Time-domain FWI in TTI media

Philipp Witte*, Mathias Louboutin, Felix J. Herrmann

FWI workshop Natal, September 2015

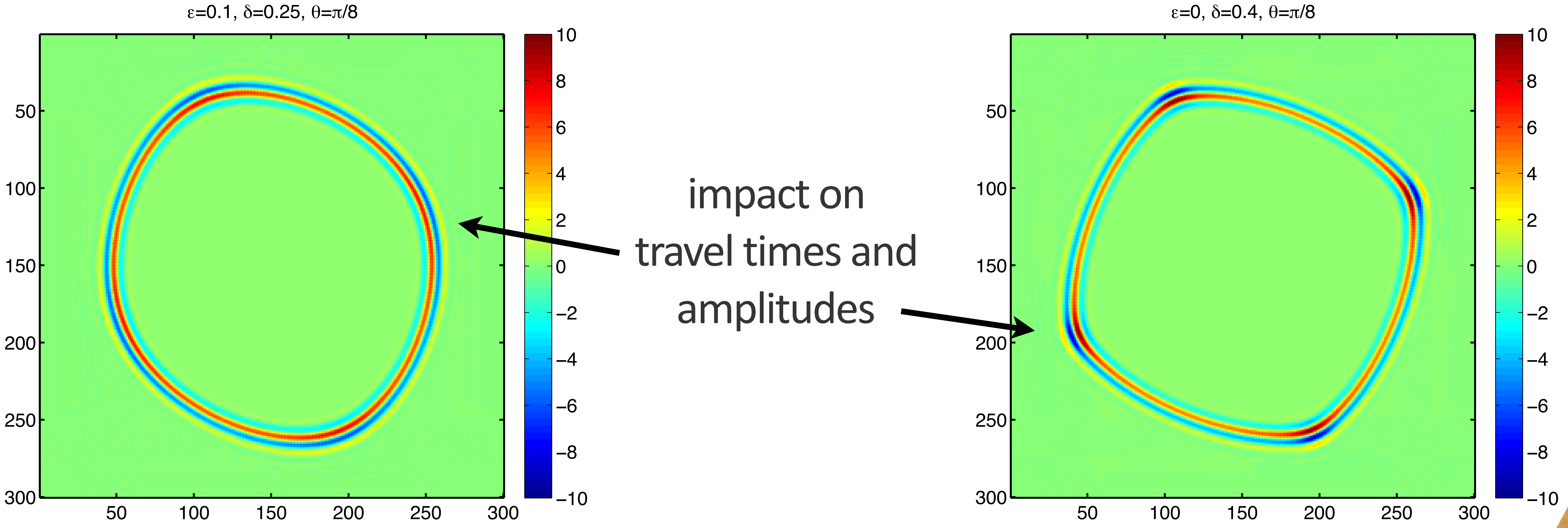
SLIM
University of British Columbia

Overview

- Introduction
- Isotropic time-domain modeling and inversion
- Extension to anisotropy
- Data examples
  - FWI of synthetic data set
  - RTM on field data set
- Alternative derivative operators
  - EPS method
  - Low-rank FD
- Conclusion
- Outlook

Introduction

- We need accurate modeling to match synthetic and observed data
- Anisotropy significantly affects wave propagation



Introduction

Two main approaches to anisotropic modeling:

- Pseudo-acoustic TTI wave equation
  - e.g. Zhang et al. (2011), CGG
  - coupled 2 x 2 system of equations with pseudo pressure field and auxiliary wave field
- Pure quasi-P wave TTI wave equation
  - e.g. Xu et al. (2014), Statoil or Chu et al. (2011), Conoco Philips
  - only one wavefield, free of S-wave artifacts

Isotropic modeling in the time-domain

Acoustic, isotropic wave equation in continuous form

$$m\frac{\partial^2 u}{\partial t^2} - \nabla^2 u = q$$

Time derivative: second order leap-frog scheme

$$\frac{\partial^2 u}{\partial t^2}(t) \approx \frac{\mathbf{u}^{k+1} - 2\mathbf{u}^k + \mathbf{u}^{k-1}}{\delta t^2}$$

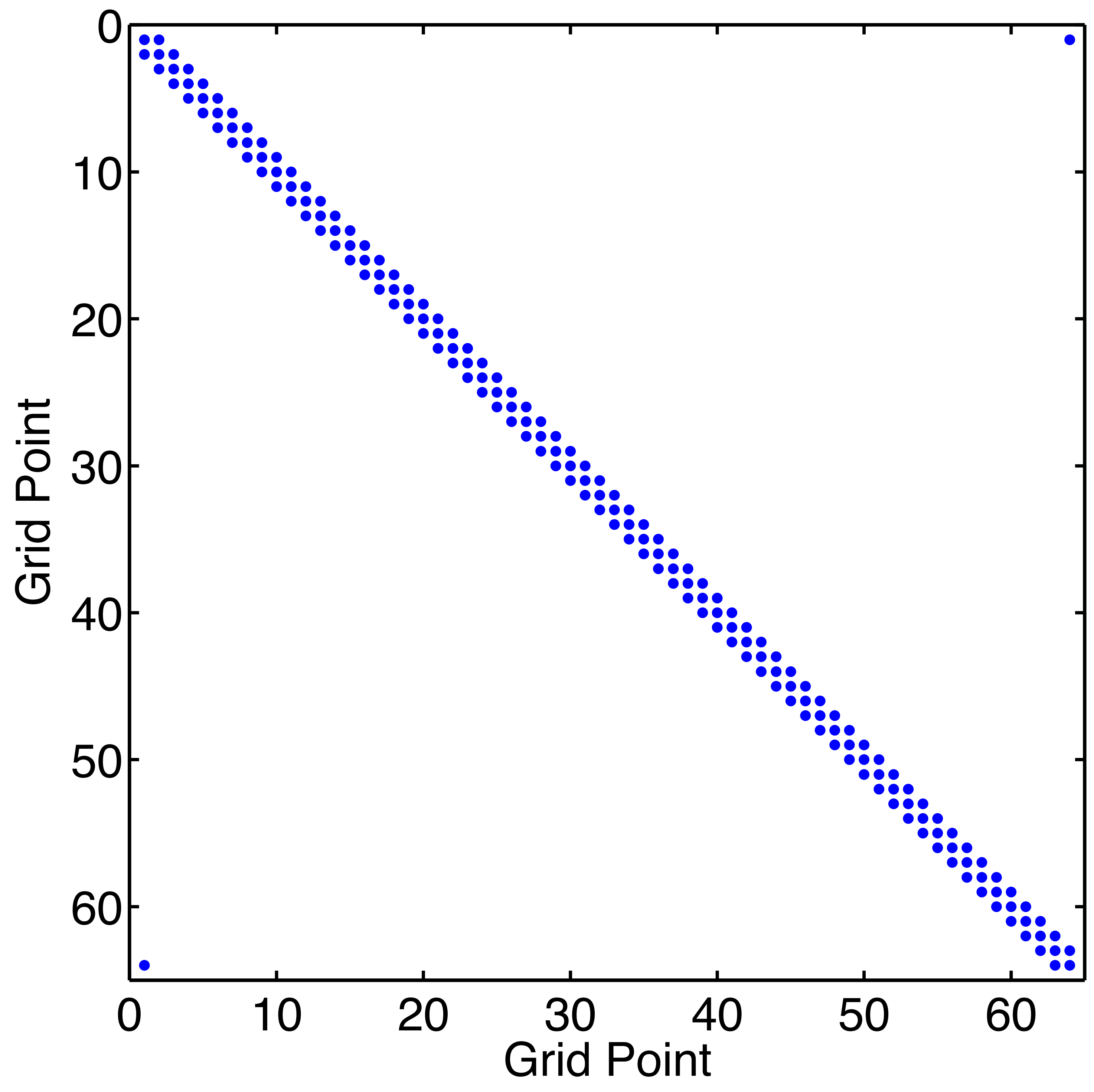
Isotropic modeling in the time-domain

1D Spatial derivative: FD schemes of various orders (2,4,6)

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{1}{h^2} \sum_{i=-\frac{p}{2}}^{\frac{p}{2}} \alpha_i \mathbf{u}_i = \mathbf{D_{xx}}$$

α_i : Coefficients of FD stencil

→ $\mathbf{D_{xx}}$ is sparse matrix



Isotropic modeling in the time-domain

Construction of 3D 2nd derivative operators:

$$\mathbf{L_x} = \mathbf{D_{xx}} \otimes \mathbf{I_{yy}} \otimes \mathbf{I_{zz}}$$

3D Laplacian is sum of 1D Laplacians

$$\mathbf{L} = \mathbf{L_x} + \mathbf{L_y} + \mathbf{L_z}$$

Advantages of operator in matrix form

- boundary conditions, FD order etc. can easily be modified without changing anything in time loop itself

Forward modeling

Fully discretized wave equation

$$\mathbf{A_1 u^{k+1} + A_2 u^{k} + A_3 u^{k-1} = q^{k-1}}$$

with:

$$\mathbf{A_1} = \mathbf{diag}(\frac{\mathbf{m}}{\triangle t^2})$$

$$\mathbf{A_3} = \mathbf{diag}(\frac{\mathbf{m}}{\triangle t^2})$$

$$\mathbf{A_2} = -\mathbf{L} - 2\mathbf{diag}(\frac{\mathbf{m}}{\triangle t^2})$$

$\mathbf{q^k}$: Source wave field at time step k

Forward modeling

Corresponds to one large system of equations $\mathbf{Au} = \mathbf{q}$

$$\begin{pmatrix} \mathbf{A_1} & 0 & 0 & \cdots & \cdots & 0 \\ \mathbf{A_2} & \mathbf{A_1} & 0 & & & \vdots \\ \mathbf{A_3} & \mathbf{A_2} & \mathbf{A_1} & & & \vdots \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{A_3} & \mathbf{A_2} & \mathbf{A_1} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \mathbf{u}^3 \\ \vdots \\ \vdots \\ \mathbf{u}^{nt} \end{pmatrix} = \begin{pmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \\ \mathbf{q}^3 \\ \vdots \\ \vdots \\ \mathbf{q}^{nt} \end{pmatrix}$$

that we solve in a time-stepping manner

$$\mathbf{u^{k+1}} = \mathbf{A_1}^{-1}(-\mathbf{A_2 u^k} - \mathbf{A_3 u^{k-1}} + \mathbf{q^{k-1}})$$

Back propagation

For back propagated (adjoint) wavefields solve: $\mathbf{A}^T \mathbf{v} = \mathbf{P_r}^T \delta \mathbf{d}$

$$
\begin{pmatrix}
\mathbf{A_1}^T & \mathbf{A_2}^T & \mathbf{A_3}^T & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & \mathbf{A_1}^T & \mathbf{A_2}^T & \mathbf{A_3}^T \\
\vdots & & & 0 & \mathbf{A_1}^T & \mathbf{A_2}^T \\
0 & \cdots & \cdots & 0 & 0 & \mathbf{A_1}^T
\end{pmatrix}
\cdot
\begin{pmatrix}
\mathbf{v}^1 \\ \mathbf{v}^2 \\ \mathbf{v}^3 \\ \vdots \\ \vdots \\ \mathbf{v}^{nt}
\end{pmatrix}
=
\begin{pmatrix}
\mathbf{P}_{\mathbf{r_1}}^{\mathrm{T}} \delta \mathbf{d}^1 \\ \mathbf{P}_{\mathbf{r_1}}^{\mathrm{T}} \delta \mathbf{d}^2 \\ \mathbf{P}_{\mathbf{r_1}}^{\mathrm{T}} \delta \mathbf{d}^3 \\ \vdots \\ \vdots \\ \mathbf{P}_{\mathbf{r_1}}^{\mathrm{T}} \delta \mathbf{d}^{nt}
\end{pmatrix}
$$

Solve backwards in time $k = nt, ..., 1$

$$
\mathbf{v^{k-1}} = \mathbf{A_1}^{-T}(-\mathbf{A_2}^T \mathbf{v^k} - \mathbf{A_3}^T \mathbf{v^{k+1}} + \mathbf{P_{r_1}}^T \delta \mathbf{d}^{k-1})
$$

Gradient for (non-) linear inversion

How does a change in the model $\delta \mathbf{m}$ affect the modeled data?

$$\frac{\partial}{\partial \mathbf{m}}\Big(\mathcal{A}(\mathbf{m})\mathbf{u} = \mathbf{q}\Big) \iff \frac{\partial \mathcal{A}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{u} + \mathcal{A}(\mathbf{m})\frac{\partial \mathbf{u}}{\partial \mathbf{m}} = 0$$

➡ Jacobian and its adjoint

$$\mathbf{J} = \frac{\partial \mathbf{u}}{\partial \mathbf{m}} = -\mathcal{A}(\mathbf{m})^{-1}\Big(\frac{\partial \mathcal{A}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{u}\Big)$$

$$\mathbf{J}^T = -\Big(\frac{\partial \mathcal{A}(\mathbf{m})}{\partial \mathbf{m}}\mathbf{u}\Big)^T \mathcal{A}(\mathbf{m})^{-T}$$

(Haber, 2013)

Gradient for (non-) linear inversion

FWI least squares objective function

$$\Phi(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s} \left\|\mathbf{d}_i - \mathbf{P_r}\mathcal{A}(\mathbf{m})^{-\mathbf{1}}\mathbf{q}_i\right\|_2^2$$

$\mathbf{d_i}$: i_{th} observed shot record

$\mathbf{P_r}$: Restriction operator

$\mathbf{q_i}$: i_{th} source

The gradient of the FWI objective is given by

$$\mathbf{J}^T\delta\mathbf{d} = -\sum_{i=1}^{n_s}\left(\frac{\partial\mathcal{A}(\mathbf{m})}{\partial\mathbf{m}}\mathbf{u}_i\right)^T \mathcal{A}(\mathbf{m})^{-T}\mathbf{P_r}^T\delta\mathbf{d}_i$$

Temporal derivative of wave field

Solution of adjoint wave equation

Gradient for (non-) linear inversion

Calculation of FWI gradient (RTM image)

- Model the wavefields $\mathbf{u}$ for all time steps $k = 1, ..., nt$
- Model the adjoint wavefields $\mathbf{v}$ in reverse order $k = nt, ..., 1$ and at each time step multiply

$$\mathbf{v^{k-1}} \text{ and } \mathbf{v^{k+1}} \text{ with } \frac{1}{\Delta t^2} \quad \text{and} \quad \mathbf{v^{k}} \text{ with } \frac{-2}{\Delta t^2}$$

(corresponds to applying a time derivative operator $\mathbf{D}$)

- Correlate wavefields and update gradient

$$g = g - (\mathbf{u^{k-1}})^T \mathbf{diag}(\mathbf{D v^{k-1}})$$

- Repeat for all sources and sum

SLIM

Extension to anisotropy

2D anisotropic wave equation with PS methods

$$
\begin{aligned}
m\frac{\partial^2 U(k_x,k_z,t)}{\partial t^2} = \Big[k_x^2 + k_z^2 + &(2\delta\sin^2\theta\cos^2\theta + 2\epsilon\cos^4\theta)\frac{k_x^4}{k_x^2+k_z^2} + \\
&(2\delta\sin^2\theta\cos^2\theta + 2\epsilon\sin^4\theta)\frac{k_z^4}{k_x^2+k_z^2} + \\
&(-\delta\sin^2 2\theta + 3\epsilon\sin^2 2\theta + 2\delta\cos^2 2\theta)\frac{k_x^2k_z^2}{k_x^2+k_z^2} + \\
&(\delta\sin 4\theta - 4\epsilon\sin 2\theta\cos^2\theta)\frac{k_x^3k_z}{k_x^2+k_z^2} + \\
&(-\delta\sin 4\theta - 4\epsilon\sin 2\theta\sin^2\theta)\frac{k_xk_z^3}{k_x^2+k_z^2}\Big]U(k_x,k_z,t)
\end{aligned}
$$

k_x: spatial wavenumber in x-direction

ϵ, δ, θ: Thomsen parameters, tilt angle

$U(k_x,k_z,t)$: wavefield in t-k domain

(Zhang et al., 2005)

Extension to anisotropy

Discretize with 2nd order leap frog scheme in time and rewrite as

$$\mathbf{diag}(\frac{\mathbf{m}}{\Delta t^2})\left(\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1}\right) - \mathbf{L}\mathbf{u}^n = \mathbf{q}^{n+1}$$

➡ exactly the same scheme as for isotropic modeling, only that

$$\mathbf{L} = \mathrm{real}\left(\mathbf{F}^*\mathrm{diag}(\mathbf{k}_{ani})\mathbf{F}\right)$$

$\mathbf{F}$: 2D Fourier matrix,

$\mathbf{k}_{ani}$: wavenumbers + anisotropy parameters

Extension to anisotropy

Extension to anisotropic modeling:

- Just Laplacian changes ➡ only change $\mathbf{A_2}$ in existing workflow
- Backpropagation is done in the same way as before

Jacobian and adjoint for anisotropy:

- Only take partial derivative with respect to squared slowness $\mathbf{m}$
- $\mathbf{m}$ does not appear in Laplacian ➡ expressions for $\mathbf{J}, \mathbf{J}^T$ do not change either!
- $\mathbf{J}$ passes adjoint tests with and without anisotropy

$$||\delta\mathbf{d}^T\mathbf{J}\delta\mathbf{m} - \delta\mathbf{d}^T\mathbf{J}^T\delta\mathbf{d}|| < \epsilon$$

Extension to anisotropy

One code for both isotropic and anisotropic modeling and inversion

- both modes use the same functions and operators
- only Laplacian changes for anisotropy
- easy to use via function overloading

$$\mathtt{data} = \mathtt{Gen\_data}(\mathtt{m0}, \mathtt{params}, \mathtt{q})$$

isotropic modeling with FD Laplacian

$$\mathtt{data} = \mathtt{Gen\_data}(\mathtt{m0}, \mathtt{params}, \mathtt{q}, \mathtt{ani})$$

anisotropic modeling with PS Laplacian

Extension to anisotropy

Function to automatically determine optimal grid spacing and time step

$$[\texttt{m0}, \texttt{params}] = \texttt{Setup\_CFL}(\texttt{m0}, \texttt{params}) \Rightarrow \Delta t = \frac{h}{\sqrt{2} v_{max}}$$

$$[\texttt{m0}, \texttt{params}, \texttt{ani}] = \texttt{Setup\_CFL}(\texttt{m0}, \texttt{params}, \texttt{ani}) \Rightarrow \Delta t = \frac{2h}{\sqrt{2} \pi v_{max}}$$

Higher time step restrictions for PS method, but less grid points per wave length possible (G=5 instead of 10)

$$h = \frac{v_{min}}{G \cdot f}$$

Synthetic example: FWI on BG model

Generate data with anisotropic modeling

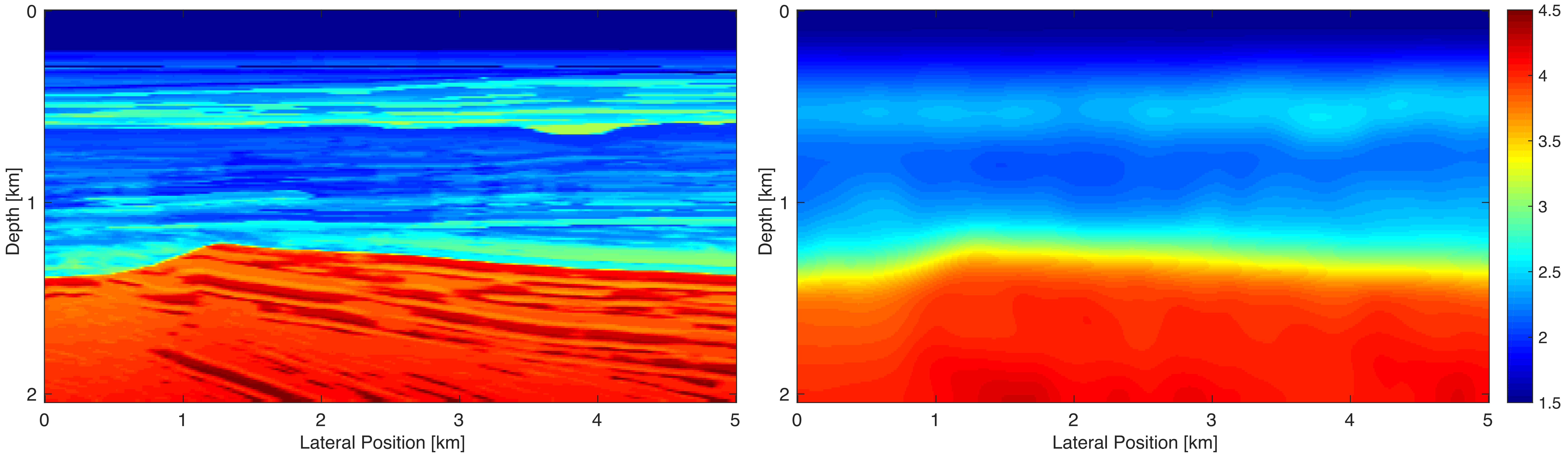
- Invert with isotropic and anisotropic modeling kernel
- Influence of anisotropy on data misfit, final velocity model?

Setup:

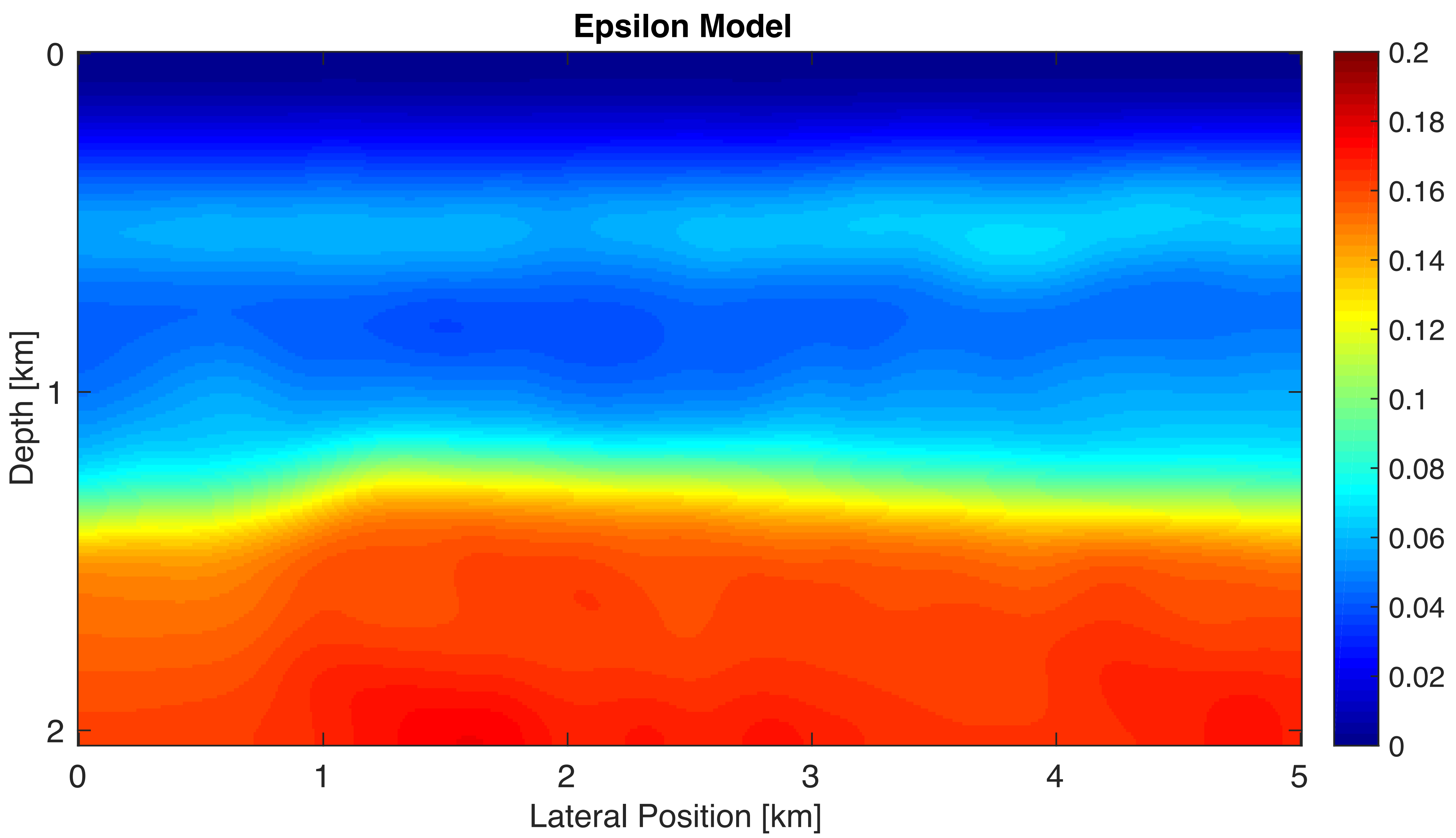
- 501 receivers (10 m apart), 99 sources (50 m apart)
- 2.4 s recording time (601 samples)
- Data generated with 15 Hz peak frequency

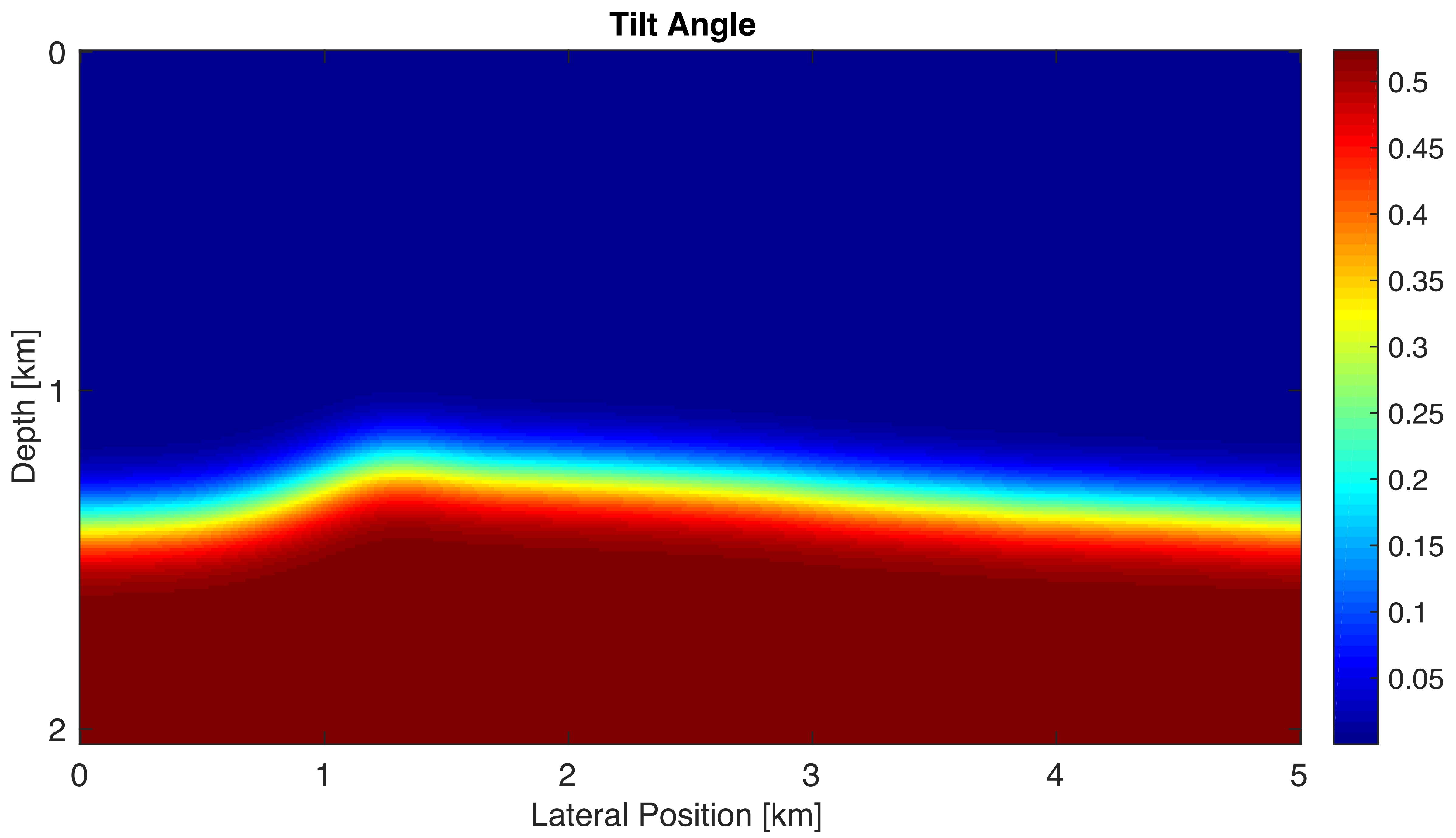
FWI BG model

True velocity model and initial model

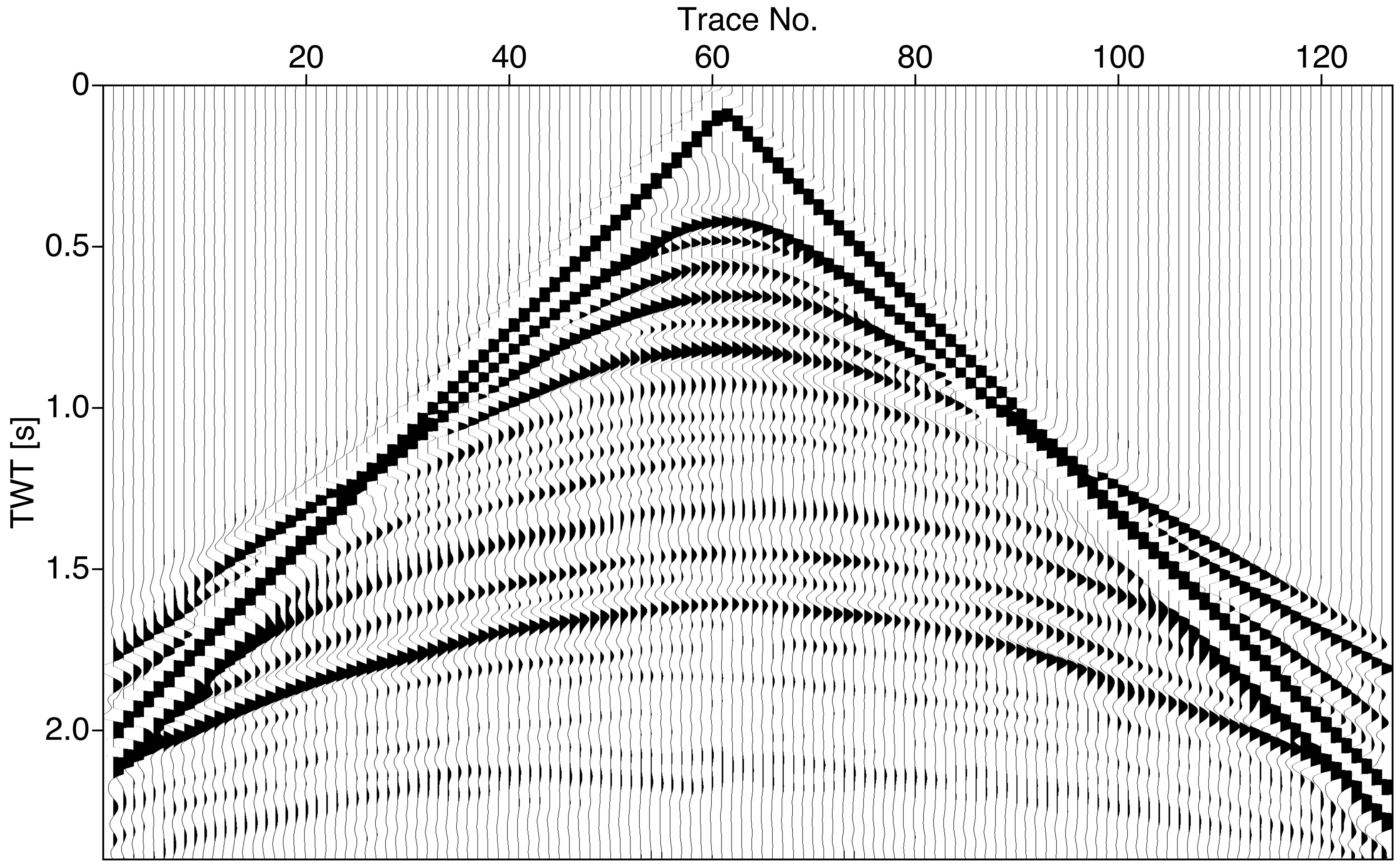


FWI BG model


Epsilon Model
Depth [km]
Lateral Position [km]
0
1
2
0
1
2
3
4
5
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0

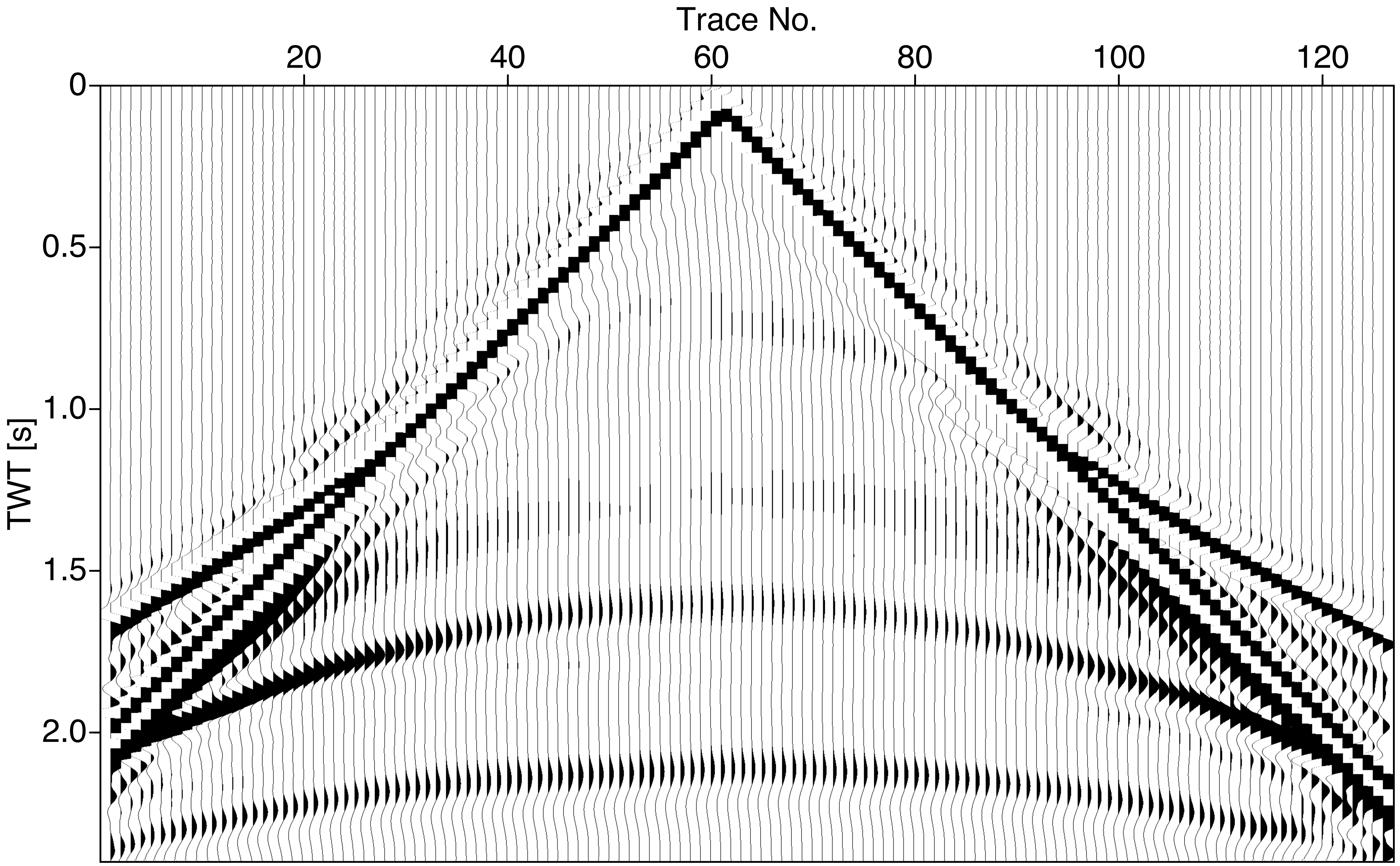

Tilt Angle
Depth [km]
Lateral Position [km]
0
1
2
0
1
2
3
4
5
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05

FWI BG model



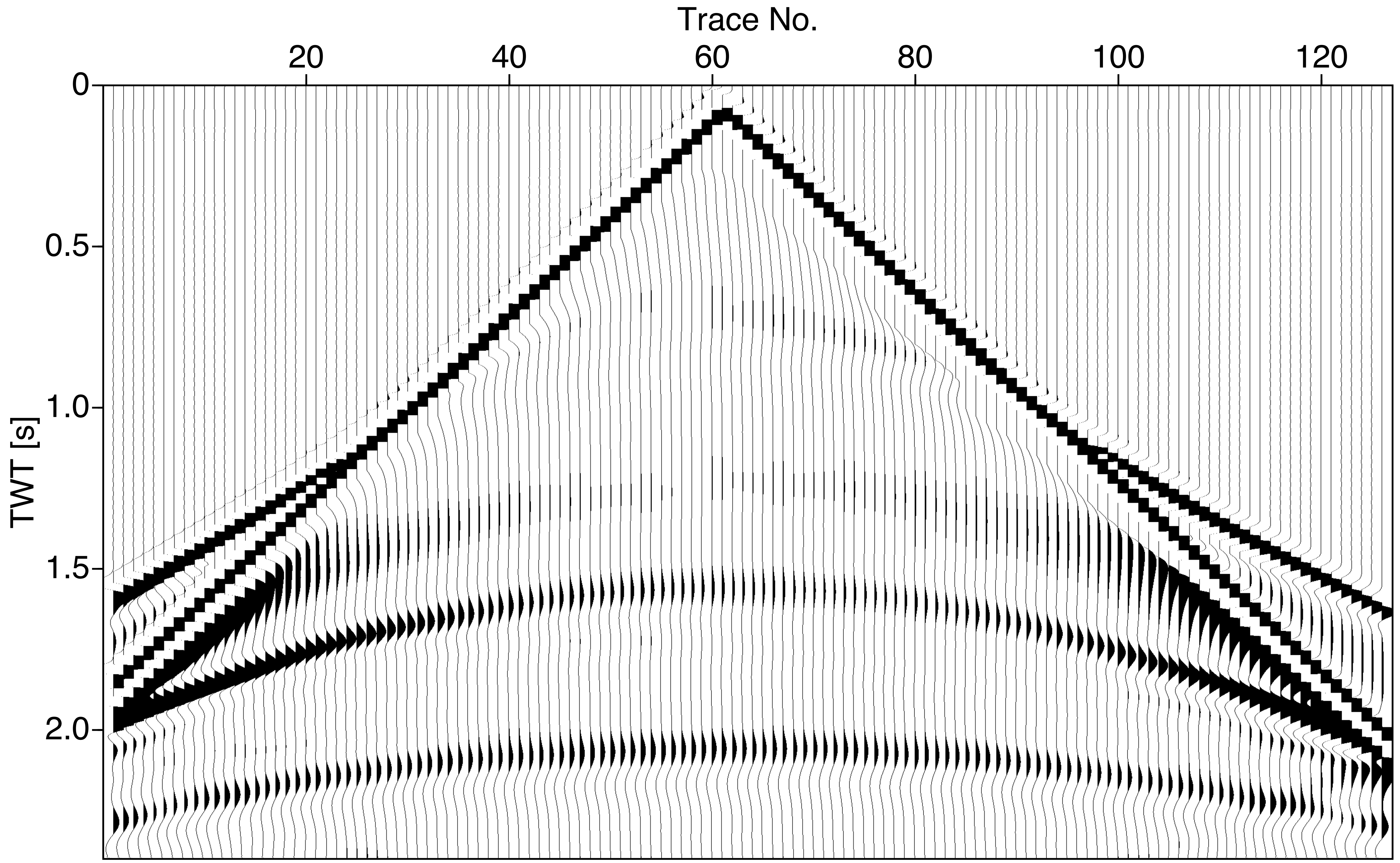
Observed data

FWI BG model



Initial data with anisotropy

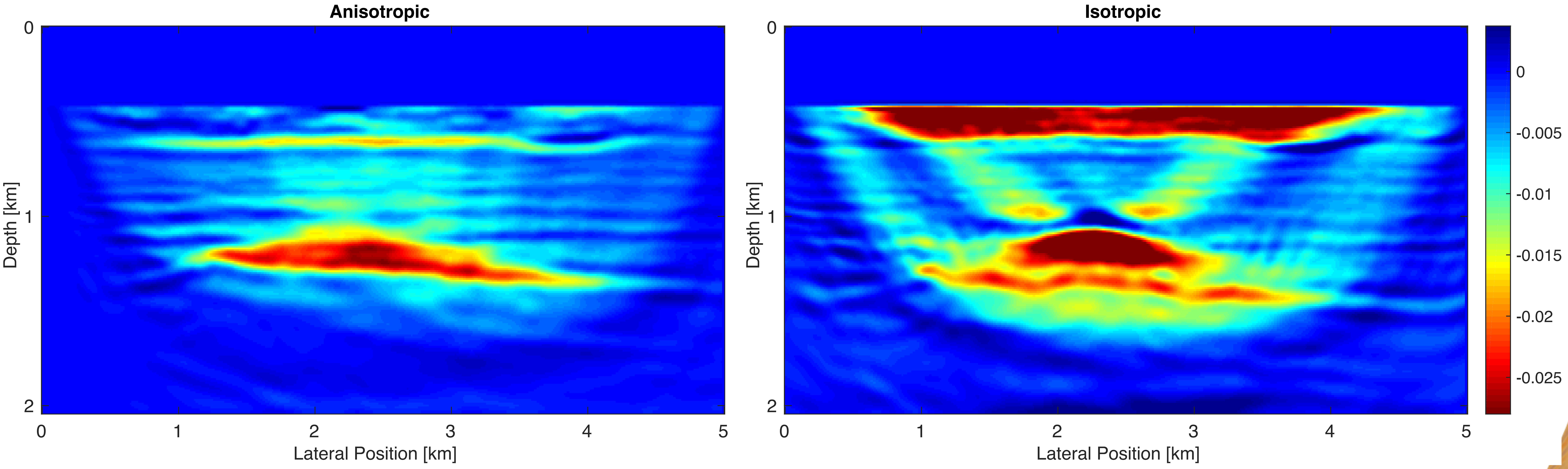
FWI BG model



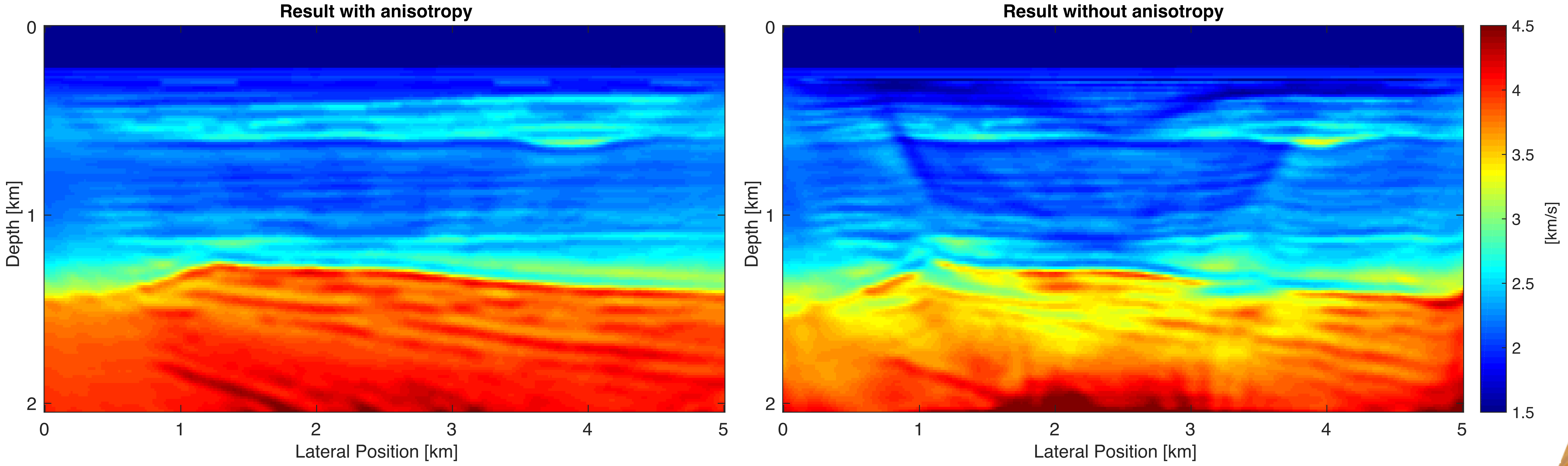
Initial data without anisotropy

FWI BG model

First gradients

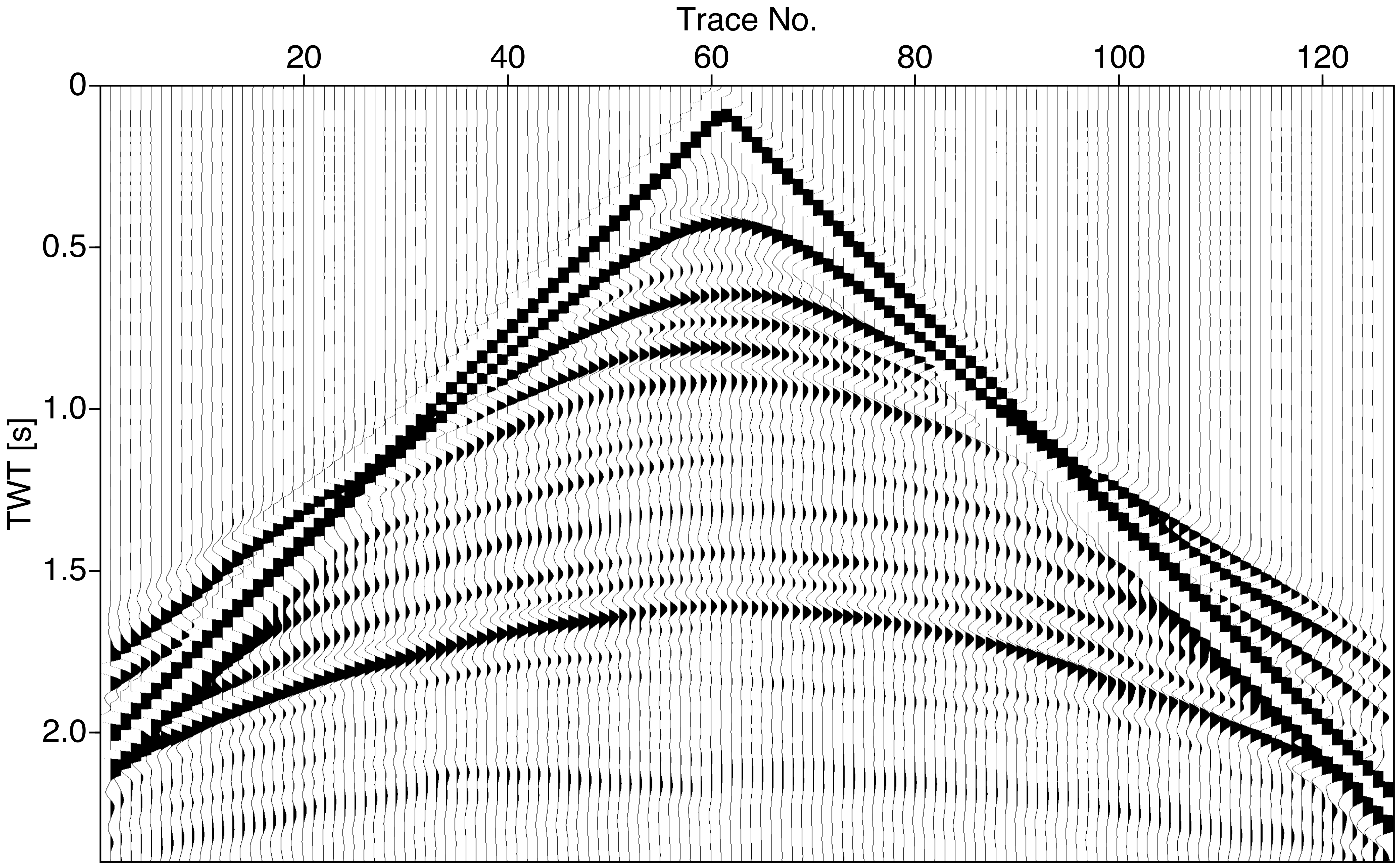


FWI BG model



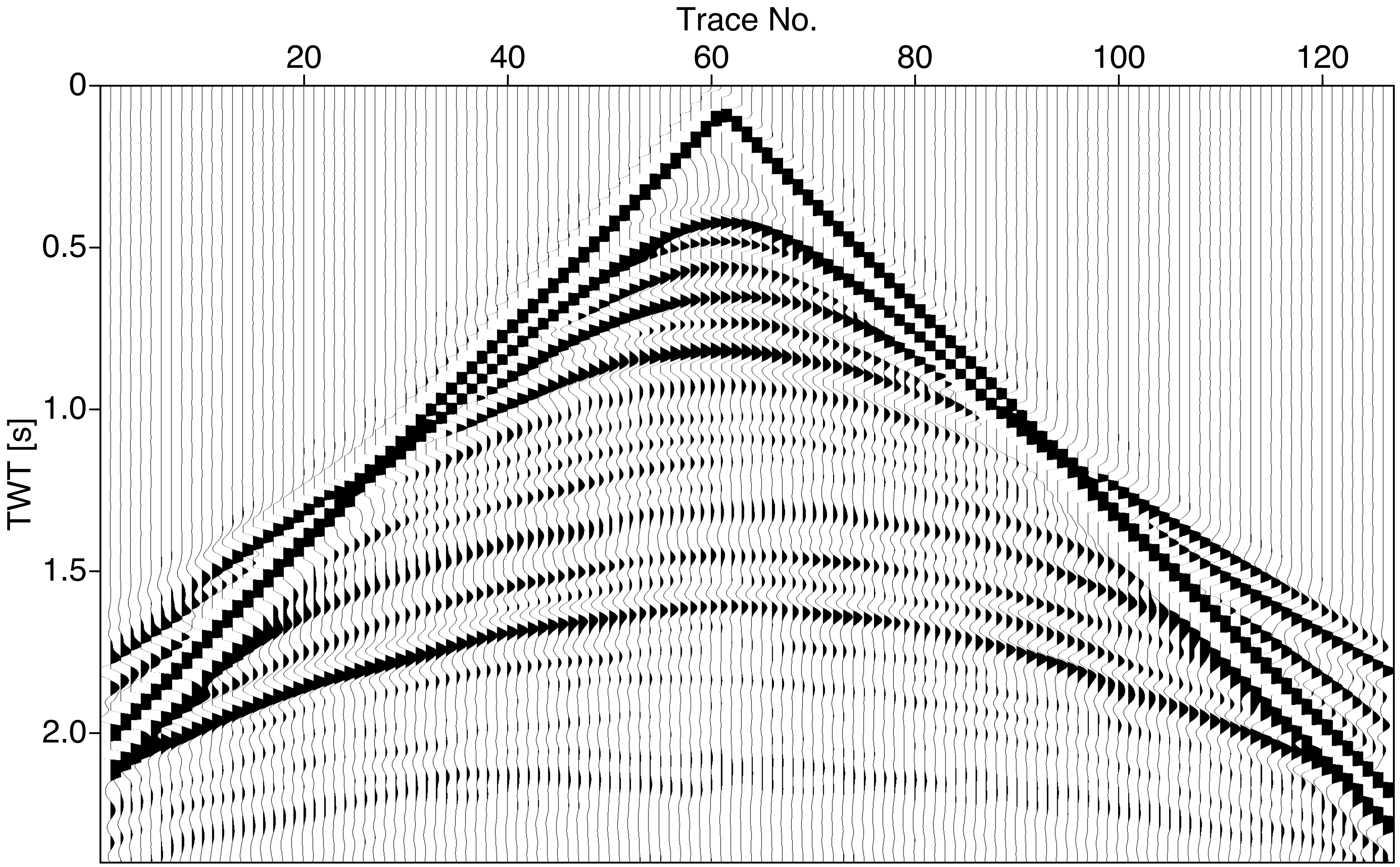
Result with anisotropy
Result without anisotropy
Depth [km]
0
1
2
Lateral Position [km]
0
1
2
3
4
5
4.5
4
3.5
3
2.5
2
1.5
[km/s]

FWI BG model



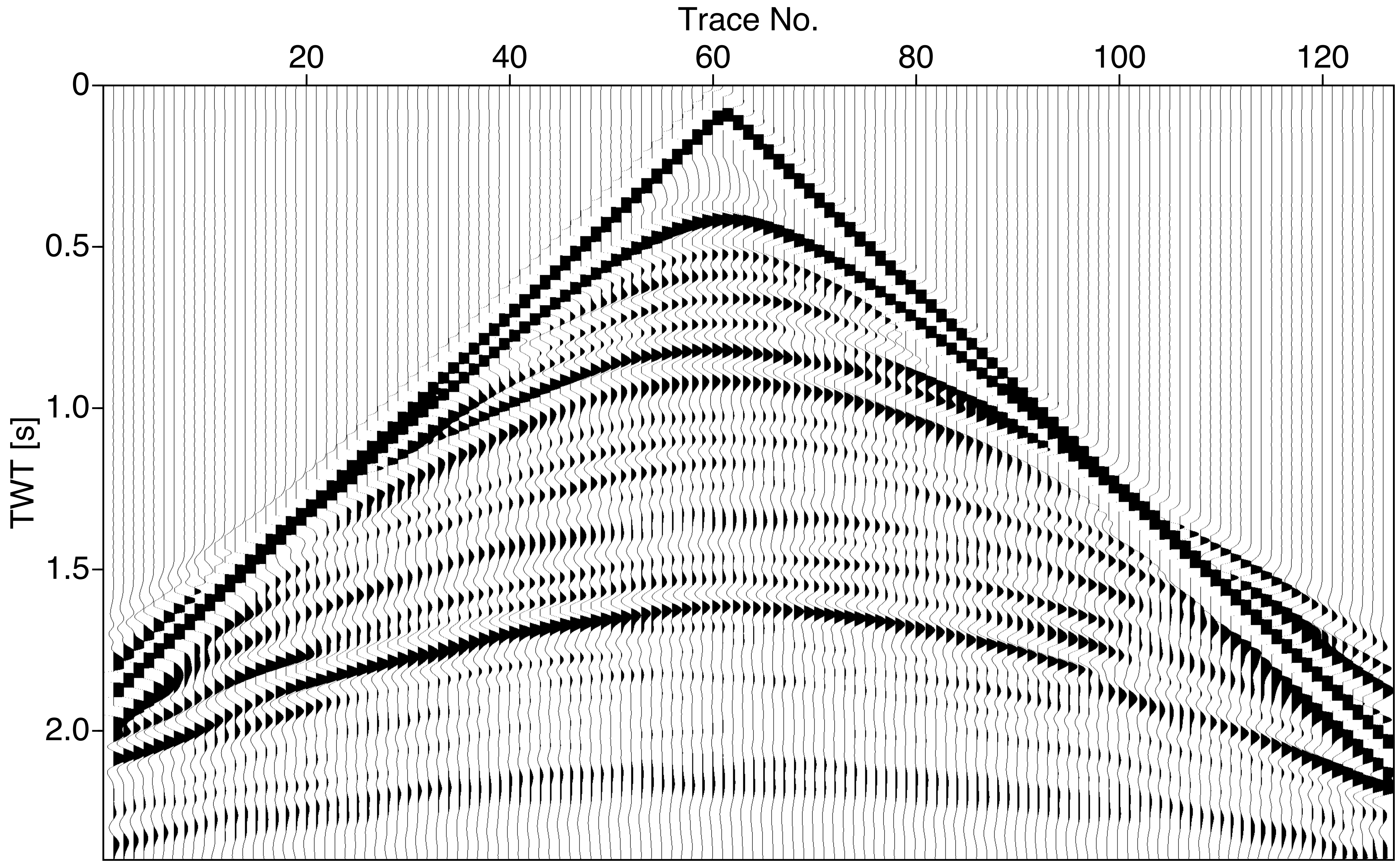
Final data with anisotropy

FWI BG model



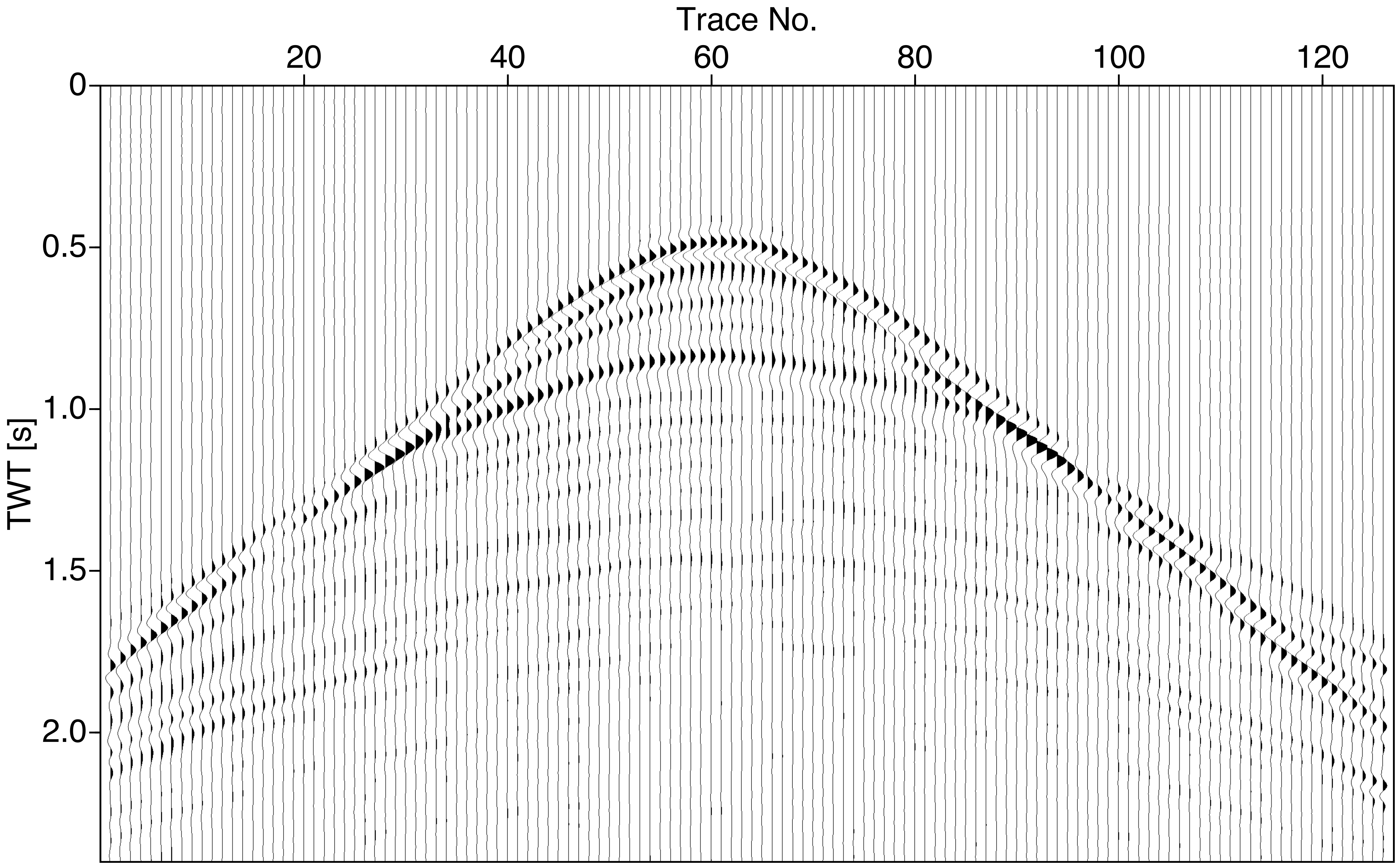
Observed data

FWI BG model



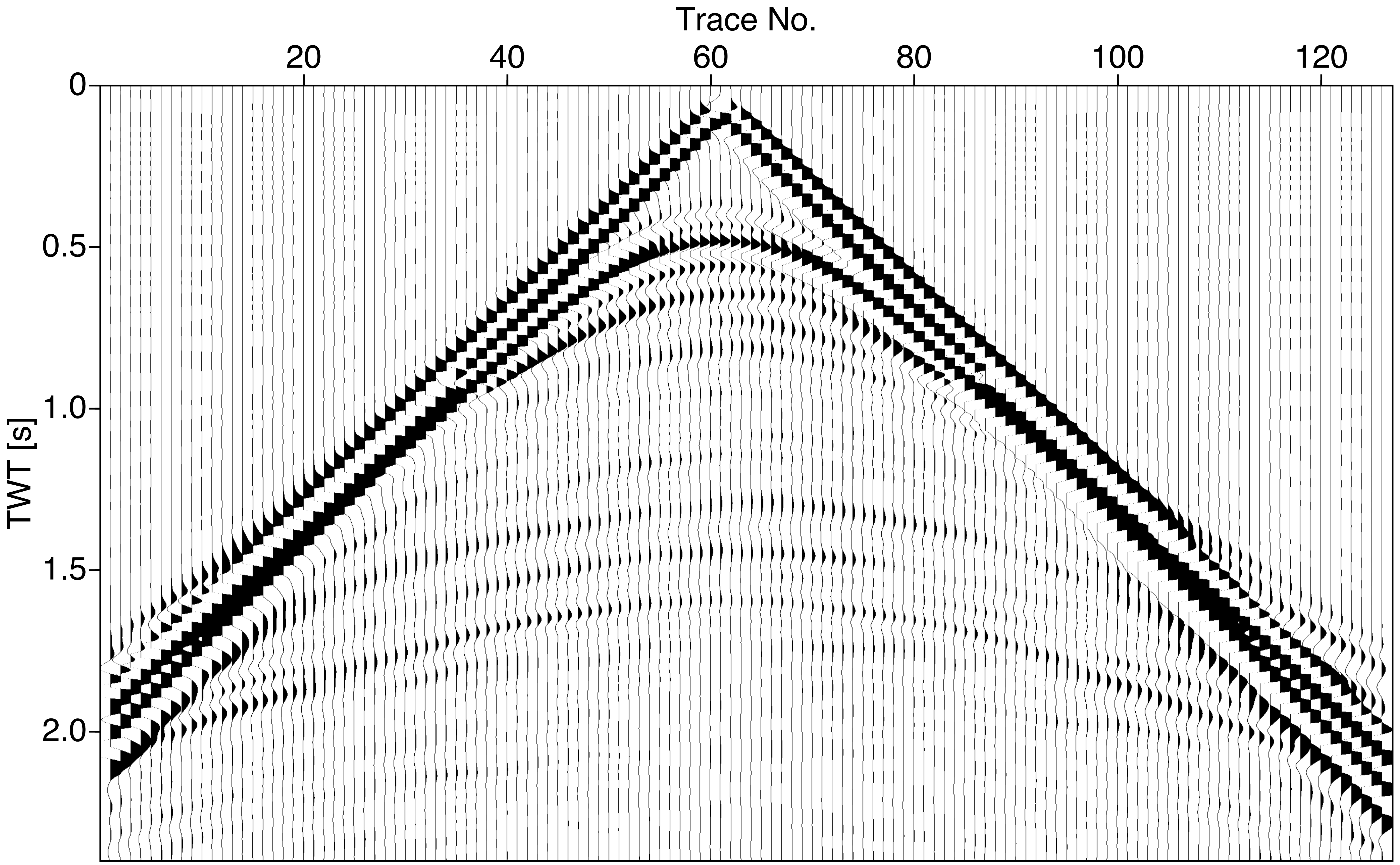
Final data without anisotropy

FWI BG model



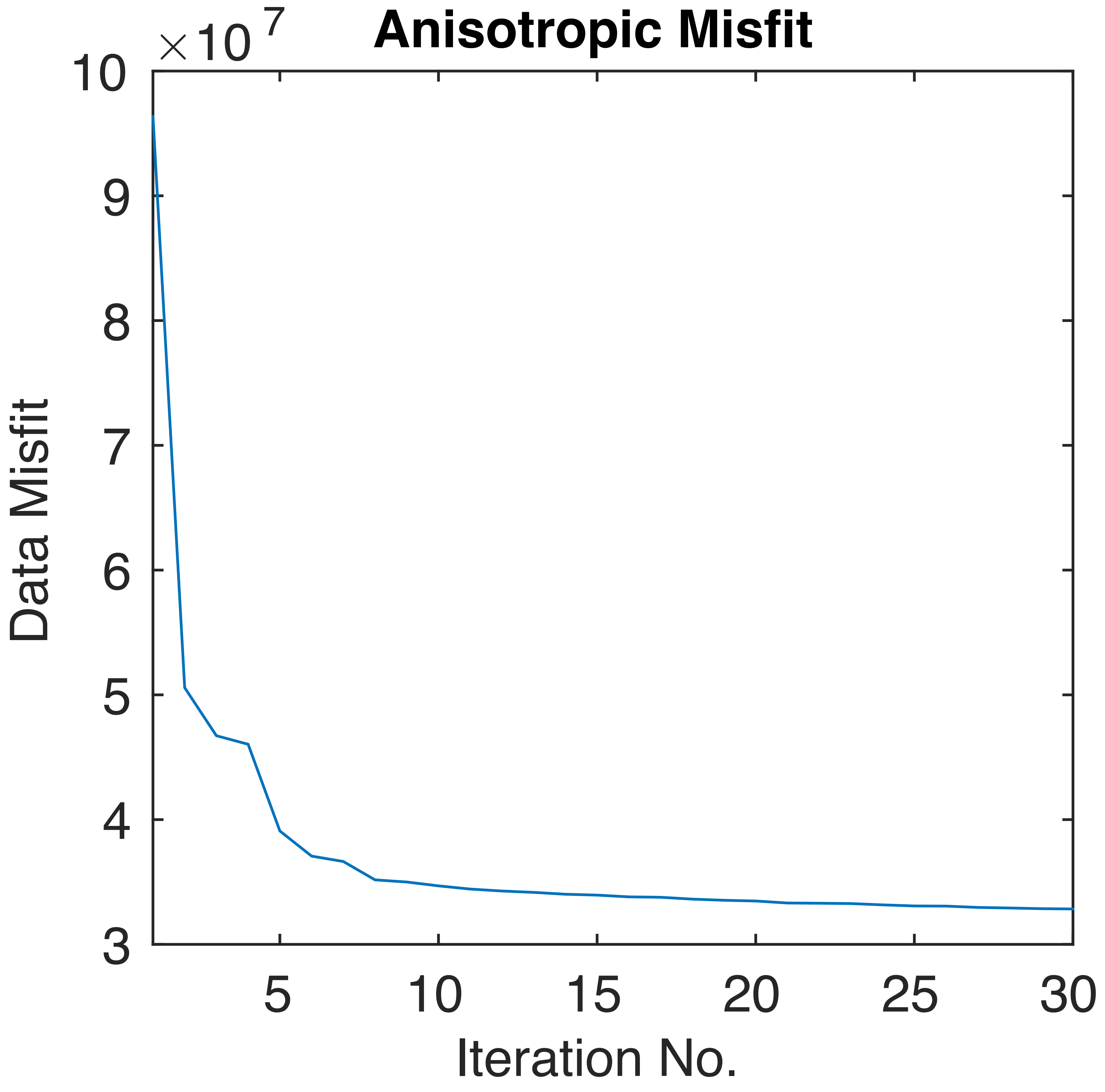
Difference with anisotropy

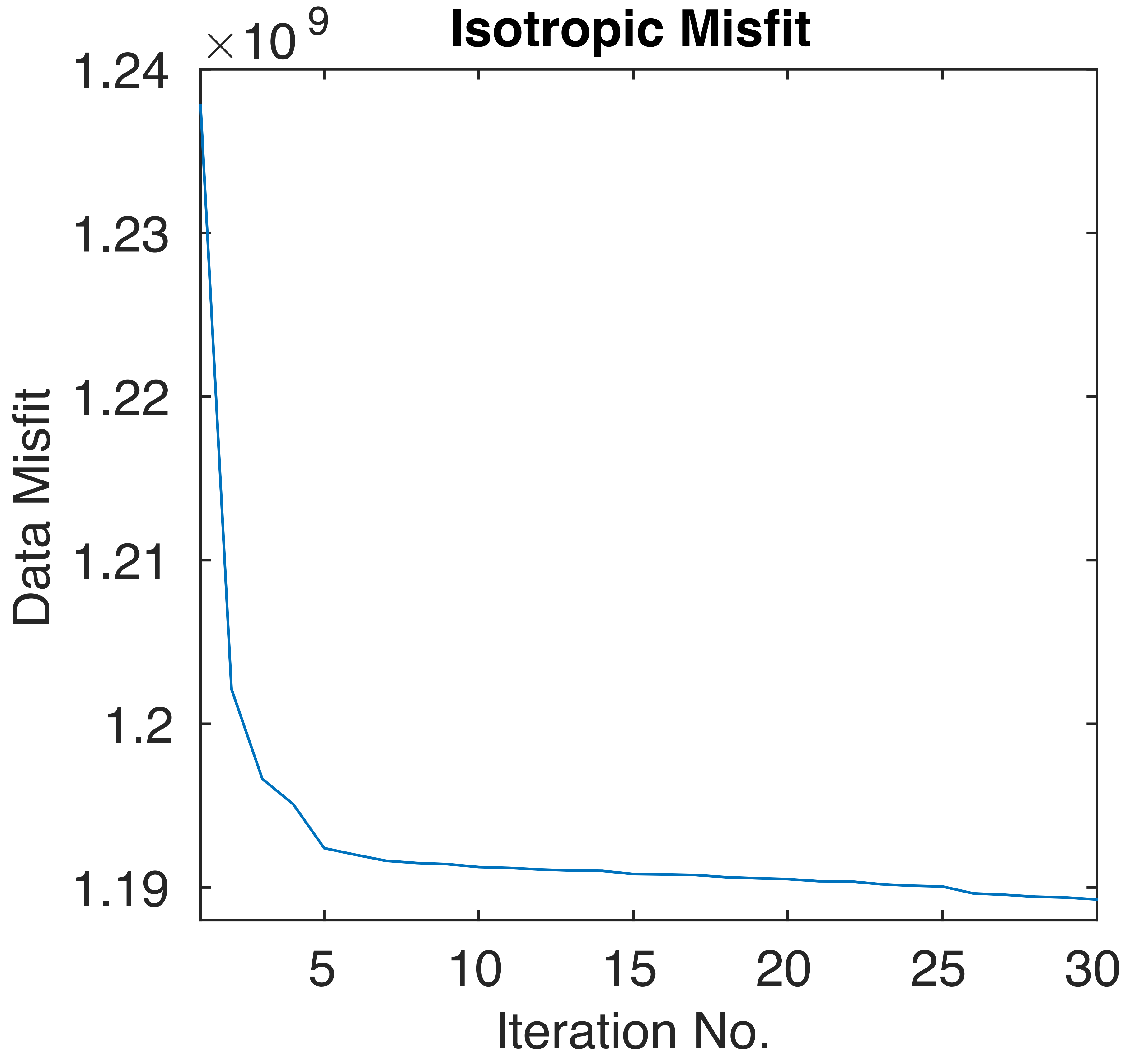
FWI BG model



Difference with anisotropy

FWI BG model


Anisotropic Misfit
×10^7
10
9
8
7
6
5
4
3
Data Misfit
5
10
15
20
25
30
Iteration No.


Isotropic Misfit
×10^9
1.24
1.23
1.22
1.21
1.2
1.19
Data Misfit
5
10
15
20
25
30
Iteration No.

RTM field data example

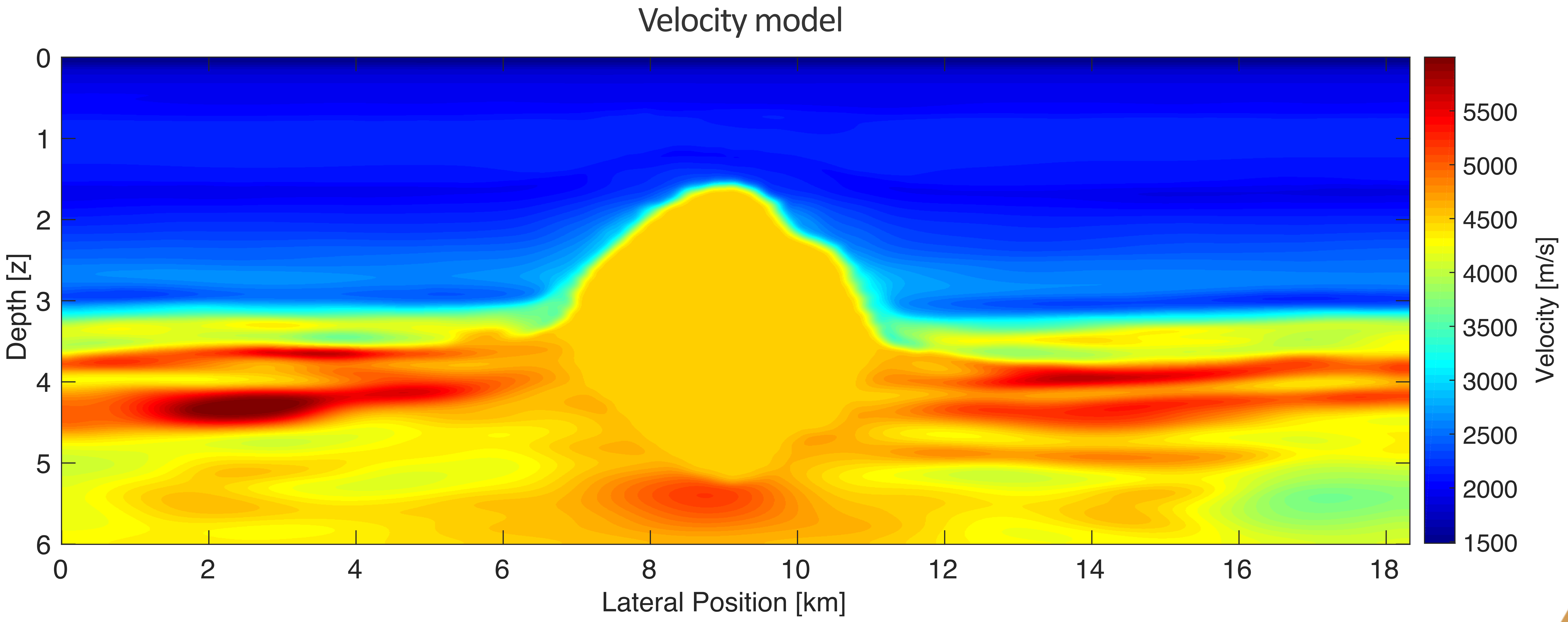
Reverse time migration of BP Machar data set

- Marine data set from North Sea
- 18.3 km x 6 km Model
- 330 shot records with 8 seconds recording time
- Streamer with 1080 receivers

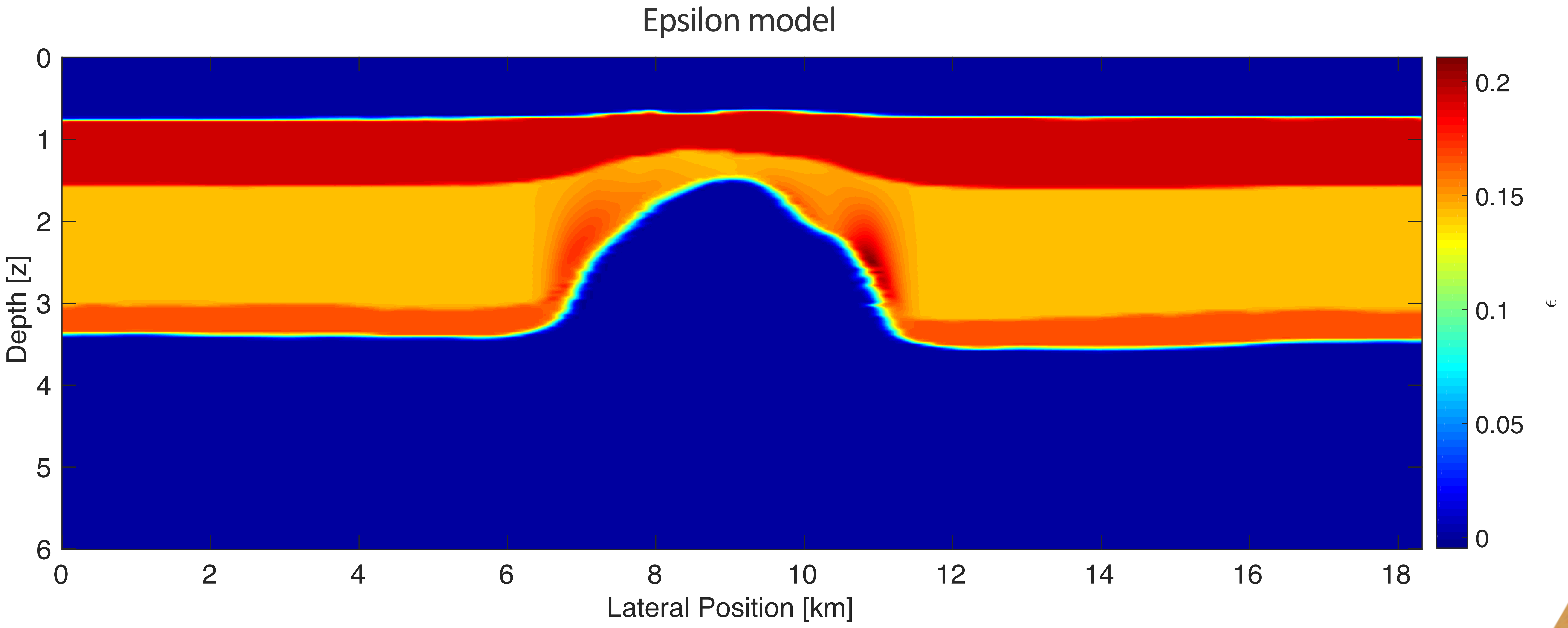
Processing parameters

- Direct wave removed
- 10 m grid spacing (602 x 1832 grid points)
- Migrate with ~30 Hz peak frequency
- Velocity model and anisotropy parameters provided by BP

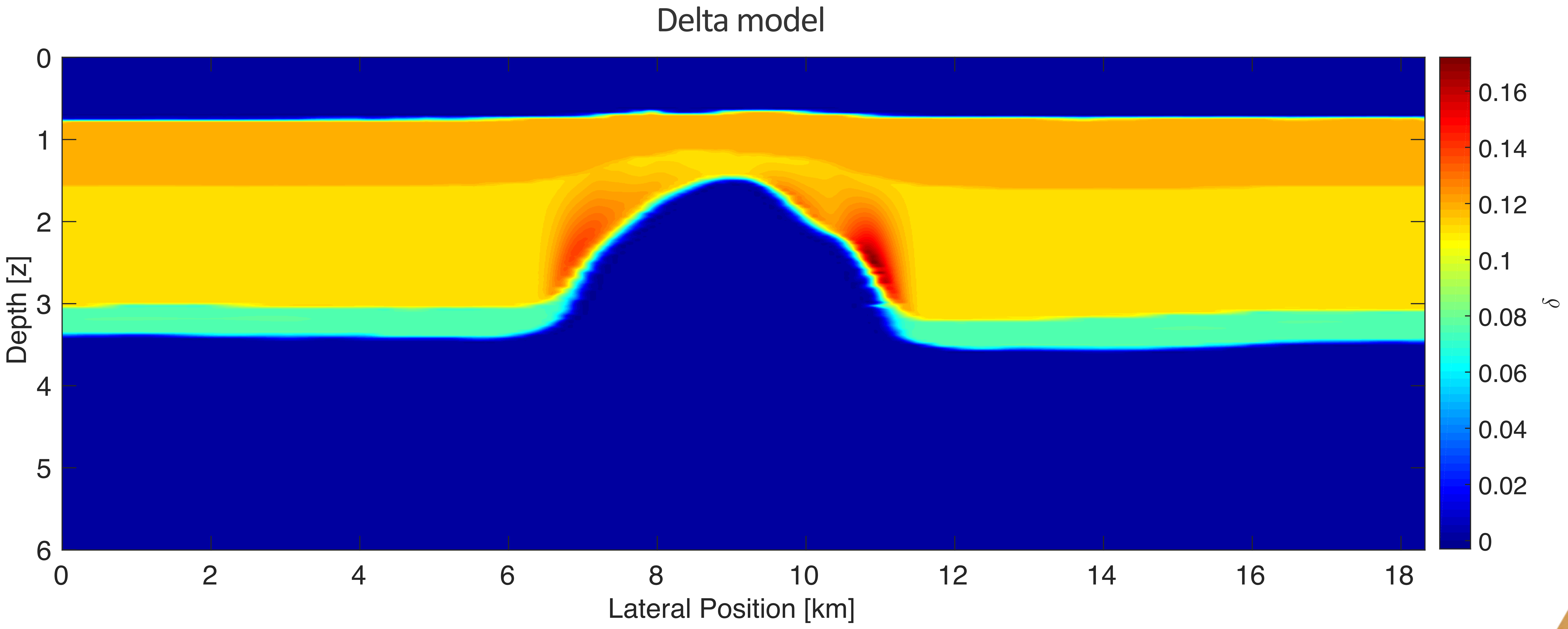
RTM field data example



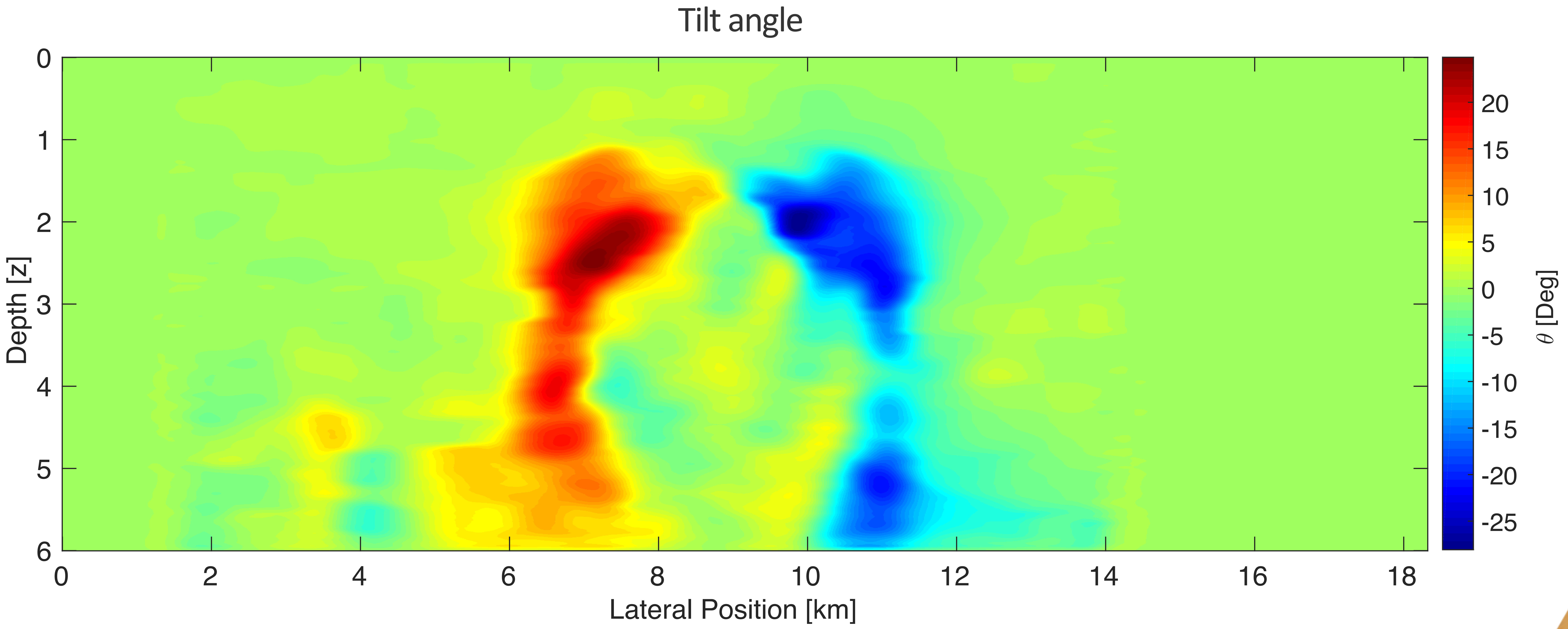
RTM field data example



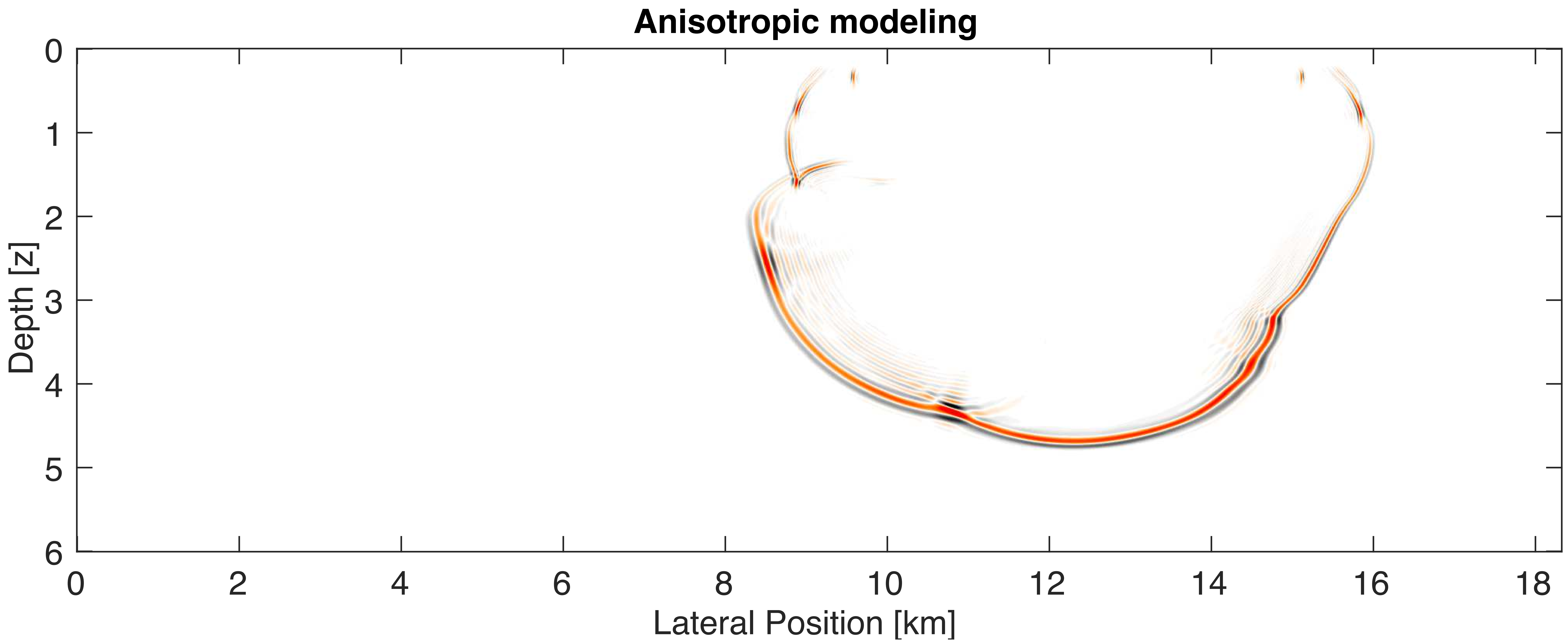
RTM field data example



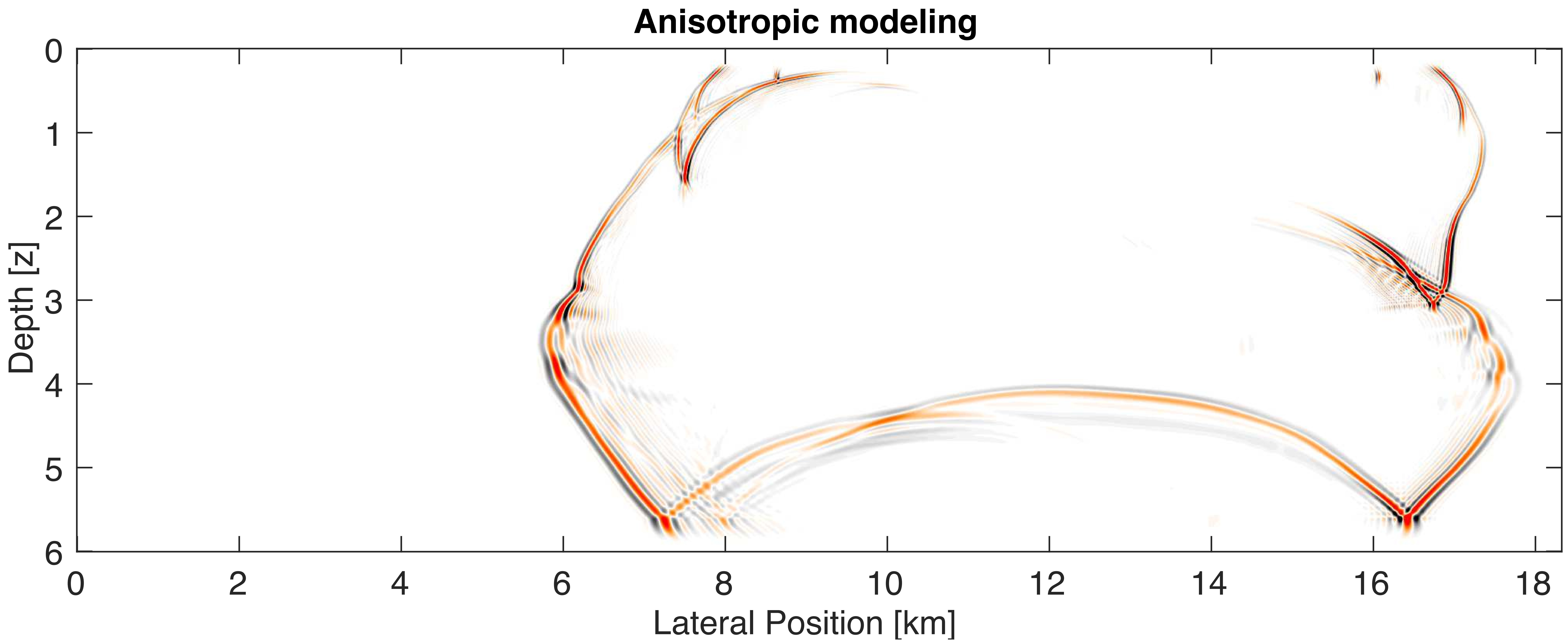
RTM field data example



RTM field data example



RTM field data example



RTM field data example

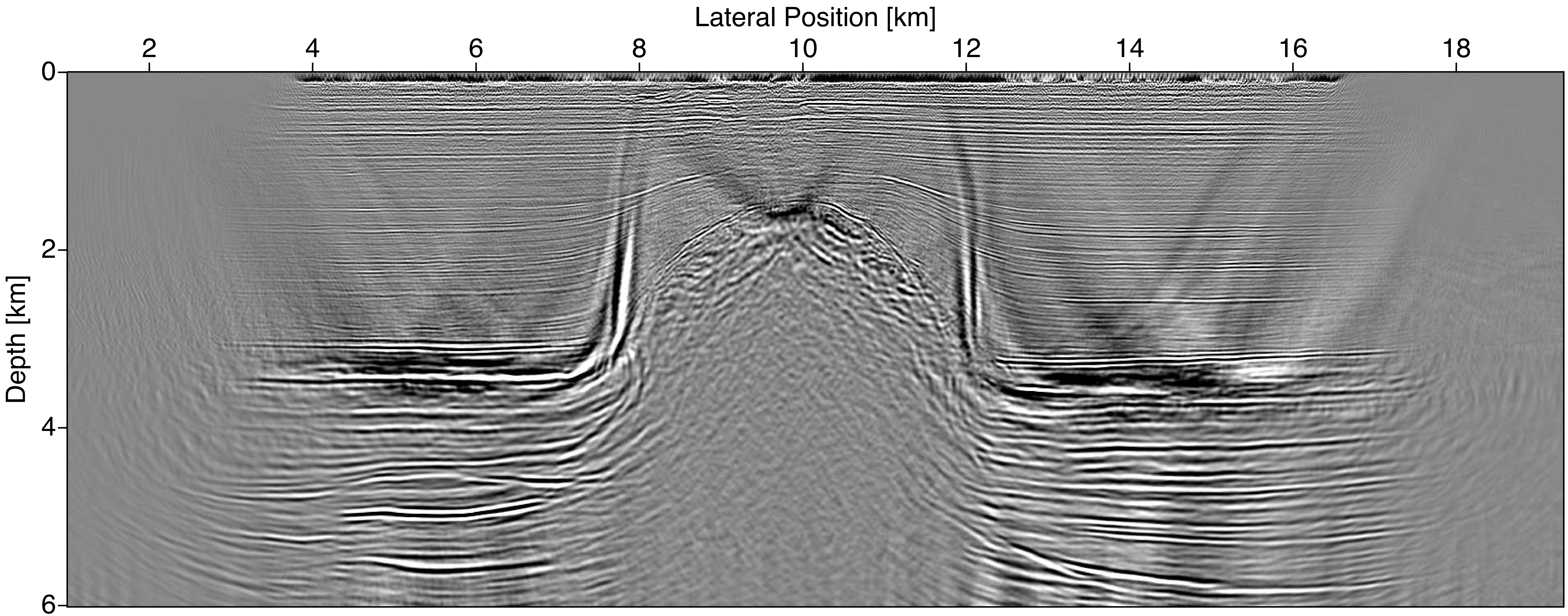


Image with depth scaling only

RTM field data example

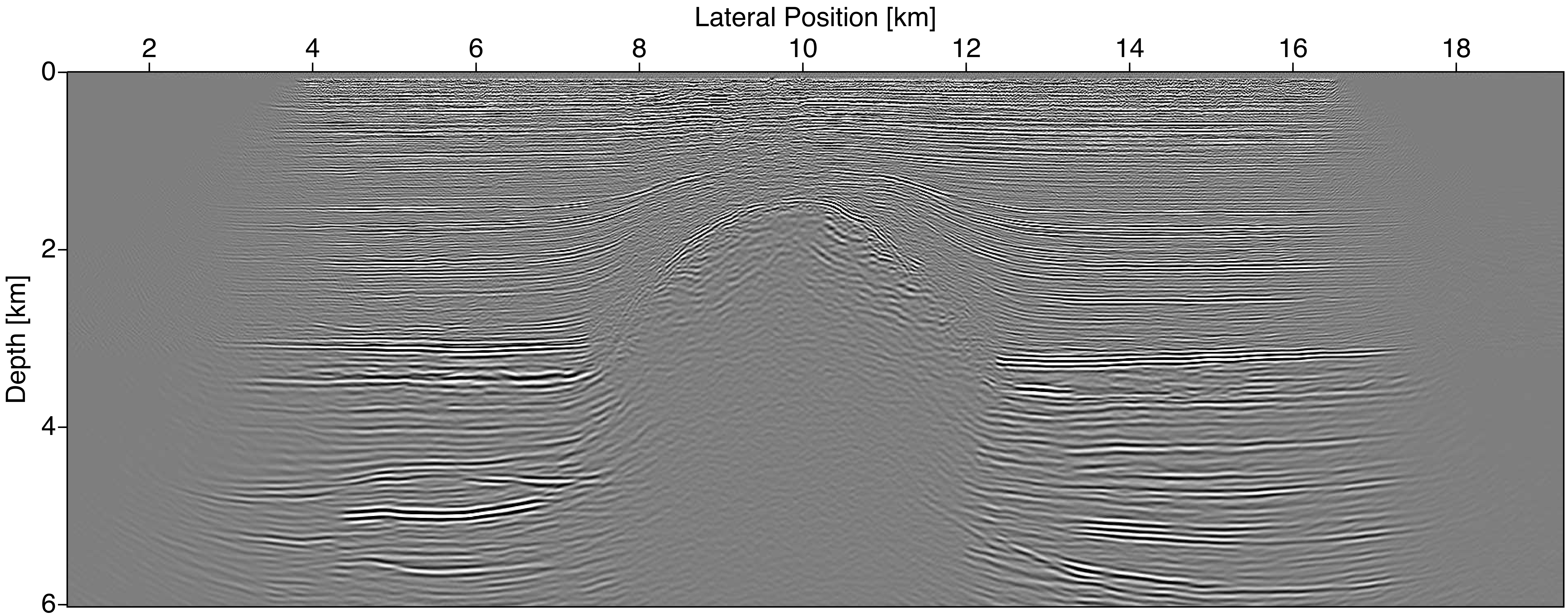


Image with depth scaling and horizontal derivative

Alternative derivative matrices

Main drawback of current implementation:

- Each time steps requires n_x x n_z forward + inverse FFTs
- becomes even more problematic for 3D

Alternatives for highly accurate derivative matrices (time-domain)?

- EPS method (Sandberg et al., 2011)
- Low-rank FD and low-rank FFD (Song et al., 2013)
- ?

EPS method

Eigen-decomposition Pseudo-Spectral (EPS) method (Sandberg et al., 2011)

- Derivative operator as integral operator

$$Lf(x) = \int_a^b K(x,y)f(y)dy$$

- where $K(x,y)$ is a Kernel function of the form

$$K(x,y) = \sum_{m=1}^{\infty} \lambda_m u_m(x) v_m(y)$$

$\{u_m\}_{m=1}^{\infty}, \{v_m\}_{m=1}^{\infty}$: Set of orthonormal functions

EPS method

Numerical representation by

- truncating the sum after N_c terms
- replacing the integral by quadratures with nodes $\{\theta_k\}_{k=1}^{N}$ and weights $\{w_k\}_{k=1}^{N}$ (Gaussian quadratures or PSWFs)

$$Lf(\theta_k) = \sum_{m=1}^{N_c} \lambda_m u_m(\theta_k) \sum_{l=1}^{N} w_l v_m(\theta_l) f(\theta_l) \qquad k, l = 1, \ldots, N$$

With discretized operator matrix

$$L_{kl} = \sum_{m=1}^{N_c} \lambda_m u_m(\theta_k) w_l v_m(\theta_l)$$

Rank completion

- Operator has rank N_c where $N_c < N$
- Truncation controls the operator norm, but:
- Truncated functions absorb high frequency oscillations
- Add "tail" to operator such that L_{kl} has full rank

$$L_{kl} = \sum_{m=1}^{N_c} \lambda_m u_m(\theta_k) w_l v_m(\theta_l) + \sum_{m=N_c+1}^{N} \lambda_m \tilde{u}_m(\theta_k) w_l \tilde{v}_m(\theta_l)$$

where $\{\tilde{u}_m(\theta_l)\}_{m=N_c+1}^{N}$ are orthogonalized vectors

Second derivative operator

Second derivative operator on interval [-1,1] has SVD

$$\left\{\frac{m\pi}{2}, \sin\left(\frac{m\pi}{2}(x+1)\right)\right\}_{m=1}^{N}$$

which leads to Kernel function

$$K(x,y) = -\frac{\pi^2}{4}\sum_{m=1}^{N} m^2 \sin\left(\frac{m\pi}{2}(x+1)\right)\sin\left(\frac{m\pi}{2}(y+1)\right)$$

Second derivative operator

Second derivative operator with Dirichlet boundary conditions

$$L_{kl} = -\frac{\pi^2}{4}\left(\sum_{m=1}^{N_c} m^2 \sin\left(\frac{m\pi}{2}(\theta_k+1)\right) w_l \sin\left(\frac{m\pi}{2}(\theta_l+1)\right) + \sum_{m=N_c+1}^{N} m\tilde{u}_m(\theta_k) w_l \tilde{v}_m(\theta_l)\right)$$

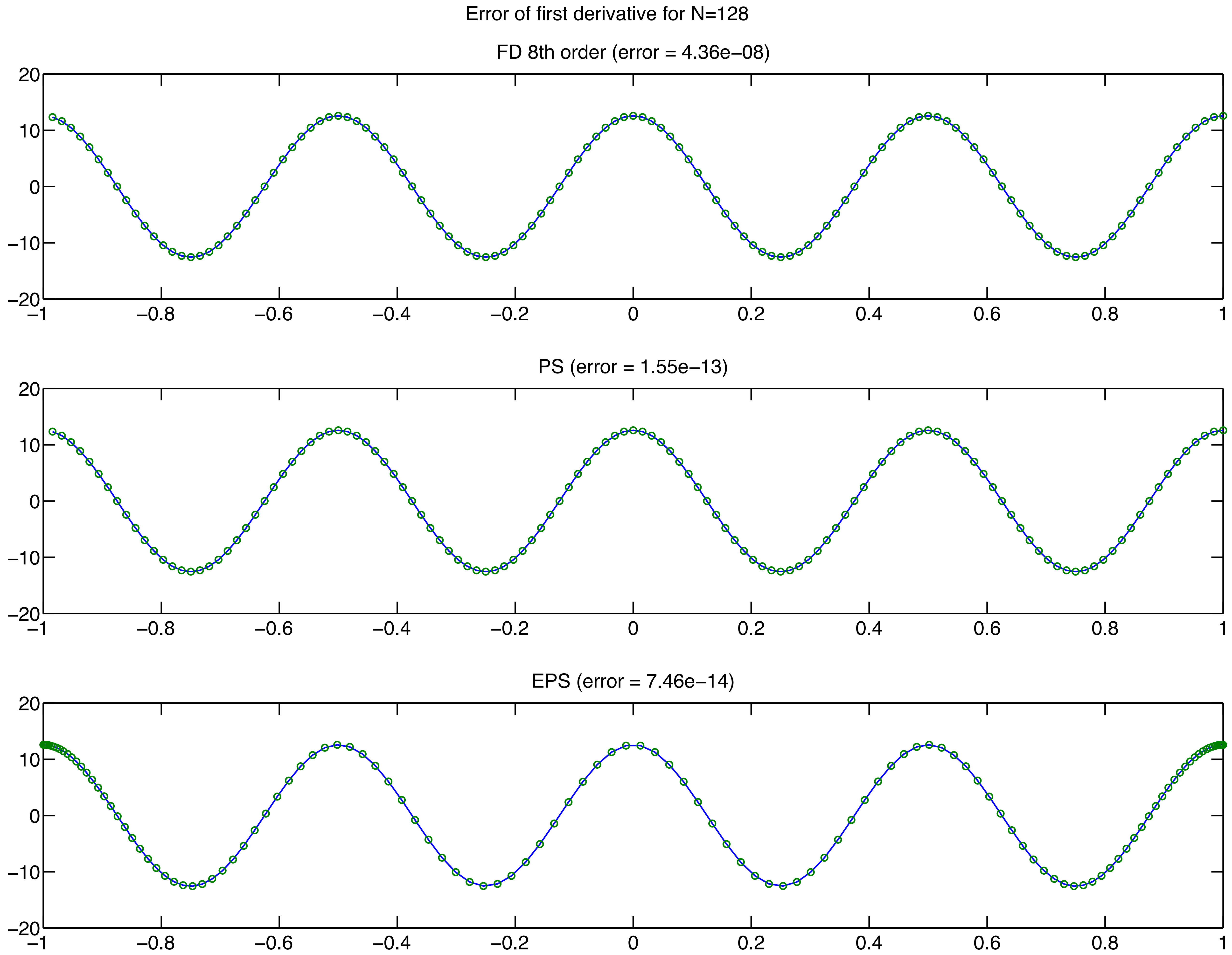
In practice choose N_c such that for functions with bandlimit $2c$

$$\frac{N_c\pi}{2} \leq c$$

Accuracy (1st derivative)

Derivative of

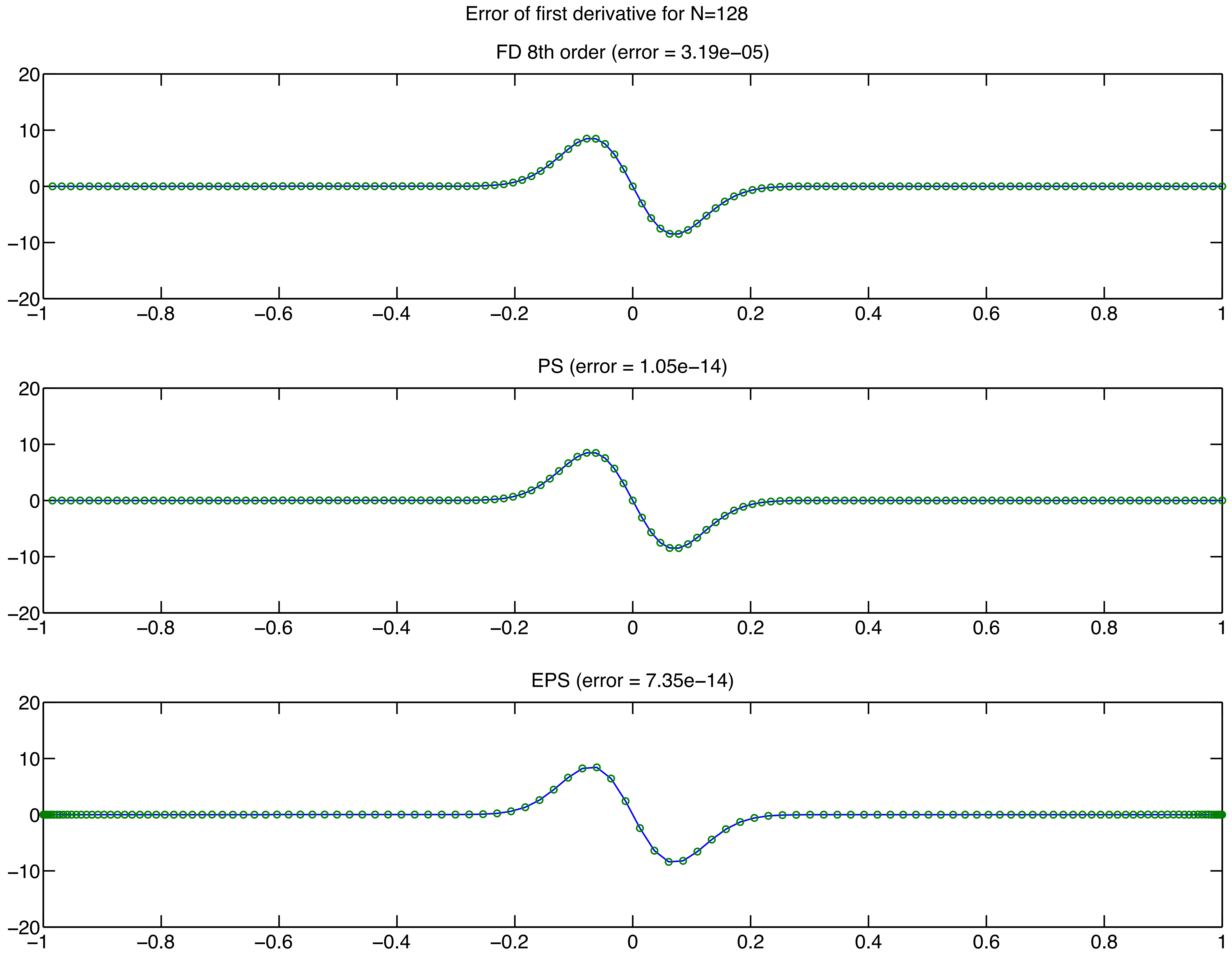
$$f(x) = \sin\left(4\pi(x+1)\right)$$



Accuracy (1st derivative)

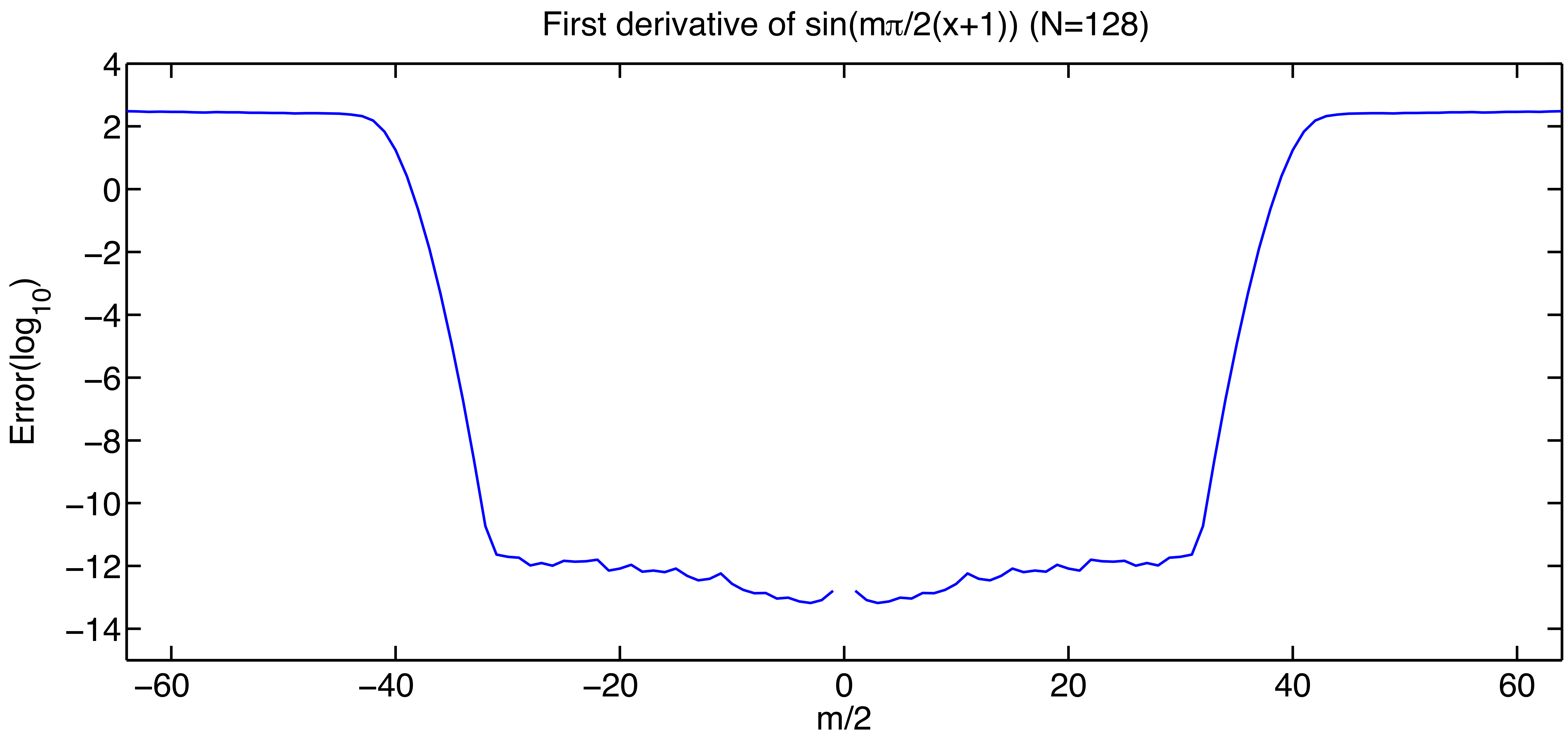
Derivative of

$$f(x) = e^{-100x^2}$$



Accuracy (1st derivative)

Accuracy as function of band limit



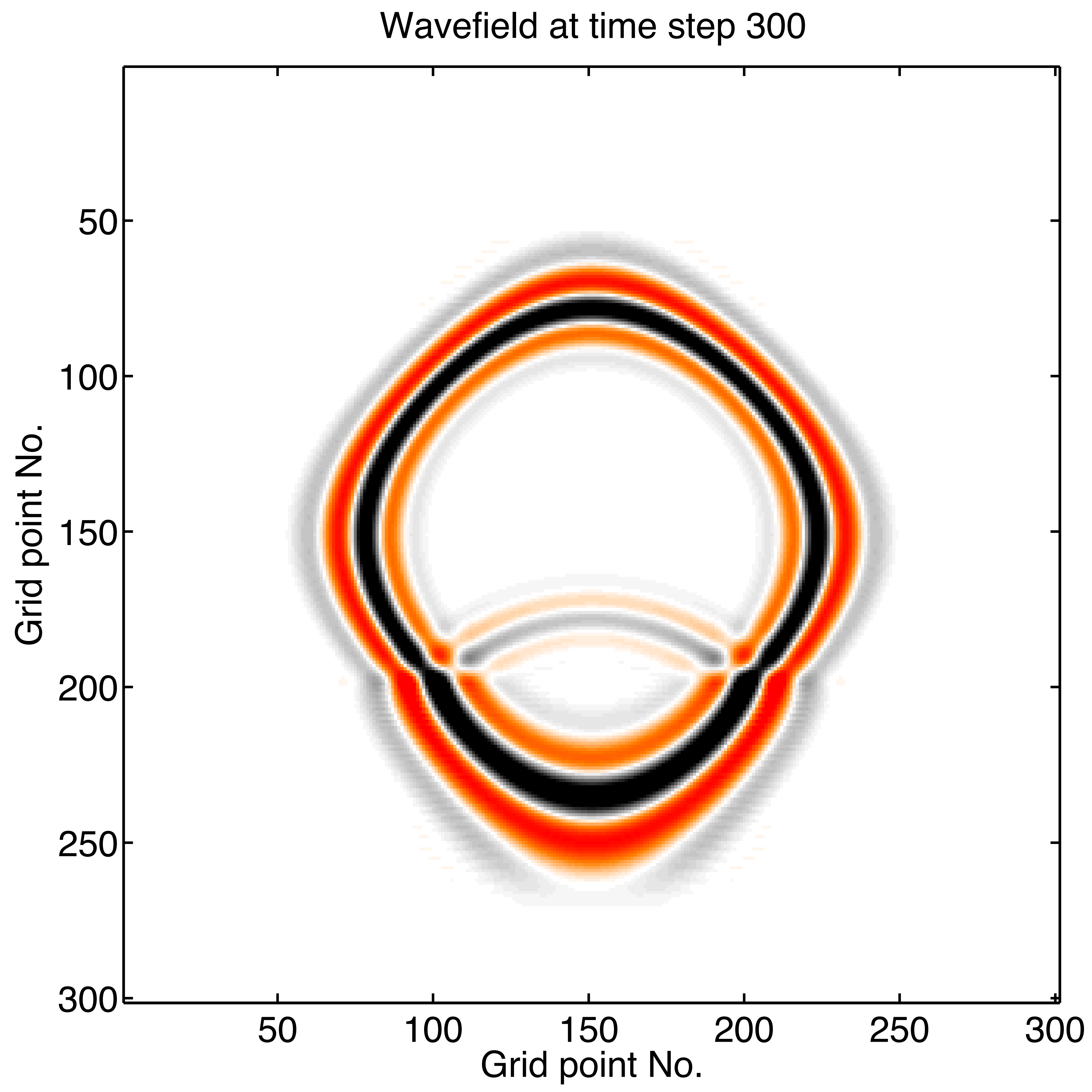
2D Wave field

2D derivative operators:

$$\mathbf{L_x} = \mathbf{D_{xx}} \otimes \mathbf{I_{zz}}$$

$$\mathbf{L} = \mathbf{L_x} + \mathbf{L_z}$$

Model 2D wave equation (non regular grid)



EPS method

EPS derivative operator:

- is a dense matrix
- requires special algorithms to be applied efficiently (e.g. partitioned low rank representation)
- small norm ➡ allows larger time steps
- accuracy in range of machine precision

We need to test whether this operator is applicable for large scale problems!

Low-rank FD

Wave extrapolation in time (Song, Fomel and Ying ,2013):

$$p(\mathbf{x}, t+\Delta t) + p(\mathbf{x}, t-\Delta t) = \int_{-\infty}^{\infty} \hat{p}(\mathbf{k}, t)\cos(|\mathbf{k}|v(x)\Delta t)e^{-i\mathbf{k}\cdot\mathbf{x}}d\mathbf{k}$$

with the mixed domain matrix

$$W(\mathbf{x}, \mathbf{k}) = \cos(|\mathbf{k}|v(x)\Delta t)$$

Low rank representation of $W(\mathbf{x}, \mathbf{k})$

- as spectral method with cost $\mathcal{O}(NN_x \log N_x)$ (N small integer)
- as FD method with cost $\mathcal{O}(LN_x)$ (L related to order of FD scheme/number of FD coefficients)

Low-rank FD

High accuracy of low-rank FD compared to normal FD

- Error of 1D extrapolator in medium with linearly increasing velocity

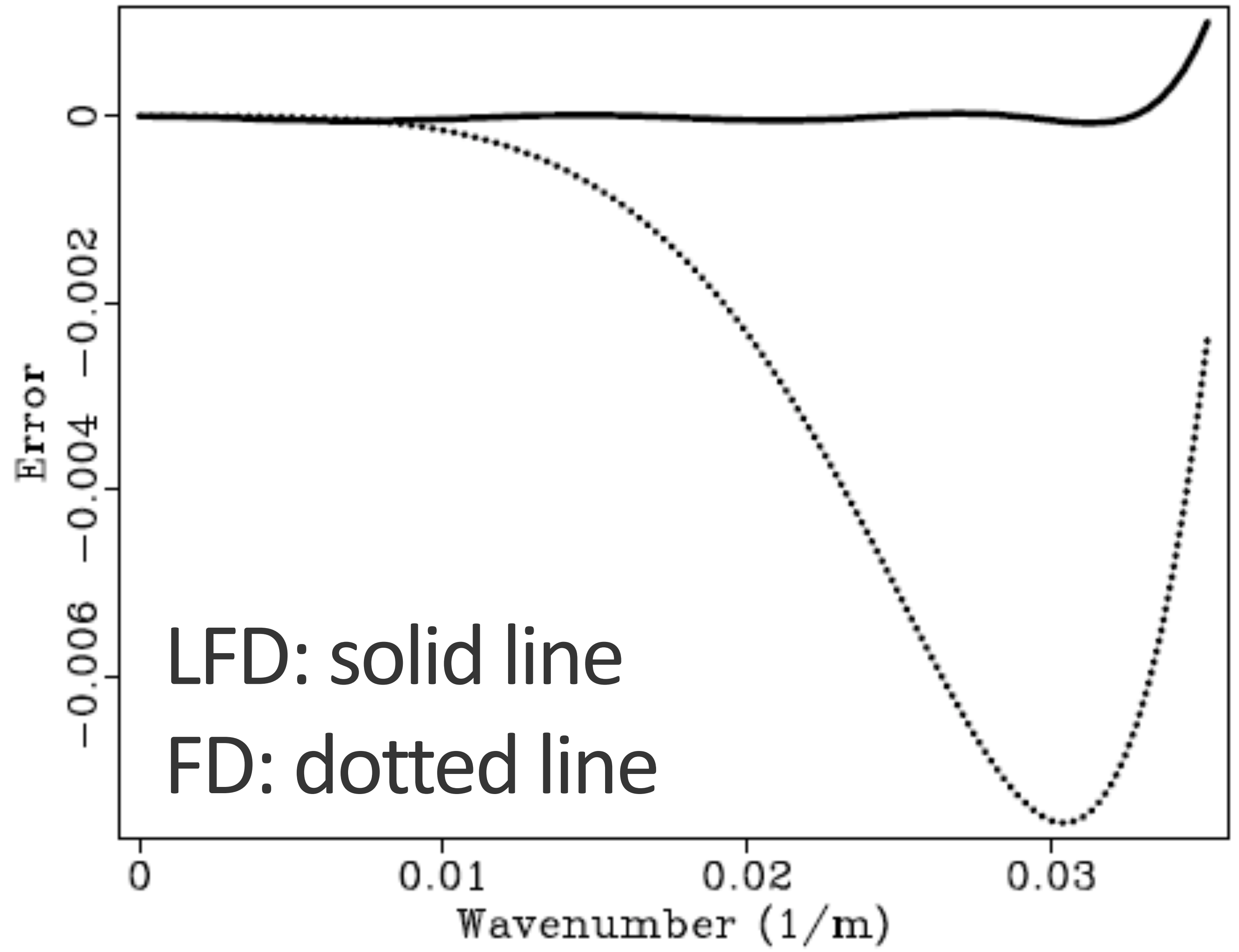


Image from Song et al., 2013

- Extension to TTI media available

Conclusion

Time-domain modeling and inversion code with matrix based operators:

- simple extension to anisotropy (2D TTI with PS method)
- Exact Jacobians and adjoint Jacobians
- Run existing codes for RTM, LSRTM, FWI in anisotropic mode by passing additional argument with Thomsen parameters

Outlook

Current problems:

- PS method is computationally expensive (multi-dimensional FFTs at every time step)
- Anisotropic mode is twice as slow as isotropic mode
- Stability and artifacts are an issue for strong abrupt changes in ϵ, δ, θ

- Investigation in alternative derivative operator (e.g. EPS method)
- LSRTM, FWI on field data sets
- Invert for anisotropy parameters

Acknowledgements

PhD students and Postdocs at SLIM

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada via the Collaborative Research and Development Grant DNOISEII (375142--08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess, ION, Petrobras, PGS, Schlumberger, Statoil, Sub Salt Solutions, Total SA, and Woodside.