Pure appl. geophys. 159 (2002) 563–596
0033 – 4553/02/030563 – 34 $ 1.50 + 0.20/0

Pure and Applied Geophysics

Using Automated, High-precision Repicking to Improve Delineation of Microseismic Structures at the Soultz Geothermal Reservoir

C. A. ROWE,*[1] R. C. ASTER,[1] W. S. PHILLIPS,[2] R. H. JONES,[3]
B. BORCHERS[4] and M. C. FEHLER[2]

Abstract—An automatic, adaptive, correlation-based algorithm for adjusting phase picks in large digital seismic data sets provides significant improvement in resolution of microseismic structures using only a small fraction of the time and manpower which would be required to re-analyze waveforms manually or semi-automatically. We apply this technique to induced seismicity at the Soultz-sous-Forêts geothermal site, France. The method is first applied to a small, previously manually repicked subset of the catalogue so that we may compare our results to those obtained from painstaking, visual, cross-correlation-based techniques. Relative centroid-adjusted hypocenters show a decrease in median mislocation from 31 to 7 m for preliminary and automatically adjusted picks, respectively, compared to the manual results. Narrow, intersecting joint features not observed in the preliminary hypocenter cloud, but revealed through manual repicking, are also recovered using the automatic method. We then address a larger catalogue of ~7000 microearthquakes. After relocating the events using automatic repicks, the percentage of events clustering within 5 m of their nearest neighbor increases form 5 to 26% of the catalogue. Hypocenter relocations delineate narrow, linear features previously obscured within the seismic cloud, interpreted as faults or fractures which may correspond to fluid propagation paths, or to changes in stress as a result of elevated pore pressures. RMS travel-time residuals for the larger data set are reduced by only 0.2%; however, phase-pick biases in the preliminary catalogue have influenced both the velocity model and station correction calculations, which will affect location residuals. These pick biases are apparent on the adjusted, stacked waveforms and correcting them will be important prior to future velocity model refinements.

Key words: Cross correlation, phase repicking, automatic algorithm, precise relocation, microearthquakes.

Introduction

The Soultz-sous-Forêts geothermal field is located in the Alsace region of eastern France, on the western edge of the Rhine Graben (Fig. 1), about 50 km

[1] Department of Earth and Environmental Science, New Mexico Tech, Socorro, NM 87801, USA.
[2] EES-4, MS D443, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
[3] ABB Offshore Systems Ltd., Rosemanowes, Penryn, Cornwall T10 9DU, United Kingdom.
[4] Department of Mathematics, New Mexico Tech, Socorro, NM 87801, U.S.A.
* Corresponding Author: C. A. Rowe, Now at Department of Geology and Geophysics, University of Wisconsin,1215 W. Dayton St. Madison, WI 53706, USA. E-mail: char@geology.wisc.edu

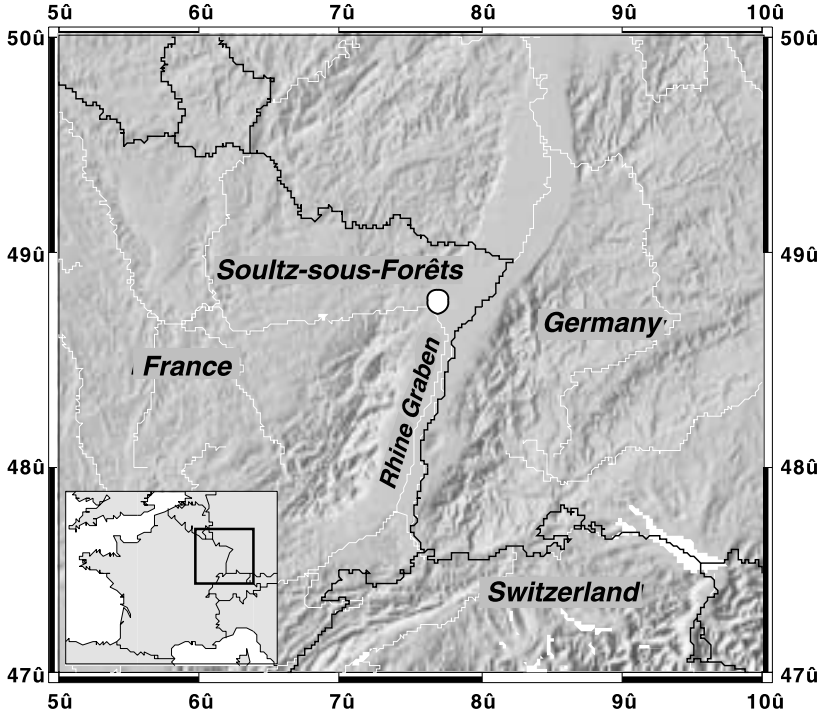


Figure 1
Location map for the Soultz-sous-Forêts reservoir (from PHILLIPS, 2000), situated on the western edge of the Rhine Graben in eastern France. This valley is an extensional rift exhibiting high geothermal heat flow.

north of Strasbourg (BARIA *et al.*, 1999). It is situated within the former Pechelbron oil field, in an area of high subsurface temperature gradients (BARIA *et al.*, 1999). The target geothermal reservoir resides within granitic rock between 2 to 4 km depth (BARIA *et al.*, 1999). This granite has been studied in detail using high-resolution borehole imaging and has been found to host two distinct populations of hydrothermally sealed fractures, striking N20°E and N30°W, respectively (GENTER and TRAINEAU, 1996). Stress estimates at Soultz indicate that the maximum horizontal stress within the reservoir is oriented roughly N10°W (JUNG, 1991; CORNET and JONES, 1994); hence many observed fractures strike subparallel to the maximum horizontal stress and are favorably oriented for hydraulic stimulation of the reservoir (BARIA *et al.*, 1999). Hydrofracturing at the reservoir in September and October, 1993 resulted in over 16,000 microearthquakes, recorded at three deep, down-hole multi-component seismometers and one down-hole hydrophone (Fig. 2). Numerous studies have been undertaken using Soultz microearthquakes to estimate growth of the newly fractured portion of the rockmass (JONES *et al.*, 1995), fracture orientations (e.g., GAUCHER *et al.*, 1998) or flowpaths (e.g., PHILLIPS, 2000; STARZEC *et al.*, 2000).

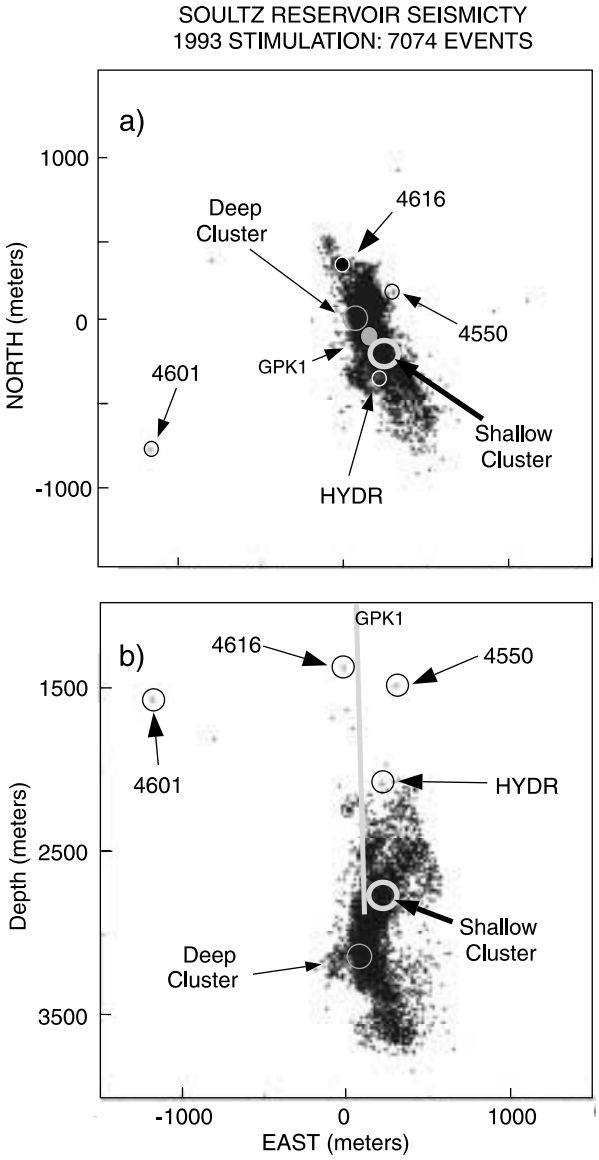


Figure 2
Soultz-sous-Forêts geothermal reservoir (after ASTER and ROWE, 2000). Shown are the four borehole sensors (multicomponent seismometers 4550, 4616 and 4601, and hydrophone HYDR) in both map view (a) and east–west cross section (b). The cloud of preliminary hypocenter locations is shown, along with the position of injection well GPK1. Gray circles indicate approximate locations for manually repicked clusters from PHILLIPS (2000).

Interpretation of seismogenic features depends critically upon the accuracy and interpretability of the source regions illuminated by the estimated hypocenters. In typical microearthquake studies at a local or regional scale, hypocenter mislocations of up to several kilometers may be tolerated, when only the identification of general seismic source regions and trends (and associated possible macroseismic effects) are the primary concern. Significant improvement in event locations for local and regional networks has been demonstrated through improved velocity modeling via joint hypocenter determination (JHD) (e.g., PUJOL, 1992) or local three-dimensional tomography (e.g., KISSLING, 1988; BLOCK, 1991; BLOCK *et al.*, 1994).

Earthquake location accuracy, however, depends not only upon the characterization of the seismic velocity structure through which seismic body waves have travelled, but also upon the ability to estimate accurately the arrival times of those waves. The most common human or computer picking approach is done one event at a time, for many recording stations. Slight to moderate picking inconsistencies are manifested ubiquitously at any particular station among even very similar events, resulting from variations in signal-to-noise ratio and other effects, as well as human subjectivity. Although one source of systematic mislocation can be successfully addressed through improved velocity modeling, fine details of the seismogenic structures are still commonly blurred by the intrinsic "fuzziness" of hypocenters arising from the remaining random pick errors and associated travel-time residuals (JONES and STEWART, 1997). The picking inconsistencies and resulting hypocentral scatter likewise affect other applications which rely upon the parametric data and associated error estimates. Such applications as reciprocal array processing (SPUDICH and BOSTWICK, 1987), spatial mapping of frequency-magnitude distribution (e.g., WIEMER and WYSS, 1997), and evolution of fractal dimension (e.g., NANJO *et al.*, 1998), as well as general understanding of spatio-temporal seismicity trends in many contexts, will benefit from correction of large numbers of picks. Certainly the reduction of pick scatter is crucial to mapping of very small-scale structures such as the fracture networks within a geothermal field (e.g., ROFF *et al.*, 1996; PHILLIPS *et al.*, 1997), repeating earthquakes (e.g., NADEAU *et al.*, 1995) or deep fault drilling (e.g., THURBER *et al.*, 2000), where precision to a scale of a few meters or less is desired.

A statistical approach to addressing hypocenter fuzziness caused by pick inconsistencies has been presented by JONES and STEWART (1997) and has been incorporated into JHD inversion by FEHLER *et al.* (2000). This technique, referred to as "collapsing," is based on assumptions of Gaussian pick errors and the observation that randomly perturbed points spread out in space. Hypocenters are permitted to move towards an event-based center of gravity, constrained by the dimensions of their error ellipsoids. Although collapsing can be prone to artifacts, synthetic tests as well as comparison to results from exhaustive, manual repicking of real data sets have demonstrated considerable success (e.g., JONES and STEWART, 1997; FEHLER *et al.*, 2000). This suggests that, in the absence of waveform data which might permit correction of the picks, collapsing, in combination with techniques for adjusting the

velocity model and/or near-surface static corrections, may be a good approach to improve delineation of seismogenic features.

It is possible, however, to address picking inconsistencies directly when digital waveform data are available, using signal processing methods to adjust the picks for consistency among events exhibiting waveform similarity (e.g., DODGE *et al.*, 1995; SHEARER, 1997, 1998; RUBIN *et al.*, 1998, 1999; ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001). Subsequent hypocenter relocation may then proceed using either single-event location, JHD, or simultaneous velocity and hypocenter inversion. Sharpening of the seismogenic source volume image has been demonstrated through relative relocation using waveform cross correlation techniques (e.g., FREMONT and MALONE, 1987; DEICHMANN and GARCIA-FERNANDEZ, 1992; GOT *et al.*, 1994; NADEAU *et al.*, 1995; GILLARD *et al.*, 1996; SHEARER, 1997, 1998; LEES, 1998; RUBIN *et al.*, 1998, 1999; WALDHAUSER *et al.*, 1999) or combinations of differential cross correlation and catalogue travel times (WALDHAUSER *et al.*, 1999). In these studies, typical errors of 1 km or more have been reduced to relative location errors of a few tens of meters or less, allowing previously undetected structures within the seismic source zone to be observed.

We have developed a new, automatic phase-repicking algorithm (ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001) that addresses the correction of scatter which arises from inconsistent picking, by using estimated relative pick lags to calculate corrected picks for consistency among similar waveforms (e.g., DODGE *et al.*, 1995) without imposing constraints from preliminary catalogue parameters on our corrected picks. This method does not address mislocations which may arise from unmodeled velocity variations; however, coupling the improved picks with better velocity models and more sophisticated location methods should further improve the accuracy of reservoir imaging in future work. This study presents the results of applying our repicking algorithm to seismicity recorded at the Soultz reservoir during September and October, 1993.

Data

The stimulated volume of the Soultz reservoir occupies roughly 0.48 km^3 in a depth range of about 2100 to 3800 m. Median distances from microearthquakes to the nearest and most distant receiver were approximately 800 m and 2100 m, respectively. Data were digitized at a sampling rate of 5 KHz. The borehole network consisted of three, 4-component (redundant) sensors and one hydrophone. Four-component data were projected into three orthogonal components prior to our analysis.

Location of the Soultz microearthquakes from preliminary analyst picking resulted in the seismicity cloud shown in Figure 2, consisting of hypocenters of $\sim$7000 events which will be used in our analysis. This subset of microearthquakes was chosen from the larger catalogue based on the availability of seven phase picks from

preliminary analysis. Unlike relocation efforts in large data sets recorded by large regional seismic networks (e.g., SHEARER, 1997, 1998; WALDHAUSER *et al.*, 1999), the Soultz data present significant additional location difficulties resulting from the sparse, four-station network. In the case of Soultz, all available phases are needed to adequately constrain hypocenters, and it was therefore not advisable to neglect arrivals which were noisy or nodal (PHILLIPS, 2000). Further, the sparse network geometry (Fig. 2) can cause asymmetric errors due to the location of the most distant station (4601), which provides much of the east-west location control.

Preliminary hypocenters determined using single event location (SEL) methods with initial arrival time picks and velocities determined by JUPE *et al.* (1994) are shown in Figure 2. These events were located with four *P*-wave and three *S*-wave arrivals (no *S* was available from the hydrophone channel, as *S* is generally unreadable there). A uniform *P*-wave velocity of 5.85 km/s and *S*-wave velocity of 3.34 km/s were used, consistent with data from a shot fired in well GPK1 at a depth of 3360 m. Individual station corrections determined from the calibration shot and applied for SEL hypocenters (JUPE *et al.*, 1994), and station corrections determined by a recent JHD inversion, are shown in Table 1.

Because detailed velocity investigation is beyond the scope of this study, we adopt the velocities and JHD station corrections determined by JUPE *et al.* (1994) and also used by PHILLIPS (2000) for the purpose of our comparative location analysis.

Previous Re-location Work at Soultz

GAUCHER *et al.* (1998) analyzed a small number of multiplets within the 1993 injection seismicity cloud. They obtained relative relocations defining structures near the borehole, which were correlated with fractures observed using acoustic imaging and well logs.

PHILLIPS (2000) relocated microearthquakes in two densely populated volumes within the seismic cloud by manually repicking relative arrival times of phases. The "shallow" and "deep" clusters, contained roughly 200 and 350 events, respectively. Approximate locations of these clusters are indicated by gray circles on Figure 2. Relocations from this manual repicking collapsed the diffuse clouds of seismicity into narrow, intersecting features which exhibited temporal seismicity patterns showing

Table 1

Station corrections (ms) for SEL and JHD Soultz microearthquake locations

Station	SEL *P*	SEL *S*	JHD *P*	JHD *S*
4550	0.0	0.0	−5.49	−8.25
4601	−17.0	−19.0	−16.4	−18.9
4616	−5.5	−7.0	0.5	1.32
HYDR	3.0	N/A	1.73	N/A

initial tight clusters or alignments of events which subsequently spread into adjacent areas. These early alignments in the sequence have been interpreted as joints or joint/fracture intersections that may represent fluid flow paths (PHILLIPS, 2000).

MORIYA *et al.* (2001) have addressed precise relative relocation of Soultz data through identification of highly similar earthquake doublets and multiplets. They have estimated relative hypocenter relocations using the cross-spectral method of POUPINET *et al.* (1984), and have observed numerous well-defined, narrow seismic lineations. MORIYA *et al.* (2001) further note that the high proportion of doublets and multiplets seen in the early period of the stimulation declines as time progresses. These observations are in good agreement with those of Phillips (2000) and Phillips *et al.* (2001), and we discuss them further, below.

Automatic Repicking Technique

We have developed a method which combines the most advantageous features of existing waveform cross-correlation-based phase adjustment methods (e.g., DODGE *et al.*, 1995; SHEARER, 1998; RUBIN *et al.*, 1998, 1999) in an automatic package which can be easily implemented even for very large (many thousands of events) data sets with a manageable amount of analyst interaction. Details of the signal processing, clustering and conjugate gradient techniques and our innovations are reported elsewhere (e.g., ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001); however, we summarize our method below.

Signal Processing Tools

Pick adjustments are estimated by comparing station-common waveforms having preliminary picks, and estimating relative lags between these picks. Lag estimation proceeds in two steps: A coarse integer-sample correlation, and a subsequent fine correlation step which provides a subsample relative lag refinement.

Coarse (integer-sample) cross-correlation provides initial lag adjustments for each waveform pair. For multicomponent data, we apply polarization filtering, derived from eigenvalue decomposition of a joint signal covariance matrix, to improve P and S signal-to-noise levels prior to waveform comparison. Prior to correlation, we calculate the cross-spectrum and cross-coherency for each pair of signals. The cross-coherency is used as a filtering function to adaptively downweight incoherent frequency bands prior to cross-correlation. We then calculate a suite of cross-correlation functions from overlapping, narrow-band-filtered representations of the (polarization- and coherency-filtered) waveform pair. The integer-rounded, energy-weighted mean correlation maximum and its associated standard deviation are calculated from these functions (ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001). Waveform pairs exhibiting sufficiently high integer correlation values (> 0.8 for the Soultz catalogue) with low associated lag standard deviations (< 2

samples) are aligned to the nearest sample, rewindowed, and passed to the subsample lag estimation algorithm.

A cross-spectral phase slope method (e.g., POUPINET *et al.*, 1984; ITO, 1995) is used to obtain subsample lag estimates following integer lag estimation. Because unbiased estimation of the cross-spectral phase slope is essential to obtaining accurate estimates of subsample lags and corresponding errors, we apply multitaper spectral estimation to this problem. A suite of eigenspectral measurements may be combined in a simple or weighted average to provide low spectral leakage, low variance spectral estimate and associated standard errors (e.g., THOMSON, 1982; PARK *et al.*, 1987; ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001). Conjugate multiplication of corresponding eigenspectra for the waveform pair provides multiple linearly independent estimates of the cross-spectral phase, yielding a mean cross-spectral phase slope estimate with associated dimensionally meaningful standard deviation.

After integer and subsample lags and their respective standard deviation estimates are obtained for each waveform pair, we sum the lags to provide a total lag estimate; the standard deviation of the pick lag is the quadrature sum of the coarse and fine lag standard deviations (ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001).

Solving for Consistent Pick Adjustments

After all possible waveform comparisons have been made, the resulting, overdetermined system of individual interevent lags may be solved to obtain consistent pick adjustments and associated uncertainties. We do this using an iterative conjugate gradient approach formulated to operate efficiently with a sparse matrix storage scheme (ASTER and ROWE, 2000; Rowe, 2000; Rowe *et al.*, 2001). The minimum L_1-norm residual solution is used, with a numerical modification for near-zero residuals to avoid gradient discontinuities, (e.g., ROWE, 2000; ROWE *et al.*, 2001). This provides improved robustness to outliers over L_2 methods (e.g., PARKER and MCNUTT, 1980; SHEARER, 1997). We further improve the solution by conservatively rejecting outliers in this massively overdetermined problem and successively resolving the system (ROWE and ASTER, 1999; ASTER and ROWE, 2000; ROWE, 2000; ROWE *et al.*, 2001). We calculate 1-sigma error bars via Monte Carlo perturbation of the data vector.

Event Association

We have found that solving the entire catalogue in one step does not provide optimal pick corrections because inconsistent waveform types are compared (ROWE, 2000; ROWE *et al.*, 2001). The choice of an appropriate event association method has proven to be crucial to improving the hypocenter locations. Because of the high degree of waveform variability within the small seismogenic volume, the Soultz

catalogue has some corresponding difficulties which may not be encountered in other settings. Our event association and pick adjustment strategy is outlined below.

We have adopted an event clustering strategy based upon waveform similarity (e.g., ASTER and SCOTT, 1993; REVIERE-BARBIER and GRANT, 1993). We treat the maximum cross-correlation value for each event pair as a measure of distance in similarity space, so that the correlation matrix may be viewed as a spatial (dis-) similarity matrix.

We obtained good intra-cluster results when we selected one of the most variable phases and used it as our clustering basis. The *S*-wave arrivals for station 4550 exhibit tremendous variation in waveform, therefore we selected an extended (0.05 s) correlation window for that phase, filtered using the signal cross-coherency function for each waveform pair. Cross-correlation of these waveforms provided a similarity matrix for a highly variable parameter in the seismogram database. We then applied a dendrogram-based hierarchical pair-group classification clustering scheme (e.g., LUDWIG and REYNOLDS, 1988; CARR *et al*., 1999; ROWE, 2000; ROWE *et al*., 2001) to build similarity groups. We have adopted a cross-correlation peak value of 0.75 as a minimum similarity criterion for any cluster.

Intra-cluster Adjustments

Within each cluster we then cross-correlate all phases and solve the respective first-difference systems to determine optimal pick adjustments. Because of significant variation in initial pick error from event to event (adjustments range from a fraction of a sample to 50 samples or more), as well as variations in waveform signal-to-noise ratio, we find that correlation window length can strongly affect waveform alignment even within highly similar clusters. The best choice can vary from cluster to cluster, as well as from phase to phase within a cluster. We therefore calculate cross-correlations for a suite of window lengths for each phase, and choose the parameters which produce the best mean cross-correlation and lag standard deviation values for the phase being addressed (ROWE, 2000; ROWE *et al*., 2001).

Once lags are estimated for all phases within a cluster, we solve for pick adjustments using our L_1-norm-minimizing conjugate gradient solver, imposing an *a postiori* zero mean constraint so that the mean of the adjusted picks corresponds to the preliminary pick mean. The corrected picks can provide very precise relative arrival times within a similarity cluster; however, inter-cluster relationships must still be addressed to achieve consistent inter-cluster picks, which may correspondingly correct the inter-cluster spatial relationships. Some means of correcting this bias without relying upon preliminary locations or traveltimes is desirable.

Inter-cluster Adjustments

We have found that a hierarchical correlation approach using stacked seismograms for each cluster provides good adjustments for inter-cluster relationships. For

each phase, we calculate vertical waveform stacks within each of the clusters, where stacks are aligned on the adjusted picks (see Fig. 3). In this way, each cluster may be viewed as having a composite catalogue location represented by the hypocenter calculated from its waveform stacks. To adjust the stacked waveforms for each phase (hence correcting for inter-cluster pick biases and, ultimately, inter-cluster relative

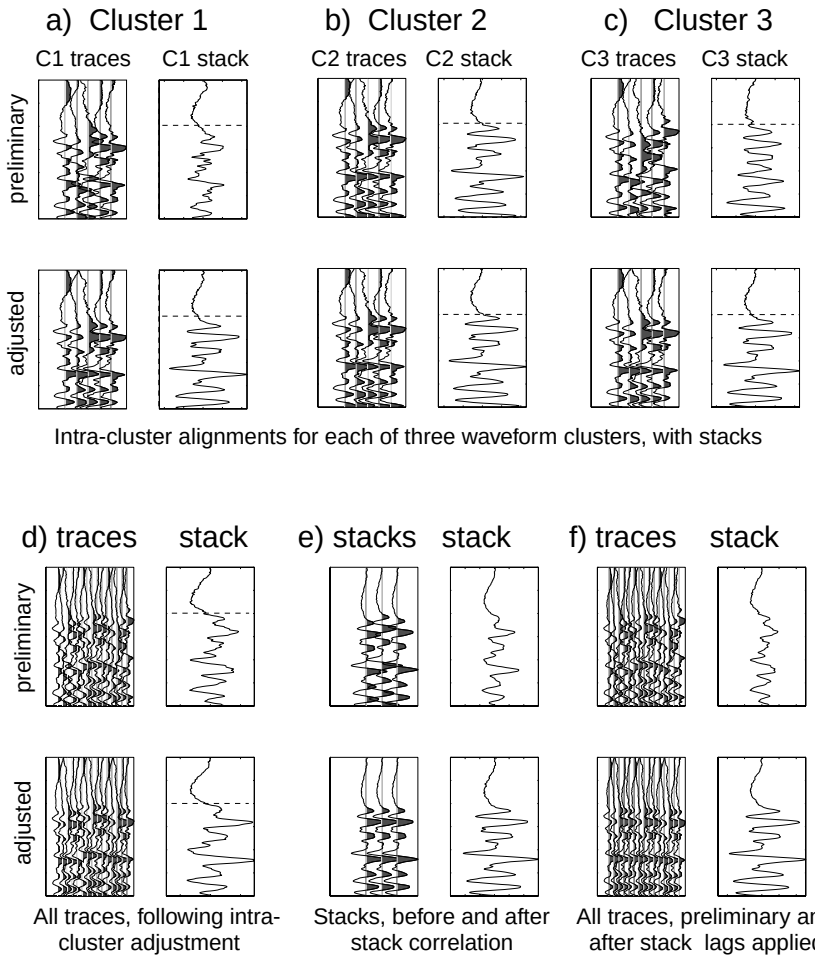


Figure 3

Hierarchical correlation, lag adjustment and stacking method. (a–c) illustrate three synthetic clusters of five events each. Upper panels show traces aligned on preliminary picks and associated waveform stack; lower panels show traces aligned on adjusted picks, with associated stack. Horizontal dashed lines indicate pick times on the stacked trace. (d) Clusters from a, b and c displayed together to show initial pick inconsistency (upper panel) and remaining biases among aligned clusters (lower panel). (e) Stacks from clusters of a, b and c showing stack mean pick misalignment (top) and aligned stacks (bottom). (f) the same three clusters showing initial misalignments (top) and final, corrected alignments (bottom) after both intra-cluster (a–c) and inter-cluster (e) lags have been applied to the individual traces.

centroid location), we apply the cross-correlation and conjugate gradient solution to the stacks. Each stack is assigned a relative pick lag which would serve to align stacks (mean picks) for every correlated cluster. We apply the stack lag to all corresponding phases for member events within each cluster, providing a final adjustment for inter-cluster relationships (Fig. 3). The tightly constrained intra-cluster relative relationships are thus preserved while the addition of the stack lag information adjusts inter-cluster centroids with respect to one another.

Final Pick Adjustments

There may still be a catalogue-wide absolute picking bias for some phases. For example, successively later picks for smaller events with lower signal-to-noise ratios – particularly for more distant stations whose arrivals are more emergent to begin with – may result in a net late pick bias. Addressing any such final absolute pick offset may be appropriately handled through adjusting to a master event, application of an autopicking routine to the second-order stacks, or application of consistent station corrections as derived through JHD or other joint location methods, although this latter approach does not serve to separate velocity delays from pick error delays. We discuss this further, below.

Automatic Repicking Test Case

The Soultz reservoir seismicity presents an excellent test case for development and application of the automated repicking algorithm, although it also presents some significant relocation issues not encountered with many larger seismic networks in tectonically active regions. The primary difficulty lies in the small number of recording stations, as discussed above. Generally, relocation studies using waveforms from large seismic networks (e.g., SHEARER, 1997, 1998; RUBIN, 1998, 1999) are able to neglect poor quality arrivals without compromising overall ability to obtain robust hypocenter locations. At the Soultz reservoir there are only four sites where downhole sensors were operating during the 1993 stimulation; hence, discarding any phase arrivals can destabilize the hypocenter solution process, particularly given the difficult network geometry (Fig. 2).

On the other hand, the sensors are situated in boreholes at depths between 1376 and 2075 m, resulting in very low background noise and good microearthquake signals. Three of the instruments are multicomponent sensors, which allow polarization filtering for improving signal correlation (particularly for *S* waves).

We first tested our method using the deep cluster events of PHILLIPS (2000) as a benchmark against which to compare our results (ROWE *et al.*, 1998, ROWE and ASTER, 1999; ASTER and ROWE, 2000; ROWE, 2000). The painstaking Phillips repicks and relocations were assumed to be the "correct" answer; step-wise comparison with these allowed us to gauge improvements in the development of our technique.

Initial correlation on the *S*-wave picks for station 4550 using a window length of 50 ms (about 5 cycles) provided a similarity matrix for clustering the 311 events within the Phillips deep cluster which were included in our catalogue. This resulted in eighty similarity groups of varying sizes, dominated by doublets (two-event clusters). Figure 4 shows a histogram of similarity cluster membership for this trial data set. Eighteen events were orphaned (unassociated) in the clustering process; the remaining 293 were addressed in the analysis.

Applying the automatic repicking algorithm to *P* and *S* phases for the deep cluster, we demonstrate the resulting waveform alignment in Figure 5. The

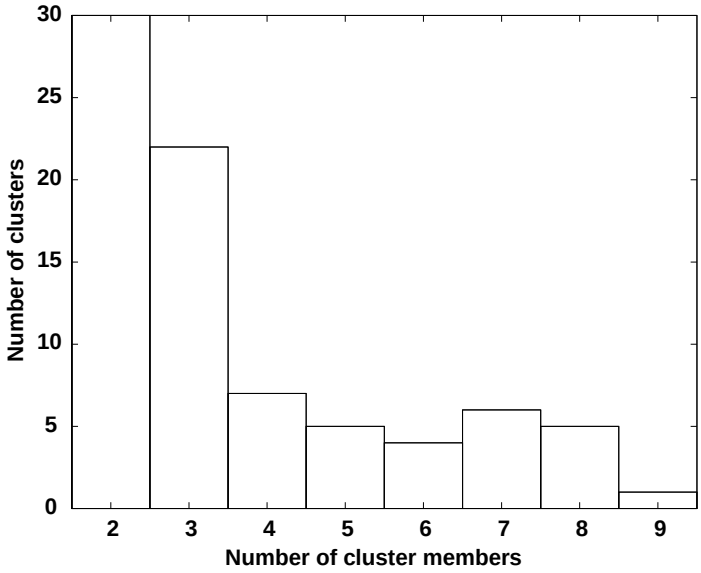
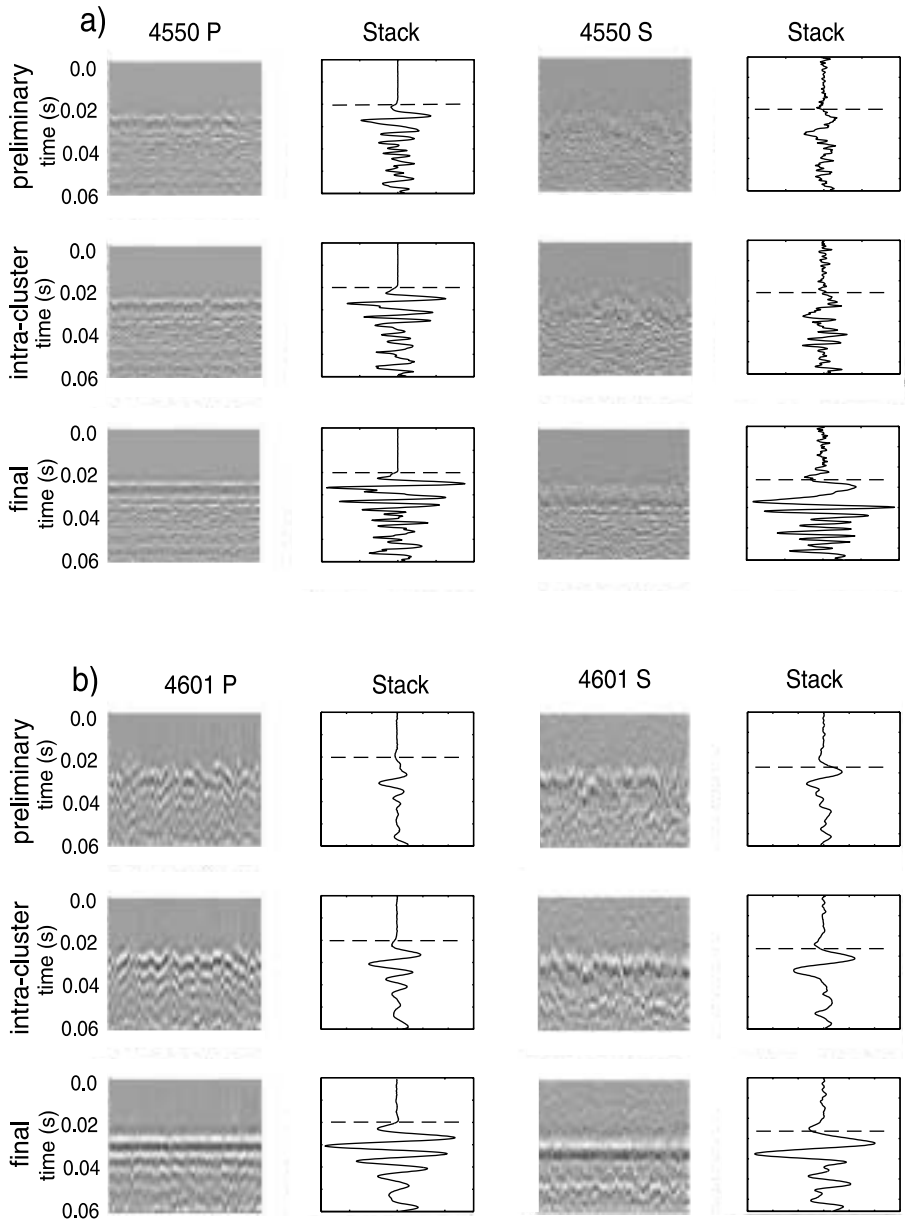


Figure 4
Earthquake cluster sizes for PHILLIPS (2000) deep cluster. Histogram shows the number of families for each category of doublet, triplet, quadruplet, and so forth. A total of eighty clusters were identified.

►

Figure 5
Waveform alignment plots for events in the deep cluster (PHILLIPS, 2000). Waveforms are plotted as a gray-scale surface with white peaks and gray troughs. Each phase is depicted in six panels. Waveform alignments are on the left, stacks of the waveforms are on the right. Top panels show waveforms aligned on preliminary picks. Center panels show waveforms (and stack) aligned on the repicks arising from intra-cluster relative repicks. Lower panels show alignments and stacks of waveforms after inter-cluster adjustments have been made. (a) *P* waves for station 4550 are shown in the two left-hand columns. *S* waves for station 4550 are shown in the two right-hand columns. (b) *P* waves for station 4601 are shown in the left-hand columns; *S* waves are shown in the two right-hand columns. (c) *P* waves for station 4616 shown on the left-hand side, *S* waves for station 4616 shown on the right. (d) *P* waves for the hydrophone. No preliminary *S* was picked for the hydrophone, as this instrument was underdamped and the waveforms were often clipped.

uppermost panels for each phase illustrate alignment on preliminary analyst picks. Center panels represent alignments after intra-cluster adjustments have been made; the lowermost panels show alignment following correction for inter-cluster consistency. Stacked waveforms for each alignment plot are shown on the right, with the



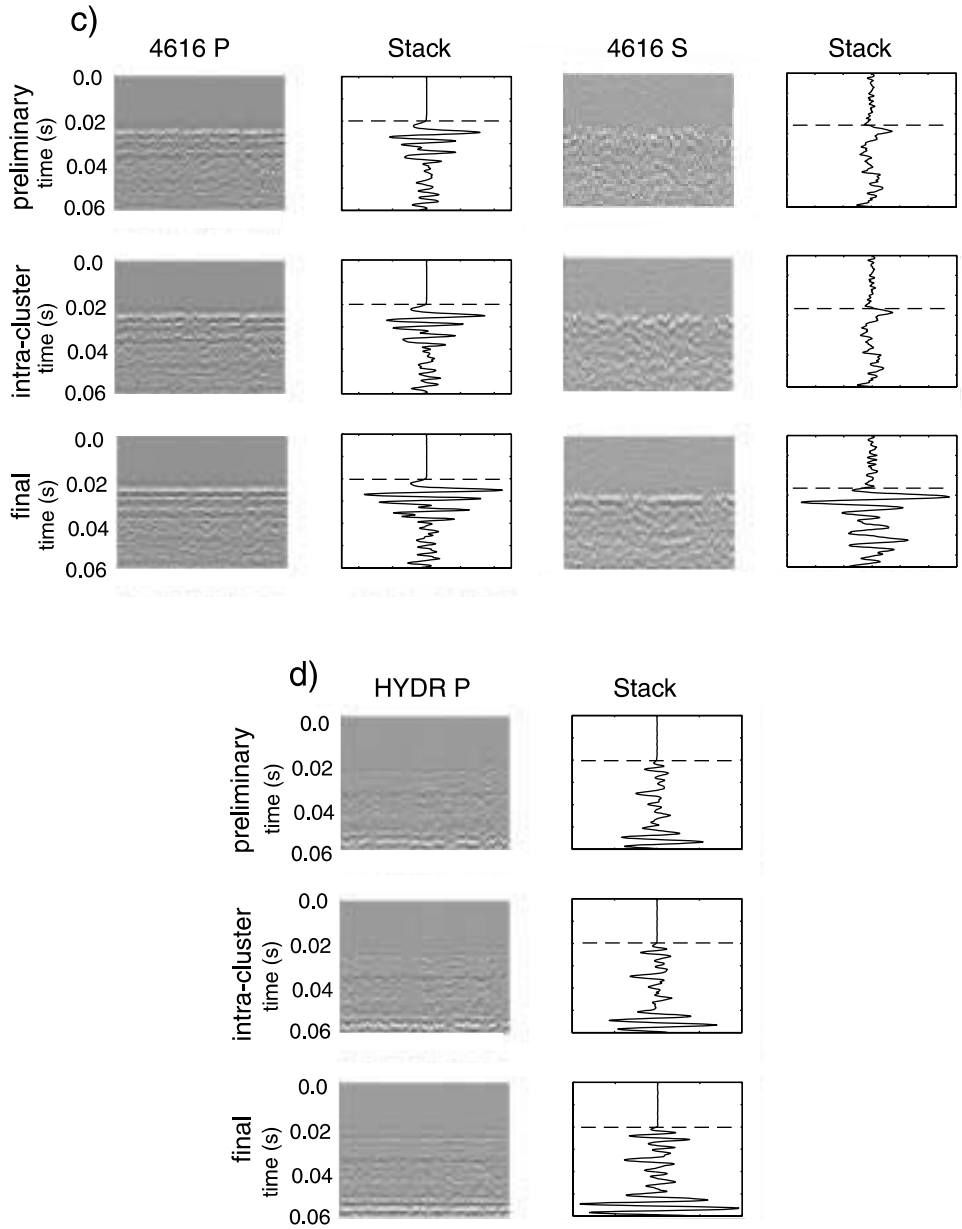


Figure 5c, d

aligned picks indicated by horizontal black bars. The increase in stack amplitude provides a qualitative measure of the repicking success.

Note that Figure 5 does not show the filtered and projected waveforms which were cross-correlated to obtain the pick lags. Because the filtering and optimal signal

rotation may be different for every interevent cross-correlation, it is not possible to show these in a single figure. We show instead alignments of unfiltered traces. For *P*-wave alignments, we show the vertical component, and for *S*-wave alignments, we display the east-west (*X*) horizontal component. Improvement in waveform alignment is especially marked for *S* waves. For instance, at station 4616 the summed *S*-wave stack amplitude after repicking is nearly three times that of the preliminary stack.

After cross-correlating and solving for consistent pick adjustments, we relocated the deep cluster earthquakes using the same location program (single event, damped least-squares) and velocity parameters ($V_p = 5.85$ km/s, $V_s = 3.34$ km/s) determined by JUPE *et al.* (1994) and used by PHILLIPS (2000). Figure 6a, b illustrates hypocenter locations for preliminary picks and manual repicks, respectively. In Figure 6c we show hypocenters calculated using automatic intra-cluster repicks, to illustrate the increased degree of clustering obtained by this step. These hypocenters correspond to the waveform alignments in the center panels for each phase shown in Figure 5. Tight intra-cluster alignments result in increased clustering among hypocenters, but the large inter-cluster biases have not yet been addressed. Note that the structures in Figure 6b are still not clear in Figure 6c, although hypocenters exhibit significantly increased clustering over the preliminary locations of Figure 6a. Figure 6d shows the relocations following inter-cluster stack correlation and adjustment (corresponding to picks aligned in the lowermost panel for each phase in Fig. 5). The clearly defined, near-orthogonal intersecting joint features illuminated by the manual repicking in Figure 6b and discussed in detail in PHILLIPS (2000) and PHILLIPS *et al.* (2001) are also clearly distinguishable in Figure 6d. Median hypocenter mislocation, compared to relocations using the Phillips picks, is reduced from 31 to 7 m after we have adjusted centroid median biases.

Because Phillips correlated peaks and troughs with the visual cross-correlation, then applied station corrections to his picks to approximate first-break consistency for locating events, it is difficult to compare directly the relocation quality in terms of hypocenter RMS traveltime errors; however, we iteratively adjusted the JUPE *et al.* (1994) JHD station corrections to minimize mean traveltime residuals for automatic repicks and preliminary picks. The resulting RMS travel-times errors are 0.68 ms, 0.25 ms and 0.14 ms for preliminary, automatic and manual picks, respectively, with associated standard deviations of 0.62, 0.34 and 0.24 ms, respectively.

Application to the Larger Data Set

When we had completed testing of the automatic algorithm using the PHILLIPS (2000) deep cluster, we turned to the task of relocating the ~7,000 events in the Soultz catalogue that had seven preliminary phase picks.

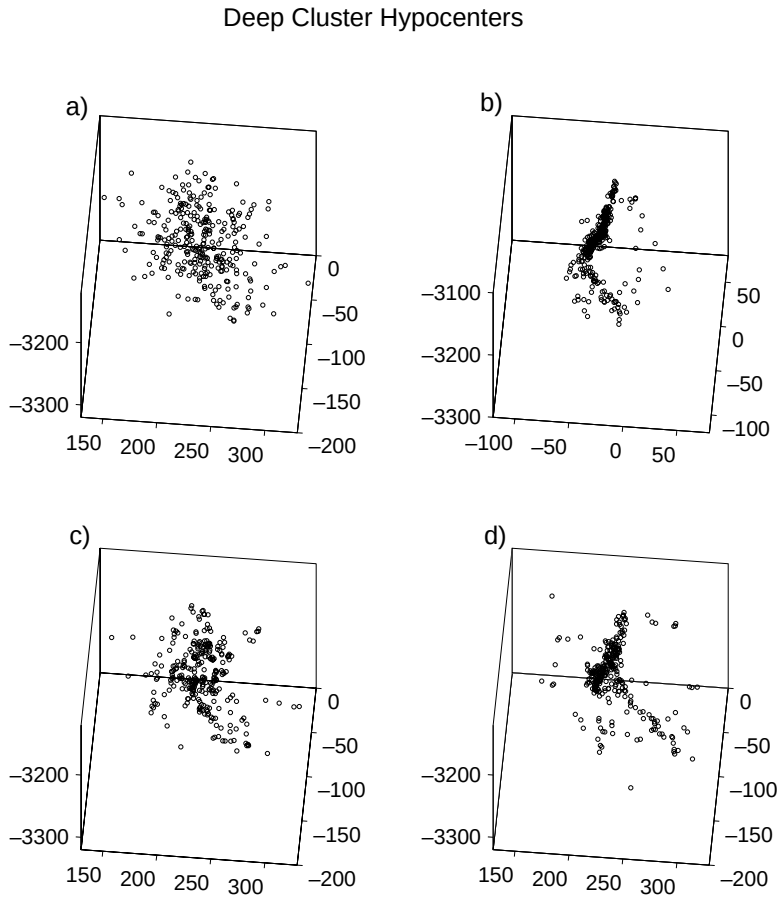


Figure 6
Hypocenter locations for Phillips' deep cluster. (a) Preliminary locations, based on initial, analyst picks. Note the random scatter in the distribution of seismicity. (b) Relocations based on PHILLIPS (2000) manual repicking. The diffuse cloud of hypocenters (a) has converged into two nearly orthogonal, intersecting joint features and a few scattered events. (c) Relocations based on intra-cluster pick adjustments (corresponding to waveform alignments in center panels of Figure 5). (d) Relocations after inter-cluster stack adjustments have been applied to individual traces (corresponds to alignments in lower panels of Figure 5).

Relocations

Processing the ~7000-event catalogue proceeded in the hierarchical correlation/clustering method as discussed for the deep cluster (above). Following cross-correlation on the *S* waves at station 4550, the algorithm segregated the catalogue into 2408 clusters, plus 661 unassociated (orphaned) events. Figure 7 shows the resulting cluster memberships. We show the number of clusters of each size (from 2 to 18 members) in Figure 7a, while plotting the cumulative proportion of the catalogue comprised of the different cluster sizes in Figure 7b. Picking inconsistencies

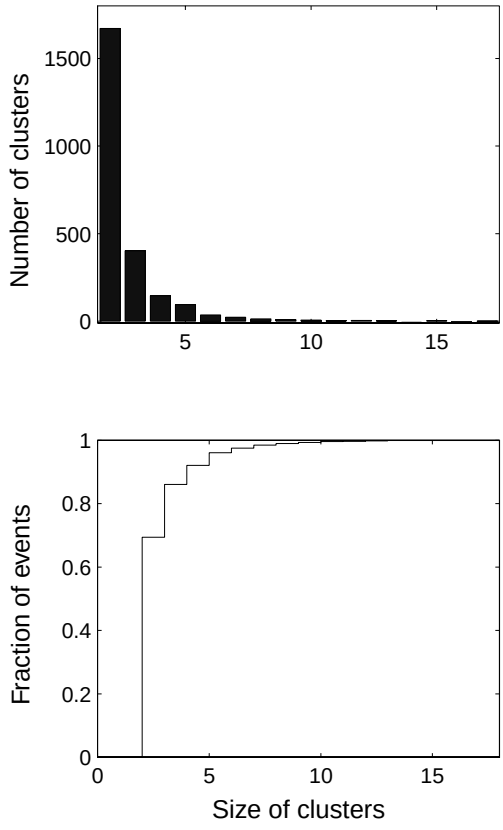


Figure 7
Cluster memberships for the ~7000-event data set. (a) Histogram shows the number of families for each category of doublet, triplet, quadruplet, and so forth. A total of eighty clusters were identified. (b) Cumulative fraction of the catalogue assigned to clusters of two, three, four and so forth. Note that doublets and triplets comprise over 87% of the data, indicating that adjusting relative pick lags among clusters is important to the final resolution of spatial relationships among earthquakes.

were resolved within similarity clusters with median lag adjustments ranging from 0.251 ms for station HYDR *P* picks to 1.6 ms for station 4601 *S* picks (Table 2). After adjusting each of the seven phase picks for consistency within each of the 2408 clusters, the aligned waveforms for each phase were stacked to provide a composite waveform (e.g., Fig. 3). These cluster stacks were then correlated with corresponding stacked waveforms for all other clusters so that inter-cluster picking inconsistencies might be resolved, and the final pick lag adjustments made. Inter-cluster stacked waveform adjustments are shown in Table 3. Maximum inter-cluster lag adjustments were 17.2 and 11.1 ms for *S* waves and *P* waves, respectively; both maxima were observed for station 4601, the most distant station from the seismogenic volume (Fig. 2).

Table 2

Intra-cluster relative pick adjustments and standard deviations

Station	Median *P* lag (ms)	*P* st.dev (ms)	Median *S* lag	*S* st.dev (ms)
4550	0.34	0.03	0.76	0.05
4601	1.1	0.06	1.6	0.06
4616	0.38	0.03	0.8	0.06
HYDR	0.25	0.03	N/A	N/A

Table 3

Inter-cluster stack pick adjustments and standard deviations

Station	Median *P* lag (ms)	*P* st.dev (ms)	Median *S* lag (ms)	*S* st.dev (ms)
4550	0.77	1.2	1.5	2.4
4601	1.8	2.7	3.0	4.7
4616	0.85	1.3	2.0	2.8
HYDR	0.55	0.96	N/A	N/A

Automatic repicking and relocation of our ~7000 event data set yielded significant improvement in illumination of interpretable seismogenic features. Median RMS traveltime error for repicked hypocenters decreased from 3.741 to 3.739 ms compared to original picks, using the station corrections derived from JHD (e.g., JUPE *et al.*, 1994). This moderate change of 0.2% indicates that overall location errors are far more heavily influenced by the oversimplified velocity model and the station correction parameters, which were derived to optimize fitting of the preliminary picks for a larger data set, than by moderate pick adjustments; however, waveform alignments and improved resolution of fracture and joint features indicate that the pick adjustments are providing improved hypocenter locations.

Figures 8–11 show the results of relocating our ~7000-event data set. We show map-view and three-dimensional projections for depth slices at 2500–2900 m (Fig. 8), 2900–3100 m (Fig. 9), 3100–3300 m (Fig. 10) and 3300–3700 m (Fig. 11). The three-dimensional projections have been chosen for optimal view angle to show the emerging structures within the seismicity cloud at each level. In each of Figures 8–11, we depict preliminary hypocenter locations in the left-hand panels, and relocated hypocenters in the right-hand panels. We have used the same SEL location algorithm (PHILLIPS, 2000) for both preliminary and repicked events to demonstrate the level of improvement which is possible solely from automatic repicking.

Relocated shallow events are shown in Figure 8. Figure 8a shows these events in map view. Black arrows (B) and (C) indicate view azimuths in Figure 8b, c. Shallow earthquakes exhibit roughly parallel lineations in map view, striking N40°W in the more southerly part of the reservoir, although the seismicity does not define these

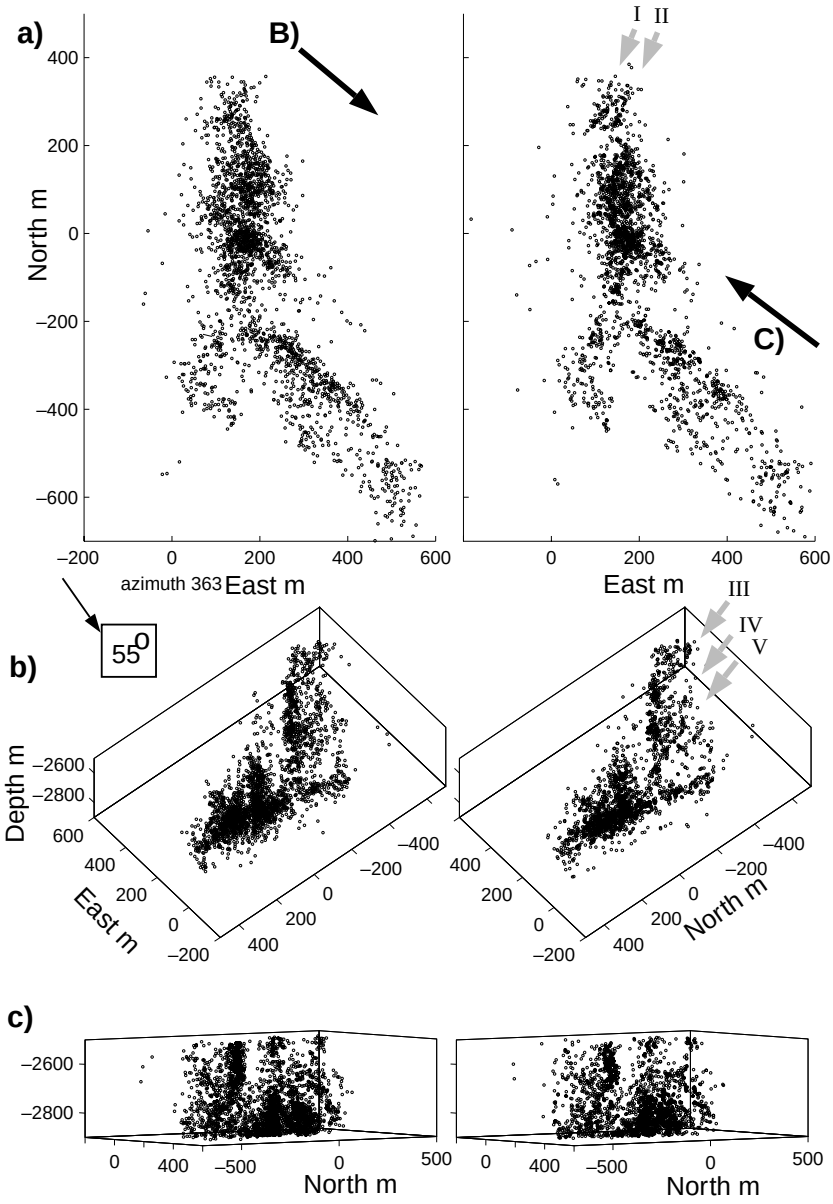


Figure 8
Preliminary (left) and relocated (right) hypocenters for shallow (2500 to 2900 m depth) Soultz seismicity. Details are discussed in the text. (a) Map view. Arrows (B) and (C) indicate the view directions for (b) and (c). (b) Steeply inclined view at 55° of shallow seismicity, looking towards S50°E. (c) Cross-section looking N40°W along strike of the segmented, southern portion of the shallow reservoir.

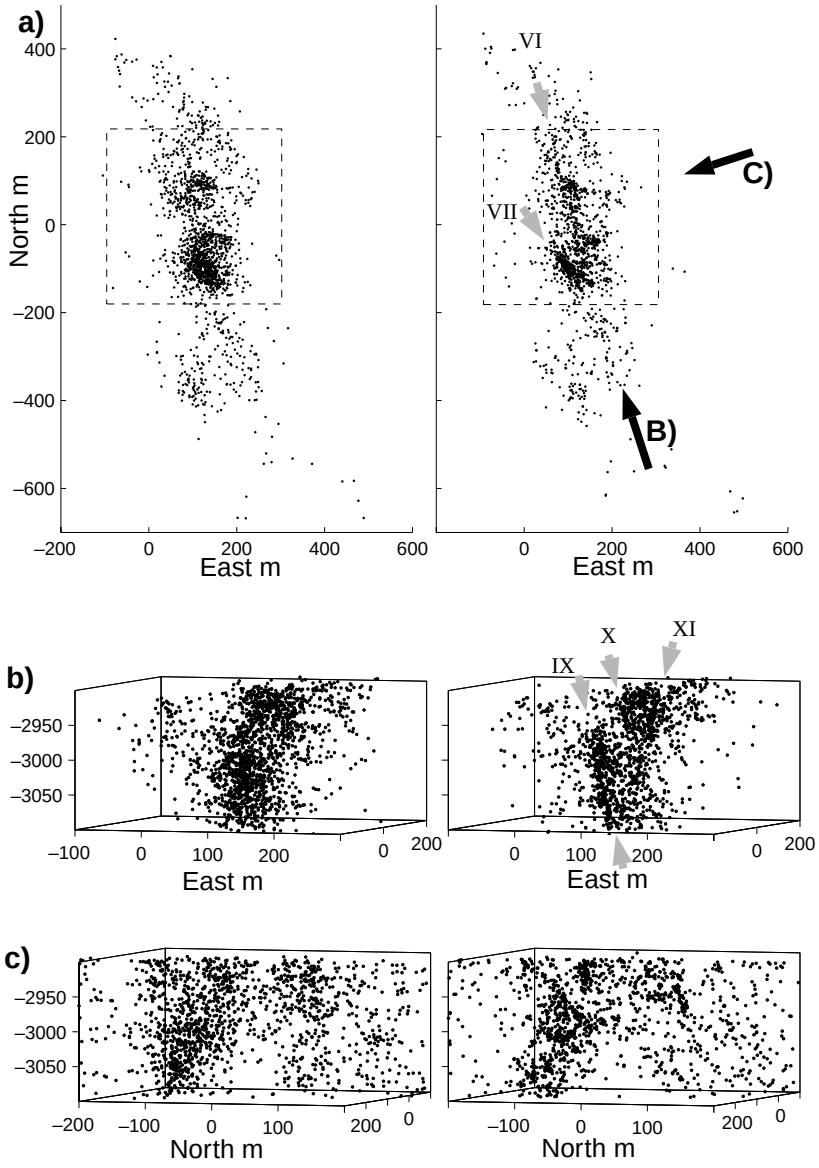


Figure 9
Intermediate depth seismicity (between 2900 and 3100 m). Details are discussed in the text. (a) Map view. Cross-section view angles in (b) and (c) are indicated by direction arrows (B) and (C). (b) Cross-sectional view looking N18°W, along strike of the clustered features seen in map view. (c) Broadside view looking S72°W.

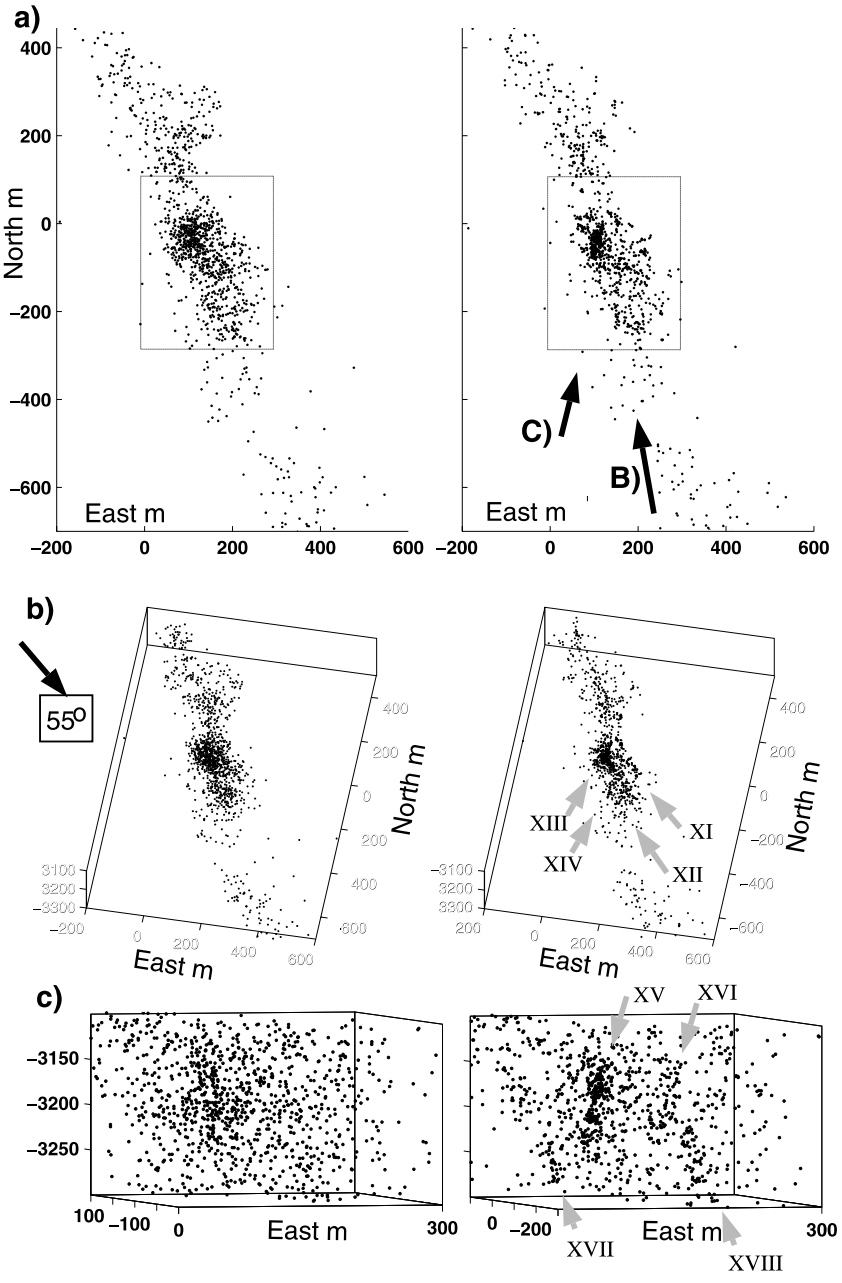


Figure 10
Deep seismicity between 3100 and 3300 m. Details discussed in the text. (a) Map view. Dashed box outlines the close-up elevation shown in c. View azimuths for (b) and (c) are indicated by arrows (B) and (C). (b) Elevated view at 55° inclination, azimuth N10°W, approximately coincident with Phillips' deep cluster intersection axis. (c) Cross-section looking N14°E, along strike of Phillips' "primary" plane.

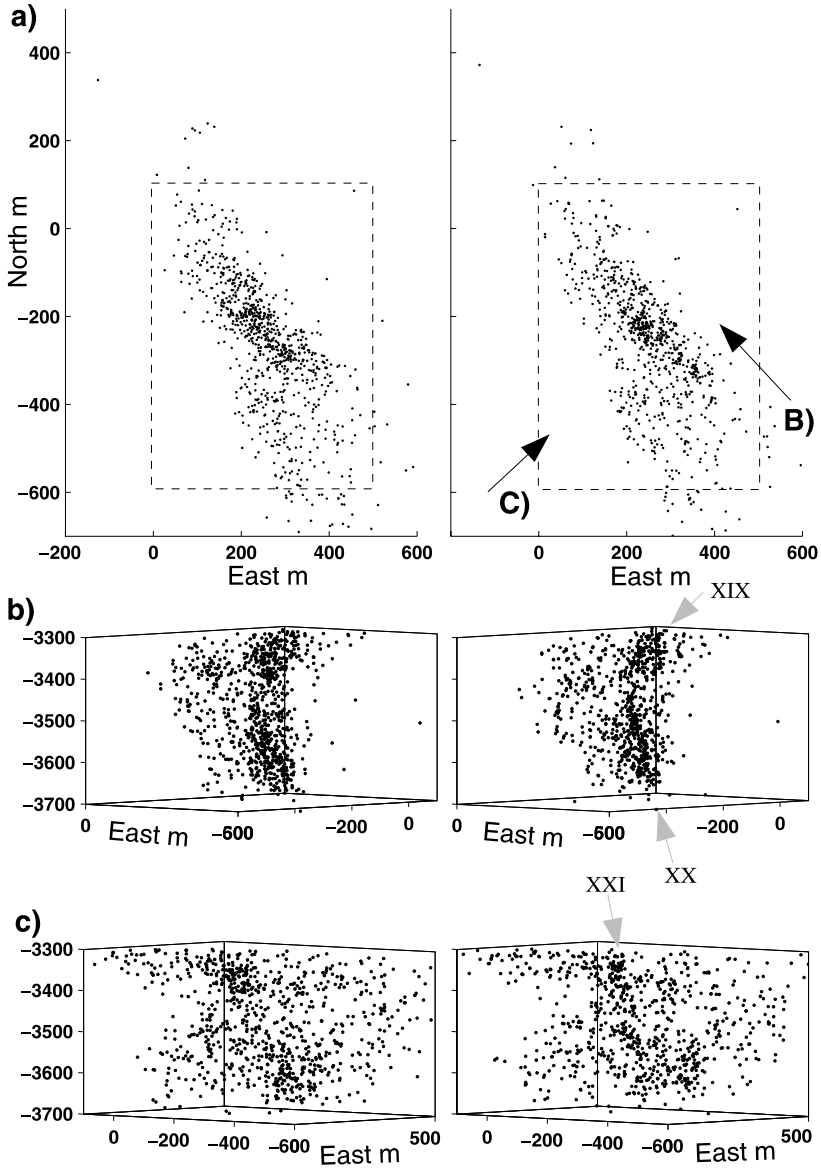


Figure 11
Deep seismicity (3300–3700 m) associated with 1993 injection. Details are discussed in the text. (a) Map view. Dashed box represents region shown in the elevations in (b) and (c), whose azimuths are indicated by arrows (B) and (C). (b) View looking N40°W along strike of the general trend in (a). (c) Broadside view, perpendicular to (b).

trends as clearly in the more northerly part. This orientation is similar to the orientation of one of the populations of hydrothermally sealed fractures identified by GENTER and TRAINEAU (1996). To the north, lineations suggest fractures or faults with an orientation of approximately N20°E (indicated by gray arrows I, II at the top of the figure). An elevated view looking S50°E with a 55° inclination (Fig. 8b) suggests that the N30°W seismicity trend dominating the southern part of the reservoir at these depths is made up of discrete, subparallel planes (indicated by gray arrows III, IV, V) steeply dipping towards the northeast. These structures are not discernible at all in the preliminary hypocenter cloud (8b, left-hand panel). In Figure 8c we show a cross-section looking towards N55°W. This shows the steeply eastward-dipping nature of the southern lineations, as well as a northerly dip for the isolated planes of seismicity seen in 10a at 200–300 m north.

Intermediate-depth events (between 2900 and 3100 m) are plotted in Figure 9. In map view the northwesterly-striking features may still be seen, although lineations are less laterally extensive. Two linear features are suggested, one striking N15°W (gray arrow VI) and one approximately N30°W (gray arrow VII). Seismicity did not expand as far to either side of the borehole at this level as it did at a shallower depth, suggesting a change in the fracture system and its connectedness at approximately 2900 m depth. The dashed boxes in Figure 11a indicate the area displayed in close-up cross-section in Figure 9b, c. Black arrows (B) and (C) indicate view azimuths for Figure 9b, c. In Figure 9b, looking N10°W, we see that after relocation the hypocenters define planar features which dip steeply to the north (gray arrows IX and X), but also a lineation suggesting an intersecting feature (gray arrow XI). Viewing these features broadside in Figure 9c (perpendicular to strike) illuminates small linear seismicity clusters within their planar surfaces.

Deeper (between 3100 and 3300 m) events are shown in Figure 10. Relocated hypocenters exhibit considerable improvement in clustering as seen in map view (Fig. 10a), particularly in the dense concentration of events at (100, −50). These events correspond to the depth of the PHILLIPS (2000) deep cluster, discussed earlier. Black arrows (B) and (C) indicate view azimuths in Figure 10b, c. In Figure 10b we see the inclined fracture (''secondary'') plane (gray arrows XI and XII) of the Phillips deep cluster (see Fig. 6) as well as the northerly-trending (''primary'') plane (gray arrows XIII and XIV) that it intersects. Relocated seismicity suggests additional lineations parallel to these major features which may indicate a series of subparallel seismogenic fractures activated by the reservoir stimulation. The cross-section in Figure 10c looks towards N14°E, along strike of the Phillips ''primary'' plane. The steeply westward-dipping lineation (gray arrow XV) is echoed by another, less well-defined plane some 100 m to the east (gray arrow XVI); two deeper lineations dipping steeply towards the east (gray arrows XVII and XVIII) are suggested in this view as well.

Figure 11 shows the deepest seismicity at Soultz associated with the 1993 injections, extending from 3300 to 3700 m depth. In map view (Fig. 11a) we see the

diffuse cloud is resolved after relocation into a dense, northwesterly-trending feature that appears to be comprised of discrete patches of seismicity. Black arrows (B) and (C) indicate view azimuths for Figure 11b, c, respectively. In a cross-section looking N40°W (Fig. 11b), this concentrated zone of seismicity is revealed to be composed of two sets of roughly planar features, striking N40°W. The shallower features, extending to ~3500 m depth, appear to be subvertical lineations which dip towards the south (gray arrow XIX), whereas the deeper feature is a well-defined plane extending from ~3500 to ~3700 m depth, dipping steeply to the north (gray arrow XX). Viewed broadside (Fig. 11c), both shallow and deep planes consist of discrete seismicity patches which may correspond to distinct rupture zones on the seismogenic planes. Quasilinear edges (gray arrow XXI) to these zones suggest truncation by some linear or planar feature intersecting the seismic plane.

Discussion

The microearthquakes comprising this data set are the result of two different periods of reservoir stimulation, in September and October, 1993. Figure 12 illustrates preliminary hypocenter locations plotted as a function of time for the stimulation. In Figure 12a we show hourly event counts (for the 7000 microearthquakes we studied). Figure 12b, c, d illustrates the seismicity distribution through time in East–West (X), North–South (Y) and depth (Z) dimensions, respectively. Approximately 80% of the microearthquakes occurred during the September stimulation, during which ~25,000 m^3 of water were injected over an open-hole depth range between 2800 and 3500 m at rates of up to 36 l/s (CORNET and JONES, 1994) and overpressures of ~10 MPa (JUPE *et al.*, 1998). The bottom portion of the injection well GPK1 was filled with sand (depths > 3480 m) to prevent flow into a previously identified, deep fracture zone. Seismicity during this first stimulation began at a depth of ~2900 m, adjacent to the injection site, and expanded over time in an elongate cloud oriented roughly N-S at shallower depths (< 2900 m) and roughly NW–SE below this (CORNET and JONES, 1994).

Histograms of hypocentral depths of the seismicity resulting from the two stimulations are shown in Figure 13. The September stimulation, shown in Figure 13a, exhibits a strong peak in seismicity at a depth of ~2900 m. The October stimulation has the greatest number of its events occurring at ~3600 m depth (Fig. 13b). The time-varying hypocentral depth plot in Figure 12d shows this, and demonstrates that the removal of sand from the final 110 m of well GPK1 prior to the second stimulation causes the primary seismic activity to be generated by new slip within the deep fracture system, while the shallower portion remains comparatively quiet until overpressures are raised sufficiently to induce further failure on the previously destressed fracture system (JUPE *et al.*, 1998).

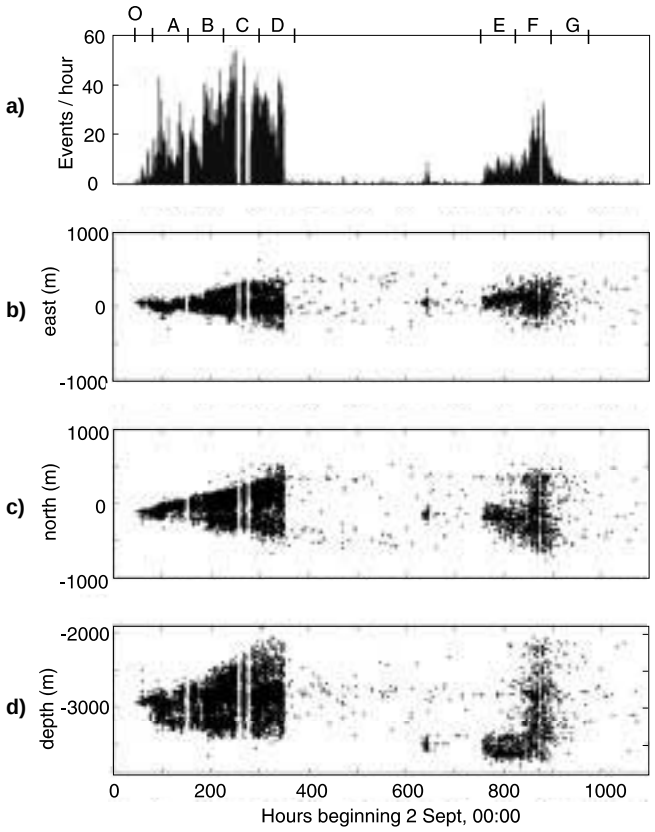


Figure 12
Spatial development of Soultz seismicity as a function of time, during the September and October 1993 injections over open-hole depths of 2800–3500 m in well GPK1 (Fig. 2). (a) Hourly event counts. Segments A–D and E–G represent 72-hour time intervals examined in Figures 14 and 15. (b) Longitude (meters east). (c) Latitude (meters north). (d) Hypocentral depth.

The greatest concentration of events within a volume might be expected to yield the greatest number of similar earthquakes. Intuitively we can look at the growing volume of seismicity and suspect that in the earlier hours of stimulation, when seismicity is concentrated within a small volume, many similar earthquakes should be observed. MORIYA *et al.* (2001) have shown through their analysis of doublets and multiplets that seismicity is dominated by doublets/multiplets during early stimulation, whereas the later activity has a significantly lower proportion of highly similar events. PHILLIPS (2000) noted a similar trend in his manual re-analysis, demonstrating that tight lineations of relocated hypocenters are formed prior to the spreading of seismicity into adjacent planar features. Phillips interprets the tight lineations of hypocenters as primary flowpaths for the injected fluid. Waveform similarity (i.e.,

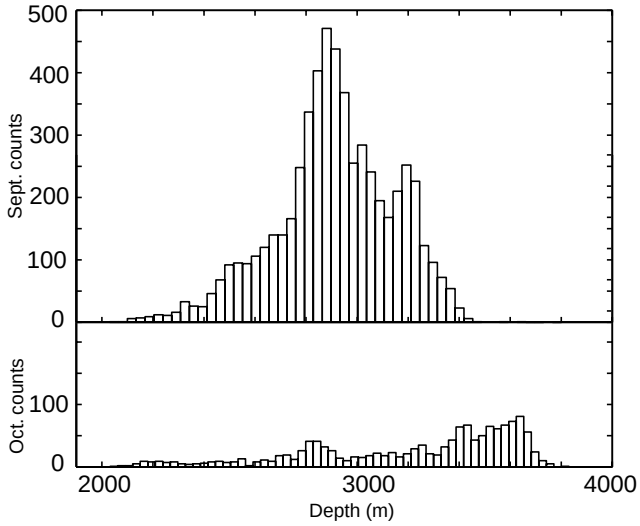


Figure 13
Histograms of hypocentral depths during the September and October, 1993 reservoir stimulations (Fig. 12) is dominated by earthquakes occurring at about 2900 m depth. After removal of the sand at the bottom of GPK1, the October stimulation (b) exhibits greatest activity at 3400–3600 m depths, with a secondary peak at the fracture zone near 2900 m depth.

membership in multiplets) is more likely to be observed within the early, tightly clustered events where source and path effects may be more consistent.

We expand upon these observations by calculating time-varying and spatially varying hypocenter density within the stimulated volume. The seismogenic volume in any time interval $d\tau$ may be estimated as the product of the maximum X, Y and Z dimensions during $d\tau$. These dimensions are determined to first order by estimating a three-dimensional envelope about the time-varying seismicity. MORIYA *et al.* (2001) report nearly 90% multiplet/doublet fraction during the first 48 hours of stimulation. Mean hypocenter density within this time period is $\sim 10^5$ events per km^3. Figure 14 shows event counts, approximate reservoir volume and estimated mean event density for the first five intervals (from Fig. 12) following the onset of stimulation. Event count per unit volume appears to decline in a roughly logarithmic fashion (Fig. 14c). The highest value of event density, 10^5/km^3, corresponds to the MORIYA *et al.* (2001) estimated 90% doublet/multiplet fraction for events. The lowest density value, $\sim$1400/km^3, corresponds to MORIYA *et al.*'s approximately 50% multiplet fraction.

Identification of planar and linear structures based on microseismicity may lead to identification of likely fluid flow directions and/or potential reservoir connectivity. FEHLER *et al.* (1987) developed an innovative method of identifying likely fracture planes along which earthquakes may have occurred. ROFF *et al.* (1996) further investigated linear and planar features through analysis of waveform amplitude

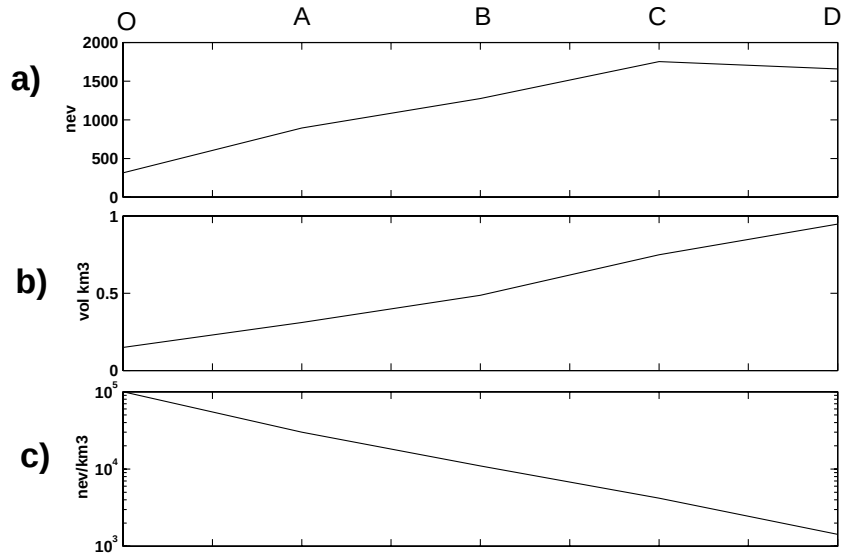


Figure 14
(a) Event counts from the first five 72-hour time intervals (A–D) for the September 1993 stimulation, as indicated in Figure 12a, beginning at point **O** (1623 hours, 2 Sept.). Vertical axis represents the number of earthquakes, horizontal axis indicates time interval (slight decrease in seismicity for the final data point is a reflection of the shorter time interval of active stimulation; reservoir deflation began during this interval). (b) Reservoir volume in each time segment. Vertical axis represents volume in km^3. The volume was approximated by estimating an envelope about the seismicity in the X, Y and Z directions as shown in Figure 14b–d, and calculating the mean volume for each time interval. (c) Event density estimate for each time interval. Vertical axis indicates estimated number of events per km^3, plotted on a logarithmic scale.

ratios, and PHILLIPS *et al.* (1997) investigated cluster structures via manual hypocenter relocations and slant stack amplitudes. STARZEC *et al.* (2000) found a correlation between the shapes of variograms of shear displacement and fracture orientation, suggesting a correlation of spatial variability of seismic slip and fracture complexity.

Tracking the highest concentrations of seismicity as a function of time may lead to the temporal mapping of fluid flow paths (e.g., PHILLIPS, 2000) or, alternatively, adjacent patches of failure on faults activated by pore pressure increases (CORNET and YIN, 1995); this may lead to better understanding of the dimensions and time development of important reservoir features. We illustrate in Figure 15 an east-west cross-section of the hypocenters in each of the seven 72-hour intervals A–G identified in Figure 12a. Upper panels depict original locations; lower panels show relocated hypocenters. Because we required the existence of seven phase arrivals to relocate events, and the period prior to time segment (a) (beginning at point O, indicated on Figure 12a) offered no hydrophone arrivals, we are unable to compare relocations for that period. We have calculated hypocentral density by counting neighbors within a sphere of 5 m radius centered on each hypocenter during each of the time intervals.

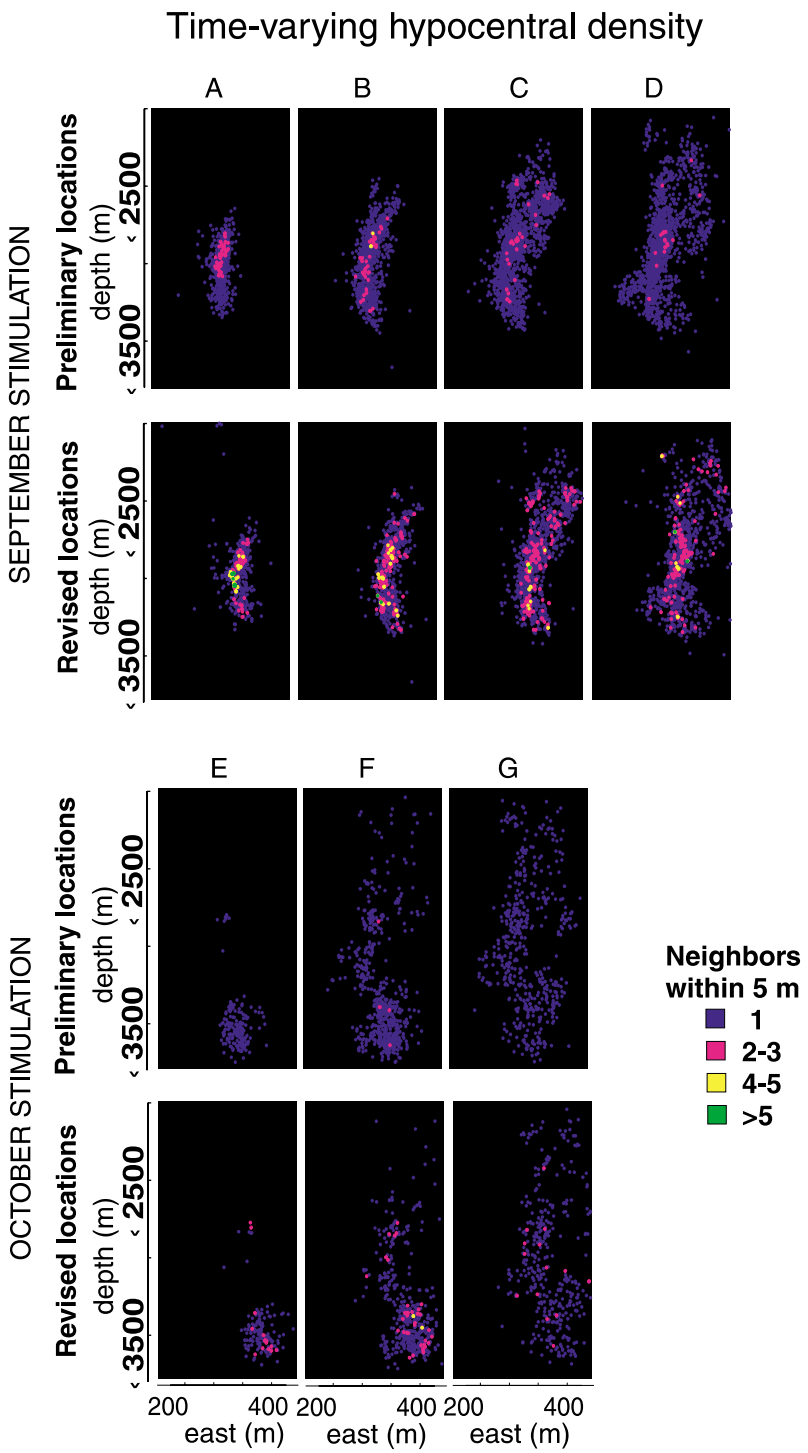

Time-varying hypocentral density
A
B
C
D
SEPTEMBER STIMULATION
Preliminary locations
Revised locations
depth (m)
2500
3500
E
F
G
OCTOBER STIMULATION
Neighbors within 5 m
1
2-3
4-5
>5
200
400
east (m)

Table 4

Percentages of singlet/multiplet events in 72-hour time segments

Time segment	a	b	c	d	e	f	g
% singlet preliminary	88	93	97	98	100	99	100
% multiplet preliminary	12	7	3	2	0	1	0
% singlet relocated	67	69	76	80	81	84	92
% multiplet relocated	33	31	24	20	19	16	8

Using the method of SNOKE (1987) to estimate source radius from corner frequencies, we have found that the median source radius for the relocated data subset to be approximately 3 m, so most events which fall within the 5 m search radius will have abutting or overlapping rupture patches. Upper panels of Figure 15 show results for the seven time intervals (A–G) using preliminary locations; lower panels illustrate relocated events. Hypocenters are color-coded to indicate spatial density, as indicated on the color key in Figure 15. Total numbers of earthquakes depicted are 6871 and 6217 for preliminary and relocated events, respectively. The discrepancy arises out of the number of events which were orphaned during the clustering process; these earthquakes were not repicked and so are not included in the relocations.

Overall, a clear increase in clustering after relocation is indicated in Figure 15. The tightest clustering remains nearest the injection point during the September stimulation (panels A–D), although the appearance of doublet events (magenta) on the edges of the developing reservoir margins suggests a concentration of activity at a migrating pressure front, although this may not necessarily reflect the locations of fluid paths (e.g., CORNET and YIN, 1995; CORNET *et al.*, 2000). A similar, albeit smaller, trend is apparent in panels (E) and (F), which correspond to the later, October, stimulation following removal of sand from the bottom of the borehole. Among events in the seven time periods (A–G) displayed in Figure 15, 74% of relocated events do not locate within 5 m of their nearest neighbor, whereas 95% of original hypocenters fail to cluster. Table 4 presents the singlet (events which have no neighbors within a 5 m radius) versus multiplet (two or more events within a 5 m radius) statistics by time interval, comparing percentage of singlet events for preliminary and relocated hypocenters during the seven time periods of Figure 15.

We note that relocating events solely through the use of adjusted phase picks may still not provide optimal relative and absolute hypocenter relationships, without

◀

Figure 15

Spatially varying hypocentral density as a function of time. Preliminary (top panels) and revised (bottom panels) depicting seismic development of the Soultz reservoir during each of the time intervals A–D, and E–G of Figure 14a. Color represents the number of neighbors within a 5 m radius. Intervals A–D correspond to the September stimulation, when the open hole extended from 2800 to 3400 m depth and the bottom of the well was sanded. Intervals E–G correspond to October injection following removal of the sand.

addressing inadequacies of the velocity model and station corrections. The Soultz velocity model and station corrections were derived from a single downhole shot (e.g., JUPE *et al.*, 1994) in combination with analysis of microseismicity using the preliminary phase picks. We observe systematic biases in these phase picks, which, if uncorrected, will contribute to the inaccuracy of preliminary velocity assumptions and subsequent analyses using those velocities. For example, pick times shown in Figure 5b demonstrate that the mean *S*-wave pick for station 4601 is clearly late, occurring more than a quarter cycle (~6 ms) after the *S*-wave arrival. Because we impose a zero mean constraint on our pick adjustments, this phase pick bias is an artifact of the preliminary pick distribution and has contributed to velocity and station correction estimates at Soultz. The bias may be accommodated in part by imposing an *a postiori* station correction, but this does not address the problem of separating actual pick bias from possible velocity heterogeneities. Removing this artifact from the catalogue using the waveform stacks is the subject of ongoing work; such a final processing step should help in future efforts to refine the very simple, uniform velocity model currently used for earthquake location at Soultz.

Summary

We have demonstrated an automatic phase-repicking algorithm to reduce the random component of earthquake location that arises from phase-picking inconsistency among similar earthquakes. We apply this technique to microearthquakes at the Soultz-sous-Forêts reservoir in the Rhine Graben of eastern France. The algorithm was tested using the results of careful, manual repicking on two isolated clusters of seismicity within the Soultz catalogue, then applied to the larger data set.

Results of hypocenter relocation on the test subset using the automatically adjusted picks demonstrate that the algorithm can recover detailed structures within the microseismicity cloud with a degree of improvement approaching that possible with careful, manual repicking. Median hypocenter discrepancies compared to PHILLIPS (2000) careful manual relocations improved from 31 to 7 m upon adjusting the preliminary phase picks with our automated algorithm. In the cases of distorted or clipped waveforms, such as with the Soultz hydrophone, human intervention may still be necessary if optimal adjustment of every phase for every earthquake is desired.

Application to a data set of ~7000 Soultz microearthquakes results in significant enhancement of seismogenic fractures and faults within the reservoir, consistent with detailed studies on smaller selected subsets. Event associations with nearest neighbors, estimated by counting the number of earthquakes within a 5 m radius, increases from 5 to 26% of the catalogue. These linear joint and fracture structures tend for the most part to be steeply dipping, with predominant northwest–southeast orientations, although northeasterly-striking fracture sets also comprise a significant portion of the resolved structure. The previously diffuse cloud also exhibits subparallel, *en echelon* linear features with a more northerly strike, among shallow

(2500–2900 m) events. Between 2900 and 3100 m we observe features which may represent intersecting faults or fractures, some dipping steeply to the east, others dipping steeply to the west. Overall, the image presented by relocated seismicity indicates that slip is occurring on fractures with varying orientations, whose characteristics change with depth. Evaluation of the relationships among these complex structures may yield better insights into reservoir flow behavior.

Correction of the phase arrival estimates to reduce random errors in the hypocenter location process can sharpen the results of subsequent applications including precise event location via SEL, JHD and/or collapsing techniques, as well as seismic tomography, reciprocal array processing, spatial b value or fractal dimension analysis and other imaging techniques which depend upon highly consistent phase picks. Additionally, the correlation, intra-cluster adjustment, stacking and subsequent stack correlation illuminate systematic biases in the preliminary phase picks which have contributed to inaccuracies in the initial velocity models. Identification and removal of these systematic data errors will allow for better modeling of the reservoir in the future.

Acknowledgements

The authors wish to thank D. Alde for help with data selection criteria. Preliminary phase picking for the full catalogue was undertaken by T. Thomas. H. Tobin and J. Schlue provided thorough and insightful reviews of the manuscript. We appreciate the helpful comments and insights of two anonymous reviewers. This work was supported by funding from the New Energy and Industrial Development Organization of Japan through the Laboratory of Hiroaki Niitsuma of Tohoku University, Sendai, Japan; support was also provided under the auspices of the U.S. Department of Energy by Sandia National Laboratory under Contract DE-AC04-94AL850000 and by the National Science Foundation Office of Polar Programs grant #9419267. Views expressed here are those of the authors and not necessarily those of the organizations for which they work.

References

Aster, R. C., and Scott, J. (1993), *Comprehensive Characterization of Waveform Similarity in Microearthquake Data Sets*, Bull. Seismol. Soc. Am. *83*, 1307–1314.

Aster, R. C., and Rowe, C. A. (2000), Automatic phase-pick refinement and similar event association in large seismic data sets. In *Advances in Seismic Event Location* (eds. Thurber, C., and Rabinowitz, N.) (Kluwer, Amsterdam, 1993) pp. 231–263.

Baria, R., Baumgärtner, J., Gerard, A., Jung, R., and Garnish, J. (1999), *European HDR Research Programme at Soultz-sous-Forêts (France) 1987–1996*, Geothermics *28*, 655–669.

Block, L. (1991), *Joint Hypocenter-Velocity Inversion of Local Earthquake Arrival Time Data in Two Geothermal Regions*, Ph.D. Dissertation, Massachusetts Institute of Technology, 448 pp.

BLOCK, L. V., CHENG, C. H., FEHLER, M. C., and PHILLIPS, W. S. (1994), *Seismic Imaging Using Microearthquakes Induced by Hydraulic Fracturing*, Geophysics *59*, 102–112.

CARR, D. B., YOUNG, C. J., HARRIS, J. M., ASTER, R. C. and ZHANG, X. (1999), *Cluster analysis for CTBT seismic monitoring*, Seismol. Res. Lett. *70*, 227–228.

CORNET, F. H., SHAPIRO, S. A., AUDIGRANCE, P., and ROYER, J. J. (2000), *Large-scale in situ Permeability Tensor of Rocks from Induced Microseismicity; Discussion and Reply*, Geophys. J. Int. *140*, 465–473.

CORNET, F. H., and JONES, R. (1994), *Field Evidence on the Orientation of Forced Water Flow with Respect to the Regional Principal Stress Directions*, Proc. 1st North Amer. Rock Mech. Symp., Austin, Texas, 61–69.

CORNET, F. H., and YIN, J. (1995), *Analysis of Induced Seismicity for Stress Field Determination and Pore Pressure Mapping*, Pure appl. geophys. *145*, 677–700.

DEICHMANN, N., and GARCIA-FERNANDEZ, M. (1992), *Rupture Geometry from High-precision Relative Hypocentre Locations of Microearthquake Cluster*s, Geophys. J. Int. *110*, 501–517.

DODGE, D. A., BEROZA, G. C., and ELLSWORTH, W. L. (1995), *Foreshock Sequence of the 1992 Landers, California Earthquake and its Implications for Earthquake Nucleation*, J. Geophys. Res. *100*, 9865–9880.

FEHLER, M., HOUSE, L., and KAIEDA, H. (1987), *Determining Planes along which Earthquakes Occur: Method and Application to Earthquakes Accompanying Hydraulic Fracturing*, J. Geophys. Res. *92*, 9407–9414.

FEHLER, M., PHILLIPS, W. S., JONES, R., HOUSE, L., ASTER, R., and ROWE, C. (2000), *A Method for Improving Relative Earthquake Location*s, Bull. Seismol. Soc. Am. *90*, 775–780.

FREMONT, M.-J., and MALONE, S. D. (1987), *High Precision Relative Locations of Earthquakes at Mount St. Helens, Washington*, J. Geophys. Res. *92*, 10,223–10,236.

GAUCHER, E., CORNET, F. H., and BERNARD, P. (1998), *Induced Seismicity Analysis for Structure Identification and Stress Field Determination*, Proc. SPE-ISPM Rock Mechanics in Petroleum Engineering Conference *1*, 545–554.

GENTER, A., and TRAINEAU, H. (1996), *Analysis of Macroscopic Fractures in Granite in the HDR Geothermal well EPS-1, Soultz-sous-Forêts, France*, J. Volc. Geotherm. Res. *72*, 121–141.

GILLARD, D., RUBIN, A. M., and OKUBO, P. (1996), *Highly Concentrated Seismicity Caused by Deformation of Kilauea's Deep Magma System*, Nature *384*, 343–346.

GOLDSTEIN, P., DODGE, D., FIRPO, M., and RUPPERT, S. (1998), *What's new in SAC2000? Enhanced Processing and Database Access*, Seismol. Res. Lett. *69*, 202–205.

GOT, J.-L., FRECHÊT, J., and KLEIN, F. W. (1994), *Deep Fault Plane Geometry Inferred from Multiplet Relative Relocation Beneath the South Flank of Kilauea*, J. Geophys. Res. *99*, 15,375–15,386.

HARRIS, M. and YOUNG, C. (1997), *MatSeis: A Seismic GUI and Tool-box for MATLAB*, Seism. Res. Lett. *68*, 267–269.

ITO, A. (1995), *High Resolution Relative Hypocenters of Similar Earthquakes by Cross-spectral Analysis Method*, J. Phys. Earth *33*, 279–294.

JONES, R. H., and STEWART, R. C. (1997), *A Method for Determining Significant Structures in a Cloud of Earthquake*s, J. Geophys. Res. *102*, 8245–8254.

JONES, R. H., BEAUCE, A., JUPE, A., FABRIOL, H., and DYER, B. C. (1995), *Imaging Induced Microseismicity during the 1993 Injection Tests at Soultz-sous-Forêts, France*, Proc. World Geotherm. Cong., Florence, 2665–2669.

JUNG (1991), *Hydraulic Fracturing and Hydraulic Testing in the Granitic Section of Borehole GPK1, Soultz-sous-Forêt*s, Geotherm. Sci. Technology *3*, 149–198.

JUPE, A. J., JONES, R. H., WILLIS-RICHARDS, J., DYER, B., NICHOLLS, J., and JACQUES, P. (1994), *Report on HDR Phase 4 – Soultz Experimental Programme 1993/1994, CSM Associates Ltd.*, IR02/12.

JUPE, A. J., JONES, R., DYER, B., and WILSON, S. (1998), *Monitoring and Management of Fractured Reservoirs Using Induced Microearthquake Activity*, Soc. Petroleum. Engineers contribution SPE 47315 in International Collaborative Project on Mapping of Engineered Geothermal Reservoirs, MTC Project Report.

KISSLING, E. (1988), *Geotomography with Local Earthquake Dat*a, Rev. Geophys. *26*, 659–698.

LANCE, G. N., and WILLIAMS, W. T. (1967), *A General Theory for Classificatory Sorting Strategies 1. Hierarchical Systems.*, Computer J. *10*, 271–276.

LEES, J. M. (1998), *Multiplet Analysis at Coso Geothermal*, Bull. Seism. Soc. Am. *88*, 1127–1143.

LUDWIG, J. A., and REYNOLDS, J. F. (1988), *Statistical Ecology, a Primer on Methods and Computing*, (John Wiley & Sons, New York) 337 pp.

MORIYA, H., NAKAZATO, K., and NIITSUMA, H. (2001), *Detailed Fracture System of Soultz-sous-Forêts HDR Field Evaluated by Using Microseismic Doublet/Multiplet*, Pure appl. geophys., submitted.

NADEAU, R., FOXALL, W., and MCEVILLY, T. (1995), *Clustering and Periodic Recurrence of Microearthquakes on the San Andreas Fault at Parkfield, California*, Science *267*, 503–507.

NANJO, K., HAGAHAMA, H., and SATOMURA, M. (1998), *Rates of Aftershock Decay and the Fractal Structure of Active Fault Systems*, Tectonophysics *287*, 173–186.

PARK, J., LINDBERG, C. R., and VERNON, F. L. III (1987), *Multitaper Spectral Analysis of High-frequency Seismograms*, J. Geophys. Res. *92*, 12,675–12,684.

PARKER, R. L., and MCNUTT, M. K. (1980), *Statistics for the One-norm Misfit Measure*, J. Geophys. Res. *85*, 4429–4430.

PHILLIPS, W. S., HOUSE, L. S., and FEHLER, M. C. (1997), *Detailed Joint Structure in a Geothermal Reservoir from Studies of Induced Microearthquake Clusters*, J. Geophys. Res. *102*, 11,745–11,763.

PHILLIPS, W. S. (2000), *Precise Microearthquake Locations and Fluid Flow in the Geothermal Reservoir at Soultz-sous-Forêts, France*, Bull. Seismol. Soc. Am. *90*, 212–228.

PHILLIPS, W. S., RUTLEDGE, J. T., HOUSE, L. S., and FEHLER, M. C. (2001), *Induced Microearthquake Patterns in Hydrocarbon and Geothermal Reservoirs: A Review*, Pure appl. geophys., submitted.

POUPINET, G., ELLSWORTH, W. L., and FRECHET, J. (1984), *Monitoring Velocity Variations in the Crust Using Earthquake Doublets: An Application to the Calaveras Fault, California*, J. Geophys. Res. *89*, 5719–5731.

PUJOL, J. (1992), *Joint Hypocentral Location in Media with Lateral Velocity Variations and Interpretation of the Station Corrections*, Phys. Earth Planet. Interiors *75*, 7–24.

REVIERE-BARBIER, F., and GRANT, L. T. (1993), *Identification and Location of Closely Spaced Mining Events*, Bull Seismol. Soc. Am. *3*, 1527–1546.

ROFF, A., PHILLIPS, W. S., and Brown, D. W. (1996), *Joint Structures Determined by Clustering Microearthquakes Using Waveform Amplitude Ratios*, Int. J. Rock Mech. Min. Sci. Geomech. Abstr. *33*, 627–639.

ROWE, C. A. (2000), *Correlation-Based Phase Pick Correction and Similar Earthquake Family Identification in Large Seismic Waveform Catalogues*, Ph.D. Dissertation., New Mexico Tech.

ROWE, C. A., ASTER, R. C., FEHLER, M. C., PHILLIPS, W. S., and Kyle, P. R. (1998), *Hypocenter Relocation in Tectonic and Volcanic Seismic Data Sets Using Automatic Phase Correlation: Application of Cross Correlation, Signal Envelope and Principal Components Methods to Improve Pick Consistency*, EOS *79*, F588.

ROWE, C. A., and ASTER, R. C. (1999). *Application of Automatic, Adaptive Filtering and Eigenspectral Techniques to Large Digital Waveform Catalogs for Improved Phase Pick Consistency and Uncertainty Estimates*, EOS *80*, F660.

ROWE, C. A., ASTER, R. C., BORCHERS, B., and YOUNG, C. J. (2001), *An Automatic, Adaptive Algorithm for Refining Phase Picks in Large Seismic Data Sets, Part I: Technique*, Bull. Seismol. Soc. Am., in review.

RUBIN, A. M., GILLARD, D., and GOT, J.-L. (1998), *A Reinterpretation of Seismicity Associated with the January 1983 Dike Intrusion at Kilauea Volcano, Hawaii*, J. Geophys. Res. *103*, 10,003–10,015.

RUBIN, A. M., GILLARD, D., and GOT J.-L. (1999), *Streaks of Microearthquakes Along Creeping Faults*, Nature *400*, 635–641.

SHEARER, P. M. (1997), *Improving Local Earthquake Locations Using the L_1 Norm and Waveform. Cross Correlation: Application to the Whittier Narrows, California, Aftershock Sequence*, J. Geophys. Res. *102*, 8269–8283.

SHEARER, P. M. (1998), *Evidence from a Cluster of Small Earthquakes for a Fault at 18 km Depth Beneath Oak Ridge, Southern California*, Bull. Seismol. Soc. Am. *88*, 1327–1336.

SNOKE, J. A. (1987), *A Stable Determination of (Brune) Stress Drops*, Bull. Seismol. Soc. Am. *77*, 530–538.

SPUDICH, P., and BOSTWICK, T. (1987), *Studies of the Seismic Coda Using an Earthquake Cluster as a Deeply Buried Seismograph Array*, J. Geophys. Res. *92*, 10,526–10,546.

STARZEC, P., FEHLER, M., BARIA, R., and NIITSUMA, H. (2000), *Spatial Correlation of Seismic Slip at the HDR-Soultz Geothermal Site: Qualitative Approach*, Bull. Seismol. Soc. Am. *90*, 1528–1534.

THOMSON, J. D. (1982), *Spectrum Estimation and Harmonic Analysis*, Proc. IEEE *70*, 1055–1096.

THURBER, C., ROECKER, S., MCEVILLY, T., HASLINGER, F., STEINBERG, D., and ELLSWORTH, W. (2000), *Earthscope at Parkfield, California: Setting the Stage for Drilling*, EOS *81*, S318.

WALDHAUSER, F. ELLSWORTH, W. L., and COLE, A. (1999), *Slip-parallel Seismic Lineations on the Northern Hayward Fault, California*, Geophys. Res. Lett. *26*, 3525–3528.

WIEMER, S. and WYSS, M. (1997), *Mapping the Frequency-magnitude Distribution in Asperities: An Improved Technique to Calculate Recurrence Times?* J. Geophys. Res. *102*, 15,115–15,128.

(Received October 30, 2000, revised/accepted February 2, 2001)