Wavenumber Recovery by 2D Frequency-Domain FWI from Slowness- and Source-Frequency-Limited Elastic Seismic Data

Sasmita Mohapatra

And

George McMechan

04/15/2025

Outline

1. Objective
2. Elastic FWI Motivation
3. Elastic FWI Examples
4. Main Ideas
5. Spectral Amplitudes
6. Model Impedances
7. Summary
8. Conclusion

Objective

- Full wave inversion (FWI) can be done in a multi-illumination pattern with different apertures.
- A few frequency components are enough to perform modeling and inversion.
- Migration provides structural geometry; its 2D FT captures reflectivity wavenumbers, not elastic parameter information.
- The inverted model updates contain higher wavenumbers associated with parameter contrasts. The background models contain the low wavenumber portion of the parameter distributions. Ideally these two together provide a continuous distribution of wavenumbers whose spectra can be compared with the spectra of the known target models.
  - Velocities affect both time and amplitude information.
  - Densities affect only amplitude information.

Elastic FWI Motivation

- Elastic Full Waveform Inversion (EFWI) is challenging because of coupling of different parameters. Here, two strategies are used to mitigate nonlinearity in elastic FWI
  - Elastic wavefield simulation is performed using an improved time-domain finite-element method. Inversion is performed in the frequency domain (via DFT), enabling frequency selection and time windowing to reduce nonlinearity.
  - Parameter decoupling is achieved via a multi-steplength gradient approach that searches for optimal steplengths and a combined gradient direction per iteration. Progressively increasing frequencies are added as iterations proceed to avoid cycle skipping.

Time-domain modeling

Improved time-domain finite element modeling

In the time domain, the elastic-wave finite-element (FE) equation can be expressed in matrix form (Marfurt, 1984)

$$\boldsymbol{K}\hat{\boldsymbol{u}}(t)+\boldsymbol{M}\frac{\partial^2\hat{\boldsymbol{u}}(t)}{\partial t^2}=\hat{\boldsymbol{f}}(t).$$

Stiffness matrix

Mass matrix

The FE code is improved in precision, and is parallelized with MPI to distribute shot simulations over cluster nodes.

At receivers, we apply time windowing on the residual wavefields to mitigate the nonlinearly of the inversion.

Calculate the frequency wavefield using DFT

To calculate the frequency wavefields, we use the discrete Fourier transform (DFT) equation

$$u(\omega) = \sum_{t=0}^{T} \hat{u}(t)\, e^{i\omega t}.$$

Frequency wavefield

Time-domain wavefield

Thus, a few selected frequency snapshots are calculated at each subsurface grid point for the gradient calculation. The gradual low-to-high frequency-selection strategy can be applied.

Frequency-domain gradient

Finite-element equation and l_2 objective function in the frequency domain

The corresponding FE modeling equation in the frequency domain (Marfurt, 1984) is

$$\left(\boldsymbol{K}-\omega^2\boldsymbol{M}\right)\boldsymbol{u}(\omega)=\boldsymbol{S}\boldsymbol{u}(\omega)=\boldsymbol{f}(\omega).$$

Impedance matrix

The least-squares (l_2) objective misfit function in the frequency domain is

$$\min(E)=\frac{1}{2}\sum_{\omega}\sum_{s}\left[\boldsymbol{u}_f(\boldsymbol{x}_R,\omega)-\boldsymbol{d}(\boldsymbol{x}_R,\omega)\right]^T\left[\boldsymbol{u}_f(\boldsymbol{x}_R,\omega)-\boldsymbol{d}(\boldsymbol{x}_R,\omega)\right]^*.$$

Predicted frequency data at the receivers

Measured frequency data at the receivers

Gradient Equations in the frequency domain

One element of the gradient vectors is the derivative of the objective function with respect to a subsurface grid parameter p_j (Shin and Cha, 2009)

Forward-propagated source wavefield

Backward-propagated residual wavefield

$$\frac{\partial E}{\partial p_j} = \sum_{\omega} \frac{\mathrm{Re}\left[\sum_{s} \boldsymbol{u}_f^T(\boldsymbol{x},\omega) \frac{\partial \boldsymbol{S}^T}{\partial p_j} \boldsymbol{u}_b^*(\boldsymbol{x},\omega) \right]}{\sum_{s} \boldsymbol{u}_f^T(\boldsymbol{x},\omega) \frac{\partial \boldsymbol{S}^T}{\partial p_j} \frac{\partial \boldsymbol{S}^*}{\partial p_j} \boldsymbol{u}_f^*(\boldsymbol{x},\omega) + \lambda}.$$

Source illumination term

Derivative of impedance matrix

Small damping term

Hence, the gradient vectors: $\boldsymbol{g}_{V_p}, \boldsymbol{g}_{V_s}, \boldsymbol{g}_{\rho}$ are constructed for V_p, V_s, and ρ, respectively.

Elastic FWI theory

Multiple steplengths for the gradient method

To update the model parameters, we utilize the steepest-descent gradient method

Iterations

$$\boldsymbol{p}^{k+1} = \boldsymbol{p}^{k} - \alpha^{k} \boldsymbol{g}^{k}.$$

For elastic FWI, three different parameters need to be updated. Only one steplength α is often not satisfactory.

We use a multi-steplength approach to search for optimal steplengths and a composite gradient direction, for multiple parameters in the elastic FWI,

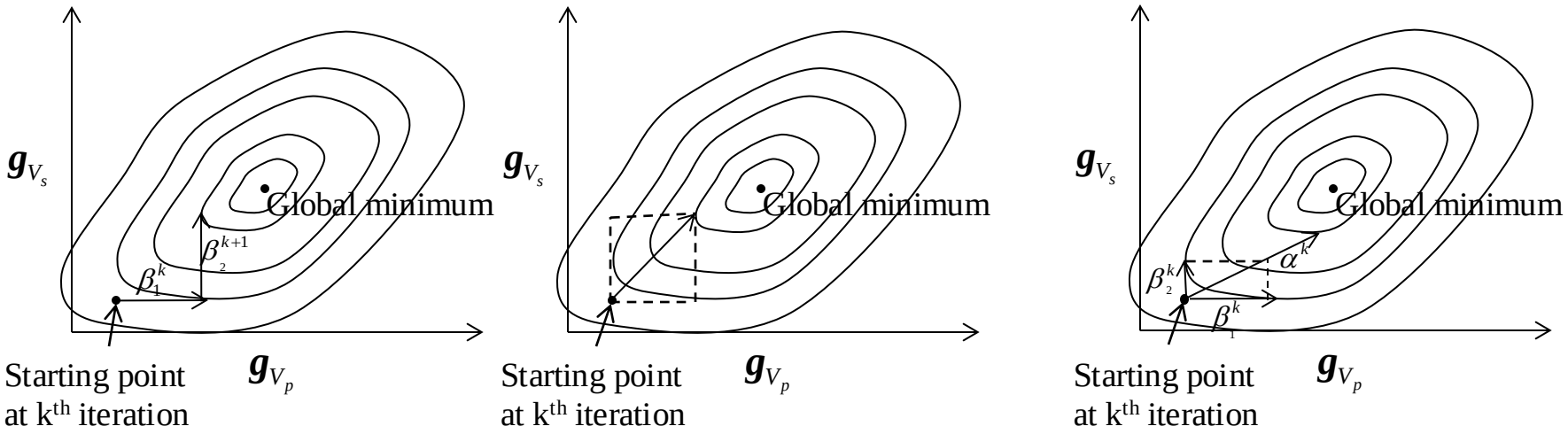
$$\boldsymbol{p}^{k+1}\left(V_p, V_s, \rho\right) = \boldsymbol{p}^{k}\left(V_p, V_s, \rho\right) - \alpha^{k}\left(\beta_1^{k}\boldsymbol{g}_{V_p}^{k} + \beta_2^{k}\boldsymbol{g}_{V_s}^{k} + \beta_3^{k}\boldsymbol{g}_{\rho}^{k}\right).$$

The overall steplength, estimated by parabolic fits.

Individual weights (or steplengths), estimated by a perturbation method one by one.

Physical meaning of multiple steplengths

For example, if we invert two parameters: V_p, and V_s, there are usually three ways to update the model using the gradient method.



(a) Update different parameters in turn (Tarantola, 1986). The convergence is slow.

(b) Ignore the difference of parameter types, and invert them together with the same weights. The convergence is fast, especially using more expensive Gauss-Newton or Quasi-Newton methods. However, it is hard to invert the density simultaneously, which behaves very differently from the velocities.

(c) Using the new scheme, treat them with different weights: β_1, and β_2 at each step. Then find the optimal steplength: α along the combined gradient direction using parabolic fits. The convergence is fast, and is not easily trapped in local minima for different parameter types.

EFWI test on the finite-difference (FD) synthetic data for the Marmousi-2 model

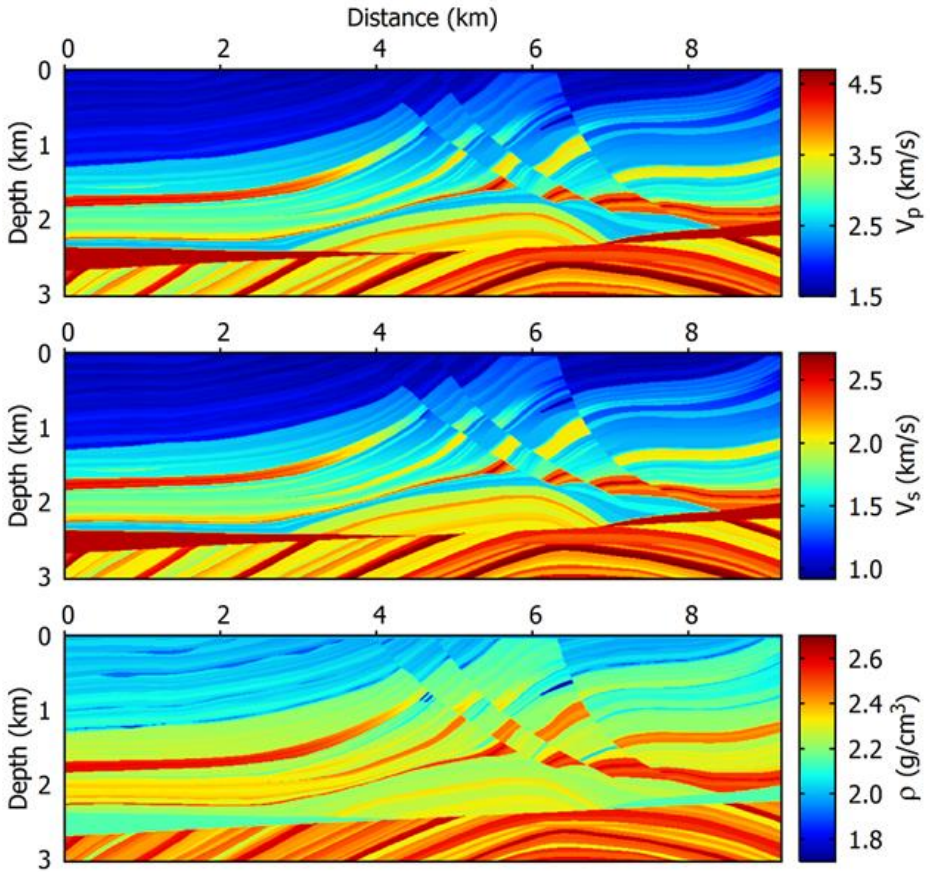
Main Ideas

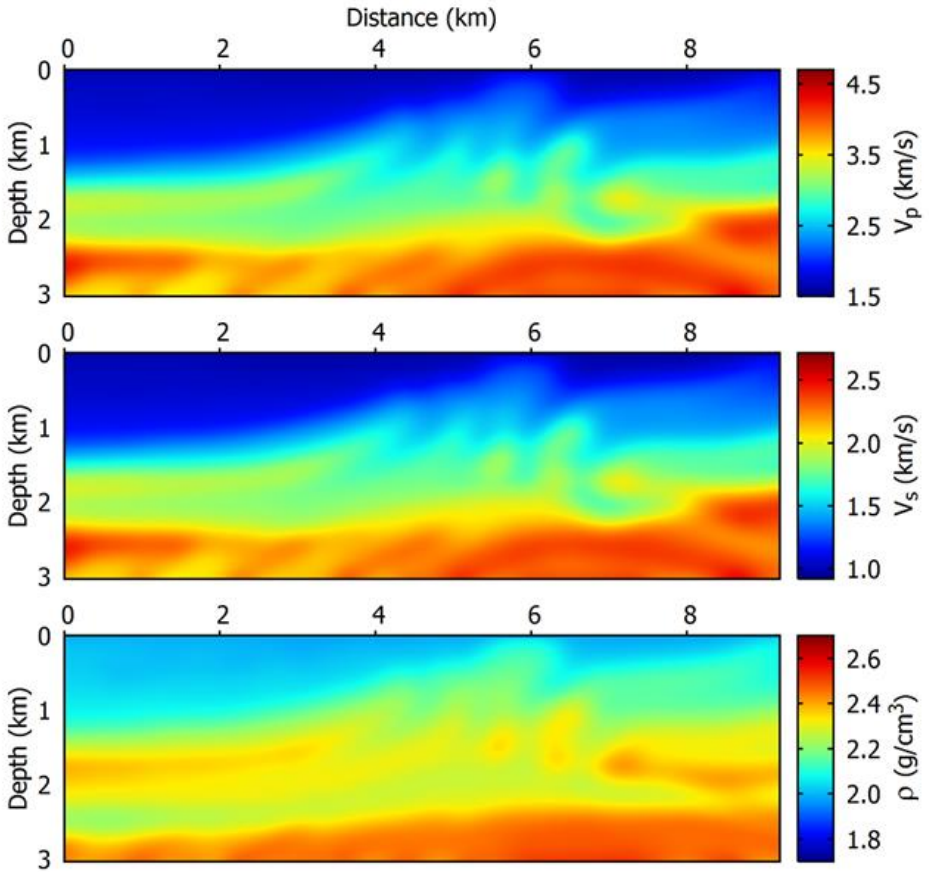
- FWI can extract features only if physical wavenumbers (k) are present in the illuminating wavefield. The fundamental relationship is $k=\omega p$.
- The phase of the data provides stronger constraints than the amplitude; thus velocities are better fitted than density. Image correlation, and calculation of the correlation coefficient (R^2) of model images quantify the behavior of decreasing model misfits as iterations proceed.
- The inverted model updates contain progressively higher wavenumbers associated with parameter contrasts. The background models contain the low wavenumber portion of the parameter distributions. Ideally these two together provide a continuous distribution of wavenumbers whose spectra can be compared with the spectra of the known target models.
  - Velocities affect both time and amplitude information. Densities affect only amplitude information.

EFWI Marmousi2 Model Parameters

- Model grids: 920 x 304 with Grid interval: 10 m or 0.01km
- Source type: 6 Hz Ricker wavelet with Time interval: 0.0015 s
- For S&R along the surface
  - Shot numbers: 111 shots on the surface.
  - Receiver number: 231 recorders along the surface.
- For S&R along 3 edges
  - Shot numbers: 111 on the surface, 33 along the sides (total 177).
  - Receiver number: 231 on the surface, 73 along the sides (total 377).
- Frequencies:
  - Group I: 3.09 Hz-7.65 Hz (15 frequencies)
  - Group II: 7.97 Hz -12.21 Hz (15 higher frequencies)

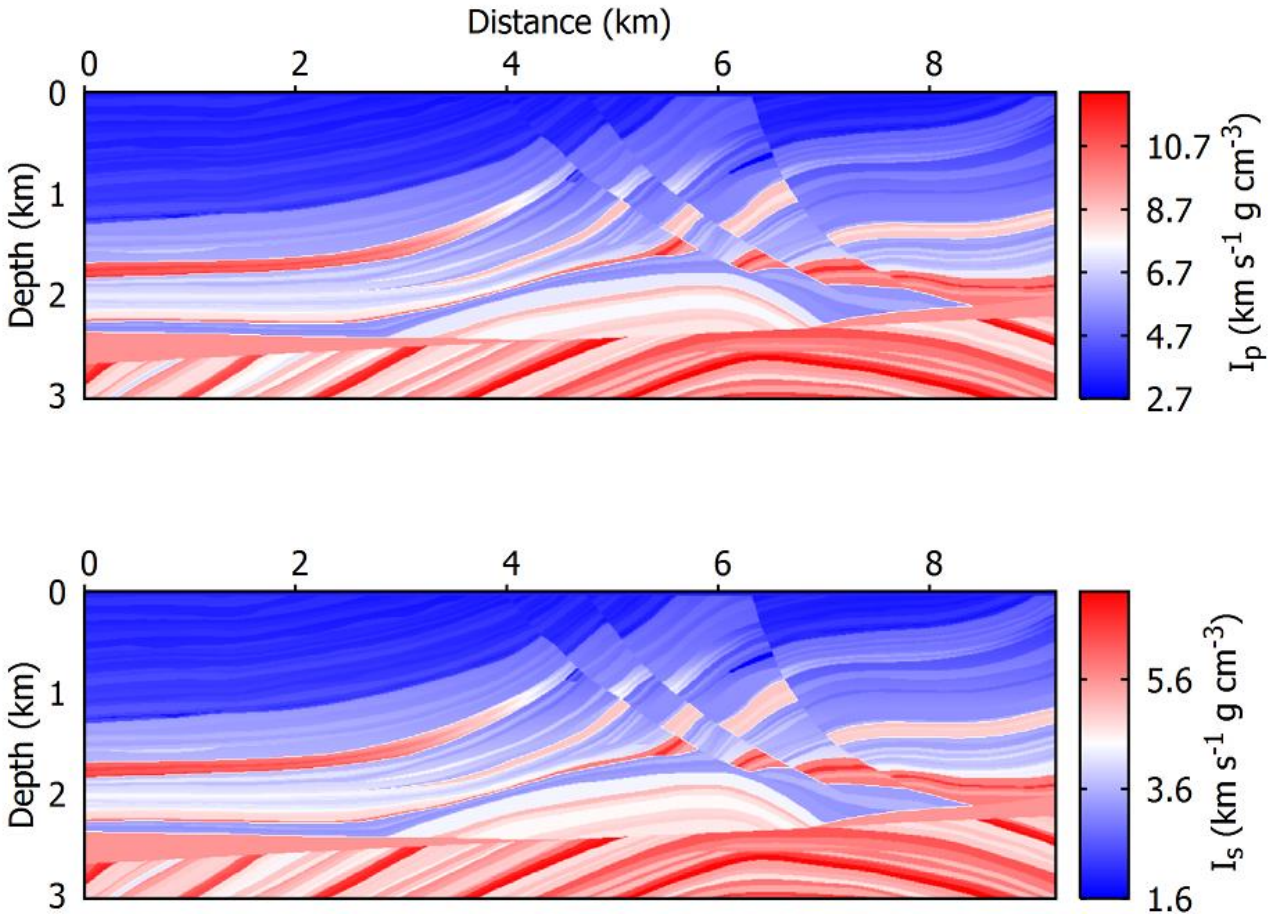
True and Starting Marmousi2 Model





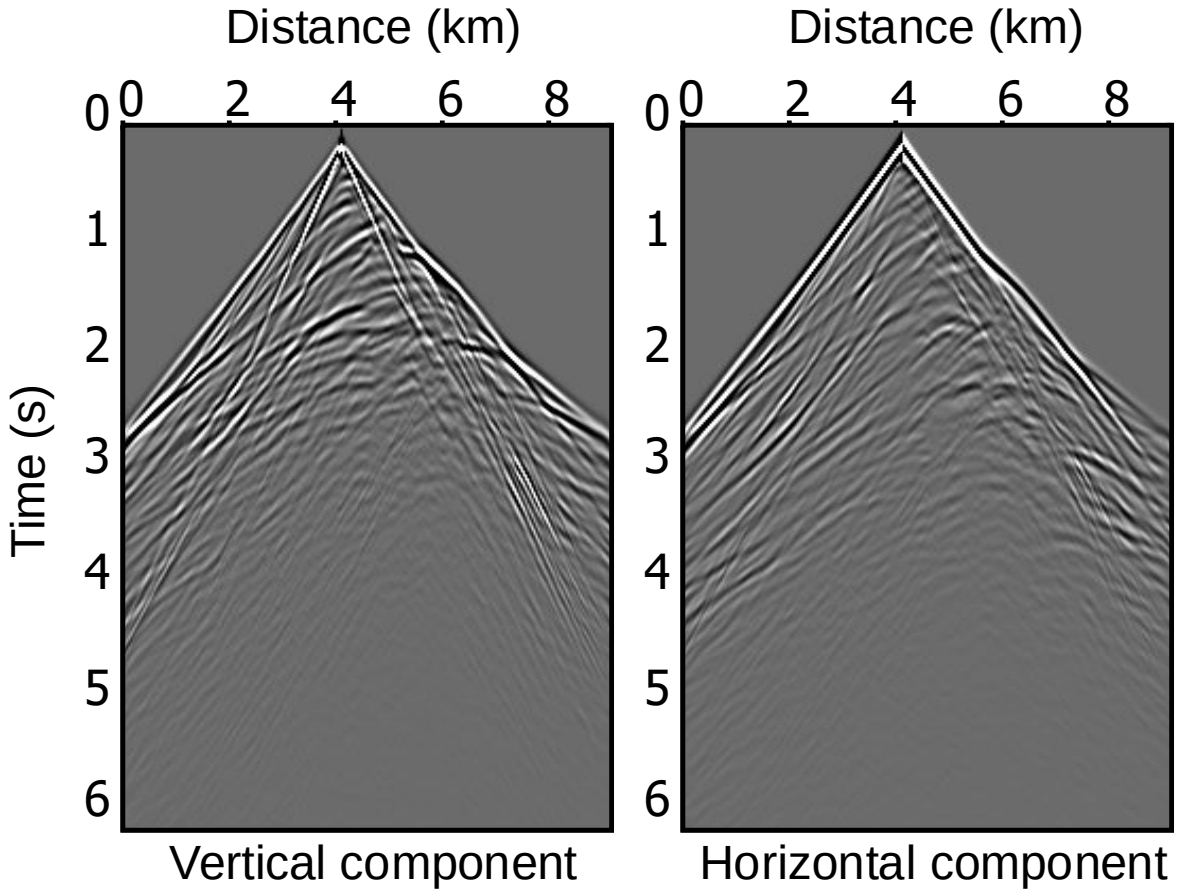
True model impedances: I_p, and I_s

The true impedance models are generated from the products of the velocities and density.

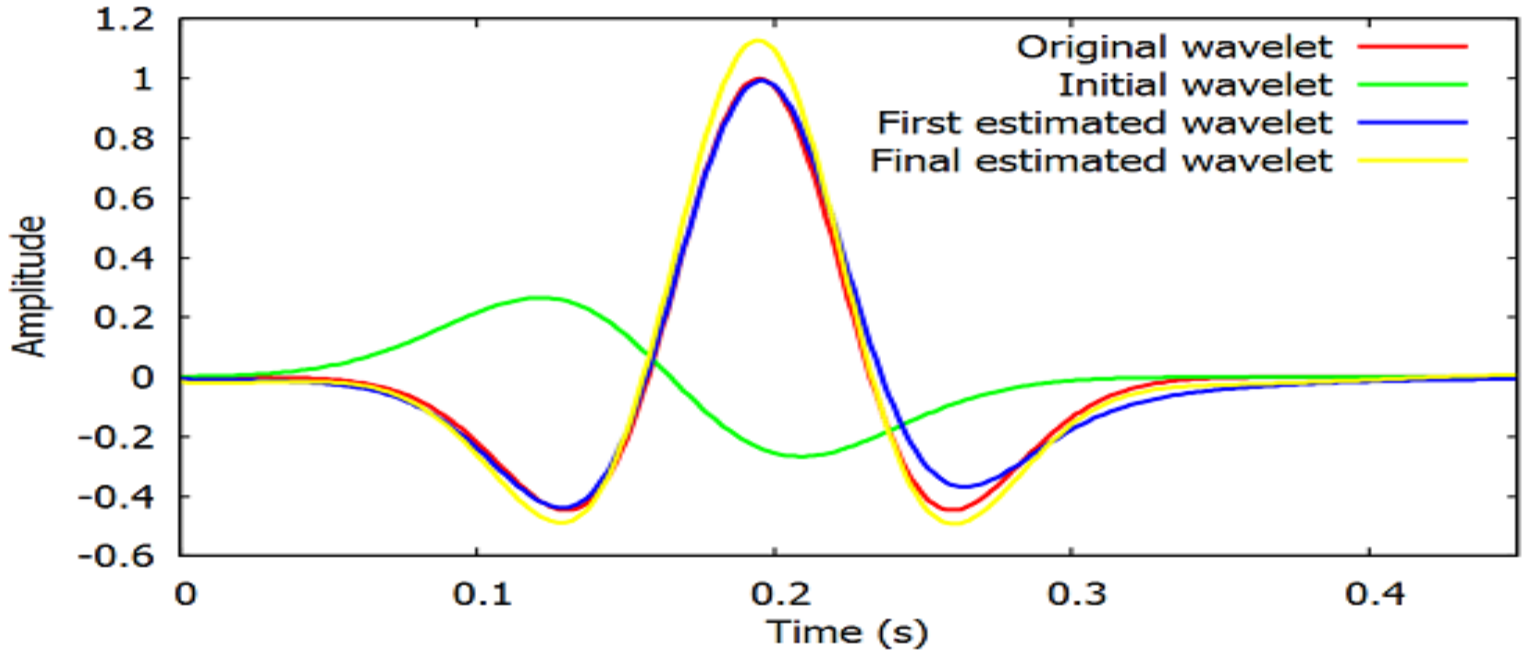


FD Synthetic Shot Gather

There are 111 explosive shot gathers, and 231 two-component receivers along the surface for each shot. The shot gathers are simulated using a 2D time-domain P-SV wave fourth-order stress-velocity staggered-grid FD modeling algorithm, and using a Ricker wavelet with the peak frequency of 6Hz.



The estimated source wavelet



The original Ricker source wavelet (the red line) used for the FD-synthetic wavefields; the initial source wavelet of the first derivative of the Gaussian function (the green line) as the input for the first iteration of the inversion; the estimated Ricker wavelet (the blue line) after the first iteration of the inversion; and the estimated Ricker wavelet (the yellow line) after the final iteration (the 24th iteration in phase II) of the inversion.

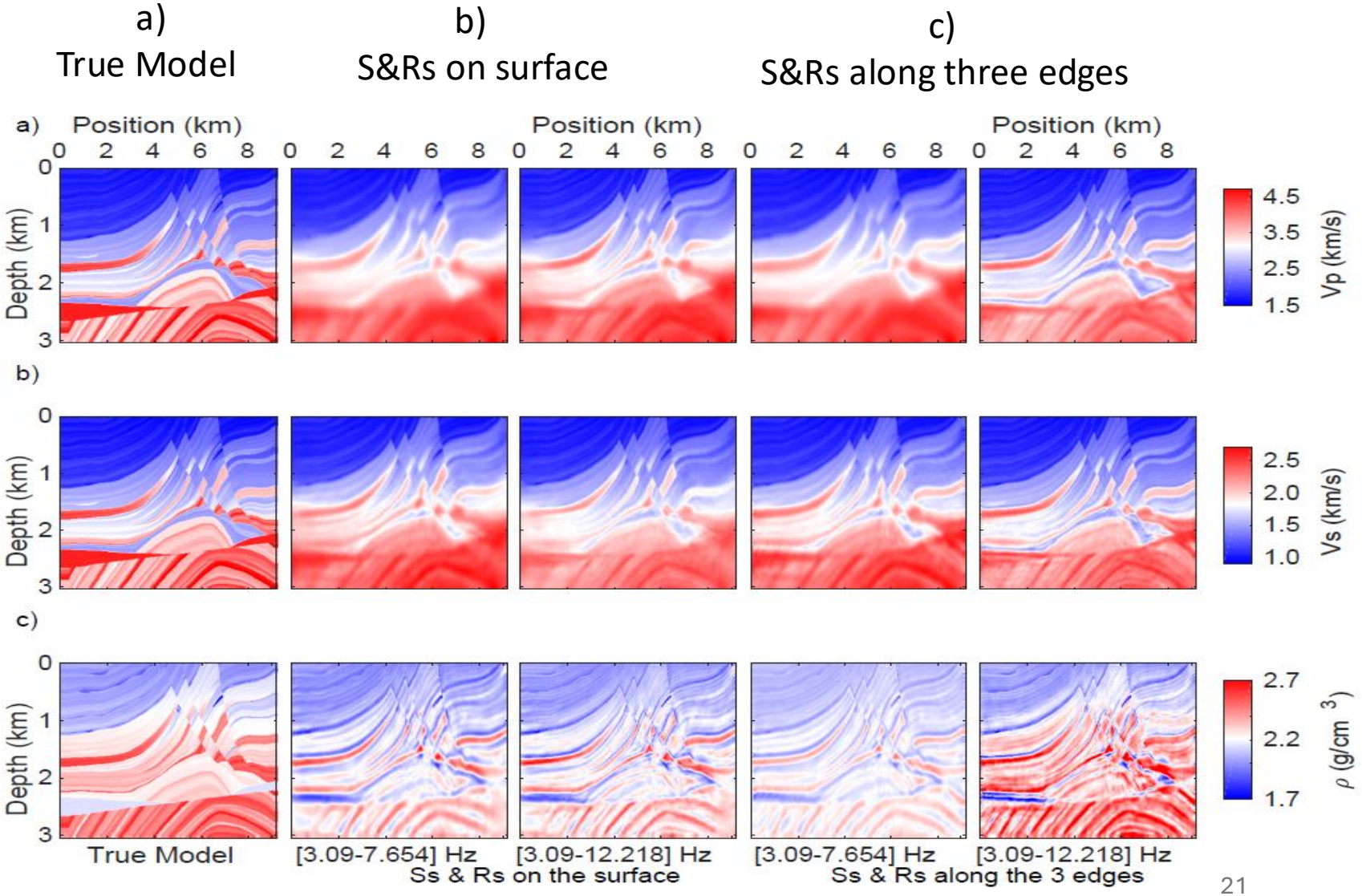
Wide Aperture Marmousi2 model elastic parameters

a) The left column contains the true Marmousi - 2 model.

b) The 2nd and 3rd columns contain the inverted images of data with Ss and Rs only on the surface. The inverted images are after 60 iterations

c) The 4th and 5th columns contain the inverted images of data with Ss and Rs along the surface and the left and right edges. The inverted images are after 60 iterations

d) All images are obtained using the 2-component synthetic gather.



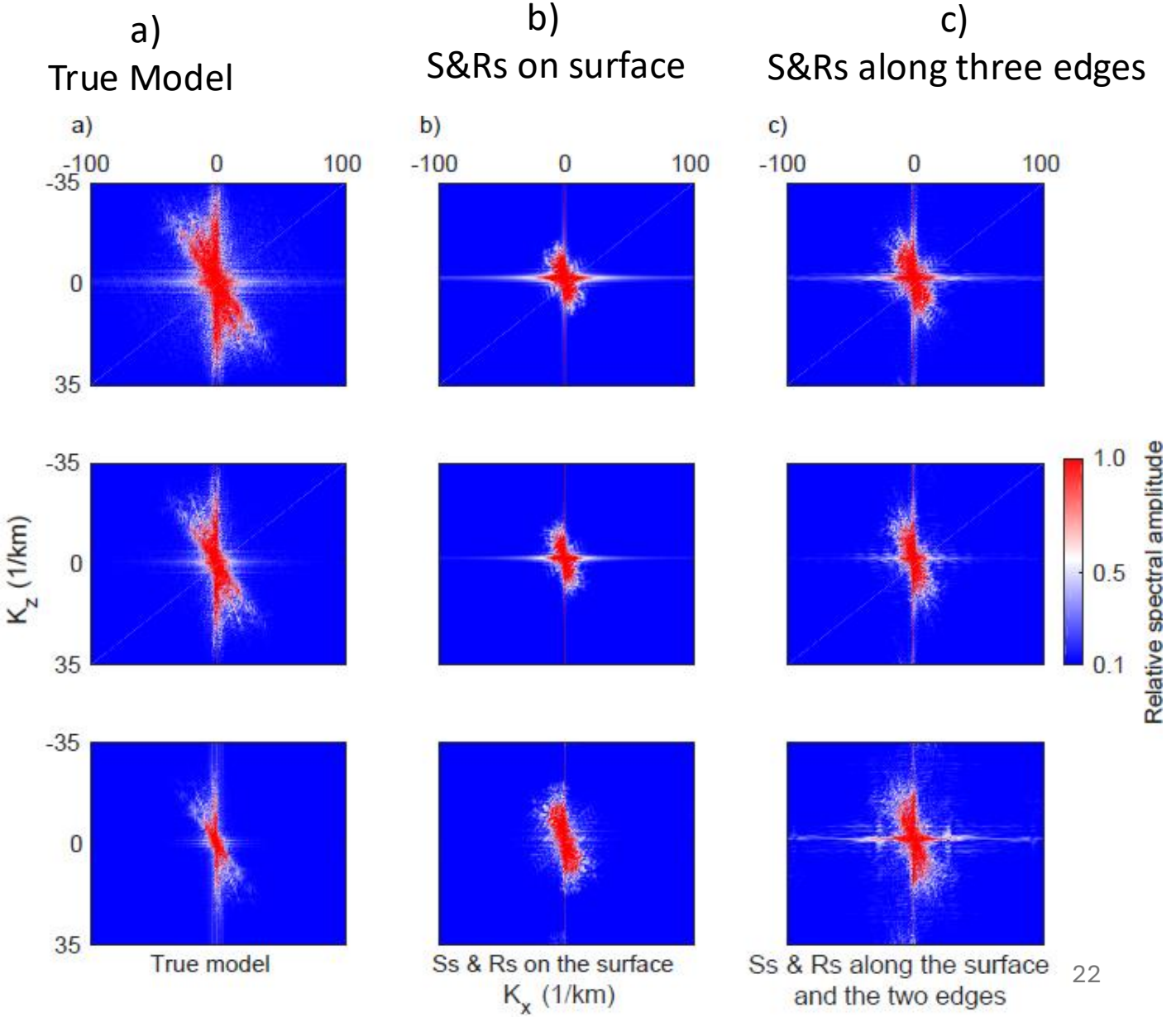
Spectral Amplitude for the Marmousi2 Model

a) Fourier spectral amplitudes as a function of wavenumbers of the images. Column (a) contains spectra of the true model.

b) Column (b) contains the spectra of the inverted model from the data for the Ss and Rs on surface.

c) Column (c) contains the spectra of the inverted model from the data for the Ss and Rs along the surface and the two edges of the Marmousi-2 model.

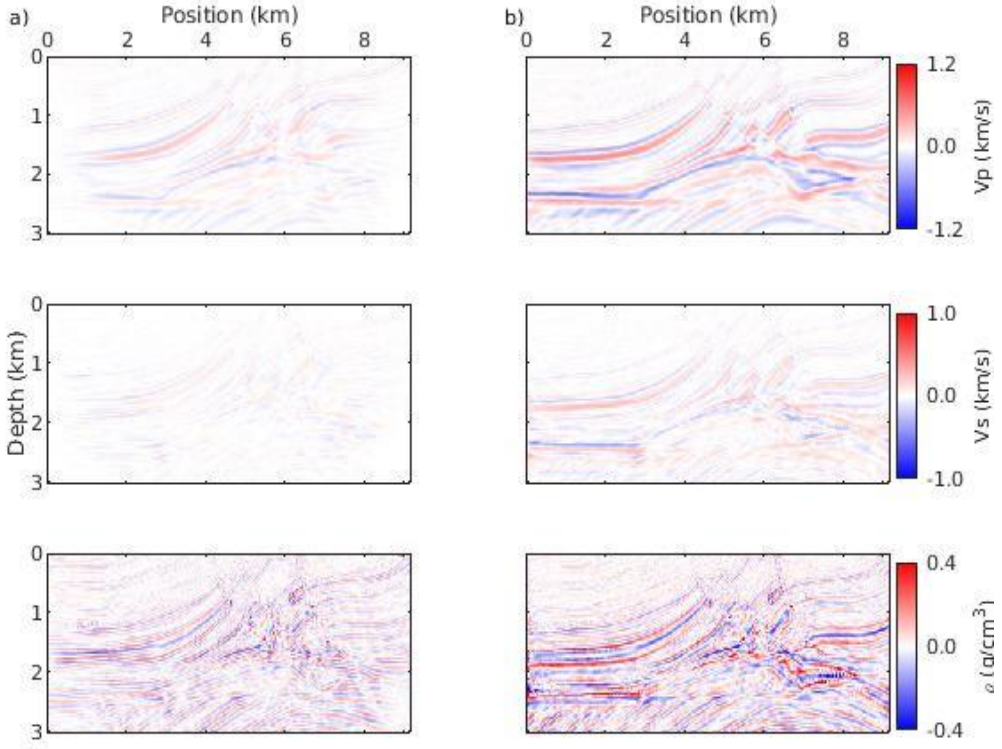
d) The upper row is for Vp, the middle row is for Vs, and the bottom row is for density



Spatial distributions in improvement in inversion (Marmousi2 model)

a) Full frequency window from 3.09 - 12.218 Hz

S&Rs on surface S&Rs along three edges

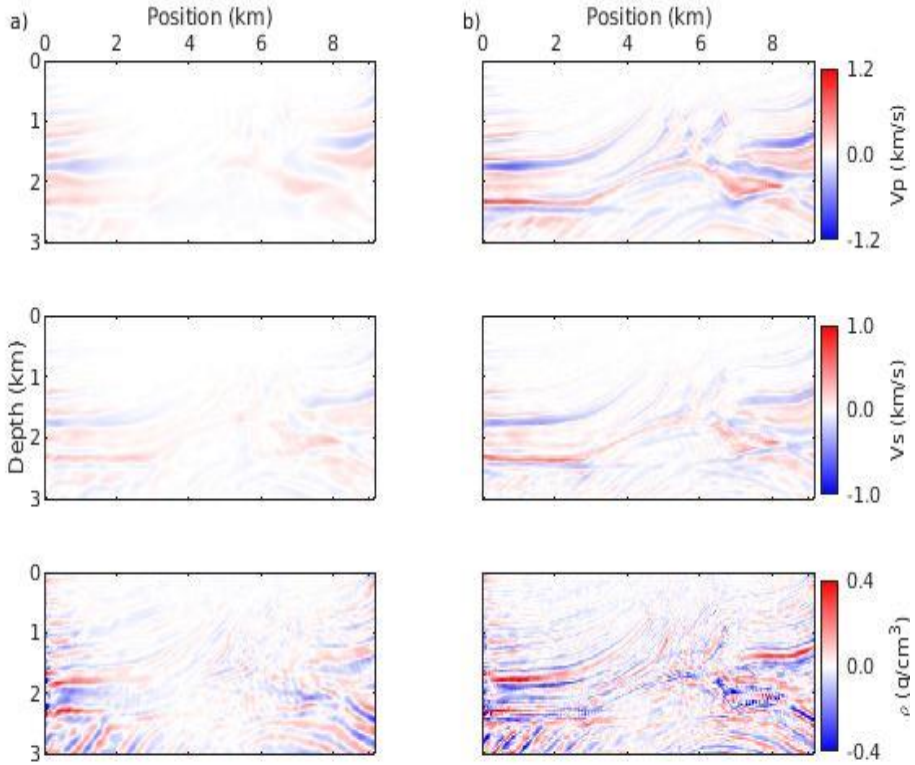


a) Left figure shows the spatial distributions of the improvement in the inversion results obtained by increasing the frequency window from 3.09-7.654 Hz to 3.09-12.218 Hz.

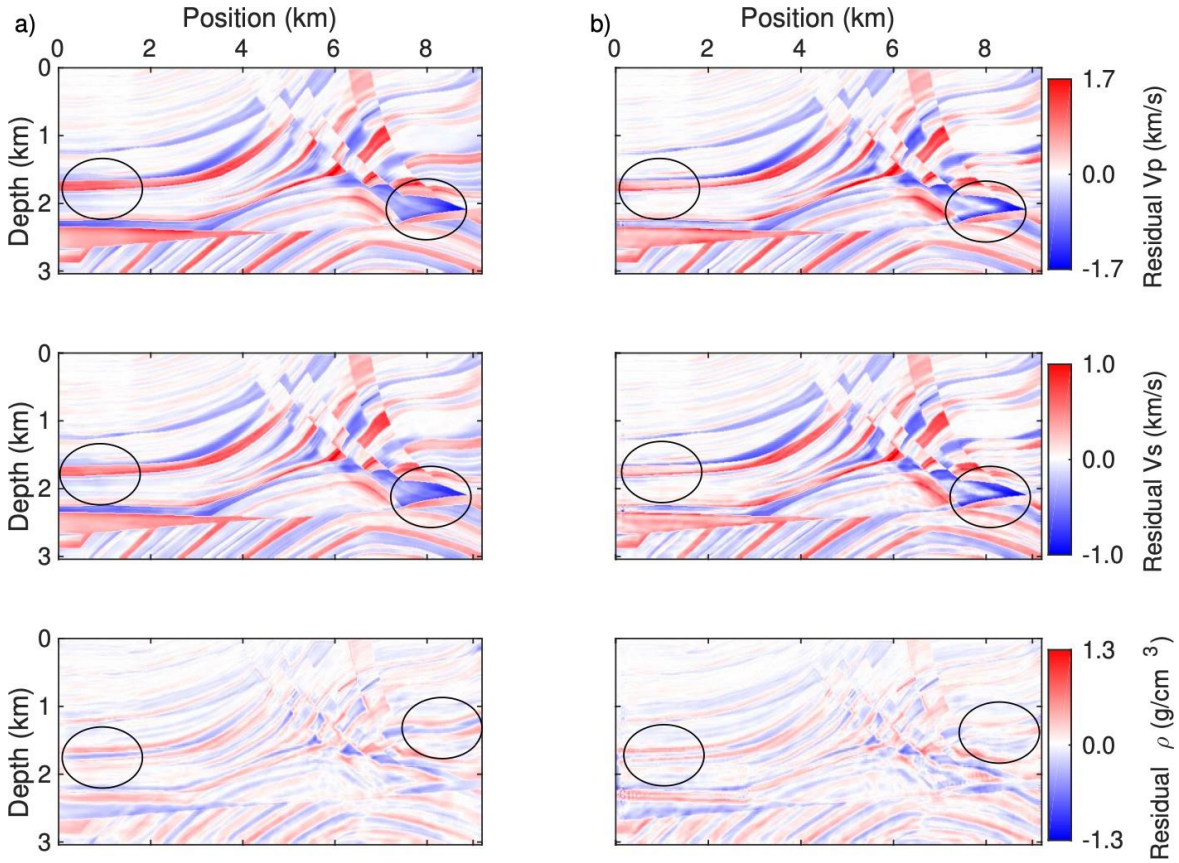
b) Right figure shows the spatial distributions of improvement in the inversion results obtained by increasing the acquisition aperture.

b) S&Rs along three edges

Low frequency window full frequency window



Cont ....

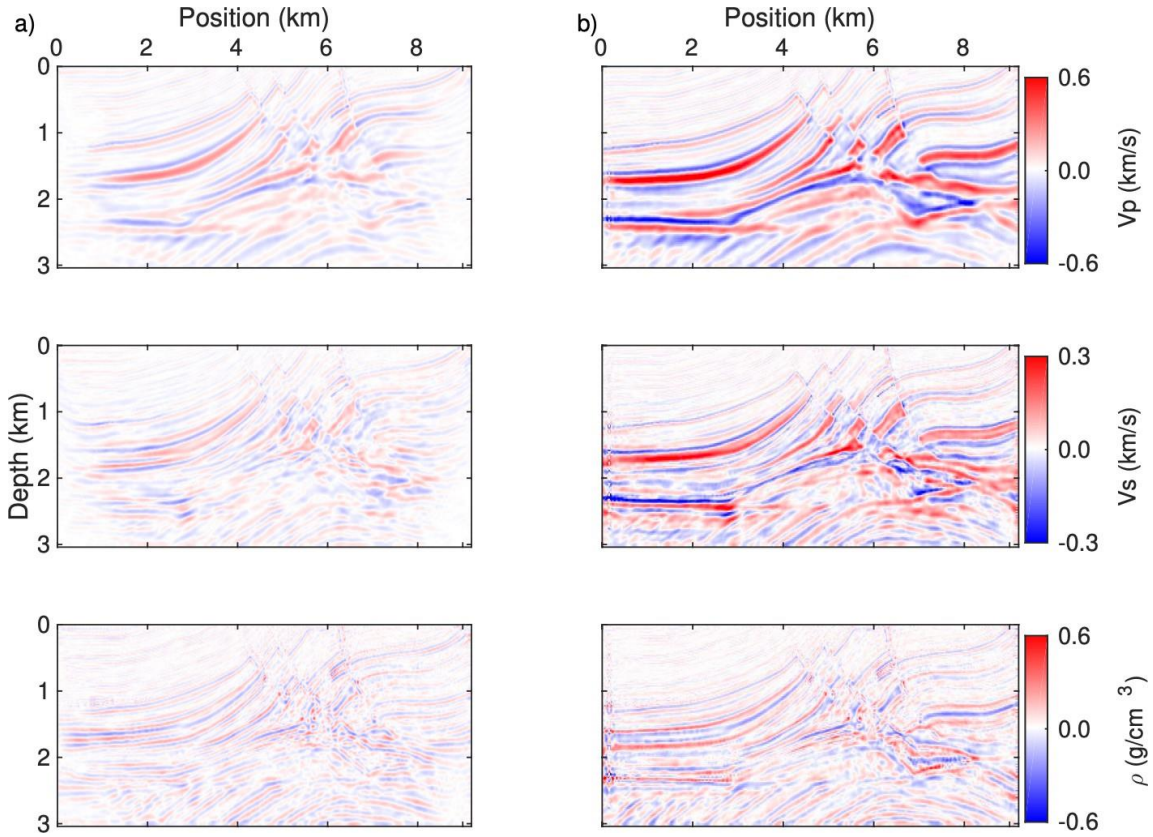


The spatial distributions of VP, VS and ρ residuals which correspond to the parts of the model that are not recovered by inversion for

(a) Ss and Rs on the surface only (the difference between columns 3 and 1 in slide 21) and for

(b) Ss and Rs along the surface and the two edges (the difference between columns 5 and 1 in slide 21). Both (a) and (b) use the full frequency range.

Cont...



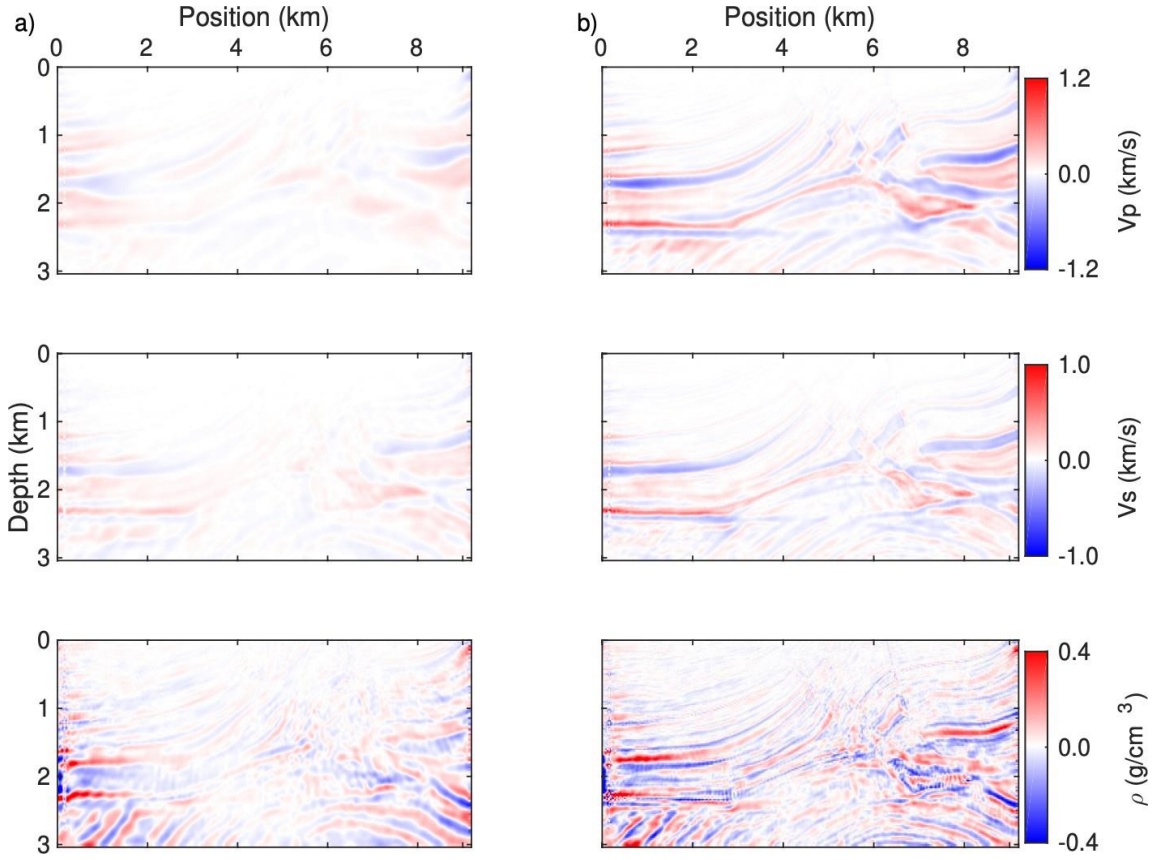
The spatial distributions of the improvement in the inversion results obtained by increasing the frequency window from 3.09-7.654 Hz to 3.09-12.218 Hz for

(a), using the Ss and Rs only on the surface (the difference between columns 2 and 3 in slide 21), and for

(b), using the Ss and Rs also along the edges (the difference between columns 4 and 5 in slide 21).

The image improvement values are larger in (b).

Cont...

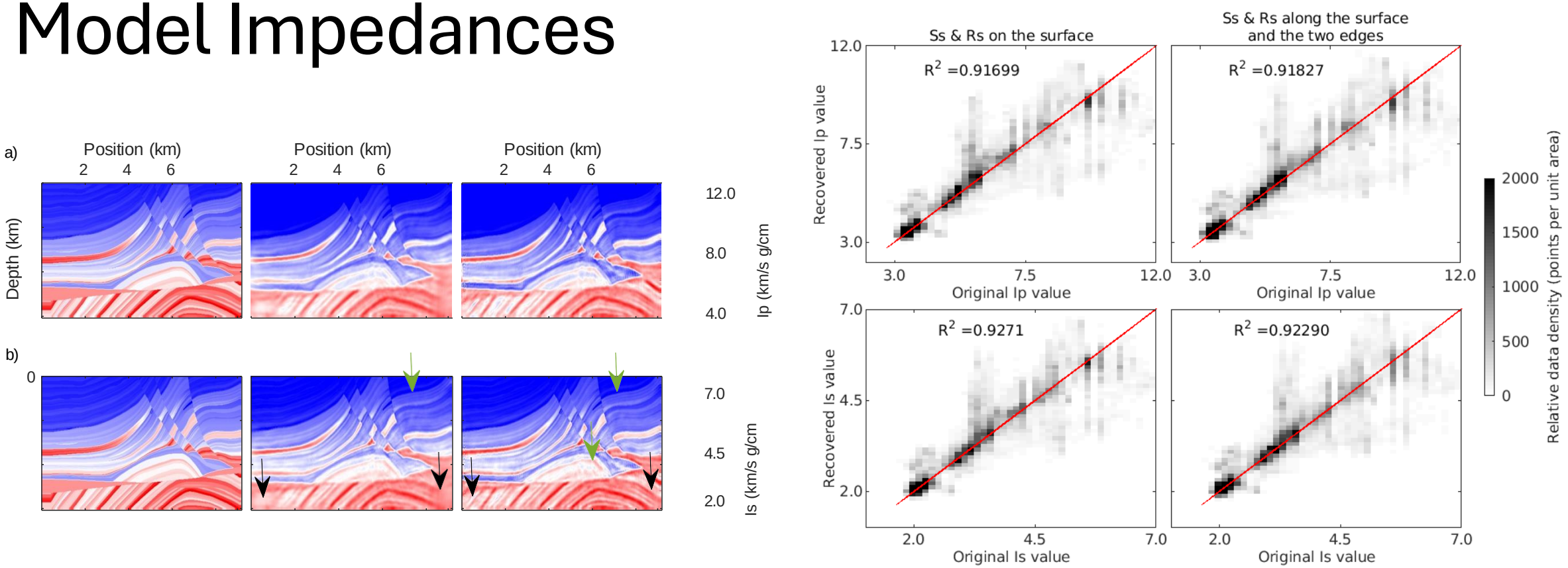


The spatial distributions of the improvement in the inversion results obtained by increasing the acquisition aperture and

(a), using only the low frequency window (i.e., the difference between columns 2 and 4 in slide 21), or

(b), using the full frequency window (i.e., the difference between columns 3 and 5 in slide 21).

Model Impedances



a &b) The true and inverted impedances for a) Ip and b) Is, from the products of velocity and density results. Increasing the aperture increases the ability to invert the deeper, dipping structures (compare at the black arrows), and the Is images have higher resolution than the Ip images from data for a given frequency bandwidth (compare at the green arrows), because range of the S-wave wavenumbers is higher than that of the P-wave wavenumbers.

c) Cross-plots between original and recovered Ip and Is wavefield amplitudes for both Ss and Rs orientations.

Summary

- In comparing the reconstructed models with the true model, the corresponding IC plots, the R^2 values are consistently slightly higher when the Ss and Rs along three edges are used. Expanding the source and recorder apertures to include the two vertical edges increases the wavenumber bandwidth in the inversion by improving the recovery of the dipping features in the model, but is still not as complete as that of the true model.
- In the Marmousi-2 model example, the impact in model fitting is much larger for increasing the acquisition aperture than for increasing the frequency bandwidth.
- The recovered wavenumbers (k) are limited by the source frequency (ω) bandwidth, and the slowness (p) in the data aperture used in the inversion (as $k = \omega p$).
- The computation for all 90 iterations is carried out on 56 CPU cores of Intel Xeon X5650 cluster at UT Dallas, using about 110 hours.

Conclusion

- This pilot study sets of stage for future evaluation of acquisition design for optimization of inversion to recover wavenumber information from targets for various wavenumber characteristics.
- Various combinations of Ss and Rs space/time or the equivalent wavenumber/frequency representations of a target, its recorded data, and its spectra, can be simulated prior to field data acquisition to maximize the information recovery, and to quantify the corresponding cost/benefit trade-offs.
- The impedances from the products of the velocities and density compensate for their under- and over-estimations of the deviations.
- We are currently analyzing the performance impact on the Frontera Cascade Lake node.

References

- *A.J. Devaney,(1982).* A filtered backpropagation algorithm for diffraction tomography. *Ultrasonic Imaging 4, 336-350*.
- *Pan and Kak, (1983).* A computational study of reconstruction algorithms for diffraction tomography. *IEEE Transactions of acoustics, speech, and signal processing, vol. ASSP-31*.
- *A.J.Devaney,(1984).* Geophysical Diffraction Tomography. IEEE Transactions on Geoscience and remote sensing. *Vol. GE-22, No.1.*
- *M.Slaney and A.C.Kak, (1986).* Diffraction Tomography.
- *Ru-Shan Wu and M.Nafi Toksoz, (1987).* Diffraction tomography and multisource holography applied to seismic imaging.
- *Ali Mohammad-Djafari and Guy Demoment, (1987).* Maximum entropy Fourier synthesis with application to diffraction tomography*. Applied Optics vol.26,No.9*
- *Liang-Zie Hu & McMechan, (1987).*Wave-field transformations of vertical seismic profiles. *Geophysics vol. 51. No. 10.*
- *Kun Xu and George A. McMechan, (2014)*. 2D frequency-domain elastic FWI using time-domain modeling and multistep-length gradient approach. *Geophysics vol. 79, No. 2.*
- *Sasmita Mohapatra and George A. McMechan, (2021).* Wavenumber Recovery by FWI from slowness-limited, and source-frequency-limited elastic seismic data. *Journal of Seismic exploration, 30, 577-600*