Influences on the location of repeating earthquakes determined from $a$ and $b$ value imaging

Masashi Kawamura¹ and Kate Huihsuan Chen¹

¹Department of Earth Sciences, National Taiwan Normal University, Taipei, Taiwan

Abstract To explore where earthquakes tend to recur, we statistically investigated repeating earthquake catalogs and background seismicity from different regions (Parkfield, Hayward, Calaveras, and Chihshang Faults). We show that the location of repeating earthquakes can be mapped using the spatial distribution of the seismic $a$ and $b$ values obtained from the background seismicity. Molchan’s error diagram statistically confirmed that repeating earthquakes occur within areas with high $a$ values (2.8–3.8) and high $b$ values (0.9–1.1) on both strike-slip and thrust fault segments. However, no significant association held true for fault segments with more complicated geometry or for wider areas with a complex fault network. The productivity of small earthquakes responsible for high $a$ and $b$ values may thus be the most important factor controlling the location of repeating earthquakes. We inferred that the location of high creep rate in planar/listric fault structures might be indicated by $a$ values of ~3 and $b$ values of ~1.

1. Introduction

A repeating earthquake sequence (RES) is a group of earthquakes characterized by nearly identical waveforms with highly similar location and magnitude. Such earthquakes have occurred in diverse tectonic environments, including convergent plate boundaries [Matsuzawa et al., 2002; Igarashi et al., 2003; Uchida et al., 2003], transform faults [Vidale et al., 1994; Ellsworth, 1995; Nadeau et al., 1995], and plate-suture boundary faults [Chen et al., 2008, 2009]. These earthquakes tend to dominate in regions with frequent microseismicity but not near regions considered locked. Numerical, analytical, and laboratory experiments indicate that the physical properties of RESs may include repetitive ruptures of small asperities surrounded by stable sliding areas [Beeler et al., 2001; Sammis and Rice, 2001; Johnson and Nadeau, 2002; Nadeau and McEvilly, 2004]. Therefore, RES observations can be used to detect fault creeping and locations where the creeping and locked areas of the faults abut [Bürgmann et al., 2000; Sammis and Rice, 2001; Malservisi et al., 2005; Chen and Lapusta, 2009; McLaskey and Kilgore, 2013]. RESs may also function as fault slip-rate indicators [Nadeau and McEvilly, 1999; Uchida et al., 2003; Nadeau and McEvilly, 2004]. Knowledge of the expected location of RESs is crucial for fault deformation monitoring and assessment of earthquake potential. However, a precise description of factors determining RES locations is lacking. Where are repeating earthquakes expected to occur? Is the occurrence of RESs simply a function of earthquake magnitude or is a specific fault property required? These questions cannot be answered without a systematic investigation using independent repeating earthquake data from different tectonic environments.

If spatial distribution of repeating earthquakes depends on fault zone heterogeneity (i.e., asperity distribution), it can be examined through frequency-size statistics of earthquakes [Turcotte, 1989; Ben-Zion and Rice, 1993, 1995]. The frequency-magnitude distribution of earthquakes is expressed by the Gutenberg-Richter relationship $\log N = a - bM$ [Gutenberg and Richter, 1954], where $N$ is the cumulative number of earthquakes with magnitude larger than $M$. The $a$ value represents the earthquake productivity in a given area. Higher $a$ values indicate larger numbers of $M \leq M_c$ earthquakes, where $M_c$ is the magnitude of completeness of the local earthquake catalog. The $b$ value differentiates the areas of larger earthquakes from those of microearthquakes. At Parkfield, $b$ values range from 0.5 to 0.7 in the asperities but are greater than 1.2 in the creeping segments [Amelung and King, 1997; Wiemer and Wyss, 1997]. If repeating earthquakes tend to occur in places where fault creeping is evident, particular threshold of $a$ and $b$ values may be expected in different areas. However, there is a lack of systematic studies using repeating earthquake data from various regions to compare differences in these frequency-size statistics.

In this study we analyze $a$ and $b$ values from California and Taiwan, where a large number of repeating earthquakes were documented. We aim to answer the following questions: (1) Is the fault zone heterogeneity...
2. Repeating Earthquakes in Taiwan and California

Taiwan is situated at the active plate boundary between the Eurasian Plate and the Philippine Sea Plate and high rates of seismicity occur in the area. With a convergence rate of 8 cm/yr, the Longitudinal Valley Fault (LVF) in eastern Taiwan shows extreme strain concentration in an oblique collision zone [Angelier et al., 1997] with a fast surface slip rate of 1−3 cm/yr [Yu and Kuo, 2001]. The 160 km long LVF is characterized by reverse and strike-slip faulting with active creep in the south, whereas frequent and active slip rates occur in the north. Of the RESs, 52% are concentrated in areas of relatively high seismic activity, mainly in two distinct segments along the LVF, the Chihshang and Hualien segments [Chen et al., 2008, 2009]. For the detailed description on repeating earthquakes and the inferred slip rate, please see Text S1.1 and Figures S1 and S2 in the supporting information.

California, located on the plate boundary between the Pacific Plate and the North American Plate, is also characterized by high rates of seismicity. RESs are abundant and widely distributed along the creeping faults but do not overlap historical major earthquakes [Nadeau and McEvilly, 2004]. It can therefore be inferred that the high creep rate region hosts a large number of repeating earthquakes [Waldhauser and Ellsworth, 2002; Bürgmann et al., 2000]. In this study, we also statistically analyzed repeating earthquakes documented from the San Andreas Fault system with strike-slip faulting: Parkfield area, the Hayward, and Calaveras Faults (Figure 1b). The detailed description on repeating earthquakes and the inferred slip rate are also provided in Text S1.2 and Figures S3–S5.

3. The a and b Value Imaging and Molchan’s Error Diagram

To determine whether spatial variation of a and b values was correlated with the location of repeating earthquakes, we investigated size distribution of background earthquakes in the different areas where repeating earthquakes were observed using dense seismographic networks. The period of background seismicity was consistent with the repeating earthquake catalog (Table S1 in the supporting information). The cross-sectional a and b values were calculated for the different regions using the procedures listed in Text S3 and Figure S6. The examples of Gutenberg-Richter diagram for the selected grids showing minimum and maximum number of earthquakes in each study area are shown in Figure S7. Here we obtained the best choice of cutoff magnitude by evaluating the goodness of fit based on the difference in the cumulative number of earthquakes for each magnitude bin between the observed (NA) and that of predicted from regression line (Si = 10a−bM):

\[
\text{Goodness of fit} = 100 - \left(\frac{100}{M_{\text{max}}} \sum_{M_i} |N_i - S_i| / \sum_{M_i} N_i \right)
\]

where M_i represents each magnitude bin. The choice of M_{\text{max}} can be determined by the highest goodness of fit for the representative a and b value measurement in each computational grid. As shown in Figure S7, the M_{\text{max}} a and b value is somewhat unstable (e.g., in Parkfield the two extreme cases reveal very different M_i of 1.1 and 1.9), indicating the strong heterogeneity in seismicity behavior along the fault. We therefore conducted a series of sensitivity tests to better determine the cutoff magnitude and grid size (please see Text S3 for the details).

The along-strike cross sections of a and b values in Parkfield were computed using 0.5 × 0.5 km grid node spacing. At each grid node we determined a and b values if at least 200 events were available. In Figure 2, we found a large portion of repeating events located within areas with a values >3.0 and b values >1.0. By testing various threshold a and b values, we attempted to obtain by visual inspection the best correlation between location of repeating events and areas showing greater than the threshold values. The threshold for a value between 3.16 and 2.93 generated a rapid growth in relative fraction of repeater areas (the fraction of high a and b value area made up by repeating events) from 30% to 65% (Figure S8). The relative fraction of repeater area therefore is plotted against the threshold to show how the percentage changes with the threshold (Figure S9). To obtain the best choice of threshold, we applied Molchan’s error diagram [Molchan, 1997] by showing miss rate as a function of a and b value thresholds in Figure 3. Here the "miss
Figure 1. (a) Map of background earthquakes (black crosses) and repeating earthquakes (red dots) with focal depths of less than 25 km in the Chihshang and Hualien segments in Taiwan. (b) Map of seismicity and repeating earthquakes in California with focal depths of less than 15 km for the Parkfield and Hayward fault segments and less than 10 km for the Calaveras fault segment. The green traces denote the active faults, and the background color represents the topography.

Figure 2. (a and b) Cross-sectional view of $a$ and $b$ values along the strike of the San Andreas Fault in the Parkfield region. The repeating earthquakes are denoted by open red circles. The $a$ and $b$ values were obtained from at least 200 earthquakes in the surrounding of each $0.5 \times 0.5$ km grid. (c and d) The same cross-sectional view of $a$ and $b$ values with black and gray showing the $a$ and $b$ values greater and smaller than threshold values (THR), respectively. Using this THR, the black areas illustrate the areas showing $a$ values equal to or larger than THR, whereas the gray areas are those with values lower than THR. Here the thresholds for $a$ and $b$ values are those values ranked in the top 44% and 50%, respectively (determined in Figure 3, corresponding to $a$- and $b$-thresholds of 3.07 and 0.88).
"rate" is defined as the number of repeating earthquakes located outside the grid of high $a$ or $b$ values (>threshold) divided by the total number of repeating earthquakes. The tendency of "repeating events occurring within high $a$ and $b$ values areas" at Parkfield can be confirmed if the observations (open black circles) show lower miss rate compared to the random miss rate (line connecting [0, 100] to [100, 0]) and 95% confidence level of random miss rate (curve connecting [0, 100] to [100, 0]). As shown in Figure 3, the best performance favoring this tendency is achieved when the vertical distance to the random behavior reaches its maximum [Zechar and Jordan, 2008], as indicated by the vertical black arrow for Parkfield data. The best threshold selection, i.e., "ranking" of $a$ and $b$ values can therefore be chosen (the filled circles in Figure 3). Here the ranking is determined by first sorting the $a$ and $b$ values computed in each grid by increasing size, and then converting them to percentage. For example, the threshold corresponds to the top 44% of $a$ values and the top 50% of $b$ values obtained for Parkfield in Figure 3. These correspond to an $a$ value of 3.07.
the areas showing $a$ values equal to or larger than the threshold, whereas gray areas are those with values lower than threshold. The distributions of $a$ and $b$ values in Figure 2 are similar in terms of (1) the fact that patches with higher $a$ values (>3.07) and $b$ values (>0.88) are mainly located in northwestern shallower areas (<7.5 km) and southeastern deeper areas (>8 km) of the fault plane and (2) repeating earthquakes tend to occur within or near these patches. The threshold remains similar in Parkfield data even if different sized square grid cells with a different selection of minimum numbers of earthquakes in the surrounding of each grid are used. This suggests that a close relationship exists between the occurrence of repeating earthquakes and abundant small-magnitude background earthquakes.

### 4. Characteristics of RES Locations in California and Eastern Taiwan

Examination of all study areas showed similar tendencies in error diagrams with the exception of the Hualien area. In Figure 3, open circles with different colors are located below the random miss rate except for the high $a$ and $b$ values. Using this threshold, black areas in the lower plots of Figure 2 illustrate the areas showing $a$ values equal to or larger than the threshold, whereas gray areas are those with values lower than threshold. The distributions of $a$ and $b$ values in Figure 2 are similar in terms of (1) the fact that patches with higher $a$ values (>3.07) and $b$ values (>0.88) are mainly located in northwestern shallower areas (<7.5 km) and southeastern deeper areas (>8 km) of the fault plane and (2) repeating earthquakes tend to occur within or near these patches. The threshold remains similar in Parkfield data even if different sized square grid cells with a different selection of minimum numbers of earthquakes in the surrounding of each grid are used. This suggests that a close relationship exists between the occurrence of repeating earthquakes and abundant small-magnitude background earthquakes.

### Table 1. The $a$ and $b$ Value Thresholds (THR) Determined Using Molchan’s Error Diagram

<table>
<thead>
<tr>
<th>LOCATION</th>
<th>Best Ranking of $a$ Value</th>
<th>Best $a$ THR</th>
<th>Best Ranking of $b$ Value</th>
<th>Best $b$ THR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkfield</td>
<td>44%</td>
<td>3.07</td>
<td>50%</td>
<td>0.88</td>
</tr>
<tr>
<td>Hayward</td>
<td>35%</td>
<td>2.83</td>
<td>29%</td>
<td>0.94</td>
</tr>
<tr>
<td>Calaveras</td>
<td>9%</td>
<td>3.39</td>
<td>22%</td>
<td>1.12</td>
</tr>
<tr>
<td>Chihshang</td>
<td>36%</td>
<td>3.82</td>
<td>21%</td>
<td>0.99</td>
</tr>
<tr>
<td>Hualien</td>
<td>27%</td>
<td>4.34</td>
<td>70%</td>
<td>0.89</td>
</tr>
</tbody>
</table>

and a $b$ value of 0.88. When the observations (open circles) are clearly located below the 95% confidence interval of random miss rate, this was treated as a statistically significant relationship between the locations of repeating earthquakes and high $a$ and $b$ values.

Using this threshold, black areas in the lower plots of Figure 2 illustrate the areas showing $a$ values equal to or larger than the threshold, whereas gray areas are those with values lower than threshold. The distributions of $a$ and $b$ values in Figure 2 are similar in terms of (1) the fact that patches with higher $a$ values (>3.07) and $b$ values (>0.88) are mainly located in northwestern shallower areas (<7.5 km) and southeastern deeper areas (>8 km) of the fault plane and (2) repeating earthquakes tend to occur within or near these patches. The threshold remains similar in Parkfield data even if different sized square grid cells with a different selection of minimum numbers of earthquakes in the surrounding of each grid are used. This suggests that a close relationship exists between the occurrence of repeating earthquakes and abundant small-magnitude background earthquakes.

### 4. Characteristics of RES Locations in California and Eastern Taiwan

Examination of all study areas showed similar tendencies in error diagrams with the exception of the Hualien area. In Figure 3, open circles with different colors are located below the random miss rate except for the green circles. This suggests that the statistical association between the locations of repeating earthquakes and large $a$ and $b$ values is significant for the Parkfield, Hayward, Calaveras, and Chihshang Faults. The best threshold selection indicated by filled circles in Figure 3 is also marked in the plots of relative fraction of repeater area as a function of threshold for different areas in Figures S10–S13. The assumption of a single, vertical planar fault for our $a$ and $b$ value determination, although not applicable for the Chihshang Fault, does not seem to impact the analysis of miss rate. The reason is that the projection scheme for background seismicity and the repeaters is the same.

Using the best performance $a$ and $b$ value threshold detailed in Table 1 (corresponding to filled circles in Figure 3), we found that despite the area differences in the $a$ and $b$ value rankings, the corresponding THR are confined to a narrow range (2.8–3.8 for $a$ values and 0.9–1.1 for $b$ values). In Molchan’s error diagram (Figure 3), the data from the Hualien area are located above the curves of random behavior, suggesting that the background seismicity does not correlate with higher $a$ and $b$ values. Unlike in the Chihshang area where the seismicity forms a narrow, SE dipping inclined fault zone from near surface to a depth of 25 km, the background seismicity in Hualien occurs in several clusters with different characteristics (Figure S14). The RESs in Hualien coincide with the cluster underneath the eastern flank of the Central Range. To test if the complex structures in the Hualien area lead to a biased $a$ and $b$ value computation, we narrowed the selection of seismicity and RESs to the cloud-like cluster underneath the Central Range (location please see Figure S15a) for analysis. As shown in Figure S15b, the resulting error diagram is similar to the green circles in Figure 3, indicating that unclear fault zone architecture may bias the frequency-magnitude distribution of earthquakes. Another complicated segment in the southernmost LVF in eastern Taiwan, the Taitung area, was also examined. In Figure S16, the data points are located above and very close to the curves of random miss rates, suggesting of the absence of a statistically significant spatial correlation between RES and high $a$ and $b$ values. The error diagram for the map of resolved $a$ and $b$ value variations for shallow earthquakes in Taiwan also shows a significant but anticorrelation between RES and high $a$ and $b$ values (Figure S17). Results for the whole of Taiwan, Hualien, and Taitung areas in particular seem to suggest that a mature, planar-like structure is needed for high correlations between RES locations and frequency-magnitude relationships. Also note that the repeating events tend to have highest occurrence rate at/near the magnitude range showing highest background seismicity rate (Figure S18). The exception occurs in the Hualien and Taitung areas, where the largest number of repeating events appears in a relatively large magnitude range (M3.5 to 4.0).
When the a and b value correlation is analyzed for each study area, to discriminate the behavior of repeating events from nonrepeating events in Figure S19, we found that the majority of repeating events tend to show relatively high a and b values, whereas the nonrepeating events reveal relatively low a and b values. The similar behavior is found in Calaveras, Hayward, and Chihshang Faults (Figures S19a–S19d) but not in Hualien and Taitung regions where the repeating and nonrepeating earthquakes are preferentially in locations with relatively low a and b values (Figures S19e and S19f). This suggests that the asperity distribution may also play a role in the controls of repeaters’ location.

5. Discussion and Conclusions

What is the physical meaning of high a and b values? Laboratory experiments and numerical modeling show that high b values are indicative of low stress, whereas low b values correspond to high accumulated and differential stress [Scholz, 1968; Goebel et al., 2013; Kun et al., 2013]. Thus, b values can be used to map heterogeneous stress conditions in the Earth’s crust [Wiemer and Wyss, 2002; Torman et al., 2014; Schorlemmer et al., 2005]. High a values observed in RES sites indicate a high seismicity level, which is useful in excluding the possibility of “locked” faults. The high a value (~3) and b value (~1) inferred here are a result of a significant number of small earthquakes. Such high rates of microseismicity are commonly observed in creeping faults [Scholz, 1990; Malservisi et al., 2005], where strain is accumulated in the transition zone from locked to creeping patches [Vidale et al., 1994; Nadeau et al., 1995; Sammis and Rice, 2001; Waldhauser and Ellsworth, 2002]. Among statistical techniques, frequency-magnitude distribution and fractal dimension (D) are also regarded as effective measures for distinguishing locked from creeping fault sections [Wiemer and Wyss, 1997; Wyss et al., 2004]. The predominantly high b value (> 1.2) and D~2b have been proposed as more suitable descriptors of the creeping portion of the fault [Wyss et al., 2004]. Amelung and King [1997] found that the relationship between the summed moment release and fault length is different for noncreeping and creeping faults. Assuming constant stress drop for all earthquakes, the number of events (N) scales with fault length (L) as N~L−2 for noncreeping faults, while for creeping faults the scaling is N~L−3. However, these statistical relationships vary by style of faulting, tectonic loading rate, and aseismic slip behavior [Amelung and King, 1997; Wyss et al., 2004; Schorlemmer et al., 2005; Torman et al., 2014].

The locations where creeping and locked fault areas abut have been commonly found to be delineated by the foci of small repeating earthquakes [Burkmann et al., 2000; Beeler et al., 2001; Sammis and Rice, 2001; Johnson and Nadeau, 2002; Uchida et al., 2003; Nadeau and McEvilly, 2004; Chen et al., 2008]. As inferred by numerical, analytical, and laboratory experiments, repeating events are driven by steady aseismic creep and are either located on the edge of larger asperities or far from them [Beeler et al., 2001; Sammis and Rice, 2001; Johnson and Nadeau, 2002; Nadeau and McEvilly, 2004]. In contrast to magnitude-frequency distribution, RES (as a small portion of background seismicity on the same fault) represents the high creep-rate location [Park et al., 2005]. High a property should be further investigated through geophysical measurements. The location of RESs in Parkfield was found to be correlated with high Vp and high Vp/Vs ratios, suggesting that high pore fluid pressure may be present and weaken the deep fault zone [Nadeau et al., 1994]. In the Chihshang area, the distribution of RESs coincides with a region of high Vp and high Vp/Vs [Huang et al., 2014]. This is similar to Parkfield, where high pore pressure may play an important role in weakening the fault and generate large numbers of small earthquakes. In the northernmost segment of the LVF (the Hualien area), the complex tectonic provinces reveal sharp along-strike variation in rheologic behavior. In this segment, the RES located beneath the eastern flank of the Central Range probably marks a transition boundary between the continental crust in the west and the oceanic crust in the east. The high Vp/Vs feature, however, is not obvious here, suggesting that other mechanisms are responsible for the occurrence of the RES. This mechanism may also control the nonplanar fault structure apparent in the Hualien area.
To conclude, in this study the combined information from two individual measurements (magnitude-frequency and repeating earthquake distribution) provides an accurate description of the creeping area on a fault. Using seismicity and repeating earthquake catalogs from five well-documented RES-rich areas, we found that regions subject to frequent repeating events and possessing linear/listric fault structures were statistically significantly correlated with higher a values (2.8–3.8) and b values (0.9–1.1), whereas other regions less qualified as simple planar fault planes were negatively correlated. A sufficient number of small earthquakes is needed for the occurrence of both high a and b values. We hypothesize that, given that the deformation conditions within a fault zone are suitable for a planar fault plane, the location of repeating earthquakes can be best described by an a value of approximately 3 and a b value of approximately 1. This feature of a and b values may be useful for foresee the location of RESs for measuring creep rate at depth. Further investigation of RES-rich areas may allow testing of this hypothesis.

Acknowledgments
We thank Robert Nadeau, Bill Ellsworth, Dennise Templeton, and Zhigang Peng who provided repeating earthquake catalogs for study. The repeating earthquake data set is from cited references listed in Table S1. Please contact the corresponding authors for accessing repeating earthquake catalogs. This work was supported by Taiwan MOST grant 103-2116-M-003-01-MY3.

References


