A GROUND TRUTH DATABASE FOR REGIONAL SEISMIC RESEARCH

Ivan Henson, Indra Gupta, and Robert Wagner, Multimax, Inc., Largo, Maryland

Sponsored by Defense Threat Reduction Agency
Contract No. DSWA01-98-C-0156

ABSTRACT

The main objective of this project is the formation of databases providing ground truth for regional seismic research conforming to the latest Center for Monitoring Research (CMR) schema. These ground truth data sets will be incorporated into the contractor's existing Ground Truth Database (GTDB), and the "event directories" will be made available for use at CMR. The effort focuses on events in the CMR Calibration Event Bulletin (CEB) from China, the former Soviet Union, and North America. Events are carefully re-analyzed by using *geotool* and following procedures established at the CMR for the Reviewed Event Bulletin (REB). Revisions consist mostly of re-timing phase arrivals and, whenever possible, adding missed phases. Waveform data from available Incorporated Research Institutions for Seismology (IRIS) stations are retrieved and analyzed; the criteria for selecting the IRIS stations are primarily to collect data within 25 degrees distance, then to reduce any azimuthal gap for distances beyond 25 degrees (up to about 50 or 60 degrees), and to obtain data at PKP distances. A modified version of *LocSAT*, called *Locate*, which uses a graphical interface and allows interactive processing to generate location results, is used to locate the CEB events. Two sets of location results per event are derived: the first by using only the revised International Data Centre (IDC) arrivals, and the second by using, both the IRIS and revised IDC data.

All (70) events in China from the CMR Calibration Event Bulletin (CEB) have been re-analyzed by using waveform data from over 140 IDC stations contributing over 4000 arrivals. By analyzing data from over 130 IRIS stations, nearly 3800 waveforms were retained and nearly 2300 arrivals added. The most important contribution of the IRIS data is the addition of a large number of arrivals from regional phases. As an example, for epicentral distances less than 17 degrees (about 1900 km), the regional phases (Pn, Pg, Sn, and Lg) in the three datasets consisting of combined IRIS and revised IDC, only the revised IDC, and the original CEB arrivals, total 1362, 225, and 140, respectively. Analysis of all CEB events in China has been completed and the results delivered to the CMR. Similar analysis of CEB events in the former Soviet Union, based on the use of both IDC and IRIS stations, is in progress.

For the Chinese database of 70 events, a comparison of the CEB locations with those derived by using the revised IDC arrivals and the combination of IRIS and revised IDC data showed generally small differences (average less than 9 km in each case). However, a similar comparison of source depths showed the CEB estimates to be significantly larger than each of the other two estimates, suggesting that the CEB source depths have probably been overestimated. This is supported by the remarkably good agreement observed between depth estimates based on the combined use of IRIS and revised IDC data and those available from the International Seismological Centre (ISC) for a large number of common events.

Key Words: database, ground truth, Calibration Event Bulletin, CMR, CTBT, regional seismic research

OBJECTIVE

Monitoring the Comprehensive Nuclear-Test-Ban Treaty (CTBT) requires the ability to detect, locate and identify seismic events effectively. Once a suspicious event has been detected, its location must be accurately determined as a precursor to event identification and potential on-site inspections. The Calibration Event Bulletin (CEB) database was initiated at the Center for Monitoring Research (CMR) in June 1996 in response to recommendations by the Group of Scientific Experts (Bondar, 1998a). The main purpose of the CEB database is to collect data routinely for the calibration of the International Monitoring System (IMS) network so that locations can be improved. The selection of CEB events is made from the Reviewed Event Bulletin (REB) events by meeting certain criteria, such as a body wave magnitude between 4 and 6. Soon after the calibration effort began at the PIDC it was recognized that that not all calibration events can be accepted as ground truth (Bondar,1998b). A primary objective of this project is the formation of databases which provide improved ground truth for events in China, former Soviet Union, and North America, as selected from the CMR Calibration Event Bulletin (CEB). Several studies have demonstrated the usefulness of such databases for improved monitoring of the CTBT (e.g., Grant *et al.*, 1997; Bondar *et al*., 1998). Re-analysis of the IDC data combined with retrieval and analysis of selected IRIS data leads to improved location and depth estimates of CEB events.

RESEARCH ACCOMPLISHED

(a) Formation of Databases

Analysis and relocation of all (70) events in China from the CMR Calibration Event Bulletin (CEB) have been completed by using waveform data from over 140 IDC stations and over 130 IRIS stations. The CEB epicentral locations of these 70 events are shown in Figure I along with the locations of both IDC and IRIS stations which recorded these events. Note the large number of IRIS stations at regional distances for most of the 70 Chinese events.

Analysis procedures followed are basically the same as for the REB bulletin. Most revisions consist of minor re-timings with emphasis on defining phase types, and missed phases are added whenever possible. At far regional distances, most later crustal phase types Sn and Lg are often made non-defining due to large residuals that result from using IASP91 travel times. Occasionally, a CEB event undergoes significant changes when, for example, improper depth phases or multiple events occur. The analysis made use of waveform data from over 140 IDC stations contributing, over 4000 arrivals. A total of 132 IRIS stations contributed nearly 3800 waveforms and 2300 arrivals. The criterion for acquiring data from IRIS stations is primarily to collect data within 25 degrees distance, then to reduce any azimuthal gap for distances beyond 25 degrees (up to about 50 or 60 degrees), and to obtain data at PKP distances. The most important contribution of the IRIS data is the addition of a large number of arrivals from regional phases. For epicentral distances less than 17 degrees (about 1900 km), the regional phases (Pn, Pg, Sn, and Lg) in the three datasets consisting of combined IRIS and revised IDC, only the revised IDC, and the original CEB arrivals, total 1374, 230, and 140, respectively. A comparison of the propagation paths available in the original CEB dataset and the IRIS and revised IDC dataset for Pn and Lg, shown in Figures 2 and 3, respectively, clearly demonstrates the significant contribution the IRIS data have made in providing additional regional phases.

A modified version of *LocSAT*, called *Locate*, which uses a graphical interface and allows interactive processing to generate location results, is used to locate the CEB events. The location time residuals for each arrival are displayed, and the residuals that exceed the CMR limit for allowable defining phases are highlighted. Relocations can be computed quickly for different defining arrivals to assess the effect of adding arrivals, such as those measured from IRIS data, to the original CMR data set.

Two sets of location results per event are derived: the first by using only the revised IDC arrivals, and the second by using both the IRIS and revised IDC data. As an example, the recording stations, which provided data used in determining the epicentral location and depth of the Chinese seismic event of 12 March 1996, are shown in Figure 4a. For this seismic event near the borders of Mongolia and former Soviet Union, the original CEB solution was based on the use of only 44 IDC stations, whereas the solution based on the

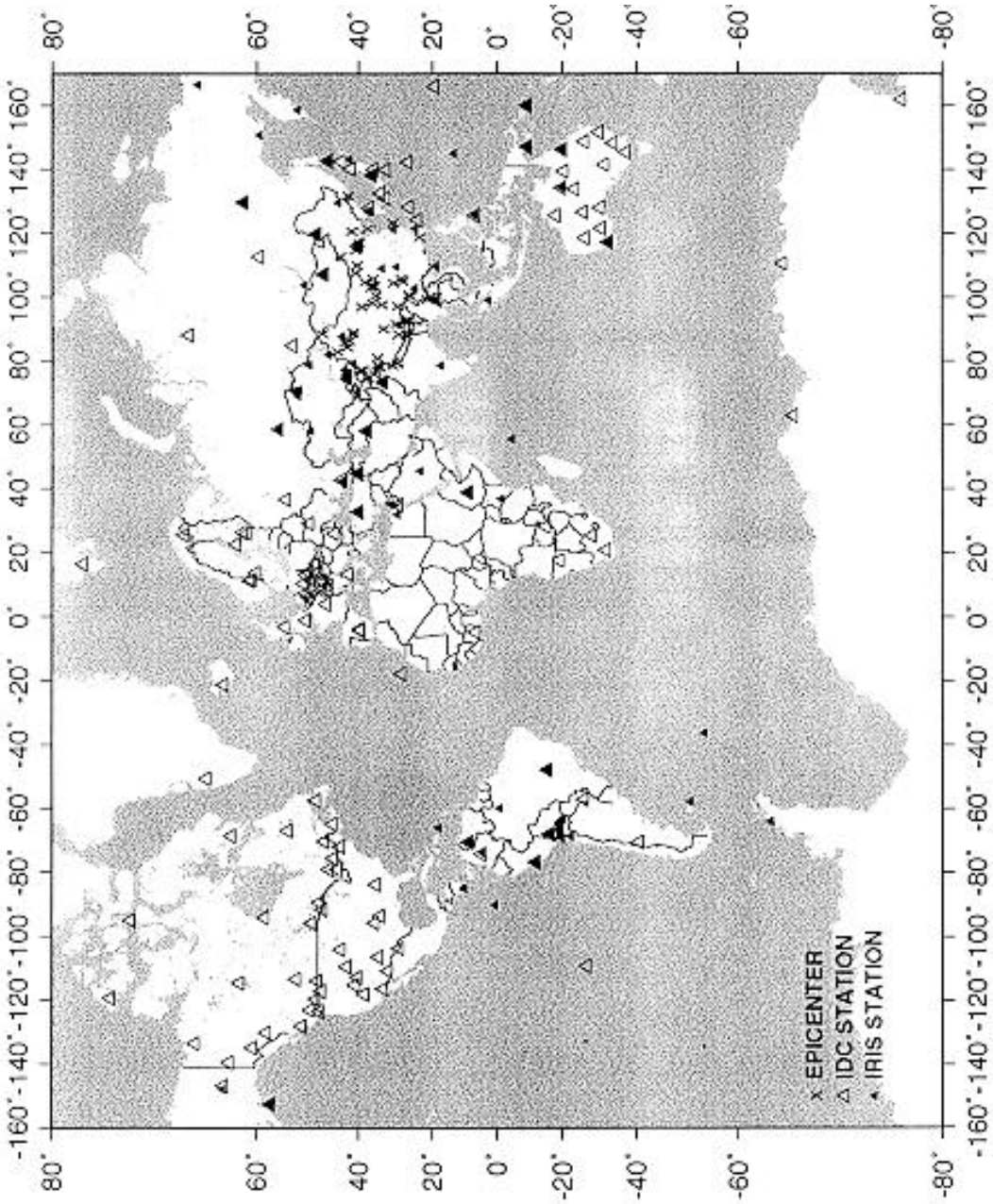

Figure 1. CEB epicentral locations of 70 events in China and the IDC and IRIS recording stations which provided waveform data analyzed in this project. Note the large number of IRIS stations at regional distances for most of the 70 Chinese events.

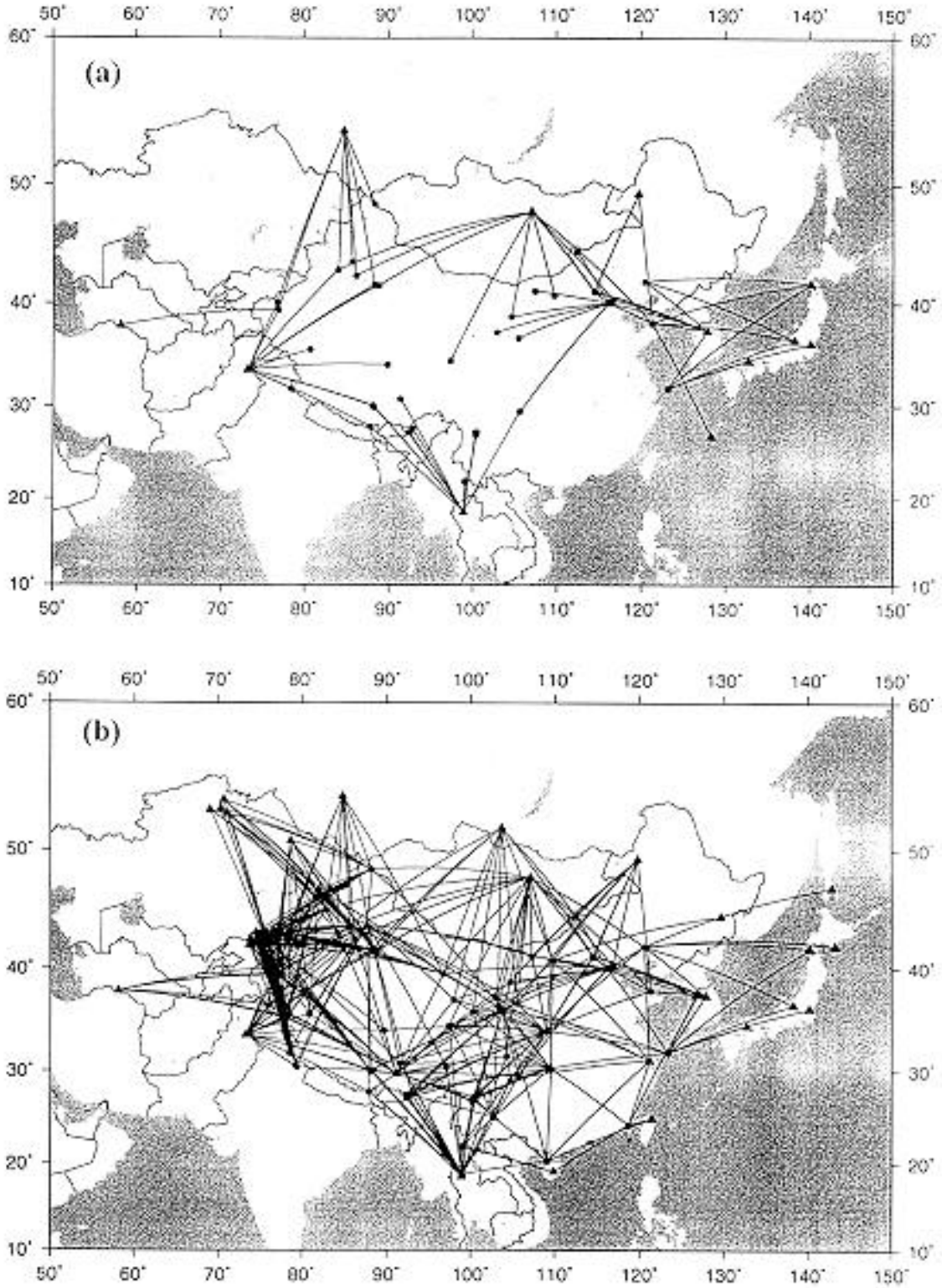

Figure 2. Comparison of Pn phases and corresponding propagation paths available in (a) original CEB dataset showing 42 events and 74 paths, and (b) IRIS and revised IDC dataset showing 66 events with 592 paths. The IRIS data provide a large number of additional Pn arrivals.

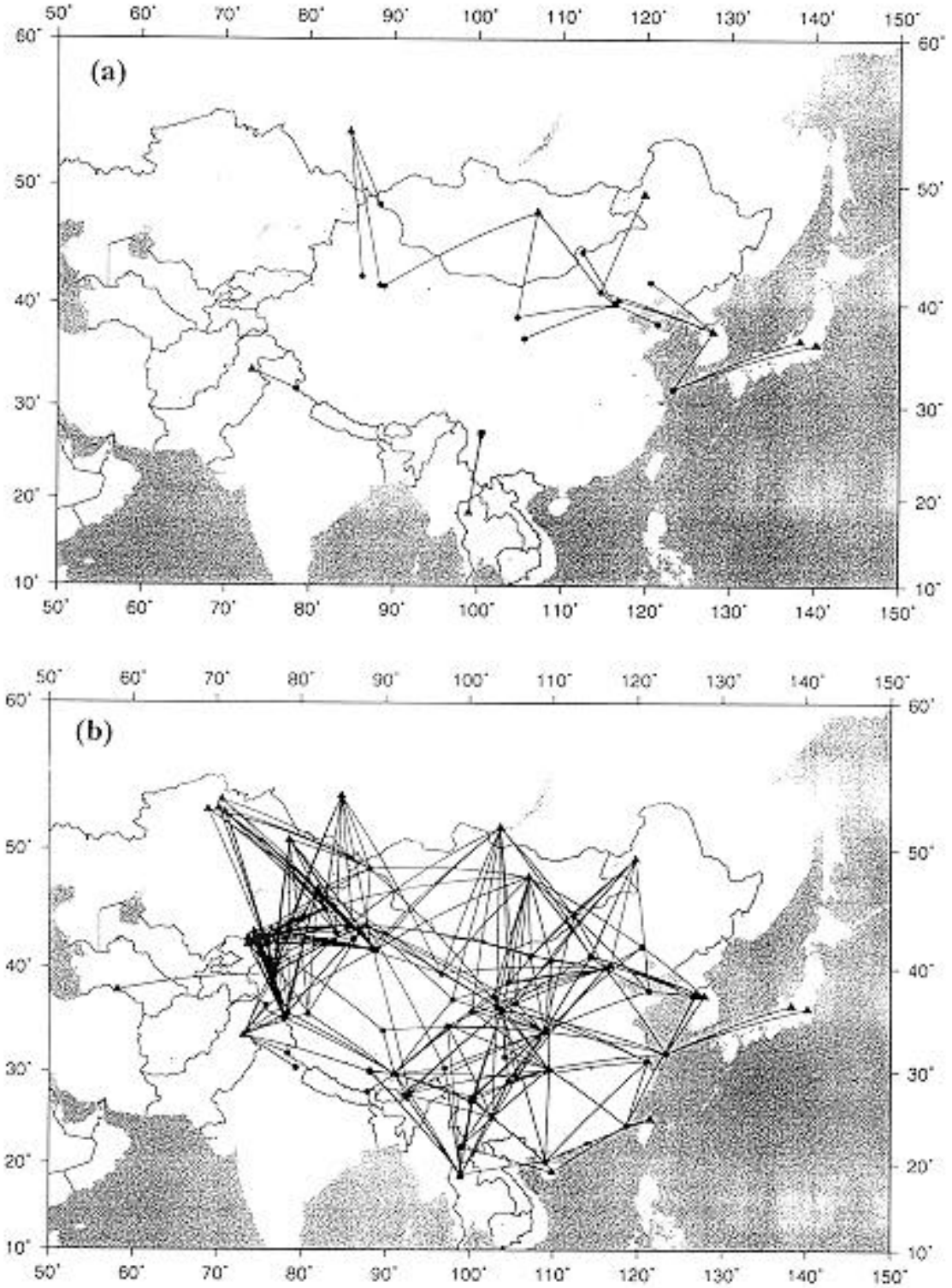

Figure 3. Comparison of Lg phases and corresponding propagation paths available in (a) original CEB dataset showing 18 events and 24 paths, and (b) IRIS and revised IDC dataset showing 62 events with 414 paths. The IRIS data provide a large number of additional Lg arrivals.

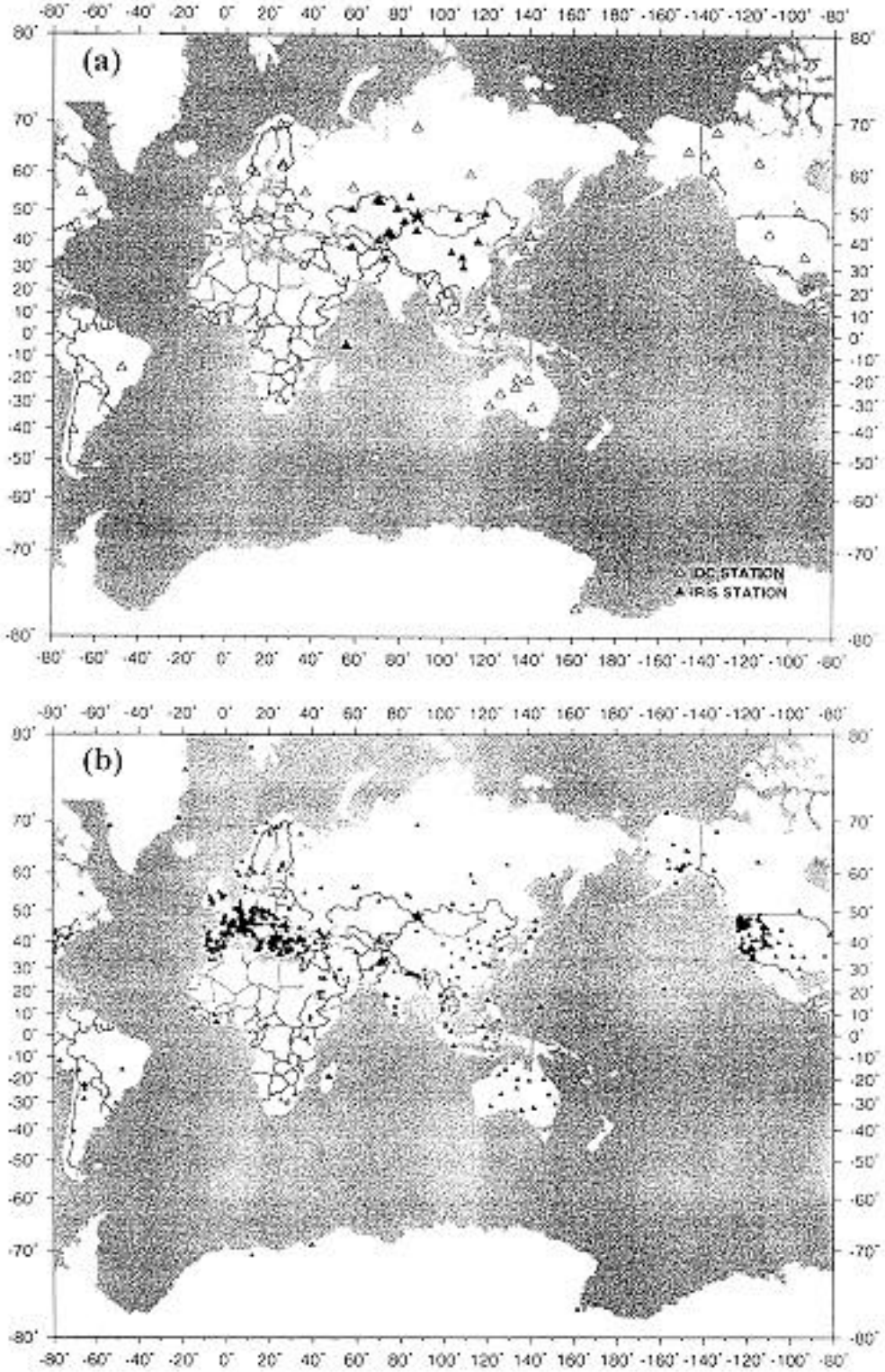

Figure 4. Comparison of recording stations providing data used in determining the epicentral location and depth of the seismic event of 12 March 1996, shown by a star; (a) 44 IDC stations used in the original CEB solution and the additional 23 IRIS stations, and (b) 586 recording stations used in the ISC solution. Note the significantly larger number of stations used by the ISC.

combined use of IRIS and revised IDC data made use of 23 IRIS stations in addition to the 44 IDC stations. For the same event, the epicentral location determined by the International Seismological Centre (ISC) used data from a much larger number (586) of recording stations (Figure 4b). A comparison of the epicentral location and depth for this seismic event, as obtained from seven different sources: (1) the original CEB, (2) revised IDC, (3) combination of IRIS and revised IDC, (4) ISC, (5) National Earthquake Information Center, USGS, (6) BJI (State Seismological Bureau, Beijing, China), and (7) MOS (institute of Physics of the Earth, Moscow, Russia), is shown in Figure 5. Location error ellipses from the first four sources are also shown. The solutions from BJI and MOS networks, based on the use of numerous stations at local and regional distances, are likely to be more accurate than that in the original CEB. It is interesting to note that the epicentral location based on the combined use of IRIS and revised IDC (No. 3 in Figure 5) lies somewhere in the middle of the BJI and MOS determinations and the CE9 location. The derived source depth also lies between the BJI and MOS determinations. It seems therefore that the solution based on the use of both IRIS and revised IDC data provides improved estimates of both epicentral location and depth in comparison to the original CEB solution.

For the entire Chinese database of 70 events, a comparison of the CEB locations with those derived by using the two methods mentioned above show no systematic trend and generally small differences (average less than 9 km in each case). Figure 6 shows differences in epicentral location between the original CEB and those derived by the combined use of both IRIS and revised IDC for all 70 Chinese events. The arrows point from the CEB location to the new location and the location differences range between 0.8 and 28.3 km with an average value of 8.5 km. However, a similar comparison of source depths shows the CEB estimates to be significantly larger than each of the other two estimates, suggesting that the CEB source depths may have been overestimated. The differences in source depth between the

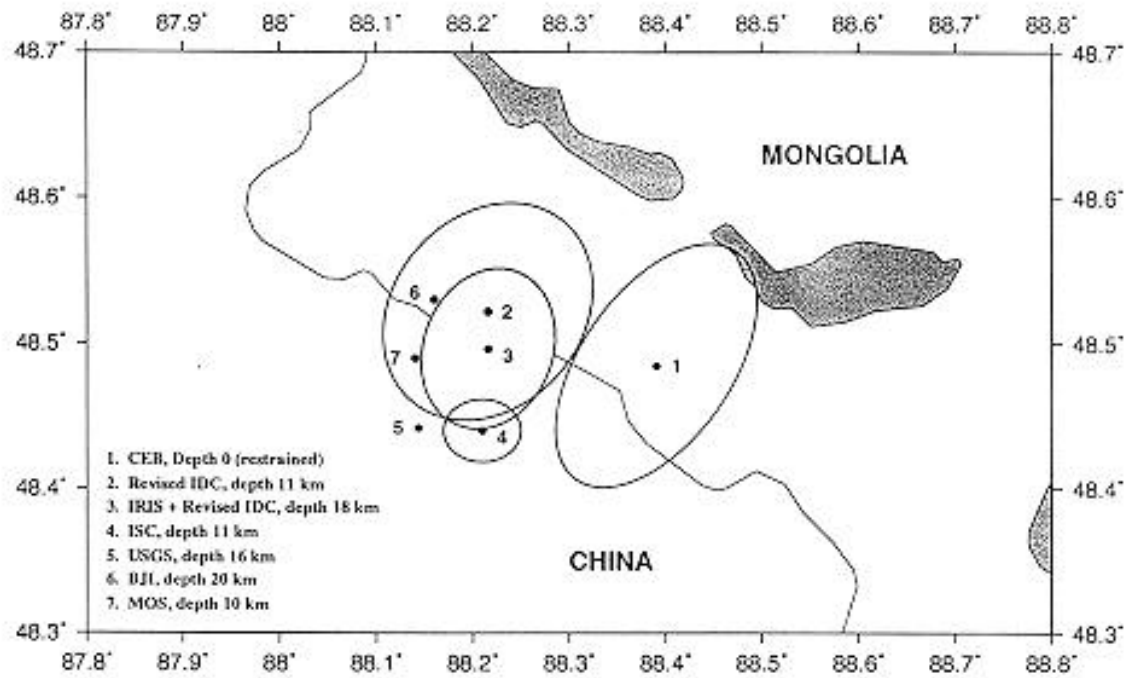

Figure 5. Comparison of epicentral location and depth for seismic event of 12 March 1996 from seven different sources: original CEB, Revised IDC, IRIS + Revised IDC, ISC, USGS, BJI, and MOS. Location error ellipses from the first four sources are also shown. Results based on the use of both IRIS and revised IDC data (No. 3) appear to provide improved estimates of both epicentral location and depth.

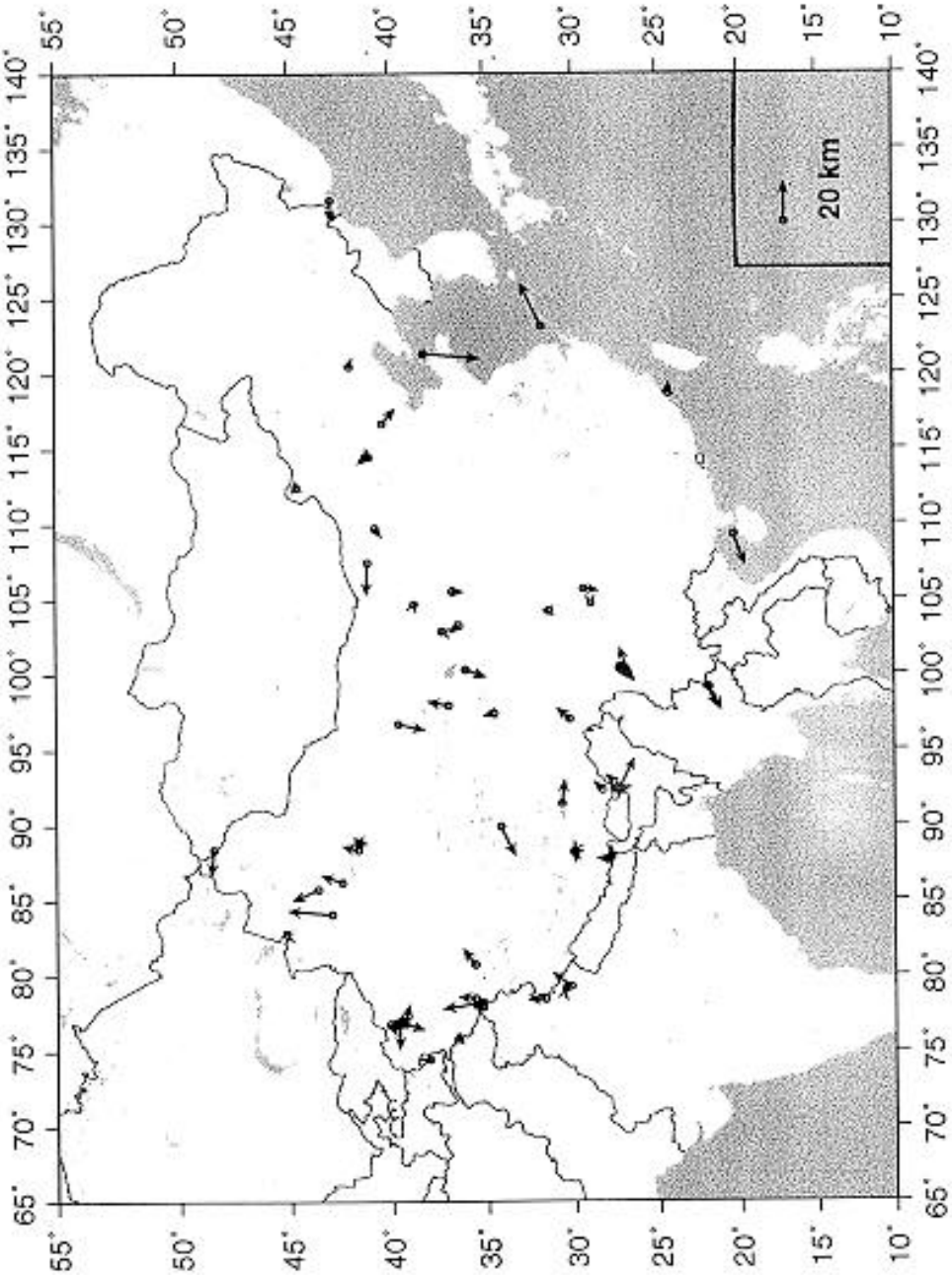

Figure 6. Differences in epicentral location between the original CEB and those derived by the combined use of both IRIS and revised IDC for all 70 Chinese events. The arrows point towards the new location and the average difference between the two locations is 8.5 km.

original CEB and those derived by the combined use of both the IRIS and revised IDC data, shown in Figure 7, have an average value of about 17 km, indicating that for most events the new depths are significantly smaller than those in the original CEB. Results from only 49 events are shown since, out of 70 Chinese events, source depths for both sets of data were available for only 49 events. These differences in source depths are supported by the remarkably good agreement observed between depth estimates based on the combined use of IRIS and revised IDC data and those available from the ISC for a large number of common events. According to Storchak (1999), epicentral locations from the ISC, derived from a much larger dataset and covering a larger range of azimuths than those by the IDC, are probably more accurate than the IDC locations. It should however be noted that the ISC determinations are generally over two years behind real time and may therefore be of only limited use for monitoring the CTBT.

The results obtained from analysis of all (70) CEB events in China were delivered to the CMR. Each event location consists of four CSS3.0 files, including the origin, origerr, arrival, and assoc database tables. Therefore, a total of 560 such database files were delivered to CMR ready for installation into their CEB database. All IRIS waveform data which contributed to the above 70 CEB events in China were also delivered to the CMR on one 8 mm archive tape in CSS3.0 format. Analysis, similar to that carried out for the Chinese events, has been completed for 43 CEB events in the former Soviet Union by using waveform data from both IDC and IRIS stations.

(b) Analysis of Several Other Datasets

We are also analyzing near-field, local, and regional data from the 1997 Depth of Burial Experiments at the former Soviet test site at Balapan in eastern Kazakhstan. More recent (1998) data from the Degelen test site are also being retrieved and examined.

CONCLUSIONS AND RECOMMENDATIONS

Formation of databases with significantly improved ground truth can be accomplished by a combination of re-analysis of the IDC data and retrieval and examination of data from other available sources such as the IRIS. For the Chinese dataset, the most important contribution of the IRIS data is the addition of a large number of arrivals from regional phases, so that solutions based on the combined use of both IRIS and IDC data provide improved estimates of both epicentral location and depth in comparison to the original CEB solution. Such analysis should also be carried out for other regions of the world.

REFERENCES

Bondar, I. (1998a). PIDC Calibration Event Bulletin Technical Reference, CMR-98/16, Center for Monitoring Research, Arlington, Virginia.
Bondar, I. (1998b). PIDC Ground Truth Event Bulletin Technical Reference, CMR-98/21, Center for Monitoring Research, Arlington, Virginia.
Bondar, I., X. Yang, J. Wang, M. Bahavar, H. Israelsson, and K. McLaughlin (1998). Tuning, and calibration activities at the PIDC, Proc. of the 20,h Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty (CTBT), U. S. Department of Energy, Washington, DC, 1-10.
Grant, L., R. Wagner, F. Ryall, and I. Henson (1997). Ground truth database for regional seismic discrimination research, Proc. of the 19th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty, U. S. Department of Energy, Washington, 867-876.
Storchak, D. A. (1999). Reasons for pIDC mislocations: a view from the ISC, oral presentation at the 1999 General Assembly of the IUGG, 30 July 1999 (EG 12/LT4).

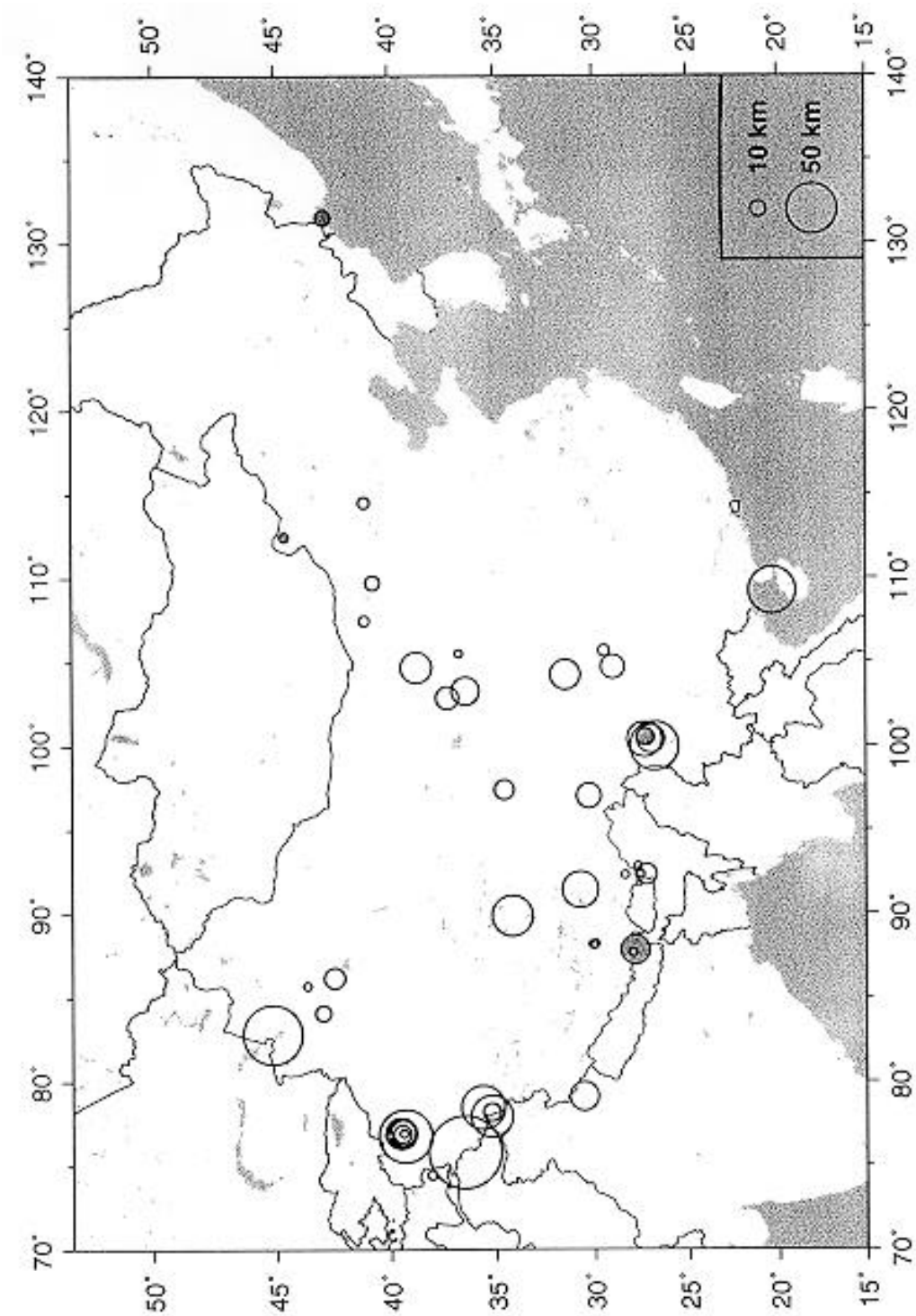

Figure 7. Differences in estimates of source depth between the original CEB and those derived by the combined use of both IRIS and revised IDC for 49 Chinese events. The open circles denote new depths smaller than those in the CEB and the filled circles the opposite. For most events, the new depths are significantly smaller than those in the original CEB.