Microearthquake cluster detection based on waveform similarities, with an application to the western Swiss Alps

Hansruedi Maurer and Nicholas Deichmann
Institute of Geophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

Accepted 1995 May 29. Received 1995 May 26; in original form 1995 February 2

SUMMARY
Highly similar waveforms of different earthquakes are due to similar focal mechanisms and common propagation paths. The relative hypocentral locations of events in clusters of similar earthquakes can provide useful insights into geometry and style of faulting at depth within the crust. The detection of such earthquake clusters within large data sets can only be accomplished efficiently by means of an automatic procedure. Therefore, we developed an algorithm based on correlation analysis that detects and associates events with similar waveforms. The algorithm has been applied to a data set recorded in the western Swiss Alps: 619 out of a total of 1497 events exhibit similarities with other events. Based on a more detailed investigation of two selected clusters with known focal mechanisms, it could be shown that the active fault planes correspond to neotectonic structures mapped in the study area. Due to their oblique orientation relative to the larger-scale epicentral alignment, these faults have been interpreted as Riedel shears.

Key words: Alps, fault-plane solutions, microearthquakes, Switzerland, waveform analysis.

1 INTRODUCTION
Foreshock and aftershock sequences as well as earthquake swarms are often characterized by numerous events with striking waveform similarity (e.g. Tsujiura 1983). Conditions for similarity are fulfilled if all factors contributing to the seismograms (source-time function, propagation path, station site and recording instrument) are essentially identical. In this context, source-time functions are regarded as similar if the events have a common focal mechanism and similar magnitudes. Identical propagation paths imply that the hypocentres must be tightly clustered in space compared to the dimensions of near-source heterogeneities and to the observed wavelengths. Such events are usually referred to as clusters or earthquake families. The waveform similarity allows a highly consistent determination of phase onsets that can be used for precise relative hypocentre locations of the events pertaining to a given cluster (e.g. Poupinet, Ellsworth & Frechet 1984).

The large amount of digital data recorded by modern seismic networks serves as an excellent data base for an analysis of similar events. With the increasing number of seismograms, however, manual analysis becomes an impossible chore. In the frame of this study we present an algorithm that detects and associates reliably events with similar waveforms. Others have recognized the usefulness of such an algorithm as well. Thus Aster & Scott (1993) implemented a method which uses the same basic approach as ours to detect similar events but differs in the way these events are then grouped into individual clusters.

The first part of this paper describes our cluster detection algorithm. To test its performance, it was applied to a data set of local earthquakes recorded in northern Switzerland. The data have been already analysed manually by Deichmann & Garcia-Fernandez (1992) and can therefore be used to check the detection algorithm.

The second part of the paper describes an application of the method to a new data set recorded in the western Swiss Alps. After a brief introduction to the tectonic setting and the data used, we present the results of the cluster detection. By means of subsequently performed relative locations of two selected swarms, we map active fault systems at depth within the crust and examine their relationship to neotectonic features at the surface.

2 CLUSTER DETECTION ALGORITHM
2.1 Method of the algorithm
The requirements for the proposed algorithm are on-line access to waveform data, phase readings and hypocentral parameters. Strictly speaking, hypocentral parameters are not necessarily required, but where there are large data sets their use significantly reduces the computational effort. Two earthquakes are
A possible solution is to calculate the mean or the median of all corresponding non-zero elements of the individual station correlation matrices. Tests with different data sets have shown that where there are only a few \( cc_{ij}^d \) values of a particular event pair, neither mean nor median is sufficiently robust against data blunders like grossly mispicked phase onsets or low signal-to-noise ratios. To find appropriate statistics, it must be considered that it is more likely that blunders decrease \( cc_{ij}^d \) than that they will accidentally raise them. Therefore, we implemented an asymmetrically trimmed 'mean', \( \bar{cc}_{ij} \), where the lowest values are removed before the mean is calculated (eq. 3):

\[
\bar{cc}_{ij} = \frac{1}{n-q-1} \sum_{q}^{n-1} cc_{ij}^d.
\]

Here \( cc_{ij}^s \) contains the \( cc_{ij}^d \) values sorted in ascending order, \( n \) is the number of stations that recorded the \( i \)th and \( j \)th events, and \( q \) is the first element of \( cc_{ij}^s \) that is not removed (\( q = nK/100 \)). The value of \( K \) is of the order of 25 per cent and must be determined by trial and error.

S-wave data often provide a more reliable measure of similarity than P-wave data. This is due to the fact that their waveforms are generally more complex than the first arriving P-waves. Their complexity allows a good discrimination of similar waveforms, whereas the first arriving P-waves may lead to high cross-correlation coefficients even when there is no similarity in the remaining parts of the seismograms. Our cluster detection algorithm is therefore based mainly on \( \bar{cc}^s \).

The decision whether an event pair is similar or not is made on the basis of a threshold value \( T_p \): \n
\[
\bar{cc}_{ij} = \begin{cases} 
cc_{ij} & \text{if } -T_p < \bar{cc}_{ij} < T_p \\
0 & \text{otherwise.}
\end{cases}
\]

It is possible that two S wavelets are similar and the P-wave onsets have reversed polarities (Fig. 1). This may happen if the station is very close to a nodal plane and if a cluster consists of events with slightly different focal mechanisms. In order to separate such subclusters, a second threshold \( T_p \) is introduced which acts in the same way as \( T_p \) and leads to the

![Figure 1. Example of two seismograms with similar S waveforms but reversed P-wave polarity.](https://academic.oup.com/gji/article-abstract/123/2/588/558979)

© 1995 RAS, GJI 123, 588–600
The definition of the network correlation matrix \( \tilde{cc}_{ij} \):

\[
\tilde{cc}_{ij} = \begin{cases} 
    \tilde{cc}_{ij}, & |\tilde{cc}_{ij}| > T_p, \\
    0, & \text{otherwise}
\end{cases}
\]

\( T_p \) can also be used to extract events with similar S wavelets and reversed P polarities. They may, for example, provide information on temporal variations of the fault-plane orientation.

In the next step an algorithm that detects clusters of similar events must be found. A possible solution is the equivalence-class (EC) approach (Press et al. 1988), which follows a simple scheme: if elements \( A \) and \( B \) are similar and elements \( A \) and \( C \) are similar it is concluded that elements \( B \) and \( C \) are also similar. Astur & Scott (1993) applied such an algorithm to local earthquake data.

EC algorithms are very fast but they have a disadvantage that may be important in earthquake cluster analysis problems. We assume that a data set contains two distinct clusters and that the waveforms of both clusters have a weak similarity. \( T_s \) is usually of the order of 0.7–0.9 and therefore it is possible that events belonging to different clusters are treated as similar events. The EC approach would then result in a large cluster that contains both groups of events. With our algorithm we take into account the structure of \( \tilde{cc}_{ij} \) in order to obtain a more reliable cluster association.

Events that belong to the same group should have the same or at least similar \( \tilde{cc}_{ij} \) row-patterns. To enhance the similarity of the rows of a cluster and to remove elements that are accidentally common to other clusters, we define a modified network correlation matrix \( mc_{cij} \), in which the \( cc_{ij} \) values are replaced by the corresponding normalized scalar product defined as

\[
mc_{cij} = \frac{\sum_{k=1}^{\text{events}} \tilde{cc}_{ik} \tilde{cc}_{jk}}{\sqrt{\sum_{k=1}^{\text{events}} \tilde{cc}_{ik}^2 \sum_{k=1}^{\text{events}} \tilde{cc}_{jk}^2}}.
\]

Then, a cluster separation threshold \( T_c \) is introduced. \( mc_{cij} \) values above \( T_c \) are set to 1 and the remaining elements are set to 0. With this procedure it is possible to 'decouple' row patterns of different clusters.

The association of events to clusters can be performed according to the following scheme.

1. Take the first row of the \( mc \) matrix and calculate the scalar product with all other rows.
2. If the scalar product is higher than an event association threshold \( T_p \), an event associated with the first event is detected. If an appropriate value of \( T_p \) is chosen, \( T_c \) can be set to 1.
3. Store the row number of the associated event and delete the row of the associated event.
4. Take rows 2, 3, ..., (last – 1) and repeat the procedure.

The remaining associated event matrix \( (aem_t) \) contains the row patterns of the different clusters. The row index \( i \) indicates the cluster number. Cluster members are characterized by non-zero elements of a specific row of the \( aem_t \).

The efficiency of the algorithm described above depends heavily on the proper choice of the various thresholds. While the determination of \( T_s \), \( T_p \) and \( T_c \) is straightforward, a proper choice of \( T_p \) may require extensive testing. In the following, we will give some guidelines for an appropriate selection of threshold values.

1. **S-wave similarity** \( (T_s) \). The S-wave CC threshold depends on the signal quality and the recording parameters (sampling rate, system transfer function, dynamic range). For short-period data \((\approx 100 \text{ Hz sampling rate}, 1–20 \text{ Hz passband and } \approx 70–80 \text{ dB dynamic range})\) of local earthquakes (recording distances up to 100 km), \( T_s \) values between 0.75 and 0.95 work well.

2. **P-wave similarity** \( (T_p) \). The P-wave similarity threshold is not as critical as the S-wave threshold. A value of 0.0 works well most of the time, but if a more restrictive detection is attempted, \( T_p \) can be chosen on the order of \( T_s \).

3. **Cluster separation** \( (T_c) \). An appropriate choice of the threshold for the normalized scalar product is somewhat tricky. A high threshold may lead to a very sparse \( mc_{cij} \) and destroyed row patterns. Conversely, a low \( T_c \) fills the \( mc_{cij} \) and the desired separation of different clusters is not achieved. Furthermore, there is a trade-off between \( T_p \) and \( T_c \). Based on our experience we propose a \( T_c \) value that is of the order of 0.5.

4. **Event association** \( (T_p) \). If the cluster separation of \( T_c \) is done successfully, \( T_c \) can be set equal to one. With the introduction of a larger \( T_p \) value, the cluster association becomes more robust against spuriously grouped events.

These rules of thumb are valid in most cases but in a few exceptional situations they may not hold. It is therefore advisable to select a small subset out of the full data set and perform a manual search over the selected data. Then the subset is re-analysed with our algorithm. This allows a fine tuning of the different thresholds.

### 2.2 Test of the algorithm

Since 1983, the seismicity of northern Switzerland has been monitored with a dense station network (e.g. Deichmann 1992). It is characteristic for this region that the epicentres are not distributed uniformly over the whole area but are concentrated in several clusters. Deichmann & Garcia-Fernandez (1992) investigated three of these clusters in detail. Due to the relatively small amount of data, it was possible to perform the cluster detection and association manually; therefore, these data are well suited for a performance test of our automatic cluster detection algorithm. In addition to the data used by Deichmann & Garcia-Fernandez (1992), we also included more recent events (1990–93) that occurred in the investigation area. As an example, seismograms of all events recorded at one particular station are shown in Fig. 2. For more seismograms from these clusters as well as for details on station and hypocentral locations, see the paper by Deichmann & Garcia-Fernandez (1992).

The results of the cluster detection are summarized in Fig. 3. The \( \tilde{cc} \) matrix is shown in Fig. 3(a) whereas \( mc_{cij} \) is shown in Fig. 3(b). Fig. 3(c) shows the associated event matrix \( aem_t \).

Two clusters in the data set, denoted as B and C, are the Laufelfingen clusters analysed by Deichmann & Garcia-Fernandez (1992). Both clusters are properly detected and associated. In addition to the events found by Deichmann & Garcia-Fernandez, the algorithm associated event no. 17 with

© 1995 RAS, GJI 123, 588–600
Microearthquake cluster detection

Figure 2. Example of one set of signals used for the cluster detection performance test. The vertical-component seismograms shown were recorded by station GEF, at an epicentral distance of about 7 km (Deichmann & Garcia-Fernandez 1992). The numbers above the seismograms correspond to the event numbers in Figs 3 and 4. Capital letters denote different clusters.

The type B events. The seismogram of event 17, shown in Fig. 2, exhibits a reversed P-wave polarity compared to the remaining signals of that series. It has nevertheless been associated with the B type series, because the signals of event 17 recorded at all other stations are highly similar to those of the other events. Such apparently erroneous associations, however, may have important tectonic implications. It might be possible that these events provide information about a possible curvature or irregularity of the fault. A further cluster, the Zeglingen swarm (denoted by A in Fig 2), has also been identified by the algorithm. In addition to the triplet found by Deichmann & Garcia-Fernandez (1992), the algorithm detected two events that also belong to the Zeglingen swarm. Moreover, in the data set recorded between 1990 and 1993, the algorithm found two previously unknown clusters that have been confirmed by visual inspection (see Fig. 2).

Assuming that the events are sorted by origin times, Fig. 3(c) not only visualizes the cluster association but shows also their temporal distribution. The occurrence of cluster A, for instance, extends over the whole nine-year-long observation period, whereas cluster B is restricted to a short time span of only two days.
Finally, we compare our algorithm to the method proposed by Aster & Scott (1993). The first part of their procedure, which leads to the individual $c_{ij}$, is similar to our solution, but in contrast to our algorithm they obtain $c_{ij}$ by calculating the median and they do not consider the $P$ wave train. For the cluster detection and association they use the EC approach. Fig. 4 shows the results of applying the two procedures to the data set from northern Switzerland for various threshold values $T$. The $T$ parameter is called the $\beta$ value in the paper of Aster & Scott (1993). A $T$ threshold of 0.85 leads to the best results. As shown in Fig. 4, the EC approach is very sensitive with respect to $T$: small variations of $T$ lead to dramatic changes in the cluster association, whereas our solution remains more stable over a wide range of $T$ values.

3 APPLICATION

3.1 Tectonic setting of the western Swiss Alps

The location and the major tectonic units of the western Swiss Alps are shown in Fig. 5. Most of the northern part of the study area (i.e. north of the Rhone Valley) is covered by the Helvetic nappes, which represent the former sedimentary cover of the external crystalline massifs. The study area is bound by the Aar Massif to the east and by the Aiguilles-Rouge Massif to the west. The Helvetic nappes and crystalline massifs are considered to have been part of the former European plate (e.g. Pfiffner 1992).

South of the Rhone Valley, the dominant Penninic nappes are interpreted to comprise slices of crystalline crustal material derived from the former Adriatic and Briançonnais microplates as well as the ancient Alpine oceans (Stampfl 1993). Due to strong deformation and metamorphism imposed during the Alpine plate collisions, it is difficult to determine with confidence the former locations and tectonic environments of the various Penninic units. Moreover, the structure of this entire region is complicated by the presence of a near 90° change in trend of the Alpine mountain chain (e.g. Escher, Masson & Steck 1993).

The western Swiss Alps are clearly the seismically most active zone in Switzerland (Pavoni 1977). During the last centuries, several moderate earthquakes, which caused significant damage to buildings, occurred in this region. The most prominent event in the 20th century is the 1946 Wildhorn event (Wanner 1955). Its epicentre was situated in the area of the Rawil Depression (Fig. 5) and, with an estimated magnitude between 5 and 6, it caused some damage to buildings in the town of Sierre. It is thus also referred to as the Rawil or Sierre earthquake. The aftershock region of this earthquake was investigated in detail by Pavoni (1980). Based on a joint interpretation of neotectonic structures and of fault planes determined from first motions of microearthquake seismog-
3.2 Earthquake data

In this study, we used digital seismograms recorded between 1983 and 1992. During the whole period, the seismicity of the western Swiss Alps was monitored by stations of the permanent network of the Swiss Seismological Service (SSS). As shown in Fig. 5, stations of the permanent SSS network are situated in and around the study area. The signals of the SSS network are telemetered in analogue form by frequency modulation to the national data centre in Zurich. After demodulation and anti-alias filtering (12.5 Hz corner frequency), the signals are digitized with a sampling rate of 64 Hz for automatic processing (Baer 1990).

In addition, data from a six-station network covering an area of approximately 3 × 3 km around the lake of Zeuzier (Fig. 5) were used. These signals were also telemetered in analogue form to a central recording site, where they were digitized with a sampling rate of 200 Hz.

Between 1989 and 1991, the permanent network was complemented with 11 additional temporary stations (Fig. 5), with the aim of investigating the seismicity in the western Swiss Alps in more detail (Maurer 1993). The Reimer MLR2 instruments, which were used for this study, continuously record analogue data on four-track magnetic tapes (e.g. Maurer 1993). The recordings on the analogue field tapes were later digitized in the laboratory at a sampling rate of 128 Hz.

Figure 4. Comparison of the cluster detection algorithm proposed in this study with the one of Aster & Scott (1993). The clusters in each panel are labelled consecutively (A, B, C,...). See the text for further explanation.
Absolute event locations were determined by joint inversion of hypocentral parameters and velocity structure (Maurer & Kradolfer 1995). The final locations of the whole data set, which comprises 1497 events, are shown in Fig. 6. The epicentral distribution north of the Rhone Valley is more structured than in the south. A conspicuous concentration of epicentres extends over the Helvetic nappes in the area of the Rawil Depression. It coincides with the epicentral region of the 1946 Wildhorn event. The north-eastern part of the investigation area, where the Aar Massif outcrops at the surface, was almost completely aseismic during the whole recording period. The epicentral distribution south of the Rhone Valley is rather diffuse and thus more difficult to characterize.

3.3 Cluster analysis

All events included in Fig. 6 are used for the cluster analysis. The search radius \( r \) was set to 5 km. A total number of 381,821 cross-correlations were calculated for \( P \) phases and 413,782 for \( S \) phases. Provided the available computer memory is sufficient to store the relevant trace segments of the whole data set, the algorithm is quite efficient: processing the 1497 events required less than two hours of CPU time on an HP9000/720 workstation.

Fig. 7 shows a histogram of the correlation coefficients. The maxima of the \( P \) distribution are found at 0.6 and \(-0.5\) while the \( S \) maxima are at 0.35 and \(-0.35\). The differences between the histograms are a consequence of differences in signal complexity. Due to the more simple shape of the \( P \) wave train, a similarity of two signals is more likely for \( P \) than for \( S \), which explains the broader distribution for the \( P \)-phase correlation coefficients. On the other hand, the coefficients for \( S \) clearly deviate from a pure normal distribution between 0.6 and 1.0. This implies that the data set contains a significant number of similar events. It should be noted that Fig. 7
contains only coefficients of event pairs that passed the proximity test. Therefore, the histograms reflect only the relative difference between P- and S-wave similarities and do not characterize general waveform similarities of a specific region.

After several tests, we used values of 0.80 for $T_s$ and 0.6 for $T_p$. $T_p$ was set to 0.5 and $T_s$ to 1.0. Figs 8(a) and (b) display reduced forms of $\tilde{c}_{ij}$ and $mcc_{ij}$: only those events with at least two non-zero $\tilde{c}_{ij}$ elements in a row or column are shown. The same subset of events is also depicted in Fig. 9. The $aem_{ij}$ matrix is displayed in Fig. 8(c). In addition to numerous small clusters with two or three events, we recognize several conspicuous patterns in Fig. 8(c). Clusters 14 and 20 contain a large number of events (74 and 47, respectively) and scatter in time over the whole observation period. Both clusters lie in the region of Ifigenalp (Fig. 9). Unfortunately, they were only recorded by the Zeuzier network (Fig. 5). Due to the large epicentral distances of the Zeuzier stations (10 km) and therefore unfavourable location geometry, a more detailed analysis of these events is not possible. Besides these two large clusters, we observe several earthquake swarms that are located in the region of Bonneveaux (Fig. 9).
Figure 8. Cluster analysis of the western Swiss Alps data. Thresholds used: $T_x=0.8$, $T_y=0.6$, $T_z=0.5$, $T_p=1$. See the text for further explanation.

Figure 9. Epicentre map of those events that could be associated with at least one other event (619 events).
In the next section we will concentrate on cluster numbers 111 and 138. They are situated near the town of Anzère (46.33°N/7.3°E), at a depth of 6–8 km within the earthquake alignment of the Wildhorn zone (Fig. 9). These clusters are of special interest because it was possible to construct well-constrained fault-plane solutions for one member of each cluster.

4 RELATIVE LOCATION OF THE ANZÈRE CLUSTERS

4.1 The events

The events associated with cluster numbers 111 and 138 are listed in chronological order in Table 1. They are subsequently denoted as type I (cluster 111) and type II (cluster 138). Their absolute epicentral locations scatter over an area of 2 × 2 km, but their signals are sufficiently similar that a slightly lower Tc threshold in the detection algorithm causes them to coalesce into a single larger cluster. Visual inspection of the seismograms, however, clearly shows that they pertain to two different event types. In fact, as will be shown later on, the two signal types correspond to two different fault-plane solutions, with similar strike but slightly different dip of the nodal planes (Table 2 and Fig. 10). Event no. 2, which was considerably stronger than the others \((M_L = 3.8)\), was actually not recognized by the automatic procedure as being part of the cluster. Many of the records of this event are clipped and the frequency content is considerably lower than that of the other events. Consequently, the resulting correlation coefficient is too low. However, analysis of first-motion polarities and inspection of the unclipped records show that it is indeed a member of the type I events.

Table 1. Events in cluster numbers 111 (Type I) and 138 (Type II). Master events are marked by M. Events with magnitudes \(M_L < 1.2\) were not recorded by the national network.

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
<th>Time</th>
<th>(M_L)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1989.09.27</td>
<td>21:56:33</td>
<td>1.6</td>
<td>I</td>
</tr>
<tr>
<td>2</td>
<td>1989.09.30</td>
<td>04:41:02</td>
<td>3.8</td>
<td>I</td>
</tr>
<tr>
<td>M 3</td>
<td>1989.09.30</td>
<td>05:26:43</td>
<td>2.2</td>
<td>I</td>
</tr>
<tr>
<td>4</td>
<td>1989.10.01</td>
<td>14:08:27</td>
<td>1.6</td>
<td>I</td>
</tr>
<tr>
<td>5</td>
<td>1989.10.02</td>
<td>03:05:59</td>
<td>1.6</td>
<td>I</td>
</tr>
<tr>
<td>M 6</td>
<td>1990.07.19</td>
<td>23:38:35</td>
<td>1.7</td>
<td>II</td>
</tr>
<tr>
<td>7</td>
<td>1990.07.20</td>
<td>05:27:24</td>
<td>1.8</td>
<td>I</td>
</tr>
<tr>
<td>8</td>
<td>1990.07.26</td>
<td>12:30:14</td>
<td>2.4</td>
<td>II</td>
</tr>
<tr>
<td>9</td>
<td>1990.07.26</td>
<td>15:45:04</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>10</td>
<td>1990.07.30</td>
<td>06:59:25</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>11</td>
<td>1990.08.02</td>
<td>08:23:53</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>12</td>
<td>1990.08.02</td>
<td>23:39:11</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>13</td>
<td>1990.08.13</td>
<td>10:52:46</td>
<td>1.8</td>
<td>I</td>
</tr>
<tr>
<td>14</td>
<td>1990.08.13</td>
<td>23:34:53</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>15</td>
<td>1990.08.14</td>
<td>11:07:26</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>16</td>
<td>1990.08.15</td>
<td>22:03:04</td>
<td>1.4</td>
<td>II</td>
</tr>
<tr>
<td>17</td>
<td>1990.08.19</td>
<td>12:00:37</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>18</td>
<td>1990.09.15</td>
<td>14:14:43</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
<tr>
<td>19</td>
<td>1991.01.14</td>
<td>03:23:47</td>
<td>2.0</td>
<td>I</td>
</tr>
<tr>
<td>20</td>
<td>1991.02.14</td>
<td>14:38:10</td>
<td>1.7</td>
<td>I</td>
</tr>
<tr>
<td>21</td>
<td>1991.04.15</td>
<td>19:16:33</td>
<td>&lt;1.2</td>
<td>II</td>
</tr>
</tbody>
</table>

Table 2. Focal mechanism parameters for the two Anzère clusters. Nodal planes: strike and dip with direction of dip. \(P\) and \(T\) axes: azimuth and plunge.

<table>
<thead>
<tr>
<th>Type</th>
<th>Fault plane</th>
<th>Aux. plane</th>
<th>(P)-axis</th>
<th>(T)-axis</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>110/90</td>
<td>200/50NW</td>
<td>163/27</td>
<td>57/27</td>
</tr>
<tr>
<td>II</td>
<td>285/80NE</td>
<td>187/51NW</td>
<td>154/35</td>
<td>50/19</td>
</tr>
</tbody>
</table>

Figure 10. Focal mechanisms and relative locations of the two Anzère clusters. The arrows labelled B and C in the epicentre map (a) point in the viewing direction for the two vertical cross-sections shown below. The straight lines through each cluster in the cross-sections correspond to the traces of the fault planes (in b) and of the auxiliary planes (in c). The size of the crosses is proportional to the calculated standard deviations of the relative locations. The origin is set to the location of the master event of cluster I. The numbers correspond to those in Table 1. Events 2 and 20 are located exactly below event 1, so that they are not visible in the map view. The fault-plane solutions are equal-area, lower-hemisphere projections with full circles corresponding to compression and empty circles to dilatation.
High-resolution relative locations of the events in each cluster and of the two clusters relative to each other were performed in order to answer two main questions.

1. Based on the relative hypocentral distribution, which of the two possible planes in the fault-plane solution was the active fault?
2. What is the relationship of the two clusters to each other and to other neotectonic features in the epicentral region?

4.2 Location method

The accuracy of absolute hypocentre locations is mainly controlled by reading errors, velocity model errors and non-linear effects (Pavlis 1986). It is possible to subdivide these error sources into systematic and random errors. Model errors and non-linear effects are known as systematic errors. Reading errors, on the other hand, are often considered random but, strictly speaking, they also contain a systematic component: based on experience, the analyst tries to pick phase onsets as consistently as possible, but there is no ultimate certainty that the chosen points in the seismograms reflect the true onset of the expected wavetype. When picking S-wave onsets on vertical-component seismograms, the problem may become very serious.

Given a high degree of signal similarity, the consistency of onset determinations can be greatly improved with the help of cross-correlation techniques, either in the time domain (e.g. Deichmann & Garcia-Fernandez 1992) or in the frequency domain (e.g. Poupinet et al. 1984). For digital waveform data, as they were used in this study, it is possible to obtain a relative timing precision of the order of a few milliseconds.

Using a location procedure, based either on a joint determination of hypocentres or on a master event technique, such precise relative arrival-time determinations allow one to calculate the location of events within a particular cluster with errors of only a few tens of metres.

In this paper we followed the procedure adopted by Deichmann & Garcia-Fernandez (1992). Onsets are determined by cross-correlation in the time domain together with a least-squares adjustment of the resulting arrival times for all possible event pairs in each cluster. Relative hypocentre locations are obtained with a master event technique. The location relative to the master events marked in Table 1 was performed in this way for all the events of type I, except for event no. 2, and for five out of the 14 events of type II. Because of several clipped records and the large difference in frequency content, event no. 2 was relocated relative to the type I master event in a separate step. Finally, using only those stations with sufficiently similar signals, the master events of the two clusters were located relative to each other.

4.3 Location results

The results of the relocation procedure and a comparison with the fault-plane solutions (Table 2) are shown in Fig. 10.

1. For both clusters the active fault plane could be determined unambiguously. Fig. 10 shows clearly that the hypocentral alignment coincides only with the approximately WNW–ESE striking planes.
2. Despite several clipped seismograms, it was possible to determine reliably the position of the $M_L$ event relative to the type I cluster (event no. 2 in Figs 10(b) and (c)). Though the accuracy of the $M_L$ event location is lower than for the other cluster members, it is evident that the cluster and the mainshock share a common fault plane.
3. Figs 10(a) and (b) show that the two clusters did not occur on the same fault plane. Although the fault orientations are nearly parallel, the clusters are about 1 km apart.
4. Table 1 shows that inter-event times within a particular cluster range from less than an hour to more than a year. Although there is no sequential correlation between the time of occurrence of an event and its relative location on the fault, the later events tend to lie deeper relative to the earlier ones and thus serve to increase the spatial extent of the cluster. This pattern is particularly evident for the type I events (Fig. 10c).

5 DISCUSSION AND CONCLUSIONS

5.1 Seismotectonic implications

As mentioned earlier and shown in Fig. 6, a zone of increased seismicity extends over the northern part of the study area (Wildhorn Zone). Fig. 11 shows an epicentre map and depth cross-section that contains only events with well-constrained hypocentre locations. The hypocentre alignment appears to define a near-vertical fault zone, which strikes in an ENE–WSW direction.

Surprisingly, the fault planes determined by first motion analysis and relative locations show a WNW–ESE strike (Fig. 10). These differences in orientation may be explained with the concept of Riedel shears. They were discovered by Riedel (1929) with clay deformation experiments in the laboratory. Riedel shears are secondary, conjugated fault systems that form at angles of about 15° and 85° to the strike of the main shear zone (e.g. Twiss & Moores 1992). Riedel shears have also been observed in connection with earthquakes: Tchalenko & Ambroseys (1970), for instance, investigated a faulting pattern in the area of the 1968 Dasht-e-Bayaz earthquake (Iran) that they interpreted as a system of Riedel shears.

Fig. 11(a) shows a horizontal cross-section through the Helvetic domain of the western Alps (Burkhard 1988). We recognize recent faults that cut all other tectonic structures and exhibit a strike that is parallel to the active fault planes of the two Anzère clusters. We therefore conclude that the earthquake clusters occurred on such recent faults. On the other hand, Pavoni (1980) described neotectonic fault patterns that are parallel to the main epicentre alignment in the Wildhorn zone (Fig. 11b). Moreover, other fault-plane solutions of recent microearthquakes in the same area also agree with this larger-scale alignment (Maurer 1993). Therefore, the contemporary seismicity in the Helvetic domain of the northern Valais appears to be the result of a reactivation of both the main shear zone and associated Riedel shears.

5.2 Methodological considerations

The preceding example demonstrates how the analysis of earthquake clusters with similar events can contribute useful information to our understanding of the seismotectonics of a region. If the epicentre distribution is sparse and the locations are reliable, clustering can be easily detected by eye from routinely drawn epicentre plots, and events with similar
locations occurring over a short time span stand out readily in most earthquake catalogues. However, in regions of increased seismicity or unfavourable station geometry, absolute hypocentre locations are not accurate enough to resolve individual clusters. Moreover, when similar events occur over an extended period of time with long intervals between them, manual detection becomes a matter of chance. One should also note that frequently the later events tend to increase significantly the spatial extent of a cluster. Therefore, detecting and integrating these later events in the relocation procedure contributes greatly to the reliable identification of the active fault. In these cases, an algorithm to identify clusters based on waveform similarity is essential.

Such a cluster detection algorithm can serve to search large data sets systematically for similar events that can be associated with earthquakes with known fault-plane solutions. High-resolution relative locations of the hypocentres will then produce systematic maps of active faults at depth within the crust.

Since families of similar earthquakes seem to be ubiquitous, cluster detection based on waveform similarities can also serve as an independent quality control for absolute hypocentre locations. In fact, earthquakes with similar signals, as recorded by the usual short-period seismographs, must be located not more than a few hundred metres apart. The distance between hypocentres pertaining to a particular cluster, as calculated from their routinely determined absolute locations, can thus be regarded as a measure of the consistency and relative precision of the routine location procedure.

ACKNOWLEDGMENTS

M. Baer contributed significantly to this project with his software for the acquisition and analysis of seismic data at the Swiss Seismological Service. C. Bärlocher's and M. Grieder's help with installation and maintenance of the recording instruments is gratefully acknowledged. Logistic support for running

© 1995 RAS, GJI 123, 588–600
the temporary stations was provided by the Centre de Recherches Scientifiques Fondamentales et Appliquées, Sion. We thank M. Burkhard, N. Pavoni, D. Mayer-Rosa and St. Mueller for fruitful discussions and active support in the course of this study and J. Ansorge and K. Holliger for their helpful comments on the manuscript. Furthermore, we owe thanks to Frank Scherbaum and Torsten Dahm who improved the clarity of the article by their careful reviews.

Contribution number 844 of the Institute of Geophysics, ETH-Zürich.

REFERENCES