A High-Resolution Earthquake Catalog for the 2004 M6 Parkfield Earthquake Sequence using a Matched Filter Technique

Miguel Neves[1,*], Zhigang Peng[1], Guoqing Lin[2]

1 - School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

2 - Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami, Miami, FL, USA

*Corresponding author: mjneves@gatech.edu

Abstract

We present the high-resolution Parkfield Matched filter Relocated (PKD-MR) earthquake catalog for the 2004 M_w 6 Parkfield earthquake sequence in Central California. We use high-quality seismic data recorded by the borehole High-Resolution Seismic Network combined with matched filter detection and relocations from cross-correlation derived differential travel-times. We determine the magnitudes of newly detected events by computing the amplitude ratio between the detections and templates using a principal component fit. The relocated catalog spans from November 6, 2003, to March 28, 2005, and contains 13,914 earthquakes, which is about 3 times the number of events listed in the Northern California Seismic Network catalog. Our results on the seismicity rate changes before the 2004 mainshock do not show clear precursory signals, although we find an increase in the seismic activity in the creeping section of the San Andreas Fault (about ~30 km NW of the mainshock epicenter) in the weeks prior to the mainshock. We also observe a decrease in the *b*-value parameter in the Gutenberg-Richter relationship in the creeping section in the weeks prior to the mainshock. Our results suggest an increased release of stress seismically in the creeping section and a decrease in the aseismic creeping rate. However, *b*-value and seismicity rates remain stable in the Parkfield section where the 2004 mainshock ruptured. This updated catalog can be used to study the evolution of aftershocks and their relations to afterslip following the 2004 Parkfield mainshock, seismicity before the mainshock, and how external stresses interact with the Parkfield section of the San Andreas Fault and the 2004 sequence.

Introduction

The Parkfield segment of the San Andreas Fault (SAF) lies in the transition zone between the mostly seismic southern section and the central creeping section of the SAF (Bakun and Lindh, 1985). With a transitional behavior, deformation in the Parkfield section (Figure 1) is partially accommodated by repeating M6 characteristic earthquakes (Bakun and Lindh, 1985). The September 28, 2004 Mw6.0 Parkfield earthquake was the latest of seven M~6 event in this region since 1857 (Bakun *et al.*, 2005; Langbein *et al.*, 2005). From the analysis of the previous 6 mainshocks until 1966 it was expected that the next rupture would occur no later than 1993 (Bakun and Lindh, 1985). As part of the Parkfield Prediction Experiment (Bakun and Lindh, 1985), a dense network of geophysical instruments, including the 13-station borehole High Resolution Seismic Network (HRSN), were deployed in the area to record in detail the rupture and preparation process of the anticipated mainshock. Thus, the later than expected 2004 Mw6.0 Parkfield earthquake is one of the best recorded earthquake sequences (Bakun *et al.*, 2005).

The prediction of the future Parkfield mainshock rupture (before 1993) was based on the similarities between the 1934 and 1966 rupture (Bakun and Lindh, 1985; Roeloffs and Langbein, 1994). Whereas both the 1934 and 1966 rupture initiated below Middle Mountain (MM), ruptured unilaterally to the SE and were preceded by M5.1 foreshocks (Bakun and Lindh, 1985), the 2004 mainshock initiated in the SE of the Parkfield section near Gold Hill (GH) and ruptured mostly to the NW (Bakun *et al.*, 2005; Langbein *et al.*, 2005). Additionally, no clear precursory signals to the 2004 M6 earthquake have been identified (Bakun *et al.*, 2005; Langbein *et al.*, 2005; Johnston, 2006), similar to the 1901 and 1922 ruptures for which no foreshocks were reported.

Given the wealth of data available, the Parkfield segment and the 2004 earthquake sequence have been the objects of many seismological studies (Michelini and McEvilly, 1991; Eberhart-Phillips and Michael, 1993; Waldhauser *et al.*, 2004; Peng *et al.*, 2006; Thurber *et al.*, 2006; Custódio and Archuleta, 2007; Nadeau and Guilhem, 2009; Peng and Zhao, 2009; Shelly *et al.*, 2011; Meng *et al.*, 2013; Delorey *et al.*, 2017; Lin, 2018; Perrin *et al.*, 2019; Lin *et al.*, 2022). For example, Thurber *et al.* (2006) derived a 3D compressional wavespeed model using double-difference tomography and relocated more than 20 years of seismicity at Parkfield, including the 2004 aftershock sequence. Their results show a mostly planar fault at Parkfield, following the Southwest Fracture Zone (SWFZ), rather than the main SAF. Lin (2018) improved earthquake locations using source-specific station term and waveform cross-correlation differential travel-time differences to constrain both absolute and relative locations for earthquakes between 2000 and 2018. Recently, Perrin *et al.* (2019) constrained absolute and relative locations using Thurber *et al.* (2006) 3D model and waveform cross-correlation differential travel-time differences for earthquakes since 1966, and they identified a fault plane being twisted between MM and GH, likely due to the long-term effects of a strong asperity surrounded by a week creeping region. In contrast, Lomax and Savvaidis (2022) apply relocation to earthquakes in the Parkfield segment using source-specific station travel-time corrections combined with waveform coherence, and found that their relocated earthquakes follow a single near-vertical fault surface that is planer than observed in previous studies.

However, standard earthquake catalogs are known to be incomplete (Kagan, 2004; Peng *et al.*, 2006; Enescu *et al.*, 2007), especially in the early times after a large mainshock due to the increased noise levels from coda waves of the mainshock and large aftershocks. Peng and Zhao (2009) used a matched filter technique (MFT) to detect additional early aftershocks, and found 11 times more events than those reported in the Northern California Seismic Network (NCSN) catalog in the 3 days following the mainshock. The detailed catalog allowed them to observe an along-strike and down-dip expansion of aftershocks, likely driven by afterslip (Barbot *et al.*, 2009; Jiang *et al.*, 2021). Peng and Zhao (2009) simply assumed the same location of the template events for the detected events. Hence, the resulting locations for the newly detected events were not precise. Also using data from the Parkfield segment, Meng *et al.* (2013) performed a matched filter detection in the month before and after the neighboring M6.5 San Simeon earthquake in 2003 and found that the imposed static stress changes by the 2003 mainshock led to a slight decrease of seismicity in the creeping section, and a slight increase in the seismicity rate around the Parkfield section. This pattern was similar to the changes in low-frequency earthquakes at larger depths in both sections following the San Simeon mainshock (Shelly and Johnson, 2011).

To our knowledge there is no long-term MFT-detected earthquake catalog in the periods preceding and following the 2004 event with precise relocation and magnitude estimations. Building upon the work of Peng and Zhao (2009) and Meng *et al.* (2013), here we present a high-resolution earthquake catalog from a systematic MFT detection using the recorded HRSN data for the months before and after the 2004 mainshock, which we designate PKD-MR catalog (for Parkfield Matched filter Relocated earthquake catalog). In the next sections, we describe detailed procedures on how we apply the MFT to detect new events, followed by magnitude calibrations and relocations. We then present initial observations of the aftershock sequence and an analysis of the seismic rate changes and *b*-value variations in the months prior to the mainshock.

Earthquake catalog compilation

Matched filter detection

Matched filter technique, also called template matching detection, is a detection approach (Gibbons and Ringdal, 2006; Shelly *et al.*, 2007; Peng and Zhao, 2009; Yang *et al.*, 2009; Ross *et al.*, 2019) that allows the detection of earthquakes missed by traditional energy-based detection methods (Allen, 1978). Here we perform a network-wide MFT detection using the 13 HRSN stations (Figure 1) following a similar procedure to Meng *et al.* (2013). The detection period is from 6 November 2003, 46 days before the neighboring M6.5 San Simeon earthquake (same starting date in Meng *et al.*, 2013 study), to 28 March 2005, 6 months after the 2004 M6 Parkfield mainshock. The BP channel data (20 sample/s) were used. We first apply a 2–8 Hz two-way fourth-order band-pass filter to the seismic data. To compile our template catalog, we use two relocated catalogs for Northern California. We use Lin (2018) catalog for Parkfield, which includes earthquakes along the SAF as our main template catalog. To include events located off the main SAF and templates from neighboring regions such as the San Simeon aftershock zone, we add earthquakes in the NCSN relocated catalog (Waldhauser and Schaff, 2008) with the Caltech-USGS Seismic Processing (CUSP) event ids not included in Lin (2018) (Figure 1a). We select catalogued earthquakes between latitude 35.3º N and 36.4º N and longitude 120.0º W and 121.0º W. We consider earthquakes located less than 20 km away from the SAF as local templates (Figure 1b).

We retrieve available P and S phase arrival times from the Northern California Earthquake Data Center (NCEDC) for the templates or attempt to pick arrivals with PSIRPicker (Li and Peng, 2016) for events and stations with no reported arrivals. The PSIRPicker uses a general 1D velocity model (Table S1) to predict the initial P or S arrivals, followed by a simple short-term-average/long-term average (STA/LTA) detector (Allen, 1978) to pick the accurate arrivals. The majority of our S phase picks (99%) are from the PSIRPicker results since only 6 S phase picks could be retrieved from the NCSN catalog for the HRSN stations. We then compute the signal-to-noise ratios (SNR) for each event at each channel (P phases on vertical channels and S phases in the horizontal channels) by comparing the sum of squared amplitudes in a 5 second window before the phase arrival with a 5 second window after the arrival. Only earthquakes with at least SNR>5 arrivals at 12 channels were included in the template catalog. This results in a catalog with a total of 10,184 template earthquakes between 6 November 2003 and 28 March 2005 (Figure 1a).

Detection was performed using waveforms from −1 to 4 seconds around the P-phase arrival in the vertical channels and around the S-phase arrival in the horizontal channels. Using only channels with SNR>5, we cross-correlate each template waveform at each channel with the corresponding channel's daily waveform (86460 s long). We then shift daily cross-correlation functions (CCFs) for each channel to the origin time of the correlating template and stack to obtain a network averaged daily CCF. Peaks above twelve times the median absolute deviation

(MAD, Figure S1) value of the averaged CCF were selected as detections (Ross *et al.*, 2019; Zhai *et al.*, 2021). Since one event can be detected by more than one template, we look for detections with overlapping windows in any of the channels used for detection. In these cases, we keep only the detection of the template with the highest cross-correlation coefficient (CCC) as in Meng *et al.* (2013) (i.e., the best detecting template). The newly detected events are initially assigned the hypocentral location of the best detecting template.

For the periods of May 13 and 14, 2004, December 29, 2004, to January 12, 2005, and February 20 to 22, 2005, the majority of the HRSN network stations do not have complete daily records. In these periods we perform detections using the channels with complete 24-hour waveforms and include seismic stations from the NC network in a 0.5º radius area from the 2004 mainshock location.

We include templates distanced up to about 50 kilometers from the region of interest to remove false detections of distant events as local detections, for example from the M6.5 San Simeon earthquake area (e.g., Meng *et al.*, 2013). The use of separate template P and S window allows detection of events with slight variations in S-P times, but it is also prone to false detections from local templates that correlate strongly with impulsive P or S arrivals of distant large earthquakes, contaminating our detection catalog. Using distant events as templates we can identify these false detections by removing events for which the best detecting template is located more than 20 km away of the SAF.

This results in a total of 115,263 detections with 28,107 events being best associated with local templates. These 28,107 detections formed our local detection catalog (Supplementary Dataset S1), which was used for subsequent relocation. In comparison, 4,393 events are listed in our template catalog in the same space-time windows (Supplementary Dataset S2).

Earthquake relocation

We constrain the relative locations of the new detections in our catalog using cross-correlation derived differential travel-times (e.g. Waldhauser, 2000; Waldhauser and Schaff, 2008; Lin, 2018). We use the XCORLOC package (Lin, 2018) to perform differential travel-time relocation. The first step in our relocation process was to improve the azimuthal coverage. We retrieve all waveforms from stations in all available seismic networks in the region (Figure 1a). These include mostly stations from the NCSN (network code NC) but also from the Southern California Seismic Network (network code CI), the Central Coast Seismic Network (network code PG), as well as the HRSN (network code BP) that were used in the initial MFT detection.

As mentioned before, we assume that the newly detected events have the same initial hypocenter as the best detecting template and assign the corresponding phase arrival times to the new detections. To calculate differential travel-times, we cross-correlate each event with all possible pairs in a 5 km radius. We use –0.5 to 1.5 seconds around the P-phase arrival for the vertical channels, truncating if S phase is included in the window (Shelly *et al.*, 2013), and –0.5 to 2.0 seconds around the S-phase arrival for the horizontal channels. In the case of the S phase, we keep the measurement of the horizontal channel with the highest CCC. To ensure millisecond precision we retrieve the waveforms from available channels with the highest sampling rate (e.g., the DP channel for HRSN with 250 sample/s) and use a spline interpolation technique to 1 ms in the samples around peak cross-correlation. Similar to the detection stage, cross-correlation is performed using 2–8 Hz band-pass filtered data. We use only differential times when the measured CCCs are greater than 0.7. Each event pair must have at least 8 differential time measurements and an event needs to pair with at least one template event in order to be considered for relocation. We then input the resulting differential travel-times into XCORLOC using the velocity model for Parkfield as reported in the NCSN catalog (Oppenheimer *et al.*, 1993). We perform relocation by keeping template events location fixed. We estimate uncertainties as in Lin (2020) with a bootstrap method by repeating 15 times (Figures S2 and S3). Our final catalog (PKD-MR catalog, Supplementary Dataset S3) includes 13,914 earthquakes (Figure 2) about 3.2 times the number of earthquakes in our local templates catalog.

Magnitude Calibration

To estimate the magnitudes of the events in our new catalog, we follow a similar procedure to previous matched filter studies (Peng and Zhao, 2009; Meng *et al.*, 2013; Shelly *et al.*, 2016; Shelly, 2020; Yao *et al.*, 2021). We estimate the magnitude of the new events (M_{New}) by comparing its amplitude with the template magnitude ($M_{Template}$) and amplitude:

$$M_{New} = M_{Template} + c \cdot log_{10}\alpha \qquad (1)$$

where *a* corresponds to the amplitude ratio between the newly detected events and the detecting template and *c* to a calibration factor. Different approaches have been used to determine *a*, such as peak amplitudes ratio (Peng and Zhao, 2009), least-squares fit (Gibbons and Ringdal, 2006) or principal component fit (Shelly *et al.*, 2016). Peak amplitude ratio considers only one data point and can be easily affected by noise. To estimate the magnitude of the newly detected events, we follow a similar method proposed by Shelly *et al.* (2016) and Shelly and Hardebeck (2019) to extend the duration magnitude M_d scaling of the templates catalog to the newly detected events by a principal component fit.

Specifically, we calibrate magnitudes for the same pairs used in the relocation procedure. That is, we compare the signals of the newly detected events with each template that pairs with at least 8 CCCs above 0.7. The amplitude ratio for each measurement is then calculated using a principal component fit using the −0.5 to 1.5 s time window for the vertical channels and −0.5 to 2 s for the horizontal channels. The amplitude ratio for each pair corresponds to the median of the amplitude ratio measurements at each channel. To avoid biases resulting from comparing different magnitude types, we limit our estimates using only templates with reported M_d. To estimate the calibration factor, we compare the amplitude ratios of the templates and their catalogued magnitude differences (Shelly *et al.*, 2016). We obtain a value of 0.828 for c (Figure S4). The final magnitude estimate for a new detection corresponds to the median of the magnitude estimates obtained from each detection-template pair (Figure 2c).

General properties of the new catalog and potential limitations

The PKD-MR earthquake catalog has 13,914 earthquakes (Figure 2). The first aftershock listed in our relocated catalog was detected 99 s after the mainshock located about 15.8 km NW of the mainshock epicenter, with a magnitude of 2.46 (Figure 3). Matched filter detection in our chosen frequency band (2-8 Hz) fails to detect early aftershocks in the first minutes after the mainshock due to its high coda wave energy (Figure S5). Still, this is a significant improvement to the NCSN catalog, in which the first catalogued aftershock is detected 222 s after the mainshock.

As noted by previous studies, large magnitude earthquakes correlate poorly with small- and moderate-size earthquakes (e.g. Lin, 2020; Shelly, 2020). Hence, large earthquakes in our catalog such as the M6 mainshock and the two largest M5 aftershocks are not relocated. For these earthquakes we add the locations listed at Lin (2018) catalog to the new catalog. Therefore, caution should be taken when comparing the locations of the larger earthquakes with the relative locations of the other events.

We show a comparison of the PKD-MR catalog with the catalogs from Lin (2018), Thurber *et al.* (2006) and Waldhauser and Schaff (2008) in Figure 4 (see also Supplementary Movie S1). Locations of earthquakes in the new catalog are similar to the previously relocated catalogs in the region delineating the same seismic structures but with additional events (Figure 4). Here we observe that most events are located below the SWFZ, and shallow events form a flower-type structure, which were also reported before by Thurber et al. (2006) (Figure 4b). Depths in the new catalog are in-line with those observed in Lin (2018) catalog that we use to build our local template catalog with average depth decreasing 14 meters after relocation (Figure S6). When comparing with Thurber *et al.* (2006) and Waldhauser and Schaff (2008), we observe slight variations in depth, in particular the deepest cluster at about 20 km NW of the mainshock epicenter along-strike extends between 12 and 13 km in our new catalog and Lin (2018) catalog. In comparison, events in Thurber *et al.* (2006) extend from about 12.5 to 14 km and even deeper depths between 13.5 to 14.5 km in Waldhauser and Schaff (2008) catalog (Figure 4c), which are likely due to the different velocity models used in relocation process.

The PKD-MR catalog includes 10,280 events for which we were able to estimate the magnitude. The remaining 3,634 events in the catalog have no estimated magnitude since they correspond to templates with no magnitude reported at the NCSN catalog or detections that correlate only with these templates. Note that we allow template magnitudes to change when they correlate with other templates. We observe on average a -0.04 change in the template magnitudes (Figure S7), therefore magnitudes can be different than those reported in the NCSN and other relocated catalogs.

It is widely observed that earthquake magnitudes follow the power-law distribution, also known as the Gutenberg-Richter or GR relationship (Gutenberg and Richter, 1944; Ishimoto and Iida, 1939) in the form of:

$$log(N) = a - bM \tag{2}$$

where N corresponds to the cumulative number of earthquakes with a magnitude equal to and larger than M. a and b are constants, where a corresponds to the expected number of earthquakes with magnitudes larger than or equal to 0, and b indicates the relative number of larger magnitude earthquakes versus smaller magnitude earthquakes in the distribution (i.e., the *b*-value).

As mentioned before, earthquake catalogs are inherently incomplete due to for example limitation in seismometer sensitivity, station coverage or overlapping arrival of seismic waves

(Kagan, 2004; Peng *et al.*, 2006; Enescu *et al.*, 2007; Peng and Zhao, 2009). Hence, a magnitude threshold is usually defined, above which the magnitudes follow the GR distribution, and the catalog is assumed to be complete. Such threshold is called magnitude of completeness (M_C). M_C has been shown to vary significantly spatially and with time (e.g., Wiemer and Wyss, 2000; Kagan, 2004) and depends on variables such as number of samples or the estimation method (Woessner and Wiemer, 2005). The choice of M_C can have a significant impact on the statistical properties derived from the catalog (e.g. Woessner, 2005; Herrmann and Marzocchi, 2021; van der Elst, 2021). We estimate the new completeness following the Entire Magnitude Range (EMR) method of Woessner and Wiemer (2005), which considers the entire data with a two-part model, a power-law model above M_C and a normal cumulative distribution below M_C.

Binning the catalog in 0.1 magnitude intervals, we estimate M_C of 0.6 for the new catalog (Figure 5a), an apparent improvement to the template catalog for which we estimate a completeness of 1.2. Caution is necessary when interpreting the M_C parameter for the PKD-MR catalog (Herrmann and Marzocchi, 2021). We observe a clear change in the GR distribution trend above and below magnitude 1.2 (Figure 5a). This is also reflected in our *b*-value estimates. Using the maximum likelihood estimator, we obtain a *b*-value of 0.65 for the new catalog and of 0.95 for the template catalog (Figure 5a). This is because the new catalog is not actually complete below 1.2 due to the limitations of the MFT method (e.g., improper mixing of different magnitude scales, spatio-temporal varying incompleteness) (Herrmann and Marzocchi, 2021). In addition, MFT only detects events similar to the templates within certain distant ranges, and with our detection parameters we do not allow detection of events with significant overlaps. Finally, even though the MFT is more robust than traditional detection methods, variations in background noise levels still impact its detection performance. We therefore use Apparent Magnitude of Completeness (AM_C) to refer to estimates using the PKD-MR catalog. As expected, M_C estimates vary with time and the same occurs with AM_C estimates with larger AM_C in the period right after the mainshock, when noise levels are higher due to the mainshock coda waves, and during the periods when several of the HRSN stations were down. In these later periods our AM_C estimates do not show significant improvements, even though we use the NC stations to perform additional detections.

Recently, van der Elst (2021) proposed a new *b*-value estimator insensitive to variations of catalog completeness with time. This *b*-positive estimator (b^+) calculates *b*-value considering positive magnitude differences ($\overline{m'}$) of consecutive earthquakes above a minimum magnitude difference (M'_C):

$$b^+ = [\ln(10)\,(\overline{m'} - M'_C)]^{-1}\,,\ \overline{m'} \geq M'_C \qquad (3)$$

We use the *b*-positive estimator with the magnitude of completeness correction, considering only earthquakes with magnitudes above our estimated AM_C. Using this method we estimate a *b*-value of 0.87 for both the template and PKD-MR catalogs (Figure 5b). These values are more in-line with previous *b*-value estimates of 0.92 at Parkfield (Schorlemmer *et al.*, 2004) and are stable with different choices of the minimum magnitude difference (Figure 5d). This result can be explained with the same argument that van der Elst (2021) proposed for the use of the positive magnitude differences. The occurrence of a larger earthquake limits the detection capability of a subsequent smaller earthquake, which also applies in the MFT detection. We therefore suggest the use of the *b*-positive estimator for statistical analysis with our PKD-MR catalog (and other catalog built with MFT) and caution with the magnitude of completeness parameters.

Aftershock expansion

We present here a general description of the earthquake sequence using the PKD-MR catalog. Peng and Zhao (2009) previously analyzed the early aftershock sequence using a matched filter improved catalog for the first three days after the mainshock, but their analysis did not include relocation of the new detections. In their study, Peng and Zhao (2009) observed a linear migration of early-aftershocks with logarithmic-time along-strike and down-dip, and they interpreted as an expansion driven by afterslip (Kato, 2007).

Here we extend the observation of aftershocks to 6 months after the mainshock. As expected, we also observe a migration with logarithmic time along-strike of the early aftershocks (Figure 3b). A similar expansion is observed in terms of depth (Figure S8). Along-strike this expansion is larger to the NW (towards the creeping section) and more limited towards the SE (towards locked section). The logarithmic-time expansion to the NW extends to about 38 km of the mainshock epicenter during the first 14 days (Figure 3). Further than this point, seismic activity in the creeping section appears to resume during the first week after the mainshock with similar rates to those observed before the mainshock (Figure S9).

In terms of depth (Figure S8), the first detected aftershock is located at 5 km depth and after this aftershock a linear expansion with logarithmic-time is observed both in the down-dip direction and up-dip direction. This depth logarithmic-time expansion occurs during the first week after the mainshock.

Precursory activity

Seismic or aseismic activity in the period leading to large mainshocks has been the object of several studies for decades (Roeloffs, 2006; Pritchard *et al.*, 2020; Kato and Ben-Zion, 2021 and references therein) with the goal of understanding the nucleation process and identify possible warning signals of the impeding earthquake. Precursory signals and earthquake initiation is usually explained in the framework of two models): the cascade model which explains mainshock triggering as a result of stress perturbations imposed by a sequence of preceding earthquakes, also known as foreshocks (Olson and Allen, 2005; Mignan, 2014); the pre-slip model that considers rupture initiates due to an aseismic process such as slow-slip that leads to the rupture of surrounding asperities (Dieterich, 1978; Mignan, 2014). Other proposed frameworks for precursory activity include progressive localization (Ben-Zion and Zaliapin, 2020) in which deformation along a distributed region progressively concentrates to primary slip zones culminating in large earthquakes, and seismic quiescence (Mogi, 1969; Wyss and Habermann, 1988) when a significant decrease in seismicity is observed prior to a mainshock. Precursory activity is still an open question in the Parkfield earthquakes. Although previous mainshocks were preceded by moderate foreshocks, no foreshocks have been found in the 2004 mainshock (Bakun *et al.*, 2005). Nonetheless, possible evidence of precursory signals has been identified retrospectively (Nadeau and Guilhem, 2009; Chun *et al.*, 2010; Shelly and Hardebeck, 2019). For example, Nadeau and Guilhem (2009) identified an unusual deep tremor episode three weeks before the 2004 mainshock, based on the envelope function of the HRSN recordings. Using low-frequency earthquakes as templates, Shelly (2009, 2017) found elevated tremor rates and southward tremor migration in the 3 months before the 2004 Parkfield mainshock. However, most tremor events occurred in Cholame, which are at least 20 km away from the initiation point of the Parkfield mainshock. Chun *et al.* (2010) observed a rise in P-wave attenuation 18 months before the mainshock. However, based on cross-correlation of ambient seismic noises, no clear change in seismic velocity was observed before the mainshock, except an abrupt velocity reduction caused by the nearby 2003 M6.5 San Simeon earthquake (Brenguier *et al.*, 2008; Zhao *et al.*, 2010). Ben-Zion and Zaliapin (2020) examined the localization processes of earthquakes in the Parkfield segment and observe a delocalization prior to the 2004 mainshock, which they suggest can be the result of increasing stress around the Parkfield asperity.

Using the PKD-MR catalog, we identify no clear signal of precursory micro earthquake activity (Figure 2). No migration pattern either towards or away from the epicenter is identified (Figure 2a and Figure 3). We use β-statistics to identify significant seismicity rate changes with time (Matthews and Reasenberg, 1988; Aron and Hardebeck, 2009). β-statistics compares the number of observed earthquakes in a period to the number of expected earthquakes in that period based on the observations of a background period. We estimate β using two adjacent moving time windows of 30 days (Figure 6) moving forward by one day and consider a significant rate change when $|\beta|>2$. Each measurement corresponds to the last point in the window, meaning only earthquakes before the measurement are considered. We consider only events distanced less than 5 km off the SAF for this analysis. Results are similar when considering different magnitude and distance thresholds and time windows (Figures S10 to S14). Additionally, we estimate the coefficient of variation of the interevent times and the ratio of maximum seismic moment to total seismic moment following a similar method to Cabrera *et al. (*2022). These parameters give information on the temporal clustering of seismicity and how moment is released during the seismic sequence, respectively (Figure S15 and S16). We use a 150-event moving window that moves forward by one event including events above AM_C.

For the entire Parkfield segment (Figure 1b), we observe two periods that show significant changes that appear to be related to seismic activity at the creeping section (Figure 6). Following the neighboring M6.5 San Simeon earthquake, there is a significant decrease in β in the creeping section and a significant increase in the Parkfield section similar to the observations by Meng *et al.* (2013). Prior to the Parkfield M6 mainshock, we observe a steady increase of the seismicity rate in the creeping section starting 49 days before the mainshock that becomes significant 17 days prior to the mainshock. In the 48 hours before the mainshock we identify a jump in β estimates to 7.1, suggesting a sudden increase of activity in the creeping section prior to the mainshock. The β increase in the days before the mainshock coincides with the peak of ratio of maximum seismic moment to total seismic moment in the creeping section (Figure S15). However, no significant changes are observed in the Parkfield section where the mainshock ruptured in the same period (Figures 3, 6, S15 and S17). We also observe no significant variations in the temporal clustering of events which shows a gradual decrease in the clustering of events following the occurrence of the neighboring M6.5 San Simeon earthquake (Figure S16).

Temporal variations of *b*-value

We also analyze the temporal variations of the *b*-value parameter of the GR distribution (Figure 7). The actual significance of the *b*-value in the GR distribution is still debated (e.g. Marzocchi *et al.*, 2019), but recent studies observe that the *b*-value is inversely related to differential stress both on laboratory settings (Amitrano, 2003) and field (e.g. Schorlemmer and Wiemer, 2005; Scholz, 2015). Tormann *et al.* (2013) analyzed the variations of *b*-values at Parkfield and connected it to surface creep rate variations with decreasing b-values correlating with decreasing creep and increasing loading stresses on the fault. More recently, Gulia and Wiemer (2019) studied 58 aftershock sequences and observed the *b*-value can be an indicator if a large earthquake is the mainshock or a foreshock. Nanjo (2020) also observed that for the 2019 Ridgecrest earthquake sequence the locations of the larger M6.4 foreshock and the M7.1 mainshock could be retrospectively identified by analyzing *b*-values in the area.

To estimate the time variations of *b*-value, we consider the methodology and findings of previous studies that perform similar time variation analysis such as Gulia and Wiemer (2019), Dascher-Cousineau *et al.* (2020) and van der Elst (2021). We also consider the findings by Marzocchi *et al.* (2019) on the best practices to estimate *b*-values. We analyze all events in the PKD-MR catalog with estimated magnitudes and distances less than 5 km of the SAF strike. We also test different distance values between 2 and 10 km (Figures S18 and S19). We define a moving window of 150 events that moves forward by one event. We then estimate the AM_C for the window using the EMR method. For event windows with at least 50 events above AM_C we estimate the *b*-value using the *b*-positive estimator (van der Elst, 2021). A *b*-positive estimate is deemed robust if the distributions had more than 3 magnitude bins with events and passes a Lilliefors test (a = 0.05) to ensure exponentiality of the magnitude differences distribution (Lilliefors, 1969; Marzocchi *et al.*, 2019). We further identify periods of significant b-positive variations analyzing the slope of a fitted linear function using 30-day length time windows (see Supplementary Information).

Variations of the *b*-positive parameter are commonly observed at Parkfield since 1970. *b*-positive varies between a minimum of 0.47 recorded 20 minutes after the 2004 mainshock and a maximum of 1.79 recorded 10 months after. Other particularly low *b*-positive estimates of 0.52 are also recorded in 1992 with the occurrence of M>4 earthquakes (Roeloffs and Langbein, 1994). Estimates of *b*-positive using the PKD-MR catalog show in general similar trends to estimates obtained using the NCSN catalog (Figure 7) but with shorter period variations, because a smaller time window is used for the PKD-MR catalog. In the aftershock period PKD-MR catalog estimates show higher variability but the long-term trends are consistent with the NCSN catalog. Using the PKD-MR catalog we can get a more detailed picture of the *b*-positive temporal variations.

Estimates of *b*-positive in the Parkfield section are generally stable prior to the 2004 mainshock considering both catalogs (Figure 7), suggesting variations in the entire segment (Figure 1) prior to the mainshock are driven by the creeping section. After the nearby M6.5 San Simeon earthquake, we identify an increase in *b*-positive in the creeping section peaking at 1.71 80 days after the San Simeon earthquake. After this peak, estimates show some variability with periods of decrease and increase until about 47 days prior to the 2004 mainshock when we start to observe a period of significant decrease. This decreasing period lasts through the first week after the 2004 mainshock and *b*-positive estimates drop from 1.4 to 0.6.

In contrast, after the mainshock, *b*-positive variations in the entire segment resemble the *b*-positive variations in the Parkfield section suggesting *b*-positive is driven by seismicity in the Parkfield section during this period. Though strong oscillations are observed, there is a general trend of increasing *b*-positive in the entire area consistent with the increased occurrence of the smaller aftershocks. Following the occurrence time of the 2004 M9.1 Sumatra earthquake there is a *b*-positive peak, also observed in the NCSN catalog. However this peak appears to initiate before the Sumatra earthquake, similar to a variation observed in the β-statistics (Figure 6), and this period also coincides with the HRSN network outage. Additional analysis is necessary to clarify what drives this variation and if there is a relation to the M9.1 Sumatra earthquake, which has been observed to impact tremor rates in Parkfield (Ghosh *et al.*, 2009).

Discussion

Using the high-quality data of the HRSN network and matched filter detection we compiled a high-resolution catalog for the Parkfield segment of the SAF spanning from November 6, 2003, to March 28, 2005, including the 2004 M6 mainshock. Events in the PKD-MR catalog have high-precision relative locations (see Supplementary Information) obtained by cross-correlation derived differential travel-times and magnitude estimations based relative amplitude measurements.

Our initial analysis of the temporal evolution of earthquakes in the months before the 2004 M6 earthquake does not indicate clear precursory signals near the mainshock epicenter. However, we find significant changes in the creeping section. β-statistics show a significant seismic activity increase about 49 days before the mainshock (Figure 6) around the same time a significant *b*-positive decrease is identified (Figure 7). A more pronounced β increase is also observed in the 48 hours before the mainshock that coincides with an increased seismic moment release in the creeping section. A decrease of the seismic rate in the creeping section after the stress changes imposed by the neighboring M6.5 San Simeon earthquake (Meng *et al.*, 2013) also relate to positive variations of *b*-positive. These results are consistent with the observations of Tormann *et al.* (2013) and Khoshmanesh and Shirzaei (2018) that connected *b*-value variations at Parkfield with creeping rates. However, this does not explain all the smaller *b*-positive variations observed in aftershock period as they do not relate to the observed seismicity rates. Regardless, our observations suggest an increased release of stress seismically with an increase in the number of earthquakes with larger magnitudes and a decrease in the aseismic creeping rate in the creeping section prior to the 2004 mainshock. In the same period, seismic activity remains stable in the rupturing Parkfield section. Since these changes are observed in the creeping section with a median along-strike distance of more than 30 km NW of the mainshock epicenter (Figure 3 and S17), we are unable to connect them to the mainshock nucleation process. These changes in the creeping section could be related to a larger-scale preparation process like the aforementioned delocalization prior to the 2004 Parkfield mainshock Ben-Zion and Zaliapin (2020) or the Slow Slip Event (SSE) identified in the Parkfield area by Khoshmanesh and Shirzaei (2018). Still, Ben-Zion and Zaliapin (2020) analysis shows delocalization more than a year prior to the mainshock and changes we observe occur only weeks before. Khoshmanesh and Shirzaei (2018) identified that SSE starts in the creeping section in 2003 and appears to migrate towards the Parkfield section with a rate decrease in the area where we identify the seismic changes. This SSE observation can explain the seismic rate increase and *b*-positive decrease we observe in the creeping section. But it is still intriguing why no seismic changes are identified in the Parkfield section due to the SSE. We also note that no other precursory changes have been observed in the Parkfield section before the mainshock like strain measurements (Borcherdt *et al.*, 2006; Johnston *et al.*, 2006). These observations add to a number of other changes observed prior to the 2004 mainshock also without a clear connection to the mainshock rupture (e.g., Bakun *et al.*, 2005).

The mechanisms of the Parkfield sequence still have many unanswered questions that the wealth of data in our new catalog could potentially address. Especially in the early aftershock period, which includes the majority of our new detections, our catalog may contain important information on the mechanisms of aftershock propagation and its relation to afterslip at Parkfield (e.g., Jiang *et al.*, 2021). From our initial observations, in the aftershock period we can identify the logarithmic-time expansion of the aftershocks in the weeks following the mainshock that extends to about 38 km NW along-strike into the creeping section. The PKD-MR catalog has also the potential to be used for further studies on external stress interactions with the Parkfield section and the 2004 sequence, for example tidal stress modulations (e.g., Delorey *et al.*, 2017) or dynamic stresses imposed by distant earthquakes such as the 2004 M9.1 Sumatra earthquake (e.g., Taira *et al.*, 2009).

Data and Resources

The PKD-MR earthquake catalog is provided as part of the supplementary information. We also include our local detections and template catalogs to allow easy reproduction. Differential travel-time measurements can be shared upon request. Seismic waveforms used for detection and relocation can be downloaded from the NCEDC and SCEDC data centers. Waveforms were retrieved using ObsPy (Beyreuther *et al.*, 2010) for networks BP (NCEDC, 2014), NC (USGS Menlo Park, 1967), CI (California Institute of Technology and United States Geological Survey Pasadena, 1926) and PG (Central Coast Seismic Network, PG&E). Supplementary information includes additional figures, the velocity model used for relocation and descriptions of the supplementary catalogs.

Declaration of Competing Insterests

The authors acknowledge there are no conflicts of interest recorded.

Acknowledgements

This project is supported by National Science Foundation Earthscope program grants EAR-1736197 (ZP and MN) and EAR-1736078 (GL). MN is supported by a PhD fellowship from the Portuguese research foundation FCT, grant SFRH/BD/139033/2018. Figures were generated using Generic Mapping Tools (Wessel *et al.*, 2019) and Matploltib (Hunter, 2007). The authors thank Rob Skoumal, one anonymous reviewer and the Data Mine editor Taka'aki Taira for constructive comments and reviews. We also thank Roland Bürgmann for providing feedback.

References

Abercrombie, R. E. (1996). The magnitude-frequency distribution of earthquakes recorded with deep seismometers at Cajon Pass, southern California, *Tectonophysics* **261**, no. 1, 1–7, doi: 10.1016/0040-1951(96)00052-2.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. seism. Soc. Am.* **68**, 1521–1532, doi: 10.1353/cch.2001.0013.

Amitrano, D. (2003). Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the bvalue, *Journal of Geophysical Research: Solid Earth* **108**, no. B1, n/a-n/a, doi: 10.1029/2001jb000680.

Aron, A., and J. L. Hardebeck (2009). Seismicity Rate Changes along the Central California Coast due to Stress Changes from the 2003 M 6.5 San Simeon and 2004 M 6.0 Parkfield Earthquakes, *Bulletin of the Seismological Society of America* **99**, no. 4, 2280–2292, doi: 10.1785/0120080239.

Bakun, W. H., B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M. J. S. Johnston, J. Langbein, J. J. Lienkaemper, *et al.* (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, 7061, *Nature* **437**, no. 7061, 969–974, doi: 10.1038/nature04067.

Bakun, W. H., and A. G. Lindh (1985). The Parkfield, California, Earthquake Prediction Experiment, *Science* **229**, no. 4714, 619–624, doi: 10.1126/science.229.4714.619.

Barbot, S., Y. Fialko, and Y. Bock (2009). Postseismic deformation due to the M w 6.0 2004 Parkfield earthquake: Stress-driven creep on a fault with spatially variable rate-and-state friction parameters, *Journal of Geophysical Research* **114**, B07405, doi: 10.1029/2008JB005748.

Ben-Zion, Y., and I. Zaliapin (2020). Localization and coalescence of seismicity before large earthquakes, *Geophysical Journal International* **223**, no. 1, 561–583, doi: 10.1093/gji/ggaa315.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010). ObsPy: A Python Toolbox for Seismology, *Seismological Research Letters* **81**, no. 3, 530–533, doi: 10.1785/gssrl.81.3.530.

Borcherdt, R. D., M. J. S. Johnston, G. Glassmoyer, and C. Dietel (2006). Recordings of the 2004 Parkfield Earthquake on the General Earthquake Observation System Array: Implications for Earthquake Precursors, Fault Rupture, and Coseismic Strain Changes, *Bulletin of the Seismological Society of America* **96**, no. 4B, S73–S89, doi: 10.1785/0120050827.

Brenguier, F., N. M. Shapiro, M. Campillo, V. Ferrazzini, Z. Duputel, O. Coutant, and A. Nercessian (2008). Towards forecasting volcanic eruptions using seismic noise, *Nature Geosci* **1**, no. 2, 126–130, doi: 10.1038/ngeo104.

Cabrera, L., P. Poli, and W. B. Frank (2022). Tracking the Spatio-Temporal Evolution of Foreshocks Preceding the Mw 6.1 2009 L'Aquila Earthquake, *Journal of Geophysical Research: Solid Earth* **127**, no. 3, e2021JB023888, doi: 10.1029/2021JB023888.

California Institute of Technology and United States Geological Survey Pasadena (1926). Southern California Seismic Network, International Federation of Digital Seismograph Networks, doi: 10.7914/SN/CI.

Chun, K.-Y., Q.-Y. Yuan, and G. A. Henderson (2010). Precursory Rise of P-Wave Attenuation before the 2004 Parkfield Earthquake, *Bulletin of the Seismological Society of America* **100**, 509–521, doi: 10.1785/0120090104.

Custódio, S., and R. J. Archuleta (2007). Parkfield earthquakes: Characteristic or complementary?, *Journal of Geophysical Research: Solid Earth* **112**, no. B5, doi: 10.1029/2006JB004617.

Dascher-Cousineau, K., T. Lay, and E. E. Brodsky (2020). Two foreshock sequences post gulia and wiemer (2019), *Seismological Research Letters* **91**, 2843–2850, doi: 10.1785/0220200082.

Delorey, A. A., N. J. van der Elst, and P. A. Johnson (2017). Tidal triggering of earthquakes suggests poroelastic behavior on the San Andreas Fault, *Earth and Planetary Science Letters* **460**, 164–170, doi: 10.1016/j.epsl.2016.12.014.

Dieterich, J. H. (1978). Preseismic fault slip and earthquake prediction, *Journal of Geophysical Research: Solid Earth* **83**, no. B8, 3940–3948, doi: 10.1029/JB083iB08p03940.

Eberhart-Phillips, D., and A. J. Michael (1993). Three-dimensional velocity structure, seismicity, and fault structure in the Parkfield Region, central California, *Journal of Geophysical Research: Solid Earth* **98**, no. B9, 15737–15758, doi: 10.1029/93JB01029.

van der Elst, N. J. (2021). B-Positive: A Robust Estimator of Aftershock Magnitude Distribution in Transiently Incomplete Catalogs, *Journal of Geophysical Research: Solid Earth* **126**, no. 2, e2020JB021027, doi: 10.1029/2020JB021027.

Enescu, B., J. Mori, and M. Miyazawa (2007). Quantifying early aftershock activity of the 2004 mid-Niigata Prefecture earthquake (Mw6.6), *Journal of Geophysical Research: Solid Earth* **112**, no. B4, doi: 10.1029/2006JB004629.

Eshelby, J. D., and R. E. Peierls (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **241**, no. 1226, 376–396, doi: 10.1098/rspa.1957.0133.

Ghosh, A., J. E. Vidale, Z. Peng, K. C. Creager, and H. Houston (2009). Complex nonvolcanic tremor near Parkfield, California, triggered by the great 2004 Sumatra earthquake, *Journal of Geophysical Research: Solid Earth* **114**, no. B12, doi: 10.1029/2008JB006062.

Gibbons, S. J., and F. Ringdal (2006). The detection of low magnitude seismic events using array-based waveform correlation, *Geophysical Journal International* **165**, 149–166, doi: 10.1111/j.1365-246X.2006.02865.x.

Gulia, L., and S. Wiemer (2019). Real-time discrimination of earthquake foreshocks and aftershocks, *Nature* **574**, 193–199, doi: 10.1038/s41586-019-1606-4.

Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California*, *Bulletin of the Seismological Society of America* **34**, no. 4, 185–188, doi: 10.1785/bssa0340040185.

Herrmann, M., and W. Marzocchi (2021). Inconsistencies and Lurking Pitfalls in the Magnitude–Frequency Distribution of High-Resolution Earthquake Catalogs, *Seismological Research Letters* **92**, no. 2A, 909–922, doi: 10.1785/0220200337.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering* **9**, no. 3, 90–95, doi: 10.1109/MCSE.2007.55.

Ishimoto, M., and K. Iida (1939). Observations of earthquakes registered with the microseismograph constructed recently, *Bull. Earthq. Res. Inst.* **17**, 443–478.

Jiang, J., Y. Bock, and E. Klein (2021). Coevolving early afterslip and aftershock signatures of a San Andreas fault rupture, *Science Advances* **7**, no. 15, eabc1606, doi: 10.1126/sciadv.abc1606.

Johnston, M. J. S. (2006). Seismomagnetic Effects from the Long-Awaited 28 September 2004 M 6.0 Parkfield Earthquake, *Bulletin of the Seismological Society of America* **96**, no. 4B, S206–S220, doi: 10.1785/0120050810.

Johnston, M. J. S., R. D. Borcherdt, A. T. Linde, and M. T. Gladwin (2006). Continuous Borehole Strain and Pore Pressure in the Near Field of the 28 September 2004 M 6.0 Parkfield, California, Earthquake: Implications for Nucleation, Fault Response, Earthquake Prediction, and Tremor, *Bulletin of the Seismological Society of America* **96**, no. 4B, S56–S72, doi: 10.1785/0120050822.

Kagan, Y. Y. (2004). Short-Term Properties of Earthquake Catalogs and Models of Earthquake Source, *Bulletin of the Seismological Society of America* **94**, no. 4, 1207–1228, doi: 10.1785/012003098.

Kato, N. (2007). Expansion of aftershock areas caused by propagating post-seismic sliding, *Geophysical Journal International* **168**, no. 2, 797–808, doi: 10.1111/j.1365-246x.2006.03255.x.

Kato, A., and Y. Ben-Zion (2021). The generation of large earthquakes, 1, *Nat Rev Earth Environ* **2**, no. 1, 26–39, doi: 10.1038/s43017-020-00108-w.

Khoshmanesh, M., and M. Shirzaei (2018). Episodic creep events on the San Andreas Fault caused by pore pressure variations, 8, *Nature Geosci* **11**, no. 8, 610–614, doi: 10.1038/s41561-018-0160-2.

Langbein, J., R. Borcherdt, D. Dreger, J. Fletcher, J. L. Hardebeck, M. Hellweg, C. Ji, M. Johnston, J. R. Murray, R. Nadeau, *et al.* (2005). Preliminary Report on the 28 September 2004, M 6.0 Parkfield, California Earthquake, *Seismological Research Letters* **76**, no. 1, 10–26, doi: 10.1785/gssrl.76.1.10.

Li, Z., and Z. Peng (2016). An Automatic Phase Picker for Local Earthquakes with Predetermined Locations: Combining a Signal‐to‐Noise Ratio Detector with 1D Velocity Model Inversion, *Seismological Research Letters* **87**, no. 6, 1397–1405, doi: 10.1785/0220160027.

Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown, *Journal of the American Statistical Association* **64**, no. 325, 387–389, doi: 10.2307/2283748.

Lin, G. (2018). The Source‐Specific Station Term and Waveform Cross‐Correlation Earthquake Location Package and Its Applications to California and New Zealand, *Seismological Research Letters* **89**, no. 5, 1877–1885, doi: 10.1785/0220180108.

Lin, G. (2020). Waveform Cross‐Correlation Relocation and Focal Mechanisms for the 2019 Ridgecrest Earthquake Sequence, *Seismological Research Letters* **91**, no. 4, 2055–2061, doi: 10.1785/0220190277.

Lin, G., Z. Peng, and M. Neves (2022). Comparisons of in situ Vp/Vs ratios and seismic characteristics between northern and southern California, *Geophysical Journal International* **229**, no. 3, 2162–2174, doi: 10.1093/gji/ggac038.

Lomax, A., and A. Savvaidis (2022). High-Precision Earthquake Location Using Source-Specific Station Terms and Inter-Event Waveform Similarity, *Journal of Geophysical Research: Solid Earth* **127**, no. 1, e2021JB023190, doi: 10.1029/2021JB023190.

Marzocchi, W., I. Spassiani, A. Stallone, and M. Taroni (2019). How to be fooled searching for significant variations of the b-value, *Geophysical Journal International* **220**, no. 3, 1845–1856, doi: 10.1093/gji/ggz541.

Matthews, M. V., and P. A. Reasenberg (1988). Statistical methods for investigating quiescence and other temporal seismicity patterns, *Pure and Applied Geophysics PAGEOPH* **126**, no. 2–4, 357–372, doi: 10.1007/bf00879003.

Meng, X., Z. Peng, and J. L. Hardebeck (2013). Seismicity around Parkfield correlates with static shear stress changes following the 2003Mw6.5 San Simeon earthquake, *Journal of Geophysical Research: Solid Earth* **118**, no. 7, 3576–3591, doi: 10.1002/jgrb.50271.

Michelini, A., and T. V. McEvilly (1991). Seismological studies at Parkfield. I. Simultaneous inversion for velocity structure and hypocenters using cubic B-splines parameterization, *Bulletin of the Seismological Society of America* **81**, no. 2, 524–552, doi: 10.1785/BSSA0810020524.

Mignan, A. (2014). The debate on the prognostic value of earthquake foreshocks: A meta-analysis, 1, *Sci Rep* **4**, no. 1, 4099, doi: 10.1038/srep04099.

Mogi, K. (1969). Some features of recent seismic activity in and near Japan (2), Activity before and after great earthquakes, *Bull. Earthquake Res. Inst. Univ. Tokyo*, no. 47, 395–417.

Nadeau, R. M., and A. Guilhem (2009). Nonvolcanic tremor evolution and the San Simeon and Parkfield, California, earthquakes, *Science* **325**, 191–193, doi: 10.1126/science.1174155.

Nanjo, K. Z. (2020). Were changes in stress state responsible for the 2019 Ridgecrest, California, earthquakes?, *Nature Communications* **11**, 3082, doi: 10.1038/s41467-020-16867-5.

Northern California Earthquake Data Center (2014). High Resolution Seismic Network (HRSN), Northern California Earthquake Data Center, doi: 10.7932/HRSN.

Olson, E. L., and R. M. Allen (2005). The deterministic nature of earthquake rupture, 7065, *Nature* **438**, no. 7065, 212–215, doi: 10.1038/nature04214.

Oppenheimer, D. H., F. W. Klein, J. P. Eaton, and F. W. Lester (1993). The Northern California Seismic Network bulletin, January-December 1992, Report 93–578, Open-File Report, doi: 10.3133/ofr93578.

Peng, Z., J. E. Vidale, and H. Houston (2006). Anomalous early aftershock decay rate of the 2004 Mw6.0 Parkfield, California, earthquake, *Geophysical Research Letters* **33**, L17307, doi: 10.1029/2006GL026744.

Peng, Z., and P. Zhao (2009). Migration of early aftershocks following the 2004 Parkfield earthquake, *Nature Geoscience* **2**, no. 12, 877–881, doi: 10.1038/ngeo697.

Perrin, C., F. Waldhauser, E. Choi, and C. H. Scholz (2019). Persistent fine-scale fault structure and rupture development: A new twist in the Parkfield, California, story, *Earth and Planetary Science Letters* **521**, 128–138, doi: 10.1016/j.epsl.2019.06.010.

Pritchard, M. E., R. M. Allen, T. W. Becker, M. D. Behn, E. E. Brodsky, R. Bürgmann, C. Ebinger, J. T. Freymueller, M. Gerstenberger, B. Haines, *et al.* (2020). New Opportunities to Study Earthquake Precursors, *Seismological Research Letters* **91**, no. 5, 2444–2447, doi: 10.1785/0220200089.

Roeloffs, E. A. (2006). Evidence for Aseismic Deformation Rate Changes Prior to Earthquakes, *Annual Review of Earth and Planetary Sciences* **34**, no. 1, 591–627, doi: 10.1146/annurev.earth.34.031405.124947.

Roeloffs, E., and J. Langbein (1994). The Earthquake Prediction Experiment at Parkfield, California, *Reviews of Geophysics* **32**, no. 3, 315, doi: 10.1029/94rg01114.

Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019). Searching for hidden earthquakes in Southern California, *Science* **364**, no. 6442, 767–771, doi: 10.1126/science.aaw6888.

Scholz, C. H. (2015). On the stress dependence of the earthquake b value, *Geophysical Research Letters* **42**, 1399–1402, doi: 10.1002/2014GL062863.

Schorlemmer, D., and S. Wiemer (2005). Microseismicity data forecast rupture area, *Nature* **434**, no. 7037, 1086–1086, doi: 10.1038/4341086a.

Schorlemmer, D., S. Wiemer, and M. Wyss (2004). Earthquake statistics at Parkfield: 1. Stationarity of b values, *Journal of Geophysical Research* **109**, no. B12, doi: 10.1029/2004jb003234.

Shelly, D. R. (2017). A 15 year catalog of more than 1 million low-frequency earthquakes: Tracking tremor and slip along the deep San Andreas Fault, *Journal of Geophysical Research: Solid Earth* **122**, no. 5, 3739–3753, doi: 10.1002/2017JB014047.

Shelly, D. R. (2020). A High‐Resolution Seismic Catalog for the Initial 2019 Ridgecrest Earthquake Sequence: Foreshocks, Aftershocks, and Faulting Complexity, *Seismological Research Letters* **91**, no. 4, 1971–1978, doi: 10.1785/0220190309.

Shelly, D. R. (2009). Possible deep fault slip preceding the 2004 Parkfield earthquake, inferred from detailed observations of tectonic tremor, *Geophysical Research Letters* **36**, no. 17, doi: 10.1029/2009GL039589.

Shelly, D. R., G. C. Beroza, and S. Ide (2007). Non-volcanic tremor and low-frequency earthquake swarms, *Nature* **446**, 305–307, doi: 10.1038/nature05666.

Shelly, D. R., W. L. Ellsworth, and D. P. Hill (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 Long

Valley Caldera, California, earthquake swarm, *Journal of Geophysical Research: Solid Earth* **121**, no. 3, 1776–1795, doi: 10.1002/2015JB012719.

Shelly, D. R., and J. L. Hardebeck (2019). Illuminating Faulting Complexity of the 2017 Yellowstone Maple Creek Earthquake Swarm, *Geophysical Research Letters* **46**, no. 5, 2544–2552, doi: 10.1029/2018gl081607.

Shelly, D. R., and K. M. Johnson (2011). Tremor reveals stress shadowing, deep postseismic creep, and depth-dependent slip recurrence on the lower-crustal San Andreas fault near Parkfield, *Geophysical Research Letters* **38**, no. 13, doi: 10.1029/2011GL047863.

Shelly, D. R., S. C. Moran, and W. A. Thelen (2013). Evidence for fluid-triggered slip in the 2009 Mount Rainier, Washington earthquake swarm, *Geophysical Research Letters* **40**, 1506–1512, doi: 10.1002/grl.50354.

Shelly, D. R., Z. Peng, D. P. Hill, and C. Aiken (2011). Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes, *Nature Geosci* **4**, no. 6, 384–388, doi: 10.1038/ngeo1141.

Taira, T., P. G. Silver, F. Niu, and R. M. Nadeau (2009). Remote triggering of fault-strength changes on the San Andreas fault at Parkfield, 7264, *Nature* **461**, no. 7264, 636–639, doi: 10.1038/nature08395.

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, *Bulletin of the Seismological Society of America* **96**, no. 4B, S38–S49, doi: 10.1785/0120050825.

Tormann, T., S. Wiemer, S. Metzger, A. Michael, and J. L. Hardebeck (2013). Size distribution of Parkfield's microearthquakes reflects changes in surface creep rate, *Geophysical Journal International* **193**, no. 3, 1474–1478, doi: 10.1093/gji/ggt093.

USGS Menlo Park (1967). USGS Northern California Network, International Federation of Digital Seismograph Networks, doi: 10.7914/SN/NC.

Waldhauser, F. (2000). A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California, *Bulletin of the Seismological Society of America* **90**, 1353–1368, doi: 10.1785/0120000006.

Waldhauser, F., W. L. Ellsworth, D. P. Schaff, and A. Cole (2004). Streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity, *Geophysical Research Letters* **31**, no. 18, doi: 10.1029/2004GL020649.

Waldhauser, F., and D. P. Schaff (2008). Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods, *Journal of Geophysical Research: Solid Earth* **113**, doi: 10.1029/2007JB005479.

Wessel, P., J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian (2019). The Generic Mapping Tools Version 6, *Geochemistry, Geophysics, Geosystems* **20**, no. 11, 5556–5564, doi: 10.1029/2019GC008515.

Wiemer, S., and M. Wyss (2000). Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan, *Bulletin of the Seismological Society of America* **90**, no. 4, 859–869, doi: 10.1785/0119990114.

Woessner, J., and S. Wiemer (2005). Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty, *Bulletin of the Seismological Society of America* **95**, no. 2, 684–698, doi: 10.1785/0120040007.

Wyss, M., and R. E. Habermann (1988). Precursory seismic quiescence, *PAGEOPH* **126**, no. 2, 319–332, doi: 10.1007/BF00879001.

Yang, H., L. Zhu, and R. Chu (2009). Fault-Plane Determination of the 18 April 2008 Mount Carmel, Illinois, Earthquake by Detecting and Relocating Aftershocks, *Bulletin of the Seismological Society of America* **99**, no. 6, 3413–3420, doi: 10.1785/0120090038.

Yao, D., Y. Huang, and J. L. Fox (2021). New Insights into the Lake Erie Fault System from the 2019 ML4.0 Ohio Earthquake Sequence, *Seismological Research Letters*, no. 0895–0695, doi: 10.1785/0220200335.

Zhai, Q., Z. Peng, L. Y. Chuang, Y.-M. Wu, Y.-J. Hsu, and S. Wdowinski (2021). Investigating the Impacts of a Wet Typhoon on Microseismicity: A Case Study of the 2009 Typhoon Morakot in Taiwan Based on a Template Matching Catalog, *Journal of Geophysical Research: Solid Earth* **126**, no. 12, e2021JB023026, doi: 10.1029/2021JB023026.

Zhao, P., Z. Peng, and K. G. Sabra (2010). Detecting remotely triggered temporal changes around the Parkfield section of the San Andreas fault, *Earthq Sci* **23**, no. 5, 497–509, doi: 10.1007/s11589-010-0748-0.

Miguel Neves
Zhigang Peng
School of Earth and Atmospheric Sciences
Georgia Institute of Technology
311 Ferst Drive
Atlanta, Georgia 30332-0340 U.S.A.
mjneves@gatech.edu
zpeng@gatech.edu

Guoqing Lin
Department of Marine Geosciences
Rosenstiel School of Marine, Atmospheric and Earth Science
University of Miami
Miami, Florida 33149 U.S.A.
glin@rsmas.miami.edu

List of Figure Captions

Figure 1. Study area in Central California. (a) Map of stations and templates used in our study. (b) Location of the local templates. Bottom panel shows an along-strike profile of the earthquakes 5 km away from the San Andreas Fault with depth. Different colors denote the fault sections with different slip behaviors, including the creeping, Parkfield and locked sections, bounded by the SAFOD pilot hole and the town of Cholame. Fault lines in the area are denoted

in gray. Red T markers the start and end of the cross-section on the bottom panel. The sizes of the earthquakes in the cross-section plot scale with their estimated source radius (Peng and Zhao, 2009), which was calculated assuming a circular crack model (Eshelby and Peierls, 1957) and a nominal 3-MPa stress drop using the moment-magnitude relationship by Abercrombie (1996).

Figure 2. General pattern of the newly relocated earthquake catalog. a) Plot of the earthquakes in the new catalog in the period before the mainshock, with color denoting time relative to mainshock. b) Same as a) but showing aftershocks. In both plots, red colors indicate events happening closer to the mainshock. c) Plot of earthquake magnitudes with time for the entire Parkfield segment and each of the fault sections. Vertical lines mark the times of three major earthquakes during the study period.

Figure 3. Along-strike distribution of the events in our relocated catalog with time. a) Occurrence time of events in days with a linear time scale. b) Occurrence time of events in seconds with a logarithmic time scale.

Figure 4. Comparison of our catalog with different relocated catalogs for Parkfield. a) Earthquakes in the new PKD-MR catalog above estimated AMc (0.6). Red and blue lines mark locations of 2 km cross-sections shown in b) and c), respectively. The gray line marks the zero cross-fault distance. The sizes of the earthquakes scale with their estimated source radius following the same method as Figure 1. Supplementary movie S1 shows cross-sections along the entire study area.

Figure 5. a) Gutenberg-Richter distribution of the template catalog (orange) and the new relocated catalog (blue). b-value maximum-likelihood and apparent magnitude of completeness estimates are denoted in gray and black for the template and the new catalog, respectively. b) Distribution of positive magnitude differences above AM_C. b-positive maximum-likelihood estimates are denoted in gray and black for the template and the new catalog, respectively. c) Variation of *b*-value estimates with assumed magnitude of completeness. Uncertainties are estimated by bootstrapping. d) Variation of *b*-positive estimates with assumed minimum magnitude difference.

Figure 6. Variation of *β*-statistics at the studied period using two 30-day moving windows and our new PKD-MR catalog. a) Three panels show variations considering the entire Parkfield segment and the different sections of the fault. b) Detailed view of the 50 days prior to the mainshock.

Figure 7. Time variations of *b*-value using the b-positive estimator. Top three panels show variations for the Parkfield segment and the creeping and Parkfield sections. Grey shaded areas indicate periods where most HRSN stations do not have complete daily records. Bottom panel shows estimates for the entire NCSN catalog. Color bars at the top in the top three panels and bottom in the bottom panel show identified periods of significant b-positive changes using our slope analysis, where red corresponds to significant increases, blue to significant decreases and gray to no significant changes.