Application of Early Arrival Waveform Inversion to Shallow Land Data

Han Yu and Sherif Hanafy

ABSTRACT

We estimate the near-surface velocity distribution by applying early arrival waveform inversion (EWI) to shallow seismic land data collected with source-receiver offsets no longer than 232 m. This data set is collected at Wadi Qudaid in western Saudi Arabia with the purpose of characterizing the shallow subsurface for its water storage and reuse potential. To enhance the accuracy of EWI, we extracted a natural source wavelet from the data, and also corrected for the attenuation effects with an estimated factor Q. Results suggest that, compared to traveltime tomography, EWI can generate a more highly resolved velocity tomogram from shallow seismic data. The EWI and traveltime tomograms are compared with the conductivity tomogram inverted from an electromagnetic (EM) data set collected at the same location. Both the seismic and EM results show good agreement with one another in indicating similar geological features. The next step is to show that EWI provides a more accurate estimate of the statics than obtained from traveltime tomography.

1

INTRODUCTION

Traveltime tomography is a ray-based geophysical imaging method that inverts the first-arrival traveltimes for the subsurface velocity distribution (Aki and Richards, 2002). It iteratively updates the velocity models by smearing the traveltime residuals along the calculated ray paths (Nolet, 1987), and provides a smooth estimate of the earth's velocity distribution. However, traveltime tomography employs a high-frequency assumption that conflicts with the finite-frequency bandwidth of seismic data, and so generates tomograms with low-to-intermediate resolution.

Full waveform inversion (FWI) (Tarantola, 1984) can overcome the high-frequency limitation in either the space-frequency (Pratt et al., 1998) or the space-time domains (Zhou et al., 1995). However, FWI is computationally expensive and its misfit function is highly nonlinear with respect to velocity perturbations. Therefore, Sheng et al. (2006) proposed an early arrival waveform tomography method for near-surface refraction data. This approach can handle land data if we have a good approximation of the subsurface velocity (Buddenseik, 2004). In this work, our goal is to apply the approach (Sheng et al., 2006) to seismic refraction data collected at Wadi Qudaid, 100 *km* north of Jeddah, Saudi Arabia. Since the subsurface soils are partially saturated, the attenuation factor Q is required to correct for attenuation effects in the data before we invert for the velocity model. Compared to FWI, EWI also avoids a high-frequency assumption but has more reliable convergence properties because it only has to explain the early arrivals in a trace, not all of them.

We also collected an electromagnetic (EM) data set at the same location as the seismic data set. The agreement between the seismic and EM tomograms will increase the confidence in the final results, and provide us with insights on how we might use combined data types to better estimate the geological properties of the subsurface.

In this paper, we first present the theory of EWI. We then introduce the data processing steps for the wadi data. Finally we apply EWI to the early arrivals of the 2D land data set and present the tomograms in the numerical results section. The last section summarizes the salient results of our research.

THEORY

Early arrival waveform inversion (Sheng et al., 2006) assumes the constant-density acoustic wave equation,

$$\frac{1}{c^2(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t|\mathbf{x}_s)}{\partial t^2} - \nabla^2 p(\mathbf{x}, t|\mathbf{x}_s) = s(\mathbf{x}, t|\mathbf{x}_s), \quad (1)$$

where $p(\mathbf{x}, t|\mathbf{x}_s)$ denotes the pressure field at position $\mathbf{x}$, the listening time is t, and the source is at $\mathbf{x}_s$ and excited at $t = 0$. The velocity model is represented by $c(\mathbf{x})$, and $s(\mathbf{x}, t|\mathbf{x}_s)$ represents the time history and amplitude of source function. Equation 1 is used to calculate synthetic seismograms by a finite-difference method (Levander, 1988), where the solution can be written in terms of its Green's function $g(\mathbf{x}, t|\mathbf{x}_s, 0)$ as

$$p(\mathbf{x}, t|\mathbf{x}_s) = \int g(\mathbf{x}, t|\mathbf{x}', 0) * s(\mathbf{x}', t|\mathbf{x}_s) d\mathbf{x}', \quad (2)$$

in which the symbol * denotes temporal convolution.

We ignore the shear wave effects in the wave equation (Zhou

et al., 1995), because they are largely absent in the early arrivals of the seismic traces. EWI estimates the velocity model by minimizing the early arrival misfit function (Boonyasiriwat et al., 2010), where the waveform data residual is defined as

$$\Delta p(\mathbf{x}_g, t | \mathbf{x}_s) = [p_{obs}(\mathbf{x}_g, t | \mathbf{x}_s) - p_{calc}(\mathbf{x}_g, t | \mathbf{x}_s)] W(\mathbf{x}_g, t | \mathbf{x}_s). \quad (3)$$

Here, $\mathbf{x}_g$ is the receiver position vector, p_{obs} and p_{calc} are, respectively, the observed and calculated data, and $W(\mathbf{x}_g, t | \mathbf{x}_s)$ is a window function that mutes all the energy except for the early arrivals. The velocity model $c(\mathbf{x})$ is iteratively updated by minimizing the misfit functional E, represented by the L_2 norm of the data residuals over time and space,

$$E = \frac{1}{2} \sum_s \sum_g \int (\Delta p(\mathbf{x}_g, t | \mathbf{x}_s))^2 dt. \quad (4)$$

A nonlinear conjugate-gradient method (Luo and Schuster, 1991) is used to minimize the gradient function. The gradient of the misfit functional E with respect to changes in the velocity $c(\mathbf{x})$ is the first variation (Logan, 1996) of E at the vector point $c(\mathbf{x})$ in the direction of $\delta c(\mathbf{x})$. This gradient $[grad(\mathbf{x}) = \frac{\partial E}{\partial \delta c(\mathbf{x})}]$ can be computed by migrating the waveform residuals in time (Tarantola, 1984), which is

$$grad(\mathbf{x}) = \frac{1}{c^3(\mathbf{x})} \sum_s \int \dot{p}_f(\mathbf{x}, t | \mathbf{x}_s) \dot{p}_b(\mathbf{x}, t | \mathbf{x}_s) dt, \quad (5)$$

where $\dot{p}$ is the time derivative of p, $p_f(\mathbf{x}, t | \mathbf{x}_s)$ represents the forward-propagated wavefields, and $p_b(\mathbf{x}, t | \mathbf{x}_s)$ represents the back-projected waveform residual wavefields given by

$$p_b(\mathbf{x}, t | \mathbf{x}_s) = \int g(\mathbf{x}, -t | \mathbf{x}', 0) * \delta p(\mathbf{x}', t | \mathbf{x}_s) d\mathbf{x}', \quad (6)$$

and

$$\delta p(\mathbf{x}', t | \mathbf{x}_s) = \sum_g \delta(\mathbf{x}' - \mathbf{x}_g) \Delta p(\mathbf{x}_g, t | \mathbf{x}_s). \quad (7)$$

Now the velocity model can be updated iteratively along the conjugate directions defined by

$$\mathbf{d}_k = -\mathbf{P}_k \mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, \quad (8)$$

where iterations $k = 1, 2, 3, ..., k_{max}$, $\mathbf{g}=[grad(\mathbf{x})]$, and $\mathbf{P}$ is the conventional geometrical-spreading preconditioner (Causse et al., 1999). For the first iteration, we set $\mathbf{d}_0 = -\mathbf{g}_0$. The parameter β_k is calculated by the Polak-Ribiére formula (Nocedal and Wright, 1999)

$$\beta_k = \frac{\mathbf{g}_k^T (\mathbf{P}_k \mathbf{g}_k - \mathbf{P}_{k-1} \mathbf{g}_{k-1})}{\mathbf{g}_{k-1}^T \mathbf{P}_{k-1} \mathbf{g}_{k-1}}. \quad (9)$$

The velocity model is updated by

$$c_{k+1}(\mathbf{x}) = c_k(\mathbf{x}) + \lambda_k d_k(\mathbf{x}), \quad (10)$$

where λ_k is the step length, which is determined by a quadratic line-search method (Nocedal and Wright, 1999), and $d_k(\mathbf{x})$ is the component of the direction vector $\mathbf{d}_k(\mathbf{x})$ indexed by $\mathbf{x}$. To compute the gradient direction at each iteration reduces to computing the reverse time migration operation. Additional forward modelings are required for the line search. Equation 10 is applied iteratively until the misfit functional E satisfies a stop-

ping criterion. In this paper, the initial velocity model $c_0(\mathbf{x})$ is the traveltime tomogram (Nemeth et al., 1997).

DATA PROCESSING

The land data are processed before applying EWI in order to reduce elastic effects in the field data. The data processing steps include accounting for 3D wave propagation effects, attenuation effects, trace normalization, extraction of the natural source wavelet, and bandpass filtering. The following describes the detailed processing procedures.

1. A bandpass filter of $15 \sim 70\ Hz$ is applied to the traces to eliminate strong noise in the data. In this way, the surface waves are partially removed since their dominant frequency is under $15\ Hz$. Noise before the picked first arrivals is also muted.

2. The 3D land data are approximately transformed to 2D by multiplying the trace spectrum by $\sqrt{i/\omega}$ in the frequency domain and by multiplying the trace by $\sqrt{t}$ in the time domain (Boonyasiriwat et al., 2010).

3. The attenuation effects in the field data should be corrected for because the forward modeling is based on the acoustic wave equation. According to Liao and McMechan (1997), the linear attenuation transfer function T is a function of f, t, and Q, such that

$$T(f) = exp\{-f\frac{\pi t}{Q}\}, \qquad (11)$$

which transforms the input signal $S(f)$, which is the spectrum of the first arrival, to the output trace spectrum $R(f)$

by $R(f) = T(f)S(f)$. In equation 11, f can be assigned to each frequency of a trace's spectrum, and t is the first-arrival traveltime. The factor Q can also be written in terms of the centroid frequencies f_r and f_s of the first arrival, respectively, at the receiver and the source as

$$f_r = f_s - \frac{2\pi\sigma_s^2}{Q}t, \qquad (12)$$

where σ_s^2 can be the variance of the source spectrum. After f_r, f_s, and σ_s^2 are estimated from the data and the first arrivals t are picked, the Q value can be estimated from equation 12. Then, equation 11 is applied to the traces to correct the attenuation effects in the frequency domain. In this work, we estimate Q as a constant because we only deal with early arrivals and a shallow part of the earth. Thus the same attenuation correction formula is used for all the traces.

4. All traces are normalized to reduce the errors resulting from previous correction steps. To retain only the early arrivals, the traces are also muted with a time window that starts about 3 to 4 periods after the first arrival. In this case, surface waves and later arrivals are not inverted by EWI.

5. To enhance the EWI results, a natural source wavelet is extracted by averaging the near-offset first arrivals within a window of $10 \sim 20$ traces. We apply zero-lag cross correlation to several consecutive traces to align them by their first arrivals.

6. The very near-offset traces with source-receiver offsets no greater than 14 *m* are muted because they contain surface waves and noise even after filtering, and it is difficult to match them with the synthetic seismograms.

NUMERICAL TESTS

The EWI algorithm is applied to the 2D field data set obtained by a seismic survey conducted at Wadi Qudaid, located along the western side of Saudi Arabia to the east of King Abdullah university of Science and Technology (see Figure 1). The goal is to verify that EWI provides a robust and more resolved estimate of the earth's velocity model compared to traveltime tomography.

2D Near-Surface Refraction Survey, Wadi Qudaid, KAUST

The 2D seismic survey is conducted at Wadi Qudaid (see Figure 1a) to the east of KAUST. The 2D acquisition geometry (Figure 1b) consists of one line of vertical component geophones. Along this line, there are 117 receivers with a 2.0 *m* spacing, and the shots are located at every receiver position so that there are $117 \times 117 = 13689$ traces. In this field experiment, we used a 200 *lb* weight drop (Figure 2) to generate the seismic source energy with $10 \sim 15$ stacks at each shot location. Each common shot gather (CSG) was recorded with a sampling interval of 1.0 *ms* for 1 *s*. Since we need only the early arrivals, arrivals after 0.25 *s* are muted. The CSG #30 and the picked first-arrival traveltimes are shown in Figure 3.

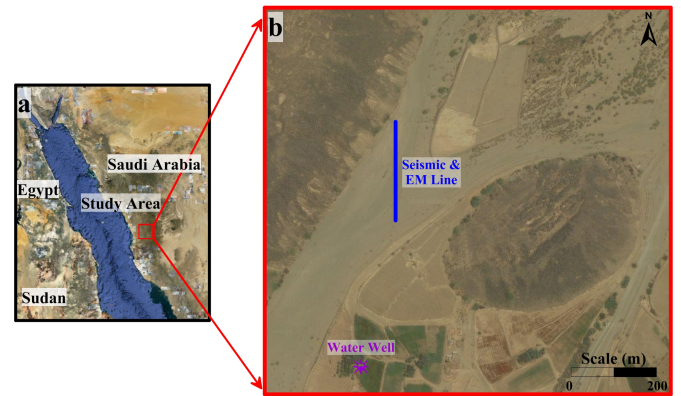For this data set, we estimate the dominant wavelength and

Figure 1: (a) A map shows the study area, Wadi Qudaid, the location of the seismic and EM profiles, and (b) the water well at the site.
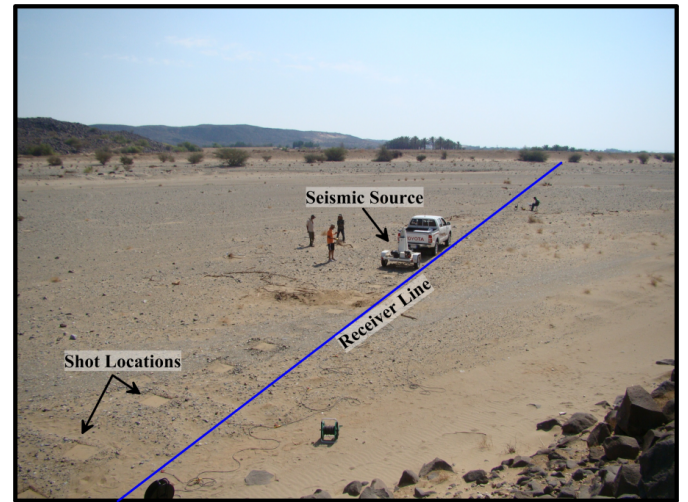

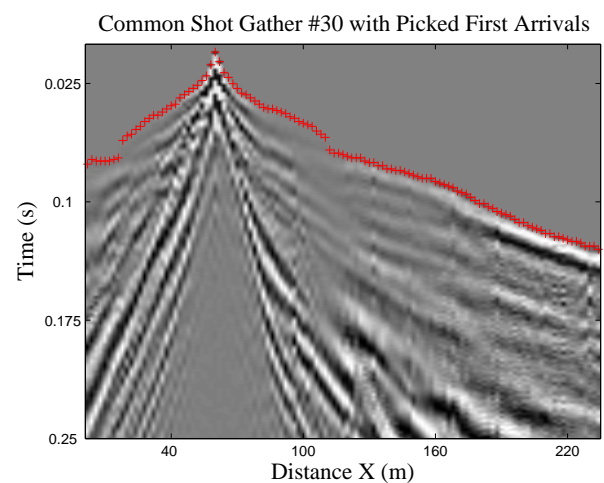Figure 2: A photo (looking south) taken during data acquisition.

Figure 3: CSG #30 with picked first arrival times.

the dominant frequency of the first arrival head waves to be 6 *m* and 60 *Hz*, respectively, where the minimum P-wave velocity is estimated to be 350 *m/s*. Figure 4 shows the picked first arrival times for all traces presented as a 2D matrix, and Figure 5 shows the spectrum of the 30th trace for CSG #30.
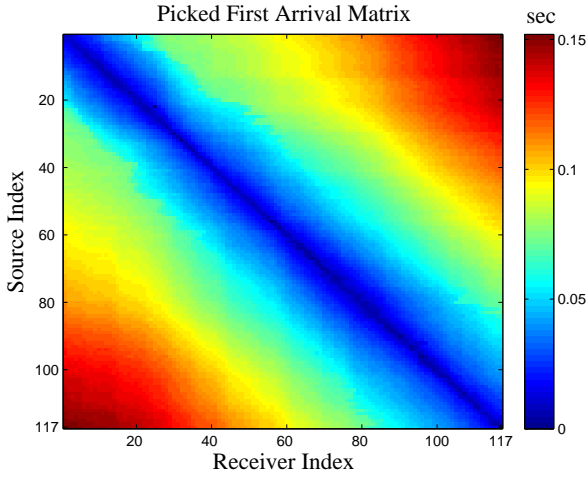

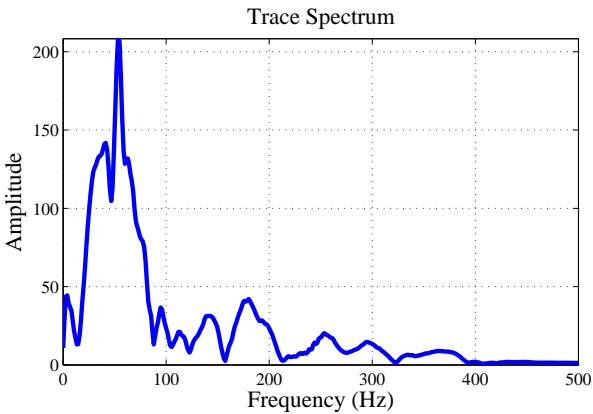Figure 4: The picked first arrivals for the Wadi data set.

Figure 5: Data spectrum of the trace recorded by the 30th receiver of CSG #30.

Prior to waveform inversion, the data are processed by the procedures mentioned earlier. Figure 6 shows the centroid frequencies f_r plotted against traveltimes of the first arrivals. The variance σ_s^2 of the source centroid spectra is equal to 312.6991 Hz^2, which is the average σ_s^2 values from all the sources (Liao and McMechan, 1997). Here, we use a subband of $0 \sim 120\ Hz$

to calculate f_s, f_r and σ_s^2 to avoid errors from the noise. By equation 12, the Q value is estimated to be about 18, which is a typical value for near-surface soil with significant absorption. According to equation 11, $1/T(f)$ is applied to the trace spectra to correct for the attenuation effects in the frequency domain. Figure 7 shows CSG #117 before and after the attenuation correction. A band pass filter of $15 \sim 70\ Hz$ is applied to the data set, and the 3D to 2D correction is also applied. Only the event that arrived within three or four periods after the first arrival are used. Figure 8 shows the raw and corrected CSG #117 before and after all the processing steps.
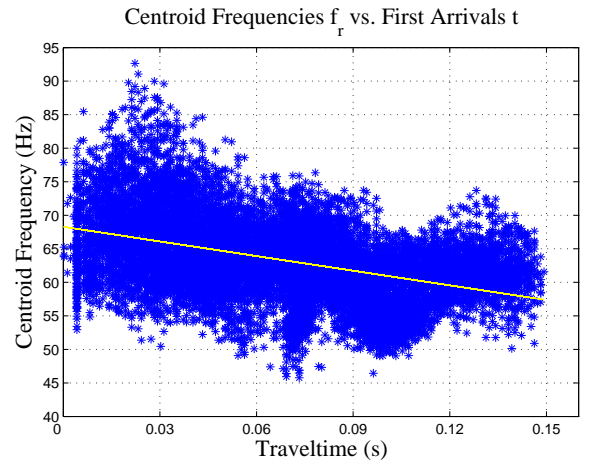

Figure 6: Pairs of the centroid frequencies and the first arrival traveltimes at the receivers. The attenuation factor Q is estimated to be 18 by the best-fit line denoted by the solid yellow line.

A natural source wavelet is extracted by averaging $10 \sim 20$ near offset first arrivals aligned by cross-correlation in the space-time domain. Figure 9b shows the calculated natural source wavelet using the rectangular window (Figure 9a) to catch the first arrival wavelets for CSG #25. Since the weight drop is controlled electronically, all shots have similar source wavelets as illustrated by Figure 10. Hence, the extracted wavelet
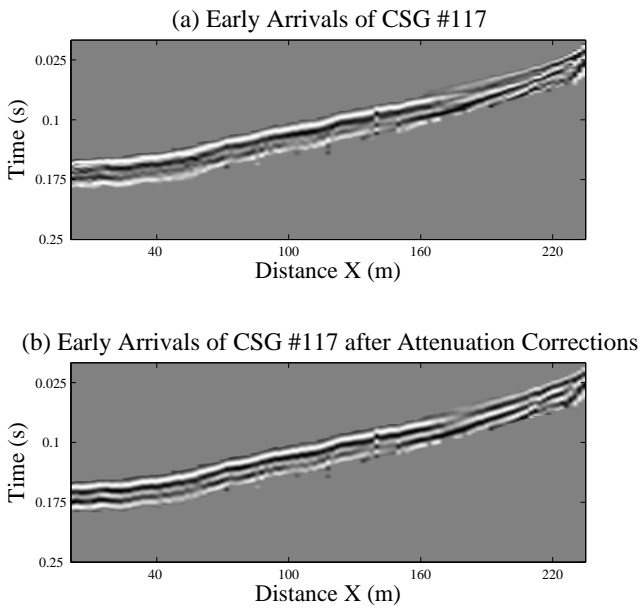
(a) Early Arrivals of CSG #117

(b) Early Arrivals of CSG #117 after Attenuation Corrections

Figure 7: (a) The original and (b) the Q corrected (bottom) CSG #117.
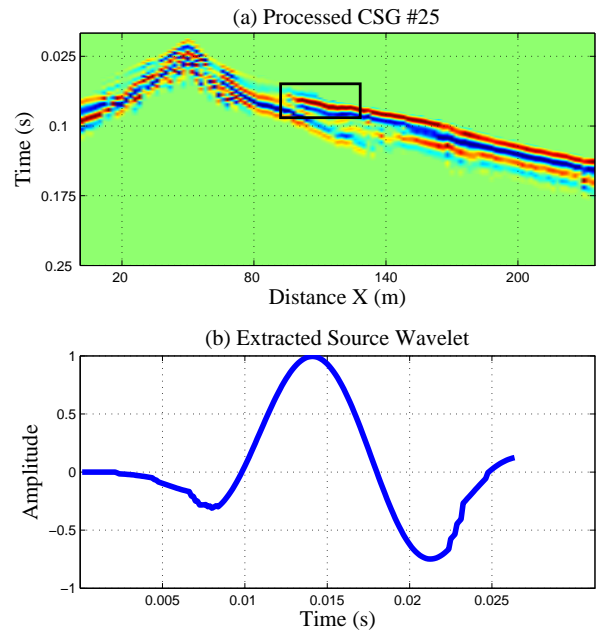
(a) Processed CSG #25

(b) Extracted Source Wavelet

Figure 9: (a) The processed CSG #25 after the application of a muting window and (b) the source wavelet extracted from the early arrivals.
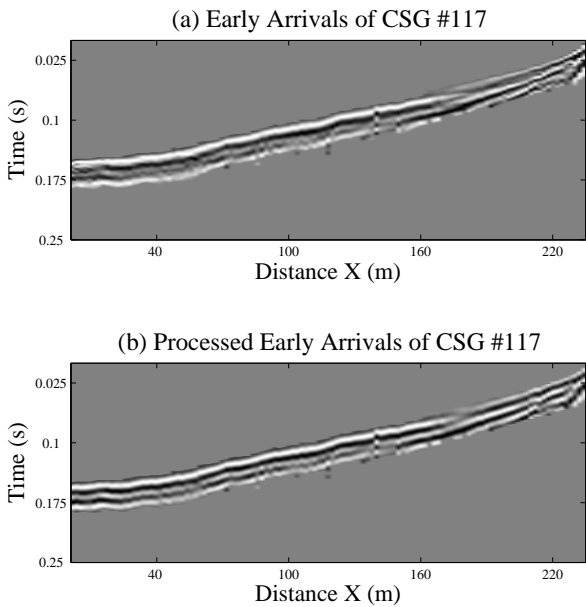
(a) Early Arrivals of CSG #117

(b) Processed Early Arrivals of CSG #117

Figure 8: Early arrivals of (a) the original and (b) the processed CSG #117.

Common Offset Gather

Figure 10: Common offset gather with an offset equal to 200 *m*.

is applied to all shot locations.

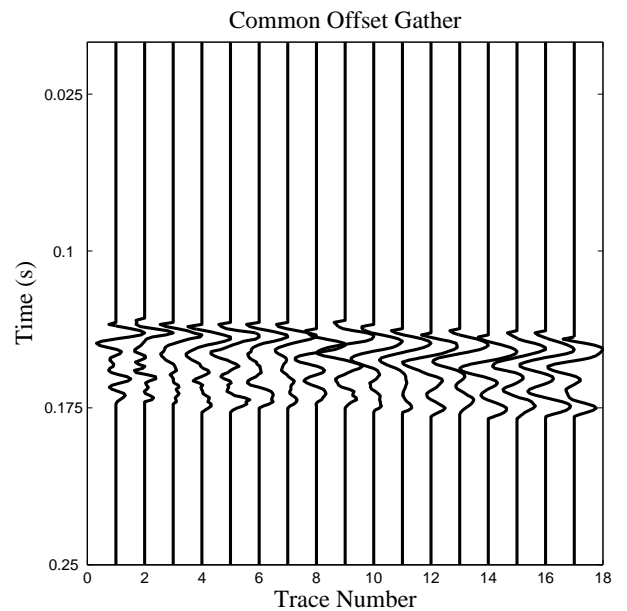Figure 11a shows the traveltime tomogram which is used as the initial velocity model for the EWI. Figure 11b shows the

waveform tomogram after 30 iterations. Compared to the traveltime tomogram, EWI provides a more highly resolved tomogram.
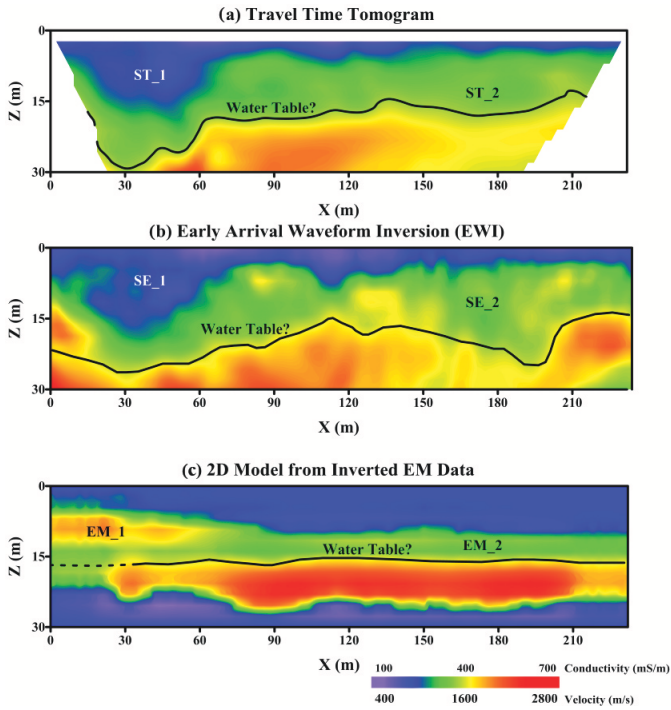
Figure 11: (a) The traveltime tomogram, (b) the EWI tomogram after 30 iterations, and (c) the resistivity tomogram by inverting the EM data.
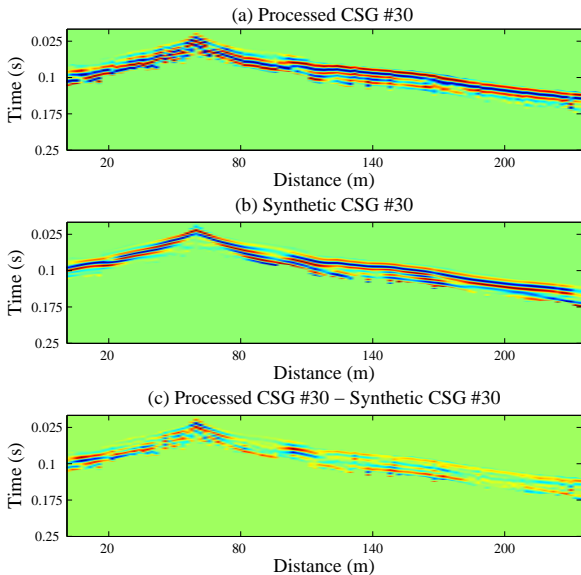

Figure 12: (a) Processed CSG #30, (b) synthetic CSG #30 based on the EWI tomogram, and (c) their difference.

Figure 12 compares the processed data with the synthetic data calculated from the EWI tomogram after 30 iterations. From Figure 12, many early arriving events in the synthetic

data correlate well with the observed data, which is illustrated by their difference (Figure 12c).

Figure 13 shows the RTM images using the EWI and traveltime tomograms as the migration velocity. Using the EWI tomogram as the migration velocity, the RTM image appears to be better focused than that obtained by using the traveltime tomogram as the migration velocity.
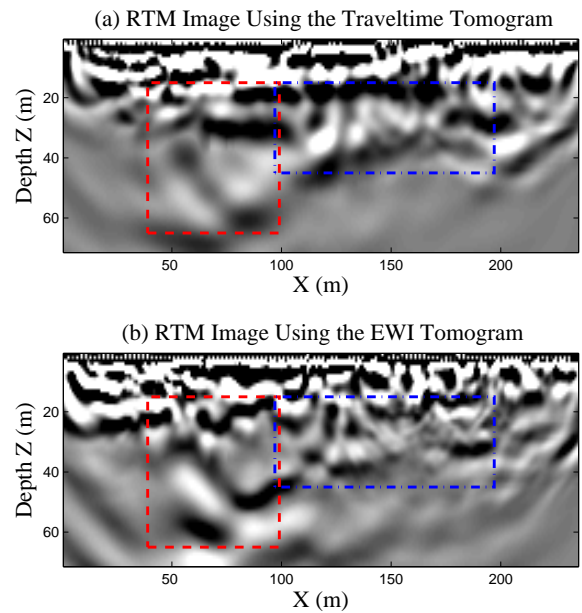

Figure 13: RTM images using the (a) traveltime, and (b) early arrival waveform tomograms as the migration velocity.

Electromagnetic Survey

An electromagnetic (EM) profile is recorded at the same location as the seismic profile. The data are collected using the EM34-3 instrument with three different coil separations (frequencies) at each station. We used the 10 m (6400 Hz), the 20 m (1600 Hz), and the 40 m (400 Hz) coil separation in the horizontal mode. We used FreqEM software to invert the raw data (Figure 14), where a 1D resistivity model is generated at each

station and then all 1D models are combined to create the 2D model shown in Figure 11c. The 2D EM model is roughly consistent with the traveltime tomogram (TT) and EWI tomogram (Figures 11a and 11b). The low-velocity anomaly shown on both the TTT (ST-1) and EWI (SE-1) tomogram corresponds to the local anomaly (EM-1) shown in the EM model between $x = 0$ - 60 m and depths less than 13 m. This anomaly corresponds to a gravel-sand layer (Figure 15a). The second layer (ST-2 in the traveltime tomogram, SE-2 in the EWI tomogram, and EM-2 in the EM model) corresponds to a sand-silt layer with some gravels (Figure 15b). Here, Figure 15b is taken at a water well located around 1 km south of the seismic/EM profile (Figure 1). The water table is at 18 m depth from the ground surface at the water well, and is shown as the solid black line in Figure 11.
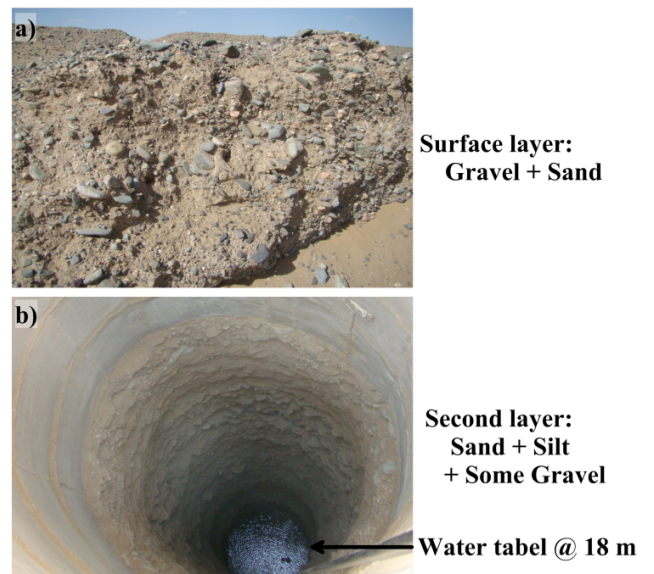

Figure 15: Two photos show the subsurface layers: (a) the first layer, composed of gravel and sand with a thickness less than 5 m. (b) The second layer as shown inside a water well near the study area. It is composed of sand and silt with some gravel and has a thickness of 10 - 15 m.
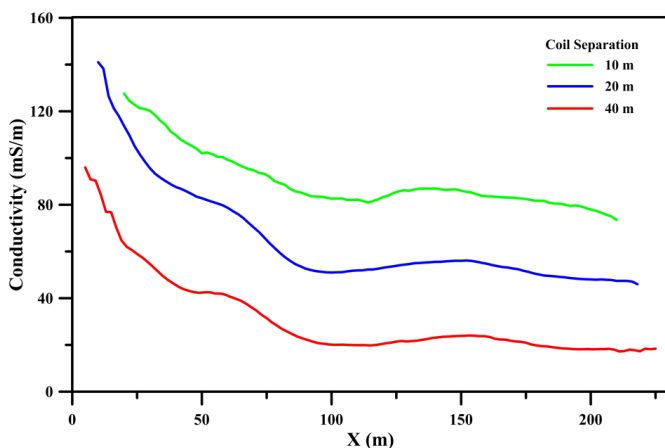

Figure 14: The raw EM data collected at Wadi Qudaid site using EM34-3 instrument and three different coil separation (frequency) at each station location.

CONCLUSIONS

The early arrival waveform inversion method is used to invert seismic data collected at Wadi Qudaid. Compared to travel-time tomography, EWI does not require a high-frequency assumption and benefits from the attenuation correction applied to the recorded traces. EWI can be successfully applied to near-surface surveys if careful processing steps are carried out before inversion. These steps include a bandpass filter and corrections for 3D geometric spreading and attenuation effects, trace normalization, and the natural source. Our results with migration suggest that EWI creates a more accurate velocity model compared to the traveltime tomogram. Comparisons of the RTM images based on EWI and traveltime tomograms show that EWI can provide more highly resolved subsurface structures and illuminate complex structures compared to migration with the traveltime velocity model.

The drawbacks of EWI compared to traveltime tomography include complicated data processing steps, higher computational cost, and slow convergence. Moreover, EWI fits com-

plex waveforms instead of arrival times, which can be prone to the problem of getting stuck in a minima. Our future work includes using the multiscale method (Boonyasiriwat et al., 2010) to carry out EWI and gradually increasing the data window to image deeper structures. We also hope to use our towed land streamer system to efficiently collect shallow seismic data, and eventually invert these data in almost real time using the EWI approach. This might provide high-resolution statics corrections for oil exploration data recorded at the same site.

REFERENCES

Aki, K. and P. G. Richards, 2002, Quantitative seismology: 2nd edition: University Science Books.

Boonyasiriwat, C., G. T. Schuster, P. Valasek, and W. Cao, 2010, Applications of multiscale waveform inversion to marine data using a flooding technique and dynamic early-arrival windows: Geophysics, **75**, R129–R136.

Buddenseik, 2004, Colluvial wedge imaging using traveltime tomography along the wasatch fault near mapleton, Utah: Master's thesis.

Causse, E., R. Mittet, and B. Ursin, 1999, Preconditioning for full-waveform inversion in viscoacoustic media: Geo-

physics, **64**, 130–145.

Levander, A. R., 1988, Fourth-order finite-difference p-sv seismograms: Geophysics, **53**, 1425–1437.

Liao, Q. and G. A. McMechan, 1997, Tomographic imaging of velocity and q, with application to crosswell seismic data from the gypsy pilotsite, oklahoma: Geophysics, **62**, 1804–1811.

Logan, J. D., 1996, Applied mathematics: 2nd edition: Wiley-Interscience.

Luo, Y. and G. T. Schuster, 1991, Wave equation traveltime inversion: Geophysics, **56**, 645–653.

Nemeth, T., E. Normark, and F. Qin, 1997, Dynamic smoothing in crosswell traveltime tomography: Geophysics, **62**, 168–176.

Nocedal, J. and S. J. Wright, 1999, Numerical optimization: Springer series in operations research and financial engineering.

Nolet, G., 1987, Seismic tomography: Springer.

Pratt, G., C. Shin, and Hicks, 1998, Gauss-newton and full Newton methods in frequency space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Extending the aperture and increasing the signal-to-noise ratio of refraction surveys with super-virtual interferometry: Geophysics, **71**, U47–U57.

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

Zhou, C., W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, 1995, Acoustic wave-equation traveltime and waveform in-

version of crosshole seismic data: Geophysics, **60**, 765–773.