Visco-acoustic full-waveform inversion in the space-frequency domain

Apirujee Nakpathom and Chaiwoot Boonyasiriwat

ABSTRACT

Full waveform inversion (FWI) is a method for estimating the physical properties of the Earth's subsurface structures from recorded geophysical data. Typically, FWI estimates a model of subsurface properties in an iterative manner by minimizing the misfit between predicted and recorded data. The predicted data are generated from a modeling process based on the visco-acoustic wave equation in the frequency-space domain. This is equivalent to solving a nonlinear inverse problem using an optimization technique. In this work, we use the L-BFGS quasi-Newton method for optimization, and develop a Fortran90 implementation of visco-acoustic FWI to estimate seismic velocity and quality factor (Q-factor) of the subsurface structure. The code was validated by the synthetic data experiment with the velocity and Q-factor models from Kamei and Pratt (2013). For the models, 2D elastic data were synthetically generated in the time domain using the Sofi2D. The frequency-domain data were then inverted for velocity and Q-factor models by visco-acoustic FWI. The numerical results show that velocity and Q-factor can be estimated from this algorithm, and the estimation is better for velocity than for Q-factor. In the future, we will conduct experiments on real seismic 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 quality 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 as the subprogram in Madagascar and uses Message Passing Interface (MPI) and OpenMP 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. The FWI algorithm is also presented as flow chart in this section.

In synthetic results section, we investigated the minimization scheme by testing with the complex model from Kamei and Pratt 2013. The synthetic data are generated with 3 difference method such as a visco-acoustic finite difference in frequency domain (same algorithm to calculate the culcated data in the inversion process), a visco-acoustic finite difference in time domain and a visco-elastic finite difference in time domain. We use the external package (Sofi2D) to generate the last two sets of data. the results show that the algorithm can estimate both velocity and quality factor from all data set, but velocity estimated form visco-elastic data set has a lower resolution compare with another data set, because it contain the s-wave in data that will be classified a noise.

In the results from real field data,

THEORY

Visco-acoustic Wave Modeling in Frequency Domain

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

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

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

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

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

$$\tilde{c}=c\left[1+i2Q\right]^{-1} \tag{3}$$

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 4^{th} order finite difference method. At the boundary of the computational domain, we use the perfectly matched to absorb wave energy at outside the computational domain.

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

$$\mathbf{A}(\mathbf{m})\mathbf{p}=\mathbf{s} \tag{4}$$

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:

$$\mathbf{LU}\left[\mathbf{p}_1\mathbf{p}_2\ldots\mathbf{p}_{ns}\right]=\left[\mathbf{s}_1\mathbf{s}_2\ldots\mathbf{s}_{ns}\right] \tag{5}$$

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

$$E(\mathbf{m})=12\triangle\mathbf{d}^\dagger\triangle\mathbf{d} \tag{6}$$

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,

$$E(\mathbf{m}_0+\triangle\mathbf{m})=E(\mathbf{m}_0)+\sum_{j=1}^{M}\partial\mathrm{E}(\mathbf{m}_0)\partial m_j\triangle m_j+12\sum_{j=1}^{M}\sum_{k=1}^{M}\partial^2\mathrm{E}(\mathbf{m}_0)\partial \tag{7}$$

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

$$\partial\mathrm{E}(\mathbf{m})\partial m_l=\partial\mathrm{E}(\mathbf{m}_0)\partial m_l+\sum_{j=1}^{M}\partial^2\mathrm{E}(\mathbf{m}_0)\partial m_j\partial m_l\triangle m_j \tag{8}$$

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

$$0 = \partial \mathrm{E}(\mathbf{m}_0)\partial m_l + \sum_{j=1}^{M} \partial^2 \mathrm{E}(\mathbf{m}_0)\partial m_j \partial m_l \triangle m_j \quad (9)$$

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

$$-\partial \mathrm{E}(\mathbf{m}_0)\partial \mathbf{m} = \partial^2 \mathrm{E}(\mathbf{m}_0)\partial \mathbf{m}^2 \triangle \mathbf{m} \quad (10)$$

Hence, the perturbation model:

$$\triangle \mathbf{m} = -\left[\partial^2 \mathrm{E}(\mathbf{m}_0)\partial \mathbf{m}^2\right]^{-1} \partial \mathrm{E}(\mathbf{m}_0)\partial \mathbf{m} \quad (11)$$

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

$$\nabla E_{m_j} = -\sum_{k=1}^{N_\omega} \sum_{l=1}^{N_s} \mathbb{R}\left[\mathbf{p}_{l,k}^T \left(\partial \mathbf{A}_k \partial m_j\right)^T \mathbf{A}_k^{-1} \triangle \mathbf{d}_{l,k}^*\right] = -\sum_{k=1}^{N_\omega} \sum_{l=1}^{N_s} \mathbb{R}\left[\mathbf{p}_{l,k}^T \left(\partial \mathbf{A}_k \partial m_j\right)^T \mathbf{r}_{l,k}^*\right] \quad (12)$$

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

$$\mathcal{H} = \mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right] + \mathbb{R}\left[(\partial \mathbf{J} \partial \mathbf{m})^T (\delta \mathbf{d}^* \cdots \delta \mathbf{d}^*)\right] \quad (13)$$

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

$$\triangle \mathbf{m} = \left\{\mathbb{R}\left[\mathbf{J}^T \mathbf{J}^* + (\partial \mathbf{J} \partial \mathbf{m})^T (\delta \mathbf{d}^* \ldots \delta \mathbf{d}^*)\right]\right\}^{-1} \nabla E_{\mathbf{m}} \quad (14)$$

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,

$$\triangle \mathbf{m} = \left\{\mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right]\right\}^{-1} \nabla E_{\mathbf{m}} \quad (15)$$

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

$$\triangle \mathbf{m} = \alpha \nabla E_{\mathbf{m}} \quad (16)$$

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

$$\mathcal{H}_a = \mathbb{R}\left[\mathbf{J}^T \mathbf{J}^*\right] \quad (17)$$

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))

$$\mathbf{J}_{k(s,r),l} = \mathbf{p}_s^T \left(\partial \mathbf{A} \partial m_j\right)^T \mathbf{A}^{-1} \delta_r \quad (18)$$

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,

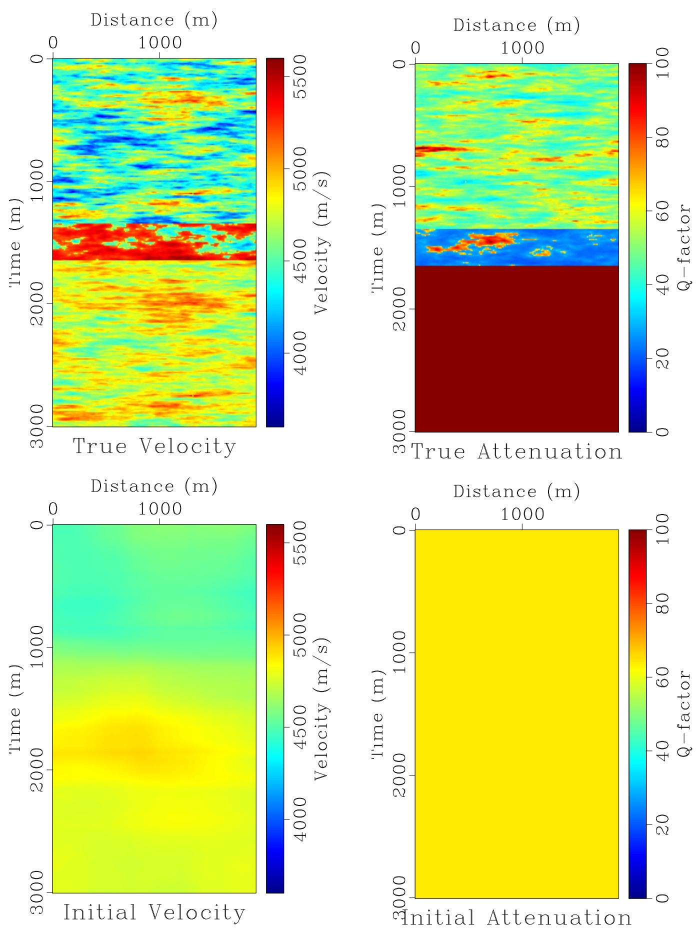
$$\mathcal{H}_{(i+1)}^{-1} = \left(\mathbf{I} - \mathbf{s}_i \mathbf{y}_i^T \mathbf{y}_i^T \mathbf{s}_i\right) \mathcal{H}_{(i)}^{-1} \left(\mathbf{I} - \mathbf{y}_i \mathbf{s}_i^T \mathbf{y}_i^T \mathbf{s}_i\right) + \mathbf{s}_i \mathbf{s}_i^T \mathbf{y}_i^T \mathbf{s}_i \tag{19}$$

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.

RESULTS

Results from Synthetic data

In synthetic experiment, we use the velocity and Q-factor models from Kamei and Pratt (2013). There models are interesting in that 3 velocity zone and non-corresponding 3 Q-factor zone. There are also a lot of detail in this model that challenge us to recover the detail.



The acquisition geometry is shown in 1. From the initial model, we inverted to find the estimated model with 24 frequency components start from 3 to 49 Hz with 2 Hz interval.

We generated 3 observed data sets from the same true model. The first one is the visco-acoustic data from frequency-domain finite difference which is a same modelling algorithm using in the inversion process. The other 2 types are the visco-acoustic and visco-elastic data that obtain from time-domain finite difference modelling from Sofi2D package.

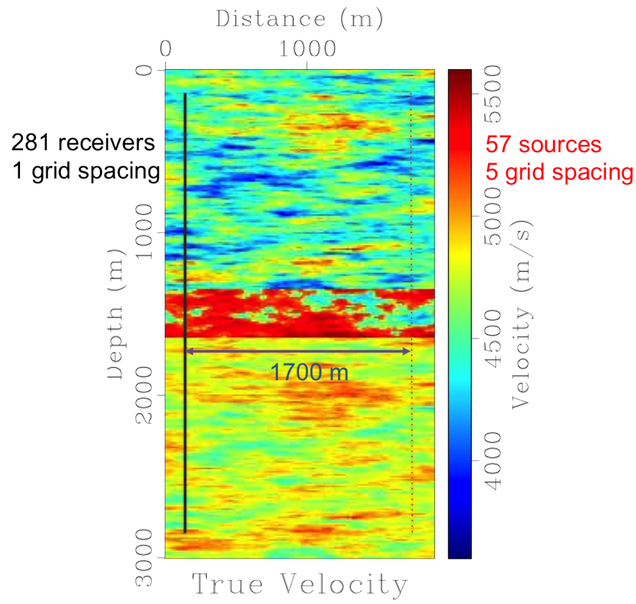


Figure 1: Acquisition

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 1.

Results from Field data

In this test, we applied the algorithm with the field data set. There are 14 total shot with the hand correct trace used. The results show in table 2.

SUMMARY

Both velocity and attenuation can be reconstructed by FWI without the constrain of maximum and minimum value of the model. From the estimated model form synthetic data, we can see that,not only the structure, but also the value of velocity and Q-factor were accurately estimated. Even though there are the shear wave component in the visco-elastic data, we still can find the rough structure of the model accurately.

ACKNOWLEDGMENTS

The authors are grateful for the support from PTTEP for the Mahidol University Center for Scientific Computing (MCSC).

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.

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.

F90-MADAGASCAR VISCO-ACOUSTIC FWI CODE FLOWCHART

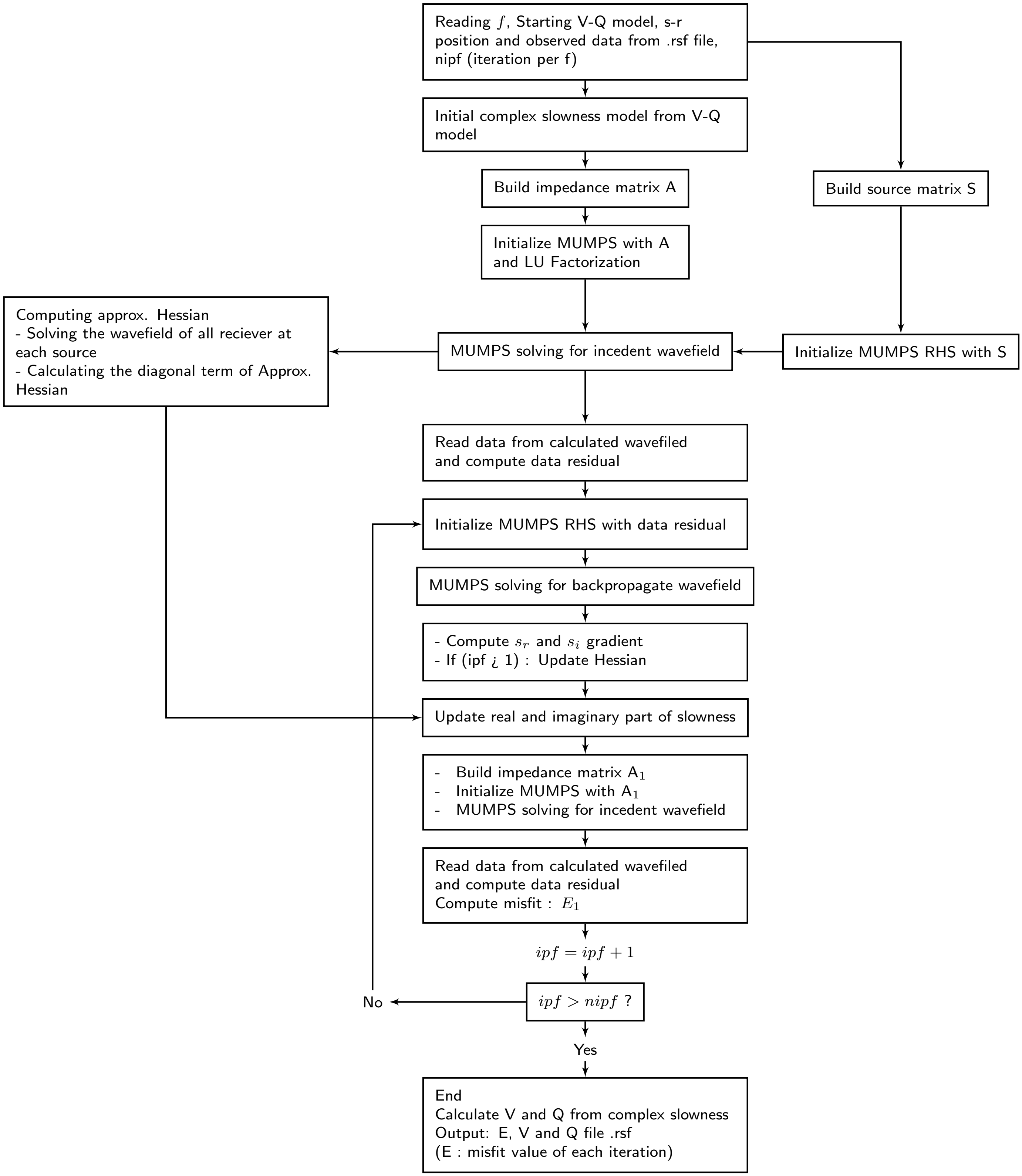
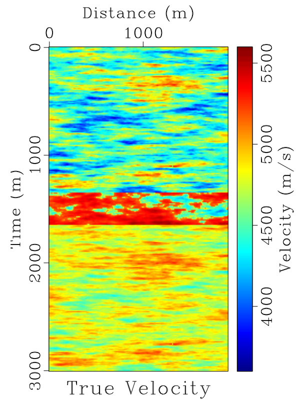
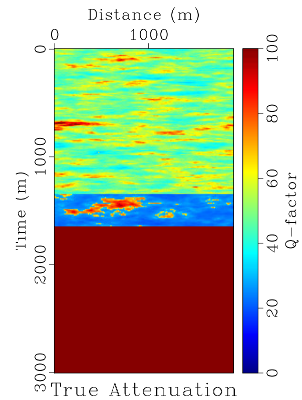
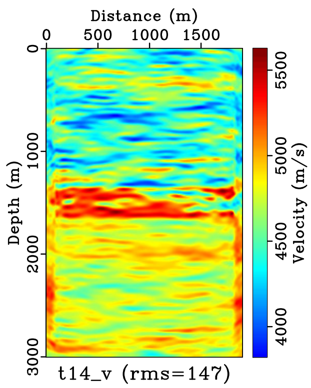


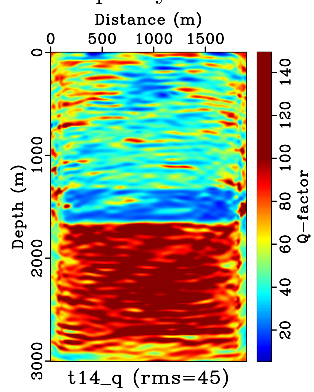
Figure 2: F90-Madagascar Inversion Code Flow Chart

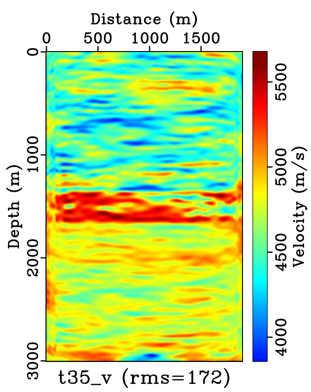
Table 1: Comparing between Steepest descent with v-q parameter and quasi-Newton with complex slowness parameter

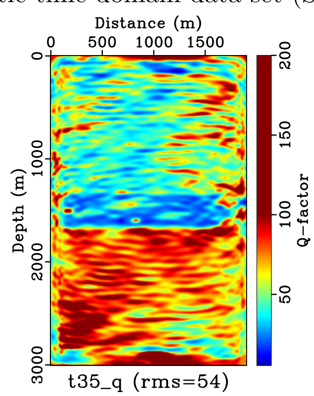
<table>
<tr><th>Velocity</th><th>Attenuation</th></tr>
<tr><td colspan="2">True model</td></tr>
<tr><td>


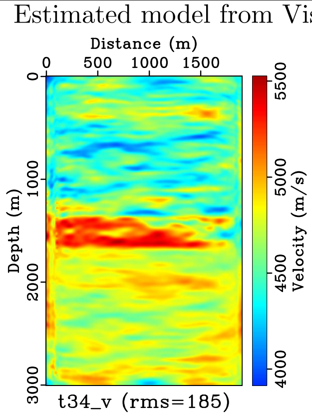
</td><td>


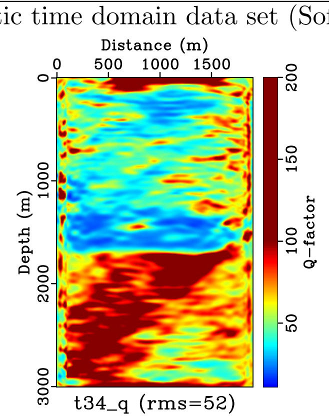
</td></tr>
<tr><td colspan="2">Estimated model from Visco-acoustic frequency domain data set</td></tr>
<tr><td>


</td><td>


</td></tr>
<tr><td colspan="2">Estimated model from Visco-acoustic time domain data set (Sofi2D)</td></tr>
<tr><td>


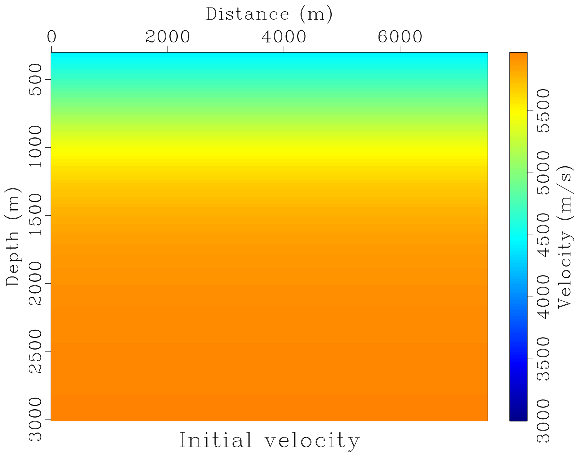
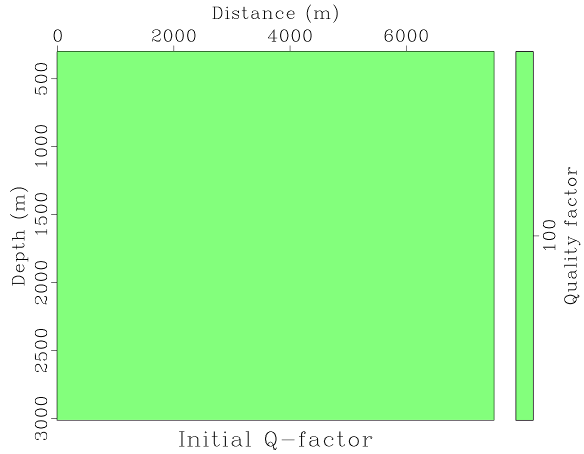
</td><td>


</td></tr>
<tr><td colspan="2">Estimated model from Visco-elastic time domain data set (Sofi2D)</td></tr>
<tr><td>


</td><td>


</td></tr>
</table>

Table 2: Comparing between Steepest descent with v-q parameter and quasi-Newton with complex slowness parameter

Velocity	Attenuation
Initial model	
	
Estimated model