Bulletin of the Seismological Society of America, Vol. 97, No. 3, pp. 804–825, June 2007, doi: 10.1785/0120060079

Ⓔ

Regional Travel-Time Uncertainty and Seismic Location Improvement Using a Three-Dimensional *a priori* Velocity Model

by Megan P. Flanagan, Stephen C. Myers, and Keith D. Koper

Abstract We demonstrate our ability to improve regional travel-time prediction and seismic event location accuracy using an *a priori* 3D velocity model of Western Eurasia and North Africa (WENA1.0). Travel-time residuals are assessed relative to the *iasp91* model for approximately 6000 *Pg, Pn*, and *P* arrivals, from seismic events having 2σ epicenter accuracy between 1 km and 25 km (ground truth 1 [GT1] and GT25, respectively), recorded at 39 stations throughout the model region. Ray paths range in length between 0° and 40° (local, regional, and near teleseismic) providing depth sounding that spans the crust and upper mantle. The dataset also provides representative geographic sampling across Eurasia and North Africa including aseismic areas. The WENA1.0 model markedly improves travel-time predictions for most stations with an average variance reduction of 29% for all ray paths from the GT25 events; when we consider GT5 and better events alone, the variance reduction is 49%. For location tests we use 196 geographically distributed GT5 and better events. In 134 cases (68% of the events), locations are improved, and average mislocation is reduced from 24.9 km to 17.7 km. We develop a travel-time uncertainty model that is used to calculate location coverage ellipses. The coverage ellipses for WENA1.0 are validated to be representative of epicenter error and are smaller than those for *iasp91* by 37%. We conclude that *a priori* models are directly applicable where data coverage limits tomographic and empirical approaches, and the development of the uncertainty model enables merging of *a priori* and data-driven approaches using Bayesian techniques.

Online material: Correction surfaces and histograms of travel-time residuals for 40 stations.

Introduction

Seismic-event location critically depends on the accuracy of predicted travel times and is often hampered by our insufficient knowledge of heterogeneous crust and mantle velocity structure. This absence of information causes systematic offsets in location estimates, which can be particularly large in aseismic regions, where our understanding of earth structure is severely limited. High-quality location estimates can be achieved by improving travel-time prediction with more accurate earth models. Event locations based on 1D radially symmetric earth models, such as *iasp91* (Kennett and Engdahl, 1991) are often biased, because they do not account for significant travel-time variations that result from crust and mantle heterogeneities. This is true in particular at regional distances (less than 25°) where the propagation paths of seismic phases often exhibit the strongest variability. Using 2D and 3D velocity models can improve travel-time prediction and hypocentral estimates by accounting for aspherical earth structure. The scenario of locating a small-magnitude event recorded by a sparse network of regional stations can benefit the greatest from accurate characterization of complex 3D crust and mantle structure. The objectives of this study are to use an *a priori* 3D velocity model and a high-quality dataset of well-located reference events (i.e., "ground truth," GT) to evaluate improvement in travel-time prediction provided by the model, develop a travel-time uncertainty model based on residuals relative to the 3D model, and assess our ability to improve both the location accuracy and location uncertainty of small-magnitude regionally recorded seismic events.

Previous work toward improving travel-time prediction and location has progressed through both empirical and model-based approaches. In data-rich regions, purely empirical approaches, such as tomographic inversion, can be used to compute 3D models. In regions of sparse coverage, a combined model and empirical approach is typically used. In regions where no data exist (i.e., aseismic), one must use an *a priori* approach to characterizing 3D earth structure. All these pursuits would benefit from a validated and well-characterized prior model as well as accurate, quantified uncertainties on the derived parameters.

Tomographic inversions often involve processing surface and/or body-wave observations to derive a 3D velocity model of the region of interest that can be validated with a separate dataset of GT events not used in developing the model itself (to avoid circularity). Recently, interest in using 3D velocity models to improve location estimates of both earthquakes and explosions in the Eurasian continent has been renewed, much of it motivated by treaty verification specifications of location uncertainty. These efforts represent two broad approaches: using relatively low-resolution global tomography models, or using higher-resolution local and regional scale velocity models. Global velocity models built from reported body-wave arrivals minimize a set of worldwide travel-time residuals and are primarily intended for prediction of teleseismic travel times. Such global 3D models can sometimes have limited use at regional to near-teleseismic distance where travel-time predictions are strongly influenced by highly variable crust and upper mantle structure. Most often this is observed at the *Pn–P* crossover distance as well as the *P* triplication distances where danger of phase misidentification is greatest. This is one reason that tomography has evolved toward producing higher resolution, continent–scale velocity models that can better predict regional phases such as *Pn*, *Sn*, and *Lg*. Global studies by Smith and Ekström (1996) and Antolik *et al.* (2001, 2003) established that 3D tomographic models of the whole mantle improve teleseismic-network location estimates relative to a 1D model when using teleseismic travel time data alone. Indeed, they reduce root-mean-square (rms) mislocation distance by 40% and 10% respectively. Global 3D models have also been used to demonstrate improvement in location when using both teleseismic *P* and regional *Pn* arrival times. Variance reductions of up to 50% have been achieved for relocations of GT0 to GT10 events by Yang *et al.* (2004) using both teleseismic and regional arrival times and by Ritzwoller *et al.* (2003) using regional phases alone. Still these global models are limited in resolution to about 500 km laterally. Higher-resolution (2° to 3°) regional scale tomographic models have also been developed (e.g., Villaseñor *et al.*, 2001; Reiter *et al.* 2005) and shown to improve the location of small regionally recorded events yielding approximately 30% rms mislocation. Alternatively, when too few data exist to perform tomographic inversion, other techniques can be pursued. A common approach has been to use a 1D velocity model and apply empirically based source-specific station corrections or "correction surfaces" to remove bias due to 3D earth structure. Where high-quality recordings of GT events of known location are available, the correlation structure of observed arrival times can be investigated with a spatial interpolation method such as kriging (e.g., Myers and Schultz, 2000) to then compute a correction surface relative to some baseline 1D model (e.g., *iasp91*).

Where recordings of regional GT events are too sparse for interpolation, travel-time calibration can be done by regionalization where the propagation of seismic phases within certain subregions is characterized separately. Such path-dependent characterization is performed by associating different types of crustal and mantle tectonics with their corresponding travel-time patterns (Tralli and Johnson, 1986) and developing the corresponding travel-time curves (Morozov *et al.*, 2005) that can be further combined using empirical rules (Bondár *et al.*, 2001; Yang *et al.*, 2001b). Armbruster *et al.* (2002) found that a combination of the empirical approach of Myers and Schultz (2000) and the regionalization approach (Yang *et al*, 2001a,b; Murphy *et al.*, 2005; Richards *et al.*, 2003) provided the best technique for decreasing mislocation in eastern Asia.

Empirical kriged corrections and other model optimization methods are well suited to improve travel-time prediction in areas where calibration events are available, and, as with tomography, improvement is expected in well-resolved, well-constrained portions of the model. In seismic monitoring the goal is to monitor the entire earth, including vast aseismic regions devoid of GT events. Therefore, we must have confidence in travel-time predictions for all potential ray paths, including paths along which existing tomographic models are poorly resolved. It is these poorly resolved areas that motivate us to explore the use of 3D *a priori* models to characterize 3D earth structure.

The approach of developing an *a priori* regionalized velocity model has been applied at scales that are both global, such as CRUST5.1 (Mooney *et al.*, 1998) or 3SMAC (Nataf and Ricard, 1996) and regional, such as EurID (Du *et al.*, 1998), WINPAK3D (Johnson and Vincent, 2002), CEA (Begnaud *et al.*, 2001), and WENA1.0 (Pasyanos *et al.*, 2004). All these models are compiled by synthesizing geophysical information from published literature that is relevant to the velocity structure, geology, seismicity, and tectonics throughout the region. We choose to test the WENA1.0 model as it covers the largest geographical extent of our study area, Western Eurasia and North Africa (WENA), and has been evaluated comprehensively for its ability to predict several diverse geophysical observations. Therefore, we believe it represents our best *a priori* assessment of geophysical structure of this area to date.

In the present study, we evaluate WENA1.0 for both travel-time prediction and regional seismic-location improvement. We present an end-to-end validation that includes comparison of model-predicted travel times with observed travel times for 39 stations throughout the WENA region, and we relocate a large set of 196 reference events with known location accuracy of 5 km and better (GT5). We also derive azimuthally invariant, distance-dependent, travel-time uncertainty models for WENA1.0 and *iasp91*, which we use to estimate coverage ellipses for the relocations. A large high-quality dataset of reference GT events in Europe, Africa, and Eurasia (Ruppert *et al.*, 2005) that comprises approximately 4800 GT25 and better locations is employed. This dataset provides more than 30,000 first *P*-wave arrivals with which to evaluate our WENA1.0 model. Evaluation of travel-time prediction of the GT25 arrivals at individual stations and the event relocation tests performed with the GT5 events are described. Improvement in travel-time prediction of regional *Pg*, *Pn*, and *P* waves as provided by the 3D WENA1.0 model is quantified, and improvement in both location and uncertainty estimates (error ellipses) are assessed. Because the WENA1.0 model construction is consistent in both data-rich and data-poor regions, we believe this 3D velocity model and distance-dependent uncertainty model are applicable in both seismic and aseismic regions and that this procedure may serve as template for validating any 3D velocity model.

Western Eurasia North Africa 3D Velocity Model (WENA1.0)

The region of interest, Western Eurasia and North Africa, covers a diverse set of continental margins, mountain ranges, and deep sedimentary basins all having complex tectonic history. This heterogeneity presents a significant challenge to accurately predict travel times in this region without a detailed 3D earth model. Realistic modeling of travel times and amplitudes of regional seismic phases requires an accurate high-resolution model of complicated crust and upper mantle structure. Global tomographic velocity models typically have low resolution on the order of several hundred kilometers, too coarse to adequately model short regional paths. Alternatively, regional 3D models image smaller scale structures (some having 1° ~ 2° resolution) and contain greater amounts of heterogeneity but are often constructed from, and optimized to, a single data type such as body-wave travel times or surface waves. Both types of models suffer from uneven data coverage, smoothing constraints, vertical smearing, and inadequate resolution of the fine-scale heterogeneity near the surface of the Earth.

Evaluating a continent-scale model where travel-time prediction uncertainty is characterized across the whole model is our goal. We select the *a priori* 3D WENA1.0 model of Pasyanos *et al.* (2004), built for its use in estimating a multitude of regional to near-teleseismic distance geophysical data. *A priori* models offer some advantages over those derived from tomographic inversion by having uniform coverage of an entire geographic area including both tectonically active and inactive areas as well as retaining sharp features that are often smeared out by smoothing constraints necessary for stable inversions. The construction of the *a priori* WENA1.0 geophysical model was partly motivated by research efforts in seismic-event detection, location, and discrimination in a regional monitoring context. The model can serve as a reference structure for methods based on the interpolation of empirical measurements, such as kriging (Schultz *et al.*, 1998; Pasyanos, 2000), and may eventually serve as an initial starting model for a multidataset tomographic inversion for structure of the Eurasian continent. A complete description of the assembly and testing of the WENA1.0 model can be found in Pasyanos *et al.* (2004), and we provide a brief summary here.

WENA1.0 is an amalgamation of geophysical and geological studies constituting a self-consistent 3D earth model for the crust and upper mantle. It is derived from a compilation of independent sediment, crust, and upper mantle velocity models, topography, seismicity, phase blockages, *Pn* tomography, surface-wave dispersion, and receiver function measurements that in turn are taken from extensive literature reviews where data are available (see Pasyanos *et al.*, 2004, and references therein). However, large regions of WENA are essentially aseismic and poorly sampled geophysically. In these regions, extrapolation of properties such as seismic velocities from data-rich areas to analogous data-poor regions of similar tectonic character and geologic history is used. This allows uniform specification of the model parameters over the entire region. Using a consistent methodology to build the model across the broad WENA region allows us to use data from the well-sampled regions for evaluation of the entire model as we present here using *P*-wave travel-time data. A consistent approach in model construction is important, as applicability of an uncertainty model in aseismic areas depends on this assumption.

The model spans 0° to 85° latitude and −20° to 75° longitude and has 1° resolution (Fig. 1). The crustal portion is divided into 45 distinct geophysical regions based on seismicity patterns, tectonic history, crustal thickness, and crustal and *Pn* velocities. Each 1° cell is composed of 24 distinct layers: a water layer, 3 sediment layers, 3 crustal layers, an uppermost mantle *Pn* layer, and 16 mantle layers extending to 700 km depth (below which we use *iasp91*). The model is primarily based on *CRUST5.1* from Mooney *et al*, (1998), the high-resolution 1° sediment model of Laske and Masters (1997), and the 2° 3SMAC mantle model of Nataf and Ricard (1996). The 3SMAC model is appealing, in particular, because it is a true *a priori* model, not derived from any seismic data. Rather, it employs chemical and thermodynamic modeling to estimate a 3D model of density, *P* and *S* velocities, and attenuation independently. 3SMAC fits well with the geologic, crustal regionalization approach by including subducting slabs and hotspots, and it allows for topography on the 410- and 660-km discontinuities in the upper mantle. While WENA1.0 specifies *P*- and *S*-wave velocity, density, Q_P and Q_S in each layer, we address only the *P*-wave travel-time predictions in this study.

Because our emphasis is on predicting regional body-wave travel times, we need relatively fine-scale regions to account for complex, short-wavelength 3D structure. Examples of WENA1.0 depth to Moho and *P*-wave velocity at different depths are also shown in Figure 1. These horizontal cross sections highlight the nature of geophysical heterogeneities such as continental roots, spreading ridges, mountain belts, collision zones, and stable platforms. The slice at 15 km depth represents continental midcrustal *P*-wave velocity whereas the *Pn* velocity represents lithospheric struc-

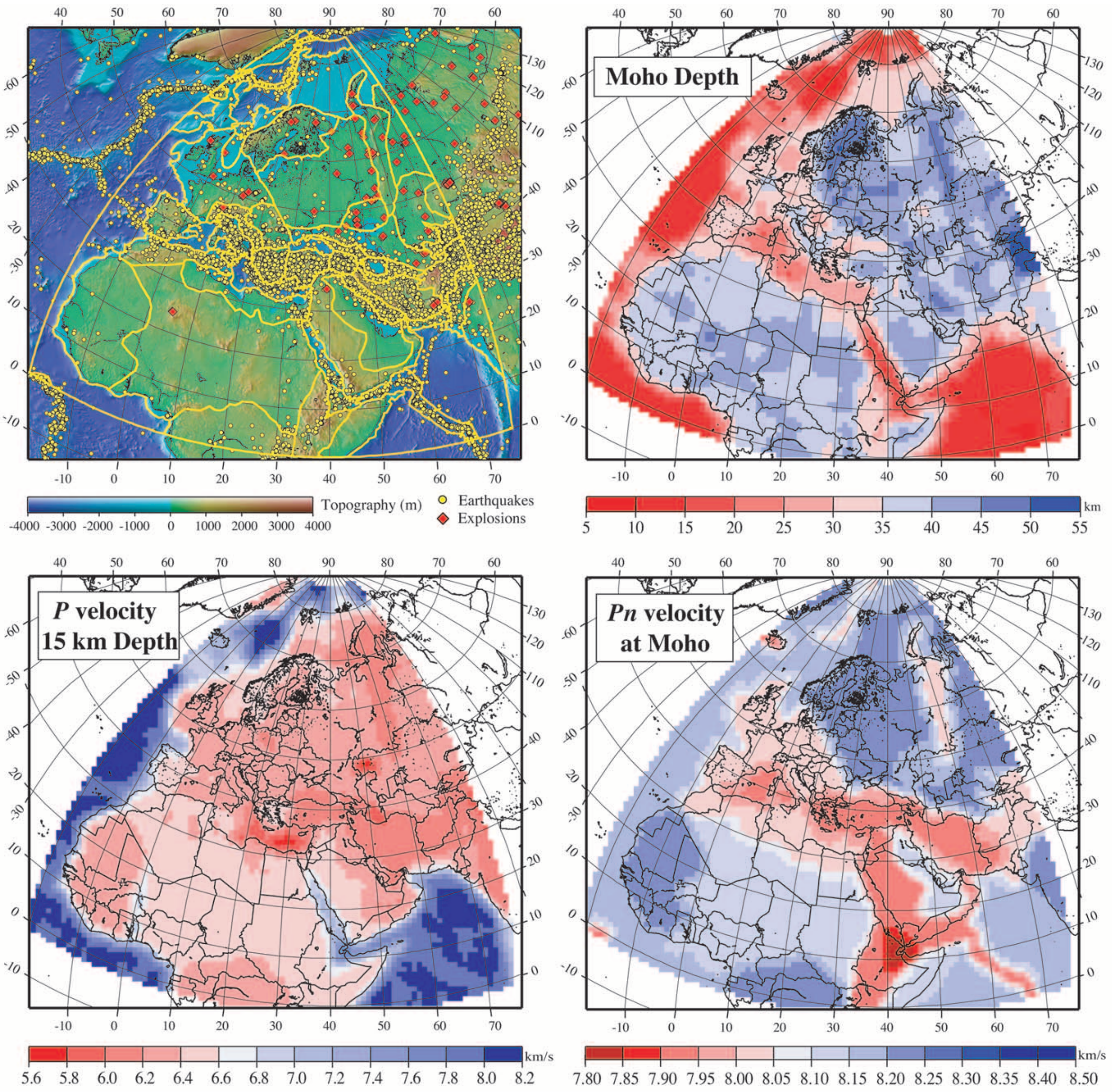

Figure 1. Azimuthal equal area projection map of the *a priori* WENA1.0 geophysical model compiled by Pasyanos *et al.* (2004) (top left). The crustal regionalization (shown by yellow lines) comprises 45 geologic areas having distinct *P*- and S-wave velocities, density, attenuation, and layer thicknesses. High-quality seismic-event location of GT25 or better, and having focal depth less than 30 km, are plotted for shallow earthquakes (yellow circles) and explosions (red diamonds). Topographic shading illustrates the diverse geometries and major tectonic features in Western Eurasia and North Africa. Example geophysical features of the WENA1.0 model are shown: depth to Moho (top right), horizontal cross section of *P*-wave velocity at 15 km depth (lower left), and *Pn* velocity directly beneath the Moho (lower right).

ture directly beneath the Moho extending to about 80–100 km depth. One of the most prominent low-velocity regions is located in the Middle East, extending from Turkey to Iran and into western Afganistan and Pakistan (the Turkish Iranian Plateau) where low *Pn* velocity (~7.85 km/sec) is prevalent.

The variety of tectonic and thermal history, deformation style, and heterogeneous crustal types (from young oceanic to Archean craton) found throughout WENA underscores the clear need for using a 3D model to accurately predict travel times in this region. A key issue with any *a priori* model remains evaluating its accuracy and uncertainty. As

WENA1.0 does not provide travel-time prediction uncertainties, we develop our own uncertainty model based on the variance structure of the travel-time residuals.

Validation Datasets

Improvement in both travel-time prediction and location performance provided by WENA1.0 is assessed by using well-recorded events with accurately known locations. We recognize the need for a high-quality dataset of reference events and invest considerable effort in collecting and evaluating such events (Ruppert *et al.*, 2005). Recent efforts of GT data collection and streamlined criteria regarding the minimum azimuthal gap and the number of stations reporting (Bondár *et al.*, 2004) have facilitated model validations by allowing researchers to quantify large GT datasets. The location of each validation event either meets the network-coverage accuracy criteria of Bondár *et al.* (2004) or is constrained by nonseismic means (e.g., explosions with known sources). This ground truth dataset was assembled over many years, and it is based on locations and arrival times reported from regional bulletins, global bulletins such as the International Seismological Centre (ISC), relocation studies using multiple phases including depth phases for single-event locations (such as the EHB catalog of Engdahl *et al.* [1998]), cluster analyses (Engdahl and Bergman, 2001), local seismic network studies such as aftershocks, and our own analyst picks and relocations. We follow the nomenclature of Bondár *et al.* (2004) to categorize the epicenter accuracy of each validation event in which an accuracy level in kilometers is appended to the GT prefix. For example, GT5 indicates an epicenter mislocation of 5 km or less.

In addition to our GT5 and GT25 locations meeting the Bondár *et al.* (2004) criteria, we also make use of GT1 explosions with published locations, GT3 explosions located relative to neighboring explosions with known location, and GT5 earthquakes that are constrained with nonseismic means, such as satellite imagery of ground displacement (e.g., InSAR). There are 45 GT1 (Sultanov *et al.*, 1999), 6 GT3 (Lilwall and Marshall, 1986), 63 GT5, and 2260 GT25 events (Fig. 2, left). Although depth is not explicitly constrained by the GT criteria, the validation events are well recorded, and we exclude fixed-depth events and select only events with depth determinations less than 30 km. The GT25 and better events are used to broadly evaluate travel-time performance while the GT5 and better events are used to evaluate location performance.

Travel times are examined on a station-by-station basis, and our goal is to collect a large amount of data with which to perform robust statistics while maintaining even sampling of the model region. Arrival times of *Pg*, *Pn*, and *P* phases at 39 stations are selected for model evaluation. We chose stations with a long recording history (i.e., a large number of arrivals) and with the largest azimuthal range of events, thus providing the fullest sampling of the WENA1.0 model possible. Typically, stations having less than 100 reported arrivals were not used unless they provided needed ray sampling near edges of the model (e.g., KAD, KHE), or they reported an arrival for a high-quality GT event which is used in the relocation test. We also chose additional stations (e.g., MBO, KUK, MFP) surrounding the southwest part of WENA because these considerably increase the coverage across North Africa, which suffers from a dearth of recording stations (see Fig. 2, left, for ray coverage).

In contrast to previous studies that include only regional *Pn* and *P* arrivals at ranges less than 20° (Ritzwoller *et al.*, 2003) or separate *Pn* ($<15°$) and *P* (25–90°) arrivals (Yang *et al.*, 2004), we use all first-arriving *Pg*, *Pn*, and *P* phases out to 50° epicentral range to include both the *P* and *Pn* crossover as well as the *P* triplications from the 410 and 660 discontinuities. Although picking errors can be quite large at these distances, we believe this represents the most realistic scenario for regional recordings of a moderate event. Using arrivals throughout this distance range sounds the crust, upper mantle, and into the lower mantle (0 km to $\sim$1000 km depth). Therefore, the whole model is evaluated, allowing us to develop one uncertainty model for all of WENA1.0.

Measurement of arrival time (i.e., picking error) dominates the travel-time residual in many cases; we cull the dataset to remove both blatant outliers and residuals that are inconsistent with their neighbors. Our procedure involves computing residuals relative to the 1D *iasp91* model, then processing them for each station using the following steps: (1) outlier removal of large residuals, (2) variogram modeling to determine spatial correlation of residuals and estimate random picking uncertainty, (3) cross-validation outlier removal. The first step is to discard residuals greater than ± 10 sec. Whereas other studies have more stringent outlier requirements with cutoffs as low as ± 3 sec (Ritzwoller *et al.*, 2003) it is our experience that a residual of 6 to 8 sec is not uncommon in Western Eurasia. Because we have numerous high-quality picks (i.e., high signal to noise, made by an analyst), we believe many large residuals may result from the high degree of velocity heterogeneity that exists throughout WENA and not solely from picking error.

Second, we use a nonstationary variogram-modeling technique to estimate the random picking error of the arrivals at each station. The variogram is a statistical parameter based on the covariance of residuals as a function of the distance between the epicenters; Myers and Schultz (2000) give a full description of the technique, and we summarize the salient points here. The variance of travel-time residuals relative to *iasp91* is plotted against the distance between the observations as shown in Figure 3 at station APA. The variances are averaged in 1° bins to reduce scatter, and an analytic curve is fit to the data variogram to determine four fundamental parameters: the *sill* is the background variance for the whole set of arrivals at that station (the value at which the variogram levels out), the *range* is the distance at which correlation between points becomes zero (the distance at which the sill is reached), the *nugget* is the value of the

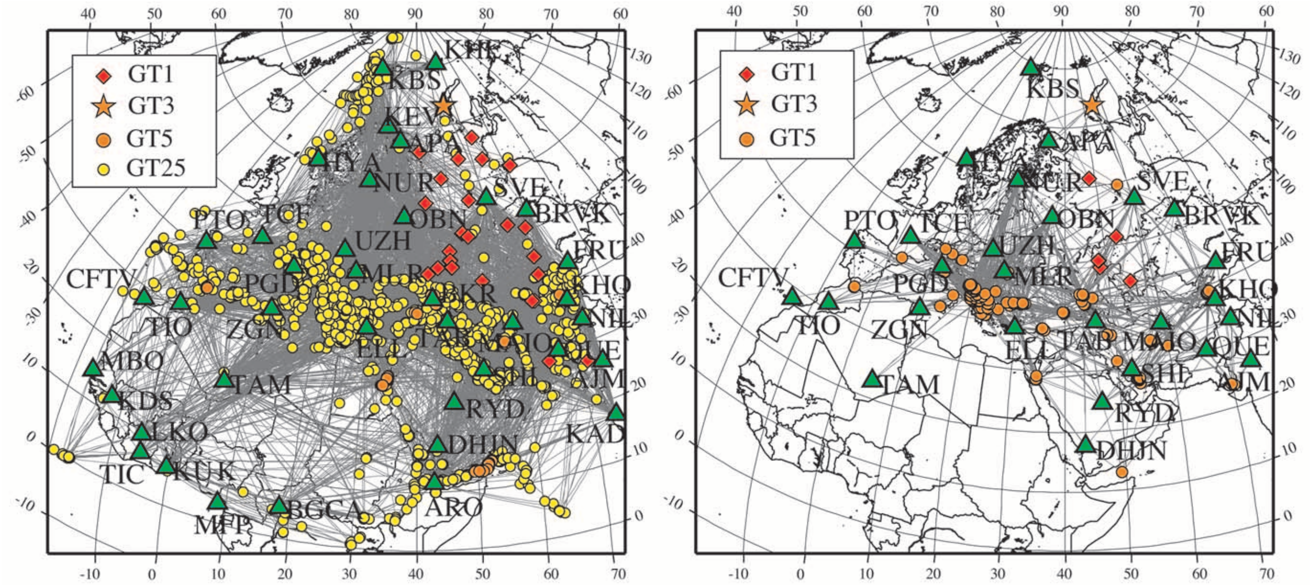

Figure 2. Maps of ray path coverage for the WENA region showing the GT25 travel-time validation dataset (left) and the GT5 location validation dataset (right). Paths (gray lines) of regional to near-teleseismic *Pg, P*n, and *P* phases from the GT25 or better events to 39 stations (green triangles) show good coverage of the entire model region. A total 2367 events (GT1, 45; GT3, 6; GT5, 63; and GT25, 2260) generate 5767 arrivals, to stations ranging from 0° to 50° epicentral distance, for travel-time validation. A network of 27 stations is used with the GT5 events for location validation. There are 196 events having accuracy: GT1(10), GT3 (2), and GT5 (184). The GT1 events are PNEs taken from the catalog of explosions scattered across the former Soviet Union and reported to have accuracy of 1 km (Sultanov *et al.*, 1999). The GT3 events are explosions at the Novaya Zemlya test site, and several of the GT5 events are generated from the cluster analyses of Engdahl and Bergman (2001).

variogram at zero distance, which is a measure of uncorrelated error (i.e., a measure of average pick uncertainty), and the *shape* of the correlation curve as a function of distance between points. The overall shape of the variogram for almost all stations is as expected in that it has a minimum at zero distance (correlation is maximum) and increases as the points become farther apart (correlation decreases). The only cases where this was not true were for stations having too few data to compute a statistically meaningful variogram (e.g., MFP, FUL). Nugget, range, and sill values varied across all 39 stations. Some had quite small nuggets of 2 sec (e.g., NIL, HYA, BGCA) but others had large nuggets up to 6 sec (e.g., TAB, UZH), and values of 3 to 4 sec were most common. The sill values were typically 2 to 4 sec higher than the nuggets and ranged from 4 to 8 sec. The range values showed the most variety among all the stations and extended from 7° to 15°. After performing the nonstationary variogram modeling and determining the nugget for each station we assigned picking uncertainty to all the data at that station as the square root of the nugget (many arrivals are reported from a bulletin and no uncertainty on the pick time is given) unless an arrival already had an uncertainty assigned manually by an analyst, in which case we kept the analyst's pick uncertainty.

Third, a leave-one-out cross-validation kriging procedure, described in Myers and Schultz (2000), is used to check the consistency of each arrival-time observation. Cross-validation entails leaving one observation out and using the rest of the observations to predict the value at the left-out point. We produce a statistical distribution of the misfit between observed and predicted values by repeating this procedure for each point in the dataset. Outliers are then removed at the 2σ level from this distribution, which reduces the original dataset from over 30,000 to 24,350 arrivals. The final ray path coverage of this validation dataset is shown in Figure 2. Ray path sampling is excellent over much of the region, and our new dataset provides considerable improvement in coverage of aseismic regions, such as North Africa, while reducing oversampling in seismically active regions. (Note that after cross-validation these 24,350 arrivals are further averaged in geographic bins with a declustering algorithm described in the next section, and the ray paths shown in Figure 2 represent our final data coverage.) This procedure is more appropriate than outlier removal based on statistics of the whole population because local trends and biases at individual stations are taken into consideration. Culling based on location accuracy and arrival-time regularity produces a self-consistent dataset of accurate travel-time measurements.

The seismic events used to test location improvement

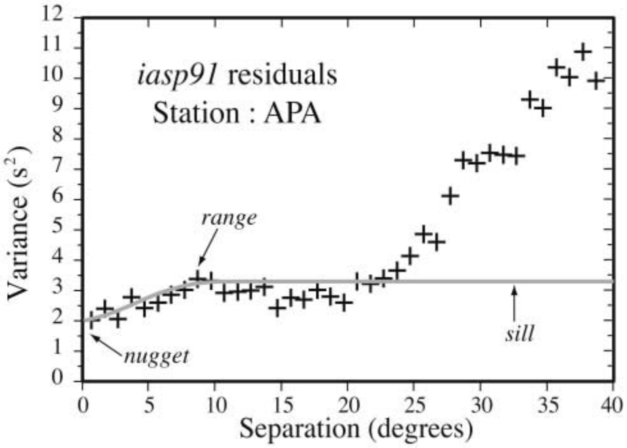

Figure 3. Example of a variogram computed for *P* residuals (relative to *iasp91*) recorded at station APA. Crosses represent the data variances computed in 1° bins of distance between events and the solid line is the variogram model determined from curve fitting. The *sill* is the background variance of the data, the *range* is the distance at which correlation between points is zero, and the *nugget* is the covariance of colocated points (at zero distance). Note the variance does not reach zero for colocated points indicating the data are not perfectly correlated due to errors in determining the arrival time. The square root of the nugget represents the random picking error (in seconds) for the GT25 dataset at this station.

provided by the 3D WENA1.0 model are a subset of the GT5 and better data described previously. For the GT5 events to be used in the relocation test we required there to be at least five defining phases, recording stations at less than 25° epicentral distance, and no travel-time residuals (relative to *iasp91*) greater than ±7 sec. The distribution of the GT5 events is shown in Figure 2 along with the stations and the corresponding ray paths; most events have both *Pn* and *P* arrivals recorded at a variety of distances and azimuths.

Declustering of GT25 Arrival-Time Data

Before computing goodness-of-fit statistics for the GT25 dataset, the residuals are declustered in an effort to further remove outliers and to minimize the redundancy of certain paths through the model space. Stations surrounding the high seismicity in the Mediterranean Sea, Turkey, Iran, and Tibet (TCF, UZH, MLR, SHI, and FRU) have numerous data and cause a disproportionate sampling of the center of the model relative to the edges (Fig. 2). We apply an epicenter-declustering algorithm (Myers, 2001) to combine closely spaced calibration events into one representative point. This entails spatial filtering to consolidate heavily clustered points into small collections of representative points with their associated uncertainties. In effect, this technique reduces spatial redundancy, which can bias summary statistics, in the extremely large set of arrival times. This declustering routine differs from the "summary ray" approach used in tomography, because it is not a simple average; instead, a weighted average as a function of GT level and pick error is made to estimate the resulting residual and uncertainty.

Declustering examines event information on a per station basis beginning by forming a 3D grid covering the entire range of available residuals for the given station. Grid resolution can be varied, and we set the latitude and longitude cell increments at 0.5° and partitioned the depth range into intervals having end points at 0, 10, 20, and 30 km. Each cell (i.e., "cluster") is assumed to consist of points possessing a high degree of association and a moderate propensity for extreme characteristics (i.e., the chance of a moderately sized cluster containing one or more outliers is assumed nonnegligible). Measurements for events falling into a single cell are replaced by a robust summary event, associated uncertainty, and variability statistics, namely, the mean and variance of the observed travel-time residuals. This summary event is positioned at the location of the observed event lying closest to the centroid of the cell. Hence, for each station, the output of the declustering procedure is a list of vectors containing the preceding statistics for each occupied cell. This process reduces the dataset from more than 24,350 to 5767 paths as shown in Figure 2.

3D Finite-Difference Travel-Time Calculations

The task of calculating travel times in 3D earth models has been approached with many algorithms including "exact" shooting and ray-bending methods along with approximate methods, which include perturbation approaches and "grid-based" methods such as finite difference. Our 3D WENA1.0 model contains sharp vertical and horizontal velocity gradients as well as thin structural layers that complicate the accurate computation of travel times. Techniques based on ray-perturbation theory or the graph method do not work well when the velocity anomalies are large relative to the reference model (Pulliam and Snieder, 1996; Bijwaard and Spakman, 1999); ray shooting and finite-difference algorithms can be more computationally expensive but provide accurate results in the presence of large velocity gradients that can produce different propagation modes (e.g., refracted and diffracted waves) in addition to direct phases (Podvin and Lecomte, 1991; Hole and Zelt, 1995; Cerveny, 2001). Finite-difference methods are efficient, yet have the limitation of only predicting the first arrival time. Two-dimensional ray shooting methods usually produce reliable results in the presence of large velocity gradients and predict phases in addition to the first arrival, but they do not account for off-great-circle propagation (see Villaseñor *et al.* [2000] and Cerveny [2001] for excellent reviews).

As we are most concerned with predicting first arrivals to use in our location algorithm, and WENA1.0 has numerous sharp velocity contrasts, we choose the finite-difference approach. We employ an algorithm originally developed by Vidale (1988, 1990) and further refined by Hole and Zelt

(1995) that applies a finite-difference (FD) approximation to the Eikonal equation to compute first-arrival travel times through regularly gridded velocity structures. The Vidale algorithm is computationally efficient but can become unstable in the presence of large, sharp velocity contrasts where wavefronts are critically refracted. These instabilities are eliminated by modifications made by Hole and Zelt (1995), and the improved code now accurately predicts head waves traveling along the boundaries of grid cells, reflected phases, and any velocity contrast with a ratio up to 10:1 even if it is vertically oriented. This technique propagates wavefronts (rather than rays) radially outward from a point source using each grid node as a secondary source for each successive grid node (Huygens' principle). Because the travel times are computed for every grid point in the volume, the code is much faster than tracing rays from a single event to a large number of receivers. In effect, we use the principle of reciprocity by placing the source point at the seismic station location and interpolate the travel times computed at each grid node to match the exact earthquake locations.

The Hole and Zelt code can read and compute times through 3D volumes of constant-velocity cubic cells. However, we modified the code with a Cartesian-to-spherical coordinate transform to convert the source and receiver locations. This modification is necessary because we need to compute travel times out to regional and teleseismic distances (0° to 50°), while the original FD algorithm operates in Cartesian coordinates, which are accurate only for distances up to 300 km in the earth. An earth-flattening transformation (Muller, 1971) is commonly used to account for the curvature of the earth; however, this is strictly valid for 1D velocity models only. Application of the spherical coordinate transformation scheme involves parameterizing a spherical surface inside the Cartesian volume of grid nodes (Fig. 4). Here, the source and receiver positions and velocity model are known in spherical coordinates, and they are transformed into the Cartesian system as input to the FD code. The FD operators remain unchanged (i.e., computation is done in Cartesian coordinates) and the output travel-time volume is converted from Cartesian to spherical coordinates as well. Although this approach means that we "waste" some nodes in the upper corners of the box where the sphere does not exist, it is more accurate than earth flattening.

The FD algorithm also requires our velocity model to be discretized into cubic constant slowness cells on a Cartesian grid; this requires fine sampling, by trilinear interpolation, of our 1° by 1° velocity model. This procedure results in the loss of some resolution with depth (e.g., some sediment layers may be thinner than the grid size chosen) and an oversampling of lateral structure as the velocities in the shallower part of the model vary more rapidly with depth than they do laterally. Our fully 3D code is run in a volume of dimensions up to 80° by 60° laterally (depending on the station location and distribution of events) and 1500 to 2200 km deep with a grid spacing of 5 km; the grid spacing is determined empirically as a trade-off between the accuracy of the travel-time prediction and computer memory limitations (see following discussion). The output from the FD code is a 3D volume of the first *P*-wave arrival time at each node. We perform two FD runs for each station of interest, one using WENA1.0 and one using the global reference model *iasp91*, and compute travel-time residuals between each of these volumes for the same source–station paths. In this way, we can minimize numerical dispersion effects of the FD code because any dispersion will be present in both calculations and hence minimized in the difference. From the output volume of travel times we can extract both individual arrival times for any source–station path as well as slices through the volume at discrete depths of 0, 10, 20, 30, 40, and 50 km which are used as 3D travel-time tables or "correction surfaces" in the location algorithm.

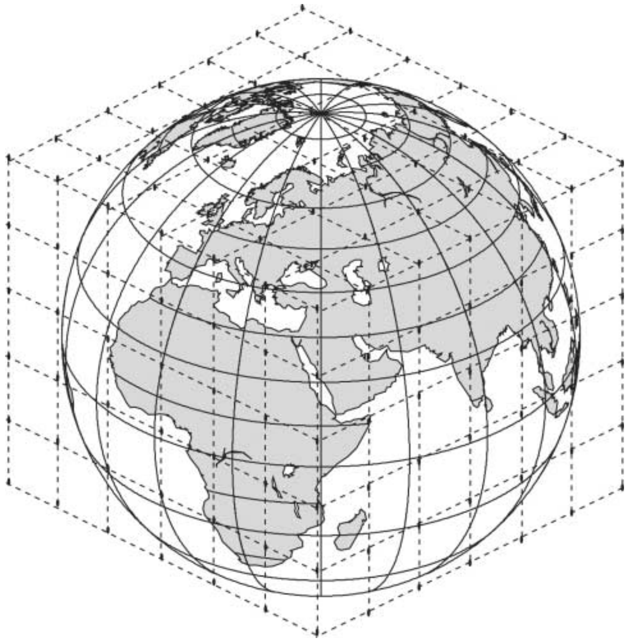

Figure 4. The "sphere in a box" parameterization for the FD code is shown in this cartoon. The source, station, and velocity models are known in spherical coordinates but the FD code operates on a Cartesian grid. We apply a spherical-to-Cartesian coordinate transform to the input and output of the code to ensure accurate travel-time calculations for a spherical earth.

Errors of the FD code arise from grid-boundary effects and the discrete approximation of the differencing algorithm (Vidale, 1988). The FD code also has some computational limitations, as the entire velocity (or slowness) grid must be stored in computer memory throughout the calculation. Thus, there is a trade-off between the size of the grid spacing and the overall size of the volume we wish to compute. We quantify the accuracy of our 3D FD method by computing the average rms residual of the FD predicted time and theoretical travel-time predictions from the 1D *iasp91* velocity model over a distance range of 0° to 40°. The theoretical times are computed using a ray tracer that employs the single-valued *tau-p* formulation (Crotwell *et al.*, 1999) similar to that of Buland and Chapman (1983). The 3D FD code is run through *iasp91* for several tests using incrementally smaller grid sizes and comparing the resulting FD travel times with the theoretical *tau-p* times. We find that a grid spacing of 10 km is computationally fast but results in average rms errors of 0.56 sec, whereas a spacing of 2 km is computationally prohibitive but produces small rms errors of about 0.21 sec. A grid spacing of 5 km is computationally acceptable and provides an average accuracy of 0.23 sec (with a maximum of 0.60 sec at 1.66° epicentral distance near the *Pg-Pn* crossover); therefore, we use this grid spacing for all subsequent computations (Fig. 5).

The accuracy of the FD calculations also varies with azimuth and distance from the source (station). Errors appear to fluctuate the greatest between 0 and 300 km from source then decrease and remain constant at a value of about 0.1 sec out to 13–15° then increase to about 0.25–0.3 sec and remain constant out to 40° distance. With regard to azimuth there appears to be a "cross" shape pattern to the errors where they are similar at all azimuths except 45°, 135°, 225°, and 315°, where they appear to be smaller and of opposite sign (Fig. 5). The magnitude and pattern of the numerical errors we observe are similar to those of the Podvin and Lecomte (1991) FD code which was used by Johnson and Vincent (2002), who report errors of 0.1 to 0.3 sec on a 5-km grid and Murphy *et al.* (2005) who report errors of 0.3 to 0.5 sec on a 5-km grid. We concur with these previous studies that numerical errors of this small size are likely to be less than the picking error associated with any given arrival time.

Although a 1D model test is useful for checking the accuracy of our FD technique, it does not address the accuracy of travel-time calculation in 3D media. Villaseñor *et al.* (2000) performed tests addressing this issue. They compared travel times computed from a ray shooting method with those from a FD method through a 2D velocity structure. They found differences between 0.5 and 2.0 sec, larger than the errors found in our test, and attributed these differences to the failure of the ray shooting method to find the minimum travel time ray path. In a different test of accuracy, Villaseñor *et al.* (2000) compared travel times computed through a 2D velocity structure with those through a 3D velocity structure using the 3D FD method of Podvin and Lecomte (1991). They found travel-time differences are small (less than 0.1 sec) and that off-great-circle propagation was not a significant factor. They and Ritzwoller *et al.* (2003) subsequently computed all their correction surfaces using the 2D version of that FD code. We choose to do the full 3D FD calculation as our model has considerably more structure than the relatively smooth surface wave model of Shapiro and Ritzwoller (2002) that was used in the tests of Villaseñor *et al.* (2000). In particular, WENA1.0 contains greater horizontal heterogeneity that can cause the *P* waves to travel off-great circle. Alternatively, Yang *et al.* (2004) used a ray shooting method to compute *Pn* travel times from 0° to 20° and their teleseismic *P* times from 25° to 97° were computed with a perturbation ray tracer that is not suitable for regions of triplication or diffraction. As a result, their dataset was limited to *Pn* arrivals between 0° and 15° and *P* arrivals between 25° and 97° epicentral distances. Our 3D FD method allows us to compute the first *P*-wave arrivals (*Pg*, *Pn*, or *P*) at any epicentral distance, so we are able to take advantage of more far-regional data. Furthermore, we do not alter our 3D velocity model (e.g., by flattening), so the primary source of error remains the resolution limited by grid size, and we conclude this will be less than 2.0 sec estimated by Villaseñor *et al.* (2000).

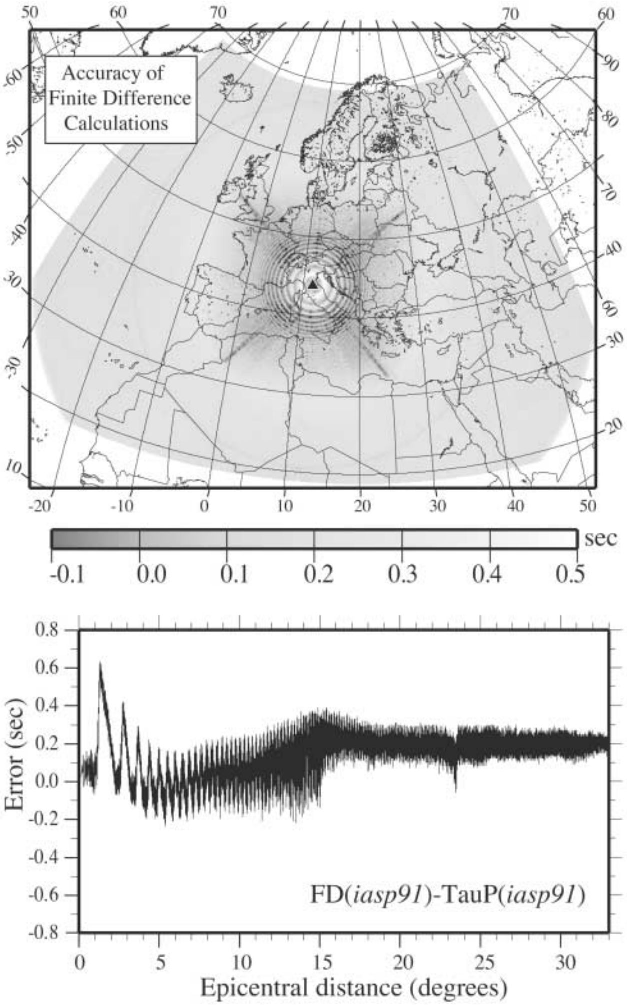

Figure 5. Estimates of computational error in the travel times predicted using the 3D FD code relative to theoretical arrival times computed using a 1D ray tracer through the global model *iasp91*. (top) Map view of errors at 10 km depth. Note the errors vary with both distance and azimuth. (bottom) Plot of errors as a function of distance from the source.

Model Evaluation

Various techniques can be used to assess model performance yet they all require a large quantity of calibration data

(GT event origins and arrivals) and robust statistical analyses. We assess the WENA1.0 model based both on its ability to predict the travel-time patterns seen in the GT25 processed dataset of *Pg*, *Pn*, and *P* arrivals and its ability to better locate the GT5 reference events. The first evaluation of WENA1.0 is to determine how well the computed regional travel times fit the GT25 travel-time observations, and we focus our efforts on the 39 stations shown in Figure 2 (left) to obtain a representative sampling of the model's predictive capability. For completeness and for comparison with other model validation studies, we examine both the *L1* norm (median) and *L2* norm (mean) statistics of misfit on a per station basis and demonstrate that WENA1.0 greatly improves the fit to regional travel times as measured by several statistical metrics. The second evaluation of the 3D WENA1.0 model is a set of relocation tests using model-based correction surfaces and the 196 events shown in Figure 2 (right) demonstrating that both the accuracy and uncertainty of the new location are improved.

Improvement in Travel-Time Prediction

To compute travel-time predictions for an exact source–station path, we use natural neighbor interpolation to calculate between the travel-time volume grid nodes. This is done for both the WENA1.0 travel-time volume and the *iasp91* travel-time volume, and each predicted time is subtracted from the observed arrival time to form residuals. For most stations, WENA1.0 reduces the residuals with respect *iasp91* as is evidenced by a summary of statistical quantities for all 39 stations presented in Table 1. Overall WENA1.0 proved to be superior to *iasp91* for predicting the observed travel-time patterns yielding a reduction in variance of 29% for the entire GT25 dataset. Variance reduction (VR) is as high as 40% to 50% for some stations (e.g., APA, ARO, KDS), but the performance at any given station is quite variable. Indeed, the results range from a significant VR of 33% to 51% (nine stations), moderate VR of 16% to 29% (13 stations), marginal VR of 7.4% to 14% (11 stations), null improvement with VR of 1.8% to −1.3% (four stations), and degradation of −4% to −18% (two stations).

Distributions of the travel-time residuals with respect to *iasp91* and WENA1.0 are shown in histogram form in Figure 6 for a subset of stations OBN, BKR, DHJN, and TAM. We report both the *L2* norm (standard deviation [σ] and variance reduction [VR]) and *L1* norm (scaled median absolute deviation [SMAD] and SMAD reduction [SMADR]) statistics for each station in Table 1. The SMAD provides an estimate of the spread of values which is less sensitive to outliers (in contrast to the standard deviation) and so is often more appropriate in the presence of non-Gaussian errors. Travel-time residual distributions often appear "heavy-tailed" and thus non-Gaussian, so the *L1* norm statistical measures are appropriate. Because there is no one single statistical measure that can definitively resolve one model as better than the other, we present a suite of metrics.

The SMADR for the entire GT25 dataset is 12% while again the performance at any given station is variable. In fact, we observe some instances where the VR is good yet the SMADR is low or negative (e.g., LKO, DHJN, TCF, KBS). This indicates that although WENA1.0 improves the overall travel-time prediction (high VR), and it predicts the outliers well (i.e., greatest magnitude residuals), the SMADR is low because it is less affected by outliers. This likely indicates that any remaining large residuals at these stations are due either to 3D structure not captured in WENA1.0 or to picking, origin time, or depth error that cannot be predicted by any 3D velocity model.

The mean amount of travel-time improvement provided by WENA1.0 is also reported (the last 2 columns in Table 1). Here the mean improvement provided by WENA1.0 for the number of travel times better predicted by *iasp91* (N_i) is negative (μ_Ni), but for the number of travel times better predicted by WENA1.0 (N_W) it is positive (μ_NW). For many stations, both the absolute value of μ_NW is greater than μ_Ni, when the VR is good and N_W is greater than N_i. However, this is not always the case (e.g., SHI, SVE, TCF in Table 1). In such instances, N_i is larger than N_W, yet the VR is favorable because, whereas WENA1.0 better predicts a smaller number of observations (N_W), their magnitude is better matched by *iasp91*.

The overall improvement we observe at the 39 stations examined is similar to that reported by other studies throughout Eurasia. Murphy *et al.* (2005) developed correction surfaces for 30 stations, mostly in Eastern Asia. For nine stations across the Former Soviet Union, including OBN, BRVK, and FRU, they observed the standard deviation of *P*-wave residuals from PNE (GT1) events to drop from 2.12 sec to 0.94 sec. They also tested the WINPAK3D velocity model of India and Pakistan (Reiter *et al.*, 2005) and find the standard deviation of *P*-wave residuals for a set of 17 GT5 reference events drops from 2.48 sec to 1.85 sec, similar to our observations at station NIL.

When we consider only the GT1–GT5 events the variance reduction is even more impressive at 49% for all stations combined (Table 1, Fig. 7). The fact that the GT level affects the fit to the regional travel times is not surprising. Accurately located events inherently contribute less to the travel-time residual and, because of our analyst prioritization, we have repicked far more GT5 and better events.

Model-Based Correction Surfaces

Another way of investigating the extent to which WENA1.0 outperforms *iasp91* in predicting travel times is to examine the spatial patterns of the travel-time residuals as seen by each station. Visually we can evaluate how well the WENA1.0 model predicts the data geographically by computing 3D travel-time correction surfaces for each of the 39 stations. To compute such model-based correction surfaces we subtract the *iasp91*-predicted time from the WENA1.0-predicted time (from the FD output) along a reg-

Table 1

Summary of Travel-Time Prediction Statistics for all Stations from the GT25 Dataset

		SMAD† (sec)		σ‡ (sec)		SMADR§	VR‖			μ_{Ni}††	μ_{NW}‡‡
Station	N*	*iasp91*	WENA	*iasp91*	WENA	(%)	(%)	N_i#	N_W**	(sec)	(sec)
AJM	141	3.383	3.418	3.506	3.288	−1.034	12.059	44	97	−0.538	0.638
APA	168	1.533	1.527	2.242	1.719	0.395	41.193	75	93	−0.574	1.063
ARO	132	2.358	1.517	2.209	1.640	35.678	44.909	36	96	−0.757	0.899
BGCA	37	1.121	1.139	1.366	1.399	−1.614	−4.787	20	17	−0.358	0.286
BKR	294	1.536	1.298	1.815	1.570	15.508	25.198	109	185	−1.515	1.733
BRVK	69	1.630	1.236	1.679	1.371	24.201	33.346	36	33	−0.869	0.874
CFTV	12	3.085	2.643	2.898	2.175	14.333	43.640	4	8	−1.153	2.263
DHJN	72	2.028	2.101	2.106	1.893	−3.590	19.172	37	35	−0.566	0.610
ELL	266	1.699	1.554	1.972	1.736	8.533	22.544	106	160	−1.228	1.881
FRU	200	1.689	1.341	1.737	1.658	20.634	8.917	73	127	−1.036	1.139
HYA	34	1.113	1.060	1.419	1.064	4.755	43.790	15	19	−0.573	0.760
KAD	87	1.815	1.534	2.642	2.345	15.479	21.216	31	56	−0.716	0.679
KBS	49	1.726	2.243	1.874	1.803	−29.977	7.454	10	39	−0.703	1.527
KDS	44	2.139	1.741	2.209	1.535	18.569	51.762	18	26	−0.602	1.401
KEV	120	1.939	1.559	1.833	1.597	19.592	24.083	48	72	−0.660	1.328
KHE	32	2.094	1.681	1.988	1.901	19.730	8.556	13	19	−1.916	2.393
KHO	185	1.676	1.634	1.870	1.704	2.470	16.968	89	96	−1.648	1.388
KUK	35	1.439	1.396	1.540	1.543	2.960	−0.393	19	16	−0.294	0.377
LKO	57	1.129	1.384	1.584	1.362	−22.512	26.055	28	29	−0.558	0.674
MAIO	173	2.180	1.559	2.056	1.670	28.468	33.956	124	49	−1.937	1.524
MBO	25	2.542	1.612	2.470	1.804	36.585	46.646	11	14	−0.278	1.301
MFP	9	2.347	0.336	1.674	1.564	85.685	12.719	5	4	−0.664	0.538
MLR	263	1.567	1.549	1.626	1.627	1.164	−0.217	91	172	−0.972	1.061
NIL	187	2.113	1.687	1.968	1.681	20.177	27.105	92	95	−0.862	1.474
NUR	459	1.580	1.417	1.831	1.471	10.321	35.438	183	276	−0.584	1.152
OBN	383	1.522	1.191	1.594	1.341	21.751	29.165	129	254	−0.505	1.155
PGD	125	1.294	1.309	1.405	1.414	−1.144	−1.352	42	83	−0.602	0.906
PTO	75	1.650	1.267	2.009	1.719	23.219	26.753	32	43	−0.710	0.991
QUE	174	1.330	1.410	1.310	1.425	−6.033	−18.423	82	92	−1.556	1.610
RYD	97	2.091	2.097	2.258	2.237	−0.277	1.833	48	49	−0.917	0.710
SHI	180	2.153	1.789	2.019	1.834	16.935	17.487	126	54	−1.996	1.439
SVE	177	1.391	1.364	1.829	1.633	1.941	20.356	112	65	−1.474	1.280
TAB	294	1.717	1.600	1.844	1.737	6.803	11.179	103	191	−1.347	1.890
TAM	187	1.464	1.515	1.798	1.658	−3.510	14.940	80	107	−0.705	0.642
TCF	263	1.014	1.141	1.180	1.129	−12.518	8.377	185	78	−0.844	0.635
TIC	100	1.404	1.364	1.520	1.427	2.903	11.783	41	59	−0.639	0.770
TIO	131	1.190	1.345	1.412	1.334	−13.040	10.808	78	53	−0.335	0.362
UZH	325	1.683	1.700	1.899	1.658	−0.994	23.731	121	204	−0.729	0.806
ZGN	106	1.956	1.858	2.271	2.114	5.002	13.301	42	64	−0.567	0.697
GT25 total	5767	1.927	1.691	2.192	1.846	12.256	29.066	2538	3239	−1.002	1.174
GT5 only	229	1.506	1.383	2.156	1.533	8.173	49.444	106	123	−0.540	0.190

*Number of travel times (events) recorded at this station.

†Scaled median absolute deviation.

‡Standard deviation.

§Scaled median absolute deviation reduction.

‖Variance reduction.

#Number of travel times better predicted by *iasp91*.

**Number of travel times better predicted by WENA1.0.

††For N_i, the mean travel-time improvement provided by WENA1.0.

‡‡For N_W, the mean travel-time improvement provided by WENA1.0.

ular grid in latitude, longitude with 25 km sampling; depth is regularly sampled at intervals of 10 km from 0 km down to 50 km. Each prediction model provides station-specific travel-time residuals with an overall range of −6.7 sec to 6.2 sec, relative to *iasp91*, among all 39 stations. Example surfaces are shown in Figure 6 for the stations OBN, BKR, DHJN, and TAM for a source depth of 10 km; we choose this depth as it represents the average focal depth of the GT25 events selected. These four stations are well distributed in WENA and give a cursory view of how the travel-time deviations from *iasp91* vary throughout western Eurasia. WENA1.0 predicts travel-time anomalies up to ±7 sec

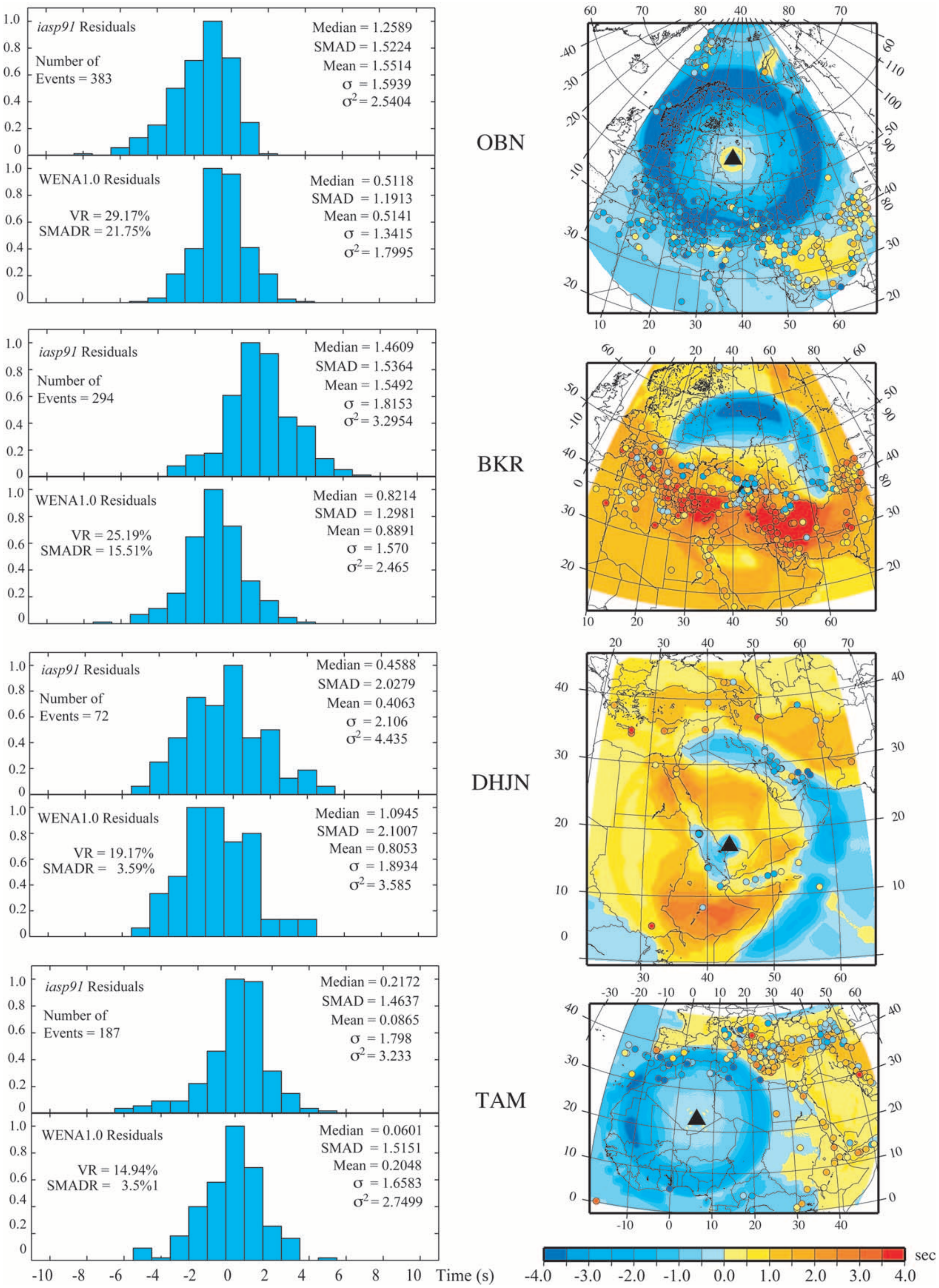

Figure 6. Caption on following page

Figure 6. Histograms of travel-time residuals for both WENA1.0 and *iasp91* at stations OBN, BKR, DHJN, and TAM along with the travel-time residual surfaces (WENA1.0-*iasp91*) computed using the 3D FD algorithm. (left) Histograms of residuals GT25-*iasp91* and GT25-WENA1.0 show the variance reduction provided (VR) by the WENA1.0 model. Although WENA1.0 does not improve travel-time prediction at all stations we find that the overall prediction is improved by 29% for the entire GT25 dataset (see Table 1 for all statistics and definitions). (right) Model-based correction surfaces representing the variation in travel time as a result of 3D velocity structure in the WENA1.0 model relative to *iasp91*. The surfaces represent the travel-time variations at 10 km depth as seen by each station (black triangle). Travel-time differences of up to ±6 sec relative to *iasp91* are seen, most in areas of very thick crust or sediment (blue indicates fast regions and red indicates slow). Circles are the empirical travel-time residuals of the GT25 dataset relative to *iasp91*, color coded using the same scale as the correction surface for direct comparison. Note that although these surfaces were derived without any of these data they reflect the trends in many places, and the remaining misfit should be due to either structure not captured in the model or conflicting picking errors. (Ⓔ See supplemental material in the online edition of BSSA.)

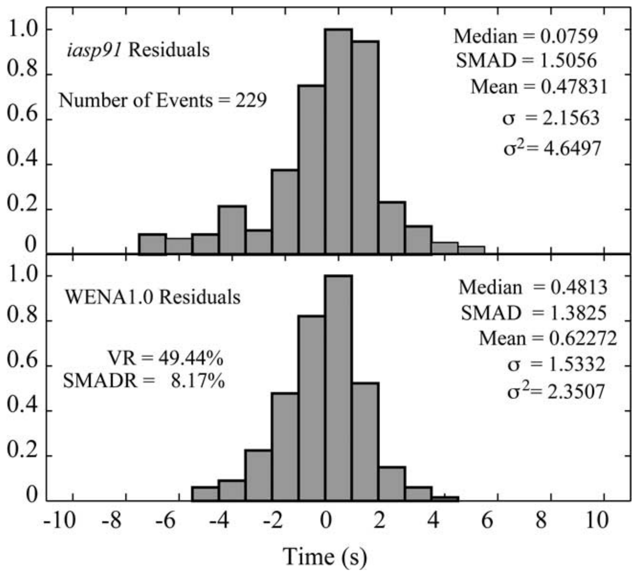

Figure 7. Histogram of travel-time residuals combined from all 39 stations for GT5 events only. The WENA1.0 model provides a variance reduction of 49% for these high-quality GT events.

(relative to *iasp91*) and we observe this same magnitude in the data; this lends credence to the accuracy of the amplitude of *P*-velocity structure present in our 3D model.

Color-coded residuals (data-*iasp91*) from events of focal depth 0 to 20 km are also plotted on top of the surfaces in Figure 6 to provide a visual assessment of the geographic agreement between WENA1.0 prediction and the data. The patterns seen in the correction surfaces correlate well with many structural features in WENA, and Moho depth, crustal age, and upper mantle velocity are all observed to be factors in the travel-time deviation from *iasp91*. The small circular pattern surrounding the stations at about 2° is the *Pg–Pn* crossover; the *Pn–P* crossover distance varies azimuthally for each station but occurs roughly between 12° and 15°. The largest deviations occur for paths through very thick crust (greater than about 45 km), deep sedimentary basins, or anomalously fast or slow upper mantle (see Pasyanos *et al.*, 2004). In fact, we find that the mantle velocity is equally as important as crustal thickness and velocity; we concur with the observation of Smith and Ekström (1996) who find that long-wavelength mantle structure can produce travel-time anomalies on the same order as station corrections. For example, note the negative residuals for events in southwest Iran as seen at DHJN (Fig. 6). Although the crust is thick and slow beneath this area, there is a patch of fast mantle velocity (Fig. 1) that contributes to these negative residuals, while late predictions in surrounding regions of Iran and the Arabian Peninsula are correctly predicted. Pasyanos *et al.* (2004) also noted the significance of mantle velocities on travel-time residual patterns as they tested and found the 3SMAC model (Nataf and Ricard, 1996) to be superior to the RUM (Gudmundsson and Sambridge, 1998) for predicting fast arrivals throughout the Russian platform.

At epicentral distances less than 10° the character of the correction surfaces is dominated by crustal velocity and Moho depth variations, whereas beyond ~13° velocity variations in the uppermost mantle become increasingly important. Fast predictions are seen on the Russian Platform, Baltic Shield, and African craton where the mantle is relatively fast (*Pn* velocity as high as 8.4 km/sec); however, the effect of the thick crust of the Urals is easily seen in the correction surface for OBN as a slower area in the northeast corner extending up to Novaya Zemlya. The residuals observed at OBN are predicted well at most azimuths and distances; however, fast residuals from events in the Zagros and Mediterranean Sea are underpredicted by about 1 sec. For BKR, the residuals for events in Hindu Kush, the Red Sea rift, Zagros, and central Mediterranean Sea are matched quite well by WENA1.0, and although WENA1.0 predicts negative travel-time deviations from the Russian Platform (blue area north of station, Fig. 6), it does not predict them far enough the south to match the fast residuals very close to the station. Slow anomalies are seen at DHJN from the East African rift and Anatolian and Iranian Plateaus where the upper mantle is quite slow (*Pn* velocity as low as 7.6 km/sec) (see Figs. 1 and 6). Also, the fast arrivals from the Zagros, near the *Pn–P* crossover distance, are well predicted by WENA1.0. Fast arrivals are predicted for TAM for most of the African Platform and Archean Cratons and slow residuals from east Africa and the Zagros are well matched, but the fast residuals from events in the Caucasus are not well predicted. (Ⓔ Maps of correction surfaces and histograms for the full set of 39 stations listed in Table 1 are available in the online edition of BSSA.)

The WENA1.0 model correctly reproduces broad travel-time residual patterns, and unlike purely empirical approaches, WENA1.0 anomalies are extrapolated far from data. As seen in Figure 6, WENA1.0 reduces the variance of residual distributions relative to *iasp91*, and in most cases the distribution of WENA1.0 residuals is closer to zero mean than *iasp91* residuals. Although great care was taken to construct WENA1.0, it is a continent-scale model and the 1° resolution may not be sufficient to capture all regional velocity anomalies. Therefore, we expect prediction to improve overall relative to any 1D velocity model, but it will not be perfect everywhere. WENA1.0 could thus be viewed as a starting model that could be refined by using empirical GT observations combined with numerous other data types in a joint inversion for 3D velocity structure.

In addition to reducing the overall bias in the travel-time observations (Table 1, Fig. 6), WENA1.0 also matches several geographic trends in travel-time anomaly. Remaining errors not attributable to inaccurate velocity structure in the 3D model are most likely the result of errors in picking the arrival times (note the variation of 1 to 3 sec in residuals with nearly identical locations, Fig. 6) and/or errors in origin time and depth. Errors in *Pg*, *Pn*, and *P* picks can range up to ±6 sec, in particular, at *Pn*–*P* crossover distances and upper mantle triplication distances for *P* where phase misidentification is typical. We examine this topic in the next section with the use of nonstationary variogram modeling.

Travel-time Uncertainty Model

Although many types of 3D models are used to improve travel-time prediction of regional phases, a key issue with any velocity model, *a priori* models in particular, remains evaluating its accuracy and uncertainty. Model uncertainty is necessary to propagate it through to travel-time prediction and hence to new location estimates to provide associated location uncertainties. Geophysical models derived from seismological data such as tomographic inversions often have uncertainties associated with the resulting velocities that can be used to assess travel-time prediction uncertainty. The WENA1.0 model does not have such uncertainties; thus we must derive an uncertainty model for it.

We implement a method to compute uncertainty estimates for the model-based correction surfaces, which will be required to obtain representative error ellipses for new event locations. The misfit between the observed travel times and those predicted from WENA1.0 is used as a measure of model uncertainty. It is commonly observed that travel-time prediction uncertainty is distance dependent. This results from velocity-model errors that cause cumulatively more bias in travel-time prediction with distance, and from the fact that longer paths sound to greater depths and therefore sample different parts of the model. Here we derive an azimuthally invariant, distance-dependent uncertainty model using our GT25 and better travel-time dataset. The fitting procedure entails calculating the mean and spread of the residual distribution in discrete distance bins. The statistics of each bin are then used to determine the uncertainty at a given confidence in each distance bin.

Deriving an uncertainty model as a function of epicentral distance begins with nonstationary variogram modeling for each station to assess the spatial statistics of the travel-time residuals. Stationary variogram modeling examines the difference between residuals solely as a function of inter-event distance. Nonstationarity describes geographic variability in the covariance structure of the residuals, so our nonstationary modeling examines covariance as a function of station–event distance as well. We compute the nonstationary variogram and record the nugget in four discrete distance ranges: 0° to 2° (local), 3° to 12° (near regional), 13° to 24° (far regional), and 25° to 96° (teleseismic). Results of the variogram analysis are shown in Figure 8 for stations OBN, KEV, MAIO, and NUR. Note that the variograms do not approach zero for points that are colocated (i.e., the nugget is not zero, and data are not perfectly correlated) due to uncorrelated errors associated with determining travel-time residuals. The nugget is similar for both *iasp91* and WENA1.0 residuals and represents random picking error. However, it is apparent that the variograms do reach minima (correlation is maximum) for points that are close together, and the variograms typically increase (correlation decreases) as points become further separated. This is evident in the variograms for the *iasp91* residuals where the variogram steadily increases or has an undulated appearance (a manifestation of azimuthal nonstationarity in this instance) as distance increases, whereas the WENA1.0 variograms appear to have similar nugget and range but the sill is often smaller and the structure is more regular as distance increases. Overall, the *iasp91* variograms are azimuthally nonstationary beyond 6° to 7° and have higher covariance than the WENA1.0 variograms at the same distance. WENA1.0 variograms appear more stationary after 6° to 7° and have a smaller sill (lower background variance). Variogram analysis illustrates how the WENA1.0 model accomplishes two goals: improving travel-time prediction by reducing the overall variance of residuals, and reducing the nonstationarity of the uncertainties. The 3D model removes the long-period structure in residuals, as ideally we want to account for all correlated structure and have only random error remain (Douglas, 2005; Myers, 2005). An azimuthally invariant uncertainty model is our ultimate goal because a simpler uncertainty model indicates that the 3D velocity model is predicting the 3D structure, so errors are more consistent and error ellipses are truly representative of location accuracy.

Individual residuals of the GT25 dataset for both WENA1.0 and *iasp91* are plotted in Figure 9 along with the uncertainty curves derived by fitting the observed spread in residuals while accounting for picking error (i.e., subtracting the nugget in each distance bin). At near- and far-regional distance, WENA1.0 uncertainty can be up to 2 sec lower

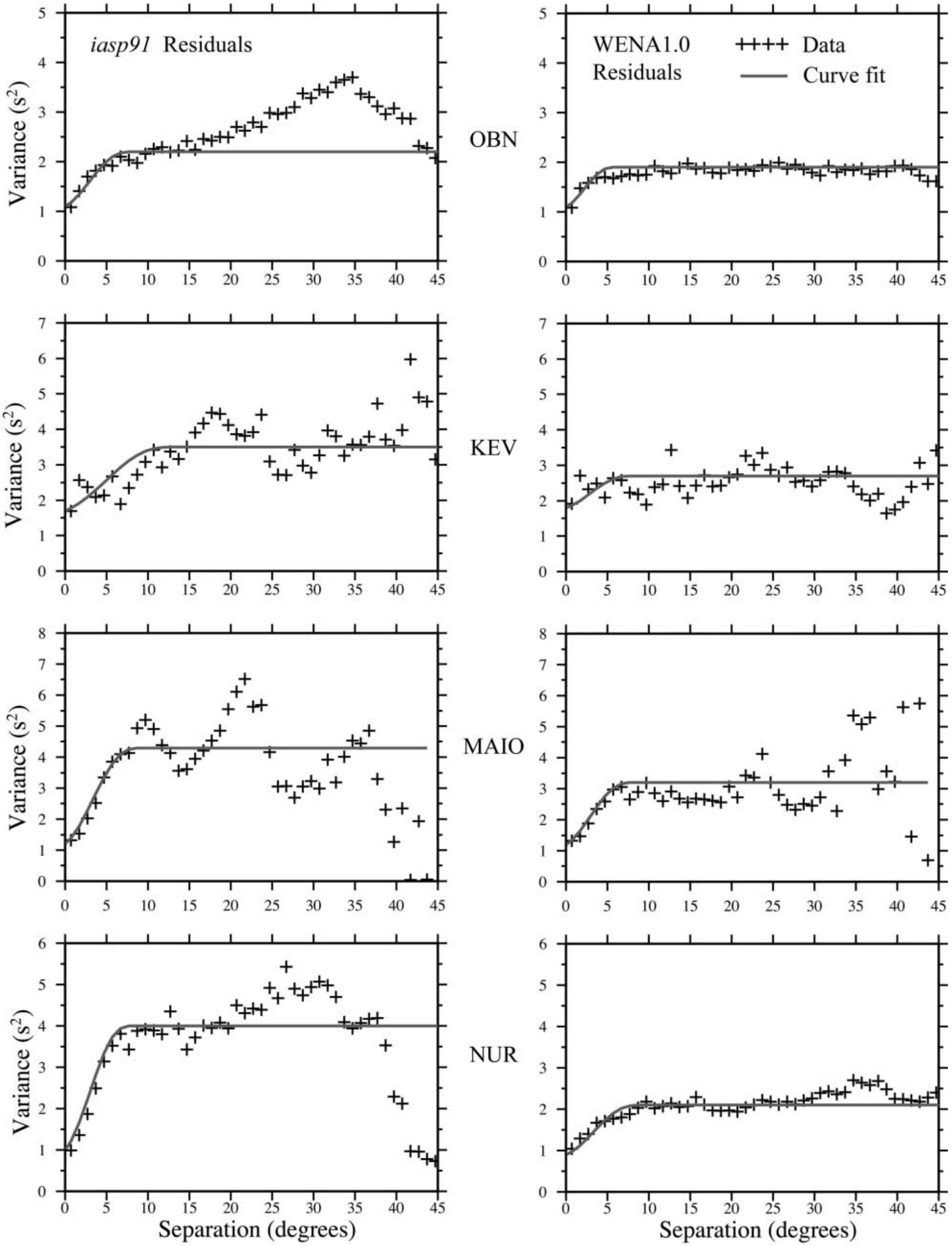

Figure 8. Variograms of travel-time residuals at stations OBN, KEV, MAIO, and NUR for both the WENA1.0 and *iasp91* velocity models. Crosses are the data variogram values in 1.0° bins; solid lines are the model variograms determined by curve fitting. The *iasp91* variogram is nonstationary (levels off at the sill, then increases again). This nonstationarity is caused by long-period features in the travel-time residual structure. WENA1.0 improves prediction of long-period residual features, and the variogram is relatively flat after the sill is reached. In some cases the sill value in the WENA1.0 variogram is reduced overall relative to the *iasp91* variogram (e.g., NUR) indicating the 3D model reduces the background variance.

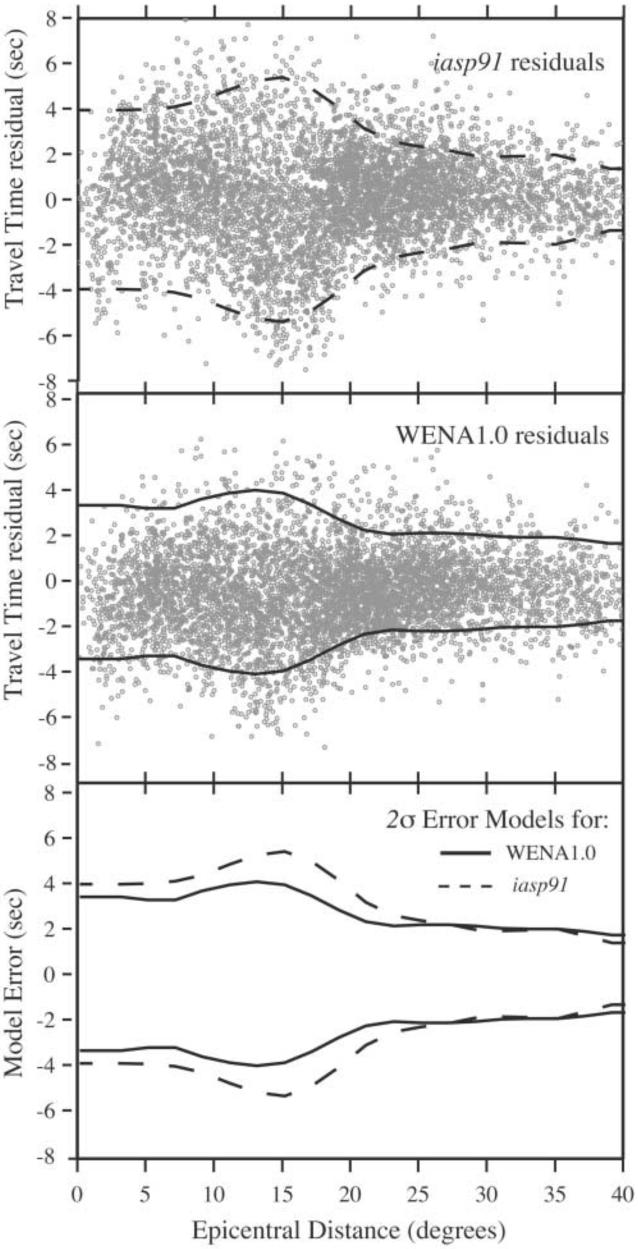

Figure 9. Residuals from the GT25 dataset for *iasp91* (top) and WENA1.0 (center) are plotted against epicentral distance, and are used to derive distance-dependent uncertainty models for each model (bottom). Note the nonstationarity and correlation in the travel-time uncertainty between 10° and 23° epicentral distance; uncertainty increases and errors are correlated.

than *iasp91*, and as distance increases to teleseismic both models show similar uncertainty. Note the nonstationarity and correlation in the travel-time uncertainty between 10° and 23° epicentral distance; uncertainty increases and errors are correlated. Not surprising, the largest errors occur in this distance range. *Pn* arrivals may have ray paths bottoming well below the crust in the lithosphere where low-velocity zones may be present and cause great variation in the arrival time. Also, at far-regional distances phase identification between various *Pn* branches or between *Pn* and *P* can be difficult and thus contribute to the spread of residuals as was observed by Ritzwoller *et al.* (2003).

Yang *et al.* (2004) also used an azimuthally invariant uncertainty model in their relocation tests that was a scaled version of the baseline *iasp91* teleseismic *P*-error model between 25° and 97°, but for the distance range 0° to 15° they derived an empirical modeling error function from their *Pn* travel-time misfits (i.e., the standard deviations of the misfits as a function of distance). They also assigned a 1-sec uncertainty to arrivals with no reported uncertainty, whereas we assign the square root of the nugget value for a given station/distance range (or used the analyst–assigned uncertainty when available). We believe that our practice better represents pick errors, resulting in a clearer distinction between pick and model uncertainties. The Yang *et al.* (2004) uncertainty model shows a shape similar to those plotted in Figure 9 but is lower in amplitude and does not exist in the critical distance range 15° to 25°. This is likely due to their stringent cutoff value for residuals and the exclusion of potentially ambiguous *Pn* or *P* data between 15° and 25° (see Fig. 2 in Yang *et al.*, 2004).

Our derived distance-dependent uncertainty model is now combined with our model-based correction surfaces to essentially provide a set of station-based 3D travel-time tables with uncertainties that can be used for computing seismic-event locations. Because the distance-dependent uncertainty model is based on a dataset that covers nearly all of WENA, it is applicable over the whole model region.

Relocation of GT5 Calibration Events

Whereas the GT25 travel-time fits provide quantitative information about WENA1.0, the GT25 dataset alone does not address the location capability of the model. As our ultimate goal is to improve regional-network location accuracy, the final evaluation of WENA1.0 is a relocation test on the 196 well-located GT5 reference events using the travel-time predictions (model-based correction surfaces, e.g., Fig. 6) and our derived uncertainty model (Fig. 9). High-quality GT5 events are selected to avoid ambiguity between model and data uncertainties. We assess the degree to which our 3D velocity model accounts for the spatial patterns in arrival times as seen in the correction surfaces (Fig. 6). Event relocations with and without the travel-time corrections are compared, and we use the statistics of mislocation, error ellipse area, 95% coverage, and location misfit vectors to evaluate location improvement.

We use the LOCOO (LOCator Object Oriented) code of Ballard (2002) that implements the traditional linearized, least-squares single-event location algorithm (Geiiger, 1912). LOCOO calculates uncertainties using the method of Jordan and Sverdrup (1981) that combines *a priori* uncertainties and *a posteriori* residuals to estimate confidence bounds on seismic-event hypocenters and origin times. In this study, we use *a priori* uncertainties, based on the uncertainty model and pick characterization described previously to determine confidence bounds (see Evernden, 1969).

Each event is located using first arriving *Pn* or *P* waves and both the *iasp91* and WENA1.0 velocity models. Relocation using *iasp91* is straightforward because the location algorithm utilizes 1D travel-time curves and the associated uncertainty curve for *iasp91* (Fig. 9, top). For relocation using WENA1.0, the correction surfaces for each station are essentially the analog of a 3D model of travel-time curves for 1D models. They amount to 3D travel-time tables for a specific station as a function of source latitude, longitude, and depth. The associated distant-dependent WENA1.0 uncertainty model (Fig. 9, center) is also specified in the same manner for each station. Measurement variance, either analyst assigned or determined by variogram modeling, and model variance comprise the uncertainty budget in this study; 95% confidence ellipses are estimated from the total *a priori* uncertainties under the assumption of Gaussian distributed independent errors.

Selection of events for the location validation required at least five defining arrival-time picks. In all our relocation tests, the depth was fixed to the GT depth, and we solve for latitude, longitude, and origin time; fixing the depth is required because the station geometry is often sparse, making the constraint on depth inherently poor. The geographic distribution of the relocated GT5 events and associated ray paths is seen in Figure 10 (left) for events that moved more than 5 km (the GT level) and had a secondary azimuthal gap (*sazgap*) of less than 270° (we further evaluate network geometry in the following section). Examining the individual ray paths along which travel time (and thus the location) is better predicted using WENA1.0 (red) or better predicted using *iasp91* (blue) allows us to qualitatively assess where WENA1.0 is performing best. We observe improved location calibration on the Russian Platform, European Arctic, Middle East, South Asia, East African Rift, Anatolian Plateau, and parts of the Mediterranean, but many individual paths show conflicting results. Most of the improvements occur in the northern and eastern parts on the model, whereas the most equivocal results are in the Hellenic Arc in the Mediterranean. The 79 GT5 events that were better located by WENA1.0 (red triangles) and the 33 events that where better located by *iasp91* (blue triangles) are plotted in Figure 10 (center), again for relocation greater than 5 km and *sazgap* of less than 270°. If we consider all relocated events that moved more than 5 km and had no secondary azimuthal gap restriction, we observe that WENA1.0 improved 103 events and *iasp91* improved only 41 (see Fig. 12, right). The amount of improvement scales with the size of the symbols, so the greatest improvement can be seen by the larger red triangles, regardless of secondary azimuthal gap criteria, demonstrating that WENA1.0 improves location accuracy in both more instances and to a greater extent.

Relocations using WENA1.0 clearly improve event location accuracy as reported in Table 2. Whereas the average *iasp91* mislocation is 24.9 km, it is 17.7 km for WENA1.0. Distributions of the relative difference in epicenter mislocation for the GT5 validation dataset using both velocity and uncertainty models are shown in Figure 11. The scatter plot (Fig. 11, top) illustrates the epicenter mislocations produced from using both WENA1.0 and *iasp91* models for those events having *sazgap* less than 270°. Symbols above the di-

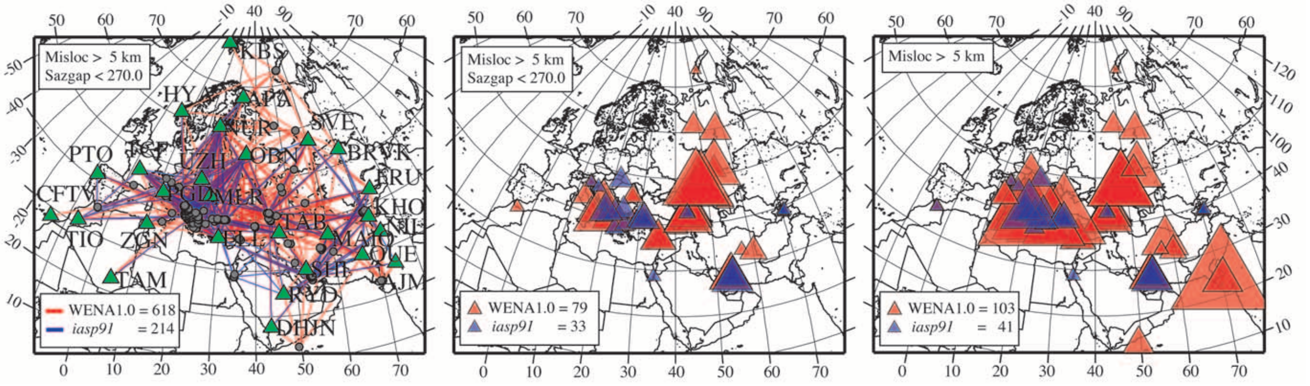

Figure 10. (left) Individual raypaths from GT5 events (gray circles) and stations (green triangles) used in relocation tests which depict specific paths along which the travel-time prediction and thus the locations are improved by WENA1.0 (red lines) or improved by *iasp91* (blue lines). We plot only those events that moved more than 5 km (the GT level) and had a secondary azimuthal gap of less than 270°, which resulted in 618 paths being better predicted, while 214 paths were degraded. (center) Geographic distribution of the 79 GT5 events which were better located by WENA1.0 (red triangles) and the 33 events which were better located by *iasp91* (blue triangles), again for relocation greater than 5 km and secondary azimuthal gap of less than 270°. The amount of location improvement scales with size of the symbols. (right) Same as center plot for relocation greater than 5 km and no secondary azimuthal gap restriction.

Table 2
Summary of Relocation Statistics for the GT5 Dataset

	iasp91	WENA1.0
Mean mislocation (km)	24.904	17.734
Median mislocation (km)	20.930	14.165
Number of events improved	33	79
Mean improvement (km)		7.169
Mean improvement (%)		28.79
Median improvement (km)		5.595
Median improvement (%)		32.22
Mean error ellipse area (km^2)	6466	4075
Median error ellipse area (km^2)	5440	3431
Mean improvement (%)		37
Median improvement (%)		37
Coverage (%)	90.45	94.09

Only events located with secondary azimuthal gap <270°.

agonal line are improvements using WENA1.0 while symbols below the line are improvements using *iasp91*. Difference in epicenter mislocation is also shown in histogram form in Figure 11 (bottom) where negative values indicate improvement with WENA1.0. This represents a mean improvement in location of 7.1 km and is similar in magnitude to that achieved by other calibration efforts that make use of 3D velocity models and GT10 or better reference events (Yang *et al.*, 2001a; Ritzwoller *et al.*, 2003; Yang *et al.*, 2004; Murphy *et al.*, 2005).

Next we consider the effect of network geometry on our relocations because it is a large factor in the resulting location accuracy (Bondár *et al.*, 2004). Mislocation as a function of primary and secondary azimuthal gap, and number of stations having a defining pick is plotted in Figure 12. When the number of stations used is small (less than five) the location is not well constrained by either velocity model. The average location error grows as the number of stations decreases, with degradation increasing rapidly at about 10 to 11 stations. Note that for any number of stations WENA1.0 almost always performs better than *iasp91* even when more than 10 stations are used. When the primary azgap is greater than 180° and or the secondary azgap is greater than 270°, the station geometry is poor and both models do an insufficient job of predicting an accurate location, as even small model and pick uncertainties are magnified into large location errors. However, the mean and standard deviation of the relocation distributions are smaller for the WENA1.0 model than for the *iasp91* model. The *sazgap* in particular proved to be an important indicator of the robustness of the final locations; thus we chose a cutoff at 270° for the remaining analysis. This reduces the GT5 dataset from 196 to 155 events, but the number of events remains high enough to draw statistically meaningful conclusions.

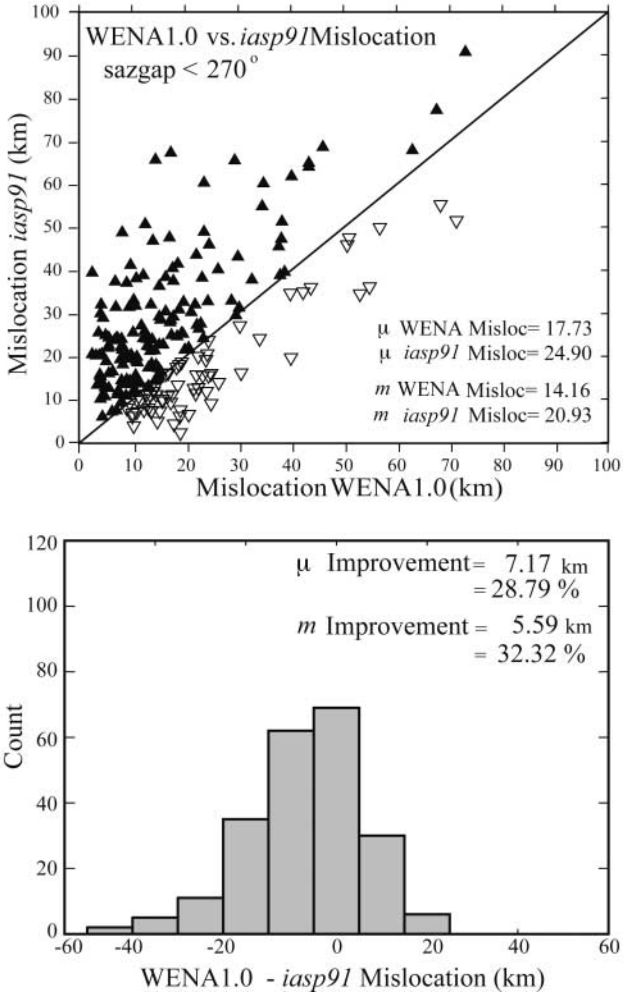

Figure 11. Scatter plot of WENA1.0 and *iasp91* epicenter mislocations (in kilometers) for the GT5 validation dataset using relocated events with secondary azimuthal gaps of less than 270° (top). Symbols above the diagonal line indicate improvements in location using WENA1.0 (filled triangles), while symbols below the line indicate improvements in location using *iasp91* (open triangles). (bottom) Histogram of the difference in epicenter mislocation (WENA1.0-*iasp91*) using the GT5 validation dataset. Note the mean (μ) and median (m) improvement in location provided by WENA1.0 is 7.1699 km and 5.595 km respectively.

Finally, we evaluate the coverage ellipses for the GT5 relocations. Determining representative error ellipses is a critical component of location accuracy, because they must adequately describe the random chance that the true location is within the bounds of the ellipse (Myers and Schultz, 2000). Error ellipse coverage is defined as the percentage of GT locations that fall within the corresponding confidence ellipse a specified percent of the time (e.g., 95%). As shown in Figure 13 and Table 2, the median area of error ellipse was reduced by 37% while the conservative modeling errors assured 94% coverage for WENA1.0. The reduction of both mean (2391 km^2) and median (2009 km^2) error ellipse area is substantial, at 37%, and is a direct consequence of the WENA1.0 uncertainty model variances used to compute the error ellipse. This is consistent with our finding that WENA1.0 both improves location accuracy and uncertainty estimates. For WENA1.0, the known location lies within the

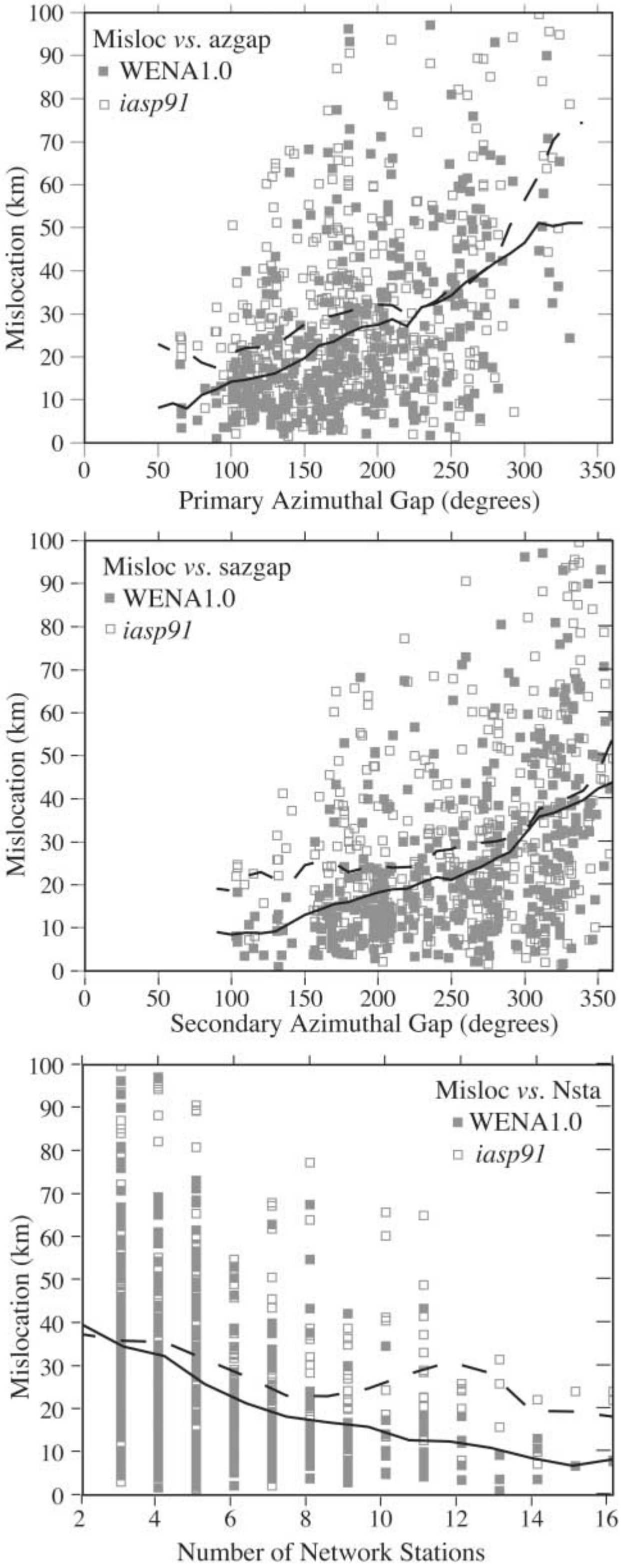

Figure 12. Mislocation is examined as a function of primary azimuthal gap (top), secondary azimuthal gap (middle), and number of stations having a defining pick (bottom). These plots collectively illustrate the network quality of the stations used for each relocation. When the primary azgap is greater than 180° and or the secondary azgap is greater than 270° the station geometry is poor and both models do a poor job of predicting and accurate location. When the total number of stations used to locate is small (less than 5) the location is also not well constrained by either velocity model. The mean mislocation for both WENA1.0 (solid line) and *iasp91* (dashed line) is also shown.

95% confidence ellipse 94% of the time and for *iasp91* it is only 90% of the time. This suggests that the WENA1.0 95% confidence ellipses are actually meaningful and that our derived error model represents true location accuracy. The observation that the *iasp91* relocation is only inside the 95% ellipse 90% of the time is noteworthy and is likely the result of correlated error unaccounted for in the propagation of uncertainties (Myers, 2005). Indeed, despite the fact that errors are correlated and the location method assumes independent errors, good error ellipse coverage is obtained.

Many researchers have also had success using model-based corrections to improve seismic-event locations in western Eurasia. Results from Fennoscandia (Yang *et al.*, 2001a) showed an improvement in location error of 9 km due to source-specific station corrections (SSSCs), and the area of the error ellipses was decreased from 3830 to 1120 km^2. Murphy *et al.* (2005) used high-resolution tomographic 3D models, such as WINPAK3D (Reiter *et al.*, 2005), to predict the travel-time corrections for International Monitoring System (IMS) stations in central and southern Asia. The Murphy *et al.* (2005) SSSCs reduced mislocation error of Soviet PNEs from 20.4 km (using *iasp91*) to 6.9 km. Johnson and Vincent (2002) and Reiter *et al.* (2005) used cross-validation of rms residual reductions to evaluate their WINPAK3D models, whereas Antolik *et al.* (2001, 2003) used randomly selected subsets of arrivals to evaluate location bias when using their global model J362. Ritzwoller *et al.* (2003) used a groomed arrival-time dataset produced from a method of multiple event locations, a set of empirical phase path anomalies, and a set of GT0–GT10 events to validate their CUB2.0 surface wave model. Yang *et al.* (2004) performed a constrained bootstrap technique that reduces the effects of data clustering to produce simulated sparse network locations to test the capability of global models CUB1.0 (Shapiro and Ritzwoller, 2002), CUB2.0, and J362 to locate both regional and teleseismic events. An important aspect of all these studies is quantifying both the model and picking uncertainty and how these are propagated through to location uncertainty.

For example, Yang *et al.* (2004) report a 75% coverage ellipse, well short of their expected 90%, and they note this

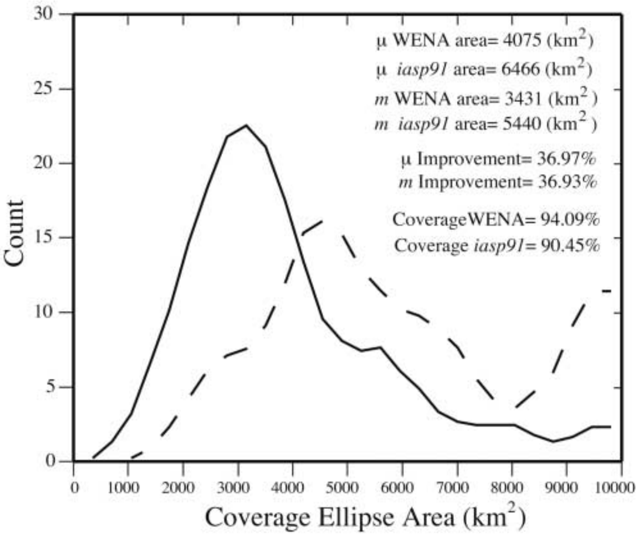

Figure 13. Distribution of WENA1.0 (solid line) and *iasp91* (dashed line) coverage ellipse areas plotted for GT5 events having *sazgap* greater than 270°. For WENA1.0 the known location lies within the 95% confidence ellipse 94% of the time and for *iasp91* it is 90% of the time. Both the mean and median area of the coverage ellipse for WENA1.0 relocations are smaller than for *iasp91* by 37%, indicating how WENA1.0 reduces the size of the coverage ellipse and thus yields a more accurate location.

could be due to events with large numbers of observations (i.e., the coverage ellipses get smaller as the number of defining picks gets larger). However, this may be a result of their likely underestimation of either their model uncertainties and/or their measurement (picking) uncertainties. They used bulletin data and assigned a measurement error of 1.0 sec by default, which may be too conservative, and they also had to increase the original scaling of their *iasp91* uncertainty model by 70% to increase the 90% coverage ellipse estimates. Through their constrained bootstrapping tests they determine that their assumption of uncorrelated errors is incorrect, as increasing the number of arrivals tends to increase the chance of raysampling redundancy, which leads to correlated arrival times and thus location bias.

Another source of uncertainty in some of these other studies is the need for baseline offsets when combining different velocity models. The CUB2.0 model used by Ritzwoller *et al.* (2003) only extends to 400 km depth and they only tested *Pn* and *P* phases out to 20° epicentral distance, whereas Yang *et al.* (2004) combined two different velocity models CUB1.0 (Shapiro *et al.*, 2002) with J362 (Antolik *et al.*, 2003). They also made some origin-time baseline adjustments as a result of the method by which the origin times were computed. Because we use a single-velocity model that was not derived from any travel-time data, we avoid the issue of baseline differences between different regional and teleseismic models when they are merged as well as the problem of circularity in testing the model with any given dataset of travel times.

Studies aimed at evaluating the location capability of certain models or techniques commonly concentrate exclusively on the accuracy of the locations relative to some benchmark. Although this is an important component of the present study, we conclude that the ability to model regional travel-time data is an equally good measure by which to assess the quality of 3D models. Evaluations based on mislocation alone may not account for the variations in network geometry from one region to another (or from event to event temporally), different mixes of (defining) regional phases, and uneven quality of reported travel times.

Conclusions

This study demonstrates that correction surfaces based on an *a priori* velocity model can both account for systematic location errors caused by crustal and upper mantle velocity structure and improve the ability to accurately locate low-magnitude events, while providing representative coverage ellipses. Empirically derived correction surfaces are applicable where GT events exist, but extrapolation of travel-time corrections from GT events is quickly damped to the statistics of the background model in regions where no data exist to constrain the corrections. Therefore, model-based corrections, as those shown here, are better suited to improve travel-time prediction and thus location capability in areas devoid of GT events.

Our test of the applicability of the WENA1.0 model for regional seismic-event location includes travel-time prediction performance, for a large GT25 validation dataset, and improvements in location accuracy, for a limited, geographically distributed set of GT5 validation events. We report several main findings. First, the 3D FD approach used here, with the Cartesian to spherical coordinate transformation, is a significant advance for travel-time prediction at regional to teleseismic distances. Second, a carefully validated set of reference events with good geographical coverage is essential for model-based location calibration. Such datasets are very time consuming to compile, as they require substantial effort to assess arrival-time quality, uncertainty, and outlier removal techniques. Third, WENA1.0 achieves a statistically significant improvement in travel-time prediction over *iasp91* with a variance reduction of 29% for all GT25 events and 49% for GT5 and better events. The clarity of improvement for GT5 events is attributed to smaller uncertainty in both epicenter error and a greater number of repicked arrival times. Improvement in WENA1.0 travel-time predictions primarily stems from removal of long-wavelength anomalies in the *iasp91* predictions. These long-wavelength features cause undesirable correlations in travel-time predictions that violate assumptions of most location algorithms.

Fourth, a distance-dependent, travel-time uncertainty model is developed for WENA1.0 and *iasp91*, and WENA1.0 travel-time uncertainties are noticeably smaller than *iasp91*, in particular, in the critical range of 0° to 25°. These representative uncertainties are key to our ability to compute realistic coverage ellipses for the relocated GT5 events. Fifth, tests of location improvement suggest that WENA1.0 improves epicenter accuracy by about 30% ($\sim$7 km) over *iasp91*, and it improves more events than does *iasp91*. Uncertainty estimates are representative of observed mislocation with a mean decrease in ellipse area of 37%, and resulting coverage ellipses are representative of true location error for WENA1.0 (i.e., our 95% confidence ellipses actually contain 94% of the location estimates).

Finally, WENA1.0 is meant as a starting point for data-driven approaches to developing velocity models (e.g., Pasyanos *et al.*, 2006); characterization of *a priori* error enables Bayesian tomographic inversion techniques, which can improve the model in data-rich areas and allow for stability of the inversion in data-poor regions.

Data Sources

Arrival times and event locations used in this study are complied from several bulletins. The EDR catalog (http://earthquake.usgs.gov/regional/neic), the ISC catalog (http://www.isc.ac.uk), the REB catalog prior to 2002 (http://www.pidc.org), EHB catalog (ftp://ciei.colorado.edu/pub/user/engdahl/EHB), and the FINNE (http://www.seismo.helsinki.fi/bul/index.html) catalogs are all publicly available. Additional arrival-time picks were made by analyst Flori Ryall and other members of the Seismology Group at Lawrence Livermore National Laboratory and are not presently available to the public.

Acknowledgments

We thank J. Hole for providing the finite-difference code and answering numerous questions to get it up and running. We gratefully acknowledge extensive and helpful conversations with M. E. Pasyanos and W. R. Walter throughout the execution of this work. We thank our analyst F. Ryall for her high-quality picking of travel times and J. Bhattacharyya for thoughtful comments. This work would not be possible without the software support of our Sandia National Laboratory colleagues S. M. Ballard, M. C. Chang, J. Hipp, R. Simons, and C. J. Young (Hipp *et al*, 2002; Ballard *et al.*, 2004). In particular, we thank J. Hipp for promptly and graciously providing the fdb and C. J. Young for his insight and abundant comments on this manuscript. M. Antolik, I. Bondár, and an anonymous reviewer provided detailed suggestions that improved and streamlined this article. All maps and many figures were generated using the Generic Mapping Tools (GMT) data processing and display package (Wessel and Smith, 1995). This research was supported by the Ground-Based Nuclear Explosion Monitoring (GNEM) Program. This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract no. W-7405-Eng-48. This is LLNL contribution no. UCRL-JC-220179.

References

Antolik, M., G. Ekström, and A. Dziewonski (2001). Global event location with full and sparse datasets using three-dimensional models of mantle P-wave velocity, *Pure Appl. Geophys.* **158,** 291–317.

Antolik, M., Y. J. Gu, G. Ekström, and A. M. Dziewonski (2003). J362D28: a new joint model of compressional and shear velocity in the Earth's mantle, *Geophys. J. Int.* **153,** 443–466.

Armbruster, J., V. Burlacu, M. Fisk, V. I. Khalturin, W-Y. Kim, I. Morozov, E. Morozova, P. G. Richards, D. Schaff, and F. Waldhauser (2002). Seismic location calibration for thirty international monitoring system stations in Eastern Asia, in *Proceedings of the 24th Annual Seismic Research Review Symposium Nuclear Explosion Monitoring: Innovation and Integration*, Ponte Vedra Beach, Florida, 231–239.

Ballard, S. (2002), Seismic event location using Levenberg-Marquardt least squares inversion, Sandia National Laboratories Report, SAND2002-3083, 25 pp.

Ballard, S., M. C. Chang, J. R. Hipp, L. A. Jensen, R. W. Simons, and L. K. Wilkening (2004). The 2004 Knowledge Base parametric grid data software suite, in *Proceedings of the 26th Seismic Research Review: Trends in Nuclear Explosion Monitoring*, Orlando, Florida, 211–220.

Begnaud, M. L., C. A. Rowe, and L. K. Steck (2004). Validating three-dimensional velocity models in China and East-Asia for use in regional seismic event location, *EOS Trans. AGU.* **85,** no. 47, (Fall Meet. Suppl.), T11C-1277.

Bijwaard, H., and W. Spakman (1999). Fast kinematic ray tracing of first- and later-arriving global seismic phases, *Geophys. J. Int.* **139,** 359–369.

Bondár, I., S. Myers, E. R. Engdahl, and E. Bergman (2004). Epicenter accuracy based on seismic network criteria, *Geophys. J. Int.* **156,** no. 3, 483–496.

Bondár, I., X. Yang, R. North, and C. Romney (2001). Location calibration data for the CTBT monitoring at the Prototype International Data Center, in *Monitoring of the Comprehensive Nuclear Test-Ban Treaty: Source Location*, F. Ringdal and B. Kennett (Editors), *Pure Appl. Geophys.* **158,** 19–34.

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

Cerveny, V. (2001). *Seismic Ray Theory*, Cambridge Univ. Press, Cambridge, UK, 713 pp.

Crotwell, H. P., T. J. Owens, and J. Ritsema (1999). The TauP ToolKit: flexible seismic travel-time and raypath utilities, *Seism. Res. Lett.* **70,** 154–160.

Douglas, A. (2005). Comment on "Improving Sparse Network Seismic Location with Bayesian Kriging and Teleseismically Constrained Calibration Events," by Stephen C. Myers and Craig A. Schultz, *Bull. Seism. Soc. Am.* **95,** 367–369.

Du, Z. J., A. Michelini, and G. F. Panza (1998). EurID: a regionalized 3D seismological model of Europe, *Phys. Earth Planet. Interiors* **103,** 31–62.

Engdahl, E. R., and E. A. Bergman (2001). Validation and generation of reference events by cluster analysis, in *Proceedings of the 23rd Seismic Research Review: Worldwide Monitoring of Nuclear Explosions*, Jackson Hole, Wyoming, 205–214.

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. Seism. Soc. Am.* **88,** 722–743.

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

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only (translated from Geiger's 1910 German article), *Bull. St. Louis Univ.* **8,** no. 1, 56–71.

Gudmundsson, O., and M. Sambridge (1998). A regionalized upper mantle (RUM) seismic model, *J. Geophys. Res.* **103,** 7121–7136.

Hipp, J. R., R. W. Simons, and L. A. Jensen (2002). The GNEM R&E parametric grid data software SuiteTools for creation, access, management, viewing, and export, in *Proceedings of the 24th Annual Seismic Research Review Symposium*, Ponte Vedra Beach, Florida, LA-UR-02-5048, 283–291.

Hole, J. A., and B. C. Zelt (1995). 3D finite difference reflection travel times, *Geophys. J. Int.* **121,** 427–434.

Johnson, M., and C. Vincent (2002). Development and testing of a 3D velocity model for improved event location: A case study for the India-Pakistan region, *Bull. Seism. Soc. Am.* **92,** 2893–2910.

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

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

Laske, G., and T. G. Masters (1997). A global digital map of sediment thickness, *EOS Trans. AGU* **78,** F483.

Lilwall, R. C., and P. D. Marshall (1986). Body wave magnitudes and locations of Soviet underground explosions at the Novaya Zemlya test site, AWRE Report No. 0 17/86, Her Majesty's Stationery Office (HMSO), London.

Mooney, W. D., G. Laske, and T. G. Masters (1998). CRUST 5.1: a global crustal model at 5° X 5°, *J. Geophys. Res.* **103,** 727–747.

Morozov, I. B., E. A. Morozova, S. B. Smithson, P. G. Richards, V. I. Khalturin, and L. N. Solodilov (2005). 3D First-arrival regional calibration model of northern Eurasia, *Bull. Seism. Soc. Am.* **95,** 1535–1560, doi 10.1785/0120030173.

Muller, G. (1971). Approximate treatment of elastic body waves in media with spherical symmetry, *Geophys. J. R. Astr. Soc.* **23,** 435–449.

Murphy, J. R., W. Rodi, M. Johnson, D. D. Sultanov, T. J. Bennett, M. N. Toksoz, V. Ovtchinnikov, B. W. Barker, D. T. Reiter, A. C. Rosca, and Y. Shchukin (2005). Calibration of International Monitoring System (IMS) Stations in Central and Eastern Asia for improved seismic event location, *Bull. Seism. Soc. Am.* **95,** 951–964, doi 10.1785/0120040087.

Myers, S. C. (2001). Methods of travel-time residual declustering for the Knowledge Base Calibration and Integration Tool (KBCIT), Lawrence Livermore National Laboratory Report, UCRL-ID-142521, 14 pp., www.llnl.gov/tid/lof/documents/pdf/245193.pdf (last accessed April 2007).

Myers, S. C. (2005). Reply to "Comment on 'Improving Sparse Network Seismic Location with Bayesian Kriging and Teleseismically Constrained Calibration Events,' by Stephen C. Myers and Craig A. Schultz," by A. Douglas, *Bull. Seism. Soc. Am.* **95,** 370–372.

Myers, S. C., and C. A. Schultz (2000). Improving sparse network seismic location with Bayesian Kriging and teleseismically constrained calibration events, *Bull. Seism. Soc. Am.* **90,** 199–211.

Nataf, H-C., and Y. Ricard (1996). 3SMAC: an a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. Interiors* **95,** 101–122.

Pasyanos, M. E. (2000). Predicting geophysical measurements: testing a combined empirical and model-based approach using surface waves, *Bull. Seism. Soc. Am.* **90,** 790–796.

Pasyanos, M. E., G. A. Franz, and A. L. Ramirez (2006). Reconciling data using Markov Chain Monte Carlo: an application to the Yellow Sea–Korean Peninsula region, *J. Geophys. Res.* **111,** doi 10.1029/2005 JB003851.

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.

Podvin, P., and I. Lecomte (1991). Finite difference computation of travel times in very contrasted velocity models: a massively parallel approach and its associated tools, *Geophys. J. Int.* **105,** 271–284.

Pulliam, J., and R. K. Snieder (1996). Fast, efficient calculation of rays and travel times with ray perturbation theory, *J. Acoust. Soc. Am.* **99,** 383–391.

Reiter, D., W. Rodi, and M. Johnson (2005). Development of a tomographic upper mantle velocity model beneath Pakistan and Northern India for improved regional seismic-event location, *Bull. Seism. Soc. Am.* **95,** 926–940, doi 10.1785/0120040139.

Richards, P. G., J. Armbruster, V. Burlacu, V. F. Cormier, M. D. Fisk, V. I. Khalturin, W-Y. Kim, I. B. Morozov, E. A. Morozova, C. K. Saikia, D. Schaff, A. Stroujkova, and F. Waldhauser (2003). Seismic location calibration for 30 international monitoring system stations in Eastern Asia: final results, in *Proceedings of the 25th Seismic Research Review: Building the Knowledge Base*, Tucson, Arizona, 280–290.

Ritzwoller, M. H., N. M. Shapiro, A. Levshin, E. A. Bergman, and E. R. Engdahl (2003). The ability of a global 3D model to locate regional events, *J. Geophys. Res.* **108,** 2353, doi 10.1029/2002JB002167.

Ruppert, S., D. Dodge, A. Elliott, M. Ganzberger, T. Hauk, and E. Matzel (2005). Enhancing seismic calibration research through software automation and scientific information management, in *Proceedings of the 27th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, Rancho Mirage, California, 937–945.

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

Shapiro, N. M., and M. H. Ritzwoller (2002). Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int.* **151,** 88–105.

Smith, G. P., and G. Ekström (1996). Improving teleseismic event locations using a three-dimensional Earth model, *Bull. Seism. Soc. Am.* **86,** 788–796.

Sultanov, D. D., J. R. Murphy, and K. D. Rubinstein (1999). A seismic source summary for Soviet peaceful nuclear explosions, *Bull. Seism. Soc. Am.* **89,** 640–647.

Tralli, D., and L. R. Johnson (1986). Lateral variation in mantle P velocity from tectonically regionalized tau estimates, *J. R. Astr. Soc.* **86,** 475–489.

Vidale, J. (1988). Finite-difference Calculation of Travel Times, *Bull. Seism. Soc. Am.* **78,** 2062–2076.

Vidale, J. E. (1990). Finite-difference calculation of traveltimes in three dimensions, *Geophysics* **55,** 521–526.

Villaseñor, A., M. H. Ritzwoller, M. P. Barmin, E. R. Engdahl, and A. L. Levshin (2000). Computation of travel times and station correction surfaces in Eurasia using three dimensional velocity models, in *Proceedings of the 22nd Seismic Research Symposium: Technologies for Monitoring the CTBT*, New Orleans, Louisiana, 453–462.

Villaseñor, A., M. H. Ritzwoller, A. L. Levshin, M. P. Barmin, E. R. Engdahl, W. Spakman, and J. Trampert (2001). Shear velocity structure of Central Eurasia from inversion of surface wave velocities, *Phys. Earth Planet. Interiors* **123,** 169–184.

Wessel, P., and W. H. F. Smith (1995). New version of the Generic Mapping Tools released, *EOS Trans. AGU* **76,** 329 [version 3.0].

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. Seism. Soc. Am.* **94,** 897–919.

Yang, X., I. Bondár, K. McLaughlin, and R. North (2001a). Source-specific station corrections for regional phases at Fennoscandian stations, *Pure Appl. Geophys.* **158,** 35–57.

Yang, X., I. Bondár, K. McLaughlin, R. North, and W. Nagy (2001b). Path dependent regional phase travel-time corrections for the International Monitoring System in North America, *Bull. Seism. Soc. Am.* **91,** 1831–1850.

Atmospheric, Earth, and Energy Department
Energy and Environment Directorate
Lawrence Livermore National Laboratory
P.O. Box 808, L-205
Livermore, California 94551
(M.P.F., S.C.M.)

Department of Earth and Atmospheric Sciences
Saint Louis University
3507 Laclede Avenue
St. Louis, Missouri 63103
(K.D.K.)

Manuscript received 31 March 2006.