APPLIED GEOPHYSICS, Vol.20, No.4 (December 2023), P. 374–384, 14 Figures.
DOI:10.1007/s11770-020-0875-x

Inexact line search method in full waveform inversion

Ma Xiaona[1], Xu Shan-hui[1]*, Ke Pei[2], Zhang Hong-tao[3]

Abstract: Full waveform inversion (FWI) is a nonlinear data fitting process that can derive high-resolution model parameters through iteration. In this process, step length is related to inversion accuracy and computational efficiency. It can be calculated efficiently with the inexact line search method, which does not require a misfit function to achieve the exact minimum. This method is aimed toward obtaining the appropriate descent using evaluation conditions and initial step length. Moreover, it does not depend on the form of the misfit function. In the inexact line search method, the evaluation condition and initial step length are obviously important factors. In this work, the classical Armijo, Wolfe, and Goldstein evaluation conditions in solving optimization problems in mathematics are studied and compared in detail. Numerical examples from the synthetic data of the overthrust model show that the convergence characteristics of Armijo and Goldstein are similar and that the computational efficiency is high and conducive to seismic FWI. In addition, the adaptive Barzilai–Borwein (ABB) method is adopted in FWI. The ABB method maximizes the changes in model parameters and gradients to adaptively calculate the initial step length. The threshold value of the ABB method for the initial step length estimation is also studied to explore a suitable threshold value that can ensure that large and small step lengths are frequently adopted in FWI. Numerical examples from the synthetic data of the overthrust model demonstrate the validity of the ABB method. Moreover, the inversion is superior when the threshold value is less than 0.5.
Keywords: Full waveform inversion; Inexact line search method; Evaluation condition; Initial step length; Threshold value

Introduction

Full waveform inversion (FWI) uses all wave information of prestack seismic data to quantitatively modify initial models by fitting the observed and simulated data and finally obtain high-precision underground model parameters (Qu et al., 2017; Rao and Wang, 2017; Liu et al., 2019). Since Tarantola (1984) proposed FWI in the time domain, its related theory has developed rapidly. With the improvement of computing power, FWI has been gradually applied to field data

Manuscript received December 3, 2020; revised manuscript received May 3, 2021
This research is financially supported by the Special Fund of the Institute of Geophysics, China Earthquake Administration (Grant Number DQJB19B25 and Grant Number DQJB21K31).
1. Institute of Geophysics, China Earthquake Administration, Beijing 100081
2. BGP Inc., China National Petroleum Corporation, Hebei Zhuozhou 072751
3. Liaoning No.1 Hydrogeological Engineering Geology Group Co.,Ltd, Liaoning Jinzhou 121000.
*Corresponding author: Xu Shanhui (E-mail: xushanhuia@126.com).

(Shipp and Singh, 2002; Shi et al., 2014; Wang et al., 2017).

FWI is an iterative method for nonlinear data fitting. The process can be summarized as follows: the initial model is set, the misfit function for measuring observed and simulated data is obtained, the gradient is calculated by the adjoint state method, the search direction is constructed by an inversion optimization algorithm, the step length is calculated, and the misfit function is reduced by continuously updating the model parameters; finally, high-precision results are obtained (Zhang et al., 2018). In the process of FWI, the calculation of the step length is a key problem, which is related to the convergence rate and accuracy of inversion. For FWI, which is a strong nonlinear inversion problem, the step length is usually obtained using the line search method along the search direction (Pratt et al., 1998; MéTivier and Brossier, 2016). Line search methods mainly include exact and inexact line search methods (Hu et al., 2014; Ma et al., 2018). A fixed step length can be used to obtain the initial step length; although the approach can improve the calculation efficiency, the inversion result is poor, and several attempts are required to obtain the fixed step length (Cao, 2015).

The exact line search method is based on a misfit function to obtain the exact step length of the current iteration. The parabolic method (Vigh et al., 2009) and direct method (Pica et al., 1990) are exact step length calculation methods that are based on the L2 norm misfit function. The parabolic method requires at least two additional forward modeling to calculate the step length in each iteration. This requirement increases the computational burden of FWI. By contrast, the direct method needs one additional forward modeling evaluation, and it is comprehensively developed at present. Bube and Nemeth (2007) presented an exact step length for the hybrid norms. Based on its principle, the exact step length calculation method depends on the form of the misfit function and, thus, suffers from poor universality. For example, the cross-correlation misfit function (van Leeuwen and Mulder, 2010), envelope misfit function (Wu et al., 2014; Wu et al., 2010; Dong et al., 2015), and other functions require to derive their corresponding exact step length calculation methods.

The inexact line search method can make the misfit function have a corresponding decrease in the search direction. Moreover, the misfit function need not reach the exact minimum in each iteration. This type of method is widely applicable to different misfit functions and offers great development potential. Relevant scholars have conducted detailed research on the application of this type of line search method to solve optimization problems and have made considerable progress (More and Thuente, 1994; Nocedal and Wright, 2006). Evaluation conditions and initial step length are two key factors in the inexact line search method. The Armijo, Wolfe, and Goldstein conditions are often used to solve optimization problems. The Armijo condition uses the information of two misfit functions for evaluation and can make each misfit function achieve a certain amount of decline. The Wolfe condition not only requires the misfit function to decline but also further considers the effect of the slope of the misfit function and uses the gradient information for evaluation. The Goldstein condition is constrained by two inequalities based on the misfit function, and it strives to obtain enough descent in the search direction (Nocedal and Wright, 2006). The current work performs a detailed study and comparative analysis of the application effect of the above three conditions and explores the evaluation conditions suitable for the inexact line search method in FWI.

For the inexact line search method, scholars have proposed different methods to calculate the initial step length; these methods, which include the first-order change method (Nocedal and Wright, 2006), the method that uses velocity and search direction (Operto et al., 2007), and the method involving a given step length interval (More and Thuente, 1994), promote the development of the inexact line search method. Zhou et al. (2006) proposed the adaptive Barzilai–Borwein (ABB) method to solve mathematical optimization problems with an inversion optimization algorithm. The ABB method adaptively calculates the initial step length using the variations in model parameters and gradients and achieves significant convergence effect and inversion accuracy (Dos Santos and Pestana, 2015). In the ABB method, the function of a small step length is to modify the search direction, and the function of a large step length is to obtain enough descent in the appropriate search direction. In view of the characteristics of the ABB method, it is used to adaptively calculate the initial step length of the inexact line search method in FWI. The threshold selection of the ABB method affects the effectiveness of inversion optimization; specifically, the threshold is usually set to 0.5 (Zhou et al., 2006). In the current work, we set the thresholds of the initial step length of the ABB method to be greater than 0.5, equal to 0.5, and less than 0.5. We also study their influence on the results of FWI. The overthrust model test is performed to verify, compare, and analyze the methods.

Theory of full waveform inversion

Take the following two-dimensional constant density acoustic equation as an example:

$$\frac{1}{v^2(x,z)}\frac{\partial^2 p(\mathrm{x,z},t)}{\partial t^2}=\nabla^2 p(\mathrm{x,z})+s(\mathrm{x,z},t), \quad (1)$$

where t represents time, (x,z) represents the horizontal and vertical coordinates, $p(x,z,t)$ represents the wave field, $v(x,z)$ represents the velocity, ∇^2 represents the Laplace operator, and $s(x,z,t)$ represents the source term. The L2 norm misfit function is constructed as follows:

$$f=\frac{1}{2}\sum_s\sum_g\left[p^{cal}(x,z,t)-p^{obs}(x,z,t)\right]^2, \quad (2)$$

where $p^{cal}(x,z,t)$ is the calculated wave field and $p^{obs}(x,z,t)$ is the observed wave field. The adjoint state method is used to calculate the gradient of the misfit function (2) (Plessix, 2006; Ren and Liu, 2015):

$$g=\frac{2}{v(x,z)^3}\int_0^T\lambda\frac{\partial^2 p(x,z,t)}{\partial t^2}dt, \quad (3)$$

where T represents the maximum computation time, p represents the forward propagating wave field, and λ represents the residual backward propagating wave field. The backpropagation residuals corresponding to the L2 norm misfit functions are expressed as follows:

$$\delta p=\sum_g p^{cal}(x,z,t)-p^{obs}(x,z,t), \quad (4)$$

The model parameter update expression can be written as follows:

$$v_{k+1}=v_k+\alpha_k d_k, \quad (5)$$

where v_{k+1} and v_k represent the velocity parameters of two iterations. d_k represents the search direction, which is usually obtained by gradient and quasi-Newton inversion algorithms. a_k is a scalar value representing the step length. Its main function is to transform the gradient vector into model parameters (Pratt et al., 1998).

Inexact line search method

The inexact line search method updates the model with a given initial step length and uses the evaluation conditions to make the misfit function converge quickly. Unlike the exact line search method, the inexact line search method is suitable for any form of misfit function, and it does not need to derive the exact step length calculation formula. Hence, its range of applicability is relatively broad. The goal of exploring the step length calculation method is to obtain enough descent in each iteration and improve the convergence rate of inversion on the premise of ensuring the accuracy of inversion results.

Evaluation condition

The Armijo, Wolfe, and Goldstein conditions are often used to solve optimization problems mathematically. These three conditions have their own characteristics (Nocedal and Wright, 2006). The Armijo condition is written as follows:

$$f(v_k+\alpha_k d_k)\le f(v_k)+c_1\alpha_k\nabla f_k^T d_k, \quad (6)$$

where $0<c_1<1$. This condition uses the gradient and search direction information of the current iteration and the misfit function of two adjacent iterations for evaluation. Hence, in this condition, the misfit function has a certain amount of decline. This condition includes all small step length values. As shown in Figure 1, some minimum step length values satisfying condition (6) can make the inversion converge gradually; however, they

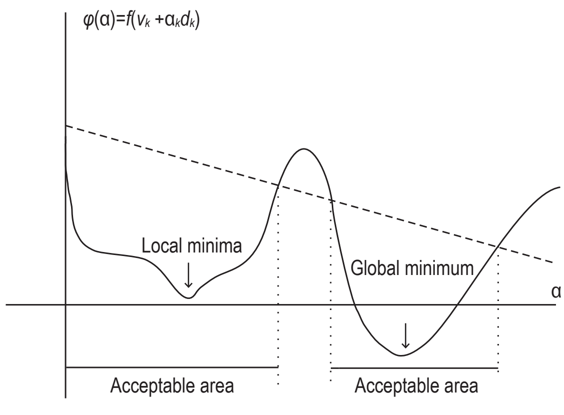

Figure 1. Armijo condition

can inevitably reduce the computational efficiency of inversion.

The Wolfe condition features the characteristics of the Armijo condition, but it also considers the slope of the misfit function. The gradient information in the case of adjacent iterations is used as follows:

$$\begin{cases} f(v_k+\alpha_k d_k) \le f(v_k) + c_1\alpha_k \nabla f_k^T d_k \\ \nabla f(v_k+\alpha_k d_k)^T d_k \ge c_2 \nabla f(v_k)^T d_k \end{cases}, \tag{7}$$

where $0<c_1<c_2<1$. The Wolfe condition can exclude the minimal step length (Figure 2), but it needs to calculate the gradient of the misfit function twice; in such a case, the amount of calculation increases significantly.

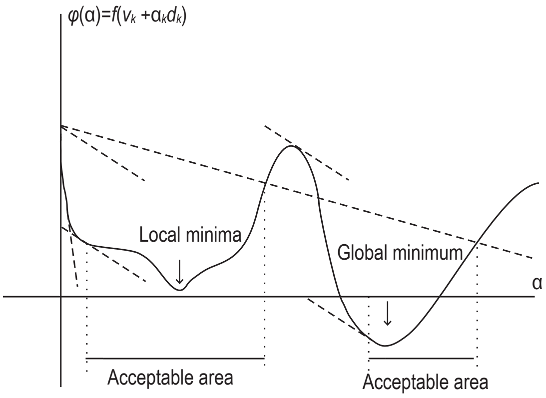

Figure 2 Wolfe condition

The Goldstein condition uses two inequalities to measure the size of the misfit function and avoid the small step length value. With this approach, the misfit function can obtain enough descent, and the computational efficiency can be improved (Figure 3). The Goldstein condition is given as follows:

$$\begin{aligned} f(v_k+\alpha_k d_k) &\le f(v_k) + c\alpha_k g_k^T d_k, \\ f(v_k+\alpha_k d_k) &\ge f(v_k) + (1-c)\alpha_k g_k^T d_k, \end{aligned}, \tag{8}$$

where $0<c<\frac{1}{2}$. Although the misfit function can obtain enough descent, some small effective step length values may be excluded from the acceptable region.

Figures 1–3 are all quoted from the introduction of evaluation conditions in the work of Nocedal and Wright (2006). The schematics of the Armijo condition (Figure 1), Wolfe condition (Figure 2), and Goldstein condition (Figure 3) indicate that with the use of gradient and other information, the evaluation condition becomes gradually strict, and the area meeting the condition is gradually narrowed. In general, these conditions are aimed at excluding small step length values and obtaining enough descent. However, for strong nonlinear problems such as FWI, the misfit functions have many local minima. If some small step lengths are excluded unilaterally, then the inversion may be unstable, and the computational efficiency may be reduced.

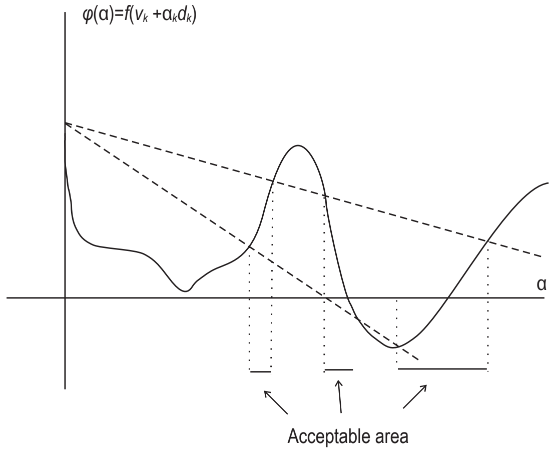

Figure 3 Goldstein condition

Initial step length

In the inexact line search method, the initial step length is a key variable. A good initial step length can reduce the number of additional forward modeling evaluations and improve the computational efficiency, while ensuring accuracy. The ABB method adaptively calculates the initial step length of a given iteration using the gradient and model parameters under the current iteration. The calculation formula is as follows (Zhou et al., 2006):

$$\alpha_k^{abb} = \begin{cases} \alpha_k^{bb2} = \dfrac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}, & \alpha_k^{bb2} / \alpha_k^{bb1} < \mu \\ \alpha_k^{bb1} = \dfrac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}}, & otherwise \end{cases} \quad \mu \in (0,1), \tag{9}$$

where $k \ge 1$, $s_{k-1}=m_k-m_{k-1}$ is the variation of the model parameters and $y_{k-1}=g_k-g_{k-1}$ is the gradient change. $\alpha_k^{bb2} \le \alpha_k^{bb1}$ is the function of a small step length, α_k^{bb2} is used to guide the search direction and find a favorable descent direction for the next iteration, while the function of large step length α_k^{bb1} is to obtain enough descent. Threshold μ can adjust the frequency of large step length α_k^{bb1} and small step length α_k^{bb2}, and its value affects the effectiveness of inversion.

Taking the Armijo condition as an example, the flow of the inexact line search method is as follows:

Step 1. The ABB formula is used to calculate the initial step length, and the value of c_1 in the Armijo condition is given (usually $c_1=10^{-4}$). Given the appropriate step length attenuation factor τ (τ=50% in this work) to reduce the initial step length appropriately, the failure of the initial step length to meet the evaluation conditions is avoided;

Step 2. The model parameters are updated with formula (5);

Step 3. Whether the Armijo condition is satisfied is evaluated. If the conditions are met, then the iteration ends. If the condition is not met, then step 4 is executed;

Step 4. The attenuation factor is used to reduce the step length appropriately, $\alpha=\tau\alpha$. Then, Step 1 is repeated.

Numerical examples

This work uses the overthrust model to test the method. The overthrust model is shown in Figure 4a; it consists of continuous sedimentary thrust faults covering the bedrock, as well as many complex small structures, such as fine layers and strata extinction (Ravaut et al., 2004). Figure 4b shows the linear increase model as the initial velocity model. The size of the model is , and the grid spacing is 12 m. The seismic wavelet is selected as the shot with the main frequency of 18 Hz. The time sampling interval is 1 ms and the number of sampling points is 3,000. The number of shot points is 60 and the spacing per shot is 120 m. The number of receiver points is 601 and the spacing per shot is 12 m.

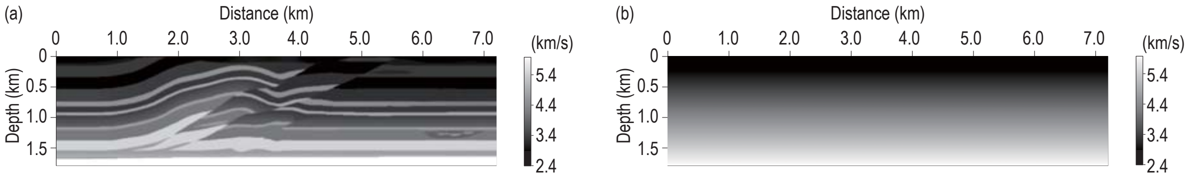

Figure 4 a) True overthrust velocity model; b) initial linear increase velocity model.

Test of evaluation condition

This work studies the validity of the Armijo, Wolfe, and Goldstein conditions in FWI. To compare the three evaluation conditions accurately, we set the other inversion settings to be the same. Figure 5 shows the inversion results of the three conditions. The Armijo condition (Figure 5a), Wolfe condition (Figure 5b), and Goldstein condition (Figure 5c) all reconstruct the overthrust velocity model. The main structure is clearly depicted, and it is close to the true velocity model (Figure 4a). Moreover, the inversion results of the three conditions present minimal differences in accuracy.

For FWI, not only the accuracy of inversion results but also computational efficiency should be considered. In the process of FWI, a large number of forward simulations comprise the main calculation (Anagaw and Sacchi, 2014). In each iteration of inversion, the gradient calculation requires a wave field forward propagation and residual wave field backward propagation, while the estimation of step length requires forward modeling. In this work, the forward modeling involved in the step length calculation is recorded as additional forward modeling. We estimate the number of iterations and the additional number of forward modeling evaluations

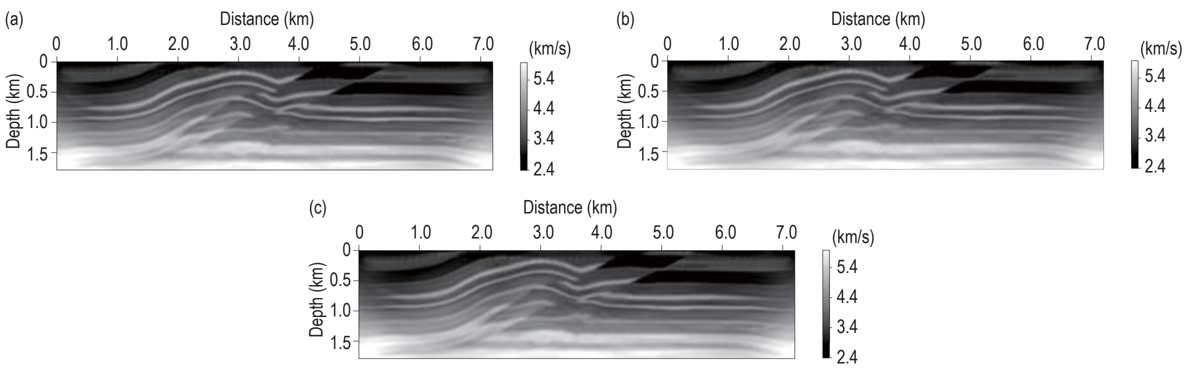

Figure 5 Inversion results of three evaluation conditions: a) Armijo; b) Wolfe; c) Goldstein

required by the step length calculation to roughly analyze the calculation efficiency. Figure 6a shows the root mean square error of FWI, and Figure 6b presents the statistical chart of iteration frequency and additional forward modeling evaluations. The convergence rates of the three evaluation conditions are similar, and the results of the inversion convergence of the Armijo and Goldstein conditions are similar (Figure 6a). Although the Armijo condition allows all small step length values to meet the condition, it does not reduce the computational efficiency of FWI. The Wolfe condition requires two gradient calculations; thus, the additional forward modeling evaluations in this condition are significantly greater than those in the other two evaluation conditions, thereby resulting in a significant increase in the amount of calculation (Figure 6b). For the inexact line search method in FWI, the Armijo and Goldstein conditions with high computational efficiency are recommended.

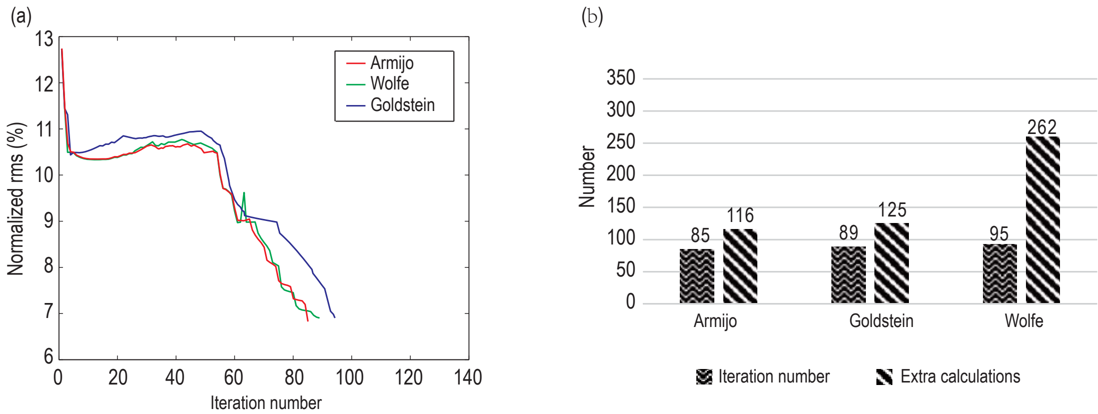

Figure 6 a) Root mean square error; b) total iteration number and number of additional forward modeling evaluations

Test of ABB method

The validity and threshold of the ABB method for initial step length calculation are tested and analyzed. In view of previous research, we use the Armijo condition as the evaluation condition in this part of the model test. The given linear increase model (Figure 4b) is taken as the initial model, and the inversion results of the ABB method are taken as 0.2, 0.5, and 0.8. In the three threshold cases, the ABB method can effectively reverse and reconstruct the overthrust velocity model. However, when the threshold of the ABB method is 0.2, the overall structure description of the overthrust model, especially the deep part of the left and right areas of the model, is relatively clear. Figure 8 shows two depth–velocity profiles randomly selected from the three threshold inversion results in Figure 7. When the threshold value is 0.2, the velocity of the inversion results is close to the real velocity value (shown by the arrow).

Figure 9 shows the step length value in the inversion process. In the initial stage of inversion iteration, when the threshold values are 0.2 and 0.5, the step length values coincide. When the number of iterations is 18, the step length values differ. Regardless of the threshold value, the step length of each iteration varies mainly

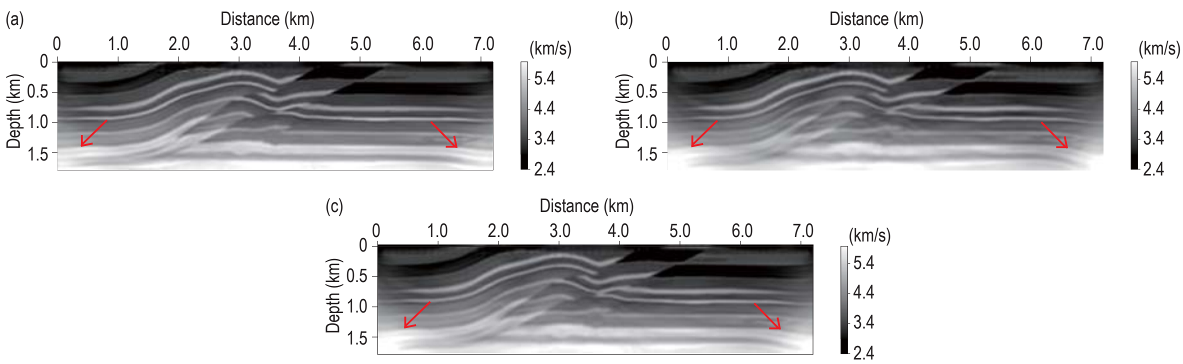

Figure 7 Inversion results of ABB method: a) threshold value is 0.2, b) threshold value is 0.5, and c) threshold value is 0.8

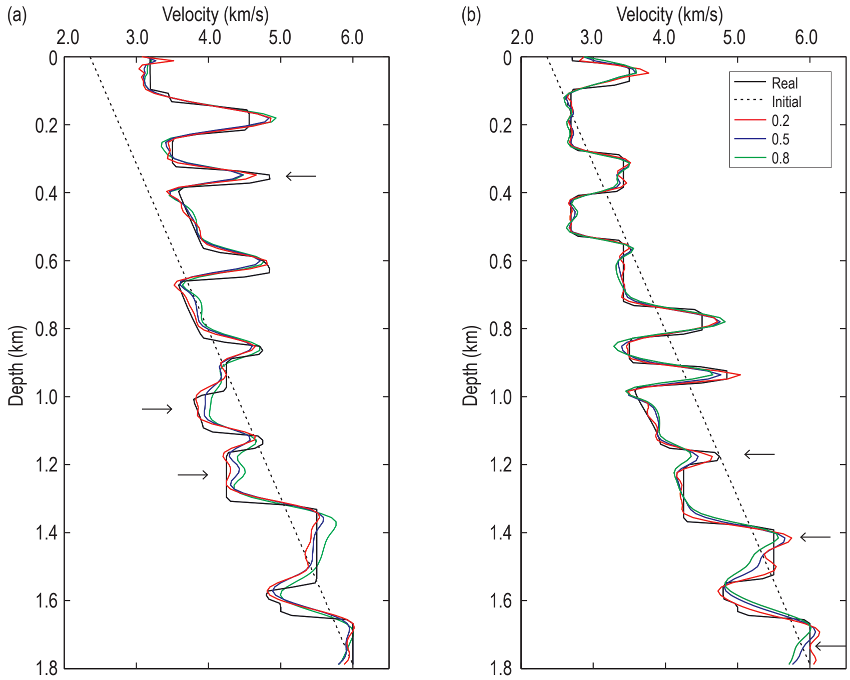

Figure 8 Comparison of depth–velocity profiles: a) 3.0 km and b) 4.8 km

because the gradient, model parameters, direction, and other information change during each iteration. As shown in Figure 9, when the threshold value is 0.2, the step length value obtained by the ABB method is generally smaller than those obtained when the threshold values are 0.5 and 0.8. Figure 10a shows the root mean square error of the inversion iteration process. The inversion accuracy of the ABB method is higher when the threshold is 0.2 than when the threshold is 0.5 or 0.8. Figure 10b shows the statistical chart of the iteration times and additional forward modeling evaluations. The iteration times of the three thresholds are similar. As for the number of additional forward modeling evaluations, it is fewer when the threshold is 0.2 than when the threshold is 0.5 or 0.8.

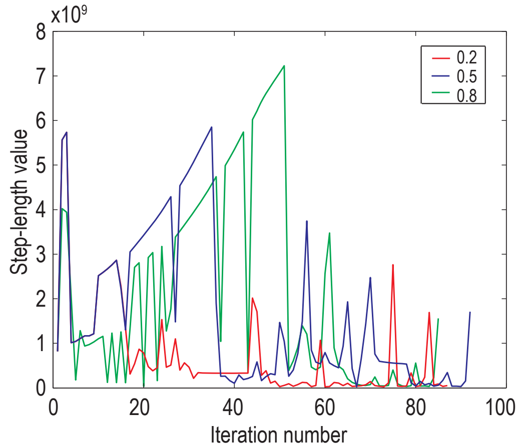

Figure 9 Step length value at each iteration of three threshold values

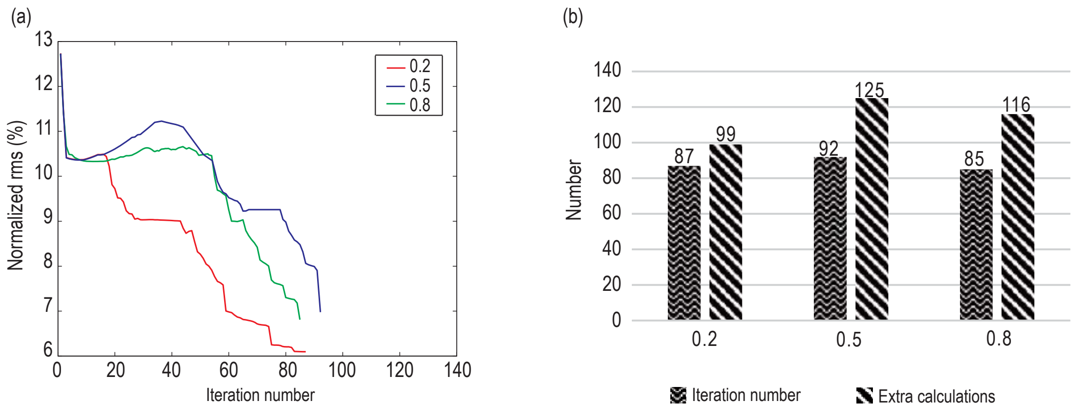

Figure 10 a) Root mean square error; b) total iteration number and number of additional forward modeling evaluations

Comparison of initial step length values

Herein, the ABB method for initial step length calculation is compared with the method using rate and search direction (Operto, 2007). The Armijo condition is still used as the evaluation condition, and the threshold value of the ABB method is 0.2. For the convenience of comparison, we temporarily record the method used by Operto et al. (2007) as the "typical method," which is different from the ABB method; the calculation formula is as follows:

$$(\alpha_k^0 = coef1 * \mathrm{maxvel}_k / \max(d_k)), \quad (10)$$

where α_k^0 is the initial step length size, $coef1$ is the constant coefficient ($coef1$=10^{-2} in this work), max(d_k) is the maximum value of the search direction, and maxvel_k is the corresponding velocity value of the maximum value of the search direction. Figure 11 shows the inversion results of the ABB method and the traditional initial step length calculation method. The ABB method (Figure 11a) can effectively invert the overthrust velocity model and describe the model clearly and accurately. The accuracy of its inversion result is also higher than that of the traditional initial step length calculation method (Figure 11b). The profile curves at 3.0 and 4.8 km in the horizontal position confirm that the inversion results of the ABB method are accurate in the spatial position and that the velocity is close to the real velocity value (Figure 12).

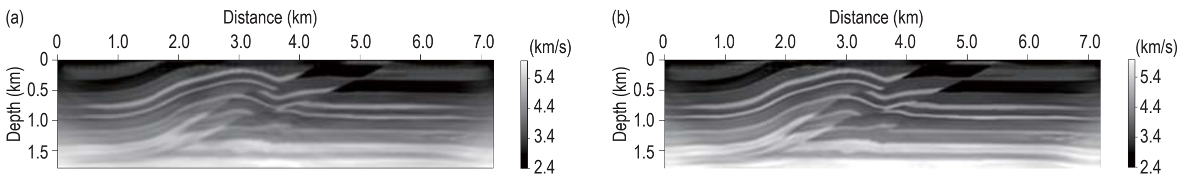

Figure 11 Inversion results: a) traditional initial step length; b) initial step length of ABB method

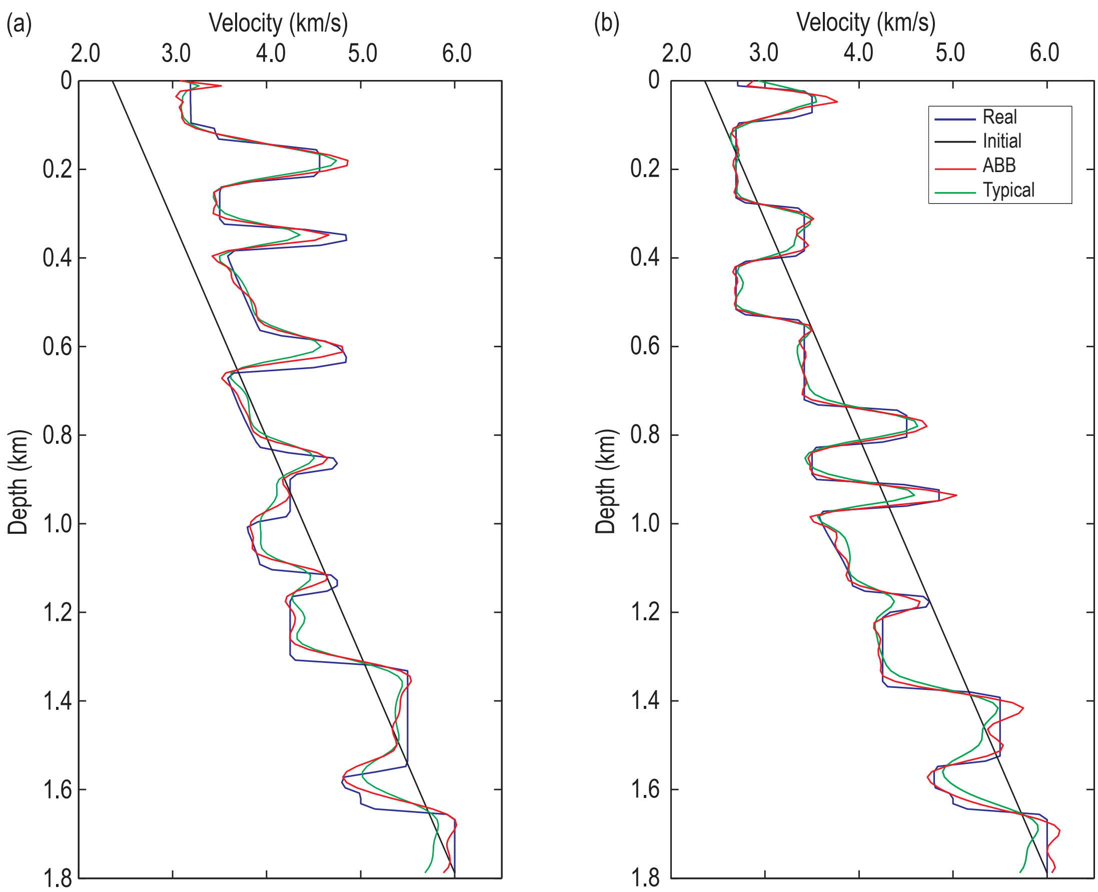

Figure 12 Comparison of depth–velocity profiles: a) 3.0 km and b) 4.8 km

Figure 13 shows the step length values obtained by the two initial step length calculation methods in the inversion process. Compared with the traditional initial step length calculation method, the ABB method obtains a larger step length in the initial stage of inversion. With the gradual convergence of inversion, the large step length and small step length are used alternately to adaptively cooperate with the inversion convergence process. The root mean square error curve in Figure 14a also proves that the inversion accuracy of the ABB method is higher than that of the traditional initial step length calculation method. The ABB method, in particular, has a fast convergence rate, and it only requires 87 iterations even if the inversion converges. Moreover, it only needs 99 additional forward modeling evaluations; such a number is ~37% higher than that of the traditional method, which requires 156 forward modeling evaluations. The ABB method not only achieves high inversion accuracy but also significantly improves inversion efficiency.

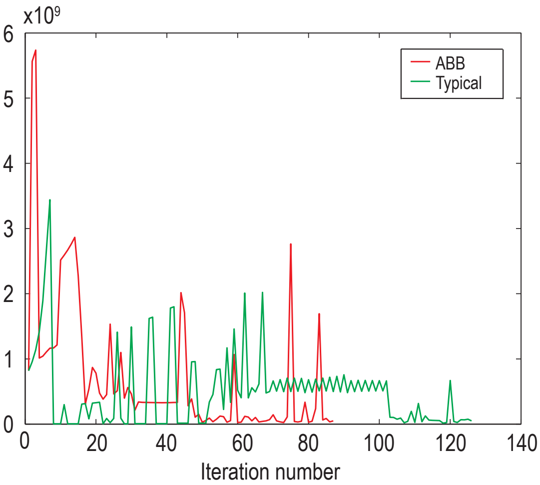

Figure 13. Step length value at each iteration of two initial step length calculation methods

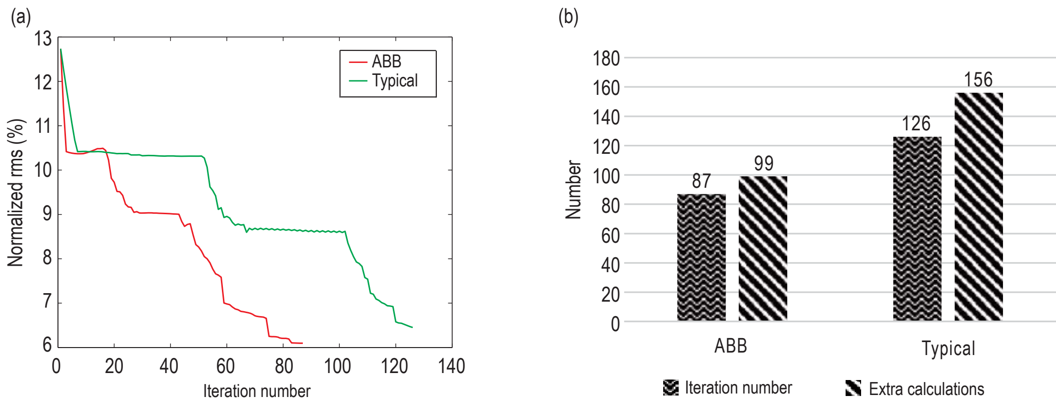

Figure 14 a) Root mean square error; b) total iteration number and number of additional forward modeling evaluations

Conclusions

This work investigates in detail the evaluation conditions and initial step length of the inexact line search method in FWI. The Armijo, Wolfe, and Goldstein evaluation conditions are studied. The overthrust velocity model example shows that the calculation efficiency of the Wolfe condition is low, because it requires an extra gradient calculation. The Armijo and Goldstein conditions have a similar convergence, few additional forward modeling evaluations, and high computational efficiency. The initial step length in the inexact line search method is also studied herein. The ABB method is used to calculate the initial step length adaptively, while fully considering changes in model parameters and gradients. The example of the overthrust velocity model proves the validity of the ABB method. The ABB method is analyzed and compared with the traditional initial step length calculation method. The ABB method improves the accuracy of the seismic data inversion result, effectively reduces the number of additional forward modeling evaluations, and improves calculation efficiency. The study of the threshold value of the ABB method shows that its threshold value range should be less than 0.5, which is conducive to obtaining high-precision FWI inversion results.

Acknowledgments

This research is financially supported by the Special Fund of the Institute of Geophysics, China Earthquake

Administration (Grant Number DQJB21K31 and Grant Number DQJB19B25).

References

Anagaw, A. Y., and Sacchi, M. D., 2014, Comparison of multifrequency selection strategies for simultaneous-source full-waveform inversion: Geophysics, **79**(5), R165–R181.

Bube, K. P., and Nemeth, T., 2007, Fast line searches for the robust solution of linear systems in the hybrid and Huber norms: Geophysics, **72**(2): A13–A17.

Cao, S. H., 2015, Frequency domain acoustic full waveform inversion and reverse time migration in anisotropic media (Doctoral Dissertation). University of Chinese Academy of Sciences.

Dos Santos, A. W. G., and Pestana, R.C., 2015, Time-domain multiscale full-waveform inversion using the rapid expansion method and efficient step-length-length estimation: Geophysics, **80**(4), R203–R216.

Dong, L. G., Huang, C., Chi, B. X., Liu, Y. Z., 2015, Strategy and application of waveform inversion based on seismic data subset: Chinese J. Geophys. (in Chinese), **58**(10),3735–3745.

Hu, G. H., Wang, L. X., Fang, W. B., et al., 2014, Full waveform inversion method and application: China Petroleum Industry Press. Beijing, China.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: SIAM Conference on Inverse Scattering, Theory and Applications, Expanded Abstracts, 206–220.

Liu, Y. Z., Huang, X. Q., Wan, X. W., Sun, M. A., Dong, L. G., 2019, Elastic multi-parameter full-waveform inversion for anisotropic media: Chinese J. Geophys. (in Chinese), **62**(5),1809–1823.

More, J. J., Thuente, D. J., 1994, Line search algorithms with guaranteed sufficient decrease: ACM Transactions on Mathematical Software, **20**(3), 286–307.

Ma, X. N., 2018, Research on Time-domain Full Waveform Inversion (Doctoral Dissertation). University of Chinese Academy of Sciences.

Métivier, L., and Brossier, R., 2016, The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication: Geophysics, **81**(2), F1–F15.

Nocedal, J., and S. J. Wright., 2006, Numerical optimization: Springer Nature, Cham, Switzerland.

Operto S, J Virieux, and F Sourbier., 2007, Documentation of FWT2D program (version 4.8): Frequency domain full-waveform modeling/inversion of wide-aperture seismic data for imaging 2D acoustic media: Technical report 007–SEISCOPE project.

Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: Geophysics, **55**(3), 284–292.

Pratt, R. G., 1998, Guass-Newton and full Newton methods in frequency domain seismic waveform inversion: Geophysics, **173**, 922–931.

Qu, Y., Li, Z., Huang, J., Li, J., Guan, Z., 2017, Elastic full waveform inversion for surface topography: Geophysics, **82**(5), R269–R285.

Ren, Z. M., and Liu, Y., 2015, Elastic full-waveform inversion using the second-generation wavelet and an adaptive-operator-length scheme: Geophysics, **80**(4), R155–R173.

Rao, Y., Wang, Y., 2017, Seismic waveform tomography with shot-encoding using a restarted L-BFGS algorithm, Scientific Reports, **7**, 2045–2322.

Rao, Y., Wang, Y., Han, D., 2019, Seismic waveform tomography with simplified restarting scheme: IEEE Geoscience and Remote Sensing Letters, **16**(1), 135–139.

Ravaut, C., Operto, S., Improta, L., et al., 2004, Multiscale imaging of complex structures from multifoldwide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: Geophysical Journal International, **159**(3), 1032–1056.

Plessix, R., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**(2), 495–503.

Shipp, R. M., & Singh, S. C., 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: Geophysical Journal International, **151**(2), 325–344.

Shi, Y. M., Zhang, Y., Yao, F. C., et al., 2014, Methodology of seismic imaging for hydrocarbon reservoirs based on acoustic full waveform inversion: Chinese J. Geophys. (in Chinese), **57**(2),607–617.

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

Vigh, D., Starr, W., & Kapoor, J., 2009, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**(5), 135–144.

Van, Leeuwen, T., Mulder, W. A., 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**(3), 1383–1394.

Wu, R. S., Luo, J. R., Wu, B. Y., 2014, Seismic envelope inversion and modulation signal model: Geophysics, **79**(3), WA13−WA24.

Zhou, B., Gao, L., and Dai, Y.H., 2006, Gradient methods with adaptive step-lengthsizes: Computational Optimization and Applications, **35**, 69−86.

Zhang, P., Han, L. G., Gong, X. B., et al., 2018, Multi-source elastic full waveform inversion based on the anisotropic total variation constraint: Chinese Journal of Geophysics, **61**(2), 716−732.