SAND2016-12410J

SALSA3D: A Tomographic Model of Compressional Wave Slowness in the Earth's Mantle for Improved Travel-Time Prediction and Travel-Time Prediction Uncertainty

by Sanford Ballard, James R. Hipp, Michael L. Begnaud, Christopher J. Young, Andre V. Encarnacao, Eric P. Chael, and W. Scott Phillips

Abstract The task of monitoring the Earth for nuclear explosions relies heavily on seismic data to detect, locate, and characterize suspected nuclear tests. Motivated by the need to locate suspected explosions as accurately and precisely as possible, we developed a tomographic model of the compressional wave slowness in the Earth's mantle with primary focus on the accuracy and precision of travel-time predictions for *P* and *Pn* ray paths through the model. Path-dependent travel-time prediction uncertainties are obtained by computing the full 3D model covariance matrix and then integrating slowness variance and covariance along ray paths from source to receiver. Path-dependent travel-time prediction uncertainties reflect the amount of seismic data that was used in tomography with very low values for paths represented by abundant data in the tomographic data set and very high values for paths through portions of the model that were poorly sampled by the tomography data set. The pattern of travel-time prediction uncertainty is a direct result of the off-diagonal terms of the model covariance matrix and underscores the importance of incorporating the full model covariance matrix in the determination of travel-time prediction uncertainty. The computed pattern of uncertainty differs significantly from that of 1D distance-dependent travel-time uncertainties computed using traditional methods, which are only appropriate for use with travel times computed through 1D velocity models.

Introduction

The goal of nuclear explosion monitoring is to detect nuclear explosions that are detonated anywhere on the Earth, including underground, in the oceans, and in the atmosphere. Signals generated by explosions in these environments are detected using seismic, hydroacoustic, and infrasound sensors, respectively. In addition to the waveform technologies, radionuclide detectors measure radioactive particles generated during a nuclear explosion and provide the ultimate confirmation that a detected explosion was, in fact, nuclear.

Since the early 1980s, all nuclear tests have been conducted underground, with their seismic signatures providing the first indication of their occurrence. When an underground nuclear test occurs, many questions arise as to its nature, including its location, yield, and the type of device detonated. The location of the event is particularly important because the Comprehensive Nuclear-Test-Ban Treaty specifies that the international community may conduct an onsite inspection in an area of 1000 km^2 (Comprehensive Nuclear-Test-Ban Treaty Organization [CTBTO], 1996), necessitating that the location of the suspected event be known to at least that accuracy and precision.

These considerations motivate the need to locate seismic events as accurately and precisely as possible. For underground explosions that do not generate a surface expression, seismic signals provide the only evidence to locate the event. Seismic event location is generally performed by starting with an initial estimate of the event origin time and location and computing residuals between observed arrival times and arrival times predicted using an Earth model of the seismic-slowness distribution in the Earth. The event origin time and location hypotheses are updated in some systematic way in order to minimize the difference between the observed and predicted travel-time data (Geiger, 1910). The accuracy of this approach relies to a significant degree on the accuracy of the travel-time predictions computed with the Earth model, and event mislocations can be significant when predictions are erroneous.

Historically, Earth models used to predict travel times for seismic event location have been 1D models, in which the seismic slowness of the materials that comprise the Earth are assumed to vary only as a function of radius, and lateral heterogeneity of the slowness distribution is ignored.

Examples of 1D models that have been used extensively for seismic event location include the preliminary reference Earth model (PREM; Dziewonski and Anderson, 1984), iasp91 (Kennett and Engdahl, 1991), and ak135 (Kennett *et al.*, 1995). To first order, the 1D slowness distribution assumption is very good, and travel-time predictions produced using this approach can be calculated very efficiently (Buland and Chapman, 1983; Crotwell *et al.*, 1999). However, the accuracy and precision of travel times computed using 1D slowness models is limited because they ignore lateral heterogeneity of seismic slowness in the Earth that is known to exist. The deleterious effects of lateral inhomogeneity are most pronounced at local and regional distances where the seismic rays spend most of their time in the highly heterogeneous shallow parts of the Earth. The effects can also be significant when network geometry is poor and incorrect slowness estimates in the region between the event location and the locations of the stations can pull or push the calculated event location closer to or further from the locations of the stations.

Many approaches have been adopted to improve the accuracy of travel-time predictions and hence event locations. One approach is to apply empirical corrections to travel-time predictions derived with 1D slowness models. The International Data Centre (IDC) of the CTBTO currently uses this approach, applying source-specific station corrections to travel-time predictions obtained with the iasp91 model in order to improve their accuracy (Firbas *et al.*, 1998). Another implementation of empirical corrections is described by Schultz *et al.* (1998), who use the kriging method to implement correction interpolation.

Although many studies have been published describing multidimensional models that were developed for the purpose of understanding Earth structure and dynamics (Li *et al.*, 2008, and Simmons *et al.*, 2010, are just two examples), fewer models have been developed expressly for improving travel-time prediction accuracy. Phillips *et al.* (2007) show that by approximating the velocity structure of the upper mantle as a simple laterally varying velocity gradient, *Pn* travel times can be predicted with good accuracy. The regional seismic travel-time model (Myers *et al.*, 2010) builds on this work to develop a comprehensive tomographic model that is used to predict *P* and *S* regional and local phases. Simmons *et al.* (2012) have also produced a tomographic model motivated by the goal of improving seismic event locations. Other notable examples include Hartog *et al.* (2003).

In addition to improving the accuracy of event locations, reliable estimates of the uncertainty of their locations are also required. An example is the scenario in which a seismic event occurs near a suspected nuclear test site. Because of measurement and model errors, the computed event location is unlikely to coincide exactly with the suspected test site, and a reliable estimate to the uncertainty of the event location is required in order to convincingly attribute the event to the site. Most methods of determining event-location uncertainties depend to a significant degree on the uncertainties of the travel-time predictions used to compute the locations (Flinn, 1965; Evernden, 1969; Jordan and Sverdrup, 1981).

It should be noted that travel-time prediction uncertainties affect not only the size and shape of the hypocenter uncertainty ellipsoids but also influence the locations of the events as well. This is because the travel-time residuals used in location are generally weighted by the combined uncertainties of the measurements and the predictions, with highly precise observations and predictions receiving more weight in the location calculation compared to those with lower precision. So, we seek to improve both the accuracy and the precision of seismic locations by improving both the accuracy and the precision of the travel-time predictions. In turn, improvements in travel-time prediction accuracy and precision can be achieved by improving the accuracy and precision of the Earth models upon which they depend.

To develop an Earth model specifically for locating the seismic events accurately and precisely, we performed a 3D tomographic inversion of an extensive seismic travel-time data set. Throughout the process, we focused on the uncertainty of all aspects of the model in order to derive the most reliable estimates of the uncertainty of the model parameters and the travel-time predictions derived from those parameters, all with the ultimate goal of producing reliable estimates of the uncertainty of seismic event locations. In this article, we describe the methodology used to compute our model and the travel-time predictions and prediction uncertainties derived therefrom. In a subsequent article, we will describe the improvements in the accuracy and precision of seismic event locations that are obtained using our model.

Data Set

The data set used for tomographic inversion comprises ~12 million *P* and *Pn* travel-time picks from 13,000 stations and 122,000 events (Begnaud, 2005; Fig. 1). All of the events used are thought to have an epicenter accuracy of 25 km or better (GT25), according to the Bondár criteria (Bondár *et al.*, 2004; Bondár and McLaughlin, 2009). The arrival-time data and the resulting locations come from a variety of sources including the International Seismological Centre (ISC), the IDC Reviewed Event Bulletin (REB), the U.S. Geological Survey Earthquake Data Report (USGS EDR), the USArray, the Deep Seismic Sounding (DSS) in the former Soviet Union (Li and Mooney, 1998), the Engdahl–van der Hilst–Buland (EHB) catalog (Engdahl *et al.*, 1998), as well as various regional networks, temporary deployments, and in-house arrival picks.

Model Domain

We invert for the compressional wave slowness in the Earth's mantle from the Moho down to the core–mantle boundary, with no limitations in the geographic dimensions. We limit our inversion to the mantle because most seismic nuclear monitoring event location is done with *P* and *Pn* phases, which bottom in the Earth's mantle. For example,

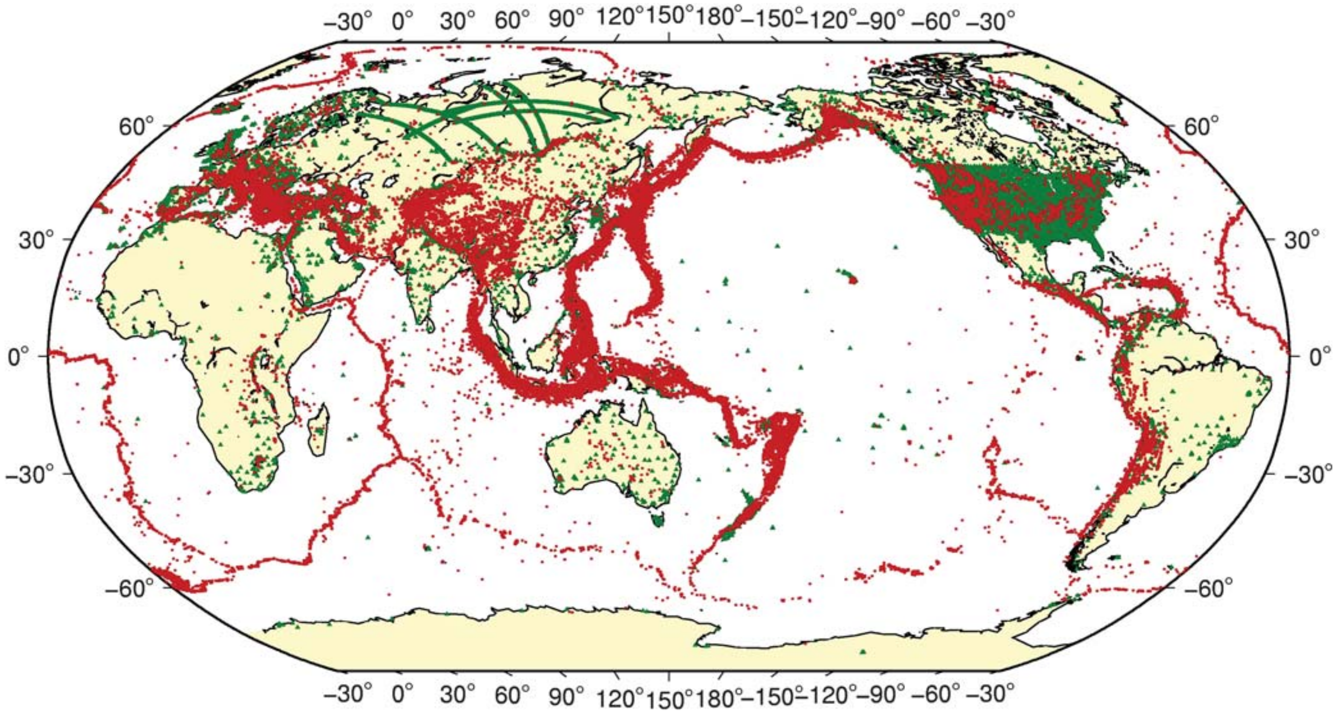

Figure 1. Stations (triangles) and events (circles) used for global tomography inversion.

75% of the location time-defining signal detections in the IDC's REB are *P* and *Pn*. We infer from this that significant improvements in travel-time predictions and prediction uncertainty for these phases will provide the bulk of the improvement that is to be readily obtained. We omit the crust from our inversion because rays for phases *P* and *Pn* travel predominantly vertically through the crust with only short components of horizontal travel. This ray geometry contributes few crossing rays for these phases in the crust, providing minimal constraint on the slowness distribution in that layer of the Earth. Inclusion of shear-wave slowness, crustal phases, core phases, and other secondary phases will be the subject of future modeling efforts.

All of the seismic rays in our data set traverse the crust on the receiver side, and most do on the source side as well. Because we do not invert for the slowness of the crust, we include in our global model a crustal model that reflects the true crust as accurately as possible. Our model of the crust consists of the National Nuclear Security Administration (NNSA) Unified Model for Eurasia (Pasyanos *et al.*, 2004; Steck *et al.*, 2004) with the CRUST2.0 model elsewhere (Laske and Masters, 1997; Laske *et al.*, 2000). The Unified Model was developed by dividing identified tectonic regions into numerous polygons and assigning 1D crustal/upper-mantle models to each polygon based on surveys of the best available regional models for that region.

Although the compressional slowness of the Earth's mantle will be optimized during tomographic inversion, it is very important to select an appropriate starting model of the slowness of the mantle from which to begin the inversion. Starting with a highly featured model, with heterogeneities hypothesized to exist in order to satisfy other constraints unrelated to *P*-wave travel time, might constrain the model to local minima and prevent the inversion from finding the compressional wave slowness distribution that optimally satisfies the *P*-wave travel-time data. For this reason, we chose to start with a very simple smooth model and heavily damp the inversion during early iterations to only allow it to move slowly away from the starting mode, at least initially. Our starting model for the mantle consists of a heavily smoothed version of the NNSA Unified Model in the upper mantle beneath Eurasia, with the ak135 model everywhere else. We chose the Unified Model for Eurasia because the preliminary tomographic models suggested that the upper-mantle velocities in the ak135 model were clearly much too slow beneath the Siberian craton as revealed by the very limited amount of *Pn* data from that region. We chose to extend the fast upper-mantle velocities to other parts of the region that were also poorly sampled by the available data.

To represent the compressional wave slowness distribution on Earth, we use a model parameterization based on the GeoTess open-source software package (Ballard *et al.*, 2016; see Data and Resources).

Tomographic Methodology

We solve for the *P*-wave model slowness distribution in the Earth's mantle using the Bayesian inference method proposed by Tarantola (2005) and Nolet (2008). This method casts the problem in terms of a normal probability distribution that uses a prior model to represent the behavior of the

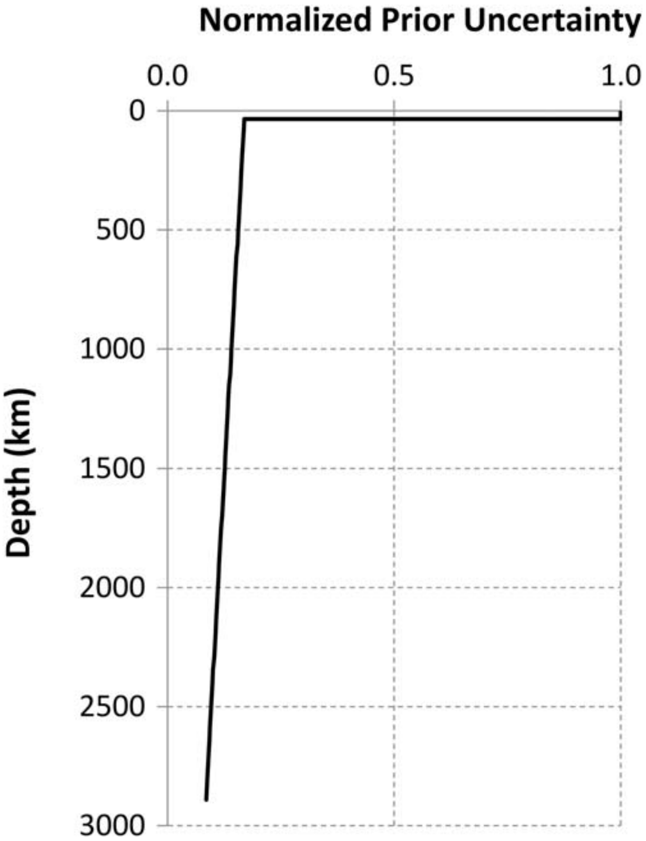

Figure 2. Normalized uncertainty of the prior covariance matrix.

Earth's model parameters as determined by existing geophysical knowledge. The associated prior covariance is used as a regulation control structure to stabilize the tomography inversion. This approach is generally more physically representative, from a geophysical point of view, over the more common Tikhonov regularization method of imposing arbitrary damping and smoothing constraints that do not necessarily include Bayesian information directly.

Tikhonov regularization is the most common method for stabilizing ill-conditioned discrete inverse problems. Aster *et al.* (2005) give a thorough and complete discussion of the approach for the interested reader. Typically, standard Tikhonov regularization may define a regularization method in which amplitude damping is handled by simply defining an identity matrix. More general approaches may go further and formulate a regularization matrix in which the first- and/or second-order derivative changes are also damped. The reader is directed to Tarantola and Valette (1982), Spakman and Nolet (1988), Spakman *et al.* (1993), and Bijwaard and Spakman (2000), for useful examples.

In our Bayesian approach, we define a shape for the amplitude damping, specifically depth-based amplitude damping. Figure 2 illustrates the shape of the *a priori* covariance matrix. The slowness uncertainty is deemed to be geographically invariant and to vary only as a function of depth in the Earth. The slowness uncertainty of the crust is relatively high, decreases sharply at the top of the mantle, and then decreases linearly with depth all the way to the core–mantle boundary where the value is a factor of 2 less than the value at the Moho. We describe only the shape of the slowness uncertainty distribution but not its magnitude because the magnitude of the prior uncertainty and the damping term in our regularization (see equation 2) are inextricably intertwined. In practice, we specify the damping term and modify it throughout the tomographic inversion as is described later, ensuring that in the end, our posterior travel-time uncertainties in poorly sampled areas of our model attain large but still reasonable values.

The following discussion presents the mathematical framework for the Bayesian approach that builds off extensive prior research dealing with travel-time prediction uncertainty in the development of the SALSA3D model (e.g., Young *et al.*, 2008, 2009; Hipp *et al.*, 2011). The details of this approach are described in significant detail in Tarantola (2005), but only the result is given here. The set of normal equations that results from the minimization that finds the maximum-likelihood solution is given by

$$\begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{A}(\mathbf{m}) \\ \mathbf{C}_m^{-1/2} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{d} \\ \mathbf{C}_m^{-1/2}\mathbf{m}_0 \end{bmatrix}. \tag{1}$$

Assume the number of data observations for our problem is N_D and that the number of grid points that we are solving for (in the mantle) is given by N_G. Then the following definitions apply: the matrix $\mathbf{A}$ is the $N_D \times N_G$ data matrix composed of the sensitivities for each data observation defined at each model grid location and depend on the model parameters ($\mathbf{m}$) in a nonlinear fashion; $\mathbf{C}_d$ is the $N_D \times N_D$ observational data covariance matrix (diagonal in our case), which could also include parameterization errors in an additive way; $\mathbf{C}_m$ is the $N_G \times N_G$ *a priori* model covariance matrix that represents the best guess variances and covariances of the prior model distribution defined at all grid positions; $\mathbf{m}$ is the $N_G \times 1$ vector of model parameters for which we are solving (in our case, P-wave slowness); $\mathbf{m}_0$ is the $N_G \times 1$ vector of initial *a priori* model parameters; and $\mathbf{d}$ is the $N_G \times 1$ vector of data observations (P-wave travel times).

We solve the above system of equations using the least-squares (LSQR) method (Paige and Saunders, 1982), which can be numerically unstable if regions of little or no data coverage are present. Although the *a priori* covariance constraint largely solves the coverage problem, its magnitude, especially at the beginning iterations of the solution, may be too small. So, similar to standard regularization approaches, additional solution damping is generally required to ensure stable convergence. Adding damping α to equation (1) yields

$$\begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{A}(\mathbf{m}) \\ \alpha\mathbf{C}_m^{-1/2} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{d} \\ \alpha\mathbf{C}_m^{-1/2}\mathbf{m}_0 \end{bmatrix}. \tag{2}$$

In addition to damping, our solution contains the sensitivity matrix $\mathbf{A}$ that is functionally dependent on the model parameter $\mathbf{m}$. Therefore, we must solve the above equation in a nonlinear fashion and iterate and update $\mathbf{m}$ several times to obtain a converged global minimum solution. We accomplish this during a single iteration by solving for small locally linear changes in the model parameters. Assuming we have a solution at iteration k, the next update is given by

$$\mathbf{m}_{k+1} = \Delta\mathbf{m}_{k+1} + \mathbf{m}_k. \tag{3}$$

Substituting this expression into the previous set of normal equations and simplifying yields

$$\begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{A}_k \\ \alpha\mathbf{C}_m^{-1/2} \end{bmatrix} \Delta\mathbf{m}_{k+1} = \begin{bmatrix} \mathbf{C}_d^{-1/2}(\mathbf{d} - \mathbf{A}_k\mathbf{m}_k) \\ 0 \end{bmatrix}, \tag{4}$$

in which we now write the data matrix with explicit dependence on the previously solved model parameters.

The solution of equation (4) produces a new model ($\mathbf{m}_{k+1}$) in a least-squares sense, which will not precisely equal the residual terms on the right side (RHS) after the model multiplies the sensitivity and constraint matrix assembly on the left side (LHS). To reduce the remaining RHS residual further, we include a set of constant travel-time corrections, in which a single correction is associated with each unique receiver defined in the input observation data set. These receiver "site terms" are calculated by modifying equation (4) as

$$\begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{A}_k & \mathbf{C}_d^{-1/2}\mathbf{R}_T \\ \alpha\mathbf{C}_m^{-1/2} & 0 \\ 0 & \alpha_{\mathrm{ST}}\mathbf{I} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{m}_{k+1} \\ T_{\mathrm{ST}} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_d^{-1/2}(\mathbf{d} - \mathbf{A}_k\mathbf{m}_k) \\ 0 \\ 0 \end{bmatrix}, \tag{5}$$

in which $\mathbf{R}_T$ is an $N_D \times N_R$ sparse matrix, N_R is the number of unique receivers in the observation data set, α_{ST} is the site term damping parameter, $\mathbf{I}$ is the identity matrix, and T_{ST} is the set of site terms for which we are solving, in addition to the model parameters $\mathbf{m}$. The matrix $\mathbf{R}_T$ is a sparse matrix with a single nonzero entry on each row, with a value of 1, defined in the column of $\mathbf{R}_T$ that represents the receiver used by that particular data ray. Equation (5) is only used after a maximum-likelihood model has been determined using equation (4). Additional information concerning the use of site terms and how they are calculated is included in the Tomographic Procedure section.

Equation (4) represents the final set of normal equations that we solve nonlinearly to obtain our new 3D model parameters $\mathbf{m}_{k+1}$. However, we still need a way of defining the *a priori* covariance matrix such that incorporating Bayesian knowledge is not overly difficult. We could begin by defining the *a priori* covariance matrix as a sparse matrix, in which nonzero covariance terms have some spatial proximity to one another and fall to zero over prescribed distances (Rodi and Myers, 2013). However, if we begin with such a definition, then the inverse square root of the *a priori* matrix in equation (4) will be a full matrix of a very large grid for which a solution of equation (4) becomes numerically intractable. Instead, we approach the problem by defining the inverse square root *a priori* covariance matrix directly. This produces a sparse, yet manageable, matrix for purposes of performing the tomography LSQR solution. We then invert this matrix to produce an *a priori* covariance matrix. By setting the diagonal of the initial inverse square root *a priori* matrix to the desired Bayesian diagonal, we can iterate the inversion several times to force the *a priori* covariance matrix to possess approximately the defined Bayesian diagonal. This process is explained in more detail in the Appendix.

Equations (3), (4), (A1), and (A2), and the iteration prescription for the prior covariance defined in equations (A5) and (A6) represent the mathematical framework with which we solve for a new model given a set of data observations, a ray tracer for populating the sensitivity matrix, an initial *a priori* set of smoothing parameter definitions, and an *a priori* best guess of the slowness uncertainty at each of our model grid points.

In addition to the framework above, we also solve for the model resolution and the posterior covariance. These expressions are given by Tarantola (2005) as

$$\mathbf{R} = \tilde{\mathbf{C}}_m\mathbf{A}^T\mathbf{C}_d^{-1}\mathbf{A} = \mathbf{I} - \tilde{\mathbf{C}}_m\mathbf{C}_m^{-1}, \tag{6}$$

in which $\tilde{\mathbf{C}}_m$ is the $N_G \times N_G$ posterior model covariance defined by

$$\tilde{\mathbf{C}}_m = [\mathbf{A}^T\mathbf{C}_d^{-1}\mathbf{A} + \mathbf{C}_m^{-1}]^{-1}. \tag{7}$$

The resolution matrix is used as part of our grid adaption algorithm, which is defined in the Tomographic Procedure section. The posterior covariance is used to produce ray-path uncertainties from our calculated Earth model, which is described in the Travel-Time Predictions and Uncertainty section.

Our model is composed of approximately 230K grid nodes, requiring about 0.2 TB of memory to store. We solve for the resolution and covariance matrices using blocked submatrix out-of-core techniques (D'Azevedo and Dongarra, 2000). The solution requires roughly 12 hrs to compute using ~400 threads distributed over 10 machines, each possessing 64–768 GB of memory per machine.

Tomographic Procedure

A single application of equation (4) is embedded in an iterative tomographic procedure consisting of three nested loops (see Fig. 3). The algorithm begins with the initial data set and a coarsely sampled version of the starting model. The mantle is represented with triangles with ~8° edge lengths and the radial grid spacing is 125, 50, and 30 km in the lower mantle, transition zone, and upper mantle, respectively. The damping parameter α is initially set to a relatively high value so that the model slowness changes very slowly with each

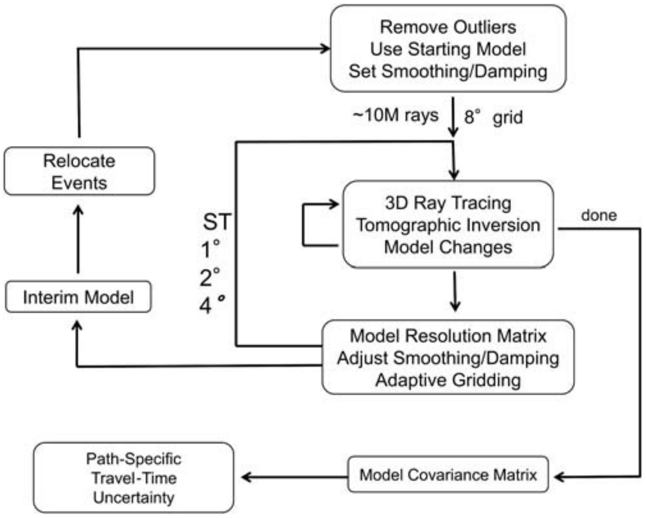

Figure 3. Flowchart defining the tomographic procedure used to generate SALSA3D.

application of LSQR, thereby reducing the probability that the solution will be trapped in local minima.

The innermost loop of the tomographic procedure consists of the following steps: (1) form the *a priori* inverse square-root covariance matrix (constraint) using the iterative equations (A5) and (A6) for the current model grid. (2) Trace all the rays in the tomography data set using the pseudobending algorithm (Um and Thurber, 1987; Zhao and Lei, 2004; Ballard *et al.*, 2009). Populate the matrix of data kernels $\mathbf{A}$ with the model parameter weights derived by interpolating the data ray paths through the model, and the vector of data residuals $\Delta\mathbf{d}$ with the difference between observed and computed travel times, weighted by the data uncertainties. (3) Execute an LSQR calculation and update the slowness values in the tomographic model using the solution vector $\Delta\mathbf{m}$. Repeat steps 2 and 3 until changes to model slowness are small. Convergence is typically achieved after about a half-dozen iterations. We refer to this inner loop as an adaption iteration.

At the conclusion of each adaption iteration, the $N_G \times N_G$ model resolution matrix of the latest model is computed (equation 6). The diagonal elements of the resolution matrix provide an indication of the ability of the available data to resolve the slowness of the model, given the current spatial resolution of the model grid. Using the GeoTess software package (Ballard *et al.*, 2016, see Data and Resources), we refine the spatial resolution of the model grid in the vicinity of grid nodes, in which the diagonal of the model resolution matrix exceeds a user-specified threshold, usually 0.25. We also reduce the damping factor α by a factor of 2 and repeat additional adaption iterations until no points in the model exceed the model resolution threshold. We refer to this sequence of adaption iterations as a relocation iteration. Figure 4 illustrates the diagonal of the model resolution matrix and the corresponding spatial grid resolution in the lower mantle as a function of adaption iteration count.

At the conclusion of the first adaption iteration, computed with uniform 8° triangles (Fig. 4a), high resolution is observed in the vicinity of the Black Sea extending eastward to the subduction zones in the western Pacific, reflecting dense data coverage in that region. High resolution is also observed in western Canada resulting from the abundant seismicity in Alaska observed by stations of the USArray in the United States (Fig. 4b). In response to the high resolution, the grid is refined from 8° to 4° in those areas (Fig. 4c) and the tomography is resumed. After another adaption iteration, the resolution in the refined areas is reduced (Fig. 4d), reflecting the inherent trade-off between model resolution and grid resolution. When the grid resolution is increased, fewer rays influence each grid node, reducing the ability of the data to constrain the model parameter values at the grid node and thereby reducing the model resolution. However, the damping is also reduced, leading to a widespread increase in model resolution over the whole model. Damping tends to reduce the influence of the data on the model grid, thereby reducing model resolution.

The widespread increase in model resolution caused by the decrease in damping justifies additional grid refinement in previously refined areas as well as additional grid refinement beneath much of the United States and the subduction zones of the western circum-Pacific subduction zones (Fig. 4e). This second round of grid refinement results in decreased model resolution in the refined areas (Fig. 4f), but again, the reduction in damping results in a competing increase in model resolution in some unrefined areas. A final round of grid refinement leads to grid refinement in South America and reduces the triangle size to ~1° in discrete areas of western North America, the Middle East, and the western Pacific (Fig. 4g). At the conclusion of the final adaption iteration, the pattern of model resolution (Fig. 4h) is more complex, the maximum value is much reduced, and areas of high resolution are much more diffuse. This procedure has allowed us to refine the grid in the parts of the model in which high grid resolution is justified by the available seismicity, without oversampling the model in regions with low seismicity. The size of the model, measured by the number of grid nodes, would have been an order of magnitude larger had we used a uniform 1° grid for the whole model.

After four adaption iterations have been performed, parts of the model have triangle edge lengths as small as ~1°, radial node spacing as small as 10 km, and the model resolution is everywhere less than the threshold of 0.25. We now perform a site term iteration (equation 5), which involves the calculation of five LSQR iterations during which site terms are computed in addition to model slowness changes. No refinement of the spatial grid resolution or reduction in damping occurs between the final adaption iteration and the site term iteration. The site terms computed in this step are stored in our final model and are applied to predicted travel times whenever travel-time predictions are requested from one of the supported stations. The set of

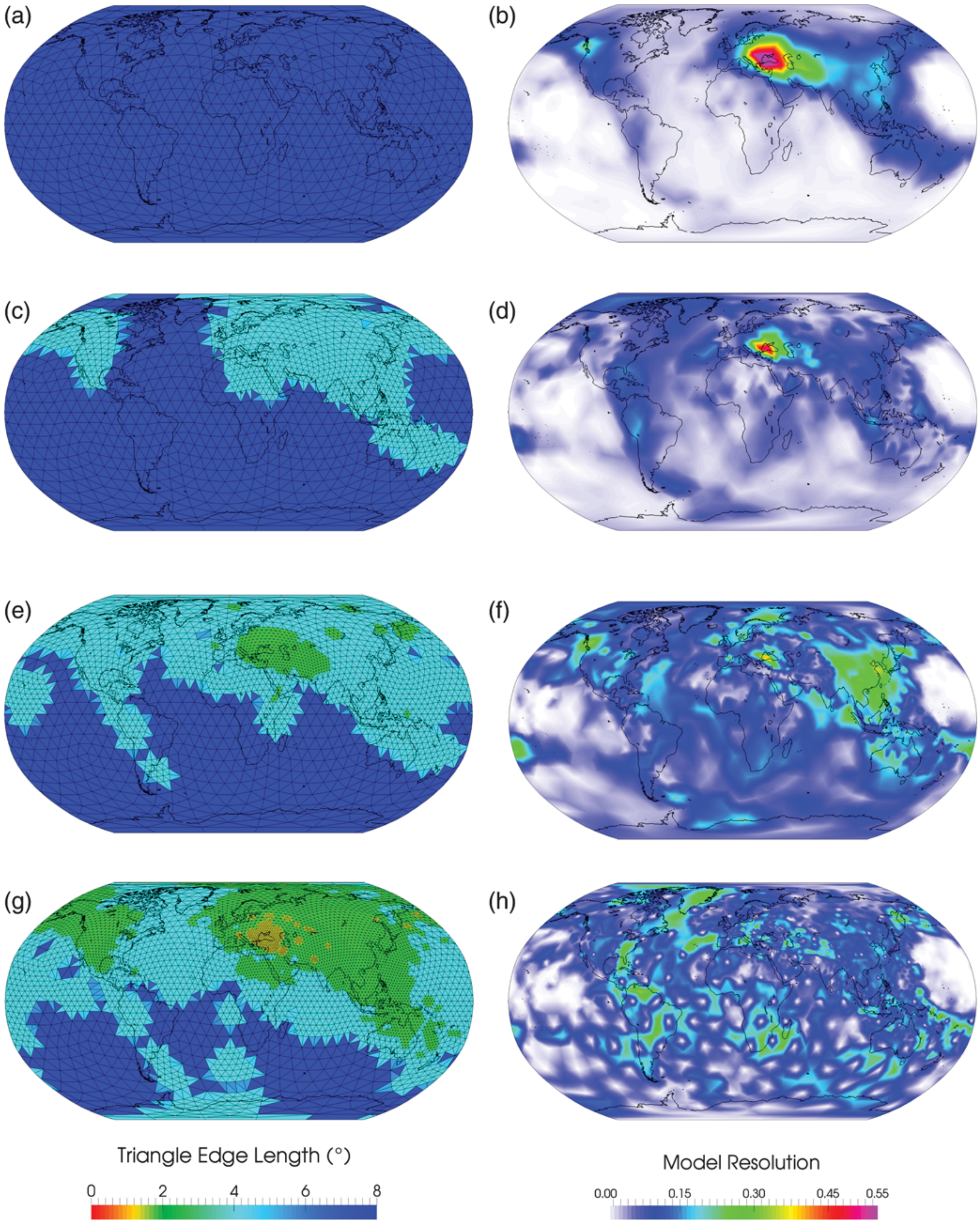

Figure 4. Maps of grid resolution and model resolution at a depth of 800 km in the lower mantle. The grid in (a) was used to compute a tomographic model with model resolution shown in (b). The model resolution in (b) was the basis for performing grid refinement that produced the grid shown in (c), etc. Each time the grid was refined, damping was also decreased, which caused an increase in model resolution.

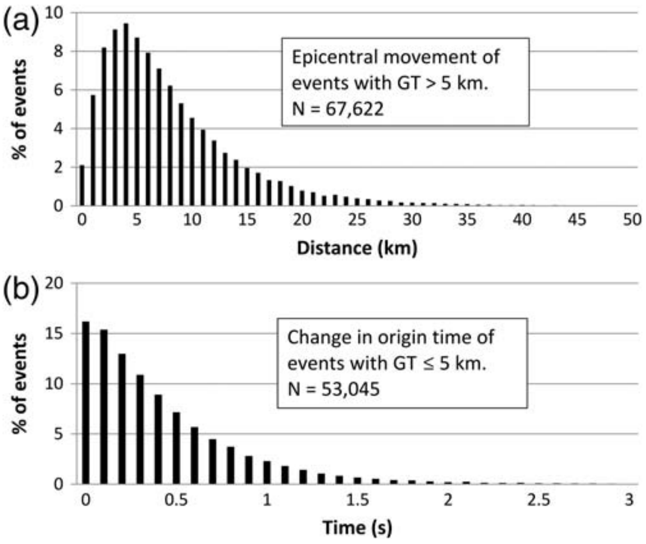

Figure 5. Effect of relocating events at the conclusion of relocation iteration 1. (a) Change in epicenter location for events with ground-truth (GT) level > 5 km. (b) Change in origin time for events with GT level ≤5 km.

adaption iterations plus the site term iteration are called a relocation iteration.

At this point, the model is not yet complete because the seismic events that comprise the tomographic data set are not optimally located with respect to the new model. The original events were likely initially located using a 1D model such as ak135, and hence their locations are inconsistent with the new model. To rectify this situation, all events in the tomography data set are relocated with the latest model. Events with ground-truth (GT) levels of GT5 or less are relocated with their hypocenters fixed, that is, only their origin times are modified. All other events are relocated with only fixed depth. The changes to the event locations and origin times are modest but significant as illustrated in Figure 5.

After relocating all events, the latest model is inconsistent with the new event locations, so the entire relocation iteration is repeated, starting with the original starting model, grid resolution, and damping parameters. Three relocation iterations are implemented, as illustrated by the three curves in Figure 6. The model changes between the first and second relocation iterations are substantial, but the changes from the second to third are negligible, indicating that the event locations and the final model are consistent with each other. The end-to-end tomography procedure reduced the root mean square (rms) residual of the tomographic data set from 1.85 to 0.94 s (50%).

Slowness Model

The major features of the SALSA3D slowness model are illustrated in Figure 7. The maps on the left side of the figure illustrate the difference in slowness between SALSA3D and ak135 at several depths in the mantle. The shallow part of the upper mantle (100 km depth) is characterized by strong contrasts in slowness between various tectonic provinces around the world. Major cratons such as the Canadian shield in North America, the Siberian craton in northern Asia, the Kaapval, Kalahari, and West African cratons in Africa, and the Archean cratons in Australia are all characterized by very fast velocities relative to ak135. Very slow velocities are found in active tectonic provinces such as western North America, the Caribbean, the north Atlantic, the Mediterranean region, the Red Sea rift, and the western and southwestern Pacific subduction zones.

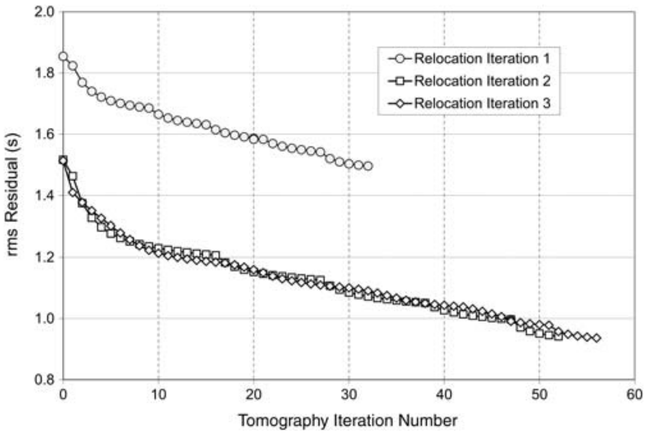

Figure 6. Root mean square (rms) residual versus tomography iteration number. The final rms residual of 0.94 represents a reduction of 50% from the starting model.

In the lower part of the upper mantle (300 km depth), the fast cratonic roots have largely disappeared, but many very slow features in tectonically active regions persist. In the shallow part of the lower mantle (800 km depth), remnants of subducted slabs are evident beneath the Caribbean and South American subduction zones, the Tethyan Seaway, and subduction zones of the southwestern Pacific. The velocities beneath the continents are generally slightly fast compared with ak135, whereas those under oceanic spreading centers tend to be slow. Deep in the lower mantle (1500 and 2500 km depths), slowness anomalies are of relatively low amplitude compared with those at shallower depths, but the general pattern of slightly fast velocities beneath the continents and somewhat slower velocities beneath the oceans persist. The exception is the African superplume, characterized by relatively slow velocities deep beneath southern Africa.

The uncertainty of the slowness of the mantle in SALSA3D is illustrated in the maps on the right side of Figure 7. In the upper mantle, areas of low uncertainty are confined to seismic regions of the Earth, which are well sampled by the data used in tomography. The circum-Pacific subduction zones exhibit very low slowness uncertainty, as do the Mediterranean and southern Asia. High uncertainties are evident in the central Pacific Ocean, northern Africa, and Antarctica, all regions of the world that are poorly represented in the tomographic data set used to generate SALSA3D. The

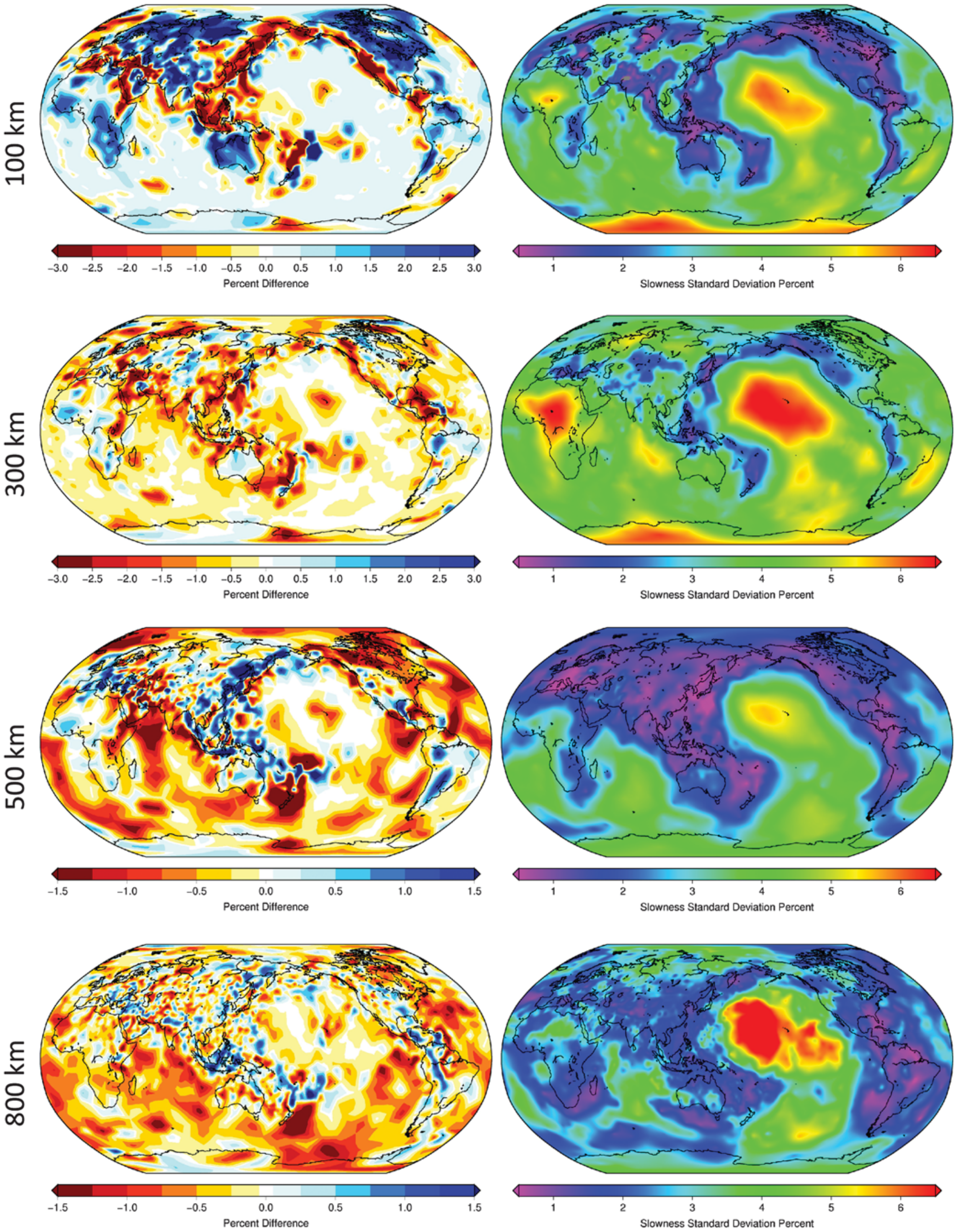

Figure 7. Map of percent slowness change from ak135 (left column) and slowness standard deviation as a percentage of slowness (right column) for depths 100, 300, 500, 800, 1500, and 2500 km. (*Continued*)

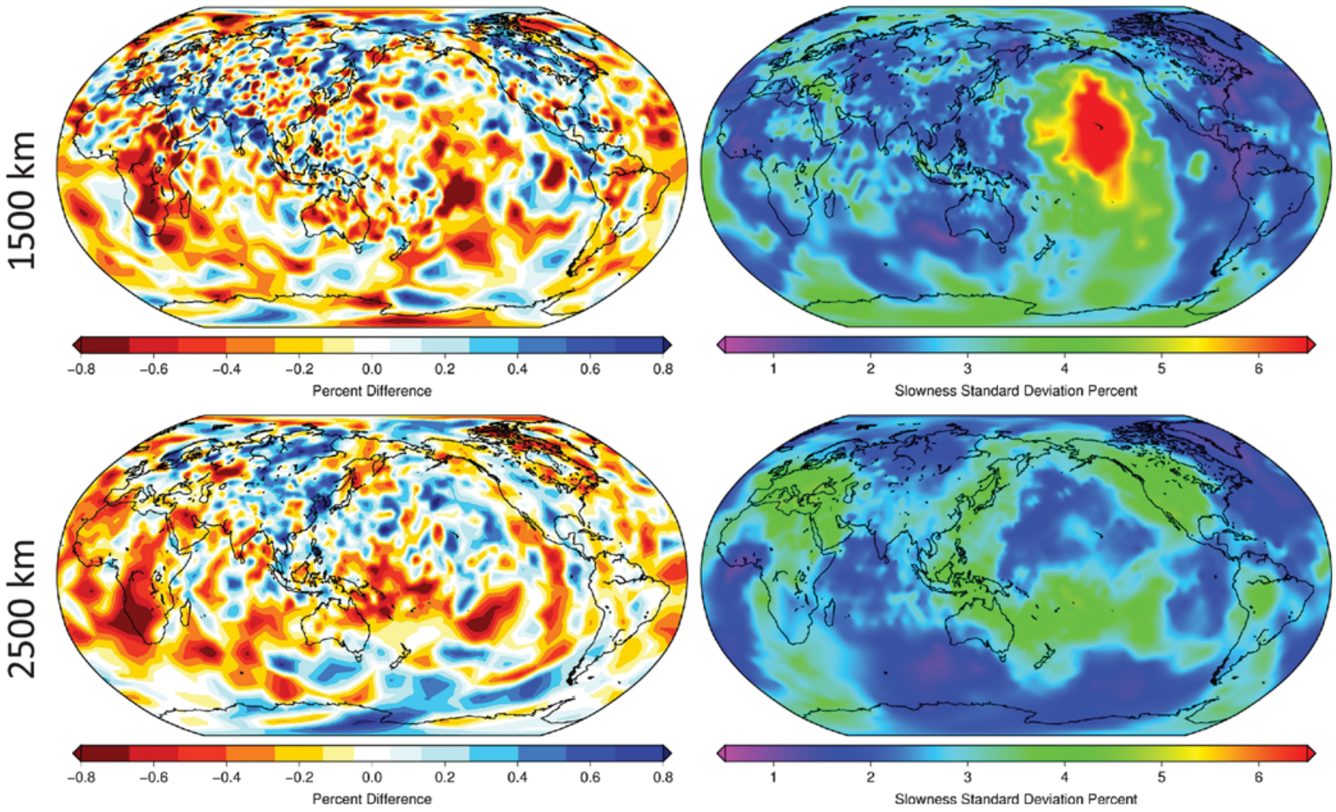

Figure 7. Continued.

slowness uncertainty of the transition zone is very low due to the great many crossing ray paths that bottom in this layer. The slowness uncertainty of the central Pacific Ocean is very high and persists to great depth reflecting the dearth of seismic rays that penetrate this part of the model. The deepest parts of the lower mantle are characterized by widespread regions of low slowness uncertainty, reflecting the abundance of long teleseismic ray paths in the tomographic data set. These teleseismic rays penetrate deep into the lower mantle and provide an abundance of crossing rays at depth, thereby constraining the slowness values in the deep mantle almost worldwide.

The slowness uncertainty images displayed on the right side of Figure 7 are extracted from the diagonal elements of the model covariance matrix described earlier. The diagonal elements are only part of the story, however. The off-diagonal elements are very important for computing travel-time prediction uncertainty. Figure 8 illustrates values of the off-diagonal elements of the covariance between a grid node at the top of the mantle beneath International Monitoring System (IMS) station MJAR in Japan and all other nodes in the model that lie near a cross section from MJAR to the island of Hawaii. Along the bulk of this path, the uncertainty of the model slowness is relatively high (Fig. 7 at 100, 300, and 500 km), reflecting the dearth of crossing rays in the tomographic data set in this region. However, the off-diagonal elements of the covariance matrix for the path from Japan to Hawaii are substantial and negative (Fig. 8), indicating significant negative correlation between the slowness in the mantle below Japan and all of the model grid nodes near the ray path connecting these two locations. This negative correlation implies that if some change in the model caused the slowness of the grid node beneath Japan to change, the slowness of other grid nodes along the path from Japan to Hawaii would have to change in the opposite direction in order to compensate and preserve the predicted source–receiver travel time. Although the uncertainty of model slowness at points along the path may be relatively high, the negative covariance terms will result in the integrated uncertainty along the path being significantly lower than would otherwise be the case.

Travel-Time Predictions and Uncertainty

Our primary objective in developing SALSA3D is to improve the accuracy and precision of travel-time predictions. Predicted travel times through the model tt are given by

$$tt = \mathbf{W}s, \tag{8}$$

in which $\mathbf{W}$ is a matrix containing sensitivity weights obtained by interpolating points along a ray path through the model, and s represents the P-wave model slowness. The matrix $\mathbf{W}$ could contain the data source–receiver sensitivities defined in the matrix $\mathbf{A}$ of equation (1). In that case, the matrices $\mathbf{A}$ and $\mathbf{W}$ are the same. However, for purposes of

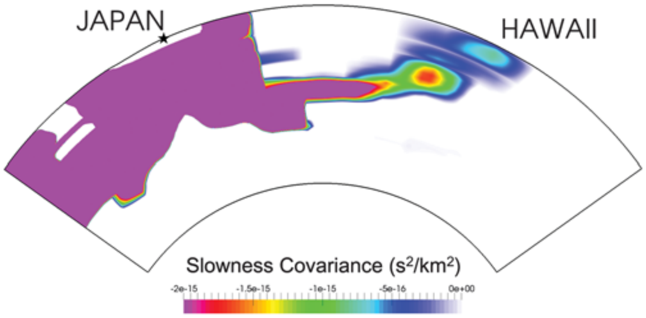

Figure 8. Values of the off-diagonal elements of the model covariance matrix for the grid node at the top of the mantle near International Monitoring System (IMS) station MJAR, marked with a star. The cross section extends from the core–mantle boundary up to the Moho. The ray path from Japan to Hawaii is defined by a tube of negative covariance.

performing arbitrary travel-time predictions through our posterior model, the matrix **W** is generally composed of synthetic regularly spaced source locations. In this way, maps of prediction uncertainties are easily acquired.

To obtain uncertainty estimates for travel-time predictions, we first examine the more general expression that gives the covariance between two distinct ray paths through an arbitrary continuous model. This expression is defined by

$$\sigma_{ij}^2 = \int_{P_i} dx\, w_i(x) \int_{P_j} dx' \mathbf{C}_M'(x, x') w_j(x'), \qquad (9)$$

in which x indicates an arbitrary position along the ith ray path (P_i), x' indicates an arbitrary position along the jth ray path (P_j), $w_i(x)$ is the data kernel, or weight, for the ith path at position x, $w_j(x')$ is the data kernel for the jth path at position x', and $\mathbf{C}_M'(x,\ x')$ is the covariance of the model between the two positions x and x' along the paths i and j. If $i = j$, then this equation gives the variance along a single ray path; if $i \neq j$, then the covariance between the ray paths is given. If we discretize the continuous integral equation to form a discrete grid, as defined for our current problem, then we can write equation (9) as

$$\sigma_{ij}^2 = \mathbf{a}_i \mathbf{C}_m' \mathbf{a}_j^T, \qquad (10)$$

in which $\mathbf{a}_i$ and $\mathbf{a}_j$ are $1 \times N_G$ data kernel vectors (usually sparse) for the ith and jth paths, respectively (these are exactly the sensitivities used in matrix **A** of equation 4), and $\mathbf{C}_m'$ is an $N_G \times N_G$ model covariance matrix.

For this article, we are only interested in the variance (or uncertainty) of a specific ray path, so we set $i = j$ and call the path P. We can now write equation (10) as

$$\sigma_P^2 = \mathbf{a}_P \mathbf{C}_m' \mathbf{a}_P^T. \qquad (11)$$

Because we only invert for slowness values in the Earth's mantle, our model covariance matrix only includes grid nodes in the mantle. However, we know that any source-to-receiver path that we trace through the Earth will pass through the crust at the receiver and generally at the source as well. We account for the additional uncertainty introduced when the ray passes through the crust by defining *a priori* uncertainties for the crust in a manner similar to that in which we defined the mantle *a priori* uncertainties. Then, we rewrite equation (11) to include the crustal covariance matrix, which gives

$$\tilde{\sigma}_P = \sqrt{\mathbf{a}_P(\tilde{m}) \tilde{\mathbf{C}}_m \mathbf{a}_P^T(\tilde{m}) + \mathbf{a}_P(m_c) \mathbf{C}_c \mathbf{a}_P^T(m_c)}, \qquad (12)$$

in which $\mathbf{C}_c$ is the crustal covariance matrix and we have taken the square root to obtain an uncertainty instead of a variance. We specified $\mathbf{C}_c$ to be a diagonal matrix with constant values equal to about 10% of typical crustal slowness values. Here, $\tilde{\mathbf{C}}_m$ is the posterior model covariance so we shall refer to $\tilde{\sigma}_p$ as the posterior uncertainty for ray P through the Earth. Note the dependence on the posterior model parameters $\tilde{m}$ in addition to those model parameters defined at crustal nodes m_c in the determination of the ray-path weights a_P.

The SALSA3D predicted travel times computed using equation (12) are shown in Figure 9 for three IMS stations: GERES, VNDA, and MJAR located in Germany, Antarctica, and Japan, respectively. Each figure illustrates the travel time computed by SALSA3D minus that computed by ak135 for a dense grid of hypothesized surface sources surrounding the station out to 95°, the distance where seismic rays begin to interact with the core–mantle boundary. The strongest travel-time anomalies occur within regional distances from the stations. Rays at these distances travel predominantly through the upper mantle, the part of the Earth's mantle thought to be characterized by relatively strong slowness heterogeneity. Station GERES in Germany is characterized by fast regional anomalies to the north in tectonically stable northern Europe and slow regional anomalies to the south in the active Mediterranean region. Station MJAR exhibits very strong contrasts at regional distances between substantial fast anomalies to points in the ocean to the east and slow anomalies to the west in continental Asia. Travel-time anomalies at teleseismic distances, while more subdued than those observed at closer distances, are not insubstantial and are sufficiently large to be important for seismic event-location calculations.

Maps of the travel-time prediction uncertainty for the same stations are also shown in Figure 9. Station GERES, located in Germany, is characterized by relatively low uncertainty because it is surrounded by regions with plentiful seismicity and seismic receivers. Uncertainties for events located in seismic areas such as the Caribbean, the Mid-Atlantic Ridge, Sumatra, and the subduction zones in the western Pacific are particularly low. In contrast, travel-time prediction uncertainties for station VNDA in Antarctica are relatively high for many regions. This is because the station is located very far south in the southern hemisphere, a part of

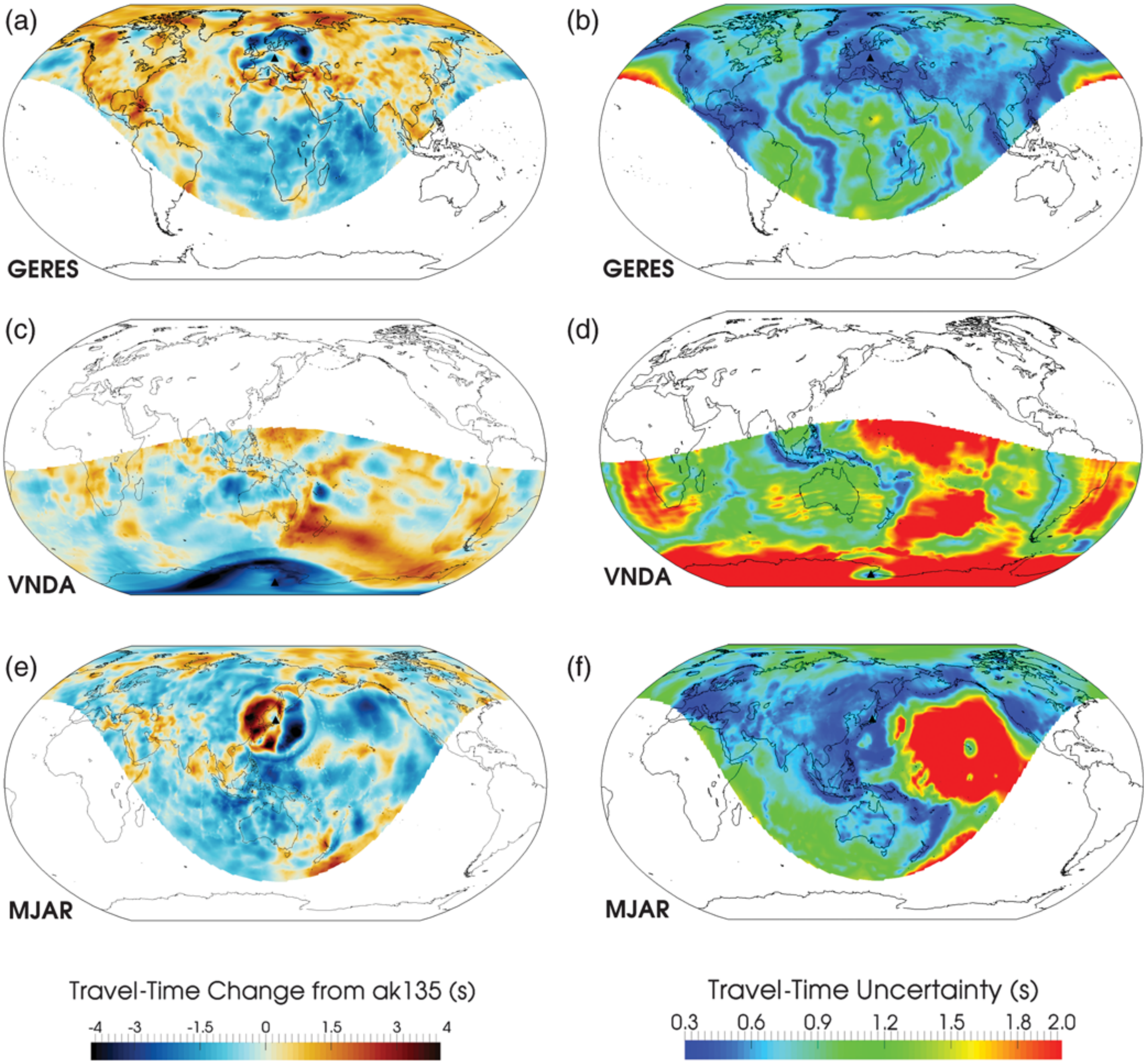

Figure 9. Travel-time prediction and travel-time prediction uncertainty for station GERES in Germany, VNDA in Antarctica, and MJAR in Japan.

the Earth that is poorly represented in our tomographic data set. But, although the uncertainty for events on the continent of Antarctica would be very high, uncertainty for events in the subduction zones of Sumatra, New Zealand, and the rest of the southwestern Pacific is low, reflecting the fact that the paths from station VNDA to these areas are well sampled in the tomographic data set.

Uncertainties for paths to station MJAR in Japan vary significantly depending on whether the event is to the west, which is very well sampled, or to the east, which is very poorly sampled by available data. Although most events in the Pacific are characterized by very high uncertainties at station MJAR, an event near Hawaii would be an exception. Because the path from Japan to Hawaii is fairly well represented in the tomography data set, that path is characterized by relatively low travel-time prediction uncertainty, at least compared with the rest of the Pacific basin.

Components of Travel-Time Prediction Uncertainty

To gain insight into how the full model covariance matrix leads to the overall pattern of prediction uncertainties displayed in Figure 9, we decompose the prediction uncertainties into discrete components by artificially manipulating some of the elements of the posterior model covariance matrix $\tilde{C}_m$ prior to computing the prediction uncertainties

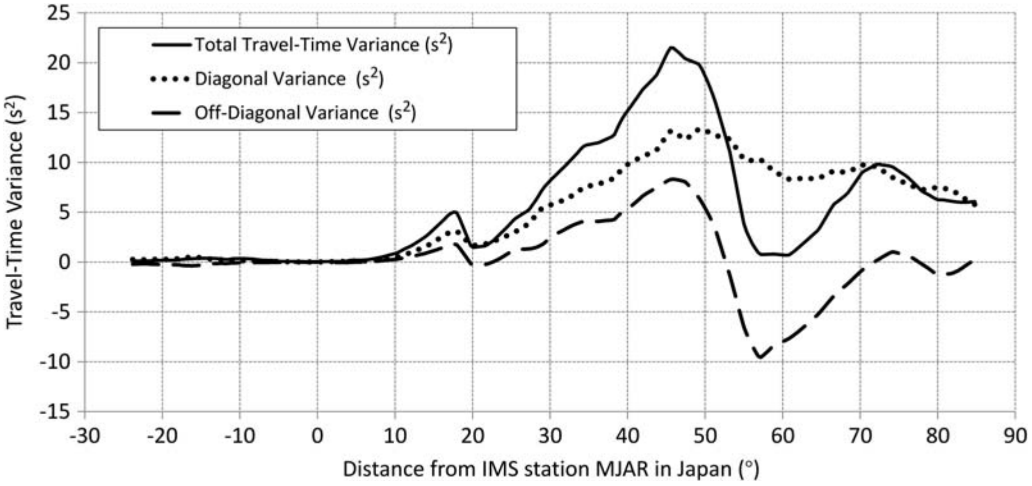

Figure 10. Values of travel-time prediction variance along a transect that passes through IMS station MJAR located in Japan, and the island of Hawaii located 60° to the east of station MJAR. The solid curve is the total travel-time variance. The dotted curve is the travel-time variance that would be computed using only the diagonal elements of the covariance matrix. The dashed curve is the travel-time variance that would be computed using only the off-diagonal elements of the covariance matrix (all diagonal elements set to zero).

using equation (12). Figure 10 illustrates these effects for a path from station MJAR in Japan, through the island of Hawaii and beyond into the eastern Pacific Ocean. This is the same path for which negative covariance terms were illustrated in Figure 8. The solid curve illustrates the unmodified prediction variance (the square of the travel-time prediction uncertainty). The dotted curve is the travel-time prediction variance that would have been calculated if only the diagonal elements of the model covariance matrix had been included in the calculation. The dashed curve is the travel-time prediction variance that would have been calculated if only off-diagonal elements of the model covariance matrix had been used (the diagonal elements set to zero).

In the distance range from about 10° to 45°, the paths for which the prediction uncertainties are calculated travel through parts of the model that are very poorly sampled by the available tomography data, and the computed uncertainties are strongly influenced by the uncertainties of the *a priori* covariance matrix. From 10° to 17° and from 20° to 45°, both components of the total travel-time variance increase monotonically. The travel-time variance due to the diagonal elements of the covariance matrix (dotted curve) rises monotonically with increasing length of the path integral through the substantial *a priori* slowness variance of the shallow mantle in the Pacific basin (Fig. 7 at 100, 300, 500, and 800 km). The component of travel-time variance due to the off-diagonal elements of the covariance matrix (dashed curve) also increases monotonically because the addition of smoothing to the tomographic model acts as a constraint on the inversion and leads to an increase in the uncertainty of travel times predicted by the model. The drop in travel-time variance exhibited by all three curves in the distance range from 17° to 20° is caused by the ray paths dropping from the upper to lower mantle at the *P*-phase triplication distances.

In the distance range from about 50° to 70°, the travel-time variance is significantly depressed by the substantial number of ray paths from Japan to Hawaii that exist in the tomographic data set. These data led to substantial negative values of covariance between a grid node at the top of the mantle under Japan and all other grid nodes near the ray path from Japan to Hawaii (Fig. 8). The reduced travel-time variance is barely evident in the travel-time variance due only to the diagonal elements of the model covariance matrix (dotted curve) but are attributable to the strong negative influence of the off-diagonal elements of the model covariance matrix (dashed curve).

At distances greater than 50°, there is a general decrease in the component of travel-time prediction variance due to the diagonal elements of the model covariance matrix (dotted curve). This long-term trend results from the fact that these long ray paths penetrate deep into the mantle beneath the Pacific Ocean, where the diagonal elements of the model covariance matrix are reduced by the relative abundance of crossing rays in the deep mantle in the tomography data set (Fig. 7 at 1500 and 2500 km).

This experiment illustrates how model covariance leads to significant reductions in prediction uncertainty for well-sampled paths through the slowness model and underscores the importance of using the full model covariance matrix, complete with all off-diagonal terms, for computing travel-time prediction uncertainty.

1D Distance-Dependent Uncertainty

The path-dependent travel-time prediction uncertainties computed from the tomographic model covariance matrix are in sharp contrast to the 1D, distance-dependent uncertainties that have traditionally been used to quantify travel-time prediction uncertainty. The standard approach begins with an

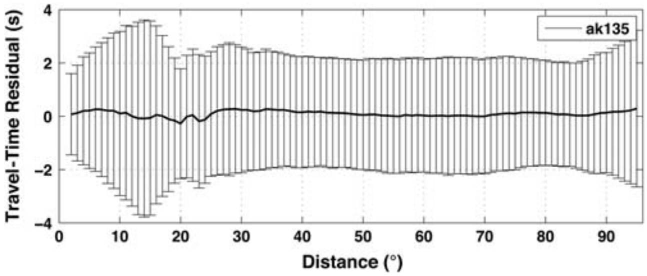

Figure 11. Mean and standard deviation of travel-time residuals as a function of source–receiver separation.

assessment of how well travel times predicted by the model fit actual travel-time observations. Travel-time residuals for a large number of travel-time observations from low GT events are sorted by source–receiver distance, and some measure of the spread of the residuals in small distance bins is computed. These spread values as a function of source–receiver distance are accepted as a measure of the travel-time prediction uncertainty (Yang *et al.*, 2004; Flanagan *et al.*, 2007; Myers *et al.*, 2010, 2015). Results for our tomography data set and the ak135 model are illustrated in Figure 11 and shown in map view for station MJAR in Japan in Figure 12.

As an example, consider any one of the distance bins illustrated in Figure 11. The uncertainty computed for the bin is based on a large number of travel-time residuals for ray paths of a similar source–receiver separation that come from source–receiver pairs all over the world. Some of the ray paths come from regions characterized by relatively high slowness in the true Earth, whereas others come from regions characterized by relatively low slowness values. The predicted travel times were all computed with a single 1D slowness model however; hence the 1D distance-dependent travel-time prediction uncertainties capture the effect of the average heterogeneity of the slowness distribution in the Earth, at least for regions of the Earth with substantial seismicity. Application of the 1D distance-dependent travel-time prediction uncertainties in aseismic regions is tantamount to assuming that the slowness heterogeneity of the seismic regions sampled by the available travel-time data accurately reflects the heterogeneity of aseismic regions as well.

Although this approach may be reasonable for 1D slowness models such as ak135, its application to 3D tomographic models is suspect. 3D models attempt to directly incorporate the slowness heterogeneity of the true Earth and, as we have shown, can significantly reduce travel-time residuals for paths through the Earth that are well sampled by the data used in tomography. If the 1D distance-dependent approach is applied to travel-time predictions computed with a 3D model, the results will be heavily biased toward well-sampled regions where residuals will be the smallest. Regions of the Earth that were poorly sampled by data during tomography will not experience the same degree of residual reduction, but that fact will not be reflected in the travel-time prediction uncertainties computed with the 1D distance-dependent uncertainty approach. Application of the uncertainties appropriate for well-sampled regions to paths through aseismic regions will significantly underestimate the uncertainty of travel-time predictions through those regions.

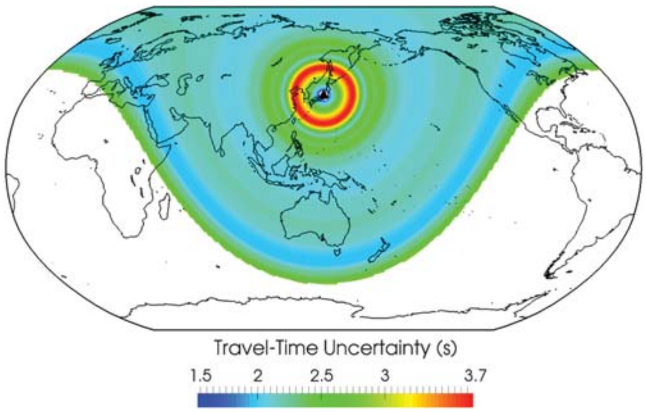

Figure 12. Map of 1D distance-dependent travel-time uncertainty for the ak135 model computed using our tomography data set. Although the uncertainty distribution is shown centered on station MJAR in Japan, the pattern would be identical for any station anywhere on the Earth.

Summary

The task of monitoring the Earth for nuclear explosions relies heavily on the exploitation of seismic data for the detection, location, and characterization of suspected nuclear tests. Motivated by the need to locate suspected explosions as accurately and precisely as possible, we developed a tomographic model of the compressional slowness in the Earth's mantle with primary focus on the accuracy and precision of travel-time predictions for ray paths through the model. Travel-time prediction accuracy is optimized by starting with a very simple, smooth starting model and allowing the model to evolve very slowly during iterative tomographic inversions. We hypothesize that this procedure helps to ensure that the model does not become trapped in local minima of the solution space, as might be the case if a more complicated model was used as the starting model. We also started our tomographic procedure with a very coarse grid in both geographic and radial directions. Periodically during tomography, we computed the model resolution matrix and refined the grid in areas where high values of the diagonal elements of the model resolution matrix suggested that higher spatial resolution was justified. Compatibility of the event locations and model slowness distribution was promoted by relocating the events used in tomography and repeating the entire tomographic procedure. Path-dependent travel-time prediction uncertainties were obtained by computing the full 3D model covariance matrix and then integrating slowness variance and covariance along ray paths from source to receiver.

SALSA3D reveals variations in seismic slowness that reflect the underlying tectonic regimes, with negative slowness anomalies relative to ak135 in stable, fast, cratonic areas and positive slowness anomalies exhibited in younger, slower, tectonically active regions. Travel-time predictions computed with SALSA3D differ by up to several seconds from predictions computed with 1D radial slowness models such as ak135. Although travel-time anomalies at regional distances are substantially larger than those at greater distances, anomalies at teleseismic distances are still substantial. Travel-time prediction uncertainties reflect the amount of seismic data that was used in tomography with very low prediction uncertainties for paths represented by abundant data in the tomographic data set and very high uncertainty for paths through aseismic regions. The pattern of travel-time prediction uncertainty surrounding a station is a direct result of the off-diagonal terms of the model covariance matrix and underscores the importance of incorporating the full model covariance matrix in the determination of travel-time prediction uncertainty. The computed patterns of travel-time prediction uncertainty differ significantly from the 1D distance-dependent travel-time uncertainties computed using traditional methods.

The impact that these improvements in travel-time prediction and prediction uncertainty have on the accuracy and precision of seismic event locations will be the subject of a subsequent paper currently in preparation.

Data and Resources

The website of GeoTess is available at www.sandia.gov/geotess (last accessed August 2016). The International Seismic Center bulletin was obtained via ftp download: ftp://colossus.iris.washington.edu/pub/ (last accessed March 2012). The Reviewed Event Bulletin from the International Data Centre (IDC) was obtained via direct database access. The Earthquake Data Report from the U.S. Geological Survey (USGS) was obtained via the USGS website. The bulletin for the USArray was obtained from the Earthscope Array Network Facility http://anf.ucsd.edu (last accessed February 2014). Various special data sets were obtained from the Incorporated Research Institutions for Seismology–Program for the Array Seismic Studies of the Continental Lithosphere (IRIS-PASSCAL) data center website at http://www.passcal.nmt.edu and the IRIS website at http://iris.edu (last accessed January 2015).

Acknowledgments

We thank Dale Anderson for helpful suggestions and discussions and Stephen Arrowsmith, Megan Slinkard, Wim Spakman, and an anonymous reviewer for constructive reviews of the article. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract Number DE-AC04-94AL85000.

References

Aster, R. C., B. Borchers, and C. H. Thurber (2005). *Parameter Estimation and Inverse Problems*, Elsevier Academic Press, New York, 301 pp., ISBN: 0-12-065604-3.

Ballard, S., J. Hipp, B. Kraus, A. Encarnacao, and C. Young (2016). GeoTess: A generalized Earth model software utility, *Seismol. Res. Lett.* **87,** no. 3, doi: 10.1785/0220150222.

Ballard, S., J. R. Hipp, and C. J. Young (2009). Efficient and accurate calculation of ray theory seismic travel time through variable resolution 3D Earth models, *Seismol. Res. Lett.* **80,** no. 6, doi: 10.1785/gssrl.80.6.989.

Begnaud, M. L. (2005). Using a dedicated location database to enhance the gathering of ground truth information, Los Alamos National Laboratory LA-UR-04-5992, 22 pp.

Bijwaard, H., and W. Spakman (2000). Non-linear global *P*-wave tomography by iterated linearized inversion, *Geophys. J. Int.* **141,** 71–82, doi: 10.1046/j.1365-246X.2000.00053.x.

Bondár, I., and K. L. McLaughlin (2009). A new ground truth data set for seismic studies, *Seismol. Res. Lett.* **80,** no. 3, 465–472.

Bondár, I., S. C. Myers, E. R. Engdahl, and E. A. Bergman (2004). Epicentre accuracy based on seismic network criteria, *Geophys. J. Int.* **156,** 483–496.

Buland, R., and C. H. Chapman (1983). The computation of seismic travel times, *Bull. Seismol. Soc. Am.* **73,** 1271–1302.

Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO) (1996). *Text of the Comprehensive Nuclear-Test-Ban Treaty*, 191 pp., available at http://ctbto.org/fileadmin/content/treaty/treaty_text.pdf (last accessed September 2016).

Crotwell, H. P., T. J. Owens, and J. Ritsema (1999). The TauP toolkit: Flexible seismic travel-time and ray-path utilities, *Seismol. Res. Lett.* **70,** 154–160, available at http://www.seis.sc.edu/TauP/ (last accessed September 2016).

D'Azevedo, E., and J. Dongarra (2000). The design and implementation of the parallel out-of-core ScaLAPACK LU, QR, and Cholesky factorization routines, *Concurrency Comput. Pract. Ex.* **12,** no. 15, 1481–1493.

Dziewonski, A. M., and D. L. Anderson (1984). Structure, elastic and rheological properties and density of the Earth's interior, gravity and pressure, in *Landoldt-Bönstein*, K. Fuch and H. Soffel (Editors), Group V, Vol. 2a, Springer, Berlin, Germany, 84–96.

Engdahl, E. R., R. van der Hilst, and R. Buland (1998). Global teleseismic earthquake relocation with improved travel-times and procedures for depth determination, *Bull. Seismol. Soc. Am.* **88,** no. 3, 722–743.

Evernden, J. F. (1969). Precision of epicenters obtained by small numbers of world-wide stations, *Bull. Seismol. Soc. Am.* **59,** 1365–1398.

Firbas, P., K. Fuchs, and W. D. Mooney (1998). Calibration of seismograph network may meet Test Ban Treaty's monitoring needs, *Eos Trans. AGU* **79,** 413–421.

Flanagan, M. P., S. C. Myers, and K. D. Koper (2007). Regional travel-time uncertainty and seismic location improvement using a three-dimensional *a priori* velocity model, *Bull. Seismol. Soc. Am.* **97,** 804–825.

Flinn, E. A. (1965). Confidence regions and error determinations from seismic event location, *Rev. Geophys.* **3,** 157–185.

Geiger, L. (1910). Herdbestimmung bei erdbeden ans den ankunftzeiten, *K. Gessel. Wiss. Goett.* **4,** 331–349 (in German).

Hartog, R., W. Spakman, C. H. Thurber, and C. Trabant (2003) The influence of path corrections and a three-dimensional global *P*-wave velocity model on seismic event location in Kazakhstan, *Pure Appl. Geophys.* **160,** 2239–2255, doi: 10.1007/s00024-003-2404-2.

Hipp, J. R., A. V. Encarnacao, C. J. Young, S. Ballard, M. C. Chang, W. S. Phillips, and M. L. Begnaud (2011). Calculating a path-dependent travel time prediction variance and covariance for a global tomographic *P*-velocity model, *Proc. of the 2011 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, Orlando, Florida, 283–291.

Jordan, T. H., and K. A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seismol. Soc. Am.* **71,** 1105–1130.

Kennett, B. L. N., and E. R. Engdahl (1991). Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.* **105,** 429–465.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.* **122,** 108–124.

Laske, G., and T. G. Masters (1997). A global digital map of sediment thickness, presented at *Fall Meeting, American Geophysical Union*, December 1997, F483.

Laske, G., G. Masters, and C. Reif (2000). *CRUST 2.0: A New Global Crustal Model at 2x2 Degrees*, available at http://igppweb.ucsd.edu/~gabi/crust2.html (last accessed September 2016).

Li, C., R. D. van der Hilst, E. R. Engdahl, and S. Burdick (2008). A new global model for *P* wave speed variations in Earth's mantle, *Geochem. Geophys. Geosys.* **9,** no. 5, Q05018, doi: 10.1029/2007GC001806.

Li, S., and W. D. Mooney (1998). Crustal structure of China from deep seismic sounding profiles, *Proc. of the 7th International Symposium on Deep Seismic Profiling of the Continents*, Asilomar, California, *Tectonophysics*, Vol. 288, 105–113.

Myers, S. C., M. L. Begnaud, S. Ballard, M. E. Pasyanos, W. S. Phillips, A. L. Ramirez, M. S. Antolik, K. D. Hutchenson, J. J. Dwyer, C. A. Rowe, *et al.* (2010). A crust and upper-mantle model of Eurasia and North Africa for *Pn* travel-time calculation, *Bull. Seismol. Soc. Am.* **100,** no. 2, 640–656, doi: 10.1785/0120090198.

Myers, S. C., N. A. Simmons, G. Johannesson, and E. Matzel (2015). Improved regional and teleseismic *P*-wave travel-time prediction and event location using a global 3D velocity model, *Bull. Seismol. Soc. Am.* **105,** no. 3, 1642–1660, doi: 10.1785/0120140272.

Nolet, G. (2008). *A Breviary of Seismic Tomography*, Cambridge University Press, New York, ISBN: 978-0-521-88244-6.

Paige, C. C., and M. A. Saunders (1982). LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software* **8,** no. 1, 43–71.

Pasyanos, M. E., W. R. Walter, M. P. Flanagan, P. Goldstein, and J. Bhattacharyya (2004). Building and testing an *a priori* geophysical model for western Eurasia and North Africa, *Pure Appl. Geophys.* **161,** 235–281.

Phillips, W. S., M. L. Begnaud, C. A. Rowe, L. K. Steck, S. C. Myers, M. E. Pasyanos, and S. Ballard (2007). Accounting for lateral variations of the upper mantle gradient in *Pn* tomography studies, *Geophys. Res. Lett.* **34,** L14312, doi: 10.1029/2007GL029338.

Rodi, W. L., and S. C. Myers (2013). Computation of traveltime covariances based on stochastic models of velocity heterogeneity, *Geophys. J. Int.* **194,** no. 3, 1582–1595.

Schultz, C. A., S. C. Myers, J. Hipp, and C. J. Young (1998). Nonstationary Bayesian kriging: A predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. Seismol. Soc. Am.* **88,** 1275–1288.

Simmons, N. A., A. M. Forte, L. Boschi, and S. P. Grand (2010). GyPSuM: A joint tomographic model of mantle density and seismic wave speeds, *J. Geophys. Res.* **115,** no. B12310, doi: 10.1029/2010JB007631.

Simmons, N. A., S. C. Myers, G. Johannesson, and E. Matzel (2012). LLNL-G3Dv3: Global *P*-wave tomography model for improved regional and teleseismic travel time prediction, *J. Geophys. Res.* **117,** doi: 10.1029/2012JB009525.

Spakman, W., and G. Nolet (1988). Imaging algorithms, accuracy and resolution in delay time tomography, in *Mathematical Geophysics*, N. J. Vlaar, G. Nolet, M. J. R. Wortel, and S. A. P. L. Cloetingh (Editors), Reidel, Dordrecht, The Netherlands, 155–187.

Spakman, W., S. van der Lee, and R. D. van der Hilst (1993). Travel-time tomography of the European–Mediterranean mantle down to 1400 km, *Phys. Earth Planet. In.* **79,** 3–74.

Steck, L. K., C. A. Rowe, M. L. Begnaud, W. S. Phillips, V. L. Gee, and A. A. Velasco (2004). Advancing seismic event location through difference constraints and three-dimensional models, *Proc. of the 26th Seismic Research Review—Trends in Nuclear Explosion Monitoring*, Orlando, Florida, 346–355.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parametric Estimation*, SIAM, ISBN: 0-89871-572-5, 978-0-89871-572-9.

Tarantola, A., and B. Valette (1982). Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.* **20,** no. 2, 219–232.

Um, J., and C. H. Thurber (1987). A fast algorithm for two-point seismic ray tracing, *Bull. Seismol. Soc. Am.* **77,** 972–986.

Yang, X., I. Bondár, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, G. Ekström, H. Israelsson, and K. McLaughlin (2004). Validation of regional and teleseismic travel-time models by relocating ground-truth events, *Bull. Seismol. Soc. Am.* **94,** 897–919.

Young, C. J., S. Ballard, J. R. Hipp, M. C. Chang, G. T. Barker, M. L. Begnaud, W. S. Phillips, L. K. Steck, and C. A. Rowe (2008). A 3D tomographic model of south central Asia based on *Pn* travel times and GT events, *Proc. of the 2008 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, Portsmouth, Virginia, 297–305.

Young, C. J., S. Ballard, J. R. Hipp, M. C. Chang, J. E. Lewis, C. A. Rowe, and M. L. Begnaud (2009). A global 3D *P*-velocity model of the Earth's crust and mantle for improved event location, *Proc. of the 2009 Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, Tucson, Arizona, 249–257.

Zhao, D., and J. Lei (2004). Seismic ray path variations in a 3D global velocity model, *Phys. Earth Planet. In.* **141,** 153–166.

Appendix

In the tomography methodology discussion above, we described how a prior covariance matrix could be used to regularize our tomographic solution. Here, we show how we construct the prior covariance so that its diagonal shape is very near to a Bayesian definition, which we provide, while its inverse square root (the regularization matrix) is sparse, enabling a tractable least-squares (LSQR) method solution.

We begin by defining the inverse square root *a priori* matrix as a roughening or penalty formulation that accounts for distance-based smoothing terms for each grid point with each of its local neighbors. Next, we normalize the matrix diagonal by pre- and postmultiplying by the inverse square root of the diagonal. Finally, we pre- and postmultiply the matrix by the inverse square root of the prescribed Bayesian uncertainty shape diagonal, which becomes a first guess at reproducing the Bayesian shape in the subsequently calculated covariance matrix. With these steps, the inverse square root *a priori* can be written as

$$ {}_0\mathbf{C}_m^{-1/2} = s_m^{-1/2}\mathbf{D}_S^{-1/2}\mathbf{S}\mathbf{D}_S^{-1/2}s_m^{-1/2}, \qquad \text{(A1)} $$

in which $\mathbf{S}$ is the $N_G \times N_G$ smoothing matrix, $\mathbf{D}_S$ is the diagonal of the smoothing matrix ($\mathrm{diag}(\mathbf{S})$), s_m is the $N_G \times N_G$ prescribed Bayesian uncertainty shape diagonal, and ${}_0\mathbf{C}_m^{-1/2}$ is the first guess at the inverse square root *a priori* matrix. This matrix is, as required, both symmetric and sparse.

Next, we define $\mathbf{S}$ in equation (A1) such that each diagonal row entry represents the sum of all of its off-diagonal near-neighbor weights. We define the weights as

$$ \begin{aligned} \mathbf{S}_{ij} &= -\beta_S K(|r_i - r_j|), \quad j \neq i; \\ \mathbf{S}_{ii} &= \sum_{\substack{j=0 \\ j \neq i}}^{n} K(|r_i - r_j|), \end{aligned} \qquad \text{(A2)} $$

in which $K(|r_i - r_j|)$ is a simple monotonic smoothing kernel that is a function of the distance between points i and j. We define $K(|r_i - r_j| = 0) = 1$ and $K(|r_i - r_j| \geq r_D) = 0$, in which r_D is some reasonably determined distance range over which the kernel is nonzero for a small local region of near neighbors about each grid point. Summing a row of $\mathbf{S}$ over all column entries produces $\mathbf{S}_{ii}(1 - \beta_s)$, in which β_s is a smoothing parameter that gives full smoothing when $\beta_s = 1$, and no smoothing (i.e., the smoothing matrix becomes diagonal) when $\beta_s = 0$.

Given the complete definition of the first guess for the inverse square root *a priori* covariance with equations (A1) and (A2), we can calculate the first guess of the *a priori* covariance matrix by squaring equation (A1), and inverting

$$ {}_0\mathbf{C}_m = [{}_0\mathbf{C}_m^{-1/2}\,{}_0\mathbf{C}_m^{-1/2}]^{-1}. \tag{A3}$$

Usually, the first guess will contain a diagonal whose values are larger, in a mean sense, than the desired Bayesian entries contained in s_m (equation A1). We now perform a scaling update of the initial inverse square root *a priori* matrix by pre- and postmultiplying the matrix by the ratio of the Bayesian uncertainty diagonal with the square root of the *a priori* covariance matrix diagonal:

$$ {}_1\mathbf{C}_m^{-1/2} = \left[\frac{s_m}{\sqrt{\mathbf{D}_{{}_0\mathbf{C}_m}}}\right]^{-1/2} {}_0\mathbf{C}_m^{-1/2} \left[\frac{s_m}{\sqrt{\mathbf{D}_{{}_0\mathbf{C}_m}}}\right]^{-1/2}. \tag{A4}$$

We can continue iterating equations (A3) and (A4) using the prescriptions

$$ {}_i\mathbf{C}_m = [{}_i\mathbf{C}_m^{-1/2}\,{}_i\mathbf{C}_m^{-1/2}]^{-1} \tag{A5}$$

and

$$ {}_{i+1}\mathbf{C}_m^{-1/2} = \left[\frac{s_m}{\sqrt{\mathbf{D}_{{}_i\mathbf{C}_m}}}\right]^{-1/2} {}_i\mathbf{C}_m^{-1/2} \left[\frac{s_m}{\sqrt{\mathbf{D}_{{}_i\mathbf{C}_m}}}\right]^{-1/2}. \tag{A6}$$

Using this method, and just a few iterations K, the diagonal of the *a priori* covariance matrix will approach the prescribed Bayesian uncertainty shape vector:

$$ \sqrt{\mathbf{D}_{{}_k\mathbf{C}_m}} \approx s_m. \tag{A7}$$

We now have an *a priori* covariance matrix whose diagonal is nearly equal to the prescribed Bayesian variance shape, and whose inverse square root matrix is sparse.

Sandia National Laboratories
P.O. Box 5800, MS 0404
Albuquerque, New Mexico 87185
sballar@sandia.gov
(S.B., E.P.C.)

Sandia National Laboratories
P.O. Box 5800, MS 0401
Albuquerque, New Mexico 87185
(J.R.H., C.J.Y., A.V.E.)

Los Alamos National Laboratory
P.O. Box 1663, F665
Los Alamos, New Mexico 87545
(M.L.B., W.S.P.)

Manuscript received 5 October 2015;
Published Online 11 October 2016