IMPROVED GROUND TRUTH IN SOUTHERN ASIA USING IN-COUNTRY DATA, ANALYST WAVEFORM REVIEW, AND ADVANCED ALGORITHMS

Eric R. Engdahl[1], Eric A. Bergman[1], Stephen C. Myers[2], and Floriana Ryall[2]

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

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Development
Office of Defense Nuclear Nonproliferation

Contract No. DE-FC52-03NA99516[1] and W-7405-ENG-48[2]

EXECUTIVE SUMMARY

A new catalog of seismicity at magnitudes above 2.5 for the period 1923-2008 in the Iran region is assembled from arrival times reported by global, regional, and local seismic networks. Using in-country data we have formed new events, mostly at lower magnitudes that were not previously included in standard global earthquake catalogs. The magnitude completeness of the catalog varies strongly through time, complete to about magnitude 4.2 prior to 1998 and reaching a minimum of about 3.6 during the period 1998-2005. Of the 25,722 events in the catalog, most of the larger events have been carefully reviewed for proper phase association, especially for depth phases and to eliminate outlier readings, and relocated.

To better understand the quality of the data set of arrival times reported by Iranian networks that are central to this study, many waveforms for events in Iran have been re-picked by an experienced seismic analyst. Waveforms at regional distances in this region are often complex. For many events this makes arrival time picks difficult to make, especially for smaller magnitude events, resulting in reported times that can be substantially improved by an experienced analyst. Even when the signal/noise ratio is large, re-picking can lead to significant differences. Picks made by our analyst are compared with original picks made by the regional networks. In spite of the obvious outliers, the median (-0.06 s) and spread (0.51 s) are small, suggesting that reasonable confidence can be placed in the picks reported by regional networks in Iran.

This new catalog has been used to assess focal depth distributions throughout Iran. A principal result of this study is that the geographic pattern of depth distributions revealed by the relatively small number of earthquakes (~167) with depths constrained by waveform modeling (+/- 4 km) are now in agreement with the much larger number of depths (~1229) determined using reanalysis of ISC arrival-times (+/-10 km), within their respective errors. This is a significant advance, as outliers and future events with apparently anomalous depths can be readily identified and, if necessary, further investigated. The patterns of reliable focal depth distributions have been interpreted in the context of Middle Eastern active tectonics. Most earthquakes in the Iranian continental lithosphere occur in the upper crust, less than about 25-30 km in depth, with the crustal shortening produced by continental collision apparently accommodated entirely by thickening and distributed deformation rather than by subduction of crust into the mantle. However, intermediate-depth earthquakes associated with subducted slab do occur across the central Caspian Sea and beneath the Makran coast.

A multiple-event relocation technique, specialized to use different kinds of near-source data, is used to calibrate the locations of 24 clusters containing 901 events drawn from the seismicity catalog. The absolute locations of these clusters are fixed either by comparing the pattern of relocated earthquakes with mapped fault geometry, by using one or more cluster events that have been accurately located independently by a local seismic network or aftershock deployment, by using InSAR data to determine the rupture zone of shallow earthquakes, or by some combination of these near-source data. This technique removes most of the systematic bias in single-event locations done with regional and teleseismic data, resulting in 624 calibrated events with location uncertainties of 5 km or better at the 90% confidence level ($GT5_{90}$). For 21 clusters (847 events) that are calibrated in both location and origin time we calculate empirical travel times, relative to a standard 1-D travel time model (ak135), and investigate event to station travel-time anomalies as functions of epicentral distance and azimuth. Substantial travel-time anomalies are seen in the Iran region which make accurate locations impossible unless observing stations are at very short distances (less than about 200 km) or travel-time models are improved to account for lateral heterogeneity in the region. Earthquake locations in the Iran region by international agencies, based on regional and teleseismic arrival time data, are systematically biased to the southwest and have a 90% location accuracy of 18-23 km, with the lower value achievable by applying limits on secondary azimuth gap. The data set of calibrated locations reported here provides an important constraint on travel-time models that would begin to account for the lateral heterogeneity in Earth structure in the Iran region, and permit seismic networks, especially the regional ones, to obtain in future more accurate locations of the earthquakes in the region.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

GOALS AND OBJECTIVES

This research seeks to improve the database of ground-truth information and velocity models useful for calibration in southern Asia with the following objectives: (1) Aggressive pursuit of in-country data acquisition, especially the collection of ground truth at GT5 level or better for events of magnitude 2.5 and larger recorded by dense local networks, including associated velocity models; (2) Expanded analyst review of relevant regional waveforms for ground-truth events by the comprehensive re-picking of phase arrival times from all available waveforms, with special attention to the regional phases Pg, Pb, Pn, Sg, Sb, and Sn; and (3) Application of advanced algorithms, specifically multiple event relocation, to refine and validate all available ground-truth data, to achieve the optimal selection of data for analysis, to better understand the uncertainties of the results, and to handle the error budget as realistically as possible.

In terms of actual accomplishments all the stated goals and objectives of the project have been met: (1) In-country data acquisition, including many event waveforms, has been actively pursued and all known available arrival-time data from regional and local networks in the Iran region acquired. In combination with data reported by global networks this data acquisition has allowed the construction of a catalog of events in the region for events of magnitude 2.5 and larger, permitting a reassessment of regional seismicity and tectonics, especially focal depth distribution. It has also allowed for the collection and validation of 624 ground truth events with location uncertainties of 5 km or better at the 90% confidence level ($GT5_{90}$); (2) An experienced analyst has reviewed several hundred regional waveforms, often complex, from stations in the Iran region and picked substantially improved arrival times of the regional phases Pg, Pb, Pn, Sg, Sb, and Sn, that have been incorporated in the seismicity catalog for the region; and (3) Much development work has been done on the multiple event location code as it is applied to the ground truth problem, with emphasis on statistical rigor and robustness. A significant advance in location accuracy that the code now makes possible is the ability to make empirical estimates of reading error for individual station-phase combinations, based on the spread of path-corrected, normalized residuals. The code also now includes estimated uncertainties of the ground truth data, which yields more accurate estimates of the uncertainties of the final estimates of absolute (calibrated) locations.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

PROJECT ACTIVITIES

Introduction

Earthquake location was the first geophysical problem to be formulated as an inverse problem (Geiger, 1910, 1912), and the first to be coded for solution by computer (Bolt, 1960; Gunst and Engdahl, 1962) but there has been surprisingly little research into the true uncertainties of routinely determined earthquake locations until about the past 10 years, largely motivated by the desire to monitor a Comprehensive Test Ban Treaty (e.g., Bondár et al., 2001). Most earth scientists have become accustomed to interpreting earthquake locations from standard catalogs in seismotectonic and tomographic studies without having a clear idea of the accuracy of those locations. The EHB catalog (Engdahl et al., 1998) uses several methods to reduce the likelihood of the largest errors, but it too suffers from uncertain scaling of the confidence ellipses that describe formal uncertainty, and systematic location bias of unknown direction and magnitude.

The three basic problems regarding location accuracy are 1) violations of the statistical assumptions underlying standard location algorithms, 2) failure of theoretical travel time models to adequately represent the true variability of travel times in the Earth (Bondár et al., 2004), and 3) erroneous arrival time measurements. Progress in both areas has been made using a multiple event relocation procedure (Bergman and Engdahl, in preparation), which leads to estimates of earthquake location, using regional and teleseismic arrival time data, with uncertainties that are often less than 5 km. We have applied this methodology extensively in the Iran region and produced a catalog of calibrated earthquake locations that can be used for a variety of purposes, notably for structural studies and for improving the capability of national seismological agencies to effectively monitor the seismic activity in a region which is considered to have one of the highest levels of earthquake hazard in the world (Tavakoli and Ghafory-Ashtiany, 1999).

With the new catalog of calibrated locations we explore the statistics of location accuracy in the study region by several standard earthquake location catalogs. Even for a carefully-reviewed procedure such as the EHB methodology, the mean location error is about 15 km. Location bias in the Iran region has a preferred direction but there is no way to adequately predict the direction of bias for any given event. Improved location accuracy in this region can only be achieved by developing travel time models that adequately describe the three-dimensional velocity structure of the crust and upper mantle in the region. Catalogs of calibrated earthquake locations play a central role in such efforts.

A New Seismicity Catalog for the Iran Region

As a first step towards an improved understanding of the location of seismic activity in the Iran region, a catalog of instrumentally recorded events during the period 1923-2008, for the region bounded by 20-44°N and 41-67°E, has been assembled. For the period prior to 1964 our catalog is based on the catalog of the International Seismological Summary (ISS) (Engdahl and Villaseñor, 2002; Villaseñor and Engdahl, 2007). Data of the International Seismological Centre (ISC) was used from 1964 through June, 2006, the latest data available at the time. Data from the Iranian national networks, ISTN and INSN, begin in the mid-1990s and early 2000s, respectively. However, our catalog is comprehensive with respect to these networks only through 2005. We have brought in data for more recent events from these networks for studies of specific calibrated clusters. Therefore the period during which our catalog can be considered comprehensive is 1964 through 2005.

In general we have considered only events with magnitude 2.5 and greater, but a few events with smaller magnitudes have been included if they are well constrained. We have not considered differences between different magnitude scales for this purpose. Engdahl et al. (2006) presented an earlier version of this catalog, with particular attention to the distribution of focal depths in the Iran region.

Our study depends critically on arrival time data from two permanent seismograph networks operating in Iran:

- Iranian National Broad-Band Seismic Network (INSN), operated by the International Institute for Earthquake Engineering and Seismology (IIEES), consists of 16 broadband, satellite-telemetered stations. Most stations have been installed since 2000.
- Iranian Seismic Telemetry Network (ISTN), operated by the Iranian Seismological Center (IRSC) at the University of Tehran's Institute of Geophysics, consists of 73 short-period, digital, telemetered stations organized in 10 sub-networks. The earliest installations occurred in 1995.

There were few seismic stations operating in Iran before 1995. The seismic stations associated with these networks and other permanent stations in the Iran region are shown in Figure 1.

We have also used arrival time data from temporary deployments by several institutions. Bulletin arrival-time data from

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Iranian and other national seismic networks in the region, as well as phase picks from an experienced analyst who reviewed waveforms of particular interest for specific events, have been compiled and, where possible, associated with arrival-time data from known earthquakes reported by international agencies. However, with the Iranian data we have also formed many new events, mostly at lower magnitudes that were not previously included in standard global earthquake catalogs.

This combined catalog of 25,722 events has been relocated using the Engdahl et al. (EHB; 1998) methodology. Epicenters from the resulting catalog are plotted in Figure 2. Significant scatter in the distribution of seismicity is evident since many events at lower magnitudes are poorly located and are not well constrained in azimuth, especially in parts of the region where there are few stations.

In order to improve the resolution of seismicity patterns shown in Figure 2 we apply a secondary azimuth gap (the largest azimuth gap filled by a single station) criteria of less than 180° to station distributions at all distances. This reduces the database to over 7,000 events and significantly reduces the scatter (Figure 3).

The new EHB catalog has been used to assess focal depth distributions throughout the Iran region. A principal result of that study is that the geographic pattern of depth distributions revealed by the relatively small number of earthquakes (~167) with depths constrained by waveform modeling (+/- 4 km) are now in agreement with the much larger number of depths (~1229) determined using reanalysis of ISC arrival-times (+/-10 km), within their respective errors. This is a significant advance, as outliers and future events with apparently anomalous depths can be readily identified and, if necessary, further investigated.

Catalog Completeness

The completeness of our new catalog of seismicity in the Iran region varies greatly as a function of time (as a proxy for station coverage). The catalog is most comprehensive, in the sense of being based on ISC data and all the available data from the Iranian national networks, for the period 1964 through 2005. The number of events in the catalog prior to 1964 (76) is too small to do a useful analysis of completeness, but the completeness magnitude is probably about 7.0 (Engdahl and Villaseñor, 2002). The major breakpoint for catalog completeness in the 1964-2005 period is the late 1990s, as the ISTN network was deployed. The effect of the introduction of the INSN in the early 2000s is smaller because it does not represent a large increase in number of stations or major improvement in station coverage. Catalog completeness also varies strongly throughout the region of the catalog, because of the uneven distribution of recording stations (Figure 1). Coverage in the southeastern part of Iran is still notably poor, although it has improved with recent installations of stations in the region by both the INSN and ISTN. We have estimated catalog completeness for two time periods (Figure 7), to illustrate the effect of the establishment of the Iranian national networks. Prior to 1998 the completeness magnitude is about 4.2, and since 1998 the catalog is complete to about magnitude 3.5.

Tectonics

The patterns of reliable focal depth distributions have been interpreted in the context of Middle Eastern active tectonics. Most earthquakes in the Iranian continental lithosphere occur in the upper crust, with the crustal shortening produced by continental collision apparently accommodated entirely by thickening and distributed deformation, rather than by subduction of crust into the mantle. In the Zagros Mountains nearly all earthquakes are confined to the upper crust (depths < 20 km), and there is no evidence for a seismically active subducted slab dipping NE beneath central Iran. By contrast, in southeastern Iran, where the Arabian seafloor is being subducted beneath the Makran coast, low-level earthquake activity occurs in the upper crust as well as to depths of at least 150 km within a northward-dipping subducting slab. Near the Oman Line, a transitional region between inter-continent collision in the Zagros and oceanic subduction in the Makran, seismicity extends to depths up to 30-45 km in the crust, consistent with low-angle thrusting of Arabian basement beneath central Iran in this region. In north-central Iran, along the Alborz mountain belt, seismic activity occurs primarily in the upper crust but with some infrequent events in the lower crust, particularly in the western part of the belt (the Talesh), where the south Caspian basin underthrusts NW Iran. Earthquakes that occur in a band across the central Caspian, following the Apsheron sill between Azerbaijan and Turkmenistan, have depths in the range 30-100 km, increasing northwards. These are thought to be connected with either incipient or remnant northeast subduction of the south Caspian Basin basement beneath the east-west trending Apsheron-Balkhan sill. Curiously, in this region of genuine mantle seismicity, there is no evidence for earthquakes shallower than 30 km.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Waveform Analysis

The workers making routine picks at regional networks usually have little experience or training and there is usually not time or motivation for a more experienced seismologist to do a careful review. To better understand the quality of the data set of arrival times reported by Iranian networks that are central to this study, we have had many waveforms re-picked by an experienced seismic analyst (Figure 4).

Waveforms at regional distances in this region are often complex. For many events this makes arrival time picks are difficult to make, especially for smaller magnitude events, resulting in reported times that can be substantially improved by an experienced analyst. Even when the signal/noise ratio is large, re-picking can lead to significant differences (Figure 5).

In Figure 6 we compare the picks made by our analyst with the original picks made by the regional networks. In spite of the obvious outliers, the median (-0.06 s) and spread (0.51 s) are small, suggesting that reasonable confidence can be placed in the picks reported by the regional networks in this region.

Multiple Event Relocation for Calibrated Locations

Our method for determining calibrated earthquake locations is fully described in Bergman and Engdahl (in preparation). It has been applied in a number of prior studies, including Ritzwoller et al., 2003; Walker et al., 2005; Parsons et al. 2006; Biggs et al., 2006; Tatar et al., 2007; and Bondár et al. 2008. We provide only a summary of the technique here.

The method is based on the Hypocentroidal Decomposition (HDC) method for multiple event relocation described by Jordan and Sverdrup (1981). The essence of the HDC algorithm is the use of orthogonal projection operators to separate the relocation problem into two parts:

- The cluster vectors, which describe the relative locations in space and time of each event in the cluster. They are defined in kilometers and seconds, relative to the current position of the hypocentroid.
- The hypocentroid, which is defined as the centroid of the current locations of the cluster events. It is defined in geographic coordinates and Coordinated Universal Time (UTC)

The cluster vectors are defined only in relation to the hypocentroid. The hypocentroid can be thought of as a virtual event with geographic coordinates and origin time in UTC. The orthogonal projection operators act on the data set of arrival times to produce a data set that includes only data that actually bears on the relative location of cluster events, i.e., multiple reports of a given seismic phase at the same station for two or more events in the cluster.

The hypocentroid is located very much as an earthquake would be, except that the data are drawn from all the cluster events. Thus it is typical for the hypocentroid to be determined by many thousands of readings. Nevertheless, the hypocentroid is subject to unknown bias because the theoretical travel times (typically ak135, Kennett et al., 1995) do not fully account for the three-dimensional velocity structure of the Earth. Geographic locations for the cluster events are found by adding the cluster vectors to the hypocentroid.

The HDC method works iteratively. At each iteration two inversions are performed, first for the cluster vectors relative to the current hypocentroid, then for an improved hypocentroid. The cluster vectors are added to the new hypocentroid to obtain updated absolute coordinates for each event. The convergence criteria are based on the change in relative location of each event (0.5 km) and the change in the hypocentroid (0.005°). The convergence limits for origin time and depth, for cluster vectors and hypocentroid, are 0.1 s and 0.5 km, respectively. Convergence is normally reached in 2 or 3 iterations.

The data sets used for the two problems need not be (and usually are not) the same. Because the inverse problem for changes in cluster vectors is based solely on arrival time differences, baseline errors in the theoretical travel times drop out and it is desirable to use all available phases at all distances outside the immediate source region. For the hypocentroid, baseline errors in theoretical travel times are more important and one may wish to limit the data set to a phase set, e.g., teleseismic P arrivals in the range 30-90°, to achieve a more stable result. The choice of data set for determining the hypocentroid has great importance in the "direct" calibration method described below.

Similarly, weighting schemes are different for the two inversions, reflecting the different natures of the two problems. Empirical reading errors for each station-phase pair are used in weighting data for estimating both the hypocentroid and cluster vectors, but the uncertainty of the theoretical travel times, which are estimated empirically for each phase from the residuals of previous runs, is relevant only to the hypocentroid.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

A considerable amount of work goes into estimation of empirical reading errors from the specific arrival time data set. This estimate is based on a robust estimator of spread (Croux and Rousseeuw, 1992) applied to the travel time residuals for a specific station and phase. In addition to their use in weighting the arrival time data for inversion, we also use these empirical reading errors to detect outliers in the data, which are flagged. Because outlier readings can cause good readings to have large residuals, the process needs to be done incrementally, starting with the largest residuals, to avoid rejection of good data. Then a new inversion is performed with the modified arrival time data set. This "cleaning" process is crucial in providing a self-consistent statistical framework for estimating location uncertainties. It is continued out until the residuals satisfy a 3σ criterion, using the current empirical reading error as the estimate of σ.

Until this point, the HDC algorithm is used only to obtain improved relative locations for the cluster events, with a geographic location for the cluster (the hypocentroid) that is biased to an unknown degree by unmodeled Earth structure convolved with the unbalanced distribution of reporting seismic stations. The calibration process attempts to remove this bias.

Calibration of a cluster is done in to ways, which we refer to as "indirect" and "direct" calibration.

Indirect Calibration

If the location and origin time of one or more of the cluster events can be specified with high accuracy from independent information, we can calibrate the entire cluster by shifting it in space and time to optimally match the known location of the calibration event(s). The most common source of such independent information are temporary seismic network deployments that capture an event with a large number of stations at very short epicentral distances, and which is also large enough to be well recorded at regional and teleseismic distances. Aftershock studies are a frequent source of such data (e.g., Tatar et al., 2007). We normally obtain the temporary network arrival time data and relocate calibration events ourselves, using a local velocity model when available, to ensure reliable locations. We have also used InSAR data for this purpose, but this requires the use of at least some seismic data at short distance to calibrate origin time (e.g., Parsons et al., 2006). In a few cases, mapped faulting from large events can be used to help constrain the location of calibration events.

When we use the direct calibration approach we take into account the uncertainty of the calibration data, and when there is more than one calibration event we also include a contribution to uncertainty to reflect any discrepancy between the relative locations of the calibration events and the cluster vectors of the corresponding events.

Direct Calibration

It is sometimes the case that there are a few permanent seismic stations close to a cluster, but that no single event is well-enough recorded to reach the level of accuracy necessary to serve as a calibration event. On the other hand, the handful of local seismic stations may have recorded many events in the cluster, so that the number of "short distance" readings is rather large, and well-enough distributed to allow the hypocentroid to be located using only these data. We have found that for clusters in Iran, good results can be obtained as long as the epicentral distance is kept less than about 180 km. At greater distances we see rapidly increasing scatter in residuals at different azimuths.

In any case where we have the data to locate one or more calibration events for the indirect method, we also have the option to use those same data in the direct calibration method. The decision on which to use is made on a case-by-case basis, because characteristics of the data sets may lead to a better result with one method than the other. Of course, if we are using InSAR data or other geological information to constrain the calibration, we must use the indirect method.

Calibration Levels

It has become common to use the GTX formulation ("Ground Truth to within X km") to describe the location accuracy of seismic events, although actual ground truth information is rarely available. As Bondár et al. (2004) point out, it is also necessary to specify a confidence level for such a designation. In our study confidence ellipses are calculated at the 90% level. We always use the confidence level subscript when making a quantitative statement or when referring to a specific calibrated data set, but it may be dropped in a more general context.

Moreover, it is necessary to be clear about what metric is used for the GT classification. We take the nearest integer of the length of the semi-major axis of the 90% confidence ellipse that includes both the uncertainty of relative location and the uncertainty of the calibration process. In our experience it is possible to achieve GT levels of 2-3 km in the most favorable circumstances.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Depending on the purpose to which one might want to apply a set of reference events, the desired level of accuracy will vary. GT5 is the most commonly discussed level for calibrated earthquake locations, and GT5 (or better) locations certainly represent a significant improvement in accuracy over what is available in standard catalogs. In our calibration studies we have generally discarded events whose relative location within the cluster cannot be determined to better than 10 km (semi-major axis length).

Focal Depth

All of our calibration studies have been performed with fixed focal depths. Considerable effort has gone into obtaining the best possible estimates of depth, but the kind of data needed to include depth as a free parameter in the relocations (direct arrivals at stations within one or two focal depths of the epicenter) is rarely available. Engdahl et al. (2006) have made a careful review of the entire catalog of earthquakes in the Iran region with special attention to focal depths derived from observed depth phases and body waveform studies. Additional review has been performed for the earthquakes in cluster calibration studies. Our general philosophy has been to establish a default depth for each cluster, based on the best-constrained events in the cluster. Tests have shown that errors of less than about 15 km in depth have negligible effect on the estimated epicenter. Of course, errors in assumed depth will lead to coupled errors in estimated origin time for calibrated events.

Engdahl et al. (2006) have shown that the vast majority of seismicity in the Iran region (except for a few special areas such as the Makran and Caspian Basin, which have subducted slabs) occurs at upper-crustal depths, less than about 25-30 km. We believe most of the depths assumed for these clusters are in error by no more than 10 km.

Calibrated Earthquake Clusters in the Iran Region

We have carried out successful calibration studies for 24 earthquake clusters, containing 901 events, in the Iran region (Figure 8). All the clusters except one (Lagodekhi) are inside Iran.

Bergman et al. (2008) discuss the analysis of each cluster in detail in the Electronic Supplement to that paper, but a summary of the clusters is presented in Table 1.

Three clusters (Bam, Bastak, and Shusf) have calibrated locations but lacked data with which to calibrate origin times. Three others (Ardebil, Firuzabad, and Nahavand) are calibrated only at levels that produce no reference events at GT5 or better, except for the single calibration event for Firuzabad. Therefore, there are 18 clusters that are calibrated at useful levels, i.e., producing events at GT5 or better with calibrated origin times.

The type of calibration used for these clusters is about evenly split between direct (11 cases) and indirect (13) methods. For 8 clusters we had one or more constraints on location from InSAR analyses or mapped faulting. Of the 10 clusters calibrated with the indirect method and seismic calibration events, six had more than one calibration event.

Of the 901 events in these clusters, 624 events (69%) qualify as $GT5_{90}$ or better, and 268 events (30%) qualify as $GT3_{90}$ or better. Note, however, that 54 of these events are not calibrated in origin time and therefore have less value for some kinds of studies. For the 847 events with calibrated origin times, 593 events (70%) qualify as $GT5_{90}$ or better and 267 events (32%) qualify as $GT3_{90}$ or better.

We have assembled a catalog of the 901 events from the 24 calibrated clusters in Table 1, which is provided in the Electronic Supplement to the Bergman et al. (2008) paper. For each event, information on the uncertainty of absolute location is given in the form of a 90% confidence ellipse.

Empirical Travel Time Anomalies

When both location and origin time can be calibrated for a cluster, we are able to estimate the unbiased travel times of all observed phases to the respective reporting stations. These estimates are the basis for improved models of the crust and upper mantle, which in the future will permit more accurate routine earthquake locations using regional and teleseismic data.

We use the arrival time data from a calibrated cluster to infer empirical path anomalies (relative to the global model ak135) from each cluster source region to surrounding seismic stations. Empirical path anomalies are referenced to the cluster's hypocentroid. In principle, even a single reading of a phase at a single station could be treated as an empirical estimate of unbiased travel time, but we place the greatest weight on empirical travel times that are estimated from multiple recordings (i.e., from multiple cluster events) of the phase of interest at a given station. We use a robust estimator of spread (Croux and Rousseeuw, 1992) to determine the uncertainty of these estimates. The path anomalies can be the result both of variations in bulk velocity and differences in ray path geometry caused by lateral heterogeneity.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

An example of the empirical path anomalies from one calibrated cluster is shown in Figure 9.

The evidence for azimuthally varying empirical path anomalies is very clear in Figure 9, and even clearer if the anomalies are plotted in map view (Figure 10). For example the group of path anomalies that are small but slightly late near 5° epicentral distance in Figure 9 are at stations southeast of the cluster (Figure 10), whereas path anomalies that are about 4 s late at the same distance in Figure 9 are located northwest and northeast of the cluster in Figure 10.

In Figures 9 and 10 it is notable that the path anomalies based on a small number of readings (gray symbols) generally show good agreement with each other and with path anomalies based on larger numbers of samples.

We have combined the empirical path anomalies of all the clusters that are calibrated in both location and origin time to produce a summary plot of Pn/P arrivals as a function of epicentral distance (Figure 11) for the Iran region.

The patterns of empirical path anomalies for individual clusters show evidence both for departures from the average earth model used for reference (ak135), and for lateral heterogeneity. When the path anomalies are combined however, most structure is lost in a cloud of impressive width, a range of 10-12 s for P phases over the regional and teleseismic distance range. Even here there is a suggestion of about 2-3 s baseline offset from ak135 in the study region. This can be accounted for with a crustal structure that is both thicker (40-45 km Moho depth, vs. 35 km for ak135) and slower in bulk velocities. It is clear, however, that accurate earthquake location in this region will require the use of crustal models that are more specific to the source regions.

Location Errors in Standard Catalogs for the Iran Region

Using our catalog of calibrated earthquake locations, we investigate the location accuracy of several standard earthquake catalogs, both regional and global. We consider only events that qualify as $GT5_{90}$ or better. Since we consider only epicenter accuracy, we use all 624 $GT5_{90}$ events in our catalog of calibrated events, even those for which origin time was not calibrated.

Regional Networks

IIEES

The IIEES catalog has been published since July 2004 and is based on arrival time data only at INSN stations, which have increased in number from 13 in July 2004 to 16 in September 2007. The IIEES catalog contains 188 events that are also contained in our set of $GT5_{90}$ events. Most (160) of the events occurred since the beginning of 2005 when network operations became routine. For each match, we have calculated the distance and azimuth of the IIEES location with respect to the $GT5_{90}$ location. These vectors are shown in Figure 12a.

The median mislocation is only 5.1 km, showing that many IIEES locations are quite accurate. Based on median mislocation, the IIEES catalog would be rated the most accurate of the catalogs analyzed. However the distribution is quite long-tailed and a significant number of IIEES locations have large mislocations (Figure 13a). At the 90% level, IIEES locations are accurate at about 22 km. The statistics of the mislocations for the IIEES catalog are summarized in Table 2.

IRSC

The IRSC catalog is based only on data from the ISTN network. Because the ISTN is organized in 10 sub-networks, the detection and location capability is quite variable, geographically. Moreover, the IRSC did not publish a unified "network" catalog until the beginning of 2006. Prior to then, locations are published for the individual sub-networks, based in most cases only on the data of that subnetwork. Thus there are many cases of multiple subnet locations of the same event. Subnet locations can be quite accurate for events inside the network, but the majority of IRSC locations before 2006 are for the case when the event is outside the network. It is well known that this inevitably leads to large errors in location. For this reason we have split the IRSC catalog into two parts for comparison with our $GT5_{90}$ locations:

- IRSC1: 538 events prior to 2006. All subnetwork solutions for the same event are retained.
- IRSC2: 82 events since 2006.

The mislocation vectors for the IRSC1 and IRSC2 data sets are shown in map view in Figure 12 and histograms of the mislocation distances is shown in Figure 13. The median mislocation for IRSC1 and IRSC2 is 31.3 and 19.6 km, respectively (Table 2), confirming our expectation that pre-2006 locations will be less accurate is strongly confirmed. At the 90% level, the location accuracy of the IRSC1 and IRSC2 is 123 and 36 km, respectively.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Even the recent IRSC2 catalog is significantly less accurate than the IIEES catalog. While some of the difference may be attributed to smaller sample size for IRSC2, which is based mainly on events from two calibrated clusters (Figure 12), we believe that the main factor is the unbalanced distribution of the ISTN stations. Because the stations are grouped tightly into sub-networks, the effective number of stations for location in many parts of Iran is far less than the total number of stations would imply, and the distribution of those sub-networks does not provide adequate azimuthal coverage for many seismic source regions. Recently, the IRSC has begun incorporating some readings from INSN stations in their processing, and this should help improve location accuracy, as should the recent installation of new stations in the east and southeast of the country.

It is ironic that the ITSN, which has been the most important source of phase arrival data for many of our calibration studies, itself provides the least accurate locations of any of the networks considered.

ISC

The ISC catalog contains 501 of the events in common with our $GT5_{90}$ data set. Twelve of the missing events occurred later than the most recent ISC Bulletin available at the time of analysis (June 2006); the remainders are smaller events that depend on data from the Iran regional networks to obtain a stable solution. Mislocation vectors are shown in Figure 12, and a histogram of mislocation distances is shown in Figure 13. The median mislocation is 10.6 km (Table 2). At the 90th percentile, the location accuracy is about 23 km, about the same as the IIEES.

EHB-Iran

The "EHB-Iran" catalog for this comparison is based on our carefully reviewed catalog of the Iran region, not the standard EHB catalog that is widely distributed. The 546 events considered are those that meet a secondary azimuth gap of 180° or better, based on all available arrivals from the ISC, PDE, and Iranian and other regional networks. They have been located with the standard EHB methodology using all arrivals. In contrast, the "standard" EHB catalog is based on arrivals only at teleseismic distances (> 28°). Based on the 90% level of accuracy, the EHB-Iran catalog provides the lowest level of location bias (18 km) of the four networks examined (Table 2), although the median mislocation of 9.1 km is not as good as that of the IIEES catalog. The map and histogram of mislocations are shown in Figures 12d and 13d, respectively.

The fact that the EHB-Iran catalog has fewer events with very large mislocations than any of the other catalogs is primarily due to the selection criterion on secondary azimuth gap.

Azimuthal Bias in Locations

To explore the possibility of a systematic direction of location bias, we made sector plots of the mislocation vectors for each catalog (Figure 14). All the networks considered exhibit a tendency for systematic bias in certain directions, driven by the convolution of station geometry with the distribution of seismicity in the region, and scaled by the accuracy of the travel time model used for location.

The IIEES catalog tends to be biased to the north, and probably represents a tendency for epicenters to be pulled toward the bulk of the stations. The IRSC1 (pre-2006, subnetwork solutions) catalog shows a strong NW-SE trend for mislocations. It may be that this pattern is the sum of two patterns such as is seen for the IIEES catalog, dominated by the frequent seismic activity in the NW-SE trending Zagros region. Some sub-networks of the ISTN see more events to the northwest and some see more events to the southeast. The sector plot for the IRSC2 catalog (post-2006) is dominated by a single calibrated cluster (Dorud), for which IRSC locations were systematically too far south, but this should not be taken to represent the performance of the ISTN network in all parts of Iran.

The two catalogs (ISC and EHB-Iran) that are based largely on teleseismic and far-regional phase readings both exhibit a strong tendency to mislocate events to the southwest. This is especially true of the EHB-Iran catalog, from which the grossest consequences of poor azimuthal coverage and large outlier readings have been removed. What remains is a signal that is the result of a significant and regionally consistent departure of real Earth structure in the Iran region from that assumed in the theoretical travel-time model used for location. All EHB-Iran locations are based on the ak135 model (Kennett et al., 1995); all ISC locations in this comparison are based on the Jeffreys-Bullen tables (ISC has recently begun using ak135).

In broad terms the systematic direction of mislocation can be summarized as a consequence of the fact that there are far more seismic stations in the northern hemisphere, relative to the Iran region, and that the true travel times are larger than the theoretical ones, yielding positive travel-time residuals (Figure 11b). The inversion for location tends to push the event further away from the bulk of seismic stations, i.e., to the south. The EHB-Iran catalog simply ignores the worst

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

cases of poor azimuthal distribution, but it cannot correct the fundamental imbalance of global station distribution.

Network Distance Range and Secondary Azimuth Gap

It is widely appreciated that earthquake location accuracy is strongly dependent on the distribution in distance and azimuth of the recording stations (e.g., Bondár et al, 2004). Single event locations of "ground truth" accuracy can only be accomplished by using only readings from short epicentral distances (typically about 200 km or less); otherwise, unmodeled Earth structure introduces too much bias. An important aspect of the improved location accuracy of the EHB algorithm (Engdahl et al., 1998) is restriction of events to those that satisfy a constraint on secondary azimuth gap. A large open azimuth or a high concentration of stations at a narrow range of azimuths exacerbates the biasing influence of unmodeled Earth structure.

We investigate these issues for earthquake locations in the Iran region by trimming the arrival time data sets of real earthquakes to match several scales of seismic network and relocating them with the EHB methodology without restriction on secondary azimuth gap. The four artificial network scales are:

- 0-2° Dense local network
- 0-5° Local / Near Regional network
- 0-10° Regional network
- 0-15° Far regional network

The mislocations of these relocations are calculated by comparison to the corresponding calibrated locations (only $GT5_{90}$ or better) and are summarized as a function of secondary azimuth gap in Figure 15. Only cases with at least 5 readings are retained.

As expected, the "dense local network" (0-2°) produces generally reliable locations, most of which have location errors of less than 10 km. The cases with large mislocations are generally those with smaller numbers of readings or very large secondary azimuth gap. The "local / near regional network" (0-5°) is also fairly well behaved but shows some very large location errors even for relatively small secondary azimuth gaps because of the biasing influence of unmodeled Earth structure at regional distances. This characteristic is very prominent with the "regional" and "far regional" networks, where the beneficial effect of good azimuthal coverage is outweighed by the biasing effect of unmodeled Earth structure. The role of azimuthal coverage in location accuracy from these four network distance ranges is summarized in Figure 16.

For networks restricted to 10° epicentral distance ranges or less, the median mislocation is generally 10 km or less for secondary azimuth gaps of less than about 200°. The use of arrival time data beyond 10° introduces substantial location bias because they bottom at depths where stronger heterogeneities exist. This can also be seen in the summary of empirical path anomalies (Figures 9 through 11).

Discussion

There is very little additional arrival time data that could be added to improve the comprehensiveness of our seismicity catalog in the time period 1964-2005. Data is, of course, becoming available from the Iranian regional networks and the ISC that would allow the catalog to be brought forward at the level of completeness of the 1998-2005 period. It would be possible to improve the catalog in the pre-ISC period (before 1964) by incorporating the arrival time data of the monthly bulletin of the Bureau Central International Séismologie (BCIS) in Strasbourg, which are not yet available in digital format.

The prospects for adding to the data set of calibrated earthquakes in the Iran region are excellent, given the on-going efforts to study current earthquakes though aftershock studies, geologic mapping, and remote sensing technologies such as InSAR, plus the routine emplacement of temporary monitoring networks for large-scale civil engineering projects and the expansion and improvement of the Iranian national seismic monitoring systems. Whenever a recent event's location can be determined with high accuracy by these systems, we have the ability to use the multiple event relocation procedure to form a cluster with nearby older events and achieve useful levels of calibration for many of them. Nevertheless, the current set of 624 events calibrated at $GT5_{90}$ or better, and even the several hundred other events calibrated at $GT10_{90}$ or better, provide a data set that is a suitable basis for efforts aimed at improving location accuracy in the region.

Our comparisons between the set of best-calibrated events and the locations of other networks and processing schemes is not intended to settle an argument over whose locations are best, but rather to reveal what steps can be taken to improve

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

the location accuracy of each of them. These steps are applicable to virtually any seismic location procedure, anywhere. The most important factor is azimuthal coverage. Like many networks that depend on their own stations to monitor seismicity within the nation's borders, both Iranian networks suffer from poor coverage of seismicity near their borders. This problem is especially severe for the highly active Zagros belt and the Makran region in the southeast. A greater emphasis on data exchange with neighboring countries is the most practical way to improve matters.

For individual event locations, a criterion on azimuth gap should be employed to distinguish between events with adequate coverage from those with a strong potential for significant location bias. For the later class of events, formal statistics on location accuracy have little significance.

For regional networks, more accurate locations can be achieved by restricting the data set to stations at smaller distances, providing there are a sufficient number of stations and they satisfy a reasonable criterion on azimuthal coverage. It is rarely advantageous to include data at far-regional distances (e.g., > 10°) in a location procedure unless exceptional confidence can be placed on the travel time model at those distances.

Conclusions

We have assembled a new catalog of seismicity at magnitudes above 2.5 in the Iran region from global, regional, and local seismic networks which covers the period 1923-2008. Magnitude completeness of the catalog varies strongly through time, reaching a minimum of about 3.6 during the period 1998-2005. Of the 25,722 events in the catalog, most of the larger events have been carefully reviewed for proper phase association, especially for depth phases, and to eliminate outlier readings, and relocated. This catalog provides an improved view of the distribution of seismicity in the region.

901 of these events have been further analyzed using a multiple event relocation technique for calibrated locations that remove most of the systematic bias in single-event locations done with regional and teleseismic data. 624 of these calibrated events have location uncertainties that qualify them as $GT5_{90}$ or better. By comparing these locations with those from the catalogs of global and regional networks (all of which are based on 1-D Earth models), we investigate the location accuracy of those networks in the region. Earthquake locations in the Iran region which are based on regional and teleseismic arrival time data (ISC and EHB) are systematically biased to the southwest and have a 90% location accuracy of 18-23 km, with the lower value achievable by applying limits on secondary azimuth gap. Iranian seismograph networks have operated both as regional networks (IIEES and the IRSC since 2006) and local networks (IRSC prior to 2006). The IIEES catalog achieves the best location accuracy, with a 90% accuracy level of 22 km. The IRSC network, even in recent years, when all the subnetwork data have been combined, suffers from a very imbalanced distribution of stations and achieves a 90% level of accuracy of 36 km. When IRSC sub-networks are used independently to locate earthquakes, as was typically the case prior to 2006, location accuracy at the 90% level exceeds 100 km. The IIEES catalog has a systematic location bias towards the north. The data for the IRSC catalog are inadequate for an analysis of systematic mislocation directions. In broad terms, location accuracy in the Iran region using 1-D travel time models and single event location procedures, is presently limited to about 20 km at the 90% level of accuracy (i.e., $GT20_{90}$).

When analyzed in terms of the distance range and azimuthal coverage of stations, our data confirm the view that the best accuracy (e.g., GT5 or better) is obtained when only stations at short epicentral distances (~200 km or less) are used and when azimuthal coverage is reasonably good (e.g., secondary azimuth < 200°). For a regional network of stations at less than 10° epicentral distance, median location accuracy of 10 km can be achieved if secondary azimuth is less than about 200°. The use of regional arrival time data at distances beyond 10° tends to degrade the quality of solutions in this region.

These results suggest that location accuracy adequate for assessing seismic hazard in the Iran region cannot be achieved without accounting for the strong lateral heterogeneity in seismic travel times through the crust and uppermost mantle. The data set of calibrated locations reported here provides an important constraint on travel-time models that would begin to account for the lateral heterogeneity in Earth structure in the Iran region, and permit seismic networks, especially the regional ones, to obtain more accurate locations of the earthquakes in the region.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

References

Bergman, E. A., E. R. Engdahl, S. C. Myers, A. M. Farahbod, and S. Tabatabai (2008), Earthquake location accuracy in the Iran region, Bull. Seism. Soc. Amer., submitted (including Electronic Supplement with details of clusters studied).

Biggs, J., Bergman, E., Emmerson, B., Funning, G., Jackson, J., Parsons, B., Wright, T. (2006), Fault Identification for Buried Strike-Slip Earthquakes using InSAR: The 1994 and 2004 Al Hoceima, Morocco Earthquakes. Geophy. J. Int., v. 166, p. 1347-1362.

Bolt, B. A. (1960), The Revision of Earthquake Epicentres, Focal Depths and Origin-Times using a High-Speed Computer, J. R. Astron. Soc., v. 3, p. 433-440.

Bondár, I., X. Yang, R.G. North and C. Romney (2001), Location Calibration Data for CTBT Monitoring at the Prototype International Data Center, Pure Appl. Geophys., v. 158(1-2), p. 19-34.

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

Bondár, I., E. Bergman, E.R. Engdahl, B. Kohl, Y-L. Kung, and K. McLaughlin (2009), A hybrid multiple event location technique to obtain ground truth event locations, Geophy. J. Int., v. 175, p. 185-201, doi: 10.1111/j.1365-246X.2008.03867.x.

Croux et al. (1992), Time-efficient algorithms for two highly robust estimators of scale, Computational Statistics, v. 1, p. 411-428.

Engdahl, E.R., and A. Villaseñor (2002), Global seismicity: 1900–1999, International Handbook of Earthquake and Engineering Seismology, v. 81A, p. 665-690.

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. Seism. Soc. Amer., v. 88, pp. 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., v. 167, p. 761-778.

Geiger, L. (1910), Herdbestimmung bei Erdbeben aus den Ankunftszeiten. Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse, p. 331-349.

Geiger, L. (1912), Probability method for the determination of earthquake epicenters from the travel time only. Bull. St. Louis Univ., v. 8, p. 60-71.

Gunst, R. H. and E. R. Engdahl (1962), Progress report of USC&GS hypocenter computer program, Earthquake Notes, v. 33, p. 93-96.

Jordan, T.H. and K.A. Sverdrup (1981), Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, Bull. Seis. Soc. Amer., v. 71, p. 1105-1130.

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

Parsons, B., T. Wright, P. Rowe, J. Andrews, J. Jackson, R. Walker, M. Khatib, M. Talebian, E. Bergman and E. R. Engdahl, 2006, The 1994 Sefidabeh (eastern Iran) earthquakes revisited: new evidence from satellite radar interferometry and carbonate dating about the growth of an active fold above a blind thrust fault, Geophys. J. Int., v. 164, p. 202–217.

Ritzwoller, M.H., N.M. Shapiro, A.L. Levshin, E.A. Bergman, and E.R. Engdahl (2003), Ability of a global three-dimensional model to locate regional events, J. Geophys. Res., v. 108(B7), p. 2353, doi:10.1029/2002JB002167.

Tatar, M., J. Jackson, D. Hatzfeld and E.A. Bergman (2007), The 28 May 2004 Baladeh earthquake (Mw 6.2) in the Alborz, Iran: implications for the geology of the South Caspian Basin margin and for the seismic hazard of Tehran, Geophys J. Int., v. 170, p. 249-261.

Tavakoli et al. (1999), Seismic hazard assessment of Iran, earth-prints.org.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Villasenor, A., and E.R. Engdahl (2007), Systematic relocation of early instrumental seismicity: Earthquakes in the International Seismological Summary for 1960-1963, Bull. Seis. Soc. Amer., v. 97(6), p. 1820-1832.

Walker, R. T., E. Bergman, J. Jackson, M. Ghorashi, and M. Talebian. (2005), The 2002 June 22 Changureh (Avaj) earthquake in Qazvin province, northwest Iran: epicentral relocation, source parameters, surface deformation and geomorphology, Geophys. J. Int., v. 160, p. 707-720.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

Name	Events	GT5	GT3	Direct	Indirect	Cal. Level	Seismic	InSAR	OT
Alborz	11	11	11	•		1.6			•
Ardebil	16	0	0	•		7.4			•
Avaj	17	16	1		•	4.3	1		•
Baladeh	25	22	20	•		1.9			•
Bam	22	12	0		•	3.4		1	
Bastak	25	17	0		•	3.5		1	
Darband	44	41	17	•		2.9			•
Dasht-e-Bayaz	101	44	0		•	4.7	3	3	•
Dorud	80	77	63	•		1.8			•
Firuzabad	36	1	0		•	13.3	1		•
Garmsar	16	16	15	•		1.3			•
Ghir	67	51	23		•	2.8	2	1	•
Igdir	71	47	1	•		3.1			•
Kahak	15	15	6	•		2.7			•
Lagodekhi	44	36	19	•		2.4			•
Nahavand	25	0	0	•		8.4			•
Qeshm	43	38	18		•	2.6	2	1	•
Rudbar	81	78	58		•	1.8	3		•
Sefidabeh	7	7	6		•	3.1	1	1	•
Shusf	7	4	1		•	4.1		1	
Tabas	30	14	2		•	4.4	2		•
Urmieh	26	25	0	•		3.5			•
Zarand	49	44	7		•	3.1	1	1	•
Zirkuh	43	10	0		•	5.1	2		•
Total	901	626	268	11	13		10	8	21

Table 1. Earthquake clusters for which a calibration analysis was made. “GT5” and “GT3” are the number of events that qualify as GT5 or better and GT3 or better. “Direct” and “Indirect” indicate the type of calibration. “Cal. Level” is the level of uncertainty of the calibration process. “Seismic” is the number of seismic calibration events used for calibration. “InSAR” is the number of InSAR analyses or instances of mapped faulting used for calibration. “OT” indicates that the origin time is calibrated.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

	#	Median	Mean	90%	Max
IIEES	188	5.1	10.1	22	87
IRSC1	538	31.3	49.3	123	275
IRSC2	82	19.6	21.0	36	67
ISC	501	10.6	12.6	23	170
EHB-Iran	546	9.1	10.3	18	38

Table 2. Statistics of mislocation vectors of several seismic networks compared with the set of corresponding GT5 locations.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

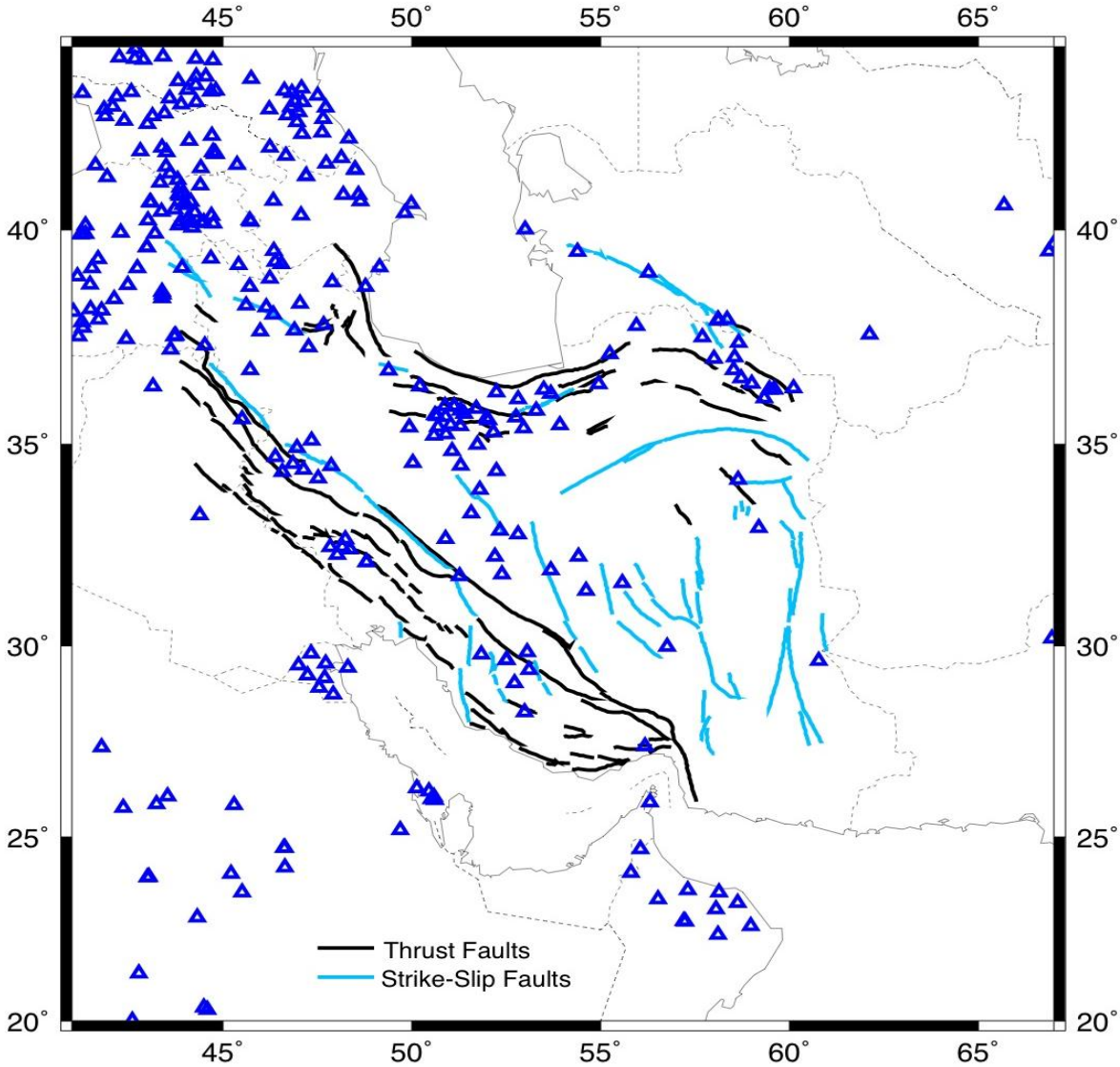


Figure 1. Stations in the Iran region reporting arrival time data during the period 1923-2008.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

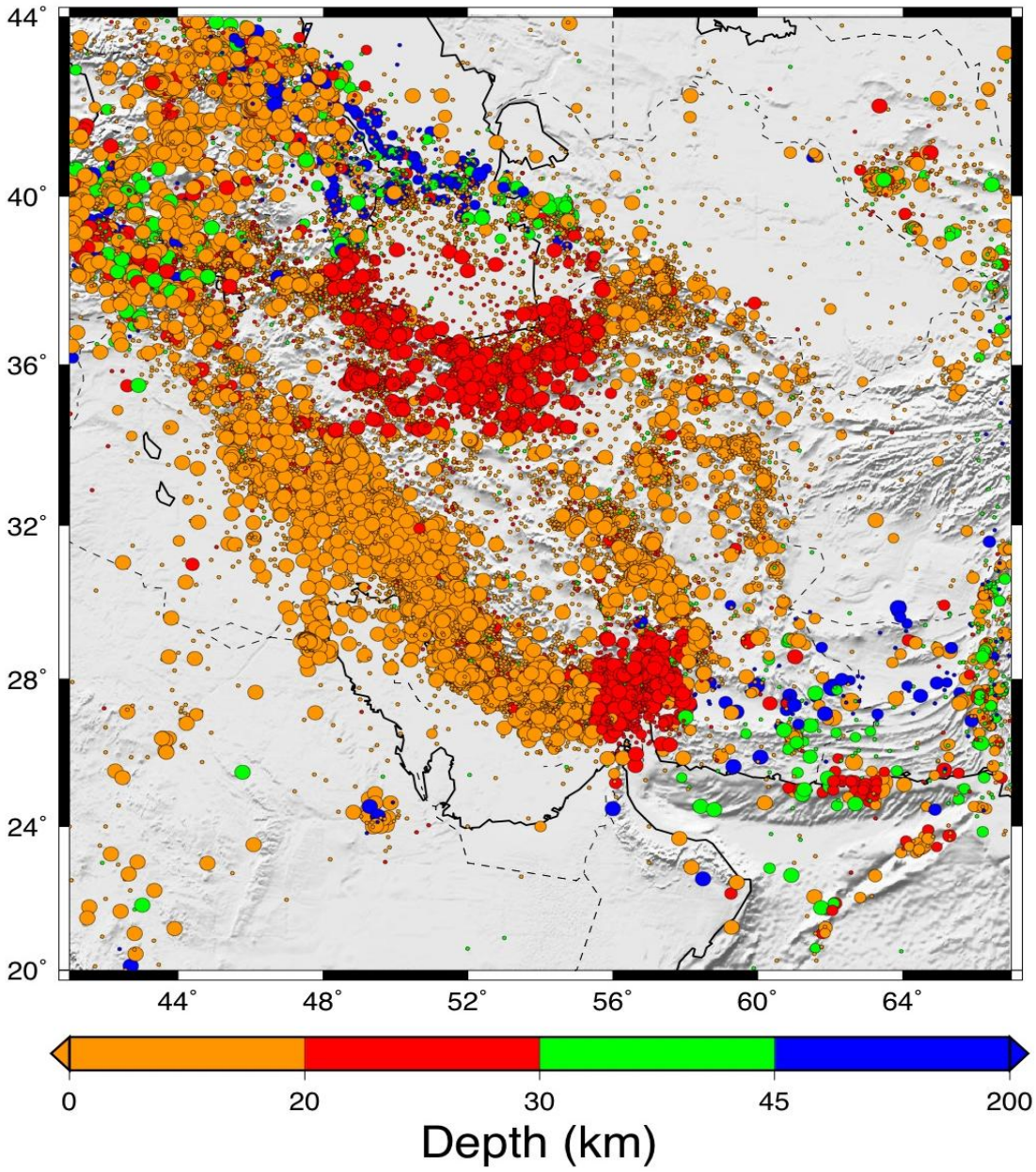


Figure 2. Seismicity map of the study region color coded by depth. Unreviewed depths are set to default regional estimates as defined by Engdahl et al. (2006).

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

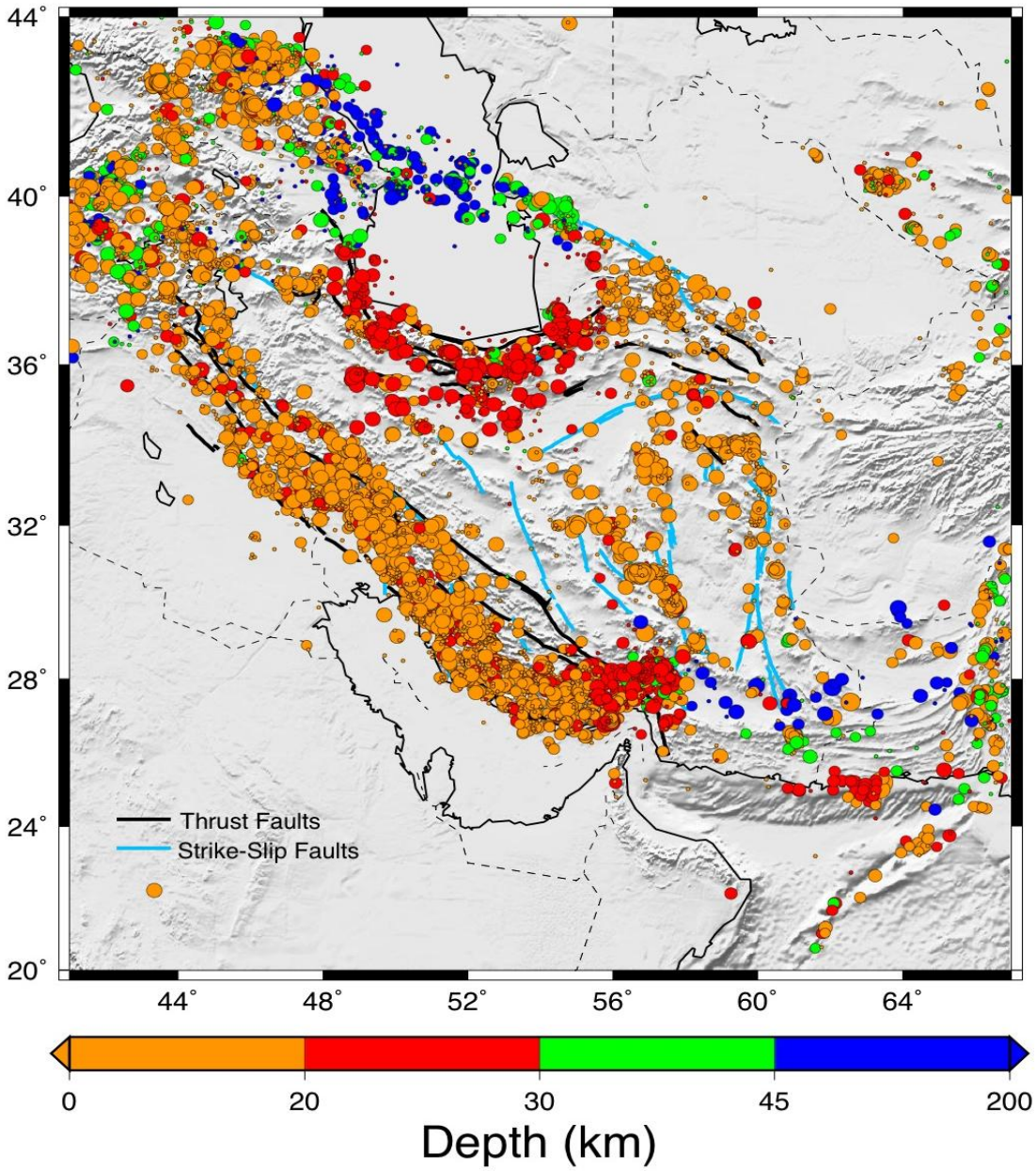


Figure 3. Events meeting a secondary azimuth gap criteria of less than 180° when applied to stations at all distances.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

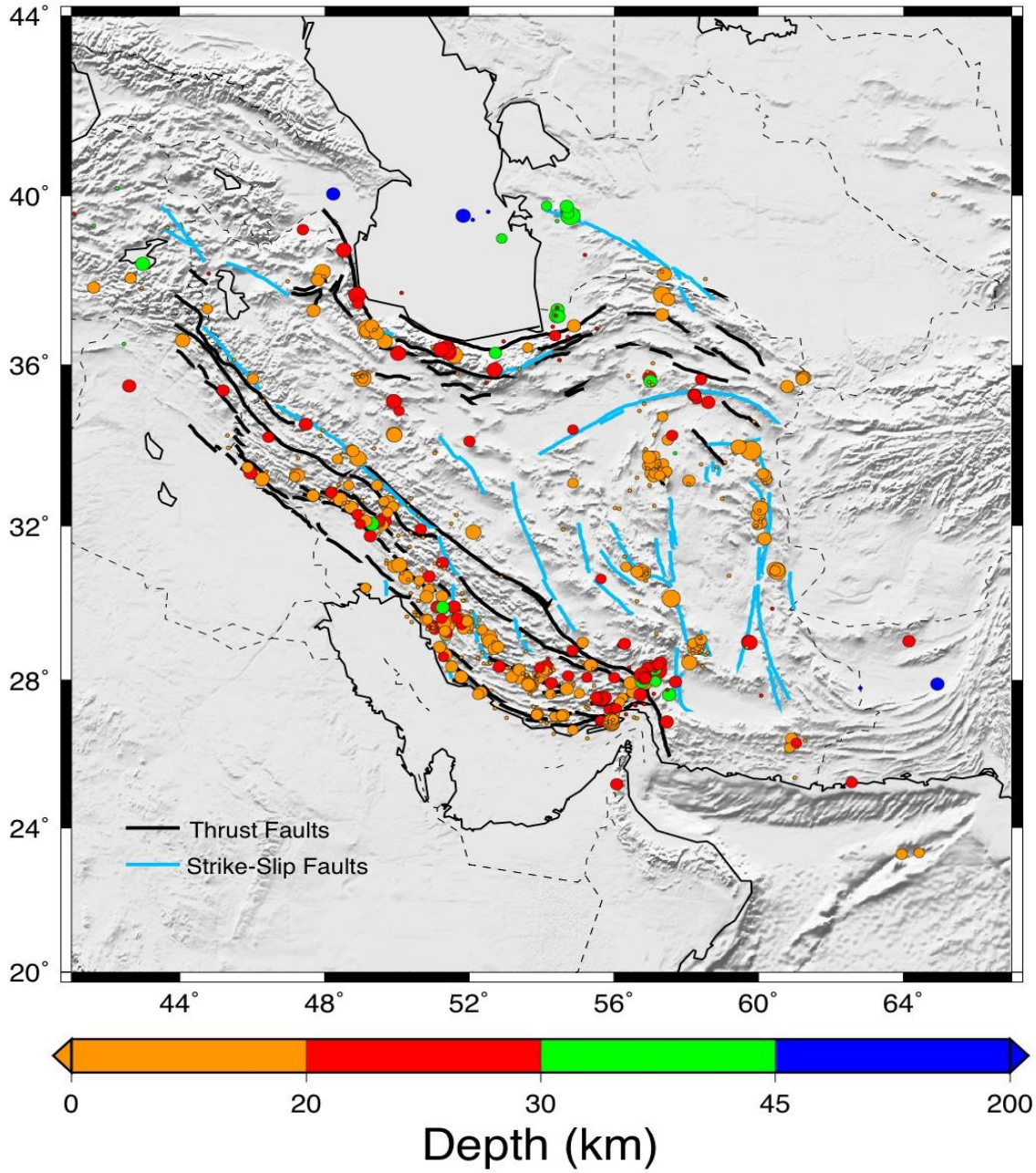


Figure 4. Events during the period 1976-2006 for which at least some arrival times were re-picked by an experienced analyst for this study.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

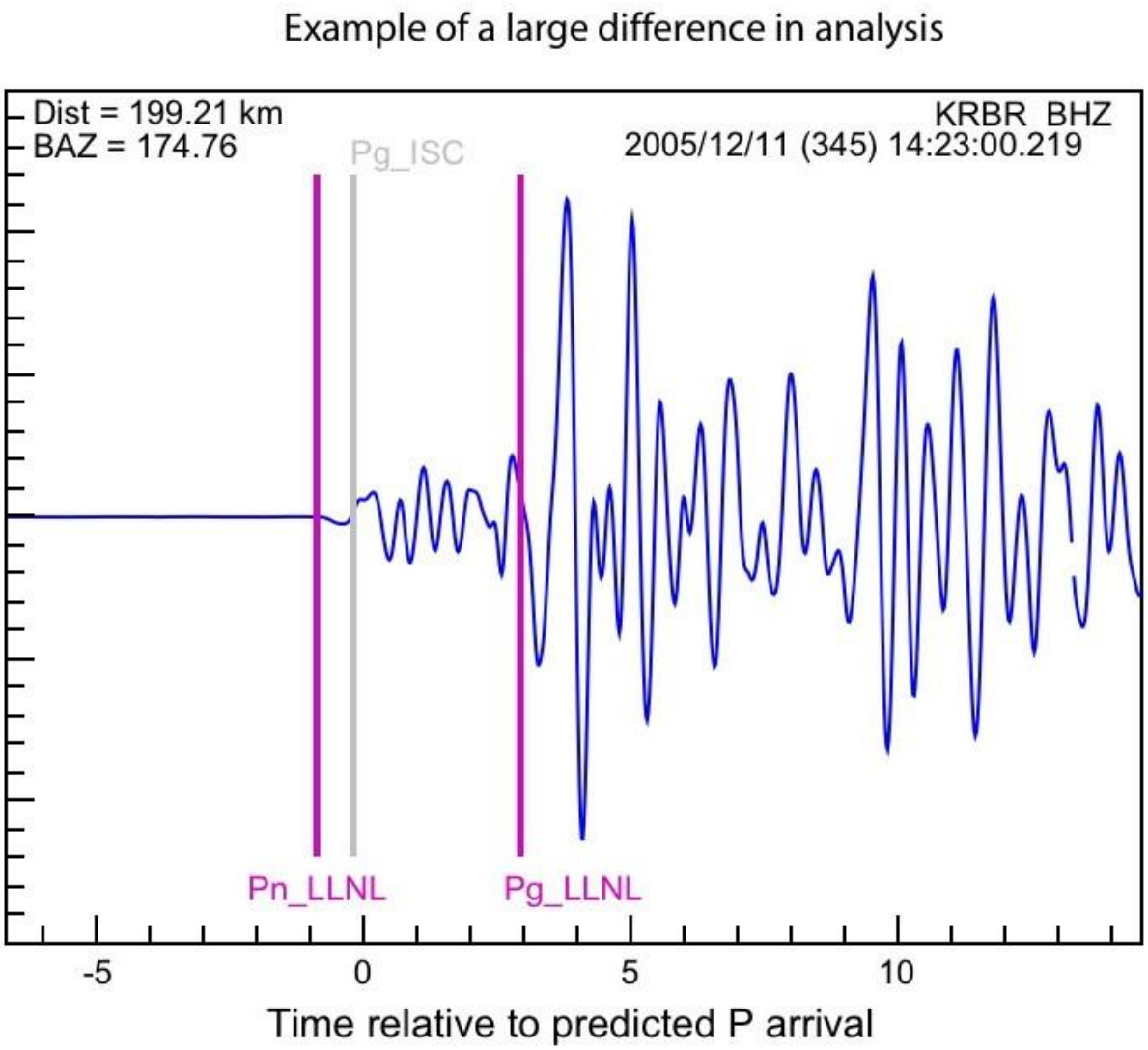


Figure 5. Regional waveforms read by an experienced LLNL analyst. In this case there is a large discrepancy between the reviewed pick and the original pick reported to the ISC.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

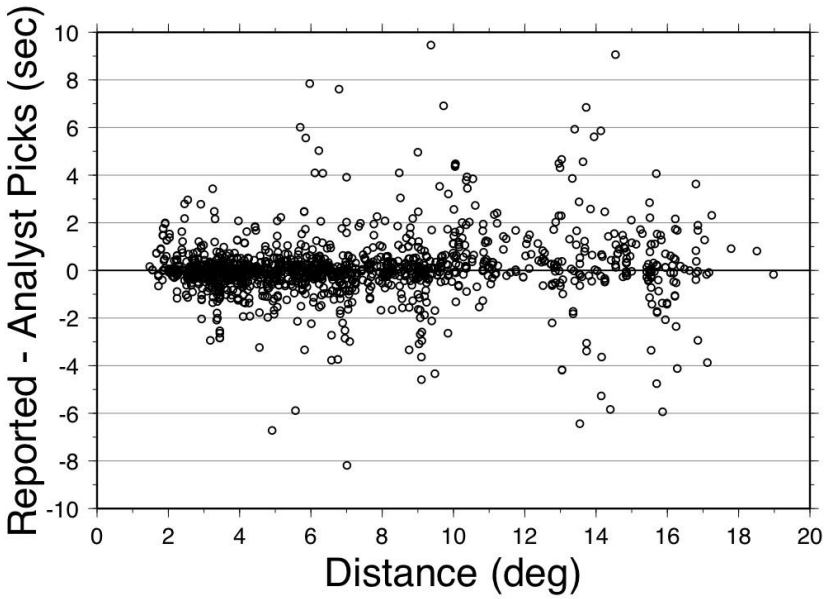


Figure 6. Difference between our analyst's picks and arrival times reported by regional networks.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

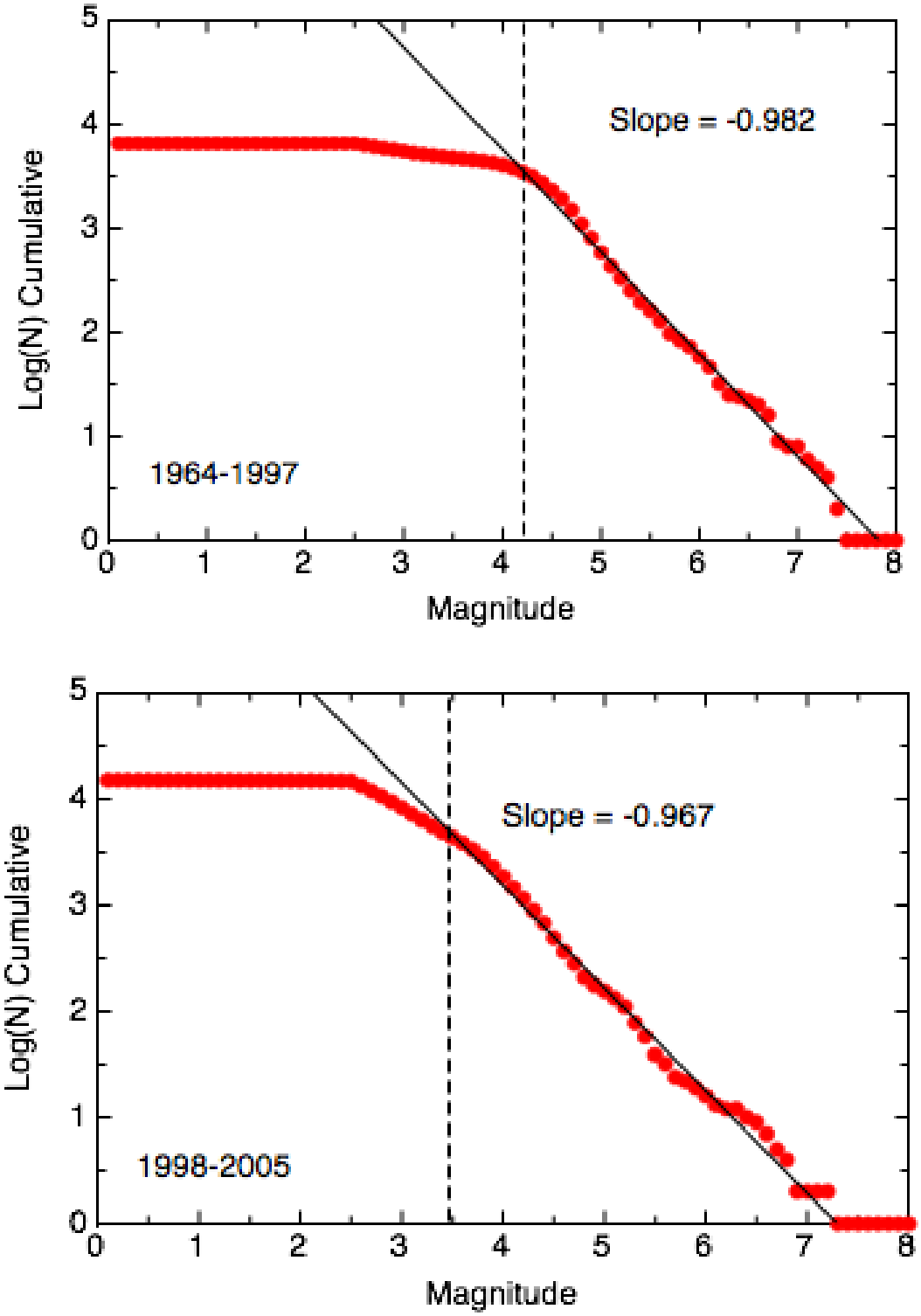


Figure 7. Completeness of the new Iran catalog through two time periods, shown in the lower left corner of each plot. Each plot is based on the cumulative number of events at each magnitude, i.e., the number of events in the catalog with a magnitude of that value or larger. A straight line is fitted to the data over a magnitude range that is appropriate to reveal the corner, which represents the magnitude completeness level. In each plot a vertical dashed line shows the approximate location of this corner magnitude.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

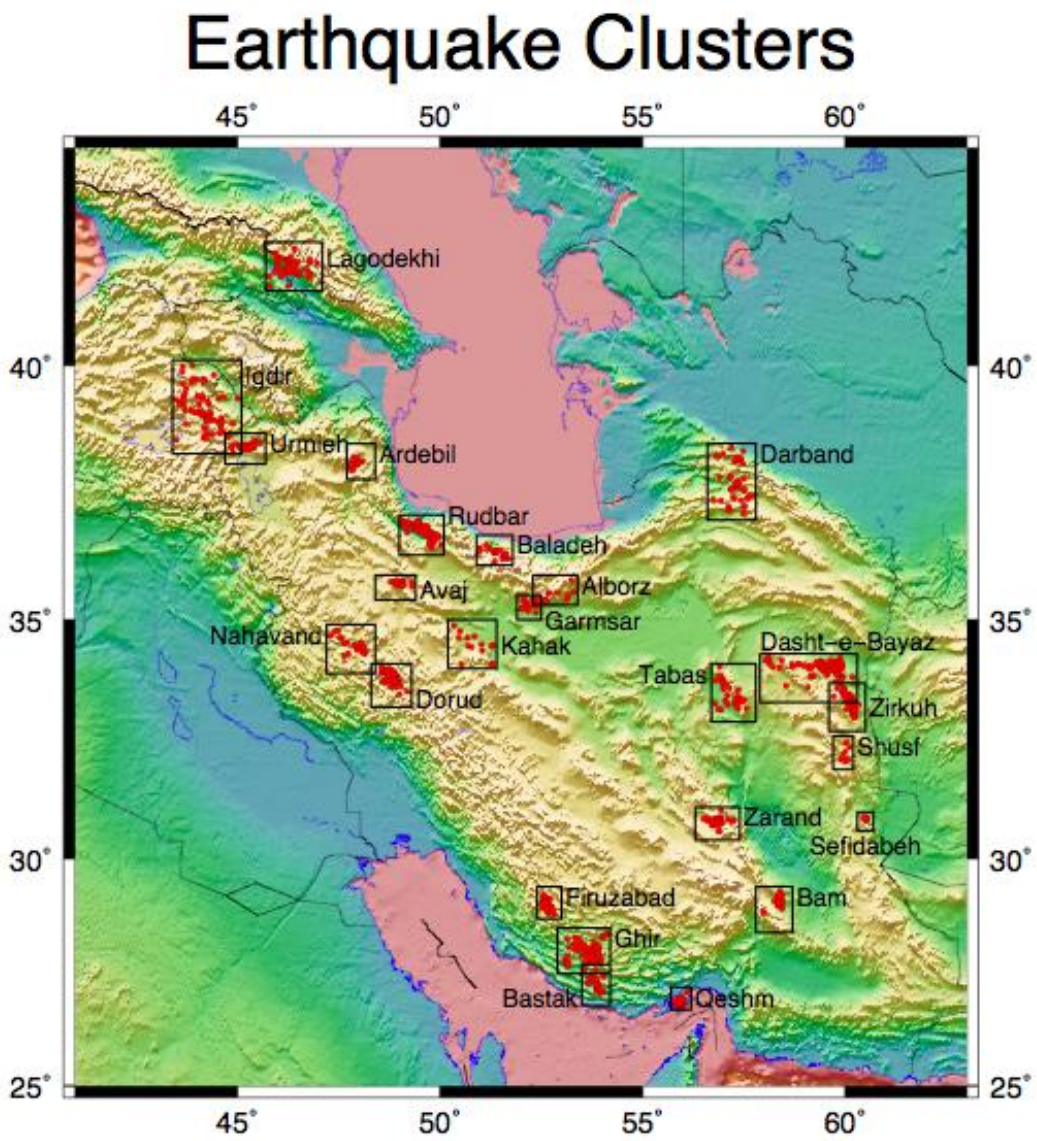


Figure 8. Locations and names of earthquake clusters in this study, with epicenters and bounding boxes of the detailed figures included in the electronic supplement.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

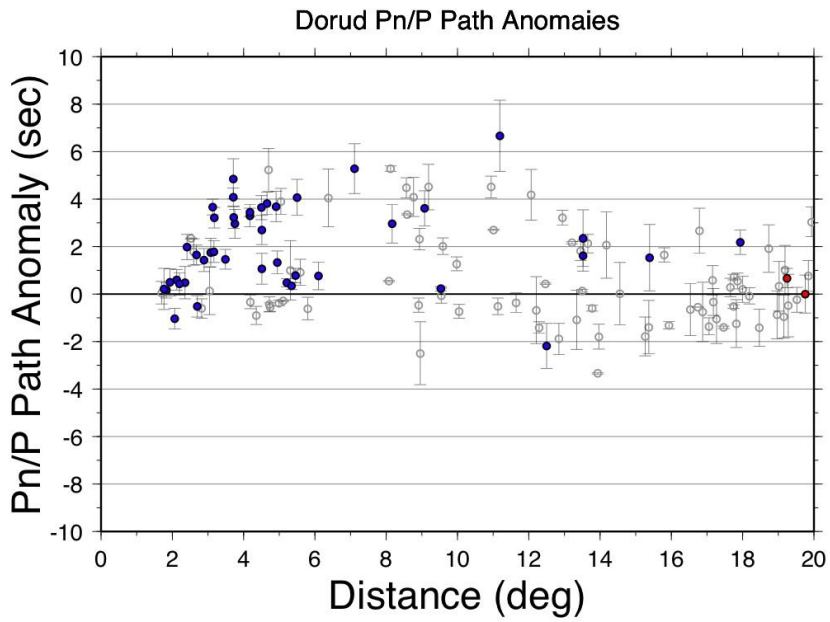


Figure 9. Empirical path anomalies for Pn and P phases at regional distances for the Dorud cluster. The uncertainty of each estimate is shown. Solid symbols represent estimates based on 5 or more readings. Open symbols are based on fewer than 5 readings. All estimates are based on at least 2 readings. Blue symbols have been identified as Pn; red symbols are identified as P.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

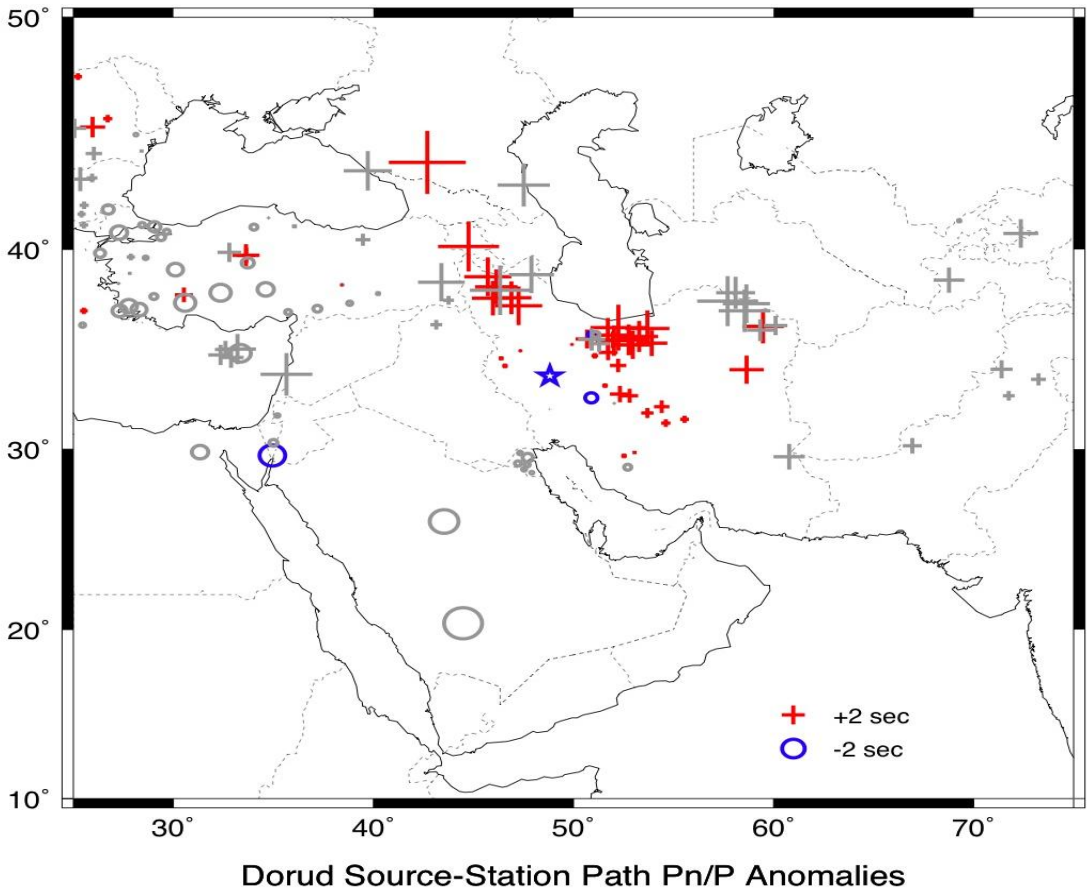
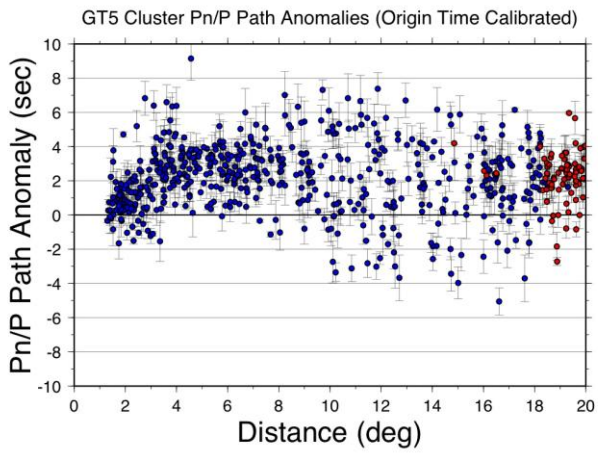


Figure 10. Map view of the empirical path anomalies for the Dorud cluster shown in Figure 9. The star indicates the hypocentroid of the cluster. Crosses are late arrivals relative to ak135, circles are early arrivals. Size of symbols is scaled to the amplitude of the anomaly. Symbols in color (red for late, blue for early) are estimates based on 5 or more readings. Symbols in gray are based on 2-4 readings.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.



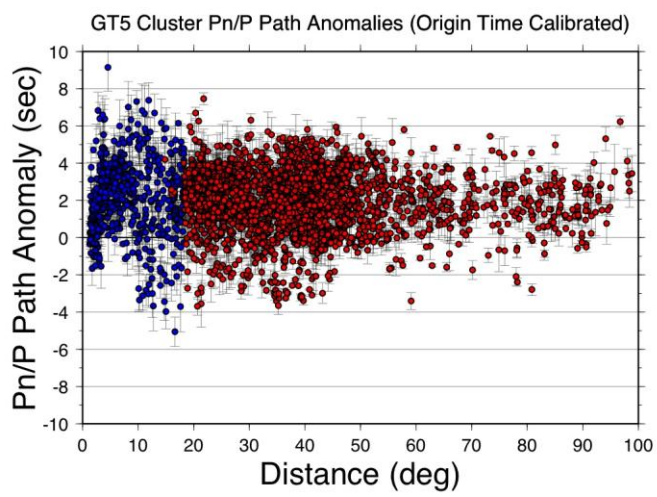


Figure 11. Empirical path anomalies, relative to ak135, and spread for Pn (blue) and P (red) phases from 21 earthquake clusters that are calibrated both in location and origin time. A) Regional distances. B) Regional and teleseismic distances.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

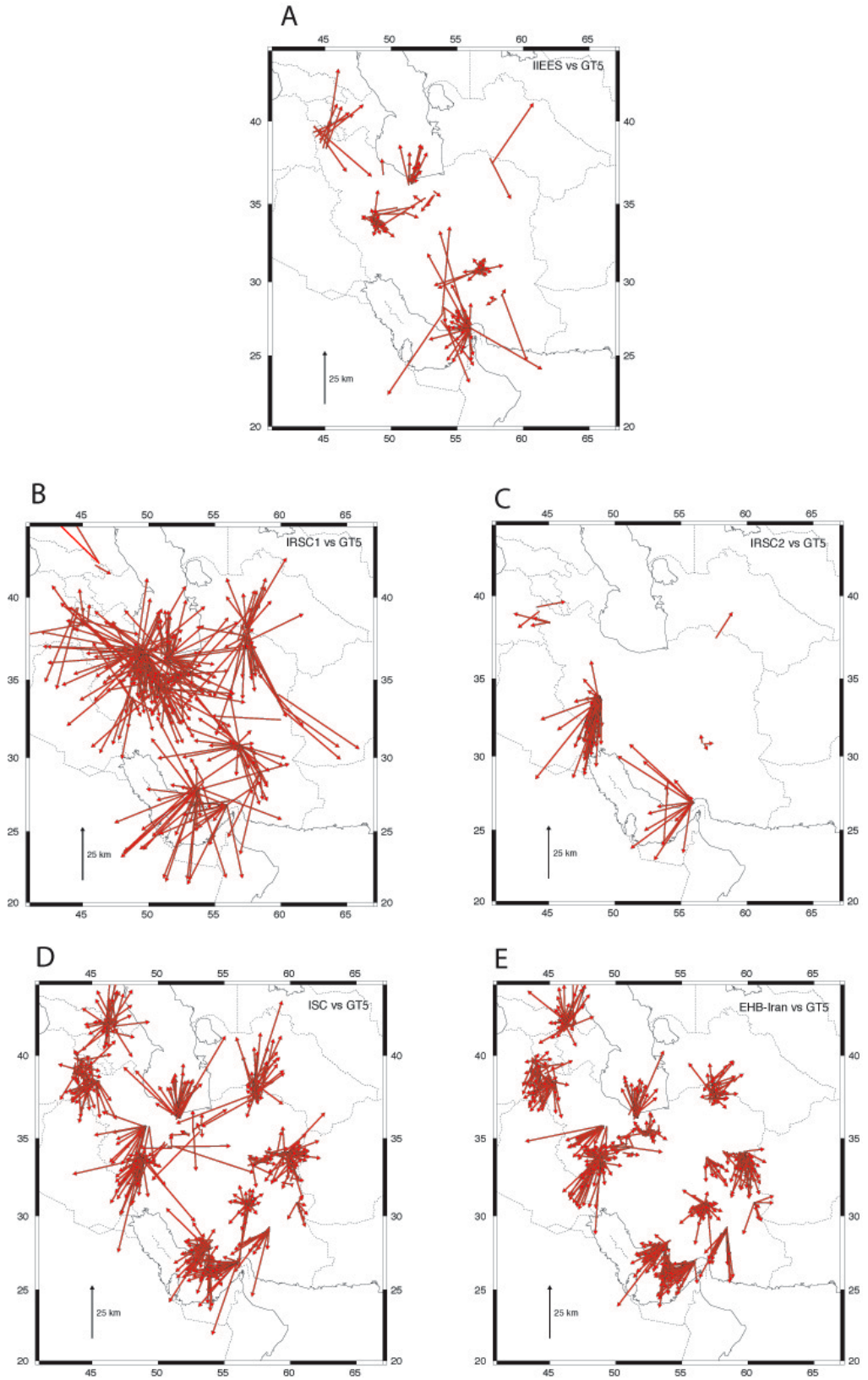


Figure 12. Mislocation vectors for locations for different catalogs, in comparison with corresponding GT5 locations. Only vectors with length less than 50 km are plotted.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

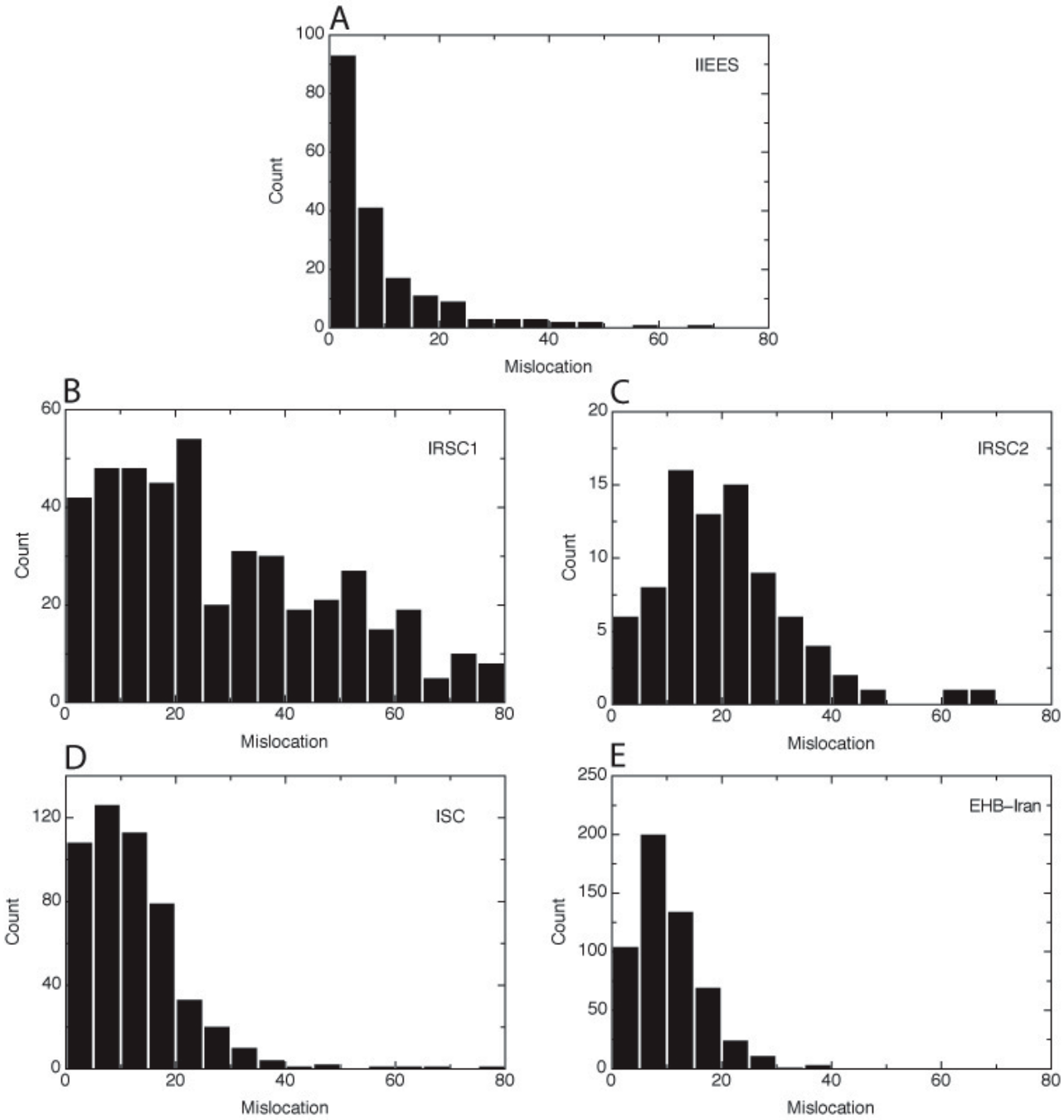


Figure 13. Histograms of mislocation distances of the different catalogs relative to the GT5 data set.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

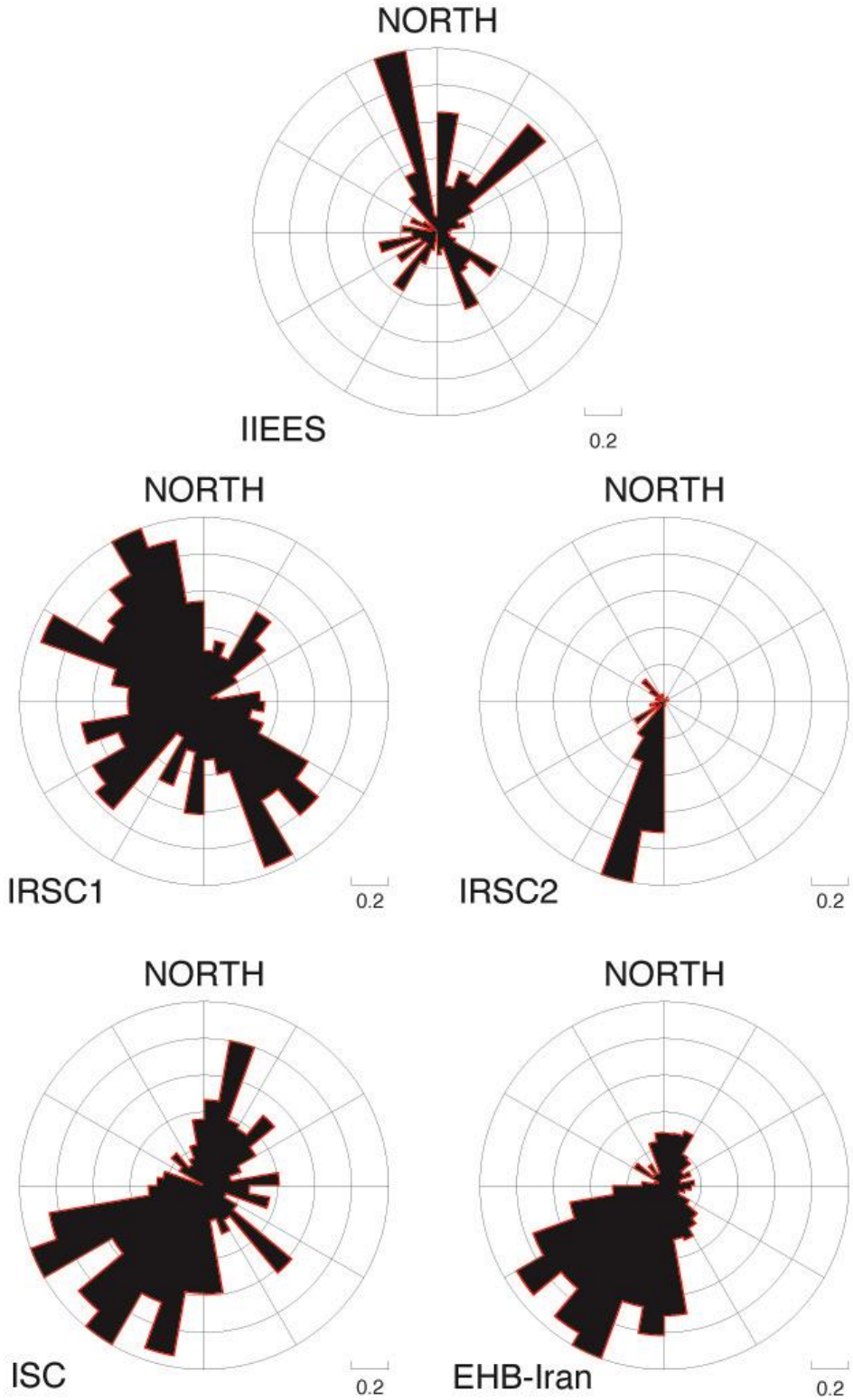


Figure 14. Sector diagrams of the azimuth of mislocation for the different catalogs. Sectors are 30° in azimuth. The number of mislocation vectors in each sector is normalized to the most populous sector.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

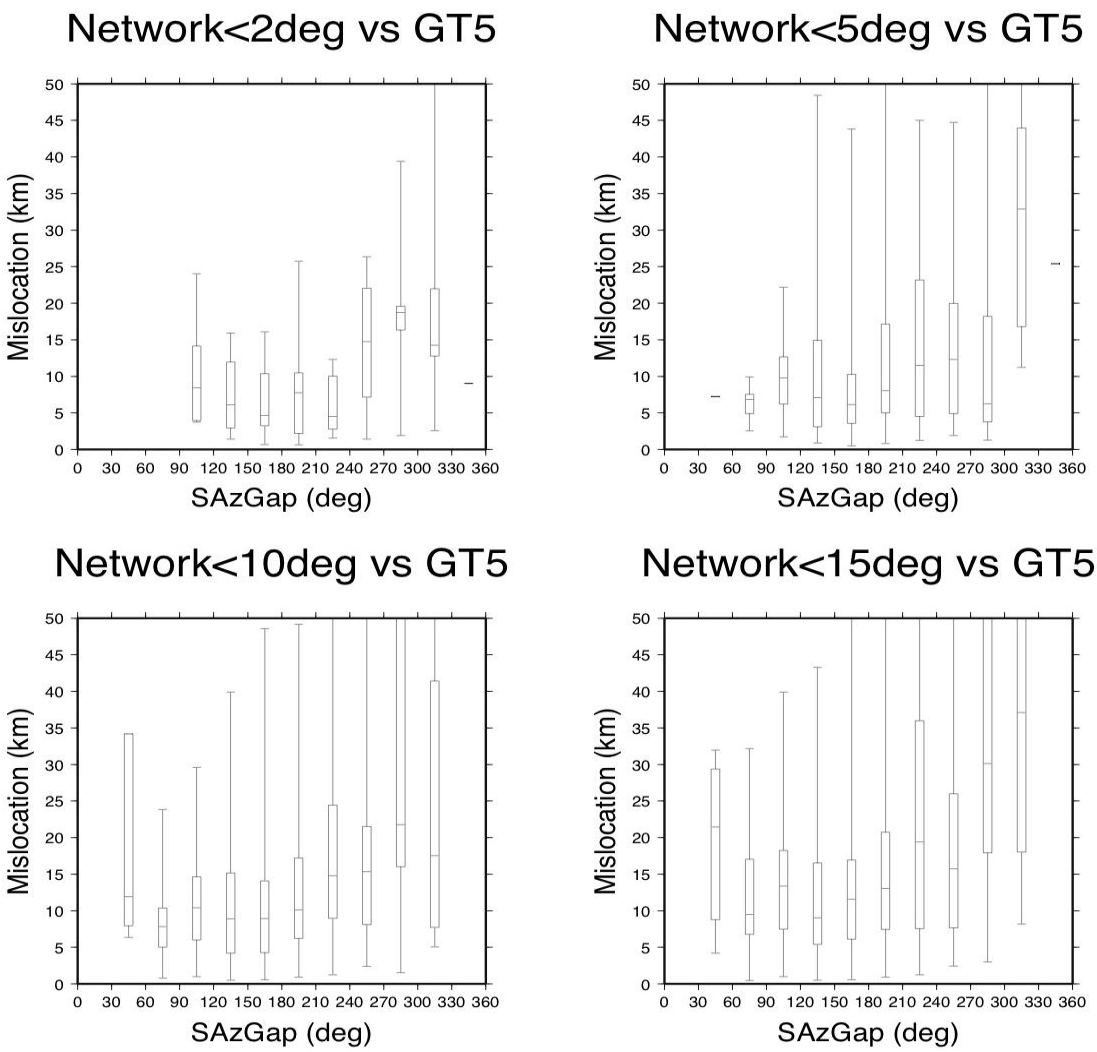


Figure 15. Box plots of mislocation as a function of network distance range and secondary azimuth gap. Arrival time data sets are trimmed to satisfy four network distance ranges (less than 2, 5, 10, and 15 degrees), with a minimum of 5 readings, located with the EHB algorithm, and compared to calibrated locations that are GT5 or better. Each box is derived from the upper and lower quartile, and the location of the median is indicated. The outer limits of each box plot are maximum and minimum observation.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

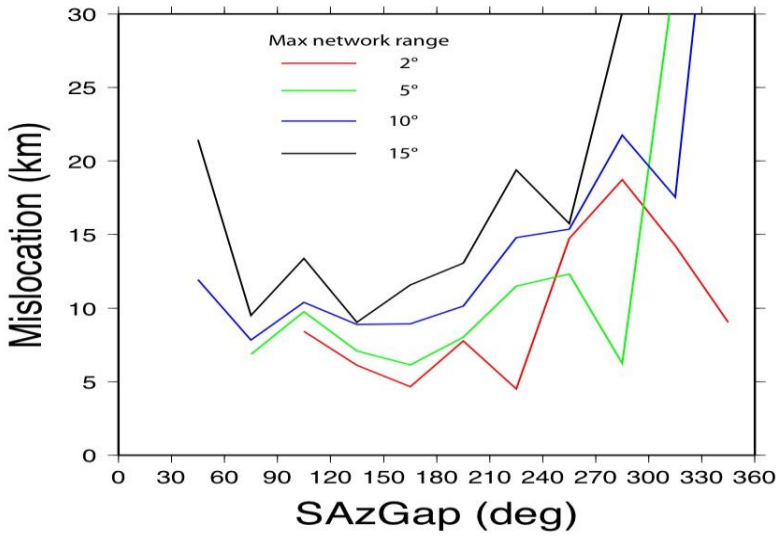


Figure 16. Summary of the medians of event mislocation as a function of network distance range and secondary azimuth gap in Figure 15.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.

PRODUCTS

Journal Publications

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.

Bergman, E. A., E.R. Engdahl, S. C. Myers, A. M. Farahbod, and S. Tabatabai (2008), Earthquake location accuracy in the Iran region, Bull. Seism. Soc. Amer., submitted.

Conference Papers

E.R. Engdahl, E. A. Bergman, S. C. Myers, and F. Ryall, "Improved Ground Truth in Southern Asia Using In-Country Data, Analyst Waveform Review and Advanced Algorithms", presented at 26th Seismic Research Review held at Orlando, Florida, during September 21-23, 2004.

E.R. Engdahl, E. A. Bergman, S. C. Myers, and F. Ryall, "Improved Ground Truth in Southern Asia Using In-Country Data, Analyst Waveform Review and Advanced Algorithms", presented at 27th Seismic Research Review held at Rancho Mirage, California, during September 20-22, 2005.

E.R. Engdahl, E. A. Bergman, S. C. Myers, and F. Ryall, "Improved Ground Truth in Southern Asia Using In-Country Data, Analyst Waveform Review and Advanced Algorithms", presented at 28th Seismic Research Review held at Orlando, Florida, during September 19-21, 2006.

E.R. Engdahl, E. A. Bergman, S. C. Myers, and F. Ryall, "Improved Ground Truth in Southern Asia Using In-Country Data, Analyst Waveform Review and Advanced Algorithms", presented at 29th Monitoring Research Review held at Denver, Colorado, during September 25-27, 2007.

E.R. Engdahl, E. A. Bergman, S. C. Myers, and F. Ryall, "Improved Ground Truth in Southern Asia Using In-Country Data, Analyst Waveform Review and Advanced Algorithms", presented at 30th Monitoring Research Review held at Portsmouth, Virginia, during September 23-25, 2008.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Energy.