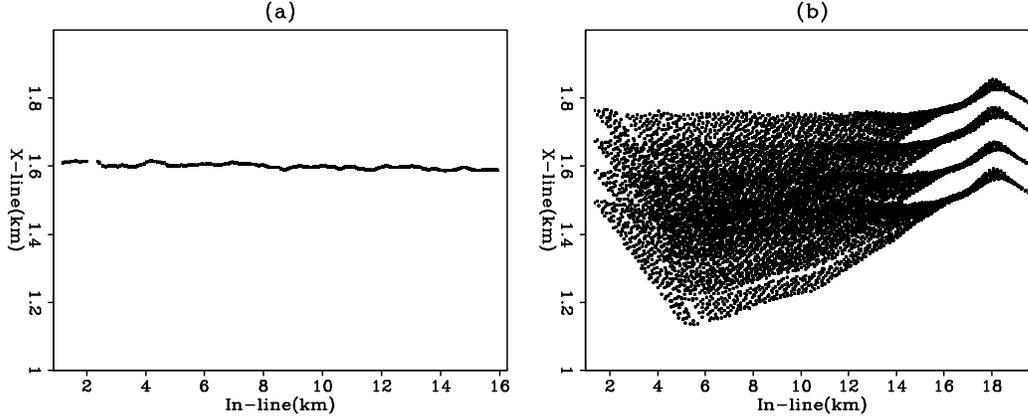


Figure 2.11: Source locations on a 2D line in (a) and corresponding receiver locations in (b). [CR] chap2/. exxongeo

only illuminated from positive angles. Although the gathers are fairly flat, small upward residual moveouts remain in the angle domain common image gathers. This agrees well with the negative residual travel time misfit in the OCS survey. Therefore, there is a chance to improve the model and the image by flattening the gathers.

To constrain the tomography, I estimate the dip field (Figure 2.14(a)) from the initial migrated image. The section shows mostly positive dips. The small negative dips correspond to the large reverse fault to the right of this section. Based on the dip field, non-stationary steering filters are constructed at each grid point. To demonstrate the steered smoothing effect, I apply the steering filter on a section filled with random numbers. The resulting field is shown in Figure 2.14(b). Notice the “synthetic reflectors” created by the steering filters follow the structural dip.

During inversion, the non-stationary steering filters are used to smooth the gradients for both NMO velocity and η . Due to the lack of rock physics information, a stationary matrix Σ with zero off-diagonal components is applied, simply to scale the relative amplitude of the slowness and η gradient. I use a large weight $\alpha = 0.8$ for the stacking power objective function, because the current VTI model is already well defined.

The inverted anisotropic model after 8 iterations is shown in Figure 2.15. Updates between the initial and the inverted VTI models are shown in Figure 2.16(a) and 2.16(b). The inversion successfully resolves a localized shallow anomaly between 13 km and 15 km at around 800 m depth. Notice that the updates in velocity are less than 10% of the background model, indicating the high quality of the initial model. The updates in η are around 25%. These positive updates in both velocity and η agree well with the negative travel time misfits in the previous OCS modeling results (Bear et al., 2005).

Comparing the initial migrated image (top panel in Figure 2.13) with the final migrated image (top panel in Figure 2.17), improved continuity and signal strength can be seen in the area highlighted by the oval. The faulting in this area is also better defined in the final image. If I correlate the inverted models with the final migrated image, I might change the interpretation of the major fault in the initial model: the emerging location on the surface might be shifted from 14.2 km to around 16 km. However, this change of interpretation can only be confirmed with other areal geological information.

The final ADCIGs produced using the inverted models are shown in the bottom row in Figure 2.17. To better illustrate the effects of the model updates, the ADCIGs are sampled more densely between $\text{CMP} = 13\text{km}$ and 16km and sparsely outside of this range. In general, we can see improved flatness for almost all the reflectors. However, DSO and stacking power objective functions are global optimization schemes. They improve the overall flatness and focusing in each iteration but cannot guarantee flatter events at each image point. For example, the ADCIGs at $\text{CMP} = 15\text{km}$ between the depth of 500m and 900m are slightly over compensated to flatten the gathers in the deeper region. More iterations should eventually flatten all the gathers in principle; however, we do not pursue higher resolution results with more iterations in order to avoid overfitting the noise in the image space.

We can verify the effect of the improvements in velocity and η on the angle gathers at different image points. Specifically, for the shallower events above 1km, most improvements happen at large angles over 35° . Therefore, I interpret the improvements

for the shallow events primarily as the contribution of the improved η model. For the deeper events at the same CMP location, both the depth and the flatness of the angle gather have been changed by inversion. The upward-curving events in the angle domain from the initial migration has been flattened by the improvements in both velocity and η model.

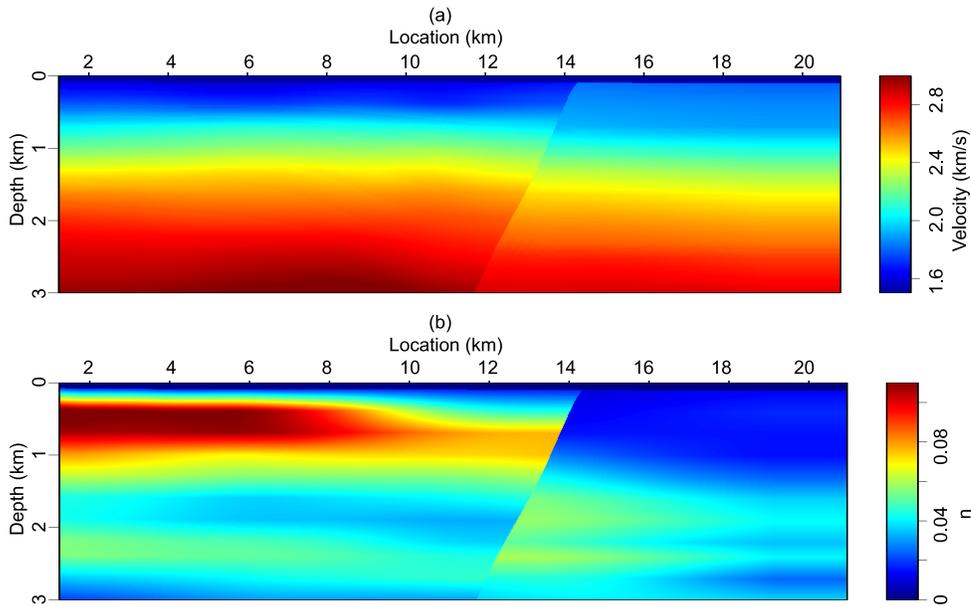


Figure 2.12: Initial velocity model (a) and initial η model (b). [CR]

chap2/.initmodel

DISCUSSIONS AND CONCLUSIONS

I have presented a methodology of wave equation migration velocity analysis for anisotropic media. By utilizing the Lagrangian adjoint-state method, I demonstrated

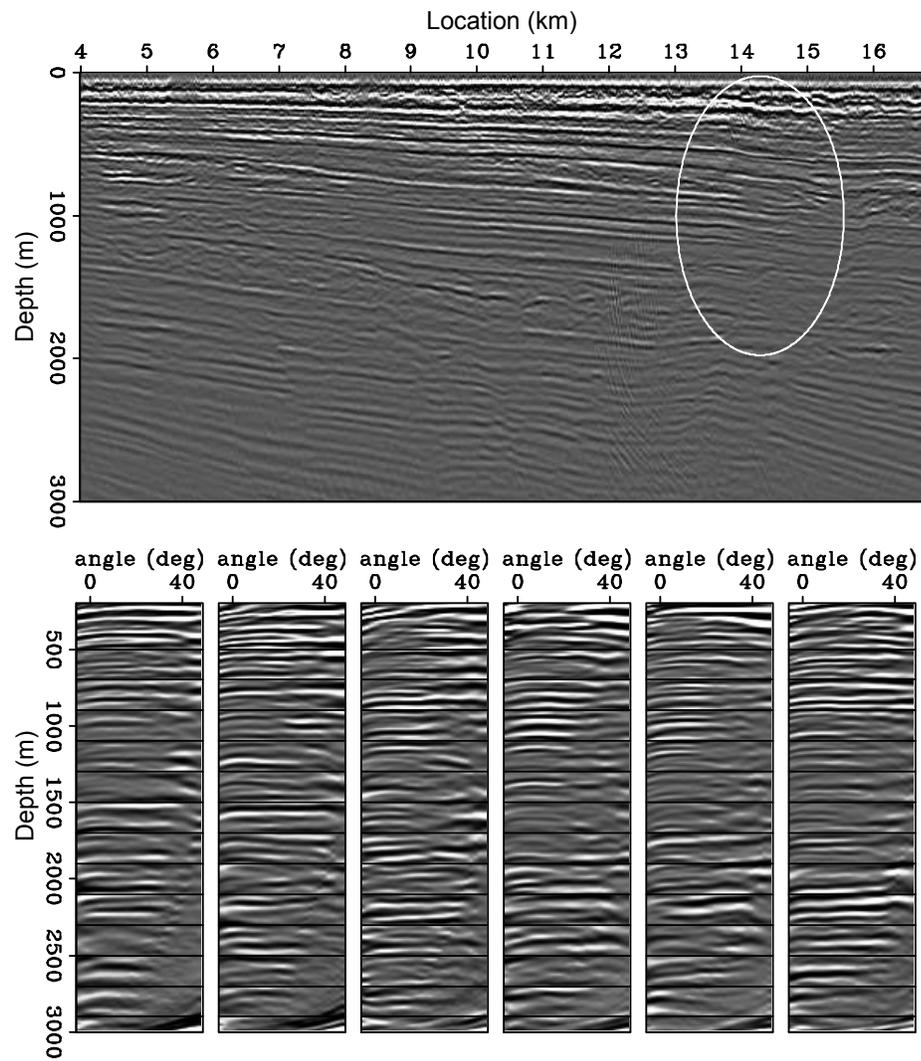


Figure 2.13: The initial stack image (Top panel) and initial angle domain common image gathers at $\text{CMP} = 7, 10, 13, 14, 14.5, 15$ km (Bottom row). [CR] chap2/. image-init-an

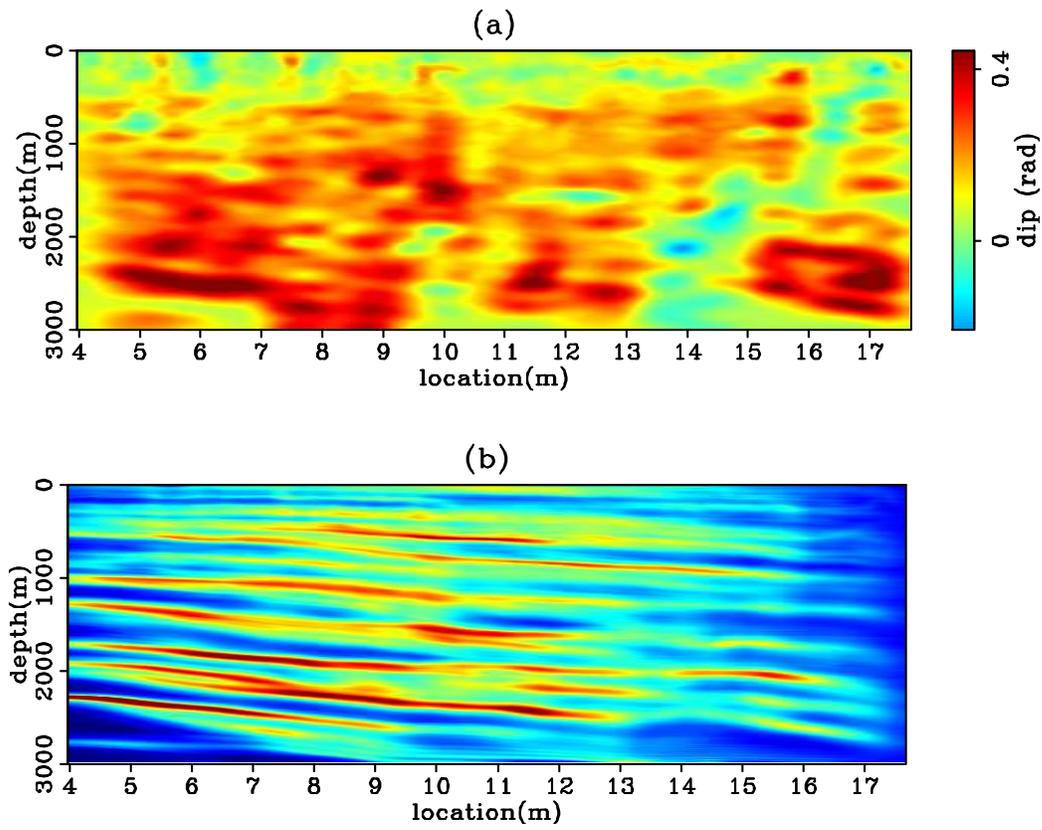


Figure 2.14: (a): Estimated dip field from the initial image on the top panel of Figure 2.13. (b): Reflectors created from a randomly distributed field. [CR]

chap2/. showsteer

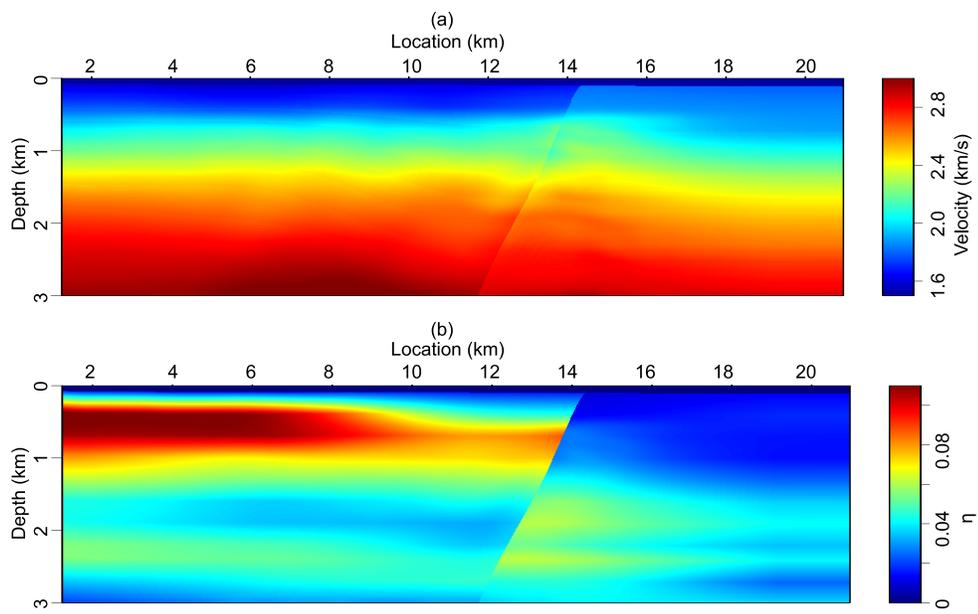


Figure 2.15: Inverted velocity model (a) and inverted η model (b). [CR]
chap2/. finalmodel

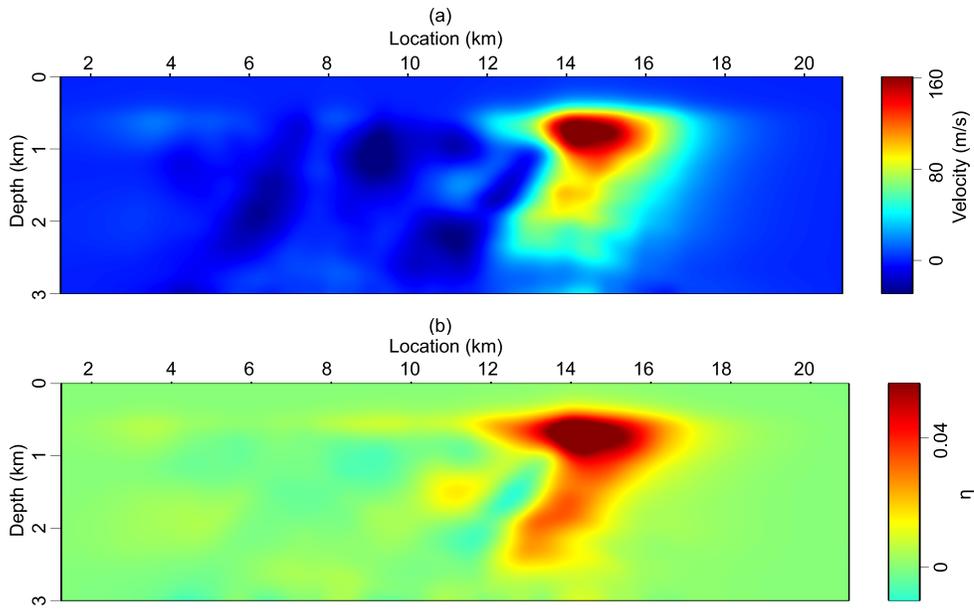


Figure 2.16: Updates in velocity model (a) and updates in η model (b) after inversion.

[CR] chap2/. updates

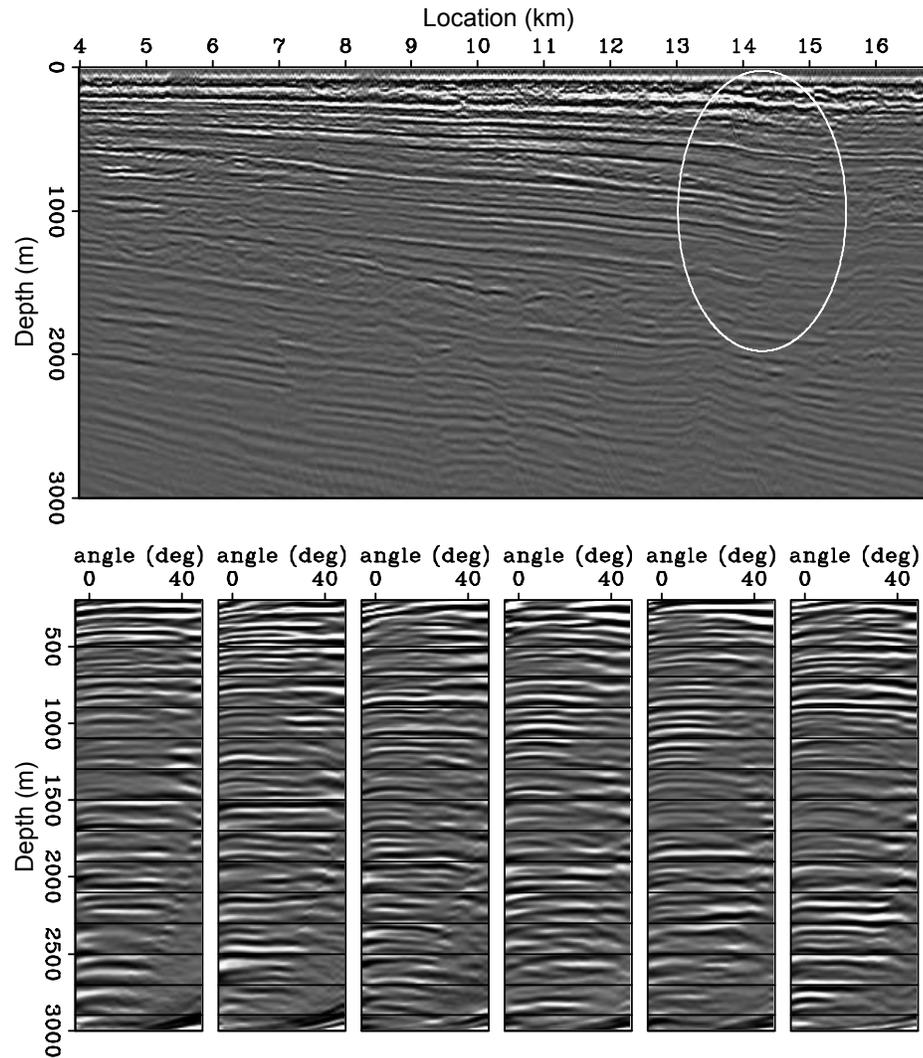


Figure 2.17: The final stack image (Top panel) and final angle domain common image gathers at CMP = 7, 10, 13, 14, 14.5, 15 km (Bottom row). Compared with Figure 2.13, improvements in continuity and enhancements in amplitude strength are highlighted by the oval. [CR] chap2/. image-fnal-an

that our anisotropic WEMVA methodology is a straightforward extension of the well-established isotropic WEMVA methodology: simply adding an additional term for the additional parameter η . Our method is fully automated and picking-free. The grid-based parameterization for the velocity and anisotropic parameters provides the flexibility to reveal perturbations with different scales.

To evaluate the accuracy of the inverted anisotropic models, I combine the DSO objective function and the stacking power objective function, benefiting from the global convergence of the former and the high resolution from the latter. In practice, I turned up the weights on the stacking power objective function as I got closer to the true solution. Nevertheless, the resolution for the anellipticity parameter η is lower than that for NMO velocity due to its weaker influence on the data.

The preconditioning scheme of our proposed method is crucial to obtain geologically meaningful results for both velocity and anisotropic parameters. This scheme can also easily incorporate human interpretation and other prior knowledge. When the data are fitted equally well, I subjectively choose to trust a smoother anisotropic model rather than a highly heterogeneous isotropic model. In addition, I assume the anellipticity parameter η , being the ratio between the directional velocities, is smoother in space than the velocities. These smoothness assumptions mitigate the ambiguities between anisotropy and heterogeneity.

When the lithological and well log information is not available, I can only estimate the diagonal elements of the cross-covariance matrix, reflecting the autocorrelation of velocity and η perturbation separately. As a result, it takes many iterations and higher weights on the stacking power term in the objective function to obtain high resolution models and sort out the ambiguities between the parameters where the data permit (namely, in the shallow region). However, had I roughly known the lithological information in a certain region, I would be able to build a cross-covariance matrix which reflects the cross-correlation between velocity and η . Consequently, I would be able to speed up the convergence and better resolve the ambiguities (Li et al., 2011). I will explain the framework in Chapter 3.

Another way to mitigate the ambiguity between the parameters is to include more data. The proposed anisotropic WEMVA methodology works in the image space with reflection data, which are most sensitive to the NMO velocity. On the other hand, most of the current full waveform inversion practices work in the data space with long-offset refractions and diving waves (e.g. Gholami et al. (2013)). These transmission components of the seismic data are most sensitive to the horizontal velocity. Jointly inverting for both reflection and transmission data should provide better constraints on the anisotropic properties of the subsurface.

ACKNOWLEDGEMENT

The authors thank BP for providing the synthetic model and ExxonMobil for providing the field data.