CRUSTAL STRUCTURE FROM CALIBRATED EARTHQUAKE LOCATIONS

Eric A. Bergman[1], Michael H. Ritzwoller[1], Mikhail Barmin[1], Stephen C. Myers[2], and Eric R. Engdahl[1]

University of Colorado[1] and Lawrence Livermore National Laboratory[2]

Sponsored by the Air Force Research Laboratory[1] and the National Nuclear Security Administration[2]

Award No. FA8718-08-C-0020[1] and LL08-BAA08-69-NDD03[2]
Proposal No. BAA08-69

ABSTRACT

An extensive and high-quality data set of earthquake locations, many of them calibrated to $GT5_{90}$ levels of accuracy, has been compiled for the region that expresses the continental collision between the Arabian and Eurasian plates, with the primary goal of using a calibrated data set to obtain an unbiased baseline model of the crustal and upper mantle velocity structure through tomographic imaging. In regions where ray density is high enough we expect to obtain improved resolution of lateral variations in structure. A specialized multiple event relocation algorithm based on the Hypocentroidal Decomposition (HDC) method has been used to calibrate clusters of earthquakes, using several kinds of available calibration data, including near-source seismic observations, results of InSAR analyses on source location and faulting mechanism, results of waveform modeling (for depth control), and mapped faulting.

Our data set is derived from 35 calibrated clusters containing 1834 events, ranging in magnitude from near 2.0 to greater than 7.0. Of these calibrated events, 1699 qualify for $GT5_{90}$ and 1134 qualify at $GT3_{90}$. All clusters except one (containing 8 events) have been calibrated for origin time. Therefore the observed arrival times from nearly all events in the data set can be taken as estimates of the true travel time of the corresponding seismic phase through the crust and upper mantle of the region. Our analysis also allows us to place constraints on average crustal P and S velocities and crustal thickness in the neighborhood of the cluster. The data set contains over 30,000 individual raypaths for Pn that provide good coverage of much of the region. If condensed into summary rays from each cluster to recording stations, the data set reduces to more than 2700 Pn rays with well-characterized uncertainties. We have conducted Pn tomography using these data in a fairly traditional approach, to obtain an unbiased baseline model for Pn propagation in the region, and to exercise the data set for quality control purposes.

We have also distributed the entire data set to workers at Lawrence Livermore, Los Alamos, and Sandia National Laboratories where it is being used to 1) validate laterally heterogeneous models for earthquake location in the region of interest, 2) test related algorithms for computing theoretical travel times through such models, and 3) conduct tomographic studies at various scales and with various approaches. The calibrated event data set is also being processed with the BAYESLOC algorithm for multiple event relocation, to further test our calibration process. The statistical model has three distinct components: travel-time predictions, arrival-time measurements, and an *a priori* statistical model for each aspect of the multiple-event problem. The prior model allows all available information to be brought to bear on the multiple event system. Prior knowledge on the probabilistic accuracy of event locations, travel-time predictions, and arrival-time measurements can be specified. We take advantage of the output of our HDC calibration methodology to specify these parameters with accuracy much greater than is normally available. Bayesian methodology merges all three components of the hierarchical model into a joint probability formulation. Through these various exercises we have produced a data set of very high quality that has been extensively tested and validated for use in the study of the crustal and upper mantle structure of a continental collision zone. We have initiated the application of this data set to such studies, but there is ample opportunity for further research based on this data set.

OBJECTIVES

This project has the goal of developing seismological data sets that can be used to develop and test more realistic travel time models for seismic phases that propagate in the crust and upper mantle. We focus on two main data sets. The first is a comprehensive, carefully reviewed bulletin of earthquakes in the region at magnitudes of 2.5 and higher. A major goal is to integrate phase readings from local seismograph networks, whose number has increased dramatically in the study region since 1995 with the best available catalog based on regional and teleseismic arrivals. The second data set is a subset of the main catalog for which calibrated locations, at $GT5_{90}$ or better, can be obtained using a multiple-event relocation methodology with near-source data of various types. Traditional time-term Pn tomography is applied to this data set to image broad-scale features of the crust and upper mantle in the region and provide a baseline model for more detailed investigations. The tomographic experiments also help reveal problems in the data sets of earthquake locations and phase identifications, and thus serve a quality control function. The resulting data sets and model results provide a solid foundation for further research.

RESEARCH ACCOMPLISHED

Main Catalog

The main earthquake catalog that has been assembled for this project is shown in Figure 1. The catalog retains all events with a likely magnitude greater than 2.5, regardless of how well they can be located. The data set contains all known arrival time data from stations at local, regional, and teleseismic distances. An effort has been made to relocate all events using the EHB algorithm (Engdahl et al., 1998), and most events have been reviewed with special emphasis on reliable focal depth (Engdahl et al., 2006). Many outlier readings have been flagged.

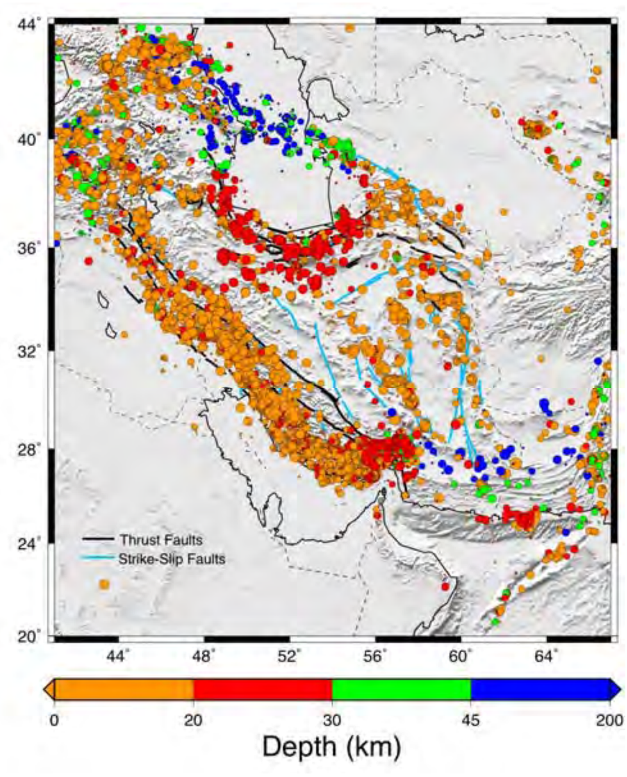

Figure 1. Events in our catalog that satisfy the condition of having secondary azimuth gap < 180°, calculated over the entire range of epicentral distances. Secondary azimuth gap is the largest azimuth gap that is filled by a single station. Focal depths are color-coded.

- The catalog contains 28,267 earthquakes in the study region, dating from 1923-2008.
- The catalog is substantially complete through 2006, in that it incorporates International Seismological Center (ISC) data through that time. Data for events in 2007 and 2008 are mainly from the United States

Geologic Survey (USGS) Earthquake Data Reports (EDR) and local networks (through May 2008). Some data from local stations is reported by the EDR in the second half of 2008, when National Earthquake Information Center (NEIC) analysts have manually extracted the data from seismic network websites for events of interest.

- Depths: Where free depth solutions are not possible and there is insufficient depth phase data (or waveform studies) to constrain a fixed depth location, focal depths are set at regionalized default depths based on the analysis of Engdahl et al. (2006). Depths typically range from about 6-20 km.
- Magnitudes: Events with reported magnitude 2.5 or greater are retained for the catalog, but some events with declared magnitude less than 2.5 are included because they are well recorded. A number of events in the ISS period (1923-1963) have no reported magnitude but are known to be larger events. In the modern period, some events have no reported magnitude but have many phase readings and are probably larger than 2.5.
- 9657 events satisfy a criterion of having secondary azimuth gap < 180°. Secondary azimuth gap is the largest gap filled by only one station, which means that that station has exceptional power in determining the location solution. Events that don't meet this criterion have less reliable locations and are probably not suitable for tomography.

The main catalog is the basis for our detailed studies on location calibration.

Calibrated Locations

Through the course of several research contracts in the ground-based monitoring program, we have developed a methodology for obtaining calibrated locations for clusters of earthquakes by leveraging the availability of near-source data that may be available for only one or a few members of the cluster (e.g., Engdahl and Bergman, 2000; Bergman et al., 2009). Near source data may include short-range seismic observations from permanent or temporary seismograph stations, post-seismic field observations of faulting and remote sensing observations, notably slip models derived from InSAR analysis.

In the study region we have conducted successful calibration studies on 35 earthquake clusters, shown in Figure 2.

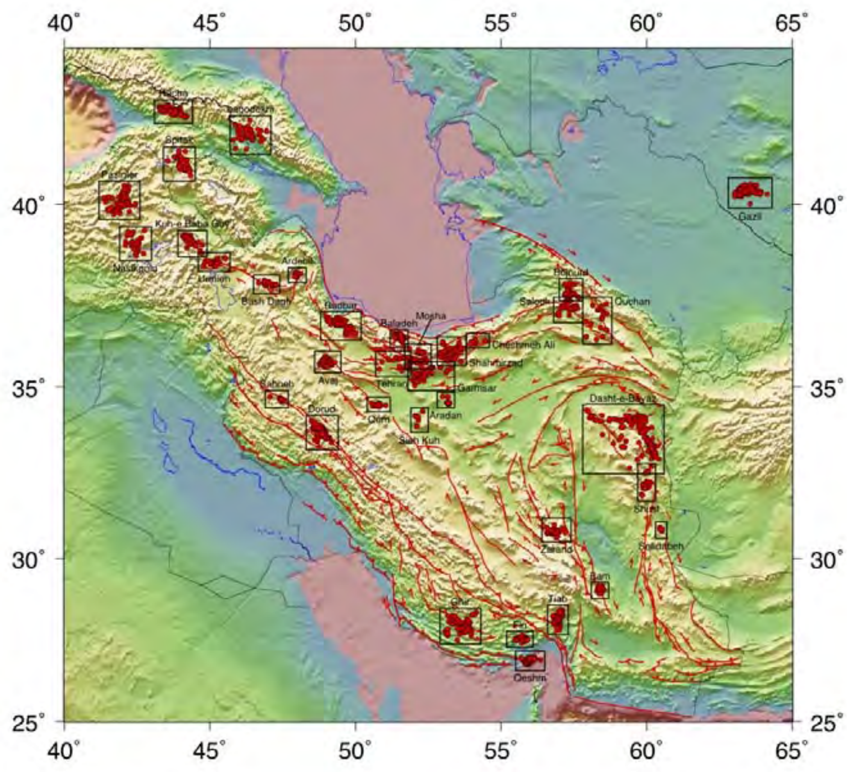

Figure 2. Location map of 35 clusters of earthquakes in the region that have been calibrated for absolute location and contain events that qualify as $GT5_{90}$ or better (~ 1700 events). Solid red circles are all earthquakes in a cluster, regardless of calibration level, but almost all events are $GT10_{90}$ or better.

The clusters contain 1834 events, of which 1699 qualify as $GT5_{90}$. Origin times are also calibrated, although they are coupled to focal depth, which is less well resolved than other source parameters. Through the calibration analysis, the arrival time data sets of these calibrated events undergo a rigorous and thorough grooming process that identifies outliers and leads to robust estimates of empirical reading error for most phase arrivals. The removal of outliers and availability of empirical readings errors is of great value when these data are to be used in tomographic studies or other kinds of analysis where a proper treatment of the error budget is important.

Although the set of GT5 events comprises only about 6% of the main earthquake catalog, they include many of the best-recorded events in the region and the associated arrival times provide quite good ray coverage over much of the region of interest.

Pn Tomography

Pn tomography is conducted with the method described by Ritzwoller et al. (2002). It would certainly be possible to carry out Pn tomography on the entire catalog, but we believe that there is much to be learned by working with the smaller but much higher quality data set of Pn arrivals from calibrated clusters.

The method assumes a constant crustal thickness, but the study region is known to have significant variations in crustal thickness in the range of 40-60 km. In the course of conducting our relocation analysis for calibrated earthquake clusters, we are able to compare observed arrival times of local and regional phase to those predicted by models of the local velocity structure. A single layer crust is often sufficient. With hypocenters and origin times calibrated using direct arrivals, the observed Pn arrivals provide strong constraint on crustal thickness and Pn velocity (Figure 3).

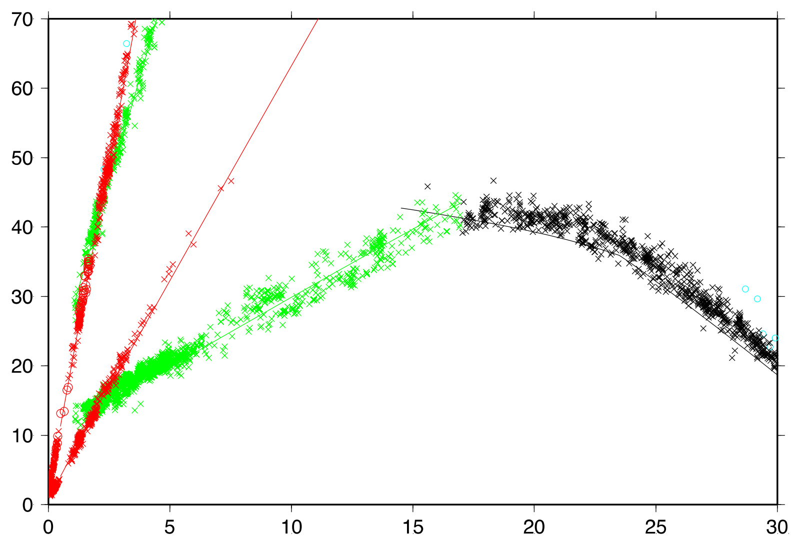

Figure 3. Observed travel times for the Dorud cluster. The hypocenters and origin times are constrained by 305 Pg and Sg readings (both in red) within 0.5° epicentral distance. The observed Pn arrivals (green) are fit by a crustal thickness of 50 km and Pn velocity of 8.05 km/s. Some Pn arrivals are seen at short epicentral distances as secondary arrivals, and many Pg arrivals are seen beyond the Pg-Pn crossover distance (~1.9°). Teleseismic P (black) arrivals are consistently later than predicted by ak135.

Clearly, this estimate of crustal thickness is a spatial average over a "doughnut" centered on the cluster, representing the region of the crust-mantle boundary sampled by the entry points of Pn phases from the cluster at all azimuths. The distribution of crustal thicknesses derived from our calibration studies is shown in Figure 4. Most of the values lie in the range 45-55 km. The thin-crust outlier (35 km) is the Gazli cluster in Uzbekistan, rather far from the continental collision zone. The thick crust outlier (63 km) is the Bam cluster in the southeast of the study area. For our Pn tomography we assume a crustal model that is a stretched version of the 2-layer ak135 crust with a thickness of 50 km. When the actual crustal thickness differs from this assumption, the associated error should be partially absorbed in the cluster static correction.

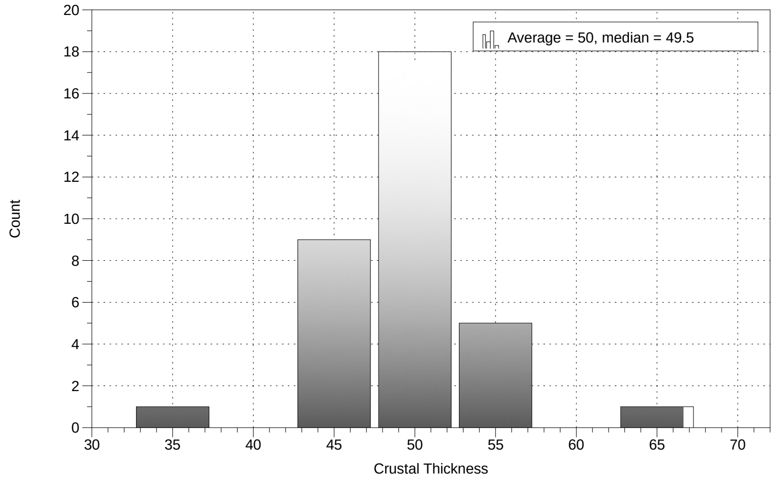

Figure 4. Histogram of crustal thickness inferred for the calibrated clusters by fitting the arrival times of Pn phases.

We have found that a sub-crustal P velocity in the range 8.05-8.15 km/s provides a good average fit to Pn arrivals for most clusters.

In all tomographic inversions, accurate imaging is strongly linked to the presence of crossing raypaths. For Pn tomography we can improve this aspect of the inversion by using longer raypaths, but the utility of this strategy is offset by the increasing scatter in Pn arrivals with distance beyond about 10°, as the rays interact with complex and laterally-variable upper-mantle structure. To investigate how much ray coverage is lost by restricting our tomography to shorter raypaths, we made plots of the ray coverage and ray density for two cases: using all Pn arrivals in the distance range 2-17° and using only data in the distance range 2-10° (Figure 5).

Certainly, the raypath coverage is improved by using Pn data out to 17°, but coverage of the region inside the "network" of calibrated clusters (Figure 2) is not greatly improved. Given the increased noise in the data at greater distances and certain concerns about the suitability of our tomographic algorithm to such cases, we have chosen to restrict our inversions for now to the distance range 2-10°.

In traditional time term tomography for Pn velocity static corrections are made at the source and receiver before the inversion for Pn velocity. These corrections are meant to account for errors in the assumed velocity structure of the corresponding crustal leg. In addition, the source term collects contributions from errors in the assumed source location, which includes depth and origin time. This process adds a large number of extra parameters to the inversion and raises concerns that much of the desired signal has been absorbed into the statics. Because the relative locations and times of all events in a calibrated cluster have been tightly constrained in the relocation exercise, we can restrict the source terms to individual clusters (Figure 6). These cluster corrections are a valuable check on the correctness of our calibration process. In theory the term should only reflect error in the assumed crustal velocity structure and crustal thickness. Since we have estimated crustal thickness for each cluster, as described above, there should be a correlation between the size of the cluster correction and the difference between the inferred crustal thickness for each cluster and the 50 km thickness assumed for tomography. As Figure 7 shows, there is a rough correlation of the expected sense, but most of the data points are close to 50 km and there is enough scatter to make the correlation a weak one. The scatter is attributable to several sources, including unmodelled variations in crustal velocities, absorption of Pn velocity variations by clusters with poor azimuthal distribution of Pn phases, and, of course, errors in calibration and associated estimates of crustal thickness.

The pattern and magnitudes of station corrections are shown in Figure 8. These corrections should reflect only departures of the true crustal velocity structure under each station from the assumed 1-D starting model. In some cases there can be a contribution from gross errors in station elevation (some stations have no entry for elevation in the standard registry). It is unlikely that there are errors in station location large enough to create a significant signal in Figure 8, but it cannot be ruled out. These matters will be reviewed. The largest source of these signals is likely to be stations for which we have Pn data covering only a limited range of azimuths. In that case, signals related to Pn velocity variations can easily be absorbed into station corrections. This is most likely to happen for stations near the edge of our study area.

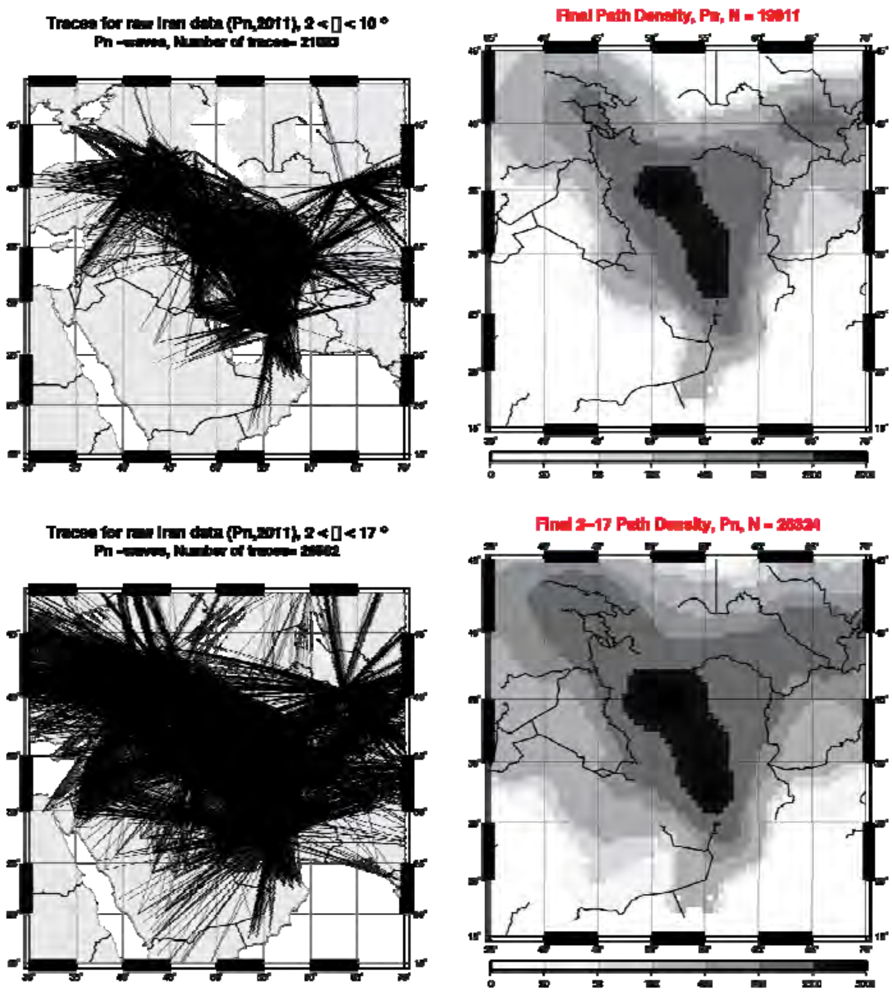

Figure 5. Raypaths and raypath density for Pn arrivals in the distance range 2-10° (top), and the distance range 2-17° (bottom). Grid size for raypath density is 0.5°. Scales for raypath density are the same.

In Figure 9 we show the results of an inversion for Pn velocity, using the arrivals from our calibrated clusters in the distance range 2-10°. Velocity variations relative to our baseline model, which uses the ak135 sub-crustal velocity of 8.04 km/s, are ±2%. We have observed reliable variations in travel time that would require variations in velocity of at least this much if they are to be explained solely by Pn velocity. However, we do not suggest that this image is a reliable image of the variations in Pn velocity in the region, for the following reasons:

1) We think it is highly likely that the propagation of Pn in this region deviates from the simple model of a refracted wave along the crust-upper mantle interface, because of strong variations in Moho depth, and that significant travel time anomalies are produced by those deviations. Any such anomalies will be partitioned in unpredictable ways between cluster statics, station statics, and the pattern of Pn velocities across the region.

2) We are still performing quality control checks on the data set, guided by the results of tomography. The most important of these checks concerns the calibration of entire clusters, but we also use the results of

tomography to investigate anomalies relates to individual events within cluster, individual stations, and individual readings.

3) We are still learning how to do tomography with a highly groomed, calibrated data set. Some of the procedures that are considered reasonable for the typical data set used in this kind of analysis, such as rejection of outliers, need to be modified or eliminated in this case. Another important aspect is weighting of the data. We have empirical reading errors for all readings that should be used in weighting the inversion.

As we refine our data set and our approach to tomography to take advantage of its unusual qualities, we expect the imagery of Pn velocity to be altered. Our main goal, however, has never been to solve the problem of Pn velocity variations in this region, but rather to use Pn tomography to refine the data set to a quality and reliability which will make it suitable for more sophisticated approaches to inferring crustal and upper mantle properties in the study region..

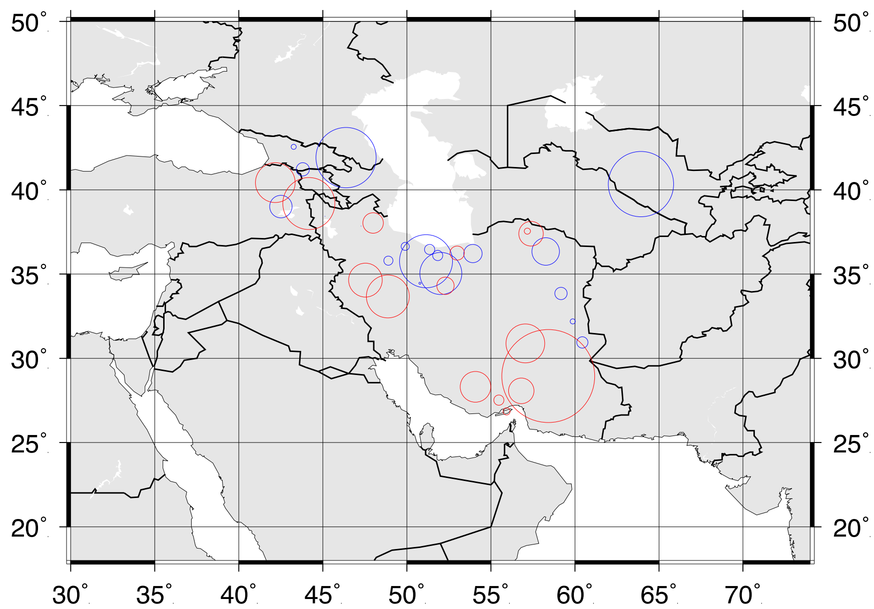

Figure 6. Cluster corrections for Pn tomography using the data from 35 calibrated clusters.

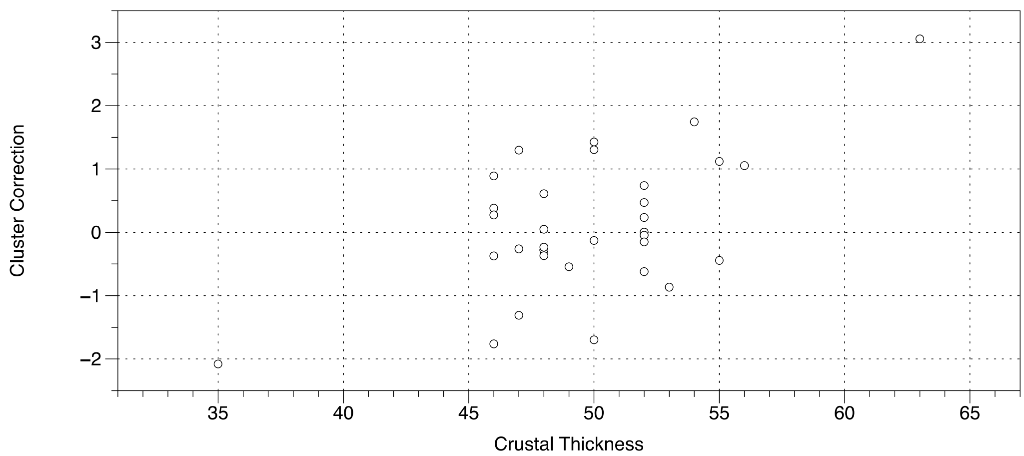

Figure 7. Cluster corrections (Figure 6) plotted against inferred crustal thickness (Figure 4). Cluster corrections are based on a constant 50 km thick crust, so clusters with inferred cluster thickness near 50 km should have small cluster corrections. We expect clusters with thicker crust to have a more positive cluster correction, those with thinner crust to have a more negative correction.

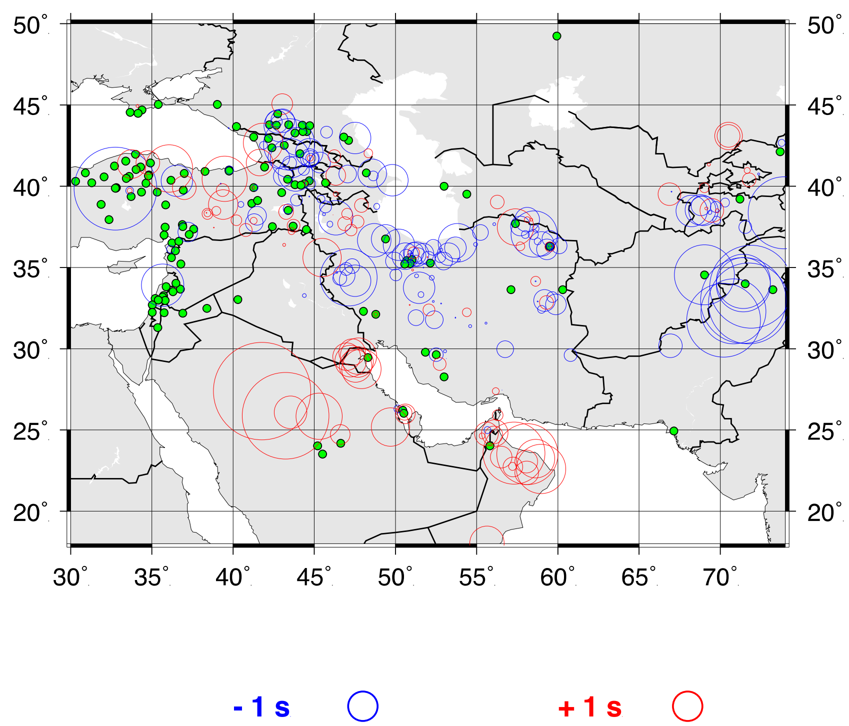

Figure 8. Station corrections for Pn tomography using data from 35 calibrated clusters. Station corrections are only calculated for N>5. Stations with fewer readings are shown as small green dots.

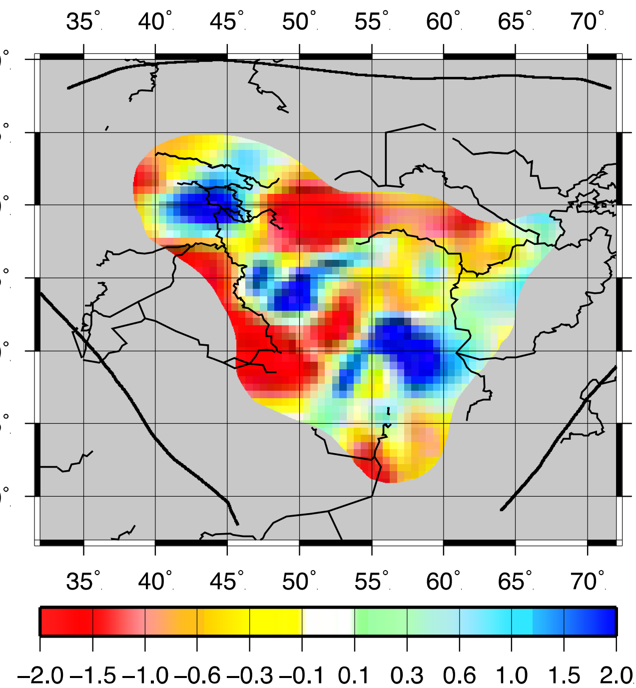

Figure 9. Pn velocity variations determined from Pn arrival time data in the distance range 2-10° from the set of calibrated clusters. Velocity is expressed as percentage relative to the baseline model, which uses a Pn velocity from ak135 (8.04 km/s).

Distribution of Data Sets
In May 2011, we distributed the entire data set to workers at Lawrence Livermore, Los Alamos, and Sandia National Laboratories where it is being used to 1) validate laterally heterogeneous models for earthquake location in the region of interest, 2) test related algorithms for computing theoretical travel times through such models, and 3) conduct tomographic studies at various scales and with various approaches.

The calibrated event data set is also being processed with the BAYESLOC algorithm for multiple event relocation, to further test our calibration process. The statistical model has three distinct components: travel-time predictions, arrival-time measurements, and an *a priori* statistical model for each aspect of the multiple-event problem. The prior model allows all available information to be brought to bear on the multiple event system. Prior knowledge on the probabilistic accuracy of event locations, travel-time predictions, and arrival-time measurements can be specified. We take advantage of the output of our HDC calibration methodology to specify these parameters with accuracy much greater than is normally available. Bayesian methodology merges all three components of the hierarchical model into a joint probability formulation.

Through these various exercises we have produced a data set of very high quality that has been extensively tested and validated for use in the study of the crustal and upper mantle structure of a continental collision zone.

CONCLUSIONS AND RECOMMENDATIONS

We have compiled a carefully-reviewed master catalog of more than 28,000 earthquakes and associated arrival time data in the study region, for the period 1923–2008. We have carried out location calibration studies of 35 clusters extracted from the main catalog, deriving almost 1700 earthquakes that qualify as $GT5_{90}$ or better. We have conducted Pn tomography on the calibrated data set to explore baseline models of crustal structure and upper mantle velocity in the region and to help reveal problems in the data set. We have distributed the entire data set to researchers at national laboratories who are exercising it in various research efforts.

We consider that this project has been very successful in demonstrating the power of high quality data sets of calibrated earthquake locations and associated arrival time data for advancing studies of Earth structure. Even where insufficient data exist to carry out tomography using only calibrated data, such data sets can provide extremely valuable constraint on models derived from larger but less reliable data sets.

ACKNOWLEDGEMENTS

This project has benefitted extensively from the cooperation and generosity of many seismologists in the region, who have provided access to data sets of great value. We have benefitted greatly from conversations about earthquake location methodologies and statistics with Istvan Bondár, Bill Rodi, and Jim Dewey.

REFERENCES

Bergman, E. A., E. R. Engdahl, M. H. Ritzwoller, and S. C. Myers (2009). Crustal structure from in-country and ground-truth data, in *Proceedings of the 31st Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies,* LA-UR-09-05276, Vol. I, pp. 22–31.

Engdahl, E. R., R. D. Van der Hilst and R. P. Buland (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.* 88 722–743.

Engdahl, E. R., J. A. Jackson, S. C. Myers, E. A. Bergman, and K. Priestley (2006). Relocation and assessment of seismicity in the Iran region, *Geophys. J. Int.*, 167: 761–778.

Engdahl, E.R., and E.A. Bergman (2000). Identification and validation of reference events within the area being regionally monitored by IMS stations in Asia and North Africa, in *Proceedings of the 22nd Seismic Research Symposium: Planning for Verification of and Compliance with the Comprehensive Nuclear-Test-Ban Treaty,* Defense Threat Reduction Agency.

Ritzwoller, M. H., M. P. Barmin, A. Villaseñor, A. L. Levshin, and E. R. Engdahl (2002). Pn and Sn tomography across Eurasia to improve regional seismic event locations, *Tectonophysics*, 358: 39–55.