Simultaneous waveform inversion for velocity, density and source moments with application to seismic-while-drilling

Jinji Li, Scott Keating, Roman Shor and Kris Innanen

ABSTRACT

Full waveform inversion (FWI) is an optimization-based approach to estimating the subsurface parameter model that minimizes the difference between synthetic and real data iteratively. In practice, incomplete acquisition and illumination of the subsurface are strong limiting factors. Adding data corresponding to new and independent ray paths as input could lead to significant increases in the reliability of FWI models. In principle, seismic-while-drilling (SWD) can supply these additional ray paths but introduce a new suite of unknowns, namely precise source locations and radiation characteristics. Here we formulate a new elastic FWI algorithm in which source positions and radiation patterns join the velocity/density values of the grid cells as unknowns to be determined. We then carry out a synthetic feasibility study in which such incompletely-known sources are included along a plausible well-trajectory through a simulated model, around which seismic receivers are placed in various configurations. This SWD-FWI is optimized with a Truncated Gauss-Newton algorithm. The subsurface model and source properties (P-wave velocity, density, and three independent 2D moment tensor values) are recovered and analyzed. The analysis suggests that the participation of SWD improves the accuracy of FWI models, especially in density inversion. The inversion of elastic properties shows a similar improvement when different drilling paths participate, which is different than expected. Such results indicate further related study is required to provide more comprehensive radiation patterns of the SWD sources.

INTRODUCTION

Full waveform inversion (FWI) is a procedure for estimating models of the physical properties of the Earth. This method relies on the prediction of seismic waves through the forward problem, which simulates the amplitude and travel times of waves (Métivier et al., 2013). This modeling procedure can be performed either in time or frequency domains (Marfurt, 1984; Gauthier et al., 1986; Pratt, 1990). By iteratively minimizing the misfit function representing the distance in the data space between synthetic and recorded data, the subsurface models of various physical parameters, such as compressional velocity, shear velocity, density, attenuation factors, and even source mechanisms, can be gradually updated to become more precise than the initial one and finally get reconstructed (Tarantola, 1984; Virieux and Operto, 2009; Fichtner et al., 2009; Pan et al., 2015). In FWI, any waves generated in the subsurface, including reflected waves, multiple waves, diving waves, and so on, can be taken advantage of, so the complete information can provide a high-resolution outcome theoretically. Therefore, this technique is recognized as promising in both academical and industrial applications.

However, when applying FWI to field data, practical issues can strongly affect the accuracy of the estimated models. One of the key factors is limited seismic data acquisition. FWI is nonlinear and ill-posed, so dense acquisition is essential (Kazemi et al., 2018). Given acquisition of data with sources and receivers that surround a model, FWI results are typically very accurate and highly resolved; but, in reality, we often only have access to surface field data, meaning ray paths interacting with the unknown medium with transmission-like geometries are missing. Transmission ray paths (available in cross-well, VSP experiments, and to a limited extent provided by diving waves at long surface offsets) are crucial in the construction of the low-wavenumber components of the FWI model. In principle, any addition of new, independent, and especially transmission-like ray paths to the dataset to be included in FWI can change the accuracy of models fairly dramatically.

Seismic-while-drilling (SWD) can generate elastic waves along a specific drilling path, and surface receivers can record them. Figure 1 illustrates schematically the ray paths and radiation patterns of an SWD source. The drill-bit is typically modelled as a dipole source of compressional and transverse waves. This is suggestive that SWD datasets can potentially contribute to FWI in a complementary manner. In recent years, the potential of SWD in seismic imaging and inversion has attracted renewed attention (Kazemi and Sacchi, 2014; Poletto et al., 2020; Nejadi et al., 2020). Auriol et al. (2020) estimated the drill bit source signature by combining force measurements, top-drive velocity, and drill-string dynamics modeling with topside and bit-rock boundary conditions. This estimate was used by Kazemi et al. (2020) in scalar acoustic FWI tests, producing evidence of positive impact SWD datasets may have on FWI. However, a complete feasibility study that incorporates several additional practical constraints is needed. For instance, as is evident in Figure 1, drill radiation patterns have strong directionality features, and indeed large swathes of surface receiver locations are likely to be sensitive primarily to S-wave energy, both of which are inherently elastic phenomena. Furthermore, the complex character of the continuously-radiating source drillbit is unlikely to be plausibly modelled in FWI as being due to static, known, and simply-radiating point sources.

Motivated by this, we adapt the frequency domain multi parameter elastic FWI scheme of Keating and Innanen (2020), in which source positions and radiation patterns join the velocity/density values of the grid cells as unknowns to be determined, to the problem of SWD-FWI as driven by surface sensor acquisition geometries.

This paper is organized as follows. Firstly, we briefly review some related theories, including the inverse problem and the forward modeling problem. In the next section of numerical tests, we demonstrate several experiments to show how much the SWD datasets can improve the quality of FWI outcomes. Finally, further analysis, as well as possible improvement on this topic, are discussed.

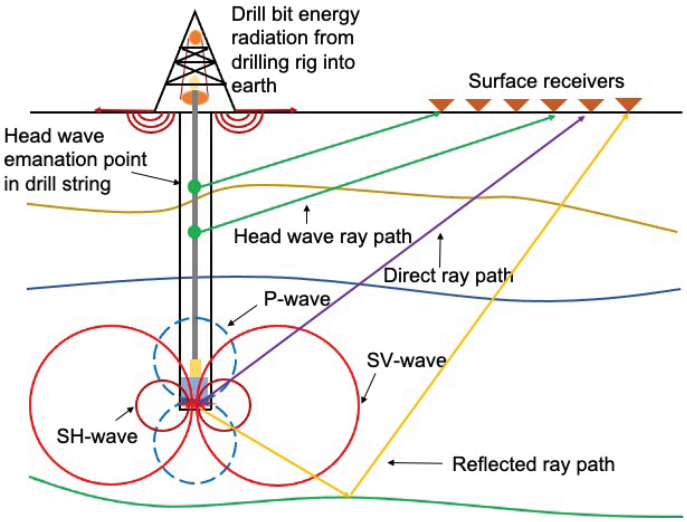

FIG. 1. Schematic figure of the radiation patterns of SWD sources.

Theory

Inverse problem

The inverse problem in FWI is a nonlinear framework for minimizing the objective function. In general, for a given frequency, the misfit function in the form of L2 norm is as follow:

$$min_{\boldsymbol{m}}\Phi(\boldsymbol{m}) = \frac{1}{2}\Sigma_{x_s}\Sigma_{x_g}\|\boldsymbol{R}\boldsymbol{u}(\boldsymbol{x_g},\boldsymbol{x_s}) - \boldsymbol{d}(\boldsymbol{x_g},\boldsymbol{x_s})\|_2^2, subject\ to\ S\boldsymbol{u} = \boldsymbol{f}, \tag{1}$$

where $\boldsymbol{m}$ is subsurface model, $\boldsymbol{R}$ is the sampling matrix which represents the measurements of receivers, $\boldsymbol{u}(\boldsymbol{x_g},\boldsymbol{x_s})$ stands for the wavefield from a source position to an arbitrary receiver, $\boldsymbol{d}(\boldsymbol{x_g},\boldsymbol{x_s})$ is the observed data between a source-receiver pair, S is the finite-difference modeling operator, and $\boldsymbol{f}$ is the source term. The calculation of objective function also includes the summation over frequencies, but we do not state it here for the convenience of the following sections. From the form of Equation 1, when $\Phi(\boldsymbol{m})$ reaches the minimum value, the difference between synthetic and observed data is minimized, which means the current model is the most consistent with reality. In this work, this optimization problem is often solved with the Newton approach, which starts by building a second-order Taylor-Lagrange expansion of $\Phi(\boldsymbol{m})$:

$$\Phi(\boldsymbol{m}+\Delta\boldsymbol{m}) \approx \Phi(\boldsymbol{m}) + \boldsymbol{g}^\dagger + \frac{1}{2}\Delta\boldsymbol{m}^\dagger\boldsymbol{H}\Delta\boldsymbol{m}, \tag{2}$$

where $\Delta \boldsymbol{m}$ means update direction, $\boldsymbol{g}$ is gradient, $g^{\dagger}$ denotes the transpose conjugate, $\boldsymbol{H}$ is the Hessian matrix, the second-order partial derivative of $\Phi(\boldsymbol{m}$ with respect to $\boldsymbol{m}$.

This quadratic approximation can be minimized iteratively. The updated model of the $(k+1)_{th}$ iteration can be calculated with the quantities of the previous iteration. This process can be written as:

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \sigma_k \Delta \boldsymbol{m}_k, \tag{3}$$

where σ_k denotes the scaling factor of $\Delta \boldsymbol{m}_k$, and it can be computed with globalization methods, usually the line search method. The decent direction $\Delta \boldsymbol{m}_k$ is the solution to a linear system, this linear system satisfies:

$$\boldsymbol{H}(\boldsymbol{m}_k)\Delta \boldsymbol{m}_k = -\boldsymbol{g}(\boldsymbol{m}_k), \tag{4}$$

The adjoint state method

In the above equation (Equation 4), $\boldsymbol{g}(\boldsymbol{m}_k)$ can be solved with the adjoint state method (Plessix, 2006). This method can construct the gradient through forward modeling and backward propagation in which the source term is replaced with residual wavefield.

Here for one source-receiver pair, the problem based on Equation 1 can be simplified as follows:

$$min_{\boldsymbol{m}} \Phi(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{R}\overline{\boldsymbol{u}} - \boldsymbol{d}\|_2^2, subject\ to\ S\overline{\boldsymbol{u}}(\boldsymbol{m}) = \boldsymbol{f}, \tag{5}$$

where $\overline{\boldsymbol{u}}$ denotes the solution of the forward problem. Then formulate the Lagrangian function:

$$min_{\boldsymbol{m}} \Phi(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{R}\overline{\boldsymbol{u}} - \boldsymbol{d}\|_2^2 + (S\overline{\boldsymbol{u}}(\boldsymbol{m}) - \boldsymbol{f}, \lambda), \tag{6}$$

where λ is a Lagrange multiplier. The derivative of L with respect to $\boldsymbol{m}$ is:

$$\frac{\partial L}{\partial \overline{\boldsymbol{u}}}\frac{\partial \overline{\boldsymbol{u}}}{\partial \boldsymbol{m}} + \frac{\partial L}{\partial \boldsymbol{m}}, \tag{7}$$

In the equation above, $\frac{\partial \overline{\boldsymbol{u}}}{\partial \boldsymbol{m}}$ is very costly to solve directly, and it is more efficient to

solve with the following way. First define the adjoint state $\overline{\lambda}$ such that $\frac{\partial L}{\partial \overline{\boldsymbol{u}}}$, this condition is equivalent to

$$S^{\dagger}\overline{\lambda} = \boldsymbol{R}^{T}(\boldsymbol{d} - \boldsymbol{R}\overline{\boldsymbol{u}}(\boldsymbol{m})). \tag{8}$$

Therefore, the gradient $\boldsymbol{g}(\boldsymbol{m})$ is obtained by:

$$\boldsymbol{g}(\boldsymbol{m}) = (\frac{\partial}{\partial \boldsymbol{m}} S\overline{\boldsymbol{u}}, \lambda). \tag{9}$$

In this work, the recovery of moment tensors is also considered. The derivative of the objective function with respect to force terms is divided into real and imaginary parts and will be the summation over all the components of the Lagrange multiplier:

$$\frac{d\boldsymbol{\Phi}}{d\boldsymbol{f}} = \Sigma_i(-\mathfrak{Re}(\overline{\lambda_i}) + \mathfrak{Im}(\overline{\lambda_i})), \tag{10}$$

where the first part in the right-hand side is the derivative of Φ with respect to the real part, and the second part is the imaginary part. However, this work is based on a 2-D problem, and these two derivatives should be extended with respect to x and z directions using Chain Rules:

$$\begin{cases} \frac{d\boldsymbol{\Phi}}{d\boldsymbol{f_x}} = \Sigma_i(\frac{d\boldsymbol{\Phi}}{df_{R_i}}\frac{df_{R_i}}{df_x} + \frac{d\boldsymbol{\Phi}}{df_{I_i}}\frac{df_{I_i}}{df_x}) = \Sigma_i(-\mathfrak{Re}\frac{df_{R_i}}{df_x} + \mathfrak{Im}\frac{df_{I_i}}{df_x}), \\ \frac{d\boldsymbol{\Phi}}{d\boldsymbol{f_z}} = \Sigma_i(\frac{d\boldsymbol{\Phi}}{df_{R_i}}\frac{df_{R_i}}{df_z} + \frac{d\boldsymbol{\Phi}}{df_{I_i}}\frac{df_{I_i}}{df_z}) = \Sigma_i(-\mathfrak{Re}\frac{df_{R_i}}{df_z} + \mathfrak{Im}\frac{df_{I_i}}{df_z}). \end{cases} \tag{11}$$

The Truncated Gauss-Newton approach

The $\delta\boldsymbol{m}$ term requires the inverse Hessian operator, which is not efficient to calculate directly. Métivier et al. (Métivier et al., 2013) illustrated by numerical experiments that the Truncated Gauss-Newton method may provide better results in the FWI context. We refer to their work and employ the truncated Gauss-Newton optimization method in this work. The Truncated Gauss-Newton method is based on the Gauss-Newton approximation of Hessian which keeps the first-order part of it and thus obtains the diagonal part of the Hessian matrix (Nash, 2000). Consequently, we use the conjugate gradient algorithm to calculate the descent direction $\delta\boldsymbol{m}$.

The first-order term of the Hessian matrix can be implicitly calculated by first building a vector product:

$$(\widetilde{\boldsymbol{H}}, \boldsymbol{\nu}) = \boldsymbol{J}^{\dagger}\boldsymbol{R}^{T}\boldsymbol{R}\boldsymbol{J}\boldsymbol{\nu}, \tag{12}$$

where $\boldsymbol{\nu}$ is an arbitrary vector, $\boldsymbol{J}$ is the Jacobian matrix. Replace $\boldsymbol{R}^{T}\boldsymbol{R}\boldsymbol{J}\boldsymbol{\nu}$ in the right-hand side of Equation 12 is with another arbitrary vector $\boldsymbol{w}$, and derive the minimization problem in this case:

$$min_{\boldsymbol{m}}\Psi = (\boldsymbol{u}, \boldsymbol{w}), subject\ to\ S\boldsymbol{u} = \boldsymbol{f} \tag{13}$$

The derivative of Ψ with respect to $\boldsymbol{m}$ is

$$\frac{d\Psi}{d\boldsymbol{m}} = (\boldsymbol{J}^{\dagger}, \boldsymbol{w}). \tag{14}$$

Similarly, this problem can be solved by the adjoint state method. First, formulate the Lagrangian function of this new minimization problem:

$$L = (\boldsymbol{u}, \boldsymbol{w}) + (S\boldsymbol{u} - \boldsymbol{f}, \mu). \tag{15}$$

where u is the solution to the forward problem, μ is the Lagrange multiplier. To apply the adjoint state method, first calculate the derivative of L with respect to $\boldsymbol{m}$, and define the adjoint state $\boldsymbol{u}$ such that the derivative of it with respect to $\boldsymbol{m}$ can be eliminated, which satisfies that

$$S^{\dagger}(\boldsymbol{m}) = -\boldsymbol{w} = -\boldsymbol{R}^{T}\boldsymbol{R}\boldsymbol{J}\boldsymbol{\nu}. \tag{16}$$

Note that there is also a Jacobian matrix. To eliminate this J term, consider deriving the forward problem with respect to the i_{th} parameter m_i in a random direction ν_i, and sum over all i provides:

$$S\boldsymbol{\alpha} = \boldsymbol{\xi}, \tag{17}$$

where

$$\boldsymbol{\alpha} = J\boldsymbol{\xi}, \tag{18}$$

is considered as the solution to forward problem in which the source term is

where

$$\boldsymbol{\xi} = -\Sigma_i(\frac{\partial}{\partial \boldsymbol{m}_i} S v_i)\overline{\boldsymbol{u}}, \tag{19}$$

Suppose $\overline{\alpha}$ is the solution to Equation 17. Replace $J\boldsymbol{\nu}$ in Equation 16 with this term, and then solve that equation for the adjoint state μ. With this μ, the gradient in Equation 14 can be solved:

$$\frac{d\Psi}{d\boldsymbol{m}} = \mathfrak{Re}(\frac{\partial}{\partial \boldsymbol{m}} S\overline{\boldsymbol{u}}, \overline{\mu}). \tag{20}$$

This gradient term is the Gauss-Newton Hessian vector product of the model of subsurface properties.

Follow the same procedure in the case of source term, it can be shown that the real and imaginary parts satisfy

$$\begin{cases} S\boldsymbol{\beta} = \boldsymbol{\eta}, \\ S\boldsymbol{\gamma} = \boldsymbol{\zeta}, \end{cases} \tag{21}$$

and

$$\begin{cases} \boldsymbol{\eta} = \Sigma_i \mathfrak{Re}(\nu_i), \\ \boldsymbol{\zeta} = \Sigma_i(-i\mathfrak{Im}(\nu_i)). \end{cases} \tag{22}$$

After getting η and ζ, Equation 21 can be solved both for the real part β and the imaginary part γ of the source term, substitute these terms into Equation 16 to replace $J\nu$, the adjoint state for both real and imaginary parts will be solved. The derivatives are as follows:

$$\begin{cases} \frac{d\Phi}{df_R} = -\mathfrak{Re}(\overline{\mu}_R) \\ \frac{d\Phi}{df_I} = -i\mathfrak{Im}(\overline{\mu}_I). \end{cases} \tag{23}$$

Therefore, the Gauss-Newton Hessian vector product of source properties can be solved. Therefore, the linear problem in Equation 4 can be solved via L-BFGS method (Nocedal and Wright, 2006). After that, the updated model can be acquired with Equation 3, and the step size is given by the Line-search method (Nocedal and Wright, 2006).

Forward problem

This work uses the anelastic finite-difference approach to solve the isotropic viscoelastic wave equations in the frequency domain. The 2-dimensional viscoelastic wave equations can be expressed as follows:

$$\begin{cases} \omega^2 \rho u_x + \frac{\partial}{\partial x}[\widetilde{\lambda}(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z}) + 2\widetilde{\mu}\frac{\partial u_x}{\partial x}] + \frac{\partial}{\partial z}[\widetilde{\mu}(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z})] + f_x = 0, \\ \omega^2 \rho u_z + \frac{\partial}{\partial z}[\widetilde{\lambda}(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z}) + 2\widetilde{\mu}\frac{\partial u_z}{\partial z}] + \frac{\partial}{\partial x}[\widetilde{\mu}(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z})] + f_z = 0. \end{cases} \tag{24}$$

where ω^2 is the square of angular frequency, ρ is the density, u_x and u_z are displacements in the x and z direction, f_x and f_z are the source terms in the horizontal and vertical direction, respectively. $\widetilde{\lambda}$ and $\widetilde{\mu}$ are complex Lamé parameters, and are frequency-dependent and related to quality factors Q_p and Q_s:

$$\begin{cases} \widetilde{\lambda} = V_p[1 + \frac{1}{Q_p}(\frac{1}{\pi} log\frac{\omega}{\omega_0} + \frac{i}{2})]^2 \rho - 2\widetilde{\mu}, \\ \widetilde{\lambda} = V_s[1 + \frac{1}{Q_s}(\frac{1}{\pi} log\frac{\omega}{\omega_0} + \frac{i}{2})]^2 \rho. \end{cases} \tag{25}$$

where ω_0 is the reference frequency, i is the imaginary unit.

Figure 2 shows how moment tensors are arranged to form source terms in a 2-D space, different colors represent different values. f_x and f_z are defined by nine grid cells surrounding the source position. The central one is zero, and the rest cells are functions of moment tensors. Specifically, f_x is the function of M_{11} and M_{12}, and f_z is the function of M_{22} and M_{12}. The normal value range for a moment tensor is between -1 to 1.

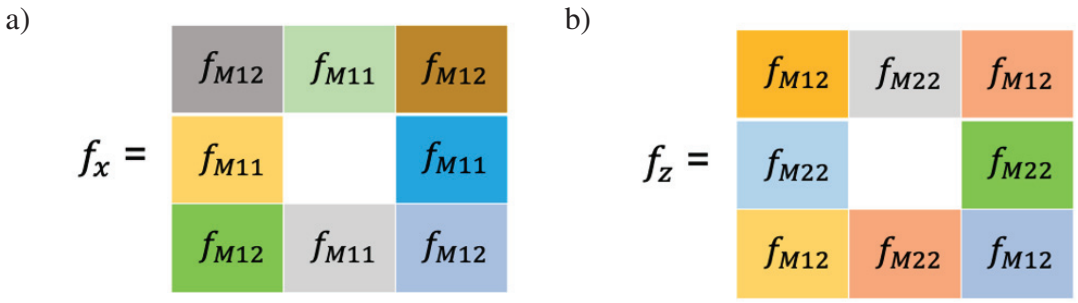

FIG. 2. The definition of source terms in a 2-D space. a) horizontal source term, b) vertical source term.

FWI-SWD FEASIBILITY ANALYSIS

The numerical tests in this section are based on synthetic models to evaluate the performances and investigate the potentiality of combining the seismic-while-drilling dataset

to the simultaneous model-source full waveform inversion. Besides, different acquisition systems, from local/all-surface seismic to vertical/deviate logging path, are applied to this model. Other parameters, such as the frequency bands, the space interval between grid cells, the numbers of iterations, and so on, remain constant throughout all the tests. In the inversions of this section, multi-scale frequency bands are used. There are ten total frequency bands, each containing six frequencies (Figure 3). The starting frequencies of every band are set to 1 Hz, while the ending frequencies linearly increase from 2.5 Hz to 25 Hz. Within each frequency band, the subfrequencies are also spaced linearly, varying from minimum to maximum frequencies of that band. In the truncated Gauss-Newton approach, two iterative processes are required. The outer loop is to solve the nonlinear system with two iterations, while the outcome of the inner loop after 40 iterations is $\delta \boldsymbol{m}$.

The model space is composed of the physical properties of subsurface and moment tensors defined on the positions of sources. According to the forward problem, subsurface properties should include V_p, V_s, Q_p, Q_s, and ρ. However, in this exploratory investigation, we initially try the inversion of V_p and ρ for convenience. It is necessary to convert the subsurface models in terms of Gaussian regions centered around each grid point to improve the convergence, and the inter-parameter confusion raised by the different changing scales of source and model parameters can be avoided (Keating and Innanen, 2020). The true models are shown in Figure 4. The horizontal size of the models is considerably larger for future numerical tests combining the far-sided SWD data, and our inversion target is a part (nearly a half horizontally) of the larger model, which is indicated with the red dashed rectangle.

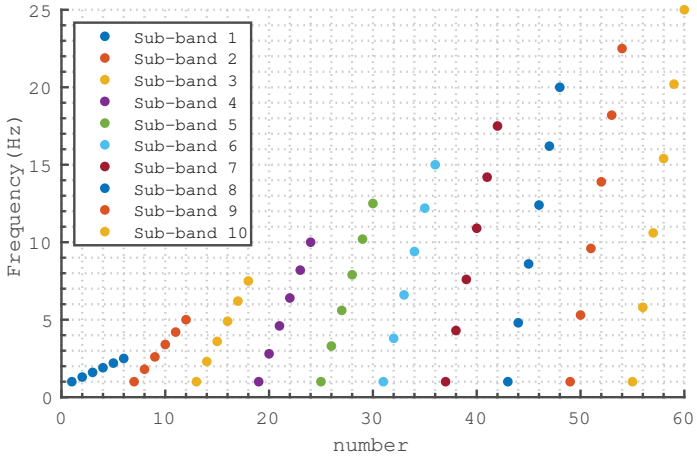

FIG. 3. Frequency components in numerical tests.

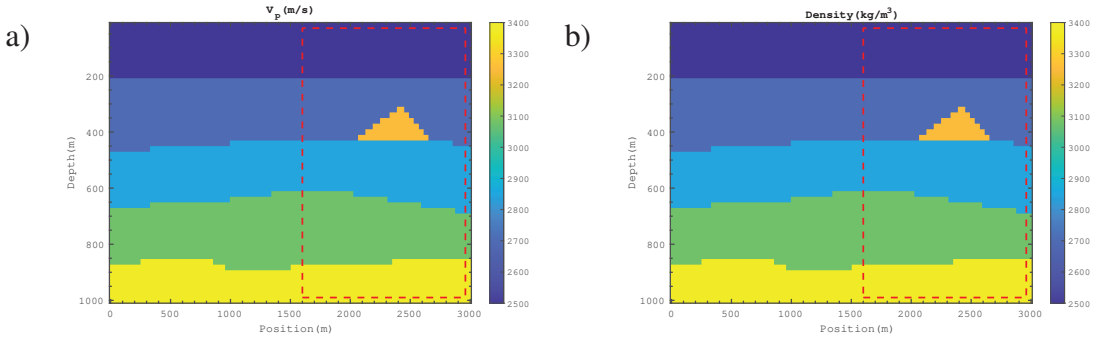

FIG. 4. True subsurface properties of the synthetic model. The inversion targets are circled by red dashed rectangles. a) Compressional velocity, b) Density.

Case 1: local-surface acquisition

In this case, the numerical examples start with a general surface seismic inversion. Then two drilling paths are added separately to measure how much contribution will the SWD datasets provide. The vertical drilling path is perpendicular to the surface, and twenty sources are evenly located from the depth of 400 meters to 700 meters of this path. The oblique drilling path is added to the same vertical range of the model, and the center of this path shares the same horizontal position. Here, for simplicity, the three cases in Case 1 are called Case 1 a), Case 1 b), and Case 1 c), respectively. Every first case in each experiment is called "raw" inversion because of no added SWD data. The acquisition systems are shown in Figure 5.

The inversion result of subsurface properties are shown in Figure 6. The left-side subplot is the result of compressional velocity, and the right-side subplot is the result of density. From Figure 6 a), which denotes for surface, or "raw" inversion, we can observe that even though the layered structure is approximately recovered, there is a bunch of detailed information for the inversion of compressional velocity being missed by us. However, when the deeper sources are added (Figure 6 c), e)), our inversion gets better. Things are similar in the inversion of density, but we can get a better resolution here in every case (Figure 6 b), d), f)).

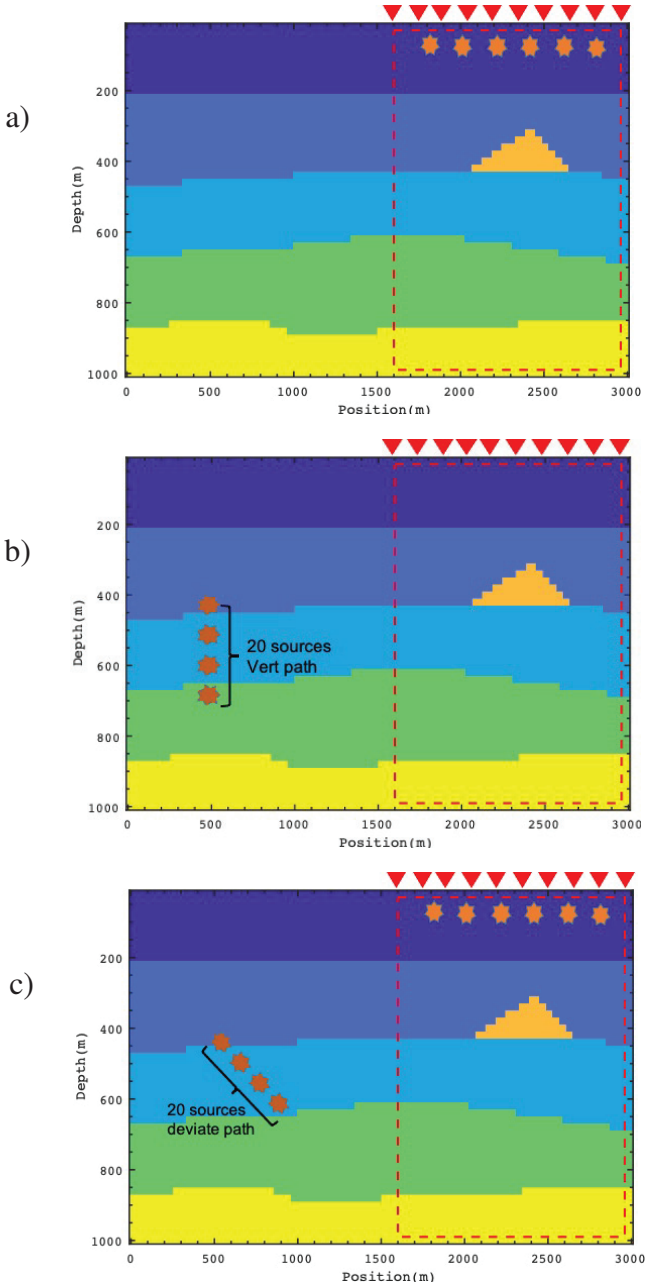

FIG. 5. Acquisition systems in Case 1. a) Local-surface acquisition, b) A vertical drilling path is added, c) A deviate drilling path is added.

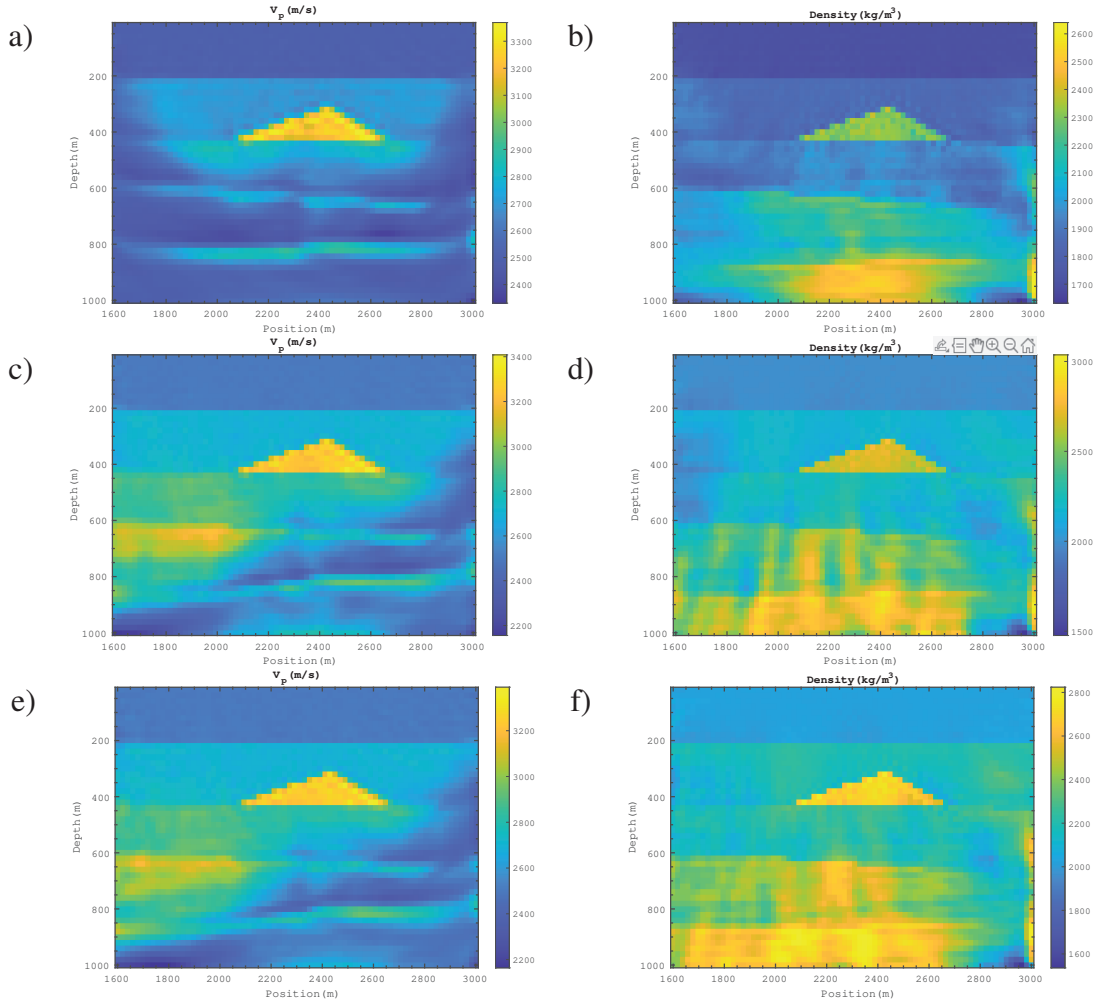

FIG. 6. Inversion result of subsurface properties. a) Compressional velocity, b) Density, c) Compressional velocity (vertical drilling path added), d) Density (vertical drilling path added), e) Compressional velocity (deviate drilling path added), f) Compressional velocity (deviate drilling path added).

Figure 7 is a set of cross plots of true and inverted moment tensors. The x-axis is true values, and the y-axis is estimated values. The subplots from left to right side are M_{11}, M_{12}, M_{22}, respectively. The red cross marks are moment tensors of surface seismic sources, while the blue circle marks are the moment tensors of drilling sources. It is shown that the recovery of surface moment tensors is quite accurate because of the consistency with the benchmark line (the red one). However, the blue circle marks show that the recovery of the M_{11} and M_{22} component of deeper sources is not so satisfying, although they all have a converging trend to the red lines.

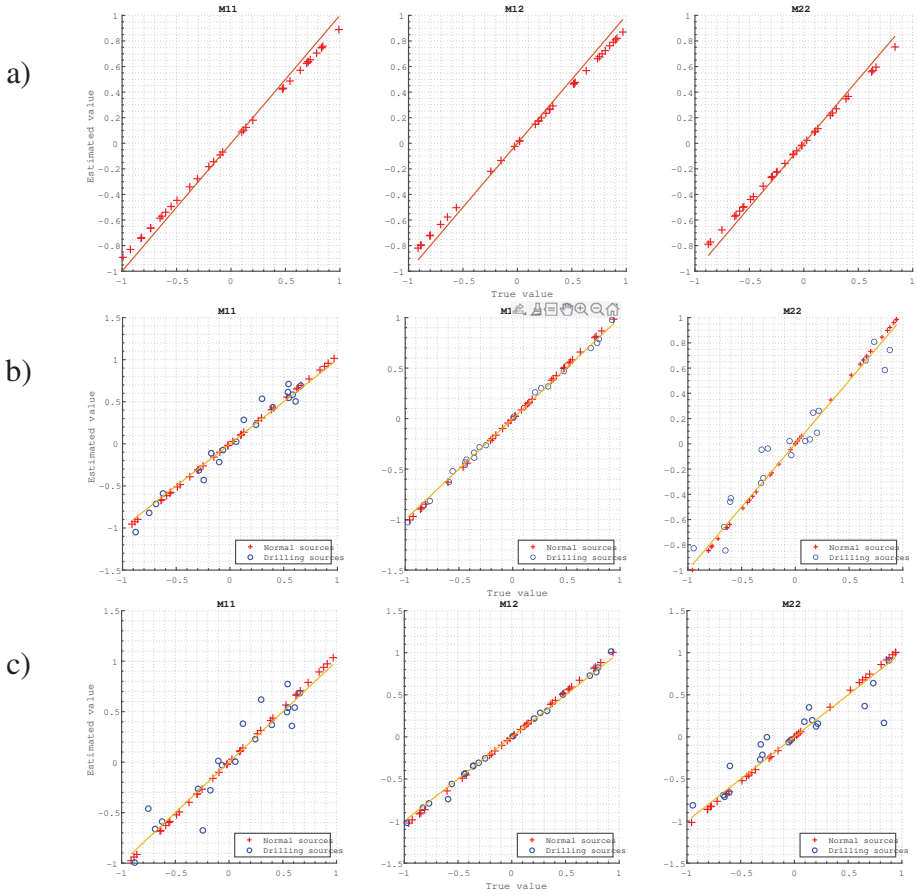

FIG. 7. Cross plots of the true and estimated moment tensors. a) Cross plots of the surface moment tensors, b) Cross plots of the surface and vertical path moment tensors, c) Cross plots of the surface and deviate path moment tensors.

A quantification section is required to assess the contribution of SWD data to the inversion of subsurface properties. Three vertical profiles located horizontally 1960 m, 2360 m, and 2760 m, are used to compare the results. Figure 8 and Figure 9 are the comparison of the recovery of compressional velocity and density, respectively. The yellow lines come from Case 1 a), the red lines come from Case 1 b), the blue lines come from Case 1 c), while the black lines are true values. The improvement seems significant when the SWD data is added to FWI, especially in the position closer to the drilling path. For the density case, we can observe that there is a considerable enhancement since the profiles of Case 1 b) and Case 1 c) are much more close to true values.

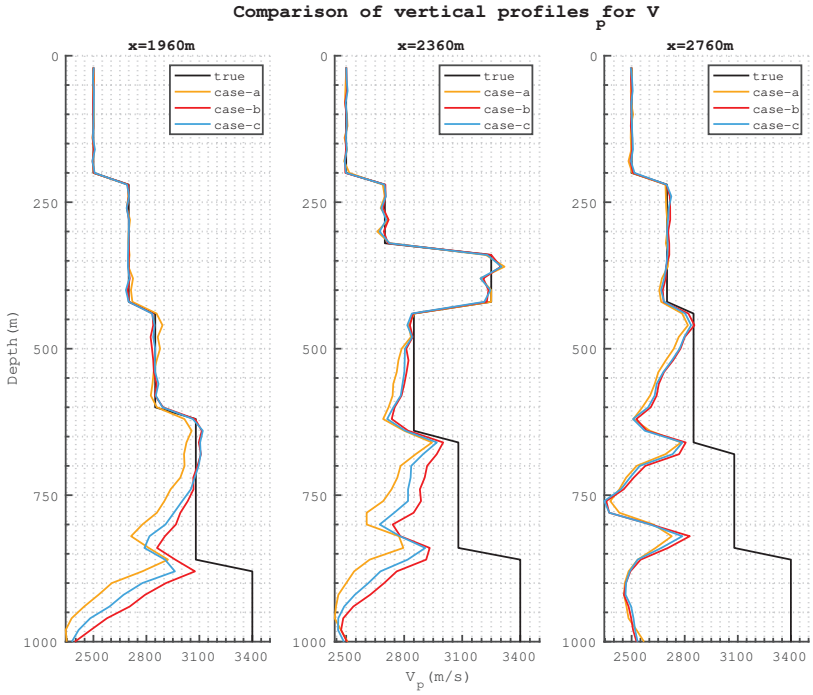

FIG. 8. Vertical profiles of compressional velocity.

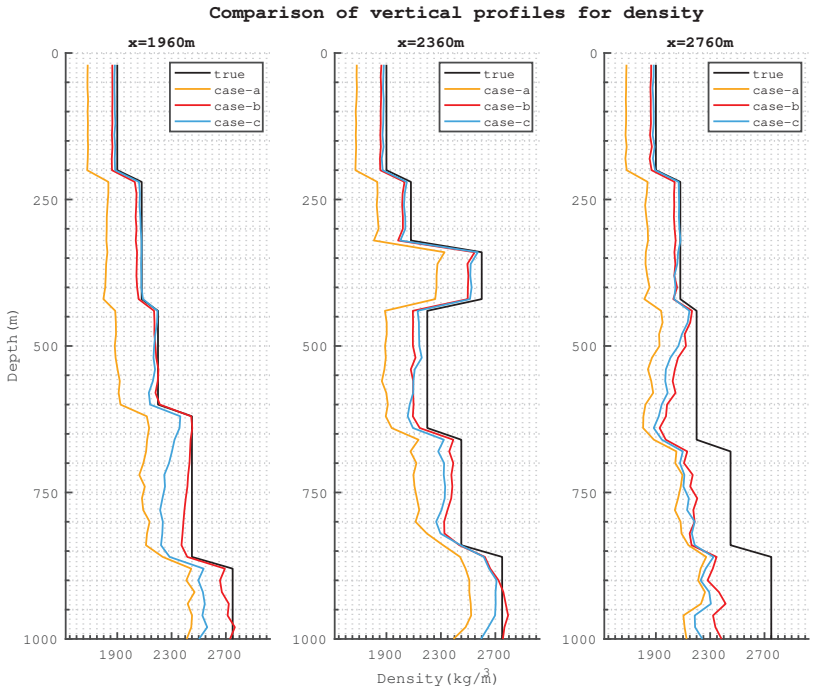

FIG. 9. Vertical profiles of density.

Case 2: all-surface acquisition

In this case, an all-surface acquisition system is implemented. Figure 10 shows the acquisition systems. 150 receivers are distributed with an interval of 20 m on the whole surface of the model. Samely, the vertical and derivate drilling paths are added. The three cases in Case 2 are called Case 2 a), Case 2 b), and Case 2 c), respectively.

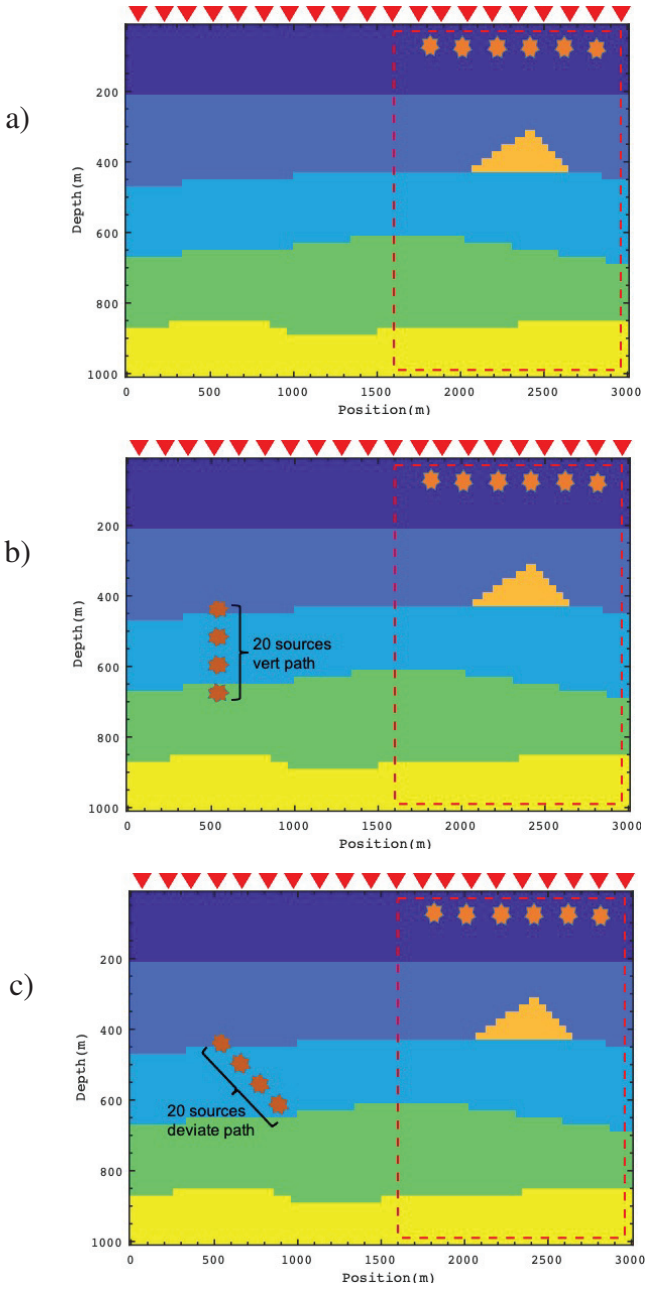

FIG. 10. Acquisition systems in Case 2. a) All-surface acquisition, b) A vertical drilling path is added, c) A deviate drilling path is added.

Figure 11 is the inversion results of this set of numerical experiments. It is observed that the energy generated by the deeper sources can enhance the inversion performance. The recovery of density is also improved, and the resolution is higher.

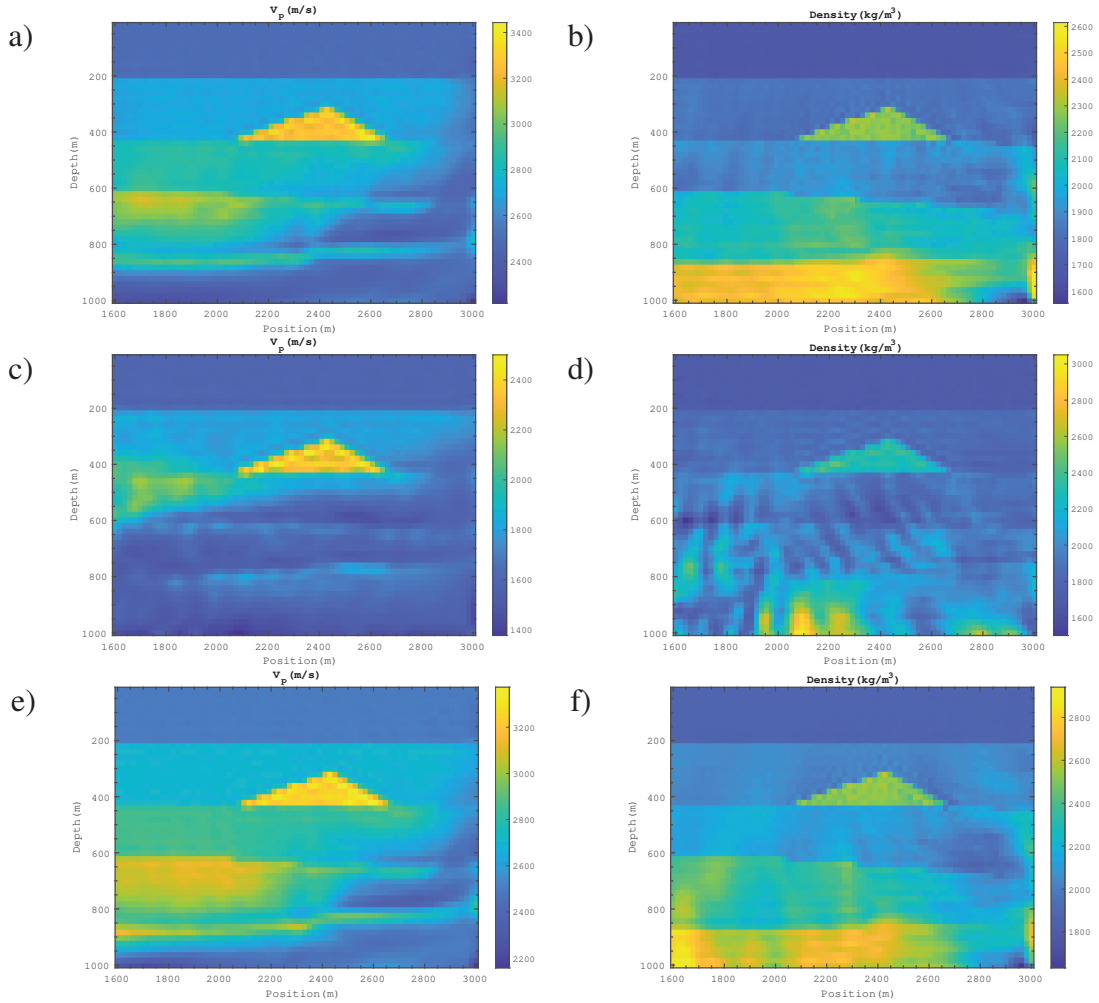

FIG. 11. Inversion result of subsurface properties. a) Compressional velocity, b) Density, c) Compressional velocity (vertical drilling path added), d) Density (vertical drilling path added), e) Compressional velocity (deviate drilling path added), f) Compressional velocity (deviate drilling path added).

Figure 12 shows the estimation of moment tensors. In the "raw" inversion (Case 2 a)), there is a small deviation of surface moment tensors. However, when we add SWD datasets, the moment tensors get perfectly inverted. In the comparison part (Figure 13, Figure 14), it is shown that when we apply the SWD data to the inversion, both Vp and Rho get better recovered. However, we still have a hard time precisely recovering the parts that are pretty far from the surface and deeper sources.

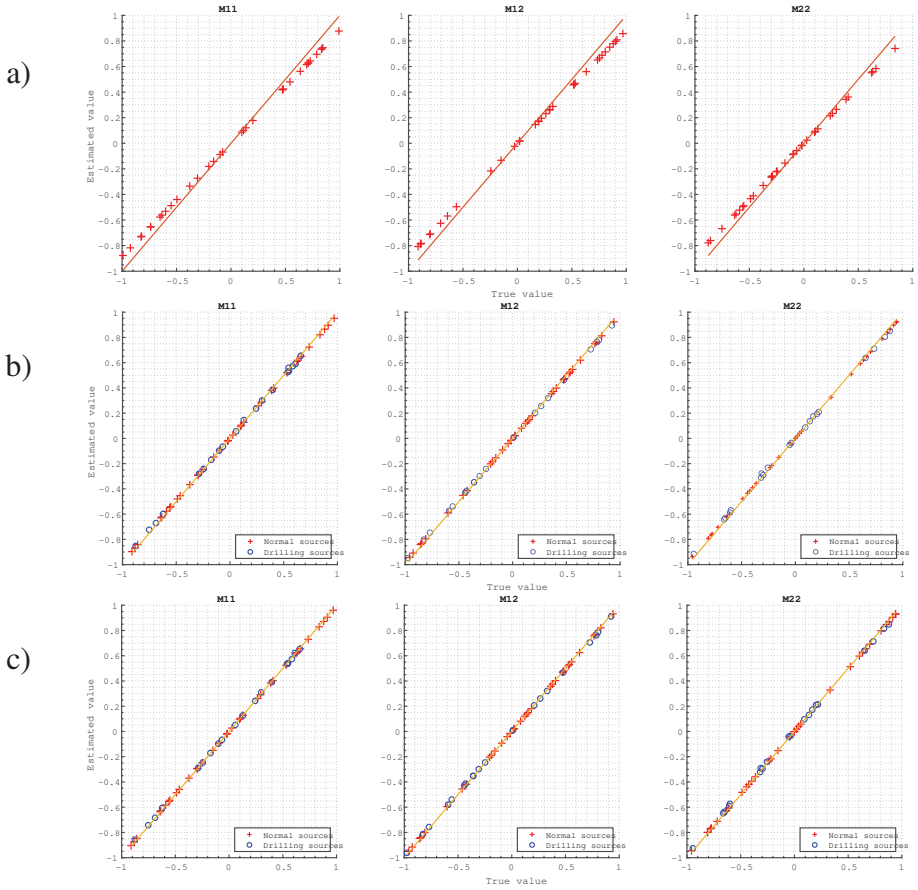

FIG. 12. Cross plots of the true and estimated moment tensors. a) Cross plots of the surface moment tensors, b) Cross plots of the surface and vertical path moment tensors, c) Cross plots of the surface and deviate path moment tensors.

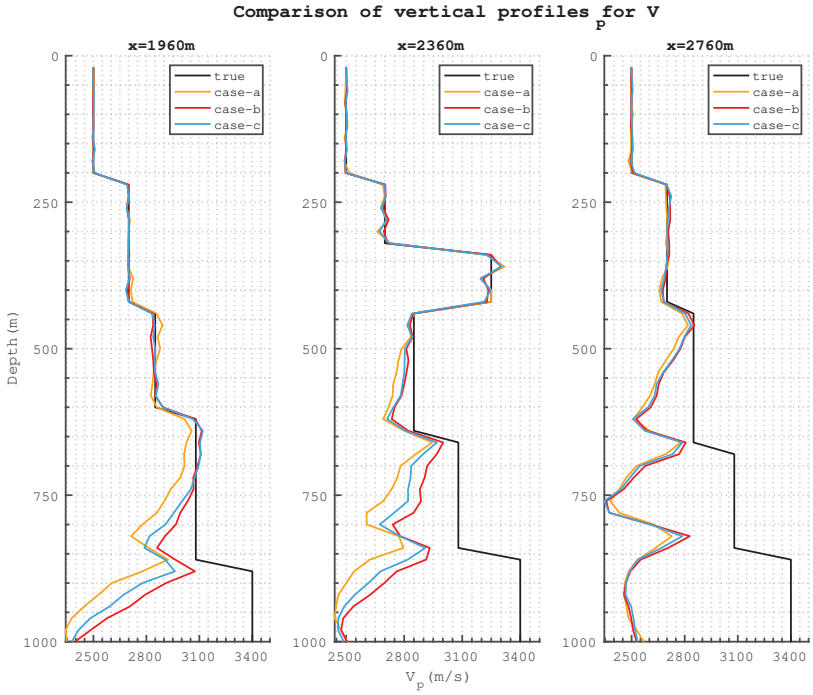

FIG. 13. Vertical profiles of compressional velocity.

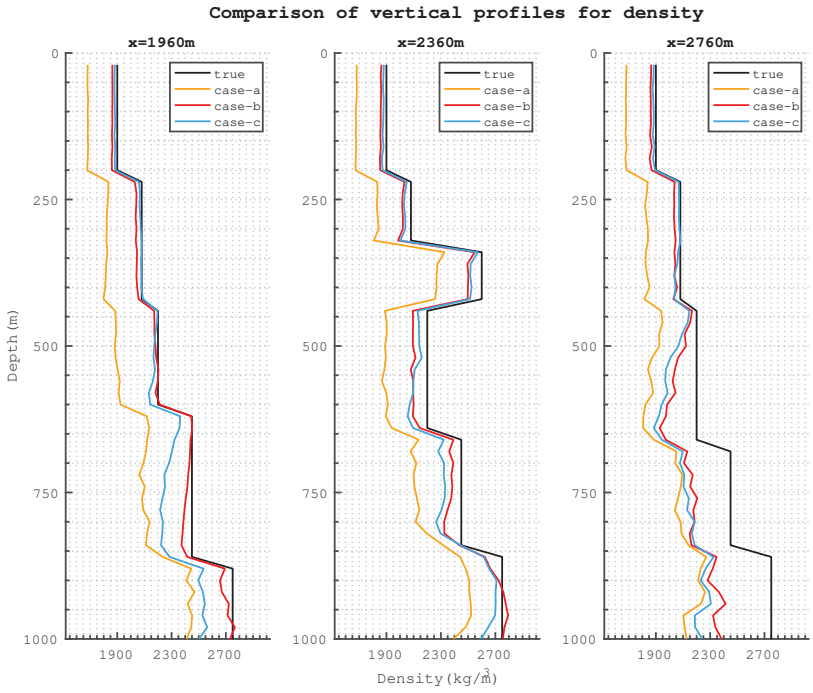

FIG. 14. Vertical profiles of density.

Case 3: local-surface and VSP acquisition

In this case, a vertical seismic profile (VSP) containing 20 receivers is added to the acquisition system of local-surface receivers. The horizontal position of this profile is 2400 m, and the vertical layout depth is from 200 m to 440 m. This acquisition systems are shown in Figure 15.

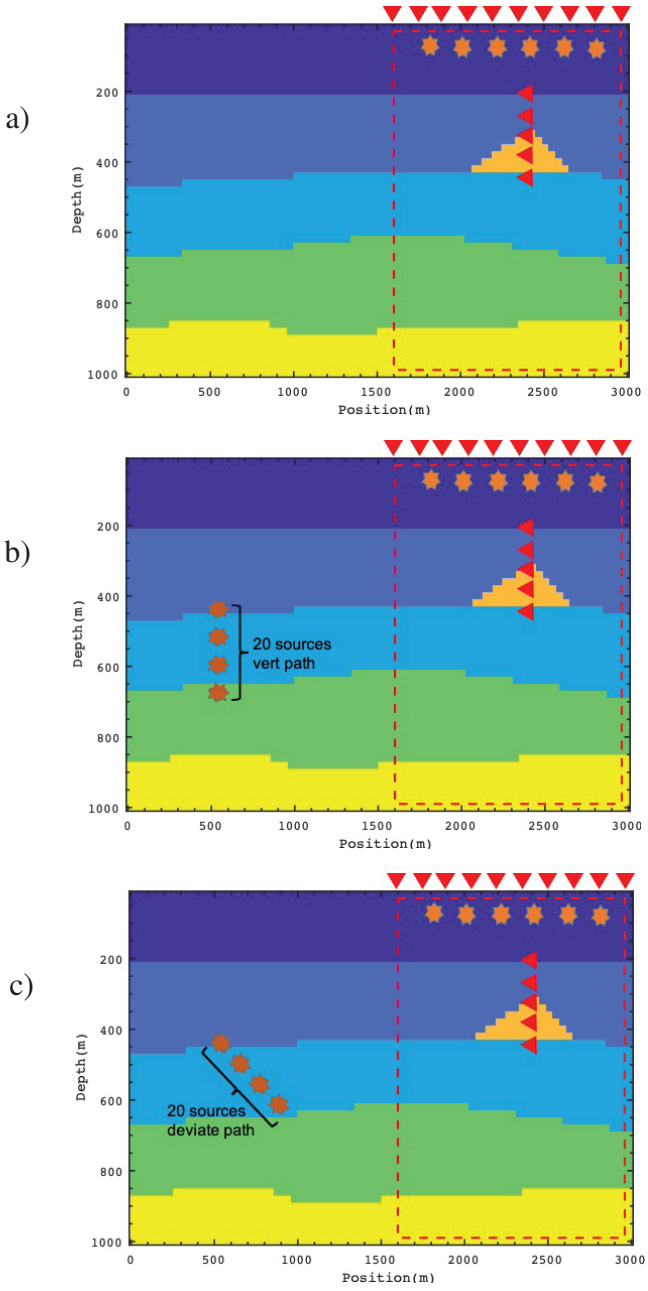

FIG. 15. Acquisition systems in Case 3. a) Local-surface acquisition and VSP, b) A vertical drilling path is added, c) A deviate drilling path is added.

The estimation of compressional velocity and density is shown in Figure 16. When SWD data is adjoined, the inversion gets better. Nonetheless, it is not easy to distinguish the differences between this case and Case 2 since the resolution for both compressional velocity and density is similar. Moment tensors here are also precisely estimated (Figure 17). However, in the comparison part (Figure 18) it can be observed that there is a contribution from SWD to FWI, especially in the density inversion.

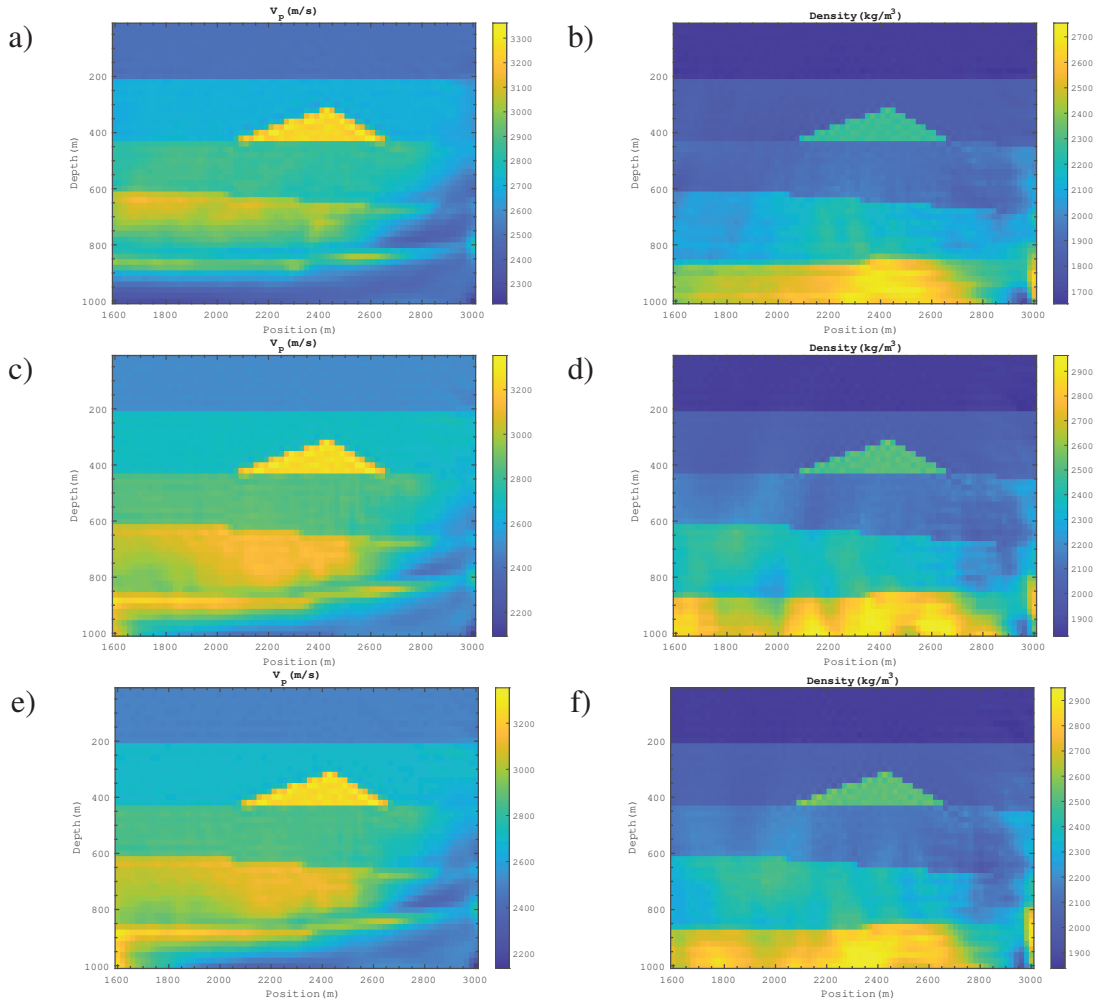

FIG. 16. Inversion result of subsurface properties. a) Compressional velocity, b) Density, c) Compressional velocity (vertical drilling path added), d) Density (vertical drilling path added), e) Compressional velocity (deviate drilling path added), f) Compressional velocity (deviate drilling path added).

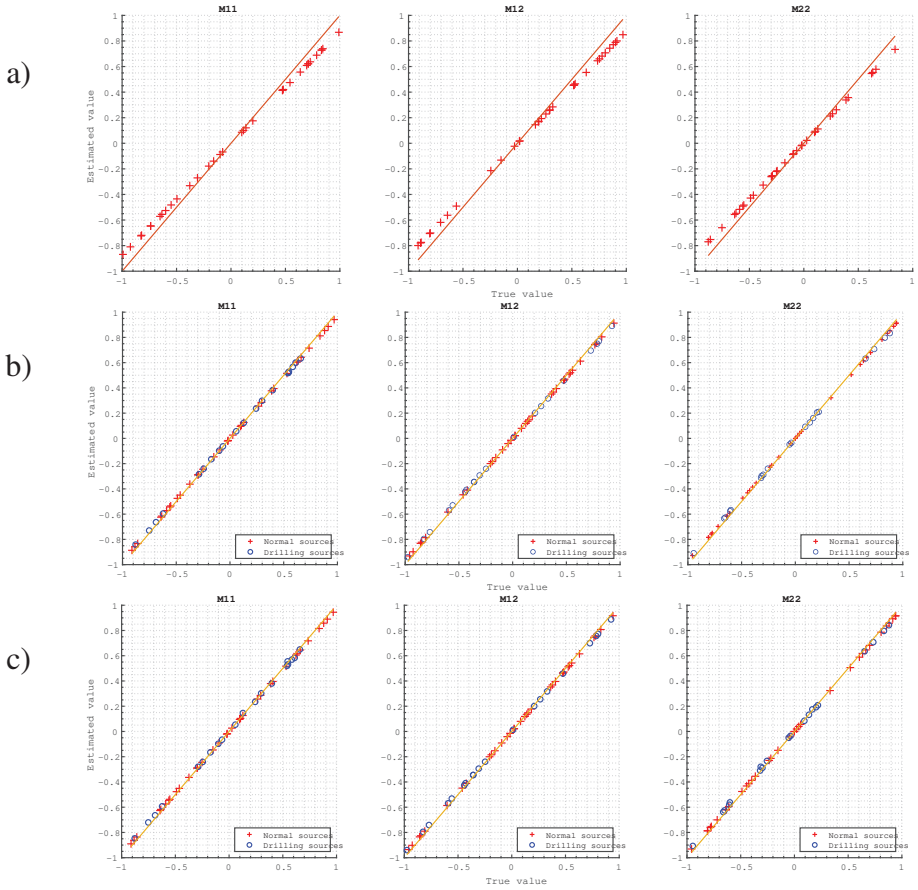

FIG. 17. Cross plots of the true and estimated moment tensors. a) Cross plots of the surface moment tensors, b) Cross plots of the surface and vertical path moment tensors, c) Cross plots of the surface and deviate path moment tensors.

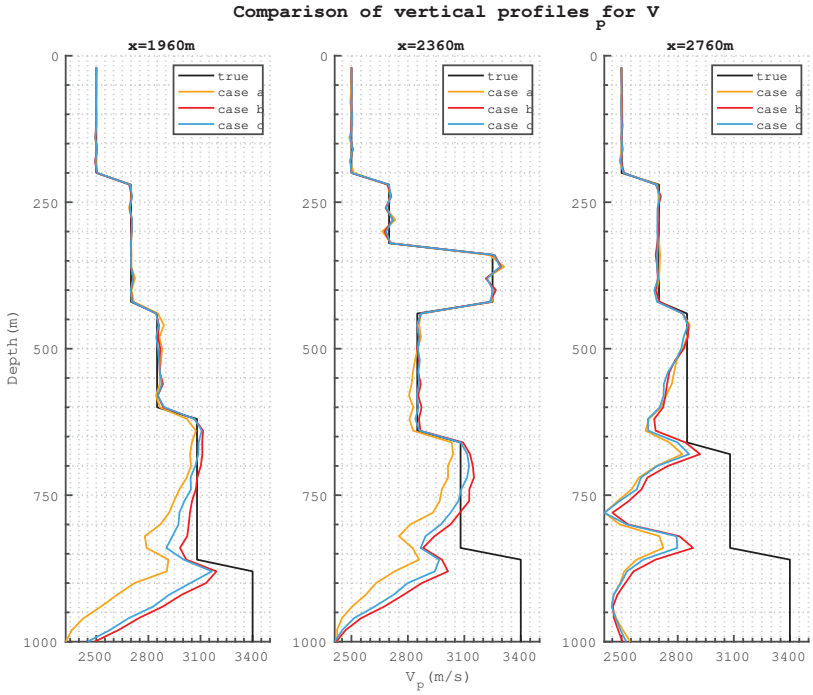

FIG. 18. Vertical profiles of compressional velocity.

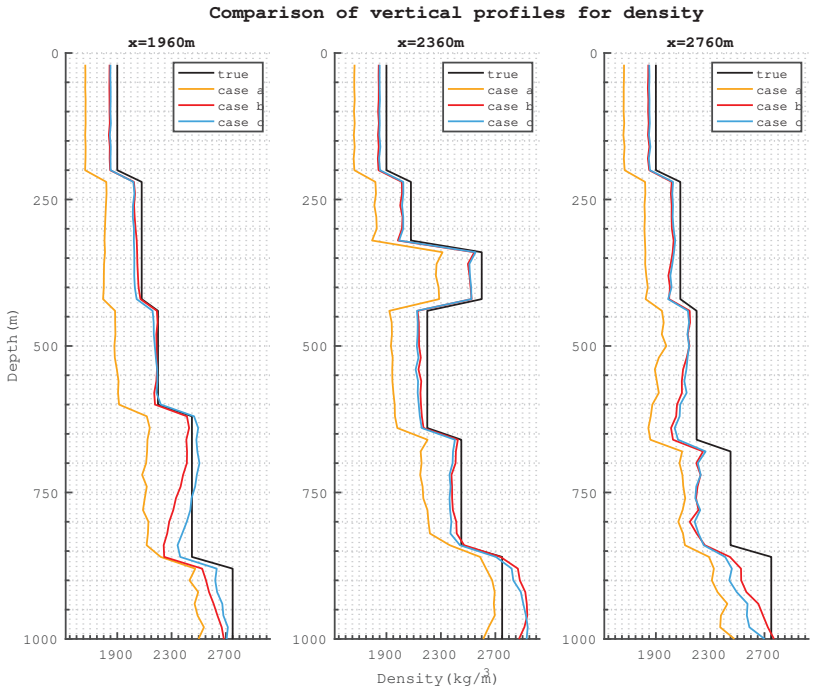

FIG. 19. Vertical profiles of density.

DISCUSSION

Table 1 is the Mean-Absolute-Percentage-Error (MAPE) of elastic properties in each case. In the first case, the most basic one, the error is quite large both for V_p and ρ. However, when the extra sources are added, the error gets an obvious reduction. In case 2 and 3, the error is smaller because we use more comprehensive acquisition system, and it is also decreased when SWD data join in. In the last case the acquisition system is the most comprehensive, and the error is significantly reduced especially when SWD datasets are added. Take the cases when vertical and deviate drilling paths into comparison. It is hard to distinguish the contribution of which dataset outweighs each other. There are two probable reasons. First, even though SWD can provide feasible low-frequency information, which is beneficial to FWI, the radiation patterns are different from those induced by conventional surface seismic. Consequently, it may not get the expected contribution of SWD if we still use viscoelastic wave equations to acquire SWD datasets. Second, attenuation influences the energy spreading although it is set to be high enough to minimize its influence. The evidence can be seen on the vertical profiles. In the locations far from both surface and SWD sources, the inversion results get very limited enhancement.

Table 1. Mean-Absolute-Percentage-Errors of each case.

	Case 1		Case 2		Case 3	
	V_p	ρ	V_p	ρ	V_p	ρ
a	18.9%	24.7%	14.4%	25.9%	11.9%	24.5%
b	14.9%	9.2%	11.4%	7.2%	9.2%	6.7%
c	15.2%	9.8%	12.7%	8.2%	9.4%	6.2%

There is another thought-provoking information we can get from the recovered properties and vertical profiles. If merely looking at the density estimation in Case a), it is natural to conclude that the inversion is successful because of a relatively higher resolution. However, when comparing the vertical profiles with those when SWD is adjoined, the density profiles have an overall bias. This phenomenon is possibly caused by the cross-talk from ρ to V_p because of the parameter weighting in simultaneous FWI. This crosstalk gets more intense when the source number grows. The magnitude of normalized elastic properties, 0 1, is different from moment tensors that are -1 1. It is easy to normalize V_p and ρ by making fractions, but it becomes harder to standardize moment tensors without influencing the gradient and approximated Hessian. Besides, the form of the updating direction of moment tensors is simpler than elastic properties and thus makes a priority when being calculated. All the factors influence the procedure when solving the linear system (Equation 4) via the L-BFGS method. The above speculations are based on several numerical testaments, but the problems are not addressed theoretically yet. Our current treatment is degrading the weight of moment tensors as the source number increases. Specifically, we use 0.25 as a weighting factor when deeper sources are added.

For future research, it is necessary but challenging to realize the radiation patterns of SWD and make good simulations with our current methods. After that, the inverse problem based on probably new forward modeling approaches should be built. Moreover, the weighting problems need to be addressed.

CONCLUSIONS

This paper explores the potentiality of taking advantage of the seismic-while-drilling data to compensate for the incomplete acquisition in simultaneous full-waveform inversion to estimate compressional velocity, density, and moment tensors. According to the numerical examples, a more comprehensive acquisition system helps to get a better FWI result. Besides, more ray paths, which are generated by deeper sources, can make an outstanding contribution to our inversion. That is why we need the SWD technique. After adding SWD data, both the inversion of elastic properties and moment tensors get considerably enhanced. It is reasonable to conclude that in our work, seismic-while-drilling data helps to improve the inversion, though the contribution of vertical and deviate drilling paths are similar. However, further related research is required to provide more comprehensive conclusions.

ACKNOWLEDGEMENTS

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

REFERENCES

Auriol, J., Kazemi, N., Shor, R. J., Innanen, K. A., and Gates, I. D., 2020, A sensing and computational framework for estimating the seismic velocities of rocks interacting with the drill bit: IEEE Transactions on Geoscience and Remote Sensing, **58**, No. 5, 3178–3189.

Fichtner, A., Kennett, B. L. N., Igel, H., and Bunge, H.-P., 2009, Full seismic waveform tomography for upper-mantle structure in the australasian region using adjoint methods: Geophysical Journal International, **179**, No. 3, 1703–1725.

Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms: Numerical results: GEOPHYSICS, **51**, No. 7, 1387–1403, https://doi.org/10.1190/1.1442188.

Kazemi, N., Auriol, J., Innanen, K., Shor, R., and Gates, I., 2020, Successive full-waveform inversion of surface seismic and seismic-while-drilling datasets without low frequencies.

Kazemi, N., and Sacchi, M. D., 2014, Sparse multichannel blind deconvolution: GEOPHYSICS, **79**, No. 5, V143–V152.

Kazemi, N., Shor, R., and Innanen, K., 2018, Illumination compensation with seismic-while-drilling plus surface seismic imaging, **2018**, No. 1, 1–5.

Keating, S., and Innanen, K. A. H., 2020, Simultaneous recovery of source locations, moment tensors and subsurface models in 2D FWI, CREWES Research Report, 32, 33, 14.

Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations: GEOPHYSICS, **49**, No. 5, 533–549, https://doi.org/10.1190/1.1441689.

Métivier, L., Brossier, R., Virieux, J., and Operto, S., 2013, Full waveform inversion and the truncated newton method: SIAM Journal on Scientific Computing, **35**, No. 2, B401–B437, https://doi.org/10.1137/120877854.

Nash, S. G., 2000, A survey of truncated-newton methods: Journal of Computational and Applied Mathematics, **124**, No. 1, 45–59, numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations.

Nejadi, S., Kazemi, N., Curkan, J. A., Auriol, J., Durkin, P. R., Hubbard, S. M., Innanen, K. A., Shor, R. J., and Gates, I. D., 2020, Look ahead of the bit while drilling: Potential impacts and challenges of acoustic seismic while drilling in the mcmurray formation: SPE Journal, **25**, No. 05, 2194–2205.

Nocedal, J., and Wright, S., 2006, Numerical Optimization, Springer Series in Operations Research and Financial Engineering: Springer New York.

Pan, W., Innanen, K. A., Margrave, G. F., and Cao, D., 2015, Efficient pseudo-gauss-newton full-waveform inversion in the tau-p domain: GEOPHYSICS, **80**, No. 5, R225–R14, https://doi.org/10.1190/geo2014-0224.1.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, No. 2, 495–503.
URL https://doi.org/10.1111/j.1365-246X.2006.02978.x

Poletto, F., Miranda, F., Farina, B., and Schleifer, A., 2020, Seismic-while-drilling drill-bit source by ground force: Concept and application: GEOPHYSICS, **85**, No. 3, MR167–MR178, https://doi.org/10.1190/geo2019-0449.1.

Pratt, R. G., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: GEOPHYSICS, **55**, No. 5, 626–632, https://doi.org/10.1190/1.1442874.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: GEOPHYSICS, **49**, No. 8, 1259–1266, https://doi.org/10.1190/1.1441754.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: GEOPHYSICS, **74**, No. 6, WCC1–WCC26, https://doi.org/10.1190/1.3238367.