A GLOBAL 3D *P*-VELOCITY MODEL OF THE EARTH'S CRUST AND MANTLE FOR IMPROVED EVENT LOCATION

Christopher J. Young[1], Sanford Ballard[1], James R. Hipp[1], Marcus C. Chang[1], Jennifer E. Lewis[1], Charlotte A. Rowe[2], and Michael A. Begnaud[2]

Sandia National Laboratories[1] and Los Alamos National Laboratory[2]

Sponsored by the National Nuclear Security Administration

Award Nos. DE-AC04-94AL85000/SL09-3D_Earth-NDD02[1] and DE-AC52-06NA25396/LA09-IRP-NDD02[2]

ABSTRACT

Effectively monitoring for nuclear tests with yields less than 1 kT using seismic data requires utilizing regional phases whose characteristics vary greatly between different geographic areas. Current approaches typically use separate models developed independently for each area or separate models for regional vs. teleseismic, causing problems when transitioning from one area to another, or from regional to teleseismic distances. Ultimately, what is needed is a single global seamless 3D model derived from a simultaneous inversion of a global data set encompassing regional and teleseismic data from a variety of areas. Several such models have been developed, but generally with the intent of providing insight into the structure of the inner Earth, not of improving treaty-monitoring capability. In this paper, we present our preliminary global *P*-velocity model developed specifically to improve event location using both teleseismic and regional distance phases.

Our base data set is the global EHB catalog, consisting of 130,000 events spanning some 46 years, which was compiled using a specialized algorithm to improve routine event hypocenter locations of global catalogs (e.g., the International Seismiological Centre's Preliminary Determination of Epicenters). We augment this with the more regionally selective Ground Truth (GT) catalog created by researchers at Los Alamos National Laboratory (LANL). The LANL catalog is less geographically-balanced than EHB (1998), but has both more and lower GT-level events for several areas of monitoring interest. Our total number of events is over 200,000, with more than 15 million *P*, *Pn*, and *Pg* ray paths. Because 3D tomography is a strongly non-linear problem, it is important to start with a high-quality initial model. We choose to use a simplified two-layer version of the Crust 2.0 model with a global mantle model recently published by Li et al (2008). To reduce the computational burden of the inversion, and to prevent overweighting due to ray path redundancy, we use representative rays for clusters of rays based on geometric similarity of the entire ray paths traced through our starting model. Using this method, we are able to reduce the number of ray paths by more than 50%. Our model is constructed using the variable resolution tessellation developed by Ballard et al (2009). For the travel time calculations, we use the robust and efficient ray bending algorithm developed by Ballard et al (2009). Sufficient damping is applied to keep the iterative velocity adjustments small enough to be stable, as the updated ray paths change for each iterative model. To produce a model with variable lateral resolution matching ray path coverage and velocity gradients, we use progressive tomographic scale refinement. Both event and site corrections are solved for, but they are introduced only after converging on an intermediate velocity model to force as much of the residual fit into the velocity model as possible. Our inversion was run using the distributed parallel computing framework developed by Sandia National Laboratories (SNL), providing us with over 150 processors, which has been shown to achieve an efficiency of better than 90% for the costly ray-tracing calculations, resulting in more than a 135x speedup over sequential algorithms.

OBJECTIVE

The overall objective of our research is to develop a single, global 3D *P*-velocity model specifically to improve seismic event location from local to teleseismic distances. Though complex to develop, the availability of such a model would considerably simplify the calculation of locations of events of monitoring interest, as well as improve the quality of the locations and the associated uncertainty estimates. Our objective for this first year of our project has been to create and evaluate a prototype model, in the process solving several of the important technical challenges that must be overcome to produce an operationally useful model.

INTRODUCTION

One of the most important characteristics to establish in evaluating a possible nuclear explosion is the location. Location is dependent both on the quality of the data available, which can generally not be controlled, and on the Earth model used, which can be. For large events that are well-recorded at teleseismic distances, an accurate location can be determined using a high-quality global radially-symmetric (1D) reference model such as AK135 (Kennet et al., 1995) because the source to receiver paths are predominantly through the deeper, more laterally homogeneous portions of the Earth and errors in predicted travel time tend to cancel out when azimuthal gap is small. For smaller events, however, some or even all of the signals will only be recorded at regional distances where the source to receiver paths traverse the crust and upper mantle, portions of the Earth that are much less laterally homogeneous. Also, azimuthal gap may be large. To get accurate travel time predictions for regional phases it is necessary to use models that include lateral heterogeneity. There are various possible approaches to doing this. Perhaps the most simple is to use different radially symmetric models for each station. Such models can often be readily obtained from local or regional network operators who use them to produce their catalogs. Any modifications to the event location software required to make use of these station-specific 1D models are generally trivial. However, applying this approach on a global scale introduces non-physical seams between the different 1D regional models and between each regional model and the background teleseismic model. The end result can be very complex and ensuring robust locator behavior across the various seams is not a trivial challenge.

A more practical approach is to model regional phases using a laterally heterogeneous model, such as has been done in the recent development of the Regional Seismic Travel Time (RSTT) model by the Ground-based Nuclear Explosion Monitoring Research & Development program (Myers et al., 2009). This approach is sometimes referred to as geometrically “2.5D” because the model varies fully in the lateral direction, but the mantle is modeled with a single (though laterally varying) gradient. This method produces significantly better travel time predictions and event locations compared to using a standard global 1D reference model, and it is more robust than the amalgamation of regional 1D models discussed above. However, because the regional representation of the Earth is not appropriate for teleseismic phases, it is still necessary to introduce a non-physical seam between the regional model and a background teleseismic model, and this seam must be properly conditioned to ensure proper locator behavior. Further, in producing the tomographic RSTT regional model, it is first necessary to “re-baseline” the known origin times of ground truth data using teleseismic data so that the resultant model will properly locate events with combined regional and teleseismic data.

Thus, while progress has been made towards a model that can produce better locations at regional and teleseismic distances and is simpler to maintain, there is still considerable room for improvement. What is ultimately needed is a model that can produce high quality locations for any data combination, regardless of distance, while possessing no problematic non-physical seams. In short, our ultimate goal is a true 3D model of the Earth, where any seams represent actual physical transitions between different materials within the body of the Earth rather than artifacts introduced by approximations in the modeling approach. Producing such a model is a difficult goal given the tomographer’s usual problems of using noisy data that provides incomplete sampling for much of the Earth and oversamples in the limited seismically active regions. In spite of this, several global models have already been produced, providing valuable insight into the overall structure of the inner Earth (e.g., Li et al., 2009). For nuclear explosion monitoring purposes, however, it is not clear that any of the currently available global 3D models provides improved locations, a topic examined in a companion paper (Rowe et al., 2009). This is not necessarily an indication of questionable quality in these models but rather the original purpose in developing them: most tomographic models have been developed to image the Earth’s interior, not to improve locations (RSTT being an exception). Our goal in this paper is to take the first step towards production of a global 3D tomographic model whose foremost purpose it to improve event locations from data recorded at distances from local to teleseismic. The model presented is the first in a series that we hope will eventually lead to an operationally-capable model, and producing it has forced us to address some of important technical challenges, which we document here.

DATA

To provide good overall coverage, we choose as our base dataset the global EHB catalog (Engdahl et al., 1998), which was compiled using a specialized algorithm to improve routine event hypocenter locations of various global catalogs (e.g., ISC, PDE). We used the recently updated version of EHB consisting of 131,000 events spanning 46 years (http://www.isc.ac.uk/EHB/index.html). Sources and receivers for the EHB catalog are shown in Figure 1. For the purposes of this study, we use only *P*, *Pn*, and *Pg* data, resulting in about 14 million ray paths. Because we are interested in developing a model with the best possible location capability in focused areas of monitoring interest, we augment EHB with the more regionally selective GT catalog created by researchers at LANL. The 77,000-event LANL catalog is less geographically-balanced than EHB, but has both more and lower GT-level events for several areas of monitoring interest. The total number of *P*, *Pn*, and *Pg* ray paths for the LANL catalog is more than 8 million.

EHB Catalog

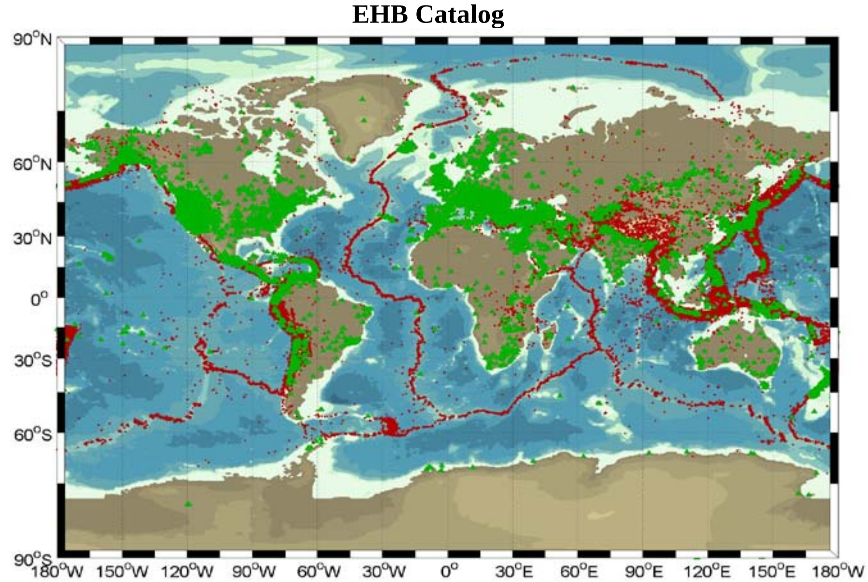

Figure 1. Origins (red circles) and stations (green triangle) for the EHB catalog.

Figure 2 shows histograms of travel time residuals relative to AK135 for the EHB catalog for both *P* and *Pn*. Because AK135 is global average model and this is a global data set, there is very little overall bias for either phase, as would be expected (0.103 s for *Pn* and 0.005 s for *P*). However, the standard deviations, especially for *Pn* (2.204 s), indicate that there is considerable misfit between the data and the model.

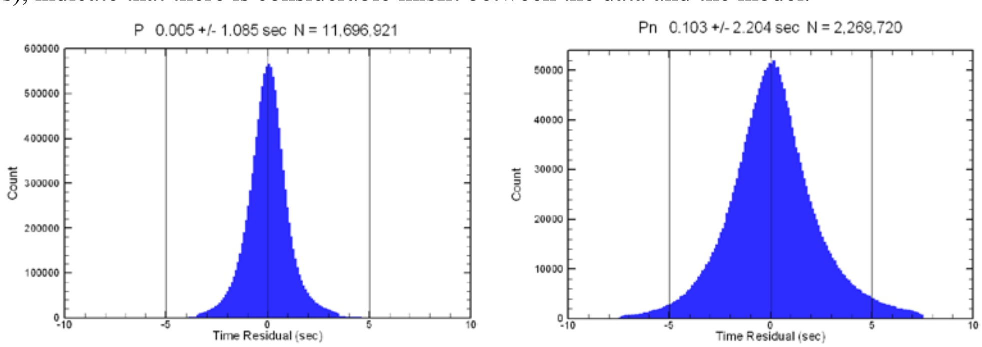

Figure 2. AK135 travel time residual histograms for (left) *P* and (right) *Pn* for all ray paths for the LANL GT catalog.

To establish whether this variance is due to inadequacy in the AK135 model or can be attributed entirely to random error in phase picks and/or in the source locations, we made maps of source-plotted residuals for individual stations. Many of these show coherent regionally varying trends in the residuals, suggesting that a large portion of the variance in the histograms is due to laterally varying velocity structure that is not modeled with the radial AK135 model. As a particularly compelling example, we show the map for station OBN in Russia in Figure 3. The most striking feature is a sharp transition from strongly negative residuals in the eastern Mediterranean and Middle East to positive residuals in the south central Asia. This pattern of residuals argues strongly for laterally varying Earth structure. Clearly there is no radially symmetric model that can fit these residuals.

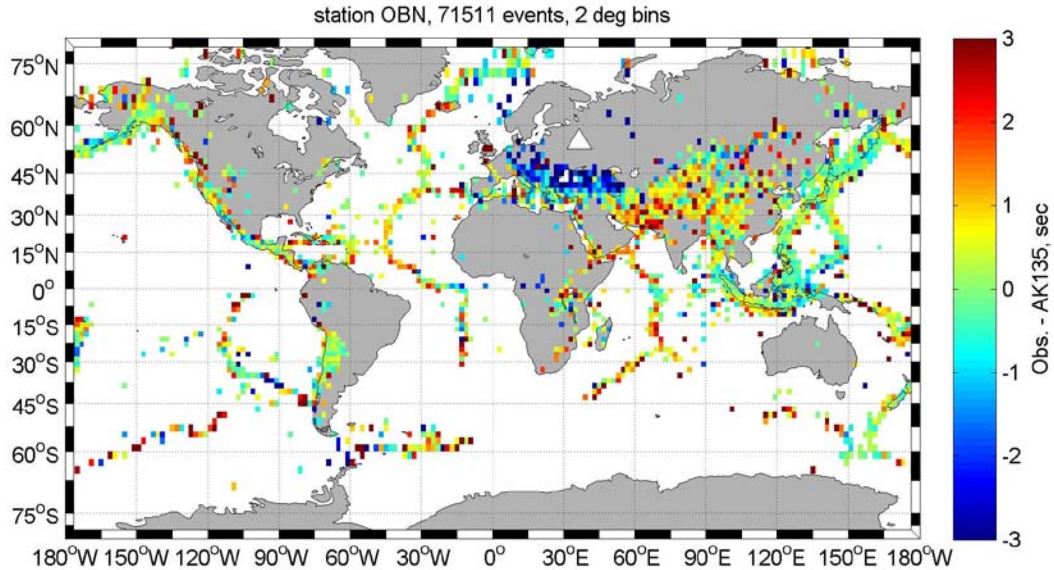

Figure 3. AK135 residuals for the EHB catalog plotted at source locations for station OBN (triangle).

MODEL REPRESENTATION AND INTERPOLATION

Our tomographic model is represented as a set of velocity/slowness nodes using the 3D grid approach described by Ballard et al. (2009). The model consists of different 1D layered radial profiles from the surface of the Earth to the core with all major velocity discontinuities represented, and each of these profiles is tied to one of a set of different geographic (latitude, longitude) nodes that are connected with an ordered set of nested tessellations. Progressively finer tessellation levels are generated by successive subdivision of the triangles of the first level tessellation, which is an icosahedron. Each subsequent level is assigned to progressively shallower discontinuities in the Earth, thereby achieving variable resolution in the radial dimension (Figure 4). An underlying assumption of this approach is that greater resolution is needed at shallower depths, which is in agreement with the conventional understanding of Earth structure. An important feature of this approach is the subdivision of triangles only occurs where the model needs additional resolution as defined both by ray hit count and by lateral variability in the velocity structure emerging during tomography, thus providing a variable scale model that is needed to reflect the uneven data sampling and the well-established laterally varying structure of the Earth. As described below (in the Tomographic Inversion section), we exploit this feature of the model representation to make the inversion process efficient and robust.

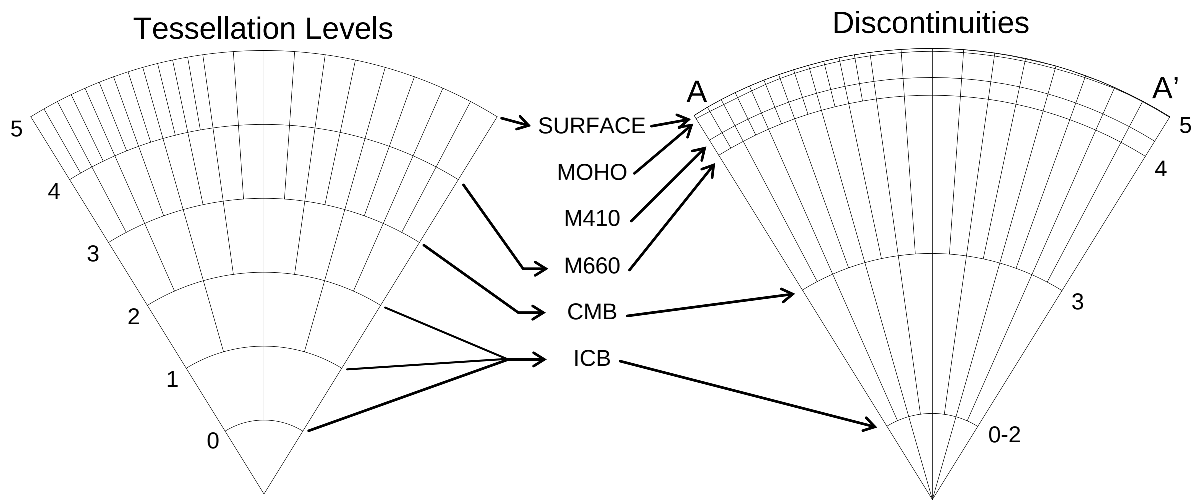

Figure 4. Progressive subdivision of icosahedron tessellation to level 6 and relationships to major velocity discontinuities.

Constructing ray paths and calculating travel times through the model requires obtaining velocity/slowness estimates at arbitrary positions in the mode. This is done through a two-stage interpolation process. First, using a walking triangle search (Lawson, 1977) we find the containing triangle in the surface tessellation for the latitude, longitude position of the point to be interpolated. This search starts at the coarsest tessellation level and proceeds to whatever finer tessellation levels are available for the position in question (i.e., we proceed to the next level if the containing triangle at the current level has been subdivided). The interpolation coefficients of the triangle nodes (the barycentric coordinates) are calculated as part of the triangle search and so they are already available when the containing triangle is found. Interpolation using these weights will produce a continuous result from one triangle to the next.

The second step is to find the layer within which the point of interest resides within the radial profiles corresponding to each of the nodes of the containing triangle. Once this is done, interpolation at the appropriate depth is completed for each of these profiles, and then these three interpolated values are combined using the containing triangle coefficients from the triangle search. In this manner, velocity at any point within our model can be quickly and accurately calculated. The process is described in much greater detail in Ballard et al. (2009).

RAY PATH AND TRAVEL TIME CALCULATIONS

Ray paths and corresponding travel times are calculated using our own version of the Um and Thurber (1987) ray pseudo-bender (Ballard et al., 2009). Our implementation includes a modified version of the Zhao and Lie (2004) method of handling discontinuities by implementing a 2D minimization algorithm that searches for the point on the velocity discontinuity surface where Snell's Law is satisfied. We reduce the likelihood that the pseudo-bending algorithm will return a local minimum by starting the ray calculation from several different starting rays. Specifically, interfaces are defined that include first order discontinuities plus additional interfaces at levels of the model where local minima are anticipated. Rays are computed that are constrained to bottom in each layer between these interfaces. The computed rays might be reflected off the top of the layer, turn within the layer, or diffract along the interfaces at the top and/or bottom of the layer. The computed ray that is seismologically valid and that has the shortest travel time is retained.

Once a ray path has been defined as a set of nodes via the pseudo-bender, including one node on each velocity discontinuity as part of the calculation to honor Snell's Law, we can calculate travel time by multiplying the inter-node segment lengths by the average slowness for the bounding nodes on each end of the segments and integrating over the entire path. Comparison with analytic tau-P calculations for a 3D version of the AK135 model demonstrate that our calculation is accurate to within a few hundredths of a second for distances from regional to teleseismic.

Though our pseudo-bender has been made as efficient as possible, calculating millions of 3D ray paths for each iteration of the tomography is time-consuming and could constrain the amount of model exploration that can be done. We overcome this problem by calculating the rays utilizing a distributed parallel computing system that we have developed based on the Java Parallel Processing Framework (JPPF), providing us with 150 processors, which have been shown to achieve an efficiency of better than 90% for the costly ray-tracing calculations, resulting in more than a 135x speedup over sequential algorithms. Utilizing this framework allows us to produce a model in less than a day that would otherwise have taken weeks, hence making it possible for us to investigate a much richer range of the total model space.

TOMOGRAPHIC INVERSION

The tomographic inversion is a standard iterative least squares solution to the linearized relation between travel time residual for each path and slowness adjustments to each node:

$$\Delta t_j = \sum_{i=1}^{N} w_i \Delta s_i \tag{1}$$

Where Δt_j is the travel time residual for the j^{th} path, Δs_i is the slowness adjustment for the ith node (of N), and w_i is the weight for the ith node. Though not immediately obvious, this travel time calculation is in fact path length multiplied by delta slowness, but the path lengths have been incorporated into the node weights.

To account for differing quality of data, for each ray path both the residual and ray path summation are inversely weighted by the estimated measurement error for each observation. Damping is applied due to the fact that the many parts of the model are under-constrained by the data, and also to control the size of the slowness/velocity adjustments to the model between iterations, which is essential for true non-linear 3D tomography where path geometries are updated after each iteration. Our previous research has demonstrated that allowing large velocity changes between iterations can lead to correspondingly large changes in ray paths which in turn lead to instability in the overall tomographic process. Increasing the damping is a simple and effective way to control this. We also invert for both receiver and source static corrections, but only introduce these parameters after a stable tomographic model has been achieved (we restart the inversion with the stable model as the starting model) to control the amount of signal that is mapped into the corrections.

Because 3D tomography is a strongly non-linear problem with many local minima, it is important to start with as good an initial model as possible. We use the mantle model of the recent 3D global tomographic model by Li et al. (2009) overlain by a simplified two-layer crustal model derived from Crust 2.0 (Bassin et al., 2000).

To achieve variable resolution in our model, we use a progressive tomographic refinement approach (Simmons et al., 2009). In our implementation, the tomography is accomplished with two major iterative loops. The outer loop is over a fixed locally adapted model that can change for each new outer iteration. The inner iteration is a standard tomography calculation that uses the current outer iteration model and its set of "active nodes" to calculate the total change in slowness. By active nodes we mean the set of nodes defined within the current outer iteration model whose velocity is modifiable. The set of active nodes may include previously adapted nodes (added in an earlier outer iteration) and new nodes that were locally added specifically for the current iteration. One of the goals of our research is to define how previously adapted nodes are included in the current outer iteration. They may be entirely excluded, or included with some amount of damping. Active nodes added several iterations before may be entirely excluded with those closer to the current generation incurring less and less damping the nearer their generation is to the current iteration.

The inner iteration is a standard tomography calculation using the active node set from the current outer iteration model. Inner iteration convergence is declared when slowness is no longer appreciably changing. What is different from standard tomography is how active nodes are included in tomography. In the new adaptive inner iteration only nodes that are still included from previous iterations (based on damping considerations described in the previous paragraph) and those new ones in the current iteration are allowed to participate in the tomography calculation. Additionally, only nodes from this set that have a non-zero ray hit count are allowed to be modified. Nodes with zero hit count are excluded, which is equivalent to processing with infinite damping. This ensures that nodes that cannot contribute to the solution are not inadvertently refined because of movement caused by insufficient damping. It also reduces the total column count in the sparse matrix ensuring faster solution times.

Once an inner iteration tomography run converges, the change in slowness is examined on each node to see which nodes exceed some predefined input tolerance. Those nodes that exceed the tolerance are provided to a model constructor, along with the current model, to build the new refined model for the next outer iteration. Only nodes that surpassed the input slowness change criteria are refined, or split, into a new set of nodes within the new model. The new model is then used for the next outer iteration.

The entire process converges when no slowness change on any active node exceeds the input tolerance or if a predefined input refinement level is reached. This avoids refining to extremely small grid spacing.

RAY REDUCTION

Because of the repeating nature of earthquakes, there is tremendous redundancy in some of the ray paths through the Earth in our data sets. Prior to inverting the data for our tomographic model we choose to remove this redundancy for two reasons. First, unless properly accounted for, redundant ray paths through the Earth will result in a higher weight in the tomographic inversion, with the significance of these rays being effectively multiplied by the number of redundant observations. For our purposes, a single path from a low GT event in area of monitoring interest is at least important (if not more) than multiple repeated paths from deep earthquakes in a subduction zone, so this is an important consideration. Second, each observation requires an additional ray bender calculation through the 3D model, a computationally expensive operation that we would prefer to make no more often than necessary. Thus, by eliminating redundancy we can correct the weighting problem as well as speed up the inversion.

We remove redundant rays by hierarchical agglomerative clustering (dendrograms) based on similarity of the ray paths. We base our similarity on the node weights calculated in Equation 1. These weights completely and uniquely characterize the ray path geometry. The distance in node space between rays *j* and *k* is:

$$\Delta_{j,k} = \sqrt{\sum_{i=1}^{N} (w_{i,j} - w_{i,k})^2} \tag{2}$$

where *N* is the number of nodes in the model, and most weights are zero for a given ray path. To convert this into a similarity that scales between 0 and 1 we normalize and subtract from 1:

$$S_{j,k} = 1 - \frac{\Delta_{j,k}}{\sqrt{\sum_{j=1}^{N} w_j^2 + \sum_{k=1}^{N} w_k^2}} \tag{3}$$

Calculating similarities between every pair of rays in such a large data set is impractical, but fortunately also unnecessary. We can make a few simplifying assumptions. First, because we expect the ray reduction to be primarily due to nearby sources, we only calculate similarities for rays to the same station. Second, we do not want to combine rays from different GT quality levels. Third, we know that if the distances between the sources are large enough, then the rays are not similar. We conservatively chose our threshold to be 3 degrees. Thus, for each station, for each GT level, we calculated similarities for all rays with sources less than 3 degrees apart.

Based on these similarities we then perform a separate agglomerative hierarchical clustering for each station/GT level set to indentify similar rays that can be replaced with single representative rays. Based on trial and error, we established that a threshold similarity of 0.7 will result in ray clusters whose source span has a radius no bigger than 0.5 degree, which we felt was appropriate given the target resolution of our model. The position of the source for the representative ray for each cluster is the mean position of the individual sources. To compute a travel time for the representative ray, we first calculate residuals for each original ray relative to AK135, then find the inverse variance (σ^2) weighted average of the M individual residual times (t_i) to come up with a representative ray residual time (σ^2):

$$t_{rr} = \frac{\sum_{i=1}^{M} \frac{t_i}{\sigma_i^2}}{\sum_{i=1}^{M} \frac{1}{\sigma_i^2}} \qquad (4)$$

which is then added back to the calculated AK135 time for the representative ray to get travel time. For the uncertainty associated with the representative time, we perform a similar inverse variance weighted average of the individual variances:

$$\sigma_{rr}^2 = \frac{\sum_{i=1}^{M} \frac{\sigma_i^2}{\sigma_i^2}}{\sum_{i=1}^{M} \frac{1}{\sigma_i^2}} = \frac{M}{\sum_{i=1}^{M} \frac{1}{\sigma_i^2}} \qquad (5)$$

Our overall reduction in ray paths is greater than 50%, providing a significant improvement in computational efficiency for generating the tomographic model.

CONCLUSIONS AND RECOMMENDATIONS

We have developed and implemented an approach to constructing a single, seamless, mutli-scale global 3D *P* velocity model that we will apply to our large data set from the combined EHB and LANL GT catalogs. We hope that the resultant model will be the first step towards an operationally capable 3D global model. In our poster presentation we will show the model itself and evaluate its effectiveness in reducing geographic variations in travel time residuals and in improving event locations.

ACKNOWLEDGEMENTS

We are grateful to Bob Engdahl, Rob van der Hilst, and Ray Buland for creating and making available the EHB catalog, which has greatly aided this and many other research efforts. We are also indebted to the staff at the ISC for providing ready access to the most recent version of the EHB catalog through their web site.

REFERENCES

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.* (in press).

Bassin, C., G. Laske, G. and Masters (2000). The current limits of resolution for surface wave tomography in North America (abstract), *EOS Trans AGU* 81: F897.

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: 722–743.

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

Lawson, C. L., 1977. *Software for C1 surface interpolation, in Mathematical Software III*, ed. J. Rice, Academic Press, New York.

Li, C., R. D. van der Hilst, E. R. Engdahl, S. Burdick (2008). A new global model for *P* wave speed variations in Earth's mantle, *Geochem. Geophys. & Geosystems* 9: 5, 21 pp.

Myers, S.C., S. Ballard, C. Rowe, G. Wagner, M. Antolik, W. S. Phillips, A. Ramirez, M. Begnaud, M. E. Pasyanos, D. Dodge, M. P. Flanagan, J. Dwyer, K. Hutchenson, D. Russell (2009). Gloal crust and upper mantle tomography for improved seismic location (abstract), presented at 2009 Seismol. Soc. Am. Annual Meeting.

Simmons, N.A., S.C. Myers, and A.L. Ramirez (2009). Multi-resolution seismic tomography of regional and mantle scale structures using tessellation-based node definitions (abstract), presented at 2009 SSA Annual Meeting.

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

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