5 Application to data from the San Andreas Fault Observatory at Depth

5.1   The San Andreas Fault Observatory at Depth

The San Andreas Fault system defines the boundary between the Pacific and the North American Plate. The Pacific Plate moves northwestward relative to the North American Plate which causes right-lateral strike-slip fault displacement (see Figure 5.1). Different sections of the San Andreas Fault system show different behaviors. Some segments of the fault are locked over several decades while building up strain which is released during great earthquakes (e.g., the great Fort Tejon earthquake in 1857 and the great San Francisco earthquake in 1906). These sections are marked as red lines in Figure 5.1. On the other hand, some segments of the fault are constantly creeping (aseismic slip) with about 20 mm/year (the yellow marked section in Figure 5.1). The constant creep does still generate earthquakes but the magnitudes of these earthquakes (M<6) are smaller than those of earthquakes originated in the fault segments that have been locked for decades. The San Andreas Fault Observatory at Depth (SAFOD) is located near Parkfield, California, and lies in the transition zone between a creeping section and a section that was ruptured a few times during a series of moderate-size earthquakes. The San Andreas Fault at Parkfield is of special scientific importance for several reasons.

Firstly, the moderate-size earthquakes mentioned above occur fairly regularly. The documented series starts in 1857 with two foreshocks (Meltzner and Wald, 1999; Sieh, 1978a) to the great Fort Tejon earthquake that ruptured the fault from Parkfield to the southeast for about 360 km (Sieh, 1978b). The earthquake sequences continued with events in 1881, 1901, 1922, 1934, 1966 (Bakun and
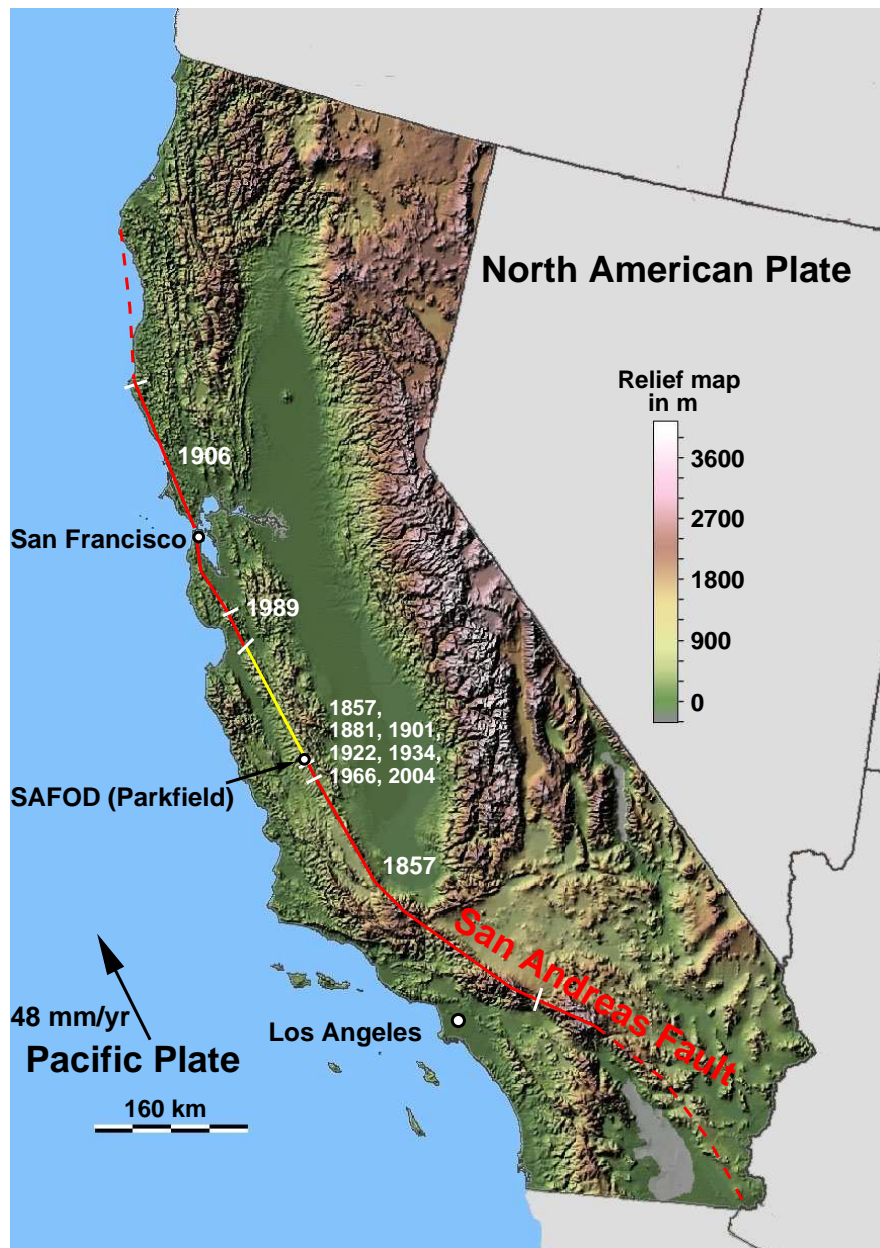
Figure 5.1: Modified Californian relief map (the original relief map is provided by the U.S. Geological Survey and can be downloaded from the following web page http://education.usgs.gov/california/maps/shaded2.htm). Ruptured segments of the San Andreas Fault are marked as solid red lines, whereas the creeping segment of the fault is indicated by a yellow solid line.

McEvilly, 1984), and with the most recent one in 2004 (Harris and Arrowsmith, 2006). In addition to the statistically significant regularity, Bakun and McEvilly (1984) found a high correlation in waveforms of the earthquakes recorded with regional seismographs in 1922, 1934, and 1966. This analysis could not be performed for the events that occurred before 1922 because of the lack of instrumental records. However, the waveform comparison of the 1922, 1934 and 1966 events with the 2004 earthquake also showed a large correlation (Bakun et al., 2005). This in turn suggests that these earthquakes have similar focal mechanisms and seismic moments as well as have ruptured approximately the same area. At the same time, differences between these events were observed, too. The hypocenter locations and the propagation mechanisms differed. The 1934 and 1966 events were initiated at the northwestern end of the ruptured segment, and propagated towards the southeast, while the 2004 event was initiated at the southeastern end, and propagated towards the northwest (Bakun et al., 2005). In addition, the 1934 and 1966 Parkfield earthquakes showed highly correlated foreshock sequences whereas the 1922 and 2004 earthquakes did not show any short-time precursory signals, neither seismic nor aseismic.

In any event, the apparent predictability of these earthquakes allows to establish a multi-stage geophysical / geochemical observatory in the vicinity of these repeating earthquakes. Such an *in situ* earthquake laboratory opens the unique possibility to understand the physical and chemical conditions before, during and after those moderate-magnitude earthquakes in order to explain the similarities and differences of those events.

In addition, numerous small earthquakes (M_w= 0 to 5) can be observed along the San Andreas Fault near Parkfield at depths of 2 km to 12 km. These small magnitude earthquakes also repeat on a regular schedule and occur in distinct spatial clusters (Nadeau et al., 1994). Some events within a cluster show nearly identical waveforms (Bakun et al., 2005) with almost as high correlation as the moderate-magnitude earthquakes. Inspired by these clustered microearthquakes one of the SAFOD goals was to drill directly into the hypocentral region of a microearthquake cluster and to measure chemical and physical properties for a better understanding of the mechanisms of these repeating microearthquakes.

In the summer of 2002 a 2.2 km deep vertical Pilot Hole was drilled adjacent to the San Andreas Fault about 1.8 km southwest of the San Andreas Fault surface trace. The observations obtained in the Pilot Hole as well as surface measurements have been used to locate the drilling target for a deep deviated well that intersects the fault within the hypocentral region of repeating M~2 earthquakes (Hickman et al., 2004). In the following these microearthquakes will be called target events. In 2004 and 2005 the Main Hole was drilled vertically to a depth of approximately 950 m below sea level (1600 m measured depth) and then deviated through the

fault zone at a approximately 50 to 60 degree inclination from the vertical to a final depth of 2400 m below sea level.

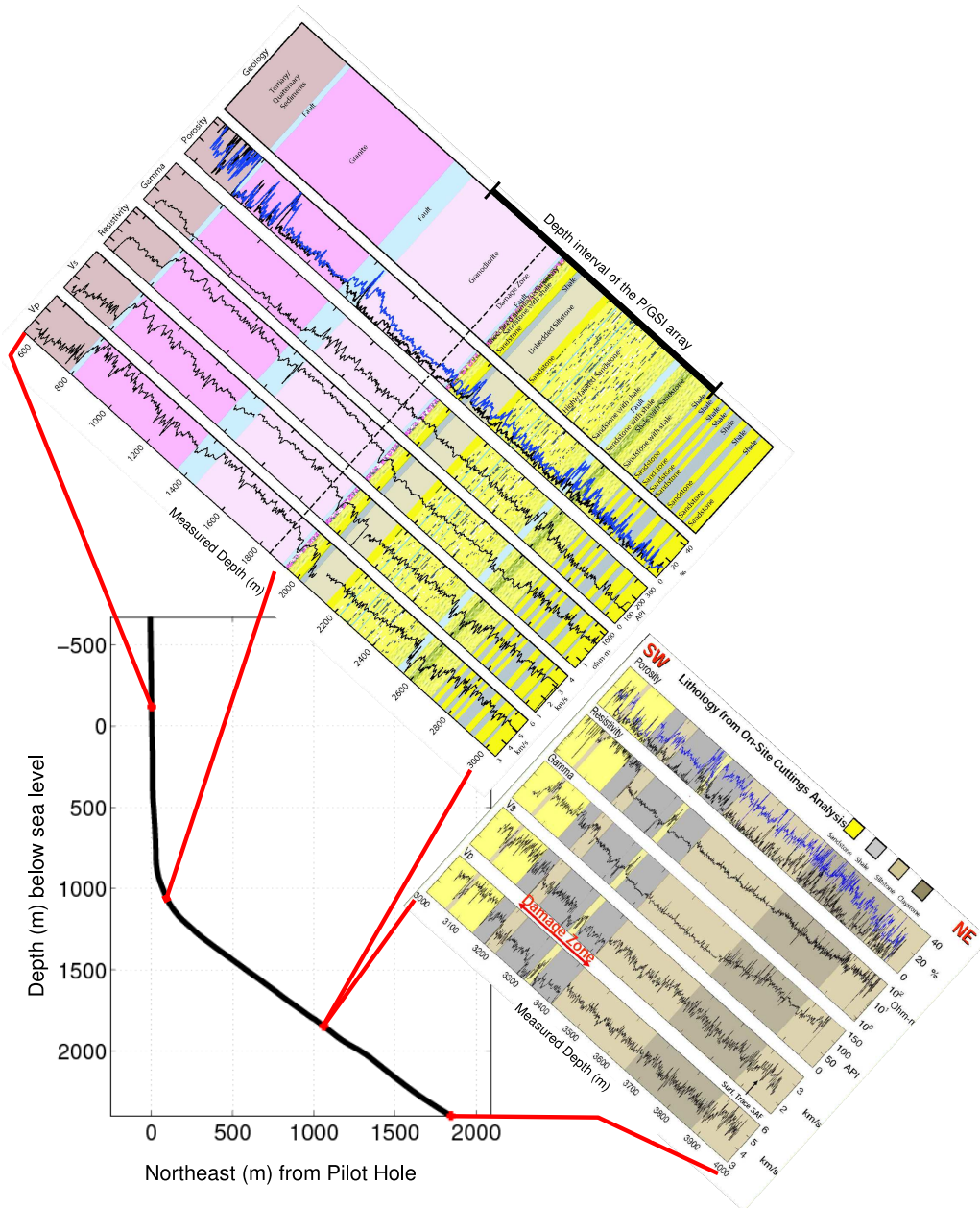## 5.2   Geological settings at the SAFOD site

Figure 5.2: Interpreted Main Hole logs from Boness and Zoback (2006) (550 m - 3000 m measured depth) and Zoback et al. (2007) (3000 m - 4000 m measured depth).

At the SAFOD drill site (which is located about 1800 m southwest from the San Andreas Fault surface trace) tertiary and quaternary sediments were encountered

by core samples for the first depth interval of about 800 m measured from the surface. For these sediments the well logs showed increasing seismic velocities with increasing depth, relatively high but altering porosity values as well as comparable constant resistivity values of about 10 Ωm and gamma log counts. The first lithological contrast can be observed at about 130 m (bsl) where the seismic velocities as well as the resistivity increased significantly and gamma log counts as well as porosity decreased. From 130 m to 680 m (bsl) granitic rocks were found in the cutting analysis of the core samples. In the first 200 m of this interval all geophysical measurements (apart from the gamma log) showed altering values which stabilized deeper in the granitic interval. However, Boness and Zoback (2006) associate the altering of the porosity log with borehole enlargement (this section of the whole was washed out). A major shear zone was identified in the Main Hole at about 700 m (bls). Below this shear zone granodiorite was found in the core samples and also the well logs indicated a different geological unit (e.g., the gamma log increased from 50 API to 100 API). In the lower part of the granodiorite interval seismic velocities as well as resistivity values decrease until the next lithological unit - a sequence of sedimentary rocks - was found at about 1200 m (bsl). The encountered sediments consisted mainly of sandstone and siltstone with small intervals of shale. At about 1950 m (bsl) the Main Hole penetrated a 250 m interval that consisted mainly of shale. However, the geophysical measurements obtained in this interval indicated a highly damaged zone (Zoback et al., 2007). Moreover, at about 3200 m measured depth the casing already started to deform indicating a creeping section of the San Andreas Fault. Below and northeast of the identified damage zone core samples consisted of siltstone and claystone. Further major shear zones were not encountered and the well logs showed a consistent increase in velocities and resistivity with increasing depth. The Main Hole ends at about 2400 m (bsl) which corresponds to 4000 m measured depth. The vertical projection of the San Andreas Fault surface trace intersects the Main Hole at 2370 m (bls or 3950 m measured depth). Consistent log values that could be associated with a shear zone were not measured at this depth. It is more likely that the San Andreas surface trace is connected with the damage zone encountered between 1950 and 2100 m (bsl) which rather reflects a steeply dipping than vertical fault zone. This interpretation is also supported by results from active seismic imaging presented by Buske et al. (2006) which show a bunch of steeply dipping reflectors in the vicinity of the San Andreas Fault surface trace as well as in the vicinity of the Buzzard-Canyon-Fault surface trace. Furthermore these reflectors converge at a depth of about 4km forming a flower structure.

5.3 SAFOD specific implementation of the location procedure

The SAFOD data analyzed in this thesis were recorded with an array deployed in the SAFOD Main Hole (details are given below). As described above the SAFOD Main Hole was drilled vertically to a depth of about 1500 m (true measured depth), at which point the borehole deviates 55 degrees towards the San Andreas Fault. Moreover, the deviated borehole also points towards the hypocentral region of target events. An array deployed in the deviating part of the borehole represents an unfavorable acquisition geometry for the presented location method in the case the event is a target event (see Figure 5.3) because the estimated P-wave polarization vectors of target events will have dips and azimuths which are similar to the borehole trajectory. Hence, the stacking of beams which are almost parallel to each other does not result in a distinct energy maximum at the hypocenter location, but it rather results in a region of maximum energy that has the shape of a beam (see Figure 5.3). To overcome this problem the use additional information like P- and S-wave arrival times is required to restrict the Gaussian beams. In detail, the width of the Gaussian beam b is set to zero for all ray segments where the difference between P- and S-wave arrivals do not approximately match the observed one. This means that no energy is propagated from this part of the ray. In turn, the Gaussian beam width is set equal to the size of the Fresnel zone for each ray segment where the calculated and observed arrival time differences matched (see Figure 5.3, green box).
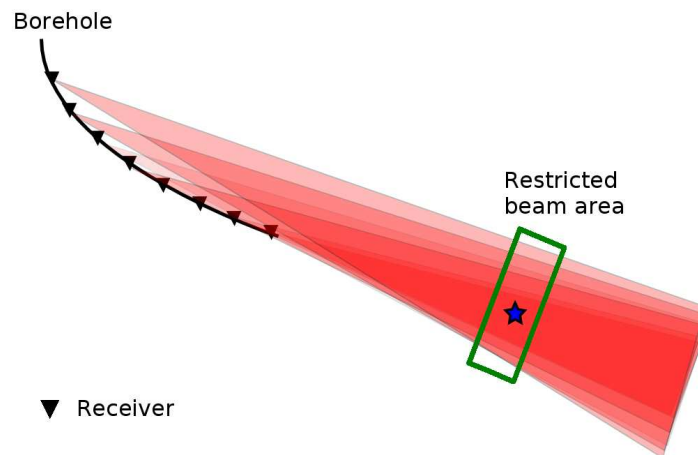

Figure 5.3: SAFOD receiver geometry and its pitfall for the presented location procedure. A target event originates somewhere along the elongated borehole trajectory. In order to locate this event we need to restrict the beams using approximate travel times (green rectangle).

The calculated travel times of P- and S-waves, and hence the arrival time differ-

ence ΔSP can be obtained in several ways. Since ray tracing at constant time steps dt_{ray} is performed through a given P-wave velocity model the P-wave travel time for each ray segment can be calculated directly with:

$$t_p(n_{ray}) = dt_{ray} \cdot n_{ray}, \tag{5.1}$$

were n_{ray} is the index of the ray segment. For a given Vp-Vs-ratio the P-wave travel time and the S-wave travel time are related in the following way:

$$t_s = t_p \cdot \frac{V_p}{V_s}. \tag{5.2}$$

Assuming a constant Vp-Vs-ratio the arrival time difference between the S- and P-wave can be obtained by:

$$\begin{aligned} \Delta SP(n_{ray}) &= t_s(n_{ray}) - t_p(n_{ray}) \\ &= t_p(n_{ray}) \cdot \left(\frac{V_p}{V_s} - 1 \right). \end{aligned} \tag{5.3}$$

Hence, the restriction of the Gaussian beam to the observed arrival time difference can be summarized as follows:

Flowchart for beam restriction

Loop over receivers

- obtain arrival time difference of P- and S-wave (automated picking; handmade picking; cross correlation)

- raytracing using bidirectional P-wave polarization and P-wave velocity model

- calculate P-wave travel time for each ray segment with equation (5.1)

- calculate arrival time difference between S- and P-wave for each ray segment using equation (5.3)

- find ray segment were the calculated arrival time difference matches the observed one

- set the width of the Gaussian beam equal to the Fresnel zone at this ray segment

End loop over receivers

In order to account for the uncertainty in the P- and S-wave detection a deviation of ± 5 ms between the calculated and observed arrival time differences was allowed. With this restriction of the Gaussian beams a summation of the back-propagated energy over all receivers yields regions of distinct stacked energy and the maximum is interpreted as the hypocenter of the target event.

5.4 Data set

From April 28th, 2005 to May 11th, 2005, Paulsson Geophysical Services Inc. (P/GSI) installed an 80 level array of 3C 15Hz seismometers in the SAFOD Main Hole to monitor the seismicity of the active fault. During this period the array recorded numerous events including one target event, tremors and surface explosions. The array was deployed along the deviated portion of the well using production tubing at depths between 878 m and 1703 m below sea level with a receiver spacing of about 15.24 m (see Figure 5.4, left). The recorded signals were sampled at 0.25 ms. Basic preprocessing of the data was done by P/GSI and consisted of geophone orientation using particle motions of twelve far-offset shots. The distribution of the far-offset shots is shown in Figure 5.4 (right). Ten of the twelve far-offset shot locations were performed at station locations of the Parkfield Area Seismic Observatory (PASO) which also has been used for the velocity model inversion by Thurber et al. (2004).
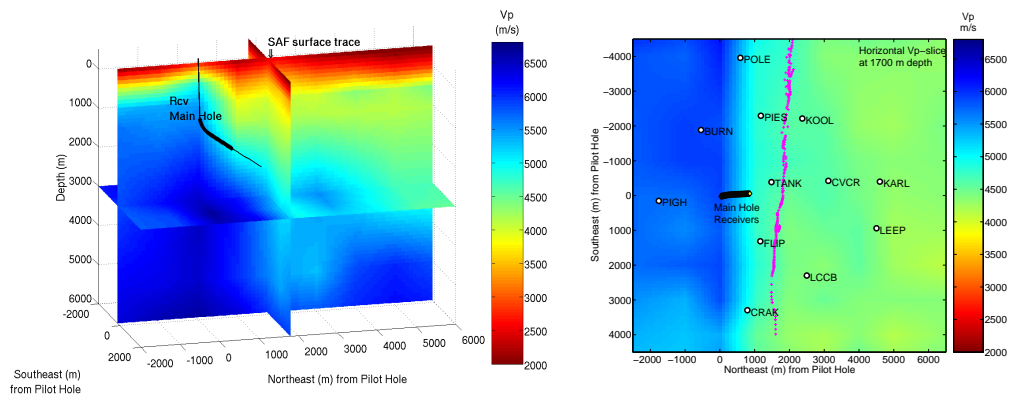
Figure 5.4: Geometry of the receivers deployed in the Main Hole plotted on top of the velocity model provided by Thurber et al. (2004). Left: Three slices through the velocity model. Right: Horizontal slice at 1700 m below sea level (depth of the deepest receiver). The magenta crosses indicate the SAF surface trace as mapped by Michael Rymer (U.S. Geological Survey). The circles mark the locations of the far-offset shots that have been used to estimate the receiver orientation.

The raw data were stored as 16 s long data intervals in *seg2* format. A reading subroutine for *seg2* was not implemented in the location algorithm. For this reason the raw data were converted into *seismic unix* format as described in Appendix C.1. During visual inspection of some raw data files some electronic noise was observed that would have triggered the detection algorithm as well as influenced the polarization information. An analysis of this electric noise as well as some details about its removal can also be found in Appendix C.2. Using the obtained receiver orientations the data were rotated into vertical-, East- and North-components. Details for this rotation (equations and rotation angles) are given in Appendix C.3. The rotated seismograms of one event are shown in Figure 5.5 (it is the target event of May 5, 2005). Both, the P- and the S-wave are clearly visible on almost all receivers.

5.5 Processing and Results

5.5.1 Event detection and polarization analysis

The first step in the location procedure is the event detection. Therefore the single receiver event detection algorithm described in section 3.1 was applied.

As a representative example the detection results obtained from the data shown in Figure 5.5 at receiver number 55 are shown in Figure 5.6. The three-component seismograms of the event are shown in (a). The corresponding short-term-average (STA) long-term-average (LTA) ratio versus time is shown in (b). The lengths
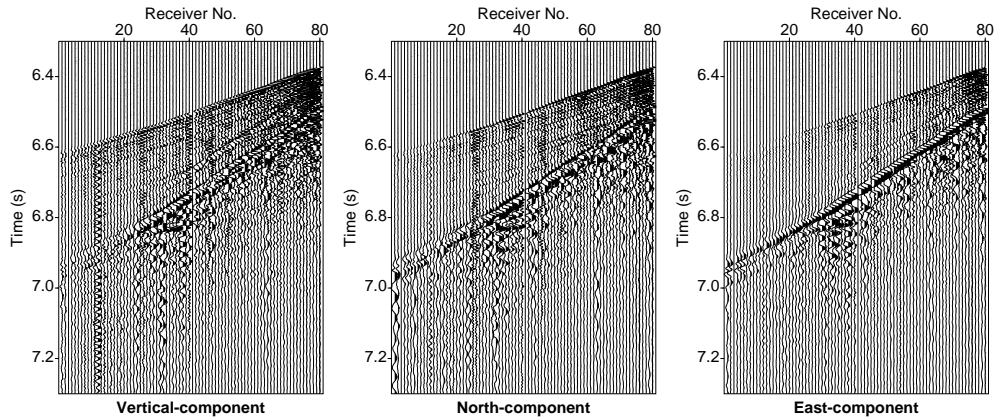
Figure 5.5: Waveforms of the target event of May 5, 2005. From left: vertical-, East- and North-component. Receiver No. 1 corresponds to the shallowest receiver and No. 80 to the deepest one, respectively.
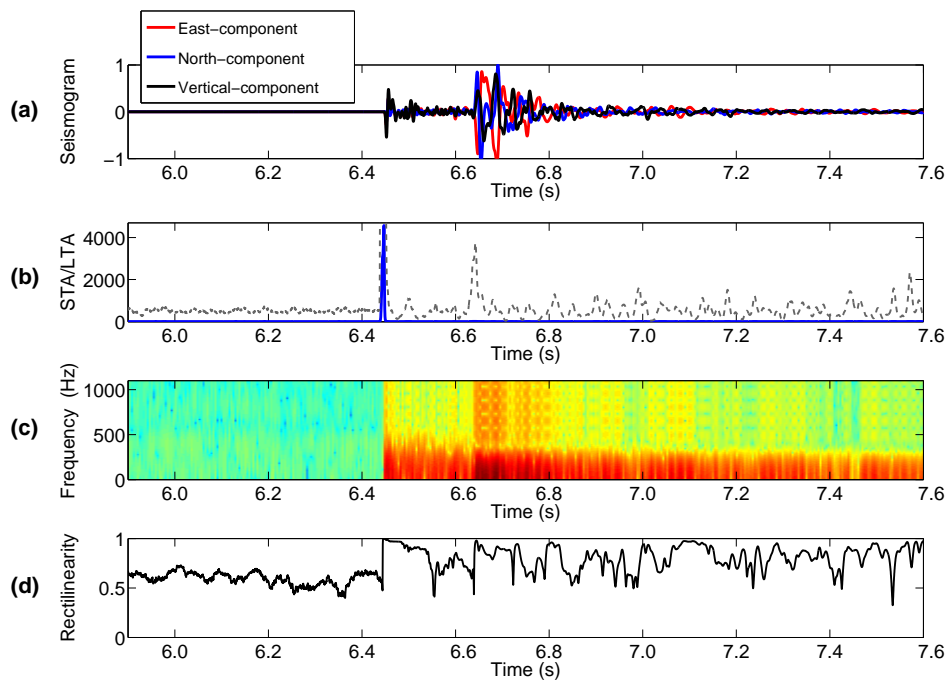

Figure 5.6: Detection of P- and S-waves. (a): 3C traces at receiver number 55 of the event in Figure 5.5. (b): STA/LTA ratio versus time of the data shown above. The gray dashed line is a 500-times increased plot of the STA/LTA ratio shown in blue. (c): Spectrogram of the data shown at the top. The intensity in dB is color-coded and increases from green colors to red colors. (d): Rectilinearity versus time obtained by equation (2.21).

of the STA analysis window was chosen as 20 ms and the corresponding LTA window was about five times longer than the STA window. The blue solid line corresponds to the true STA/LTA ratio whereas the gray dashed line is a 500 times increased version of the true STA/LTA ratio. The blue line clearly has a maximum that corresponds to the P-wave arrival at about 6.45 seconds at the traces shown above. Looking at the zoomed STA/LTA ratio a second significant rise of the ratio occurs at about 6.63 seconds. This rise corresponds to the S-wave arrival. In Figure 5.6 (c) the results of the frequency-content versus time analysis are shown. The color-coded frequency intensity shows a sharp onset of higher intensities (yellow and red colors) over a broad frequency range when the P-wave arrives at about 6.45 seconds. Furthermore, a second less sharp but clearly visible onset can be seen at the S-wave arrival time. The rectilinearity values scatter around 0.55 for the first 500 ms of the shown data (see Figure 5.6 (d)) . This indicates that this time interval does mainly contain noise. A very sharp rectilinearity rise indicates the P-wave arrival. A second sharp increase occurs at about 6.63 seconds which corresponds to the S-wave arrival.
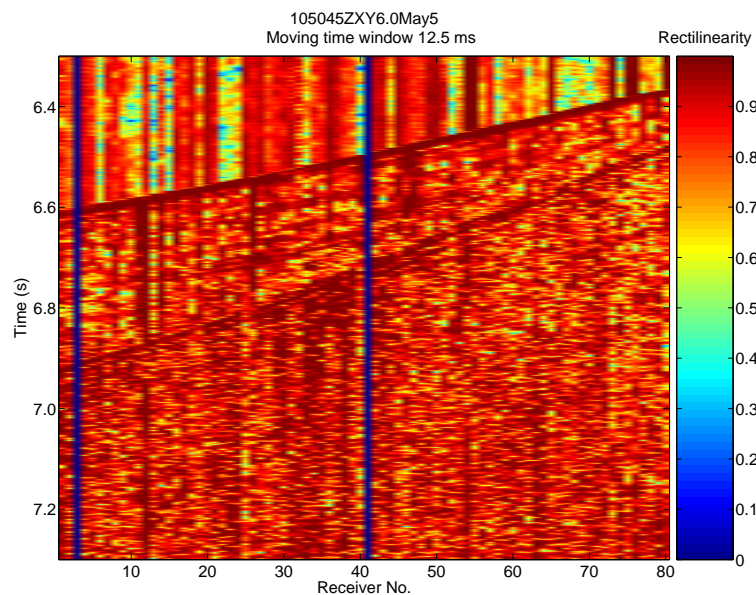

Figure 5.7: Rectilinearity analysis of the data shown in Figure 5.5 using a moving time window of 50 samples which corresponds to 12.5 ms length.

Polarization estimates were compared over the whole 80-level array in order to increase the robustness of the event detection as well as for testing the polarization reliability. The rectilinearity estimates of the data shown in Figure 5.5 are shown in Figure 5.7. Two distinct onsets of remarkably high rectilinearity values are clearly visible. These onsets correlate with the P-wave and S-wave arrival, respectively. As described in section 3.1 low rectilinearity values (below 0.5) should be observed before the P-wave arrives. As shown in Figure 5.7 rectilinearity values higher than 0.6 can be observed for about half of the receivers. This might be

caused by some directed noise as described in De Meersman et al. (2006), by electronic noise or by sensitivity/coupling issues. The latter causes are excluded for two reasons. First, high rectilinearity can only be induced by sensitivity/coupling issues if two components are much less coupled/sensitive than the third one. This is very unlikely for the considered geometry. If sensitivity/coupling issues occur for an array with receivers clamped against the casing of a deviating well one would expect it on the component in line with the borehole. This means that only one component is biased which would still lead to an elliptical or circular polarization (values around 0.5). The second reason is that visual inspection of the data did not indicate sensitivity/coupling issues.

Directed noise or electronic noise can be a major problem in polarization analysis for low signal-to-noise ratios. However, STA/LTA ratios as well as the visual inspection of the data have shown that the signal-to-noise ratio is fairly high. For this reason the distortion of the estimated polarization vector should be very small.

Nevertheless, it was tested whether the orientation of the polarization vector was consistent for the dominant arrivals over the array or whether directed noise interfered significantly with the signal polarization. Therefore the dip ϕ (equation (2.22)) and azimuth θ (equation (2.23)) of the largest eigenvector p_1 (from equation (2.20)) of the polarization ellipsoid was calculated utilizing a moving time window of 12.5 ms. The resulting dip and azimuth values observed over the whole array are shown in Figure 5.8.

For the P-wave arrival the estimated dip (here positive upward from the horizontal plane) was consistent with the geometry of the array. It increased from the lower part of the array to the upper part. The dip values associated with the S-wave arrival were very low (around zero) over the whole length of the array. For this reason the S-wave was interpreted as an SH-wave. The azimuth of the P-wave arrival as well as of the SH-wave arrival is also very consistent over the whole array. Furthermore, the SH-waves azimuth is perpendicular to the P-wave azimuth which would not be the case if a component was significantly less sensitive/coupled than the others. For this reason it is possible to conclude that the estimated polarization vectors show significantly high consistency over the array and can be considered as highly reliable.

The next processing step was the selection of a time window around the P-wave onset. Therefore the single receiver detection algorithm was applied as described above. The automatic P-wave detection selected only a time window for the location algorithm when all three key characteristics (local maximum in the STA/LTA ratio, sharp onset in the frequency-content versus time and sharp rise in the rectilinearity values) were fulfilled and if the same key features were observed by the neighboring receivers in the array. The selected time window was in the order of two dominant periods of the detected event around its onset. In order to avoid any disturbance in the polarization analysis for the selective raytracing any interfering arrival was excluded from the selected time window. This means
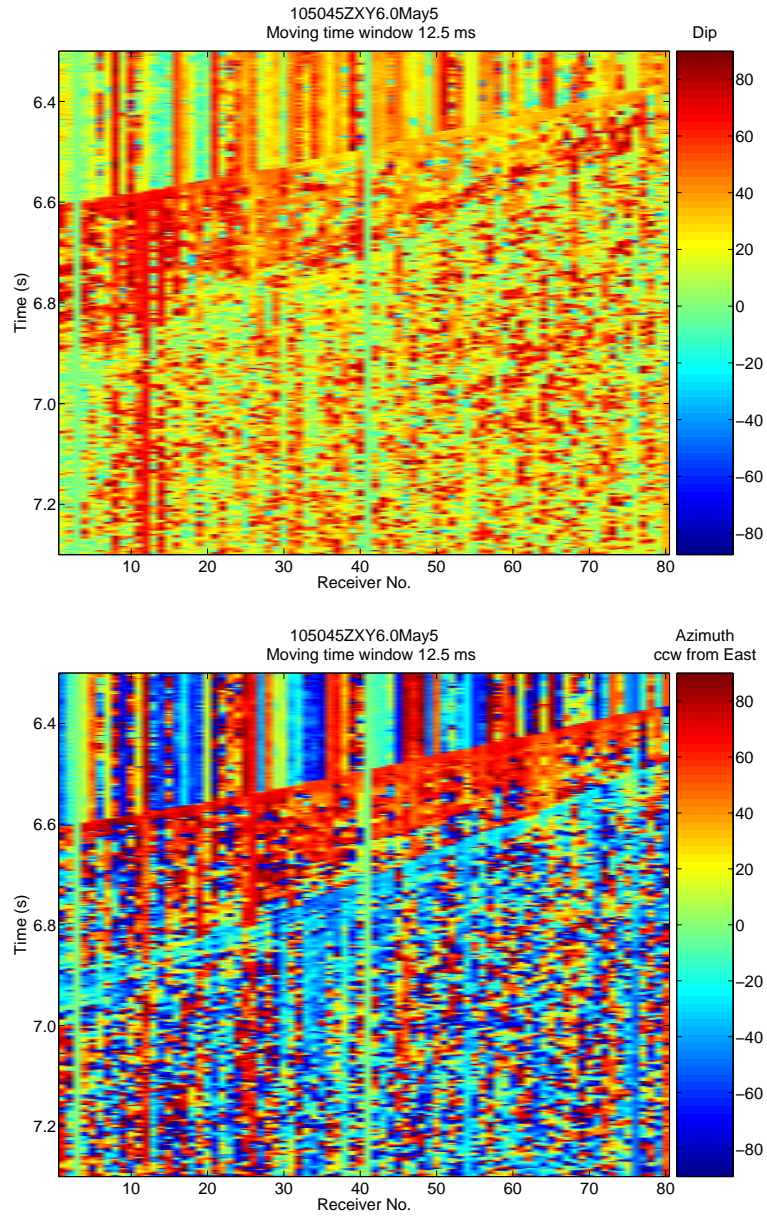
Figure 5.8: Polarization analysis of the data shown in Figure 5.5 using a moving time window of 12.5 ms length. Top: Dips of the dominant polarization vector in the moving time window. Bottom: Azimuths of the dominant polarization vector.

that the window length was automatically shortened if any interfering arrival was indicated by one of the key features. In the example given in Figure 5.6 the STA/LTA ratio rises again at about 6.5 seconds and hence, the time window had to be cut at 6.5 seconds.

Prior to locating the events the receiver coordinate system was transformed into a system perpendicular to the San Andreas Fault to match the reference system of the velocity model by Thurber et al. (2004) shown in Figure 5.4. This model has its origin at the Pilot Hole (Lat WGS 84, Long WGS 84: 35.9742579, -120.5521071 or UTM NAD27 Easting, Northing (m): 720810.16 3983669.97). The x-axis of the velocity model pointed positive N40E and the y-axis positive N130E. It was not necessary to calibrate depth values of the receivers since they were given in meters below sea level in both coordinate systems. The transformation of the data was straightforward. The UTM NAD27 coordinates of the receivers were reduced by the UTM NAD27 coordinates of the Pilot Hole. Afterwards, the horizontal receiver coordinates and data components were rotated utilizing a matrix of eigenvectors

$$\begin{pmatrix} u_x \\ u_y \end{pmatrix} = \begin{pmatrix} \sin(40°) & \sin(130°) \\ \cos(40°) & \cos(130°) \end{pmatrix} \cdot \begin{pmatrix} u_e \\ u_n \end{pmatrix} \qquad (5.4)$$

where the first column of the eigenvector matrix represents the (x,y)-coordinates of the x-axis which points 40° clockwise from North and the second column represents the (x,y)-coordinates of the y-axis which points 130° clockwise from North.

Then a more detailed polarization analysis was performed for the data in the selected time interval in order to analyze dips and azimuths of the detected P-waves over the whole receiver array. The results obtained for the event in Figure 5.5 are shown in Figure 5.9. The top plot in Figure 5.9 shows rectilinearity values obtained at different receiver depths. For the P-wave (blue crosses) almost perfect linear polarizations ($L > 0.97$) can be observed on most of the receivers. Moreover, all P-waves clearly fulfilled the rectilinearity threshold $L > 0.7$ which is marked by the solid red line. The S-wave rectilinearities (green crosses) also fulfilled the rectilinearity threshold apart from one receiver. Nevertheless, the S-wave is not as linearly polarized as the P-wave and the values scatter much more between the rectilinearity threshold and a perfect linear polarization.

The middle plot in Figure 5.9 shows the dip of the P-wave and the S-wave versus receiver depths as well as the dip of the borehole in that depth interval (magenta solid line). The estimated dip of the P-wave polarization vector follows the dip of the borehole for almost every receiver whereas the S-wave does not. Moreover, the dip of the S-wave polarization vector varies around a 10 degree dip over the whole receiver array apparently independent from the receiver depth. The azimuth of the P-wave polarization vector also follows the azimuth of the
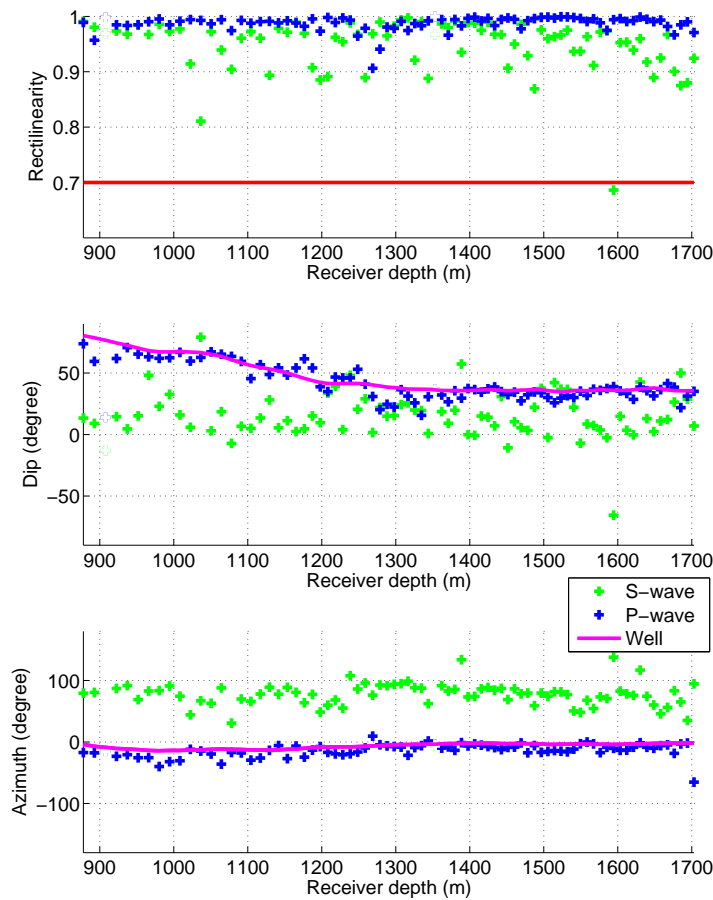
Figure 5.9: Polarization analysis of the microearthquake shown in Figure 5.5: rectilinearity (top), dip (middle) and azimuth (bottom). Blue colors correspond to results obtained from the P-wave and green colors to the S-wave. The red line in the top plot corresponds to the threshold set for quality control. The magenta lines in the middle and bottom plot represent the borehole dip and azimuth at the receiver locations. The azimuths are defined positive clockwise from Northeast.

borehole over the whole depth interval as shown in Figure 5.9 (bottom). This observation together with the observation that the estimated dip is consistent with the dip of the borehole implies that the currently analyzed event (shown in Figure 5.5) lies along the trajectory of the well.

5.5.2 Vp-Vs ratio analysis

It was noted that this particular event represents one of the target events and for the location it is necessary to restrict the Gaussian beams with approximate travel time information as described in section 5.3. For the calculation of the differences
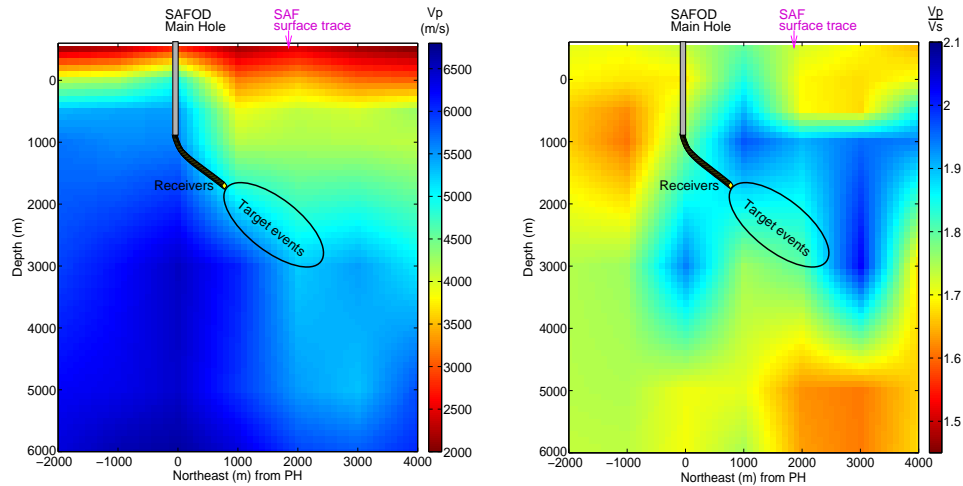
5.5. Processing and Results

Figure 5.10: Vertical slice of the P-wave velocity model (left) and the Vp-Vs-ratio model (right) through the hypocentral region of target-events.

of P-and S-wave arrival times (see equation (5.3)) information about the Vp-Vs-ratio are required. The Vp-Vs-ratio model given by Thurber et al. (2004) (see Figure 5.10, right) shows that the Vp-Vs-ratio in the area of the expected target events does not vary much. It can be approximated to be in the order of 1.83 to 1.85. Nevertheless, a further analysis as proposed by Chavarria et al. (2004) was performed to estimate an effective Vp-Vs-ratio from the arrival times of this target event. In order to obtain the highest accuracy for the target event location P-wave and S-wave arrivals were picked and cross-correlated. Using these arrival times the moveouts with distance to a reference receiver were analyzed. Such an analysis provides estimates of apparent[1] P- and S-wave velocities along the instrumented portion of the Main Hole. Figure 5.11 (left) shows moveout delays for P- and S-waves from receiver 20 to 77. One of the deepest (or closest to the event) receivers was selected as the reference receiver (number 77) and distances were calculated from it to every receiver.

The picked data seemed more reliable than the results obtained from cross-correlation, especially for the S-wave. The moveout curves clearly show changes in the slope. For this reason the array was divided into 6 segments and a linear fit to the moveout delays was applied to each of these segments (solid lines in Figure 5.11, left). Afterwards apparent velocities of the P- and S-wave were calculated. The result is shown in Figure 5.11 (middle). Note that values obtained at small distances from the reference receiver correspond to the deeper part of the array. Apparent velocities obtained at small receiver distances (up to 50 m) are not very reliable since the picking precision limits the resolution. The local decrease of the apparent P- and S-wave velocity correlates with a sudden velocity decrease in the

---

[1]For true velocities a correction of the obtained values for their associated emergence angles is required.

90
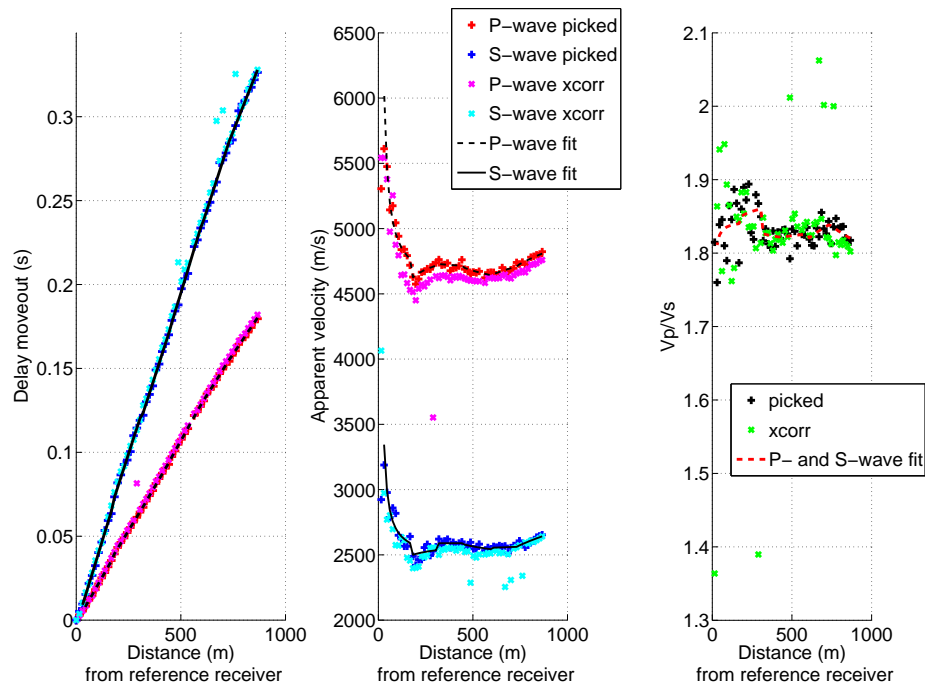
Figure 5.11: Estimation of apparent velocities and effective Vp-Vs ratio from P- and S-wave arrival time data of the target event from May 5.

sonic log data (at a measured depth of 2550 m) as shown in Boness and Zoback (2006) (see also Figure 5.2). As described in Chavarria et al. (2004) the emergence angle dependence of the apparent P- and S-wave velocities is the same as long as P- and S-waves have a similar travel path. Hence the angle dependence can be removed by taking the ratio of the apparent velocities and a local Vp-Vs-ratio can be estimated. The corresponding result is shown in Figure 5.11 (right) and it can be seen that the Vp-Vs ratio is higher than 1.8 for the whole investigated portion of the array. From this analysis as well as from the Vp-Vs-ratio model given by Thurber et al. (2004) an effective Vp-Vs ratio of 1.83 is estimated.

5.5.3 Target event location and uncertainty estimates

For the location of this target event an effective constant Vp-Vs ratio of 1.83, the P-wave velocity model from Thurber et al. (2004), the selected P-wave interval as well as the picked P-and S-wave arrival times were input into the location procedure that uses polarizations averaged in the selected P-wave interval. In order to account for possible errors in the P-and S-wave picks a deviation of ±5 ms was allowed between the calculated and observed arrival time differences. This means that the restricted area of the Gaussian beams along the rays has its center where the calculated and observed arrival time differences are the same and is extended to segments of ±5 ms along the traced ray. The obtained image of the target-event

using the velocity model from Thurber et al. (2004) is shown in Figure 5.12. The back-propagated energy clearly focuses at the hypocenter location at a depth of z=1970 m below sea level, x=1410 m northeast and y=-120 m southeast from the Pilot Hole.
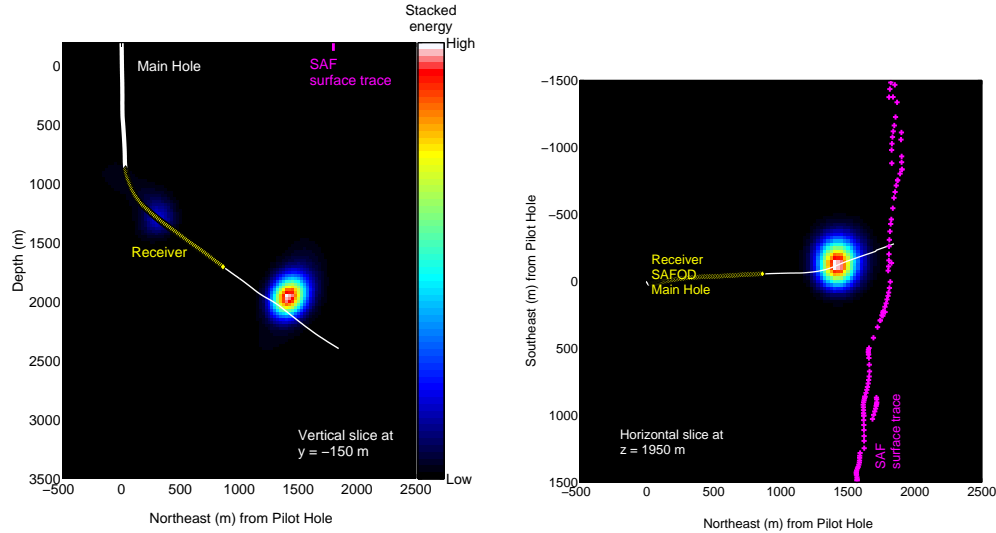

Figure 5.12: Resulting image for the target event. Black and blue colors correspond to small values of stacked energy and red to white colors in the center of this image to large values. The maximum value marks the event location. Yellow diamonds mark the receivers used for location.

The accuracy of this location depends on the accuracy of the velocity model as well as on uncertainties in the data set. The latter consist of two possible errors. The first one is defined by the accuracy of the arrival time picks. As stated above our location method for the target accounts for ± 5 ms picking uncertainty. Assuming a P-wave velocity of about 5000 m/s in the target event region a propagation time of 5 ms corresponds to a ray lengths of about 25 m. Hence, the beam restriction of ± 5 ms limits the location accuracy along the rays to about ± 25 m. The second data related error is induced by the error in the receiver orientation and defines the uncertainty in two dimensions perpendicular to the ray. This angle dependent uncertainty increases with increasing distance of the hypocenter from the array. The uncertainty of the orientation of the vertical component is assumed to be smaller than of the horizontal components because the original vertical component was always in-line with the borehole and hence is limited by the casing. The error was approximated to be in the order of ± 3 degrees. With a mean distance of 850 m between the lower third of the array and the located target event an uncertainty of ± 50 m can be estimated in the vertical direction perpendicular to the ray. For the orientation of the horizontal components an uncertainty of about ± 5 degrees is assumed. Consequently, the location uncertainty of the target event in the horizontal plane perpendicular to the ray is about ± 75 m. Note that these estimates are only valid for target events (less than 850 m away from the lower third of the array) and the values increase with increasing distance.

The location uncertainty related to the accuracy of the velocity model is much more difficult to estimate. In the following section six different velocity models have been used to locate the target event.

(a) Thurber 2006

(b) Thurber 2007

(c) Zhang 2007

(d) Roecker 2007 Altmod2

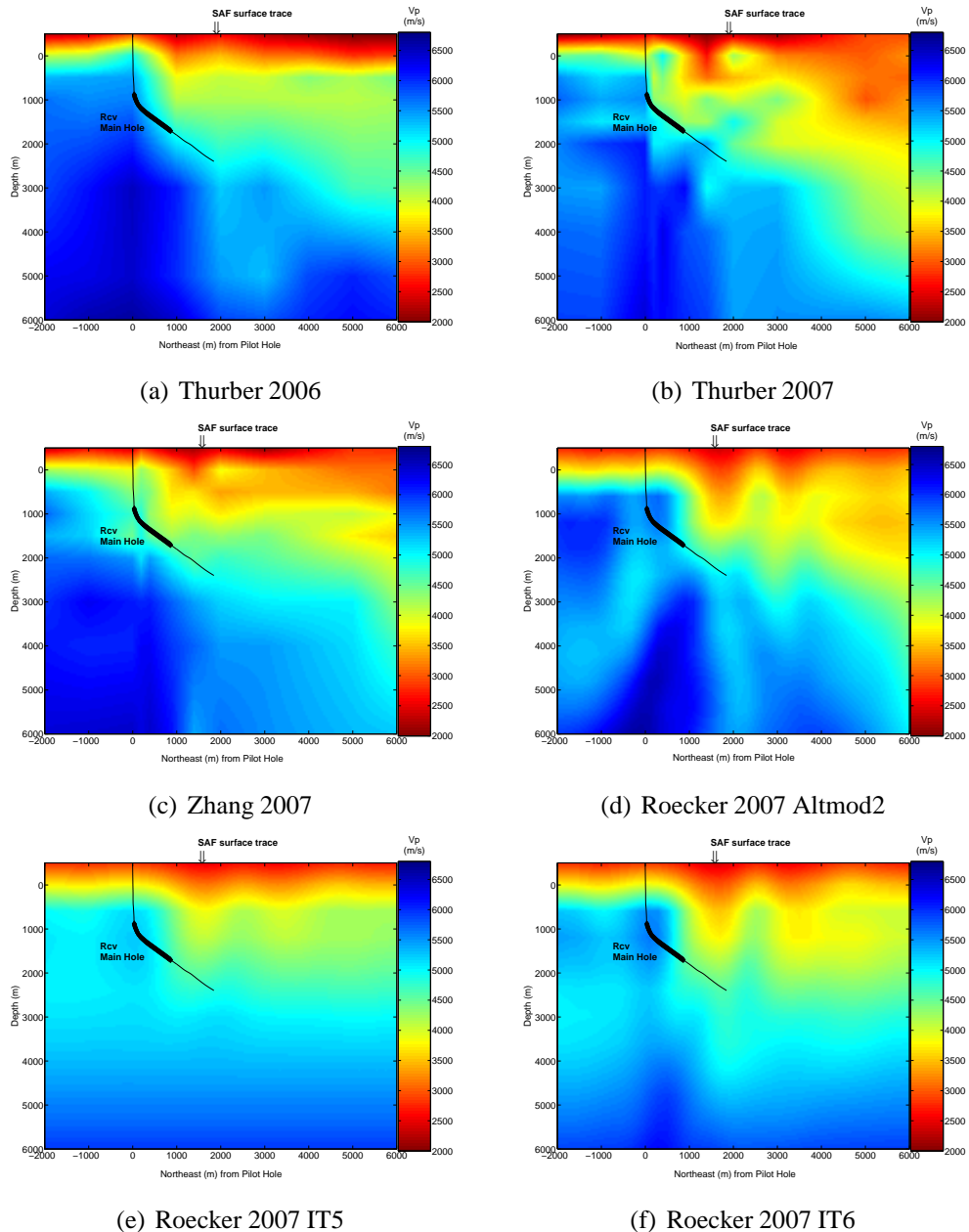(e) Roecker 2007 IT5

(f) Roecker 2007 IT6

Figure 5.13: Vertical slice of different P-wave velocity models across the San Andreas Fault (N40E) at 200 m distance N50W from the Pilot Hole. The black line represents the SAFOD Main Hole and the black diamonds the 80-level P/GSI array, respectively.

5.5.4 Target event location with different velocity models

Many different velocity models are available for the SAFOD site. However for this work only 3D models were used to test the robustness of the hypocenter location; one from Haijiang Zhang, two from Cliff Thurber, and three from Steven Roecker. The models from H. Zhang and C. Thurber were given in a coordinate system almost perpendicular to the San Andreas Fault surface trace, oriented the same way as described above with its origin at the SAFOD Pilot Hole. Grid sizes and spacing differed and details about the data extraction as well as for the interpolation on a regular grid are described in Appendix C.4. The velocity models from S. Roecker differ in orientation as well as origin from the other three velocity models. Details about the data transformation and coordinate rotation are given in Appendix C.4. A vertical slice across the San Andreas Fault 200 m northwest of the Pilot Hole of each velocity model is shown in Figure 5.13. The model named *Thurber 2006* was provided by Cliff Thurber in 2006 and corresponds to the model published in Thurber et al. (2004). A newer model was provided in 2007 which differs especially northeast of the Pilot Hole from the former one. These two models were obtained using both, earthquake and shot data (pers. comm., C. Thurber, 2007). The velocity model named *Zhang 2007* was provided by H. Zhang in 2007. The models provided by S. Roecker were also obtained using both, earthquake and shot data, although the model named *Roecker 2007 Altmod2* is more weighted towards earthquake data and the other two models are weighted towards shot data (pers. comm., S. Roecker, 2007). The difference between the latter two is the iteration step, with IT6 being the latest one. The shot data weighted P-wave velocity models show lower velocities compared to the other models which might be due to the different frequency content of shot and earthquake data. Also the P-wave velocity model from H. Zhang shows lower velocities especially in the region where the target events are expected.

A list of the obtained hypocenter coordinates using the different models is given in Table 5.1. The estimates for the location uncertainty given above apply to all the location results. As shown in Figure 5.14 the hypocenters are within the estimated 3D error boxes.

Velocity model	x (m) N40E from Pilot Hole	y (m) N130E from Pilot Hole	z (m) below sea level	UTM (m) NAD83 Northing	UTM (m) NAD83 Easting	Lat WGS84	Long WGS84
Thurber2006	1410	-120	1970	3985021.18	721529.86	35.9844969	-120.542759
Thurber2007	1430	-110	1920	3985030.07	721550.38	35.9845724	-120.542529
Zhang2007	1370	-130	1940	3984996.96	721496.49	35.9842864	-120.543135
Roecker2007altmod2	1400	-120	1970	3985013.52	721523.43	35.9844294	-120.542832
Roecker2007IT5	1360	-120	1980	3984982.88	721497.72	35.9841592	-120.543126
Roecker2007IT6	1350	-100	1960	3984962.36	721506.61	35.9839724	-120.543033

Table 5.1: Hypocenters of the target event from May 5 2005 using the P-wave velocity models shown in Figure 5.13.
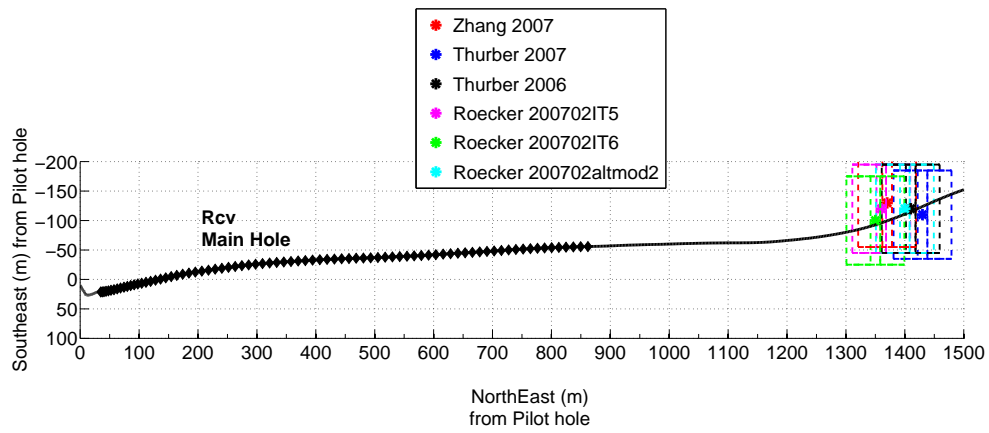
(a)
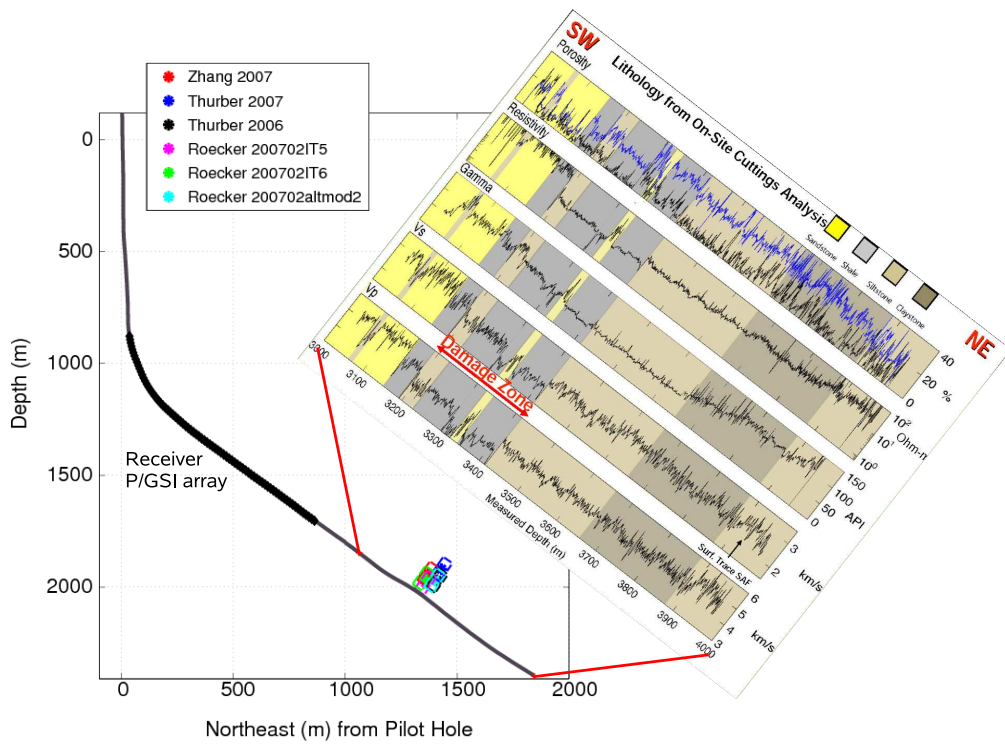
(b)

Figure 5.14: Locations of the target event from May 5, 2005 given in Table 5.1. The rectangular boxes represent the 3D error estimates. (a) Map view. (b) Side view including borehole lithology information from Zoback et al. (2007)

5.5.5   Correlation with Main Hole logging data

As indicated by the seismic imaging results from Buske et al. (2006) as well as by
the previous event locations from Thurber et al. (2004) it is possible to assume a
subvertical orientation of the different branches of the fault system. A vertical pro-
jection of the obtained hypocenters would intersect the Main Hole at about 2080 -
2100 m below sea level which corresponds to 3400 - 3410 m measured depth. The
Main Hole logs show a significant decrease in P- and S-wave velocities as well as
a resistivity increase at this measured depth (see Figure 5.14(b)). These observa-
tions might indicate a connection between the damage zone at 3400 m measured
depth and the located target event. However, depending on the projection angle
from 90° (vertical) to about 70° the target event can be associated with any fault
branch indicated in the well log at measured depths between 3200 m and 3400 m.
It is not possible to assign the target event to an individual fault branch in the
damage zone as long as the dip of these branches is not exactly known.

5.5.6   Complex waveforms

After the basic preprocessing (that consisted of raw data conversion and data ro-
tation into geographical coordinates) the files were scanned for event detection.
For all detected events a rectilinearity analysis was performed in order to test,
whether the detected event passed the rectilinearity threshold. When inspecting
the detected events some very complex waveforms were observed. An example is
shown in Figure 5.15. Overall, a change in the slope of the moveout can be clearly
observed for both, the P- and the S-wave. In fact, three distinct moveout slopes
can be estimated: (1) from receiver 1 to 30, (2) from receiver 30 to 50 and (3)
from receiver 50-80. A different moveout slope for the upper 30 receivers could
be explained by the change in the geometry of the array (the kink in the borehole).
Nevertheless, the move out changes in the lower part of the array were completely
unexpected and were not observed in the waveforms of the target event (see Fig-
ure 5.5). Comparing the polarization attributes (dip and azimuth) of the P-wave of
this detected event with the orientation of the borehole showed that this event has
its origin somewhere north or northwest of the array. This means that this event
is not a target event. For this reason a comparison with the waveforms from the
target event cannot help to understand the complexity of these waveforms.

The application of the location procedure to the event in Figure 5.15 provided a
distinct energy maximum which was interpreted as the hypocenter. Some further
investigations showed that mainly rays traced from the upper part of the array con-
tributed to the hypocenter location. In contrast, the rays traced from the deeper part
of the array did not intersect each other but rather the San Andreas Fault itself at
different coordinates. This can be caused by wrong estimates of the starting direc-
tion of the ray (wrong polarization) or by errors in the velocity model (responsible
for ray bending). Another possibility is that the wave forms that were interpreted
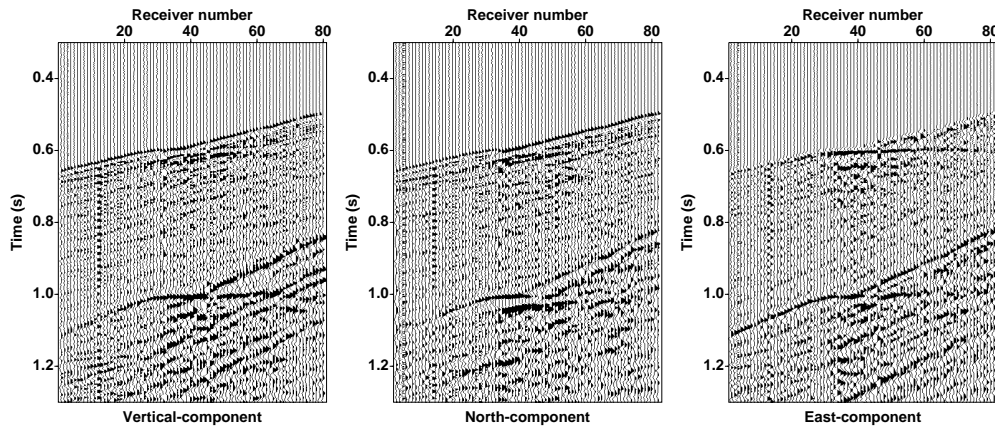
Figure 5.15: Seismograms of a non-target event.

as the direct P-wave are actually reflections from the fault. The obtained location of the events hypocenter was used as a source in order to simulate reflections from a vertical reflector at horizontal coordinates of the San Andreas Fault surface trace. The reflections were simulated using the seismic 3D ray modeling software from *NORSAR* and the same velocity model as used for the location procedure.[2] From the simulated reflections arrival times were hand-picked (to provide most reliable information) and the picks were converted into three moveout slopes (1) from receiver 1 to 30, (2) from receiver 30 to 50 and (3) from receiver 50-80. These slopes can be compared with the moveout slopes observed in the data. In Figure 5.16 such a comparison is shown.

It can be seen that the moveout slopes from receiver 1 to 30 as well as from receiver 30 to 50 observed in the data do not perfectly match the moveout of the simulated reflection. Consequently, the P-wave moveout observed at the upper part of the array does not correlate with the P-wave moveout of the simulated reflection (at least not for the assumed source - reflector - receiver geometry). In fact, this comparison was performed for three different events with completely different located hypocenters. For all three events of very different source locations (one horizontally in-line with the array but much deeper, one northwest of the array and one southeast of the array) the simulated moveout slopes of reflections did not match the observed moveout at the upper receivers. In contrast, the moveout observed in the data and obtained from the reflection simulation shows some consistency for the deeper 30 receivers (at least for two of the three simulations). Even if this analysis does not provide a prove that the first arrivals at the deeper receivers belong to a reflection from a certain branch of the San Andreas Fault some conclusions can be drawn. Firstly, the arrivals at the upper part of the array do not match the behavior of a simulated reflection. Moreover, the rays traced away from these receivers intersect and provide a hypocenter location. For this reason it is reasonable to assume that the arrivals at the upper receivers repre-

---

[2]The simulation was performed in cooperation with Dr. Alex Goertz from P/GSI.
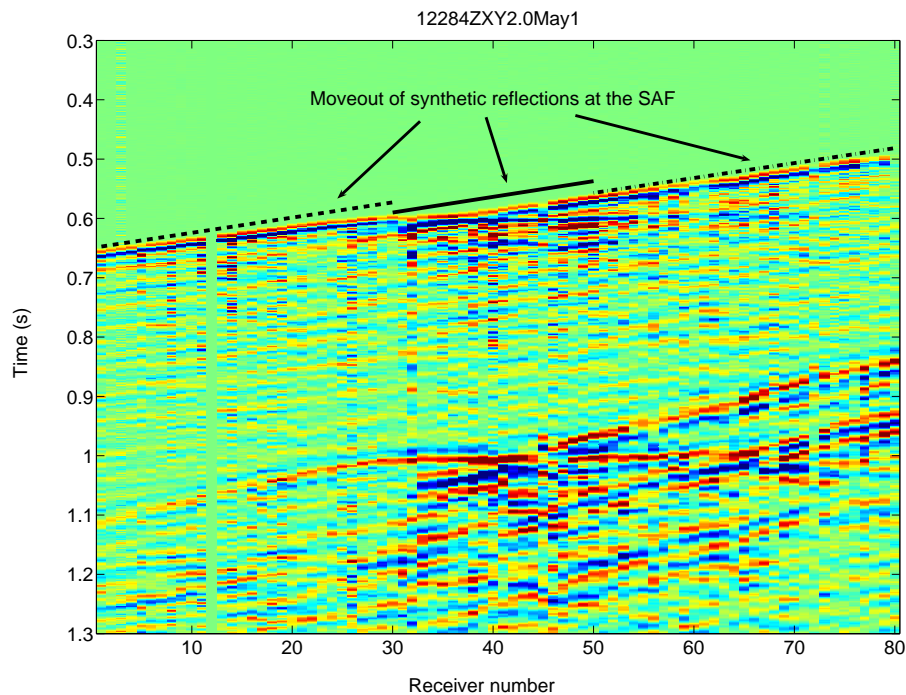
Figure 5.16: Comparison of the moveout observed in the data and of a simulated reflection at the San Andreas Fault.

sented the direct P-wave of the event. Secondly, it was found that the arrivals at the deeper part of the array do not contribute to the hypocenter location because their polarization information does not lead to intersecting rays (or Gaussian beams). The cause of the complexity of these arrivals is still under discussion; e.g., it is still possible that they correspond to a direct P-wave which propagated through a complex velocity structure which is not captured by the velocity model. Two further tests were performed to ensure that the obtained locations were robust and were not disturbed by back-propagated energy from the deeper array part. For the first test, the deeper receivers were simply excluded from the location process. The resulting hypocenter estimation was the same as for the initial location. For the second test the SAFOD specific implementation of using P-and S-wave arrival times to restrict the Gaussian beams was utilized. Again, the hypocenter correlated with the initial location. Hence it is reasonable to conclude that the location algorithm provides good estimates of the hypocenters.

5.5.7 Location of detected events

The location procedure was applied to all detected events which passed the polarization rectilinearity threshold as mentioned above. In order to ensure robust location results automatically picked P-and S-wave arrival times were used to restrict the Gaussian beams for all event locations. The obtained locations are

shown in Figure 5.17 projected onto a Vp-Vs-ratio slice perpendicular to the fault (left) as well as on a map view at about 3000 m depth (right). The map view shows that the majority of events is located near to the San-Andreas Fault surface trace (magenta crosses) mapped by Michael Rymer (U.S.G.S.). The map view indicates that the Southeast axis of the migration model is not 100 % parallel to the mapped surface trace. For this reason the event distribution projected onto the vertical slice appears more scattered around the San-Andreas Fault surface trace than it actually is.
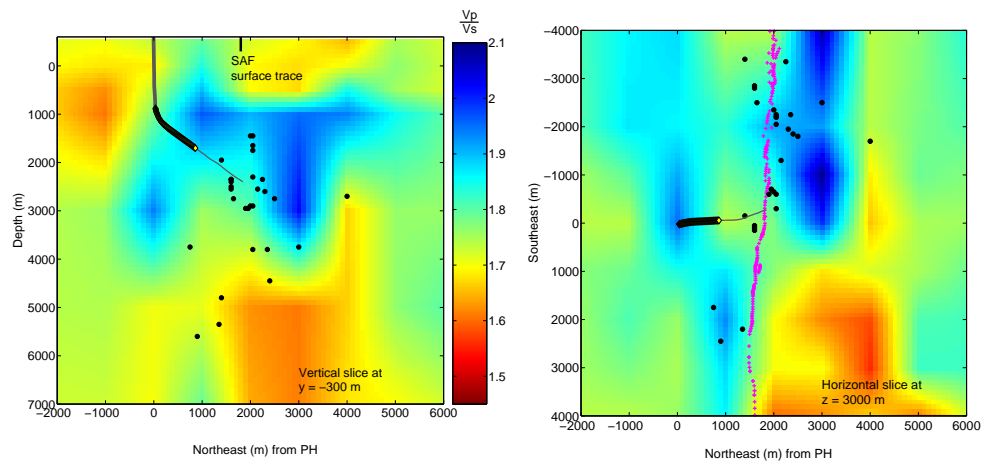

Figure 5.17: Obtained event locations projected on slices of the Vp-Vs-ratio model. Left: Vertical slice about 300 m Northwest of the Pilot Hole. Right: Horizontal slice in 3000 m depth.

5.5.8 Waveform correlation

Furthermore, a waveform correlation analysis was applied to the located events in order to identify multiplets similar to those described in Bakun et al. (2005). The resulting cross-correlation coefficients are shown in Figure 5.18. Two subsets of events were observed that occurred within hours in two distinct regions have highly correlated waveforms. Firstly we can identify a doublet with a cross correlation coefficient as high as 0.95 on May 1 which occurred at about x = 1600 m, y = 150 m from the Pilot Hole and z = 2500 m below sea level. Secondly, a triplet occurred on May 8 about x = 2000 m, y = -650 m from the Pilot Hole and z = 2900 m below sea level. The cross-correlation coefficients for this triplet are 0.85 and higher. The three-component waveforms of this triplet are shown in Figure 5.19. Even if the waveforms of this triplet are highly correlated it was found that the amplitudes are smaller by a factor 10 and 40 for the later events, respectively. This in turn suggests that these events have similar focal mechanisms and have ruptured approximately the same area but that the seismic
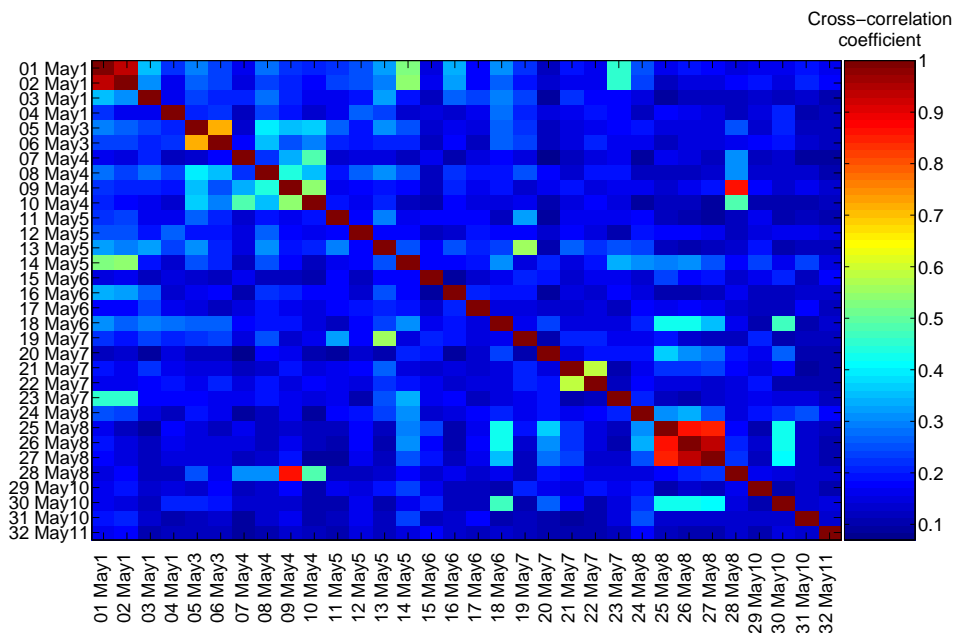
moment of the events decreased with time.

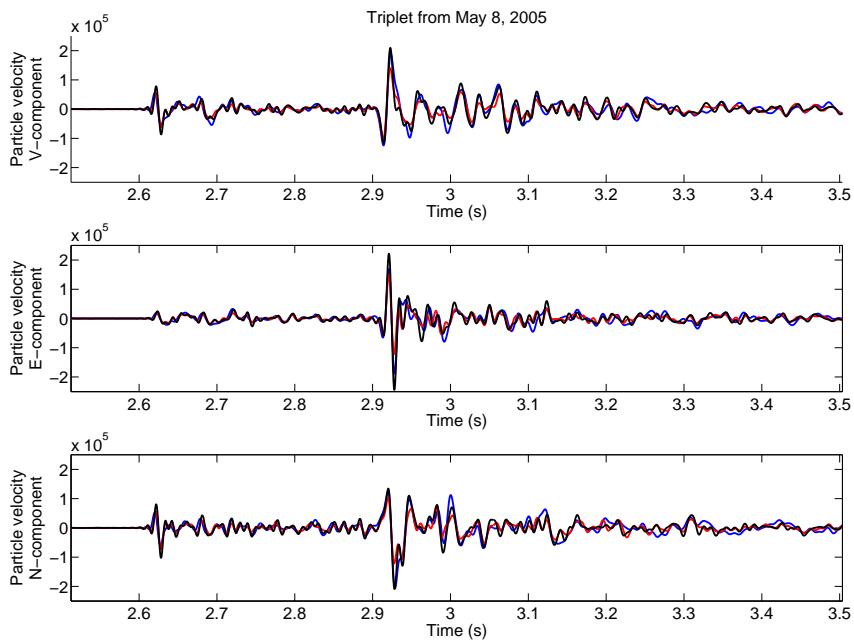Figure 5.18: Waveform correlation of the located events.

Figure 5.19: Three-component waveforms of the triplet May 8 recorded at receiver no. 78. The data are filtered with a zero-phase 100 Hz Butterworth lowpass filter. The waveforms of the event shown in blue are original, whereas the waveforms of the event shown in red (black) are 10 (40) times increased.

On the other hand, Figure 5.18 also shows that doublets can occur with larger time delays. The highly correlated event pair of May 4 and May 8 occurred with a delay of 4 days and did even show the same amplitude ranges for both P- and S-waves indicating that this doublet had not only the same focal mechanism but also the same seismic moment. This doublet was located very close to the event pair with similar waveforms from May 3 in a region below the San Andreas Fault surface trace where the Vp-Vs-ratio from Thurber et al. (2004) shows clearly increased values which is probably related to fluid filled fractures in the shear zones (Chavarria et al., 2004). Interestingly, both event pairs differ in their waveforms so much that the cross-correlation between them gives correlation coefficients below 0.4 (see Figure 5.18). This implies that the doublet of May 3 differed in focal mechanism and/or rupture plane from the event pair of May 4 and May 8. Unfortunately, the aperture is not suitable to resolve focal mechanisms or rupture planes and consequently it was not possible to interpret the observed differences in the waveforms of the two doublets any further. However, the occurrence of doublets and triplets during two weeks of continuous recording in May 2005 supports the observation of Bakun et al. (2005) that clusters of microearthquakes repeatedly rupture small, fixed parts of the complex fault system.

5.6   Discussion

The principle aim of the SAFOD project was to provide locations of events recorded with an 80 level borehole seismic receiver array from Paulsson Geophysical Services Inc. deployed in the SAFOD Main Hole in 2005. The data set contained several events including the target event of May 5, 2005. The location of the target event required special attention because an accurate location was aimed for the precise determination of the next SAFOD drilling target (PHASE3). Target events provide a pitfall for the location method developed in this work. Nevertheless, the described SAFOD specific implementations overcame the problem and allowed a precise location of the target event. Moreover, the robustness of the location was evaluated using six different available 3D velocity models. The uncertainty of the target event locations was estimated in three dimensions taking into account the picking errors, and errors in the receiver orientation. It was found that the six hypocenter estimates from the different velocity models match within the estimated 3D uncertainty.

The data set also contained events that occurred outside the target region. These events were also located and the obtained locations are consistent with the San-Andreas Fault surface trace. Some further investigations also allowed the identification of some doublets and even one triplet in the located earthquakes. The consistency in waveform similarity and earthquake location not only confirms the functional capability of the developed location procedure. It also supports earlier multiplet observations but this time the data quality is much higher due to the

small distance between the recording network and earthquake sources as well as the fact that the borehole recordings provide a much better signal-to-noise ratio than the surface stations used for earlier analyses.

From the observation of the complex waveforms as well as from the reflections often observed after P- and S-wave arrivals the conclusion follows that future work should include further investigations of these data characteristics. For example, the located earthquakes could be treated as active sources for reflection seismic imaging in order to illuminate the complexity of the San Andreas Fault System in great detail and refine large scale seismic imaging results (Buske et al., 2006).