Bulletin of the Seismological Society of America, Vol. 102, No. 5, pp. 2239–2244, October 2012, doi: 10.1785/0120120010

*Short Note*

# Waveform Relocated Earthquake Catalog for Southern California (1981 to June 2011)

by Egill Hauksson, Wenzheng Yang, and Peter M. Shearer

**Abstract** We determine a new relocated catalog, HYS_catalog_2011, for southern California from 1981 through June 2011. About 75.3% of the hypocenters are calculated with absolute and differential travel-time picks, and 24.7% could be relocated only by using absolute travel-time picks with 3D or 1D velocity models. The total catalog consists of more than 502,000 earthquakes in the region extending from Baja California in the south to Coalinga and Owens Valley in the north. The catalog consists of three *M* 7.1, *M* 7.2, and *M* 7.3 mainshocks; their foreshocks and aftershocks; and background seismicity caused by tectonic and other processes in the southern California crust. Hypocenters in the new relocated catalog exhibit tighter spatial clustering of seismicity than does the routinely generated catalog, and the depth distribution is tighter and reflects the thickness of the seismogenic zone more accurately. Compared to the standard catalog, the relocated hypocenters are more easily related to other data sets, such as mapped late Quaternary faults.

## Introduction

We have relocated the southern California seismicity recorded by the Southern California Seismic Network (SCSN) from 1981 through June 2011. First, we applied both 1D and 3D velocity models from Hauksson (2000) to improve the SCSN catalog locations using the methods of Klein (2002) and Thurber (1993). Second, we applied an approach similar to that of Hauksson and Shearer (2005), Shearer *et al.* (2005), and Lin *et al.* (2007) by clustering the events and using the differential travel times to perform relative relocations within each cluster. The resulting catalog, HYS_catalog_2011, contains more than 502,000 earthquakes, including 386,277 events that were relocated using differential travel times. The magnitude range for the SCSN catalog is from *M* 0.0 to 7.3, the largest event being the 1992 Landers earthquake (Hutton *et al.*, 2010).

The quality of each calculated hypocenter can vary depending on data availability at the time the earthquake occurred. For instance, if the background seismic noise was high or if nearby stations were not reporting, the data quality may be insufficient to allow cross correlation, and in rare cases relocation using a 3D model is not possible. Similarly, variability in the focal mechanisms or low event density may affect the quality of the cross correlation. The 3D velocity model accounts for most of the travel-time anomalies that are not included in a 1D model and significantly improves the event depth distribution. The differential travel times are used to improve the relative location accuracy of spatially clustered hypocenters.

Similar earthquake relocations have been carried out by Waldhauser and Schaff (2008) in northern California. They showed that 90% of the earthquakes in northern California could be successfully correlated, and the alignment with late Quaternary faults was significantly improved. In a different study, Rubin *et al.* (1999), using data from both northern California and Hawaii, showed that relocated earthquakes often align into concentrated streaks in the direction of fault slip. These and many other studies demonstrate that the relocated seismicity can be applied to understanding both tectonic and seismic source processes.

To facilitate spatial-, temporal-, and magnitude-based studies of the catalog, we have included all of the events in the new catalog and made it available from the Web site provided in the Data and Resources section. The relocated hypocenters include both absolute and relative uncertainties, which seismologists prefer to refer to as absolute or relative horizontal and depth errors. The absolute errors refer to the estimated distance between the true location and a corresponding calculated hypocenter. The relative errors refer to the distance between nearby hypocenters in the cluster sizes analyzed in this catalog, which have 2.5-km maximum separations. Usually the relative locations of hypocenters can be determined more accurately than the absolute locations.

Compared to the standard catalog, the relocated hypocenters are more easily related to other data sets, such as mapped late Quaternary faults in southern California (Hauksson, 2011), and they are also better suited to carry out seismicity studies such as the temporal and spatial evolution of seismicity clusters (e.g., Vidale and Shearer, 2006).

## SCSN Recorded Data and Refining Methods

The SCSN recorded data from more than 240 seismic stations in 1981 and has grown to more than 400 stations in 2011. The SCSN station coordinates and thus the hypocenters are provided in the latest revision of the World Geodetic System (WGS-84) reference ellipsoid. Hutton *et al.* (2010) provide the details of the network and the routine data processing. As part of standard processing, data analysts review all of the arrival-time picks and relocate each earthquake using a 1D velocity model. We have used these arrival-time picks, determined by the SCSN data analysts, to relocate the earthquakes using a 3D velocity model from Hauksson (2000).

We downloaded the digital seismograms for all of the events from the Southern California Earthquake Data Center (SCEDC; see the Data and Resources section). We reformatted the seismograms, resampled the data at 100 samples/second, and filtered them between 1 and 10 Hz before performing the cross correlations. We selected *P*-wave windows of 0.5 s and *S*-wave windows of 1.5-s duration. For the cross-correlation calculation, we included up to 150 nearest neighbors and required that pairs of events be separated by no more than 2.5 km. If fewer than 150 nearest neighbors existed within 2.5 km, we used Delaunay tessellation to add up to 150 more distant events to each cluster (Richards-Dinger and Shearer, 2000). To define a similar event pair, we required that at least eight cross-correlation coefficients be larger than 0.6 and the minimum average of the maximum cross-correlation coefficients be larger than 0.4. We attempted to calculate 54 million cross correlations, of which about 24 million achieved the minimum required values for a similar event pair.

To calculate the relocated hypocenters using multicore central processing units, we divided southern California into five polygons as used by Hauksson and Shearer (2005). Depending on what polygon the events in the HYS_catalog_2011 are located, they are labeled Poly1 through Poly5. First, we applied the clustering method by Shearer *et al.* (2005) and Lin *et al.* (2007) to identify clusters of events correlated to each other. Second, we used the cross-correlation differential times to relocate the 386,277 events that group into about 5000 different clusters. About 95% of the events are located within the SCSN reporting area and are labeled le for local events. Events that are located outside the SCSN reporting area are labeled in the HYS_catalog_2011 re for regional events. In general, more data are associated with the le events, and their hypocenters are better constrained. Catalog users may prefer to use data from northern California for events located north of the SCSN reporting area.

We have assembled a complete catalog of both local and regional events as processed by the SCSN. The preferred hypocenter (labeled ct in the HYS_catalog_2011) is the one determined with travel-time picks and differential travel times. If a differential travel-time hypocenter is not available, we prefer the hypocenter determined with the 3D velocity model (labeled 3d in the catalog). If the 3D velocity model is not available, we insert a hypocenter determined with HYPOINVERSE-2000 (Klein, 2002) and a 1D velocity model (labeled 1d in the catalog) or the SCSN catalog hypocenter (labeled xx in the catalog).

For the local events, about 75.3% of the hypocenters are determined with differential travel times, and 24.3% could be relocated only by using a 3D velocity model. About 1400 events (or 0.3%) could be relocated only with a 1D model, and about 800 catalog events (or 0.1%) could not be relocated because of insufficient data.

### Relocated Catalog

The most accurate hypocenters of similar event clusters (labeled ct in the HYS_catalog_2011) are plotted as black dots in Figure 1. In numerous places the events form elongated spatial clusters with high aspect ratios. These clusters often coincide with late Quaternary faults or form alignments that are suggestive of unmapped faults, sometimes referred to as seismic zones. The three other categories of hypocenters consist of those determined by using a 3D velocity model (3d), a 1D velocity model (1d), and the events that are not relocated (xx). The hypocenters of the 3d events are not as well constrained but are scattered throughout the region. In most cases the 1d or xx events have few arrival times because they may be small, or their waveforms may be complex, caused by overprinting of other events that occurred closely in time. The 1d and xx events are either early aftershocks or located mostly along the edges of the network in eastern California (including the Mammoth Lakes area), the Coalinga region (central California), and the southwestern Baja California region.

The waveform cross correlation is more effective for small events than for large ones. We have selected a subset of the catalog containing only le events within the network between latitudes of 32° N and 36.1° N to illustrate the magnitude range in which the cross correlation is most effective (Fig. 2). For events of magnitude up to $M$ 4.1, the number of events that qualify for cross correlation is consistently larger than the number of events located only with the 3D or 1D velocity models or nonrelocated xx events. Above $M$ 4.1, the number of cross-correlated events is similar to the number of 3d, 1d, and xx events. For still larger events of $M \geq 5.1$, the number of cross-correlated events is smaller than the number of 3d, 1d, and xx events. As expected, for $M \geq 6$ events there is no cross correlation. Thus, finite source effects, as reflected in the larger width of the *P*- or *S*-wave pulses, appear to affect the cross correlations of only a small subset of the catalog or the largest events of $M \geq 4.1$.

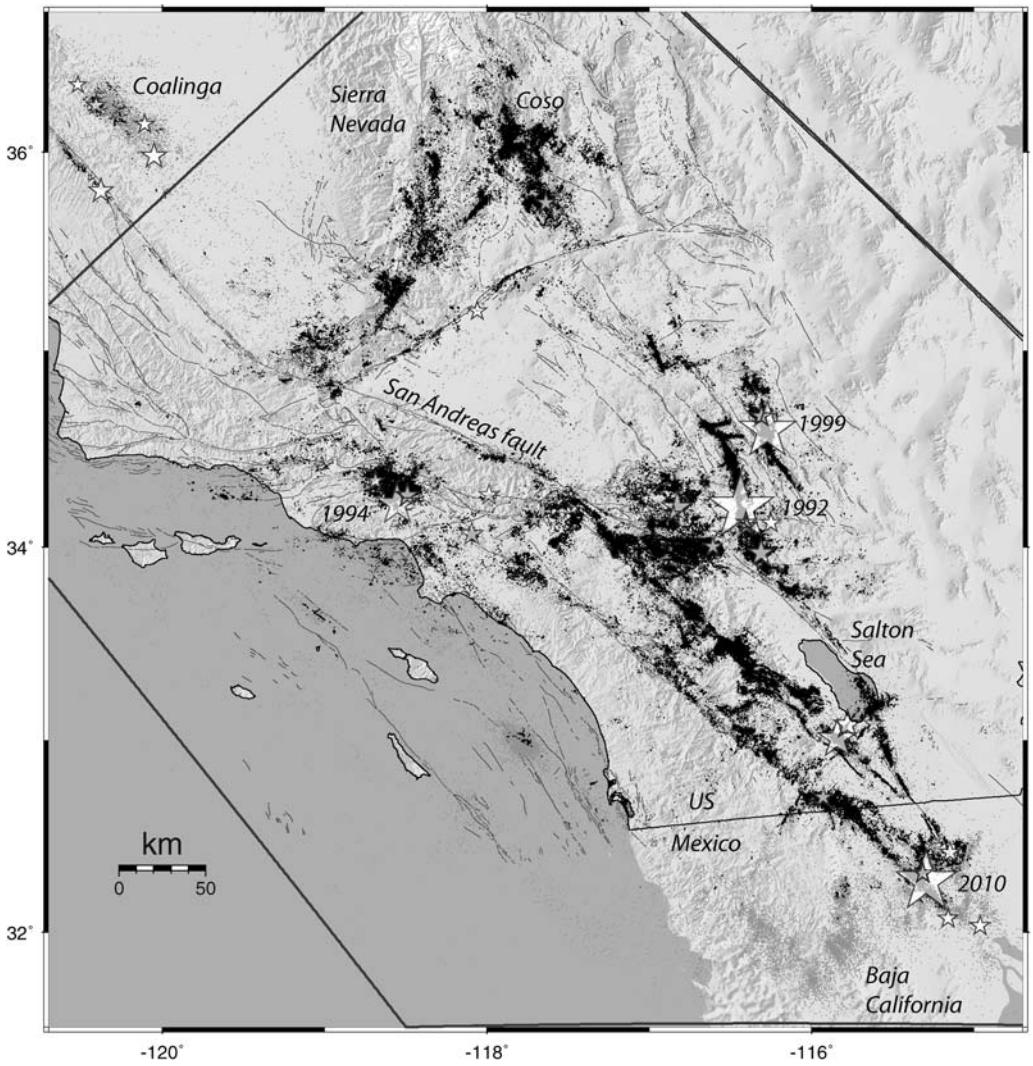


**Figure 1.** Southern California seismicity from 1981 through June 2011. The waveform-relocated epicenters are shown as black dots, the nonrelocated seismicity (events of type 3d, 1d, and xx) is shown as light gray dots, and earthquakes of $M \geq 5.5$ are shown as stars. Quaternary faults are depicted as light gray curves. The large earthquake sequences are indicated by year: 1992 $M_w$ 7.3 Landers; 1994 $M_w$ 6.7 Northridge; 1999 $M_w$ 7.1 Hector Mine; 2010 $M_w$ 7.2 El Mayor–Cucapah. The polygon is the SCSN reporting area for local events. Events located outside this polygon are called regional events. US, United States of America.

## Depth Distribution of Hypocenters

The earthquake depths are referred to average sea level and extend from 0.8 km above sea level down to about 20 km, with a few events as deep as ~30 km. About 2900 events are located above sea level and are assigned negative depths, which is permissible for shallow events located within a mountain range. A negative depth of −0.55 km, however, is the minimum depth allowed in the 3D-model location procedure. In most cases we interpret the negative depths to be very shallow and thus not well constrained unless there is a station located within a distance less than twice the actual focal depth. This lack-of-depth constraint is also reflected in the large absolute depth error. Previously, in the Lin *et al.* (2007) catalog, events with negative depth were purged before distributing the catalog, but here they are included for completeness.

The depth distribution of the seismicity reflects the thickness of the brittle part of the crust (Fig. 3). The average depth histogram peaks in the depth range of 3–7 km, with median depth at 7.2 km. The number of earthquakes increases rapidly with depth from the surface until the median depth is reached. Below ~10 km there is a rapid decrease in the number of events, followed by a second rapid decrease at ~17 km. Only a very few earthquakes are deeper than 20 km.

We have parsed out two magnitude ranges to illustrate that the smallest earthquakes tend to be shallow whereas earthquakes larger than $M$ 4.0 tend to be deeper. In general, deep events are more easily detected because $Q$ increases rapidly with depth (Hauksson and Shearer, 2006). Because there are fewer larger earthquakes, the average for the whole data set coincides more with the distribution of the smaller earthquakes. The small earthquakes exhibit a primary peak at

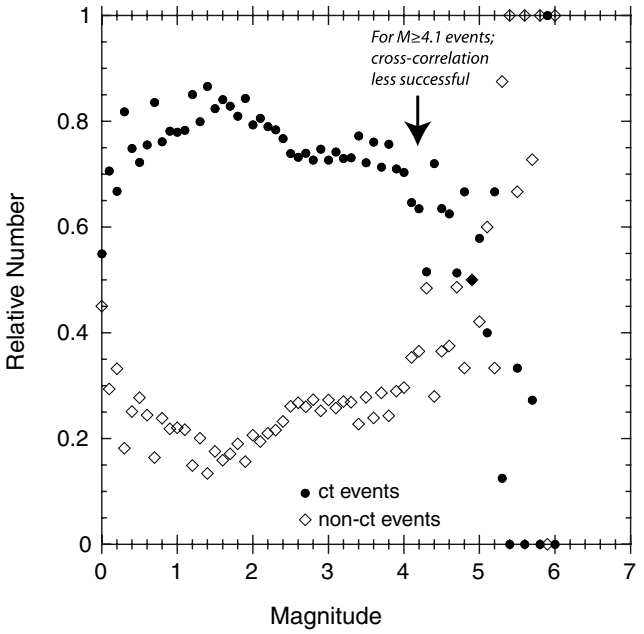


**Figure 2.** The relative number of ct events (solid circles, events relocated using waveform cross correlation) located with cross correlation and non-ct events (open diamonds, events located with other methods such as 3D or 1D models) plotted versus magnitude. All local events between 32° and 36.1° are included.

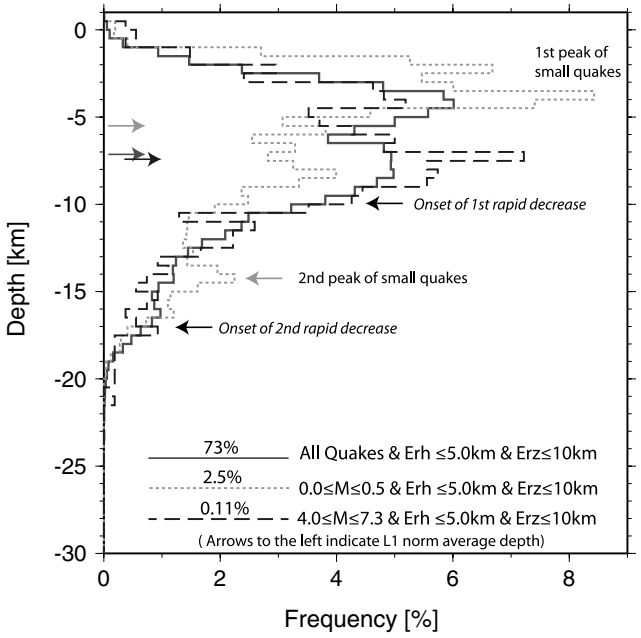


**Figure 3.** The average frequency–depth distribution of seismicity in southern California for hypocenters with horizontal errors ≤ 5.0 km and depth errors ≤ 10 km. We show the depth distribution of the whole data set (solid black curve), for only large earthquakes $M \geq 4.0$ (long dashes), and for small earthquakes of $M \leq 0.5$ (dotted curve). The percentages of the total catalog that are represented in each of the histograms after the error filtering are 73%, 0.11%, and 2.5%, respectively. The median depths are indicated to the left with arrows; top arrow is for $M \leq 0.5$ quakes; middle arrow is for whole data set; and bottom arrow is for $M \geq 4.0$ quakes (Erh, horizontal error estimate; Erz, vertical error estimate; $M$, magnitude).

~3 km depth and a secondary peak at 14 km depth near the bottom of the seismogenic crust. In the vicinity of the San Jacinto fault, small earthquakes are common in the depth range from 13 to 17 km. The detection capability of the network is very good along the San Jacinto fault because of its high density of stations and hard-rock ground conditions.

The depth distribution of large earthquakes also averages ~7 km, but the distribution is narrower, in the range of ~7 to 9 km, than the whole distribution. The relative absence of large earthquakes at shallow depth suggests that larger events on average tend to nucleate at greater depths than the smallest earthquakes (Nazareth and Hauksson, 2004).

### Absolute and Relative Relocation Errors

The location errors provided in the new catalog are one-sigma errors. The absolute horizontal and depth errors were determined using HYPOINVERSE (Klein, 2002) or SIMULPS (Thurber, 1983), which uses a similar routine for the calculation. The vertical errors are expected to be somewhat larger than the horizontal errors, depending on the station spacing and availability of analyst-reviewed *P* and *S* picks. The availability of *P* and *S* picks varies through time, but in 2010 there were about 900,000 picks available, of which 40% were *S* picks. The cumulative plots of the absolute and relative errors in Figure 4 show that 90% of the absolute horizontal errors are less than 0.75 km and 90% of the vertical errors are less than 1.25 km. Both Hauksson (2000) and Lin *et al.* (2007) tested the quality of the absolute locations using quarry blasts and were able to demonstrate that the size of the absolute errors are reasonable when compared to the true locations of the quarries.

The relative errors that were determined using the method described by Lin *et al.* (2007) are an order of magnitude smaller than the absolute errors (Fig. 4). About 90% of the relative errors are smaller than 0.1 km. Thus, the data set as a whole has reasonable accuracy.

## Discussion

The Pacific–North America plate tectonic deformation is the main process that causes small and large earthquakes in southern California. Secondary processes such as geothermal exploitation, extensional gravitation collapse, or crustal delamination also cause ongoing background seismicity. The overall pattern of waveform-relocated seismicity is dominated by the plate tectonic deformation (Hauksson, 2011). It shows the familiar prominent features such as mainshock–aftershock sequences of the 1992 Landers, 1999 Hector Mine, 1994 Northridge, and 2010 El Mayor–Cucapah earthquakes superimposed on a high level of background seismicity near the late Quaternary faults. In contrast, the background seismicity along the San Andreas fault, capable

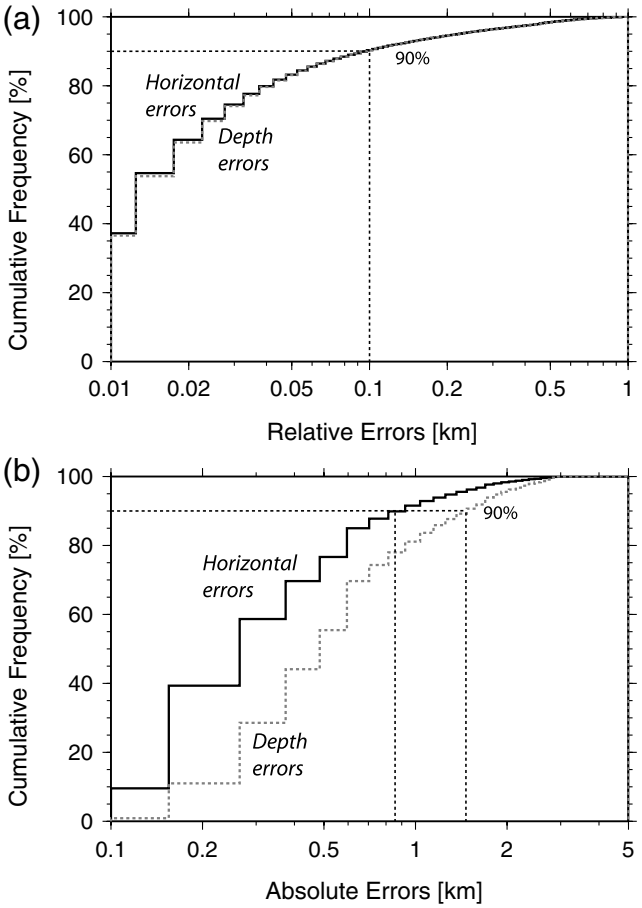


**Figure 4.** (a) Cumulative frequency of one-sigma relative horizontal (solid line) and depth errors (dashed line). (b) Cumulative frequency of absolute and depth errors for the whole catalog. The dashed vertical lines indicate the cumulative 90% values of the errors.

of generating the largest earthquakes, is diffuse, and the background rate is low.

For the first time, we are able to resolve new detailed seismicity patterns at the north end of the 2010 El Mayor–Cucapah aftershock zone, where the majority of the aftershocks occurred. The occurrence of aftershocks is not related simply to the principal slip surfaces of faults that accommodate the bulk of the slip. As two other examples, the 1992 Landers and 1999 Hector Mine earthquake sequences, also show, aftershocks exhibit heterogeneous spatial distributions, which may be related to heterogeneous mainshock stress release in the region. As mainshock slip transfers from one fault segment to the next, geometrical complexities such as step-overs or push-ups also contribute to complexity of aftershock distributions.

Adjacent to the U.S.–Mexico border, the background seismicity patterns exhibit interaction of northeast- and northwest-trending fault segments. Particularly in the Salton trough region, the relocated seismicity exhibits spatial distributions that are much tighter than what is observed with the routine catalog. These refined distributions often align with late Quaternary faults or other tectonic features. In some cases they suggest the presence of previously unmapped faults.

Seismicity in other regions, such as the southern Sierra, Banning Pass, and Coso regions and the Salton trough, is driven by secondary processes and often exhibits high levels of temporally clustered activity. As an example, the Coso and Salton Sea regions undergo swarmlike sequences, related to both tectonic and anthropogenic activities in the geothermal fields.

## Conclusions

The new relocated HYS_catalog_2011 of more than 502,000 events exhibits tighter spatial clustering of seismicity than the routinely generated catalog. The cross correlation works well for events of $M \leq 4.1$, but for larger events the effects of source finiteness make successful cross correlations less likely. In particular, the depth distribution is tighter and more accurately reflects the thickness of the seismogenic zone. The average absolute and relative one-sigma errors are less than 1.0 and 0.1 km, respectively. The catalog consists of three $M$ 7.1, $M$ 7.2, and $M$ 7.3 mainshocks, their foreshocks and aftershocks, and background seismicity caused by tectonic and other processes in the southern California crust.

## Data and Resources

We downloaded the digital seismograms for all of the events from the Southern California Earthquake Data Center (SCEDC), available at www.data.scec.org (last accessed June 2012). The relocated catalog is referred to as HYS_catalog_2011 and is available from http://www.data.scec.org/research-tools/alt-2011-dd-hauksson-yang-shearer.html (last accessed June 2012).

## Acknowledgments

This research was supported by the U.S. Geological Survey Grants G11AP20032 and G12AP20010 and by the Southern California Earthquake Center (SCEC), which is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number is 1528. The publication number of the Seismological Laboratory, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena is 10069. We are grateful to the operators and analysts who maintain the USGS/Caltech Southern California Seismic Network and who pick and archive the seismograms. Some of the computer codes used in this project were originally written and/or modified by G. Lin. We thank L. Jones and M. Boese for comments on the manuscript.

## References

Hauksson, E. (2000). Crustal structure and seismicity distributions adjacent to the Pacific and North America plate boundary in southern California, *J. Geophys. Res.* **105,** 13875–13903, doi 10.1029/2000JB900016.

Hauksson, E. (2011). Crustal geophysics and seismicity in southern California, *Geophys. J. Int.* **186,** 82–98, doi 10.1111/j.1365-246X.2011.05042.x.

Hauksson, E., and P. Shearer (2005). Southern California hypocenter relocation with waveform cross-correlation: Part 1, Results using

the double-difference method, *Bull. Seismol. Soc. Am.* **95,** 896–903, doi 10.1785/0120040167.

Hauksson, E., and P. Shearer (2006). Attenuation models ($Q_P$ and $Q_S$) in three dimensions of the southern California crust: Inferred fluid-saturation at seismogenic depths, *J. Geophys. Res.* **111,** no. B05302, doi 10.1029/2005JB003947.

Hutton, K., J. Woessner, and E. Hauksson (2010). Earthquake monitoring in southern California for seventy-seven years (1932–2008), *Bull. Seismol. Soc. Am.* **100,** 423–446, doi 10.1785/0120090130.

Klein, F. W. (2002). User's guide to HYPOINVERSE-2000, a Fortran program to solve for earthquake locations and magnitudes, *U. S. Geol. Surv. Open-File Rept. 02-171*, 123 pp.

Lin, G., P. M. Shearer, and E. Hauksson (2007). Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. Geophys. Res.* **112,** no. B12309, doi 10.1029/2007JB004986.

Nazareth, J. J., and E. Hauksson (2004). The seismogenic thickness of the southern California crust, *Bull. Seismol. Soc. Am.* **94,** 940–960, doi 10.1785/0120020129.

Richards-Dinger, K. B., and P. M. Shearer (2000). Earthquake locations in southern California obtained using source-specific station terms, *J. Geophys. Res.* **105,** no. B5, 10939–10960, doi 10.1029/2000JB900014.

Rubin, A. M., D. Gillard, and J.-L. Got (1999). Streaks of microearthquakes along creeping faults, *Nature* **400,** 635–641, doi 10.1038/23196.

Shearer, P., E. Hauksson, and G. Lin (2005). Southern California hypocenter relocation with waveform cross correlation: Part 2, Results using source-specific station terms and cluster analysis, *Bull. Seismol. Soc. Am.* **95,** 904–915, doi 10.1785/0120040168.

Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. Geophys. Res.* **88,** 8226–8236.

Thurber, C. H. (1993). Local earthquake tomography: Velocities and $V_P/V_S$ theory, in *Seismic Tomography: Theory and Practice*, H. M. Iyer and K. Hirahara (Editors), Chapman and Hall, London, 563–583.

Vidale, J. E., and P. M. Shearer (2006). A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers, *J. Geophys. Res.* **111,** no. B05312, doi 10.1029/2005JB004034.

Waldhauser, F., and D. P. Schaff (2008). Large-scale relocation of two decades of northern California seismicity using cross-correlation and double-difference methods, *J. Geophys. Res.* **113,** no. B08311, doi 10.1029/2007JB005479.

California Institute of Technology
Division of Geological and Planetary Sciences
Seismological Laboratory
Pasadena, California 91125
hauksson@caltech.edu
wenzheng@gps.caltech.edu
(E.H., W.Y.)

Institute of Geophysics and Planetary Physics
Scripps Institution of Oceanography
University of California at San Diego
La Jolla, California 92093-0225
pshearer@ucsd.edu
(P.M.S.)

Manuscript received 25 November 2011