Explicit control of numerical dispersion and instability in wavefield modeling and inversion

Odette Aragao & Paul Sava

Center for Wave Phenomena and Dept. of Geophysics, Colorado School of Mines, Golden CO 80401
aquinodearagao@mines.edu

ABSTRACT

Full waveform inversion (FWI) is a nonlinear inverse problem that refines an initial model by iteratively matching the recorded with the modeled data. The success of FWI relies on accurate and efficient wavefield modeling. The finite-difference method (FDM) is one of the most popular techniques used for this purpose because of its easy implementation and small memory requirement. However, depending on the spatial and temporal discretization, FDM may encounter numerical dispersion and instabilities, which degrade the quality of the simulated wavefields. The conventional FWI objective function, formulated based on data misfit, may lead to situations where the updated models violate conditions that are necessary to guarantee modeling stability. We propose to address this problem by explicitly incorporating a stability constraint into the inversion through a penalty term based on a probability density functions (PDF) that represents the range of velocities in which numerical dispersion and instabilities do not occur. This penalty term prevents updated models from entering into a petrophysical configuration that leads to dispersion and instability for specific choices of sampling parameters. Through elastic synthetic examples, we demonstrate the benefits of including this stability constraint into the FWI framework, as it ensures the efficiency of the FDM engine, while at the same time increasing the quality of the recovered models.

Key words: full-waveform inversion, modeling, dispersion, finite difference

1 INTRODUCTION

Full waveform inversion (FWI) is one of the most promising technique among all wavefield tomography methods, as it uses most of the information provided by seismic data to deliver high resolution models of subsurface properties. The inverse problem is constrained by a wave equation and it is solved by iteratively minimizing an objective function, which computes the misfit between observed and modeled data (Tarantola, 1984; Virieux and Operto, 2009).

The computational cost of methods based on a wave equation is dominated by the numerical wavefield extrapolation, i.e., simulating the propagation of seismic waves through earth models. Additionally, the accuracy and efficiency of the FWI engine relies on the quality of wavefield modeling. In order to calculate both spatial and time derivatives needed for numerical solutions to a wave equation, one may use the finite-difference method (FDM), which replaces derivatives with finite-difference approximations of limited accuracy. FDM is preferred over other numerical methods for wave equation simulation because of its simple implementation and reduced computational cost (Kelly et al., 1976; Virieux, 1986; Dablain, 1986; Moczo et al., 2004).

The accuracy of FDM depends on spatial and temporal discretization. On the one hand, a fine spatial grid increases the precision of computed derivatives while significantly increasing the computational cost. On the other hand, coarse grids reduce the modeling costs but introduce numerical errors in the simulations, causing phase distortions in the modeled wavefield. Additionally, large time steps may cause instabilities of the wavefield extrapolation (Courant et al., 1967), thus hampering FWI. Therefore, it is important to determine appropriate grid sizes in order to avoid numerical distortions and instabilities and increase the computational efficiency, while assuring that the modeling is accurate (Courant et al., 1967; Boore, 1972; Alford et al., 1974; Dablain, 1986; Wu et al., 1996).

Courant et al. (1967) propose the CFL condition, which defines an upper time step limit for stable finite-difference solutions of partial differential equations. If this condition is violated, large errors are produced during the simulation. The CFL condition is directly related to the spatial grid size and to the maximum velocity of the medium. Dablain (1986) shows that numerical dispersion can only be avoided by sampling with enough number of grid points per minimum wavelength, i.e., the computation grid depends on the minimum velocity of the medium and on maximum frequency. Therefore, the space and time discretization can not be arbitrarily chosen as they depends on the medium properties.

During FWI, the gradient of the objective function purely based on the misfit between observed and predicted data may lead to velocity updates that violate the CFL condition or cause undesirable dispersion, which results in inaccurate models.

We propose to avoid numerical instabilities during the inversion by explicitly incorporating a stability constraint into the FWI objective function, based on a probability density function (PDF), defining the range of parameters for which the wavefield simulations remain stable. We use the formulation proposed by Aragao and Sava (2020a,b) in order to recover models that do not violate the CFL condition or create numerical dispersion. The performance of the proposed methodology is illustrated with synthetic elastic examples.

2 THEORY

We consider the isotropic elastic wave equation

$$\ddot{\mathbf{u}} - V_P^2 \left[\nabla(\nabla \cdot \mathbf{u})\right] + V_S^2 \left[\nabla \times (\nabla \times \mathbf{u})\right] = \mathbf{f}, \tag{1}$$

where $\mathbf{u}(e, \mathbf{x}, t)$ is the elastic wavefield, $\mathbf{f}(e, \mathbf{x}, t)$ is the source function, and e, $\mathbf{x}$ and t are, respectively, the experiment index, space coordinates and time. The model parameters, V_P and V_S, are P- and S-waves velocities, respectively.

We solve the full-waveform inversion problem by minimizing an objective function $\mathcal{J}(\mathbf{u}, V_P, V_S)$ consisting of a term that estimates the misfit between the predicted and observed data $\mathcal{J}_D(\mathbf{u}, V_P, V_S)$ using the L-2 norm, and a stability penalty $\mathcal{J}_S(V_P, V_S)$ that prevents numerical instabilities and grid dispersion during the FWI process:

$$\mathcal{J}(\mathbf{u}, V_P, V_S) = \mathcal{J}_D(\mathbf{u}, V_P, V_S) + \mathcal{J}_S(V_P, V_S). \tag{2}$$

During the numerical modeling, the phase velocity becomes a function of the spatial grid and the phase velocity may be different than the true velocity when using an insufficient number of grid points per wavelength, affecting the high-frequency content of the data and, consequently, creating numerical dispersion (Boore, 1972; Alford et al., 1974). For implementing fourth-order finite difference methods, if the grid size in any dimension is coarser than a quarter of the minimum wavelength, numerical grid dispersion may occur, leading to wavefield extrapolation instability. Therefore, the efficiency of the wavefield modeling can be quantified by the relationship (Dablain, 1986):

$$\Delta x \leq \frac{\lambda_{min}}{4} = \frac{V_{min}}{4 f_{max}}, \tag{3}$$

where Δx represents the spatial grid in any direction, λ_{min} and V_{min} correspond, respectively to the minimum wavelength and the minimum velocity, and f_{max} is the maximum frequency.

Additionally, numerical instabilities may occur when one uses a propagation timestep that does not respect the CFL condition (Courant et al., 1967), which is given by

$$\Delta t \leq \mathcal{C} \frac{\Delta x}{V_{max}}, \tag{4}$$

where Δt represents the time step, $\mathcal{C}$ is the Courant number and V_{max} corresponds to the maximum velocity. The CFL condition is a necessary condition for avoiding dispersion when numerically solving partial differential equations.

Furthermore, for multiparameter inversion that uses only the data misfit, model parameter updates may produce unphysical models (e.g., negative velocities), which also lead to instabilities in wavefield extrapolation. In order to purposefully prevent numer-

ical instabilities and numerical dispersion, we design a stability penalty $\mathcal{J}_S$ in the elastic FWI (EFWI) objective function. We adapt the formulation of the probabilistic petrophysical constraint proposed by Aragao and Sava (2020a,b), which is not calculated in the physical space domain, but in the space formed by the model parameters. We define a probability density function (PDF) that represents the range of model parameters for which wavefield simulations remain stable. The PDF is represented by coordinates $\mathbf{p}$ and probability density $b(\mathbf{p})$, such that $b = 0$ for combinations of parameters that cause wavefield simulation instability. Representing the elastic model components by the vector $\mathbf{m}(\mathbf{x}) = \{V_P(\mathbf{x}), V_S(\mathbf{x})\}$, we define the stability penalty as

$$\mathcal{J}_S = a \sum_{\mathbf{m}} D(\mathbf{m}), \tag{5}$$

such that the scalar parameter a determines the relative strength of the stability term in the objective function and $D(\mathbf{m})$ represents the distance from a particular combination of model parameters $\mathbf{m}$ to the entire distribution and it is given by

$$\frac{1}{D(\mathbf{m})} = \sum_{\mathbf{p}} \frac{b(\mathbf{p})}{\|\mathbf{m} - \mathbf{p}\|_2}. \tag{6}$$

The gradients of $\mathcal{J}_S$ with respect to the model parameters V_P and V_S are, respectively,

$$\frac{\partial \mathcal{J}_S}{\partial V_P} = a\, D^2(\mathbf{m}) \sum_{\mathbf{p}} \frac{b(\mathbf{p})\,(V_P - V_{Pp})}{(\,\|\mathbf{m} - \mathbf{p}\|_2\,)^3}, \tag{7}$$

and

$$\frac{\partial \mathcal{J}_S}{\partial V_S} = a\, D^2(\mathbf{m}) \sum_{\mathbf{p}} \frac{b(\mathbf{p})\,(V_S - V_{Sp})}{(\,\|\mathbf{m} - \mathbf{p}\|_2\,)^3}. \tag{8}$$

The value of $\mathcal{J}_S$ is determined by the distance to high probabilities in the model parameter space, i.e., by the distance to the combinations of V_P and V_S such that the wavefield extrapolation is stable. Therefore, outside the stability region defined by the PDF, the gradient of $\mathcal{J}_S$ explicitly forces the models towards the stable region. For points $\mathbf{m}(\mathbf{x}) = \{V_P(\mathbf{x}), V_S(\mathbf{x})\}$ inside the stability region, the data misfit $\mathcal{J}_D$ fully controls the total gradient (Aragao and Sava, 2020a).

3 EXAMPLES

We illustrate our EFWI method with two synthetic examples and compare inversions using just the data misfit term $\mathcal{J}_D$ with inversion that also use stability constraints $\mathcal{J}_S$.

3.1 Gaussian Anomalies

The synthetic models depicted in Figure 1 contain two positive Gaussian anomalies centered at $(1.2, 0.75)$ and $(1.2, 1.75)$ km. There are 40 vertical displacement sources in a well at $x = 0.05$ km and a line of geophones at $x = 2.35$ km. For this experiment, we use the following parameters: $f_{max} = 15$ Hz; $\lambda_{min} \approx 33.3$ m; and the minimum velocity for the S-wave and maximum velocity for the P-wave are, respectively, $V_{S\,min} = 500$ m/s and $V_{P\,max} = 2000$ m/s. Therefore, the maximum spatial grid is 8 m and, for $\mathcal{C} = 0.5$, the time step needs to be smaller than 0.0023 s.

For $\Delta x = 8$ m and $\Delta t = 0.002$ s, the inversion is stable when velocities are in the range $[480,\ 2000]$ m/s, which conforms to the true models (Figure 1). Inverted velocities outside this range produce wrong models and also cause wavefield simulation instabilities. Therefore, this example illustrates how we can avoid numerical dispersion and also prevent the inversion to generate wrong models by using the stability penalty, which constrains the range of updated velocities.

We start the inversion with a model with constant $V_P = 1500$ m/s and $V_S = 500\,m/s$. Considering the range of velocities for which the wavefield simulations are stable, the inversion becomes unstable when we have V_S updates below -20 m/s or when V_P updates above 500 m/s. Therefore, we define a PDF for calculating the stability penalty consisting of the distribution shown in Figure 2a. Note that the probability density is zero for velocities that produce wave propagation instability. Figure 2b displays

the distance $D(\mathbf{m})$ and the gradient of $D(\mathbf{m})$, which are both zero in the region of stability. If the inverted velocities are inside the range defined by the PDF (Figure 2a), the data misfit $\mathcal{J}_D$ fully controls the inversion gradient as the gradients of $\mathcal{J}_S$ are zero. Otherwise, the stability constraint produces a gradient that drives the updated models to the region of stability defined by the PDF.

Without the stability constraint, we observe numerical dispersion in the inversion after few iterations. The inversion updates in Figure 3 show that, without imposing stability constraints, the V_S model has updates smaller than -20 m/s, producing velocities that are smaller than 480 m/s, thus creating wavefield extrapolation instability. Additionally, the inversion that uses just the data misfit produces negative anomalies (Figure 3) at the locations of the positive anomalies of the true models (Figure 1). Therefore, inversion based only in $\mathcal{J}_D$ produces wrong velocity models in addition to numerical dispersion.

When we include the stability constraints into the inversion objective function, inversion proceeds without any instability. The recovered models after 50 iterations are shown in Figure 4. Although the inverted models have artifacts related to the interparameter crosstalk between the elastic parameters and to the limited acquisition aperture, both anomalies have amplitude and shape closer to the true models (Figure 1). The inversion updates (Figure 5) demonstrate that the stability constraint secures the FWI stability and efficiency.

3.2 Marmousi model

The second synthetic example uses a portion of the Marmousi 2 model (Martin et al., 2002). Figure 6 shows the correct V_P and V_S models. We simulate a multicomponent ocean bottom seismic survey (OBS) with a line of receiver at the water bottom ($z = 0.2$ km) and 100 evenly spaced pressure sources located at $z = 0.05$ km. Figure 7 shows the initial V_P and V_S models.

Without the stability penalty, the inversion becomes unstable after 3 iterations. The updated models are in Figure 8, showing that V_S has negative values, which prevent further updates. Then, conventional inversion is not able to produce long-wave- length updates which deteriorates the quality of the recovered models.

The stability penalty eliminates this problem. The recovered models after 50 iterations are shown in Figure 9. Even though the recovered V_P and V_S models are contaminated with high-wavenumber artifacts in addition to the artifacts due to the inherent interparameter cross-talk of EFWI, the inversion delivers long-wavelength updates and produces high resolution models. In order to further enhance the quality of the inverted models, one can also include petrophysical constraints or a regularization term in the objective function (Aragao and Sava, 2020a,b).

4 CONCLUSIONS

We develop a strategy for incorporating a stability constraint into elastic full waveform inversion through a penalty term based on a probability density function that defines the range of velocities for which wave propagation is stable. When the updated models are inside the stability region defined by the PDF, the inversion is fully controlled by the misfit between the observed and predicted data. We demonstrate that imposing this penalty avoids numerical dispersion and instabilities during the wavefield simulations and helps guide the inverted models towards an accurate solution. Through synthetic examples, we verify the effectiveness of the proposed methodology in leading inversion towards the true models, while preventing updates that hamper the accuracy and stability of the finite difference method.

5 ACKNOWLEDGMENTS

We acknowledge CAPES (Brazilian Federal Agency for Support and Evaluation of Graduate Education) for financial support. We thank the sponsors of the Center for Wave Phenomena (CWP) for additional support of this research. The synthetic examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

REFERENCES

Alford, R. M., K. R. Kelly, and D. M. Boore, 1974, Accuracy of finite-difference modeling of the acoustic wave equation: Geophysics, **39**, 834–842.

Aragao, O., and P. Sava, 2020a, Elastic full-waveform inversion probabilistic petrophysical model constraints: Geophysics, **85**, no. 2, R101–R111.

———, 2020b, Elastic full waveform inversion with probabilistic petrophysical clustering: Geophysical Prospecting, **68**, no. 4, 1341 – 1355.

Boore, D. M., 1972, Finite Difference Methods for Seismic Wave Propagation in Heterogeneous Materials: Methods in Computational Physics: Advances in Research and Applications, **11**, 1–37.

Courant, R., K. Friedrichs, and H. Lewis, 1967, On the partial difference equations of mathematical physics: IBM Journal of Research and Development, **11**, 215–234.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, no. 1, 54–66.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**.

Kelly, K. R., R. W. Ward, S. Treitel, and R. M. Alford, 1976, Synthetic seismograms: a finite-difference approach: Geophysics, **41**, 2–27.

Martin, G., K. Marfut, and S. Larsen, 2002, Marmousi-2, an updated model for the investigation of avo in structurally complex areas: SEG Technical Program Expanded Abstracts, 2478–2482.

Moczo, P., J. Kristek, and L. Halada, 2004, The finite-difference method for seismologists. an introduction: Comenius University.

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

Virieux, J., 1986, P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method: Geophysics, **51**, no. 4, 889–901.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26.

Wu, W. J., L. R. Lines, and H. X. Lu, 1996, Analysis of higher-order, finite-difference schemes in 3-D reverse-time migration: Geophysics, **61**, 845–956.

[!]

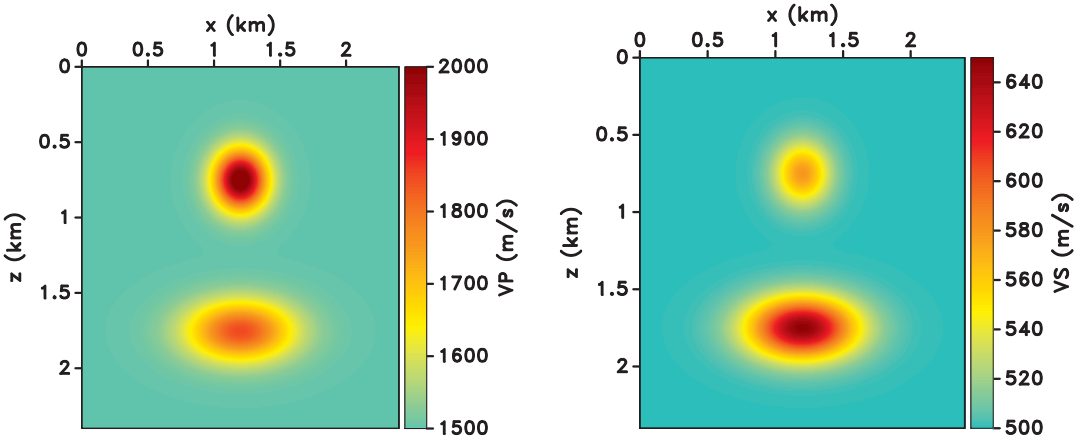
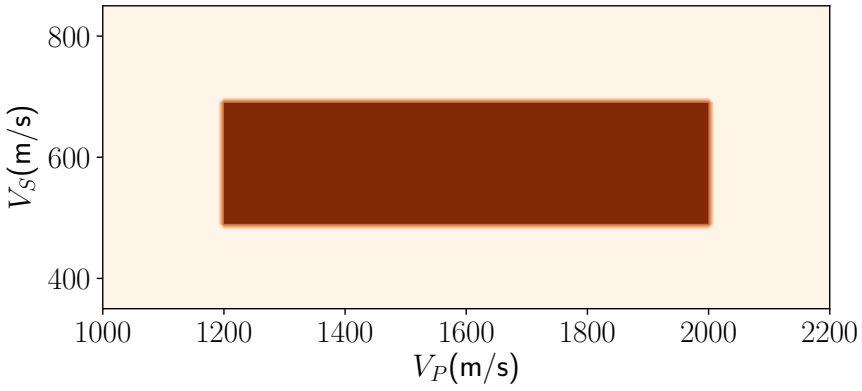
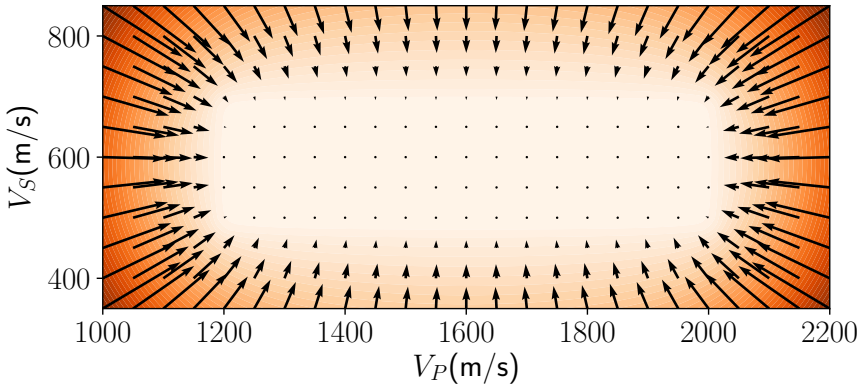


Figure 1. Synthetic models with two positive Gaussian anomalies for the V_P and V_S models centered at $(1.2, 0.75)$ km and $(1.2, 1.75)$ km.



(a)



(b)

Figure 2. (a) Example of a PDF consisting of boxcar function representing the range of velocities in which the forward simulations remain stable. (b) The distance $D(\mathbf{m})$ and its gradient. Models inside the boxcar function are subjected to zero gradients from the stability constraint, and thus inversion in this region is driven solely by the data misfit.

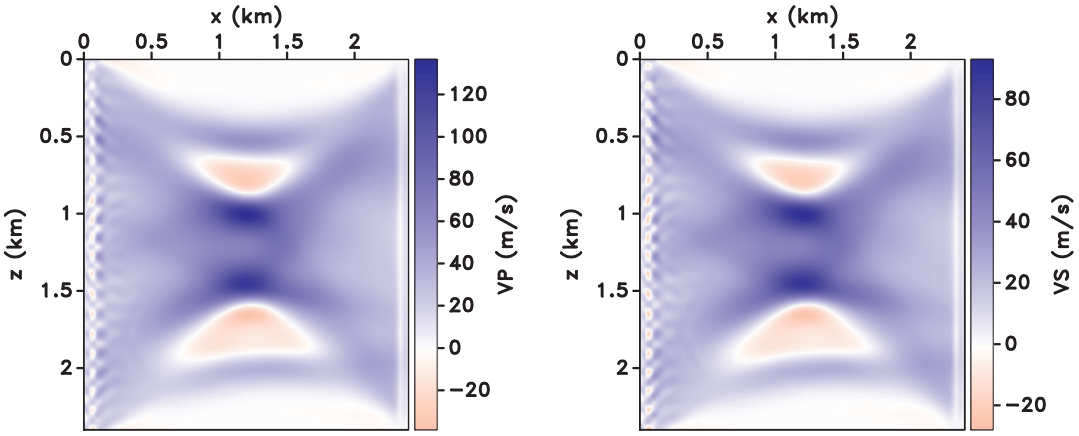


Figure 3. Inversion updates for the V_P and V_S models using the objective function $\mathcal{J}_D$ after 5 iterations. The V_S updates below -20 m/s cause inversion instabilities.

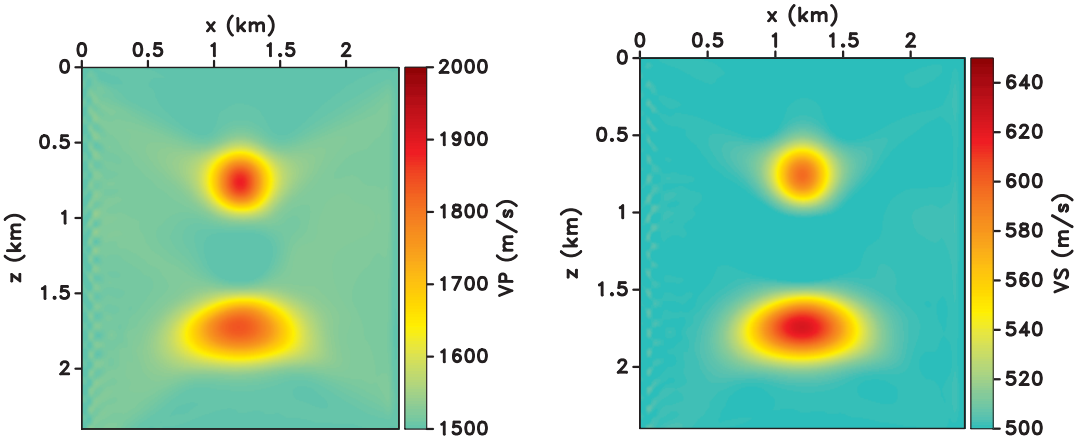


Figure 4. Recovered V_P and V_S models using the objective function $\mathcal{J}_D + \mathcal{J}_S$ after 50 iterations.

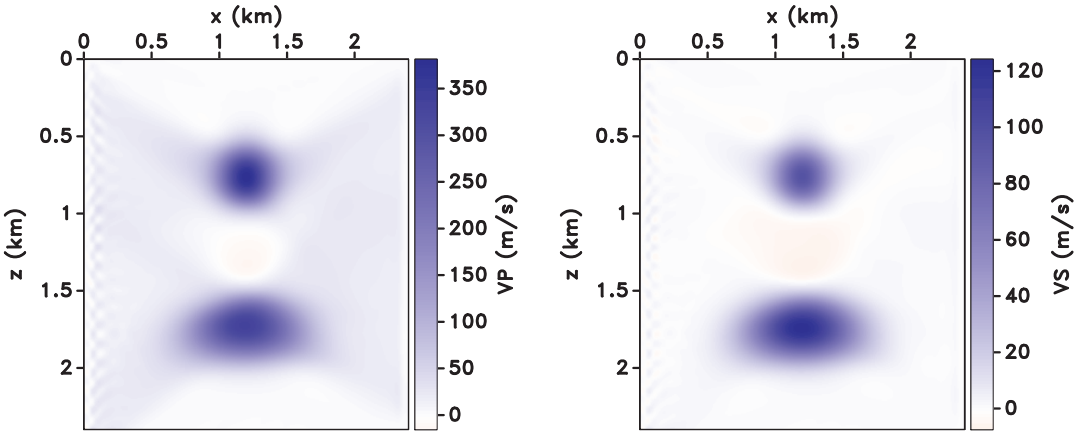


Figure 5. Inversion updates for the V_P and V_S models using the objective function $\mathcal{J}_D + \mathcal{J}_S$ after 50 iterations. Note that the inversion does not produce any updates below -20 m/s for the V_S model, thus avoiding modeling instability.

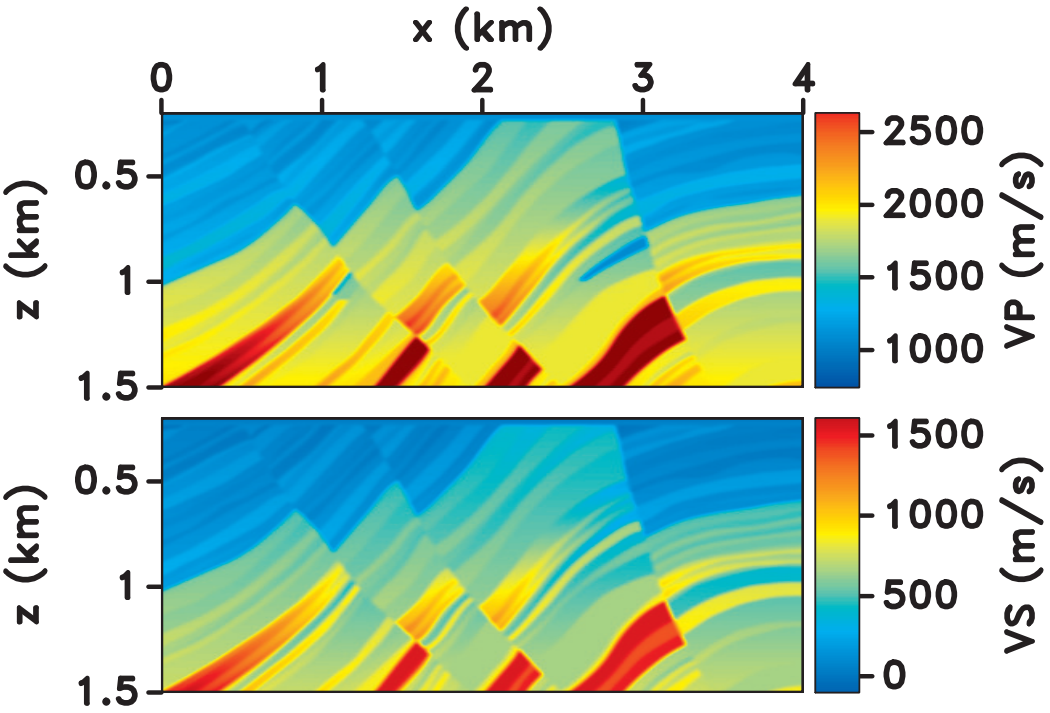


Figure 6. Marmousi 2 model: true V_P and V_S models.

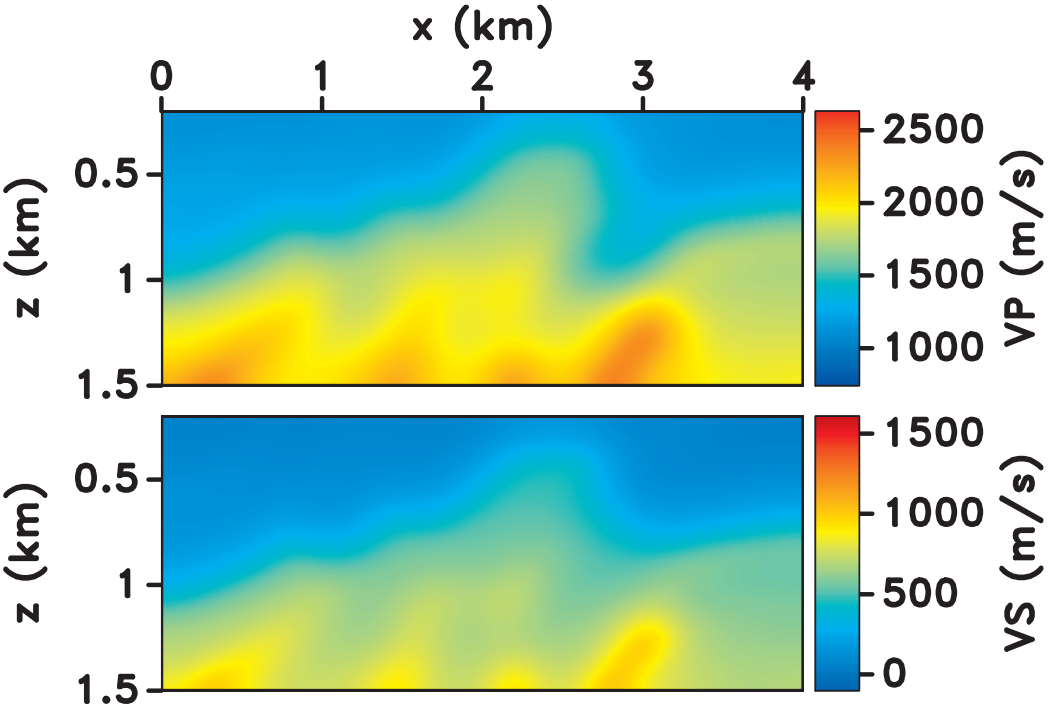


Figure 7. Marmousi 2 model: initial V_P and V_S models, obtained by smoothing the true models (Figure 6).

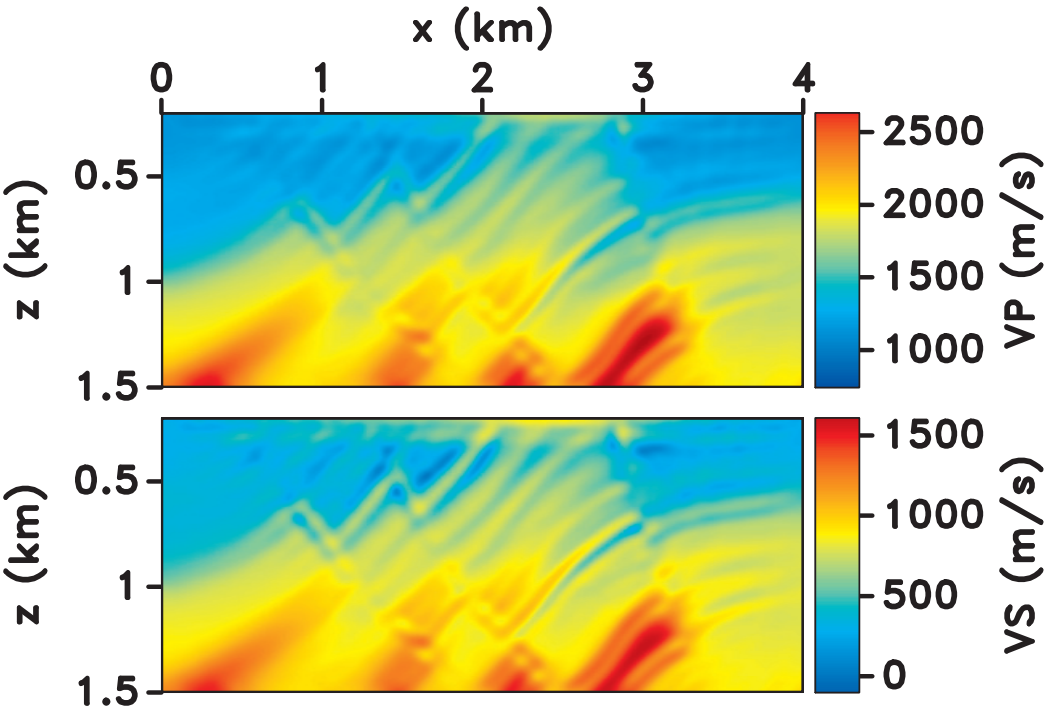


Figure 8. Marmousi 2 model: recovered V_P and V_S models using objective function $\mathcal{J}_D$. The inverted V_S model has negative values, which cause wavefield simulation instability.

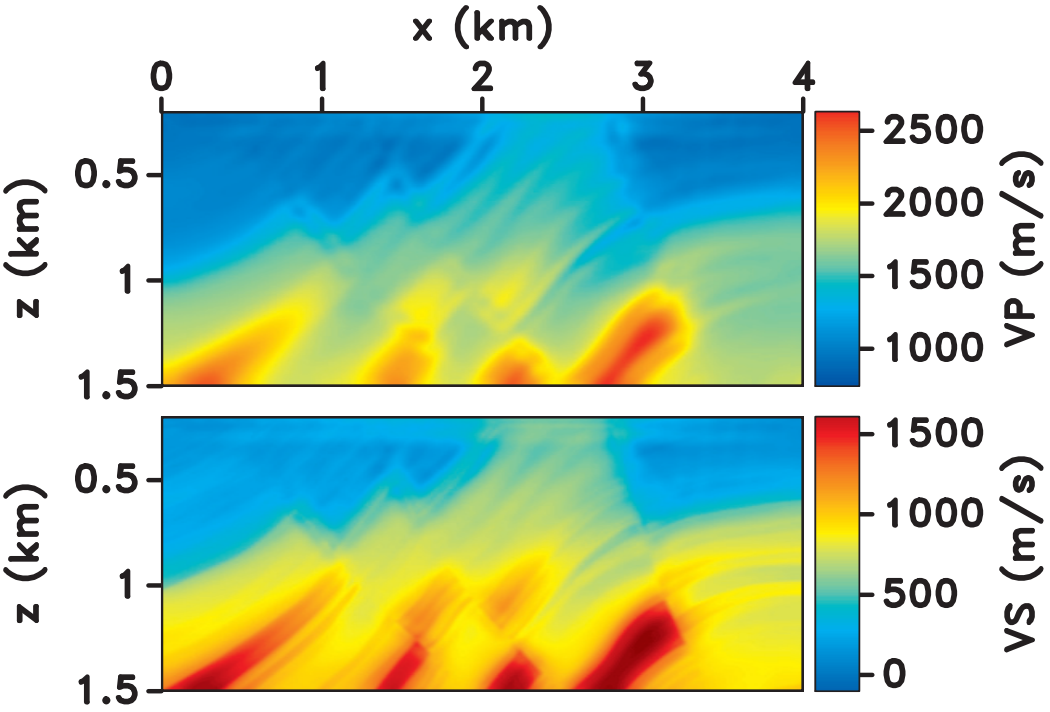


Figure 9. Marmousi 2 model: recovered V_P and V_S models using the objective function $\mathcal{J}_D + \mathcal{J}_S$. The stability penalty prevents negative V_S velocities, thus guaranteeing modeling stability.