

A Short Analysis of San Andreas Fault Geometry in Parkfield, California

Adam Margolis

December 11, 2025

1 Introduction

For my final project, I wanted to do an inversion (or two) to determine the strike and dip of the creeping Parkfield segment of the San Andreas Fault (SAF), compare those two approaches to one another, and compare it to the known value of strike and dip from the SCEC-CFM [2]. The fault, which is a right-lateral strike-slip fault, is actively creeping between Cholome and San Juan Batista at a rate of 26.2 mm/yr [6]. See Figure 1 for an overview of my study area.

Accurate strike and dip measurements are crucial components in dynamic rupture models, where fault geometry can control rupture propagation and termination. On a future project, I will need to constrain the strike and dip of the fault in this region to further study how rupture propagates between the transition from creep in Parkfield in the north to fully locked in the Carrizo Plain in the south.

On the morning of 9 January 1857, the largest earthquake documented to strike California occurred as a magnitude 7.9, rupturing the SAF from Cholome (just south of Parkfield) to the Cajon Pass about 350 km away [3]. The epicenter of this earthquake was estimated to occur extremely close to the transition between the creep and locked segments of the fault [1].

Since about 1857, there has been a documented magnitude 6.0 earthquake that repeats about every 22 years [7]. The most recent "Parkfield earthquake" occurred on 28 September 2004, and the next one is due to occur soon as it has been just over 21 years.



Figure 1: A map of California, showing my study area (indicated by the orange triangle). Also noted are other points of interest in red circles

It has been hypothesized that these "regularly repeating" magnitude 6 events can load the SAF with stress over decades of moderate earthquakes. Eventually, the fault is likely loaded to its yield stress and the fault ruptures. Large (magnitude 7 or greater) events south of Parkfield occur on average every 150 to 200 years [4]. It is highly likely that both the Parkfield and Carrizo Plain segments are both near the end of their current seismic cycle.

I hypothesize, and will test in subsequent studies, that repeating moderate (M 6.0) earthquakes in the Parkfield segment may periodically cause an 1857-like major rupture of the San Andreas Fault. Are these two unique earthquake cycles (22 years vs 150 years) ever coupled? To do this accurately, I will require a well constrained fault geometry to run dynamic rupture models. That is where this final project comes into play, to determine the geometry of the fault (the strike and dip). The first part of that larger project, is the final project for this class.

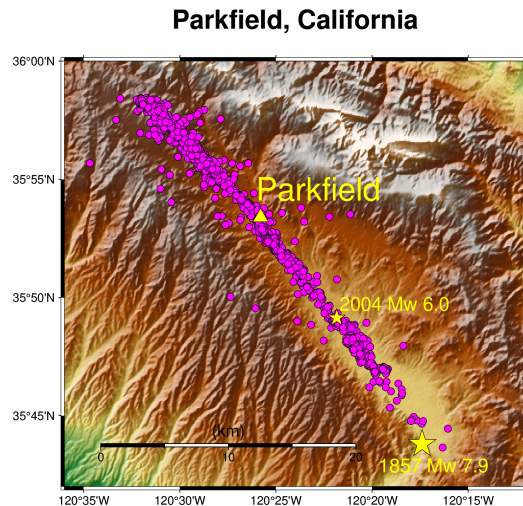


Figure 2: A map of events (shown as magenta dots) of the Parkfield section of the San Andreas Fault in the 30 days after the 2004 Parkfield earthquake. The town of Parkfield is indicated by the yellow triangle, and the 2004 Parkfield and 1857 Fort Tejon epicenters are marked by yellow stars.

2 Methods

2.1 Data downloading and Organization

To obtain a generalized fault geometry, I need to calculate the strike and dip of the fault. Since earthquakes occur on faults, I should be able to use a catalog of aftershocks which would then provide the location in 3d space where that fault lies. Given many data points, or in this case hypocentral locations of aftershocks, it is possible to invert the data to calculate strike and dip of a fault.

Eventually, I will split the fault into smaller segments, to get more localized variations in strike and dip. However, for this project I wanted to see if I could get an "average" strike and dip for the Parkfield segment and compare that to the average strike and dip values from Statewide California Earthquake Center - Community Fault Model (SCEC-CFM) [2]. I downloaded earthquake data from the USGS earthquake catalog [5] from the day of to 30 days after the 2004 Parkfield earthquake. The exact time was: 2004-09-28 00:00:00 UTC to 2004-10-28 23:59:59 UTC. The minimum and maximum magnitudes are -5 to 10, respectively. The geographic region I chose had

north at: 35.874, west at: -120.58, east at: -120.265, and south at 35.722. This would include a few foreshocks, the mainshock, and mostly aftershocks. There were a total of 2003 unique events from this catalog, which can be seen in Figure 2.

USGS creates a CSV file of the data. I then loaded it into Jupyter Notebook where I did my calculations and inversions. Next, I removed the header and extracted the three columns of data I needed. The longitude, the latitude, and depth of the event. Then I removed the comma separators to make things a bit easier to find.

Next I removed the no longer needed files and I am left with a txt file that contains the three columns of the long, lat, and depth of events. Then I extract those out to variables north, -west and depth. I also changed the west coordinates to be positive, so that all my numbers are positive.

2.2 Inverting for Best Fit Plane

Once I have my data organized, I set up my first inversion which will solve for a, b, and c in the generalized plane equation:

$$0 = ax + by + cz \tag{1}$$

To do the inversion, I need two matrices (A and d) to solve for the model matrix (m): $A = md$. A is the observation matrix (north and west coordinates) and d is the data matrix (the depths of the events). I then need to solve for m so that its components contain a, b, and c for the generalized plane equation.

Next, I set up my A matrix by taking the west coordinates and north coordinates as the first and second row, respectively. I also have a third row of ones that is of length equal to the number of earthquake events. In this case, I made it equal to the number of unique west coordinates, which is also the same number of earthquakes from the catalog. I then reshaped them into columns before stacking them as the A matrix. I set the d matrix to have the depths of the events. All that is left is to do the actual inversion.

I solve for the inversion using the same equations and formulas we were shown in class. I use matrix multiplication to multiply the A matrix to the transpose of the A matrix. Then I take the transpose of A and matrix multiply that to the d matrix. All that is left is to solve for m which can be done by inverting the first result I got and matrix multiplying it to the

second. Now I just solved for m and inside contains the values of a, b, c of the generalized plane equation.

The plane equation I inverted for yields: $34.9x - 44.2y - 2608.7z = 0$. This may not seem very useful in its current form, however. Yet, from this I can calculate the normal vector, or the vector that is perpendicular to the plane. Using the normal vector, I will be able to calculate the strike and dip of the plane.

2.3 Solving for Strike and Dip from Best Fit Plane

To solve for strike and dip, I need to first solve for the normal vector to the generalized plane equation I initially solved through the inversion. To create the normal vector, I take the first and second values (a and b) and set c to -1. I then normalize the normal vector using its magnitude. And from that, I take the first, second, and third values as the normalized normal vector components (a_n, b_n, c_n). All that is left is to use simple geometry to solve for the dip:

$$dip = arctan(\frac{\sqrt{a_n^2 + b_n^2}}{c_n}) \quad (2)$$

From this equation, I calculated a dip of about 89 degrees.

Calculating strike is more straightforward. First I need to calculate the azimuth, or the compass orientation of the normal vector from before:

$$strike = arctan(\frac{a_n}{b_n}) \quad (3)$$

Next I need to convert it to degrees from north. From this, I calculated a strike of N 38 W, or 322 degrees.

2.4 Inverting for Best Fit Line

The next inversion I did is to solve for the best fitting line. This ignores the depth component of the events, and only considers the epicenter information (latitude and longitude). This way, I can easily plot the line on the plot of earthquakes (end result is Figure 3). From this, I'll be able to determine if the line more or less follows the "linear trend" of earthquakes along the fault as a check for the data. The Parkfield section is one of the straighter portions

of the fault, so approximating this portion of the fault as a line should be a good, simple approximation on the tens of kilometer scale.

This makes inverting for the strike a little simpler. Also, for this inversion I am only considering a 2-dimensional fault geometry, which does not include the full 3d picture. Therefore, some differences in my strike calculation are expected, and hence a second dip calculation is absent.

To invert for the best fit line, I use a similar method as inverting for the best fit plane. Although I followed what we did in class for the best fitting line from the Keeling curve exercise as a template. The general equation for the best fitting line is, where m is the slope and b is the y-intercept:

$$y = mx + b \quad (4)$$

My A matrix has the the first column as the west coordinates for the epicenters from the catalog and the second column is a column of ones. The d matrix contains the north coordinates of the epicenters. Just like before, I solve for m which contains the slope of the line as the first value and the y-intercept as the second. The equation of the best fit line I calculated is $y = 0.94x - 77.24$. This line is plotted as the green line in Figure 3.

2.5 Solving for Strike from Best Fit Line

Similar to before in section 2.3, I use the slope I calculated to solve for the strike in radians using \arctan , then I convert that value to degrees. I calculated a strike of N 43 W or 317 degrees. The equation for strike is:

$$strike = \arctan(m) \quad (5)$$

3 Discussion

The strikes I calculated using the best fit plane and best fit line methods are 322 and 317 degrees, respectively. These values represent an average across all 2003 events along the Parkfield section, in about a 25 km long segment of the fault. These values are close to each other, and do not vary greatly which implies that the methods used here are good approximations. That said, the best fit plane method is more accurate because it includes the depth component of the events. Therefore, 322 degrees is what I believe is more correct. The SCEC-CFM has an average value of 142 for the strike of the

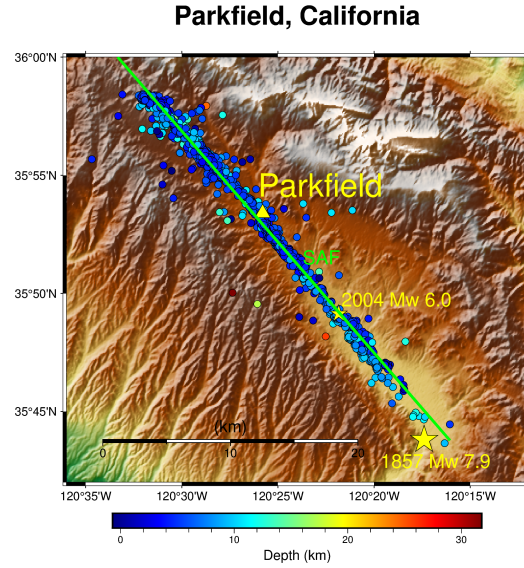


Figure 3: A map of events (shown in various colors by depth) of the Parkfield section of the San Andreas Fault in the 30 days after the 2004 Parkfield earthquake. The town of Parkfield is indicated by the yellow triangle, and the 2004 Parkfield and 1857 Fort Tejon epicenters are marked by yellow stars. In green is the best fit line, solved by the inversion. This is approximately the San Andreas Fault trace at the surface

Parkfield section, which is exactly 180 degrees off from my calculated value [2]. These values are geometrically identical and therefore are in agreement with one another.

Both of my calculated values for strike are similar, but also not exactly the same because the best fitting plane method considers the depth component of events. So the plane will need to be orientated properly to factor in the depth of the events. The best fitting line method, however, is independent of depth and therefore is a quicker way to obtain the strike, but as shown here is off by several degrees and should be used with some caution.

The dip I calculated using the best fitting plane method, about 89 degrees, is in agreement with the SCEC-CFM which has the average dip of 84 degrees [2]. And this also agrees with the fact that since the San Andreas Fault is a strike-slip fault, its dip should be nearly vertical at 90.

However, I feel I would be able to get a dip closer to 90 if I removed a few outlier events from the data. As shown in Figure 3, a handful of events

are not located on the SAF trace and are exceeding 30km in depth, which is much deeper than the 5-10 km range they are occurring along the fault trace. This is something I will consider moving forward with expanding this project next quarter.

4 Conclusion

In this short study, I was able to invert for the best fitting plane and best fitting line to then calculate strike and dip of the San Andreas Fault within the Parkfield segment. The values I calculated are averages across, roughly 25 km length of fault. I will further use this analysis to slice the fault into smaller and smaller sections where I can invert and calculate more localized variations in the fault strike and dip using an even larger catalog that is several decades long (i.e. 2000 to 2025). I will then be able to construct a well constrained fault geometry that can be used in dynamic rupture models to further analyze the seismic hazard risk posed by the repeating moderate Parkfield earthquakes and the longer earthquake cycle of the locked Carizzo Plain region.

References

- [1] J.J. Lienkaemper. 1857 Slip on the San Andreas Fault Southeast of Cholame, California. *BSSA*, (6):1659–1672, 2001.
- [2] A.; Shaw J. Marshall, S.; Plesch. Scec-cfm. 2022.
- [3] K.E. Meisling, K.E.; Sieh. Disturbance of trees by the 1857 Fort Tejon Earthquake, California. *JGR*, (B6):3225–3238, 1980.
- [4] T. Rockwell. The Non-Regularity of Earthquake Recurrence in California: Lessons from Long Paleoseismic Records from the San Andreas and San Jacinto Faults in Southern California, and the North Anatolian Fault in Turkey. *International Conference on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, (5), 2010.
- [5] United States Geological Survey. Usgs earthquake catalog. 2025.

- [6] J.R.; Rymer M.J.; Landgraf A.; Haddad D.E.; Busch M.; Coyan J.; Hannah A. Toke, N.A.; Arrowsmith. Late Holocene slip rate of the San Andreas fault and its accommodation by creep and moderate-magnitude earthquakes at Parkfield, California. *Geological Society of America*, (3):243–246, 2011.
- [7] R.; Rice J.R. Tse, S.T.; Dmowska. Stressing of locked patches along a creeping fault. *BSSA*, (3):1706–736, 1985.