

PERFORMANCE OF ACOUSTIC FULL WAVEFORM TOMOGRAPHY FOR DIFFERENT ACQUISITION GEOMETRIES

A. Kurzmann, D. Köhn, A. Przebindowska, N. Nguyen and T. Bohlen

email: andre.kurzmann@geophysik.tu-freiberg.de

keywords: full waveform tomography, reflection seismic imaging

ABSTRACT

For better parameter estimation we develop imaging methods that can exploit the richness of full seismic waveforms. Full waveform tomography (FWT) is a powerful method to reach this goal. In this study, we demonstrate the performance of our new parallel acoustic time-domain code. We present the results for two acquisition geometries: a transmission (random medium model) and a reflection example (the Marmousi model). Using the crosshole configuration, we compare our time-domain inversion results with the frequency-domain results calculated using the FULLWV code by G. Pratt et. al. Both time- and frequency-domain codes produce similar results. Another important aspect is the choice of an adequate starting model. Both examples are used to investigate the role of the starting model. In contrast to the crosshole geometry, the success of the FWT for a reflection geometry is very sensitive to the starting model. In case of the transmission example the FWT can handle rough starting models. Even a homogeneous model can be enough to succeed. However, the inversion of a reflection geometry problem requires a more detailed starting model. Additionally the applicability of different acquisition geometries will be presented for the Marmousi model. This is a test of the robustness of acoustic FWT with a decreasing number of ray paths. After the reduction of the number of sources or receivers the FWT is still able to reconstruct the subsurface properly. It is rather a matter of iterations. The less ray paths the more iterations are necessary. Furthermore, the results for application of a free surface or the usage of a marine streamer geometry show that the FWT can deal with more complex data, which also include multiple-waves.

INTRODUCTION

Although first implementations of the FWT in the 1980s were conducted in the time-domain by (Tarantola, 1984), the frequency-domain version of FWT developed in the 1990s by (Pratt, 1999) has now emerged as an efficient imaging tool. The main advantage of the frequency-domain approach is the possibility of starting the inversion at low frequencies (large-scale structures) and then moving to higher frequency compounds (small-scale structures), thereby realizing a multi-scale approach. The main advantage of the time-domain method is the efficient parallelization by domain decomposition leading to a significant speedup on parallel computers (Bohlen, 2002).

In this work we focus on the applicability of the FWT to different acquisition geometries. We will analyze the performance for a transmission and a reflection geometry. The first section describes the basic steps of the iterative optimization algorithm of the FWT. The second chapter concentrates on a crucial part of this algorithm - the estimation of an optimal step length. Section three deals with the application of FWT to a transmission geometry. A comparison of the results of time- and frequency-domain implementations completes this part. In section four we use a realistic geological model to focus on the investigation of the reflection geometry. We compare the results for different starting models and acquisition geometries

(e.g. a marine streamer geometry). This part will be finished with a summary of the performance of our time-domain code by using the Marmousi model (Versteeg, 1994).

1. THE BASIC WAVEFORM TOMOGRAPHY ALGORITHM

The aim of the FWT-algorithm is to minimize the residual of observed and modelled data. As described in (Tarantola, 1984), waveform tomography in the time-domain is an optimization problem which can be expressed by the general equation

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \mu_n \mathbf{d}_n.$$

This iterative algorithm comprises several steps:

1. Setup of the source receiver geometry,
2. The choice of an appropriate acoustic starting model \mathbf{m}_0 , i.e. a P-wave velocity model,
3. Apply acoustic forward modelling to the model \mathbf{m}_n (e.g. starting model) at iteration step n for each source.
4. Compute the residual of observed data and forward modelled seismograms,
5. Apply the backpropagation of the wavefield by using the residuals as source signals at receiver positions,
6. Compute the zero-lag cross-correlation of forward modelled and backpropagated wavefields for every source,
7. Sum up all sensitivities to get the gradient \mathbf{d}_n of the entire acquisition geometry and apply the steepest descent gradient method,
8. By multiplying the gradient with an optimal step length μ_n the change in material parameters can be calculated, which yields the updated and improved model \mathbf{m}_{n+1} ,
9. Continue with iteration step $n + 1$ by repeating steps 3 to 8,

The determination of the step length is crucial for the success of the inversion process, i.e. both fast convergence of the data misfit function and model adaption after a low number of iterations.

2. STEP LENGTH OPTIMIZATION

There are different ways to perform the model update (Figure 1(b)). The simplest one is the usage of a constant step length. It is recommended to use a model change of less than one percent (Pica (1990)). The inversion will be successful, but the minimization of the data residual requires a lot of iteration steps. However, the application of a large step length will cause a turbulent behavior of the misfit function. The algorithm permanently jumps over the desired minimum and fails to find an adequate model. Thus the usage of constant step lengths is very inefficient. This can be improved by reducing the step length (for example by a factor of 0.5) when 'hitting' a hill of the objective function. This method is able to get close to the minimum after a lower number of iteration steps. A far more efficient strategy is the application of an adaptive step length control. Every iteration three additional test runs are performed, yielding data misfits $L_{2,i}(\mu_i)$ for three different step lengths μ_i ($i = 0, 1, 2$) around an initial value. A parabolic curve fitting - applied to $L_{2,i}(\mu_i)$ - can be used to estimate an optimal step length (Figure 1(a)). The evolution of the misfit function is very smooth and depending on the local nature of the objective function it is possible to use large step lengths. This involves the reduction of both the number of necessary iterations and the computation time. This step length optimization requires more computational efforts per iteration, because for each shot three additional forward modellings are necessary. But the usage of an optimal step length yields an intensified decrease of the data misfit function. Hence, it is possible to reduce the number of iterations and a huge amount of forward modellings, respectively.

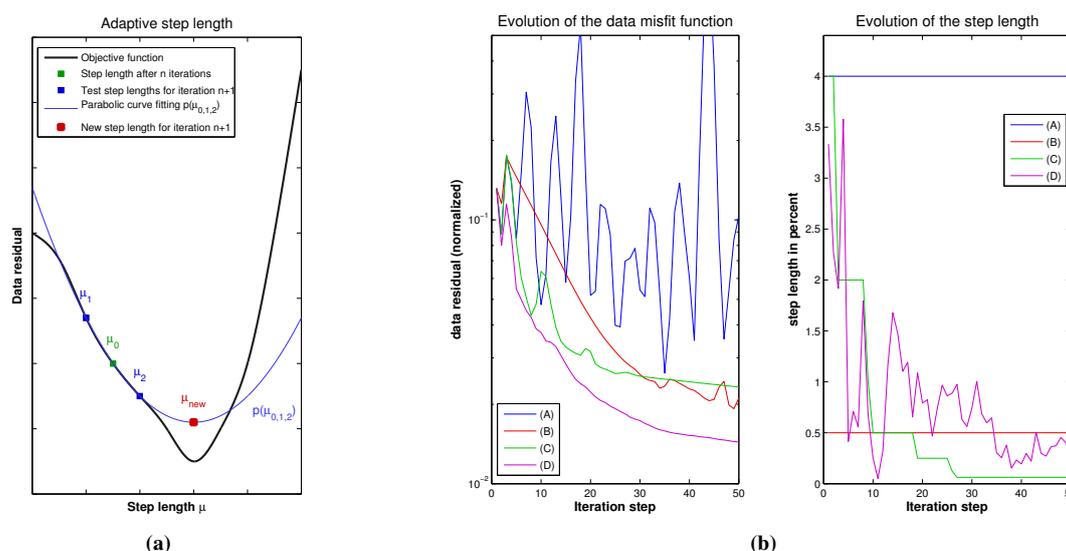


Figure 1: Step length optimization. Figure 1(a) shows the principle of an optimal adaptive step length algorithm. The step length can be obtained by minimizing a parabola, which was fitted to the data residuals of three test step lengths. The evolutions of the data misfit function and the according step lengths are illustrated in Figure 1(b). Graphs (A) and (B) show the results for constant step lengths of 4 % and 0.5 %. (C) represents a step length algorithm using a halving of the step length in case of increasing data residual from iteration step n to $n + 1$. (D) is the result of the adaptive step length algorithm shown in Figure 1(a).

3. TRANSMISSION GEOMETRY

3.1 Model setup

The following inversion problem consists of a random medium, which was generated by applying the von Karmann covariance function (Figure 2). The velocity of the P-wave varies between $1640 \frac{\text{m}}{\text{s}}$ and $2340 \frac{\text{m}}{\text{s}}$. The resulting model is a self-similar medium and thus includes structures on different length scales. The acquisition geometry is arranged in form of a crosshole configuration. It includes 91 receivers and 91 explosive sources with a dominant frequency of 125 Hz. For FD forward modelling we use a record length of 1.4 seconds, a time discretization of $\Delta T = 0.7$ ms and a grid spacing of 0.5 m, which results in a grid size of 520×320 grid points.

3.2 Comparison of time- and frequency-domain results

In order to fit the very long wavelength part, we use a smoothed version of the true model as our starting model (Figure 2). After 25 iterations the result of the time-domain implementation (Figure 3(a)) shows a lot of small-scale features, which can also be found in the true model. Note how well the seismic sections are fitted by the time-domain code (Figure 4). The adaptive step length algorithm was applied to the random medium model. The conditions were made more complicated by using a homogeneous starting model. The results of time- and frequency-domain inversion are shown in Figure 3(b). They are satisfactory but not as close to the true model as the results in Figure 3(a).

Figures 3(a) and 3(b) also show the inversion results, which are computed in the frequency-domain using the FULLWV code by G. Pratt et al. This method only inverts discrete frequencies at a time. The dominant frequency of the source signal is 125 Hz. Hence, the inversion starts at a frequency of 10 Hz, moving to 250 Hz in 10 Hz steps and with 3 iterations at a time. The results of the time- and frequency-domain inversions are comparable, even though the time-domain code inverted the whole frequency information at once, while the frequency-domain inverts for the low frequencies first and added higher frequency content later on.

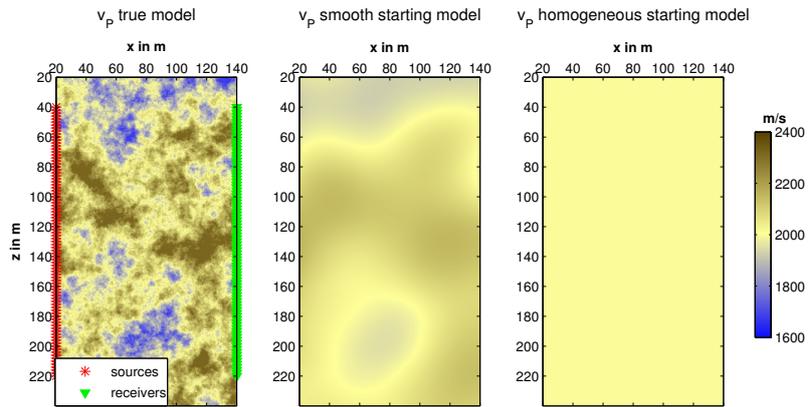


Figure 2: Crosswell-configuration: True v_P model; a smooth and a homogeneous starting model for full waveform inversion (from left to right). The smooth starting model was generated by the application of a 2D averaging filter (size: 60×60 m).

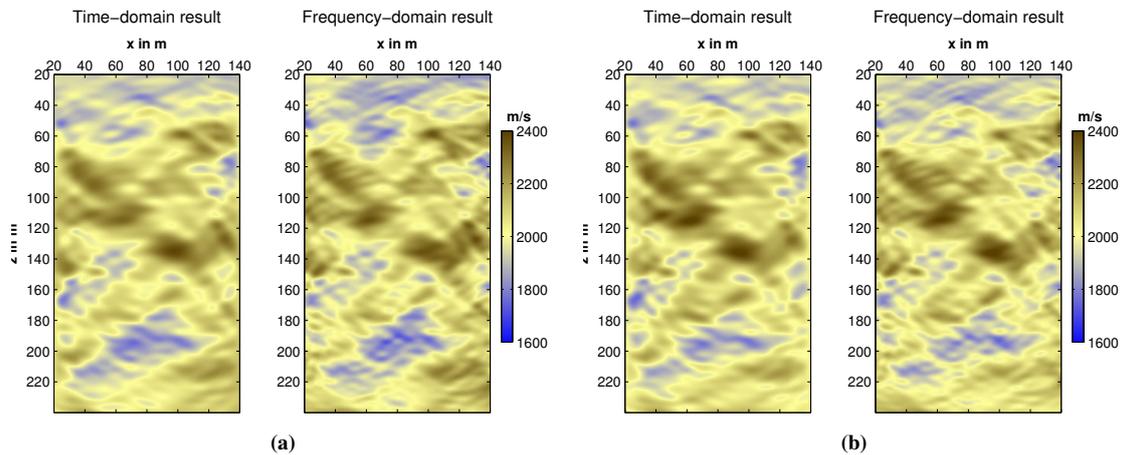


Figure 3: Compilation of random medium P-velocity models, which comprises the inversion results for the smooth starting model (Figure 3(a)) and the homogeneous starting model (Figure 3(b)). The results were obtained after 25 iterations in the time-domain and 75 iterations in the frequency-domain (25 frequencies from 10 Hz to 250 Hz with 3 iterations for each frequency).

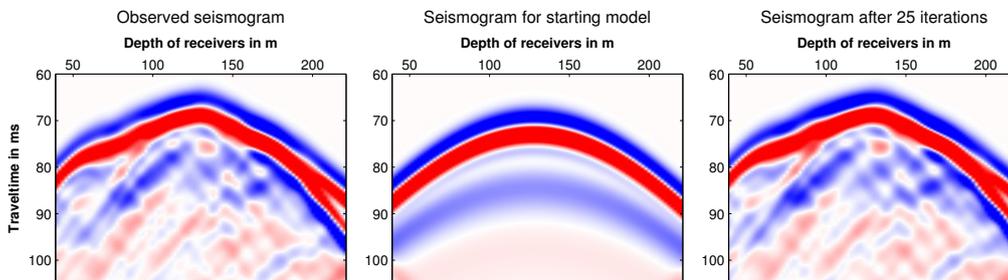


Figure 4: Seismograms for the FWT in the time-domain. The images show the observed data (left), the forward modelled seismogram for the homogeneous starting model (center) and the data, which belongs to the inversion result (right). These examples correspond to the central shot 45.

4. REFLECTION GEOMETRY

4.1 Model setup

In this case the inversion will be demonstrated on the basis of a more realistic model, the Marmousi model (Figure 5(a)). It contains a lot of different geological structures, whose P-wave velocities vary in a range from $1500 \frac{\text{m}}{\text{s}}$ to $5500 \frac{\text{m}}{\text{s}}$. This synthetic example contains an absorbing boundary instead of a free surface at the top of the model. Sources and receivers are located in a depth of 300 m below the absorbing frame. We used 87 receivers and 32 explosive sources with a dominant frequency of 10 Hz. The record length is 3 seconds, the time discretization $\Delta T = 1.0$ ms, the grid spacing 18 m and the grid size 512×178 points.

4.2 Results for different starting models

The first starting model has been calculated from the true model by choosing vertical velocity profiles from three different boreholes (see Figure 5(a)). The areas between these boreholes were interpolated linearly and smoothed by applying a 2D Gaussian filter (standard deviation $\sigma = 10$, size: 990×990 m). The image at the top of Figure 5(b) illustrates the inversion result after 200 iterations. Due to the acquisition geometry and the low ray coverage we can see a loss of resolution with increasing depth. But on the whole the result demonstrates the efficiency of full waveform tomography. Especially in the upper parts it is possible to resolve structures at subwavelength-scale. However, there are still artefacts in the central part of the model, which is just a matter of iterations. Another 100 iterations would be necessary to reconstruct the relevant structures. The satisfactory progress of the inversion is represented by the well-fitted seismogram sections in Figure 6.

The second starting model is a 1D model, which has been derived from the true model and smoothed by the filter, mentioned above. In the image at the bottom of Figure 5(b) the final velocity model after 200 iterations is shown. Just the right part has been reconstructed satisfactorily. In this case there might be too high velocity deviations between true and starting model, which preferably affects areas with low ray coverage. Thus, the algorithm falls into a local minimum of the objective function and is unable to fit low wavelength structures at the very beginning of the iteration process.

The success of the inversion depends on the choice of the starting model. The usage of an adequate smooth starting model (for example the result of a traveltimes tomography) is essential for an efficient full waveform tomography. Apart from the starting model issue, the application of a free surface may cause additional problems. In comparison to the previous velocity models, Figure 7(a) shows the result for a Marmousi model, which includes a free surface. Due to the occurrence of multiple-waves there is no doubt that the observed data becomes more complex. Obviously this affects the quality of the inversion result negatively and in contrast to the examples without using a free surface, a quite large data misfit (Figure 7(b)) can be observed. Without any further optimizations of the algorithm many iterations will be necessary to get a comparable result as shown in Figure 5(b).

4.3 Comparison of different acquisition geometries

This section will focus on the reflection experiment, i.e. the Marmousi model. The aim is the verification of the robustness and stability of the acoustic full waveform algorithm in dependence of the source-receiver-geometry. In a first test - shown in Figures 5(b) and 8(a) - we vary the number of sources (32, 16, 8, 4). The inversion results are compared after 200 iterations. In the second test the number of sources remains unchanged (32) and the number of receivers will be reduced, i.e. 87 and 22. The results are also obtained after 200 iterations (Figure 8(a)). The number of both sources and receivers are considered to be equivalent. The total number of ray paths is much more important and crucial for the success of the inversion. Furthermore it is also a matter of iterations. The less ray paths the more iterations are necessary to get a result with comparable quality. The last test deals with a marine streamer geometry. 28 sources and 40 receivers per source with an offset of 500 m are used. Due to the addition of the water layer the modelling parameters have to be modified: The record length is set to 3.5 seconds and the new grid size is 512×212 points. In fact, this is a small acquisition geometry, but it is the very first experiment of this type. Figure 8(b) shows the inversion result after 200 iterations. The dominant large-scale structures and even some subwavelength structures are resolved satisfactorily.

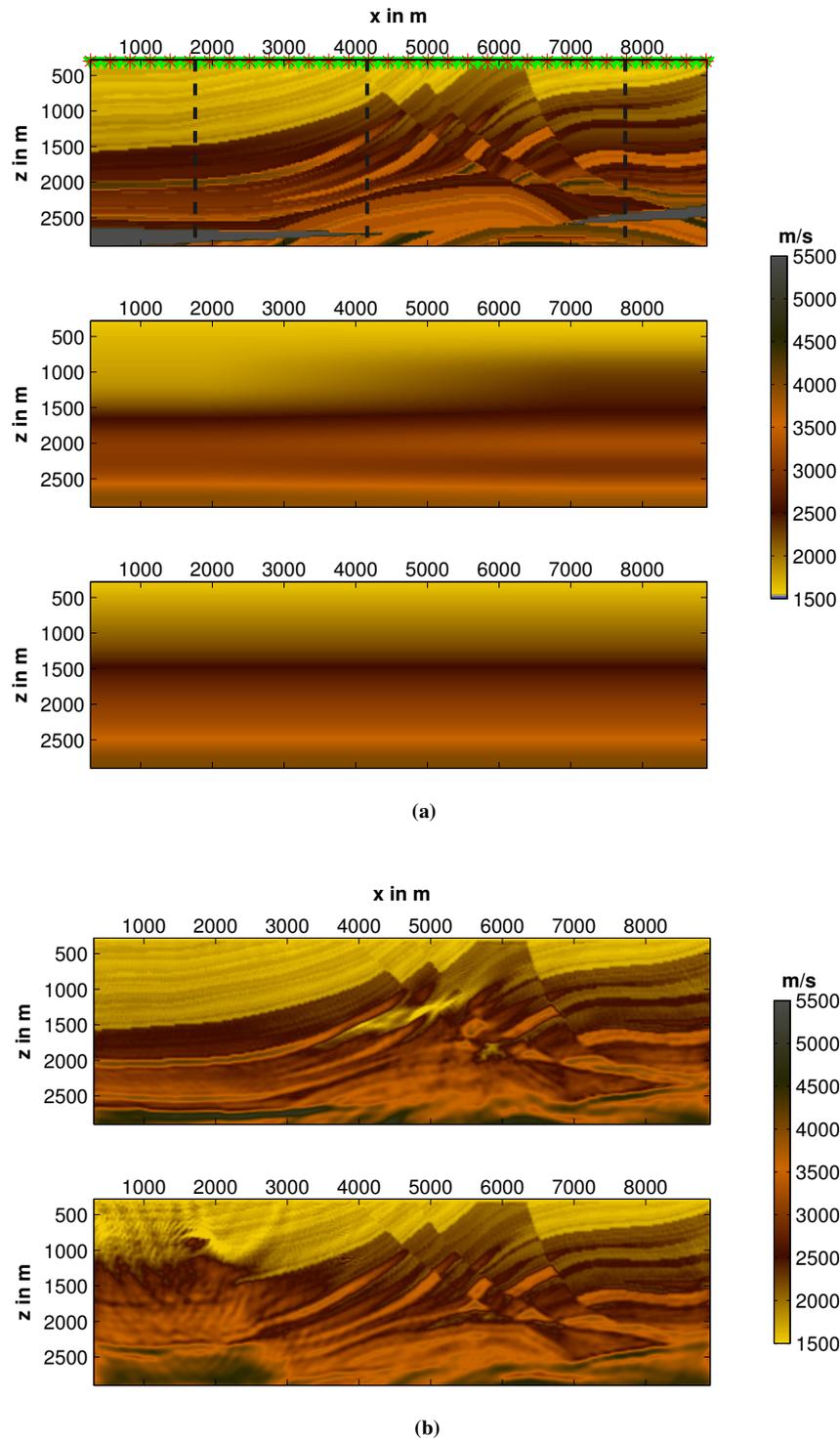


Figure 5: Figure 5(a) shows the true Marmousi model (top) and two different starting models (center and bottom). The upper starting model was generated by using three vertical seismic profiles (dashed lines in the true model image). The interjacent areas were interpolated and smoothed. The lower one is a simple smoothed 1D model. The acquisition geometry is illustrated by red asterisks (sources) and green triangles (receivers). The upper image in Figure 5(b) shows the inversion result for the smooth starting model and the lower image the velocity model for the 1D starting model. All results were obtained after 200 iterations.

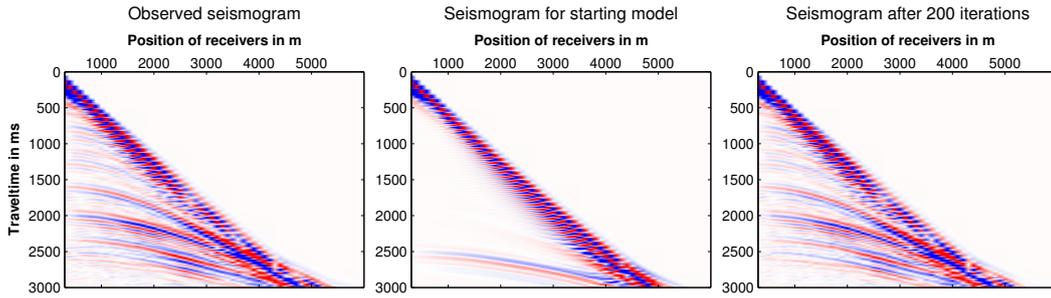


Figure 6: Seismogram sections for the full waveform inversion in the time-domain. The images show the observed data (left), the forward modelled seismogram for the smooth starting model (center) and the data, which belongs to the inversion result (right). These examples correspond to shot 1 at the left boundary of the model area.

4.4 Computational aspects

The computational efforts of time-domain full waveform inversion are very expensive. This relates to both time and memory consumption. This can be demonstrated by using a quite small example. Consider a model grid size of 460×146 grid points and 3300 time steps. As mentioned above it is necessary to compute forward modelled and backpropagated wavefields. Then the cross-correlation of those will be calculated. Thus, for every time step of the FD modelling the forward wavefields have to be saved in memory. In this case they require around 845 MB of memory. There are different possibilities to save memory and time. The memory usage can be reduced by omitting time steps, which also affects the computation time positively. In general it is unavoidable to use parallel computing. This comprises the domain decomposition (Bohlen, 2002) and the shot parallelization. Forward modellings and backpropagations for each shot of a full waveform tomography can be computed separately. Hence, on condition of the availability of computers with enough random access memory, it is sensible to distribute shots among all available computers. In case of multi-core computers domain decomposition can be done internally. As a consequence the tremendous reduction of network traffic and a speedup of the inversion algorithm (see Figure 9) can be observed. For instance we can use a cluster of 8 computers with a total number of 32 CPUs. The first possibility is the exclusive application of domain decomposition, which is the low-memory option. The domain is divided into 32 subdomains, i.e. a memory usage of 106 MB per computer. However, due to large network traffic the FWT will be done at the expense of the performance. The second possibility is the combination of domain decomposition and shot parallelization. The domain can be divided into 4 parts on every single computer, i.e. a memory usage of 845 MB per computer and in total 6760 MB. Thus, the modellings for 8 shots can be computed at once. This high memory consumption comes along with an increased performance of the FWT. Table 1 contains detailed performance information for a small Marmousi example.

Model and inversion parameters	model size: 460×146 grid points 3300 time steps for every forward modelling 150 iterations 16 sources in total: 7446 forward modellings
Hardware	32 CPUs (8 quadcores á 2.6 GHz) gigabit network
Performance	computation time per forward modelling: 0.4 seconds computation time per iteration: 19.5 seconds in total: 49 minutes

Table 1: Detailed performance information of the acoustic time-domain code for a small Marmousi model. For this purpose we used a Linux cluster at the Institute of geophysics of the TU Bergakademie Freiberg.

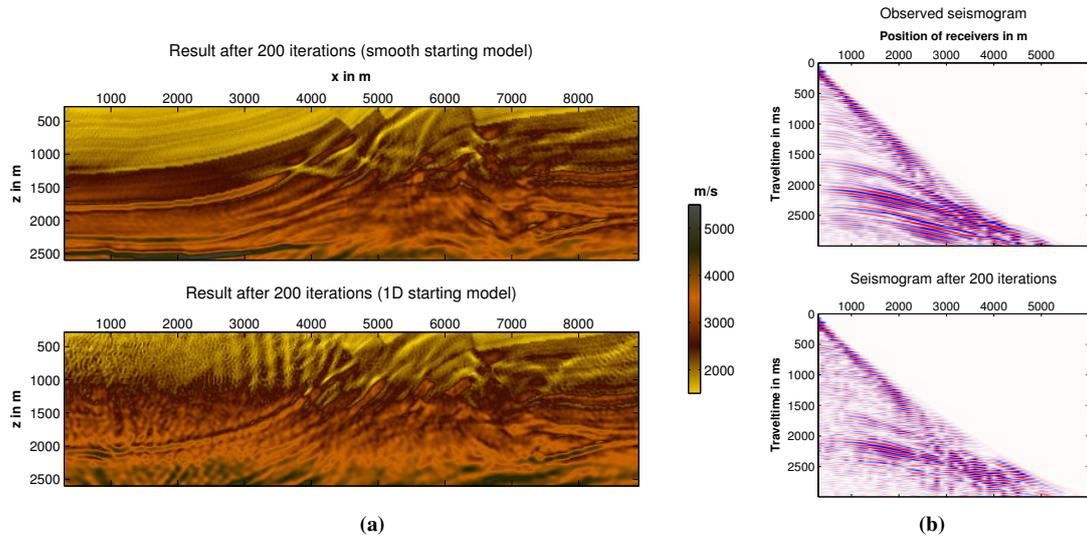


Figure 7: Results with application of a free surface. Figure 7(a) shows the inversion results for the smooth and 1D starting model. The results were obtained after 200 iterations. Figure 7(b) illustrates the observed and forward modelled data, which belong to the smooth starting model.

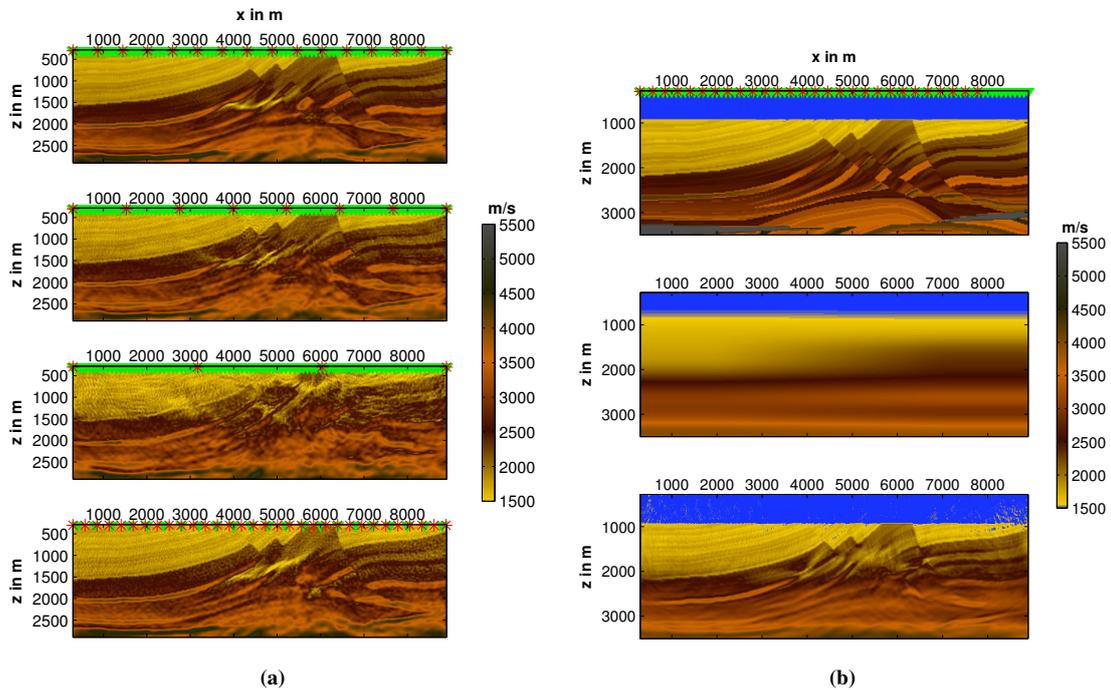


Figure 8: Figure 8(a) contains a compilation of the results for a variable number of shots and receivers ($N_{TX} = 32$; $N_{RX} = 87$; 2784 ray paths). From top to bottom: In comparison to Figure 5(b) the number of shots and receivers varied as follows: $N_{TX} = 16, 8, 4, 32$; $N_{RX} = 87, 87, 87, 22$; 1392, 696, 348 and 704 ray paths. The results were obtained after 200 iterations. Figure 8(b) shows the inversion result for the application of a marine streamer geometry. From top to bottom: true velocity model, smooth starting model and the result after 200 iterations. The acquisition geometry consists of 28 sources and 40 receivers per source (except close to the right boundary), which involves 915 ray paths.

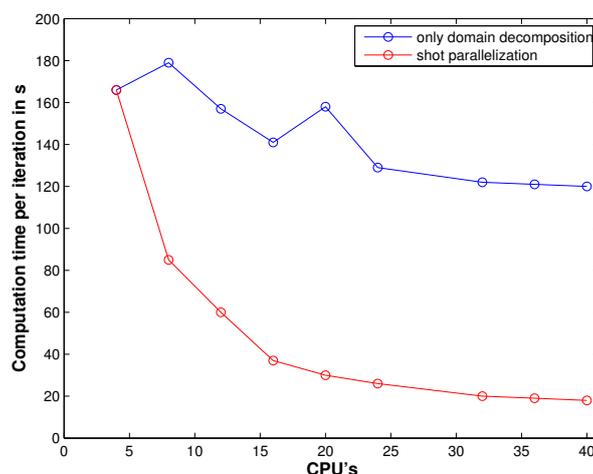


Figure 9: Comparison of the performance of conventional parallelization and a combination of domain decomposition and shot parallelization.

CONCLUSIONS

In this study we tested the performance of our time-domain waveform tomography code and compared the results with the frequency-domain implementation by G. Pratt. The time-domain and frequency-domain codes produce similar results, even for very complex model geometries. The computational cost of the time-domain code is much more expensive than that of the frequency-domain code. However, the runtime for the random medium problem of both codes is comparable (around 40 minutes). This demonstrates the effectiveness of the frequency-domain code on 1 CPU and the parallelized time-domain code on 40 CPUs. Note: Due to stability reasons different model sizes (time-domain: 320×520 , frequency-domain: 160×260 gridpoints) were used. Of course, they also affect the computation times. Especially the reduction of network communication by introducing the shot parallelization and the reduction of iterations by applying an adaptive step length algorithm improve the performance of the time-domain code.

Additionally we have shown that acoustic FWT is applicable to reflection seismic problems. The synthetic Marmousi model is the first step to invert real data later on. For this purpose it will be necessary to find an adequate starting model, which can be estimated by using traveltime tomography. Furthermore we have to take into account different physical effects like geometrical spreading. Another issue will be the source signal inversion and the examination of the applicability of acoustic full waveform tomography to elastic data.

REFERENCES

- Bohlen, T. (2002). Parallel 3d viscoelastic finite difference seismic modelling. *Computers and Geosciences*, 28:887–899.
- Pica, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, 55:284–292.
- Pratt, R. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, 64:888–901.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49:1259–1266.
- Versteeg, R. (1994). The marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge*, 13:927–936.