RECENT CALIBRATION EVENTS IN THE UNITED STATES

James W. Dewey[1], Ellen Herron[2], and John O. Kork[1]
[1]Adjunct US National Data Center at U.S. Geological Survey (USGS)
[2] U.S. National Data Center at Air Force Technical Applications Center (AFTAC)

Sponsored by United States Air Force
Contract Number: N12FY99000785

ABSTRACT

The United States has agreed to provide calibration information to the Comprehensive Nuclear-Test-Ban-Treaty Organization (CTBTO). As part of this contribution, the U.S. National Data Center (NDC) has identified 64 Calibration seismic events that occurred under U.S. Territory from 1996 -- 1998, a period during which many stations of the International Monitoring System (IMS) have been operating.

The U.S. Calibration events meet three selection criteria: (1) their epicenters are accurate to within 10 km in absolute terms; (2) their magnitudes (USGS m_b or local magnitudes as reported by the USGS) are larger than 3.5; (3) data from IMS stations are excluded from the hypocenter calculation. In order to be judged accurate to within 10 km in absolute terms, the epicenters must be known either independently of arrival-time data (e.g., large mining events) or be associated with well-recorded events for which location errors, due to mistaken phase interpretations or inaccurate velocity models, are minimized (e.g., events occurring within a local or regional seismograph network). Included among the 64 selected Calibration events are 38 events for which we believe the calculated focal depth as well as the epicenter are accurate to within 10 km.

For geographic regions in which many events meet the selection criteria, we selected approximately five Ground Truth events per square degree or per seismotectonic province. These events were selected to provide a variety of source types and event sizes.

Data provided to the CTBTO will include event origins, formal statistical estimates of hypocentral precision, the arrival-time data used to compute the hypocenters, a station-coordinate file, documentation of the velocity models with which the hypocenters are determined, and documentation of how local magnitudes are determined.

In addition, we have used sparse subsets of data randomly selected from a widely recorded event to explore criteria for being confident that events occurring in geographic regions covered by less dense local or regional networks have epicenters that are accurate to within 10 km.

Key Words: location, calibration, seismic regionalization

OBJECTIVE

The purpose of this study is to identify seismic events that have occurred in the United States during the time-period in which a significant fraction of International Monitoring System (IMS) seismographs have been in place, that are large enough that they might be expected to have been detected on IMS stations, and that are well enough located that they may be used to test and improve hypocenter accuracies of events that are located by the International Data Center (IDC). Contribution of such data was suggested in CTBT/Working Group B Sixth Session (WGB-6)/ Conference Room Paper .26 (CRP.26) published in June 1998. The data will be submitted to the Comprehensive Nuclear-Test-Ban Treaty Organization/Provisional Technical Secretariat (CTBTO/PTS).

We required that selected seismic events have the following characteristics: (1), they occurred in 1996-1998; (2), they were assigned magnitudes (m_b or local magnitude) larger than 3.5; (3), their epicenters are accurate to within 10 km when computed without using data from the IMS. For large areas of the U.S., there are no seismic events that have these desired characteristics, either because no sufficiently large seismic events occurred in the areas in 1996-1998 or because there were no means to assure epicenter accuracies to within 10 km. For a few areas, however, there are many seismic events that have the desired characteristics. For the latter areas, we have provided about five events per tectonic province or 1° by 1° area: for such areas, our intention is to provide data from a representative range of seismic-source types and source sizes, without overwhelming the IDC with essentially redundant data. We refer to the selected seismic events as “Calibration” events.

For the US Calibration events, we did not require that focal depths have accuracies within 10 km. The problem of determining accurate focal depths is somewhat decoupled from that of determining accurate epicenters. Highly accurate focal depths of shallow-focus events can commonly be determined with teleseismic waveform modelling and only approximate knowledge of the velocity structure in the hypocentral region, even if the teleseismically determined epicenters are severely biased by lateral velocity heterogeneity. Conversely, highly accurate epicenters can be determined with local arrival-time data and regional velocity models, even in cases when the regional velocity models are not well-enough known to ensure highly accurate focal depths. We will identify events with focal depths that are probably accurate to within 10 km, but these focal depths are more susceptible than the epicenters to bias by minor inaccuracies of the assumed velocity model.

We will provide the CTBTO/PTS with Calibration event origins and formal confidence intervals on epicenter and focal depth, arrival-time data used to locate the events, station-coordinate files, documentation of velocity models used in hypocenter computations, and documentation of procedures used to compute local magnitudes. The event origins and confidence intervals will be based entirely on data from non-IMS stations.

RESEARCH ACCOMPLISHED

Our research focussed on the requirement that the calibration epicenters be within 10 km of the true epicenters. Almost all of the events that we have selected as Calibration events are earthquakes. Their epicenters and focal depths were determined from arrival-time data, and therefore have some uncertainty associated with them. The possible sources of location error are diverse, and the possibility of location error had to be evaluated from several perspectives. We have used a three-stage process to select the Calibration events that we will present to the CTBTO/PTS. First, we selected and relocated candidate Calibration events on the basis of the azimuthal and distance distribution of stations that recorded the events. Second, we required that the epicenter confidence ellipses of the selected events be smaller than a threshold value. Third, we required that the epicenters of the selected events be robust under changes in the assumptions used to calculate the epicenters. The first stage yielded 80 candidate events: the second and third stages resulted in some of the candidates being judged not reliable, so that 64 Calibration events were finally selected. In addition, we use randomly selected, sparse subsets of data from well-recorded events to explore the effects of nonlinearities in the location process, non-Gaussian errors in arrival-time data, and lateral variations in seismic-wave velocities.

Selection of candidates for US Calibration events

The U.S. Geological Survey/National Earthquake Information Center (USGS/NEIC) and the institutions that operate regional networks recorded a large number of shocks of magnitude greater than 3.5 during 1996-1998 (Figures 1 -2). We used formal and informal criteria to select a small subset of these events as candidates for Calibration events. The process of selecting candidates for Calibration events, discussed in this section, resulted in 80 candidate Calibration events. We concluded that 64 of the 80 are located with sufficient confidence that they can be presented to the CTBTO/PTS as Calibration events.

To select most candidate Calibration events, we reviewed data listed in the Earthquake Data Reports (EDR's) of the USGS/NEIC and applied the following "stringent station-distribution criteria:"

A. The candidate Calibration events were recorded by at least 10 stations (excluding IMS Primary or Auxiliary stations) within an epicentral distance of 250 km. The use of ten stations gives redundancy of data, so that a gross error in arrival-time data from one or a few stations can be detected by comparison with data from the other stations. Local and near-regional P-wave arrivals from moderate-sized earthquakes tend to be readable to high precision. For most of the regions studied, local and regional velocity models are available that more closely reflect the true velocities than would global-average velocity models. Finally, for a given percentage error in the velocity model, the absolute error in predicted epicenter-to-station distance will be smaller for a short raypath than for a long raypath.
B. The azimuthal gap in the distribution of non-IMS stations was 90 degrees or less. A wide azimuthal range of observations makes the calculated epicenter less dependent on second-order differences between the true velocity structure and the velocity model used in calculating the hypocenters. In addition, arrival-time observations from all quadrants around the epicenter provide sufficient redundancy of azimuthal coverage that a gross arrival-time error at one of the stations can be detected by its effect on the calculated standard error
C. At least one non-IMS station had to be situated at an epicentral distance of 30 km or less. The presence of arrival-time data from close-in stations is critical for calculating reliable focal depths of earthquakes with arrival time data alone.

The stringent station-distribution criteria drastically reduced the area of the United States for which we were able to identify candidate Calibration events (Figures 1-2). In an effort to broaden the area from which candidate Calibration events might be selected in the future, we have also identified (Figures 1-2) events that, on the basis of the data listed in the EDR's, meet the following "relaxed station-distribution criteria":

i. The events were recorded by at least 5 non-IMS stations situated within 250 km.
ii. The azimuthal gap in the distribution of non-IMS stations within 250 km is 180° or less.

The events satisfying the relaxed criteria are of interest because it may be possible to acquire data for some of these shocks that were not reported to the USGS/NEIC but that, together with the data that are reported to the USGS/NEIC, enable the events to satisfy the stringent criteria. In addition, as will be discussed later, it appears possible to find events that fall short of the stringent criteria but that can still be confidently located to within 10 km. In the present study, we did not select Calibration events from among events that satisfied only the relaxed station-distribution criteria.

Independently of station-distribution criteria that enable a reliable epicenter to be computed from seismographic arrival-time data, we searched the EDR's for human-induced seismic events that occurred in the U.S. in 1996-1998, had magnitudes greater than 3.5, and had locations that were known independently of arrival-time data. We found one such event, a rockburst in northern Idaho.

For territories under U.S. control in the western Pacific, we found no seismic events that met stringent or relaxed station-distribution criteria, or that met criteria for human-induced Calibration events. We identified several events near Puerto Rico that met the relaxed station-distribution criteria, but not the stringent criteria.

In some seismically active areas of California, there were many seismic events that met the stringent station-distribution criteria, and we selected only a few of these events from each tectonic province or 1° by 1° area as candidates for Calibration events. For these regions, we did not apply formal criteria for selecting the candidate Calibration events from among all events that met the station-distribution criteria. Instead we selected events so as to include a variety of source-types and sizes, so that users of our Calibration events could explore the effects of local source-variations or differing source size on IMS capabilities.

Calculating hypocenters and final selection of US Calibration events

We relocated all candidate Calibration events using the program HYPOELLIPSE (Lahr, 1989), a program for computing hypocenters from local and regional data that is widely used by operators of regional seismograph networks and that is available free of charge. We located earthquakes in each region with one-dimensional velocity models that have been determined for the region in which the earthquakes occurred or for a tectonically similar region nearby. Theoretical travel-times to individual stations were modified by elevation corrections, calculated according to the default procedure of Lahr (1989), but otherwise station-specific adjustments were not applied to the theoretical travel-times. Observational data consisted primarily of first-arriving P-waves recorded at local and regional stations; some S-wave data were used, exclusively for stations situated within 100 km of the epicenters. Arrival-time data were taken from the EDR's of the USGS/NEIC (for earthquakes outside of California), from the Northern California Earthquake Data Center (for earthquakes in Northern California), and from the Southern California Earthquake Data Center (for earthquakes in Southern California).

Confidence regions, reflecting the precisions of the calculated epicenters and focal depths, are based on the assumption that the standard deviation of the P-wave arrival-time observations for a given earthquake is equal to either, (a), the standard error of the travel-time residuals for that earthquake or, (b), a lower-bound standard error (.5 for Alaskan events and .4 for other events), whichever is larger. The lower-bound values for P-wave standard error were taken to be the median values of P-wave arrival-time standard errors that resulted from a preliminary location of the candidate Calibration events in each region. The confidence regions normally computed by HYPOELLIPSE are associated with a 68% level of confidence. We have adjusted the sizes of confidence regions to correspond to a 90% level of confidence.

Hypocenters of most of the events had been earlier computed by the institutions that run regional networks within whose boundaries the events occurred. We relocated the events in order to prepare data files to transmit to the CTBTO/PTS, to be sure that we were correctly describing how the contributed hypocenters were computed, to exclude data from IMS stations that might have been used in the original computations, and to obtain a sense of the stability of each computation.

In our relocation of the candidate Calibration events, we used two different implementations of the "distance-weighting" option of HYPOELLIPSE (Lahr, 1989). We examined the confidence regions computed by HYPOELLIPSE, compared the results of applying the two different distance-weighting procedures, and compared our locations with prior regional network locations when the latter were available.

Our two implementations of the HYPOELLIPSE distance-weighting option apply this weighting to the data after four iterations of the "calculate hypocenter/adjust hypocenter" cycle. Both implementations assign full weight to P-wave readings out to distances of 100 km from the epicenter. Our preferred implementation uses only the close-in data; weights assigned to P-wave readings decrease from full weight at 100 km to zero weight at 110 km. A shallow-focus event epicenter that has been calculated under this weighting scheme and that has a small confidence ellipse is likely to be highly accurate, because the short lengths of the raypaths minimize the amplitude of travel-time biases produced by errors in the assumed velocity model. However, some events, widely enough recorded by stations within 250 km to meet the stringent station-distribution criteria, are not widely enough recorded by stations within 110 km that they can be reliably located with the preferred implementation of the distance-weighting option. In the second implementation of the HYPOELLIPSE distance-weighting option, weights assigned to P-wave readings decrease from full weight at 100 km to zero weight at 250 km. By including the data between 110 km and 250 km, the second implementation of the distance-weighting option greatly extends the azimuthal coverage of some events, but the longer ray-paths make the location more susceptible to bias by errors in the assumed velocity model. In this study, we found that epicenters calculated with the second weighting option were nearly identical to epicenters calculated with the first weighting option, when the epicenter computed with the first weighting option was well-located (as inferred from the size of the epicenter confidence ellipse). However, for some events the focal depths calculated with the second weighting option differed significantly from the focal depths calculated with the first weighting option. In most cases, these depth differences were apparently due to bias allowed by the second distance-weighting option. A distance-dependent structure in the P-wave residuals, and a consistency in this structure from events within the same small source region, suggested that the Pn-arrivals at distances between 110 km and 250 km were the source of the depth discrepancy. Pn arrivals would be sensitive to incorrect depth to the M-discontinuity in the assumed velocity model.

From the candidate Calibration events, the 64 events that we will present to the CTBTO/PTS will be those for which the 90 percent confidence ellipses on epicentral coordinates have semi-axes less than 5.0 km. In the case of earthquakes lying within a regional network, we also require that our relocated epicenters of the Calibration events be within 5.0 km of the epicenters computed by regional networks. The hypocenters and confidence-regions of events calculated with the preferred distance-weighting option will given if the 90 percent confidence ellipses have semi-axes less than 5.0 km: otherwise, we will give the epicenters, focal depths, and confidence-regions calculated with the weighting option that incorporates observations collected out to distances of 250 km.

For 38 of the 64 selected Calibration events, the 90 percent confidence intervals on focal depth were smaller than 5 km, and the focal depths that we have calculated are within 5 km of the focal depths calculated by the institutions running the regional networks in which the events occurred. The focal depths of most of these events are probably accurate to within 10 km, and we will so label them in the files that we give to the CTBTO/PTS. Focal depths calculated with arrival-times from regional networks are, however, generally more vulnerable than epicenters to biasing effects of slightly inaccurate velocity models. A substantially erroneous focal-depth is more likely than a substantially erroneous epicenter to escape detection by our three-stage procedure for identifying Calibration events.

A random-sampling exploration of station-distribution criteria for selecting candidate Calibration events

Most of the 64 events that we will present as Calibration events were initially identified as candidate events on the basis of the stringent station-distribution criteria discussed in the section, " *Selection of candidates for US Calibration events*." In the present section, we explore the robustness of the station-distribution criteria. It would be well to know if the criteria are too stringent as we currently define them. Demonstration that less stringent criteria are adequate to ensure 10-kilometer accuracy would permit Calibration events to be selected from more of the U.S.

Our testing procedure started with events that were so well recorded that they will be accurately located even in the presence of non-Gaussian errors or large velocity heterogeneity. We randomly selected stations to form sparse subsets of the overall station sets that were used to locate the well-recorded events. We calculated hypocenters for each of the sparse subsets and evaluated how well the calculated hypocenters and confidence regions agreed with the hypocenter that was calculated from all the data.

Figure 3 illustrates epicenters calculated with sparse subsets of stations that recorded a magnitude (mb, USGS) 4.8 earthquake that occurred on the San Andreas fault in Central California. The earthquake occurred in the midst of one of the most spatially dense networks of local seismographs in the world and was recorded by over 80 stations within 110 km. The source region of the earthquake has high lateral variations in seismic wave velocities. If a one-dimensional velocity model is used in computing the hypocenter, these velocity variations can result in hypocenter biases of several kilometers even when all of the local data are used in the location process (Uhrhammer et al., 1999). It might therefore be expected that some epicenters calculated with few stations might be severely biased. Three groups of sparse subsets were selected for testing: subsets consisting of eight stations, subject to the requirement that the azimuthal gap in station distribution be less than 120°; subsets consisting of five stations, subject to the requirement that the azimuthal gap in station distribution be less than 120°; subsets consisting of five stations, subject to the requirement that the azimuthal gap in station distribution be less than 180°. Stations selected were required to be situated within 250 km of the epicenter calculated with all data.

None of the three station-distribution criteria used to select the sparse subsets were sufficient, of themselves, to ensure that errors in epicenters were less than 10 km. However, 31 of the 100 epicenters calculated from the eight-station subsets, with azimuthal gaps less than 120°, had 90% confidence ellipses with semi-axes shorter than 5.0 km; all fall within 10 km of the likely true epicenter (Figure 3). Similarly, 6 of the 100 epicenters calculated from five-station subsets, with azimuthal gaps less than 120°, had 90% confidence ellipses with semi-axes less than 5.0 km; all of these also fall within 10 km of the likely true epicenter (Figure 3). None of the 100 epicenters calculated from five-station subsets with azimuthal gaps as large as 180° had 90% confidence ellipses with semi-axes shorter than 5 km. These results suggest that candidate Calibration events can be identified using station-distribution criteria that are less severe than the stringent station-distribution criteria we have used: it is necessary that additional criteria based on the hypocenter computations, such as the size of confidence intervals, be used to select the Calibration events from candidate Calibration events.

CONCLUSIONS AND RECOMMENDATIONS

The United States National Data Center has identified 64 seismic events that occurred in United States Territory to be provided to the CTBTO/PTS as U.S. Calibration events. The Calibration events occurred in 1996-1998; they were assigned magnitudes (m_b or local magnitude) larger than 3.5; and their epicenters are accurate to within 10 km when computed without using data from the IMS. For areas in which there were many events that met the preceding time, size, and accuracy requirements, we selected a limited number of calibration events so as not to provide redundant data to the CTBTO/PTS. The location of one of the events is known, independently of arrival-time data, because it was a rockburst in a mine. Of the 64 events selected as calibration events, 38 have focal depths that are probably accurate to within 10 km.

Most of the work conducted in this study addresses the requirement that the event epicenters be accurate to within 10 km. The locations of earthquake epicenters that are calculated with phase arrival-times are always susceptible to some error, and, by postulating a bizarre velocity structure and bizarre timing errors, we could probably always develop a scenario under which the epicenter of a well-recorded local earthquake with a small epicenter confidence ellipse might be in error by more than 10 km. We used three general approaches to be confident that our locations would be accurate to within 10 km in the presence of typical sources of location error. First, we identified candidate Calibration events by using stringent station-distribution criteria, which required that the events be computed with data from local and regional stations distributed over a wide range of azimuths. Second, after relocating the candidate Calibration events with the HYPOELLIPSE program and a regional one-dimensional velocity model, we applied conservative criteria to the HYPOELLIPSE-produced confidence intervals to select Calibration events from among the candidate Calibration events. Third, for each candidate Calibration event, we have available at least two hypocenters computed by HYPOELLIPSE with different weighting assumptions, and for most candidate Calibration events we have in addition the hypocenters computed independently by the institutions that run the regional seismic networks in which the events occurred. We compared our hypocenters with the regional-network hypocenters to check the stability of our computed hypocenters under changes in the assumptions used in their calculation. We identified 80 candidate Calibration events on the basis of the station-distribution criteria. The 64 events that will be given to the CTBTO/PTS are those candidate Calibration events that also satisfied both the confidence-interval criteria and the requirement that their epicenters computed with differing starting assumptions be mutually consistent.

The station-distribution criteria that we used to define candidates for Calibration events may be unnecessarily conservative. This possibility is suggested by study of epicenters computed from sparse subsets of data randomly selected from a widely recorded event. Because criteria on the size of confidence regions are also used in the final selection of Calibration events, candidates for Calibration events can probably be identified using less stringent criteria on the number of recording local stations and the size of the azimuthal gap in station distribution.

REFERENCES

Lahr, J. C., 1989, HYPOELLIPSE/Version 2.0: A computer program for determining local earthquake hypocentral parameters, magnitude, and first-motion pattern, U. S. Geological Survey Open-File Report 89-116, 92 p.

Uhrhammer, R., Gee, L. S., Murray, W., Dreger, D., and Romanowicz, B., 1999, The M_W 5.1 San Juan Bautista, California earthquake of 12 August 1998: Seismological Research Letters, v. 70, p. 10-18.

Figure 1. Seismic events of M>3.5 in 1996-1998 in Alaska (a) and Hawaii (b). Black circles denote events whose epicenters and focal depths are both judged accurate to within 10 dm. Gray circles denote events whose epicenters are judged accurate to within 10 km. White circles denote other events for which USGS/NEIC data files show five or more stations that have epicentral distances less than 25o km and that together have an azimuthal gap less than 180 deg.

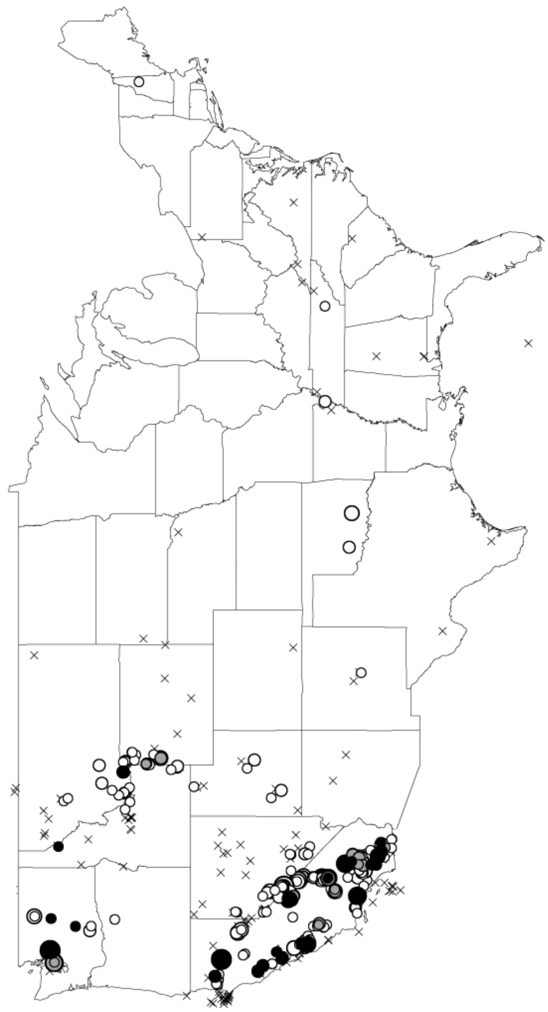

Figure 2. Seismic events of M > 3.5 in 1996-1998 in the 48 contiguous states. Black circles denote events whose epicenters and focal depths are judged accurate to within 10 km. Gray circles denote events whose epicenters are judged accurate to within 10km. White circles denote events for which USGS/NEIC data files show five or more stations that have epicentral distances less than 250 km and that together have an azimuthal gap less than 180 deg, and x's denote other events.

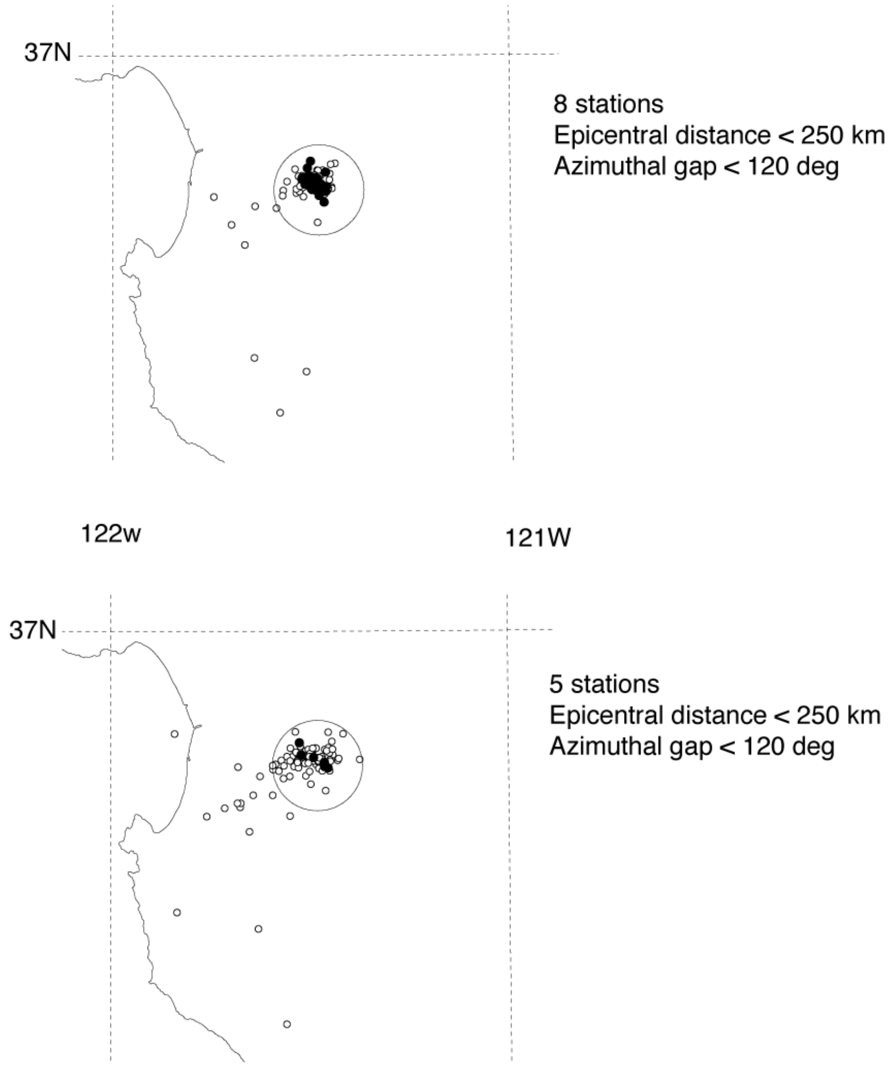

Figure 3. Epicenters of the California earthquake of 1998 August 12, 14:10 UTC, calculated with different, randomly selected, sparse subsets of the entire set of local arrival times. Filled epicenters are those whose 90 % confidence ellipses have semi-axes less than 5 km. Large circle has 10 km radius and is centered on epicenter that is calculated with all data.