A two-stage Markov chain Monte Carlo method for velocity estimation and uncertainty quantification

Georgia K. Stuart, Weihua Yang, Susan E. Minkoff, and Felipe Pereira, The University of Texas at Dallas*

SUMMARY

Bayesian seismic inversion can be used to sample from the posterior distribution of the velocity field, thus allowing for uncertainty quantification. However, Bayesian techniques like Markov chain Monte Carlo (McMC) can be extremely computationally expensive. We propose a two-stage McMC method where an upscaled wave equation solver is used to quickly filter out unacceptable velocity fields, thus reducing the computational expense. We have found that the data generated by the full fine-grid solver is correlated with the data generated by the upscaled solver, so the upscaled solver is a valid filter.

INTRODUCTION

Deterministic seismic inversion uses global or local optimization to minimize the misfit between data predicted by a model (e.g., the wave equation) and data measured at receivers. It produces a single estimate of a subset of the Earth's mechanical parameters (e.g., the velocity field) without including information about uncertainty. In contrast, Bayesian inversion incorporates receiver data in a framework to quantify uncertainty that also takes into account prior independent information about the parameter to be estimated such as well log data. This more costly framework produces a probability distribution for the estimated parameter which naturally provides information about the reliability of the solution (Mosegaard and Tarantola, 1995).

Gouveia and Scales (1998) assume all uncertainties in model or data can be described by Gaussian probability distributions, and the resulting a posteriori probability is optimized to find a maximal model for velocity via gradient-based methods. Xuan and Sava (2010) use Bayesian inversion to estimate microseismic events. They construct maps of probability indicating the likelihood that a microseismic source exists at a given position, time and with a given orientation. Gesret et al. (2015) again use Bayesian inversion to estimate earthquake location, but they include velocity uncertainty in their procedure. The Bayesian approach has some clear advantages over deterministic inversion since uncertainty is quantified in the estimation process. However, issues of accuracy and efficiency remain which has limited its use in seismic studies. One major stumbling block is the fact that McMC requires solution of the wave equation hundreds or thousands of times and is thus extremely computationally intensive (Tarantola, 2005). Furthermore, it is estimated that 90% of these velocity proposals are rejected in the McMC process (Akbarabadi et al., 2015). Hong and Sen (2009) use a hybrid genetic algorithm to explore the model space for optimal parameters. They run multiple McMC chains at different scales simultaneously in parallel and combine these chains to explore the model space.

In this work we outline a method for estimating acoustic wave velocity via a Bayesian procedure. However, we dramatically reduce the computational expense of the process by using a two-stage McMC algorithm. To date, multi-level McMC has been applied primarily in the context of reservoir simulation (Ginting et al., 2011; Akbarabadi et al., 2015; Efendiev et al., 2006). Our two-stage McMC algorithm samples from the posterior distribution of the velocity field by first solving a less computationally expensive surrogate problem on a coarse grid. The algorithm accepts or rejects the velocity perturbation based on the Metropolis-Hastings criterion. This surrogate solve, therefore, acts as a filter to remove unacceptable velocity perturbations from consideration. If accepted, we then solve the full fine grid problem to finalize whether the velocity perturbation should be accepted or rejected in the posterior distribution.

The multi-level algorithm we describe in this paper reduces the computational cost of McMC in two ways. First, a truncated Karhunen Loève Expansion is used to generate the velocity perturbations (Huang et al., 2001), and second we use a surrogate coarse-scale problem as a proposal filter. For the surrogate problem, we use operator upscaling (Vdovina et al., 2005) applied to the acoustic wave equation. It is important to highlight that the coarse scale/surrogate models in a two-stage McMC method should be computationally inexpensive but not necessarily very accurate. However, there must be a strong correlation between the results on the coarse and fine scales (Efendiev et al., 2006). Thus, here we investigate this correlation for the coarse scale (upscaling) procedure of Vdovina et al. (2005). We show that the fine and coarse grid simulations exhibit a strong correlation in a numerical study that takes into account a layered structure for the subsurface consistent with measured well log data.

THEORY AND METHODS

Bayesian Framework

A Bayesian approach to seismic inversion can incorporate independent prior information into the model estimate, thereby producing an updated posterior distribution of the velocity field perturbation, C, given observed receiver data, d_o. The goal is to construct a group of selected velocity samples from the posterior distribution for velocity. We model the velocity field as follows:

$$c(\vec{x}) = M(z) + C(\vec{x}) \tag{1}$$

where $c(\vec{x})$ is the modeled velocity field at location $\vec{x} = (x,z)$, $M(z)$ is the deterministic background velocity which for this paper is assumed to depend only on depth, z, and $C(\vec{x})$ is a stochastic perturbation of the model.

Bayes' Rule states that the posterior probability of $C(\vec{x})$ is given by

$$P(C|d_o) \propto P(d_o|C)P(C) \tag{2}$$

where $P(C)$ is the prior probability of the velocity field perturbation, C, and $P(d_o|C)$ is the likelihood of the observed data given a proposed velocity field perturbation. We assume that the likelihood function follows a Gaussian distribution, as in Tarantola (2005) , i.e.

$$P(d_o|C) \propto \exp\left(-\frac{||d_o - d_C||^2}{2\sigma^2||d_C||^2}\right) \tag{3}$$

where d_C is the simulated data for a velocity, c, with given perturbation, C, and σ is the precision parameter.

The *McMC* Process and the Metropolis-Hastings Criterion

The Metropolis-Hastings Markov chain Monte Carlo (McMC) method is used to sample from the posterior distribution $P(C|d_o)$. The standard single-stage McMC method tests a proposed sample velocity via fine-grid simulation (Hastings, 1970; Gamerman and Lopes, 2006). Generally, the McMC algorithm examines thousands of proposed samples to determine the distribution. This process is computationally very expensive as most of the samples are rejected (Akbarabadi et al., 2015). A two-stage McMC method substantially speeds up this process of sample selection.

Two-Stage *McMC* Process

The two-stage McMC process uses a surrogate (or coarse scale) model to test a sample perturbation before a fine-grid simulation is performed. In this work we use an upscaled form of the acoustic wave equation as the surrogate model. If the generated proposal is rejected in the coarse-grid simulation, then the algorithm proceeds to generate a new proposal. If accepted, then the velocity sample is used in a fine grid simulation. In this way, we can significantly reduce the run time of the McMC process. The **Two-Stage McMC Algorithm**, modeled after Efendiev et al. (2006) and Christen and Fox (2005), can be described as follows:

1. Generate a proposed velocity field perturbation C from a previous perturbation C_n and a random walk sampler $q(C,C_n)$ (Cotter et al., 2013).
2. Solve the wave equation using acoustic upscaling.
3. Accept C with probability:
$$\rho_U(C_n,C) = \min\left\{1, \frac{q(C_n|C)P_U(C|d_o)}{q(C|C_n)P_U(C_n|d_o)}\right\}. \tag{4}$$
4. If C is accepted, solve the forward problem on the full fine grid.
5. Accept C with probability:
$$\rho_f(C_n,C) = \min\left\{1, \frac{P_f(C|d_o)P_U(C_n|d_o)}{P_f(C_n|d_o)P_U(C|d_o)}\right\}. \tag{5}$$
6. Repeat until a given number of accepted proposals is reached.

In Equations (4) and (5), P_f and P_U refer to posterior distributions with likelihood functions computed from fine and upscaled simulations, respectively.

The Karhunen-Loève Expansion for Generating Stochastic Perturbations

In the numerical experiments discussed here, the computational domain is quite small (1280×1280 grid points). Nonetheless, to generate stochastic perturbations of the velocity field needed for the McMC process we would need to consider more than 10^6 random variables, which is computationally prohibitive. The Karhunen-Loève Expansion (KLE) provides a tool for generating Gaussian perturbations of the velocity by representing the stochastic process as an infinite series (Loève, 1977; Wong, 1971). The velocity perturbation, $C(\vec{x})$, is generated by a truncated KLE as follows:

$$C(\vec{x}) = \sum_{j=1}^{N} \sqrt{\lambda_j}\xi_j\theta_j(x) \tag{6}$$

where $\{\lambda_j\}_{j=1}^N$ and $\{\xi_j\}_{j=1}^N$ are the eigenvalues and eigenfunctions of a covariance function $R(x_1,x_2)$, and $\{\theta_j\}_{j=1}^N$ are independent Gaussian variables with mean 0 and standard deviation 1. In this work we used a squared-exponential covariance function and assumed the correlation lengths l_x and l_z are isotropic:

$$R(\vec{x_1},\vec{x_2}) = \sigma_C^2 \exp\left(-\frac{(x_1-x_2)^2}{2l_x^2} - \frac{(z_1-z_2)^2}{2l_z^2}\right). \tag{7}$$

Here σ_C is the standard deviation of the velocity perturbation.

We need to generate a random vector, θ, of dimension equal to the number of terms used in the KLE. The vector $(\theta_1,...,\theta_N)$ is generated by a random walk:

$$\theta_j^{n+1} = \beta\theta_j^n + \sqrt{1-\beta^2}\varepsilon_{n+1}, \qquad n = 0,1,2,\cdots \tag{8}$$

where β is a tuning parameter, $0 \le \beta \le 1$, and ε_{n+1} is chosen from a standard Gaussian distribution.

This KLE approach reduces the dimension of the random variables needed to generate a perturbation from 10^6 to between 10^2 and 10^3.

Acoustic Upscaling

We use the 2D constant-density acoustic wave equation to model wave propagation through the Earth, namely

$$\frac{1}{c^2(x,z)}\frac{\partial^2 p}{\partial t^2} + \Delta p = f \tag{9}$$

where c is the wavespeed, p is the acoustic pressure, and f is the seismic source.

However, since we must solve the wave equation thousands of times for McMC, solving it on the fine grid each time is computationally expensive. In this work we focus on operator upscaling (Vdovina et al., 2005; Korostyshevskaya and Minkoff, 2006; Vdovina and Minkoff, 2008; Vdovina et al., 2009) to solve the acoustic wave equation in a more computationally efficient way.

Following Vdovina et al. (2005), we rewrite the second-order acoustic wave equation as a first-order system in space by introducing acceleration, $\vec{v}$. Equation (9) becomes

$$\begin{aligned} \vec{v} &= -\nabla p, \\ \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} &= -\nabla\cdot\vec{v} + f. \end{aligned} \tag{10}$$

The two-dimensional domain Ω is decomposed into a coarse mesh, and each coarse element is further decomposed into a subgrid mesh. Only acceleration is upscaled in this work. The acceleration unknowns can be decomposed uniquely by $\mathbf{V} = \mathbf{V}^c + \delta\mathbf{V}$. Here the coarse acceleration is denoted by $\mathbf{v}^c$ and the fine-grid acceleration internal to each coarse-grid cell is $\delta\mathbf{v}$. Both of these spaces consist of vector functions living on the edges of the grid blocks (see Figure 1).

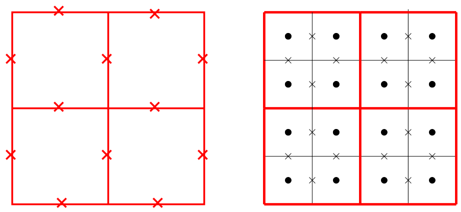

Figure 1: A 2×2 coarse grid with coarse acceleration unknowns and a corresponding 4×4 fine grid with subgrid unknowns. Pressures (denoted by dots) live at the centers of the fine cells. Acceleration (denoted by $\times$'s) lives on cell edges.

The subgrid upscaling technique is based on a mixed finite element variational formulation of the equations. However, by exploiting the equivalence between lowest order mixed finite elements and cell-centered finite differences, the more computationally intensive part of the algorithm reduces to a finite difference discretization (Vdovina et al., 2005). Rewriting System (10) in weak form, we solve for pressure and acceleration:

$$\begin{aligned} \langle \vec{v}, \vec{u} \rangle &= \langle p, \nabla\cdot\vec{u} \rangle \\ \left\langle \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, w \right\rangle &= -\langle \nabla\cdot\vec{v}, w\rangle + \langle f, w\rangle \end{aligned} \tag{11}$$

where $\vec{u}$ and w are test functions.

The algorithm proceeds in two steps:

1. Solve a series of independent subgrid problems.
2. Use the subgrid solutions to solve for acceleration on the coarse grid.

We construct the subgrid problem by restricting System (11) to the test functions $\delta\vec{u}$ which live on the fine grid. Thus we solve

$$\begin{aligned} &\langle \rho(\vec{v}^c + \delta\vec{v}), \delta\vec{u} \rangle = \langle p, \nabla\cdot\delta\vec{u} \rangle, \\ &\left\langle \frac{1}{\rho c^2}\frac{\partial^2 p}{\partial t^2}, w \right\rangle = -\langle \nabla\cdot(\vec{v}^c + \delta\vec{v}), w\rangle + \langle f, w\rangle . \end{aligned} \tag{12}$$

System (12) completely determines pressure.

In Step 1 of the algorithm, we assume that the subgrid problems do not communicate across coarse blocks by imposing the simplifying assumption that $\delta\vec{v}\cdot\nu = 0$ on the boundary of each coarse cell. The expensive part of the algorithm, therefore, exhibits near-perfect speedup with no communication between processes. (See, for example, Table 1 in Vdovina et al. (2005)).

The second step of the algorithm then uses this solution from the subgrid problems to solve the small coarse problem, namely, find the coarse acceleration, $\vec{v}^c$, such that

$$\langle \rho(\vec{v}^c + \delta\vec{v}), \vec{u}^c \rangle = \langle p, \nabla\cdot\vec{u}^c \rangle, \quad \text{for all} \quad \vec{u}^c \in V^c. \tag{13}$$

NUMERICAL RESULTS

In order to use upscaling in a two-stage McMC procedure, we must validate that the wave equation solution produced by the upscaling algorithm is an acceptable surrogate for the fine-grid simulation; i.e., we must show that the data from upscaled simulations is well correlated with the data from fine-grid simulations for velocity perturbations. Selecting a single perturbation as our "true" velocity, we generated our recorded data by running a fine-grid simulation with this reference velocity field. Errors will be calculated as the difference between the reference data and the data from different perturbations. If the errors for the fine-grid simulations are correlated with the errors from the upscaled simulations, then the upscaled problem is a valid filter for McMC.

For our numerical experiments, we used a domain of size 1280 m by 1280 m with 1 m spacing between grid points. The computational region is 960 m x 960 m with 160 m on each edge of the 2D domain designated as an absorbing layer. The layered background velocity (M in Equation (1)) is based on the 960 m of measured P-wave velocity log data provided by Pioneer Natural Resources (see Figure 2). The well log was measured at a field in the Midland Basin, and the logging interval covers portions of the Permian Spraberry, Wolfcamp, and Strawn formations.

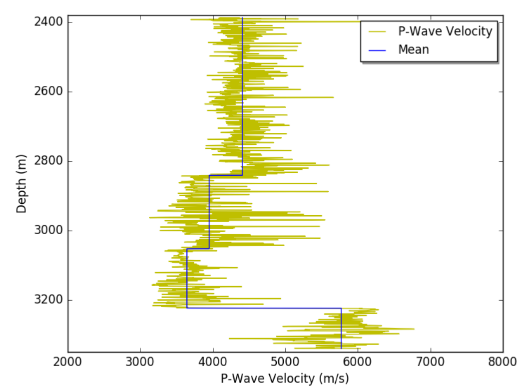

Figure 2: The yellow curve is the well log data (courtesy of Pioneer Natural Resources), and the blue curve is the layered velocity model used as our background model.

To generate perturbations to the velocity, we used the KLE process with standard deviation of 330 m/s and a correlation length of 100 m. The total velocity field we used as our reference velocity (the sum of the layered medium background velocity and a perturbation) is shown in Figure 3.

The source is a Gaussian in space located at the center of the domain. The wave is recorded at 480 receivers placed 90 m be-

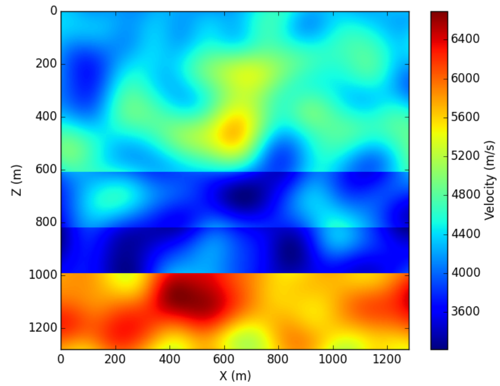

Figure 3: The reference (or “true”) velocity field.

low the surface and at every other grid point within the computational domain. For the upscaled problem, each coarse block contains 8 fine blocks in x and 8 in z. Figure 4 (fine grid simulations) and Figure 5 (upscaled simulations) show the recorded data for all 480 receivers.

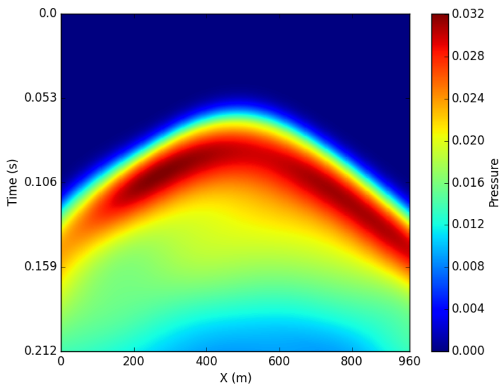

Figure 4: Receiver data (recorded at all 480 receivers) for the full fine grid simulation using the velocity field shown in Figure 3.

Finally, we compared the difference between data simulated from 200 velocity perturbations to data from a reference field simulation (see Figure 3). Specifically, we examined this difference in the case of both fine grid and coarse grid simulations as in $\|d_{\text{ref}} - d_{\text{fine}}\|$ vs. $\|d_{\text{ref}} - d_{\text{upscaled}}\|$. We used linear correlation and found good agreement between the fine and upscaled errors, with correlational coefficient $r = 0.625$. This indicates that there is a moderate linear relationship between the errors and suggests that the upscaled surrogate is a valid filter. Figure 6 shows the scatterplot of the errors. The red line is a hypothetical, perfect 1 to 1 correlation. The blue line is the line obtained by applying linear regression to the data points.

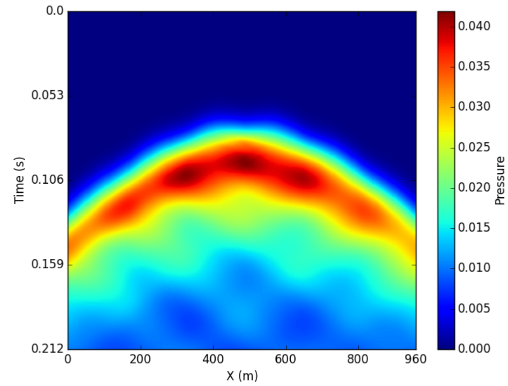

Figure 5: Receiver data for the upscaled simulation using the velocity field shown in Figure 3.

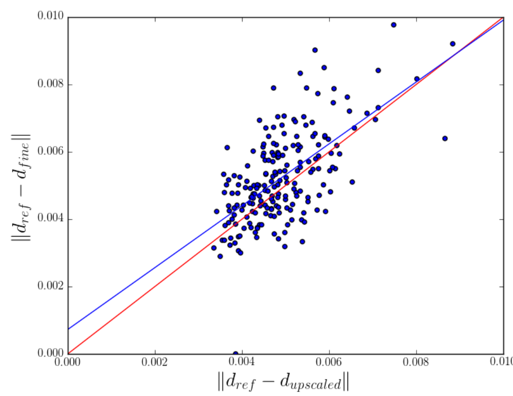

Figure 6: Comparison of errors on the fine grid vs errors from upscaling. The red line shows perfect correlation, the blue line is the regression line.

CONCLUSIONS

While Bayesian techniques have previously been used for seismic inversion, sampling from the posterior distribution using these techniques is computationally expensive. We present a two-stage McMC process using operator upscaling as a filter to reject unacceptable velocity proposals. The fine-grid solution and the upscaled solution are shown to be correlated, thus the upscaled solution is a valid filter. Our next step is to analyze a full two-stage McMC run.

ACKNOWLEDGEMENTS

This work was supported with funding from the National Science Foundation’s Enriched Doctoral Training Program, DMS grant #1514808. The well log and technical advice were generously provided by Pioneer Natural Resources. The authors would like to specifically acknowledge help from Matt McChesney and Rob Meek at Pioneer who consulted on this project. We are also grateful to Dr. Marcio Borges (LNCC-Brazil) for making his KLE code available to us.

EDITED REFERENCES

Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2016 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Akbarabadi, M., M. Borges, A. Jan, F. Pereira, and M. Piri, 2015, A Bayesian framework for the validation of models for subsurface flows: Synthetic experiments: Computational Geosciences, **19**, 1231–1250, http://dx.doi.org/10.1007/s10596-015-9538-z.

Christen, J. A., and C. Fox, 2005, Markov chain Monte Carlo using an approximation: Journal of Computational and Graphical Statistics, **14**, 795–810, http://dx.doi.org/10.1198/106186005X76983.

Cotter, S. L., G. O. Roberts, A. M. Stuart, and D. White, 2013, MCMC methods for functions: Modifying old algorithms to make them faster: Statistical Science, **28**, 424–446, http://dx.doi.org/10.1214/13-STS421.

Efendiev, Y., T. Hou, and W. Luo, 2006, Preconditioning Markov chain Monte Carlo simulations using Coarse-Scale models: SIAM Journal on Scientific Computing, **28**, 776–803, http://dx.doi.org/10.1137/050628568.

Gamerman, D., and H. Lopes, 2006, Markov chain Monte Carlo: Stochastic simulation for Bayesian inference (2nd ed.): Taylor & Francis, Chapman & Hall/CRC Texts in Statistical Science.

Gesret, A., N. Desassis, M. Noble, T. Romary, and C. Maisons, 2015, Propagation of the velocity model uncertainties to the seismic event location: Geophysical Journal International, **200**, 52–66, http://dx.doi.org/10.1093/gji/ggu374.

Ginting, V., F. Pereira, M. Presho, and S. Wo, 2011, Application of the two-stage Markov chain Monte Carlo method for characterization of fractured reservoirs using a surrogate flow model: Computational Geosciences, **15**, 691–707, http://dx.doi.org/10.1007/s10596-011-9236-4.

Gouveia, W. P., and J. A. Scales, 1998, Bayesian seismic waveform inversion: Parameter estimation and uncertainty analysis: Journal of Geophysical Research, **103**, 2759–2779, http://dx.doi.org/10.1029/97JB02933.

Hastings, W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications: Biometrika, **57**, 97–109, http://dx.doi.org/10.1093/biomet/57.1.97.

Hong, T., and M. K. Sen, 2009, A new MCMC algorithm for seismic waveform inversion and corresponding uncertainty analysis: Geophysical Journal International, **177**, 14–32, http://dx.doi.org/10.1111/j.1365-246X.2008.04052.x.

Huang, S. P., S. T. Quek, and K. K. Phoon, 2001, Convergence study of the truncated Karhunen–Loeve expansion for simulation of stochastic processes: International Journal for Numerical Methods in Engineering, **52**, 1029–1043, http://dx.doi.org/10.1002/nme.255.

Korostyshevskaya, O., and S. Minkoff, 2006, A matrix analysis of operator-based upscaling for the wave equation: SIAM Journal on Numerical Analysis, **44**, 586–612, http://dx.doi.org/10.1137/050625369.

Loeve, M., 1977, Probability theory: Springer-Verlag New York.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431–12447, http://dx.doi.org/10.1029/94JB03097.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: Society for Industrial and Applied Mathematics, http://dx.doi.org/10.1137/1.9780898717921.

Vdovina, T., and S. Minkoff, 2008, An a priori error analysis of operator upscaling for the acoustic wave equation: International Journal of Numerical Analysis and Modeling, **5**, 543–569.

Vdovina, T., S. Minkoff, and S. Griffith, 2009, A two-scale solution algorithm for the elastic wave equation: SIAM Journal on Scientific Computing, **31**, 3356–3386, http://dx.doi.org/10.1137/080714877.
Vdovina, T., S. E. Minkoff, and O. Korostyshevskaya, 2005, Operator upscaling for the acoustic wave equation: Multiscale Modeling and Simulation, **4**, 1305–1338, http://dx.doi.org/10.1137/050622146.
Wong, E., 1971, Stochastic processes in information and dynamical systems: McGraw-Hill.
Xuan, R., and P. Sava, 2010, Probabilistic microearthquake location for reservoir monitoring: Geophysics, **75**, no. 3, MA9–MA26, http://dx.doi.org/10.1190/1.3417757.