Interpretation®

Ray-tracing Traveltime Tomography vs Wave-equation Traveltime Inversion for Near-Surface Seismic Land Data

Journal:	*Interpretation*
Manuscript ID	INT-2016-0210.R2
Manuscript Type:	2016-05 Skeletonized/sparse/multiscale geophysical inversion for the interpreter
Date Submitted by the Author:	13-Apr-2017
Complete List of Authors:	Fu, Lei; King Abdullah University of Science and Technology, Earth Science and Engineering Program Hanafi, Sherif; KAUST, Earth Science
Keywords:	head waves, water table, traveltime
Subject Areas:	Application examples (applying a relatively new technique or concept), Near-surface, ground penetrating radar, environmental, and engineering applications
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
INT_WT_final _table.tex	

Ray-tracing Traveltime Tomography vs Wave-equation Traveltime Inversion for Near-Surface Seismic Land Data

Lei Fu[1*], *Sherif M. Hanafy*[1,2]

[1]*Division of Physical Science and Engineering, King Abdullah University of Science and Technology (KAUST),Thuwal, Saudia Arabic*

[2]*Geophysics Departmen, Cairo University, Cairo, Egypt*

Corresponding author: Lei Fu (lei.fu@kaust.edu.sa)

ABSTRACT

Full waveform inversion (FWI) of land seismic data tends to get stuck in the local minima associated with the waveform misfit function. This problem can be partly mitigated by using an initial velocity model that is close to the true velocity model. This initial starting model can be obtained by inverting traveltimes with ray-tracing traveltime tomography (RT) or wave-equation traveltime inversion (WT). We now show that WT can provide a more accurate tomogram than RT by inverting the first-arrival traveltimes, and RT is more sensitive to the additive noise in the input data than WT. We present two examples of applying WT and RT to land seismic data acquired in western Saudi Arabia. One of the seismic experiments investigated the water-table depth, and the other one attempted to detect the location of a buried fault. The seismic land data were inverted by WT and RT to generate the P-velocity tomograms, from which we can clearly identify the water table depth along the seismic survey line in the first example and the fault location in the second example.

INTRODUCTION

For land seismic surveys, the near-surface velocity model is important for computing an accurate migration image (Vermeer, 2002). The near-surface velocity model is obtained by traveltime tomography which inverts the first-arrival traveltimes for the P-velocity model. These arrivals are those of either diving waves or refractions. Ray-tracing traveltime tomography (RT) has been used for earthquake imaging of the earth's interior from the early 1970s (Aki et al., 1977; Humphreys et al., 1984), which was mostly implemented using the high-frequency approximation of ray tracing to calculate the traveltimes. To avoid the high-frequency assumption of RT, Luo and Schuster (1991a,b) proposed a wave-equation traveltime inversion (WT) method to obtain the near-surface P-wave velocity model. The WT method can be more accurate than RT, but it is at least an order-of-magnitude more costly because the wave equation must be numerically solved for each source.

Full wave inversion (FWI) is an underdetermined problem, where many different models yield synthetic data that match observed data within a reasonable tolerance. The local minima and cycle-skipping problems in FWI are partly caused by an inaccurate estimate of the source wavelet, an inaccurate amplitude modeling, elastic effects, random noise and an inaccurate initial velocity model. In order to suppress the source wavelet effect, Frazer and Sun (1998) developed a source-independent algorithm using convolved wavefields. Bai et al. (2014) presented visco-acoustic waveform inversion in the time domain for velocity estimation, where visco-acoustic wave equations are used to generate accurate amplitudes. To reduce the harmful effects of noise

in the data, Neelamani et al. (2008) applied a curvelet-based algorithm to attenuate random and coherent linear noise in a 3D data set from a carbonate environment. To partly migrate the local minima problem of FWI, Yu and Hanafy (2014) applied a multiscale early arrival waveform inversion (MEWI, which gradually lengthens the time window for muting early arrivals) to shallow seismic land data. However, an accurate initial velocity model is very important for obtaining an accurate tomogram in both MEWI and FWI. Both WT and RT inversion methods can provide a reasonable initial velocity model. However, WT (Luo and Schuster, 1991b) does not require a high-frequency assumption and can account for body-wave, diffraction, reflection and headwave traveltimes. Hence, the WT (Luo and Schuster, 1991a,b) method should provide a more accurate initial velocity model than the RT method for MEWI or FWI.

In this paper, one synthetic and two field data examples are used to demonstrate the advantages of WT over the RT method. We first describe the theory of the WT and RT methods, and then describe how WT and RT is implemented using a workflow. We then present the numerical results for applying both WT and RT to one synthetic and two field data examples. The final section presents the discussion and conclusions.

WAVE-EQUATION TRAVELTIME INVERSION

For seismic data, refraction waves can be modeled by solving the 2D constant-density acoustic wave equation (Yilmaz, 2001)

$$\nabla^2 p(\mathbf{x},t) - \frac{1}{\mathbf{v}(\mathbf{x})^2}\frac{\partial^2 p(\mathbf{x},t)}{\partial t^2} = f(\mathbf{x},t), \tag{1}$$

where, $\mathbf{x}$ is the model space, $\nabla^2 = \left(\frac{\partial}{\partial x^2}, \frac{\partial}{\partial y^2}\right)$ is the 2D Laplacian, $\mathbf{v}(\mathbf{x})$ is the P-wave velocity model, $p(\mathbf{x},t)$ is the pressure field, and $f(\mathbf{x},t)$ is the source term. The wave-equation traveltime inversion (WT) method is designed to minimize the following objective function:

$$\Phi(\mathbf{v}) = \frac{1}{2}\|\Delta\tau_{rs}\|^2, \tag{2}$$

where, $\mathbf{v}$ is the P-wave velocity, $\Delta\tau_{rs} = \tau_{cal}(\mathbf{x}_r,\mathbf{x}_s) - \tau_{obs}(\mathbf{x}_r,\mathbf{x}_s)$ is the refraction traveltime residual for a source at $\mathbf{x}_s$ and a receiver at $\mathbf{x}_r$ (Luo and Schuster, 1991b,a). In practice, the traveltime difference can be found by cross-correlating the predicted and the observed seismograms (Luo and Schuster, 1991b). The P-velocity distribution can be iteratively updated using any gradient-based method (Scales, 1985) such as the conjugate gradient method:

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \alpha_k \mathbf{d}_k, \tag{3}$$

where $\mathbf{v}_{k+1}$ is the P-velocity model at the $(k+1)^{th}$ iteration; α_k is the scale step length, which can be found using a line-search algorithm; and $\mathbf{d}_k$ is the updated direction:

$$\mathbf{d}_k = -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, \tag{4}$$

$$\beta_k = \frac{(\mathbf{g}_k - \mathbf{g}_{k-1}, \mathbf{g}_k)}{(\mathbf{g}_{k-1}, \mathbf{g}_{k-1})}. \tag{5}$$

Here, β_k is a scalar chosen to insure that $\mathbf{d}_k$ is conjugate to $\mathbf{d}_{k-1}$, the current search direction (Wright and Nocedal, 1999) $\mathbf{d}_k$ is given by the current k^{th} gradient $\mathbf{g}_k = \frac{\partial \Phi}{\partial \mathbf{v}_k}$ and the former search direction $\mathbf{d}_{k-1}$. Since WT is a non-linear inversion method, we use a quadratic line search method to compute the step length (Wright and Nocedal, 1999). Several reasonable values of the step length are selected, e.g., $\alpha_1 = 0.5$ and $\alpha_2 = 1$, and the objective functions $\Phi(\mathbf{v}_k + \alpha_1 \mathbf{d}_k)$ and $\Phi(\mathbf{v}_k + \alpha_2 \mathbf{d}_k)$ are computed to determine if they are less than $\Phi(\mathbf{v}_k)$, and if $\Phi(\mathbf{v}_k + \alpha_1 \mathbf{d}_k) < \Phi(\mathbf{v}_k + \alpha_2 \mathbf{d}_k)$. Then the following quadratic formula is used to determine the interpolated value of the objective function,

$$\begin{aligned} \Phi(\mathbf{v}_k + \alpha \mathbf{d}_k) \;\; &= \tfrac{(\alpha-\alpha_1)(\alpha-\alpha_2)}{\alpha_1 \alpha_2} \Phi(\mathbf{v}_k) + \tfrac{\alpha(\alpha-\alpha_2)}{\alpha_1(\alpha_1-\alpha_2)} \Phi(\mathbf{v}_k + \alpha_1 \mathbf{d}_k) \\ &+ \tfrac{\alpha(\alpha-\alpha_1)}{\alpha_2(\alpha_2-\alpha_1)} \Phi(\mathbf{v}_k + \alpha_2 \mathbf{d}_k) \qquad 0 \le \alpha \le \alpha_2. \end{aligned} \tag{6}$$

The scalar value α is found by differentiating $\Phi(\alpha)$ of equation 6 with respect to α, setting the result to zero, and solving for α (see Schuster (2017) for further details).

In wave-equation traveltime inversion, the gradient of the misfit function with

respect to the P-wave velocity is given by (Luo and Schuster, 1991b)

$$\mathbf{g} = \frac{1}{(\mathbf{v})^3} \sum_s \int \dot{p}(\mathbf{x}, t|\mathbf{x}_s)_{cal} \dot{p}(\mathbf{x}, t|\mathbf{x}_r) dt, \tag{7}$$

where

$$\dot{p}(\mathbf{x}, t|\mathbf{x}_s) = \sum_r G(\mathbf{x}, -t|\mathbf{x}_r, 0) * \delta\tau(\mathbf{x}_r, t|\mathbf{x}_s), \tag{8}$$

and the symbol $*$ represents temporal convolution. The $p(\mathbf{x}, t|\mathbf{x}_s)_{cal}$ means the source at $\mathbf{x}_s$ is excited at $t = 0$, and the receiver at $\mathbf{x}$ (model space) records the pressure field at t. $\dot{p}$ represents the time derivative of p, and $G(\mathbf{x}, -t|\mathbf{x}_r, 0)$ is the Greens function associated with equation 1 for the velocity field $\mathbf{v}$, and $\delta\tau(\mathbf{x}_r, t|\mathbf{x}_s)$ is the backpropagating residual. The gradient $\mathbf{g}$ is formed by taking the zero-lag temporal correlation of the source-side wavefield $\dot{p}(\mathbf{x}, t|\mathbf{x}_s)$ with the backpropagated wavefield $\dot{p}(\mathbf{x}, t|\mathbf{x}_r)$. The source-side wavefield is calculated by exciting the source at $\mathbf{x}_s$ for each time step. As for the backpropagated wavefield, the observed seismogram $p(\mathbf{x}_r, t|\mathbf{x}_s)_{obs}$ at receiver $\mathbf{x}_r$ is weighted by its associated traveltime residual $\Delta\tau_{rs}$ (the difference between the observed and predicted traveltime), and then backpropagated in reverse time, and recorded at $\mathbf{x}$ for each time step.

The WT workflow is shown in Figure 1a, it is used to iteratively find the velocity model that minimizes the misfit function in equation 2 until the misfit change percentage falls below a predefined tolerance, and here we set the tolerance for WT is 0.01 according to our experience.

Ray-tracing Traveltime Tomography

The Ray-tracing traveltime tomography is based on the idea that the traveltime of a ray is the discretized integral of the slowness along the ray:

$$t_i = \sum_{j=1}^{N} l_{ij} s_j, \tag{9}$$

where t_i is the traveltime of the i^{th} ray, l_{ij} is the segment length of the i^{th} ray that intersects the j^{th} cells, s_j is the slowness (reciprocal velocity) in the j^{th} cell. This results in an $M \times N$ system of equations, denoted in matrix-vector notation as $\mathbf{Ls} = \mathbf{t}$, where $\mathbf{t}$ represents the recorded $M \times 1$ traveltime data vector, and $\mathbf{s}$ is the $N \times 1$ vector of unknown slownesses in the N cells. The $M \times N$ matrix $\mathbf{L}$ contains the segment lengths of the rays.

The least-squares solution of equation 9 is by minimizing the following objective function:

$$\begin{aligned} \epsilon &= \frac{1}{2} \left[\mathbf{Ls} - \mathbf{t}\right]^{\top} \left[\mathbf{Ls} - \mathbf{t}\right], \\ &= \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{N} (l_{ij} s_j - t_i)^2, \\ &= \frac{1}{2} \sum_{i=1}^{M} r_i^2, \end{aligned} \tag{10}$$

where $r_i = \sum_j (l_{ij} s_j - t_i)$ is the i^{th} traveltime residual between the predicted and observed i^{th} traveltimes. If the rays bend then the matrix elements (l_{ij}) depend on the unknowns s_j. In this case the traveltime equations are non-linear and so that

the rays must be recomputed for the new slowness model after each iteration. The slowness $\mathbf{s}$ can be iteratively updated using the conjugate gradient method:

$$\mathbf{s}_{k+1} = \mathbf{s}_k + \alpha_k \mathbf{d}_k, \tag{11}$$

where $\mathbf{s}_{k+1}$ is the slowness at the $(k+1)^{th}$ iteration, α_k is the scale step length, $\mathbf{d}_k$ is the updated direction determined by the derivative $\frac{\partial \epsilon}{\partial \mathbf{s}_k}$ of the objective function with respect to the model parameter. The workflow for implementing RT (Zhang and Toksöz, 1998) is described by the following steps (Figure **??**b):

1. The first-arrival traveltimes (τ_{obs}) of the recorded data are picked.
2. An initial velocity model ($\mathbf{v}_k = \frac{1}{\mathbf{s}_k}, k = 0$) is given.
3. The first-arrival traveltimes (τ_{cal}) are calculated using the given initial velocity model.
4. If the RMS error between the observed and calculated traveltimes ($\tau_{obs} - \tau_{cal}$) is larger than a predefined tolerance (a quarter of the dominant period), then the velocity model is updated.
5. Steps 3 and 4 are iteratively repeated until the RMS traveltime error is less than the given tolerance.

NUMERICAL RESULTS

The WT and RT methods are now tested on one synthetic and two data sets recorded over the western side of Saudi Arabia. One of the seismic surveys was carried out at Wadi Qudaid for the purpose of determining the water-table depth. The other survey was carried out near the Gulf of Aqaba to detect the location of a hidden fault.

Synthetic Example

The synthetic P-velocity model we created is shown in Figure ??a, which is a complex model with an irregular interface. It is discretized into 60 x 300 grid-points with the grid size of 1.5 m. A finite-difference solution to the 2D acoustic wave equation is used to compute 60 shot gathers with 60 geophones evenly deployed on the surface with a 7.5 m interval. The source wavelet is a Ricker wavelet with a peak frequency of 50 Hz. The goal of this synthetic test is to show that the WT method is more accurate than the RT method when the high-frequency assumption of RT is violated. This violation occurs when the characteristic wavelength [1] of the velocity model is about the same or less than the dominant wavelength of the refraction wavefield (Luo and Schuster, 1991b).

The initial velocity model is the gradient velocity model shown in Figure ??b. The traveltimes of the first arrivals are picked and inverted, and the resulting WT and RT tomograms are shown in Figures ??c and ??d, respectively. The dashed

[1]The characteristic wavelength $\lambda^{char} = 2\pi/k^{min}$ of the earth's velocity variations is obtained by taking a 2D Fourier transform of the velocity model and identifying the minimum wavenumber k^{min} with significant amplitude.

yellow and white lines denote the value of 1550 m/s and 1800 m/s respectively in the actual tomogram, and are used as the reference markers for the tomograms. In this regard, the WT tomogram is closer to the true model than the RT tomogram.

The horizontal velocity profiles at the depth of 25 m are shown in Figure **??**, where the black, blue, green and red curves represent the velocity of the true model, initial model, RT tomogram, and WT tomogram. We can see that the velocity profile from the WT tomogram is more accurate than the RT profile. This suggests that the WT method can provide a much more accurate velocity model than the RT method when the high-frequency assumption is violated.

Table 1 shows the computational metrics for the WT and RT methods. The root mean square (rms) traveltime differences for WT and RT are 5.1 ms and 5.8 ms at the final iteration, respectively. The WT method is 120 times more computationally expensive than the RT method, when those codes are executed with a 12-core workstation.

To test the noise sensitivity of these methods, 10% random noise is added to the raw wiggle traces for WT, and 10% random noise is added to the picked traveltimes for RT. The corresponding WT and RT tomograms with the noisy data are shown in Figures **??**e and **??**f, respectively. We can see that the WT tomogram with noise (**??**e) is almost the same as the WT tomogram without noise (**??**c), while the RT tomogram with noise (**??**f) is different from the RT tomogram without noise (**??**d), which means that the RT is more sensitive to the level of noise in the input data.

Data Processing for Field Data

The first-arrival traveltimes are manually picked, and a reciprocity test [2] is applied to reject the traveltimes with a picking errors larger than one quarter of the dominant period. Then the traveltimes are used for the input data for RT. For WT inversion, in order to suppress elastic effects such as surface waves, and account for geometrical spreading, the seismic land data are processed by the following steps:

1. We applied a 3D to 2D approximation by multiplying the trace spectrum by $\sqrt{i/\omega}$ in the frequency domain (Boonyasiriwat et al., 2010).

2. Correct for geometrical spreading by multiplying the traces by $\sqrt{t}$ in the time domain.

3. The traces are windowed to only admit the first-arrival refraction events.

4. The traces recorded within one dominant wavelength from the source are muted because they contain surface waves and noise even after filtering.

Wadi Qudaid Field Data

A seismic survey was carried out at Wadi Qudaid (red box in Figure **??**a). The survey line is indicated by the double blue arrow, where 117 vertical-component geophones are deployed at 2.0 m intervals along the survey line. A 90-kg accelerated-weight drop (Figure **??**e) is used as the source at every geophone position to record 117 common

[2]Reciprocity means that the traveltime from the shot location to the receiver location should be the same as the traveltime from the receiver location to the shot location.

shot gathers (CSG). A water well near the study area shows that the near-surface sediments consist of two layers above the water table. The first layer is composed of gravel and sand with a thickness of less than 5 m (Figure **??**c), the second layer is composed of sand and silt with some gravel with a thickness of around 10 - 15 m, and the water table is around 18 m in depth (Figure **??**d).

CSG #30 is shown in Figure **??**a, where the air, Rayleigh and refraction arrivals are marked. In our study, we only focus on the refraction arrivals, so all other events are muted (Figure **??**b) and the result is used as the input to the WT inversion. The WT method can also be extended to inverting the traveltimes of reflection arrivals Zhang et al. (2011).

The initial model for WT is a velocity-gradient model, which is discretized into 160 x 928 grid-points with the grid size of 0.25 m. The RT, WT, and interpreted WT tomograms are shown in Figures **??**a, **??**b, and **??**c, respectively. For the near-surface zone from $0 < z < 25$ m, both the RT and WT tomograms have a similar velocity structure, but the depth to the water table in the RT tomogram is deeper than in the WT tomogram. The black dashed curve (Figure **??**c) denotes the estimated water table according to the velocity contour value of 1550 m/s, which is consistent with the water well information in Figures **??**c and **??**d. The RT tomogram shows a high-velocity local anomaly at $20 < x < 110$ m and $z > 25$ m, which is not the case in the WT. In the WT tomogram a continuous bedrock interface is shown at $25 < z < 35$ m.

Figure **??**a displays the traveltimes for shots 15, 65 and 110. The black, red

and blue curves represent the observed, WT, and RT traveltimes, respectively. We can see that WT traveltimes are in better agreement with the observed traveltimes. Figure **??**b shows the normalized traveltime misfit functions for the WT and RT methods. The RT misfit function only decreases to 20% after 30 iterations, however, the WT misfit function decreases to about 13% after 10 iterations. In this example, the WT method has a better convergence rate than the RT method.

Aqaba Field Data

A seismic survey was carried out near the Gulf of Aqaba (the black box area in Figure **??**a), where 120 vertical-component geophones are deployed at 2.5 m intervals along the black double-arrow line in Figure **??**b. A 90-kg accelerated-weight drop is used for a source at every geophone position to record 120 common shot gathers. A 2D resistivity profile is acquired at the same location parallel to the seismic profile. The acquisition parameters of the resistivity profile are:

1. Number of electric nodes: 64
2. Node interval: 5 m
3. Configuration array: Schlumberger-Wenner
4. Total profile length: 315 m
5. Both seismic and resistivity profiles share the same starting point of the profile

A sample CSG is shown in Figure **??**a. The air, Rayleigh, reflection and refraction arrivals are indicated in Figure **??**a, and a processed CSG is shown in Figure **??**b. The processed CSGs are the input data for the WT method.

We re-arranged the raw CSGs into the common offset gathers (COGs) with different offsets. Figure **??**a is the zero-offset COG, which shows a continuous horizontal event. Figures **??**b and **??**c are the COGs with the offsets of 20 m and 40 m, respectively, where the events are no longer continuous everywhere, especially at $x = 145$ m (the area indicated by the black dashed ellipse). The time delays of the surface waves in the distorted area (ellipse in Figures **??**b and **??**c) implies that the shear velocity is slower, and suggests the location of a possible fault.

The initial P-velocity model for WT inversion is a gradient velocity model, which is discretized into 150 x 598 grid-points with the grid size of 0.5 m. The RT and WT tomograms after 10 iterations are shown in Figures **??**a and **??**b. Both the RT and WT tomograms are similar to each other. Both the RT and WT tomograms suggest the hidden fault location (indicated by the pink ellipse and dashed red line) is at $125 < x < 145$ m, which is consistent with the resistivity tomogram shown in Figure **??**c. However, compared with the RT tomogram (Figures **??**a), the bedrock interface indicated by the black dashed curve in WT tomogram (Figures **??**b) is more consistent with the bedrock interface revealed by the resistivity tomogram (Figures **??**c).

The recorded resistivity data were inverted using the Res2DInv software to generate the resistivity tomogram (Ostrowski et al., 2010) in Figure **??**c. Two distinct

layers are visible in the resistivity tomogram, the first layer has resistivity values ranging from 250 to 500 Ohm.m with the layer thickness ranging between 6 m and 10 m. The second layer extends to the bottom of the section and has low resistivity values ranging between 10 and 50 Ohm.m, except between offsets 130 m and 145 m, where the resistivity values appear to increase to about 250 Ohm.m. The location of the fault is shown on the resistivity tomogram as a vertical anomaly (between offsets 130 and 145 m) with higher resistivity values (250 Ohm.m) shown in Figure **??**c.

Figure **??**a show the traveltimes for shots 15, 65 and 110. The black, red and blue curves represent the observed, WT, and RT traveltimes, respectively. We can see that the WT traveltimes are in better agreement with the observed traveltimes. Figure **??**b shows the normalized misfit functions for the WT and RT methods, The misfit function for the RT method only decreases to 25% after 30 iterations, however, the WT misfit function decreases to about 16% after 10 iterations. In this example, the WT method has a better convergence rate than the RT method.

DISCUSSION AND CONCLUSIONS

The WT and RT methods are applied to one synthetic example and two seismic land data sets acquired in western Saudi Arabia. The goal of the Wadi Qadid experiment is to determine the topography of the water table and the goal of the other survey is to locate the Aqaba fault. The seismic field data sets are carefully processed using geometrical spreading corrections, shot normalization, window muting along the first arrivals, and muting near-offset traces. The WT method generates P-velocity

tomograms, from which we can clearly identify the water-table depth in the Wadi Qudaid experiment, and the fault location in the Aqaba experiment.

Compared to the RT tomogram, the WT tomogram from Wadi Qudaid indicates a more accurate water-table depth that is consistent with the water-well information. The WT tomogram from the Aqaba data suggests a more accurate hidden fault location, which is in closer agreement with the resistivity tomogram. Although the WT method requires more than an order-of-magnitude more computational time compared with the RT method, it can provide a more accurate initial velocity model for MEWI or FWI.

ACKNOWLEDGEMENTS

The research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST) in Thuwal, Saudi Arabia. We are grateful to the sponsors of the Center for Subsurface Imaging and Modeling (CSIM) Consortium for their financial support. For computer time, this research used the resources of the Supercomputing Laboratory at KAUST and the IT Research Computing Group. We thank them for providing the computational resources required for carrying out this work.

REFERENCES

Aki, K., A. Christoffersson, and E. S. Husebye, 1977, Determination of the three-dimensional seismic structure of the lithosphere: Journal of Geophysical Research, **82**, 277–296.

Bai, J., D. Yingst, R. Bloor, and J. Leveille, 2014, Viscoacoustic waveform inversion of velocity structures in the time domain: Geophysics, **79**, R103–R119.

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.

Frazer, L. N., and X. Sun, 1998, New objective functions for waveform inversion: Geophysics, **63**, 213–222.

Humphreys, E., R. W. Clayton, and B. H. Hager, 1984, A tomographic image of mantle structure beneath southern california: Geophysical Research Letters, **11**, 625–627.

Luo, Y., and G. T. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophysical Journal International, **105**, 289–294.

———, 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Neelamani, R., A. I. Baumstein, D. G. Gillard, M. T. Hadidi, and W. L. Soroka, 2008, Coherent and random noise attenuation using the curvelet transform: The Leading Edge, **27**, 240–248.

Scales, L., 1985, Introduction to non-linear optimization: Springer-Verlag New York, Inc.

Schuster, G. T., 2017, Seismic inversion: SEG Publishing, Tulsa Oklahoma.

Vermeer, G. J., 2002, 3-d seismic survey design: Society of Exploration Geophysicists.

Wright, S., and J. Nocedal, 1999, Numerical optimization: Springer Science, **35**, 67–68.

Yilmaz, Ö., 2001, Seismic data analysis: Processing, inversion, and interpretation of seismic data: Society of Exploration Geophysicists.

Yu, H., and S. M. Hanafy, 2014, An application of multiscale early arrival waveform inversion to shallow seismic data: Near Surface Geophysics, **12**, 549–557.

Zhang, J., and M. N. Toksöz, 1998, Nonlinear refraction traveltime tomography: Geophysics, **63**, 1726–1737.

Zhang, S., G. Schuster, and Y. Luo, 2011, Wave-equation reflection traveltime inversion, *in* SEG Technical Program Expanded Abstracts 2011: Society of Exploration Geophysicists, 2705–2710.

FIGURE & TABLE CAPTION

Figure 1: The flow chart for, a) WT inversion, where $\Delta\epsilon$ for WT is the misfit change percentage, and we set the tolerance is 0.01, and b) RT inversion, where $\Delta\epsilon$ for RT is the RMS residual, and the tolerance we set is a quarter of the dominant period.

Figure 2: a) The true velocity, b) initial-velocity models for both the WT and RT methods, c) WT, d) RT, e) WT tomograms with 10% random noise, and f) RT tomogram with 10% random noise. For this velocity model, the characteristic horizontal variation of the velocity is $\approx 20\ m$ along the interface for 25 m $< z <$ 45 m. Thus the high-frequency assumption of ray tracing is violated for a 50-Hz source wavelet having a wavelength about 40 m along this interface. The yellow and white dashed lines represent the velocity of 1550 m/s and 1800 m/s, respectively, in the true model.

Figure 3: The comparison of the WT and RT velocity profiles at the depth of 25 m, where the black, blue, green and red curves represent the true, initial, RT, and WT velocity. This profile suggests that the WT method is more accurate than the RT method.

Figure 4: a) Google map shows the location of the Wadi Qudaid experiment, b) seismic survey line indicated by the blue double-arrow line, c) photograph showing the first layer composed of gravel and sand, with a thickness less than 5 m, d) a water well near the study area showing the second layer composed of sand and silt with

some gravel, with the thickness of about 10-15 m, and the water table is at $z = 18$ m, and the e) seismic source is excited by a 90-kg accelerated-weight drop.

Figure 5: a) CSG #30, where the air, Rayleigh, and refraction waves are indicated, and b) the processed CSG where only the refraction waves are extracted and processed.

Figure 6: The RT and WT tomograms inverted from the Wadi Qudaid data: a) RT, b) WT, and c) interpreted WT tomograms. The black dashed curve is the interpreted water table and the pink dashed curve is the interpreted bedrock interface.

Figure 7: a) The traveltimes for shots 15, 65, 110, where the black, red and blue curves represent the observed, WT, and RT traveltimes, respectively, and b) the normalized misfit functions for the WT and RT methods.

Figure 8: a) Google map shows the location of the Aqaba survey, b) photo shows an earthquake scarp (indicated by the red dashed line) on the surface, and the black arrow indicates the location of the seismic and resistivity survey lines.

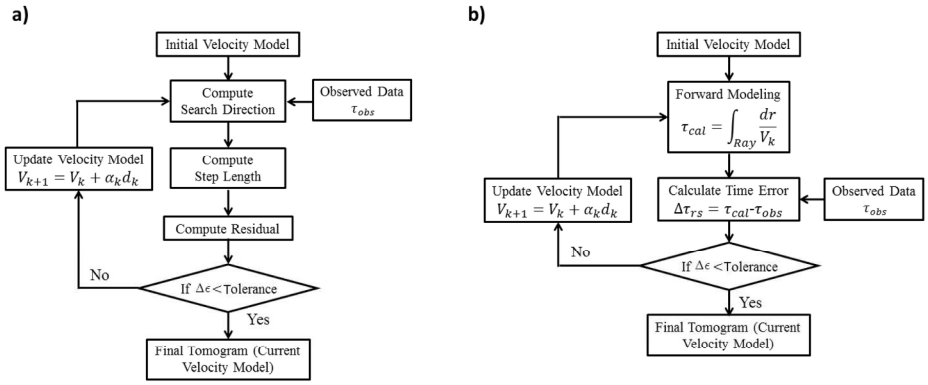
Figure 9: a) CSG#1, where the air, Rayleigh, reflection and refraction arrivals are indicated, and b) the processed CSG, where only the refraction arrivals are extracted and processed.

Figure 10: Common offset gathers for the source-receiver, a) offset = 0 m, b) offset = 20 m, and c) offset = 40 m.

Figure 11: The tomograms inverted from the Aqaba data: a) RT and b) WT tomograms, where the pink ellipses indicate possible fault locations and the black dashed curves represent the bedrock interfaces, c) resistivity tomogram inverted by Res2DInv, a low-resistivity anomaly is shown at $130 < x < 145$ m, where the white dashed curve represents the bedrock interface.

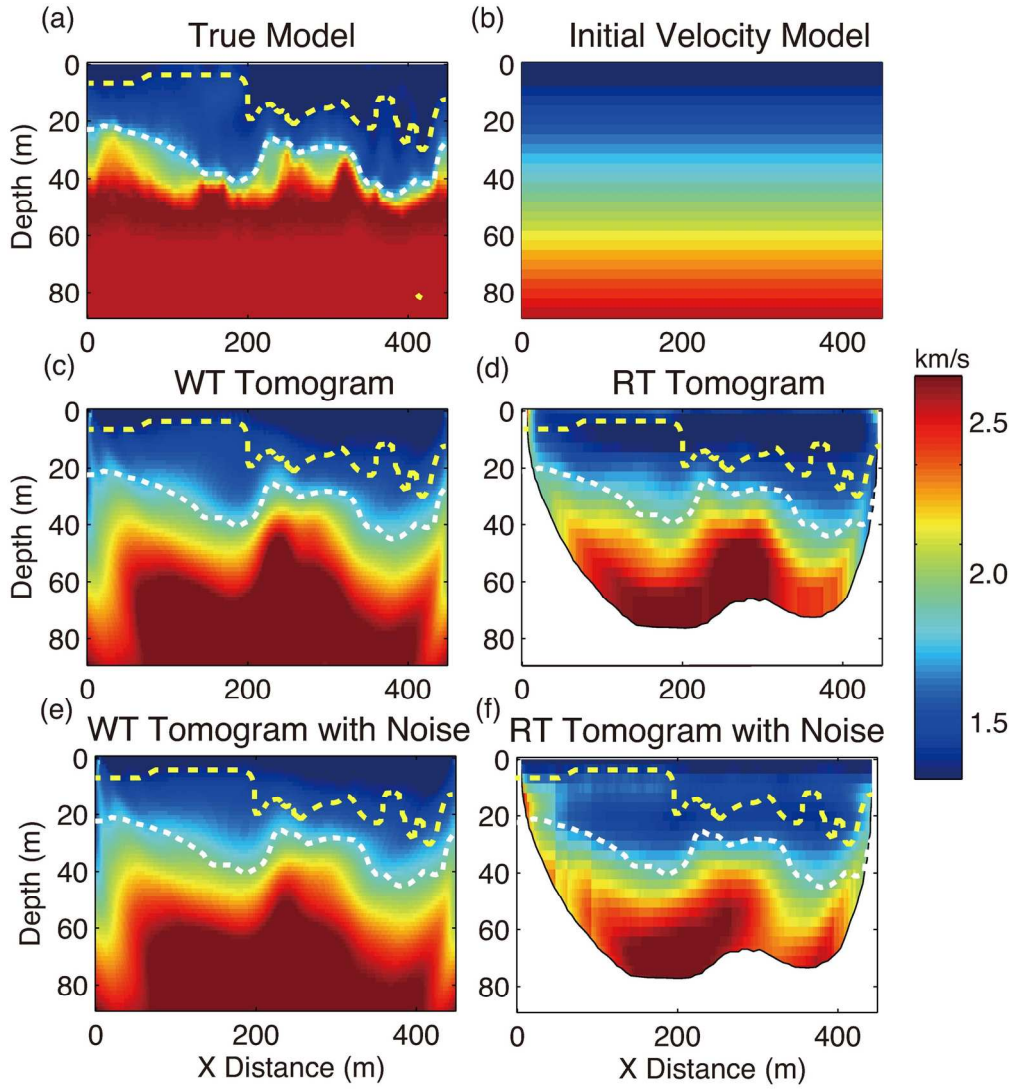
Figure 12: a) The traveltimes for shots 5, 65, 105, where the black, red and blue curves are the observed, WT, and RT traveltimes, respectivly. b) The normalized misfit function for the WT and RT methods.

Table 1: The computational metrics for the WT and RT methods.



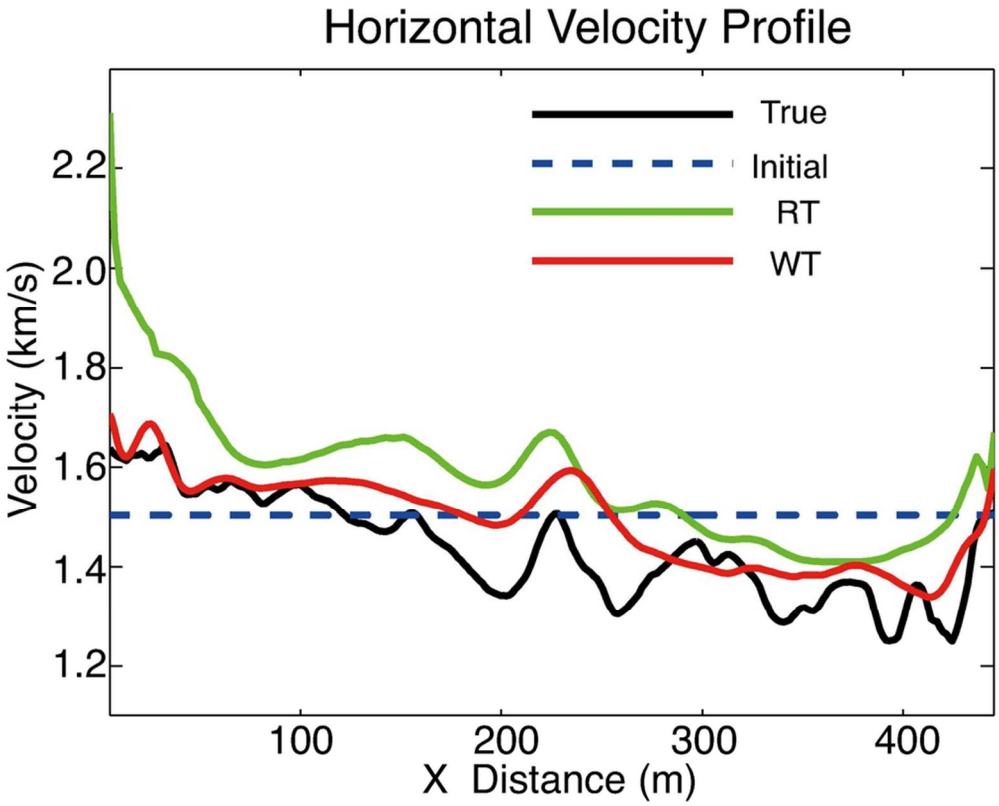
1. The flow chart for, a) WT inversion, where $\Delta \epsilon$ for WT is the misfit change percentage, and we set the tolerance is 0.01, and b) RT inversion, where $\Delta \epsilon$ for RT is the RMS residual, and the tolerance we set is a quarter of the dominant period.

190x90mm (300 x 300 DPI)



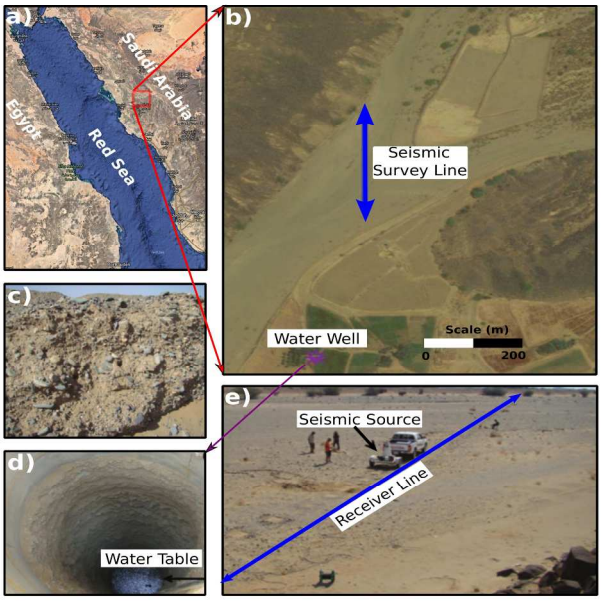
2. a) The true velocity, b) initial-velocity models for both the WT and RT methods, c) WT, d) RT, e) WT tomograms with 10\% random noise, and f) RT tomogram with 10\% random noise. For this velocity model, the characteristic horizontal variation of the velocity is $\approx 20~m$ along the interface for 25 m $ < z < $ 45 m. Thus the high-frequency assumption of ray tracing is violated for a 50-Hz source wavelet having a wavelength about 40 m along this interface. The yellow and white dashed lines represent the velocity of 1550 m/s and 1800 m/s, respectively, in the true model.

161x176mm (300 x 300 DPI)



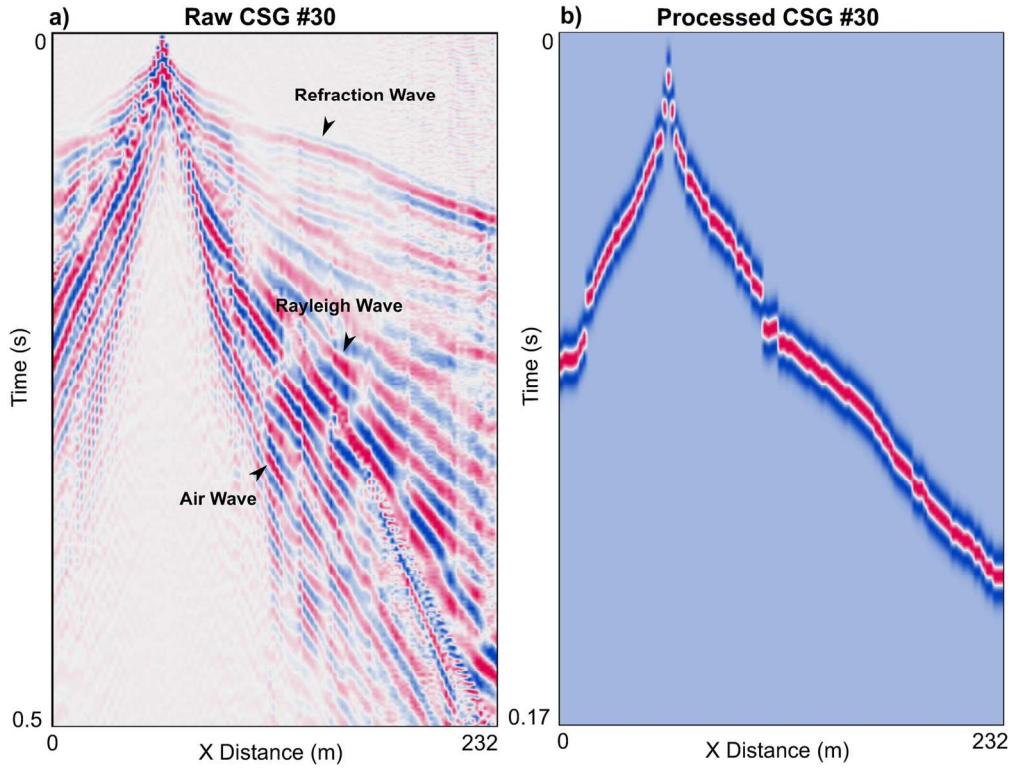
3. The comparison of the WT and RT velocity profiles at the depth of 25 m, where the black, blue, green and red curves represent the true, initial, RT, and WT velocity. This profile suggests that the WT method is more accurate than the RT method.

85x69mm (300 x 300 DPI)



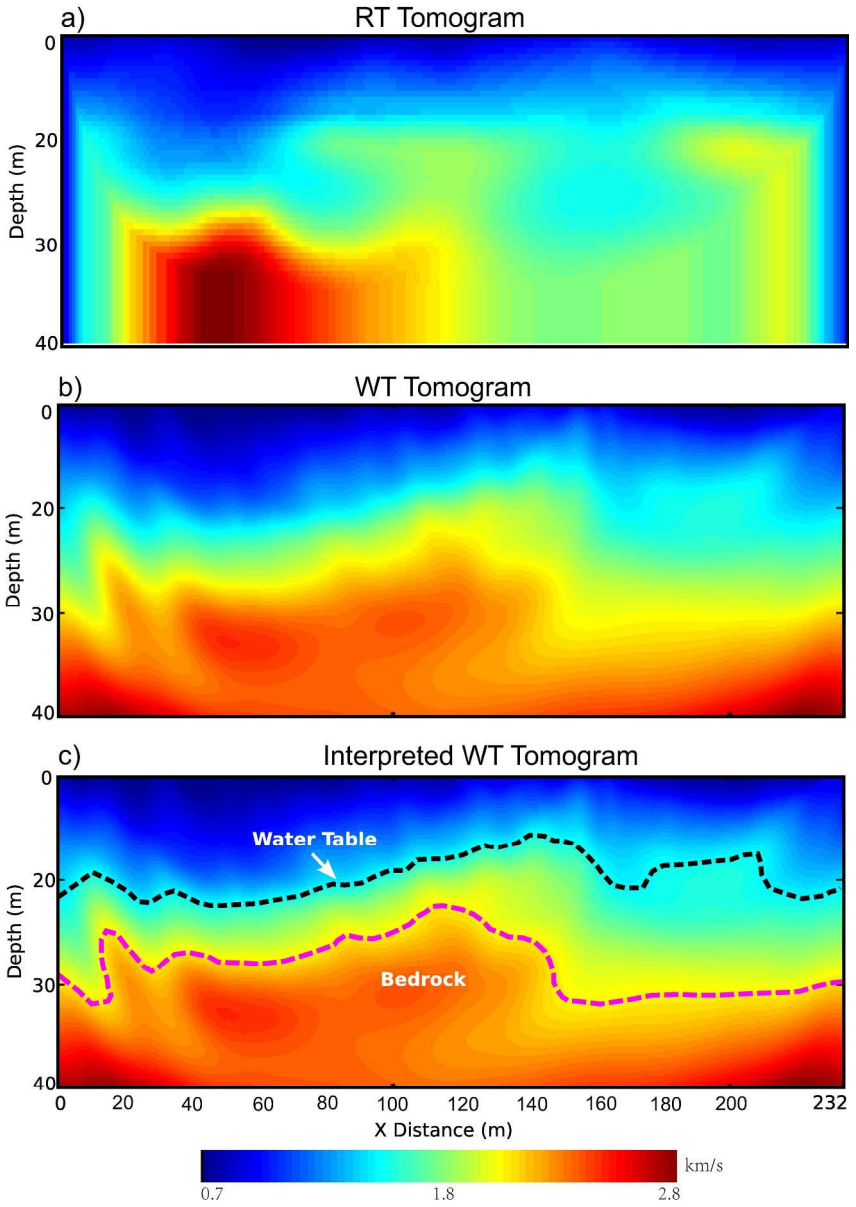
4. a) Google map shows the location of the Wadi Qudaid experiment, b) seismic survey line indicated by the blue double-arrow line, c) photograph showing the first layer composed of gravel and sand, with a thickness less than 5 m, d) a water well near the study area showing the second layer composed of sand and silt with some gravel, with the thickness of about 10-15 m, and the water table is at $z = 18 $ m, and the e) seismic source is excited by a 90-kg accelerated-weight drop.

297x420mm (300 x 300 DPI)



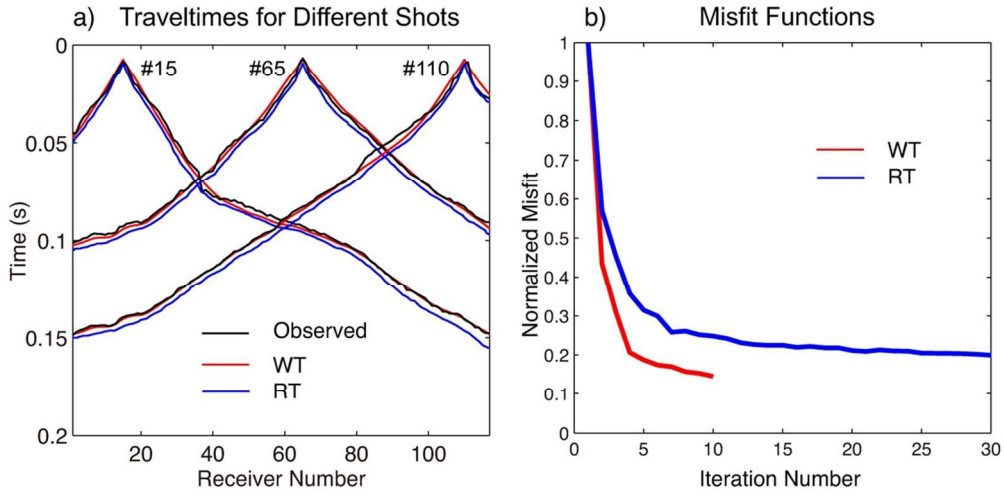
5. a) CSG \#30, where the air, Rayleigh, and refraction waves are indicated, and b) the processed CSG where only the refraction waves are extracted and processed.

148x112mm (300 x 300 DPI)



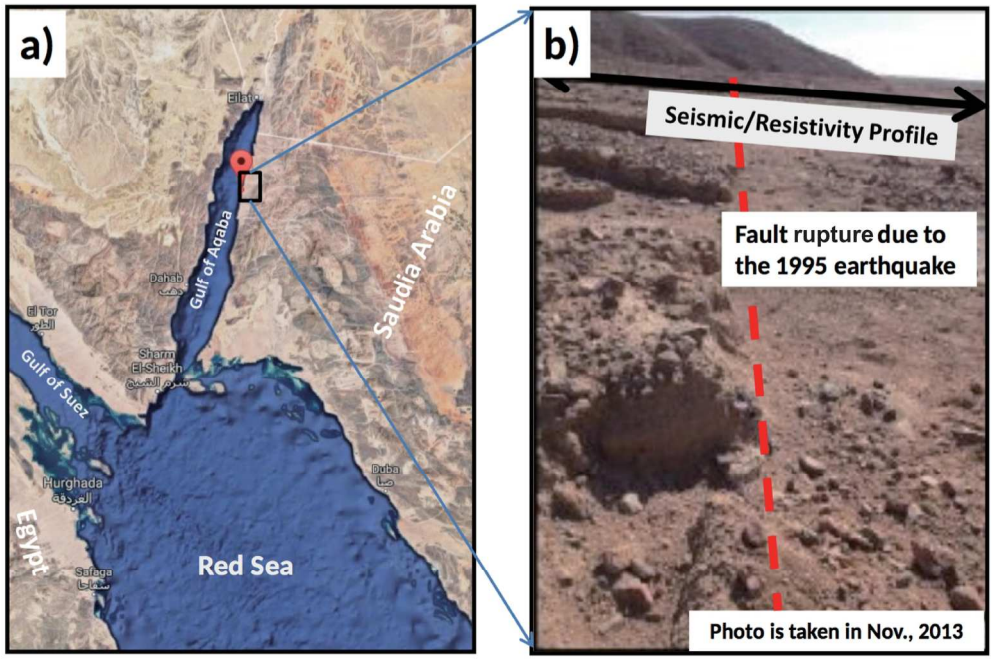
6. The RT and WT tomograms inverted from the Wadi Qudaid data: a) RT, b) WT, and c) interpreted WT tomograms. The black dashed curve is the interpreted water table and the pink dashed curve is the interpreted bedrock interface.

218x311mm (300 x 300 DPI)



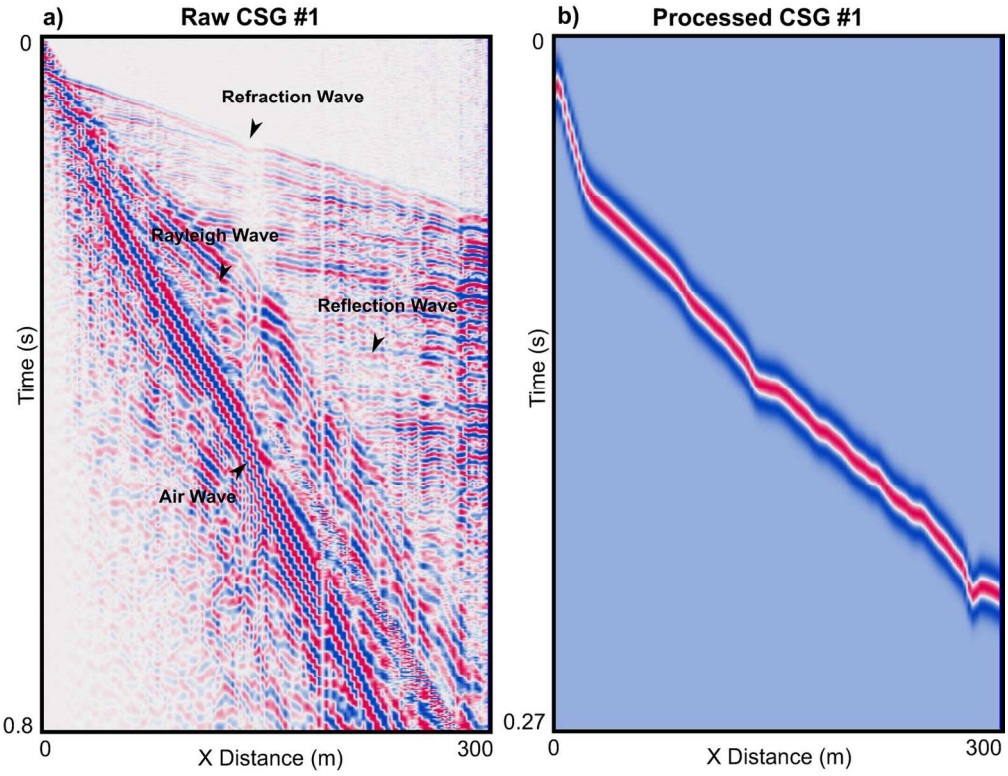
7. a) The traveltimes for shots 15, 65, 110, where the black, red and blue curves represent the observed, WT, and RT traveltimes, respectively, and b) the normalized misfit functions for the WT and RT methods.

102x54mm (300 x 300 DPI)



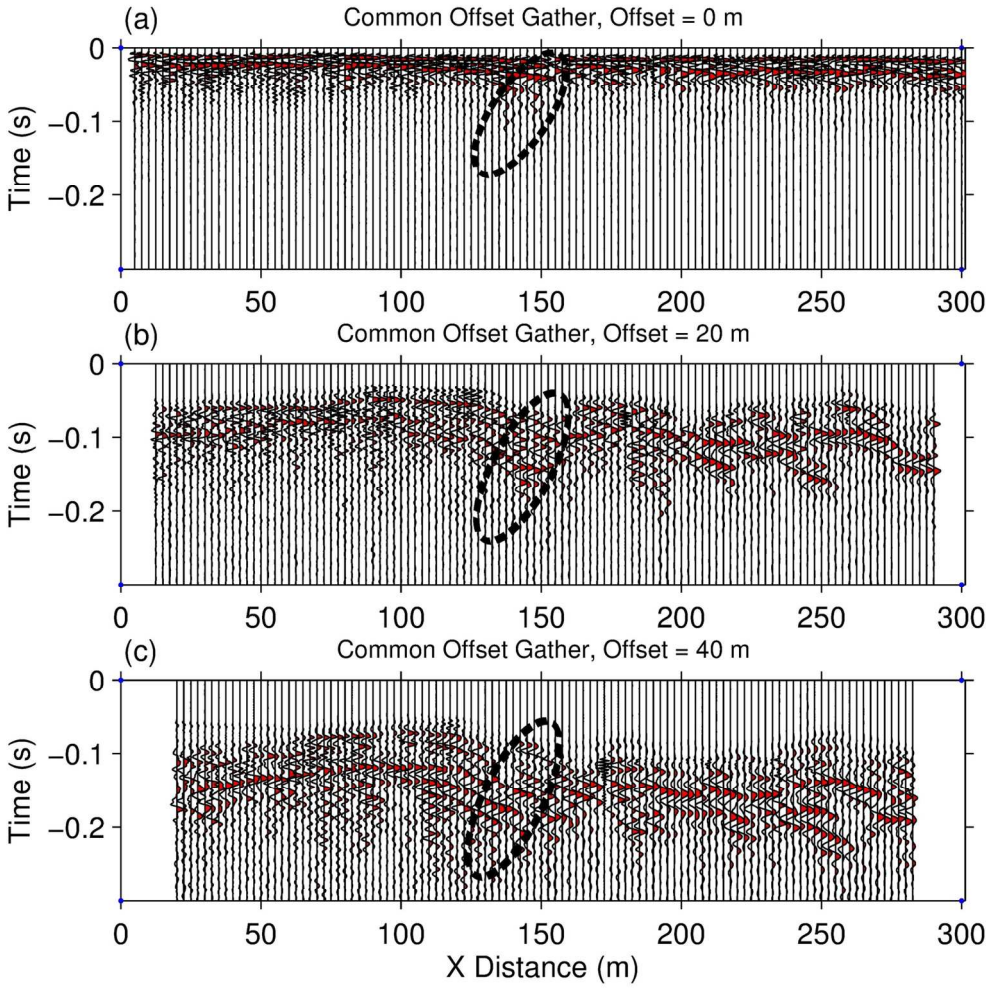
8. a) Google map shows the location of the Aqaba survey, b) photo shows an earthquake scarp (indicated by the red dashed line) on the surface, and the black arrow indicates the location of the seismic and resistivity survey lines.

163x110mm (300 x 300 DPI)



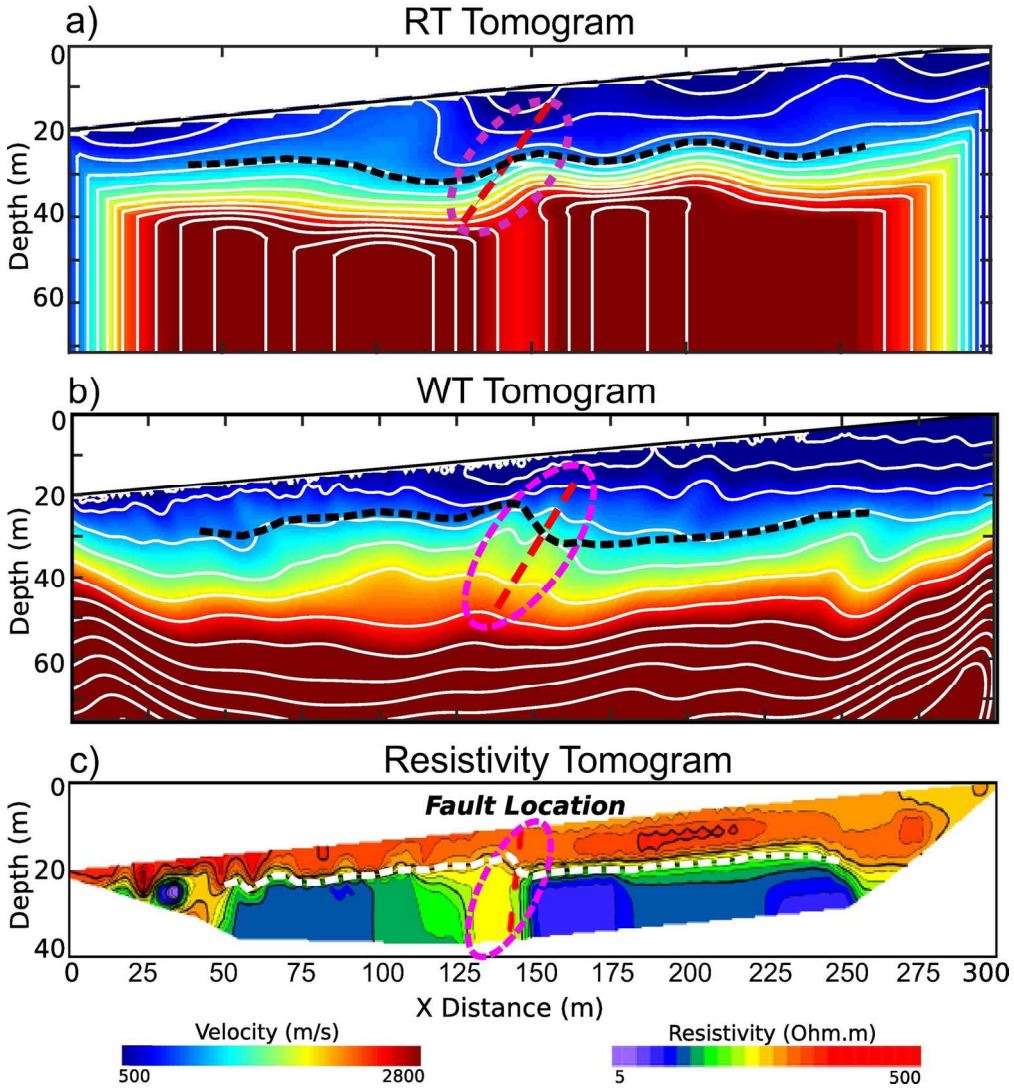
9. a) CSG\#1, where the air, Rayleigh, reflection and refraction arrivals are indicated, and b) the processed CSG, where only the refraction arrivals are extracted and processed.

148x114mm (300 x 300 DPI)



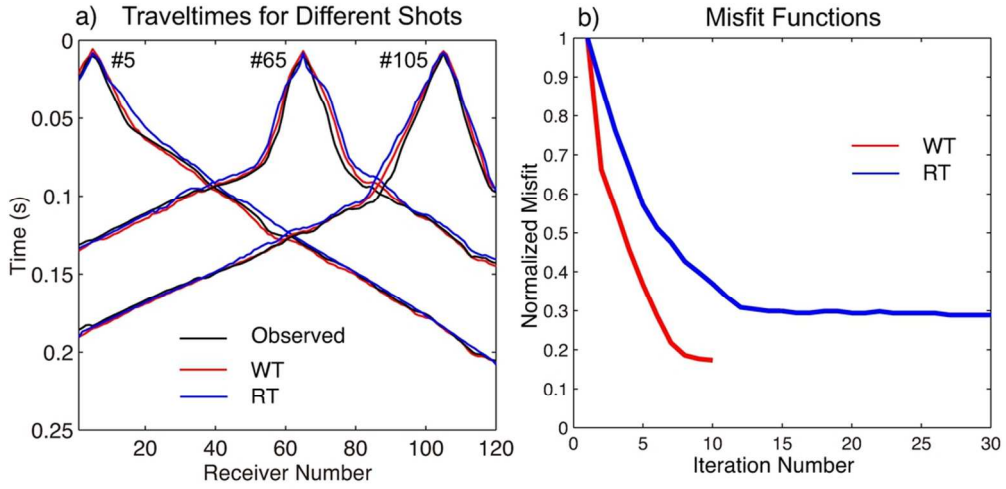
10. Common offset gathers for the source-receiver, a) offset = 0 m, b) offset = 20 m, and c) offset = 40 m.

145x145mm (300 x 300 DPI)



11. The tomograms inverted from the Aqaba data: a) RT and b) WT tomograms, where the pink ellipses indicate possible fault locations and the black dashed curves represent the bedrock interfaces, c) resistivity tomogram inverted by Res2DInv, a low-resistivity anomaly is shown at $130 < x <145$ m, where the white dashed curve represents the bedrock interface.

147x161mm (300 x 300 DPI)



12. a) The traveltimes for shots 5, 65, 105, where the black, red and blue curves are the observed, WT, and RT traveltimes, respectivly. b) The normalized misfit function for the WT and RT methods.

95x46mm (300 x 300 DPI)