Trace interpolation and elevation statics by conjugate gradients

Marcus R. Wilson, Robert J. Ferguson

ABSTRACT

We present a conjugate-gradient based inversion to correct for surface statics and irregular trace spacing. The algorithm returns a rough solution to the extrapolated wavefield with complexity $\mathcal{O}(n^{2.5})$. Convergence is fast in the wavelike region, and very slow in the evanescent region. Decimated traces are reconstructed even though no smoothing operator is applied, but recovered wavefields do not approach known source wavefields at low frequencies. We suggest that speed and accuracy of inversion by conjugate gradients can be improved through careful smoothing, or separate treatment of the wavelike and evanescent regions. Computing operators by series expansion for fast application during conjugate gradient iterations is suggested to optimize runtime.

INTRODUCTION

A wave equation inversion for acquired seismic data described in Ferguson (2006) recursively computes the extrapolated wavefield at depth using non-stationary phase shift operators (Ferguson and Margrave, 2002). The operator matrix is computed using an assumed velocity model and the wavefield at depth is derived using weighted damped least squares. This method is used to correct common shot gathers for topography and receiver statics, downward propagate the receiver wavefield through a heterogeneous near surface to a flat datum, and to correct for irregular spatial sampling in one inversion.

Full computation of the extrapolation matrix has complexity $\mathcal{O}(n^2)$, where n is the number of spatial co-ordinates (Ferguson, 2006). Computing the least squares Hessian matrix requires multiplication of the extrapolation matrix by a weight matrix, followed by the adjoint extrapolation matrix, with complexity $\mathcal{O}(n^3)$. Inversion of this matrix by Gaussian elimination also has complexity $\mathcal{O}(n^3)$. Ferguson (2009) develops a series expansion to compute the operator and Hessian simultaneously in $\mathcal{O}(n)$ operations. Smith et al. (2009) inverts the Hessian using conjugate gradients, with complexity $\mathcal{O}(kn^2)$, where k is the number of iterations required for an acceptable approximation.

In this paper we analyse the inversion of the Hessian by conjugate gradients. We assume that the phase shift operator perfectly models wave propagation, and observe the effects of trace decimation, number of iterations, and the accuracy of our velocity model on the inversion.

THEORY

A wave equation inversion for seismic data given by Ferguson (2006) simultaneously corrects for velocity variation in the near surface and irregular trace spacing using non-stationary phase shift operators. First we discuss here the development of these operators, and the application to statics and trace regularization. We will then discuss the conjugate gradient method as a means to speed the algorithm.

Non-stationary Phase Shift Operators

The phase-shift migration method of Gazdag (1978) models the propagation of a monochromatic wavefield through the subsurface as a function of a homogeneous velocity model. It gives a fast and exact solution to the scalar wave equation in homogeneous media (Gazdag and Sguazzero, 1984). To accommodate velocity variation in depth, the algorithm is run recursively on a sequence of constant velocity depth steps. That is, for each frequency ω, and each depth z, the extrapolated wavefield $\varphi_{z+\Delta z}$ is computed from φ_z by

$$\begin{aligned} \varphi_{z+\Delta z}(x) &= PSPI_{\Delta z}(\varphi_z)(x) \\ &= \frac{1}{2\pi}\int \alpha_{\Delta z}(k_x, v_z)\int \varphi_z(x')exp(ik_x\cdot(x'-x))dx'dk_x, \end{aligned} \tag{1}$$

where $\alpha_{\Delta z}$ is a function of spatial wavenumber k_x and layer velocity v_z given by

$$\alpha_{\Delta z}(k_x, v_z) = \begin{cases} \exp\left(i\Delta z\sqrt{(\omega/v_z)^2 - k_x\cdot k_x}\right) & \text{if } |\frac{\omega}{v_z}| \geq |k_x| \\ \exp\left(-|\Delta z|\sqrt{k_x\cdot k_x - (\omega/v_z)^2}\right) & \text{if } |\frac{\omega}{v_z}| < |k_x|. \end{cases} \tag{2}$$

Equation 1 is a Fourier transform of the source wavefield, followed by a multiplication, then an inverse Fourier transform. Equation 2 applies the phase shift operator in the wavelike region, where $|\frac{\omega}{v_z}| \leq |k_x|$, and attenuates energy in the evanescent region, where $|\frac{\omega}{v_z}| > |k_x|$. In practice we will represent these wavefields as vectors in $\mathbb{C}^n$, so let us consider these operators as matrices. In this case PSPI becomes

$$PSPI_{\Delta z}(\varphi_z) = [IFT][\alpha_{\Delta z}][FT]\varphi_z. \tag{3}$$

The matrix form of $\alpha_{\Delta z}$ is diagonal in Fourier co-ordinates, so using the fast Fourier transform, the cost of applying equation 3 is $\mathcal{O}(n\log n)$, where n is the length of the vector φ_z. Ferguson and Margrave (2002) accommodates lateral velocity variation using a set of constant velocity windows. The window function is defined for a given reference velocity v_j by

$$\Omega_j(x) = \begin{cases} 1 & \text{if } v(x) = v_j \\ 0 & \text{if } v(x) \neq v_j \end{cases} \tag{4}$$

and PSPI becomes

$$PSPI_{\Delta z}(\varphi_z) = \sum_j [\Omega]_j[IFT][\alpha_{\Delta z}]_j[FT]\varphi_z. \tag{5}$$

Ferguson and Margrave (2002) further describes two similar phase shifting operators, NSPS and SNPS, given by equations 6 and 7. It is trivial to show that SNPS is symmetric in the wavelike region, and Hermitian in the evanescent region, which makes it preferable as a wave extrapolation operator due to reciprocity conditions (Ferguson, 2006). For brevity,

we will denote by $P_{\Delta z}$ any of these one-way operators that shifts a wavefield downward by Δz.

$$NSPS_{\Delta z}(\varphi_z) = [IFT]\sum_j [\alpha]_j [FT][\Omega]_j \varphi_z \quad (6)$$

$$SNPS_{\Delta z}(\varphi_z) = PSPI_{\Delta z/2}(NSPS_{\Delta z/2}(\varphi_z)) \quad (7)$$

Statics and Trace Regularization

Ferguson (2006) presents an application of these phase-shift operators to correct for surface statics and irregular trace spacing. Acquired seismic data is modelled recursively as follows: given a recorded wavefield φ_z at depth z, we assume that $\varphi_z = W_e P_{-\Delta z}\varphi_{z+\Delta z} + \epsilon$, where $P_{-\Delta z}$ is an upward phase shift, as in equation 5, 6 or 7, W_e is a weighting operator that models irregular trace spacing and topography, as in Reshef (1991), and ϵ is an additive noise term. This is a mixed-determined linear system (Menke, 1989), so the least-squares approximation of $\varphi_{z+\Delta z}$ can be recovered by minimizing the misfit function

$$M(\varphi) = \|P_{-\Delta z}\varphi - \varphi_z\|^2_{W_e} + \varepsilon\|\varphi - \varphi_m\|^2_{W_m}. \quad (8)$$

Here W_m is a smoothing operator, φ_m is the a priori information on the model parameters (Tarantola, 2005), and ε is a user parameter that controls the amount of smoothing (Menke, 1989). The norms here are induced by the scalar products with respect to W_e and W_m respectively, as defined in Tarantola (2005)*. Since M is minimized when the normal equations are satisfied, we can recover $\varphi_{z+\Delta z}$ by solving

$$[S_{-\Delta z}]\varphi_{z+\Delta z} = \left[P^*_{-\Delta z} W_e P_{-\Delta z} + \varepsilon W_m\right]\varphi_{z+\Delta z} = P^*_{-\Delta z} W_e \varphi_z + \varepsilon W_m \varphi_m, \quad (9)$$

where $P^*_{-\Delta z}$ is the adjoint of $P_{-\Delta z}$. If we consider our operators as matrices here, then the cost of recovering $\varphi_{z+\Delta z}$ is dominated by the cost of computing then inverting the matrix $S_{-\Delta z}$. Ferguson (2006) derives a series approximation of $S_{-\Delta z}$ to speed up computation of the matrix, and we consider using conjugate gradients to speed inversion.

Conjugate Gradients

The conjugate gradient method is an iterative algorithm used to approximate a solution x to a linear system $Ax = b$. In our case it can be used to recover the source wavefield $\varphi_{z+\Delta z}$ from equation 9. Inverting an $n \times n$ matrix by Gaussian elimination has complexity $\mathcal{O}(n^3)$ (Strassen, 1969), whereas solving the system by conjugate gradients can return an acceptable approximation in about $\sqrt{n}$ iterations, provided the matrix is well-conditioned (Burden and Faires, 2001). A matrix is well-conditioned if it is not sensitive to rounding errors, which are likely to occur as our computations will be performed by a computer using floating-point arithmetic.

*Since W_e will be positive semidefinite, a seminorm is induced

Computationally, an iteration of the conjugate gradient method is dominated by the cost of computing the residual vector $r = Ax - b$, which measures the error between our guess x and a solution to the system. If A is a matrix, this operation is $\mathcal{O}(n^2)$. So if the algorithm converges quickly, we can expect a total complexity of $\mathcal{O}(n^{2.5})$.

METHOD

Synthetic data is modelled to satisfy the forward operator exactly, up to a random additive noise term ϵ, set to 40db below the signal level. An arbitrary source wavefield of $n = 256$ traces with n temporal samples (Figure 1a) is Fourier transformed in time, and synthetic data is generated from the resultant monochromatic wavefields according to $\varphi_0 = W_e P_{-100}\varphi_{100} + \epsilon$, where P_{-100} is NSPS (Margrave and Ferguson, 1999) computed with respect to the 'true' reference velocity model. We use a large depth step ($\Delta z = 100m$) to exaggerate the visual impact of the phase shift (Figure 2b), as we will be restricting our attention to a single depth step. The effects of smoothing are not considered here, so ε in equation 9 is set to 0. We apply Matlab's pcg algorithm to equation 9, where P_{100} is computed with respect to the reference velocity model. No preconditioning is applied, and we use the zero vector as an initial guess. We run the pcg algorithm until the prescribed maximum number of iterations is reached, or the relative residual error falls below a tolerance of 10^{-6}.

We consider three cases. First a base case, with no trace decimation and an exact velocity model. Then we allow for trace decimation, then add uncertainty to our velocity model. For each case we display the results of the conjugate gradient algorithm after n and $\sqrt{n}$ iterations. For each of these six examples, the original wavefield (a) is displayed along with the forward modelled data (b), the recovered wavefield (c) and the absolute error (d). We also compute the norms associated with the misfit function for the original, and recovered wavefields, with the norms of the modelled data included for contrast. We refer to the value of the first norm in equation 8 as the model misfit of the wavefield (e), which indicates how well the recovered wavefield agrees with the given data, and the value of the second norm as the smoothness of the wavefield (f), which indicates how close the recovered wavefield is to any a priori information we might have about the source wavefield. In our example, smoothness is calculated with respect to the true source wavefield, with smoothing operator $W_m = I$.

EXAMPLES

As a baseline, we set $W_e = I$, the $n \times n$ identity matrix, to model perfectly uniform trace spacing, and assume the true velocity model is known to be the simple velocity model given by Figure 1b. We allow the pcg algorithm to run for a full n iterations (Figure 2) and for $\sqrt{n}$ iterations (Figure 3).

When the algorithm is run for n iterations, all monochromatic wavefields converge to within the tolerance, except in the 5 - 15Hz range. Rate of convergence is fast in the high frequencies - less than 6 iterations - but decreases rapidly starting at 90Hz. We note that given our input parameters: 10m trace spacing, and reference velocities between 800 and 1800m/s, 90Hz corresponds to the uppermost edge of the evanescent boundary. We can

therefore attribute this slow convergence to the evanescent region, where the phase shift operator is a real exponential with large negative exponent.

Running the algorithm for $\sqrt{n}$ iterations, convergence is observed only in the high frequencies, above 85Hz. The effect of the phase shift is still effectively inverted, the model misfit of the recovered wavefield agrees strongly with that of the source (Figure 3e), and the image is focused (Figure 3c), but the solution does not agree with the expected solution, as the value of the smoothness norm is much larger than 0 in the low frequencies (Figure 3f). Note also that increasing the number of iterations does not improve the smoothness, as the smoothness norm of the recovered model in Figure 3f is smaller for each frequency than the corresponding value in Figure 2f, where more iterations were applied. This is unexpected, as when $W_e = I$, this problem should not be underdetermined, so no smoothing should be necessary. However, the wavefield extrapolator $\alpha_{\Delta z}$, has very small eigenvalues in the evanescent region, which may cause $P_{-\Delta z}$ to be almost singular.

Next, we set a random selection of approximately 30% of the diagonal elements of W_e to zero, and re-compute the synthetic data φ_0. We then invert as above, again using the velocity model in Figure 1b as both the true and reference velocity models, for n and $\sqrt{n}$ iterations. The results are displayed in Figures 4 and 5.

Similar to the previous case, if the algorithm is run for the full n iterations, we achieve convergence almost everywhere, with the number of iterations required increasing as frequency decreases to 0. Again, increasing the number of iterations does not seem to improve the quality of the solution: the model misfit norms of the recovered wavefields (Figures 4e and 5e) agree with those of the source wavefield, but the smoothness norms (Figures 4f and 5f) are much greater than 0 in the low frequencies. Also note that, even without a damping term, the conjugate gradient algorithm manages to fill in some of the information from the missing traces. This smoothing is less effective for large continuous gaps in trace coverage. However, it is surprising that the algorithm does any smoothing at all, since the search directions prescribed by the algorithm are computed from the residual vectors, which should always be zero in the co-ordinates corresponding to the dead traces.

Finally, we assume incomplete knowledge of the near surface velocity variation. To model this, we compute the data with respect to the more complex velocity model (Figure 1c), while computing the inverse operator with respect to the simpler model (Figure 1b). We then invert as above, and the results are displayed in Figures 6 and 7.

We observe that the rate of convergence is similar to the previous cases, however modelling errors are apparent in the model misfit plots (Figures 6e and 7e), as the curve corresponding to the source misfit is nonzero, indicating that the source wavefield is not an approximate solution to equation 9. That is, the computed operator does not map the known source wavefield to the data we are given, since the data was computed with respect to the 'true' velocity model (Figure 1c), and the operator with respect to the reference model (Figure 1b). These errors do not affect the rate of convergence, but they do affect the accuracy of the final recovered image. Once again, increasing the number of iterations above $\sqrt{n}$ has little effect on the end result.

CONCLUSION

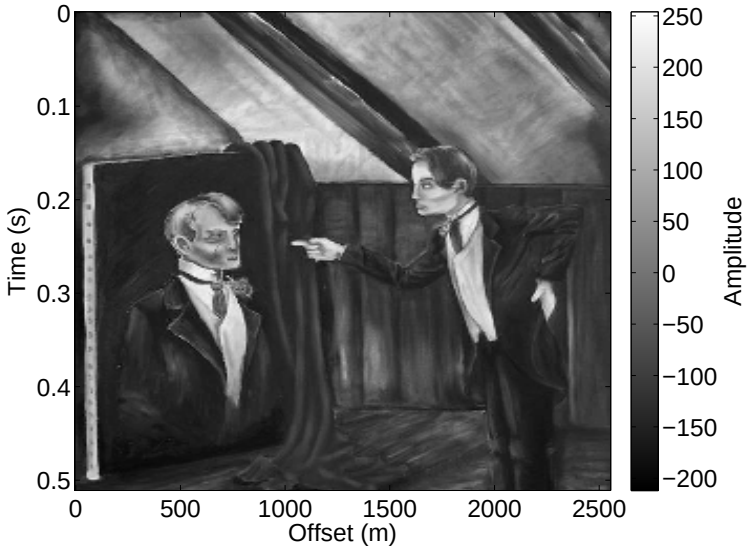
Using NSPS as our model of wavefield propagation, we find that the conjugate gradient algorithm applied to the least squares minimization problem gives a rough solution to the extrapolated wavefield in $\sqrt{n}$ iterations, and no significant improvement is gained from subsequent iterations. We note that convergence is fast in the wavelike region, and slow in the evanescent region, and postulate that the slow convergence is caused by very small operator eigenvalues from the evanescent part of the wave extrapolator $\alpha_{\Delta z}$, causing the Hessian to be almost singular. Solution damping could be achieved through the use of a nontrivial smoothing operator W_m, as in Smith et al. (2009) and Ferguson (2006), or we might attempt to treat the wavelike and evanescent regions separately.

We find that the $\sqrt{n}$ speed up in inversion of the Hessian leaves this step as the largest bottleneck in the method - much slower than $\mathcal{O}(n)$ for computation of the Hessian - and suggest that we might be able to speed up computation of the residual vectors if we can apply the Hessian as an operator in the main loop of the conjugate gradient algorithm, provided the cost of applying the operator is lower than that of computing a matrix vector product ($\mathcal{O}(n^2)$).

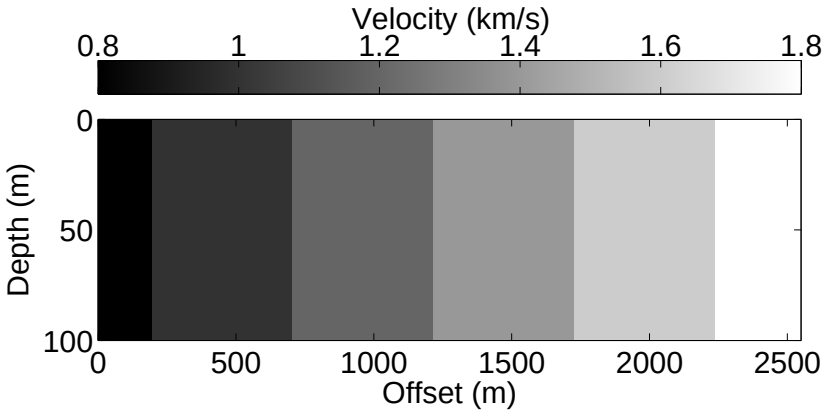
ACKNOWLEDGEMENTS

The authors wish to thank the sponsors, faculty, and staff of the Consortium for Research in Elastic Wave Exploration Seismology (CREWES), and the Natural Sciences and Engineering Research Council of Canada (NSERC, CRDPJ 379744-08) for their support of this work.

a)



b)



c)

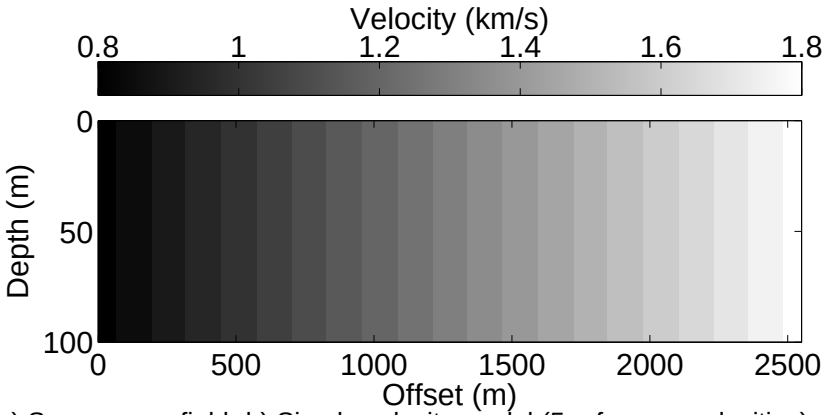


FIG. 1. a) Source wavefield. b) Simple velocity model (5 reference velocities). c) Complex velocity model (20 reference velocities)

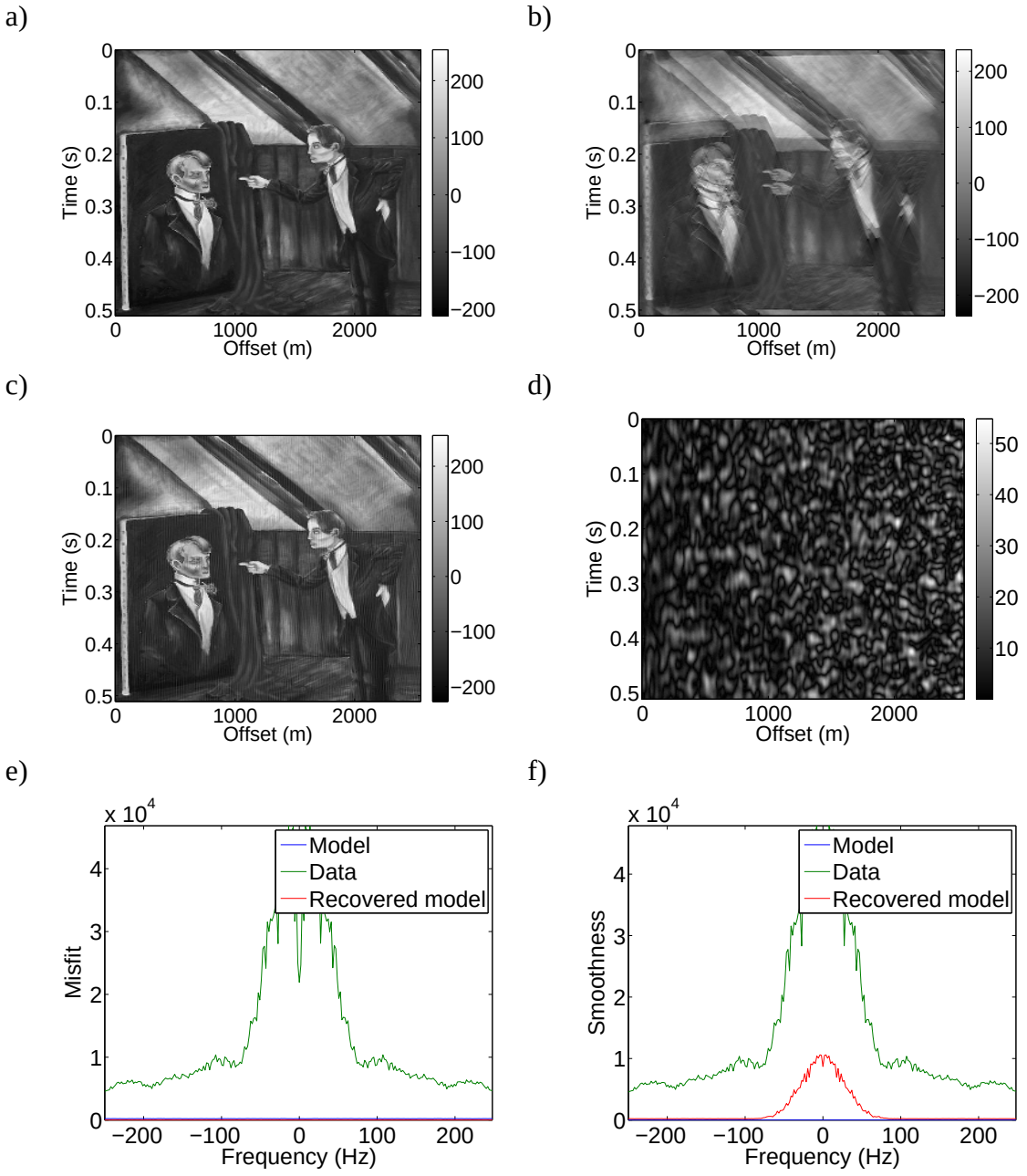


FIG. 2. Case 1: Regular trace spacing, known velocity model, n iterations. a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

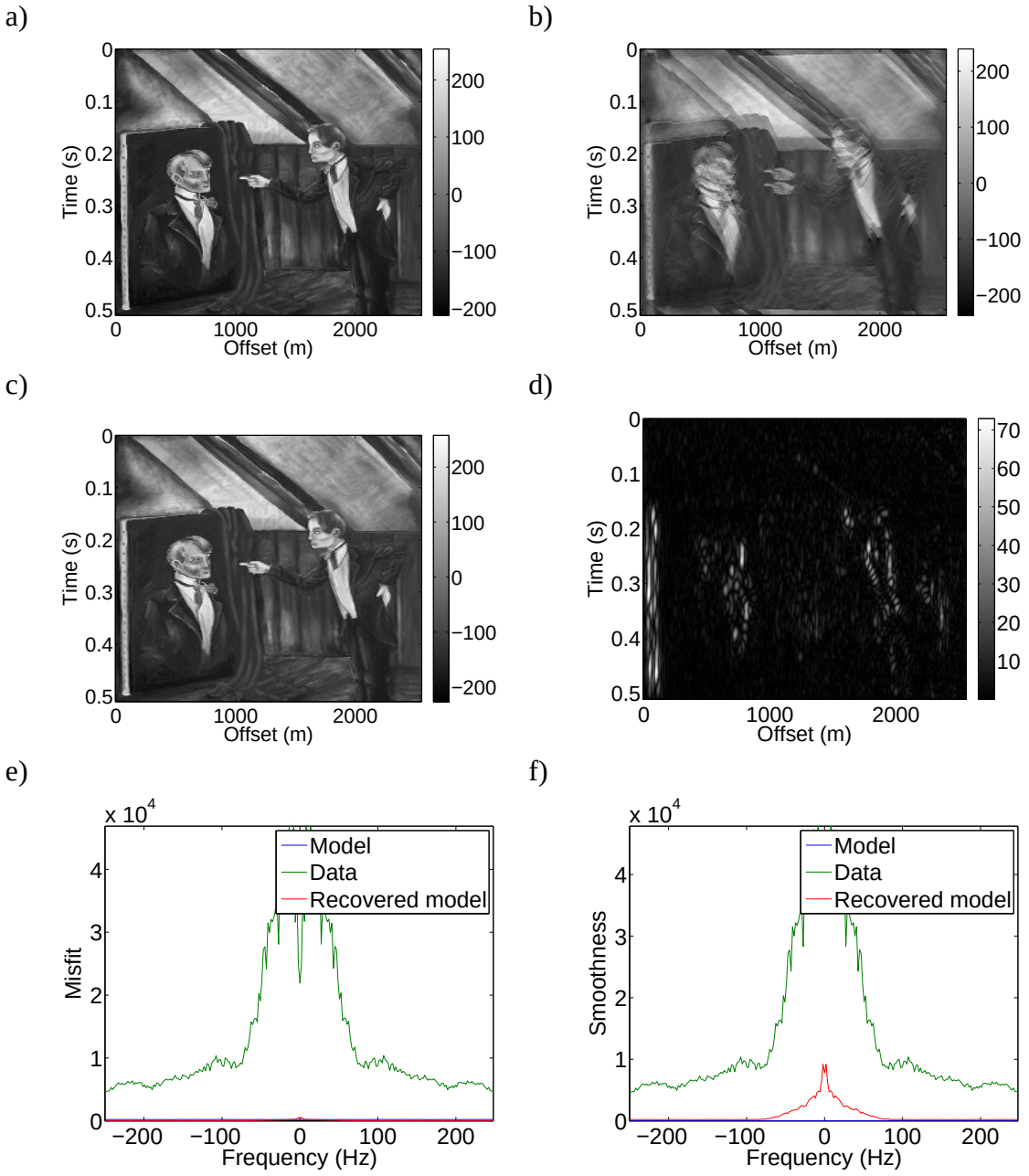


FIG. 3. Case 2: Regular trace spacing, known velocity model, $\sqrt{n}$ iterations. a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

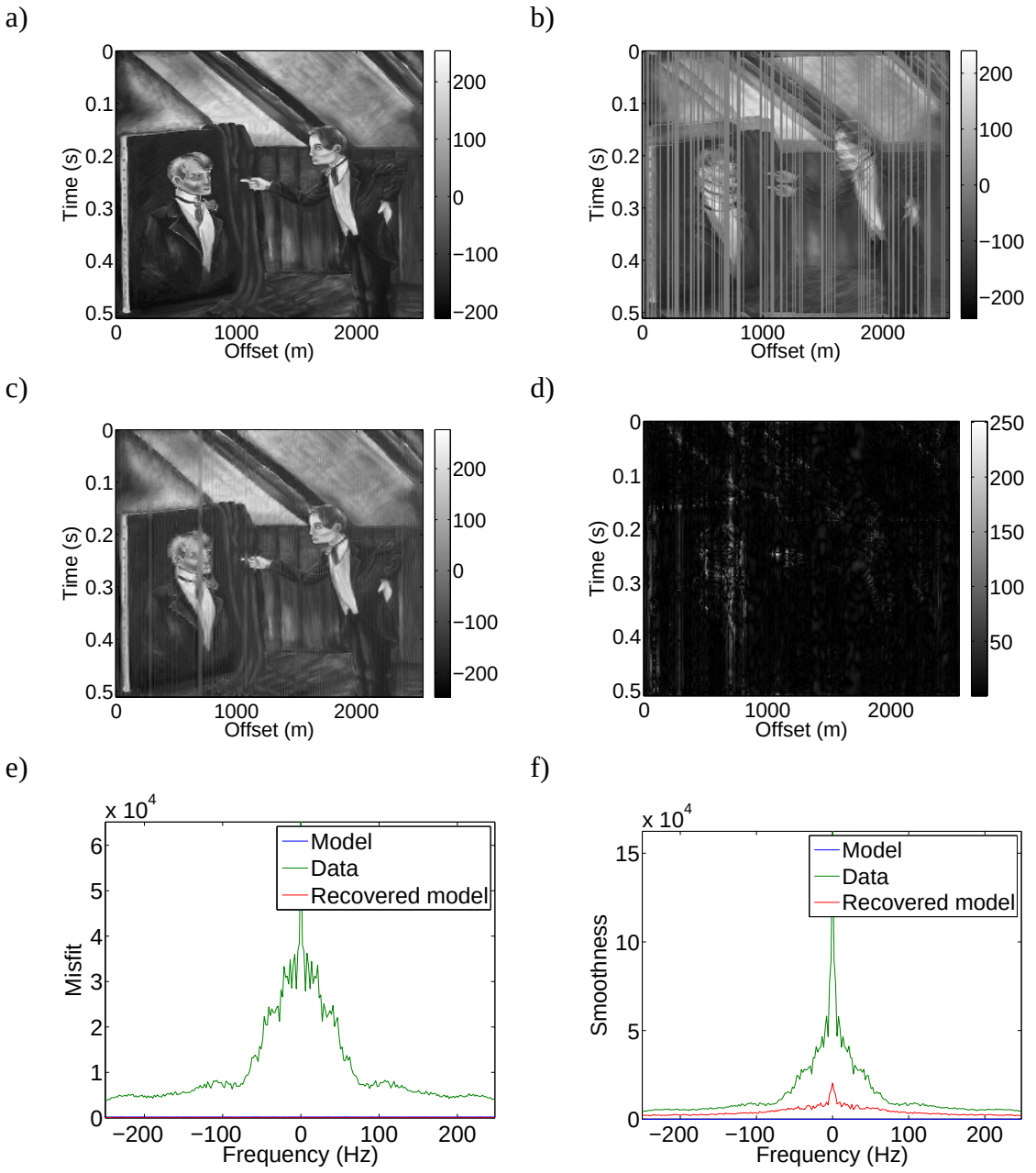


FIG. 4. Case 3: 30% trace decimation, known velocity model, n iterations a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

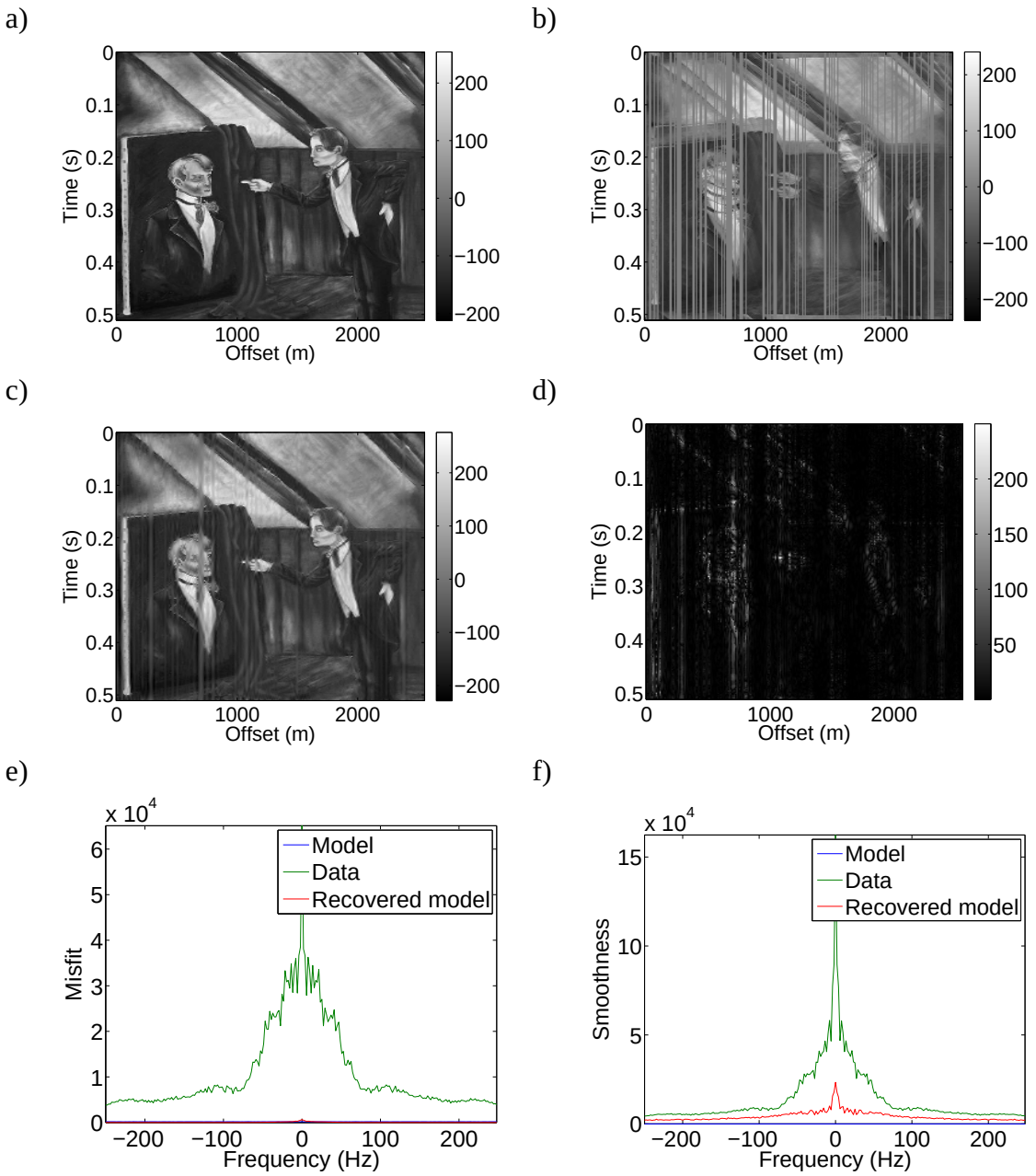


FIG. 5. Case 4: 30% trace decimation, known velocity model, $\sqrt{n}$ iterations a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

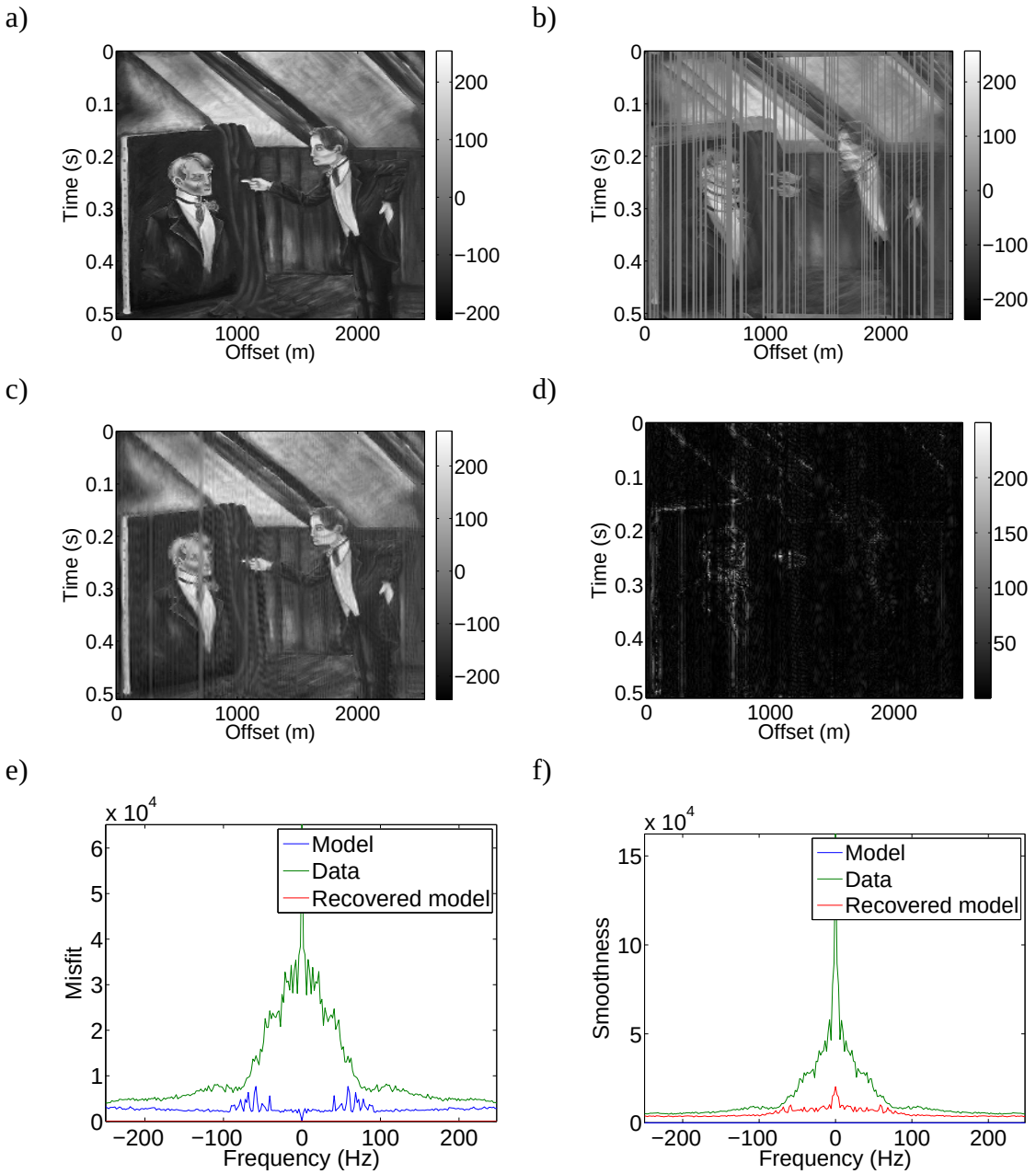


FIG. 6. Case 5: 30% trace decimation, uncertain velocity model, n iterations a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

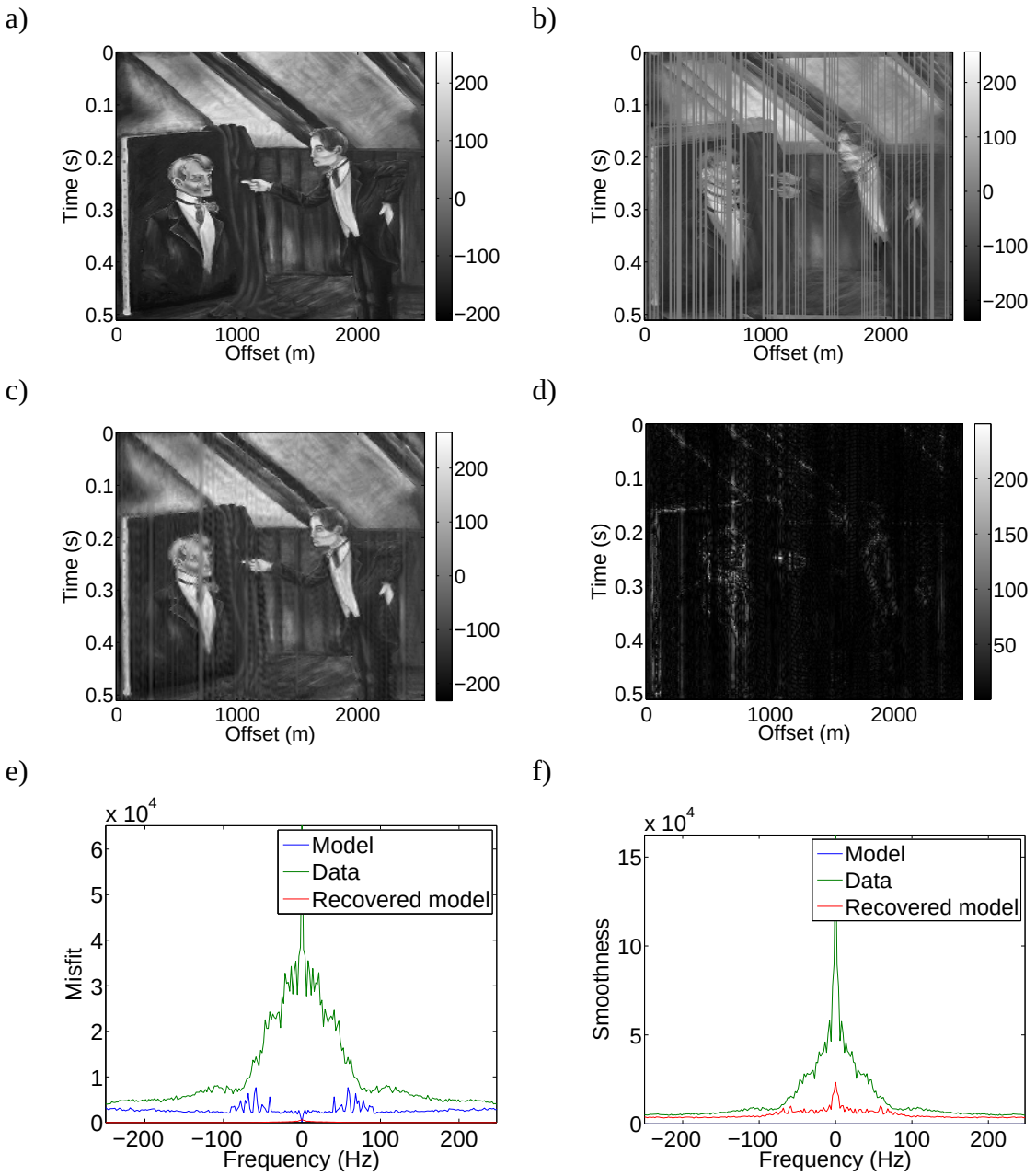


FIG. 7. Case 6: 30% trace decimation, uncertain velocity model, $\sqrt{n}$ iterations a) Source wavefield. b) Forward modeled image. c) Recovered image. d) Absolute error. e) Model misfit. f) Smoothness.

REFERENCES

Burden, R., and Faires, J. D., 2001, Numerical Analysis: Brooks/Cole, 511 Forest Lodge Road, Pacific Grove, CA 93950 USA.

Ferguson, R. J., 2006, Regularization and datuming of seismic data by weighted, damped least squares: Geophysics, **71**, No. 5, U67–U76.

Ferguson, R. J., 2009, Efficient simultaneous wave equation statics and trace regularization by series approximation: Geophysics, Submitted.

Ferguson, R. J., and Margrave, G. F., 2002, Prestack depth migration by symmetric nonstationary phase shift: Geophysics, **67**, No. 2, 594–603.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, No. 7, 1342–1351.

Gazdag, J., and Sguazzero, P., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, No. 2, 124–131.

Margrave, G. F., and Ferguson, R. J., 1999, Wavefield extrapolation by nonstationary phase shift: Geophysics, **64**, No. 4, 1067–1078.

Menke, W., 1989, Geophysical Data Analysis: Academic Press, 525B Street, Suite 1900, San Diego, California 92101-4495.

Reshef, M., 1991, Depth migration from irregular surfaces with depth extrapolation methods: Geophysics, **56**, No. 1, 119–122.

Smith, D. R., Sen, M. K., and Ferguson, R. J., 2009, Optimized regularization and redatum by conjugate gradients: 71st EAGE Conference & Exhibition, **71**.

Strassen, V., 1969, Gaussian elimination is not optimal.: Numerische Mathematik, **14**, No. 3, 354–356.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104-2688.