Improved earthquake locations and estimation of Pn and Sn path anomalies for India, using multiple event relocation and reference events

B. K. Rastogi[1,*], E. A. Bergman[2] and E. R. Engdahl[2]

[1]National Geophysical Research Institute, Hyderabad 500 007, India
[2]Center for Imaging the Earth's Interior, Department of Physics, University of Colorado, Boulder, CO 80309-0390, USA

We have combined local seismic network data and a groomed ISC/NEIC database of phase arrivals at regional and teleseismic distances to determine accurate hypocenters and origin times for 224 earthquakes that are clustered in three regions of India (Koyna, Chamoli–Uttarkashi and Kutch). Multiple event relocation of these clusters of events, combined with highly accurate reference event locations from local network data, is used to improve relative and absolute locations of all cluster events. The improved accuracy of the locations obtained for mainshocks, early aftershocks and other earthquakes in these regions offers insights into the regional seismotectonics. The observed arrival time data for these clusters, calibrated by the reference events, provide a valuable data set for estimating Pn and Sn travel time anomalies beneath India, which in turn are used to develop constraints on models of the crust and upper mantle for that region. In particular, we observe large (up to 12 seconds early) path anomalies for Sn on paths across the Indian shield, suggesting significantly higher shear velocities than predicted by the average global model ak135.

IN the well-known 'boot-strap' problem of seismology, progress in earthquake location is dependent on improvements in knowledge of earth structure, which depends on improved knowledge of the locations and origin times of seismic sources. Tomographic methods of investigating earth structure have reached limits imposed by the inherent level of accuracy in standard global earthquake catalogs, even those especially groomed for such studies[1]. One area in which most progress has been made recently is in earthquake location, where methods have been developed to significantly improve location accuracy using 'ground truth' or 'reference' events[2]. The main challenge to such a strategy is that reference events are difficult to establish. In this article, we present the results of earthquake location studies in three regions of India for which reference events are available. We use a multiple event location method to co-locate other earthquakes in these areas with the reference events. In addition to improved absolute and relative locations of clustered events which are of great value in seismotectonic studies of these earthquake sequences, these studies yield reliable estimates of absolute travel times at regional and teleseismic distances that will help in developing and testing 3-D models of the crust and upper mantle in the region.

Since the 1993 Latur earthquake, monitoring of earthquakes in the stable continental region (SCR) of India or Peninsular India has considerably improved with the installation of over twenty three-component digital stations. The location errors in the interior SCR are now often thought to be less than 10 km. However, outside the network or near its outer margin, the locations could be biased by 30 km or more. Locations made with poor network geometry are highly susceptible to bias from inadequate velocity models, and this problem is especially severe for regional networks. A fuller understanding of the seismotectonics and seismic hazard of India is severely compromised by the much poorer location accuracy for events prior to the late 1990s. This study shows one way in which these problems may be overcome.

Methodology

Our methodology for improving earthquake location accuracy depends on the analysis of clusters of moderate-sized earthquakes, using a multiple event relocation technique to obtain improved relative locations and arrival time data at regional and teleseismic distances. We also exploit 'reference' earthquakes in the clusters that have been well located by local networks to obtain highly accurate absolute locations and origin times. In order to tie the reference events to the remaining events in the clusters, it is necessary that the reference events be well recorded at regional and teleseismic distances.

We use the Hypocentroidal Decomposition (HDC)[3] method of multiple event relocation to refine the relative locations and origin times of all events in the cluster. This method has been used extensively for similar studies in recent years and is well-suited for such studies for several reasons. The method is mathematically rigorous and uses all available information in the data set on relative locations of events;

*For correspondence. (e-mail: brastogi@ngri.res.in)

no master events are used and no data discarded. All statistical properties of the data are carried through the analysis in a rigorous manner. Relative to other methods of multiple event relocation, HDC is computationally efficient, requiring the solution of smaller matrices. HDC naturally separates the location problem into two distinct parts, first estimating the relative locations and origin times of events in the cluster, and then estimating the absolute location in space and time (geographical coordinates and UT) in a procedure analogous to standard single event location. This second step is where most of the problems from un-modeled lateral heterogeneity are encountered, and where we make use of reference events to overcome the resulting bias.

To gauge the level of accuracy of reference event locations from local network data, we use the following criteria[4]:

GT5 at the 95% confidence level should have at least 10 stations, all within 250 km, with azimuth gap ≤ 110 degrees, secondary azimuth gap of ≤ 160 degrees and at least one station within 30 km.

It is also necessary that a ground truth event be at a shallow depth (< 35 km) and be recorded at least at regional distances (typically, magnitude > 3.5) so that it can be co-located with the other cluster events. The secondary azimuth gap is defined as the largest azimuth gap filled by a single station.

The events in the cluster may be widely distributed in time, as long as arrival time data at common stations are available. We assemble the clusters from the EHB catalog[1] that has already been closely scrutinized, and additional grooming is done to ensure accurate source depths and phase associations. The clusters typically are 50–100 km across and comprised up to 100 events of magnitude 3.5 or greater that have occurred since 1964 and that are well recorded at regional and teleseismic distances. For the HDC analysis we hold depths fixed at the depths determined in the EHB analysis because the regional and teleseismic data do not provide useful constraints on depth. Tests have shown that errors in assumed depth of up to several tens of kilometers have only a small effect on the accuracy of the epicenter[4].

The HDC analyses produce new locations that are defined by 'cluster vectors' in space and time coordinates relative to the hypocentroid (the geometrical mean of the cluster vectors), which is then located in the traditional manner to yield absolute locations and origin times. The relative locations defined by the cluster vectors are significantly more accurate than those defined by single event locations because the biasing effect of un-modeled path anomalies at regional and teleseismic distances is largely removed.

An important part of this analysis is to use the multiple observations of phases from many cluster events at a single station to remove outliers in the arrival time data. Normally outliers can only be identified by having large absolute residuals against the travel time model used for location, but path anomalies are largely removed in the HDC analysis and we can examine the scatter of arrivals that is due only to reading error. In the process of identifying and removing such outliers, we obtain reliable empirical estimates of reading error for each station/phase combination, which is used to improve the statistical rigor of the analysis. Reading errors so estimated are typically in the range 0.3–0.5 s for P phases, 0.5–1.0 s for Pn phases, and up to 2.0 s for S phases.

The hypocentroid of the cluster is located in an absolute sense, as if all the arrival time data were from a single event, using the 1-D model ak135 (ref. 5). Obviously, this process is subject to location bias, because of departures of the real Earth from the 1-D model used for location. To remove the bias, we shift the hypocentroid in space and time to provide an optimal match on average to the reference events that are included in the cluster. For each cluster event for which a local network location is available, we calculate the shift vector in space and time between the reference (local network) hypocenter and the HDC-derived hypocenter for that event. These individual shift vectors are averaged to obtain the cluster shift vector, which is then applied to the cluster as a whole. Through this process all events in the cluster have absolute locations and origin times largely devoid of systematic bias. Shifts in epicenter and origin time (to best match the reference event locations) are typically in the range of 5–15 km and ± 2 seconds, respectively.

The degree of consistency between the relative locations as determined by regional and teleseismic arrival time data in the HDC analysis, and the relative locations specified by the reference event data is an important test for validating candidate reference events. Discrepancies may be resolved by determining that the cluster vector is biased in the HDC analysis for some reason, or by rejecting the candidate reference event. For this reason, it is very valuable to have several reference events for a cluster.

Once the absolute locations and origin times of the cluster events have been calibrated by comparison with the reference events, the arrival time data at regional and teleseismic distances may be used to calculate absolute travel times for different phases. We make use of the repeated observations from many cluster events at a given station to make the estimate of absolute travel times using robust statistics to measure the mean and spread of the distribution, requiring a minimum of 5 observations with a spread ≤ 1.4 s for P phases, and ≤ 2.8 s for S phases. Individual travel times are calculated relative to the shifted (calibrated) cluster event locations, but the mean path anomaly is referenced to the hypocentroid of the cluster.

Indian clusters

We have studied earthquake clusters in three regions of India: Koyna, Chamoli–Uttarkashi and Kutch (Bhuj). The Koyna cluster is composed of moderate-sized events associated with induced seismicity that began in 1967 and continues today. The Chamoli cluster is composed mainly of the M_w 6.6 mainshock on 28 March 1999 and its aftershocks, plus the 1991 Uttarkashi mainshock and its aftershocks. The Bhuj cluster is composed mainly of the M_w 7.7 mainshock on 26 January 2001 and its aftershocks. For each cluster, some events have been accurately located with local networks of 3-component digital and vertical component analog stations that meet the criteria for GT5 accuracy noted above.

After the Chamoli and Bhuj earthquake sequences, aftershock monitoring began about a week after the mainshocks. During this early aftershock period, when there was no local monitoring, aftershocks are sometimes mislocated by 20–50 km using data from regional distance Indian stations and teleseismic stations. Regional phase data cannot be used to obtain accurate location unless a reliable 3-D model of the crust and upper mantle is available. Next we review the relocation studies for the three regions (Table 1).

Koyna

Induced earthquakes began occurring near the Koyna reservoir after its filling started in 1962. Seventeen earthquakes of magnitude ≥ 5 and over 150 events of magnitude ≥ 4 have been recorded from the area. However, due to inadequate station coverage, locations of the earlier events were not accurate enough to demarcate faults. To improve location accuracy around Koyna, NGRI began operating a network of up to 15 digital and 4 analog stations around the reservoir in 1993. From 1993 through 2000, ten moderate-sized earthquakes were located well enough from 3-dimensional tomography so that they could be treated as reference events. P-readings typically have an RMS residual of 0.01 to 0.1 s. Accuracy in the epicentral locations is ≤ 0.2 km and that of focal depth ≤ 0.5 km[6]. A 7-layered model that was obtained from deep seismic sounding in Koyna was used as an input 1-D model for the tomographic study. A search of the EHB catalog yielded a cluster of 31 earthquakes from the Koyna region that were well recorded at teleseismic distances since 1967. The pattern defined by the reference locations matched well with the pattern of the corresponding events in the HDC cluster analysis. The average shift required to align the HDC locations with the reference locations is 12.7 km at an azimuth of 232°, and an origin time shift of –0.29 s. This shift was applied before calculating travel times for all reported phases (minimum 5 observations) at each station.

Table 1. Basic parameters of some earthquake clusters

Cluster	Dates	Number of events	Reference events
Koyna	1967–2000	31	10
Chamoli–Uttarkashi	1967–2001	86	8
Bhuj	2000–2002	107	6

The Pn and Sn travel times (shown as path anomalies relative to ak135) from the centroid of the cluster are shown in Figure 1. We observe large negative (up to –4.9 s) Pn anomalies (shorter travel times, relative to ak135) on paths from the Koyna source area to stations in Nepal and elsewhere along the Himalayan front. Very large negative (up to –9.0 s) Sn anomalies (shorter travel times) are observed at stations in northern, eastern, and southern India.

Chamoli–Uttarkashi

Garhwal Himalaya is one of the more seismically active regions of India. Five earthquakes of $M \geq 6$ (including the 1803 earthquake of intensity IX or magnitude 8) and 12 earthquakes of magnitude 5 to 6, and many of magnitude 4 to 5 have occurred there. During the past three decades two large earthquakes have occurred, the M_w 6.8 Uttarkashi earthquake of 19 October 1991, and the M_w 6.6 Chamoli earthquake of 28 March 1999. However, due to the lack of regional seismic stations, the seismicity pattern around this region (in particular, the locations of the two mainshocks) is not well known.

After the Chamoli mainshock NGRI operated a temporary network of nine three-component digital stations from 3 April 1999 to 21 May 1999. The distance between stations and epicenters varied between 1.5 and 47 km. Sampling rates were 100 SPS or more. A velocity model and station corrections for the region were obtained by inverting the P-arrival times for 110 events. The locations are characterized by uncertainty in epicenter of < 1.5 km, uncertainty in depth of < 2 km and RMS residuals of < 0.4 s. The epicenters were located in an area of 10 km by 20 km, trending NW–SE. The depth sections suggest a northeast-dipping focal zone extending from a depth of 8 km to a depth of 16 km.

This network located 8 aftershocks with sufficient accuracy that they could be treated as reference events. A cluster of 86 events between 1967 and September 2001 was compiled from the EHB catalog, including events of the nearby Uttarkashi earthquake series. 56 of these clustered events, including the Uttarkashi and Chamoli mainshocks, are now located to better than 5 km accuracy and all but one of the remaining events have location accuracies of better than 10 km. Agreement between the HDC relative locations and the reference locations is excellent, from which we estimate a shift of 15.2 km at an azimuth of 343° and an origin time shift of –0.18 s. This shift was applied before calculating travel times for all reported phases (minimum 5 observations) at each station.

Travel times for Pn and Sn phases to many stations in the region were estimated (Figure 2) from the shifted epicenters. Relative to the Chamoli–Uttarkashi source area, which is along the Himalayan front, path anomalies for Pn at stations along the Himalayan front show small posi-

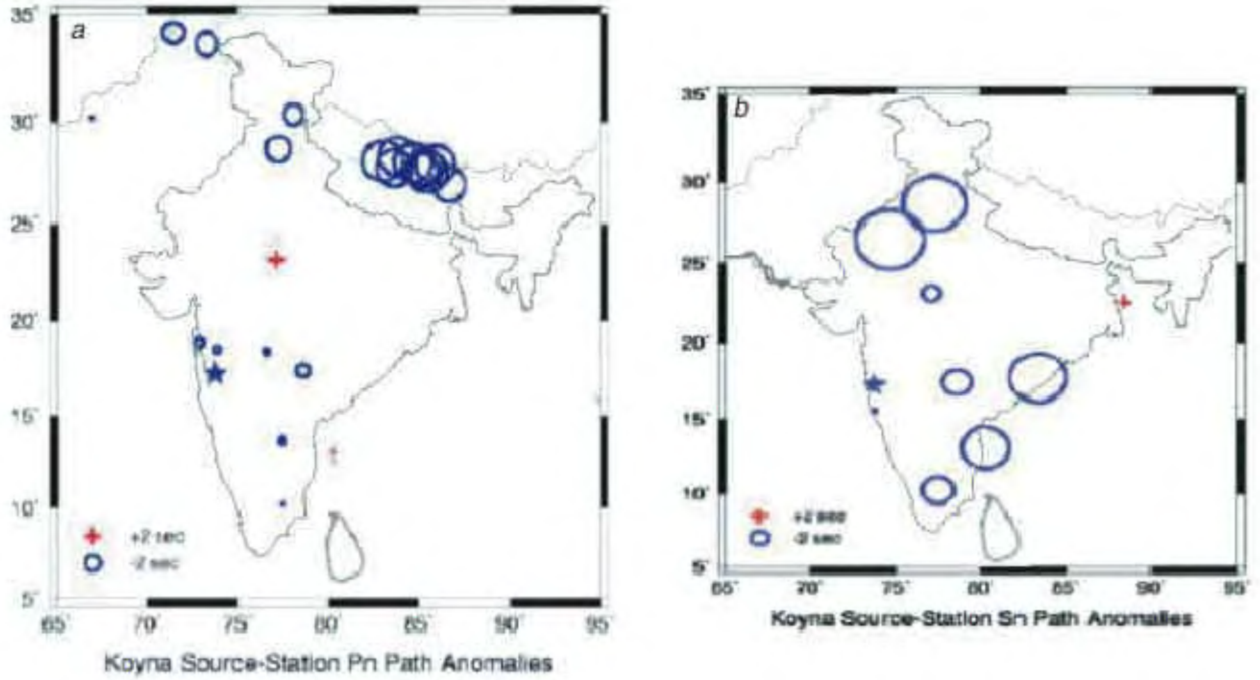

Figure 1. Map of the Indian region showing Pn (***a***) and Sn (***b***) source-station path anomalies from the Koyna cluster centroid (star). Anomalies are relative to ak135. Circles are early arrivals, pluses are late arrivals. Symbol size is proportional to the anomaly according to the scale shown. Anomalies are calculated as the median, using at least five observations, with spread ≤ .4 s for Pn and ≤ 2.8 s for Sn.

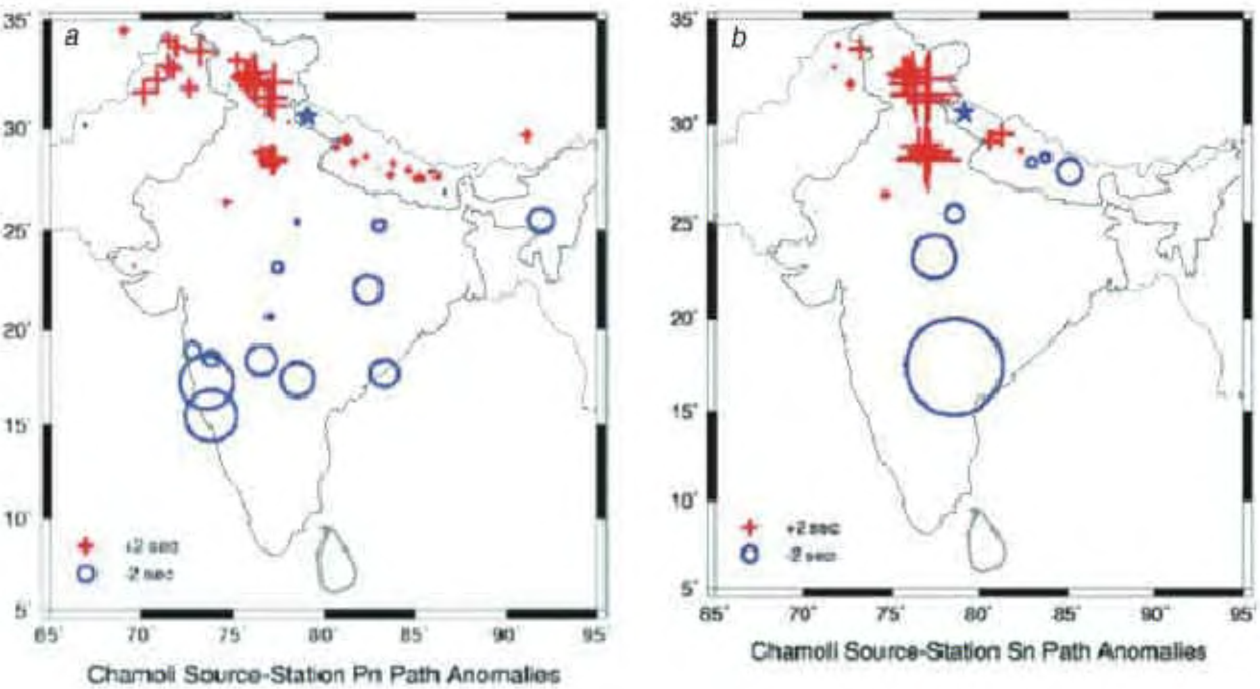

Figure 2. Map of the Indian region showing Pn (***a***) and Sn (***b***) source-station path anomalies from the Chamoli–Uttarkashi cluster centroid (star). See Figure 1.

tive anomalies, indicating longer travel times relative to ak135. Pn path anomalies to stations in central India show large negative anomalies (shorter travel times), consistent with the results from Koyna. The Sn path anomalies from Chamoli are even more striking, showing a huge negative anomaly (–12.2 s) at station HYB (Hyderabad) in south-

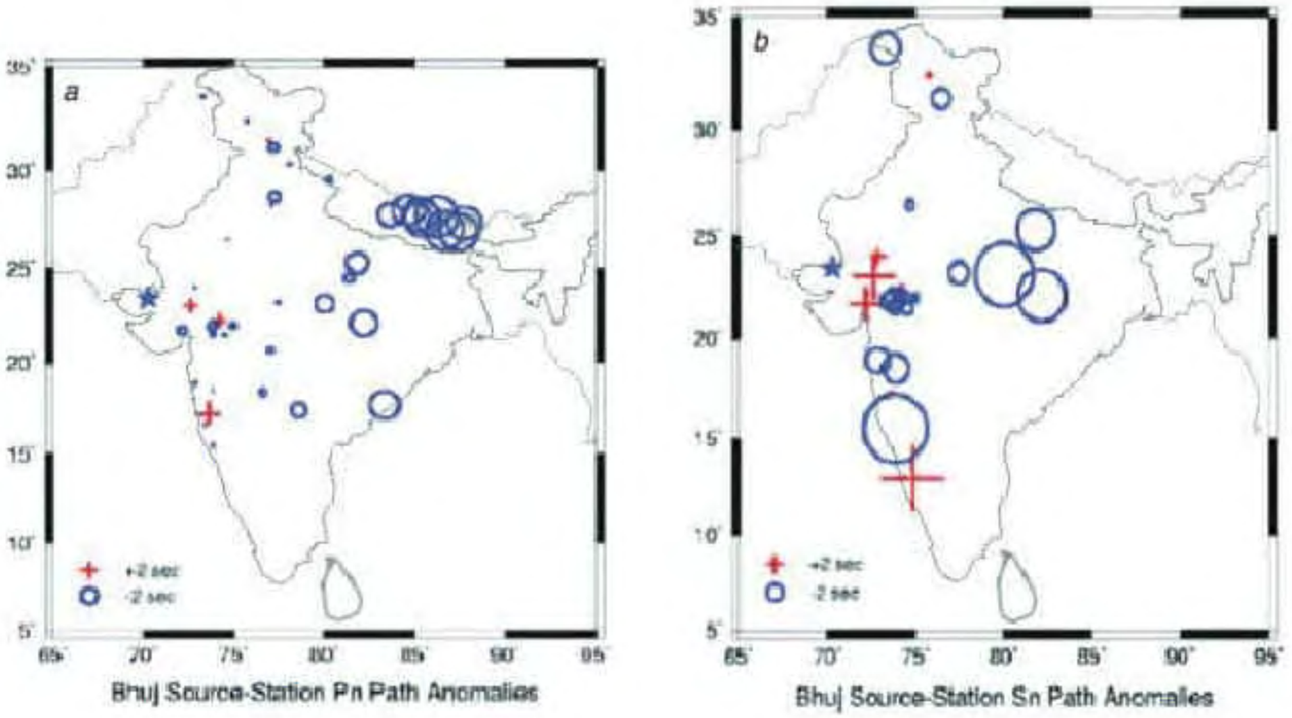

Figure 3. Map of the Indian region showing Pn (***a***) and Sn (***b***) source-station path anomalies from the Bhuj cluster centroid (star). See Figure. 1.

central India. Stations nearer to Chamoli in northern India show large positive anomalies, indicating longer travel times than predicted by ak135.

Bhuj

The M_w 7.7 earthquake in Bhuj on 26 January 2001 occurred in a region poorly covered by permanent seismic networks of India. Many temporary deployments of seismographs were made to locate aftershocks, but the earliest did not start for a week after the mainshock and therefore the locations of the mainshock and early aftershocks are not well known. Locations from an 8-station aftershock deployment by Memphis University, which are considered accurate to within 1–2 km (pers. commun., Arch Johnston), have been made available for this study. However, integration of arrival time data from other deployments and permanent network stations operated by Indian institutions, as well as improvements to the local velocity model, need to be made before final reference event locations can be determined. The results reported here are based on 6 events from the original Memphis data set that were recorded well enough at regional and teleseismic distances to be treated as reference events.

The cluster of 107 earthquakes includes one event a month before the mainshock (24 December 2000), the mainshock itself, and the larger aftershocks up to March 2002. With respect to the reference events, HDC locations were offset (biased) in a very systematic manner, allowing the entire cluster to be calibrated by a shift of 6.2 km at an azimuth of 341° and an origin time shift of 0.36 s. In all, 62 of these locations have accuracies of 5 km or better, and all the remaining events are located to better than 10 km. Further improvement can be expected with additional data, especially regional data from India and nearby countries. This cluster has a wide range of depths (8–30 km), confirmed not only by local network hypocenters but by teleseismic depth phases as well. Hence, a further enhancement to the processing would be to assign (for events which were set at an optimal depth of 18 km) more appropriate depths as indicated by the distribution in space of the local network hypocenters.

The Pn and Sn travel times from the shifted cluster are shown in Figure 3. The path anomalies are similar to those from Koyna, showing large negative Pn anomalies (shorter travel times, relative to ak135) at stations in Nepal. We see mostly large negative path anomalies for Sn at stations in central and southern India. The large positive anomaly for Sn at station Mangalore (MNGI) in southern India warrants further investigation. It may be caused by a phase association error.

Conclusion

The combination of high-precision multiple event relocation analysis and reference events is an important new tool for studying the seismotectonics in regions of high seismicity. We have used this approach to obtain accurate locations for 224 of the larger earthquakes in four important earthquake sequences in India. In particular, we have obtained accu-

Table 2. Mislocation vectors for some earthquake clusters

Cluster	Distance (km)	Azimuth	Origin time (s)
Koyna	12.7	232	–0.29
Chamoli	15.2	343	–0.18
Bhuj	6.2	341	0.36

rate locations of the mainshocks in the 2001 Bhuj, 1999 Chamoli, and 1991 Uttarkashi sequences, with respect to the aftershocks. The calibration of these epicenters using reference events makes possible an association of specific earthquakes with mapped faults and other geological structures. Absolute locations of the 1991 Uttarkashi sequence were determined by joint inversion with data for the nearby 1999 Chamoli earthquake, for which reference events were available.

Through HDC analysis and the use of reference events, we have located all of the 224 larger events of these sequences with a bias of no more than 1–2 km and 90% confidence ellipses of less than 10 km (semi-major axis). Over half of the events are located with accuracies better than 5 km, and the largest events, including the Chamoli, Uttarkashi and Bhuj mainshocks, are located with accuracies on the order of 1.5–3.5 km (90% confidence ellipse semi-major axis).

For each cluster a mislocation vector is calculated which represents the shift in epicenter and origin time of the HDC solutions needed to achieve the best agreement with the reference events that are included in the HDC clusters.

The shifts needed to bring the HDC locations into alignment with the reference event locations are of course related to the departures of the real Earth from the assumed 1-D earth model ak135 that is used to estimate absolute locations in the HDC analysis. It would be incorrect to assume that these shifts have general applicability as 'correction factors' for other earthquakes in these regions, because the shifts are sensitive to the particular data sets used in the HDC analysis. For example, if we had not used regional data in the HDC analysis, the mislocation vectors would be different. For the same reason it is not useful to compare closely the mislocation vectors of different clusters. Their main utility is to remove location bias from the corresponding cluster so that unbiased estimates of true travel times may be calculated. We have done this, using a robust statistical method, for Pn, P, Sn, and S phases observed from these clusters.

Upper-mantle travel-time anomalies, revealed by Pn and Sn arrival time data from these calibrated clusters, offer very strong signatures of lateral heterogeneity in the Indian subcontinent, characterized by early arrivals in the stable continental shield and late arrivals in the Himalayas, relative to a standard 1-D global model, ak135. Observed arrival times of Sn phases from Chamoli, at the base of the Himalayas, to stations in the stable continental interior of India are as much as 12 seconds earlier than predicted by ak135. Pn path anomalies range between +4.9 and –6.5 seconds. The upper mantle velocity structure under India has significant departures from the average global model ak135. In general, the stable continental shield is characterized by faster velocities than ak135 and the Himalayan front by slower velocities, but there are strong variations over smaller distances that severely compromise earthquake location efforts using regional distance (250–2000 km) data. The only solution is to better understand the details of these variations and incorporate such knowledge into earthquake location algorithms used to routinely locate earthquakes in this region.

With increasing numbers of high-quality local and regional seismic networks in India, and the on-going high rate of seismicity, we expect that further studies of this kind will be possible, allowing a very detailed map of the upper mantle velocity structure in the Indian subcontinent to be constructed.

1. Engdahl, E. R., Van der Hilst, R. D. and Buland, R. P., Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bull. Seismol. Soc. Am.*, 1998, **88**, 722–743.
2. Ritzwoller, M. H., Shapiro, N. M., Levshin, A. L., Bergman, E. A. and Engdahl, E. R., The ability of a global 3-D model to locate regional events. *J. Geophys. Res.*, 2003, **108**, ESE9-1–ESE9-24.
3. Jordan, T. H. and Sverdrup, K. A., Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific. *Bull. Seismol. Soc. Am.*, 1981, **71**, 1105–1130.
4. Bondár, I., Myers, S. C., Engdahl, E. R. and Bergman, E. A., Epicenter accuracy based on seismic network criteria. *Geophys. J. Int.*, 2004 (in press).
5. Kennett, B. L. N., Engdahl, E. R. and Buland, R. P., Constraints on seismic velocities in the Earth from travel times. *Geophys. J. Int.*, 1995, **122**, 108–124.
6. Mandal, P., Rastogi, B. K. and Gupta, H. K., Recent Indian earthquakes. *Curr. Sci.*, 2000, **79**, 1334–1346.