Bulletin of the Seismological Society of America, Vol. 99, No. 1, pp. 389–395, February 2009, doi: 10.1785/0120080942

Ⓔ

Short Note

Quantifying Natural Fault Geometry: Statistics of Splay Fault Angles

by Ryosuke Ando,* Bruce E. Shaw, and Christopher H. Scholz

Abstract We propose a new approach to quantifying fault system geometry, using an objective fit of the fault geometry to a test function, specifically here a fault branch. Fitting a Y-shaped object using a cost function to dextral faults in California, we find a number of significant results arising from use of a a systematic, objective, quantitative approach. (1) The largest angle of the branch structure is generally very close to 180°, implying that the branch is a splay fault off the primary through-going fault. (2) The distribution of the smallest angle, the splay angle, has a peak near ±17°, symmetric about the primary fault. (3) These features appear independent of scale. These results are not yet explained by any theory, and they pose new questions and constraints for the physics of fault system formation and behavior.

Online Material: Event information and color figures.

Introduction

Faults do not act as isolated objects, but rather as parts of complex fault systems. Quantifying the geometry of fault systems remains an important and, in many ways, an unsolved problem. Here, we propose a new approach to the problem, using objective criteria to match geometrical objects to mapped faults. Specifically, we focus on a geometrical object of a branching fault and numerically fit with a cost function the mapped fault system to this geometrical object. This approach allows for a systematic, unbiased, quantitative measure of a significant aspect of fault system geometry.

We examine the particular geometrical object of fault branches for a few reasons. One reason is that it is a geometrical object that does not have an explicit scale, and because fault structures appear similar from ranges of hundreds of meters to hundreds of kilometers, we would like to find scale-independent measures of geometry—and, moreover, to test the apparent scale invariance.

Fault branches are not, of course, the only relevant geometrical irregularities—fault stopovers (e.g., Wesnousky, 2006) and bends (e.g., King and Nabelek, 1985) are other examples. These geometrical irregularities play an important role in earthquake behavior and are used to segment faults to try to understand fault and earthquake behavior.

Fault segmentation is seismologically important in controlling earthquake rupture processes including nucleation, propagation, and termination of ruptures (e.g., King and Nabelek, 1985; Shaw, 2006). Splay faults are particularly important for dynamic fault segment interactions and earthquake rupture path selectivity. There are examples in recent earthquakes that show ruptures branched off into splay faults from primary faults from the evidence of surface rupture traces such as the 1992 Landers earthquake (Sowers *et al.*, 1994), the 2001 Kunlun earthquake (Xu *et al.*, 2006), and the 2002 Denali earthquake (Eberhart-Phillips *et al.*, 2003). A difference in a rupture path leads to differences in seismic hazards and, further, splay fault ruptures may cause disastrous damage; for example, it is thought that the rupture propagation into an offshore upward splay fault branching from a subduction plate boundary was responsible for intensive Tsunami generation during the 1944 Tonankai earthquake (e.g., Park *et al.*, 2002). Several recent theoretical studies suggest that one of the key parameters in determining the rupture extension into splay faults is the splay angle (e.g., Kame *et al.*, 2003; Ando and Yamashita, 2007; Bhat *et al.* 2007). However, the general characteristics of splay fault geometry are poorly understood. Here, we quantify splay angles for the case of strike-slip faults.

The quantification of natural fault geometries using mapped surface traces has been widely conducted; however, these analyses mostly dealt with lengths or separation distances of fault segments (e.g., Cowie *et al.*, 1996) or the roughness (topography) of fault surfaces (e.g., Brown and Scholz, 1985).

The difficulties in the analysis of splay faults should be in the definition of objective criterion to geometrically quantify the structure. For instance, regarding the splay angle, the value may change depending on the observer's visual perception: one could lay a protractor along different portions of a

*Also at Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan 305-8567.

rough fault segment. The size of the protractor may also affect the results. Dealing with a large number of data is another difficulty because the fault geometry is usually analyzed by visual and manual procedures.

In order to overcome these difficulties, we propose an objective method to quantify splay structures based on a pattern-matching analysis using a computer. The pattern matching is a standard methodology in computer-based geometrical pattern recognition and data mining algorithms where we search for a certain geometry that is similar to a test function prepared in advance to mimic a targeted geometry. To model the splay fault geometry, we employ a set of three line segments connected to each other at one end as the test function in this study. The degree of matching can be evaluated in terms of an index similar to a correlation coefficient so that the analysis becomes quantitative and objective.

In this article, we concentrate on the presentation of our methodology and our observational results; modeling of what we observe will be done in subsequent papers. The purpose of this article is to present geometrical characteristics that are found for the first time.

Method

We employed a pattern-matching algorithm for the automatic detection of splay fault structures. In our method, any part of fault traces in the data is compared with a test function prepared in advance, and the degree of their similarity is evaluated; finally, test functions well fitted to the fault traces are retained. For the geometry of the test function, we simply implement a set of three line segments with the same length that are connected with each other at a point. To adopt any splay angles in the fault traces, we allow any orientation for the assumed line segments. Examples of test functions fitted to fault traces are shown in Figures 1 and 2.

The shape of the test function is defined by the length of each line segment R and the two angles Φ_p and Φ_s, which respectively denote the angles made by the most separated and the closest two line segments. (The subscripts p and s are chosen as they will be found to be associated with the primary and splay faults, respectively.) For instance, if the test function is T shaped, $\Phi_p = 180°$ and $\Phi_s = 90°$. As shown later, we can measure not only the splay angles through Φ_p and Φ_s but also the scale length of the splay faults through R.

The degree of similarity between the test function and the fault traces is quantified in terms of the index F defined as follows. For simplicity, we define F for an individual line segment, not for the whole test function consisting of the three line segments. In this sense, F is defined as

$$F = \frac{1}{R}\int_0^R \max_{\mathbf{r}'}\left[\exp\left\{-\left(\frac{\mathbf{r}-\mathbf{r}'}{R}\right)^2/\sigma_r - \left(\frac{\theta-\theta'}{\pi}\right)^2/\sigma_\theta\right\}\right] dr, \qquad (1)$$

where $\mathbf{r}(r)$ and $\mathbf{r}'$ are respectively the location on this line segment and a point on the fault trace. θ and θ' represent the orientations of the normal vectors for this line segment and the fault trace at locations $\mathbf{r}$ and $\mathbf{r}'$, respectively. The max function searches over the fault traces to find the value closest in both distance and orientation to the line segment, thereby maximizing the function (a Gaussian) inside. That is, for each value of $\mathbf{r}$ along the line segment, we search over the fault traces $\mathbf{r}'$ to find the trace which best matches both the location and the orientation of the line segment at $\mathbf{r}$. The integration is taken over $\mathbf{r}$, the distance along the line segment. Thus no extra credit is given for having more than one nearby fault trace, just credit for the best fitting fault trace. Because the integrand is the shape of the Gaussian function, if the line segment completely overlaps with a fault trace, F takes 1; on the contrary, if both are far apart, F becomes 0. The width of the Gaussian is controlled by σ_r and σ_θ so that the sharpness of the fit is also controlled, where smaller or larger values give tighter or looser fits. Allowing a loose fit to some extent with σ_r and σ_θ is, in fact, essential in this study because the actual fault geometry is not a straight line but can still be reasonably approximated by the line. In other words, the fault roughness is absolved by a lobe existing around the line segment of the test function, which corresponds with the skirt of the Gaussian.

For the sake of implementation, we need to expand on the previously mentioned matching procedure supposing a single line segment to a procedure assuming the test function with three arms. We follow two stages in the procedure. As the initial stage, we first choose an arbitrary point on the fault in the data to locate the center of the test function. Next, we pick up one line segment whose one end is fixed at this chosen point, and we evaluate its matching with the fault geometry at any orientation. In this way, we can obtain the index $F(\varphi)$ as the function of the orientation $\varphi(0 < \varphi < 360°)$. The orientations corresponding to the three largest peaks in the $F(\varphi)$ distribution become the possible orientation of the three arms of the test function. Then, the threshold F_c is applied to the three largest peaks for their values P_1, P_2, and P_3. If all of these peak values exceed F_c, then this test function is qualified in the initial stage. In the second stage, we first iterate the initial stage to provide the candidates of the qualified test functions along the fault traces. Second, we select the most qualified candidates within certain neighborhoods, which are defined as the test functions having the largest values of $P_1 + P_2 + P_3$ with them; specifically, to determine the size of the neighborhood, we assume the minimum distance between the qualifiers R_o taken typically to be a fraction of R. This selects the optimal fit of a branch while excluding nearby local poorer fits to the same branch. We find that the results are not sensitive to R_o. Finally, we can obtain a set of test functions well fitted to the fault trace geometry in the data.

In addition, we would like to mention that this method is readily applicable to raster formatted data as well as the vector formatted data set used in this study. Rasterized data

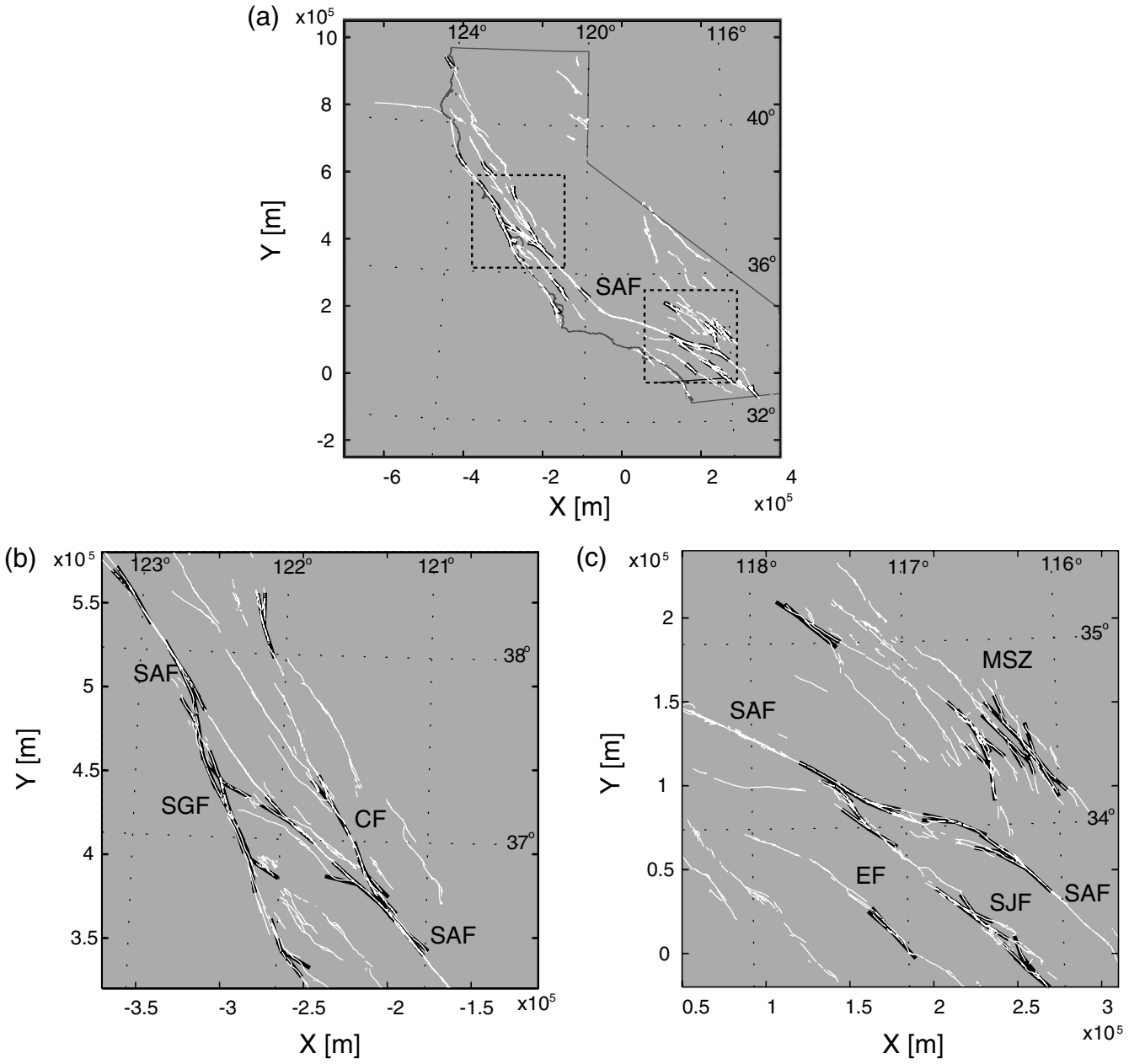

Figure 1. Surface faults and fitted fault geometry. Plots for (a) the entire California and the zooms of the boxed areas for (b) northern and (c) southern California. White and black lines represent catalogued faults and fitted fault geometry (test function), respectively. $R = 20$ km. Ⓔ A high-resolution color version of this figure is available in the electronic edition of *BSSA*.

are more easily obtained by digital images. Thus, it is interesting to apply this method to such data, including digitized hand-drawn sketches and pictures taken in the field or laboratory.

Results

We analyzed the splay angle of faults in California using the digital database of quaternary active faults (Bryant, 2005). This database includes the location of 18,730 fault segments represented in vector format data points with spatial resolution of a few hundred meters. Although the data are only about the surface trace, they can be used to represent the geometry of strike-slip faults at depth through the seismogenic layer because their dip angles tend to be 90°. We choose to use the catalogued dextral faults, which are dominant due to the tectonic setting in this region. To avoid near surface characteristics in the fault geometry, we focus on fault structures larger than 5 km by choosing R values that are a substantial fraction or larger of the seismogenic layer (Scholz, 2002) and can be considered as major persistent structures throughout this layer.

We begin with the case of $R = 20$ km, which is the length of each line segment of the test function. The surface

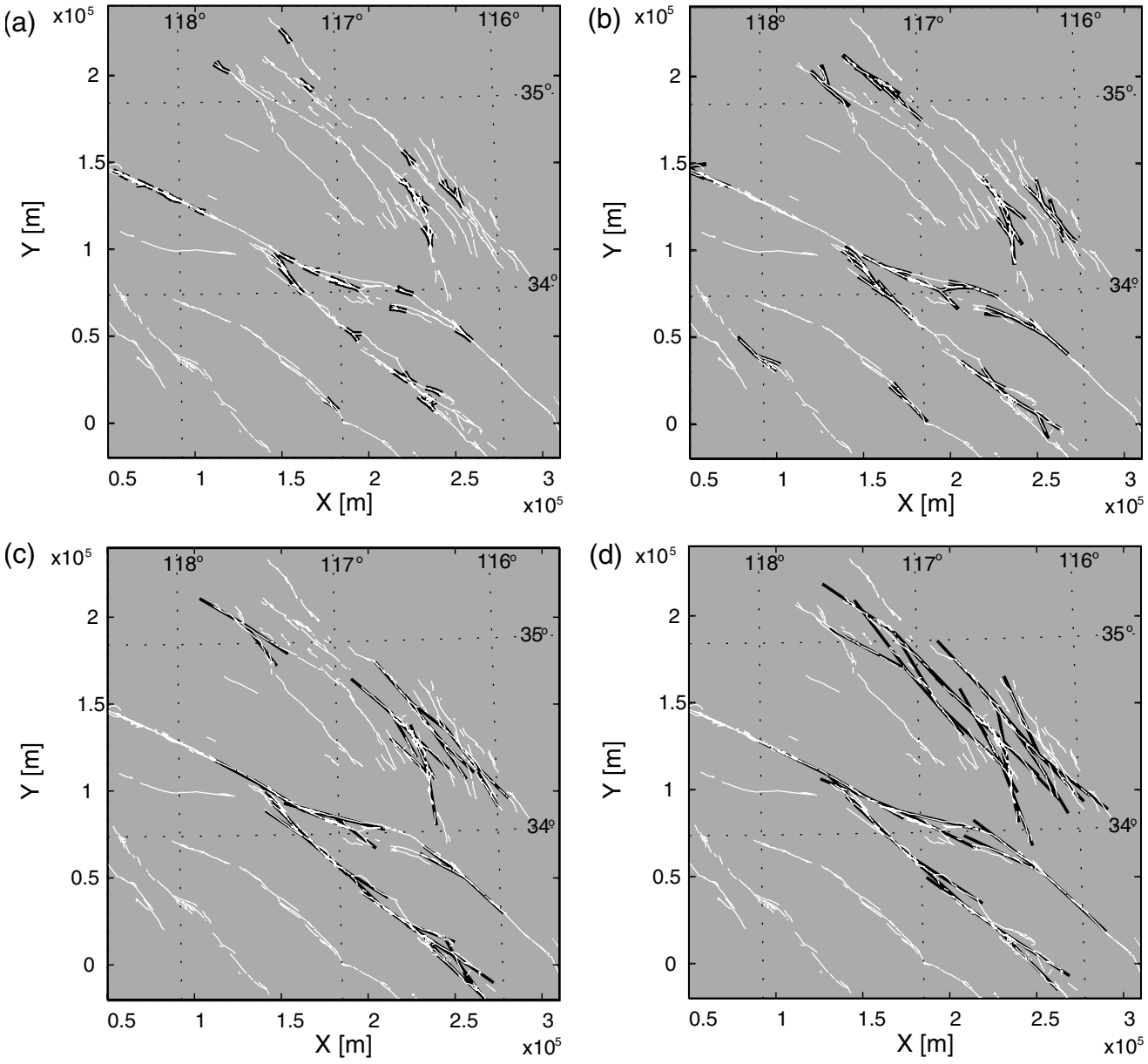

Figure 2. Surface faults and fitted fault geometry. Plots for different scales (a) $R = 5$, (b) 15, (c) 30, and (d) 50 km in southern California. Ⓔ A high-resolution color version of this figure is available in the electronic edition of *BSSA*.

traces of these dextral faults in the catalog are plotted in Figure 1 with fitted fault geometries or test functions obtained by the present pattern-matching analysis. We select a set of parameters $F_c = 0.4$, $\sigma_r = 0.001$, and $\sigma_\theta = 0.05$. These values of the parameters are confirmed to return the most appropriated fitting results after a systematic parameter study, which we will discuss later.

First of all, we demonstrate the reliability of the method and analysis results. The quality of the fitting can be confirmed in the zoom shown in Figure 1b,c for central and southern California, respectively. We see that each line segment of the fitted fault geometry follows corresponding fault traces in the catalog, and splay points in the fault traces are centered by the junctures of the fitted test functions. As we stated in the previous section, it is also confirmed that the fitted line segments are located on the center of oscillating fault traces, with the amplitudes of the oscillation being reasonably small, as seen in the figures. The performance in detecting splay fault locations, which can track nearly every splay fault supposed to be marked, is satisfactory as well. Note that the fitted fault geometries are exclusive within the area $R_o = 0.25R$ (throughout the analysis); some splays in densely distributed areas are therefore not picked due to this restriction.

Before detailing the investigations, it is worth characterizing the overall splay fault locations and geometry. Figure 1a shows that most of the fitted fault geometries are closely placed near the San Andreas fault system (SAF). Moreover,

the splay faults are particularly localized in the central and southern California regions along the SAF as boxed in Figure 1a. On the contrary, the rest of the areas have much less splay faults and the dextral fault geometry appears to be much simpler in these areas.

In central California (Fig. 1b), the fitted fault geometries are closely located near the junctions of SAF and its major splays: the Calaveras fault (CF) and San Gregorio fault (SGF). There are interesting characteristics observed in this area; these major splays are oblique to the SAF and parallel to each other, but smaller splays are not. These splays do not seem to cut the primary faults. In contrast to northern California, a cluster of minor fault segments exists apart from the SAF in southern California, which is called the Mohave shear zone (MSZ). The distributions of the fitted fault geometries are shown to be denser in the MSZ than in the region along the SAF and its major splays: the San Jacinto fault (SJF) or the Elsinore fault (EF).

Scale Dependence and Independence of Splays

The analysis is further applied to different scales over a decade from 5 km to 50 km, assuming various lengths of the test function R. Examples in southern California are shown in Figure 2. First, we should emphasize that the fault structures of specific length scales can be detected by the proper choice of R. For instance, at the junction of SAF and SJF located near $(X, Y) = (1.5, 1)$, this position is not recognized as the location of a splay for the cases of $R = 5$ km (a) and 15 km (b) due to the existence of a gap of about 4 km width between them, while this gap is small enough for the cases of $R = 30$ km (c) and 50 km (d) where the splay faults are marked. On the contrary, smaller structures are captured by smaller R cases. The four small splay faults, located along SAF within $X = 0.5$ to 1.0×10^5 m and $Y = 1$ to 1.5×10^5 m, are large enough for the case of 5 km but too small for the cases of $R > 5$ km. These observations demonstrate that this method can capture the fault structures of desired scales with the adjustment of R.

The scale dependence observed in the number of splay faults N_R is a significant characteristic as roughly recognized in Figure 2a–d. By counting the number for the entire state of California, we found that N_R is a decreasing function of R as $N_R = 137$, 97, 67, 50, 39, 34, and 27, respectively, for $R = 5$, 10, 15, 20, 30, 40, and 50 km. It is interesting that N_R appears to follow a power law function over a decade in scale length, but this is too limited a scale range to really determine such functional forms. We can also observe in Figure 2 that splays themselves can have smaller splays. This appears to coincide with the well-known fractal geometry of natural faults (Aviles *et al.*, 1987; Okubo and Aki, 1987).

Statistics of Primary and Splay Fault Angles

For angles made by three arms of each fitted test function in Figures 1 and 2, we can recognize a quite common characteristic on the geometry of the splay fault structure regardless of the scale and the region. That is, two of three arms fall into almost a straight line. This is supported by a statistical analysis as shown in Figure 3; we can see that Φ_p (see the inset and the Method section for the definition) has a well-defined peak at about 180° in the statistical distribution over the entire state of California. This implies that Y-like shapes of symmetric branches are very rare in natural fault geometry. So far, we use the term of splay faults without the definition. But our analysis clearly demonstrates that the majority of bifurcated structures in natural faults are actually what has been inferred from splay faults by experiences in the field, which consist of planar primary faults and subsidiary oblique splay faults. This geometric asymmetry is an important clue to modeling the formation of the splay faults.

The statistical analysis is finally applied to evaluate the obtained angle of the splay faults for the entire state of California. The splay fault angle technically corresponds to the angle between the two closest line segments of the test function, Φ_s. The probability distributions of the obtained splay angles Φ_s are shown in Figure 4 for various R. The sign of the angle distinguishes the orientation of a splay; Φ_s is measured clockwise from the primary fault so that the plus and minus signs denote splays located on the right- and left-hand sides of a primary fault (see the inset).

In Figure 4, there are four significant properties in the distribution of the splay angles Φ_s:

1. The splay angle distributions show sharp peaks at about $-/+17°$ (refer the caption for the median of each case).
2. Unexpectedly, the distributions of the left- and right-hand sides are remarkably symmetric considering the entire fault system, with the splays symmetrically distributed on either side of the primary faults.
3. The splay angle distributions for any R cases are similar in their shapes and peak values. This means that the geo-

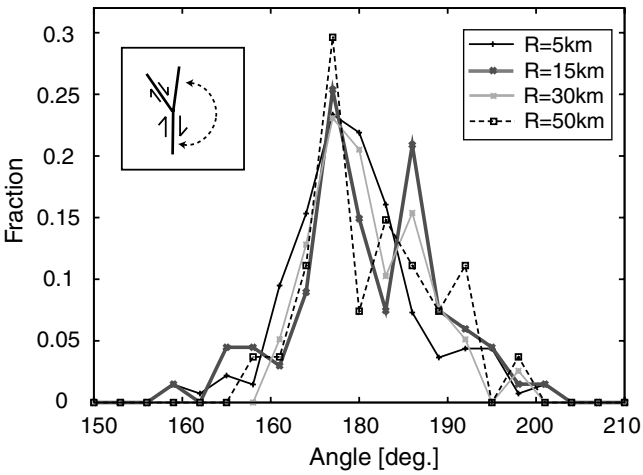

Figure 3. Probability distribution of primary fault bend angles for different scales. The plot was made at intervals of 3°. The medians of the cases of $R = 5$, 15, 30, and 50 km are 178°, 179°, 180°, and 181°.

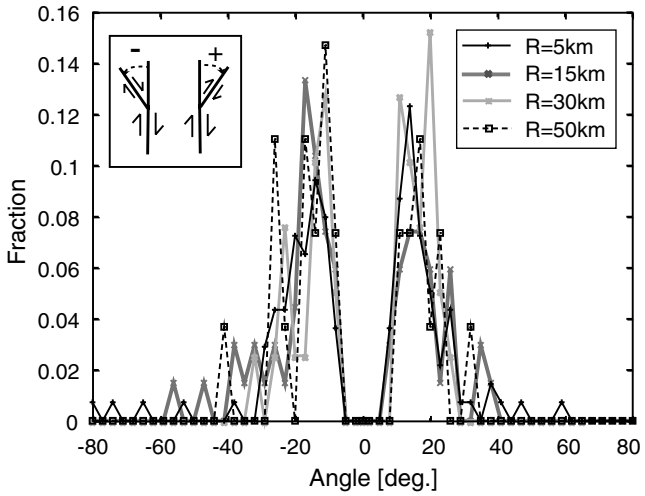

Figure 4. Probability distribution of splay fault angles for different scales. The medians of the cases of $R = 5$, 15, 30, and 50 km are $-17/16°$, $-17/18°$, $-12/16°$, and $-17/17°$ for the left- and right-hand-side branches defined in the inset.

metrical relations between a primary fault and a splay fault are self-similar, independent of the length of the splay faults.

4. Looking at the distribution of the absolute values of the angles, the distribution is asymmetric, with the peak angle close to its minimum value. In other words, the minima are sharply defined but the maxima are not.

Some roughness seen in the distribution of large R cases in Figures 3 and 4 should mainly be due to the relatively small sampling number. It is obviously confirmed that the distributions tend to be much smoother as the sampling number increases. On the other hand, it should be noted that the analysis covers the entire state of California, which includes various tectonic settings from the SAF as a mature plate boundary to the Mohave shear zone as a cluster of immature faults. Further investigation with division regarding local tectonic conditions may give us more detailed statistics behind the current overall features. However, the previously mentioned features in the distribution of the splay angles are still considered to be quite robust and tell us about the fundamental characteristics of splay fault structures.

The first and second features in the list of significant properties imply that the fault tip stress field may not play a predominant role in the formation of the splay faults, because its normal stress component is antisymmetric against the fault plane. Regarding the third feature, it should be mentioned that this can add a different perspective and data to the long-standing discussion of whether the fault geometry is self-similar or self-affine and whether it is fractal or not; there is a limitation in those discussions, as they have been mostly based on the observation of fault surface topography (e.g., Brown and Scholz, 1985; Sagy *et al.*, 2007) or the length of fault segments (e.g., Cowie *et al.*, 1996). Deeper understandings should be necessary for the individual physics of the initiation and evolution of those characteristic structures. In terms of dynamic rupture path selectivity, referring to Kame *et al.* (2003), the observed splay angle of 17° suggests that a rupture tends to propagate to either primary or splay faults. But if the rupture speed is almost at the terminal (Rayleigh-wave speed), the rupture can be bifurcated to both.

Validation of the Results

We validate the robustness of the obtained results in terms of the parameters for the threshold F_c and the sharpness of the fit σ_r and σ_θ. Figure 5a–c shows these effects on the obtained splay angles, where the sign of the angles is not considered.

We first find from Figure 5a that the choice of F_c does not strongly affect the obtained splay angles as long as we apply proper F_c that gives a statistically meaningful number of splays. For instance, the cases of $F_c < 0.7$ show quite similar results with sampling 55 splays in the case of $F_c = 0.4$ and $R = 20$ km; on the contrary, $F_c = 0.7$ shows a case of poor statistics picking only eight splays. The choice of threshold value can sometimes be a problem in pattern matching; however, this analysis shows quite stable behavior over a wide range of the parameter values.

For the sharpness parameters, we find that, as seen in Figure 5b, a variation of σ_r within half of and double most optimal values could lead the relatively large differences up to 10°. Fortunately, however, the variation in the quality of the fitting is also significantly noticeable when visually checking the fits on the map: values that are too large (loose fit) recognize parallel strands as splays by mistake, which also causes higher splay angles to appear in the statistics, while values that are too small (tight fit) tend to miss obvious splays due to their roughness in the geometry. In contrast, the variation in σ_θ (Fig. 5c) has a smaller effect although its implementation is of great help to reduce mistakes in densely distributed areas. Thus, the parameter tuning to obtain the most optimal fit is quite straightforward, and the obtained results become reliable once σ_r is properly chosen. It should be emphasized that the obtained results are quite robust because of the stableness enabled by this method.

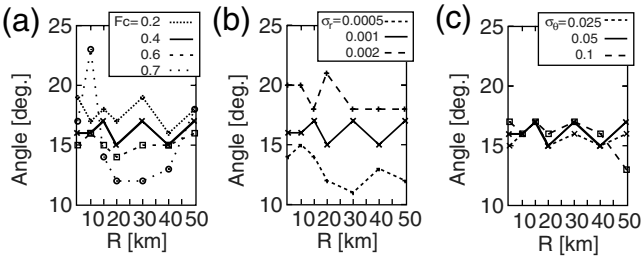

Figure 5. Medians of splay angles for different (a) F_c, (b) σ_r, and (c) σ_θ ($R = 5$ km).

Data and Resources

Fault traces used in this article came from the Digital Database of Quaternary and Younger Faults from the Fault Activity Map of California, version 2.0 (compiled by W. A. Bryant in 2005). Data can be obtained from California Geological Survey web site at http://www.consrv.ca.gov/CGS/information/publications/QuaternaryFaults_ver2.htm (last accessed September 2006).

Acknowledgments

The authors thank B. Duan and D. D. Oglesby for critically reviewing the manuscript. R. Ando was supported by a postdoctoral fellowship from the Japan Society for Promotion of Science during 2005–2007. B. E. Shaw was supported by National Science Foundation Grant Number EAR03-37226 and by the Southern California Earthquake Center.

References

Ando, R., and T. Yamashita (2007). Effects of mesoscopic-scale fault structure on dynamic earthquake ruptures: dynamic formation of geometrical complexity of earthquake faults, *J. Geophys. Res.* **112,** B09303, doi 10.1029/2006JB004612.

Aviles, C. A., C. H. Scholz, and J. Boatwright (1987). Fractal analysis applied to characteristic segments of the San-Andreas fault, *J. Geophys. Res.* **92,** no. B1, 331–344.

Bhat, H. S., M. Olives, R. Dmowska, and J. R. Rice (2007). Role of fault branches in earthquake rupture dynamics, *J. Geophys. Res.* **112,** no. B11, B11309, doi 10.1029/2007JB005027.

Bryant, W. A. (Compiler) (2005). Digital database of Quaternary and younger faults from the fault activity map of California, version 2.0, California Geological Survey web page, available at http://www.consrv.ca.gov/CGS/information/publications/QuaternaryFaults_ver2.htm (last accessed November 2008).

Brown, S. R., and C. H. Scholz (1985). Broad bandwidth study of the topography of natural rock surfaces, *J. Geophys. Res.* **90,** no. B14, 2575–2582.

Cowie, P. A., R. J. Knipe, and I. G. Main (1996). Special issue: scaling laws for fault and fracture populations—analyses and applications—introduction, *J. Struct. Geol.* **18,** no. 2–3, R5–R11.

Eberhart-Phillips, D., P. J. Haeussler, J. T. Freymueller, A. D. Frankel, C. M. Rubin, P. Craw, N. A. Ratchkovski, G. Anderson, G. A. Carver, A. J. Crone, T. E. Dawson, H. Fletcher, R. Hansen, E. L. Harp, R. A. Harris, D. P. Hill, S. Hreinsdottir, R. W. Jibson, L. M. Jones, R. Kayen, D. K. Keefer, C. F. Larsen, S. C. Moran, S. F. Personius, G. Plafker, B. Sherrod, K. Sieh, N. Sitar, and W. K. Wallace (2003). The 2002 Denali fault earthquake, Alaska: a large magnitude, slip-partitioned event, *Science* **300,** no. 5622, 1113–1118.

Kame, N., J. R. Rice, and R. Dmowska (2003). Effects of prestress state and rupture velocity on dynamic fault branching, *J. Geophys. Res.* **108,** no. B5, 2265, doi 10.1029/2002JB002189.

King, G., and J. Nabelek (1985). Role of fault bends in the initiation and termination of earthquake rupture, *Science* **228,** no. 4702, 984–987.

Okubo, P. G., and K. Aki (1987). Fractal geometry in the San-Andreas fault system, *J. Geophys. Res.* **92,** no. B1, 345–355.

Park, J. O., T. Tsuru, S. Kodaira, P. R. Cummins, and Y. Kaneda (2002). Splay fault branching along the Nankai subduction zone, *Science* **297,** no. 5584, 1157–1160.

Sagy, A., E. E. Brodsky, and G. J. Axen (2007). Evolution of fault-surface roughness with slip, *Geology* **35,** no. 3, 283–286.

Scholz, C. H. (2002). *The Mechanics of Earthquakes and Faulting*, Cambridge University Press, New York.

Shaw, B. E. (2006). Initiation propagation and termination of elastodynamic ruptures associated with segmentation of faults and shaking hazard, *J. Geophys. Res.* **111,** no. B8, B08302, doi 10.1029/2005JB004093.

Sowers, J. M., J. R. Unruh, W. R. Lettis, and T. D. Rubin (1994). Relationship of the Kickapoo fault to the Johnson Valley and Homestead Valley faults, San-Bernardino County, California, *Bull. Seismol. Soc. Am.* **84,** no. 3, 528–536.

Wesnousky, S. G. (2006). Predicting the endpoints of earthquake ruptures, *Nature* **444,** no. 7117, 358–360.

Xu, X. W., G. H. Yu, Y. Klinger, P. Tapponnier, and J. Van der Woerd (2006). Reevaluation of surface rupture parameters and faulting segmentation of the 2001 Kunlunshan earthquake (M_w 7.8), northern Tibetan Plateau, China, *J. Geophys. Res.* **111,** no. B5, B05316, doi 10.1029/2004JB003488.

Lamont-Doherty Earth Observatory
Columbia University
Palisades, New York 10964-8000

Manuscript received 15 February 2008