Rupture Passing Probabilities at Fault Bends and Steps, with Application to Rupture Length Probabilities for Earthquake Early Warning

Glenn P. Biasi*1 and Steven G. Wesnousky2

ABSTRACT

Earthquake early warning (EEW) systems can quickly identify the beginning of a significant earthquake rupture, but the first seconds of seismic data have not been found to predict the final rupture length. We present two approaches for estimating probabilities of rupture length given the rupture initiation from an EEW system. In the first approach, bends and steps on the fault are interpreted as physical mechanisms for rupture arrest. Arrest probability relations are developed from empirical observations and depend on bend angle and step size. Probability of arrest compounds serially with increasing rupture length as bends or steps are encountered. In the second approach, time-independent rates among ruptures from the Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3), are interpreted to apply to the time-dependent condition in which rupture grows from a known starting point. Length probabilities from a Gutenberg–Richter magnitude–frequency relation provide a reference of comparison. We illustrate the new approach using the discretized fault model for California developed for UCERF3. For the case of rupture initiating on the southeast end of the San Andreas fault we find the geometric complexity of the Mill Creek section impedes most ruptures, and only ∼5% are predicted to reach to San Bernardino on the eastern edge of the greater Los Angeles region. Conditional probabilities of length can be precompiled in this manner for any initiation point on the fault system and thus are of potential value in seismic hazard and EEW applications.

KEY POINTS

• Once a rupture has started, little current research bears on how far it may continue or why it stops.
• The probability of rupture stopping increases at each bend and step it must pass to grow longer.
• Conditional length estimates can be used in scenario earthquake and early warning alert applications.

INTRODUCTION

Earthquake early warning (EEW) systems are designed to warn of impending strong shaking from a large earthquake by exploiting the speed advantage of electronically transmitted signals over seismic waves (Cooper, 1868; Heaton, 1985; Kanamori et al., 1997). Efforts to develop, formalize, and apply EEW methodologies in California have moved forward in concert with advances in seismic instrumentation, telemetry, computers, data storage, and real-time seismological analysis (Heaton, 1985; Kanamori et al., 1997, 1999; Wu and Teng, 2002; Allen and Kanamori, 2003; Kanamori, 2005; Allen et al., 2009; Kohler et al., 2018; Chung et al., 2019; Cochran et al., 2019). Methodologies generally entail the rapid estimation of the magnitude of an earthquake from observations of peak displacement, velocity, and acceleration (Wu and Kanamori, 2005, 2008; Wu et al., 2007) or the predominant period and frequency content (Nakamura, 1988; Allen and Kanamori, 2003; Kanamori, 2005) of the first seconds of the first recorded P wave.

The actual moment released in the first 3–5 s of a large earthquake normally corresponds to an M 6–6.5 earthquake. Early work suggested that the eventual magnitude of an earthquake that continues to grow could be known from how it starts (Olson and Allen, 2005). Later studies have questioned this conclusion and find instead that reliable estimates of final magnitude require more data from extended P-wave displacements (Yamada and Ide, 2008; Noda and Ellsworth, 2016), up
to half or more of the duration of the rupture itself (Meier et al., 2016; Trugman et al., 2019). To estimate magnitude and rupture extent of larger earthquakes, the ShakeAlert system (Given et al., 2018) includes an algorithm named Finite-fault rupture Detector algorithm (FinDER; Böse et al., 2012). FinDER estimates event size based on a finite-fault model of rupture and ground-motion template matching to observed ground motions. The alternative Propagation of Local Undamped Motion algorithm (PLUM, Kodera, 2018) avoids magnitude estimation altogether and instead predicts alert areas from locations of observed strong ground motions and a forward model of ground motion for growth of the alert area. Originally developed in Japan, PLUM is under evaluation for the ShakeAlert system (Cochran et al., 2019). None of the present early warning algorithms develop probabilities of eventual size or magnitude early in the rupture process, a condition addressed in the present work.

Here, we present a probabilistic approach for estimating the eventual length of a growing earthquake rupture given the starting location, knowledge of the fault structure, and the assumption that rupture starts on the fault structure. A realistic scenario can be offered. If the alert earthquake starts on the southeast San Andreas fault (SAF) and already near M 6, are the probabilities of rupture growth high enough to recommend that all of Los Angeles should be alerted? What about an alert on the southern end of other large faults in southern California? Are the differences in probability large enough to support fault-specific policies? The methods outlined here can be used to support such decisions. Probabilities conditioned on alert location can be computed based on knowledge of the fault system geometry in advance for all discrete elements in the fault model. Besides EEW, the method can be used to develop rupture length probabilities for other applications including seismic hazard assessment, scenario earthquake planning, and as a reference for computer-based modeling of complex fault systems. We also develop an alternative approach to integrate the Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3; Field et al., 2014, 2017), into EEW. Probabilistic estimates of rupture length cannot take the place of direct measurement of the rupture under way. Rupture length probabilities instead offer complementary information about areas likely to experience strong shaking and surface rupture. Compared to waiting for the algorithms to develop a final magnitude (Trugman et al., 2019), the fault-based complementary information of rupture length probabilities could improve accuracy in alert area and expected intensity and reduce time to alert.

**ESTIMATING PROBABLE LENGTH OF FUTURE EARTHQUAKES: DISCRETIZED FAULT MODEL**

On a long-term basis, a fault-based rupture forecast such as UCERF3 in California can be used to estimate the likelihood that a rupture of a given length will occur. However, once a rupture has started, the a priori probabilities of earthquake occurrence no longer apply. The length estimate becomes conditional on the starting location itself and on the properties of the faults connected to it.

To introduce our approach to estimating the probability of eventual rupture length conditioned on knowledge of initial location, we begin with a simplified discrete fault model (Fig. 1). The area nominally ruptured by the time an EEW point-source algorithm could alert for a rupture under way is assumed to be roughly one subsection. The fault consists of nine subsections, and we assume that rupture initiates in the middle, as rupture of panel \(S_0\). Given rupture initiation in \(S_0\) and the nine-element discrete model shown, there are 24 possible rupture extensions (Fig. 1). If all rupture extents are equally likely (i.e., \(p_1 = p_2 = p_3\) etc.), then by total

---

**Figure 1.** Illustration of single fault composed of nine panels (subsections) illustrating possible rupture extents for an earthquake initiating in central panel \(S_0\). The probability of any given rupture is \(p_i\). The color version of this figure is available only in the electronic edition.
probability one may simply count the ruptures with the extent of interest as a fraction of all possibilities. For example, ruptures 1–4 have unilateral rupture to the right (ur) of panel \( S_0 \), so \( P_{ur} = \sum_{i=1}^{4} p_i \). Unilateral rupture to the left (ul) of panel \( S_0 \) is \( P_{ul} = \sum_{i=9}^{12} p_i \). Other cases such as starting in \( S_0 \) and ending in panel \( S_0 \) (either bilateral or unilateral) follow by summing the probabilities of the individual ruptures. Thus, in this simplest model in which all ruptures are equally likely, given a rupture initiates in \( S_0 \), one may simply count the ruptures involving each of the other subsections (Fig. 2a) and translate to probabilities by dividing by the total number of ruptures (bar heights, Fig. 2b).

**MODIFYING THE DISCRETIZED MODEL—MAGNITUDE–FREQUENCY DISTRIBUTION**

A problem with the simple fault model of Figure 1 is that, observationally, larger magnitude and thus longer ruptures occur less frequently than shorter ones. One path forward for adjusting rupture length expectations is to apply a fault magnitude–frequency distribution (MFD). The exact form of the MFD appropriate to describe the recurrence of large (\( M > 6.5 \)) earthquakes on long faults remains a topic of discussion. Field et al. (2017) find that it cannot be assumed that individual faults in California all follow the Gutenberg–Richter (GR) MFD. Nevertheless, because of its familiarity and success on some faults, it provides a relevant reference. In a GR distribution, the number of earthquakes equal to or exceeding some magnitude \( M \) is given by \[ N(M) = a - b \times M. \]

Typically, and in California, the value of \( b \) is found to be near 1. We convert model lengths to magnitudes using \( M - L \) relationships of Anderson et al. (2017). An independent or regionally determined value of \( a \) is not required because of the condition that the event has initiated. \( N(M) = 1 \) for the smallest \( M \) (here, in the mid-M 5 range) and only the relative frequency of larger events is of interest. The effect of assuming the power-law frequency distribution is to progressively decrease probabilities with increasing rupture length (Fig. 2b).

Table 1 lists the predicted relative frequencies of larger events when the GR distribution applies. \( M \) is scaled from discrete fault elements that are 7 km in length. We assume for EEW application that the alert corresponds to the rupture of a single subsection. A 7 km subsection scales to \( M_{min} = 5.33 \) using Anderson et al. (2017). The GR ratio, \( b \times 10^a(M_{min} - M) \), is the predicted rate of larger events in which \( b = 1 \). For a given magnitude, \( N(M) \) includes all events of a given length that include \( S_0 \). For example, three ruptures including \( S_0 \) have length 21 km (\( S_2 - S_1 - S_0, S_1 - S_0 - S_{-1}, \) and \( S_0 - S_{-1} - S_{-2} \)).

### Table 1

<table>
<thead>
<tr>
<th>Length (km)</th>
<th>( M )</th>
<th>GR Ratio</th>
<th>( N(M) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>5.33</td>
<td>1.000</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>5.94</td>
<td>0.246</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>6.29</td>
<td>0.110</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>6.54</td>
<td>0.062</td>
<td>4</td>
</tr>
<tr>
<td>35</td>
<td>6.73</td>
<td>0.040</td>
<td>5</td>
</tr>
<tr>
<td>42</td>
<td>6.89</td>
<td>0.028</td>
<td>4</td>
</tr>
<tr>
<td>49</td>
<td>7.02</td>
<td>0.020</td>
<td>3</td>
</tr>
<tr>
<td>56</td>
<td>7.14</td>
<td>0.016</td>
<td>2</td>
</tr>
<tr>
<td>63</td>
<td>7.24</td>
<td>0.012</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Figure 2.** (a) Histogram showing the number of ruptures each subsection could participate in. (b) Probability of a subsection being involved in rupture given rupture initiates in \( S_0 \), and each possible rupture is considered equally likely. Dashed line and stars illustrate reduction in probabilities if a power-law distribution exists among likely rupture lengths on the model fault. The color version of this figure is available only in the electronic edition.
frequency of any one of the three (absent other information) is the GR ratio/$N(M) = 0.110/3$. Table 1 extends this calculation to all possible rupture lengths in Figure 1. Table 1 immediately provides a useful reference. For example, only 24.6% of ruptures are predicted to grow to occupy a second subsection, and only 2% that start as an $M$ 5.3 single subsection rupture would go on to become an $M$ 7.0 event.

MODIFYING THE DISCRETIZED MODEL—FAULT GEOMETRY

Faults and bends

In the simple fault model of Figure 1, rupture can proceed from one panel to the next without penalty. Empirical observations and physics-based models of rupture indicate that geometrical complexities such as steps and bends affect the probability that rupture will stop (e.g., Sibson, 1985; Wesnousky, 1988; Harris et al., 1991; Lettis et al., 2002; Lozos et al., 2011, 2015; Biasi and Wesnousky, 2016). To illustrate the effect, we modify the simple fault model of Figure 1 to include steps and bends in the fault trace (Fig. 3a). The condition that the earthquake has started is expressed as probability $P_0 = 1$ at panel $S_0$. To grow in length beyond panel $S_0$, geometrical complexities at panel boundaries represent serial “challenges” to propagation. We qualitatively illustrate the reductions in probability arising from each step and/or bend challenge with the dashed lines in Figure 3b. Each incremental probability $P_i$ describes the event “start at $S_0$ and stop at the end of subsection $i$.” Probabilities on the left side are lower than on the right because three subsection connections on the right have no bend or step to reduce the probability of continuing.

To quantify the effects of steps and bends, we draw on the results of Biasi and Wesnousky (2016, 2017). Biasi and Wesnousky (2016) measured steps in mapped historic surface ruptures. Where faults were mapped beyond the ends of surface rupture, step widths at the ends of ruptures were also measured. For a given step width, Biasi and Wesnousky (2016) defined the ratio of the number of ruptures that passed to the number of that size that stopped rupture at an end as the passing ratio (Fig. 4a). An approximately linear dependence of this ratio on step width is observed for steps from 1 to 6 km. Ruptures are observed to stop or pass through steps of 3 km with approximately equal frequency. A similar passing ratio relationship was observed for bends in surface ruptures, for which the size of the angle in the surface trace is observed (Fig. 4c). For bends, observations show that bends in a fault trace $<15^\circ$ are passed over twice as often as they stop rupture whereas bends of $31^\circ$ are twice as likely to stop rupture as to be passed. Details about the original measurements of bends and steps are documented in the source publications.

Passing ratios for steps and bends in Figure 4a,c are converted to probabilities of passing in Figure 4b,d, respectively. $P_{ab}$ and $P_{as}$ are the probabilities that a bend or step, respectively, will arrest rupture. The complementary probabilities, $P_{pb} = 1 - P_{ab}$ and $P_{ps} = 1 - P_{as}$, respectively, are interpreted as the probability that a rupture will pass beyond the bend or step. For steps smaller than 1 km, a linear extrapolation is applied in Figure 4b. The discontinuity in slope at a width of 1 km is considered to be an artifact of insufficient data (Biasi and Wesnousky, 2016) that might be resolved with further study. If there is no step, no step penalty is applied. For the probability of stopping at steps shown in Figure 4d, a smoother extrapolation has been used because the range of estimates in passing ratio for angles smaller than $10^\circ$ is less well defined. If there is no bend, there is no basis for a bend penalty. Passing probability values for bends and steps shown in Figure 4b,d are listed in Table 2. A linear extrapolation is applied for step widths and bend angles that fall between widths or angles in the table.

The probability curves of Figure 4b,d provide the means to quantify consequences of bends and steps of a discrete fault model such as is shown in Figure 3. The probability of a rupture longer by one subsection is smaller by the “penalty” from the step or bend, applied as a product. The cumulative effect of these penalties for bends and steps means that complex ruptures should be rare compared to similar length ruptures on geometrically simpler faults.

![Figure 3](image-url)
Expanding to consider the UCERF3 fault model

The model in Figure 4 can be extended to the active fault system of California using the fault model in UCERF3 (Fig. 5). The discrete fault elements are called “subsections.” They extend in depth to the base of the local seismogenic zone, and half that (i.e., 5–7 km) in strike length. Fault subsections in UCERF3 can have multiple subplanes, but to be consistent in scale size with the measurements in Biasi and Wesnousky (2016, 2017), orientations are represented by an average single dip and dip direction. We estimate the dip direction using the strike defined by end points of the subsection. In UCERF3, ruptures consist of a sequence of two or more subsections. Ruptures are limited to single paths with no discontinuities greater than the maximum step size of 5 km, and no bifurcations (“Y”-shaped ruptures). The degree of interconnectedness of California faults (Fig. 5) led to the need for further screening on the basis of geometric rules including net direction change, total strike change, and maximum individual strike change at subsection intersections. See Milner et al. (2013) for the full set of rules for minimum geometric compatibility for a rupture to be considered viable for inclusion in the UCERF3 model. The Milner et al. (2013) screening makes no judgment about the relative probabilities among possible ruptures. Ruptures either pass and become part of the inversion or fail and are excluded. Rupture occurrence rates were estimated using a Monte-Carlo-based inversion (Field et al., 2014). The relative geometric complexity among ruptures was not applied as a constraint or initial value in the UCERF3 inversion.

The UCERF3 fault model has the geometric information needed to calculate step offsets and fault bends. As a model, we assume the earthquake triggering the early warning alert occupies one subsection of the fault model. The effects of bends and steps on rupture extension can be calculated using the probabilities in Figure 4b,d. Step distances between subsections are calculated from the separation of fault panels based on the

Figure 4. Passing ratios versus (a) step width and (c) bend angle, adapted from Biasi and Wesnousky (2016) and Biasi and Wesnousky (2017), respectively. Bend and step complexities are measured between fault sections of at least 5–7 km in length. (b) Probability of passing or stopping at a step versus step width ($P_{pass}$ and $P_{stopping}$, respectively). (d) Probability of passing or stopping at a bend of given angle in fault trace ($P_{pass}$ and $P_{stopping}$, respectively). The color version of this figure is available only in the electronic edition.
latitudes and longitudes of the ends of the subsections. The angle between fault subsections is computed in 3D using the average dip and computed dip direction parameters of the subsections. The conditional probability $P_k(L)$ of rupture length $L$ under step and bend effects, given that the alerting earthquake starts by rupturing subsection $k$, is product is over pairs of subsections that compose length $L$:

$$P_k(L) = \prod P_{sh,i},$$  

in which the $P_{sh,i}$ is the step or bend probability connecting adjacent subsections in the rupture (Table 2). Equation (1) applies to unilateral rupture from the initial subsection. For any specific bilateral rupture, equation (1) is applied once in each direction to cover the full rupture extent, and the probabilities associated with the two directions are multiplied. With application of equation (1) to successively longer ruptures, the accumulation of step and bend penalties produces a monotonically declining probability of rupture length.

We illustrate the application of step and bend passing probabilities to estimate rupture length probabilities with two examples from southern California (Figs. 6, 7, and 8). The first example assumes the earthquake starts at the southeastern end of the SAF at Bombay Beach (star), and rupture extends unilaterally northwest (Fig. 6). We assume the rupture roughly occupies a single subsection, corresponding to a mid-M 5 event, at the time of an alert. In Figure 6, subsection intersections for the SAF and SJF are shown as dots. From the alert location, the individual bend and step penalties for rupture are computed separately using the geometries of each subsection intersection. The individual bend and step passing probabilities are shown in Figure 7a (circles and + symbols, respectively), and the solid line shows their joint application. Cumulative applications of each using equation (1) are shown in Figure 7b. We take probabilities of length from the cumulative joint probability curve. A GR probability of length is also shown for reference. The SAF northwest from Bombay Beach is relatively straight and smooth. Probabilities of length weakly decrease in subsections 2–12 because there is little basis in fault geometry to reduce them. The first significant bend and step complexities are encountered 13 subsections northwest where rupture is assumed to transition to the Mill Creek SAF fault section. Other SAF section transitions are indicated in Figure 6. The decline in propagation probabilities north of the Coachella section is consistent with the progressive counter-clockwise rotation of earthquake starts by rupturing subsection $k$, is product is over pairs of subsections that compose length $L$:

$$P_k(L) = \prod P_{sh,i},$$  

in which the $P_{sh,i}$ is the step or bend probability connecting adjacent subsections in the rupture (Table 2). Equation (1) applies to unilateral rupture from the initial subsection. For any specific bilateral rupture, equation (1) is applied once in each direction to cover the full rupture extent, and the probabilities associated with the two directions are multiplied. With application of equation (1) to successively longer ruptures, the accumulation of step and bend penalties produces a monotonically declining probability of rupture length.

We illustrate the application of step and bend passing probabilities to estimate rupture length probabilities with two examples from southern California (Figs. 6, 7, and 8). The first example assumes the earthquake starts at the southeastern end of the SAF at Bombay Beach (star), and rupture extends unilaterally northwest (Fig. 6). We assume the rupture roughly occupies a single subsection, corresponding to a mid-M 5 event, at the time of an alert. In Figure 6, subsection intersections for the SAF and SJF are shown as dots. From the alert location, the individual bend and step penalties for rupture are computed separately using the geometries of each subsection intersection. The individual bend and step passing probabilities are shown in Figure 7a (circles and + symbols, respectively), and the solid line shows their joint application. Cumulative applications of each using equation (1) are shown in Figure 7b. We take probabilities of length from the cumulative joint probability curve. A GR probability of length is also shown for reference. The SAF northwest from Bombay Beach is relatively straight and smooth. Probabilities of length weakly decrease in subsections 2–12 because there is little basis in fault geometry to reduce them. The first significant bend and step complexities are encountered 13 subsections northwest where rupture is assumed to transition to the Mill Creek SAF fault section. Other SAF section transitions are indicated in Figure 6. The decline in propagation probabilities north of the Coachella section is consistent with the progressive counter-clockwise rotation of

### Table 2

<table>
<thead>
<tr>
<th>Step Width (km)</th>
<th>Step Penalty</th>
<th>Bend Angle (°)</th>
<th>Bend Penalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.999</td>
</tr>
<tr>
<td>0.18</td>
<td>0.97</td>
<td>2</td>
<td>0.909</td>
</tr>
<tr>
<td>0.51</td>
<td>0.868</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>0.76</td>
<td>0.72</td>
<td>10</td>
<td>0.683</td>
</tr>
<tr>
<td>1</td>
<td>0.645</td>
<td>15</td>
<td>0.615</td>
</tr>
<tr>
<td>1.5</td>
<td>0.588</td>
<td>20</td>
<td>0.545</td>
</tr>
<tr>
<td>2</td>
<td>0.559</td>
<td>25</td>
<td>0.474</td>
</tr>
<tr>
<td>2.5</td>
<td>0.527</td>
<td>30</td>
<td>0.375</td>
</tr>
<tr>
<td>3</td>
<td>0.49</td>
<td>35</td>
<td>0.231</td>
</tr>
<tr>
<td>3.5</td>
<td>0.446</td>
<td>40</td>
<td>0.153</td>
</tr>
<tr>
<td>4</td>
<td>0.394</td>
<td>45</td>
<td>0.083</td>
</tr>
<tr>
<td>4.5</td>
<td>0.331</td>
<td>50</td>
<td>0.043</td>
</tr>
<tr>
<td>5</td>
<td>0.254</td>
<td>55</td>
<td>0.034</td>
</tr>
<tr>
<td>5.5</td>
<td>0.156</td>
<td>60</td>
<td>0.029</td>
</tr>
<tr>
<td>6</td>
<td>0.029</td>
<td>65</td>
<td>0.024</td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>70</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>85</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Figure 5.* Discrete fault model FM3.1 from Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3). All faults except as marked with a dashed line are connected within the fault model by steps no larger than 5 km. Figure from Field et al. (2014). The color version of this figure is available only in the electronic edition.
fault strike on the Mill Creek to a less favorable orientation for through rupture. Only 5% of ruptures starting on the Coachella section are predicted to get past the Mill Creek section to reach eastern San Bernardino, only 2.5% continue to the near-southeast end of the Mojave South section (Fig. 7b), and only 0.2% would rupture “wall-to-wall” from Bombay Beach to Parkfield. Based on fault geometry, ruptures that start in the southeast end of the SAF should rarely reach to the eastern edge of metropolitan Los Angeles at San Bernardino.

In the second example, the rupture starts on the San Jacinto fault (SJF) at the Casa Loma stepover (Figs. 6 and 8) as a single subsection rupture scaling to a mid-M 5. In this case, rupture might extend northwest or southeast. Because probabilities in equation (1) are conditioned on the alert location, probabilities of the northwest and southeast extents are independent, and thus can be considered separately. In the UCERF3 fault model, the SJF can connect northwest to the San Bernardino North SAF two ways, over three subsections of the Lytle Creek fault (Fig. 8a, b) or continue three subsections farther on the SJF (Fig. 8c). Based on fault geometry, the direct connection is a more likely path for through ruptures, though neither is very likely to actually continue on the Mojave South section (5.8% vs. 2.7%). For rupture lengths up to 300 km, the GR relation predicts significantly lower probabilities than either prediction from fault geometric complexity. Lozos et al. (2015) and Lozos (2016) have studied rupture propagation through this intersection and found that it is sensitive to poorly resolved details of the fault system geometry. For rupture extending to the southeast on the SJF, decrements in probability correspond to recognized section boundaries (Fig. 8d). Anza and Coyote Creek sections are relatively straight, with little geometric basis for rupture arrest, whereas curvature of the Borrego fault (Fig. 6) causes a progressive decrease in probability of through rupture. The probability of any given bilateral rupture extent given a starting alert near the Casa Loma stepover would be the product of probability of the corresponding northwest and southeast extents.

In Figures 7 and 8, we so far have discussed conditional probabilities of length on a single rupture path. This may be sufficient for some purposes. However, if conditional probability of length or magnitude given an alert at mid-M 5 is required regardless of path, an accounting must be made of probabilities at branch points. As long as the paths are independent alternatives, probabilities of a given rupture length or magnitude can be combined by weighting by their relative geometric probabilities at the branch point. Using the example in Figure 8 of connection from the SJF to the SAF directly versus by Lytle Creek, the last common point is on the San Jacinto San Bernardino strand (SJSB, Fig. 8b, c). Staying on the SJF involves a bend probability of 0.76, and no step penalty. Jumping to the Lytle Creek fault involves a slightly larger bend penalty of 0.64 and a small step with penalty 0.91. Combining gives probabilities of 0.76 versus 0.58, respectively. Thus, based on fault geometric parameters, the direct connection is preferred. Probabilities of length on the direct connection path would be weighted by 0.76/(0.76 + 0.58) = 57% versus 43% for connection by Lytle Creek. Weighting of this sort applies to length or magnitude accumulated on distinct branches. In this case, the alternate paths meet on the Mojave South section. Northwest of that intersection, the probabilities of length in Figure 8b, c can be summed. Alternative weighting approaches are discussed in a later section.
UCERF3 RUPTURE LENGTH PREDICTIONS

If rupture probabilities are available for all possible ruptures and paths, these probabilities can provide a third basis for the conditional probability of rupture length given EEW initiation as a fault model subsection. For California, time-independent rupture probabilities are available from UCERF3 (Field et al., 2014). The UCERF3 rupture set was constructed to give rates for all possible ruptures in the discretized fault model. A subset with an end at a given subsection can then be selected and be a total probability for ruptures with that geometry. That is, once we know a given subsection has ruptured and comprises one end of something that could grow, then according to the UCERF3 model, with probability 1, the final rupture will be one from the set. To proceed to the EEW problem, we must assume that the relative a priori probabilities among ruptures in the subset apply to the highly time-dependent condition of a rupture in progress.

We illustrate the subsetting process for the same SJF starting point considered previously. We extract all ruptures in the UCERF3 fault model 3.1 that have one end at the Casa Loma step and that extend to the northwest. Figure 9a (star symbols) shows individual annual rates of occurrence. There are 769 ruptures with one end at the Casa Loma step. The smallest rupture among them consists of two subsections, the Casa Loma step plus one to the northwest. It has a UCERF3-
might have started. The fault geometric CCD is replotted from Figure 8c. The dashed line shows the length CCD for the fraction of ruptures starting at the step (west star, Fig. 6). (b) The solid line is corresponding complementary cumulative distribution function (CCDF) for rupture length for the rupture subset. The median length of a rupture with one end at the Casa Loma stepover is about 380 km (solid line). For the EEW case, the rupture starts at a known point and grows from there. Thus, the nucleation point will be one specific subsection from a rupture. Probabilities as though it started anywhere must be adjusted downward for this specific starting condition. To make this adjustment, we assume the earthquake might nucleate in any subsection of a rupture with equal likelihood. We thus reduce the annual probability of occurrence for each rupture in Figure 9a by 1/n, in which n is the number of subsections in the rupture. The dashed line of Figure 9b incorporates this reduction and so represents the UCERF3-based CCDF of rupture length (L) for unilateral rupture northwest from the Casa Loma stepover. A corresponding CCDF for earthquake magnitude is shown in Figure 9c. Normalizing by the number of subsections in ruptures reduces probabilities of long ruptures more severely than shorter ones such that the reduced CCDF(L) is not controlled to the same degree by M > 7.5 events (Fig. 9a).

For comparison of the UCERF3 and geometric models, we replot the fault-geometric estimated CCDF for length from Figure 8c and corresponding magnitudes on Figure 9b,c, respectively. The fault-geometric and UCERF3 estimates are equal at a rupture length of 62 km and 25% remaining probability. Thus, 75% of ruptures starting at the Casa Loma stepover have lengths of 62 km or less. Length estimates of the fault-geometric model in this range are longer at a given probability than predicted from UCERF3. This length relationship switches for longer ruptures in the remaining 25% of cases. For example, under UCERF3, 18% exceed 80 km, whereas only 6% are predicted by the fault-geometric estimate to exceed that length. In terms of the faults themselves, the UCERF3 model does not penalize ruptures connecting the San Jacinto to the SAF, despite severe mechanical problems at the intersection (Fig. 8b,c, probability drop at subsection 11). This greater freedom allows UCERF3 ruptures to extend through the intersection at a rate roughly three times higher than under the fault-geometric model.

The UCERF3 rupture model also can be used to track probability of length or magnitude through bifurcations in the fault. In Figure 9, we considered probability of length without specifying exactly which fault(s) the rupture might occupy. Thus, in the set shown, some ruptures join the SAF from the SJF both

Figure 9. UCERF3-based rupture length probabilities for rupture with one end at the San Jacinto Casa Loma step. (a) Individual annual rupture rates (probabilities) (stars) and incremental magnitude–frequency distribution (MFD, solid line, binned at 0.1 magnitude units) of all ruptures in UCERF3 fault model 3.1 having one end at the Casa Loma step of the SJF (west star, Fig. 6). (b) The solid line is corresponding complementary cumulative distribution (CCDF) (solid line) of rupture length for ruptures in (a). Dashed line shows the length CCD for the fraction of ruptures starting at the step. Compared to the solid line, individual rupture probabilities are reduced by the number of alternate initiation points (subsections) in which rupture might have started. The fault geometric CCD is replotted from Figure 8c. The fault geometric approach predicts longer ruptures than UCERF3 for 75% of ruptures starting at the step. The CCD curves cross at 62 km and 25% remaining probability. 18% of UCERF3 ruptures grow to 80 km or longer, compared to 6% for the fault geometric estimate. (c) CCD for rupture magnitude for the reduced CCD and fault geometric curves in (b). A GR conditional probability of magnitude is shown for reference (dotted). The fault geometric approach predicts larger median magnitudes for ruptures starting at the step than either UCERF3 or the GR. The color version of this figure is available only in the electronic edition.
directly to the SAF north San Bernardino section, and alternately on the Lytle Creek section. Where it is desirable to track such distinctions, the process with Figure 9 is repeated, but with the rupture set separated by fault branch. Probabilities for each branch at the “Y” are estimated according to the total UCERF3 probability of ruptures that continue. Similarly, bilateral length probabilities conditioned on the initiation point are formed by gathering the southeast and northwest sets separately in the example of Figure 8, then multiplying the probabilities of length on either side. The eventual magnitude probabilities, however, must be scaled from the combined lengths using a relationship such as in Anderson et al. (2017), or by finding the corresponding magnitudes from the UCERF3 model for each contributing rupture.

DISCUSSION

Fault-geometric passing probabilities provide an empirical basis for estimating potential rupture lengths given rupture initiation as a mid-M5 rupture on the fault system. Probabilities of length depend only on the geometry and connectivity of the fault model and on empirical step and bend observations organized as passing probabilities. And although we have motivated the research by its application to EEW, conditional length estimates are equally applicable in other contexts in which a fault model is available, and probabilities of rupture extent are needed for hazard scenarios and response planning.

The role of fault discretization on probabilities deserves some discussion. We used the fault model from UCERF3 because it is well vetted and has an internally consistent rupture rate forecast. The fault models in UCERF3 were constructed using the best available science at the time of its compilation (Dawson, 2013). Attention was paid to 3D fault geometry where known, and to the potential for fault connectivity. UCERF3 faults were discretized into subsections with depth to the base of the seismogenic zone, and half that in length. Biasi and Wesnousky (2016, 2017) used a similar scale length to analyze historical surface ruptures. Matching scale lengths allowed more meaningful comparisons between the two studies. Inspection of surface rupture maps shows that features smaller than a few kilometers in length are often superficial and not reflective of fault slip at scales governing moment release. Kinematic models also find that coherent motion of a few kilometers scale is required to develop momentum and trigger continuing rupture. Thus, the current discretizations of the fault model, the bend and step passing probabilities, and dynamic rupture are roughly consistent in their data basis and physical meaningfulness. Discretization at a finer scale could be entertained if a reliable fault description could be developed, but the application of kinematic and dynamic modeling to the finer scale model would need to be demonstrated. Bend and step probabilities would also have to be revisited at the finer scale.

We find for representative examples on southern California’s most active faults that realistic conditional probabilities of rupture length can be formed directly from probabilities at geometric complexities. We find relatively low probabilities for a rupture starting as a subsection rupture of mid-M5 then extending from the southernmost SAF into San Bernardino or beyond (Fig. 7). These low probabilities are consistent with geologic and dynamic modeling assessments that such a rupture should be rare. For rupture northwest from the northern SJF (Fig. 8), we find about a factor of 2 lower conditional probabilities from fault geometry than from UCERF3 that rupture should extend from the northern SJF onto the SAF.

For straight faults, there is no basis in fault geometry to decrease probabilities. At face value, this just means that conditional probabilities of rupture are equal among the range of lengths on the straight section. This has the effect of elevating relative probabilities of longer ruptures than would be predicted by a power law such as the GR. The San Jacinto step example presented earlier (Fig. 9b) is not inconsistent with kinematic models such as by Lozos et al. (2011), which indicate that under physically realistic conditions, ruptures can propagate indefinitely on straight faults.

For long ruptures that span fault intersections, conditional probability estimates of rupture length or eventual rupture magnitude will require either picking a single fault rupture path or combining probabilities across fault branches. We illustrated an approach using relative weights based on geometric favorability at the intersection providing alternate paths northwest from the SJF (Fig. 8). If there were further branches, this procedure could be applied recursively. One might alternatively weight branch probabilities on the basis of relative slip rates of the branches. Using the UCERF3 fault model, slip rates on the SJF and Lytle Creek where they split are 9.0 mm/yr and 1.8 mm/yr, respectively. On this basis, a weighting is found of 83% versus 17%, respectively, compared to 57% versus 43% found from geometry alone. A related division might be calculated by summing rupture rates on each branch from the UCERF3 time-independent model.

For specific branch points, paleoseismic data might also provide a basis to adjust respective weightings of branches. Schwartz et al. (2012) show that the eastern extent of the Denali fault had a more recent large surface rupture earthquake on it than the Totchunda fault near their intersection. When the Denali earthquake rupture propagated east, it took the less geometrically favored branch, they infer, because of the more recent previous Denali event. Although potentially useful at individual branches, the application of paleoseismic data in this way would be situation-specific. For California, a generalization of this type of data is available through the time-dependent version of UCERF3 (Field et al., 2015). Its use in estimating conditional probabilities of rupture length is reserved for future research.

Besides conditional probability of length or magnitude, other questions might be asked, such as the probabilities of...
magnitude for ruptures that could reach a certain point, such as an urban area. For a conditional probability question such as this, one must consider all combinations in either direction that affect the city. This would require a certain level of bookkeeping, as illustrated with Figure 1, but not comprise an entirely new approach.

For EEW applications, probabilities of length and/or magnitude after starting as a subsection-sized rupture in the fault model could readily be precompiled. If precompiled, then during an EEW alert, length probabilities can be accessed very quickly by means of a lookup table. Such a lookup will not take the place of dynamic estimates of magnitude such as are provided by the FinDer algorithm (Böse et al., 2012, 2015). To use conditional length probabilities to update the alert area, rules would be required relating continuous probabilities to discrete alert decisions. For example, the alert area might be increased if the rupture has a 50% chance of outgrowing the current bound.

We motivated this research by considering probabilities of rupture length from an EEW initial alert. The value added by our approach is the capability to have, in a fraction of a second, a probability distribution for length of eventual rupture. There is nothing like this in current ShakeAlert software, so no comparative value can be offered. During an EEW rupture, however, the main need is to quickly extend the alert area for a growing rupture. Policies for this must be worked out in advance, including the probabilities at which to use the rupture length estimates. A first term in probability discussions is whether the EEW alert is on a fault in the fault model. One could also ask whether the fault model can be trusted. If the alert earthquake is accepted to be on the fault system, and an alert earthquake reaches M 6, say, on the southeast SAF, are the growth probabilities high enough to recommend that all of Los Angeles should be alerted? What about an alert on the southern SJF? Are the differences in probability large enough to have fault-specific policies? The methods outlined here can provide insight to those decisions.

Beyond application to rupture length estimates, fault-geometric passing probabilities provide complementary model evaluation metrics for a future UCERF model. UCERF3 ruptures start with no a priori probability per se. If a rupture passes basic geometric compatibility tests (Milner et al., 2013), nothing downstream in the rupture rate inversion distinguishes simple versus geometrically complex ruptures. Mathematical relationships implementing fault geometric passing probabilities might be formulated, for example, to constrain the ratio of through ruptures to ruptures that stop at a geometric feature. Alternatively, fault-geometric probabilities could be used as a complementary tool to identify ruptures that pass the Milner et al. (2013) screening but include multiple, unfavorable geometric intersections and thus could be culled from the rupture set. An exploratory study by Biasi (2016) found that a strongly reduced rupture set could fit UCERF3 data constraints at least as well as the full rupture set and produce very similar hazard estimates. The smaller input rupture set also improved computational performance of the rupture rate inversion. Finally, instead of using fault geometric probabilities as inputs to the inversion, they could be used to compare with inversion results. The UCERF3 model has been difficult for geologists to evaluate (e.g., Schwartz, 2018) because virtually all available geologic data are used as inputs to the inversion. Once the data are fit by the inversion, little independent data remain to evaluate the resulting model. Geometrically based passing probabilities cannot directly replace a rupture rate inversion, but they do make specific, physically grounded predictions of the relative rates of long and short ruptures, and these data are not inputs to the UCERF3 inversion. Summarizing, step and bend complexities model geometry well, without reference to slip rate, and UCERF3 fits slip rate without reference to geometric complexity.

**CONCLUSION**

A fault-geometric approach is presented to estimate the conditional probabilities of rupture length and/or magnitude, based on initiation as an ~7 km, mid-M 5 earthquake. We derive empirical probabilities of passing through fault bend and step structures that depend on the angle of the bend and size of the step, respectively. When translated to probabilities of rupture arrest, fault geometric complexities comprise challenges that a rupture encounters serially if it is to increase in length. The probability of length is thus the product of the complementary probabilities of continuing through each feature. Long and complex ruptures have lower probabilities, conforming to empirical observation. Considering the specific example of rupture starting at the southeast end of the SAF, the resulting probability of length estimates is severely reduced by the strong change in fault strike onto the Mill Creek and eastern north San Bernardino fault sections and only a few percent would be expected to get through both. Though small, this fraction is still larger than the 0.4% predicted for a GR MFD we consider as a reference.

We also show that conditional length probabilities can be estimated directly based on rupture rates from UCERF3. To do this, we must assume that time-independent rupture rates from the forecast apply to the conditional probability of rupture length given the rupture starting location. Results from a comparison case for rupture initiating on the northern SJF are representative. For straight faults, individual fault-geometric passing probabilities are high and the conditional length probabilities are larger than predicted from either UCERF3 or the GR model. Probabilities switch to become higher for the UCERF3 model for the fraction of ruptures that extend through the complex transition onto the SAF. Fault-geometric probabilities might also be considered in a future UCERF model, either as a data constraint, as a complement to model construction, or as a tool to evaluate inversion results.

In an EEW context, the methods developed here provide a basis to estimate where a rupture may go given initiation as a
mid-M5 subsection-scale rupture, and with what probabilities. These probabilities are readily compiled in advance for any given starting subsection in the fault model, in effect covering likely nucleation locations for large earthquakes anywhere in the California fault network. These probabilities could be used to advise policy about alerting extent for different faults. Operationally, precompiled probabilities, suitably adjusted for confidence that the event is on the fault system, could quickly be accessed by the EEW system when an earthquake has initiated. As an example result, we find that an earthquake that initiates as a mid-M5 event at Bombay Beach on the southeast end of the SAF only reaches San Bernardino and the eastern edge of urban Los Angeles about 5% of the time.

**DATA AND RESOURCES**

All data used in this article are from published sources in cited references.

**ACKNOWLEDGMENTS**

This research was supported by the Southern California Earthquake Center (SCEC) Award Numbers 17064 and 12012 (SCEC Contribution Number 10941). SCEC is funded by the National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and U.S. Geological Survey (USGS) Cooperative Agreement Number G17AC00047. Center for Neotectonics Studies Contribution Number 83. Detailed USGS internal reviews by J. McGuire and E. Cochran and two anonymous external reviews led to improvements in clarity and presentation.

**REFERENCES**


Manuscript received 23 November 2020
Published online 18 May 2021