Annual modulation of seismicity along the San Andreas Fault near Parkfield, CA

Lizet B. Christiansen,1 Shaul Hurwitz,1 and Steven E. Ingebritsen1

Received 2 November 2006; revised 9 January 2007; accepted 22 January 2007; published 23 February 2007.

[1] We analyze seismic data from the San Andreas Fault (SAF) near Parkfield, California, to test for annual modulation in seismicity rates. We use statistical analyses to show that seismicity is modulated with an annual period in the creeping section of the fault and a semiannual period in the locked section of the fault. Although the exact mechanism for seasonal triggering is undetermined, it appears that stresses associated with the hydrologic cycle are sufficient to fracture critically stressed rocks either through pore-pressure diffusion or crustal loading/unloading. These results shed additional light on the state of stress along the SAF, indicating that hydrologically induced stress perturbations of ~2 kPa may be sufficient to trigger earthquakes. Citation: Christiansen, L. B., S. Hurwitz, and S. E. Ingebritsen (2007), Annual modulation of seismicity along the San Andreas Fault near Parkfield, CA, Geophys. Res. Lett., 34, L04306, doi:10.1029/2006GL028634.

1. Introduction

[2] Microearthquakes may be triggered or modulated by climatic forces that have an annual period, implying a causal link between the hydrologic cycle and the mechanical behavior of the upper crust [Gao et al., 2000; Saar and Manga, 2003; Christiansen et al., 2005; Kraft et al., 2006]. This relation implies that stresses induced by the annual hydrologic cycle are sufficient to fracture near-critically stressed rock either through pore-pressure diffusion [Talwani and Acree, 1984; Shapiro et al., 2003; Hainzl et al., 2006] or loading/unloading of the elastic crust [Heki, 2003]. In this study, we use a suite of statistical tests [Christiansen et al., 2005] to explore whether seismicity on the San Andreas Fault (SAF) in the vicinity of Parkfield, California, is annually modulated.

[3] The Parkfield region is an ideal location to search for a connection between seismicity and precipitation because the SAF here is seismically active [Bakun et al., 2005]; an extensive seismic network provides detailed earthquake data [Bakun and Lindh, 1985; Bakun et al., 2005; Roeloffs and Langbein, 1994]; precipitation rates are relatively low (>0.5 m/y on average); the fault is believed to be extremely weak [Zoback et al., 1987; Rice, 1992; Hickman and Zoback, 2004; Townend and Zoback, 2004]; and the characteristics of microearthquakes in the region have been studied extensively [Poley et al., 1987; Rubin et al., 1999; Nadeau and McEvilly, 2004]. Any relation between rainfall and earthquake occurrence in this environment would imply a stress threshold for triggered seismicity that is lower than commonly accepted [Harris, 1998].

2. Methods

[4] We analyze a 21-year seismic catalog (January 1984 to January 2005) from the SAF near Parkfield (http://quake.geo.berkeley.edu/) (Figure 1a) using a suite of statistical analyses [Christiansen et al., 2005]. During the past 20 years seismicity has generally increased (Figure 1b). No significant changes have been made to the seismic network to change the detection of earthquakes of M > 1.25. The seasonal modulation we explore for is a small perturbation overlain on this long-term trend and other periodic signals [Nadeau et al., 1995; Nadeau and McEvilly, 2004]. We separately examine 2284 events in the locked/transition section of the fault (south of 36.0°N; hereafter referred to as locked) and 3093 events in the creeping section (north of 36.0°N) because of postulated mechanical differences between these sections [Schorlemmer and Wiemer, 2005].

[5] To ensure that instrument variability or inconsistent catalog completeness do not affect our results, we use a frequency-magnitude distribution to compute the minimum magnitude for catalog completeness. Using a b-value of 1 in a Gutenberg-Richter analysis, the minimum magnitude for reliable completeness is M = 1.25 (Figure 1c). The magnitude distribution for each season is calculated to test for seasonal measurement biases. Frequency-magnitude trends for each season match those of the full dataset, and Mc values are identical (Figure 1c). Based on these results, we remove all earthquakes from the catalog below M = 1.25.

[6] We also remove spatially and temporally clustered sequences of aftershock earthquakes from the time series using established methods [Reasenberg, 1985]. Because we examine the locked and creeping sections separately, we decluster the entire data set, as well as the locked and the creeping sections independently. The number of earthquakes removed by declustering from the entire data set is slightly greater than the number of earthquakes removed by declustering from the locked and creeping sections separately (Table 1). The magnitude of each main shock is adjusted to represent the removed aftershock events in each sequence. Declustering reduces the number of events by ~30% (Table 1), largely by filtering out the seismic activity following the September 2004 M = 6.0 earthquake in the locked section of the fault.

[7] To determine the statistical significance of the observed earthquake distribution pattern, we apply a series of five statistical tests following the methodology in Christiansen et al. [2005], using a combination of ANOVA tests and a Kruskal-Wallis test [Dixon and Massey, 1983] on both unprocessed and normalized data. Often Fourier series...
and power spectra are used for these types of analyses; however, the time-series is too short to provide a robust result for annual cycles. We use three normalization schemes to ensure that the normalization process does not bias our results. Data are normalized using logarithmic and square-root transformations. In addition, in a third normalization scheme, the number of earthquakes each month is divided by the maximum number of earthquakes recorded in any month of that year, so that in each year the number of earthquakes per month varies between 0 and 1. ANOVA tests are performed on the non-normalized data as well as the three normalized data sets, and the Kruskal-Wallis test is preformed on non-normalized data. By using multiple statistical methods on a declustered dataset, we reduce the likelihood that the normalization process will bias the results. We consider the number of earthquakes per month and the number of earthquakes binned over 2-, 3-, and 6-month intervals. The intervals are rotated through the year using a moving window to determine when the greatest difference between earthquake numbers is achieved. For example, with 6-month intervals, we compare January-June with July-August, then February-July with August-January, and so forth. The statistical tests determine the probability that the timing of seismicity differs significantly from a random distribution. We require that all five tests have a significance of \( >95\% \) \((p\text{-value} < 0.05)\) for data to be considered non-random.

3. Results

Before invoking formal statistics, it is useful to explore the processed and unprocessed data for a visible structure. Figure 2 shows various subsets of the data binned by month. In the locked section, earthquake occurrence peaks semi-annually in March-May and September-November (Figure 2a). Relatively few earthquakes occur in January and during June-July. In the creeping section, peak earthquake numbers are more broadly distributed over a 6-month interval from August - January (Figure 2b). Earthquakes with \( M > 2 \) have a bimodal distribution in both sections of the fault.

It is possible that a few years with anomalously high seismicity rates could create the structure that is visible in Figure 2, yet not reflect the overall trends of the data set. To explore this possibility we calculate the average number of earthquakes in two ways. In the first approach, we divide the total number of earthquakes by the total number of months in the 21-year data set. For example, for the locked section, 996 earthquakes are divided by 252 months, giving an average of \( \approx 4 \) earthquakes per month. In the second approach, the average number of earthquakes per month is calculated separately for each year. For example, if 24 earthquakes occur in a particular year, the average number of earthquakes/month for that year is two. For each averaging method, we count the number of months that

<table>
<thead>
<tr>
<th>Number of EQs</th>
<th>1-Month</th>
<th>2-Months</th>
<th>3-Months</th>
<th>6-Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declustered</td>
<td>3734</td>
<td>0.10–0.23</td>
<td>0.03–0.07</td>
<td>0.02–0.03</td>
</tr>
<tr>
<td>Creeping</td>
<td>2817</td>
<td>0.14–0.23</td>
<td>0.03–0.05</td>
<td>0.01–0.02</td>
</tr>
<tr>
<td>Creeping (w/o ’04)</td>
<td>2614</td>
<td>0.19–0.37</td>
<td>0.04–0.10</td>
<td>0.02–0.05</td>
</tr>
<tr>
<td>Locked</td>
<td>996</td>
<td>0.04–0.20</td>
<td>0.01–0.02</td>
<td>0.01–0.02</td>
</tr>
<tr>
<td>Locked (w/o ’04)</td>
<td>921</td>
<td>0.09–0.26</td>
<td>0.02–0.10</td>
<td>0.01–0.03</td>
</tr>
<tr>
<td>Not declustered</td>
<td>5377</td>
<td>0.05–0.42</td>
<td>0.01–0.05</td>
<td>0.002–0.02</td>
</tr>
</tbody>
</table>

*Bold values indicate a significance of \( >95\% \) for all 5 tests.
exceeds the average (Figure 3). In the locked section, the majority of months with above-average seismicity occurs in spring and fall (Figure 3a), consistent with the pattern in Figure 2a. In the creeping section, the trend is less clearly defined (Figure 3b); however, there is a greater number of years with above-average seismicity during August - January, consistent with the pattern seen in Figure 2b.

The p-values for the five formal statistical tests show strong evidence for seasonal variations in seismicity in both the locked and creeping sections (Table 1). In the locked section, earthquake distributions are significantly non-random for 2- and 3-month intervals, and in 4 of 5 tests for 1-month intervals. Data from the creeping section show significant seasonality in 2-, 3-, and 6-month intervals, with the greatest significance in 6-month binning (p-values ≤ 0.003), consistent with the broad peak in seismicity seen in Figures 2b and 3b. Removing year 2004 (when the largest recent earthquake occurred) from the dataset reduces the significance for 2-month intervals to 90% in both sections.

We have also compared the binned earthquake data with 10,000 randomly distributed earthquake datasets, each with 3734 events (Table 2). Values in bold indicate when the actual number of earthquakes for a given interval in the actual dataset exceeds the 95th percentile of the 10,000 random data sets. In the locked section, the number of earthquakes is above the 95th percentile in 1-, 2-, and 3-month intervals. In the creeping section, seismicity is elevated for 2-, 3-, and 6-month intervals.

Similarly, we compare the actual number of earthquakes during a given interval with the expected number of earthquakes, assuming a uniform distribution over time (Table 2). The ratio between the actual number of earthquakes and the expected number for a uniform distribution is termed anomalous earthquakes (AE); AE = 1 when the actual number of earthquakes is the same as the expected number of earthquakes in a given interval. In the locked section, the actual number of earthquakes is higher than average for 1-, 2-, 3-, and 6-month intervals (Table 2), although only slightly higher in 6-month intervals (AE = 1.06). The deviation from the average is largest for 1-month intervals, with 30 earthquakes above the 80 expected earthquakes (AE = 1.36) for the month of November. In the creeping section, the increase in seismicity is broadly distributed over the same 6-month interval defined by the other tests. There are approximately 20 additional earthquakes per month during the 6-month interval of increased seismic activity.

4. Discussion

Each statistical approach to assess the timing of seismicity gives similar results, and all are consistent with the initial, visual inspections of the data (Figure 2). The locked section exhibits narrowly-defined, semiannual peaks in seismicity in spring and fall. The creeping section exhibits a broad increase in seismicity from August - January.

We see some evidence of correlation between periods of heavy rainfall and increased seismicity in the creeping section, but not in the locked section. Rainfall in Parkfield follows an annual cycle, with the onset of rainfall typically occurring in November and largest storms typically occurring from February through April [Roeloffs, 2001] (Figure 4). We compiled rainfall data from the Parkfield station [Roeloffs, 2001] and four stations near Parkfield: Paso Robles, Black Mountain, Bradley, and Santa Margarita (http://cdec.water.ca.gov/); where data were missing or erroneous, averages of the remaining stations were used. Between 1984 and 2004, 13 months had >15 cm of rainfall. In the creeping section, each of the high rainfall months is followed by a peak in seismicity. The lag time varies from 2 to 9 months and averages ~5 months.

Figure 2. Number of earthquakes per month for full dataset, declustered dataset, and earthquakes with M > 2 for (a) locked and (b) creeping sections. Solid horizontal lines show average monthly seismicity rate for declustered dataset; dashed horizontal lines show average monthly seismicity rate for M > 2.

Figure 3. Number of years with above-average seismicity, binned monthly for (a) locked and (b) creeping sections. Dashed line = monthly average based on total number of earthquakes over length of the time series. Solid line = monthly average based on the number of yearly seismic events.
In the locked section, we see a correlation between shallow groundwater levels and periods of increased seismicity (Figure 4). The highest water level occurs in April and the lowest water level occurs in October; strain data follow a similar trend. Seismicity peaks approximately one month after the seasonal low in water level.

The differences in timing of seismicity between the locked and creeping sections of the fault may relate to mechanical differences [Schorlemmer and Wiemer, 2005]. However, there may also be significant hydrologic differences between these sections. Inspection of USGS EROS Data Center imagery (http://edc.usgs.gov/) shows that there is typically more irrigated agriculture south of Parkfield (~locked section) than north of Parkfield (~creeping section). Thus the locked section may experience two recharge pulses annually, one natural and one artificial, whereas the creeping section experiences only the single, natural recharge season (Figure 4). Further, it has long been suggested [Irwin and Barnes, 1975; Kharaka et al., 1999] that there are “deep” (metamorphic) fluid sources in the creeping section that are absent in the locked section; these may increase pore pressure at depth.

Though some aspects of the hydrologic cycle would seem to be the most obvious driver for (semi-) annual variations in seismicity, the relations between precipitation, recharge, stress/strain, and seismicity (Figure 4) are not well-constrained. Pore-pressure diffusion is one possible causal mechanism that has received considerable recent attention [cf. Saar and Manga, 2003; Christiansen et al., 2005; Hainzl et al., 2006]. Hydrologic recharge increases pore pressure and thereby decreases effective stress at depth. For Parkfield, the hydraulic diffusivity calculated based on the average depth of earthquakes (5 km) and an apparent lag time of ~5 months is 2 m²/s. This value is similar to the value invoked to explain precipitation-induced earthquakes in Bavaria (3.3 ± 0.8 m²/s) [Hainzl et al., 2006]. However, it is much higher than that inferred from the water-level response of Parkfield wells to barometric pressure (~10⁻⁴ m²/s) [Roeloffs, 1998] or from a pumping test in another active fault (7 ± 1 · 10⁻⁵ m²/s) [Doan et al., 2006]. Further, such a large value of diffusivity would correspond to a permeability of roughly 10⁻¹⁵ m², much larger than the value of ~10⁻¹⁸ m² measured in the Cajon Pass well further south on the SAF [Townend and Zoback, 2000], or the values used in numerical models to match the observed thermal structure [Saffer et al., 2003].

The extremely high apparent diffusivity may imply that the lag is greater than 1 year or that the annual modulation is induced by inelastic relaxation of the crust associated with decreasing groundwater levels. It should be noted that our analysis cannot distinguish between a lag time of ~5 months and multiple years + 5 months (i.e. 17, 29, 41 months). If the actual lag time is ~3.5 years instead of ~5 months, then the inferred permeability (diffusivity) is reduced to 10⁻¹⁸ m², the value inferred from the Cajon Pass well [Townend and Zoback, 2000].

Any hydrologic triggering of earthquakes in the Parkfield region would imply that the associated stress triggers are small, supporting previous indications that SAF is very near failure stress. Assuming 100% recharge and a minimum porosity of 0.1, the effective stress change exerted by 15 cm of rain – an unusually large monthly total – is ~2 kPa. This is below typically accepted levels for external triggering [Harris, 1998]. However, Ziv and Rubin [2000] found that static stress changes of less than 10 kPa in central California have noticeable triggering effects, and Hainzl et al. [2006] invoked much lower values (<2 kPa) to explain precipitation-induced earthquakes in Bavaria.

---

**Table 2. Number of Earthquakes for Given Intervals for the Parkfield Data Set**

<table>
<thead>
<tr>
<th>Expected</th>
<th>1-Month</th>
<th>2-Months</th>
<th>3-Months</th>
<th>6-Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declustered</td>
<td>311</td>
<td>355 (1.14)</td>
<td>668 (1.10)</td>
<td>996 (1.07)</td>
</tr>
<tr>
<td>Creeping</td>
<td>235</td>
<td>266 (1.13)</td>
<td>526 (1.12)</td>
<td>781 (1.11)</td>
</tr>
<tr>
<td>Locked</td>
<td>83</td>
<td>113 (1.36)</td>
<td>206 (1.24)</td>
<td>295 (1.18)</td>
</tr>
<tr>
<td>Not declustered</td>
<td>448</td>
<td>893 (1.99)</td>
<td>1659 (1.84)</td>
<td>2244 (1.67)</td>
</tr>
</tbody>
</table>

*Bold values exceed the 95th percentile when expected earthquakes are compared with 10,000 randomized data sets. Numbers in parentheses represent anomalous earthquakes (AE) as defined in text.*

---

**Figure 4.** Monthly averages normalized to the January value of strain (at D101 and FR01; see Figure 1a for locations), shallow groundwater levels (at WBV), and average rainfall in the Parkfield region (see text for description). Prior to normalization, strain data show an annual signal of 0.2–1 μstrain (15–70 kPa, assuming Young’s modulus ~70 GPa); average-annual water-level changes are about 2 meters; and rainfall ranges up to ~18 mm/month. (b) Monthly average number of earthquakes in the creeping section for the period 1984–2005 and 2σ error bars. (c) Monthly average number of earthquakes in the locked section for the period 1984–2005 and 2σ error bars.
[20] It is possible that alluvial valleys along the strike of the fault concentrate runoff and create small zones where crustal loading is enhanced. In the shallow WBV1 well (Figure 4), average annual water-level changes are ~2 m, with highest water levels in April [Roelofs, 1998]. Drainage data from the nearby DL01 and FR01 dailometers (Figure 4) show an annual signal of 0.2–1 microstrain (15–70 kPa, assuming Young’s modulus = 70 GPa) from March - May. If runoff and groundwater recharge are focused such that change in the water table is locally enhanced, stress change at seismogenic depths may exceed 2 kPa. In addition, some creepmeters show increased creep rates during the rainy season, while others show accelerated creep following individual storms [Roelofs, 2001], similar to the relationship we infer between precipitation and earthquakes in the locked and creeping sections of the fault, respectively.

[21] The evidence for seasonal variations in seismicity at Parkfield is strongly supported in all statistical tests, as well as being visually evident in data sets, and the timing of seismicity seems linked to the hydrologic cycle. The associated stress triggers are small, supporting previous indications that SAF is near failure stress. However, further work is required to examine the mechanism by which meteoric fluids trigger seismicity.

[22] Acknowledgments. L. B. Christiansen was funded by the U.S. Geological Survey through an associateeeship from the National Research Council. We thank Jeanne Hardebeck, Ruth Harris, Evelyn Roelofs, Mark Zoback, and an anonymous reviewer for constructive comments.

References