Wave-equation traveltime tomography by global optimization

Ali Ameen Almomin

ABSTRACT

Wave-equation traveltime tomography is conventionally done by picking the maximum cross-correlation lags between the modeled and observed data. However, a trace-by-trace method of picking makes the velocity update more susceptible to local noise in the correlation as well as inconsistencies in the data. In this paper, I compare the local method of picking the maximum correlation to a global method based on maximizing the stacking power along an interpolated spline surface in the correlation window. The results show that the global scheme is more robust to local noise but sacrifices accuracy and convergence rate.

INTRODUCTION

Conventional full waveform inversion (FWI), which was first introduced by Tarantola (1984), has an objective function that is highly non-linear. The forward operator is linearized around the background velocity, which makes the initial model a determining factor for the convergence of the inversion. Therefore, a lot of previous work (Luo and Schuster, 1990; Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Biondi, 2009) has focused on finding more tractable objective functions that have stronger dependence on the kinematics of the wavefield than on the amplitude of the waveform. One attractive method that uses such an objective function is wave-equation traveltime inversion (WT), which was first introduced by Luo and Schuster (1990). In this inversion, the objective function depends on the lag of maximum cross-correlation between the observed and modeled data. Conventionally, these lags are picked in a trace-by-trace scheme, which produce errors due to correlating noise, multiple events, and inconsistencies in the observed data. To overcome this problem, I cast the picking procedure as a global optimization problem in order to avoid local errors by making use of the redundancy of the data.

THEORY

The objective function of FWI can be written as follows:

$$\mathbf{J}_{FWI}(\mathbf{v}) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} ||\mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}) - \mathbf{d}_{obs}(t, \mathbf{x}_g, \mathbf{x}_s)||_2^2, \quad (1)$$

where $\mathbf{x}_s$ and x_r are the source and receiver locations, d_{cal} is the modeled data with velocity $\mathbf{v}$, and d_{obs} is the observed data. By setting the first derivative of equation (1) around the velocity $\mathbf{v}_0$ to zero, the velocity update can be expressed as follows:

$$\Delta \mathbf{v} = \frac{-s}{\mathbf{v}_0^3} \sum_{\mathbf{x}_s} \sum_t \mathbf{L}(\mathbf{v}_0) \mathbf{S}(\mathbf{x}_s, t) \frac{\partial^2}{\partial t^2} \mathbf{L}^\dagger(\mathbf{v}_0)(\mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}) - \mathbf{d}_{obs}(t, \mathbf{x}_g, \mathbf{x}_s)), \qquad (2)$$

where s is the step size, $\mathbf{S}$ is the source signature, and $\mathbf{L}$ and $\mathbf{L}^\dagger$ are the forward wave propagation operator and its adjoint, respectively.

The objective function wave-equation traveltime tomography can be written as follows:

$$\mathbf{J}_{\Delta\tau}(\mathbf{v}) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} ||\Delta\tau(\mathbf{x}_g, \mathbf{x}_s; \mathbf{v})||_2^2, \qquad (3)$$

where $\Delta\tau$ is the lag of the maximum cross-correlation between the observed data and the data modeled by a velocity model $\mathbf{v}$. Again, the first derivative of equation (3) around the lags $\Delta\tau$ is set to zero to get the velocity update, which can be expressed as follows:

$$\Delta \mathbf{v} = \frac{s}{\mathbf{v}_0^3} \sum_{\mathbf{x}_s} \sum_t \mathbf{L}(\mathbf{v}_0) \mathbf{S}(\mathbf{x}_s, t) \frac{\partial^2}{\partial t^2} \mathbf{L}^\dagger(\mathbf{v}_0) \frac{\Delta\tau}{\xi} \frac{\partial}{\partial t} \mathbf{d}_{obs}(t + \Delta\tau, \mathbf{x}_g, \mathbf{x}_s), \qquad (4)$$

where ξ is defined as follows:

$$\xi = \sum_t \frac{\partial}{\partial t} \mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}_0) \frac{\partial}{\partial t} \mathbf{d}_{obs}(t + \Delta\tau, \mathbf{x}_g, \mathbf{x}_s), \qquad (5)$$

By examining equations (2) and (4), it can be shown that (WT) can handle much larger velocity errors than (FWI).

Now, I cast the picking procedure of the lags $\Delta\tau$ as a global optimization problem with an objective function as follows:

$$\mathbf{C}(\Delta\tau) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} f(\mathbf{A}\Delta\tau(\mathbf{y}_g, \mathbf{y}_s)), \qquad (6)$$

where $\mathbf{y}_g$ and $\mathbf{y}_s$ are a sparse representation of the source and receiver locations, $\mathbf{A}$ is a bicubic spline interpolation operator that maps the sparse coordinates $\mathbf{y}_g$ and $\mathbf{y}_s$ to the original coordinates $\mathbf{x}_g$ and $\mathbf{x}_s$, and f evaluate the correlation value at $\Delta\tau(\mathbf{x}_g, \mathbf{x}_s)$.

The goal of the global optimization is to maximize the function described by equation (6), which is to maximize the stacking power along the interpolated spline surface. The searching procedure is a simulated annealing algorithm, which varies the spline points along the time axis in a stochastic sense until a satisfying solution is reached. In the following section, I show the results of using such global scheme to pick the correlation lags.

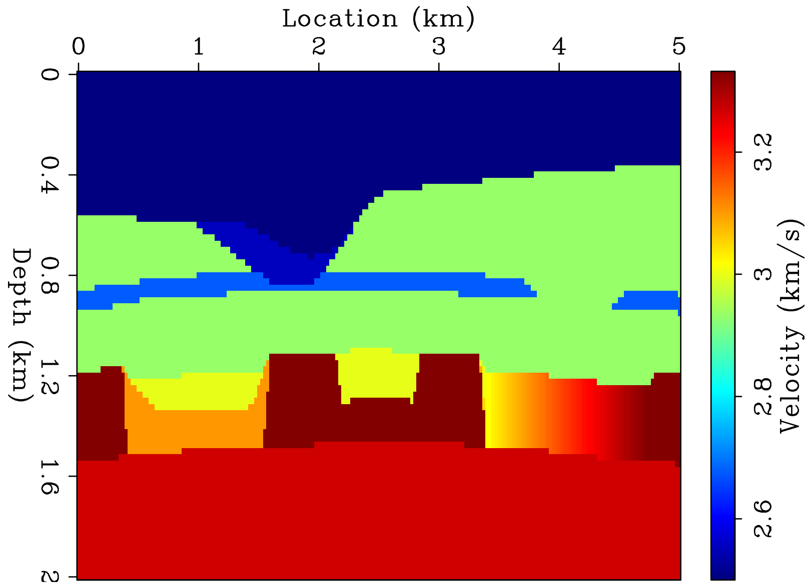


Figure 1: The true velocity model used to create the data. [**ER**]

SYNTHETIC EXAMPLE

I use the velocity model shown in Figure 1 to create the synthetic data. The sampling interval of the model is 25 m for both the x-axis and z-axis. The data is modeled using a Ricker wavelet with a fundamental frequency of 15 Hz and a sampling interval of 3 ms. The acquisition geometry is a cross-well configuration, where sources are located in a vertical well on the left side of the model and the receivers are located in a vertical well on the right side of the model. Both the sources and the receivers start at the surface and cover the full depth of the model with a sampling inteval of 25 m.

I start the velocity inversion with a constant background velocity of 2.9 km/s, which is very far from the true velocity model. After modeling the data with the background velocity, I cross-correlate the modeled data with the observed data. Figure 2(a) shows the lags picked by maximum correlation at each trace. There is an overall trend from the top left corner to the bottom right corner, which is caused by the gradient in the original data. However, there is also some large anamolies in the picked lags with sharp discontinuities around them. These anomalies are caused by the events refracted by the large velocity contrasts in Figure 1. Figure 2(b) shows the lags picked by the global method, which are significantly different from those in Figure 2(a). The global algorithm ignored the local maximum caused by refracted energy and picked lags that are more consistent with their surroundings.

Now, I use the lags estimated by both methods to find a velocity update. The scale of the update is estimated by performing a line search. Figure 3 shows the updated velocity model using both methods after one iteration. The results of the local method shows some noise in the velocity model update, especially around the edges. However, since the velocity update is the sum of the updates of all the shots,

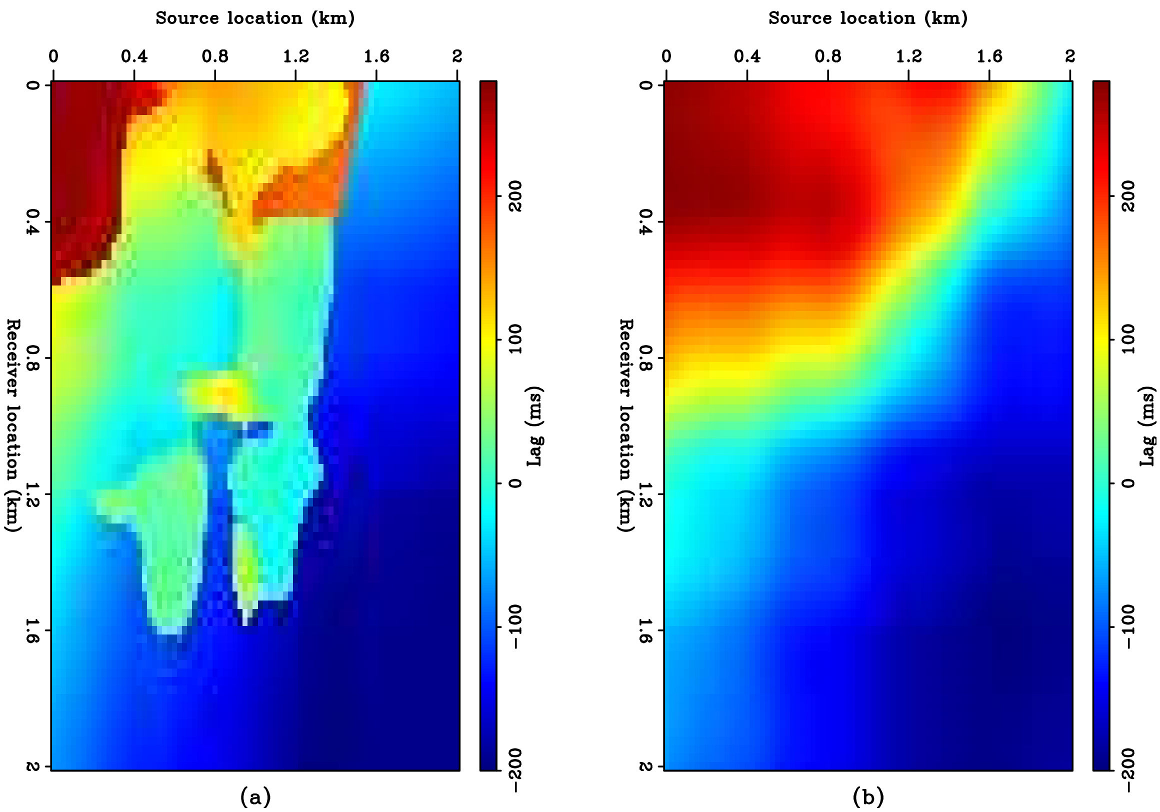


Figure 2: The correlation lags picked by: (a) the trace-by-trace maximum correlation method, and (b) the global spline-fitting method. [**ER**]

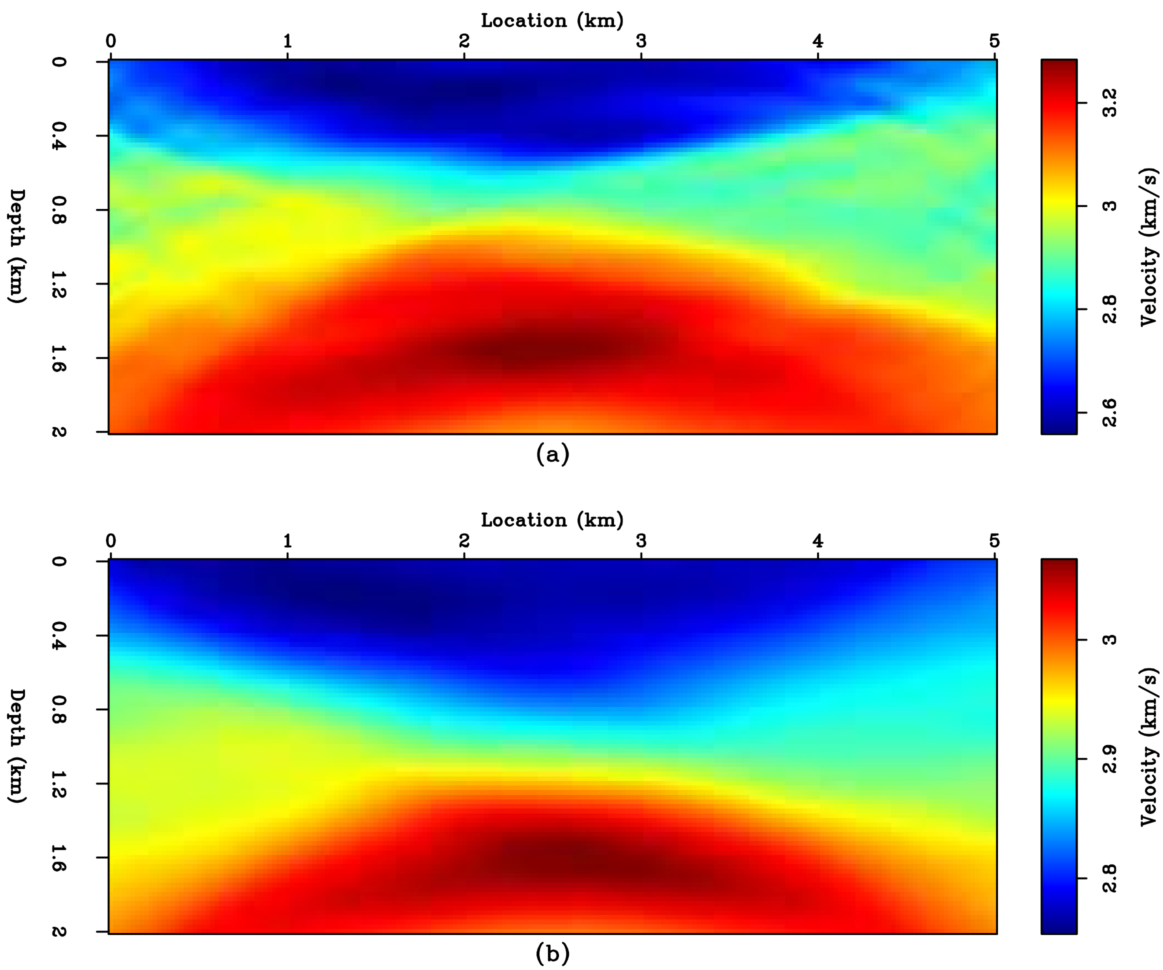
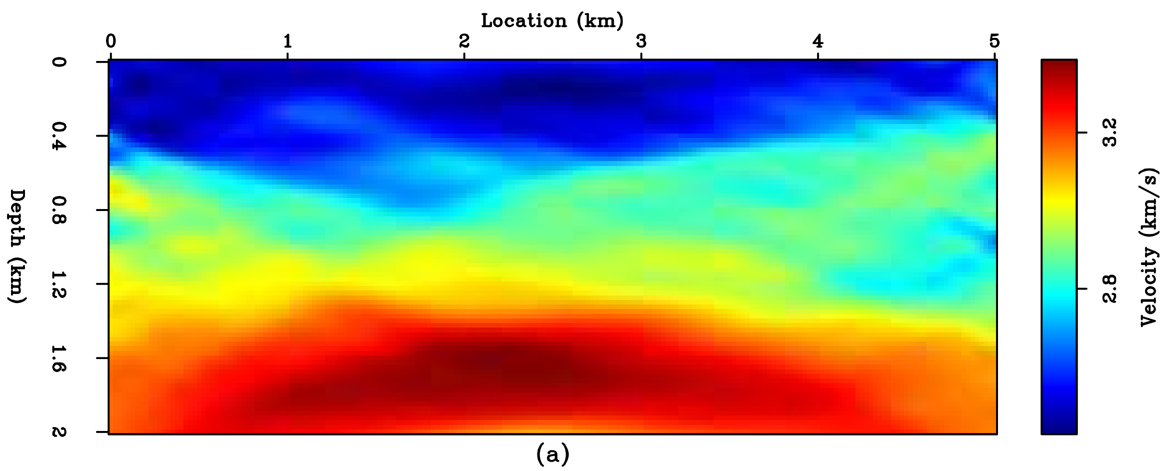


Figure 3: The velocity model after 1 iteration using: (a) the trace-by-trace maximum correlation method, and (b) the global spline-fitting method. [**ER**]

the total results are satisfactory. On the other hand, the velocity model estimated by the global method looks cleaner and more consistent laterally. However, the velocity update by the global method shows less detail than that of the local method. In addition, the global method converges at a slower rate than the local method.

Finally, I run both inversions for 10 iterations to further show the difference between the two methods. The results of both inversions are shown in Figure 4. The noise in the local method estimate grows even larger as I run more iterations, but the accuracy seems to be consistent with the true velocity model. On the other hand, the global method is still too smooth and did not pick the small details as well as the local method. Finally, there seems to be some bias in the global method toward lower velocities.



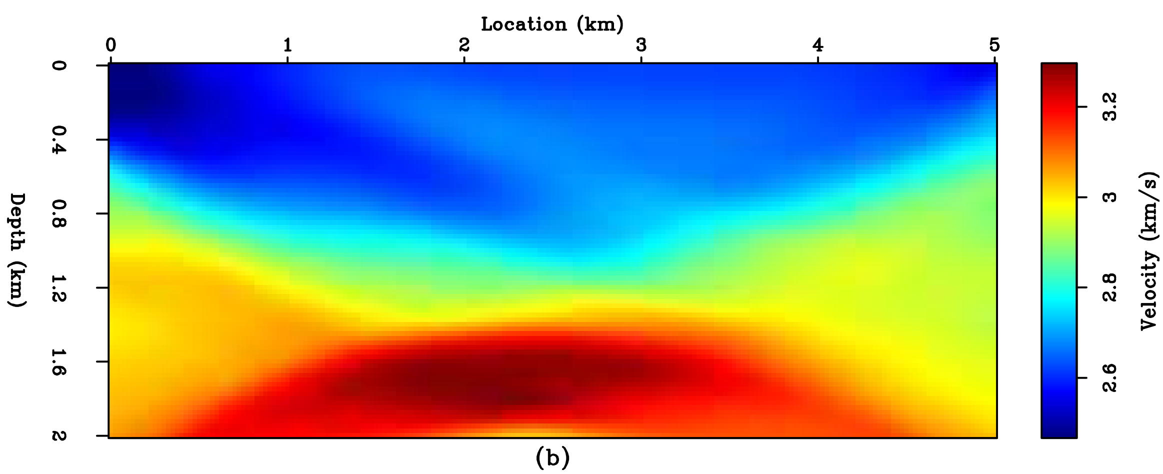


Figure 4: The inverted velocity model after 10 iterations using: (a) the trace-by-trace maximum correlation, and (b) the global spline-fitting method. [**ER**]

CONCLUSIONS AND DISCUSSIONS

We showed that using a global scheme for picking the correlation lags can better detect the correct events and is more robust to noise and multiple events than a trace-by-trace method. However, since the global picking is done on sparse spline points, the picking surface is too smooth, which causes the result to have less detail than that of the local method and converges at a lower rate. Nevertheless, the estimated velocity model did correctly approach the true velocity model, even though we started with a constant velocity.

Moreover, the accuracy and smoothness of the global method is strongly influenced by the number of spline points specified by the user. Therefore, accurately estimating the smoothness of the picked cross-correlation surface can help in determining a good spline geometry.

FUTURE WORK

As we have seen, the global optimization is less accurate than the local fitting. One way to improve the results is to use a hybrid optimization method in which the algorithm either alternates between a global and a local iteration or runs a global search, followed by a local one. In addition, all testing has been done on transmission seismology. The next step is to extend the method to reflection seismology.

Finally, the tests case in this paper are synthetic data. In field data, the noise level is an issue that might degrade the local method even more, in which case the global method can show even more improvement.

ACKNOWLEDGMENTS

We thank Biondo Biondi, Robert Clapp and Xukai Shen for their valuable discussions and help throughout this project.

REFERENCES

Biondi, B., 2009, Measuring image focusing for velocity analysis: SEP-Report, **138**, 59–79.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Luo, Y. and G. T. Schuster, 1990, Wave-equation traveltime inversion: SEG Expanded Abstracts.

Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.

Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

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