Visco-acoustic full-waveform inversion in the space-frequency domain using the MUMPS direct solver

Apirujee Nakpathom, Jeff Shragge, and Chaiwoot Boonyasiriwat

ABSTRACT

Full-waveform inversion (FWI) has been demonstrated to be an effective technique to recover the property of the subsurface based on mathematically minimizing the difference between observed and calculated data from full-wavefield modeling. In this study, we present the FWI based on visco-acoustic wave modeling in frequency domain to reconstruct P-wave velocity (V_p) and P-wave attenuation factor (Q). The two difference algorithms are investigated and applied to both cross-hole and onshore data.

INTRODUCTION

In seismic exploration, seismic waves were generated by artificial energy sources, and travel through the earth's subsurface. The waves reflect at boundaries where there are contrasts in acoustic impedance. These reflected and other waves are recorded by receivers on the surface - seismic data.

Full waveform inversion (FWI) is a method to estimate the subsurface property by matching the observed data and predicted data which are generated from a modeling process based on the visco-acoustic wave equation. Visco-acoustic FWI can estimate both the velocity and attenuation factor of the subsurface. Attenuation is a mechanism that decreases the amplitude of seismic waves propagating through the medium. The real earth attenuates waves due to the conversion of wave energy into heat. This an-elastic behavior decease amplitude of the wave field and thus can have effect on FWI by reducing the resolution of the estimated model. So it is important to compensate for an-elastic behavior to make the estimated results more reliable. Attenuation properties of the subsurface can be expressed as quality factor (Q), a dimensionless value that involves with wave velocity as the complex velocity.

Visco-acoustic wave field that propagate through the earth can be generate by applying finite difference method (FDM) to acoustic wave equation, in this case, in frequency-space domain (FD). There are the advantages of the FDFDM wave modeling approach. For example, FDFDM modeling can be specified as a discrete matrix formulation, so it is possible to derive matrix formulas for the iterative non-linear inverse method, i.e. steepest descent method and Gauss-Newton method. It also allows us to sequential reconstruct model parameters from low to high frequency component of the data. Different from time domain inversion methods that use all frequency of the data, frequency domain methods use only few frequencies to invert the accurate model. By using the low frequency estimated model as a starting model for later inversions, the chance of convergence to a local minimum is somewhat reduced. Furthermore, the attenuation can be included at no increased computational cost by straight forward using the complex velocity. In contrast, time domain methods require convolution operation with a response function to model visco-acoustic effects. Another advantage is the simultaneously multisource modeling by using LU decomposition. Use of LU decomposition in FDFDM allows the fast calculation of the wavefield due to multiple sources once an initial factorization step is completed.

In this study, we developed the modeling and inversion code. The code is written in Fortran 90 and uses Message Passing Interface (MPI) for parallelism. We also use the MUlti frontal Massively Parallel direct Solver (MUMPS) for solving LU factorization.

In theory section, we present the mathematical method used in visco-acoustic modeling and inversion. The two-difference minimization method, i.e. steepest descent and Quasi-Newton, will be introduced along with the parameter classed. The FWI algorithm is also presented as flow chart in this section.

In synthetic results section, we investigated both minimization schemes together with two-difference parameter classes by testing with the simple boxes model. From previous synthetic study, we found some drawback of the

slowness parameter class. So we investigated further by using quasi-Newton method with v-q parameter class with both cross-hole and onshore synthetic data. Then we present numerical results from FWI.

THEORY

Visco-acoustic Wave Modeling in Frequency Domain

From the two-dimensional acoustic wave equation in space-time domain,

$$\frac{1}{c(x,z)^2}\frac{\partial^2}{\partial t^2}p(x,z,t)+\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right]p(x,z,t)=s(x,z,t) \tag{1}$$

where $p(x,z,t)$ is the pressure wavefield, $s(x,z,t)$ is source, and $c(x,z)$ is acoustic wave velocity. We apply Fourier transform to equation (1), and then we obtain

$$\frac{\omega^2}{c(x,z)^2}P(x,z,\omega)+\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right]P(x,z,\omega)=S(x,z,\omega) \tag{2}$$

Above equation is the two-dimensional acoustic wave equation in the space-frequency domain. Where $P(x,z,\omega)$ and $S(x,z,\omega)$ is the pressure wavefield and source in the space-frequency domain. ω is the angular frequency.

One of the advantage of working in the frequency domain is we can directly implement of the attenuation property through the relation

$$\tilde{c}=c\left[1+\frac{i}{2Q}\right]^{-1} \tag{3}$$

where $\tilde{c}$ is the complex velocity and Q is the attenuation factor.

In order to solve Eq. (2), we discretize equation (2) using an implicit finite difference scheme based on 2^{nd} accurate centered difference operators with constant grid spacing in both x and z directions. The centered difference operator for the second derivative can be written as,

$$\frac{\partial^2 P_{i,j}}{\partial x^2}=\frac{P_{i-1,j}-2P_{i,j}+P_{i+1,j}}{\triangle x^2} \tag{4}$$

where $\triangle x$ is the grid spacing. After substituting the difference operator (4) into our wave equation (2), we obtain,

$$\begin{aligned}\frac{\omega^2}{{c_{i,j}}^2}P_{i,j}+&\left[\frac{P_{i-1,j}-2P_{i,j}+P_{i+1,j}}{\triangle x^2}\right]\\&+\left[\frac{P_{i,j-1}-2P_{i,j}+P_{i,j+1}}{\triangle z^2}\right]=S_{i,j}\end{aligned} \tag{5}$$

At the boundary of the computational domain, we use the Engquist-Majda absorbing boundary condition,

$$\frac{\partial P}{\partial n}-i\frac{\omega}{c}P=0 \tag{6}$$

Which can be discretize by using 1^{st} order difference operator as,

$$\frac{P_{i,1}-P_{i,2}}{\triangle z}-i\frac{\omega}{c_{i,1}}P_{i,1}=0 \tag{7}$$

For the case of the top boundary where $j=1$.

From the discretization, acoustic wave equation (2) can be written in a matrix form as,

$$\mathbf{A}(\mathbf{m})\mathbf{p}=\mathbf{s} \tag{8}$$

Where $\mathbf{A}$ is the sparse differential operators matrix with dimension $N\times N$, where $N=nz\times nx$, depends on frequency ω and on model properties $\mathbf{m}$ that includes both velocity and attenuation variation. Two-dimensional pressure $\mathbf{p}$ and source $\mathbf{s}$ fields at frequency ω are stored as vectors of dimension N.

The 2D pressure field is obtained by solving the system of linear equation, eq.8. In this study, MUMPS (MUltifrontal Massively Parallel sparse direct Solver) is used to solve the equation. Therefore, multiple sources' solutions can be obtained once the matrix $\mathbf{A}$ was factorized using a $\mathbf{LU}$ decomposition scheme:

$$\mathbf{LU}\left[\mathbf{p}_1\mathbf{p}_2\dots\mathbf{p}_{ns}\right]=\left[\mathbf{s}_1\mathbf{s}_2\dots\mathbf{s}_{ns}\right] \tag{9}$$

This is the other advantage of working in the frequency domain, because we can simulation the wavefield from multi-source simultaneously.

Full-waveform Inversion Algorithm

The least-squares misfit function that is minimized is given by

$$E(\mathbf{m})=\frac{1}{2}\triangle\mathbf{d}^{\dagger}\triangle\mathbf{d} \tag{10}$$

where $\triangle\mathbf{d}$ is the data residual vector, $\triangle\mathbf{d}=\mathbf{d}_{obs}-\mathbf{d}_{cal}(\mathbf{m})$, the difference between observed and calculated data. The symbol † denotes the complex conjugate operator.

From a starting model $\mathbf{m}_0$, we are going to search for a local minimum of the misfit function $E(\mathbf{m}_0)$ by iterative nonlinear local optimization. By following the Born approximation (Born and Wolf (1993); Beydoun and Tarantola (1988)), The estimated model $\mathbf{m}$ can be written as $\mathbf{m}=\mathbf{m}_0+\triangle\mathbf{m}$, where $\triangle\mathbf{m}$ is a perturbation model.

The misfit function of model $\mathbf{m}$, $E(\mathbf{m})$, can be expanded by Taylor series as,

$$\begin{aligned}E(\mathbf{m}_0+\triangle\mathbf{m})=E(\mathbf{m}_0)&+\sum_{j=1}^{M}\frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial m_j}\triangle m_j\\&+\frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M}\frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial m_j\partial m_k}\triangle m_j\triangle m_k+\mathcal{O}(\|\triangle m\|^3)\end{aligned} \tag{11}$$

Where M denotes the number of inverted parameter. The first order derivative respect to the model parameter m_l can be written as,

$$\frac{\partial \mathrm{E}(\mathbf{m})}{\partial m_l} = \frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial m_l} + \sum_{j=1}^{M} \frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial m_j \partial m_l} \triangle m_j \qquad (12)$$

The misfit function reaches its minimum when the first order derivative equal zeros. Therefore, the left hand term must be set to zero,

$$0 = \frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial m_l} + \sum_{j=1}^{M} \frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial m_j \partial m_l} \triangle m_j \qquad (13)$$

And implies this expression respect to model parameter $\mathbf{m}$,

$$-\frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}} = \frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}^2} \triangle \mathbf{m} \qquad (14)$$

Hence, the perturbation model:

$$\triangle \mathbf{m} = -\left[\frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}^2}\right]^{-1} \frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}} \qquad (15)$$

where $\frac{\partial \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}}$ is called the gradient $\nabla \mathrm{E}(\mathbf{m}_0)$ and $\frac{\partial^2 \mathrm{E}(\mathbf{m}_0)}{\partial \mathbf{m}^2}$ the Hessian $\mathcal{H}$.

Gradient expression

For multiple sources and frequencies, The gradient of the misfit function for the model parameter m_j is computed with the adjoint-state method (Chavent (1974); Tarantola (1984); Plessix (2006); Chavent (2009)), which gives

$$\begin{aligned} \nabla E_{m_j} &= -\sum_{k=1}^{N_\omega} \sum_{l=1}^{N_s} \mathbb{R}\left[\mathbf{p}_{l,k}^T \left(\frac{\partial \mathbf{A}_k}{\partial m_j}\right)^T \mathbf{A}_k^{-1} \triangle \mathbf{d}_{l,k}^*\right] \\ &= -\sum_{k=1}^{N_\omega} \sum_{l=1}^{N_s} \mathbb{R}\left[\mathbf{p}_{l,k}^T \left(\frac{\partial \mathbf{A}_k}{\partial m_j}\right)^T \mathbf{r}_{l,k}^*\right] \end{aligned} \qquad (16)$$

where ^{T} denotes the transpose operator, * the complex conjugate,N_s the number of sources, and N_ω the number of frequencies simultaneously inverted, in this case, $N_\omega = 1$. $\mathbb{R}$ defines the real part of the complex value $\mathbf{p}_{l,k}$ is the monochromatic incident wavefield associated with frequency k and source l. $\mathbf{r}_{l,k}$ is the back-propagated residual wavefield. Note that the residuals associated with one source are assembled to from one vector.

Hessian expression

The Hessian reads

$$\mathcal{H} = \mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right] + \mathbb{R}\left[\left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}}\right)^T (\delta \mathbf{d}^* \cdots \delta \mathbf{d}^*)\right] \qquad (17)$$

where $\mathbf{J}$ is the Frechet derivative or the sensitivity matrix. The first term $\mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right]$ is called the approximate Hessian. It is the zero-lag correlation between the partial derivative of wavefields with respect to different parameters. Therefore, it represents the spatial correlation between the images of different point scatterers. It can be view as a resolution operator resulting from limited bandwidth of the source and the acquisition geometry. Indeed, applying the inverse of the Hessian is equivalent to applying a spiking deconvolution of the gradient misfit function. (e.g. Ali et al. (2009))

The term $\mathbb{R}[\mathbf{J}^T \mathbf{J}^*]$ is diagonal dominant since the diagonal terms are defined by zero-lag auto-correlation. This diagonal term reduces the effects of the geometrical spreading. Therefore, in the frame of surface acquisition, it helps to scale the deep perturbations (large offsets / small amplitudes) with respect to the shallow perturbations (near offsets / high amplitudes)

The second term $\mathbb{R}\left[\left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}}\right)^T (\delta \mathbf{d}^* \cdots \delta \mathbf{d}^*)\right]$ is the zeros-lag correlation between the second-order partial derivative of the wavefields with data residuals. Since first-order partial derivative os related to single scattering, it can be expected that second-order partial derivative is related to double or multiple scattering.

The perturbation model (descent direction) equation (15) reads

$$\triangle \mathbf{m} = \left\{\mathbb{R}\left[\mathbf{J}^T \mathbf{J}^* + \left(\frac{\partial \mathbf{J}}{\partial \mathbf{m}}\right)^T (\delta \mathbf{d}^* \ldots \delta \mathbf{d}^*)\right]\right\}^{-1} \nabla E_{\mathbf{m}} \qquad (18)$$

This expression is generally referred as the Newton method, which is locally quadratic convergence.

Generally, the second term of Hessian is neglected since in the framework of the Born approximation multiple scattering are neglected (Pratt et al. (1998)). This leads to a quasi-Newton direction called Gauss-Newton,

$$\triangle \mathbf{m} = \left\{\mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right]\right\}^{-1} \nabla E_{\mathbf{m}} \qquad (19)$$

If the Hessian is replaced by a scalar α, the expression gives the steepest descent direction,

$$\triangle \mathbf{m} = \alpha \nabla E_{\mathbf{m}} \qquad (20)$$

Consider the approximate Hessian $\mathcal{H}_a$ (Pratt et al. (1998)) defined by

$$\mathcal{H}_a = \mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right] \qquad (21)$$

The element of the Frechet derivative matrix associated with the source-receiver pair k and the parameter m_j is given by (Ali et al. (2009); Malinowski et al. (2011))

$$\mathbf{J}_{k(s,r),l} = \mathbf{p}_s^T \left(\frac{\partial \mathbf{A}}{\partial m_j}\right)^T \mathbf{A}^{-1} \delta_r \qquad (22)$$

where δ_r is an impulse source located at the receiver position r. As shown by the equation above, we have to simulate one forward problem for each source and each receiver pair. This means that the computational cost of the approximate Hessian depends on the acquisition geometry and number of sources and receivers. In the

huge acquisitions, the large space of the storage must be required for the approximate Hessian.

To mitigate the problem, only diagonal terms of the approximate hessian are computed at first iteration of the inverted frequency (Ravaut et al. (2004); Operto et al. (2006)).

To update the approximate Hessian or its inverse $\mathcal{H}^{(i)}$ at each iteration of inversion, the BFGS algorithm is applied by taking into account the additional knowledge of $\nabla E^{(i)}$ at iteration i. The BFGS formula for the inverse of the quasi-Hessian is given by,

$$\mathcal{H}^{-1}_{(i+1)} = \left(\mathbf{I} - \frac{\mathbf{s}_i\mathbf{y}_i^T}{\mathbf{y}_i^T\mathbf{s}_i}\right)\mathcal{H}^{-1}_{(i)}\left(\mathbf{I} - \frac{\mathbf{y}_i\mathbf{s}_i^T}{\mathbf{y}_i^T\mathbf{s}_i}\right) + \frac{\mathbf{s}_i\mathbf{s}_i^T}{\mathbf{y}_i^T\mathbf{s}_i} \quad (23)$$

where $\mathbf{s}_i = \mathbf{m}_i - \mathbf{m}_{(i-1)}$ and $\mathbf{y}_i = \nabla E_{(i)} - \nabla E_{(i-1)}$. In this study, we used only the diagonal part of the hessian.

Model parameters

From the complex velocity equation (3), we define the complex slowness as,

$$\begin{aligned} \tilde{s} &= \frac{1}{\tilde{c}} \\ \tilde{s} &= s\left(1 + \frac{i}{2Q}\right) \end{aligned} \quad (24)$$

and

$$\begin{aligned} c &= \frac{1}{s^{(r)}} \\ Q &= \frac{s^{(r)}}{2s^{(i)}} \end{aligned} \quad (25)$$

where $s_r = \mathbb{R}[\tilde{s}]$ and $s^{(i)} = \mathbb{I}[\tilde{s}]$. $\mathbb{R}$ and $\mathbb{I}$ refer to real and imaginary part of complex value respectively.

Now, we assume two real-value parameter classes as $\mathbf{s}^{(r)}$ and $\mathbf{s}_i$ which are the real and imaginary component of complex slowness $\tilde{\mathbf{s}} = \mathbf{s}^{(r)} + i\mathbf{s}^{(i)}$. Both parameters have N element, where $N = nz \times nx$. Then, the model parameters can be written as

$$\mathbf{m} = \begin{bmatrix} \mathbf{s}^{(r)} \\ \mathbf{s}^{(i)} \end{bmatrix} \quad (26)$$

which the number of model parameter is $2N$.

The gradient is partition into two vectors as (Kennett and Sambridge (1998); Williamson (1990))

$$\nabla_{\mathbf{m}} E = \begin{bmatrix} \nabla_{\mathbf{s}^{(i)}} E \\ \nabla_{\mathbf{s}^{(r)}} E \end{bmatrix} \quad (27)$$

Following Kamei and Pratt (2013), we precondition the model gradient by pre-multiplication,

$$\nabla_{\mathbf{m}} E = \begin{bmatrix} \nabla_{\mathbf{s}^{(i)}} E \\ \lambda\nabla_{\mathbf{s}^{(r)}} E \end{bmatrix} \quad (28)$$

λ is a scalar that scales and balances two gradient classes, and is referred to as an 'attenuation penalty term' which illustrates that $\lambda \neq 1$ performs an implicit regularization: The condition $\lambda < 1$ penalizes the attenuation component of the model gradient, and $\lambda > 1$ penalizes the velocity component of the model gradient.

RESULTS

The acquisition geometry of the synthetic data includes

- Model size: 101×201 grid point ($nz = 101$ and $nx = 201$) with grid spacing 5 m.
- 39 sources locate at 25 m depth and 25 m spacing.
- 191 receivers with 475 m depth and 5 m spacing.
- True model: two simple boxes model as shown in table 1
- Initial model: homogeneous as shown in table 1
- 15 iterations are preformed at each inverted frequency.

Comparison between sfinversion01 and sfinversion02 algorithm

In this test, 17-frequency components between 3 and 20 Hz are fully inverted. We used the algorithms that present in Figure 1 and 2, and had results as shown in table 1. From the results:

- Velocity shape: Both method give an accurate position of boxes.
- Velocity amplitude: Quasi-Newton give a better amplitude.
- Attenuation shape: Quasi-Newton give a better shape.
- Attenuation amplitude: Quasi-Newton with complex slowness parameter model give an inaccurate amplitude of the right box that should be less than background. There are the stronger coupling in this parameter class. Steepest descent with v-q parameter model give a very low attenuation amplitude of the left box (about Q=20) but it give a correct trend of the right box.
- Steepest attenuation result gives obviously artifacts at the source position. By using Hessian in Quasi-Newton method, the artifacts is decreasing.

Cross-hole and onshore results from Quasi-Newton and v-q model parameter

In this test, In case of onshore, receivers locate at 25 m depth. Frequency components between 1 and 30 Hz are fully inverted. We used the algorithms that present in Figure 2 but instead of converting complex slowness, v-q parameters are inverted, and had results as shown in table 2.

Realistic Model Example

In this experiment, we use the more realistic Q and velocity models used in Kamei and Pratt (2013) with the permission from Rie Kamei. Figure 3 shows the experimental set up of this crosswell test. There are 57 sources

located in a well and 281 receivers located in another well. The true velocity and Q models are shown in Figures 4 and 5. Figures 8 and 7 show the initial models used in thise experiment. The estimated models are shown in Figures 8 and 9. The inversion results from visco-acoustic full-waveform inversion show very promising results of both velocity and Q estimation. In the future, we will apply the inversion to field data.

SUMMARY

Both velocity and attenuation can be reconstructed by FWI without the constrain of maximum and minimum value of the model. But in order to get the better resolution, we have to apply the precondition and regularization to the algorithm which are the next step of this work.

ACKNOWLEDGMENTS

The authors are grateful for the support from PTTEP and would like to thank Rie Kamei for allowing us to use the realistic models for our experiment.

REFERENCES

Ali, H. B. H., S. Operto, and J. Virieux, 2009, Efficient 3d frequency-domain full waveform inversion (fwi) with phase encoding, *in* 71st EAGE Conference & Exhibition, SPE and EAGE.

Beydoun, W. and A. Tarantola, 1988, First born and rytov approximations: Modeling and inversion conditions in a canonical example: Journal of the Acoustical Society of America, **83**.

Born, M. and E. Wolf, 1993, Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light; sixth edition with corrections: Pergamon Press.

Chavent, G., 1974, Identification of functional parameter in partial differential equations: Identification of Parameters in Distributed Systems, 31–48.

———, 2009, Non linear least squares for inverse problems. Scientific Computation: Springer.

Kamei, R. and R. Pratt, 2013, Inversion strategies for visco-acoustic waveform inversion: Geophysical Journal International, **194**, 859–884.

Kennett, B. L. N. and M. Sambridge, 1998, Inversion for multiple parameter classes: Geophysical Journal International, **135**, 304–306.

Malinowski, M., S. Operto, and A. Ribodetti, 2011, High-resolution seismic attenuation imageing from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion: Geophysical Journal International, **186**, 1179–1204.

Operto, S., J. Virieux, J.-X. Dessa, and G. Pascal, 2006, Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern nankai trough: Journal of Geophysical Research: Solid Earth, **111**.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

Pratt, R., C. Shin, and Hicks, 1998, Gauss–newton and full newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: Geophysical Journal International, **159**, 1032–1056.

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

Williamson, P., 1990, Tomographic inversion in reflection seismology: Geophysical Journal International, **100**, 255–274.

Algorithm

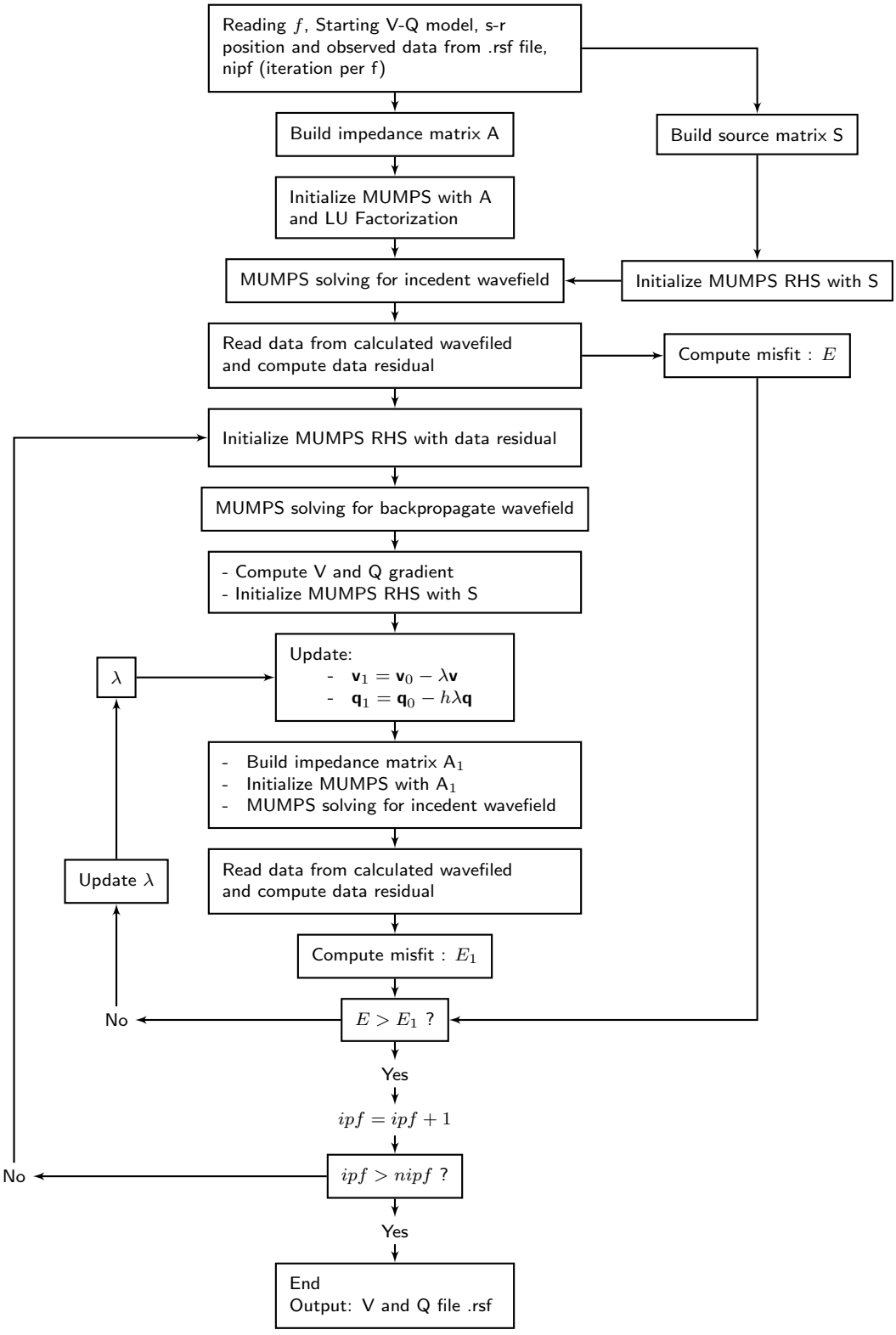


Figure 1: F90-Madagascar Inversion Code Flow Chart: Code named "**sfinversion01**"

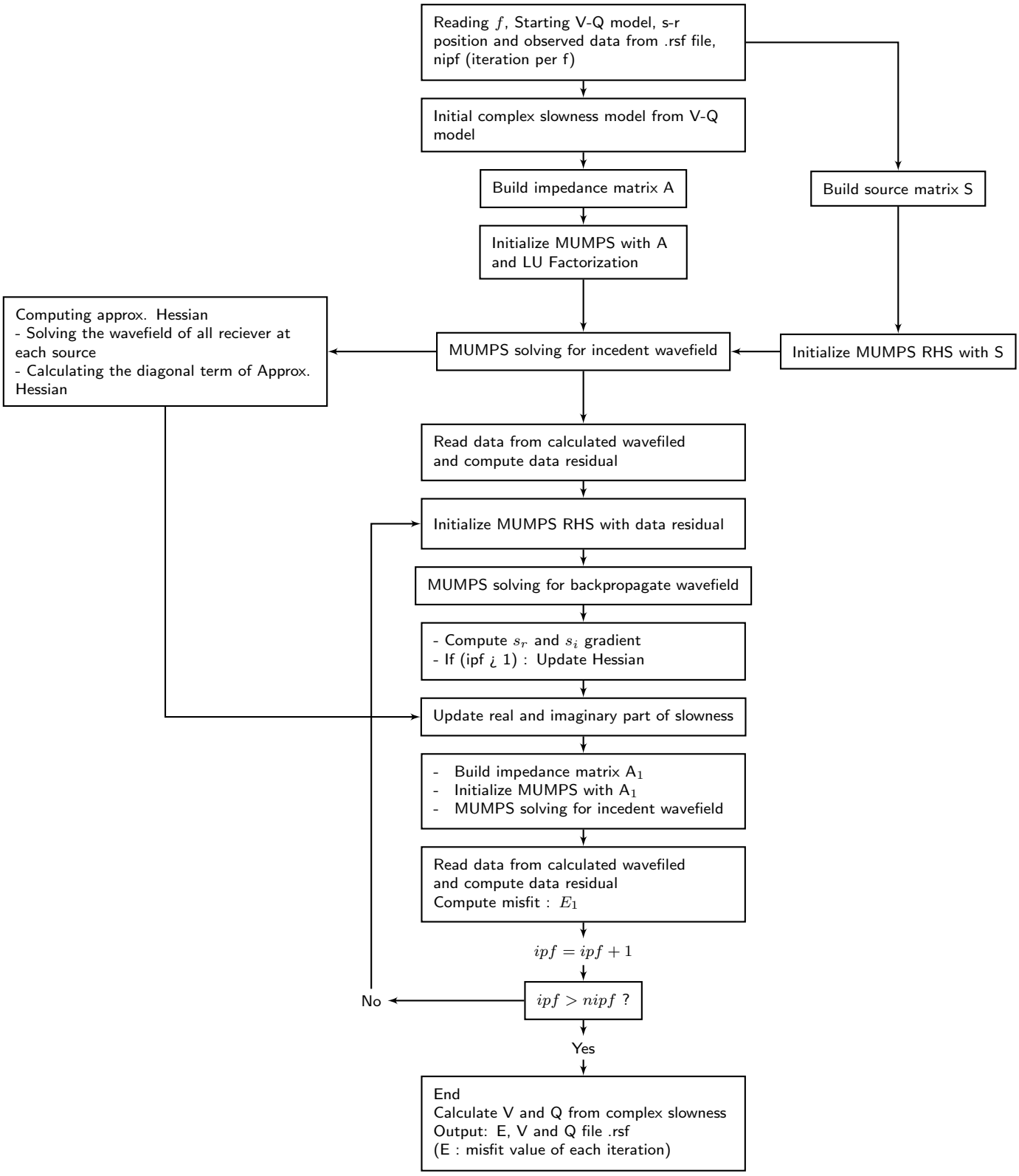


Figure 2: F90-Madagascar Inversion Code Flow Chart: Code named "**sfinversion02**"

Table 1: Comparing between Steepest descent with v-q parameter and quasi-Newton with complex slowness parameter

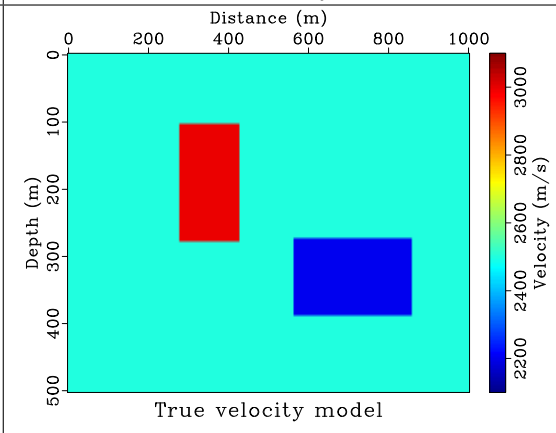
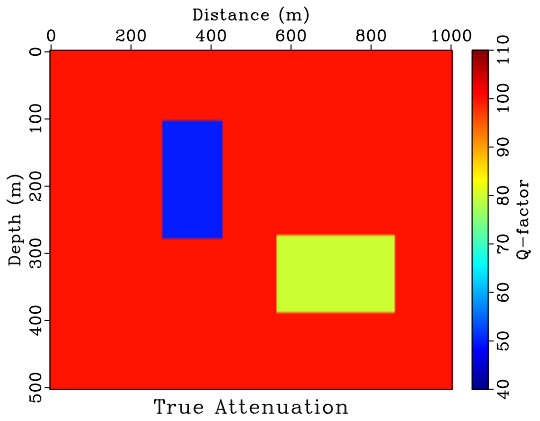
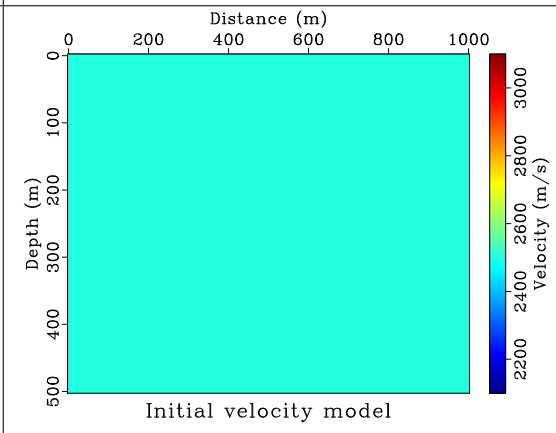
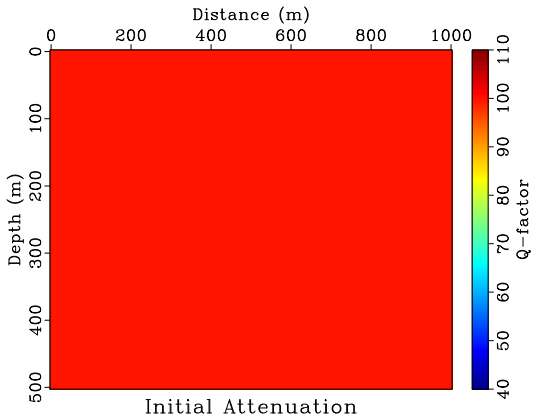
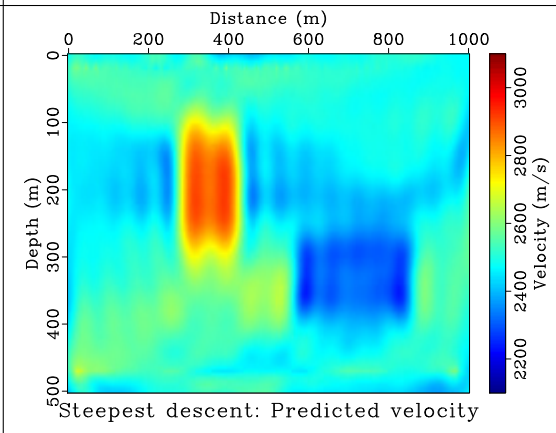
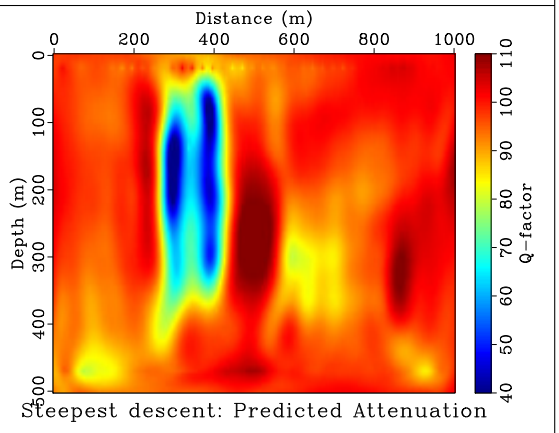
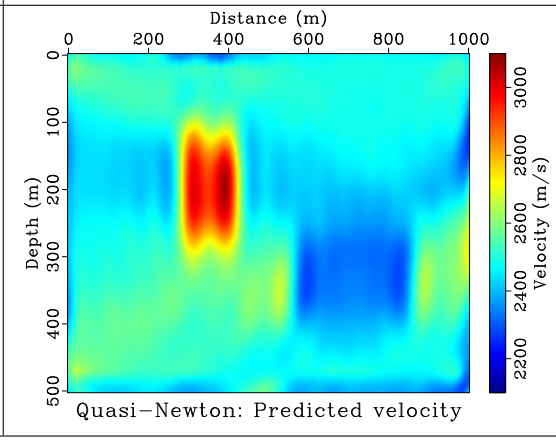
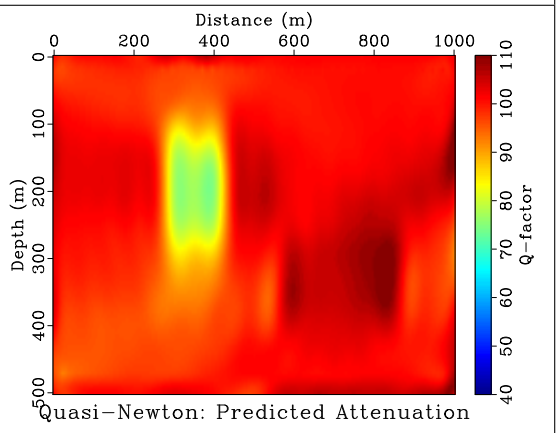
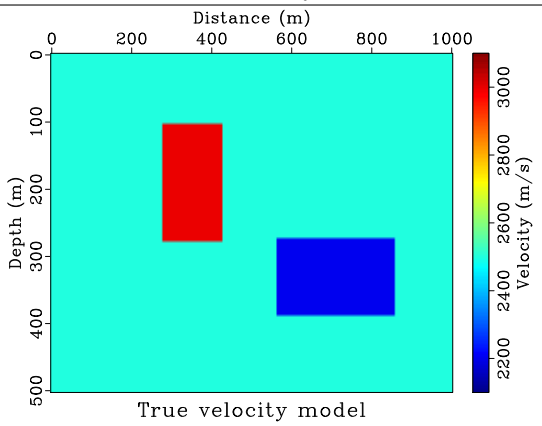
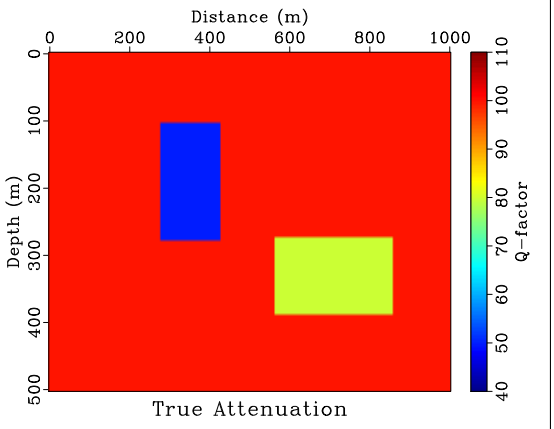
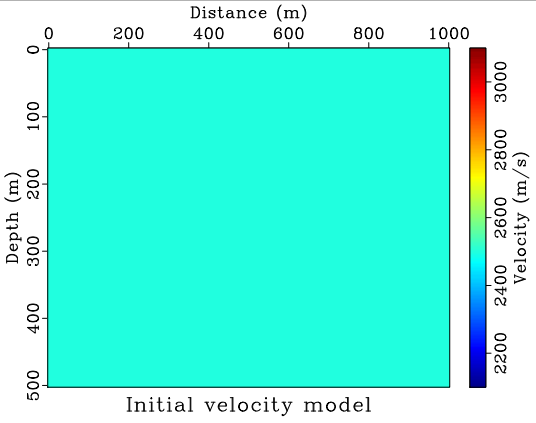
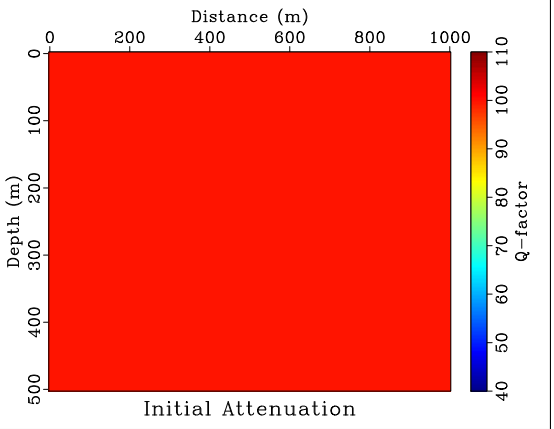
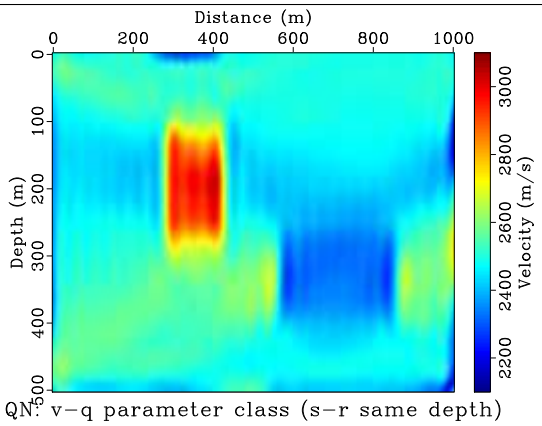
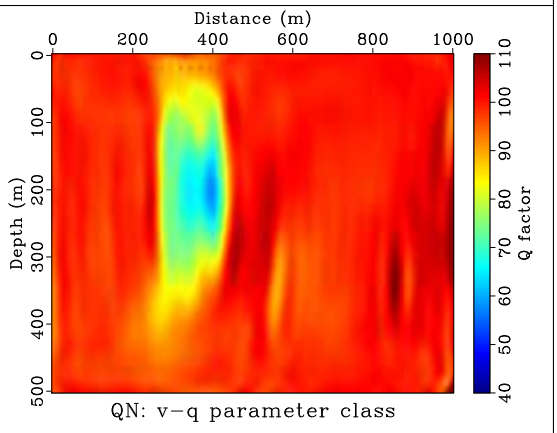
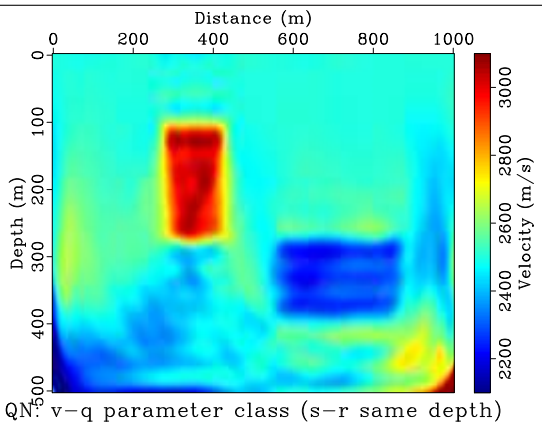
Detail	Velocity	Attenuation
True model		
Initial model		
Steepest descent		
quasi-Newton		

Table 2: FWI result cross-hole and onshore data

Detail	Velocity	Attenuation
True model		
Initial model		
Cross-hole		
Onshore		

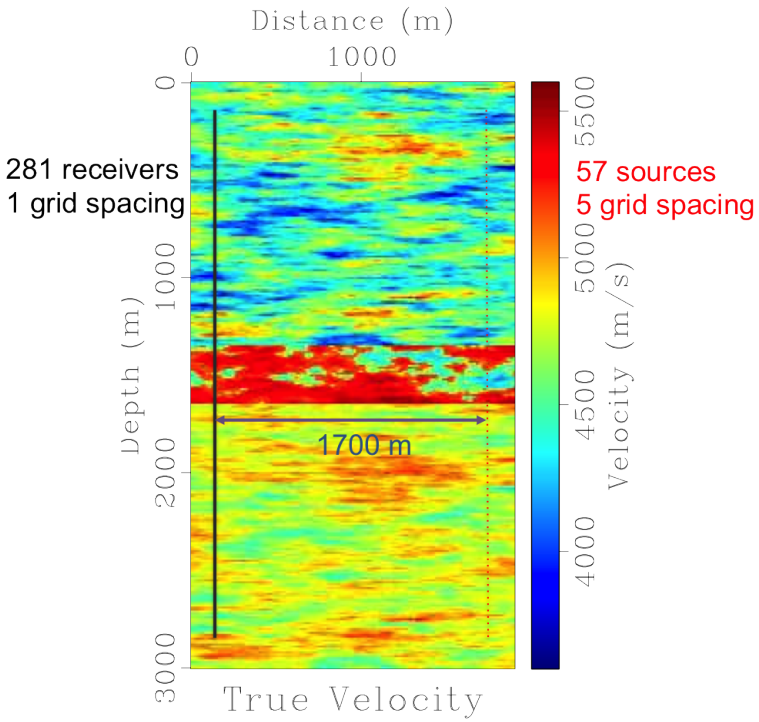


Figure 3: The experimental set up of the realistic model experiment.

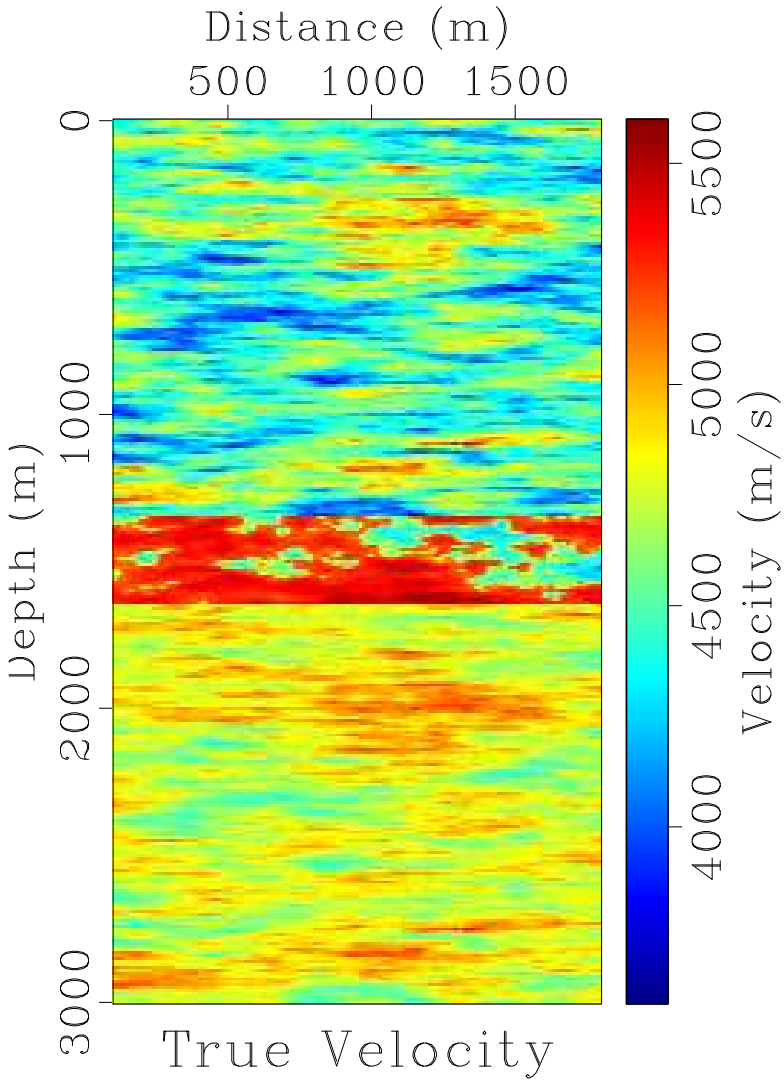


Figure 4: True velocity model of the Kamei-Pratt's model.

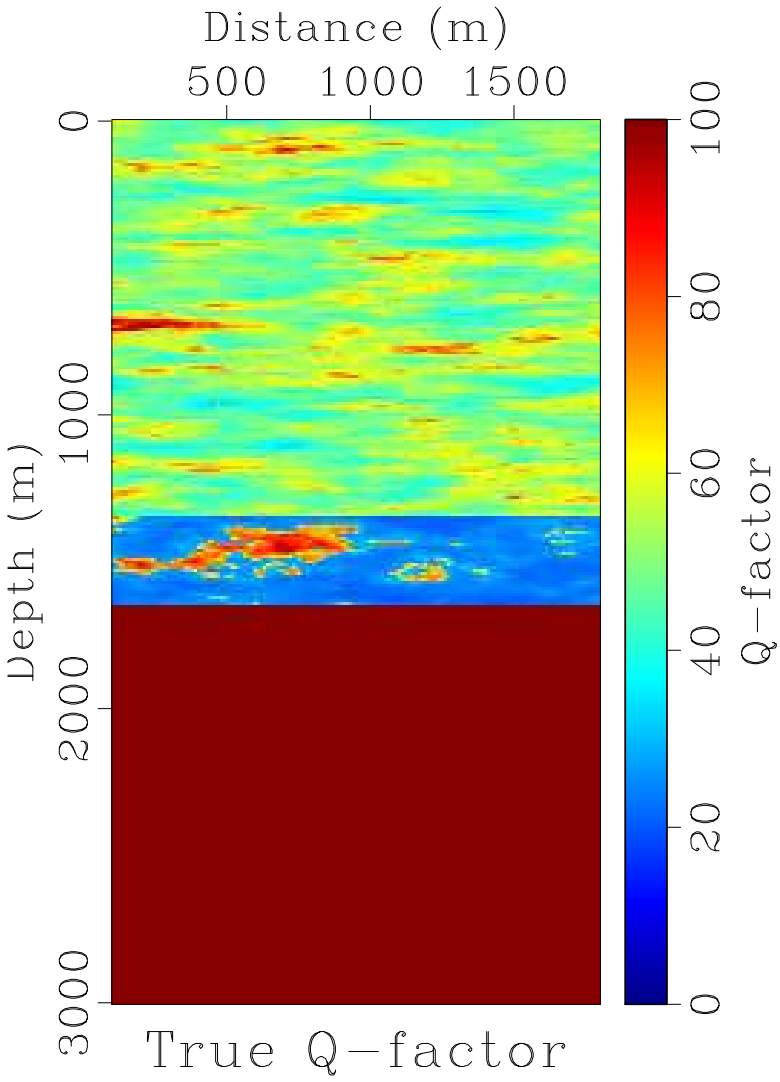


Figure 5: True Q model of the Kamei-Pratt's model.

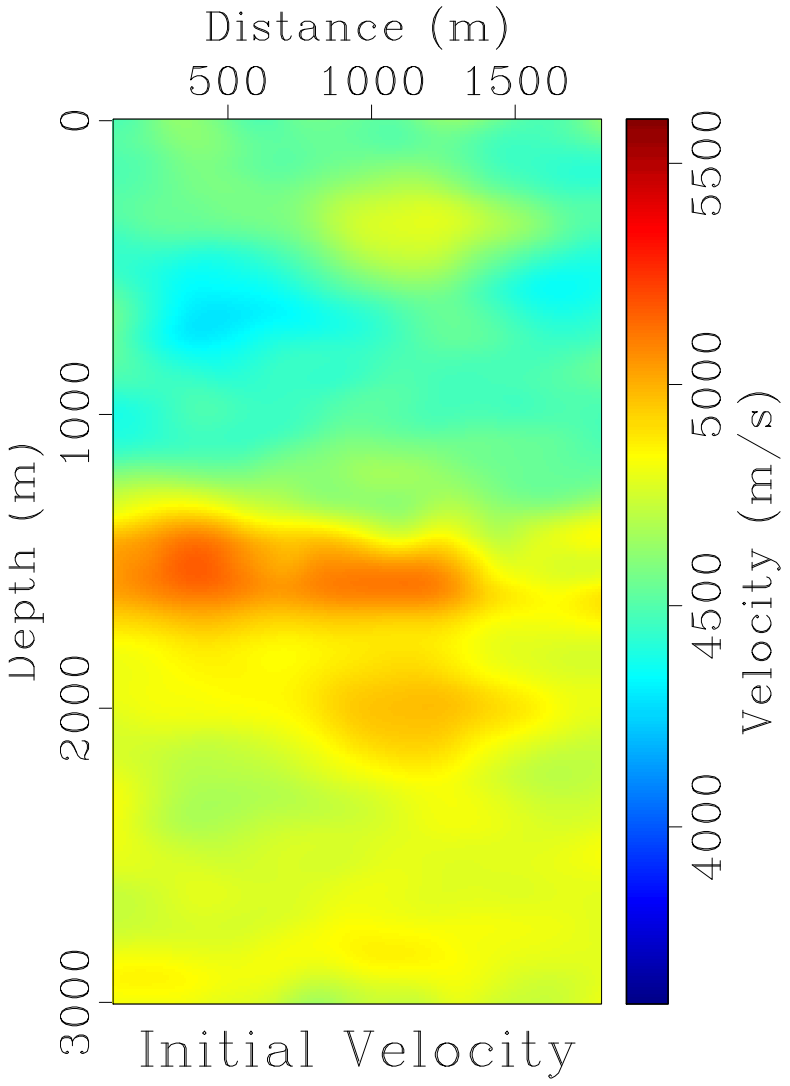


Figure 6: Initial velocity model for the realistic model experiment.

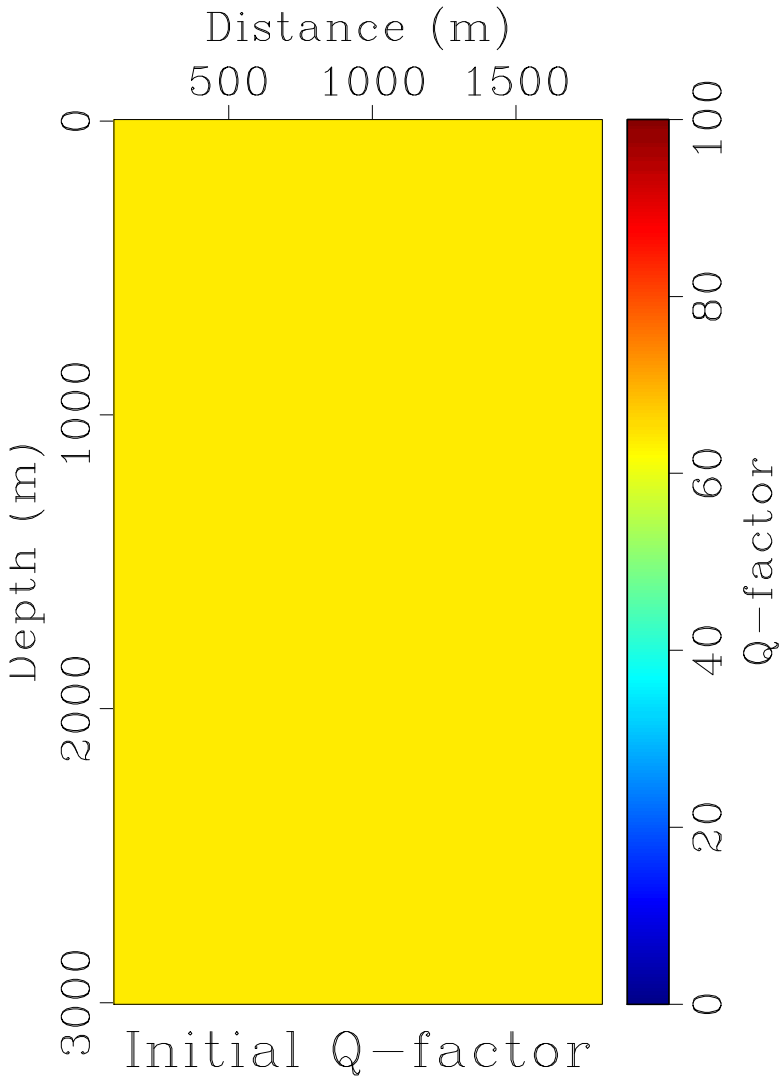


Figure 7: Initial Q model for the realistic model experiment.

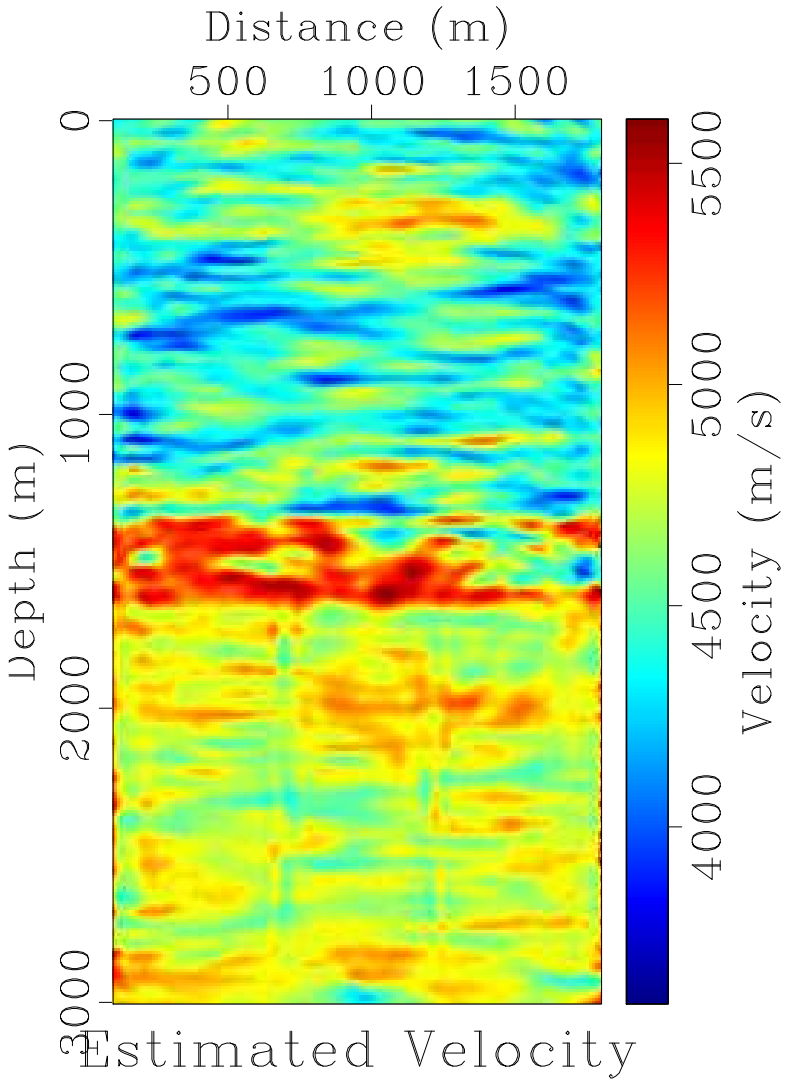


Figure 8: Estimated velocity model for the realistic model experiment.

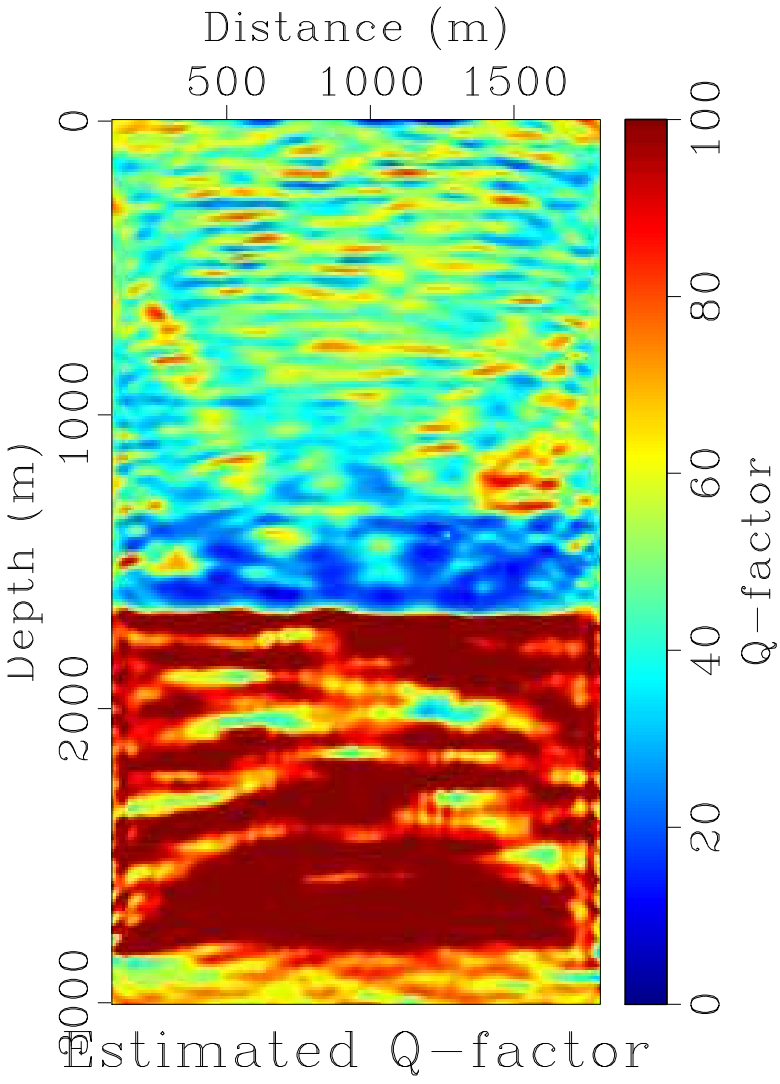


Figure 9: Estimated Q model for the realistic model experiment.