AGU PUBLICATIONS

Geophysical Research Letters

Supporting Information for

Repeating Earthquakes with Remarkably Repeatable Ruptures on the San Andreas Fault at Parkfield

Rachel E. Abercrombie[1], Xiaowei Chen[2], and Jiewen Zhang[2]

[1]Boston University, [2]University of Oklahoma

Contents of this file

Text S1 to S6
Figures S1 to S8

Additional Supporting Information (Files uploaded separately)

Excel File: pfield_repeating_directivity_2019_SI.xlsx

Table S1. Hypocentral parameters for events in Sequences:

The NSCN ID is from the NCSN web page, the origin time from the Waldhauser and Schaff double difference catalogue (2008), and M_w from spectral stacking. LM = Sequence number in Lengline & Marsan (2009). McL = sequence in McLaskey *et al.* (2012). RECU = sequence number in Rubinstein *et al.* (2012). Note that these ended prior to this study so do not include most recent repeats. Sequence 9 was identified as two separate sequences by Lengine & Marsan (2009). A number of their sequences end and start at the 2004 M6 event suggesting that the perturbation in attenuation and velocity at the time was sufficient to upset their automatic discrimination. Events from both of their sequences in Sequence 9 were also identified as in a single sequence by Rubinstein *et al*. (2012). The results of this analysis, including detailed observation and analysis of the individual seismograms suggests that these should be considered one sequence.

* indicates event not included in analysis as occurred before current HRSN recording system

Introduction

This Supplementary Information includes a description of the plotting information for Figures 1 and 2, a Table with the Hypocentral information of all earthquakes used, and text (S1-S6) and figures (S1-S8) providing additional description of the methods, and examples of the modeling.

Text S1. Plotting Information

For precision and reproducibility, we provide the exact information plotted in Figure 1. Origin: all cross sections are plotted with an origin of 35.9841, -120.5427, the location of earthquake ID 30227076 in the (NCAeqDD, Waldhauser and Schaff, 2008) catalog. This earthquake is one of the SAFOD repeaters.

The cross sections are at an angle 140 deg, to North, consistent with the earthquake lineation along the San Andreas fault at Parkfield (see Thurber *et al*., 2006).

The 2004 M6 earthquake hypocenter is from the NCSN online catalog:

35.81816, 120.366, depth 8.136 km

The 1966 M6 earthquake hypocenter is from the NCSN BSL online catalog:

35.95, -120.5, and we fix the depth at 9 km

The earthquake hypocentral information for all events used in this analysis is included in Table S1.

Text S2. Identification of Repeating Sequences and Empirical Green's Function Analysis

To identify repeating earthquake sequences, we require all the hypocenters to be within 150 m of the approximate centroid of the sequence, similar to the resolution of the NCAeqDD catalog combined with the effects of the changes in near-surface velocity structure following the 2004 M6 earthquake (Rubinstein and Beroza, 2005; Wu *et al.*, 2016). We then check for waveform consistency and coherency – over 0.9 to 60 Hz for many station-event pairs. We also confirm that all sequences have been previously identified by other researchers, meeting their particular systematic criteria for similarity (McLaskey *et al.*, 2012; Lengliné and Marsan, 2009; Rubinstein *et al.*, 2012; see Table S1).

We follow the procedures described by Abercrombie *et al.* (2017a) and use small earthquakes as empirical Green's functions (EGFs) to remove the path effects and calculate spectral ratios and relative source time functions (STFs) at each station for each earthquake. We set the time windows to the S-P time of the closest station, 0.7 s for Sequence 2, and 0.6 s for Sequences 5 and 9. These are shorter than time windows calculated following Abercrombie *et al.* (2017a), but at least five times longer than the expected source duration.

To account for location uncertainties, we initially search for EGF earthquakes within 0.75 km horizontally and 1.2 km vertically of the sequence centroid, and restrict their magnitude to between 1.3 and 2.5 M units smaller than the sequence mean. We select the most appropriate EGFs from these following initial analysis.

We deconvolve the EGF events from the main events (Prieto *et al.*, 2009) and obtain spectral ratios and source time functions (STFs) for each event at each station and component.

We use only ratios and STFs for seismogram pairs (per station and component) with a cross-correlation >0.8 and a depth difference of less than 150 m, or a cross correlation of at least 0.9 at frequencies below the expected corner frequency (Abercrombie *et al.*, 2017a). 73 EGFs for Sequence 2, 2 for Sequence 5 and 172 for Sequence 9 meet these criteria (Table S1).

Text S3. Resolution and Comparison of the Source Time functions:

To ensure that we have the resolution to detect azimuthal and temporal variation in the earthquake source time functions (STFs) we compare examples of the source time functions for each sequence with a delta function filtered to the same frequency range as the data. Figure S1 compares an earthquake in each sequence affected by the M6 and one 6-8 years later after the post-seismic slip has decreased to the background level. All the source time functions in Sequence 2 are significantly longer than the minimum resolution indicated by the delta function, showing that azimuthal variation is real, and the rupture velocity reasonable, but small details in the STFs, of the duration of the delta function, are not. There is no obvious difference between the two example events. Sequence 5 shows a clear difference in average duration for the earthquakes with time, but the STFs are little longer than the delta function suggesting that azimuthal variation could be lost making any measurements of directivity highly uncertain. Sequence 9 shows a slight difference with time, and STFs that are sufficiently longer than the delta function to resolve rupture direction, but the rupture velocity could be underestimated (Abercrombie *et al.,* 2017b).

Text S4. Stretching and Fitting the Source Time Functions for directivity

In Figures S2 - S4 we show examples of fitting the source time functions of one earthquake from each of the three sequences. The directivity (including dip) and velocity of the earthquakes in Sequence 2 are well constrained (Figure S2). The direction of earthquakes in Sequence 9 (Figure S3) is well constrained, but not the velocity as a range of azimuths are close to the minimum resolvable STF (see Figure S1). The source time functions in Sequence 5 are too short compared to the resolution limit to resolve directivity with confidence (Figure S4).

We also compare the results from fitting using a unilateral line source, a symmetrical bi-lateral line source, or a bilateral line source which extends twice as far in the rupture direction as in the back direction (“2-to-1”, in no case is this resolvable from the unilateral line source). The results of these comparisons are shown in Figure S5, together with those for the EGF selection tests described below. For Sequence 2 (Figure S5a) the unilateral model has a lower misfit than the bilateral model for all events, with consistent direction and rupture velocity. There is more variability for the EGF selections that involve fewer data, but the principle conclusions do not change. For Sequence 5 (Figure S5b) the bilateral or unilateral models have a smaller misfit, indicating that the directivity is poorly resolved. This is expected from the short length of the STFs compared to the resolvable minimum. The earthquakes in Sequence 9 (Figure S5c) all exhibit slightly lower misfit for the unilateral model than the bilateral, though this is less clear for the EGF selections using fewer data. This is consistent with the duration of the STFs lying between those of Sequence 2 and Sequence 5.

Text S5. Tests of Stability and Temporal Resolution with Varying Stations and EGF Selection.

To ensure that our results are not an artefact of the available stations, and selection of EGFs we repeat the analysis frequency and time domain analyses with different EGF choices. Figures 3 and 4 show the initial results which use all available stations and EGFs that meet our strict criteria, regardless of the time of the target event or the EGF. These results include the most data, but different sets of EGFs for each event and station, and no correction for any temporal variation in site or path effects. We therefore perform four tests to investigate whether this is biasing our results, and whether temporal variation in path effects is causing an apparent variation in source process. The results of the directivity modeling for these tests are included in Figure S5.

For our first test, we continue to use all stations, and all available EGFs, but divide the EGFs by time. To compare the time when the region is affected by the M6 earthquake to that when it is relatively healed, we divide both main earthquakes and EGFs into two time intervals: (a) 2004 M6 to 28 September 2005 (1 year), and (b) all other times, see Figures S6 and S7. Note that Sequence 5 has only one EGF in each time period (in May and September 2005), and so the ratios and STFs are less stable. The timing of these EGFs does not fully represent the variation in attenuation and spectral ratios during the first year after the 2004 M6 earthquake.

For our second test, we identify consistent sets of EGFs and stations for all earthquakes in a sequence. We restrict this analysis to the vertical component which contains the best quality data. We individually select the combination of main earthquakes, stations and EGFs to obtain the maximum consistent number for all. For sequence 2 we find 43 EGFs for 7 main events at 5 stations with minimum cross-correlation for at least one mainshock-EGF pair at one station, and a median correlation of 0.7 for all EGFs at all stations for all events. Sequence 9 we find 69 EGFs for 8 mains at 6 stations. For Sequence 5, we only have two EGFs; they are consistent across the majority of stations (6) and events (6) but we do not analyze this sequence in this manner, as the results are little different from those of including all data and less stable.

Our third test (for Sequences 2 and 9) uses the same consistent sets of stations and EGFs but further divides the EGFs into the two time intervals described in the initial test.

Using only the consistent stations and EGFs decreases the number of events and stations we can analyze (Figures 3, 4, S6b and S7b), but we do not observe any systematic differences between these results and those using all data, implying that small variations in EGFs and components between stations and events has a negligible effect.

Using different temporal sets of EGFs (compare Figures 3, 4, to S6a and S7a, and S6b and S7b to S6c and S7c) removes some of the temporal variation in the spectral ratios (bringing the red and black ratios closer to the yellow), implying that a significant part of this variation is caused by attenuation. We try further dividing the sequences into shorter time periods to address the changes in frequency content during the months following the M6 earthquake (compare red to black ratios in Figures), but the numbers of EGFs become too small to obtain any reliable trends. For Sequence 5 the difference between the two EGFs, at 6 and 12 months following the M6, is likely to represent the variation attenuation less well than the larger number of EGFs for Sequences 2 and 9 that better sample the time period immediately after the M6.

We perform one further test of the spectral ratios to investigate the relative effects of temporal variation in source and attenuation. We use the same set of EGFs for all events in each sequence, independent of time (our second test), and then apply the attenuation increase calculated by Kelly *et al*. (2015) to those target earthquakes that occur within a year of the 2004 M6 (Figure S7). We try values of Δt^* = 0.001, 0.0025 to represent the range that they observe throughout the first year, and 0.005 to consider a higher value; Figure S7 shows Δt^*=0.0025. This extra attenuation is able to account for most, if not all, of the temporal variation in frequency content in the ratios. However, we note that these attenuation values were calculated under the assumption that there was no temporal change in source properties during this time period.

Text S6. Corner Frequency and Stress drop as a function of time

For comparison with previous work we fit the spectral ratios to obtain Brune-style corner frequencies and stress drops for the earthquakes we study. We calculate a station-weighted estimate for each event by stacking the ratios from each station. We fit the ratios with:

$$\frac{\dot{M}_1(f)}{\dot{M}_2(f)} = \frac{M_{01}}{M_{02}}\left(\frac{1+\left(f/f_{c1}\right)^{\gamma n}}{1+\left(f/f_{c2}\right)^{\gamma n}}\right)^{\frac{1}{\gamma}} \qquad (1)$$

where f is frequency, f_{c1} and f_{c2} are the corner frequencies of the large and small earthquakes (target and EGF), respectively, M_{01} and M_{02} are the seismic moments of the large and small earthquakes, respectively, n is the high-frequency fall off (we assume n=2), and ⍰ is a constant controlling the shape of the corner (we assume γ=2), Brune 1970; Boatwright 1980. We place constraints on f_{c2} by limiting the stress drop of the stacked EGFs to between 1 and 50 MPa to constrain the solution (Shearer *et al.*, 2019).

To calculate the M_w we perform a spectral decomposition analysis (Allmann and Shearer, 2007) using HRSN recordings of over 5000 earthquakes located along a 80 km stretch of the San Andreas fault centered on SAFOD. We calculate the relative amplitudes of the event terms in the frequency range 2-4 Hz, which are proportional to the seismic moment, and convert them to Mw by assuming that M3 = M_w 3.

We calculate the stress drop ($\Delta\sigma$), following Eshelby (1957)

$$\Delta\sigma = \frac{7M_0}{16}\frac{f_c^3}{k^3\beta^3} \qquad (2)$$

where β is the S wave velocity, and k is determined by the source model, we assume k=0.32 (Kaneko and Shearer, 2014; Abercrombie *et al.*, 2017a). We use $\beta = 3$ km/s for Sequence 2, $\beta = 3.3$ km/s for Sequence 5 and $\beta = 3.2$ km/s for Sequence 9 based on the velocity model at their respective locations (Thurber *et al.*, 2006). We fit the spectral ratios obtained from the four different sets of EGF analyses described above. The results are shown in Figure S8, for all events with well-constrained fits. Most stress drops are in the typical range of ~5-50 MPa, and correspond to rupture dimensions of ~80 m, ~30m and ~50 m for Sequence 2, 5 and 9 earthquakes, respectively. Sequence 2 earthquakes have essentially constant stress drop and corner frequency within resolution. Sequence 5 show a clear decrease in corner frequency and stress drop following the 2004 M6 event, indicating an increase in rupture area, then gradual recovery to pre-2004 levels. The corner frequencies of many earthquakes in this sequence were too high to be resolved within the bandwidth of the data – especially when only subsets of the EGFs were included. From our tests of time-dependent attenuation, it is possible that some, if not all, of this apparent source variation is an artifact of attenuation changes following the 2004 M6 earthquake. Sequence 9 shows a small decrease in corner frequency, following the 2004 M6 earthquake, for time-independent EGF selections, but this disappears when correction for the time varying attenuation is included, suggesting that no resolvable change in source properties occurred.

We show the resulting values in Figure S8. Since we know that the earthquakes in at least two sequences show clear directivity, the simple circular approximation is not ideal but we include the numbers to show that these earthquakes exhibit stress drops well within the typical range, when calculated under the same assumptions. It also enables us to place quantitative constraints on any decrease in stress drop and increase in rupture area (decrease in corner frequency) observed for the earthquakes in Sequence 5 following the M6 2004 earthquake. We include the results from using the different selections of EGFs to show how much variability this source of uncertainty contributes.

References:

Abercrombie, R. E., S. Bannister, J. Ristau, and D. Doser, 2017a. Variability of Earthquake Stress Drop in a subduction setting, the Hikurangi Margin, New Zealand, *Geophys. J. Int*., 208, 306-320, doi:10.1093/gji/ggw393.

Abercrombie, R. E., Poli, P. & Bannister, S. 2017b. Earthquake Directivity, orientation and stress drop within the subducting plate at the Hikurangi margin, New Zealand. *Journal of Geophysical Research: Solid Earth, 122*. https://doi.org/10.1002/2017JB014935

Allmann, B. P., and P. M. Shearer (2007), Spatial and temporal stress drop variations in small earthquakes near Parkfield, California, *J. Geophys. Res.*, **112**, B04305, doi:10.1029/2006JB004395.

Boatwright, J., 1980. A spectral theory for circular seismic sources: simple estimates of source dimension, dynamic stress drop, and radiated seismic energy, *Bull. Seism. Soc. Am. 70*, 1–28.

Brune, J., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, *75*, 4997-5009.

Eshelby, J. D., 1957. The determination of the elastic field of an ellipsoidal inclusion and related problems, *Proc. Roy. Soc. Lond., A, 241*, 376-396.

Lengliné, O., and Marsan, D. (2009), Inferring the coseismic and postseismic stress changes caused by the 2004 M_w = 6 Parkfield earthquake from variations of recurrence times of microearthquakes, *J. Geophys. Res.*, 114, B10303, doi:10.1029/2008JB006118.

McLaskey, G. C., A. M. Thomas, S. D. Glaser, and R. M. Nadeau (2012), Fault healing promotes high-frequency earthquakes in laboratory experiments and on natural faults, *Nature*, **491**, 101– 105, doi:10.1038/nature11512.

Prieto, G. A., R. L. Parker, I F. L. Vernon, 2009. A Fortran 90 library for multitaper spectrum analysis, *Computers and Geosciences*, 35, 1701-1710, doi:10.1016/j.cageo.2008.06.007.

Rubinstein, J. L., and G. C. Beroza (2005), Depth constraints on nonlinear strong ground motion from the 2004 Parkfield earthquake, Geophys. Res. Lett., 32, L14313, doi:10.1029/2005GL023189.

Rubinstein, J. L., Ellsworth, W. L., Chen, K. H., and Uchida, N. (2012), Fixed recurrence and slip models better predict earthquake behavior than the time- and slip-predictable models: 1. Repeating earthquakes, *J. Geophys. Res.*, 117, B02306, doi:10.1029/2011JB008724.

Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006), Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. Seismol. Soc. Am.*, **96**4B), S38– S49.

Waldhauser, F. and D.P. Schaff (2008), Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, NCAeqDD v201112.1 *J. Geophys. Res.*, 113, B08311, doi:10.1029/2007JB005479.

Wu, C., Delorey, A., Brenguier, F., Hadziioannou, C., Daub, E. G., and Johnson, P. (2016), Constraining depth range of S wave velocity decrease after large earthquakes near Parkfield, California*, Geophys. Res. Lett., 43,* 6129– 6136, doi:10.1002/2016GL0691.

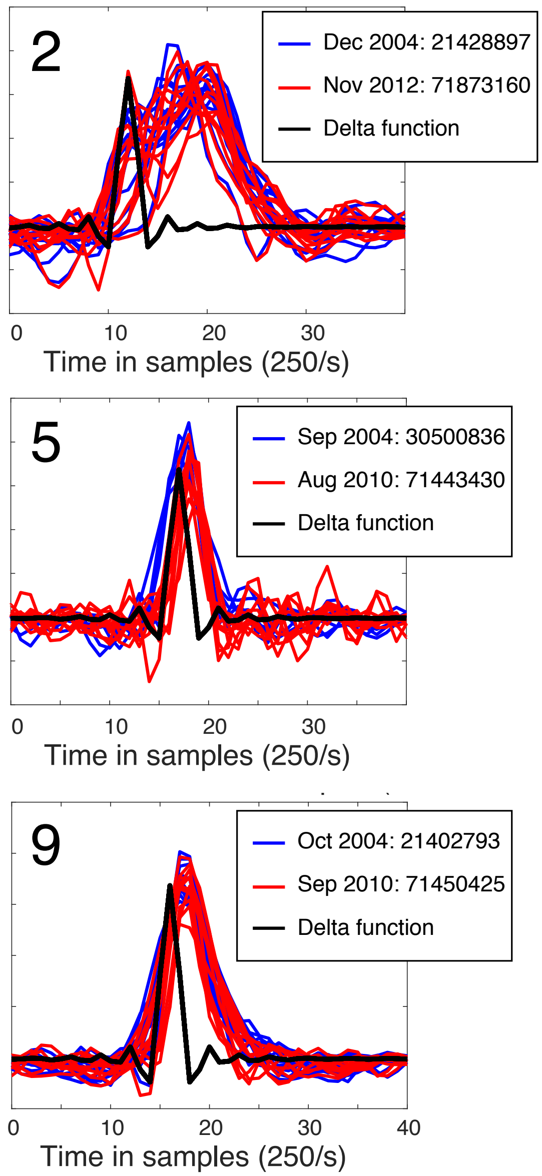


Figure S1. Resolution of source time functions (STFs). For each sequence the STFs at all stations are compared for (1) an event strongly affected by the 28 Sep 2004 M6 earthquake (blue), (2) one 6-8 years later after post-seismic slip has decreased to background level (red) , and the resolvable delta-function – a spike that is sampled and filtered (<80 Hz) in the same manner as the data. Sequence 2 shows no obvious difference between earthquakes at different times, and the STFs are significantly longer than the resolvable delta function showing that azimuthal variation is real, but small details in the STFs, of the duration of the delta function, are not.

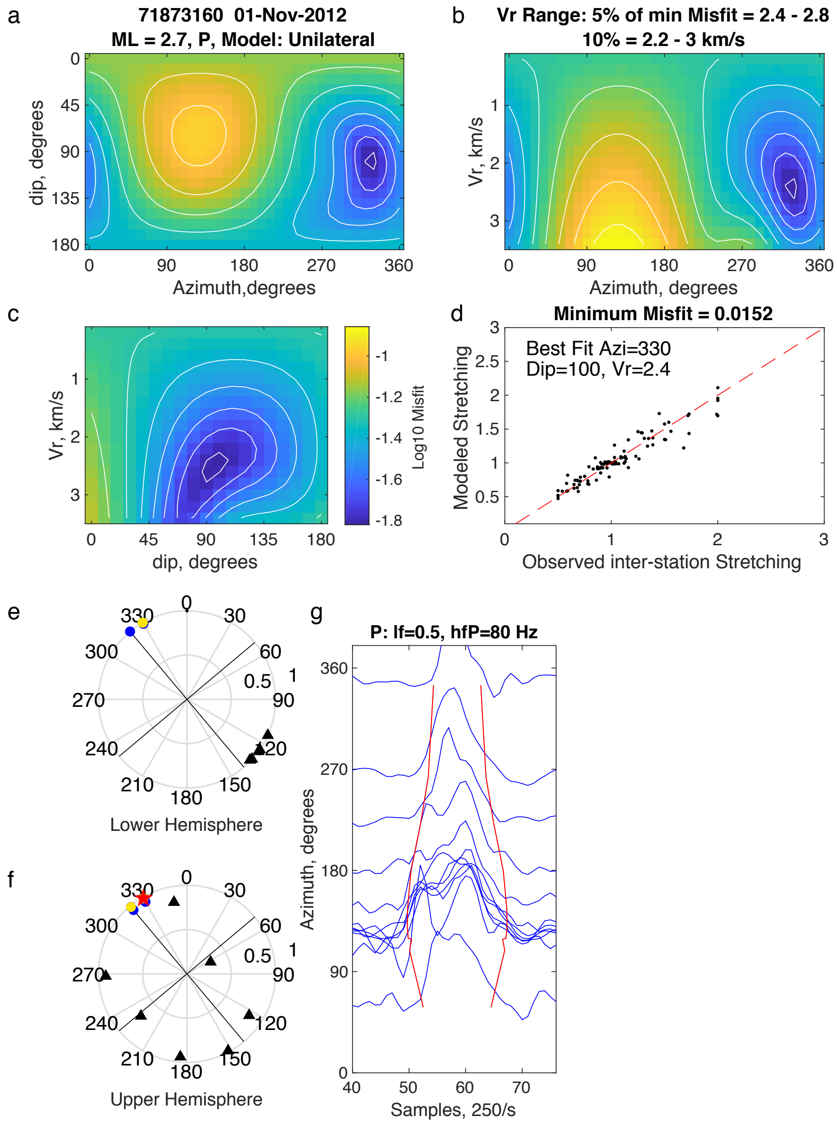


Figure S2. Stretching and Modelling earthquake ID 71873160 in Sequence 2, using all EGFs and stations and assuming a unilateral line source. (a), (b), and (c) the misfit between the azimuthal variation in stretching of the STFs and the unilateral line source, as a function of line azimuth, dip (0 = vertically up, 90 = horizontal and 180 = vertically down), and rupture velocity. (d) relative stretching between each pair of stations from the STFs compared to that predicted by the best fitting line source. The red line shows 1:1 correspondence. (e) the lower and (f) the upper hemisphere focal spheres; a strike-slip mechanism aligned with the San Andreas Fault is shown in both plots, but otherwise they only show upper and lower hemisphere angles, respectively. The black triangles are the stations used, the red star is the best fitting line source (minimum misfit), the yellow region shows the directions with misfit within 5% of the minimum, and the blue the directions within 10% of the minimum. (g) the P wave (blue, cyan if not included) source time functions used in the analysis, with filter frequencies. The red lines indicate how the duration varies for the best fitting line source. They are centered on the STFs, with a mean duration of the approximate duration of the STF estimated by stacking all stations.

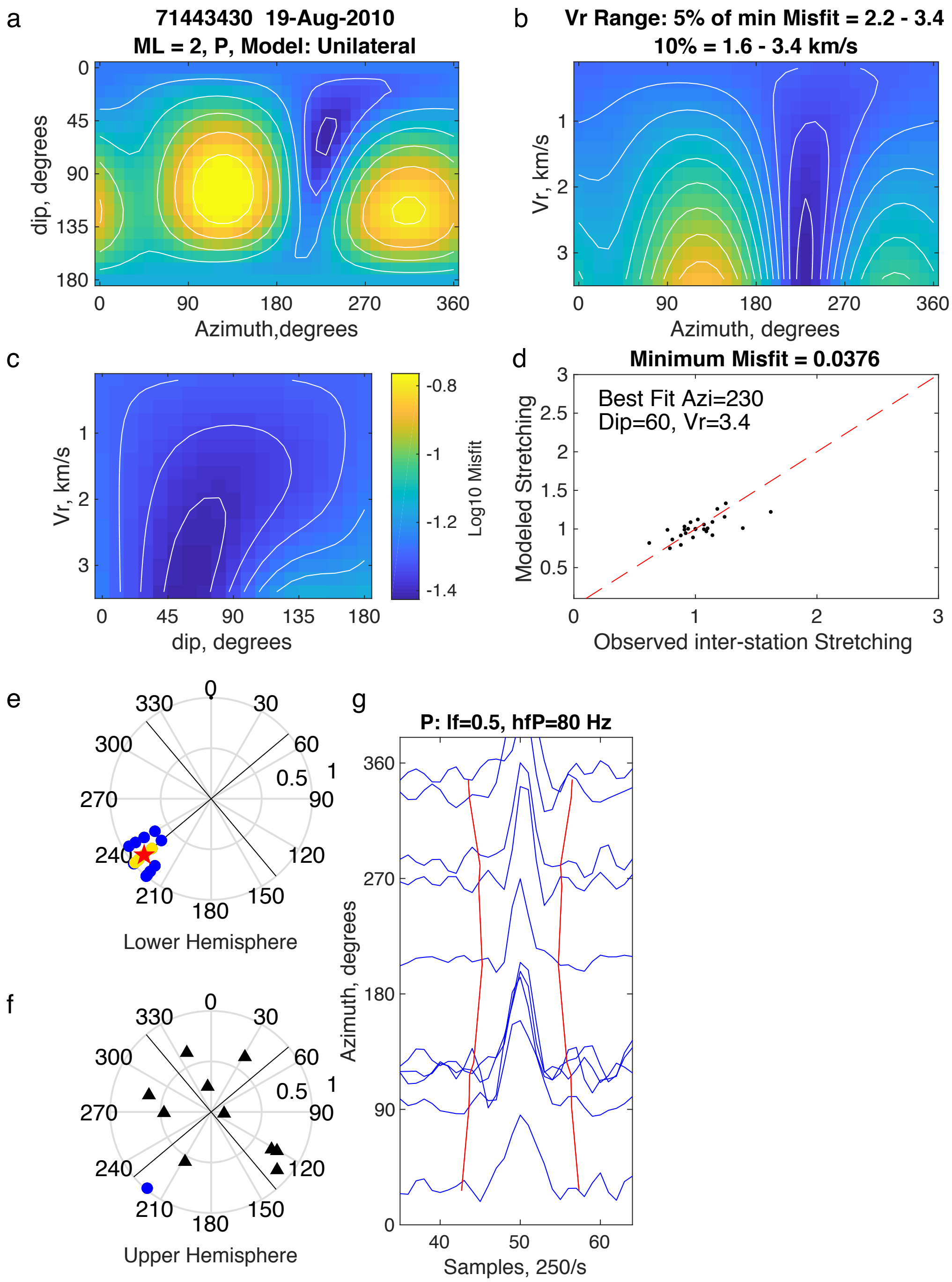
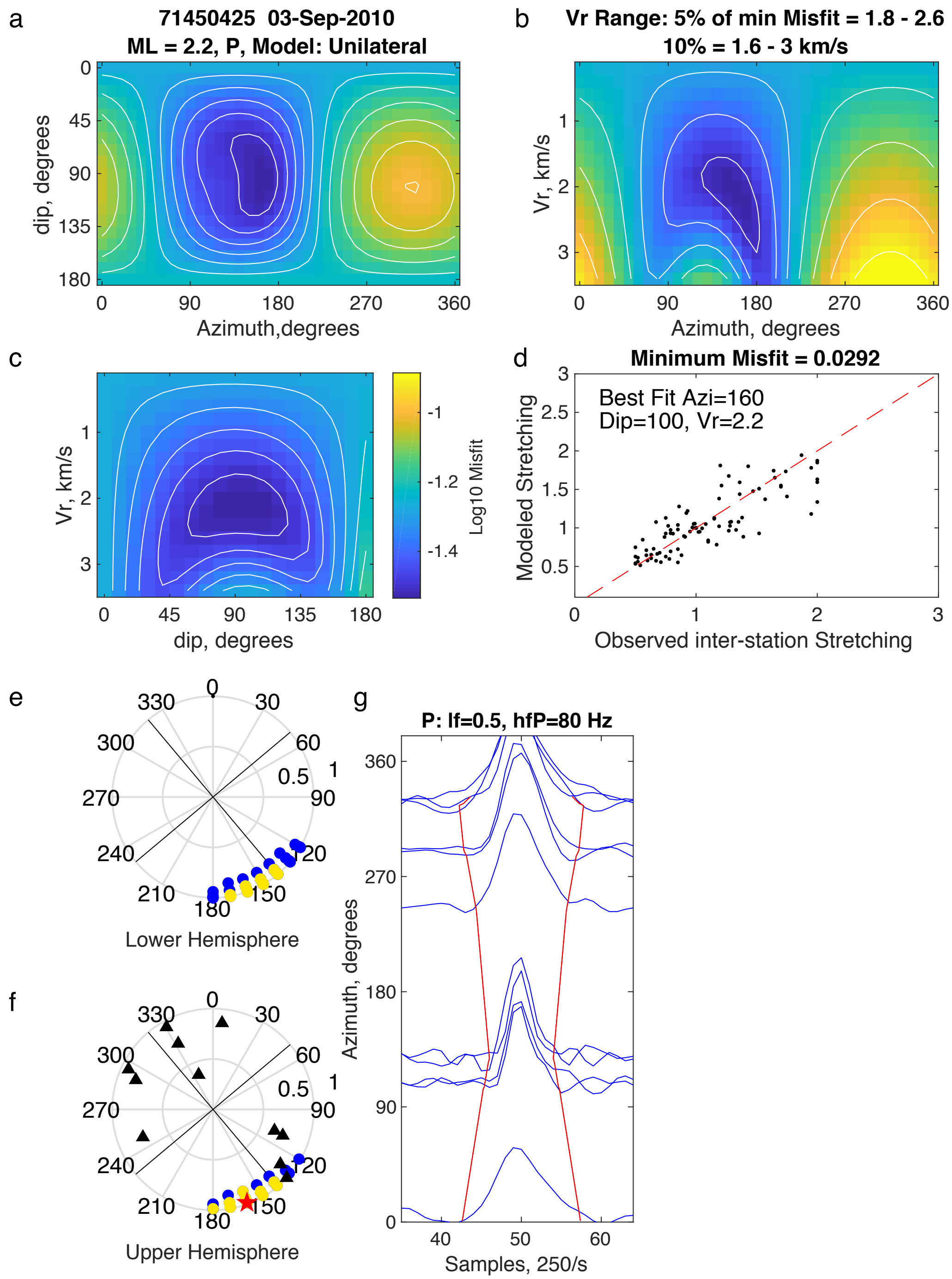


Figure S3. Stretching and Modelling earthquake ID 71443430 in Sequence 5. See Figure S2 for explanation.



282

283 **Figure S4.** Stretching and Modelling earthquake ID 71450425 in Sequence 9. See Figure S2 for
284 explanation

285

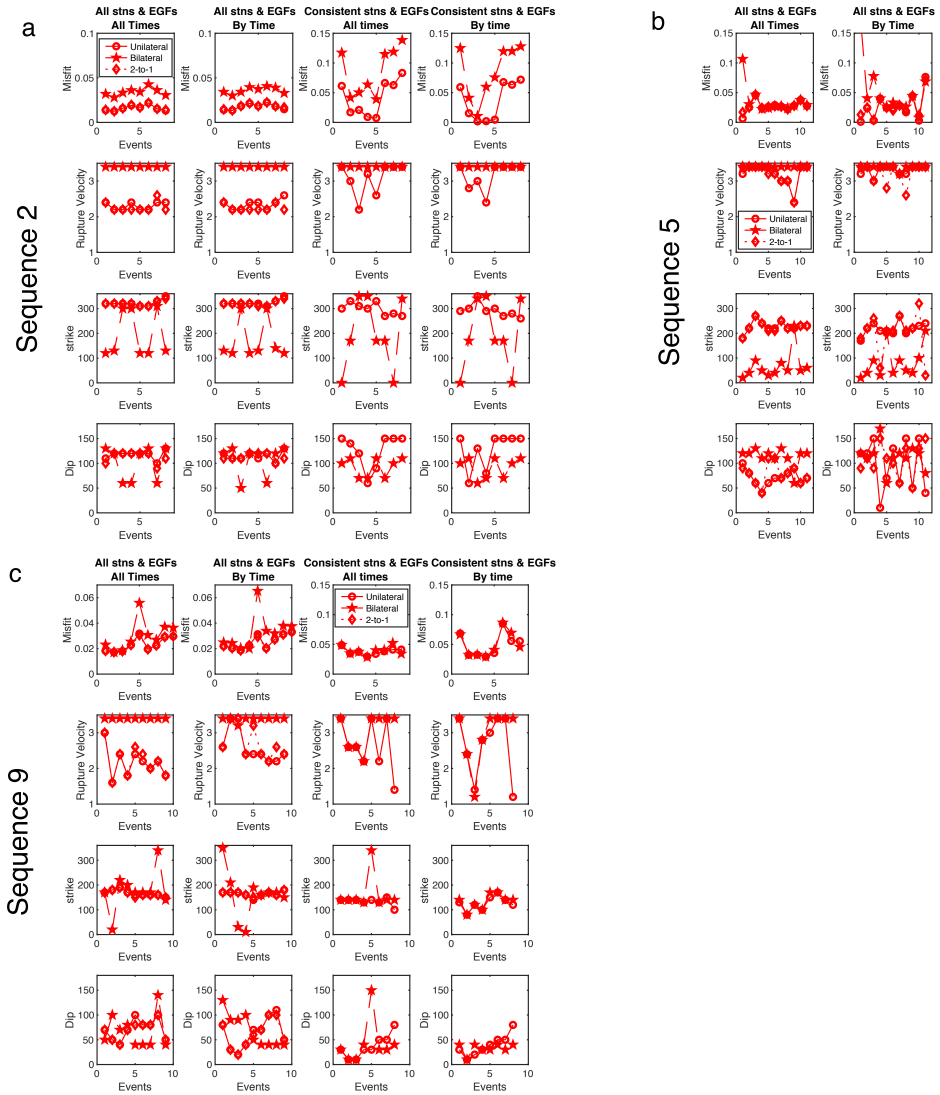


Figure S5. Directivity Measurement and Resolution for Varying EGF selection for (a) Sequence 2, (b) Sequence 5 and (c) Sequence 9. Results are compared for the best fitting model using either a unilateral line source, a symmetrical bi-lateral line source, or a bilateral line source which extends twice as far in the rupture direction as in the back direction ("2-to-1", in no case is this resolvable from the unilateral line source). Each column shows the results for a different set of stations and EGFs. Column 1 is all available stations and EGFs, on all 3 components, for each individual earthquake, regardless of occurrence time. Column 2 is same as 1, but Sequence and EGF earthquakes are divided into two time periods: one covering 28-Sep 2004 to 1 Oct 2005, and the other including both the time before the M6, and after 1 Oct 2005. Column 3 is as Column 1, but using only a consistent set of stations and EGFs – each main event has exactly the same set of EGFs at exactly the same set of stations, all on only the vertical component. Column 4 is as Column 3, but with the EGFs further divided by time as in Column 2. Hence moving from left to right involves decreasing amounts of stations and EGFs.

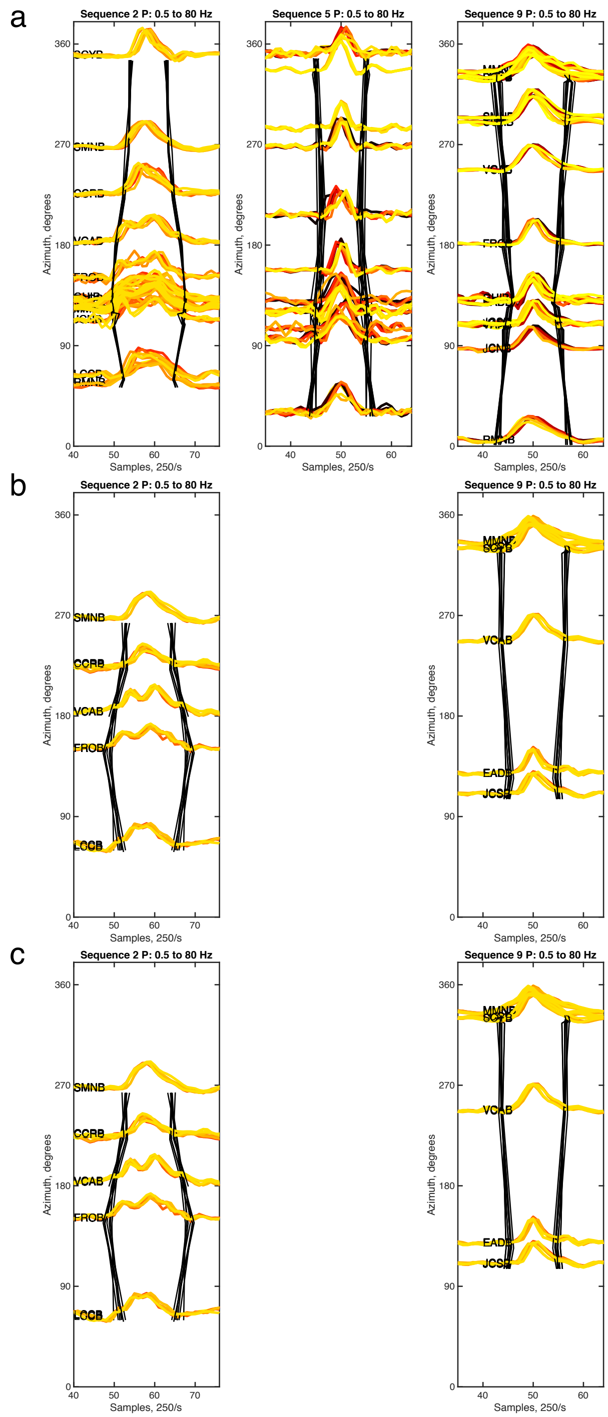


Figure S6. Source Time functions as Figure 3 in paper, for different sets of conditions shown in the columns in Figure S5; Figure 3 is for all stations, all EGFs (Figure S5: column1). (a) all stations, all EGFs, EGFs divided by time (Figure S5: column2), (b) consistent stations and EGFs (Figure S5: column3), (c) consistent stations and EGFs, EGFs divided by time (Figure S5: column 4).

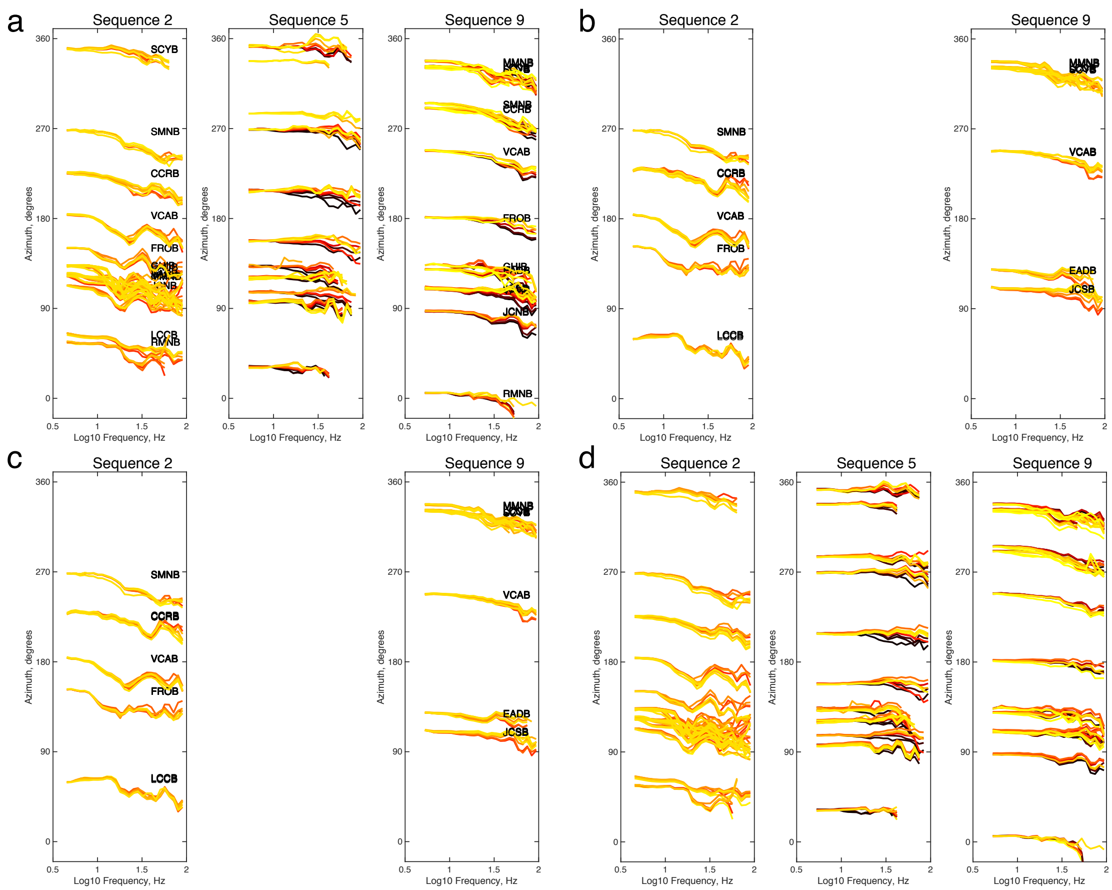


Figure S7. Spectral Ratios as Figure 4 in paper, for different sets of conditions shown in columns in Figure S5; Figure 4 is for all stations, all EGFs (Figure S5: column1). (a) all stations, all EGFs, EGFs divided by time (Figure S5: column2), (b) consistent stations and EGFs (Figure S5: column3), (c) consistent stations and EGFs, EGFs divided by time (Figure S5: column 4). Note consistency, but decreasing quantity of data, and consequent lack of resolution moving from (a) to (c). (d) is for the corrections in Column 1 (all stations, all EGFS: Column 1), but with the addition of the extra attenuation (Δt^*=0.025) calculated by Kelly *et al.* to events within 1 year following M6.

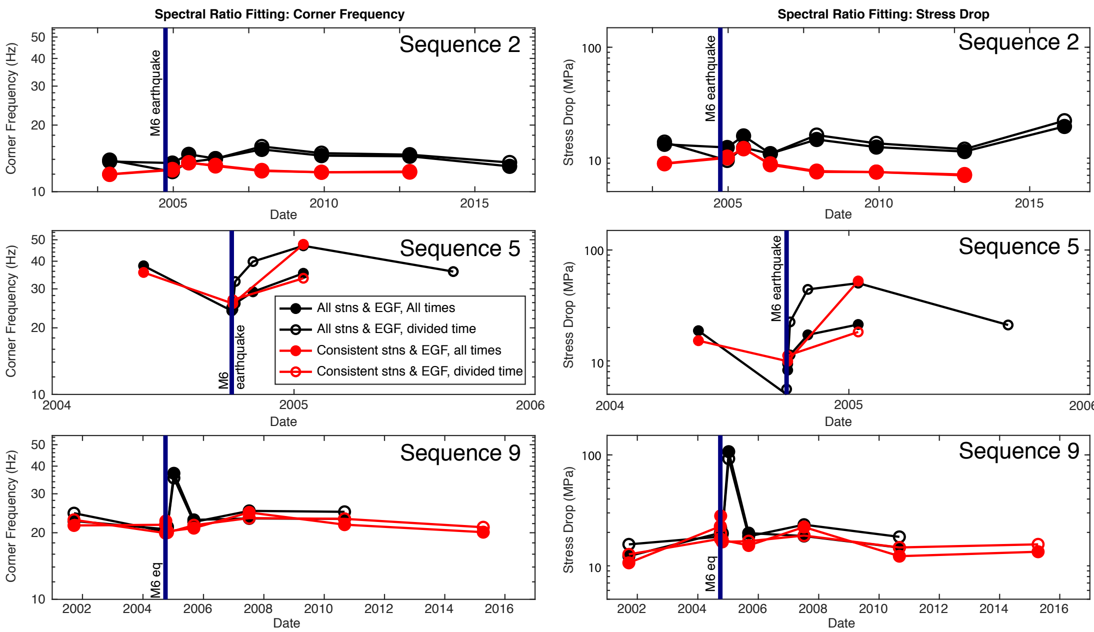


Figure S8. Corner Frequency and Stress drop from spectral ratio fitting. Left hand column: corner frequencies, and Right hand column: stress drop. Top row = Sequence 2, middle row = Sequence 5 (many events have fc too high to resolve – especially after number of ratios included is decreased for consistency), and bottom row = Sequence 9. For each sequence the results for different selection of EGF are shown; variation that is least dependent on selection is the most reliable. Selections correspond to those described in the paper and Figure S5: black filled symbols - All available stations and EGFs are used for each event, regardless of date, black open symbols - All available stations and EGFs but only from the same time period as the respective target earthquake, red solid symbols - consistent EGFs and stations for each target event, from all time period, and red open symbols - consistent stations and EGFs only from the same time period as the respective target earthquake.