Validation of Regional and Teleseismic Travel-Time Models by Relocating GT Events

Xiaoping Yang[1], István Bondár[1], Joydeep Bhattacharyya[1], Michael Ritzwoller[2], Nikolai Shapiro[2], Michael Antolik[3], Göran Ekström[3], Hans Israelsson[1], and Keith McLaughlin[1]

[1] Science Applications International Corporation

[2] University of Colorado at Boulder

[3] Harvard University

Abstract

Recently three-dimensional global seismic velocity models of crust and mantle have been developed and location improvements have been demonstrated with regional (CUB1 and CUB2 models) and teleseismic (J362 model) calibrations, respectively (Shapiro and Ritzwoller, 2002; Ritzwoller et al., 2002; Antolik et al., 2003). In this study, we validated event location improvements from these regional and teleseismic models, separated and combined, for Europe, the Mediterranean, North Africa, the Middle East, and Western Eurasia using a large set of high quality GT0-GT10 events (Bondár et al., 2003). To ensure effective comparisons between the models, we only used events that could be reasonably located, selected by secondary azimuthal gap less than 160º by Pn phase (epicentral distance less than 15º) and/or teleseismic P phase (epicentral distance between 25º and 97º). Besides relocating events using all station arrivals, a subset of the GT events was also relocated using controlled station geometries generated from "constrained bootstrapping". The advantages of this approach include: (1) simulating sparse network (named Simulated Sparse Network Bulletin or SSNB), (2) increasing the statistical power of the testing, (3) reducing the effect of correlated errors to ensure valid 90% error ellipse coverage statistics, and (4) measuring location bias due to un-modeled 3D Earth structures. We compared the location results, with and without calibrations, on

mislocation, error ellipse area, 90% coverage, origin time bias, origin time errors, and misfit. Relocation results of more than 1000 GT0-GT10 events show that combined regional and teleseismic calibration provided the best location improvements, which achieved the GT5 accuracy level. The event relocation using all stations has deficient 90% error ellipse coverage due to correlated errors. In contrast, the coverage is about 90% for SSNB locations with 10 stations. In event location using a large number of observations, location algorithms are required to account for non-linearity, and non-Gaussian, non-zero mean, and non-independent errors.

Introduction

In recent years significant effort and progress have been made in seismic event location calibration, particularly for sparse seismic networks such as the International Monitoring System (IMS). The goal of calibration is to improve location accuracy and reduce the uncertainty while retaining 90% coverage, i.e. the true locations are inside their error ellipses for 90% of events. To accomplish this, more accurate travel times and error estimates are needed, and ground truth (GT) data are vital for validating the models and location improvements.

The IMS event location process involves a hierarchy of corrections for travel times (Yang et al., 2001a). The corrections include ellipticity, elevation, station bulk, and Source-Specific Station Corrections (SSSCs). SSSCs are travel time corrections relative to the baseline global 1D IASPEI91 model (Kennett and Engdahl, 1991). For a given station, SSSCs are calculated from the station out to regional or teleseismic distance, specified on rectangular latitude and longitude grids. Associated modeling errors are also specified in the same manner for each station to reflect the model uncertainty. Another error, measurement error, is given for each arrival and both errors are used in weighting the arrivals in the location inversion. As the last step of the location procedure, error ellipses are estimated from the *a priori* total errors (including modeling and measurement errors), assuming independent Gaussian statistics and linearity.

Both model-based and empirical approaches have been used to improve regional travel time predictions. One-dimensional models were initially employed in developing the

SSSCs for IMS stations in Fennoscandia and North America (Yang et al., 2001). Tomography inversion was applied to refine resolution of the 3D model in Eastern Asia (Murphy et al., 2002). GT data were used in developing empirical SSSCs for Caucasus mountain using kriging (Myers and Shultz, 2000a). In regions where both velocity models and GT data are available, a hybrid approach of combining both model predictions and travel time observations has been used in developing SSSCs for Eastern Asia (Armbruster et al., 2002). Three-dimensional seismic velocity models (CUB1 and CUB2 models) and raytracing have been developed and demonstrated significant location improvement (Shapiro and Ritzwoller, 2002; Ritzwoller et al., 2002a).

Previous efforts have been mostly focused on developing regional SSSCs in various regions, since regional travel times deviate from the IASPEI91 more severely than teleseismic travel times. Regional calibration has reduced regional travel time model errors close to the level of the teleseismic model errors. Teleseismic phases comprise an important fraction of RMS residual data and therefore strongly influence event location, and there are still significant biases in IASPEI91 teleseismic travel times compared to the real 3D earth structure. Because of the slowness differences between teleseismic and regional phases, mislocation due to one-second error in a teleseismic travel time is more severe than that caused by the same error in a regional travel time. Global 3D velocity models have been developed (J362 model and its predecessors, e.g. SP12 model) and shown large impact in improving event location (Antolik et al., 2001; 2003).

Model validation requires robust statistical analyses and location procedures, along with a large amount of quality GT origins and arrival data. Various techniques have been used to assess model goodness, including comparisons between predicted vs. empirical path corrections (e.g. Ritzwoller et al., 2002a),
(e.g. Johnson and Vincent, 2002), and evaluations of location bias using randomly selected subsets of arrivals (e.g. Antolik et al., 2001). Effective testing depends on the accuracy, the amount, and the geographic coverage of GT data. Location results may fluctuate with outliers, particularly at critical azimuth (e.g. Johnson and Vincent, 2002). Outlier analysis is useful in providing more robust results (e.g. Ritzwoller et al., 2002a). Requirements on minimum azimuthal gap and the number of observations (referred to as the number of defining phases, ndef) help eliminate poor events that are unreliable in

evaluating location improvements (e.g. Ritzwoller et al., 2002a). However, the number and geographic coverage of high quality GT events have been limited. In particular, only data independent from the model development should be used in validation testing. This requirement is particularly difficult to fulfill for the empirical approach of model development, especially kriging and body wave tomography, in which the best travel time data are often already included. Typically a leave-one-
validating kriged SSSCs. Since many events are clustered, more objective testing would require leave-one-path-out inst
contrast, the model-based approach, particularly the surface wave models such as CUB, has the advantage of exploiting additional data.

A large set of high quality GT events has been collected recently by the Group-2 Location Calibration Consortium (Bondár et al., 2003). The Group-2 database includes GT0-GT10 events obtained from cluster analyses, Group-2 GT selection criteria, or origins with known or well-estimated location accuracy. Outliner analyses were performed for quality control of the data. In this database both GT arrival/origin
inform
information and empirical path corrections are available for use in model validation testing. Comparisons between predicted and empirical path corrections show significant agreement between the two (Ritzwoller et al., 2002a; Bhatacharryya et al., 2003). Model-based path corrections represent variance reductions between 20-50% w.r.t. the empirical path corrections. This paper describes event relocation testing that utilizes the new data set to test and validate the regional and teleseismic velocity models.

Many GT events used in the previous relocation studies (e.g. Ritzwoller, et al., 2002a; Antolik et al., 2001) and in the Group-2 Consortium database have large numbers of observations. In such cases the assumption of uncorrelated errors may no longer be valid. In addition, because of the good azimuthal coverage, these events are generally well located, mostly within 25 km uncertainties. Regardless of whether calibrated travel-
are applied or not, the locations do not change significantly. Furthermore, arrivals tend to be clustered on the focal sphere and it is now becoming generally understood that these clustered observations are not statistically independent. Therefore, tests that simply relocate these events with all arrivals have poor statistical power to demonstrate calibration improvement/deteriorations.

ead, which will significantly reduce the test data set. In

event-out strategy is used in

Calibrated travel-times have the largest impact on sparsely recorded events. Monte Carlo techniques have previously been used to examine the statistics of sparsely recorded events using large events with many arrivals (e.g. Antolik et al., 2001). To increase the statistical power of the experiment we generate Simulated Sparse Network Bulletin (SSNB) from well-recorded events by taking subsets of stations that satisfy constraints on the number of stations and the secondary azimuthal gap. The subsets of stations meeting these constraints thus simulate sparse networks (e.g. IMS), and yet provide reasonable azimuthal coverage to avoid extremely poor locations due to degenerate network geometries. It has recently been argued that the size of the so-called "secondary azimuthal gap" (sgap), the largest azimuthal gap filled by a single station, is a good indicator of the sensitivity of the location to individual arrival time outliers. Therefore, since our goal is to test the travel-time calibrations, it is desirable to design our tests so that they are less sensitive to measurement errors in the test data set.

In this study we demonstrate location improvements in the Group 2 Consortium region of interest, including Europe, the Mediterranean, North Africa, the Middle East, and Western Eurasia. We validate regional and teleseismic calibrations (both independent and combined) that were developed from 3D global velocity models, CUB1, CUB2, and J362. Testing is conducted by relocating a large data set of GT0-GT10 events with regional and teleseismic SSSCs. Only GT events with secondary azimuthal 160º are used in relocation to ensure stable location results for calibration evaluation. We relocated GT events using all available quality arrivals as well as SSNB. The issue with correlated errors is directly addressed later in this paper.

3D global regional and teleseismic models

Regional models CUB1 and CUB2

The 3D model CUB1 and its update CUB2 were constructed using a Monte-Carlo inversion method (Shapiro and Ritzwoller, 2002) applied to group (Ritzwoller and Levshin, 1998) and phase velocity dispersion curves (Trampert and Woodhouse, 1995; Ekström and Dziewonski, 1998). Both models are given on a 2°x2° grid to a depth of 400 km. Below 400 km both models revert to the Harvard 3D model S20a (Ekström and Dziewonski, 1998). There are three principal differences between CUB1 and CUB2.

First, the models differ in the crustal reference model used. The crustal reference for CUB1 is CRUST5.1 of Mooney et al. (1998), whereas CUB2 uses CRUST2.0 (Bassin et al., 2000). Second, the models are based on different surface wave tomography methods. CUB1 uses Gaussian tomography that is based on geometrical ray-theory with intuitive Gaussian smoothing constraints to simulate surface wave sensitivities (Barmin et al., 2001). CUB2 is based on diffraction tomography that uses a simplified version of the scattering sensitivity kernels that emerge from the Born or Rytov approximations (Ritzwoller et al, 2002b). Diffraction tomography accounts for path-length dependent sensitivity, wave-front healing and associated diffraction effects, and provides a more accurate assessment of spatially variable resolution than traditional tomographic methods. Third, the models differ in the way in which Vp is computed from Vs. CUB1 uses a simple empirical logarithmic scaling relation $d\ln(Vp)/d\ln(Vs) \sim 0.5$, in which perturbations are taken relative to the Vs and Vp velocities in ak135 (Kennett et al., 1995). CUB2 uses a theoretical conversion based on mineralogical partial derivatives for a hypothetical composition of the upper mantle. The method is based on the work of Goes et al. (2000), as described in detail by Shapiro and Ritzwoller (2002b). In essence, given the mineralogical composition, the bases for the anelastic correction are partial derivatives of the elastic moduli with respect to the independent variables at infinite frequency, a mixing law, and a relation between temperature and shear Q. This Vs model is anchored to tectonic IASPEI91 travel times. In CUB2 this transformation has not yet been regionally tuned. Mineralogical composition is homogeneous across the region of study, there has been no account for the possible effects of fluids in the mantle beneath tectonically deformed regions, and shear Q is purely a function of temperature.

To calculate CUB Pn SSSCs for use in event relocation, we used a 2-D raytracer that handles refracted and reflected P wave in 3D laterally inhomogeneous media along 2D cross section of a spherical earth by ray shooting. Regional travel times were computed along a set of profiles radiating from each station up to distances of 20° for an upper-crustal source depth of 10 km. We chose an azimuthal spacing of 3° between the profiles and along each profile, travel times were computed at a set of points spaced at every 25 km. After sub

corrections were interpolated to a 1°×1° rectangular, geographic grid centered on each

station. Figure 1 shows an example of regional CUB Pn SSSCs for station ABKT (Alibek, Turkmenistan).

We derived an empirical modeling error for the CUB Pn SSSCs, as shown in Figure 2, using travel time misfits obtained from the EHB events (Engdahl et al., 1998). Ray tracing was performed to predict the CUB travel times for over 1,000,000 Pn rays. Misfits were obtained between the predicted corrections (w.r.t. IASPEI91) and the EHB travel time residuals (w.r.t. ak135; Kennett et al., 1995) as the two 1D travel times are the same for regional phases. Modeling errors were estimated from the standard deviations of the misfits as a function of epicentral distance. In lieu of detailed empirical error maps or error surfaces produced by a robust theory of error propagation from uncertainties in the 3-D model, we chose this simple and conservative approach to estimate model errors. This model error estimate is azimuthally invariant. It is similar in form to the baseline currently used in routine IMS location calculations (Figure 2). Typically, the model errors for the CUB models are about half of the corresponding values for IASPEI91.

Teleseismic model J362

Recently Antolik et al. (2003) have developed a joint compressional and shear velocity model of the Earth's mantle, J362. This is a spherical harmonic degree-18 model with a horizontal length scale of 1000 km. Absolute and differential body-wave travel times as well as surface wave dispersion measurements have been used to develop J362. Using a data set of GT0-10 earthquakes and explosions, Antolik et al. (2003) have shown that J362 achieves an improvement of about 10% in RMS mislocation for explosions, relative to SP12. J362 also decreases the origin time error by an average of 0.05 sec over the SP12 model.

Teleseismic P-wave SSSCs were computed with a perturbation theory based ray-tracer (ray bending) in the distance range 25º to 97º, in order to diminish the biasing effects of upper mantle triplications and the core-diffracted phases on the travel-times. We included the effect of a realistic crustal structure by combining this model with the CRUST2.0 model of Bassin et al. (2000). The ray paths were calculated by using the 1-D velocity model PREM (Dziewonski and Anderson, 1981). Variation in the raypaths due to the large-scale 3-D structure in the model is expected to be small. The ray tracer

unstable when computing travel times for the upper mantle triplication phases. Identifying the appropriate phase branch can be problematic. Therefore, we limited our computations to distance ranges greater than 25°. Similarly, to avoid the effect of strong lateral heterogeneity at D", we limited the maximum distance of our calculations to 97°. We used a representative upper-crustal source depth of 10 km for the computations. We generated the SSSCs for the Group 2 Consortium study region, i.e., in a rectangular grid with latitude of -15º to 80º N and longitude of -40º and 100º E, for a global distribution of stations. To balance computational efficiency and numerical accuracy, we chose to compute the teleseismic SSSCs in a 2° grid and accept 0.15 s as an estimate of numerical noise. Figure 1 shows an example of teleseismic P SSSCs for ABKT. For modeling errors, based on our experience with the regional calibrations we adopt a scaled version of the IASPEI91 P error model (Figure 2):

$$\text{Model\_error}^2 = (0.5*\text{PIDC\_error})^2 + 0.15^2 \text{ travel-times}$$

For areas within the study region but less than 25° or more than 97° away from a station, the corrections are equal to zero and the modeling errors are similar to those used for IASPEI91.

Baseline differences between regional and teleseismic

Since the regional and teleseismic models were developed independently, it is possible that baseline differences exist between models. Therefore, before relocating GT events using both Pn and teleseismic P, we need to examine if bias exists between the regional and teleseismic calibrations and whether any baseline shifts may are warranted. To assess the baselines of each model, comparisons were made for time residuals of over 700 GT events in the Group 2 database between regional and teleseismic phases using IASPEI91, CUB1, CUB2, and J362. We only included reasonable Pn and P arrivals (with absolute time residuals less than 5 sec) and reliable GT events (with 5 or more Pn phases and 5 or more P phases). In general, there are more P phases than Pn phases. For each GT event, time residuals of Pn and P phases were calculated using observed GT arrivals and predicted travel times from each model for fixed GT locations. We then computed the mean and standard deviation of Pn and P time residuals, respectively, for each GT event.

Comparisons were made for several combinations of regional and teleseismic models. Table 1 reveals that significant biases exist between several of the travel-time tables and the standard deviations reveal which models are poorly matched. In particular, IASPEI91 regional and IASPEI91 teleseismic phases show a statistically significant baseline offset of –0.2 sec and a standard deviation of 1.1 sec. The combination of CUB2 and J362 has the smallest baseline difference and the smallest standard deviation, but an adjustment can be made to CUB1 of about 0.75 seconds to bring it into alignment with J362. It is noteworthy that combinations of CUB1 and CUB2 with J362 or IASPEI91 reduce the standard deviation with respect to combinations of Pn IASPEI91 with respect to teleseismic IASPEI91. The results do suggest that teleseismic P calibrations derived from J362 combined with uncalibrated IASPEI91 Pn may be a bad match.

A baseline correction can be applied to any of the models to reduce the mean differences, but the standard deviations reveal event-by-event baseline variations that cannot be remedied by simple bulk adjustments to the travel-time tables. The biases are systematic and reflect the broad tectonic provinces of shield and platform vs. tectonically active regions. In this work we chose to reduce the baseline shift by simply applying a bulk correction of 0.75 sec to the regional phase. Event location may benefit from more sophisticated scheme in reducing the bias, such as incorporating the shift in the SSSCs for different stations, azimuths, and distances.

Methodologies and data sets for validation testing

We validate the global models by relocating a large set of GT events in the study region. Only reliable GT data were selected for use in event location to avoid ambiguity between model and data uncertainties. Events were relocated using all stations as well as SSNB using regional (Pn phase) and teleseismic (P phase) calibrations, separately and combined. For the combined (Pn and teleseismic P) case a baseline shift of 0.75 sec was applied to Pn to reduce inherent bias between regional and teleseismic calibrations. We evaluated event location improvements by comparing the results with those without calibration using a set of

depth was fixed to zero since these events are mostly shallow.

Data selection for all-station relocations

We selected all GT0-GT10 events in the Group-2 database that can be reliably relocated (1) using only regional phases, (2) using only teleseismic phases, and (3) using both regional and teleseismic phases. To minimize ambiguity in arrival data, we only used Pn arrivals within 15º and P arrivals between 25º and 97º. We selected events in three groups with sgap less than 160º. This guarantees the maximum azimuthal gap is less than 160º and the number of stations is at least 5.

In the IMS event location algorithm, *GT10* events with sgap < 160º for Pn only, and with epicentral distances calculate error ellipses. The errors are partitioned into modeling errors and measurement errors. Estimates of measurement errors based on signal-to-noise ratio are not available from bulletin data. A measurement error of 1.0 sec was therefore assigned for most to arrivals by default. The model errors assigned to the regional and teleseismic arrivals are described in the previous Section.

Pn data set

There are 526 GT0-
within 15º (Figure 3). This data set provides a fairly good geographic coverage across the region. They are well-recorded events, with a median number of defining phase of 51 and a minimum ndef of 6. Most of the events are GT5 (87%), and 64% of them are from the EHB that were selected as GT5 events at the 95% confidence level. Another 26% of the GT events are earthquake clusters, mostly GT5 events generated from cluster analyses (Engdahl and Bergman, 2001). The remaining 10% of events are mostly nuclear, chemical explosions, or mining events (GT0-GT2). Pn SSSCs were calculated for all 1098 stations for the CUB models.

P data set

There are 793 GT0-GT10 events with sgap < 160º for P only at epicentral distances between 25º and 97º (Figure 4). These events are also well recorded, with median ndef of 100. P SSSCs are calculated for 2821 stations. Large portions of the events are GT1 (40%) and GT5 (35%). About 70% of the events are clusters; 30% of them are nuclear explosion clusters. For the remaining events, about 17% are nuclear explosions, and about 10% are EHB GT5 events.

each of the

Pn and P data set

There are 1234 GT0-GT10 events with sgap < 160º for Pn within 15º and P between 25º and 97º. They are 28 GT0, 328 GT1 (27%), 14 GT2, 650 GT5 (53%), and 214 GT10 (17%) events. These include 111,498 P from 2823 stations and 39017 Pn phases from 1166 stations. The events were well recorded, with a median ndef of 84. Most of the events (80%) have more teleseismic P than Pn arrivals. The median ratio of the numbers of defining P to Pn is 3. The maximum azimuthal gap is 154º and 87% of the events have an azimuth gap less than 100º. A large number of the events (24%) are explosions at historic test sites or PNE explosions scattered across the former Soviet Union (7%). The majority of the remaining events are earthquakes or mining related events.

Constrained bootstrapping

In order to maximize the discriminatory power of the relocation test, we devised a technique that for a fixed number of stations (1) samples azimuth as uniformly as possible to avoid statistical dependence on correlated ray paths, (2) minimizes the largest azimuthal gap at each realization for stable locations, and (3) avoids over emphasizing a single key station/arrival. We chose network geometries with 10 stations and secondary azimuthal gaps less than 160º. We find acceptable SSNB geometries by building the 'sgap tree', illustrated by the fictitious example in Figure 5. The sgap tree is a hierarchical binary tree where each node represents a station that splits the remaining azimuthal gap in a way that the secondary gap closed by the station is maximized. If there are several stations situated at the same azimuth, we pick one of them randomly. The sgap tree stemming from a particular station is deterministic and unique. Cutting the tree at the

smallest secondary azimuthal gap for the given number of stations. If this secondary gap is smaller than the prescribed secondary gap limit, we accept the geometry. The procedure is repeated for each station in the network, which yields all network geometry meeting the constraints defined above. However, because of possible symmetries, different root stations may generate the same station configurations. To avoid repeating geometries we retain only the unique station configurations. We wish to select from this set of sparse networks a balanced subset that does not overly emphasize any single station.

The set of acceptable station geometries are typically large. Some stations, situated at strategic azimuths, are sampled all over again, while less important stations appear only once or twice. The individual station configurations can therefore be considered as the basis vectors spanning the 'network geometry space'. We orthogonalize this space using an algorithm analogous to the Gram-Schmitt orthogonalization of a matrix and select the 20 most representative network geometries. The selected station configurations will then represent the SSNB realizations of an event. Since we select the most representative subnetworks as opposed to just blindly taking random subsets of stations as in traditional bootstrapping, we call this approach constrained bootstrapping. This procedure provides a controlled relocation experiment and allows better estimations of calibrated vs. uncalibrated mislocations and their statistical uncertainties. It also provides some quality control over the test data set, as some events are found to contain outlier arrivals that produce unstable populations of SSNB locatio

SSNB data sets

To generate SSNBs we selected well-recorded GT0-GT10 events in a way that they provide good geographical coverage. We limited the number of events from the same event cluster to 10 to avoid over-representing clusters with a large number of well-recorded events. Sampling a cluster by several events allowed us to examine the consistency of location bias estimates and identify outliers. The SSNB "seed" events were selected for the validation of regional and teleseismic calibrated travel times, respectively, from the events shown in Figures 3-4. The common events, 116 events in total, provided the basis for the evaluation of combined regional and teleseismic calibration travel times.

For direct comparisons SSNB seed events were also relocated using all stations, denoted as "all-station (seed)". Note that in general the all-station cases discussed in this paper refer to the entire data sets. The all-station (seed) results are for the SSNB "seed" events only, which use subsets of the entire data sets (282 vs. 526 events for Pn data sets; 359 vs. 793 events for P data sets; and 116 vs. 1234 events for Pn and P data sets).

Location evaluation metrics

For location improvement comparisons, we evaluated statistics on mislocation, error ellipse area, 90% error ellipse coverage, origin time bias, origin time error, and standard deviation of observations. GT uncertainties (GTX) are included when evaluating location improvement. Events are considered indecisive when either the calibrated and uncalibrated locations are known to only a finite accuracy, if both the uncalibrated and calibrated solutions are located within this range we cannot really decide if the event location was improved or deteriorated. An event location may move within the range of GT accuracy using calibrated travel-times, while the uncalibrated solution remains outside the GT accuracy range. This case represents clear improvements. The opposite case, when the uncalibrated location is within the GT accuracy while the calibrated location moves out of the GT accuracy range represents clear deterioration. Finally, both the calibrated and uncalibrated solutions may be located beyond the GT accuracy range. Events in this category are considered improved if the mislocation is smaller for calibrated case than the uncalibrated. These four categories result in the overall statistics on events improved, deteriorated, and indecisive.

Comparisons using the aforementioned four categories provide direct assessment on mislocation relative to GTX with and without calibration. Note that these categories are constrained by the IASPEI91 location, since our major goal is to compare location improvement w.r.t. IASPEI91. Such analysis can also be done for direct comparisons between two calibration models, e.g. CUB1 and CUB2. In fact, simply comparing the numbers in the categories w.r.t. IASPEI91 may be misleading when trying to evaluate two calibration models. Locations from two calibration models can be easily much better than IASPEI91, so the numbers in each categories for such a case may drastically differ from the IASPEI91 based categories. For example, a significantly larger number of events may be indecisive for the case with two calibration models, since the events may be better located. Therefore, such analyses are only meaningful when the categories are constrained to the models in comparisons, instead of simply based on the IASPEI91.

We calculate the 90% coverage by incorporating the GT uncertainty, GTX (Figure 6). The coverage parameter E is defined as:

ted locations are within the GT uncertainty, or they are identical. Since the GT

$$E = x^2/(smajax^2+GTX^2)+y^2/(sminax^2+GTX^2)$$

E follows a χ^2 distribution with two-degrees of freedom. The 90% coverage corresponds to E=1.0, and an event is not covered when E>1.0.

Additional metrics is also developed to measure the performance of the SSSCs. For SSNB, location bias is evaluated since each event provides a number of realizations. In such a case the distance between the GT location and the centroid of the realizations serves as an estimate of the location bias (with estimated uncertainty) due to un-modeled lateral heterogeneities.

In general travel-time calibration should significantly reduce median mislocation and error ellipse area as well as variances:

- Median mislocation of GT events should be significantly reduced
- Median error ellipse area should be reduced
- Coverage (% of GT events inside error ellipses) should be the same or better
- Fit, as expressed by residuals or their variance, should be similar or better

Event relocation using regional Pn and teleseismic P calibration

All-station relocation

We relocated all 526 GT events in the Pn data set (Figure 3), 793 GT events in the P data set (Figure 4), and 1234 GT events in the Pn and P data set using all stations. These events were well recorded by Pn and P phases, with median ndefs of 51-100, gaps of 58°-76°, and sgaps of 76°-99°. We compared the relocations results of each data set from different models w.r.t. IASPEI91. Table 2 compares the number and percentage of event mislocations for Pn only (CUB1 and CUB2), P (J362) only, and the combined Pn (CUB1 and CUB2) and P (J362) cases. Mislocation is significantly improved for all cases except that there is only marginal improvement using CUB2. In general, there are about 20% more events improved than deteriorated, and about 10%-20% of total events are located within GTX uncertainty (last three rows in Table 2). This fraction of the population is well located with or without calibration (about 5 km) or the GTX uncerta
to measure a significant mislocation (row #2 in Table 2). GT events with location

accuracy better than 5 km are essential in relocation testing. There are more events located outside GTX uncertainty without calibration but within GTX with calibration (row #3 in Table 2), compared to the events located within GTX without calibration but outside GTX with calibration (row #4 in Table 2). The majority of events are located outside GTX with or without calibration. Of these, significantly more events are improved (row #5) than deteriorated (row #6) with calibration.

Relocation statistics on median mislocation, error ellipse area and coverage, origin time and error, and sdobs are summarized in Table 3. Figure 7 shows the cumulative mislocations with and without calibration for each data set. For all 526 GT0-GT10 events in the Pn data set using CUB1, the median mislocation is 6.6 km. The largest improvement occurred at the 95th percentile, with mislocation reduced by 27% (Figure 7a). These poorly located events are distributed throughout the region. For all 793 GT0-GT10 events in the P only data set using J362, the median mislocation is 6.1 km. The largest improvements occurred at the 30th percentile, with mislocation reduced by 29% (Figure 7b). The poorly located events, mostly GT10 events along the Mid-Ocean Ridge and Transforms, did not improve using J362. For all 1234 events in the Pn and P data set using CUB1 Pn and J362 P, there is significant location improvement as those in separate calibration, but the mislocation is further reduced by joint calibration (Figure 7c). As in the P only case, with joint calibration the largest improvement occurred at the 25th percentile, while the poorly located events did not improve.

The GT events in all three data sets are generally well located, given the good azimuthal distributions of stations and the large number of observations. The median mislocation is 6-8 km, with and without calibration. There is large mislocation improvement using CUB1 while CUB2 does not show improvement. Overall, the median mislocation was reduced by 10-20%. By far teleseismic calibration has delivered the largest location improvements (27%). Using both regional and teleseismic SSSCs the mislocation has been reduced to the level of GT5 uncertainties (5.7 km). Unlike regional calibration alone, for the case of joint calibration location results from CUB1 are only somewhat better than those with CUB2 since teleseismic phases play the dominant role. Table 3 also shows that the relocation results using the bulk correction are slightly better than that

without baseline shift. Direct comparisons on CUB1 vs. CUB2 and bulk vs. no bulk correction are discussed in the next Section.

We also compared the 227 events in common between the three data sets, Pn only, P only, and joint Pn and P. As expected, there are more events improved than deteriorated with combined calibrations, comparing with events improved using regional or teleseismic calibration alone. Mislocation is the smallest using joint regional and teleseismic calibration (5.4 km vs. 7.2 km for Pn only and 8.1 km for P only), all better than the corresponding IASPEI91 results. The reduction in mislocation using calibration is similar in each case (13%-16%). Similar to the results in Table 2, again it appears that, with regional or teleseismic calibration alone, more events may be improved than the joint calibration case. This does not infer that combined calibration is less effective but rather an effect from the very large ndef (median of 157). With a large number of observations, events are already well located even without calibration (median of 6.2 km).

As shown in Table 3, using calibration there are also significant reductions in error ellipse area, origin time errors, and sdobs. However, the median origin time bias is increased and mostly positive, indicating that these models are biased or the GT event origin times. Note that, while the GT events possess fairly accurate epicenter estimates (less than 10 km), the origin times may not be very reliable. Using calibration the error ellipse coverage is significantly lower than 90%. Figure 8 shows the coverage parameter vs. ndef for combined regional Pn and teleseismic P calibration using all stations. The 90% coverage is true for ndef<=50 for uncalibrated case and for ndef<=27 for calibrated case.

SSNB relocation

The SSNB realizations were relocated by both calibrated and uncalibrated regional travel-
times. As in the all-station locations, the CUB1 Pn travel times were also adjusted
to accommodate the 0.75 sec baseline shift between CUB1 and J362.

Figure 9 shows examples of SSNB relocations for the 02/14/1977 GT5 earthquake in Pakistan and the 08/20/1972 GT1 PNE in Russia. The open triangles represent the relocations using the SSNB sub-networks with calibrated and uncalibrated regional Pn

travel-times predicted by the CUB1 and IASPEI models. The vector connecting the GT location with the centroid of SSNB realizations represents the location bias due to un-modeled lateral heterogeneities. The uncertainty in location bias, shown as ellipses around the head of the bias vector, is derived from the individual SSNB sub-network locations. As expected, calibrated travel-times reduce location bias.

Figure 10 shows the mean mislocation from the all-station (seed) locations (top) compared to the mean location bias from the 10-station sparse network solutions (middle) using regional and teleseismic calibrated (CUB1+J362) and uncalibrated (IASPEI91) travel-times for event clusters with multiple SSNB seeds. The event clusters are sorted by their GT accuracy (thick line). The mean mislocation and bias across all the clusters are about 5-7 km. Therefore for clusters of GT10 quality most of the events are located within the GT accuracy, regardless of using calibrated or uncalibrated travel-times. This result indicates that GT10 events have only limited use for location calibration purposes. While using all-stations (seed) 57% of the clusters show location improvements due to calibrated travel-times, the SSNB bias estimate exhibits improvement for 78% of the clusters. The improvements in all-station (seed) locations and the location bias obtained from the SSNBs are shown at the bottom of Figure 10. In general, the measurable improvements are larger and more consistent for SSNBs, indicating that the constrained bootstrapping provides a more sensitive environment to evaluate calibrated travel-times and indeed increase the statistical power of the relocation tests.

Table 4 summarizes the overall percent of events improved, deteriorated or remained indecisive due to the application of calibrated travel-times. The statistics on the SSNB centroids represent the reduction in location bias. The SSNB 10-station locations represent the reduction in mislocation obtained from the constrained bootstrapping. For the sake of comparison the results from the all-station (seed) locations of the same SSNB seeds are also given. More events are improved than deteriorated. Calibrated travel-times reduce both bias and mislocation. When mixing calibrated travel-times with uncalibrated ones (last three columns in Table 4) the improvements tend to fade away. Thus it is important to calibrate both regional and teleseismic phases.

Table 5 gives the principal metrics on location bias, mislocation, coverage and the area of the 90% error ellipse for the cases of regional, teleseismic and combined regional and

teleseismic calibration. As in Table 4, both the SSNB and all-station (seed) location results are listed. The location bias and mislocation are reduced in all cases. Similarly to Table 2,

calibrated travel-times. Note the striking difference between the median area of error ellipses obtained from the SSNB and all-station (seed) locations. The ellipses are significantly smaller when using all stations (seed), resulting in much poorer coverage statistics. Since the SSNB seed events are typically recorded by more than 100 stations, many stations sample similar ray paths. Thus the assumption of independent errors made by the location algorithm is violated, leading to unrealistic error ellipses. We further discuss the issue of correlated errors in the next Section.

We compared the percentage of events with improved/deteriorated locations due to calibrated travel-times for the regional (CUB1, CUB2), teleseismic (J362), and the combined regional and teleseismic (CUB1+J362, CUB2+J362) cases. In all cases calibration improved 60-70% of events. CUB1 improved somewhat more events t CUB2 when only regional phases were used or when combined with J362 teleseismic travel-time predictions. On the other hand, CUB2 yields somewhat larger reduction in location bias. The statistics for the area of error ellipse and coverage do not significantly differ for CUB1 and CUB2 either alone, or combined with J362.

Comparisons for the 116 common events in all combinations show that in general calibrated travel-times reduce location bias. The most profound reduction in bias was achieved when both regional and teleseismic calibrated travel-times were used. The median calibrated location bias for the combined regional and teleseismic case is about 6 km, approaching the uncertainty in the majority of GT event locations. In the regional case CUB2 provides somewhat larger improvements than CUB1. The median area of the error ellipse is significantly reduced in all cases, dropping below or approaching 1000 km^2. The reduction in the area of error ellipse is achieved without significant deterioration in coverage. Calibration provides improvements in all cases, and combining regional and teleseismic phases further improves event locations, using either uncalibrated or calibrated travel-times. Since constrained bootstrapping increases the statistical power of the test, it shows more significant improvements between uncalibrated and calibrated travel-times than the all-station relocations. Improvements

the improvements are significantly smaller when mixing uncalibrated and

due to the combination of regional and teleseismic phases are comparable to those from the all-station results.

Discussions

Baseline shift between regional and teleseismic calibrations

In the event relocation, we applied a bulk correction of 0.75 sec to regional Pn to correct the baseline shift between regional and teleseismic calibration. To assess the effect of this practice, we also relocated events using joint CUB1 Pn and J362 P SSSCs without the bulk correction. As described previously, direct comparisons should be made for this case (column #1 in Table 6), instead of using the results w.r.t. IASPEI91 (Table 2). The results are somewhat better when using the bulk correction. Slightly more events are improved than deteriorated, but about half of all events cannot resolve one way or the other. The overall metrics is also slightly better when using the bulk correction (Table 3).

Comparisons between CUB1 and CUB2

The Pn travel times predicted by CUB1 and CUB2 were compared with empirical path corrections for both explosions and earthquakes (Ritzwoller et al., 2002; Bhatacharryya et al, 2003). The general distribution of anomalies is similar for both models but there are some regional variations. With respect to empirical path corrections, CUB1 is generally better correlated and shows greater variance reduction than CUB2 in the Mediterranean. Correlations for CUB1/CUB2 and J362 are poorest in Greece. The correlations of predicted SSSCs between CUB1 and CUB2 are low in the Mediterranean highlighting where these two models are the most different. In this area the CUB1 SSSCs are faster but CUB2 SSSCs are slower (Figure 11).

Relocation results using all stations indicate that overall CUB1 does better than CUB2 in improving event location. Compared to CUB2, there are significantly more events improved than deteriorated using CUB1, both with and without teleseismic calibration (last two columns in Table 6). The events improved using CUB1 are mostly in the Mediterranean (Figure 12). It is plausible that, the mineral geology differ between Asia and the Middle East, so the S to P mapping for CUB1 and CUB2 work better at different regions. For best effect of location calibration, the CUB1 model should be used in this region.

The Mediterranean is a tectonically complex region. Besides CUB2, J362 also performs poorly in this region, compared to the IASPEI91 (column #4 in Table 7). Unlike CUB1 (column #1 in Table 7), in both cases more events are deteriorated than improved (columns #2 and #4 in Table 7). Again, direct comparisons between CUB1 and CUB2 show that CUB1 is better (column #3 in Table 7). When joining with teleseismic calibration, the results are also poor, even for the joint CUB1 case, due to the effect of large number of teleseismic phases (columns #5 and #6 in Table 7). As expected, direct comparisons show that the results are still better when joining CUB1 than CUB2 with teleseismic calibration (last column in Table 7). Unlike the statistics for all events (Table 3), in this region the median mislocation is similar for all models, including the IASPEI91.

Note that in this region the GT events have lower quality. Most of them are GT5 (91%) and GT10 (3%), and 66% of all events in this region were selected from the EHB bulletin based on the Group-2 validated using cluster analysis as in this region to improve the models and data.

Correlated errors vs. 90% coverage

The 90% coverage is very low using all stations for each case with separated or combined calibration (Figure 8). While this problem may be due to under *a priori* errors, SSBN results detect that the problem is related to the breakdown of the assumption of uncorrelated errors.

Herrin and Taggart (1968) have shown by analyzing the Long Shot nuclear explosion that a large number of arrivals sampling similar ray paths along unmodeled three-dimensional Earth structure may introduce location bias. In the Long Shot case a large number of ray paths travel through a subducted oceanic slab with high seismic velocity, thus arriving systematically earlier than the predicted arrival times. The systematic travel-time prediction bias results in a 26 km mislocation, way outside the confidence ellipse (139 km²), which assumes uncorrelated errors. More recently Myers and Schultz (2000b) have pointed out that tr

Although in the past few years several non-linear hypocenter location methods were developed to account for non-linearity and non-Gaussian error distributions (e.g. Billings, 1994; Billings et al., 1994; Sambridge and Gallagher, 1993; Sambridge and Kennett, 2001; Rodi et al., 2002), most routinely used location algorithm (including the one we used in this study) assumes Gaussian, independent errors. One of our major motivations for developing the constrained bootstrapping technique was to minimize the effect of correlated errors.

To demonstrate how correlated errors may deteriorate coverage, introduce location bias and produce unrealistic error ellipses, we performed an SSNB study on the 1994/10/7 03:26:00 Lop Nor, China underground nuclear explosion (GT1). This event was recorded by 562 stations in the teleseismic (25º-97º) distance range. The station is far from uniform, dominated by stations in Europe, Japan and California. We generated SSNBs ranging from 6 to 400 stations, each represented by 20 realizations. The median location bias increases with the number of stations, using calibrated and uncalibrated travel-times (Figure 11). As more and more stations are added to the solution, the location is driven away from the GT location. Although J362 decreases the location bias, the pattern of the location changes between the two is quite similar, indicating that the relative importance of some station clusters (possibly the Californian network) steadily increases because the location algorithm ignores the fact that travel-time predictions are correlated along similar ray paths. Note that even if the J362 model precisely accounted for all three-dimensional heterogeneities in the Earth, the false assumption of independent errors would still resulted in biased location.

As shown in Figure 13, as the information carried by the network geometry is exhausted relatively early (the azimuthal gap stabilizes after about 20 stations, the secondary gap after 100 stations), adding further stations simply increases redundancy. However, because of the covariance matrix is calculated incorrectly by assuming that the errors are uncorrelated, the area of the error ellipse shrinks indefinitely. This results in an approximately linear increase in the coverage parameter, meaning that after a while the true epicenter lies outside the error ellipse. On the other hand, the median misfit (RMS residual) does not offer any clue about the location quality either: it remains basically constant once the secondary gap information is exhausted. In this particular example the

location bias due to correlated errors increases with the number of stations. This may not always be the case. Location bias may or may not increase as more arrivals are added to the solution. The actual behavior depends on how the net weights of closely clustered stations are balanced with respect to the unmodeled 3D Earth structures.

Correlated errors will always produce unrealistic error ellipses for events located by a large number of arrivals. The constrained bootstrapping was developed to minimize the effect of correlated errors. Figure 14 shows the cumulative histograms of the coverage parameter for the all-stations (seed) locations and the 10-station sparse network locations for the same events using calibrated and uncalibrated regional and teleseismic travel-times. While the error ellipses of the all-station (seed) locations cover only 70% of the true epicenters, the SSNB calibrated locations are covered about 90% of the time by the 90% coverage ellipse.

The 90% coverages are improved using calibration but the 90th percentile of coverage is still low for the low-ndef SSNB tests for the teleseismic case. A conservative re-scaling of the IASPEI91 model errors may be estimated from the square root of the coverage ratio as a function of cumulative percentage. To ensure 90% coverage at the 90th percentile the model error needs to be increased to approximately 70% of the IASPEI91 model error. This corresponds to a net 50% variance reduction with respect to the IASPEI91 model errors.

A fundamental approach in addressing the 90% coverage problem is to construct the covariance matrix in a way that it takes into account the correlation between similar ray paths. Chang et al. (1983) proposed a scheme to estimate the covariance matrix in the presence of inter-correlated errors. Bayesian kriging (Myers and Schultz, 2000a) routinely deals with this problem. A modern location algorithm should take into consideration non-linearity, as well as non-independent errors. Sparse, uniformly spaced station networks, such as the IMS network, are less vulnerable to correlated errors. However, the formal location uncertainties published by international agencies that use large number of stations to locate events, such as the ISC (International Seismological Centre) and NEIC (National Earthquake Information Center), are likely to be underestimated.

Conclusions

We have validated the regional CUB models and teleseismic J362 model using a large set of high quality GT events. Relocation results using all-station and SSNB show that there are large location improvements using regional or teleseismic calibration. This is consistent with cluster analysis as well as other previous relocation studies that demonstrated significant improvement using either regional or teleseismic model-based calibration alone. Since in practice regional and teleseismic phases are used jointly, we relocated more than 1000 GT0-GT10 events with Pn and P SSSCs. Our results show that the largest location improvement is achieved by combined regional and teleseismic calibration, which has reached the GT5 accuracy level. In combining the regional and teleseismic calibration, the baseline differences need to be adjusted, at least using a simple bulk correction.

Constrained bootstrapping not only provides a controlled experiment designed to increase the statistical power of relocation results, but also allows the removing location bias due to unmodeled 3D Earth structures. In general, the improvements are larger and more consistent for SSNBs than those for the all-station locations. A corollary of the SSNB approach is that for any well-recorded event a set of sub-networks exists that locate the event better than the entire network.

We have shown that the 90% coverage error ellipses obtained from the all-station locations are unrealistically small and do not cover the true locations 90% of the time. This is due to the assumptions of Gaussian and independent errors that are built in most routinely used location algorithms. While constrained bootstrapping minimizes the effect of correlated errors, the ultimate remedy would be to improve the location algorithm, e.g. using non-Gaussian/non-
matrix.

Acknowledgements

This work was sponsored by the Defense Threat Reduction Agency under contract DTRA01-00-C-0013.

References

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

Antolik, M., Y.J. Gu, G. Ekström and A. Dziewonski (2003), J362D28: A new joint model of compressional and shear velocities in the Earth's mantle, Geophys. J. Int., in press.

Armbruster, J., V. Burlacu, M. Fisk, V. Khalturin, W.-Y. Kim, I. Morozov, E. Morozova, P. Richards, D. Schaff, and F. Waldhauser (2002), Seismic location calibration for thirty International Monitoring System stations in Eastern Asia, Proceedings of 24th Seismic Research Review- Nuclear Explosion Monitoring: Innovation and Integration, Ponte Vedra Beach, Florida, September 17-19.

Barmin, M.P., M.H. Ritzwoller, and A.L. Levshin (2001), A fast and reliable met surface wave tomography, Pure Appl. Geophys., 158, 1351 – 1375.

Bassin, C., G. Laske and G. Masters (2000), The Current Limits of Resolution for Surface Wave Tomography in North America, EOS Trans AGU, 81.

Bhattacharyya, J., H. Israelsson, K. McLaughlin, E.R Engdahl, and E. Bergman (2003), Validation of a global and a regional model using empirical path corrections derived from event clusters, in preparation.

Billings (1994), Simulated annealing for earthquake location, Geophys. J. Int., 118, 680-692.

Billings, S.D., M.S. Sambridge, B.L.N. Kennett (1994), Errors in hypercentral location: picking, model, and magnitude dependence, Bull. Seism. Soc. Am. 84, 1978-1990.

Bondár, I., S. Myers, E.R. Engdahl, and E. Bergman (2002), Epicenter accuracy base on seismic network criteria, submitted to Geophy. J. Int.

Bondár, I., E.R. Engdahl, X. Yang, H. Ghalib, A. Hofstetter, V. Kirichenko, R. Wagner, I. Gupta, G. Ekström, E. Bergman, H. Israelsson, and K. McLaughlin (2003), Collection of a reference event set for regional and telseismic location calibration, in preparation.

Chang, A., R. Shumway, R. Blandford, and B. Barker (1983), Two methods to improve location estimates- preliminary results, Bull. Seism. Soc. Am. 73, 281-295.

Dziewonski, A.M., and D.L. Anderson (1981), Preliminary reference earth model, Phys. Earth Plannet. Inter., 25, 297-356.

Ekström, G. and A.M. Dziewonski (1998), The unique anisotropy of the Pacific upper mantle, Nature, 394, 168-172.

Engdahl, E.R., R. van der Hilst, and R. Buland (1998), Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, Bull. Seism. Soc. Am. 88, 727-743.

Engdahl, E.R. and E.A. Bergman (2001), Validation and generation of reference events by cluster analysis, Proceedings of 23rd Seismic Research Review, Jackson Hole, WY, October 15.

Goes, S., R. Govers, and R. Vacher (2000), Shallow mantle temperatures under Europe from P and S wave tomography, J. Geophys. Res., 105, 11,153-11,169.

Herrin E. and J. Taggart (1968), Source bias in epicenter determinations, Bull. Seism. Soc. Am., 58, 1791-1796.

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

Kennett, B. and E.R. Engdahl (1991), Travel Times for Global Earthquake Location and Phase Identification, Geophys. J. Int., 105, 429-465.

Kennett, B.L.N., E.R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from

Mooney, W.D., G. Laske, and G. Masters (1998), CRUST5.1: A global model at 5 degrees by 5 degrees. J. Geophys. Res., 102, 727-748.

Murphy, J., W. Rodi, M. Johnson, J. Sultanov, T. Bennett, M.N. Toksoz, C. Vincent, V. Ovtchinnikov, B. Barker, A. Rosca, and Y. Shchukin (2002), Seismic calibration of Group 1 International Monitoring System (IMS) stations in Eastern Asia for improved event location, Proceedings of 24th Seismic Research Review- Nuclear Explosion Monitoring: Innovation and Integration, Ponte Vedra Beach, Florida, September 17-19.

Myers, S.C. and C. Shultz (2000a), Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, Bull. Seism. Soc. Am.,

Myers, S.C and C.A. Schultz (2000b), Calibration of seismic travel time using events with seismically determined locations and origin times, EOS Trans. AGU, 81, F845.

Ritzwoller, M.H. and A.L. Levshin (1998), Eurasian surface wave tomography: Group velocities, J. Geophys. Res., 103, 4839 – 4878.

Ritzwoller, M.H., N.M. Shapiro, A. Levshin, E.A. Bergman and E.R. Engdahl (2002a), Assessment of global 3-D models based on regional ground truth locations and travel times, submitted to J. Geophys. Res.

Ritzwoller, M.H., N.M. Shapiro, M.P. Barmin and A. Levshin (2002b), Global surface wave diffraction tomography, submitted to J. Geophys. Res.

Rodi, W., C. Shultz, W. Hanley, S. Sarkar, and H. Kuleli (2002), Grid-search location methods for ground-

Sambridge, M., and K. Gallagher (1993), Earthquake hypocenter location using genetic algorithms, Bull. Seism. Soc. Am. 83, 1467-1491.

Sambridge, M., and B.L.N. Kennet (2001), Seismic event location: non-linear inversion using a neighborhood algorithm, Pure appl. Geophys. 158, 241-257.

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

Shapiro, N.M. and M.H. Ritzwoller (2002b), Thermodynamic constraints on seismic inversions, submitted to Geophys. J. Int.

Trampert, J., and J. Woodhouse (1995), Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, Geophys. J. Int., 122, 675-690.

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

Yang, X., I. Bondár, K. McLaughlin, and R. North (2001b), Source-Specific Station Corrections for regional phases at Fennoscandian stations, Pure appl. Geophys. 158, 35-57.

Table 1. Pn-P travel-time residual differences for more than 700 GT0-10 events with at least 5 Pn and 5 P phases

Pn	P	Pn-P standard deviation	Pn-P mean	Number of events
IASPEI91	IASPEI91	1.12	-0.23	730
CUB2				
CUB2				
IASPEI91	J362	1.21	0.18	765
CUB1				
CUB1	IASPEI91	1.08	0.38	765

Table 2. Mislocation statistics (w.r.t. IASPEI91) for models CUB1, CUB2, and J362, separated and combined, using all stations

	Pn (CUB1)	**Pn (CUB2)**	**P (J362)**	**PnP (CUB1+J362)**	**PnP (CUB1' and J362)**	**PnP (CUB2+J362)**
Number of events	526	526	793	1234	1234	1234
Model<GTX; IASP91<GTX or model=IASP91	117	105	70	222	210	214
Model<GTX; IASP91>GTX	57	53	74	130	133	129
Model>GTX; IASP91<GTX	48	60	28	107	119	116
Model>GTX; IASP91>GTX and model<IASP91	198	161	444	512	509	492
Model>GTX; IASP91>GT X and Model>IASP91	106	147	177	263	263	283
% of events improved	48%	41%	65%	52%	52%	50%
% of events deteriorated	29%	39%	26%	30%	31%	32%
% of events indecisive	22%	20%	9%	18%	17%	17%

relocation using CUB2 Pn and J362 P SSSCs.

Columns:
1. Pn (CUB1)- relocation using CUB1 Pn SSSCs
2. Pn (CUB2)- relocation using CUB2 Pn SSSCs
3. P (J362)- [illegible] number of GT events used in relocation
4. PnP (CUB1+J362)- relocation using CUB1 Pn and J362 P SSSCs
5. PnP (CUB1' and J362)- relocation using [illegible] events improved: mislocations within GTX with calibration, but beyond GTX without calibration, and J362 P SSSCs
6. PnP (CUB2+J362)- [illegible]

Rows:
1. # of [illegible]
2. # of events indecisive: mislocations both within GTX or equal to each other, with and without calibration
3. # of well-[illegible] without calibration
4. # of well-located events deteriorated: mislocation beyond GTX with calibration, but within GTX without calibration
5. # of other events improved: mislocations both beyond GTX with and without calibration, but mislocation with calibration smaller than that without calibration
6. # of other events deteriorated: mislocations both beyond GTX with and without calibration, but mislocation with calibration larger than that without calibration
7. % of events location improved with calibration compared to IASPEI91 (#3 and #5 above)

[illegible] IASPEI91 [illegible]

8. % of events location deteriorated with calibration compared to IASPEI91 (#4 and #6 above)
9. % of events location indecisive: mislocation within GTX or the same calibration (#2 above)

Table 3. Other relocation statistics (w.r.t. IASPEI91) for models CUB1, CUB2, and J362, separated and combined, using all stations

	Pn (CUB1)	**Pn (CUB2)**	**P (J362)**	**PnP (CUB1+J362)**	**PnP (CUB1'+J362)**	**PnP (CUB2+J362)**
Median ndef	51	51	100	84	84	84
Median gap (°)	76°	76°	68°	58°	58°	58°
Median sgap (°)	99°	99°	85°	76°	76°	76°
Number of events	526	526	793	1234	1234	1234
IASP91 median mislocation (km)	7.1	7.1	8.3	7.1	7.1	7.1
model median mislocation (km)	6.6	7.2	6.1	5.9	5.7	5.9
% relative mislocation change	7%	-1%	27%	17%	20%	17%
% of events ellipse area improved						
IASP91 median ellipse area (km^2)	448	448	464	394	394	394
model median ellipse area (km^2)	253	253	322	233	233	232
% relative						
IASP91 90% coverage	83%	83%	65%	75%	75%	75%
model 90% coverage	76%	70%	75%	76%	77%	75%
% of events OT improved	44%	41%	27%	37%	34%	33%
IASP91 median OT (sec)						
Model median OT (						
% of events OT error improved						
IASP91 median OT error (sec)						
model median OT error (sec)						
% of events sdobs improved	65%	61%	84%	67%	79%	79%
IASP91 median sdobs (sec)						
model median sdobs (sec)						

Columns:
1. Pn (CUB1)- relocation using CUB1 Pn SSSCs
2. Pn (CUB2)- relocation using CUB2 Pn SSSCs
3. P (J362)- relocation using J362 P SSSCs
4. PnP (CUB1+J362)- relocation using CUB1 Pn and J362 P SSSCs
5. PnP (CUB1'+J362)- relocation using CUB1 Pn and J362 P SSSCs
6. PnP (CUB2+J362)- relocation using CUB2 Pn and J362 P SSSCs.

Rows:
1. # of events:
2. IASP91 median mislocation (km): without calibration
3. model median mislocation (km): with calibration
4. % relative mislocation change: median mislocation changes between calibrated and uncalibrated w.r.t. to
5. % of events with ellipse area improved with calibration
6. IASP91 median ellipse area (km2): without calibration
7. model median ellipse area (km2): with calibration

with and without

8. % of relative ellipse area change: median ellipse area changes between calibrated and uncalibrated

9. IASP91 90% coverage: without calibration

10: model 90% coverage: with calibration

11: % of events OT improved: % of events origin time improved with calibration compared to IASPEI91

12. IASP91 median OT (sec): without calibration

13. model median OT (sec): with calibration

14. % of events OT error improved: % of events origin time errors improved with calibration compared to IASPEI91

15. IASP91 median OT error (sec): without calibration

16. model median OT

17. % of events sdobs improved: % of events standard deviation of errors improved with calibration compared to IASPEI91

18. IASP91 median sdobs (sec): without calibration

19. model median sdobs (sec): with calibration

Table 4. SSNB relocation improvements in bias and location due to calibrated travel times

	Percent events	CUB1	CUB2		+J362	CUB2 +J362	CUB1 +IASP	CUB2 +IASP	IASP +J362
SSNB centroids	Improved	51	50	62	58	56	49	45	48
	Deteriorated	30	33	28	25	28	35	40	35
	Indecisive	19	17	10	17	16	16	15	17
	U_{within}, C_{within} GT	19	17	10	17	16	16	13	17
	U_{beyond}, C_{within} GT	11	10	9	9	9	6	8	6
	U_{withn}, C_{beyond} GT	6	8	4	3	5	5	8	4
	U_{beyon}, C_{beyond} GT	64	65	77	71	70	73	71	73
	N_{events}	276	273	331	355	356	359	357	355
SSNB 10-station locations	Improved	52	52	57	57	56	47	45	48
	Deteriorated	34	35	34	33	35	34	37	41
	Indecisive	14	13	9	10	9	19	18	11
	U_{within}, C_{within} GT	14	13	9	10	9	10	9	11
	U_{beyond}, C_{within} GT	11	13	6	8	9	7	7	6
	U_{withn}, C_{beyond} GT	7	7	4	5	6	5	7	5
	U_{beyon}, C_{beyond} GT	68	67	81	77	76	78	77	78
	N_{events}	4856	4856	6836	7145	7142	7175	7158	7127
All-station Locations (seed)	Improved	56	48	58	48	49	46	41	45
	Deteriorated	27	36	28	33	34	35	44	35
	Indecisive	17	16	14	19	17	19	15	20
	U_{within}, C_{within} GT	17	16	14	19	17	19	15	20
	U_{beyond}, C_{within} GT	12	12	8	10	11	8	9	10

	U_{withn}, C_{beyond} GT	6	8	3	9	10	8	12	7
	U_{beyon}, C_{beyond} GT	65	64	75	62	62	65	64	63
	N_{events}	276	273	331	355	356	359	357	355

SCs.

Columns: 1. CUB1- relocation using CUB1 Pn SSSCs
2. CUB2- relocation using CUB2 Pn SSSCs
3. J362- relocation using J362 P SSSCs
4. CUB1+J362- relocation using CUB1 Pn and J362 P SSSCs
5. CUB2+J362- relocation using CUB2 Pn and J362 P SSSCs
6. CUB1+IASP- relocation using CUB1 Pn and IASPEI91 P SS
7. CUB2+J362- relocation using CUB2 Pn and J362 P SSSCs
8. IASP+J362- relocation using IASPEI91 Pn and J362 P SSSCs

Rows:

1. % of events location improved with calibration compared to IASPEI91 (#3 and #5 below)
2. % of events location deteriorated with calibration compared to IASPEI91 (#4 and #6 below)
3. % of events location indecisive: mislocation within GTX or the same with and without calibration (#2 below)
4. U within, C within GT: % of events with mislocations both within GTX or equal to each other, with and without calibration
5. U within, C beyond GT: % of events with mislocation within GTX with calibration, but beyond GTX without calibration
6. U within, C beyond GT: % of events with mislocation beyond GTX with calibration, but within GTX without calibration
7. U beyond, C beyond GT: % of events with mislocations both beyond GTX with and without calibration
9. N events: total number of GT events used in relocation

Table 5. SSNB relocation comparison of median bias, mislocation, area of error ellipse, and actual coverage of the 90% error ellipse using uncalibrated (U) and calibrated (C) travel-

	Metrics	CUB1		CUB2		J362		CUB1+J362		CUB2+J362							
		U	C	U	C	U	C	U	C	U	C	U	C	U	C	U	C
SSNB	Bias [km]	8.4	7.2	8.2	7.0	9.2	7.9	7.7	6.6	7.7	6.4	7.8	7.3	7.8	7.5	7.7	7.1
	Mislocation [km]	8.8	7.6	8.8	7.6	11.1	9.9	9.9	8.7	9.9	8.6	9.9	9.1	9.9	9.2	9.9	9.5
	Error ellipse area [km^2]	1522	745	1522	730	1953	1133	1815	971	1815	959	1816	1420	1816	1461	1814	1300
	Coverage [%]	99	96	99	93	92	85	97	92	97	92	96	95	97	95	97	93
All-stations (seed)	Mislocation [km]	7.9	6.8	7.8	7.3	8.2	7.1	6.6	5.7	6.6	5.9	6.6	6.4	6.6	6.9	6.6	6.0
	Error ellipse area [km^2]	518	293	529	288	411	274	250	171	251	167	254	206	254	200	253	203
	Coverage [%]	80	79	81	71	63	64	71	69	71	68	70	64	71	60	71	74

Columns: U: uncalibrated; C: calibrated
1. CUB1- relocation using CUB1 Pn SSSCs
2. CUB2- relocation using CUB2 Pn SSSCs

3. J362- relocation using J362 P SSSCs
4. CUB1+J362- relocation using CUB1 Pn and J
5. CUB2+J362- relocation using CUB2 Pn and J362 P SSSCs
6. CUB1+IASP- relocation using CUB1 Pn and IASPEI91 P SSSCs.
7. CUB2+J362- relocation using CUB2 Pn and J362 P SSSCs
8. IASP+J362- relocation using IASPEI91 Pn and J362 P SSSCs

Rows: 1. Bias (km): location bias
2. mislocation (km): mislocation
3. error ellipse area (km2): error ellipse area
4. coverage (%): % of events inside their error ellipses

Table 6. Mislocation statistics (w.r.t. IASPEI91) for direct comparions of models C vs. CUB2, CUB1'+J362 vs. CUB1+J362, and CUB1+J362 vs. CUB2+J362 using all stations

	PnP (CUB1+J362 CUB1'+J362)	**Pn (CUB1 vs. CUB2)**	**PnP (CUB1+J362 vs. CUB2+J362)**
Number of events	1234	526	1234
Model1<GTX; model2<GTX or model1=model2	607	117	441
Model1<GTX; model2>GTX	41	57	86
Model1>GTX; model2<GTX	50	41	85
Model1>GTX; model2>GTX and model1<model2	291	205	360
Model1>GTX; model2>GTX and Model1>model2	245	106	262
% of events improved	27%	50%	36%
% of events deteriorated	24%	28%	28%
% of events indecisive	49%	22%	36%

events: total number of GT events used in relocation
Columns: 1. PnP (CUB1+J362 vs. CUB1'+J362)- relocation using joint CUB1 Pn and J362 P SSSCs, without (CUB1) and with
2. Pn (CUB1 vs. CUB2)- relocation using CUB1 and CUB2 Pn, respectively (model1 and model2, respectively).
3. PnP (CUB1+J362 vs. CUB2+J362)- relocation using CUB1 or CUB2 Pn joint with J362 P SSSCs

Rows: 1. # of
2. # of events indecisive: mislocations both within GTX or equal to each other, with model1 and model2.
3. # of well-
model2
4. # of well-located events deteriorated: misl
model2

362 P SSSCs

5. # of other events improved: mislocations both beyond GTX with both model1 and model2, but mislocation with model1 smaller than that with model2

6. # of other events deteriorated: mislocations both beyond GTX with both model1 and model2, but mislocation with model1 larger than that with model2

7. % of events location improved with model1 compared to model2 (#3 and #5 above)

8. % of events location deteriorated with model1 compared to model2 (#4 and #6

9. % of events location indecisive: mislocation within GTX or the same with both model1 and model2 (#2 above)

Table 7. Relocation statistics for the Mediterranean region (w.r.t. IASPEI91) for models CUB1, CUB2, and J362, separated and combined, using all stations

	Pn (CUB1)	**Pn (CUB2)**	**Pn (CUB1vs. CUB2)**	**P (J362)**	**PnP (CUB1+J362)**	**PnP (CUB2+J362)**	**PnP (CUB1+J362 vs. CUB2+J362)**
Number of events	415	415	508	179	508	508	508
Model1<GTX; model2<GTX or model1=model2	110	100	131	17	132	122	131
Model1<GTX; model2>GTX	42	47	50	9	39	56	50
Model1>GTX; model2<GTX	46	56	57	9	63	73	57
Model1>GTX; model2>GTX and model1<model2	131	99	157	63	137	113	157
Model1>GTX; model2>GTX and Model1=model2	41	63	113	81	137	144	113
Model1 median mislocation (km)	6.1	6.1	6.5	6.8	6.4	6.4	6.4
% of events improved	41%	35%	41%	40%	35%	33%	41%
% of events deteriorated	32%	41%	33%	50%	39%	43%	33%
% of events indecisive	27%	24%	26%	10%	26%	24%	26%
Model1 median mislocation							
Model2							

relocation using J362 P SSCs, compared to IASPEI91 (model1 and model2,

Columns: 1. Pn (CUB1)- relocation using CUB1 Pn SSSCs, compared to IASPEI91 (model1 and model2, respectively).

2. Pn (CUB2)- relocation using CUB2 Pn SSSCs, compared to IASPEI91 (model1 and model2, respectively).

3. Pn (CUB1 vs. CUB2)-

respectively).

4. P (J362)-

respectively).

5. PnP (CUB1+J362)-

(model1 and model2, respectiv

6. PnP (CUB1'+J362)- relocation using CUB1 Pn (with a 0.75 sec bulk correction) and J362 P SSSCs

results; total number of GT events used in relocation

7. PnP (CUB2+J362)-

(model1 and model2, respectively)

location within GTX with model1, but beyond GTX with

8. PnP (CUB1+J362 vs CUB2+J362)- relocation using CUB1 and CUB2 Pn joint with J362 P SSSCs

Rows: 1. # of

2. # of events indecisive: mislocations both within GTX or equal to each other, with both model1 and model2

3. # of well-

model2

above)

4. # of well-
with model2

5. # of other events improved: mislocations both beyond GTX with both model1 and model2, but mislocation with model1 smaller than that with model2

6. # of other events deteriorated: mislocations both beyond GTX with both model1 and model2, but mislocation with model1 larger than that with model2

7. % of events location improved with model1 compared to model2 (#3 and #5 above)

8. % of events location deteriorated with model1 compared to model2 (#4 and #6 above)

9. % of events location indecisive: mislocation within GTX or the same with both model1 and model2 (#2 above)

10. Model1 median mislocation (km).

11. Model2 median mislocation (km).

located events deteriorated: mislocation beyond GTX with model1, but within GTX

Figure captions

Figure 1. An example of SSSCs for station ABKT (Alibek, Turkmenistan), calculated from 3D global models using raytracying. (a) Regional Pn SSSCs from CUB1. (b) Teleseismic P SSSCs from J362. Triangles show the station locations.

Figure 2. Modeling errors for regional Pn and teleseismic P phases. (dashed line) IASPEI91. (solid line) CUB. (dot-dashed line) J362.

Figure 3. (a) 526 selected GT0-GT10 events for use in relocation testing of regional Pn calibration. They are 10 GT0 (square), 36 GT1 (circle), 1 GT2 (triangle), 450 GT5 (invert triangle), and 19 GT10 (diamond).
arrivals) with CUB Pn SSSCs for this data set. Triangles show station locations.

Figure 4. (a) 793 selected GT0-GT10 events for use in relocation testing of teleseismic P calibration. They are 17 GT0 (square), 316 GT1 (circle), 10 GT2 (triangle), 279 GT5 (invert triangle), and 171 GT10 (diamond). (bottom) 2821 stations (108,172 arrivals) with J362 P SSSCs for this data set. Triangles show station locations.

Figure 5. SSNB illustration of sgap tree for station "A" of the fictitious network shown in the inset. The sgap tree is a hierarchical binary tree where each subsequent station is selected in a way that it splits the remaining azimuthal as evenly as possible. The selection order of stations is shown on the right axis, and the numbers at the left and right of the nodes indicate the remaining azimuthal gaps. Cutting the sgap tree at any level of the hierarchy (dashed line) provides the sub-network with the most uniform azimuthal coverage and the minimal secondary azimuthal gap (in this example 130º, provided by station B for a 10-station sub-network). Different root stations may generate different sgap trees.

Figure 6. Coverage parameter E. E should be χ^2 distribution with two-degrees of freedom. E=1.0 corresponds to the 90% coverage. When E>1, events are not covered. In this diagram the event is not covered.

Figure 7. Cumulative mislocations with (red) and without (green) SSSCs. (a) Relocation using CUB1 Pn SSSCs compared to IASPEI91. (b) Relocation using J362 P SSSCs compared to IASPEI91. (c) Relocation using CUB1 Pn and J362 P SSSCs compared to IASPEI91.

Figure 8. Coverage parameter vs. ndef using joint CUB1 Pn and J362 P SSSCs (triangle) and using IASPEI91 (diamond).

Figure 9. SSNB example for the 1997/02/14 00:22:37 earthquake in Pakistan (top), and the 1972/08/20 03:00:00 PNE Region-3 (bottom). Open triangles represent the relocations with the 20 most characteristic SSNB 10-station sub-networks with calibrated (CUB1, red) and uncalibrated (IASPEI, blue) regional Pn travel-times. The location bias is estimated by the vector (solid line) connecting the GT location (star) with the centroid of the SSNB realizations (solid triangle). The uncertainty in the location bias estimate (ellipses) is derived from the individual SSNB locations.

Figure 10. (top) Mean mislocations using all stations (seed) with (red inverted triangles) and without (blue triangles) calibration. (middle) Mean bias from SSNB locations with (red inverted triangles) and without (blue triangles) calibration. (bottom) Mean improvement in bias (red circles) from SSNB locations and mean improvement in location (blue squares) from all-station (seed) locations due to calibrated regional and teleseismic (CUB1+J362) travel-times for event clusters sampled by more than 3 SSNB seed events. Event clusters are sorted by GT category. The thick green line indicates the GT accuracy of the reference events in the cluster.

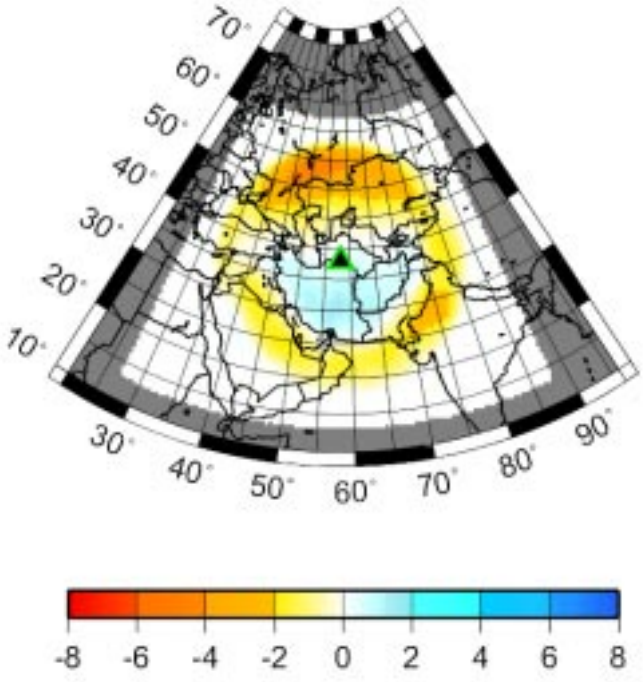
Figure 11. CUB1 vs. CUB2 Pn SSSCs used in relocating the 526 GT events. (a) CUB1 Pn SSSCs. (b) CUB2 Pn SSSCs. Green shows faster travel time and red shows slower travel time.

Figure 12. Mislocation differences between CUB1 and CUB2 in km. Green shows CUB1 better, red shows CUB2 better, and blue shows indecisive. Using CUB1 there are 262 events improved (mostly in the Mediterranean), 147 events deteriorated, and 117 events within GT uncertainty (Table 6).

Figure 13. Median gap (orange), sgap (green), area of the 90% coverage ellipse, coverage parameter, misfit and location bias plotted as a function of number of phases used in the SSNB locations with calibrated (red) and uncalibrated (blue) travel-times.

Figure 14. Cumulative histograms of the coverage parameter with all-station (left) and the SSNB 10-station sparse networks (right) with calibrated (red) and uncalibrated (blue) regional and teleseismic travel-times. The theoretical χ^2 distribution (assuming Gaussian and independent errors) of the coverage parameter is shown as a black line. The arrows indicate the actual coverage of the nominal 90% coverage ellipse.

(a)


70°
60°
50°
40°
30°
20°
10°
30°
40°
50°
60°
70°
80°
90°
-8 -6 -4 -2 0 2 4 6 8

(b)

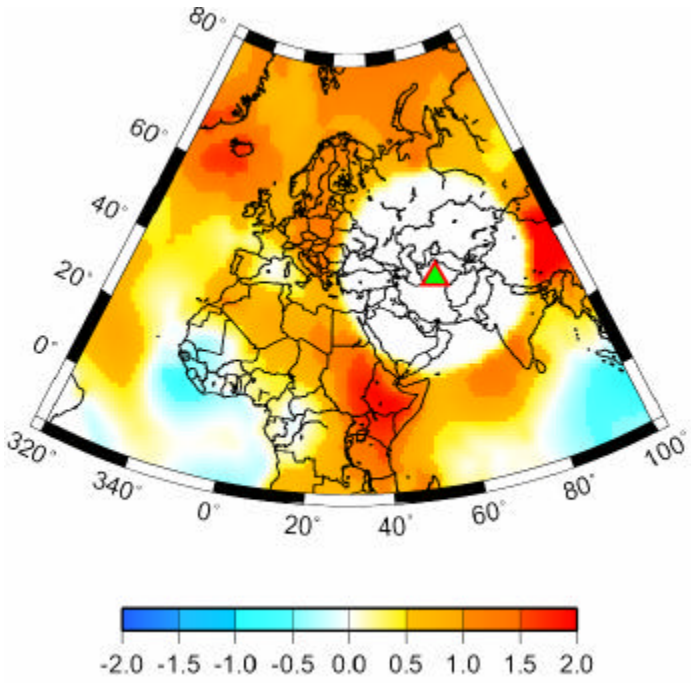

80°
60°
40°
20°
0°
320°
340°
0°
20°
40°
60°
80°
100°
-2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0

Figure 1.

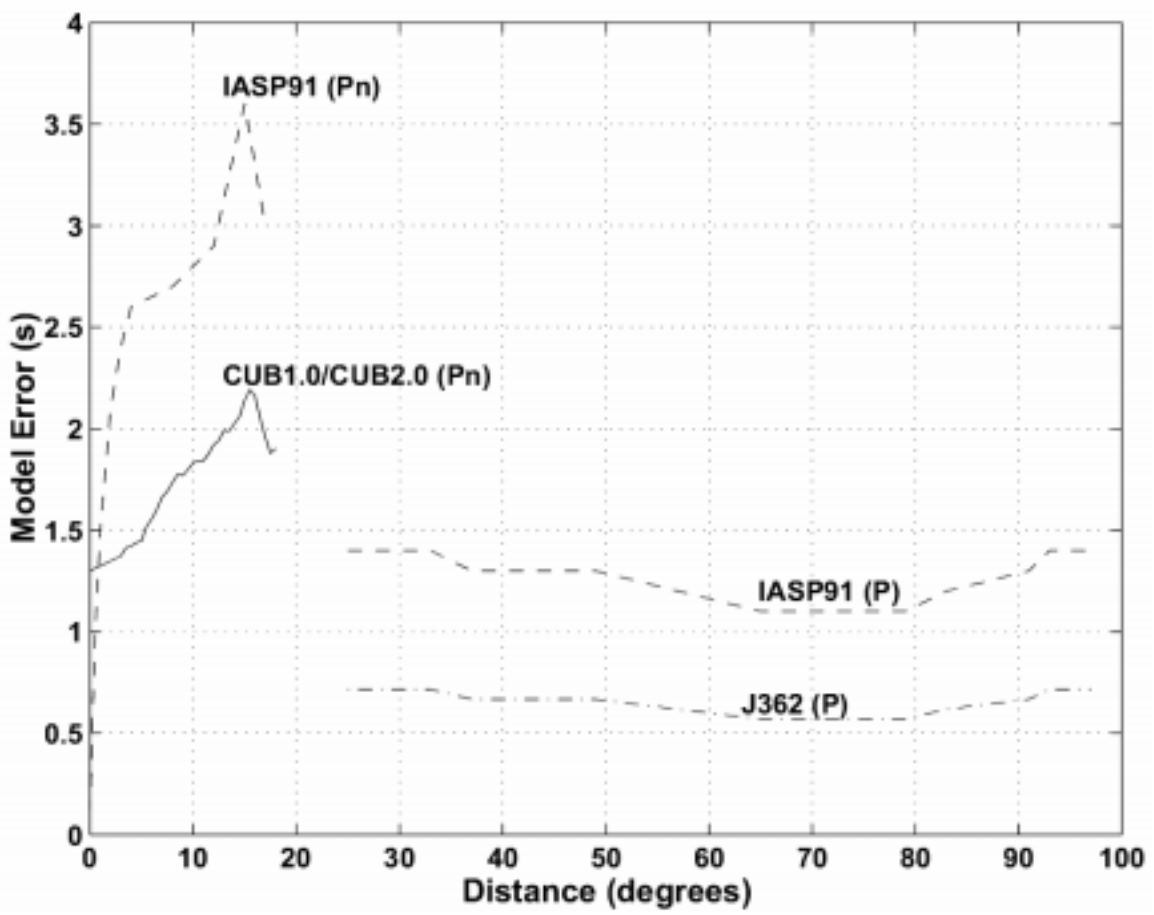
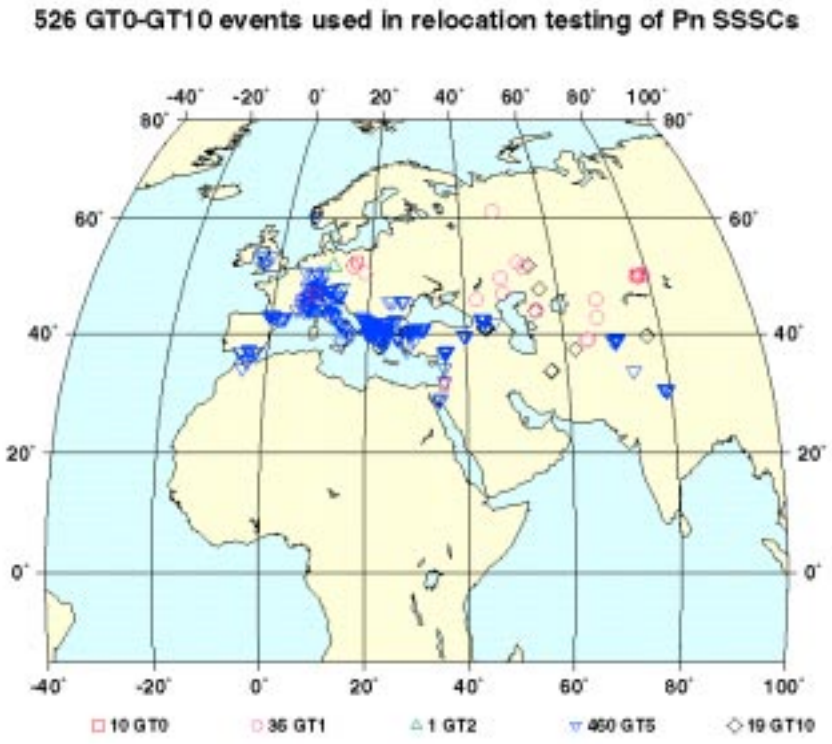

IASP91 (Pn)
CUB1.0/CUB2.0 (Pn)
IASP91 (P)
J362 (P)
Model Error (s)
Distance (degrees)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
10
20
30
40
50
60
70
80
90
100

Figure 2.

(a)


526 GT0-GT10 events used in relocation testing of Pn SSSCs
-40° -20° 0° 20° 40° 60° 80° 100°
80° 60° 40° 20° 0°
10 GT0
35 GT1
1 GT2
450 GT5
19 GT10

(b)

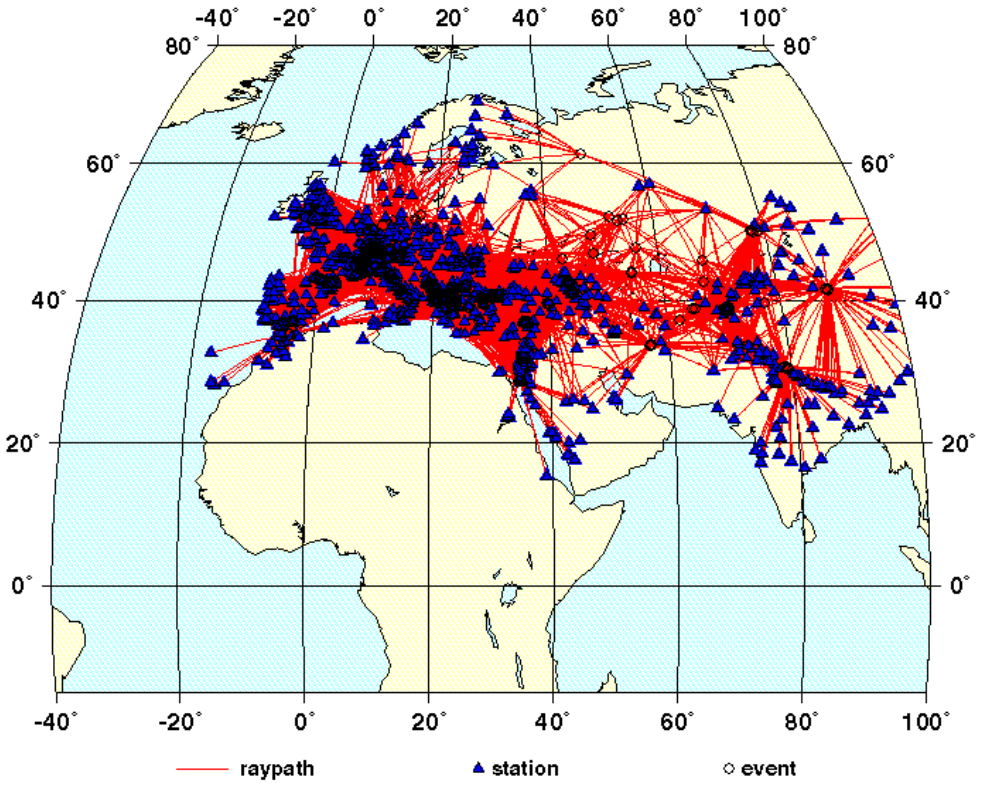
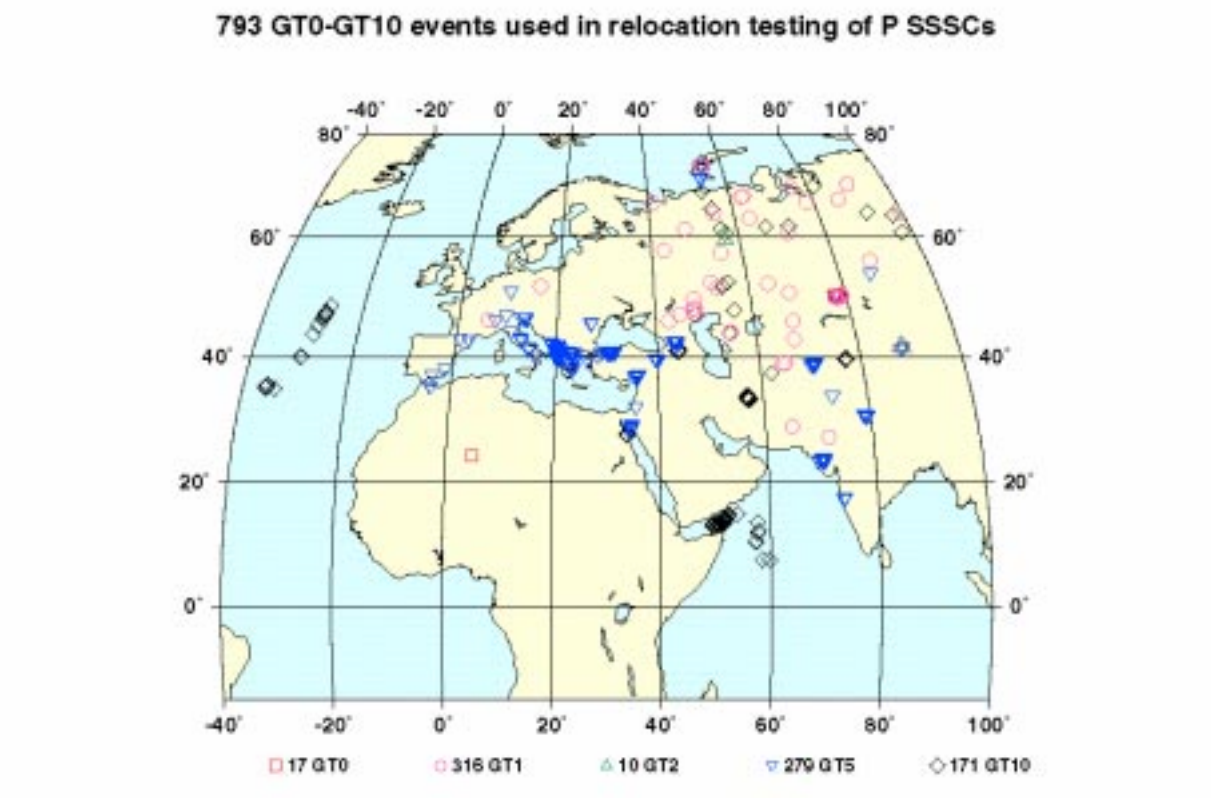

-40° -20° 0° 20° 40° 60° 80° 100°
80° 60° 40° 20° 0°
raypath
station
event

Figure 3.
(a)


793 GT0-GT10 events used in relocation testing of P SSSCs
17 GT0
316 GT1
10 GT2
279 GT5
171 GT10

(b)

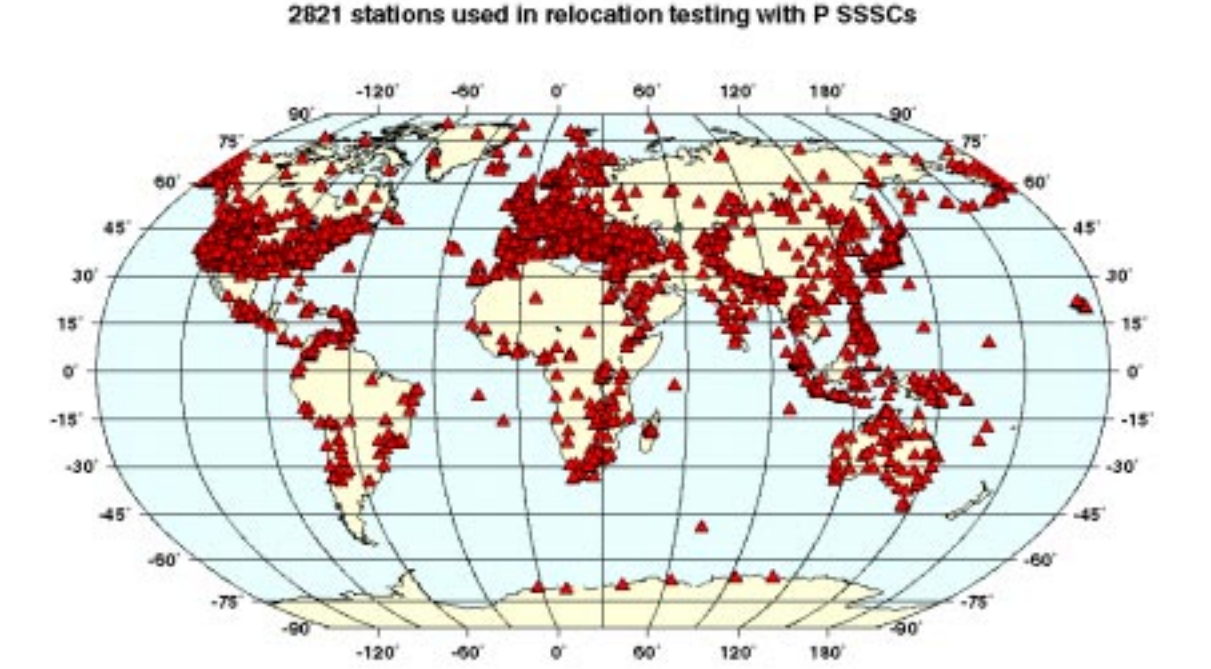

2821 stations used in relocation testing with P SSSCs

Figure 4.

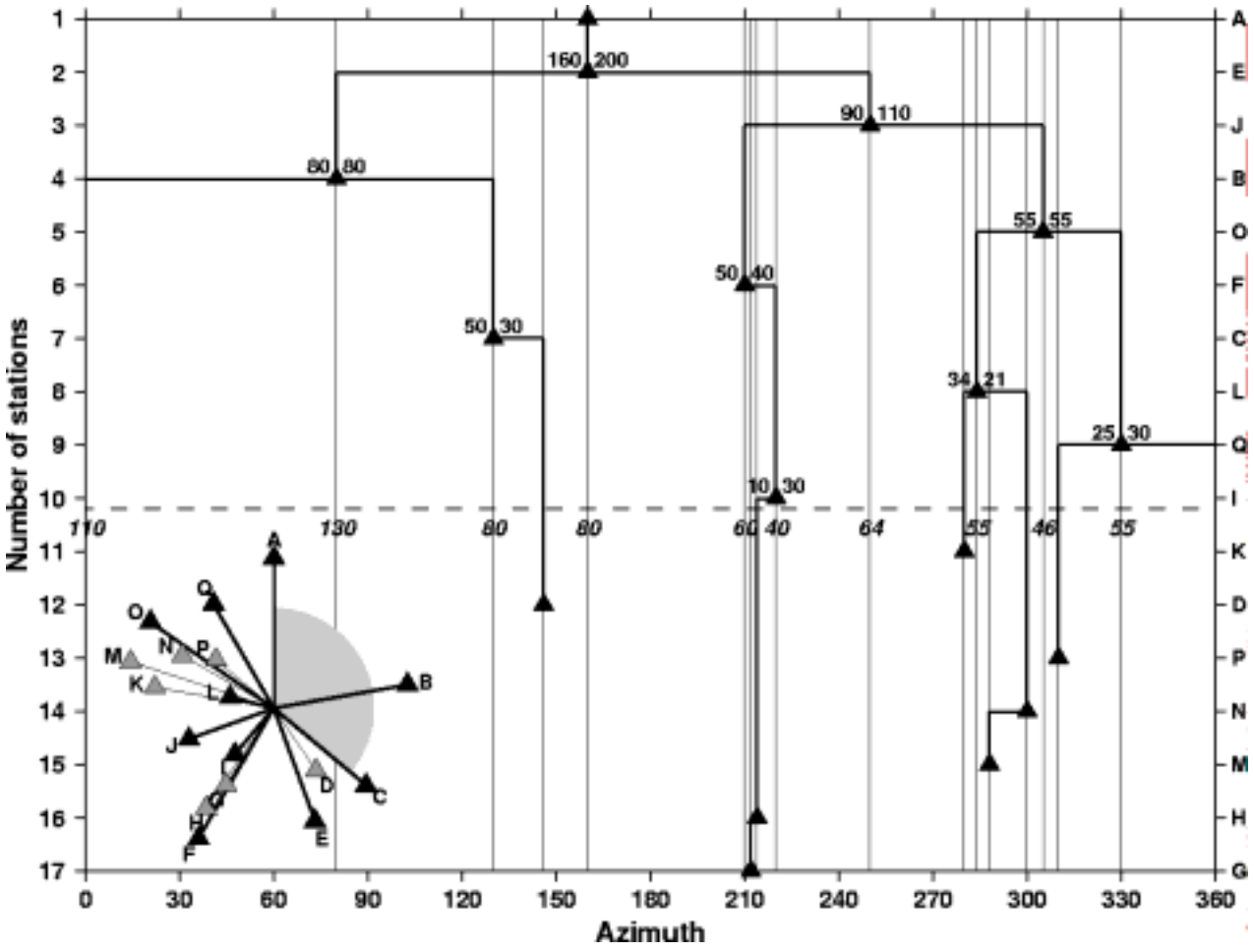


Figure 5.

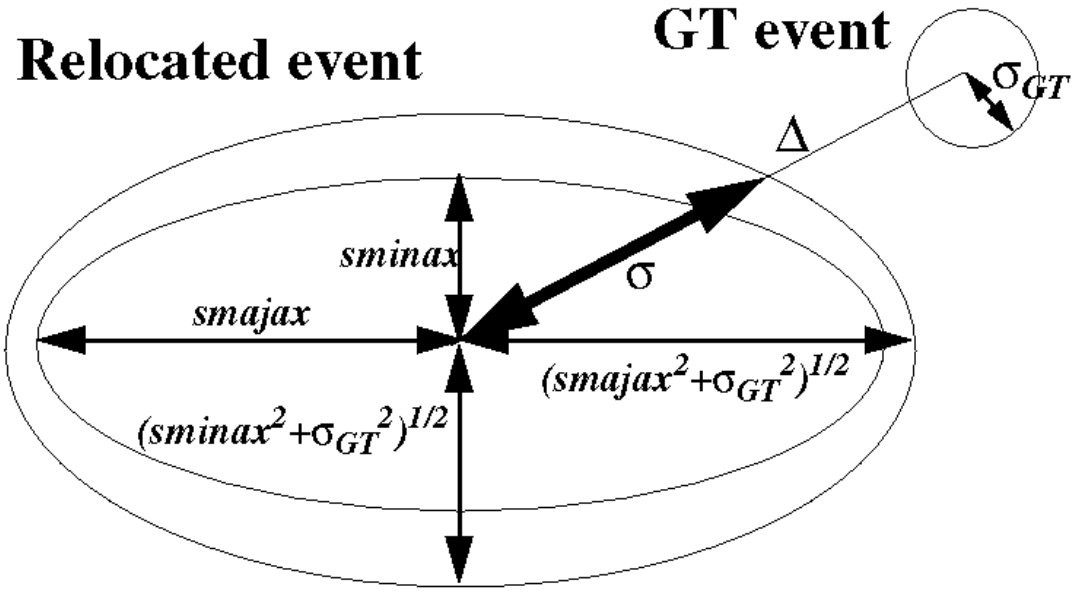
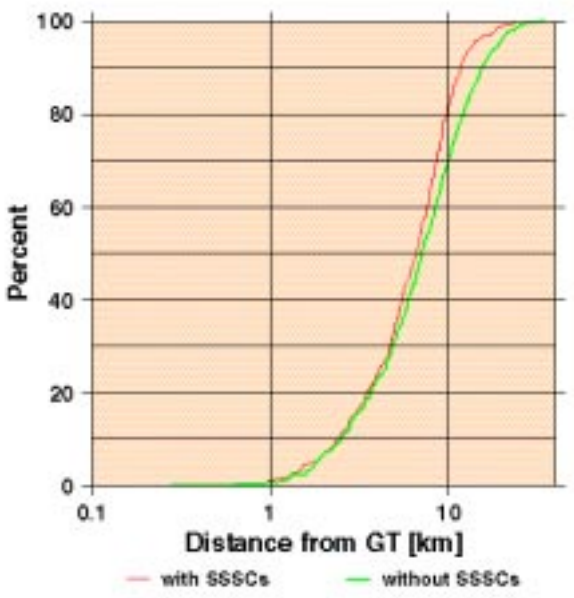
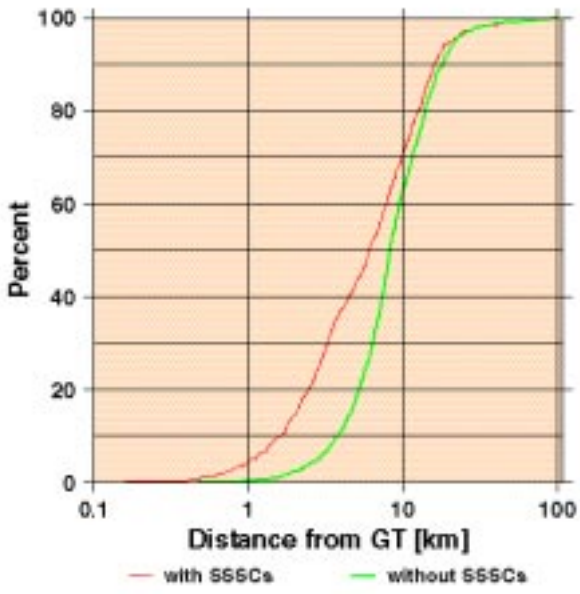

Relocated event
GT event
σ_{GT}
Δ
sminax
σ
smajax
$(smajax^2+\sigma_{GT}^2)^{1/2}$
$(sminax^2+\sigma_{GT}^2)^{1/2}$

Figure 6.

(a)


100
80
60
40
20
0
Percent
0.1
1
10
Distance from GT [km]
with SSSCs
without SSSCs

(b)


100
80
60
40
20
0
Percent
0.1
1
10
100
Distance from GT [km]
with SSSCs
without SSSCs

(c)

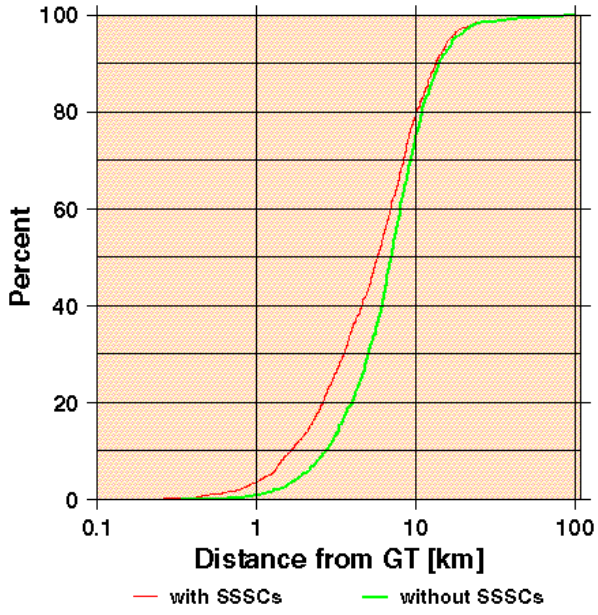

100
80
60
40
20
0
Percent
0.1
1
10
100
Distance from GT [km]
with SSSCs
without SSSCs

Figure 7.

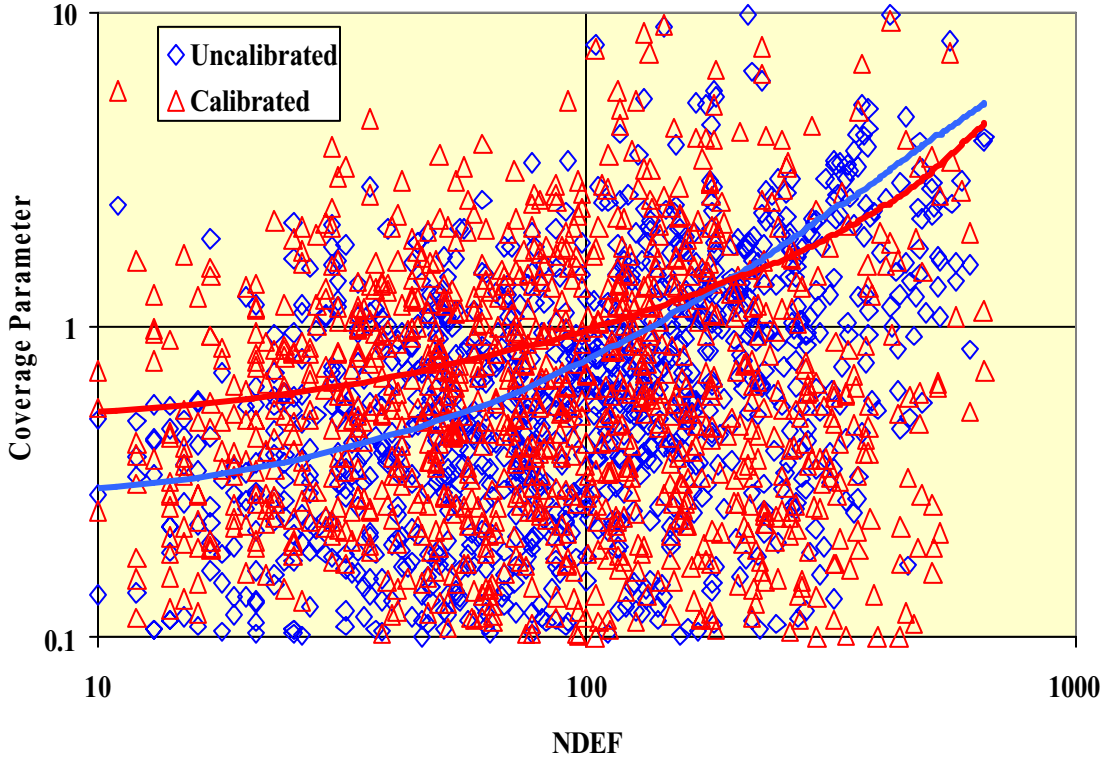


Figure 8.

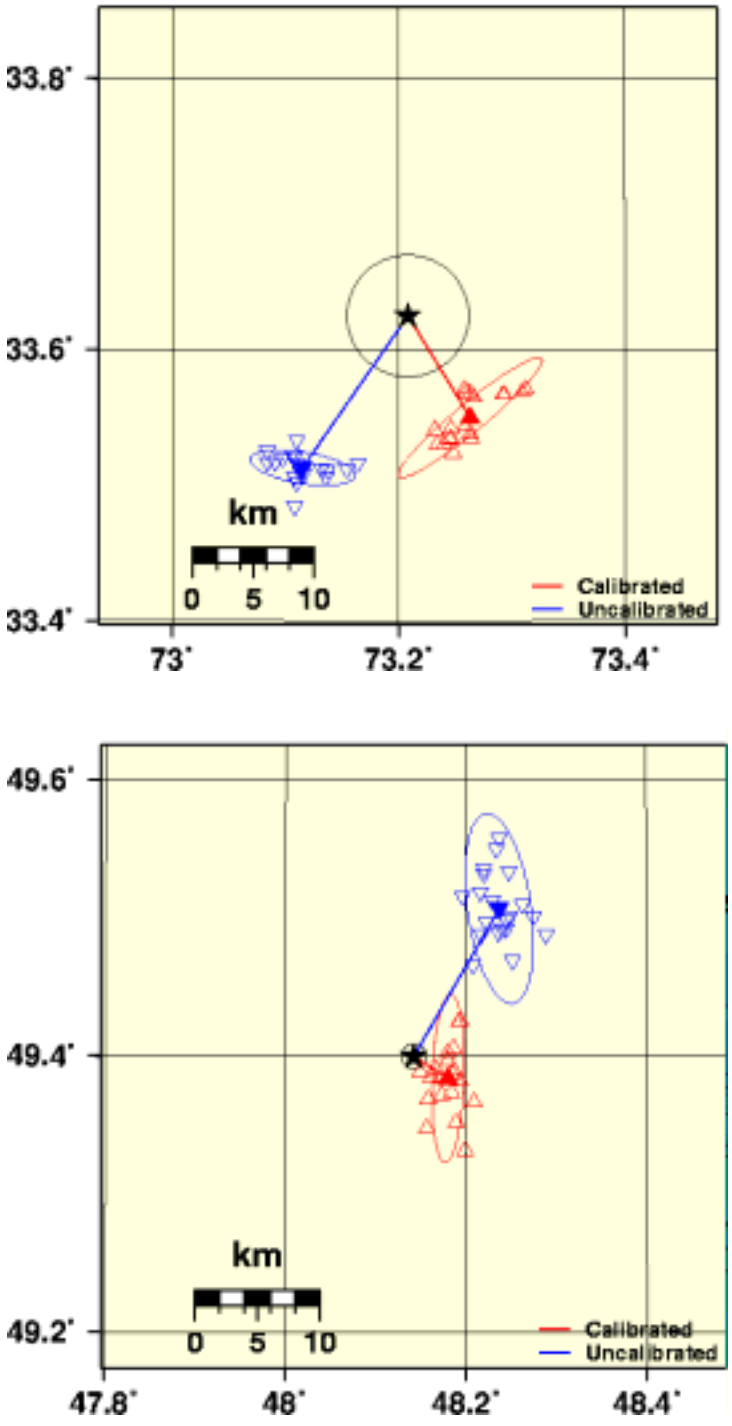
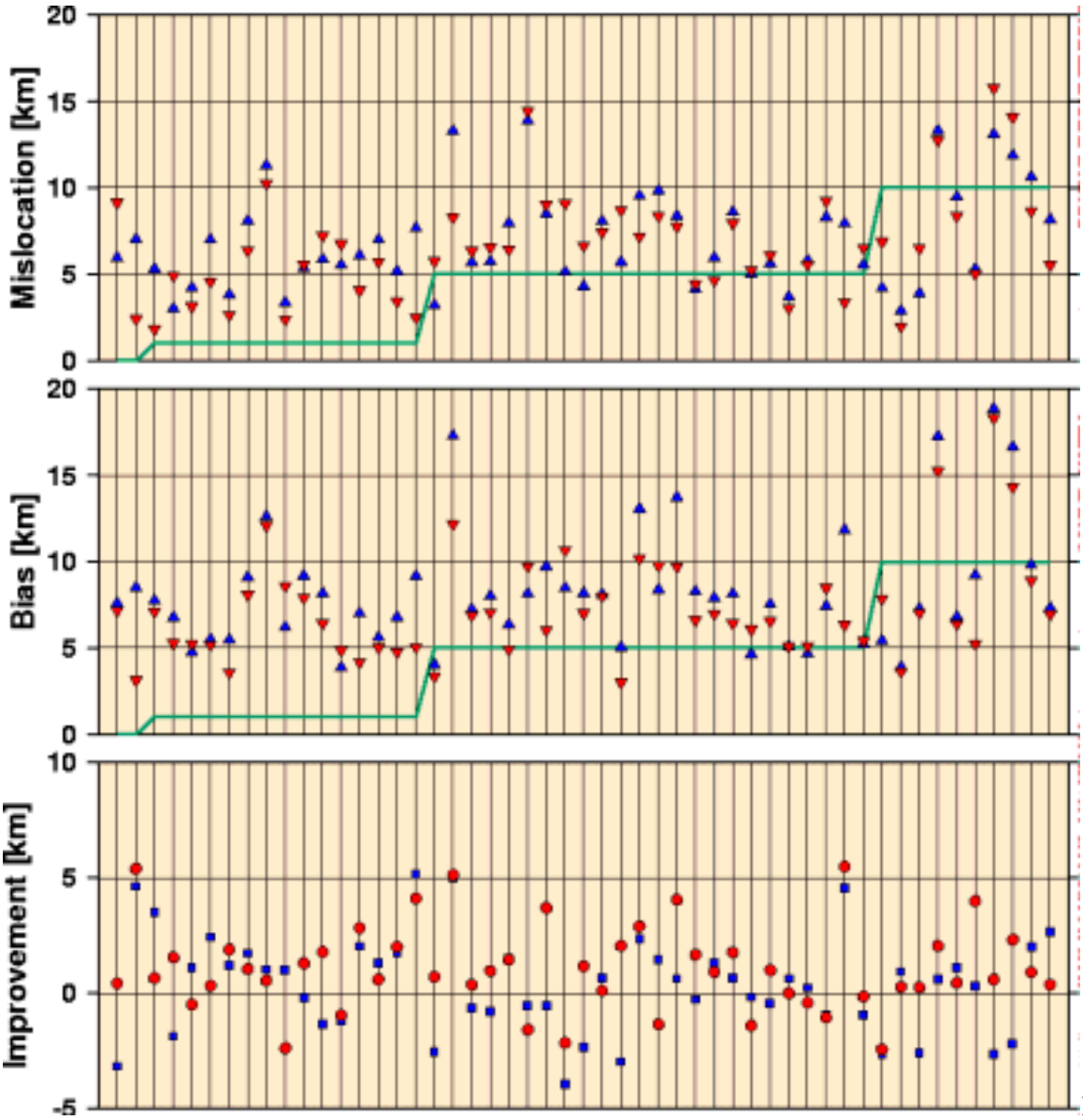

33.8˚
33.6˚
33.4˚
73˚
73.2˚
73.4˚
km
0 5 10
Calibrated
Uncalibrated
49.6˚
49.4˚
49.2˚
47.8˚
48˚
48.2˚
48.4˚
km
0 5 10
Calibrated
Uncalibrated

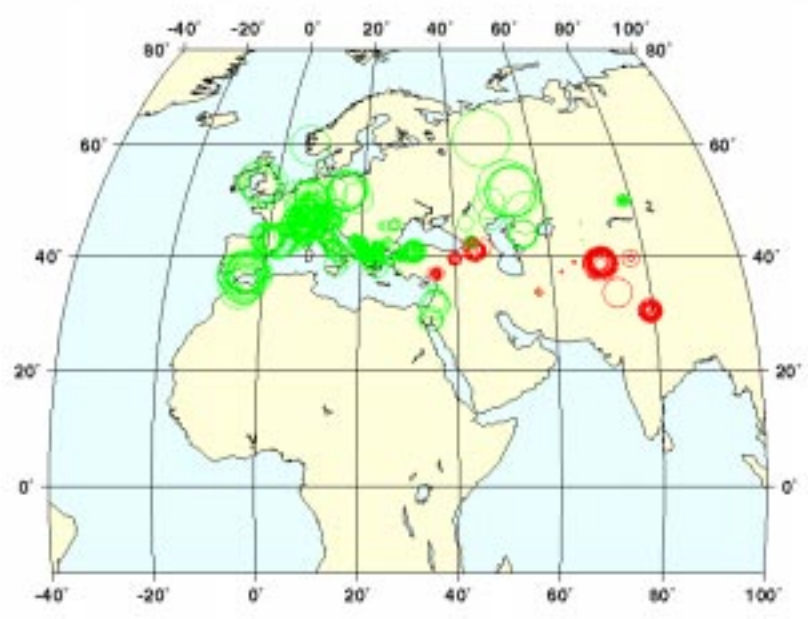
Figure 9.



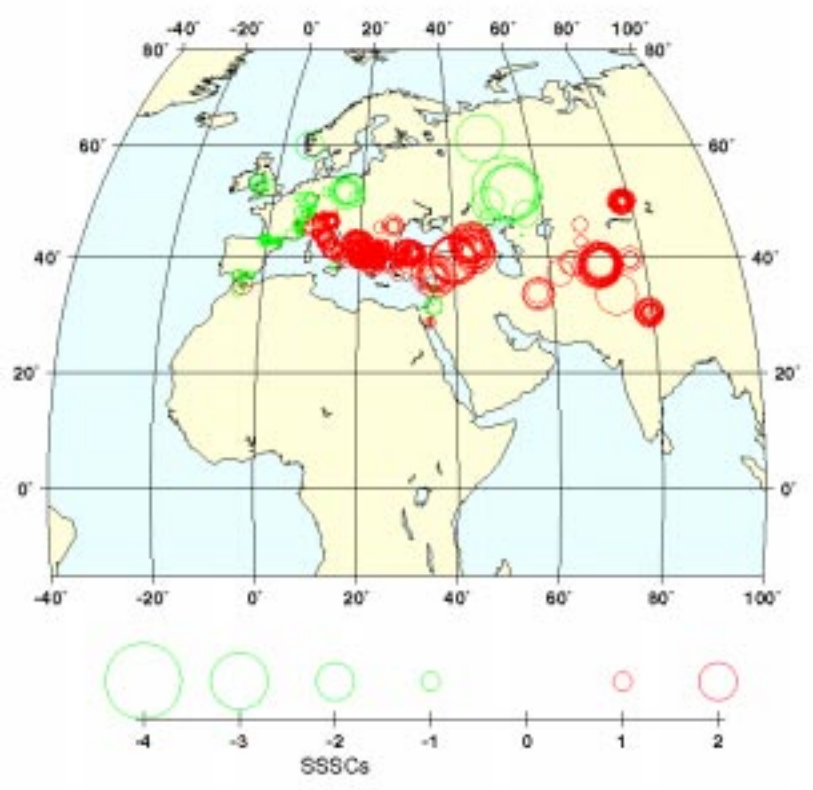
10.

Figure

1(a)


-40°
-20°
0°
20°
40°
60°
80°
100°
80°
60°
40°
20°
0°

(b)


-40°
-20°
0°
20°
40°
60°
80°
100°
80°
60°
40°
20°
0°
-4
-3
-2
-1
0
1
2
SSSCs

Figure

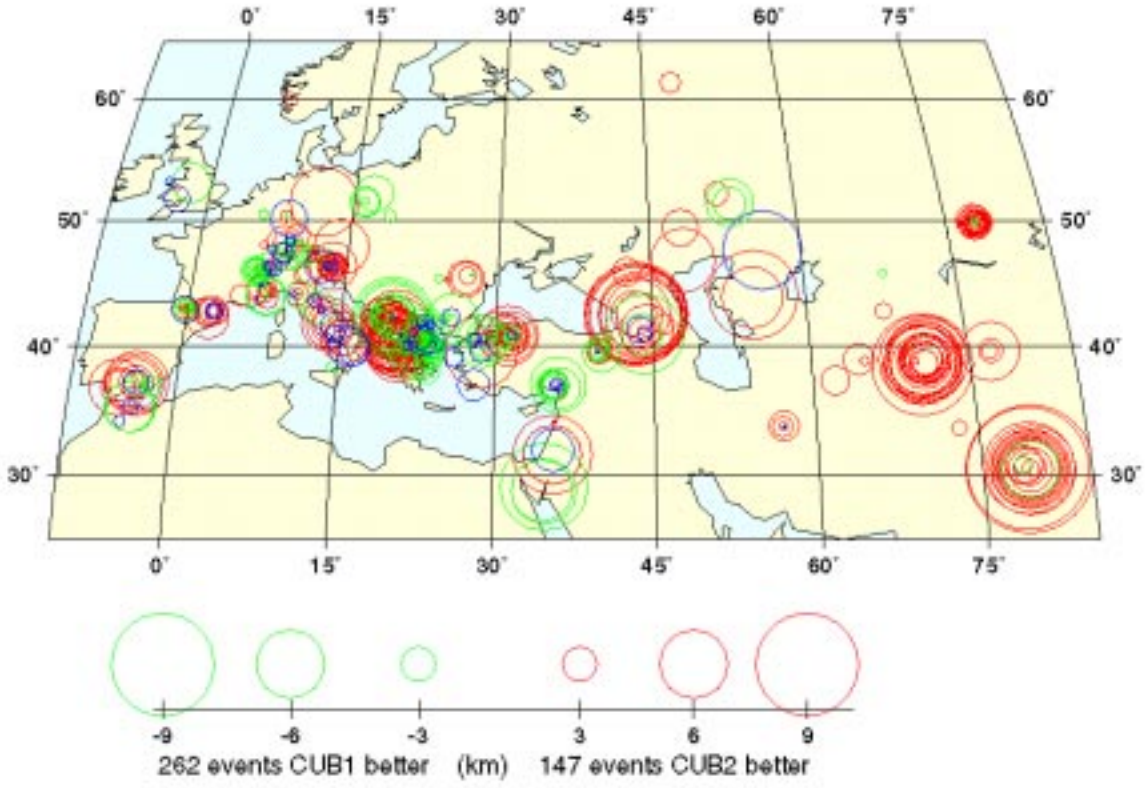
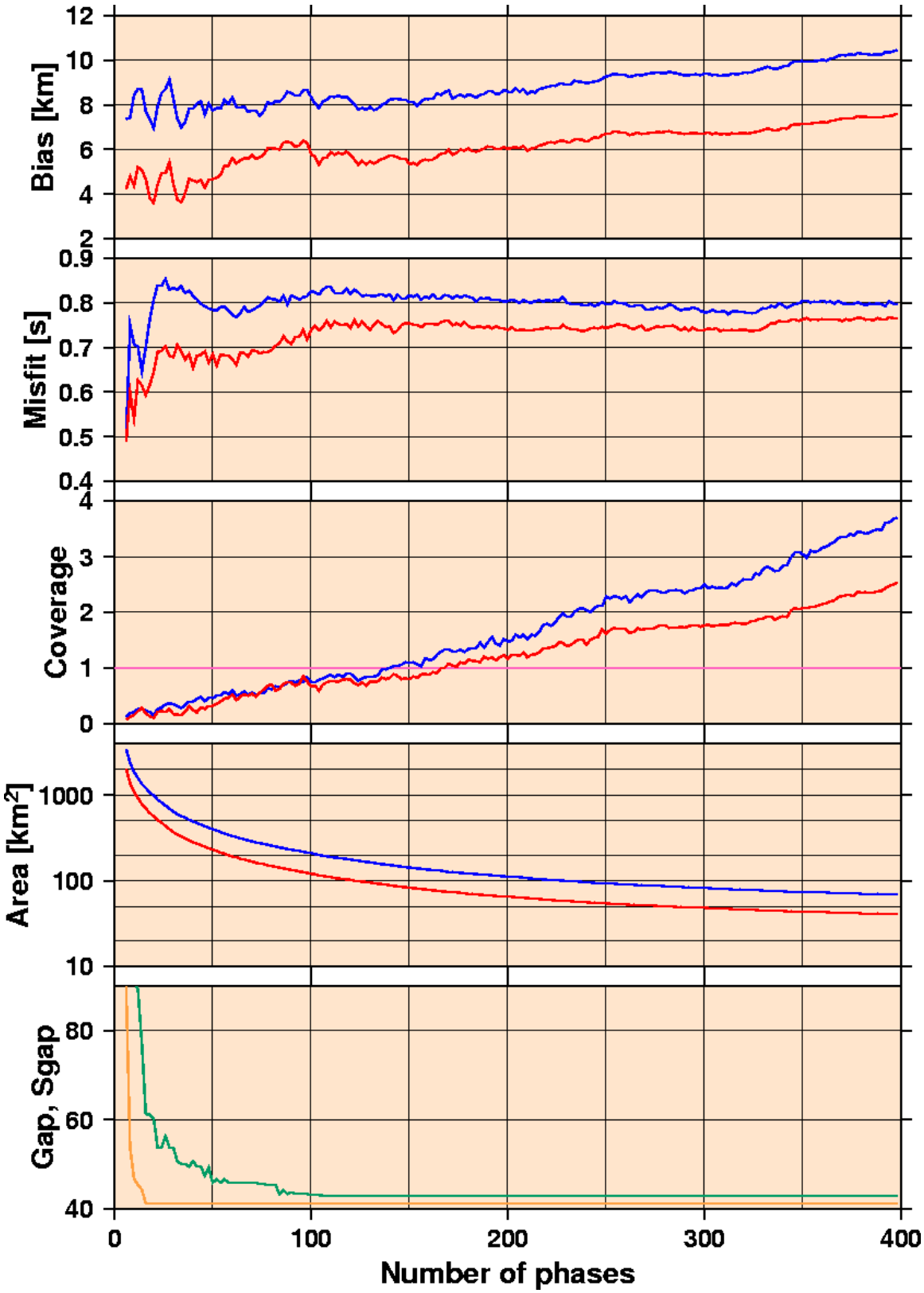

0°
15°
30°
45°
60°
75°
60°
50°
40°
30°
-9
-6
-3
3
6
9
262 events CUB1 better
(km)
147 events CUB2 better

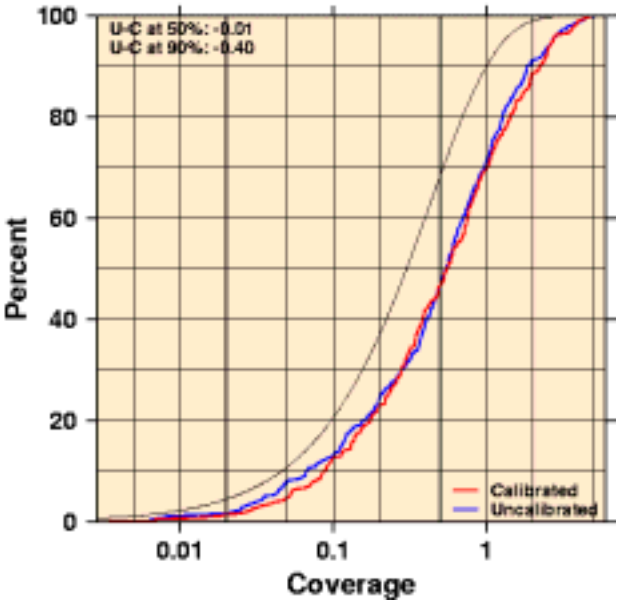
Figure 12.



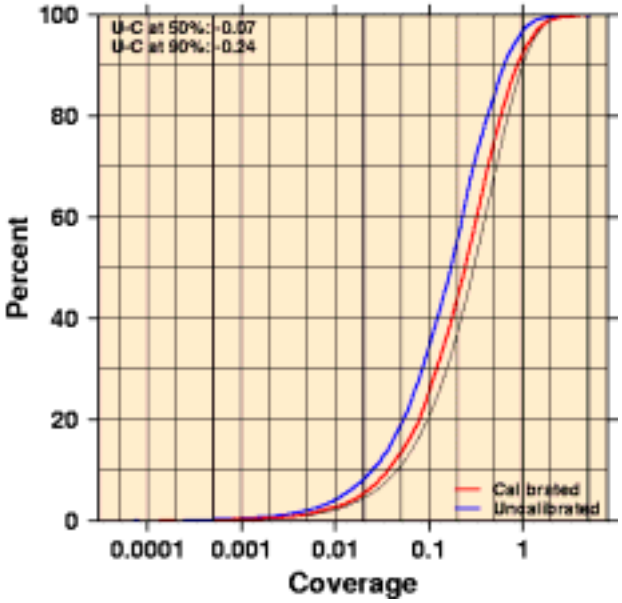
Figure

13.

(a)


U-C at 50%: -0.01
U-C at 90%: -0.40
Percent
Coverage
0.01
0.1
1
Calibrated
Uncalibrated

(b)


U-C at 50%: -0.07
U-C at 90%: -0.24
Percent
Coverage
0.0001
0.001
0.01
0.1
1
Calibrated
Uncalibrated

14.

Figure