Sorting source parameters to produce coherent record sections;
a geophysical jigsaw puzzle

Master thesis project by:
Nineke Oerlemans

Supervisors:
Prof. T. Heaton, Caltech
Dr. H. Paulssen, Utrecht University

July 1999

Abstract

In this research, cross-correlation is used to determine the similarities in waveforms between local events recorded at the same station. From the theory of reciprocity and knowledge of finite strains in the earth it is expected that the recordings of two identical events that occur very close in space will be the same at one recording station. As the location of one of the two events begins to differ gradually, the recorded waveform will change gradually until the two events are so far apart that the waveforms are no longer coherent. A similar gradual change will take place if the locations are kept fixed and the focal mechanisms, including magnitude, change gradually. The possibility of changing the location and/or the focal mechanism gives a total of seven degrees of freedom in making gradually changing record sections: three dimensions spatially; strike, dip and rake; and the magnitude.

Data from Long Valley recorded at station MLAC and data from the Northridge aftershock sequence recorded at station PAS are used. The waveforms from MLAC form very coherent clusters, but it is not yet possible to get a gradual change from one characteristic cluster to the next. More events will have to be recorded in order to make gradually changing record sections.

For the Northridge sequence, some cross-correlation coefficients indicated that the vertical location from the used datasets (E. Hauksson and J. Mori) were not correct. The full waveform solutions from L. Zhu turned out to give more consistent results. This is to be expected as the cross-correlation used in this research also looks at the complete waveform.

From this research it can be concluded that in order to solve the jigsaw puzzle of gradually changing record sections, more data has to be available and an algorithm has to be developed that combines the information coming from cross-correlations, noise level, focal mechanisms and locations, and their error bars, in getting a better, more coherent set of earthquake parameters.

Index

1 Introduction

When an earthquake occurs and is big enough to trigger several stations, the first parameters to be calculated with the retrieved data are the location, origin time, magnitude and focal mechanism of the event. Location and origin time are calculated by a travel-time inversion. First arrival-times of different phases are picked, and a velocity structure between the hypocenter and the different stations is assumed. Shearer (1997) gives an overview of the current methods. The magnitude of an earthquake is determined using the amplitudes of different parts of the seismic signal; the focal mechanism is often determined by using first-motion polarities from P-waves or by performing moment tensor inversion of the waveforms. None of these methods use the similarity of waveforms as a tool to determine whether or not events should be alike in the calculated parameters.

A future project professor T. Heaton has in mind is to make an animation of propagating waves in the subsurface. This animation will not be based on a velocity structure and a synthetic event, but it will be based on waveforms recorded by a local (broadband) station. In order to calculate the exact path a wave traveled before reaching the station, the exact location of the event has to be known. The animation will require many local events that have exact (relative) locations to reconstruct a full propagating wavefront. Current location methods do not use similarities in waveforms in their calculations. The main purpose of this research is to explore the possibilities of the use of maximum cross-correlation values in determining the relative locations of local events.

For the last 15 years, the cross-correlation technique has been used frequently to improve the precision of P- and S-wave picks (Time-domain cross-correlation: Rowe et al. 1998, Augliera et al. 1992, Dodge et al. 1995, Dodge et al. 1996, Shearer 1997, 1998, Mezcua 1994, Frémont 1987, Frequency-domain cross-correlation: Ito 1985). These picks are then used for a more precise travel-time inversion for earthquake locations. The cross-correlation is calculated for a short (~1 sec) time-window around the arrival-time. The time shift to get the maximum of cross-correlation is used to line up the arrivals, and this way more coherent picks can be made for the events that do not have a very clear arrival onset. This technique results in smaller errors in the relative locations of events.

Vsevolozhsky et al. (1993) used cross-correlation of envelopes of waveforms between data and a catalog of Green's functions to line up phases like Pn, Pg, Sn and Rayleigh. The distance of the events is then determined from the minimized misfit between data and synthetics.

In this research, maximum cross-correlation values are used directly to give an indication of the relative locations of the events considered. Using the similarities in waveforms to locate the events relative to each other, the theorem of reciprocity and the knowledge of finite strain in the earth's crust are considered. From this it can be derived that two identical events that occur very close to one another in space will give almost the same surface motion at one recording station (chapter 2). As the distance between the two events increases, the traces will change gradually and give a lower maximum cross-correlation value, until the events are so far apart that the waveforms are no longer coherent. Along with the distance between two events, the focal mechanisms also play a role in the degree of similarity of waveforms. For two events occuring at the same location, the waveforms will be identical if the focal mechanisms (moment tensors), including magnitude, are identical. However, the degree of waveform similarity will decrease as the events take on different mechanisms.

The combination of change in location, focal mechanism and magnitude brings the total set of degrees of freedom to seven: three dimensions spatially (x,y,z); strike, dip and rake of the focal mechanism; and the magnitude. The goal of this research is to use waveform similarities as the basis for locating events in this seven-dimensional space. First the focal mechanisms will be held constant, reducing the degrees of freedom with three. Then events will be chosen in sets with a particular distance from the station varying in depth, or at a certain depth with varying distances from the station, reducing the degrees of freedom even more. The magnitude is not considered further in this research.

Starting with the locations and focal mechanisms from a catalog, the events are regrouped in such a way that coherent, gradually changing record sections can be made. For this the maximum cross-correlation between the different events is used. By comparing given focal mechanisms and locations from the catalog with the calculated maximum cross-correlation coefficients, it is attempted to comment on how well these parameters agree in the

indication of similarities between the events and also about where inconsistencies in the catalog can be detected.

In order to be able to solve the jigsaw puzzle of gradually changing record sections, many events that are evenly spread over the area of interest are needed. In this research the data considered are from two areas with recordings at TriNet stations. The first is the Long Valley area, a seismically active volcanic region in the eastern part of central California with events recorded at station Mammoth Lakes (MLAC). The second dataset is the Northridge aftershock sequence recorded at station Pasadena (PAS). Long Valley has many events in a wide range of magnitudes. The Northridge sequence has fewer events, but they are, on average, higher in magnitude, and the spacing of the events is larger. Long Valley has events that are much closer in space, but usually lower in magnitude. It is clear that the order of magnitude of the two data sets differs, but this will give an indication of the type of region for which the method might work best.

2 Theory

Two events identical in location, magnitude and focal mechanism will produce identical seismograms at a given station, assuming that nothing changes in the medium along the path the waves travel during the time between the two events (Shearer, 1998). The theory of reciprocity states that if the source and receiver are interchanged, the recording at the station will not change. Extending this idea, one can imagine an event at the surface and two recording stations close to one another in the earth's subsurface (see figure 2.1). Because of finite strains in the earth, the motions at both stations cannot differ too much, and therefore the recordings will be almost identical. Going back to the normal situation, one can state that if two identical events happen close to one another in the subsurface, the recordings at the surface will have to be almost identical as well. The seismograms are expected to change gradually as the distance between the two events increases until the distance is such that the waveforms are no longer coherent. The same reasoning holds for the focal mechanism. If two events at the same location have identical focal mechanisms and magnitude, the recorded seismograms will be identical. With a (gradual) change of focal mechanism, the degree of waveform similarity will (gradually) decrease. It is unclear how similar the events must be in location and focal mechanism to get this gradual change (Shearer, 1997).

The magnitude of an event will alter the frequency content mostly (and will make the wave sensitive to larger-scale heterogeneities when the magnitude increases). In this research the magnitude is not considered.
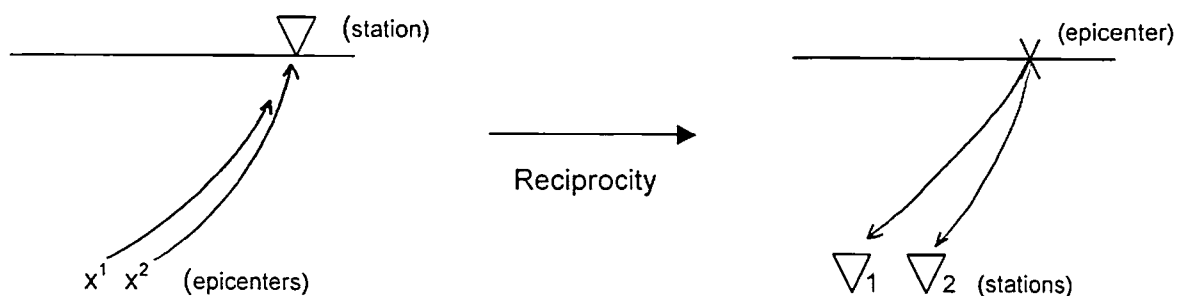

Figure 2.1: *Reciprocity*

The cross-correlation function is a measure of the similarity between waveforms. If two events are similar in focal mechanism, the maximum cross-correlation coefficient of their waveforms can tell you something about the distance between these two events. Spieth et al. (1981) stated that if waveforms of different events (same station) are identical when filtered below a threshold frequency, it can be said that the sources are within one-quarter wavelength of that threshold frequency. Maybe one should better say that the distance will be larger than a quarter wavelength when the correlation is below a certain threshold value instead of the other way around. It is possible to make a similar statement about focal mechanisms. If the waveforms (and locations) of two events are highly similar, then the mechanism of the two events must be similar. So, when the focal mechanism of one of the two events is known, one will know the mechanism of the other event if the maximum cross-correlation is above a certain threshold value. This would be very helpful for smaller events, for which it is difficult to determine the focal mechanisms accurately.

In this research, cross-correlation is used to determine the similarities in waveforms between local events recorded at the same station. Based on the maximum cross-correlation coefficients and the locations and focal mechanisms from the catalogs, the events are regrouped in such a way that coherent, gradually changing record sections appear. In general, one expects the cross-correlation coefficients to decrease as events become more distant from each other in any of the seven degrees of freedom of the problem. By determining how well higher cross-correlation coefficients agree with similarities in focal mechanism and location, one can comment on the consistency between these parameters and possible errors in the catalogs.

For example, in the one-dimensional case one expects identical events, which are on a straight line with increasing hypocentral distance from the station, to give cross-correlation coefficients with the closest event that become smaller as they are located farther away from the station (see figure 2.2). If the normalized maximum cross-correlation coefficients are put in a matrix with increasing (or decreasing) distance of the events, the cross-correlation coefficients become lower as one moves away from the main diagonal. The main diagonal contains the normalized auto-correlation (=1) of each event, and on the sub-diagonals are the normalized maximum cross-correlations ($0 < \phi < 1$), which form decreasing series away from the main diagonal in both the left and right horizontal directions. The lining-up of events with a correlation matrix of this form (decreasing series away from the main diagonal) will be called the 'ideal' case.

To place a set of events in the most 'ideal' order, an algorithm has to be developed that, after input of a matrix made with the events in random order, finds the order of events that gives a matrix that comes closest to the 'ideal' case: a matrix with most of its energy around the main diagonal.

$\triangledown$ station

		1	2	3	4
x_1	1	1.0	0.89	0.73	0.55
x_2	2	0.89	1.0	0.88	0.69
x_3	3	0.73	0.88	1.0	0.87
x_4	4	0.55	0.69	0.87	1.0

x = hypocenter locations

Figure 2.2: *Geometry and an example cross-correlation matrix for the 'ideal' 1-D case with a horizontal or a vertical line*

Of course in real life the spacing will be 3-dimensional, and the lining-up will not be so evident. Also, the real earth has 3-dimensional heterogeneity which alters the path (and consequently the waveform) more than for the case of a uniform halfspace velocity model. In order for this method to work fully for the 3-dimensional space, one will need many events with very similar focal mechanisms, evenly spaced over the area of interest. This is hard to find in nature. A fault

plane with a lot of fore- and aftershocks, or a seismically active volcanic area are good places to start.

3 Methodology of waveform analysis

The steps explained in this section are illustrated by a waveform example from two events from the Long Valley area (event-identification numbers: 9039416 & 9038488).

3.1 Data selection

First the data has to be retrieved from the data-center and selected for events that have *decent* traces. This means that the waveform includes the direct P-wave arrival (and the origin time), and the trace is long enough to be used in a cross-correlation comparison (at least 20 seconds after the P-wave arrival). Also, three components are needed for each event. In the chapters about the different data-sets (chapters 4-6), more will be explained about the data used (see figure 3.1 for the example raw data).
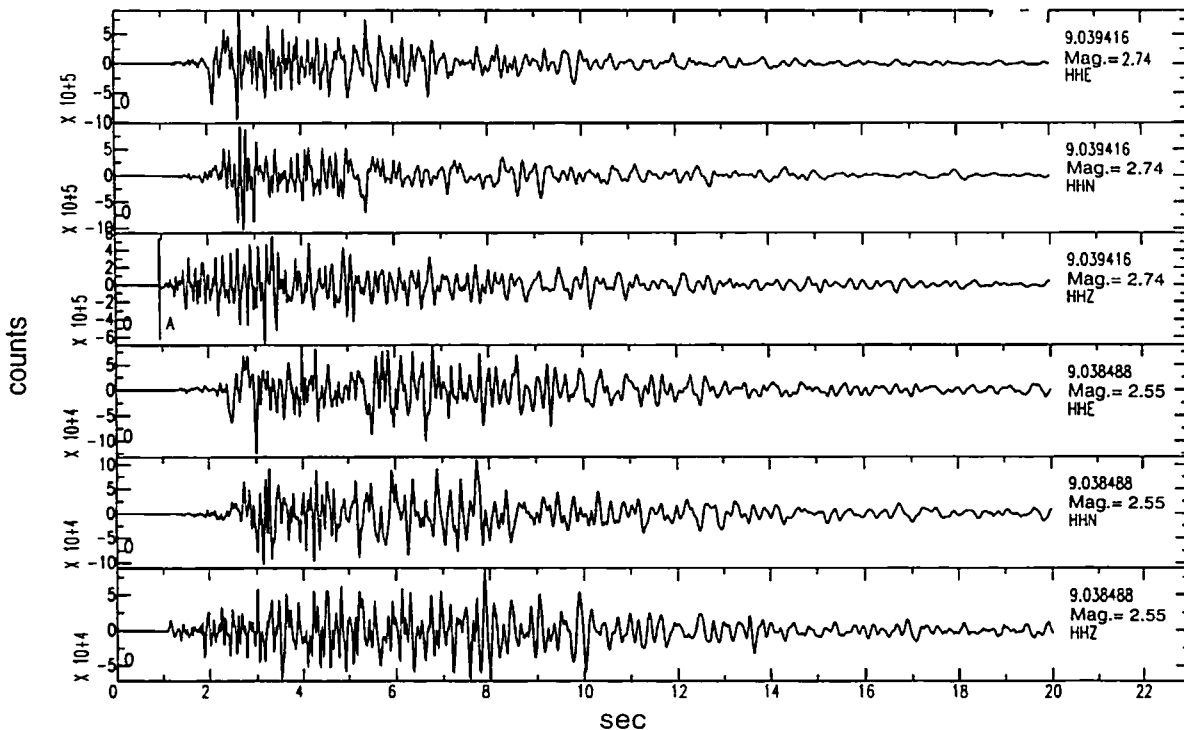
Figure 3.1: *Raw data from events 9039416 and 9038488. East (HHE), north (HHN) and vertical (HHZ) components.*

3.2   Rotation

The horizontal (N & E) waveform components are rotated in clockwise direction such that one of the components points in the direction away from the source (radial), along the great circle from source to station (see figure 3.2).
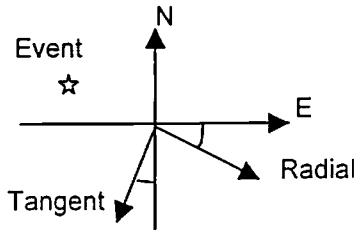

Figure 3.2 : *example of geometry of rotation*

3.3   Integration

Integration of the velocity trace to get displacement, if desired. Displacement has naturally a lower frequency content than velocity.

3.4   Filtering

A bandpass butterworth filter in the lower frequency range can be applied to get an even lower frequency range if desired. The filter is primarily used to eliminate higher frequencies. A very low frequency corner is used to eliminate long-period drift.
Events will have a more similar signature in the lower frequency range because high frequency waveforms are more affected by short-wavelength material heterogeneity.

Both integration and filtering lower the frequency content of the signal. This will generally result in a higher similarity in waveforms. What is wanted for this research is not the highest similarity at all costs, but high similarities for the widest frequency range possible. This will allow for more detail and therefore a better relative relocation in the end.

3.5 Cut the traces to the desired length for correlation

Not the whole trace is used for correlation. In practice, the length of the shortest recorded trace determines the length of the time window used for the whole (sub)set looked at. Within a subset the length of the traces used will be the same, but it can change from subset to subset.

If the events are further away from the station, the part of the waveform before the direct P-wave onset is not used (8 seconds or more before the P-wave arrival). Figure 3.3 shows the example traces after rotation, filtering and cutting.
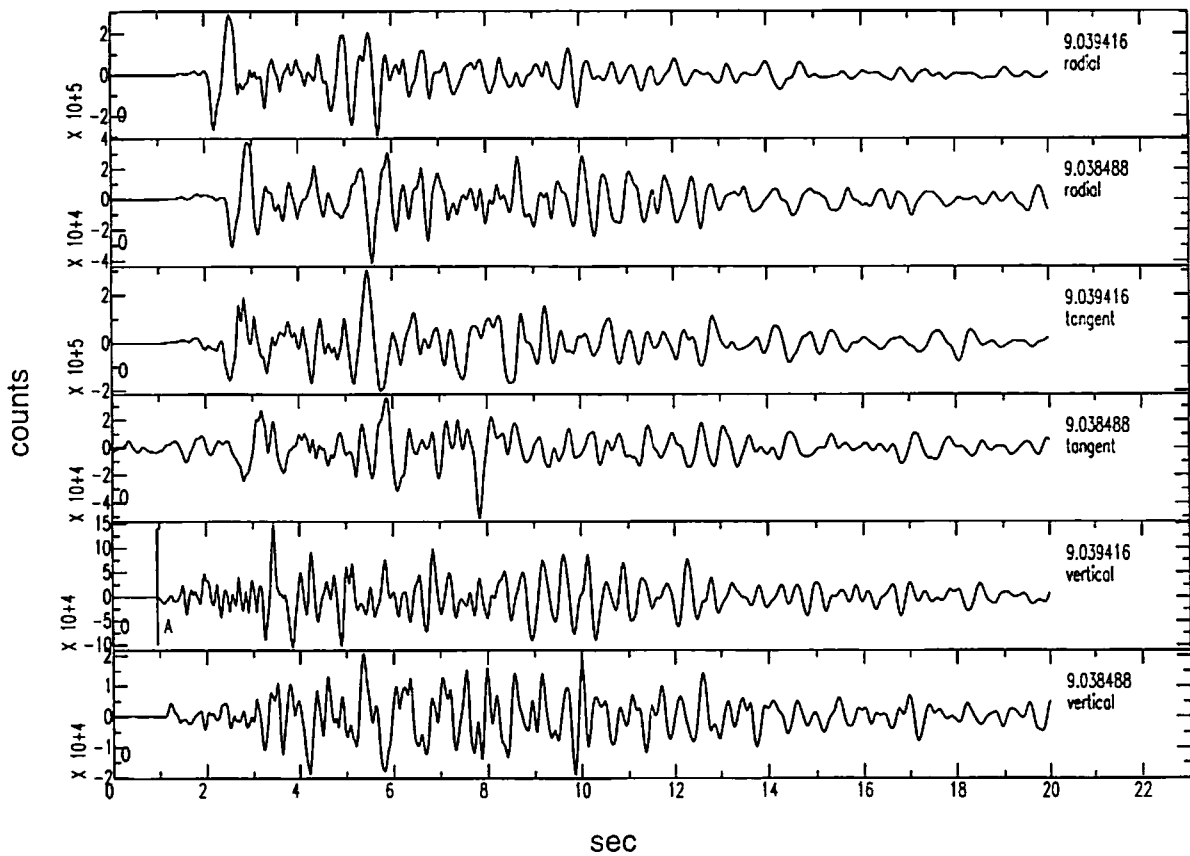
Figure 3.3: *Traces after rotation, butterworth filter (0.01 - 2 Hz) and cutting. No integration is applied.*

3.6 Cross-correlation

First the maximum cross-correlation coefficient (ϕ_{ij}) for the three separate components of all recorded events is determined [using SAC (1998), see references, Ch. 9] (see equation 3.1)

$$\phi_{ij} = \max_{-t_1 < \tau < t_1} |\phi_{ij}(\tau)|$$

$$where : \phi_{ij}(\tau) = \int_t u_i(t)u_j(t + \tau)dt$$

Equation 3.1: *Maximum cross-correlation value picked from the cross-correlation function. The variables u_i and u_j are the seismograms in time.*

The maximum absolute cross-correlation coefficient is picked within a certain time window ($-t_1 < \tau < t_1$) around the zero axis of the cross-correlation trace. Depending on the difference in distance and depth between the traces within the subset, this window will be 4 to 10 seconds wide. For each individual subset the same time-window will be used, but it can change from subset to subset. Figure 3.4 shows the auto-correlation and cross-correlation traces for the example data. The maximum will be picked in the 4 second window with $t_1=2$. As they become further away from one another in space, the shift between the traces to get maximum correlation will become larger than 2 seconds, and the time window in which the maximum will be picked will have to be larger.
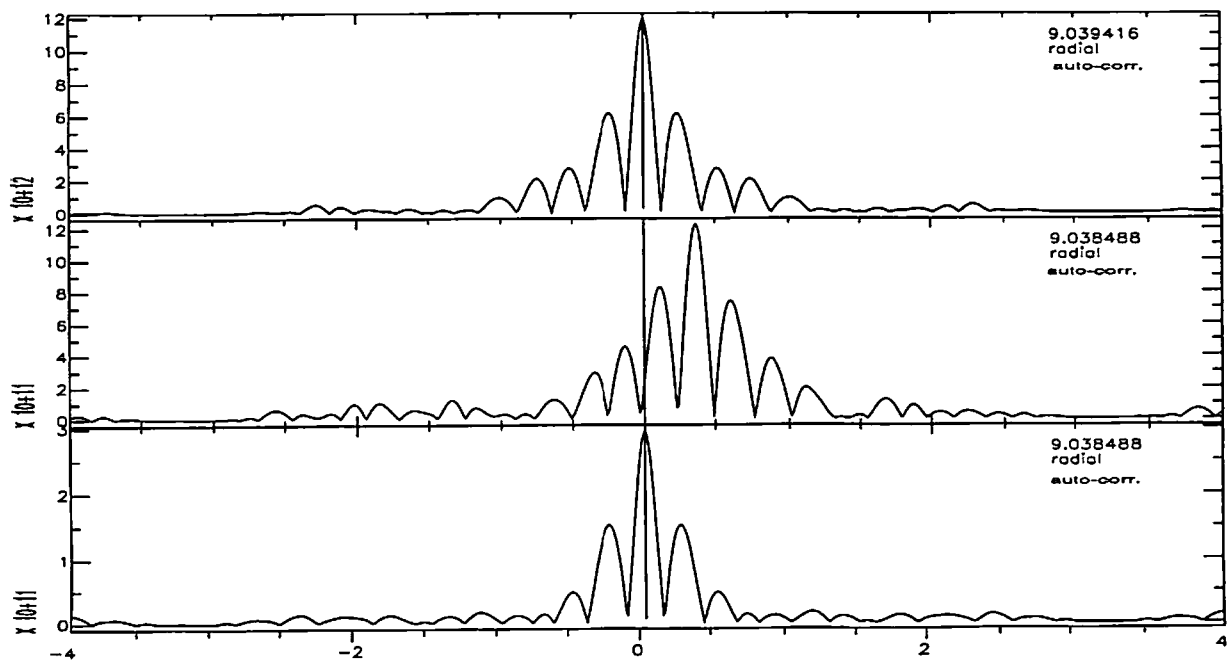

Figure 3.4: *Auto- and cross-correlation traces from events 9039416 and 9038488*

The maximum cross-correlation is normalized such that the auto-correlation of the first (i^{th}) trace used is 1 (see equation 3.2).

$$\phi_{ij} = \phi_{ij}(\tau)_{norm} = \frac{\phi_{ij}}{\sqrt{\phi_{ii}(0) \cdot \phi_{jj}(0)}}$$

Equation 3.2: *Normalized cross-correlation coefficient*

All the normalized cross-correlation coefficients are placed in a squared matrix. The result is three symmetric matrices (one for each component) with the auto-correlation (=1) on the main diagonal and the normalized maximum cross-correlations on the subdiagonals.

3.7 Average for the three components

When from two events all three traces recorded at a particular station are highly similar, then we expect the focal mechanisms to be highly similar as well. Therefore, the average of the cross-correlation matrices from the three components can be used (see equation 3.3). If all three components from two different events are highly similar, the focal mechanisms will be the same, and the distance between the two events will be small.

$$\phi_{ij}^{av} = \frac{\phi_{ij}^{vert} + \phi_{ij}^{tan} + \phi_{ij}^{rad}}{3}$$

Equation 3.3: *Calculation of average matrix*

3.8   Determine the most 'ideal' order

The major part of this research was the development of an algorithm which determines the order of events that comes closest to the 'ideal' order as discussed in the introduction (chapter 1). This algorithm therefore reads in the starting average cross-correlation matrix and calculates a closeness number that indicates how close the matrix is to the 'ideal' situation. Next it changes the order of the events, reconstructs the matrix and again calculates the closeness number. This way, for each order of events (systematic search), the matrix is reconstructed (if one changes the order of columns, one also has to change the order of rows to keep track of the coefficients), and a closeness number calculated.

In chapter 1 we saw that the 'ideal' order of events is the one in which the series of maximum cross-correlation coefficients decreases from the main diagonal to both sides on the (horizontal) rows. To get the closeness number, the differences between neighboring coefficients are calculated, starting from the main diagonal ($i = j$) and moving to the left and right on the horizontal rows. If the difference is negative (i.e., the coefficient further away from the main-diagonal is bigger than the closer one), the difference is squared and added to the closeness number. If the difference is positive, nothing is done with the value. For each order of events the closeness number is calculated, and the order with the smallest closeness number is called the 'minimized' order. This order will be used in further analyses. There is no other use for the closeness number. It is not normalized and therefore one cannot compare between different subsets.

By using the least-squares, the larger differences will contribute more to the closeness number. This way the more 'smooth' orders will have smaller closeness numbers and a better chance of being the 'minimized' order. Equation 3.4 gives the mathematical expression for the calculation of the closeness number (C_1), and in figure 3.5 the flow diagram of the calculation of the closeness number is given.

$$R = \sum_{P_R}(\phi_{ij} - \phi_{ij+1})^2 .....where......P_R = \{\forall i,j : j > i, \phi_{ij} < \phi_{ij+1}\}$$

$$L = \sum_{P_L}(\phi_{ij+1} - \phi_{ij})^2 .....where.......P_L = \{\forall i,j : j < i, \phi_{ij+1} < \phi_{ij}\}$$

$$C_1 = R + L$$

Equation 3.4: *Mathematical expression for the calculation of the closeness number C_1*
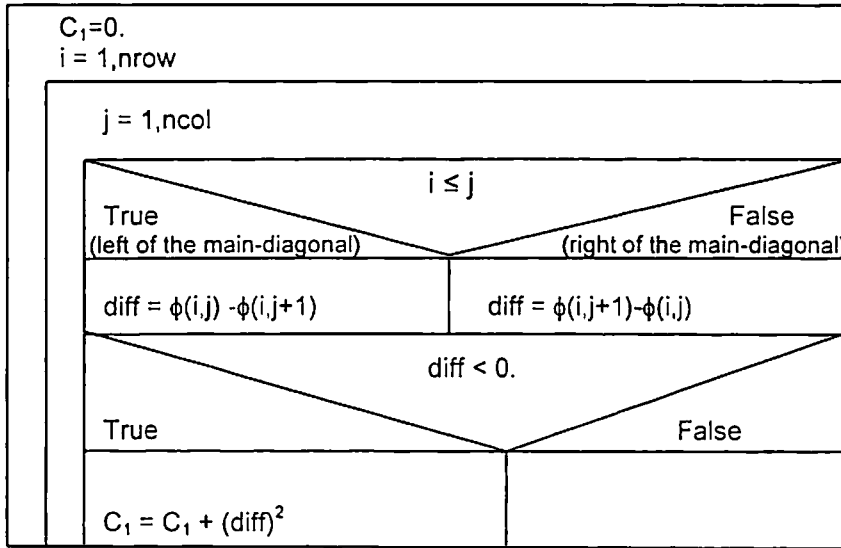
```
C_1=0.
i = 1,nrow

    j = 1,ncol

                              i ≤ j
        True                                    False
        (left of the main-diagonal)             (right of the main-diagonal)

        diff = φ(i,j) -φ(i,j+1)          diff = φ(i,j+1)-φ(i,j)

                              diff < 0.

        True                                    False

        C_1 = C_1 + (diff)^2
```

Figure 3.5: *Flow diagram of the calculation of the closeness number. Only the coefficients that are out of place, i.e. those that violate a decreasing series in both directions away form the main diagonal, will be added to the closeness number in a least-squares sense.*

The developed algorithm searches through all the possible orders of events in a systematic way and calculates the closeness number for each order.

The algorithm can find the most 'ideal' order for:

1. the average matrix

2. the average matrix plus the S-P time series (see paragraph 3.9)

For each separate set of events, the order with the smallest closeness number is called the 'minimized' order.

3.8.1 Limitations of the algorithm

In the calculation of the 'minimized' order a brute force method is used. The algorithm is limited in the amount of events it can take in one run. The limiting factor is that all possible orders of events are considered. For eleven events this is a total of 11!=39.916.800 possibilities. In order to be able to use more events at a time, the algorithm has to be optimized. Partially this is done by stopping the calculation of the closeness number for a particular matrix (order) if the previous minimum (for the same subset) is exceeded.

3.9   S-P time series

As mentioned before, the S-P times can be used in the minimization as well. For local ranges the distance D can be estimated as follows [where α is the P-wave velocity, t_s is the arrival time of the S-wave and t_p is the arrival time of the P-wave(Lay, 1995)]:

$$D = \frac{\alpha \cdot (t_s - t_p)}{\sqrt{3} - 1}$$

For crustal events this reduces to be D=8.0·(t_s - t_p), assuming a homogeneous halfspace and $\alpha = 3\sqrt{3}$.

One has to keep in mind that there is a trade-off in the S-P times between shallower events in slower media and deeper events in faster media. Therefore the S-P time series are not used as the only parameters in determining the minimized order of events. But by including the S-P times, an extra constraint is placed on the (relative) position of the events, rather than using similarities in waveforms only.

For each order of columns for which a closeness number is calculated, the least-squares contribution of the S-P times series (T) is added to the closeness number as it has the same order of magnitude as R and L. Again only the terms that are violating the decreasing series will contribute to the closeness number. See equation 3.5 and figure 3.6 for the mathematical expression and the flow diagram.

$$= \sum_P (t_s - t_p)_j^{2} .....where.......P = \{\forall j : (t_s - t_p) < (t_s - t_p)\}$$

$$C = R + L + T$$

Equation 3.5: *Mathematical expression for the calculation of the contribution of the S-P time series to the closeness number C.*

```
┌────────────────────────────────────────────────────────────┐
│  T=0.                                                        │
│  i = 1,nrow                                                  │
│  ┌──────────────────────────────────────────────────────┐   │
│  │  sp(i) = t_s(i) - t_p(i)                              │   │
│  │  diff = sp(i) - sp(i+1)                               │   │
│  │  ┌──────────────────────────────────────────────────┐ │   │
│  │  │                diff < 0.                          │ │   │
│  │  │  True                              False          │ │   │
│  │  ├─────────────────────────┬────────────────────────┤ │   │
│  │  │  T = T +(diff)²          │                        │ │   │
│  └──┴─────────────────────────┴────────────────────────┘ │   │
└────────────────────────────────────────────────────────────┘
```
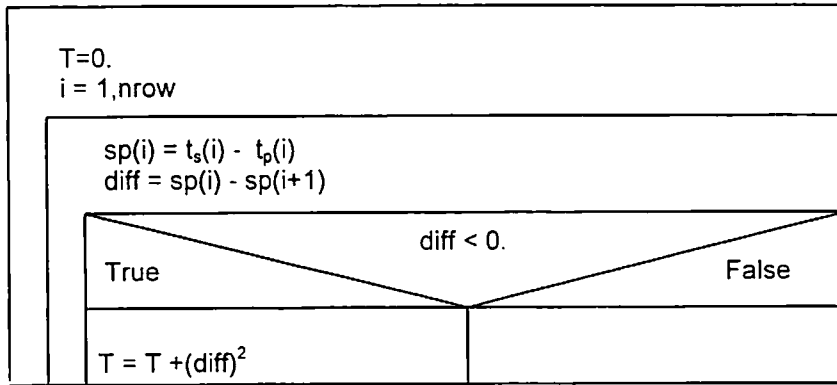
Figure 3.6: : *Flow diagram of the calculation of S-P time series part of the closeness number*

So again only the terms that are violating the decreasing series contribute to the closeness number in a least-squares sense. The total closeness number (C) is calculated by adding R and L (=C_1) (fig 3.5) and T (fig 3.6).

4 Synthetic data-set

4.1 Introduction

Before working with real data, a brief experiment is first performed using a synthetic data-set. The synthetics are Green's functions used for an inversion to determine the slip history of the Northridge, California, USA, earthquake that occurred on January 17, 1994 (M_w=6.7) (Wald et al (1996)). The Green's functions used all have the same focal mechanism and therefore this experiment will only look at the changes in waveform due to different locations and not due to different focal mechanisms. Three different subsets of Green's functions are investigated: a horizontal line with the subfaults along the strike of the fault, a vertical line with the subfaults along the dip of the fault and a set of subfaults randomly picked. The goal of this experiment is to see if the developed algorithm lines up the Green's functions belonging to the different subfaults in the same order as we expect from the geometry of station and subfaults. For the vertical line of subfaults the minimized order is also calculated after adding noise to the traces to analyse the influence of noise on the developed algorithm.

4.2   Data selection and pre-processing

Wald et al. (1996) determined the slip history of the Northridge earthquake by representing the slip on the fault with numerous subfaults. In this procedure the slip on each subfault is calculated by the summation of many point sources over the point source area. In this manner synthetic Green's functions with identical fault rupture models are generated. In their inversion scheme Wald et al. use a faultplane with a strike of 122° and a dip of 40°. The top of the fault is 5 km below the surface. The fault has a length of 18 km along strike and a width of 24 km down dip. The total fault surface is divided into 196 subfaults (14x14). The velocity model used to calculate the Green's functions is a horizontally layered structure (see table 4.1).

V_p (km/s)	V_s (km/s)	density (g/cm)	thickness (km)	depth (km)
1.9	1.0	2.1	0.5	0.0
4.0	2.0	2.4	1.0	0.5
5.5	3.2	2.7	2.5	1.5
6.3	3.6	2.8	23.0	4.0
6.8	3.9	2.9	13.0	27.0
7.8	4.5	3.3		40.0

Table 4.1: *Northridge regional velocity structure for rock sites (Wald et al. 1996)*

For this research the synthetic Green's functions as calculated for station VSQ (~ 30 km) are used. Figure 4.1[a] illustrates the geometry of fault, station, and subfaults. Columns of subfaults are represented by vertical lines along the dip. The numbering of columns and rows starts at northwest corner. For the horizontal line of subfaults the first row is used, and the subfaults are numbered from 1 to 11 running from NW to SE (the algorithm work with eleven events at a time). For the vertical line of subfaults, the second column is used and the subfaults are numbered from 1 to 11 running down dip. For the set of randomly picked subfaults the first three columns are used and the numbering is as given in figure 4.1[b]. Only the subfault numbers given in bold are used in the minimization.

c1	c2	c3
1	15	29
2	16	30
3	17	31
4	18	**32**
5	19	33
6	**20**	**34**
7	**21**	35
8	22	36
9	23	37
10	24	**38**
11	**25**	39
12	26	40
13	27	**41**
14	**28**	42

b)
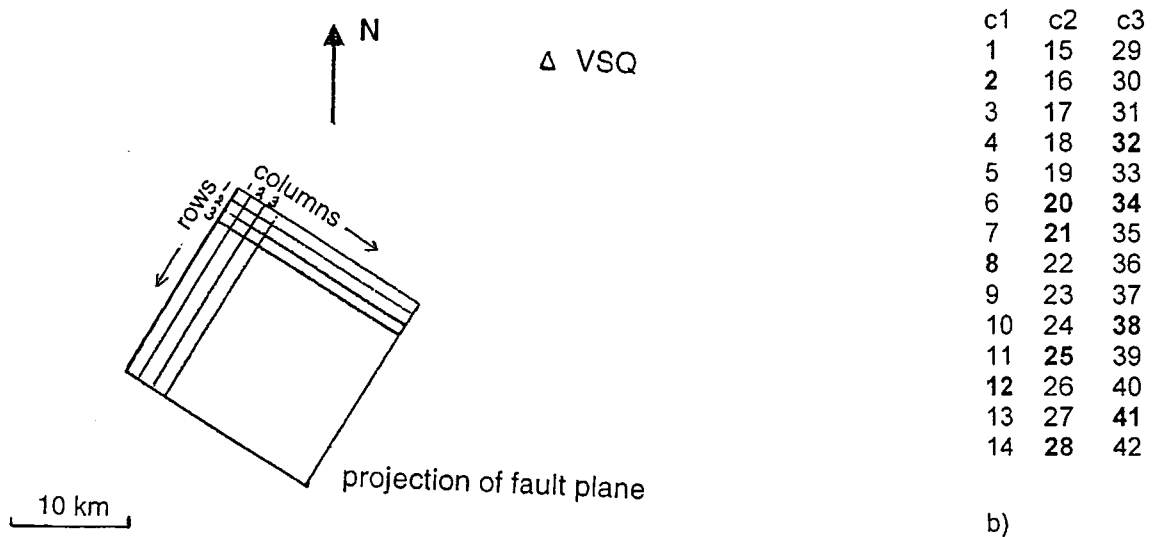
a)

Figure 4.1: a) *Geometry of station, fault plane, subfaults and event lines.* b) *Numbering of the subfaults in the first three columns of the fault, as used in the random chosen subset. Used subfaults are given in bold.*

Synthetics: vertical line along the dip, three components,
vertical, north and east, minimized order
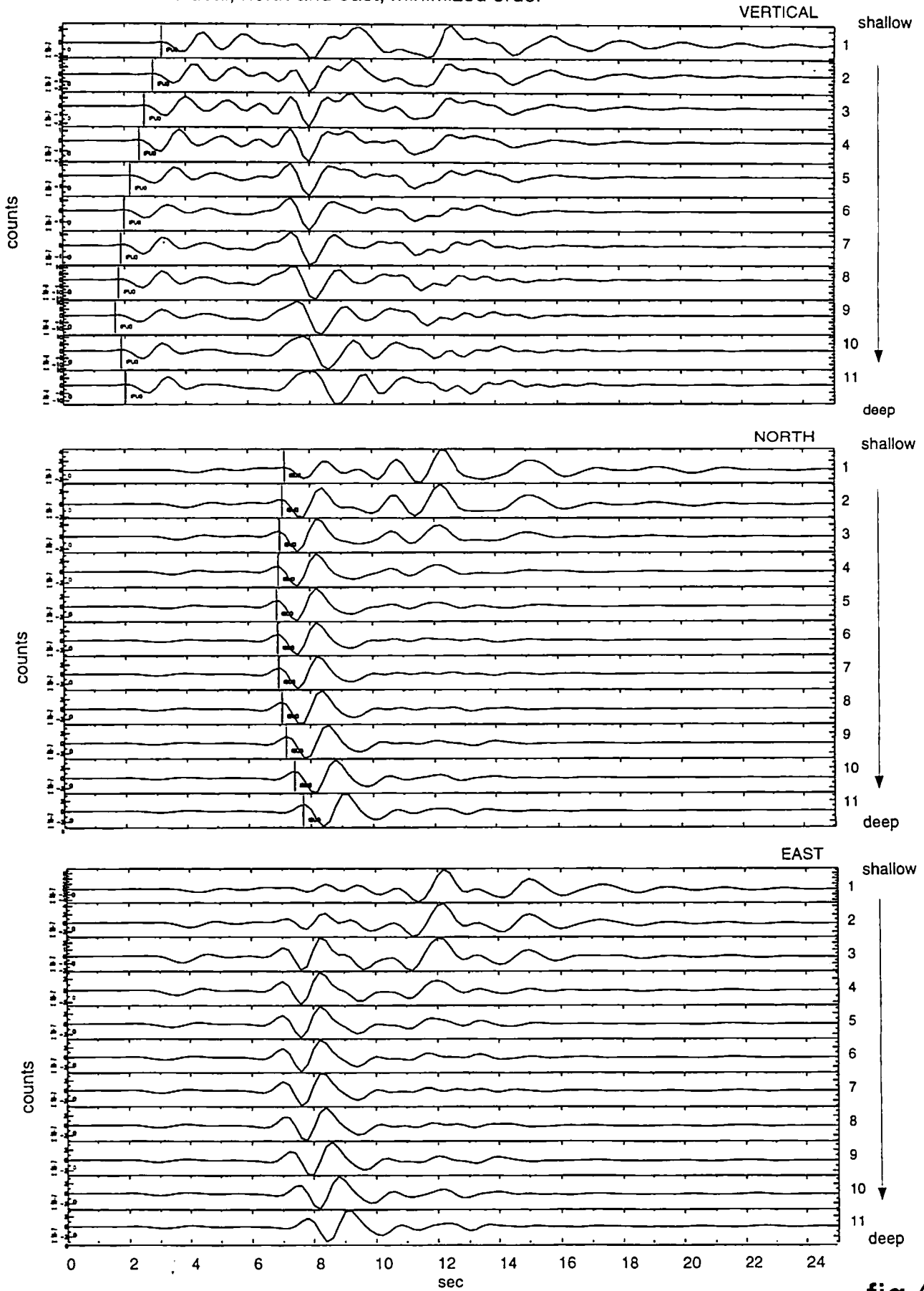
fig 4.2

No rotation is applied to the waveforms, although all three components are present. This should not affect the calculations as the rotation is merely used to focus the energy onto one component and improve the signal to noise ratio. The traces are bandpass filtered with a butterworth filter (two-pass, 4th order) in the frequency range of 0.1 to 1 Hz. Due to the lower frequency content, the velocity traces are not integrated before the calculation of the cross-correlation matrix. For the cross-correlations the first 25 seconds of each trace are used.

4.2.1 Noise

In order to see how stable the method is under noise, 'real' noise has been added to the synthetics from the vertical line. The three component noise-record comes from a random station (VCS) in a quiet period (no local earthquakes). After filtering of the noise in the same frequency band as the synthetics, the traces are divided into parts of 25 seconds. The noise is added to the synthetics after adapting the amplitude of the noise to get a signal to noise ratio of 3 for the maximum amplitudes.

4.2.2 S-P time series

For the vertical line with noise and the set of random picked subfaults, S-P times are also used in the minimization to analyse how large the influence of the S-P time series on the minimized order is. The S- and P-times are hand-picked. By comparing the traces within the same set, the picks have been made more coherent from trace to trace. It would even have been better if the arrivals were first lined up by use of the maximum cross-correlation time lag for a time window around the direct arrival only (see Chapter 2, Theory).

4.3 Results

Minimized orders are calculated for the following cases:

1. vertical line (11 traces, starting at the top of the second column of subfaults);
2. horizontal line at 5km depth (11 traces, starting at the second column at the NW end of the fault);
3. a) vertical line as in 1 with added noise;

   b) vertical line with added noise plus the S-P time series;
4. a) random picks from the first three columns of subfaults at the NW-side (11 traces, vertical separation max. 2.56km);

   b) random picks plus S-P times.

The minimization gives the 'perfect' answer for the vertical line (case 1). One expects the traces to change gradually as you go further away from (or come closer to) the station (the fault dips away from the station). And indeed the minimized order, which is the order as calculated by the algorithm described in paragraph 3.8, runs along the dip. The closeness number for this case is 2×10^{-3} . Figure 4.2 contains the traces for this subset in the minimized order; table 4.2 shows the average cross-correlation matrix after minimization.

entry   number of event

entry												number of event
1	1.00	0.84	0.73	0.63	0.60	0.58	0.51	0.48	0.41	0.32	0.23	11
2	0.84	1.00	0.92	0.81	0.71	0.67	0.61	0.55	0.50	0.38	0.25	10
3	0.73	0.92	1.00	0.90	0.83	0.75	0.69	0.63	0.53	0.39	0.28	9
4	0.63	0.81	0.90	1.00	0.93	0.82	0.75	0.67	0.59	0.41	0.24	8
5	0.60	0.71	0.83	0.93	1.00	0.96	0.86	0.74	0.62	0.40	0.24	7
6	0.57	0.67	0.75	0.82	0.96	1.00	0.96	0.85	0.68	0.45	0.29	6
7	0.51	0.61	0.69	0.75	0.86	0.96	1.00	0.95	0.80	0.54	0.38	5
8	0.48	0.55	0.63	0.67	0.74	0.85	0.95	1.00	0.92	0.68	0.48	4
9	0.41	0.50	0.53	0.59	0.62	0.68	0.80	0.92	1.00	0.87	0.69	3
10	0.32	0.38	0.39	0.41	0.40	0.45	0.54	0.68	0.87	1.00	0.88	2
11	0.23	0.25	0.28	0.24	0.24	0.29	0.38	0.48	0.69	0.88	1.00	1

closeness number = 1.78×10^{-3}

Table 4.2: *Final average matrix for minimized order , vertical line (shallower event is on bottom of matrix!)*

Also in case 2, the horizontal line along the strike (at 5 km depth), a 'perfect' answer was obtained after minimization which runs along the strike from northwest to southeast (or vice versa). The closeness number for this case is 3×10^{-4}. Matrix and waveforms are not shown.

When the noise is added to the events on the vertical line (case 3^a & 3^b, S/N=3), the minimization does not give the 'perfect' answer anymore, but is still rather close. The minimized order is calculated with and without the S-P time series.

The minimized orders are:

1. 1,2,3,4,5,8,6,7,10,9,11 (without P-S time series)

2. 1,2,3,4,5,7,6,8,9,10,11 (with P-S time series)

With the addition of the noise, the arrivals times of the P-waves are less clear. For this analysis the P- and S-picks from the traces without the noise are used. With non-synthetic traces the higher frequency components can be used to pick the arrivals better. Figure 4.3 gives the traces after the noise is added in the minimized order, and table 4.3 contains the average cross-correlation matrix after minimization.

entry												t(s)-t(p)	number of event
1	1.00	0.52	0.67	0.55	0.49	0.49	0.43	0.45	0.32	0.24	0.22	5.7	11
2	0.52	1.00	0.56	0.55	0.56	0.63	0.52	0.45	0.36	0.33	0.23	5.6	10
3	0.67	0.56	1.00	0.65	0.58	0.58	0.51	0.56	0.47	0.30	0.27	5.5	9
4	0.54	0.55	0.65	1.00	0.76	0.62	0.56	0.56	0.56	0.47	0.19	5.3	8
5	0.49	0.56	0.58	0.76	1.00	0.64	0.67	0.52	0.62	0.37	0.22	4.9	6
6	0.49	0.63	0.58	0.62	0.64	1.00	0.68	0.52	0.47	0.32	0.23	5.1	7
7	0.43	0.52	0.51	0.56	0.67	0.68	1.00	0.75	0.56	0.47	0.31	4.7	5
8	0.45	0.45	0.56	0.56	0.52	0.52	0.75	1.00	0.59	0.50	0.39	4.5	4
9	0.32	0.36	0.47	0.56	0.62	0.47	0.56	0.59	1.00	0.75	0.56	4.4	3
10	0.24	0.33	0.30	0.47	0.37	0.32	0.47	0.50	0.75	1.00	0.78	4.2	2
11	0.22	0.23	0.27	0.19	0.22	0.23	0.31	0.39	0.56	0.78	1.00	4.0	1

closeness number = 0.12

Table 4.3: *Final average matrix for minimized order vertical line plus noise (shallower event is on bottom of matrix!)*

Synthetics: vertical line + noise, three components, vertical, north and east, minimized order
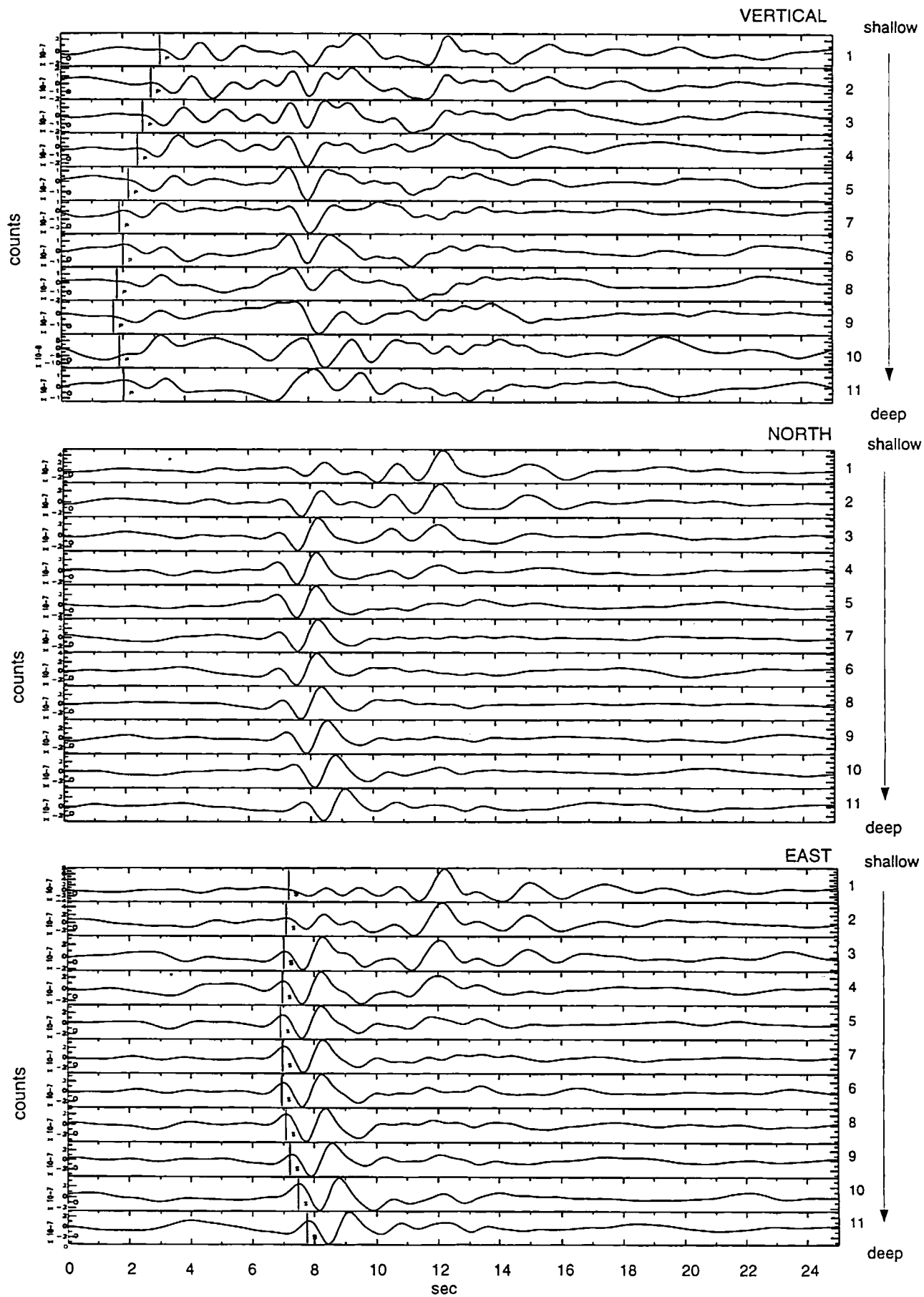
fig. 4.3

In the last cases (4[a] & 4[b]), where the subfaults are picked randomly from the first three columns of the total fault, the minimized order is the same for the calculations with and without the S-P time series (see figure 4.1[b] for the numbering of the events in the first three columns of the fault plane).

Minimized order:

1. 2,32,34,20,21,38,8,25,41,12,28

Because the difference in the S-P times is rather close, the order is not changed by adding the P-S time series minimization. The traces and correlation-matrices are shown in figure 4.4 and table 4.4. The order according to the hypocentral distance to the station is: 2,32,34,20,21,8,38,25,12,41,28. So numbers 8 and 38 as well as 12 and 41 have changed places with respect to this order. Apparently the algorithm does not sort fully according to the hypocentral distance.

entry												number of event
1	1.00	0.77	0.61	0.52	0.47	0.33	0.32	0.32	0.25	0.32	0.40	28
2	0.77	1.00	0.79	0.64	0.56	0.41	0.39	0.39	0.33	0.34	0.38	12
3	0.61	0.79	1.00	0.79	0.71	0.59	0.54	0.57	0.50	0.47	0.43	41
4	0.52	0.64	0.79	1.00	0.80	0.69	0.60	0.57	0.47	0.43	0.32	25
5	0.47	0.56	0.71	0.80	1.00	0.73	0.78	0.74	0.57	0.53	0.42	8
6	0.33	0.41	0.59	0.69	0.73	1.00	0.87	0.84	0.70	0.62	0.29	38
7	0.32	0.39	0.54	0.60	0.78	0.87	1.00	0.96	0.77	0.65	0.35	21
8	0.32	0.39	0.57	0.57	0.74	0.84	0.96	1.00	0.85	0.76	0.41	20
9	0.25	0.33	0.50	0.48	0.57	0.70	0.77	0.85	1.00	0.86	0.34	34
10	0.32	0.34	0.47	0.43	0.53	0.62	0.65	0.76	0.86	1.00	0.52	32
11	0.40	0.38	0.43	0.32	0.42	0.29	0.35	0.41	0.34	0.52	1.00	2

closeness number = 5.22×10^{-2}

Table 4.4: *Final average matrix for minimized order, random chosen subfaults (shallower event is on bottom of matrix!)*

Synthetics: random events from three vertical lines, three components, vertical, north and east, minimized order.
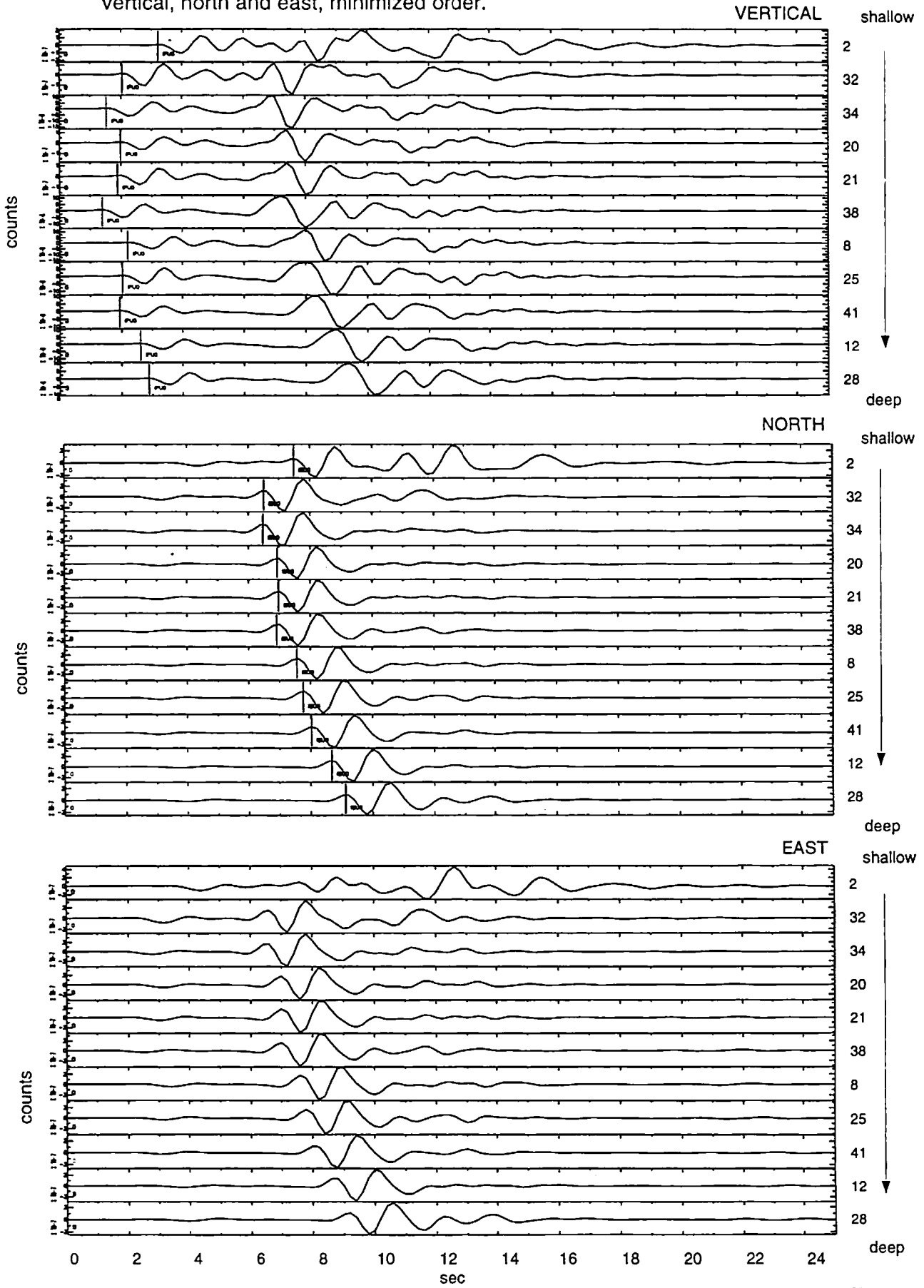
fig. 4.5

4.4 Synthetics conclusions

Vertical and horizontal line of subfaults, without S-P time series:

Although both closeness numbers are fairly small, the horizontal closeness number is a factor of ten smaller. This can be explained by the horizontally layered structure that is used in the calculations of the Green's functions. On the horizontal line all waves travel approximately the same distance in each layer. The waves from the vertical line however travel a different distance in the thick layer in which they all lie (for the velocity model see table 4.1). This way the horizontal case is closer to the one-dimensional 'ideal' case. The velocity structure is the same for all the traces and only the location changes. In general waveforms change more rapid with depth than with distance.

Vertical line of subfaults plus added noise:

For a signal to noise ratio of 3 the algorithm works well. It is not clear at what signal to noise ratio the algorithm stops working. With real data of course noise is always present. One should pick the S-P time series consistent (by lining up the traces on the arrivals) in all events in a subset.

Randomly picked subfaults from the first three columns:

The line-up as given by the algorithm is not fully in agreement with the hypocentral distance. The reason why is not clear. Is it that the horizontal distance between the first and third column is too large to have the algorithm work, or is there something else? One should have a closer look at randomly chosen subfaults from the first two columns only and see if that gives a hypocentral line-up yes or no.

Further synthetic experiments can be done on Green's functions generated for slowly changing focal mechanisms, coming from the same location and a combination of small changes in location and focal mechanisms. Then one can see what makes the biggest difference: the location (horizontal or vertical) or the focal mechanism changes.

Based on the paper by Spieth (1981), we can say something about the relative separation of the events by looking at the wavelength used for correlation and the cross-correlation coefficients.

wavelength = 1 second * 6km/s = 6 km

1/4 wavelength = 1.5 km

A separation of 1.5 km seems a bit too small for this geometry (the dimension of the elements is 1.29 km horizontal and 1.71km vertical), but is in the right order of magnitude if you use it the way she wrote it. Better would be to say that the events will lie further apart than the 1.5 - 2 km if the maximum cross-correlation coefficient is lower than ~0.8.

5 Long Valley

5.1 Introduction

Long Valley is a volcanic area that has experienced around 90,000 earthquakes (M>1.0) since 1980. The density of earthquakes in the entire region is large, therefore Long Valley is a good candidate for the method we are using. MLAC is a 3-component, broad-band station operated by the California Institute of Technology and located at the southern edge of the caldera. Figure 5.1 shows the Long Valley caldera and station MLAC. The dots are the epicenters of events recorded by MLAC, between July 1, 1993, and September 8, 1998, for which location parameters could be found in the Northern California Earthquake Data-Center catalog (see next paragraph for details). The fan indicated by two straight lines running in a W-E direction contains the subset of events that will be used in this analysis.
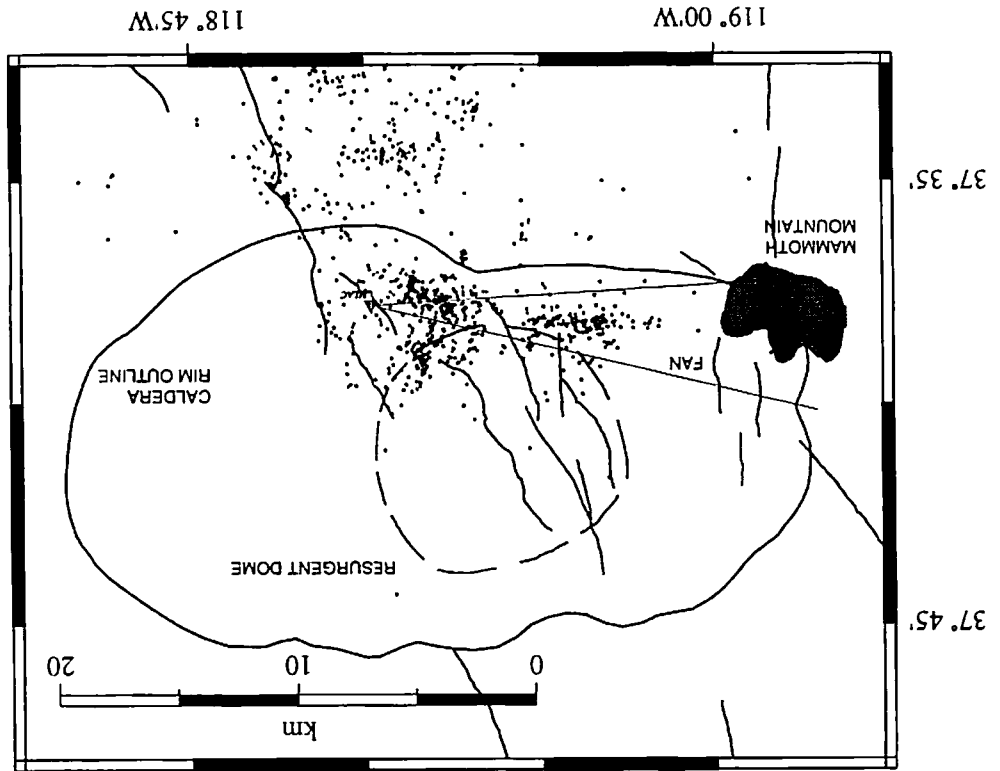
Figure 5.1: Long Valley caldera. Indicated are station MLAC and the events that were recorded by MLAC and could be correlated to an event in the catalog of the NCEDC (see text for more detail). Within the area of the indicated fan (angle of 10°) lie 256 events. Some of those events are used for this analysis.

5.2 Data Selection

The Southern California Earthquake Center (SCEC) database has been searched for the entire period station MLAC was active (July 1, 1993 - September 8, 1998). The station is still in use, but new data from after September 8, 1998 are not incorporated in the analysis. Waveforms are retrieved for events centered within a circle with a radius of 25 km around MLAC. As MLAC is situated at the northern edge of the Southern California Seismic Network (SCSN), the location parameters in this area are of a very poor quality in the SCEC database, because the events are only recorded by a few stations. To solve this problem the events from the SCEC database had to be correlated with the events recorded by the Northern California Network. As this network has a better station density in the Long Valley area, the location and origin time parameters are more accurate in the catalog of the Northern California Earthquake Data-Center (NCEDC). An inconvenience is that the event identification numbers are different in both data-sets. To be able to match two events the next filter is used:

> "Two events are considered to be the same if the origin time of the NCEDC-event lies within the recording time of the trace from the SCEC-event, and the origin times of the two events are not further than 2 seconds apart."

This results in a set of 1089 events that could be correlated, and 37 events recorded by MLAC that could not be correlated with an event recorded by NCEDC (epicenters are given in figure 5.1). The magnitude range is from 1.7 to 5.1. In order to make the method work, the events that are going to be compared should not come from azimuths that are too far apart. The events will then be too far apart and possibly in two different media and will not correlate well no matter what. For this analysis, the events within a fan with an angle of 10 degrees pointing west from MLAC are used (256 events - see figure 5.1) This chapter covers 105 events from this set which have been put in subsets. Figure 5.2 gives a distance-depth section for these events. Unfortunately, a number of events are not very well recorded or processed (see appendix I). Therefore 38% of the events looked at could not be used in the analysis.

The information for the focal mechanisms also comes from the NCEDC.
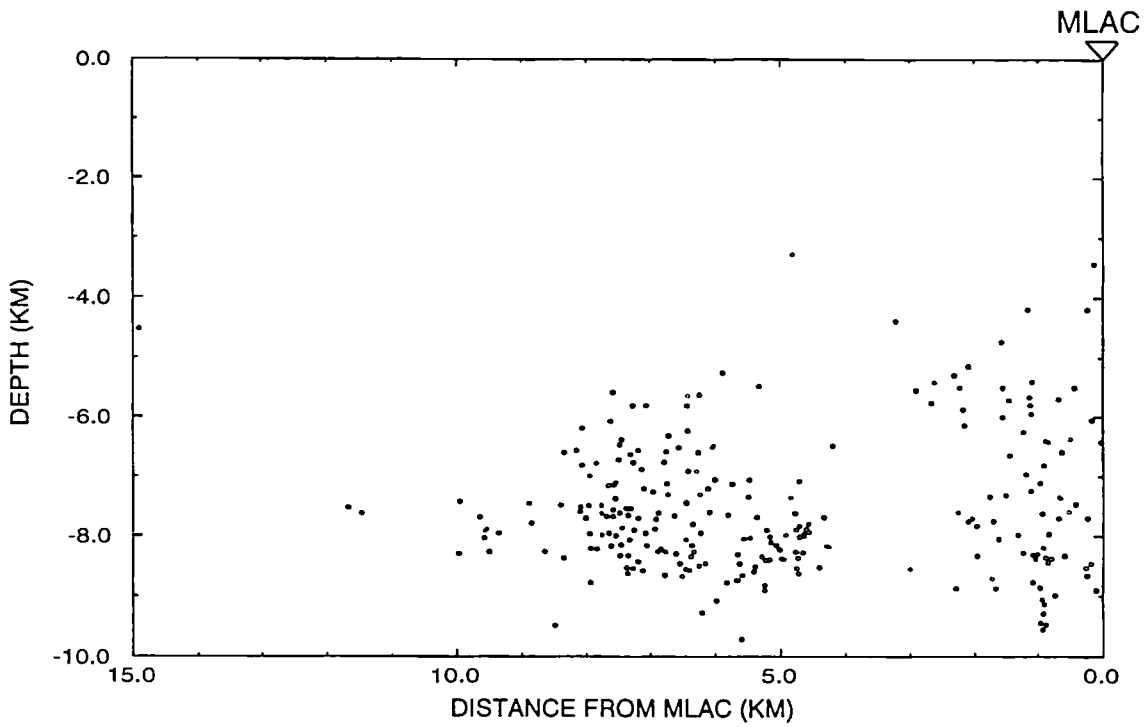
Figure 5.2: *Distance from MLAC versus depth for the 256 events within the fan pointing west from MLAC.*

5.3   Data processing

The two horizontal components are rotated to get the tangential and radial components. The traces used are velocity waveforms, and no integration is applied before the calculations. All components are filtered with a butterworth bandpass filter (corner frequencies 0.01, 2 Hz, 2 poles). For the cross-correlation, different time windows are used for different subsets, depending on how far away from MLAC the events occured. The correlation maximum is picked within 2 seconds at two sides of the zero axis of the correlation trace.

5.4 Results

The analysis was performed on different subsets within the 256 events in the previously discussed 10° fan. Subsets consist of events at a particular distance range from the station with varying depths. The different figures will show where the subsets are located.

5.4.1 Subset 1

Subset 1 contains the events within the distance range of 7.29 to 7.65 km from MLAC. Of the 27 events, 5 cannot be used in the analysis because of bad traces (see appendix I). The 22 events are split up in set 1^a and 1^b. Both sets containing 11 events, which is the maximum number that can be used by the algorithm. The division into two sets is based on the average cross-correlation coefficients. The two sets consist of events that correlate reasonably well with each other and not with the events in the other set. The first 18 seconds of each trace are used for cross-correlation. Minimization without the P-S time series gives the following minimized orders:

subset 1^a: 1,2,4,3,8,9,12,15,20,22,21
subset 1^b: 6,13,7,5,10,11,17,16,18,19,14
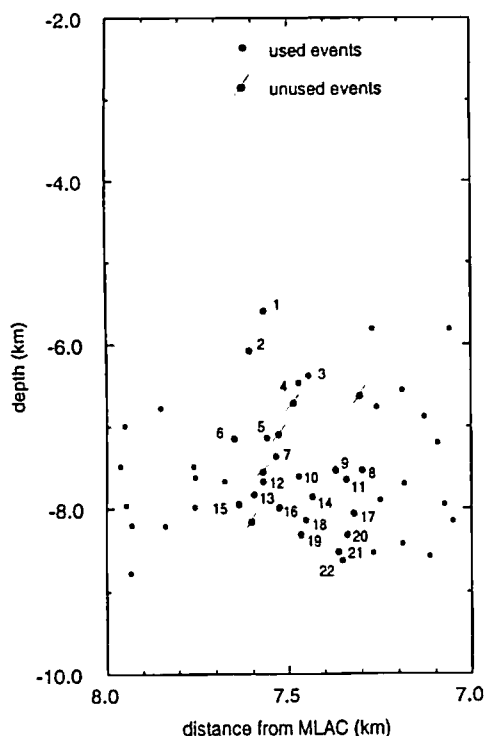
Figure 5.3: *Depth section subset* 1^{a+b}

The numbers correspond with the numbers in the figures and tables. Figure 5.3 shows the depth section of this part of the dataset. The used events are the black-filled dots; the events that could not be used are marked by a line through the dot. Figure 5.4 contains the waveforms of the three components for subset1^a and figure 5.5 shows the focal mechanisms. Table 5.1 contains the average cross-correlation matrix for subset 1^a after minimization. Figures 5.6 & 5.7 and table 5.2 contain the same information but for subset1^b.

Subset 1a: three components, vertical, tangent and radial, minimized order

fig 5.4

entry

												number of event
1	1.00	0.53	0.48	0.44	0.45	0.52	0.55	0.40	0.31	0.28	0.27	21
2	0.53	1.00	0.48	0.53	0.53	0.56	_0.61_	0.59	0.39	0.46	0.32	22
3	0.48	0.48	1.00	_0.75_	0.69	0.50	0.65	0.58	0.53	0.50	0.28	20
4	0.44	0.53	0.75	1.00	_0.81_	0.78	0.72	0.70	0.65	0.61	0.29	15
5	0.45	0.53	0.69	0.81	1.00	0.80	0.72	0.76	0.69	0.56	0.29	12
6	0.52	0.56	0.50	0.78	0.80	1.00	0.84	0.74	0.63	0.57	0.38	9
7	0.55	0.61	0.65	0.72	0.72	0.84	1.00	0.71	0.54	0.53	0.32	8
8	0.40	0.59	0.58	0.70	0.76	0.74	0.71	1.00	0.72	0.70	0.39	3
9	0.31	0.39	0.53	0.65	0.69	0.63	0.54	0.72	1.00	0.77	0.36	4
10	0.28	0.46	0.50	0.61	0.56	0.57	0.53	0.70	0.77	1.00	0.57	2
11	0.27	0.32	0.28	0.29	0.29	0.38	0.32	0.39	0.36	0.57	1.00	1

closeness number = 0.07

Table 5.1: *Final average matrix for minimized order, subset 1*[a]
(shallower event is on bottom of matrix!)

Figure 5.5:
Focal mechanisms, subset 1[a]

The focal mechanisms cannot be found in the catalogs for all of the events. Subset 1[a] has no focal mechanism for event #15 (mag=2.4) and 20 (mag=1.8). Subset 1[b] has no focal mechanism for event # 13 (mag=1.7) and #19 (mag=1.9).

From the average cross-correlation coefficients we can conclude that event #15 will have a focal mechanism very similar to the one calculated for event #12 (ϕ=0.81). (The cross-correlation values mentioned in the text are underlined in the different tables.)

Event #20 is in turn similar to event #15 (ϕ=0.75) and will have a similar focal mechanism. Of course, this cannot be verified for these events, but for the other events no contradictions in the information given by focal mechanisms and cross-correlation coefficients occur.

For example event #22 has a rather low cross-correlation coefficient with all the other events (highest ϕ=0.61 with event #8), and also a focal mechanism that is not very similar to any others (event #1 and #8 are closest). This indicates that event #1 and #22 are too far apart to have similarities (in this region), and that, although event #8 is closer spatially, the distance between the events and/or the difference in focal mechanism is too big to give a recognizable similarity in the waveforms.

31

Subset 1b: three components, vertical, tangent and radial, minimized order

fig 5.6

number
of event

1	1.00	0.61	0.12	0.08	0.14	0.30	0.10	0.11	0.16	0.13	0.06	6
2	0.61	1.00	0.39	0.45	0.59	0.63	0.46	0.49	0.31	0.19	0.15	13
3	0.12	0.39	1.00	0.87	0.59	0.49	0.41	0.35	0.27	0.60	0.28	7
4	0.08	0.45	0.87	1.00	0.69	0.56	0.53	0.40	0.27	0.29	0.30	5
5	0.14	0.59	0.59	0.69	1.00	0.80	0.80	0.76	0.54	0.24	0.22	10
6	0.30	0.63	0.49	0.56	0.80	1.00	0.18	0.71	0.42	0.24	0.22	11
7	0.10	0.46	0.41	0.53	0.80	0.83	1.00	0.74	0.46	0.28	0.27	17
8	0.11	0.49	0.35	0.40	0.76	0.72	0.74	1.00	0.75	0.28	0.30	16
9	0.16	0.31	0.27	0.27	0.54	0.42	0.46	0.75	1.00	0.57	0.59	18
10	0.13	0.19	0.30	0.29	0.24	0.24	0.28	0.28	0.57	1.00	0.83	19
11	0.06	0.15	0.28	0.30	0.22	0.21	0.27	0.30	0.59	0.83	1.00	14

closeness number = 0.13

Table 5.2: *Final average matrix for minimized order, subset 1[b]*
(shallower event is on bottom of matrix!)

Figure 5.7:
*Focal mechanisms,
subset 1[b]*

6
7
5
10
11
17
16
18
14

Event #13 has the highest similarity with event #6 (ϕ=0.61) and therefore is expected to have a focal mechanism in the direction of the one of event #6, but they won't be the same. Event #19 has a high correlation (ϕ=0.83) with event #14 and therefore will have a very similar focal mechanism. Event #16 (mag=2.5) has been assigned a very different focal mechanism from the others in subset1[b]. Based on the cross-correlation coefficients, this does not seem correct. A high correlation (ϕ~0.74) with events #18, #17, #11and #10 suggests that the focal mechanism for event #16 should be closer to those from events #18, #17, #11 and #10 which are similar.

5.4.2 Subset 2

Subset 2 contains the events within a distance range of 0.85 to 1.2 km from MLAC. Of the 28 events, 2 cannot be used at all, and another 9 events have only two recorded (see appendix I). For the 17 events that remain, the first 12 seconds of each trace is used for cross-correlation, and the 11 events with the highest cross-correlations are used in the minimization. The average cross-correlation coefficients show a separation into two groups. The minimized order is: 2,5,6,13,15,16,9,7,11,8,10

The numbers correspond with the numbers in the figures and tables. Figure 5.8 shows the depth section of this part of the dataset, and figures 5.9 and 5.10 contain the waveforms and focal mechanisms of the eleven events with three components used in the minimization. Table 5.3 contains the average cross-correlation matrix.
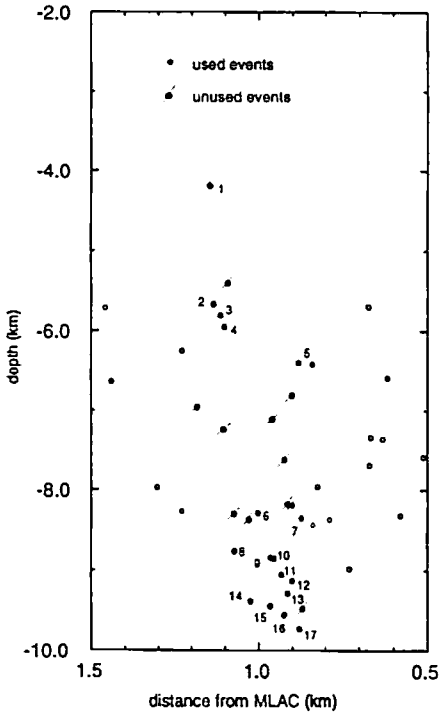
Figure 5.8: *Depth section subset 2*

Subset 2: three components, vertical, tangent, and radial, minimized order

VERTICAL

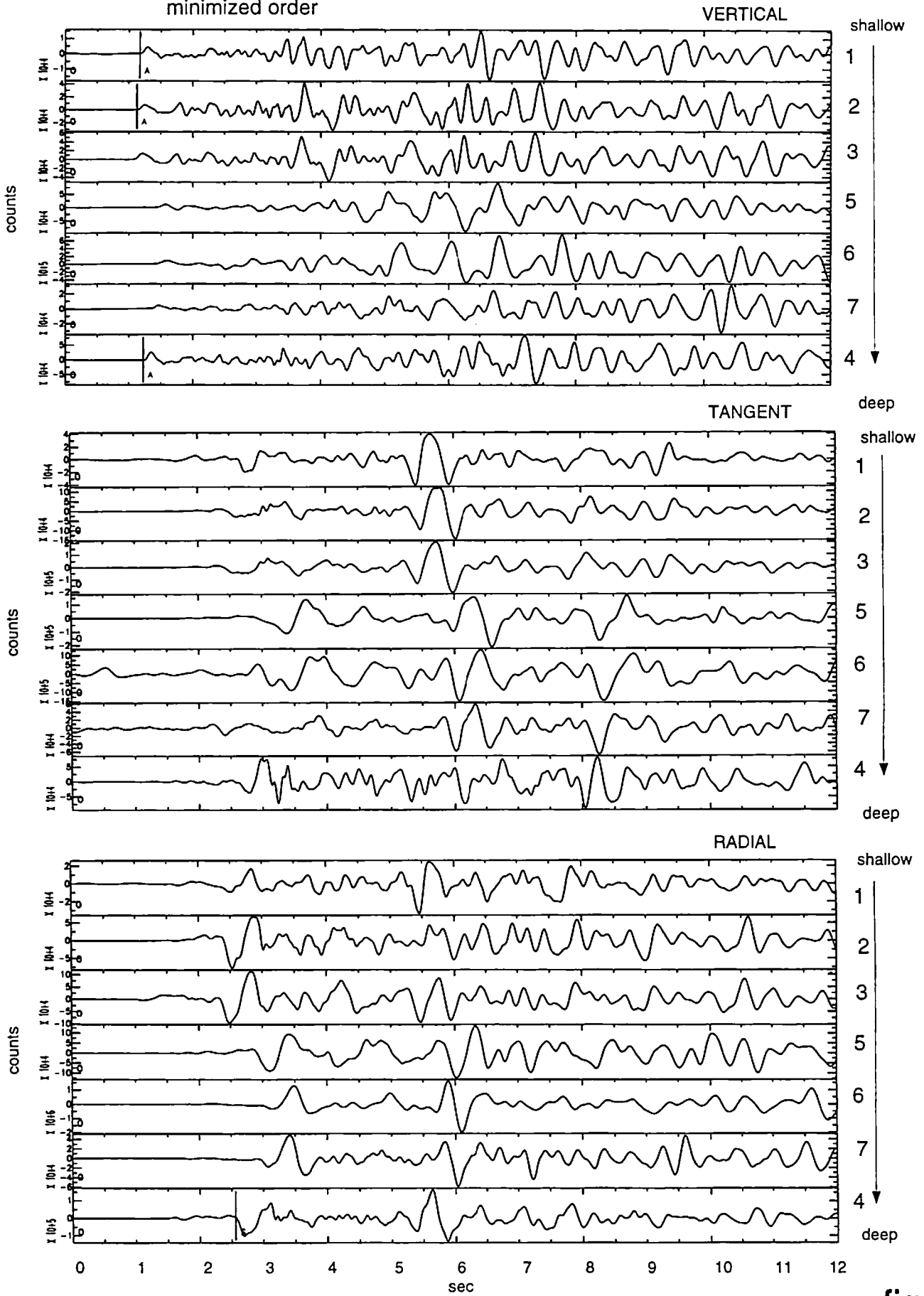fig 5.9

35

entry

												number of event
1	1.00	0.65	0.59	0.56	0.55	0.56	0.51	0.46	0.34	0.27	0.23	10
2	0.65	1.00	0.88	0.76	0.79	0.67	0.62	0.45	0.40	0.34	0.30	8
3	0.59	0.88	1.00	0.84	0.86	0.71	0.67	0.56	0.51	0.32	0.30	11
4	0.56	0.76	0.84	1.00	0.82	0.72	0.69	0.62	0.52	0.31	0.29	7
5	0.55	0.79	0.86	0.82	1.00	0.75	0.60	0.59	0.51	0.26	0.25	9
6	0.56	0.67	0.71	0.72	0.75	1.00	0.83	0.79	0.49	0.27	0.25	16
7	0.51	0.62	0.67	0.69	0.60	0.83	1.00	0.82	0.51	0.30	0.32	15
8	0.46	0.45	0.56	0.62	0.59	0.79	0.82	1.00	0.60	0.32	0.32	13
9	0.34	0.40	0.51	0.52	0.51	0.49	0.51	0.60	1.00	0.29	0.24	6
10	0.27	0.34	0.32	0.31	0.26	0.27	0.30	0.32	0.29	1.00	0.85	5
11	0.23	0.30	0.30	0.29	0.24	0.25	0.32	0.32	0.24	0.85	1.00	2
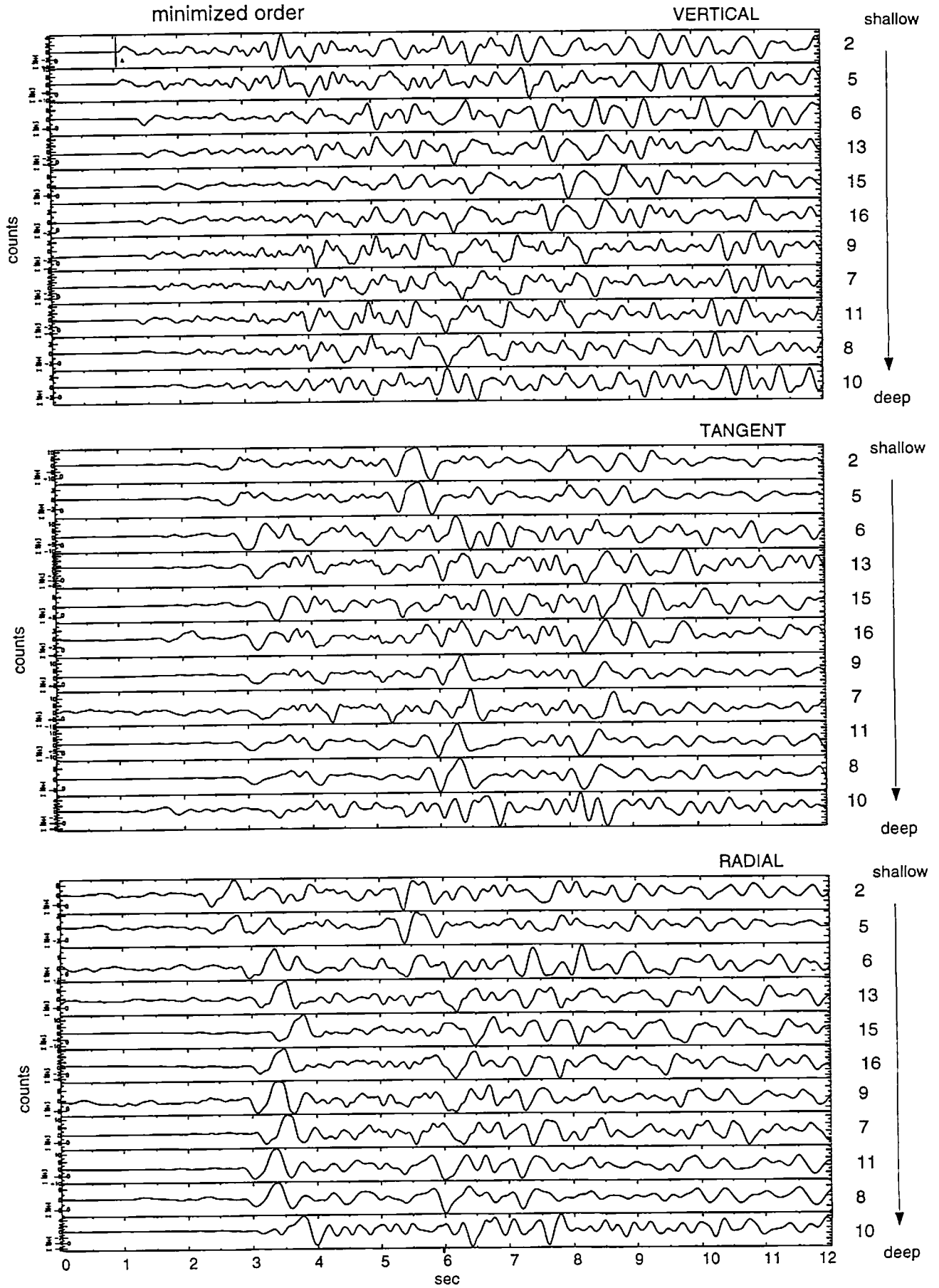
closeness number = 0.03

Table 5.3: *Final average matrix for minimized order, subset 2
(shallower event is on bottom of matrix!)*

Figure 5.10:
*Focal mechanisms,
subset 2*

The separation into two groups lies between events #5 and #6. The cross-correlation between these two is only 0.29, and there are no good correlations with any of the other events. The separation seems to be depth dependent, as events #2 and #5 are separated from the other events by more than 2 km. The focal mechanisms do not really give a clearer explanation of this phenomenon. The focal mechanism of event #8 seems to be wrong, as the cross-correlation coefficients with events #11, #7 and #9 are high (0.88, 0.76 and 0.79), but the focal mechanisms do not correspond. Events #16, #15 and #13 have similar focal mechanisms and correspondingly high cross-correlation coefficients

($\phi_{16,15}$ = 0.83, $\phi_{15,13}$ = 0.82) . The focal mechanism of event #6 is very different from event #13, but still these events give a correlation of 0.60.

In order to really say something about locations, more events with similar focal mechanisms will be needed. For this moment we can conclude that the events with high correlations are (according to the catalog) within 1.5 km of each-other. The events not used in the minimization (because a selection of 11 had to be made) are #1, #3, #4, #12, #14 & #17. From these, #3 and #4 correlate very well with events #2 and #5 (ϕ>0.8). Events #12, #14 and #17 correlate very well with event 16 (ϕ>0.8). Event #1 does not correlate well with any of the other events.

5.4.3 Subset 3

Subset 3 contains the events within the distance range of 1.2 to 1.75 km from MLAC. Of the 14 events, 7 cannot be used (see appendix I). The first 12 seconds of each trace are used for cross-correlation.

The minimized order is: 1,2,3,5,6,7,4
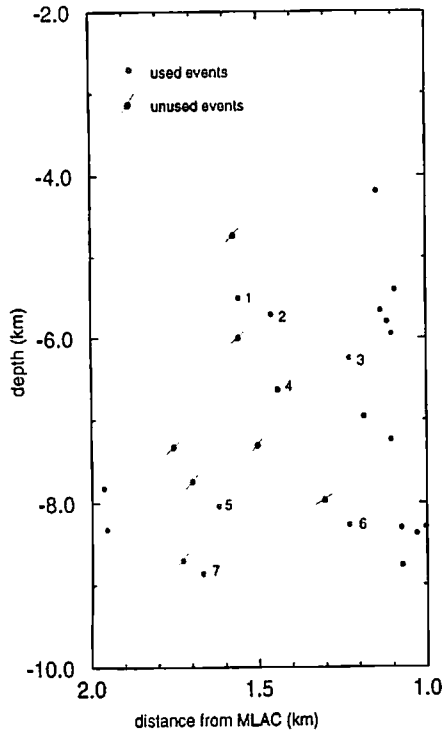
The numbers correspond with the numbers in the figures and tables. Figures 5.11 to 5.13 show the depth section, traces and focal mechanisms for subset 3. Table 5.4 contains the average cross-correlation matrix after minimization.

Figure 5.11: *Depth section subset 3*

Subset 3: Three components, vertical, tangent and radial, minimized order

fig 5.12

entry

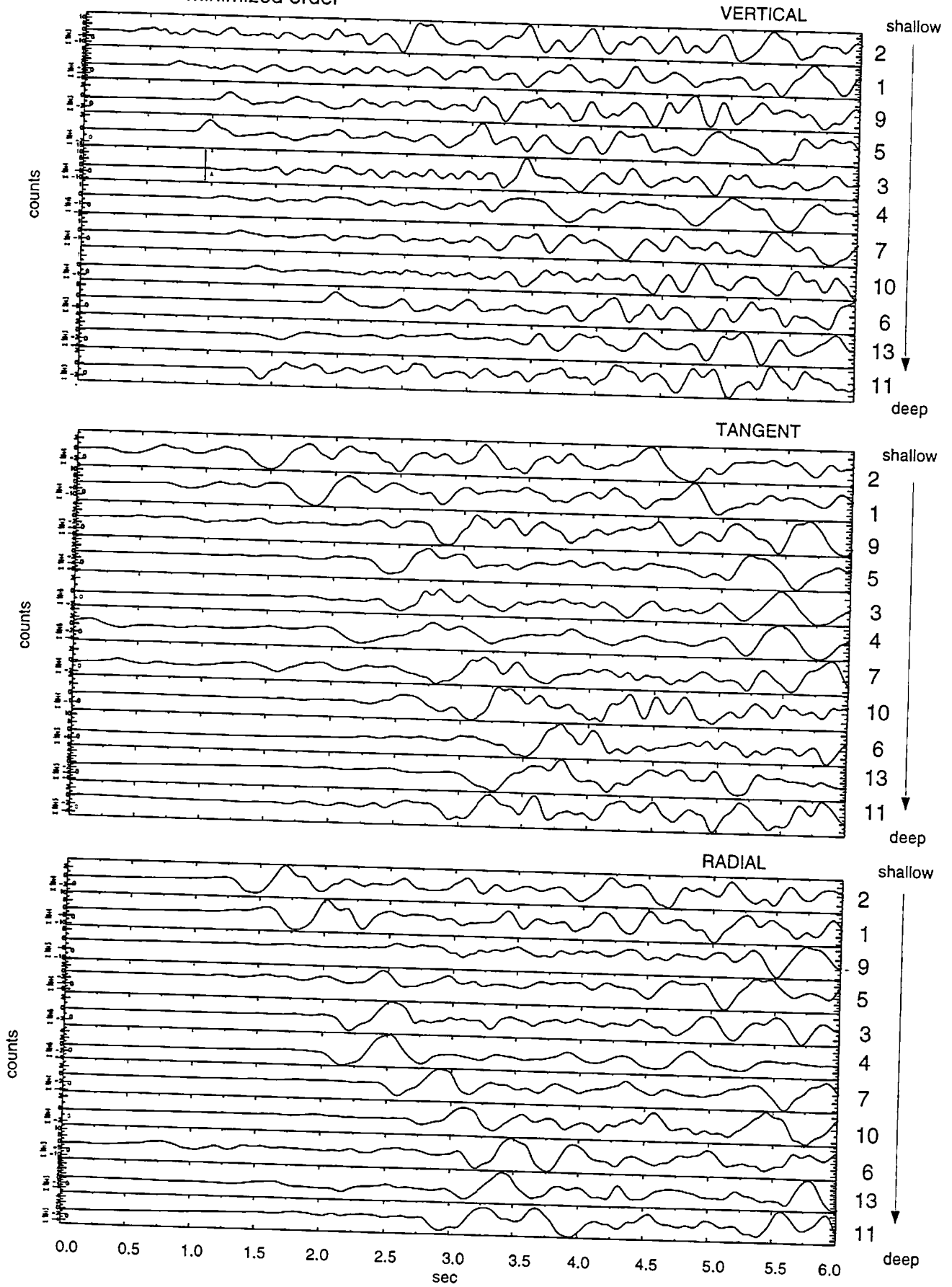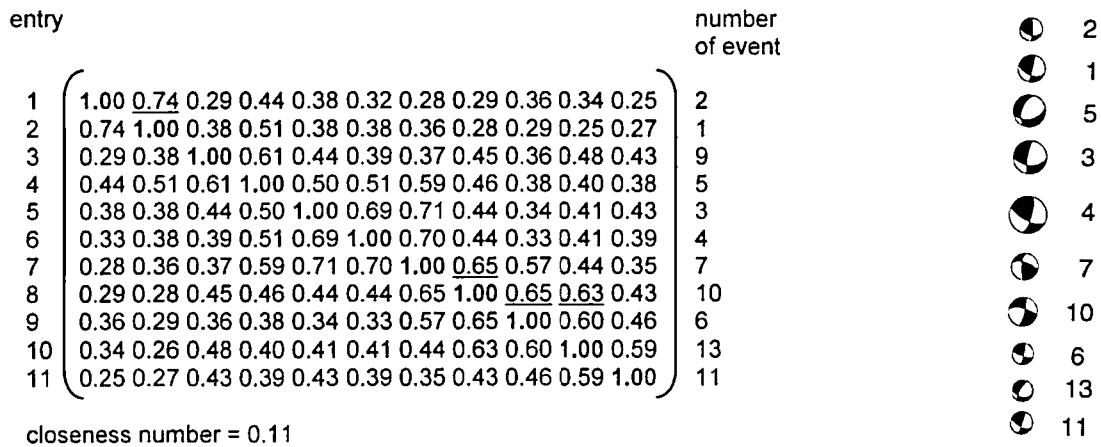| entry								number of event
1	1.00	0.59	0.63	0.41	0.30	0.31	0.28	1
2	0.59	1.00	0.88	0.46	0.28	0.33	0.36	2
3	0.63	0.88	1.00	0.48	0.34	0.33	0.34	3
4	0.41	0.46	0.48	1.00	0.62	0.50	0.32	5
5	0.30	0.27	0.34	0.62	1.00	0.71	0.32	6
6	0.30	0.32	0.33	0.50	0.72	1.00	0.28	7
7	0.28	0.36	0.34	0.32	0.32	0.28	1.00	4

closeness number = 0.01

Table 5.4: *Final average matrix for minimized order, subset 3*
(shallower event is on top of matrix!)
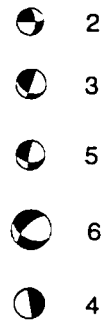
Figure 5.13:
Focal mechanisms, subset 3

The average cross-correlation coefficients show that the traces of events #2 and #3 are highly similar ($\phi_{2,3}$ = 0.88), but the focal mechanisms as given by NECDC are different. Event #4 does not correlate with any of the other events, which can be explained by the completely different focal mechanism it has.

The cross-correlation coefficient between #3 and #5 (ϕ=0.48) is very low. The focal mechanisms are fairly similar, so the distance in this case must have been caused by the difference in the recorded traces, if we assume that the focal mechanism is right. Event #7 has no focal mechanism given, but has a reasonable correlation with event #6 (ϕ=0.72), so the mechanism will be similar.

5.4.4 Subset 4

Subset 4 contains the events within the distance range of 1.9 to 2.35 km from MLAC (there are no events between 1.75 and 1.9km). Of the 12 events, only 6 have three components. The first 9 seconds of each trace are used for cross-correlation. Five of the events are within 1 km (depth) in the location-set used.

The minimized order for the 6 events is: 1,2,3,4,5,6.
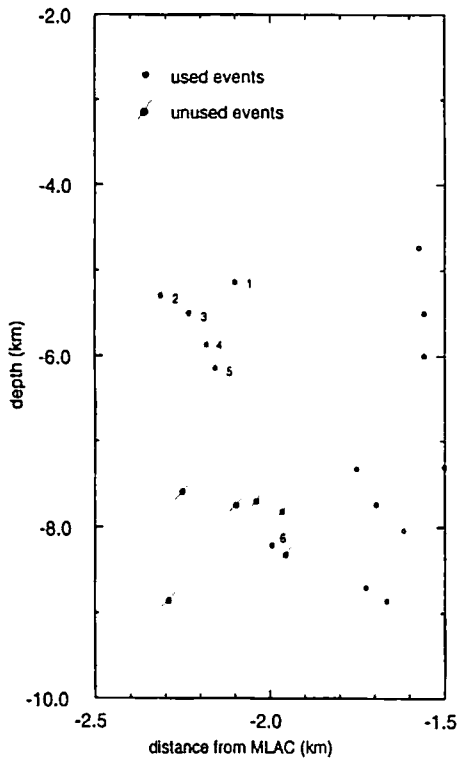
Figure 5.14: *Depth section subset 4*

The numbers correspond with the numbers in the figures and tables. Figures 5.14 to 5.16 show the depth section, traces and focal mechanisms for subset 4. Table 5.5 contains the average cross-correlation matrix after minimization.

Subset 4: three components, vertical, tangent and radial, minimized order

fig 5.15

entry							number of event
1	1.00	0.49	0.44	0.41	0.37	0.39	6
2	0.49	1.00	0.89	0.76	0.61	0.47	5
3	0.44	0.89	1.00	0.77	0.60	0.47	4
4	0.41	0.76	0.77	1.00	0.85	0.69	3
5	0.37	0.61	0.60	0.85	1.00	0.78	2
6	0.39	0.47	0.47	0.69	0.78	1.00	1
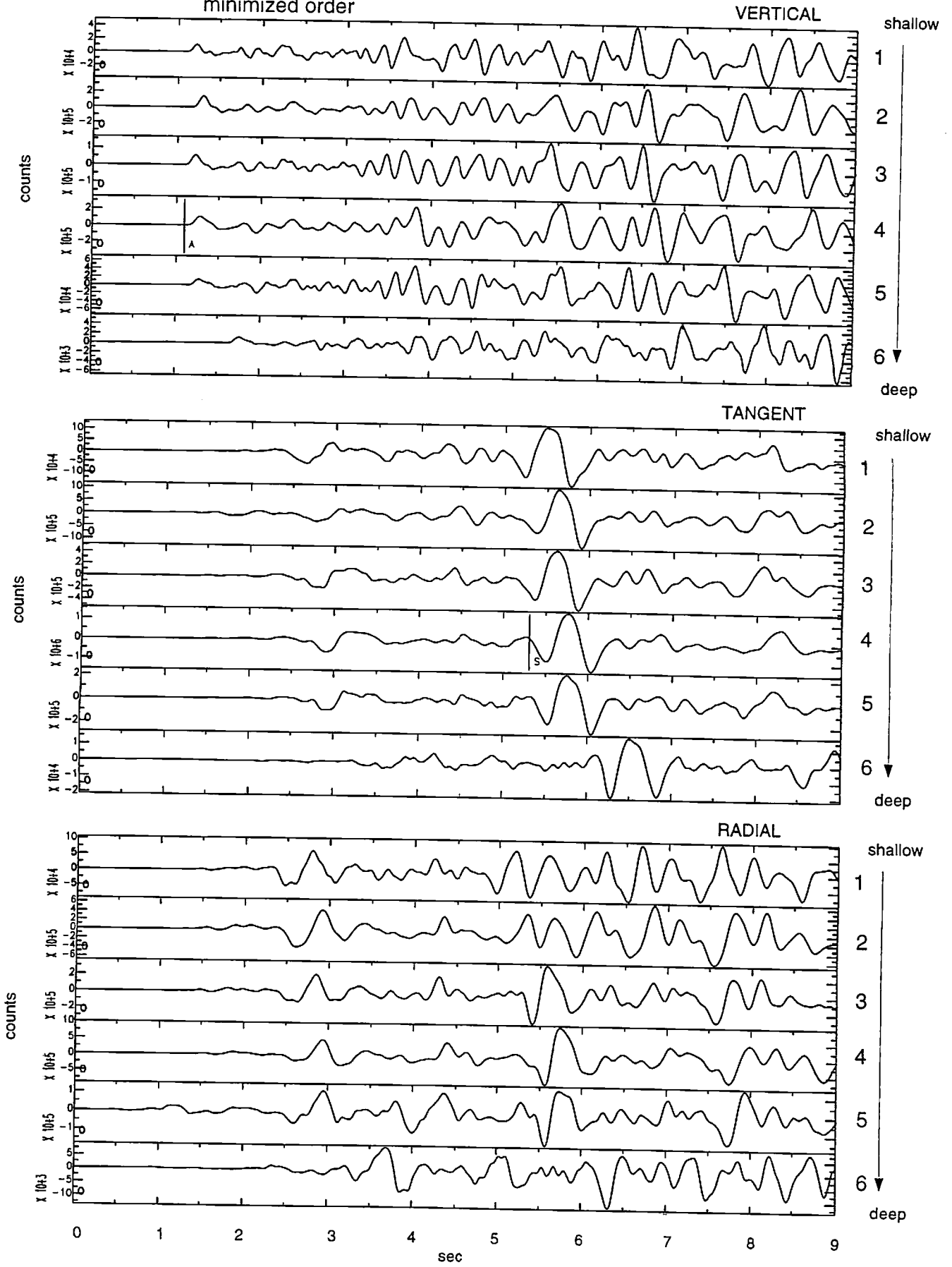
closeness number = 5×10^{-3}
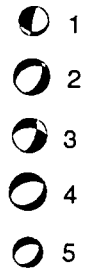
Table 5.5: *Final average matrix for minimized order, subset 4 (shallower event is on top of matrix!)*

Figure 5.16: *Focal mechanisms, subset 4*

Because of the similar focal mechanisms and the small separations, the cross-correlation coefficients are high for five of the six events. Event #6 does not correlate well with any of the other events (ϕ_{max}=0.49). Unfortunately there is no focal mechanism available for event #6, so one cannot say if the fact that the cross-correlation coefficients are low is due to the distance only, or if the focal mechanism is very different as well.

5.4.5 Subset 5

Subset 5 contains the events within the distance range of 0 to 0.85 km from MLAC. Of the 24 events, 11 cannot be used (see appendix I). This leaves 13 events to work with. The first 6 seconds of each trace are used for cross-correlation. This is rather short, but it is limited by the recorded time window for some events. Minimization is done for the eleven events that have the highest correlation coefficients.

The minimized order is: 2,1,9,5,3,4,7,10,6,13,11.

The numbers correspond with the numbers in the figures and tables. Figures 5.17 to 5.19 show the depth section, traces and focal mechanisms for subset 5. Table 5.6 contains the average cross-correlation matrix after minimization.
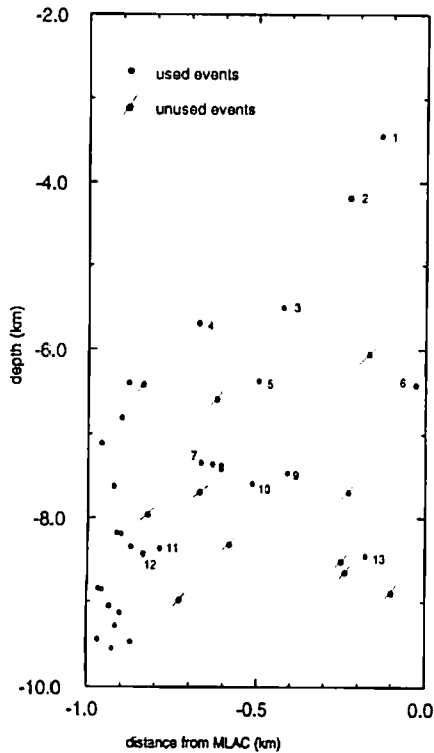
Figure 5.17: *Depth section subset 5*

Subset 5: three components, vertical, tangent and radial,
minimized order

fig 5.18

entry												number of event
1	1.00	0.74	0.29	0.44	0.38	0.32	0.28	0.29	0.36	0.34	0.25	2
2	0.74	1.00	0.38	0.51	0.38	0.38	0.36	0.28	0.29	0.25	0.27	1
3	0.29	0.38	1.00	0.61	0.44	0.39	0.37	0.45	0.36	0.48	0.43	9
4	0.44	0.51	0.61	1.00	0.50	0.51	0.59	0.46	0.38	0.40	0.38	5
5	0.38	0.38	0.44	0.50	1.00	0.69	0.71	0.44	0.34	0.41	0.43	3
6	0.33	0.38	0.39	0.51	0.69	1.00	0.70	0.44	0.33	0.41	0.39	4
7	0.28	0.36	0.37	0.59	0.71	0.70	1.00	0.65	0.57	0.44	0.35	7
8	0.29	0.28	0.45	0.46	0.44	0.44	0.65	1.00	0.65	0.63	0.43	10
9	0.36	0.29	0.36	0.38	0.34	0.33	0.57	0.65	1.00	0.60	0.46	6
10	0.34	0.26	0.48	0.40	0.41	0.41	0.44	0.63	0.60	1.00	0.59	13
11	0.25	0.27	0.43	0.39	0.43	0.39	0.35	0.43	0.46	0.59	1.00	11

closeness number = 0.11

Table 5.6: *Final average matrix for minimized order, subset 5 (shallower event is on top of matrix!)*

Figure 5.19: *Focal mechanisms, subset 5*

Events #1 and #2 (ϕ = 0.74) do not correlate very well with the rest of the events. They seem to be spatially isolated from the other events, because the focal mechanisms are not that very much different.

Looking at events #7, #10, #6 and #13, we see that the cross-correlation coefficients are reasonable (ϕ>0.63). The focal mechanism of events #7, #10 and #6 all are close enough, considering that ϕ does not come above 0.65. Event #13 seems to be a bit further off.

5.5 Conclusions Long Valley

This dataset shows that the use of cross-correlations in determining relative locations for sure has a future. Within the total set of events very coherent clusters of events are present and the algorithm as presented in this thesis does recognize these clusters and places them together. For the method to work fully as desired, more events will have to be recorded in order to show a gradual change of signature between the different characteristic clusters.

6 Northridge

6.1 Introduction

After having looked at the Long Valley events, an area is next considered where, hopefully, the events will group a little less in very coherent clusters that do not correlate with other clusters. The Northridge sequence has a larger spacing, but at the same time higher magnitudes than the set in from Long Valley. Again it is attempted to say something about the consistency in the information · coming from cross-correlations, locations and focal mechanisms.

6.2 Data Selection

The events used occured between January 17th, 1994 (mainshock) and July 1994. The waveforms are obtained via the SCEC-datacenter. By searching the SCEC catalog, 55 events larger than magnitude 4 are found as well as 64 events between magnitude 3.6 and 3.9 recorded at station Pasadena (PAS), maintained by Caltech. Events with lower magnitudes cannot be used at the frequency band for which the traces are filtered. The set is sorted for decent waveforms, which is defined as one which has at least 33 seconds of clean record (no clipping or anything else) for all three components. After this selection, 19 events of magnitude larger than 4 and 14 events in the magnitude range of 3.6 to 3.9 remain. This brings the total set to 33 events. The events cover an area of 30 km (width) by 20 km (depth) in depth section (figure 6.1), and in planview they cover an area of 20 km (NS) and 30 km (EW) (figure 6.2). The location parameters are from the relocated sets by Egill Hauksson and Jim Mori (simultaneous inversion of hypocenters and 3-D velocity: D. Eberhart-Phillips, 1993). Focal mechanisms are calculated by Lupei Zhu (full waveform: L. Zhu 1996) and Egill Hauksson (first motion).
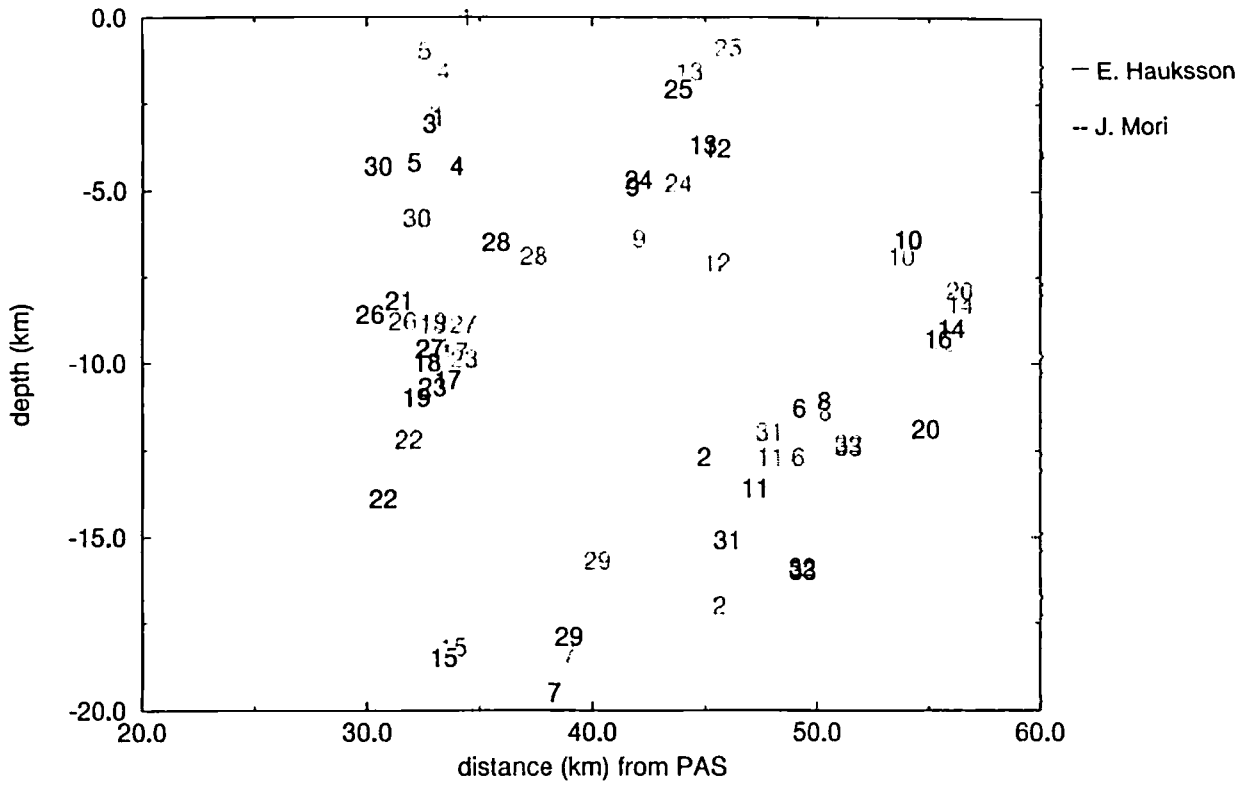
Figure 6.1: *Depth section of used events for Northridge aftershock sequence. Both E. Hauksson and J. Mori's dataset are shown.*
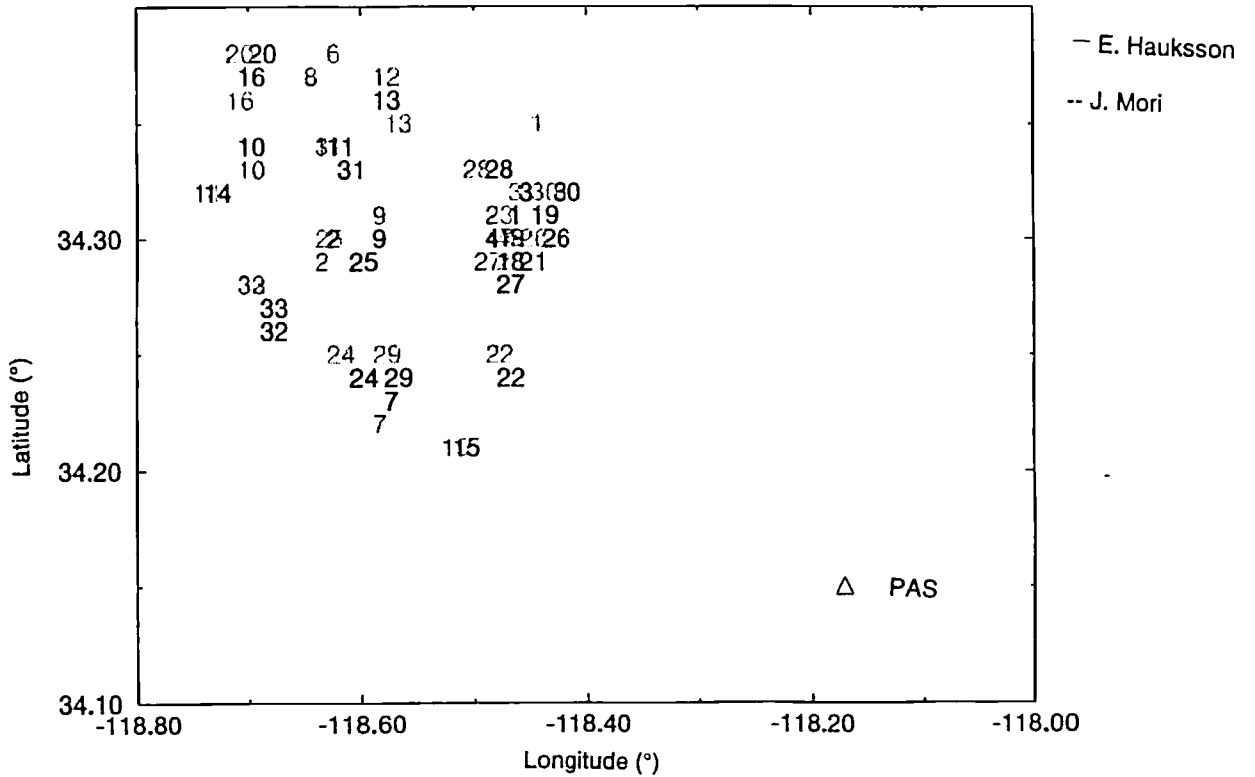
Figure 6.2: *Plan view of used events for Northridge aftershock sequence. Both E. Hauksson and J. Mori's dataset are shown.*

47

6.3 Data processing

The horizontal components are rotated to the tangential and radial direction. The retrieved traces are velocity waveforms, and integration to displacement is applied before the calculations. All components are filtered with a butterworth filter (corner frequencies 0.05, 0.4 Hz, two poles). For the cross-correlation the first 33 seconds of each trace is used. The maximum correlation value is picked within 5 seconds at two sides of the zero axis of the (cross-) correlation trace.

6.4 Results

6.4.1 Total set correlations

First the complete set of events is looked at to see how well they correlate. The cross-correlation matrices are calculated and averaged for the three components. The result is given in table 6.1.

	$\phi \geq 0.8$	$0.7 \leq \phi < 0.8$	$0.6 \leq \phi < 0.7$
average of 3 comp.	5	5	9

Table 6.1: *Overview of average cross-correlation coefficients for the Northridge data-set*

For the events that have a high average cross-correlation coefficient, an analysis has been made in which the cross-correlation coefficients, the focal mechanisms given by L. Zhu and E. Hauksson and the depth and horizontal locations from E. Hauksson and J. Mori are compared. The result of the analysis can be found in table 6.2. The numbers of the events correspond to the numbers in the depth section given in figure 6.1. Figure 6.1 contains both the locations from E. Hauksson and J. Mori. The different subsets, which will be discussed later are shown by broken lines. The focal mechanisms for all the events are plotted in figure 6.3.

48

Event n°	average φ	focal mech.	depth	horizontal distance	remarks
4,27	0.90	+	--	+	
18,19	0.82	+(L)	+	+	
11,31	0.82	+-	+	+	noisy 31, radial
32,33	0.81	++	++	++	noisy 32
6,8	0.80	+	+	+	
======	======	======	======	======	======
10,14	0.75	+(L)	+	+	mag. 5.6 <-> 4.1
1,18	0.74	+(L)	--	++	noisy 1
7,29	0.73	+-	+-	+-	noisy 29
1,19	0.71	+(L)	--	++	noisy 1
19,27	0.70	++(L)	+	+	
18,23	070	++	+	+	
======	======	======	======	======	======
6,11	0.67	+	+	+(J)	mag. 4.9 <-> 4.0
12,13	0.64	+	+(E)	+	
2,31	0.64	+(L)	+-	+-	noisy 31, radial
14,20	0.63	--	+(J)	+	fm 3.6 well determined?
10,20	0.62	--	+	+(J)	mag. 5.6 <-> 3.6
4,19	0.62	+(L)	--	+	see 4,27
17,19	0.61	+(L)	+	+	
21,26	0.60	??	+	+	no focal mech. for 21

The scale is as follows: ++ the same → agreement between catalog and cross-correlation
+   close
+-   similar
--   different → no agreement between catalog and cross-correlation
(L) focal mechanism better from Lupei Zhu
(J) relative location better from Jim Mori or (E) from Egill Hauksson

Table 6.2: *Comparison of cross-correlation coefficients, focal mechanisms and location parameters for different events.*
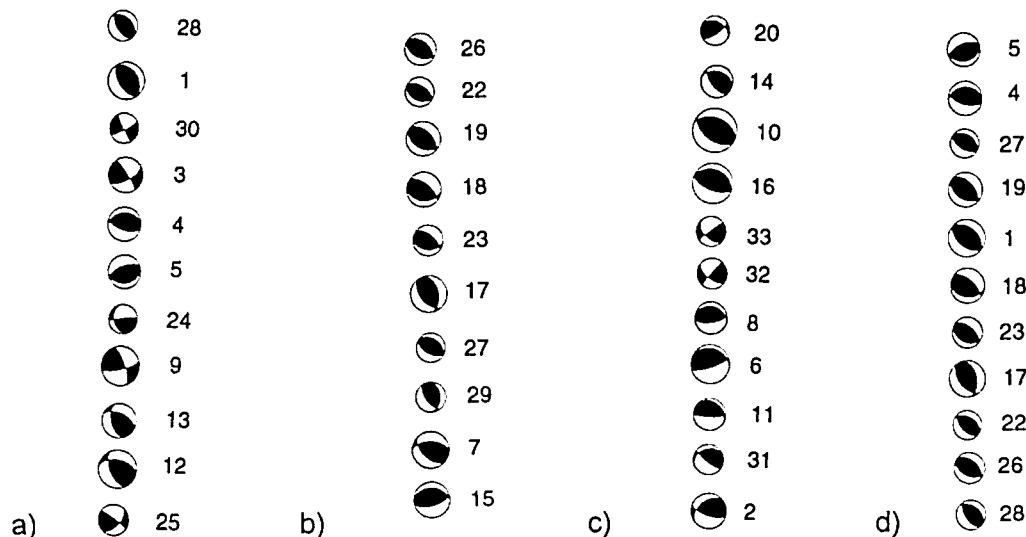
Figure 6.3[a+b+c+d]: *Focal mechanisms the Northridge aftershock sequence*

For all the events shown here one can say that a high (>0.6) cross-correlation coefficient indicates a similarity in focal mechanism. As the magnitude of the event gets lower, the focal mechanism can be less well constrained. This is what might be happening with event #20.

The distances and depths are mostly close or the same. Only a few times does the depth give an inconsistent picture. This can be explained the fact that the depth is the least accurately determined parameter in other inversions.

- Event pair #4 and #27 is a clear example of inconsistency between the different parameters. The waveforms are almost the same, the focal mechanisms and horizontal distances are close, yet the depth is considerably different (>5km) for both location catalogs. For a few events the depths calculated by L. Zhu in his full waveform inversion are considered. For this pair of events, #4 is shifted down to a depth of 11 km. Unfortunately, L. Zhu's set does not have a depth determined for #27, but as it is a full waveform inversion, the events should end up closer together.

- The same is true for events #4 and #19. The cross-correlation coefficient is smaller than for the previous pair, but one would expect them to be closer together than they are placed by the two data-sets used. For this pair L. Zhu gave a depth of 11 km for both events. There is no value given for the horizontal location, so one can say nothing about the actual distance between the events, but it is promising.

- Also for the pairs #1 & #18 and #1 & #19 the depth does not make sense. Unfortunately, #1 does not show up in L. Zhu's list of focal mechanisms.

6.4.2 Subsets

The total set of 33 events is split up into several subsets as shown in figure 6.1. For each subset of 11 events, the minimized order is calculated with and without the S-P time series. The results for each subset are described below.

6.4.2.1 Subset 1

Subset 1 contains the 11 shallowest events. At first the minimized order seems to show a separation between the events further away from and the events closer to PAS.
Minimized order: 28,1,30,3,4,5,24,9,13,12,25.
The numbers correspond with the numbers in the figures and tables. Figure 6.4 and 6.5 contain the focal mechanisms and waveforms of the eleven events, with three components used in the minimization. The averages of the cross-correlation coefficients, are rather low (see table 6.3).

entry												t(s)-t(p)	number of event
1	1.00	0.39	0.22	0.28	0.22	0.24	0.25	0.25	0.26	0.26	0.34	6.2	25
2	0.39	1.00	0.65	0.46	0.29	0.26	0.24	0.25	0.24	0.21	0.25	6.2	12
3	0.22	0.65	1.00	0.37	0.33	0.27	0.26	0.26	0.27	0.20	0.18	6.3	13
4	0.28	0.46	0.37	1.00	0.46	0.28	0.27	0.23	0.18	0.20	0.44	5.7	9
5	0.22	0.29	0.33	0.46	1.00	0.22	0.28	0.25	0.22	0.19	0.27	5.6	24
6	0.24	0.26	0.27	0.28	0.22	1.00	0.60	0.27	0.25	0.36	0.33	4.6	5
7	0.25	0.24	0.26	0.27	0.28	0.60	1.00	0.57	0.30	0.37	0.47	4.5	4
8	0.25	0.25	0.26	0.23	0.25	0.27	0.57	1.00	0.56	0.40	0.44	4.8	3
9	0.26	0.24	0.27	0.18	0.22	0.25	0.30	0.56	1.00	0.32	0.30	4.4	30
10	0.26	0.21	0.20	0.20	0.19	0.36	0.37	0.40	0.32	1.00	0.33	4.4	1
11	0.34	0.25	0.18	0.44	0.27	0.33	0.47	0.44	0.30	0.33	1.00	4.0	28

closeness number = 0.29

Table 6.3: *Final average matrix for minimized order, subset 1*

Figure 6.4:
Focal mechanisms, subset 1

Only events 5 & 4 and 12 & 13 have a cross-correlation coefficient of 0.60 and 0.65. All the others are lower. This makes it a rather poor set with which to actually work. The traces are shown in figure 6.5. By looking at the focal mechanisms, one can see that the set is rather diverse, and the potential for this algorithm to work with this subset is rather low. The S-P time series do not change the separation between events further away and the ones closer. The minimization without the S-P times gives the same two clusters: 30,3,4,5,1,28 versus 9,24,13,12,25. The distinction is not clear from the cross-correlation coefficients, but shows up spatially in the minimized order.

Northridge, subset 1: three components, vertical, tangent and radial, minimized order
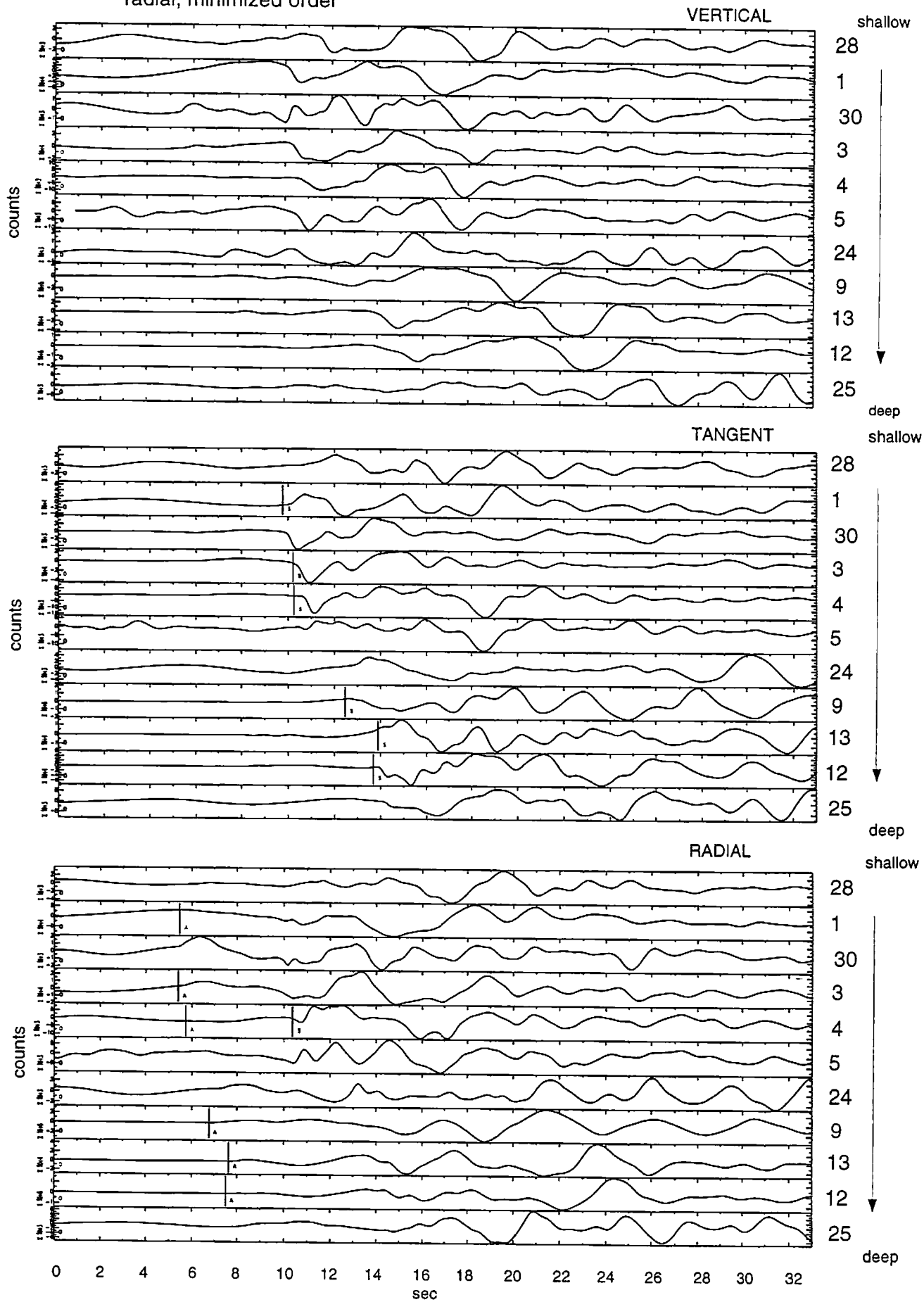
fig 6.5

6.4.2.2 Subset 2

Subset 2 contains 11 events in a square of 10 X 10 km.

The minimized order including S-P time series is: 26,22,19,18,23,21,17,27,29,7,15.

The numbers correspond with the numbers in the figures and tables. Figure 6.6 and 6.7 contain the focal mechanisms and waveforms.

entry												t(s)-t(p)	number of event
1	1.00	0.34	0.33	0.19	0.37	0.29	0.32	0.41	0.29	0.30	0.16	4.7	15
2	0.34	1.00	0.73	0.32	0.42	0.47	0.35	0.38	0.34	0.41	0.33	5.1	7
3	0.33	0.73	1.00	0.40	0.51	0.32	0.35	0.42	0.35	0.45	0.33	4.4	29
4	0.19	0.32	0.40	1.00	0.56	0.35	0.42	0.51	0.72	0.58	0.46	4.5	27
5	0.37	0.42	0.51	0.56	1.00	0.39	0.46	0.57	0.60	0.53	0.35	4.5	17
6	0.29	0.47	0.32	0.35	0.39	1.00	0.42	0.59	0.51	0.31	0.33	4.7	21
7	0.32	0.34	0.35	0.42	0.46	0.42	1.00	0.69	0.53	0.32	0.33	4.5	23
8	0.41	0.38	0.42	0.51	0.57	0.59	0.69	1.00	0.82	0.40	0.46	4.6	18
9	0.29	0.34	0.35	0.72	0.60	0.51	0.53	0.82	1.00	0.51	0.46	4.2	19
10	0.30	0.41	0.45	0.58	0.53	0.31	0.32	0.40	0.51	1.00	0.29	4.2	22
11	0.16	0.33	0.33	0.46	0.35	0.33	0.33	0.46	0.46	0.29	1.00	4.1	26

closeness number = 0.51

Table 6.4: *Final average matrix for minimized order, subset 2*

26
22
19
18
23
17
27
29
7
15

Figure 6.6:
Focal mechanisms, subset 5

By looking at the average cross-correlation of the three components (table 6.4), one can say that events #26 and #15 do not fit in the set, as they do not correlate well with any of the other events. Events #7 and #29 also do not correlate very well with the other events, so the spatial isolation as shown in figure 6.1 seems valid.

Northridge, subset 2: three components, vertical, tangent and radial, minimized order

VERTICAL

shallow

26
22
19
18
23
21
17
27
29
7
15

deep

TANGENT

shallow

26
22
19
18
23
21
17
27
29
7
15

deep

RADIAL

shallow

26
22
19
18
23
21
17
27
29
7
15

deep

counts

sec

0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32

54
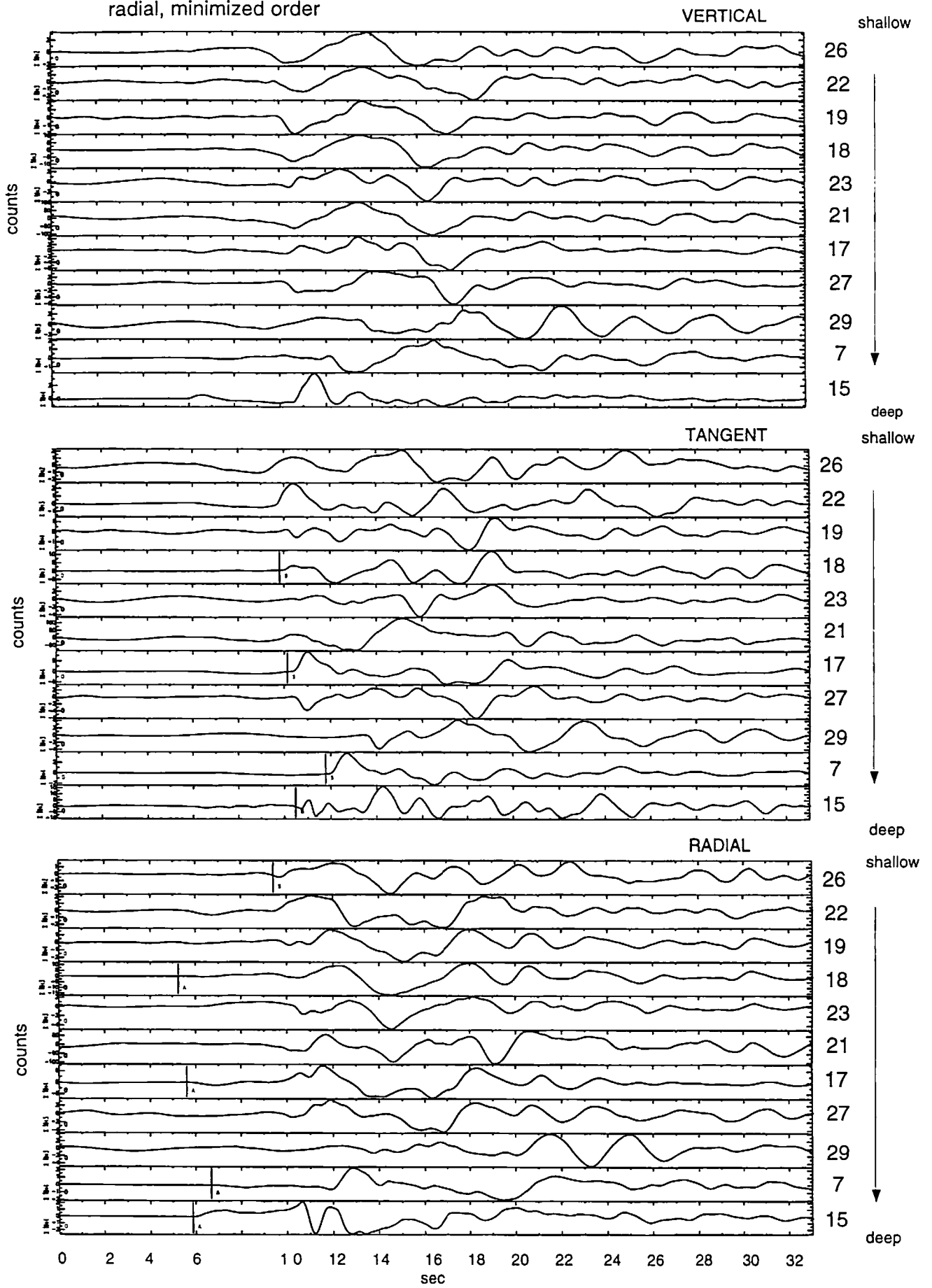
fig 6.7

6.4.2.3 Subset 3

Subset 3 contains the 11 events furthest away from PAS, in a depth range of 5 to 20 km. For correlation, the first 8 seconds after the origin time of each trace are left out.

The minimized order including S-P time series is: 20,14,10,16,33,32,8,6,11,31,2.

If the S-P time series is not used in the minimization, events #32 and #33 shift to after event #2. The numbers correspond with the numbers in the figures and tables. Figure 6.8 and 6.9 contain the focal mechanisms and waveforms.

entry												t(s)-t(p)	number of event
1	1.00	0.64	0.63	0.45	0.24	0.25	0.39	0.26	0.23	0.27	0.29	7.3	20
2	0.64	1.00	0.76	0.52	0.29	0.24	0.33	0.35	0.37	0.41	0.29	7.3	14
3	0.63	0.76	1.00	0.50	0.23	0.21	0.34	0.28	0.27	0.32	0.26	7.1	10
4	0.45	0.52	0.50	1.00	0.22	0.23	0.40	0.45	0.47	0.42	0.38	7.1	16
5	0.24	0.29	0.23	0.22	1.00	0.81	0.26	0.34	0.37	0.36	0.35	6.4	33
6	0.25	0.24	0.21	0.23	0.81	1.00	0.30	0.30	0.34	0.30	0.38	6.4	32
7	0.39	0.33	0.34	0.40	0.26	0.30	1.00	0.80	0.56	0.56	0.49	6.4	8
8	0.26	0.35	0.28	0.45	0.34	0.30	0.80	1.00	0.68	0.60	0.49	6.6	6
9	0.23	0.37	0.27	0.47	0.37	0.34	0.56	0.68	1.00	0.82	0.50	6.0	11
10	0.27	0.41	0.32	0.42	0.36	0.30	0.56	0.60	0.82	1.00	0.63	6.1	31
11	0.29	0.29	0.26	0.38	0.35	0.38	0.49	0.49	0.50	0.63	1.00	5.9	2

closeness number = 0.22

Table 6.5: *Final average matrix for minimized order, subset 3*

Focal mechanisms (subset 3): 20, 14, 10, 16, 33, 32, 8, 6, 11, 31, 2

Figure 6.8:
Focal mechanisms, subset 3

By looking at the average cross-correlation of the three components, one can see that this is a very similar set (see table 6.5). All the events have an average cross-correlation coefficient of 0.5 or higher with one or more other event(s). For a threshold of 0.6, only one event (#16) has no similarity with any of the other events. By looking at which trace correlates well with each other, one can divide this subset into two parts:

1. events 10, 20, 14, 16
2. events 8, 31, 11, 6, 32, 33, 2

This is basically a separation according to the distance away from PAS.

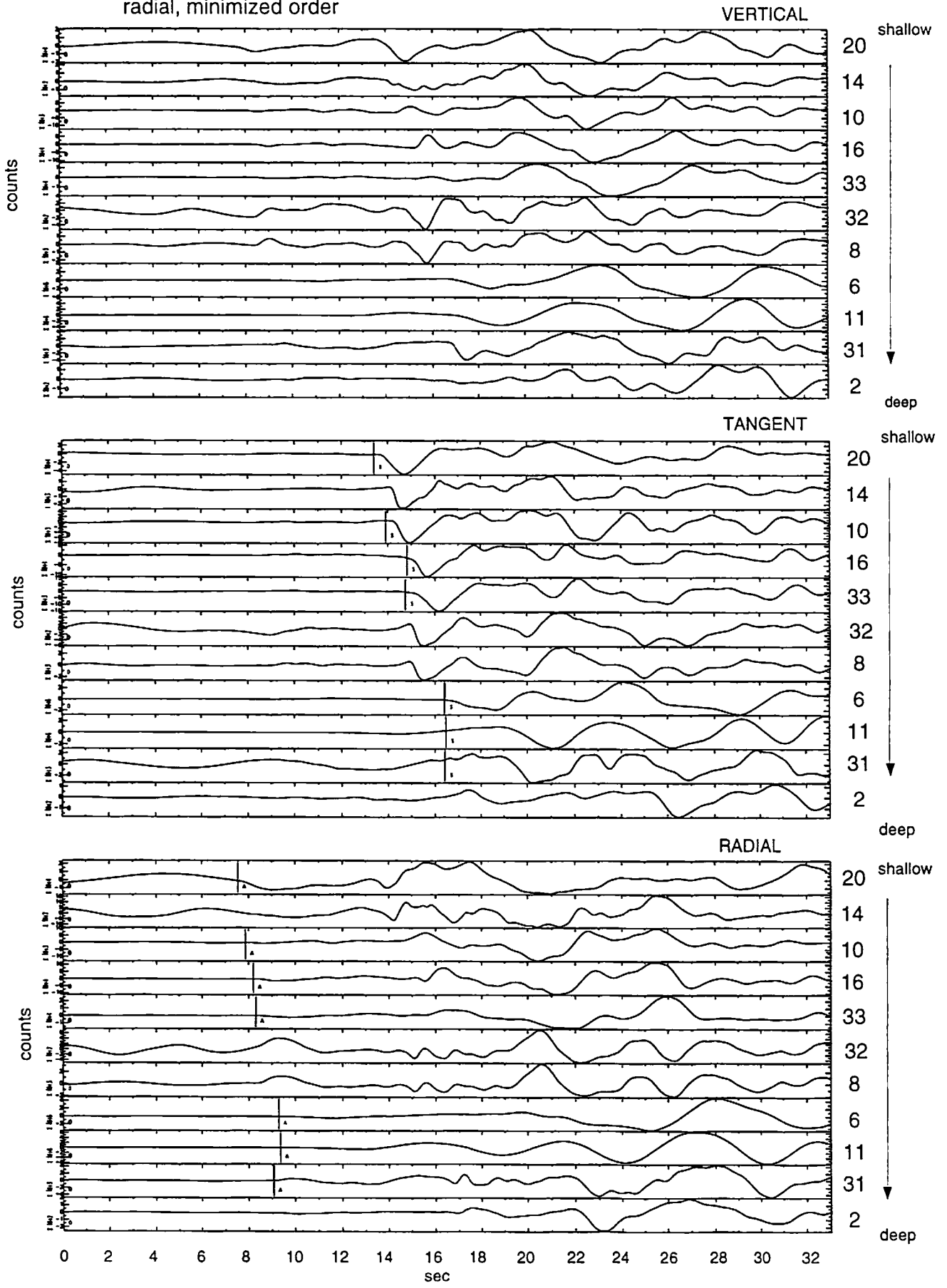# Northridge, subset 3: three components, vertical, tangent and radial, minimized order

VERTICAL

RADIAL

fig 6.9

6.4.2.4 Subset 4

Subset 4 contains 11 events with shorter distances, but changing depths. The minimized order is: 28,26,22,17,23,18,1,19,27,4,5 including the S-P time series without the S-P time series the minimized order is: 26,28,5,23,19,18,1,17,27,4,22.

The numbers correspond with the numbers in the figures and tables. Figure 6.10 and 6.11 contain the focal mechanisms and waveforms.

entry												t(s)-t(p)	number of event
1	1.00	0.56	0.58	0.39	0.33	0.40	0.44	0.43	0.32	0.31	0.32	4.6	5
2	0.56	1.00	0.90	0.61	0.35	0.43	0.41	0.44	0.48	0.46	0.47	4.5	4
3	0.58	0.90	1.00	0.70	0.43	0.47	0.39	0.53	0.55	0.44	0.42	4.5	27
4	0.39	0.61	0.70	1.00	0.70	0.82	0.51	0.60	0.49	0.42	0.33	4.3	19
5	0.33	0.35	0.43	0.70	1.00	0.72	0.54	0.59	0.36	0.49	0.33	4.4	1
6	0.40	0.43	0.47	0.82	0.72	1.00	0.70	0.56	0.39	0.42	0.35	4.6	18
7	0.44	0.41	0.39	0.51	0.54	0.70	1.00	0.45	0.30	0.30	0.39	4.6	23
8	0.43	0.44	0.53	0.60	0.59	0.56	0.45	1.00	0.51	0.32	0.36	4.6	17
9	0.32	0.48	0.55	0.49	0.36	0.39	0.30	0.51	1.00	0.27	0.33	4.1	22
10	0.31	0.46	0.44	0.42	0.49	0.42	0.30	0.32	0.27	1.00	0.39	4.1	26
11	0.32	0.47	0.42	0.33	0.33	0.35	0.39	0.36	0.33	0.39	1.00	4.0	28

closeness number = 0.22

Table 6.6: *Final average matrix for minimized order, subset 4*

Figure 6.10:
Focal mechanisms, subset 4

By looking at the average cross-correlation coefficient, one can see that events #4, #27, #19, #1, #18, #23 and #17 correlate fairly well with one another (see table 6.6), that event #5 correlates slightly (~0.57) with event #27 and #4, and that events #22, #26 and #28 correlate considerably less well with all of the other events. For events #4 and #1, the depths as given by the used data-sets are questionable, because of the high correlations with #18, #19 and #27.

For the events that are high in cross-correlation and closer together in the data-sets used, one can say little about the actual lining up of the events. They are all grouped in a depth section of less than 5 by 5 km with events #22, #26 and #28 outside of the cluster.

The focal mechanisms are fairly coherent throughout the whole set, even for the events that have less good cross-correlations. The results indicate that the algorithm puts together the events that are similar and puts aside the events with fewer similarities, but not in any specific order. To get the right positions something else will be needed.

Northridge, subset 4: three components, vertical, tangent and radial, minimized order
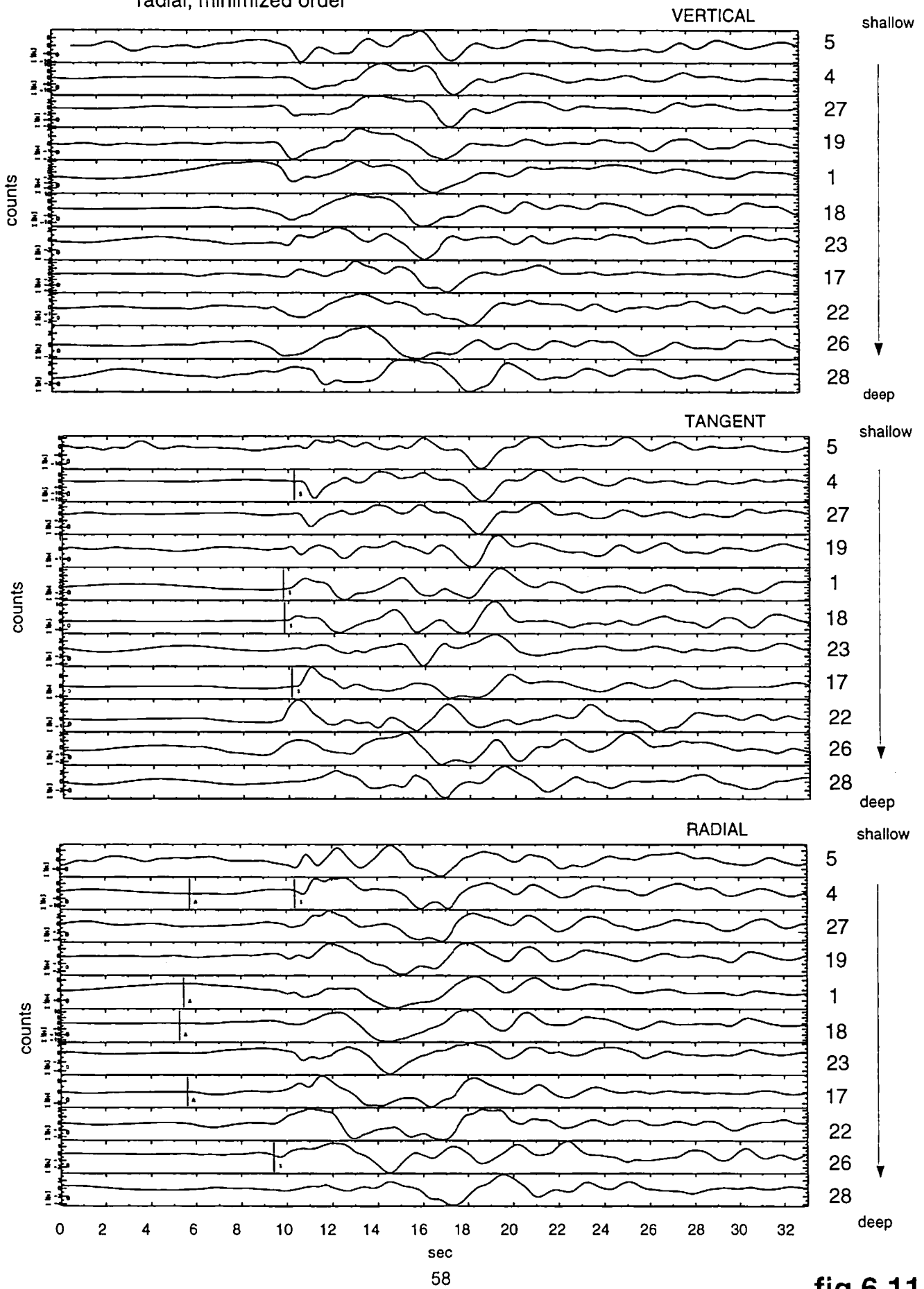
fig 6.11

6.5 Conclusions Northridge

The Northridge dataset has a different order of magnitude compared to the dataset from Long Valley. The distances between the events and to the station are larger and the magnitudes of the events are higher in the Northridge dataset. Inherently the frequency of the recorded traces is therefore dominated by lower frequencies. Looking at the recordings shows directly that the traces are much less coherent and by consequence the method seems less promissing. This does not mean that the method cannot be used for a dataset with a larger order of magnitude than the one from Long Valley. By enlarging the timewindow considered for cross-correlation the order of magnitude of the calculations is also altered and the result might be better. For this dataset the timewindow used is limited by the recordings and it is not possible to compare traces of longer than 33 seconds.

7 Discussion

This research does not end with a relocated set of events by use of cross-correlation coefficients, but it does show that there is a future to using cross-correlations in an algorithm that determines focal mechanism and location parameters. In this stage of development the biggest problem are the recordings. More events are needed to give a gradual change between characteristic clusters and when the order of magnitude of distances and magnitudes of events increases also the timewindow used for cross-correlation has to be enlarged.

Because a full relocation is not possible with the data available an attempt has been made to give comments on the given catalog locations and focal mechanisms by comparing them with the calculated maximum cross-correlation values between the different events. Sometimes the cross-correlation really gives good reasons to believe that the focal mechanisms should be more similar than given in the catalog. In order to make definitive remarks about it, one should consider more than only one cross-correlation coefficient. By taking only cross-correlations from waveforms recorded at one single station, you are basically just looking at one part of the focal mechanism 'beach-ball'. More constraints can be put on the conclusions if the same calculations are done for more stations at different azimuths.

By comparing the cross-correlations values one can see if the events are really similar in location and/or focal mechanism. Before doing this, one can already look at the focal mechanism and find out how well constrained the focal mechanism is. If the focal mechanisms can be considered to be trustworthy, one can make comments on the location parameters. Unfortunately there was no time to look at the relocations made by Peggy Johnson for the Long Valley area to see if things would really change and what the implications for the results would have been.

The amplitude distribution in the radiation pattern of a focal mechanism shows a decrease in amplitude going towards the nodal planes (where the amplitude is zero). In the configuration where the recording station is close to the nodal plane for two events that have different focal mechanisms, it is very well possible that the events give a high cross-correlation coefficient, without the focal mechanisms being wrong. Because the amplitudes are low, the difference in signature will be small, and the correlation therefore can be high.

The algorithm developed here is far from ideal. In the future people can try to implement the method in current location inversions. One can think of a big inversion scheme for focal mechanism, location and waveform similarities at the same time. Therefore, it first has to be investigated more thoroughly how the waveforms (and cross-correlation coefficients) change with a certain change of location and focal mechanism parameters. Only then can we say how far apart focal mechanism and location parameters will be for such and such cross-correlation coefficient. One way to do this, besides synthetic datasets, is to use borehole seismic experiments. Locations are exactly known and the focal mechanism stays the same, so the change with location can be investigated carefully.

Also one can adapt the current algorithm such that more events can be used at a time and to enlarge the degrees of freedom allowed. To work in more dimensions than along a one-dimensional line can be achieved by putting up a frame (spatially) of characteristic events and adding events in their correct positions within the frame by considering the cross-correlation coefficients. An other idea is to line-up along more than one line. The lines will have to be picked in different directions in the multi-dimensional space. The method can be repeated until a stable configuration is reached.

8 Conclusions

In this research a first attempt is made to solve the jigsaw puzzle of gradually changing record sections in order to get better relative locations for local events. Cross-correlations between the different waveforms are used as the main parameter to solve the puzzle. It is considered that if the location and/or focal mechanism change slowly, the waveforms will show a gradual change as well. The algorithm developed uses a starting matrix with all the normalized cross-correlation coefficients and finds the matrix that fits best the 'ideal' situation of decreasing series away from the main diagonal by changing the order of events.

The used synthetic dataset has a constant focal mechanism for all the waveforms and therefore only considers a change in location parameters. From this experiment it can be said that if waveforms (events) indeed belong on one line, the order given by the algorithm will basically be an order with distance from the station. Even with a signal to noise ratio of 3 the algorithm gives a good result. When the waveforms are taken from a broader line of about 3 km wide (randomly chosen subset), the result is not a line-up with distance anymore. The global line-up is good, but some neighboring waveforms have changed places. More experiments should be done to see what determines the linewidth of hypocenters for which the method works.

The Long Valley data set recorded at MLAC contains sets of waveforms that are very coherent and similar, but from one coherent set to the other there exists no gradual transition yet. This is a result of not having enough events. Partly to blame is the fact that 38% of the 105 events looked at could not be used because of bad recordings or because just two components were available in the catalog.

The dataset for the Northridge aftershock sequence is completely different from the one from Long Valley. The spacing is much larger, and the magnitudes of the events are bigger. This results in different waveform signatures, including more surface waves and lower frequencies. For some shallow events, the cross-correlation coefficients indicated that the vertical location was wrong. The cross-correlations with several deeper events was high. L. Zhu's catalog of focal mechanisms and vertical locations has given more consistent results. L. Zhu uses a full waveform inversion to calculate his source mechanisms and depths. J. Mori and E. Hauksson, on the other hand used first motion traveltime inversions to determine their parameters. As the cross-correlation technique used in the developed algorithm is a measure of similarity of a larger time window of the waveform, it is not a surprise that L. Zhu's catalog does give more consistent results. It should be possible to implement some sort of a cross-correlation coefficient weigh-factor in the first motion algorithm to incorporate whole waveform information in the inversion for the focal mechanism.

The dataset for the Northridge aftershock sequence is completely different from the one from Long Valley. The spacing is much larger, and the magnitudes of the events are bigger. This results in different waveform signatures, including more surface waves and lower frequencies. For some shallow events, the cross-correlation coefficients indicated that the vertical location was wrong. The cross-correlations with several deeper events was high. L. Zhu's catalog of focal mechanisms and vertical locations has given more consistent results. L. Zhu uses a full waveform inversion to calculate his source mechanisms and depths. J. Mori and E. Hauksson, on the other hand used first motion traveltime inversions to determine their parameters. As the cross-correlation technique used in the developed algorithm is a measure of similarity of a larger time window of the waveform, it is not a surprise that L. Zhu's catalog does give more consistent results. It should be possible to implement some sort of a cross-correlation coefficient weigh-factor in the first motion algorithm to incorporate whole waveform information in the inversion for the focal mechanism.

7 Discussion

This research does not end with a relocated set of events by use of cross-correlation coefficients, but it does show that there is a future to using cross-correlations in an algorithm that determines focal mechanism and location parameters. In this stage of development the biggest problem are the recordings. More events are needed to give a gradual change between characteristic clusters and when the order of magnitude of distances and magnitudes of events increases also the timewindow used for cross-correlation has to be enlarged.

Because a full relocation is not possible with the data available an attempt has been made to give comments on the given catalog locations and focal mechanisms by comparing them with the calculated maximum cross-correlation values between the different events. Sometimes the cross-correlation really gives good reasons to believe that the focal mechanisms should be more similar than given in the catalog. In order to make definitive remarks about it, one should consider more than only one cross-correlation coefficient. By taking only cross-correlations from waveforms recorded at one single station, you are basically just looking at one part of the focal mechanism 'beach-ball'. More constraints can be put on the conclusions if the same calculations are done for more stations at different azimuths.

By comparing the cross-correlations values one can see if the events are really similar in location and/or focal mechanism. Before doing this, one can already look at the focal mechanism and find out how well constrained the focal mechanism is. If the focal mechanisms can be considered to be trustworthy, one can make comments on the location parameters. Unfortunately there was no time to look at the relocations made by Peggy Johnson for the Long Valley area to see if things would really change and what the implications for the results would have been.

The amplitude distribution in the radiation pattern of a focal mechanism shows a decrease in amplitude going towards the nodal planes (where the amplitude is zero). In the configuration where the recording station is close to the nodal plane for two events that have different focal mechanisms, it is very well possible that the events give a high cross-correlation coefficient, without the focal mechanisms being wrong. Because the amplitudes are low, the difference in signature will be small, and the correlation therefore can be high.

9 References

P. Augliera, M. Cattaneo, C. Eva (1992), *Seismic multiplets analysis and its implication in seismotectonics*, European Seismological Commission, 23rd general assembly, symposium on study of seismic source, theory and observations; and symposium on seismotectonic analysis, applications to Europe, Elsevier, Amsterdam.

D. Dodge, G. Beroza (1995), *Foreshock sequence of the 1992 Landers, California, earthquake and its implications for earthquake nucleation*, JGR, v100, no B7, p9865-9880

D. Dodge, G. Beroza (1996), *Detailed observations of California foreshock sequences: implications for the earthquake initiation process*, JGR, v101. no B10, p22,371-22,392

D. Eberthart-Phillips (1993), *Seismic tomography: Theory and practice, chapter 2: Local earthquake tomography: earthquake source regions*, edited by H. Iyer and K. Hirahara, Chapman & Hall 1993, London, p 613-643

M-J. Frémont, S. Malone (1987), *High precision relative locations of earthquakes at Mount St. Helens, Washington*, JGR, v92, no B10, p10,223-10,236

A. Ito (1985), *High resolution relative hypocenters of similar earthquakes by cross-spectral analysis method*, J.Phys.Earth, v33, p279-294

T. Lay, T.Wallace (1995), *Modern global Seismology*, Academic Press.

J. Mezcua, J. Rueda (1994), *Earthquake relative location based on waveform similarity*, Tectonophysics, v233, p253-263

C. Rowe, R. Aster, M. Fehler, W. Phillips, R. Kyle (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, AGU fall meeting 1998

SAC, *New Seismic Analysis Code*, [may 30,1994 (version 00.57)], Joseph E. Tull, Lawrence Livermore National Laboratory, Copyright 1995 Regents of the University of California.

P. Shearer (1997), *Improving local earthquake locations using the L1 norm and waveform cross-correlation: Application to the Whittier Narrows, California, aftershock sequence*, JGR, v102, no B4, p8269-8283

P. Shearer (1998), *Evidence from a cluster of small earthquakes for a fault at 18 km depth beneath Oak Ridge, Southern California*, BSSA, v 88, No 6, p1327-1336

M. Spieth, R. Geller (1981), *Precise relative locations of local earthquakes near San Juan Bautista, California*, EOS, v62, no 45.

G. Vsevolozhsky, T. Wallace (1993), *Souce location form a very broad band single station*, EOS, AGU fall meeting 1993

D. Wald, T. Heaton, K. Hudnut (1996),*The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS and Leveling data*, BSSA, v86, S49-S70.

L.Zhu, D. Helmberger (1996*), Advancement in source estimation techniques using broadband regional seismograms*, BSSA, v86, no5, p1634-1641

10 Appendix I

While working on data from the Long Valley area, recorded by MLAC, I was really shocked by the amount of data that could not be used because recordings were bad. It is a shame that two components come in perfectly well, but the third component cannot be used. This doesn't mean that the station didn't record but that either the one component is malfunctioning, or that the retrieval of the data and the further processing is not correct. I am not aware of what is and what isn't possible, but if the one component is malfunctioning it seems to me that it can be repaired. If the retrieval of the data is not working the way it should, resulting in data of at least one component that cannot be used, and people know it, it should be able to solve this problem. Therefore this appendix. A list of events that I encountered that are not well recorded, and couldn't be used by me. Hopefully somebody can do something with it...

The next section first shows all the eventid's, as used by the SCEC-Data Center. After that, a few waveforms are shown, with characteristic recording errors.

SUBSET 1:	SUBSET 2:	SUBSET 3:	SUBSET 4:	SUBSET 5:
9032441	9038239	3260069	3260163	3132722
9033195	3260174	3260134	3260191	3137560
9034591	3260263	3260187	3260311	3297728
9032268	3260529	3260569	3260432	3260117
9033179	3260522	3260762	3260095	3259875
	3260159	3261110	3260314	3259880
	3260269		3260392	3259866
	3260070		3260393	9037180
	3260139			9038670
	3260182			
	3260244			

66

E

date : NOV 23 (327), 1997
mag : 2.87
evid : 9.032441e+06

N

NOV 23 (327), 1997
2.87
9.032441e+06

Z

NOV 23 (327), 1997
2.87
9.032441e+06

date : NOV 30 (334), 1997
mag : 5.16
evid : 9.033195e+06

E

NOV 30 (334), 1997
5.16
9.033195e+06

N

NOV 30 (334), 1997
5.16
9.033195e+06

Z

E

date: DEC 11 (345), 1997
mag: 3.71
evid: 9.034591e+06

N

DEC 11 (345), 1997
3.71
9.034591e+06

Z

DEC 11 (345), 1997
3.71
9.034591e+06

date: NOV 22 (326), 1997
mag: 3.05
evid: 9.032268e+06

NOV 22 (326), 1997
3.05
9.032268e+06

NOV 22 (326), 1997
3.05
9.032268e+06

date : JAN 03 (003), 1998
mag : 2.85
evid : 9.038239e+06

E

X 10+5

JAN 03 (003), 1998
2.85
9.038239e+06

N

X 10+5

JAN 03 (003), 1998
2.85
9.038239e+06

Z

X 10-7

date : MAR 31 (091), 1996
mag : 3.20
evid : 3.260174e+06

MAR 31 (091), 1996
3.20
3.260174e+06

date: NOV 20 (324), 1993
mag; 2.46
evid: 3.132722e+06

NOV 20 (324), 1993
2.46
3.132722e+06

NOV 20 (324), 1993
2.46
3.132722e+06

11 Acknowledgements:

Furthermore I would like to thank the 'Molengraaf-fonds', the Free Movers fund from Utrecht University, Amoco and the 'Studiefonds' for their financial support that made my stay in the US possible.

Last but not least I would like to thank my parents for always having stimulated me in exploring the world around me and to see that there is more than the beautiful lowlands of Holland.