# Chapter 2

# Building initial models by WTI

In this chapter I discuss a procedure of initial model building for early-arrival waveform inversion. Waveform inversion is a highly non-quadratic inverse problem because the objective function is highly non-linear with respect to the model perturbation. This means that the waveform inversion objective function contains many local minima which can easily make a gradient based inversion method converge to the wrong solution. The only way to ensure convergence to the global minimum with a gradient based method is to have a good starting model that is close to the correct model. Such a good starting model can be provided using many methods. Based on the physics used in tomography, these methods can be categorized as ray and wave-equation based methods. Each category can be further divided into data-domain methods and image-domain methods based on whether the objective function is measured in the data space or measured in the image space. In this chapter, I will show that for near-surface model building, especially in geologically complex areas, a good starting model can be provided by a wave-equation based, data domain method called wave-equation traveltime inversion.

## INTRODUCTION

It is well known that good starting models are essential for successful waveform inversion applications (Virieux and Operto, 2009). As a data-domain tomographic method, waveform inversion updates velocity using mismatches between observed data and modeled data. This is usually conducted in an iterative fashion (Crase et al., 1990; Akcelik, 2002; Brossier et al., 2009), where data mismatches become smaller as waveform inversion results get closer to the true model. The data mismatches include traveltime mismatches and waveform mismatches. Traveltime mismatches tend to dominate data mismatches for velocity models that are far from the true model. However, traveltime mismatches start to decrease as velocity models get closer to the true model. When the velocity models become very close to the true models, the waveform mismatches become the dominant data mismatch. The change in the dominant data mismatch occurs because traveltime mismatches are usually caused by errors in the long-wavelength components of the velocity model, whereas waveform mismatches usually reflect errors in the short-wavelength components of the velocity model. Waveform Inversion (WI) (Tarantola, 1984; Mora, 1987; Pratt et al., 1998) mainly corrects for waveform mismatches, hence the final velocity model has high-resolution. However, WI implicitly assumes small traveltime errors when correcting for waveform mismatches. This assumption leads to the stringent requirement of having a starting model with the correct long-wavelength component.

Acceptable starting models can be obtained by traveltime-based tomography methods, which only minimize the traveltime mismatches between observed data and modeled data. Because forward modeling of traveltimes can be done by either shooting rays or propagating waves using wave-equations, both ray based and wave-equation based traveltime inversion (Hampson and Russell, 1984; Olson, 1984; White, 1989; Luo and Schuster, 1991) can be used to build starting models for the waveform inversion. In many practical applications, traveltime inversion is the preferred choice of building initial models for waveform inversions (Sirgue et al., 2009; Brenders, 2011; Shen et al., 2012). On the other hand, the complexity of the geological settings determines the applicability of ray and wave equation traveltime inversion methods.

Because rays are only good approximations for wave-paths in simple geological settings, ray-based traveltime inversion can produce a good enough starting model for WI in those areas. However, in geologically complex areas, where rays fail to approximate wave-paths, Wave-equation Traveltime Inversion (WTI) (Woodward, 1990; Luo and Schuster, 1991) is the better tool for obtaining a good starting model for WI.

WTI uses the wave-equation to forward model the synthetic data, and aims at minimizing the traveltime differences between observed data and modeled data. This is different from minimizing the waveform differences in WI. The key difference is the data mismatch minimization, and it is what makes WTI much more tolerant of the starting model than WI of its starting model. In the model space it is mainly the tolerance of errors in the long-wavelength component. This is illustrated in the following simple comparison among three gradients: WI gradient for short-wavelength model error, WI gradient for long-wavelength model error, WTI gradient for long-wavelength model error. In this example, long-wavelength model errors and short-wavelength model errors are represented by large and small Gaussian velocity anomalies, respectively. The background model has laterally invariant velocity that increases vertically. Two circular Gaussian velocity anomalies were used as velocity perturbations. One anomaly has a radius of 100m, the other has a radius of 450m (Figure 2.2). True models consist of the background model with each of the two Gaussian velocity anomalies, while the background model is the initial model used for inverting each of the two true models. Synthetic surveys were performed on the two true models and the initial model, with a total of 60 shots and 100m shot spacing. The source wavelet is a Ricker wavelet of 10 Hz peak frequency. The peak frequency of the source wavelet causes in the extent of the big anomaly being three to four times the spatial wavelength of the dominant signal — a long-wavelength anomaly — and the extent of the small anomaly being the same spatial wavelength as the dominant signal — a short-wavelength anomaly. By comparing modeled data from the three surveys, the influence of the short and long wavelength Gaussian anomalies can be observed in the shot gathers of the same source location (Fig 2.3). Their effects are obvious when comparing recorded traces at 4 km horizontal location (Figure 2.4). The short-wavelength anomaly (acted as a diffractor) mainly resulted in a waveform mismatch

while the long-wavelength anomaly mainly resulted in a traveltime mismatch. For such observed data, WI gradient has the correct sign for the short-wavelength velocity error (Figure 2.5 top), yet the WI gradient for the long-wavelength error has the wrong sign (Figure 2.5 middle). However, WTI gradient has the correct sign even with the long-wavelength velocity error (Figure 2.5 bottom).

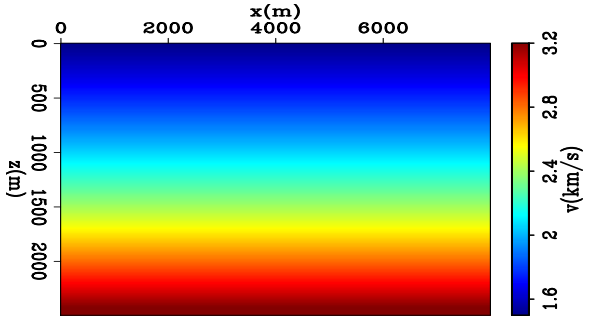


Figure 2.1: Background velocity model for gradient calculation. [**ER**] chap2/. vbg

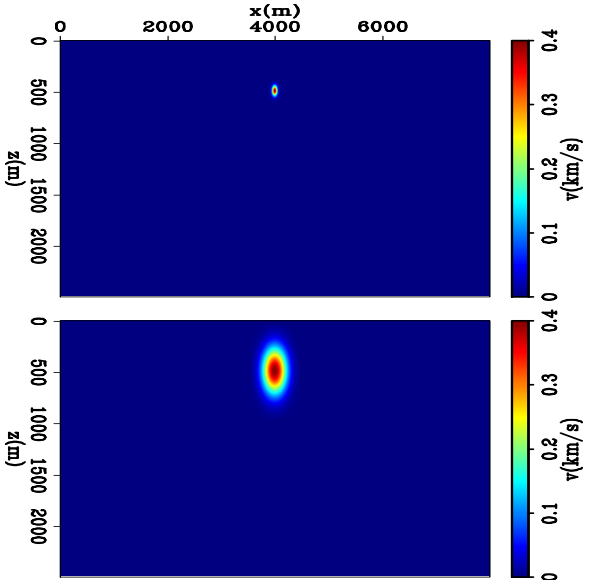


Figure 2.2: Gaussian velocity anomalies, max value at 400 m/s with radius of: top, 100 m; bottom, 450 m. [**ER**] chap2/. vanomcomp

This example suggests that the WI gradient is capable of correcting the short wavelength errors in the model, while the WTI gradient can correct the errors in the

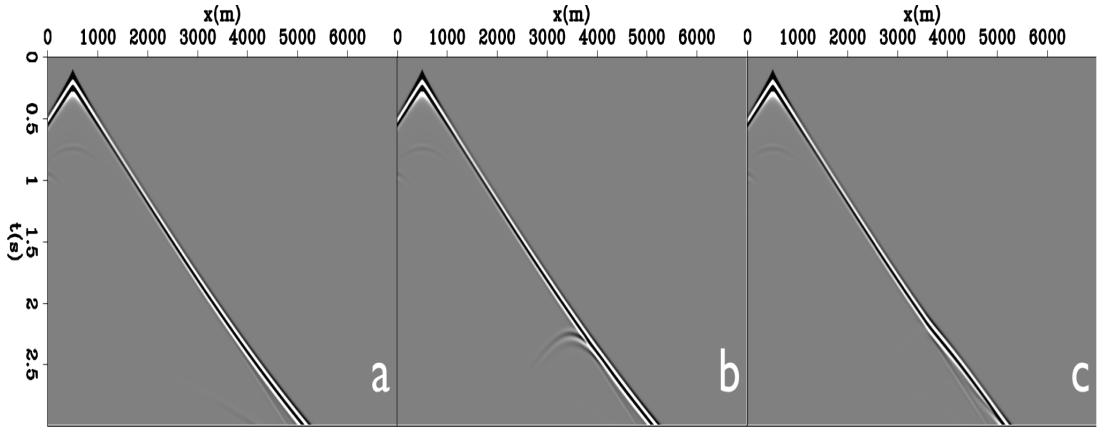


Figure 2.3: Same shot modeled from: a) background model; b) background model with the small Gaussian velocity anomaly; c) background model with the large Gaussian velocity anomaly. [**ER**] chap2/. divrcd

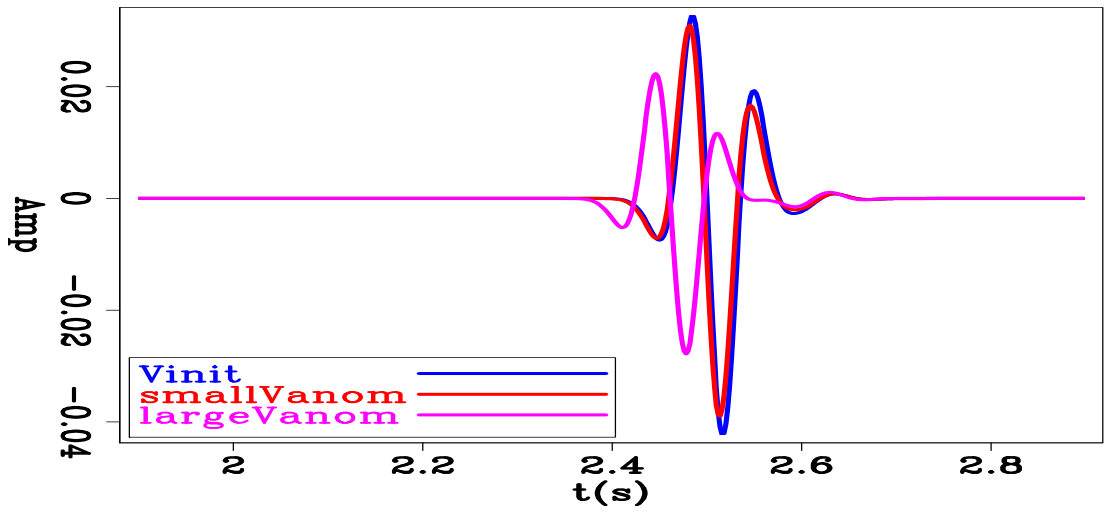


Figure 2.4: Same trace of the same shot location modeled from: a) background model; b) background model with the small Gaussian velocity anomaly; c) background model with the large Gaussian velocity anomaly. [**ER**] chap2/. divrcdtr

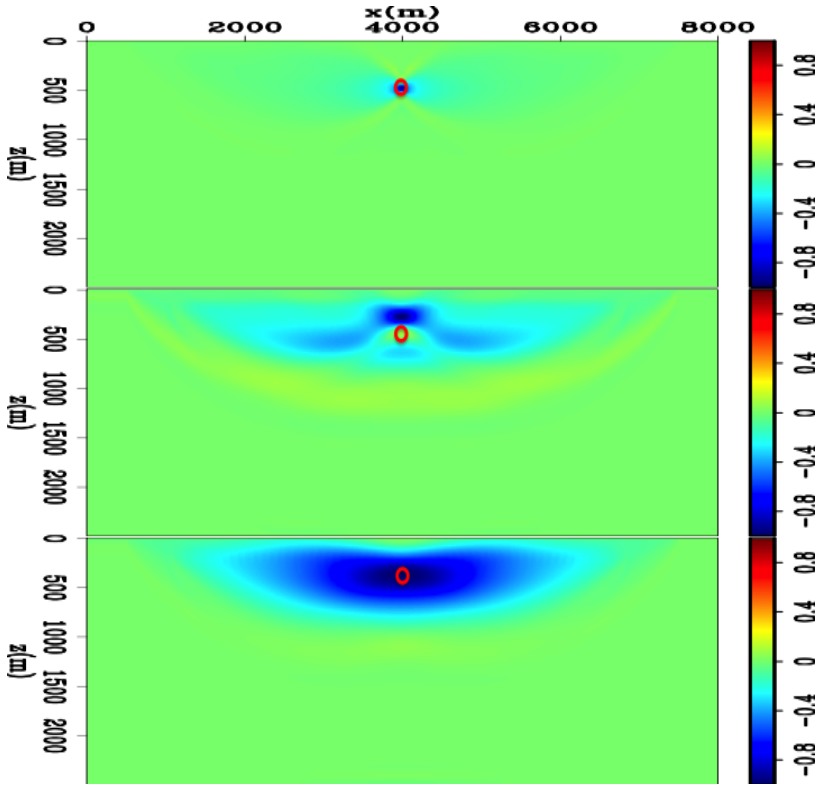


Figure 2.5: First iteration squared slowness gradient (red circles denote the center of the anomalies) from: top, waveform inversion of velocity model with short-wavelength error; middle, waveform inversion of velocity model with long-wavelength error; bottom, wave-equation traveltime inversion of velocity model with long-wavelength error. [**NR**] chap2/. gradnote

long wavelength component in the model where WI fails. However, similar to other tomography methods, WTI can not converge to the true solution from an arbitrary starting model.

# TWO-STEP TOMOGRAPHY

As mentioned before, data-domain wave equation based inversion methods update the velocity models by minimizing the differences between forward modeled data and recorded data. The differences can be traveltime, amplitude, or a combination of the two. Mathematically, a generalized inversion objective function for the purpose of near-surface velocity estimation can be written as:

$$f\left(\mathbf{d_{obs}}, \mathbf{D}(\mathbf{m})\right) \approx \mathbf{0}, \tag{2.1}$$

where $f$ is a function of $\mathbf{d_{obs}}$, observed data, and $\mathbf{D}(\mathbf{m})$, forward modeled synthetic data from $\mathbf{m}$, the near-surface velocity model. The function $f$ plays an important role in the inversion because it determines how non-linear the inversion problem is, how accurate the forward modeling must be, and what resolution is possible in the inversion results.

If only traveltime mismatches are used, the inversion is relatively linear in the sense that there is a more or less linear relationship between the model errors and data misfits. In such cases, the starting model does not have to be very close to the true model, nor must data contain low frequency for the inversion to converge. In addition, the wave equation modeling engine in the inversion does not have to model data amplitudes accurately. WTI achieves this by taking the $f$ as the sum of the L2 norms of the traveltime difference between $\mathbf{D}(\mathbf{m})$ and $\mathbf{d_{obs}}$. The traveltime difference is defined as the time lag that maximizes the cross-correlation of $\mathbf{D}(\mathbf{m})$ and $\mathbf{d_{obs}}$. However, since WTI only corrects long-wavelength model errors, the resulting models have low resolution.

High resolution inversion results can be obtained if full amplitude and traveltime

information is used. In such cases, the inversions are extremely non-linear, and convergence to the correct model requires either very accurate starting models or the inclusion of very low frequency data. Also the wave-equation modeling engine in the inversion must model amplitude with accuracy, otherwise the amplitude difference due to modeling insufficiency will lead to potentially erroneous results. This full amplitude and traveltime matching can be done by taking $\mathbf{f}$ as the L2 norm of $(\mathbf{d_{obs}} - \mathbf{D}(\mathbf{m}))$, the objective function of the conventional waveform inversion (Tarantola, 1984; Pratt et al., 1998). The resulting high resolution earth model is accurate, assuming that we have a good starting model (or low-frequency data) and a complicated wave-equation modeling engine that can model all the physics in the recorded data. In practical applications, amplitude of the entire observed data in a typical survey is very hard to reproduce with reasonable computational expense. Thus an objective function that de-emphasizes amplitude matching and still retains the resolution of WI results is favorable. One of such objective function is called kinematic objective function, which I will use in the examples in this chapter and discuss in more details in the next chapter. Overall, the proposed workflow is as follows:

1) Wave-equation traveltime inversion from the ray tomography result.

2) Kinematic Waveform Inversion (KWI) from step 1) result.

In the first step of the workflow, the geophysical meaning of the WTI is minimizing traveltime differences between corresponding events in the observed early-arrivals and the modeled early-arrivals. One extreme case is where the early-arrivals only contain the first-breaks, in which WTI takes the form of first-break traveltime inversion. In real data applications, first-breaks are usually less contaminated by noise than later events. In addition, advances in data processing have resulted in effective methods for noise elimination of the first-breaks, further increasing the signal-to-noise ratio of the first-breaks. As a result, the strong signal-to-noise ratio makes first-breaks a good candidate for WTI. Moreover, good first-break matching usually indicates a good near-surface model as far as long-wavelength components are concerned (Lawton, 1989; Adams et al., 1994). Taking all these into consideration, a feasible implementation of the first step in the workflow is wave-equation first-break traveltime inversion

starting from the model estimated with ray tomography. Next, I demonstrate the effectiveness of the new workflow with synthetic and field data examples.

# EXAMPLES: ESTIMATING NEAR-SURFACE LOW VELOCITY LAYERS

With more and more onshore data acquisition in various geological environments, it is not uncommon to see shingling in data ( Figure 2.6). Shingling is where first-break traveltimes are only continuous piecewise in a shot gather. This phenomenon is often caused by a low velocity layer in the near-surface ( Figure 2.7, Tong and Yi, personal communication). Such data poses a serious problem for ray-based methods. Through wave-equation modeling, it becomes evident that shingling in the shot gather (Figure 2.6) is caused by the fact that first-break amplitudes decay so fast with offset that after a certain offset, the only reliable pick becomes the next arrival. In other words, the discontinuity is caused by tracking different events as receiver offset increases in the same shot gather.

The shingling phenomenon is further studied by comparing wave-equation modeled shot gather with ray-based traveltime modeling result. A synthetic shot was constructed where the true velocity model has a low velocity layer with a background velocity that smoothly increases with depth. The velocity is laterally invariant. The synthetic data are generated using a constant density acoustic modeling engine. The shot gather (Figure 2.8 left panel) shows shingling as we move further away from the source location. An eikonal equation based first-break traveltime calculation was performed using the same velocity model, and the resulting traveltime was overlaid on top of the wave-equation modeled shot gather (Figure 2.8 right panel). The theoretical first-break traveltime is continuous. However, the amplitude of the first-break decays so much at far offset that the only reliable pick of first-break traveltime is the event that comes later. The amplitude decay is also reflected in the ray density when I trace rays through the model (Figure 2.9). The density of diving rays from the top of the low-velocity layer decreases significantly as the source-receiver offset increases,

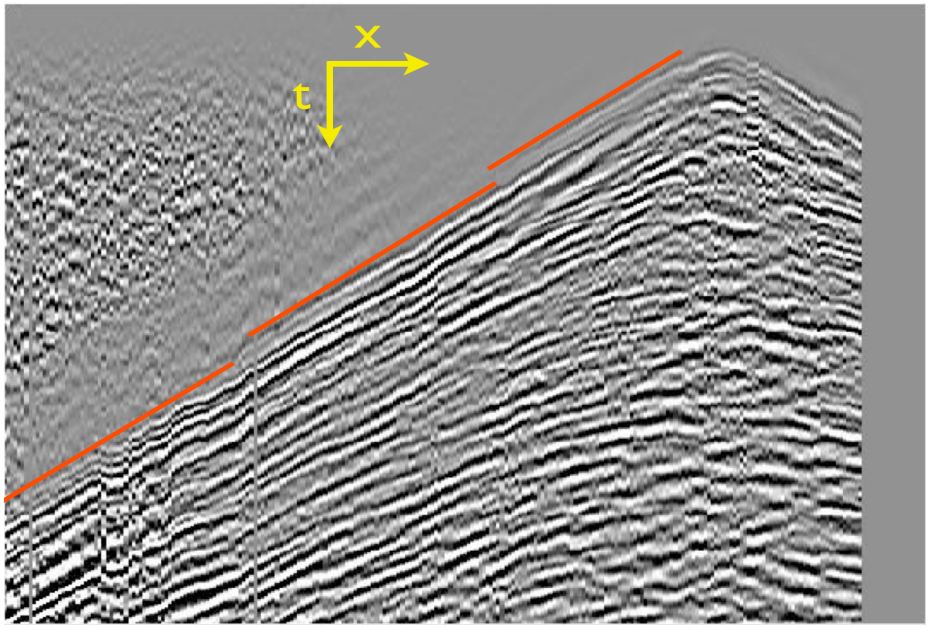


Figure 2.6: Shot gather of 2D land dataset showing the shingling phenomenon. Red lines denote approximate first-break picks that are discontinuous as offset increases. [**NR**] chap2/. realshot

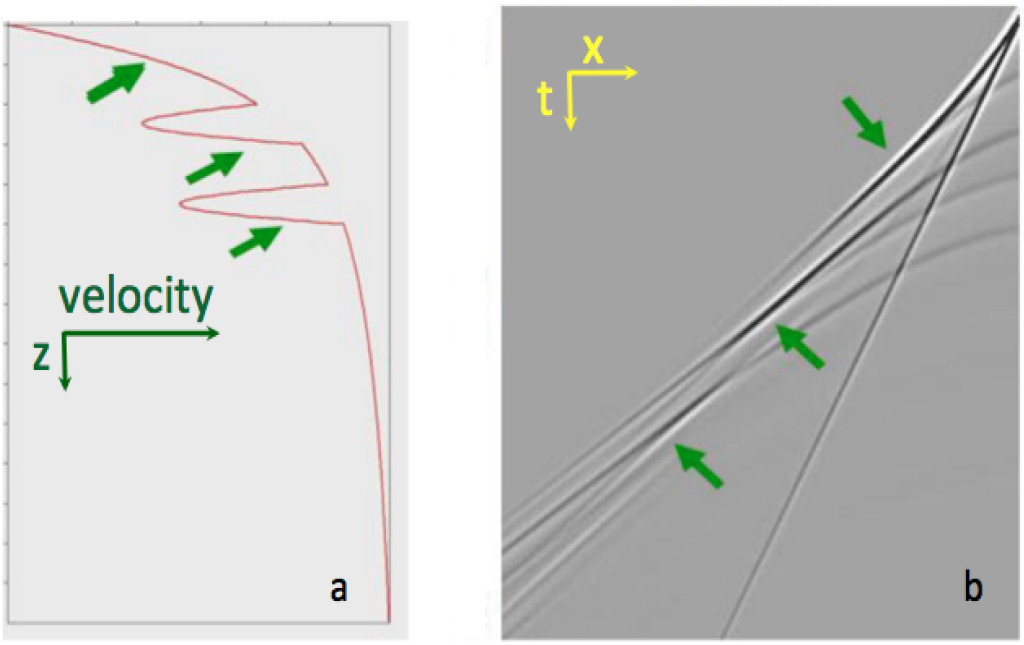


Figure 2.7: a) 1D velocity model with a low velocity layer in the near-surface. b) early-arrivals by acoustic modeling using lateral invariant 1D model in a). The green arrows denote decaying early-arrivals and where they occurred in the near-surface. [**NR**] chap2/. shingling

while the density of diving rays from the bottom of the low-velocity layer increases.

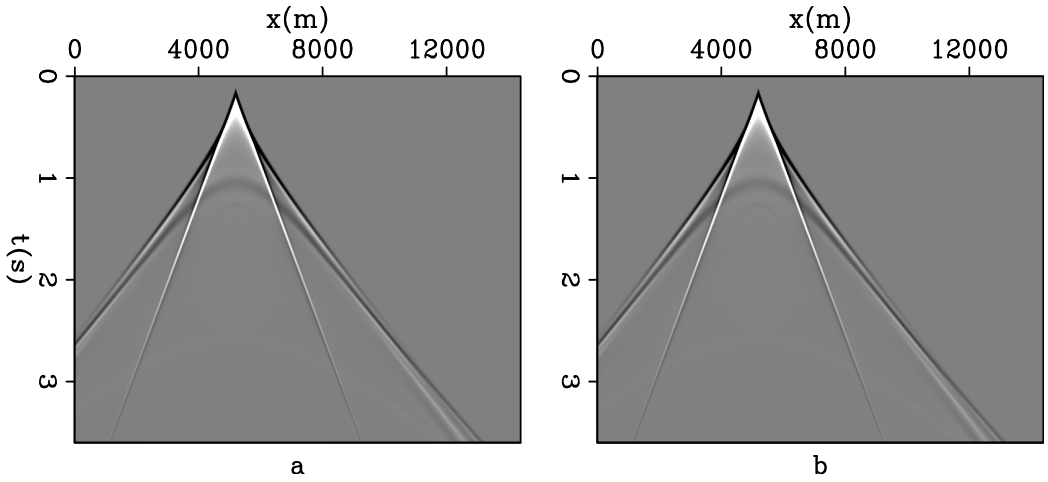


Figure 2.8: Left: synthetic shot gather by acoustic modeling using true velocity model; right: same shot gather with modeled first-break traveltime overlay. [**ER**] chap2/. synshot

In the following, I first show synthetic examples where I invert a synthetic shingling data set using the workflow. I then show results of the same workflow applied to a 2-D field land data set.

## SYNTHETIC DATA APPLICATION

The true model has laterally invariant velocity, with the depth profile shown as the solid blue curve in Figure 2.11. The starting model is also laterally invariant, with the depth profile shown as the solid red curve in Figure 2.11. A total of 56 shots with 240 m shot spacing was generated for the survey. All shots were modeled using the same Ricker wavelet of 10 Hz peak frequency. Receivers were everywhere on the surface. In the inversion, I assumed a known source wavelet and employed the workflow described above. Input early-arrivals and early-arrivals modeled from the starting model, from the traveltime inversion result and from the final result are shown

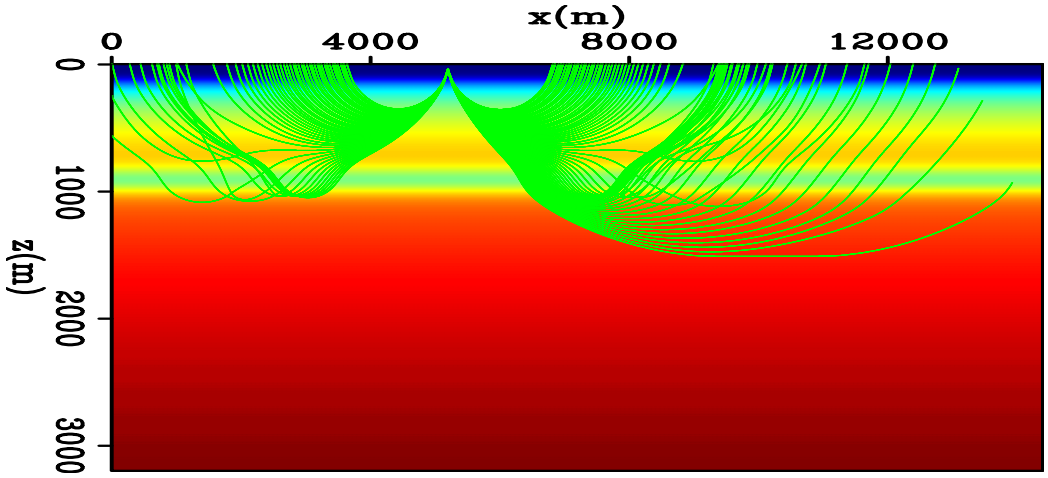


Figure 2.9: Ray-tracing through the true velocity model from the same source location of the shot gather in Figure 2.8. [**ER**] chap2/. synraytracing

in Figure 2.10. Early-arrivals modeled from the starting model matched true early-arrivals at near offsets. The two completely cycle skipped each other at far offsets. The WTI successfully matched both the near and the far offset early-arrivals, but was unable to produce the 'jump' from near offset traveltime to far offset traveltime. However, from the WTI result, early arrival KWI was able to match the 'jump' part of the early-arrivals, thus matching the early-arrivals of all offsets. As mentioned before, the solid blue and red curves shown in Figure 2.11 denote the true velocity model and the starting velocity model, respectively. The solid magenta curve is the traveltime inversion result, and the dashed green curve is the final inversion result. Using the two-step workflow, I was able to recover the low velocity layer with 10 Hz peak frequency source wavelets and data which is generally considered as high frequency for KWI applications. Even with such high frequency signal, the WTI correctly updated the long wavelength component of the near-surface model. The subsequent early-arrival KWI only added the short wavelength component of the low velocity layer. I also tried direct waveform inversion without WTI using the same models. In this case, convergence to the correct solution occurred only with data peak frequencies of 7 Hz or lower.

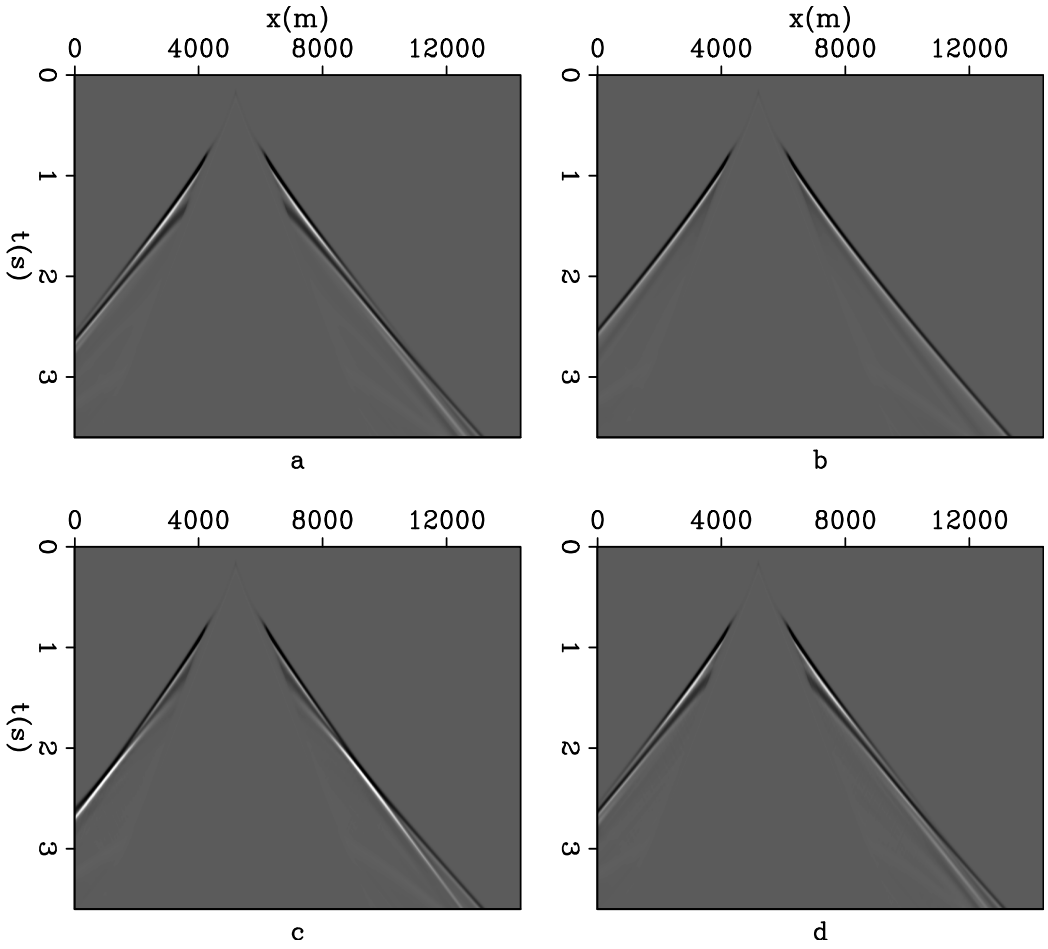


Figure 2.10: early-arrival modeled from : a) true model; b) starting model; c) travel-time inversion result; d) final inversion result. [**CR**] chap2/. synrefracomp

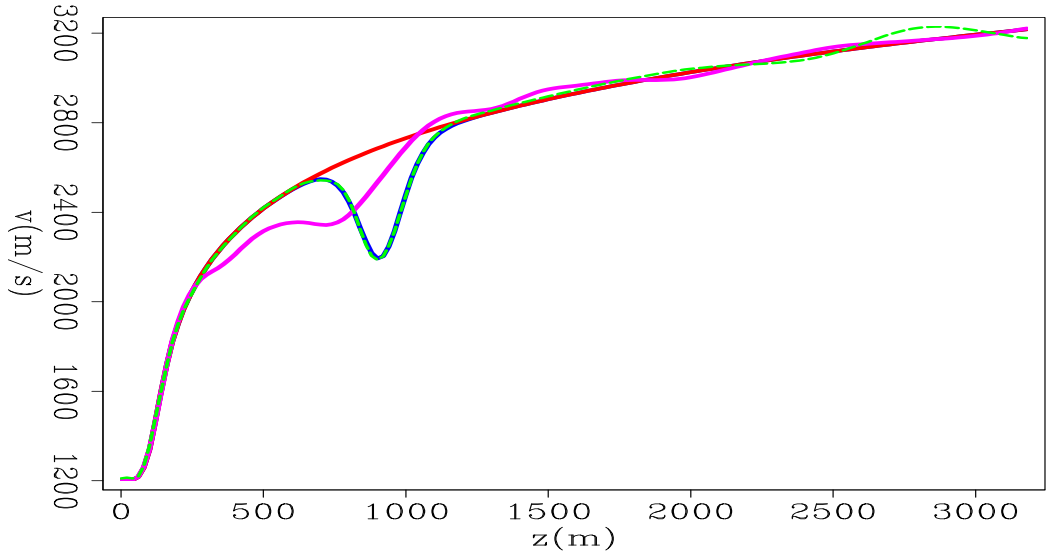


Figure 2.11: Inversion results with blue being true model, red being starting model, magenta being the intermediate inversion result and green being the final inversion result. [**CR**] chap2/. syninv

## FIELD DATA EXAMPLE

The new workflow was also applied to a 2D land dataset acquired in Saudi Arabia. The line geometry and the stacked section of the 2D data are shown in Figure 2.12. The section of the line used for waveform inversion is between the two vertical red lines. The acquisition geometry was not strictly 2D, but it was converted to 2D by a simple projection of the line onto the x-axis. The starting model was obtained by a ray-based method using picked first-break traveltimes of the near offset - hence the shallow depth where the model was estimated. The starting source wavelets were obtained by stacking part of the moved-out near offset early-arrivals, and they varied from shot to shot. The source wavelets were updated at each iteration of the inversion. I used a total of 110 shots, with 180 m shot spacing and 30 m receiver spacing. Offset used for inversion ranged from -4000 to -400 m for each shot. The lowest frequency in the data was 10 Hz, which makes direct application of waveform-based inversion for the long wavelengths of velocity difficult. Two waveform inversion runs were performed for comparison. The first run used KWI directly from the

ray-based tomography result, and the second run used the aforementioned two-step workflow.

The starting model and final results of the two runs are shown in Figure 2.13. The inversion result after the two-step workflow shows a continuous low velocity event in the upper middle portion of the section. This is confirmed by the regional geology where the survey was conducted. On the other hand, direct KWI was unable to detect the low velocity layer.

The reverse time migration (RTM) images in Figure 2.14 clearly show that wave-equation traveltime inversion has improved the final image overall. With better velocities, we get more spatial coherence, particularly in the neighborhood of the low velocity layer where reflections are seen both in the RTM result and on the stacked section in Figure 2.12 (indicated by the red ellipses). However, the image of the very near-surface (down to 200 m depth) become worse because the exclusion of early-arrival less than 400 m source-receiver offset, which explicitly constrain the very near-surface velocity model. Without these early-arrivals, the deeper near-surface velocity model was improved at the expense of the worse very near-surface model.

Figure 2.15 shows ray tracing through the final velocity model within the data offsets used for inversion. These rays suggest first-breaks did travel through, and were able to recover the low-velocity layer. Because waveform matching was conducted in the KWI step, it is also important to compare the modeled data with the input early-arrival data (Figure 2.16). Waveforms and traveltimes match quite well despite differences in absolute amplitude. This demonstrated the importance of matching kinematics over matching both kinematics and amplitude.

## CONCLUSIONS

I proposed to use WTI to build a better starting model for waveform inversion. The WTI is capable of correcting the long-wavelength errors in the ray-tracing tomography result, such errors can not be handled by ray-based traveltime inversion methods or waveform inversion alone. I combined the WTI and KWI into a two-step workflow

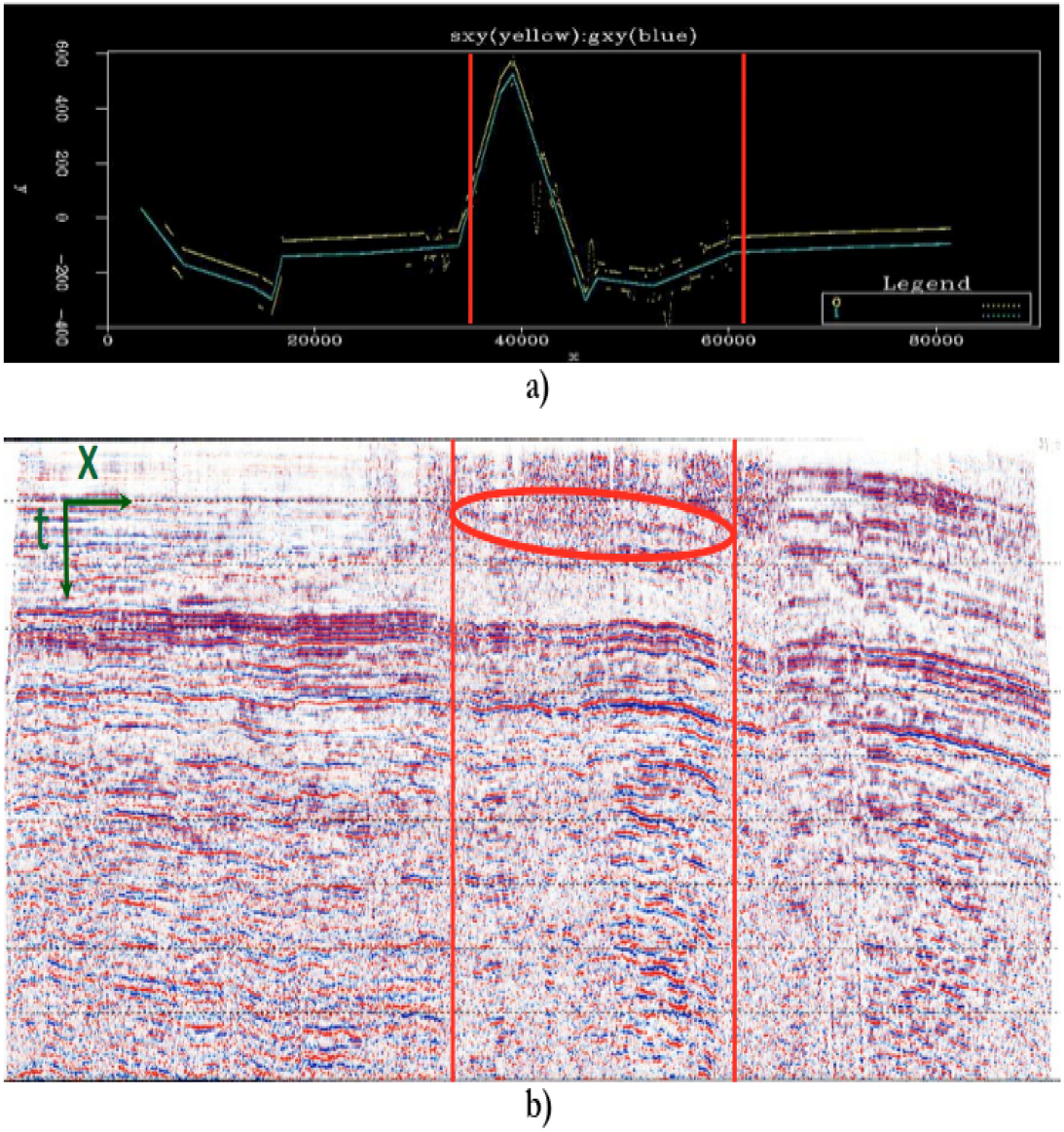


Figure 2.12: A land 2D data case showing a) the x,y source receiver geometry, and b) the stacked. The area for waveform inversion is between the vertical red lines. Blue indicates source, and yellow indicates receiver locations. [**NR**] chap2/. geostack

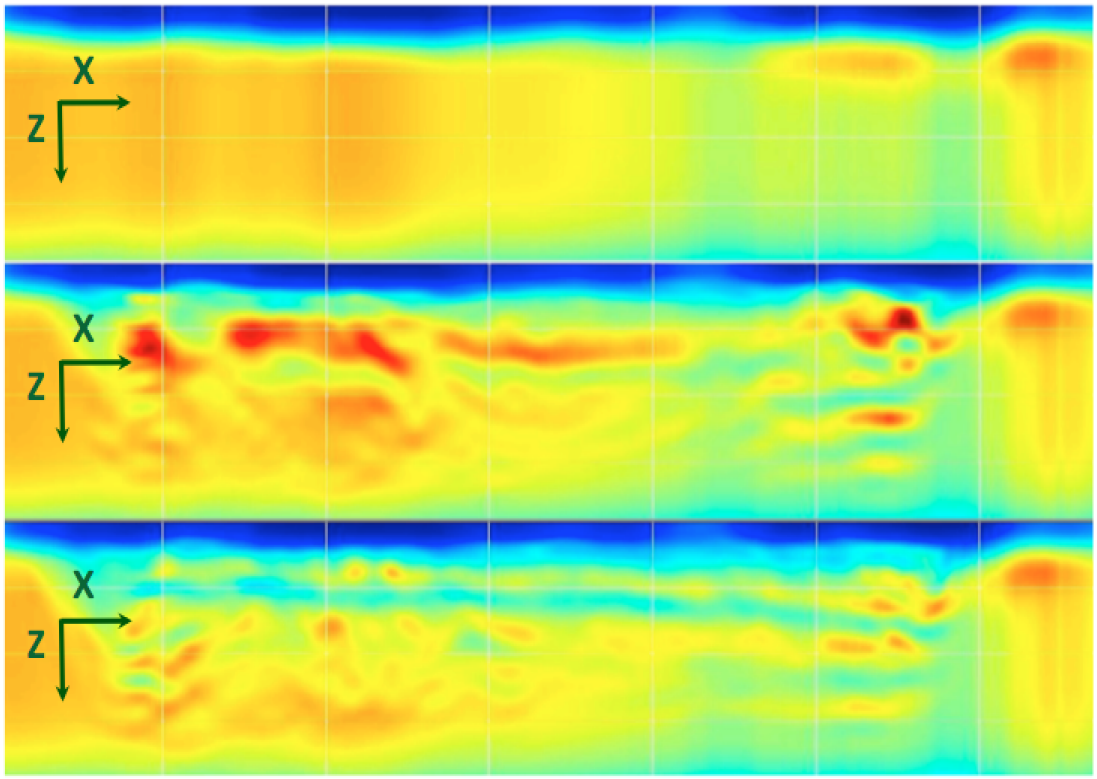


Figure 2.13: Starting depth velocity model using ray-based tomography (top), and waveform inversion results without (middle) and with (bottom) wave-equation traveltime inversion. [**NR**] chap2/. threevel

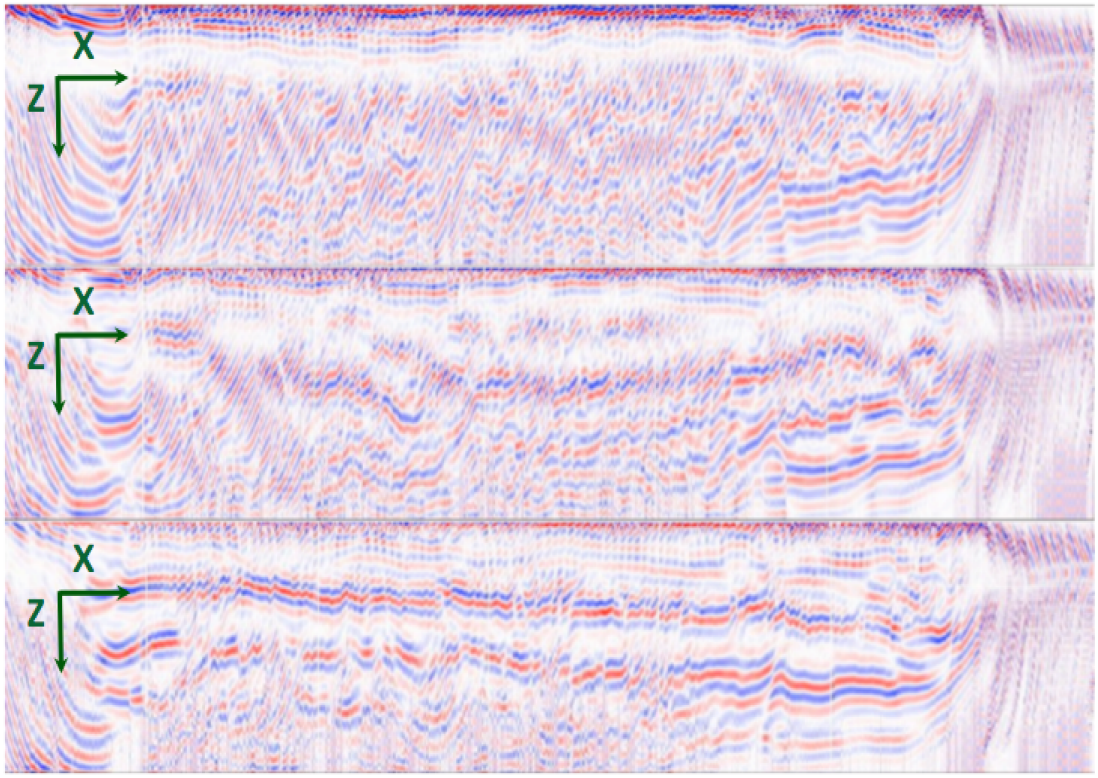


Figure 2.14: RTM images corresponding to the starting depth velocity model using ray-based tomography (top), and to the waveform inversion results without (middle) and with (bottom) wave-equation traveltime inversion. [**NR**] chap2/. threeimg

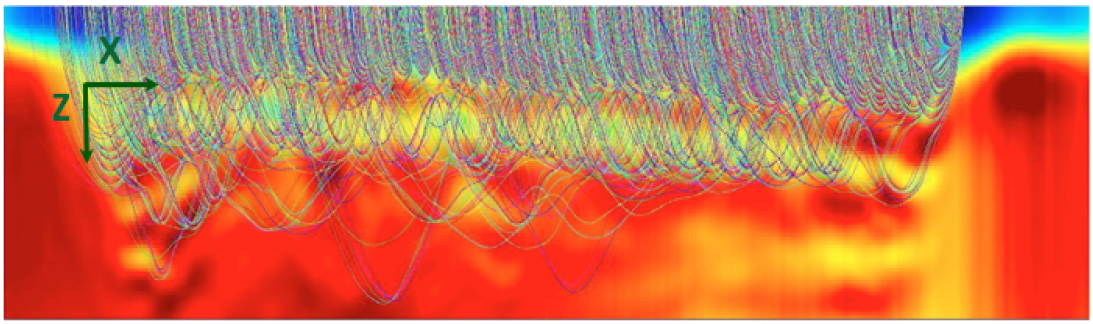


Figure 2.15: Ray tracing through the final velocity model. This shows that the low velocity layer is from inversion instead of being an artifact. [**NR**] chap2/. realray

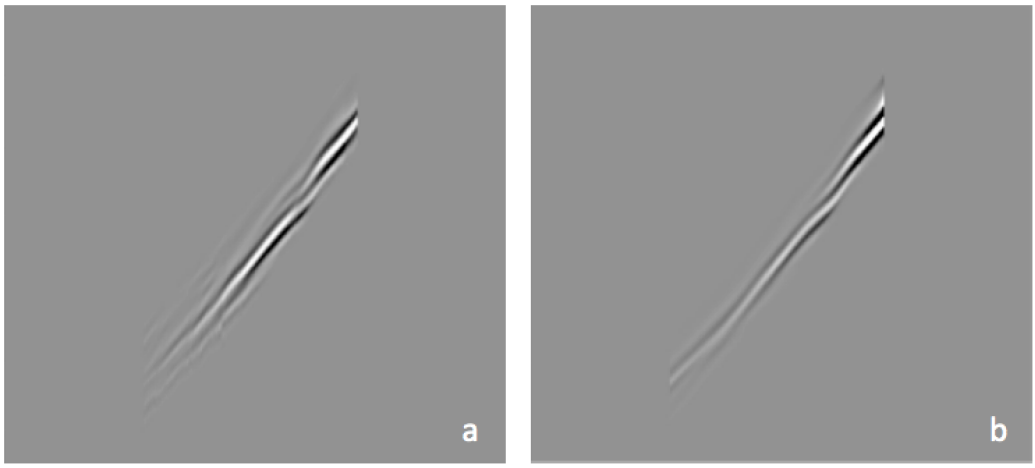


Figure 2.16: Comparison of a) input and b) modeled early-arrivals. Note the similar kinematics despite minor differences in absolute amplitudes. [**NR**] chap2/. realcomp

that is capable of correcting both the long-wavelength and short-wavelength errors in the models. This was demonstrated using both a synthetic data example and a 2D field dataset from Saudi Arabia. The wave-equation traveltime inversion was able to correct the long-wavelength errors in the ray-tracing tomography result, such that subsequent waveform inversion can resolve the low velocity layer in the near-surface. The long-wavelength errors in the starting model is significant enough that waveform inversion alone can not recover the low velocity layer.

## ACKNOWLEDGMENTS

I would like to thank Saudi Aramco for providing me with the intern opportunity to work on the 2D real data and publish the result. I would like to thank Dr. Luo for suggesting using WTI for KWI initial model building.