A More Realistic Earthquake Probability Model Using Long-Term Fault Memory

James S. Neely[*1,2,3], Leah Salditch[1,2,4], Bruce D. Spencer[2,5], and Seth Stein[1,2]

ABSTRACT

Forecasts of the probability of a large earthquake occurring on a fault during a specific time interval assume that a probability distribution describes the interevent times between large earthquakes. However, current models have features that we consider unrealistic. In these models, earthquake probabilities remain constant or even decrease after the expected mean recurrence interval, implying that additional accumulated strain does not make an earthquake more likely. Moreover, these models assume that large earthquakes release all accumulated strain, despite evidence for partial strain release in earthquake histories showing clusters and gaps. As an alternative, we derive the necessary equations to calculate earthquake probabilities using the long-term fault memory (LTFM) model. By accounting for partial strain release, LTFM incorporates the specific timing of past earthquakes, which commonly used probability models cannot do, so it can forecast gaps and clusters. We apply LTFM to the southern San Andreas fault as an example and show how LTFM can produce better forecasts when clusters and gaps are present. LTFM better forecasts the exceptionally short interevent time before the 1857 Fort Tejon earthquake. Although LTFM is more complex than existing models, it is more powerful because (unlike current models) it incorporates fundamental aspects of the strain accumulation and release processes causing earthquakes.

KEY POINTS

- Earthquake probability models do not fully reflect the strain processes that drive earthquakes.
- We derive a new earthquake probability model that mimics the strain process.
- Because our model replicates the observed strain processes, we expect more realistic forecasts.

Supplemental Material

INTRODUCTION

For decades, seismologists have tried to predict when, where, and how large the next earthquake on a major fault would be. However, to date, these attempts have been unsuccessful (Hough, 2016). Instead, seismologists develop forecasts of the probability of an earthquake occurring in a region over a given timeframe (e.g., Field *et al.*, 2015; Schorlemmer *et al.*, 2018). These estimates are incorporated in mitigation policies, notably via hazard maps that predict the shaking levels that structures should withstand and raise public awareness of seismic hazards.

The forecasts are based on the concept of the earthquake cycle, in which the strain that accumulates between large earthquakes caused by motion between the two sides of a locked fault is released by slip on the fault when an earthquake occurs (Reid, 1910). However, the current forecast methodology does not include fundamental aspects of the strain accumulation and release process. Here, we present a new method of deriving probability estimates using the recently introduced long-term fault memory (LTFM) model, which is designed to reflect the strain accumulation and release processes and allows temporal clusters of earthquakes and gaps between them (Salditch *et al.*, 2020). We apply this method to the Mojave section of the San Andreas fault in California—a region with a well-documented earthquake history at Pallett Creek (Weldon *et al.*, 2005; Scharer *et al.*, 2010)—and show how it can produce more accurate forecasts.

LIMITATIONS OF CURRENT EARTHQUAKE PROBABILITY METHODS

The Mojave section of the San Andreas is of concern because large earthquakes occur on average every 135 yr, most recently

1. Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois, U.S.A., https://orcid.org/0000-0003-4029-2979 (JSN); https://orcid.org/0000-0002-4478-1836 (LS); 2. Institute for Policy Research, Northwestern University, Evanston, Illinois, U.S.A., https://orcid.org/0000-0001-6155-7249 (BDS); 3. Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois, U.S.A.; 4. Now at Geologic Hazards Science Center, U.S. Geological Survey, Golden, Colorado, U.S.A.; 5. Department of Statistics and Data Science, Northwestern University, Evanston, Illinois, U.S.A.

*Corresponding author: jneely@uchicago.edu

Cite this article as Neely, J. S., L. Salditch, B. D. Spencer, and S. Stein (2022). A More Realistic Earthquake Probability Model Using Long-Term Fault Memory, *Bull. Seismol. Soc. Am.* **XX**, 1–13, doi: 10.1785/0120220083

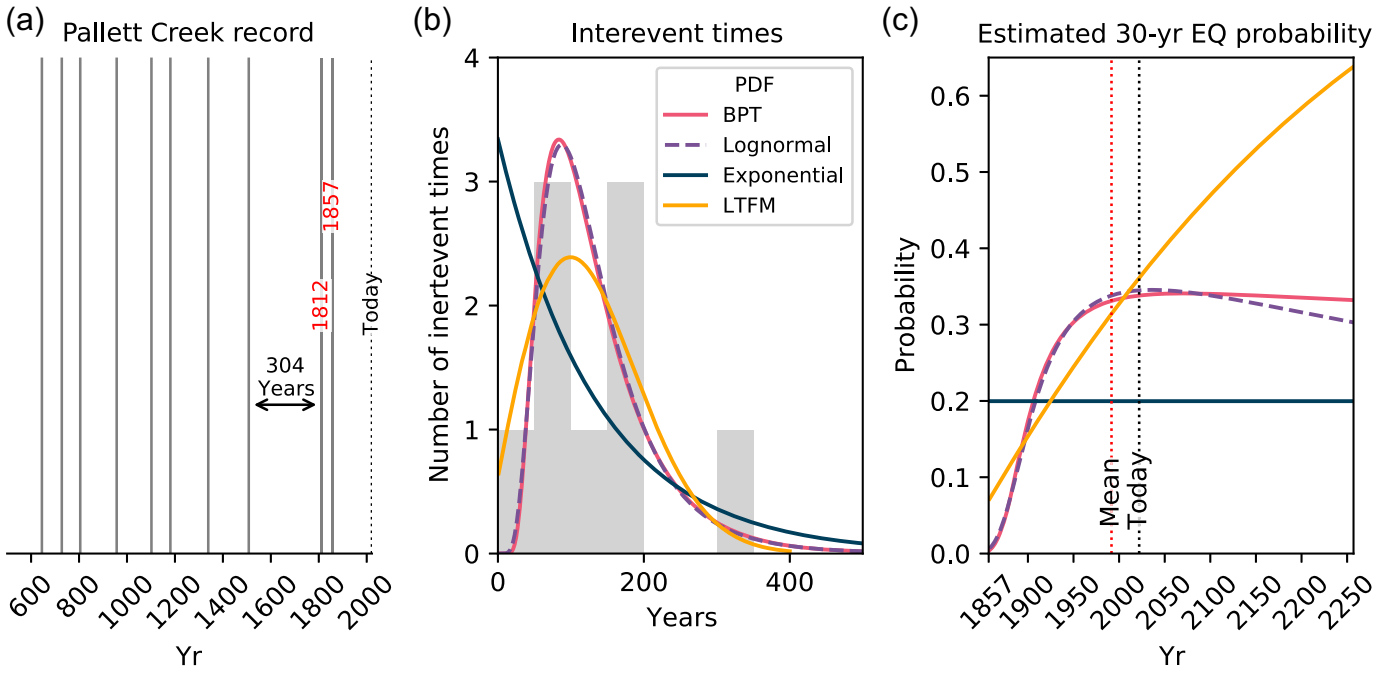


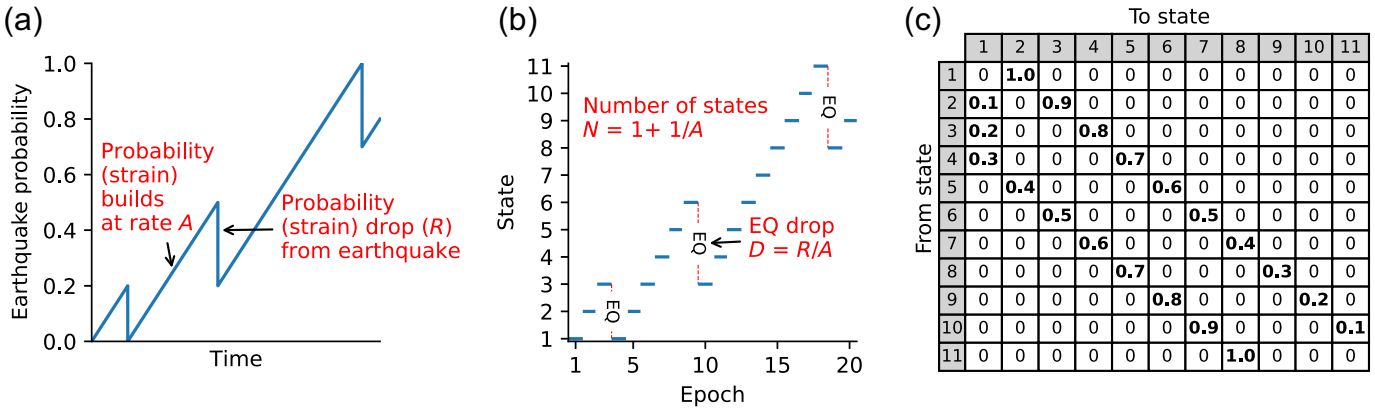
Figure 1. Estimating earthquake probabilities. (a) Pallett Creek paleoseismic record of past earthquakes (Scharer *et al.*, 2011, listed in Table S1). (b) Histogram of interevent times (gray bars) and estimated probability distributions for interevent times for exponential, lognormal, and Brownian passage time (BPT) models. The lognormal and BPT models have a coefficient of variation (CV) of 0.54. Also shown is the long-term fault memory (LTFM) presented in this article. (c) Conditional probability of an earthquake (EQ) in the next 30 yr given the 1857 date of the last earthquake. The color version of this figure is available only in the electronic edition.

the 1857 **M** 7.9 Fort Tejon earthquake. Prior to 1857, an **M** 7.5 earthquake occurred in 1812, which was preceded by a long 304 yr quiescent period (Scharer *et al.*, 2011). Studies suggest that the probability of a large earthquake here in next 30 yr is 20%–25% (Biasi *et al.*, 2002; Field *et al.*, 2015). How such probabilities are calculated varies across studies but follows a general methodology. Studies start with a paleoseismic record giving dates of large past earthquakes (Fig. 1a) and fit some probability density function (PDF) to the distribution of interevent times (Fig. 1b). Here we show three models, the time independent (which assumes the probability is constant with time) exponential—commonly referred to as the Poisson model in seismological literature—and the time-dependent (which assumes the probability changes with time) lognormal and Brownian passage time (BPT) models. From these PDFs, one can calculate the probability of a large earthquake during a time period, typically the next 30 yr, given that one has not occurred since 1857 (Fig. 1c).

For the time-independent exponential model, the 30 yr probability stays constant at 20% (Fig. 1c). The 30 yr probabilities change with time for the time-dependent lognormal and BPT models. In 2022, the lognormal and BPT models produce similar 30 yr estimates (34%) but diverge in the future. The lognormal and BPT probabilities have been relatively flat since the expected mean recurrence time (~1990s), but the lognormal estimated probability decreases more rapidly than the BPT, which levels off. The varying shapes of these curves reflect different assumptions about how earthquake probability changes with time.

The shortest interevent time (1812–1857) was preceded by the longest interevent time (304 yr) in the Pallett Creek paleoseismic record, suggesting that the long quiescent period may have influenced the exceptionally short subsequent interevent time. However, the commonly used earthquake probability models assume that the interevent times between earthquakes are independent, treating earthquakes as a renewal process (Cornell and Winterstein, 1988). In the time-independent exponential model, this interevent time independence is reflected in the constant earthquake probability with respect to time. The time-dependent models incorporate this independence assumption by resetting the probability to zero after each earthquake. Hence, the calculated probability of the next earthquake depends only on the time since the most recent earthquake and the known distribution of interevent times but not the specific sequence of interevent times for previous earthquakes.

Because earthquakes release strain accumulated on a fault, the independence assumption implies that an earthquake releases all strain accumulated since the previous one. Although assuming independence simplifies the probability calculation, it ignores the wealth of geologic observations showing long temporal gaps followed by clusters of earthquakes, suggesting that earthquakes often only partially release the accumulated strain (Wallace, 1970; Rockwell *et al.*, 2000; Friedrich *et al.*, 2003;



From state \ To state	1	2	3	4	5	6	7	8	9	10	11
1	0	**1.0**	0	0	0	0	0	0	0	0	0
2	**0.1**	0	**0.9**	0	0	0	0	0	0	0	0
3	**0.2**	0	0	**0.8**	0	0	0	0	0	0	0
4	**0.3**	0	0	0	**0.7**	0	0	0	0	0	0
5	0	**0.4**	0	0	0	**0.6**	0	0	0	0	0
6	0	0	**0.5**	0	0	0	**0.5**	0	0	0	0
7	0	0	0	**0.6**	0	0	0	**0.4**	0	0	0
8	0	0	0	0	**0.7**	0	0	0	**0.3**	0	0
9	0	0	0	0	0	**0.8**	0	0	0	**0.2**	0
10	0	0	0	0	0	0	**0.9**	0	0	0	**0.1**
11	0	0	0	0	0	0	0	**1.0**	0	0	0

Figure 2. The LTFM as a hidden Markov process. (a) Example with three earthquakes. (b) Representation of (a) as a Markov model with the probability space discretized into 11 states. (c) Transition probability matrix indicating probability of transition from state j to state k for this set of model parameters. The color version of this figure is available only in the electronic edition.

Weldon *et al.*, 2004; Sieh *et al.*, 2008; Goldfinger *et al.*, 2013; Salditch *et al.*, 2020; Hecker *et al.*, 2021).

In this article, as discussed next, we present an LTFM earthquake probability model allowing partial probability drops, reflecting partial strain releases in earthquakes. This behavior allows LTFM to incorporate more available information—the specific order of past earthquakes—in its forecasts. LTFM also is designed to require that earthquake probability always increases between earthquakes, reflecting monotonic strain accumulation.

CALCULATING EARTHQUAKE PROBABILITIES WITH LTFM

The LTFM model (Fig. 2a) builds on previous earthquake probability models but with modifications to model the temporal patterns of strain accumulation and release. LTFM, like some earlier models (e.g., Lomnitz-Adler, 1983), assumes that the probability of a large earthquake is linearly proportional to the strain accumulated on the fault. The accumulated strain, and hence probability, increases with time until an earthquake happens, after which probability decreases, but not necessarily to zero. The system thus retains long-term memory of earthquakes prior to the most recent, so the probability of an earthquake depends on prior earthquakes and so can remain relatively high over multiple cycles. Unlike the renewal models, the probability does not necessarily reset after an earthquake, so interevent times are not independent.

As originally formulated by Salditch *et al.* (2020), LTFM has two basic parameters: A, the rate at which strain, described by earthquake probability, increases, and R, the drop in probability (strain) from an earthquake. The drop R can be variable to simulate earthquakes of different magnitudes or constant for simplicity because the paleoearthquake record generally shows that past earthquakes were large enough to provide a clear record, but their specific magnitudes are unknown. Salditch *et al.* (2020) generally assumed a fixed R but indicated that it could be variable. An additional probability (strain) threshold that must be met or exceeded before an earthquake can occur can be included but is not for this analysis. How much memory is retained in the system after an earthquake depends on the size of the probability drop (R) relative to the rate of probability accumulation (A). If $R \gg A$, the probability usually resets to zero after an earthquake, as in renewal models. For smaller relative R values, some residual probability (also known as memory) could be (but not necessarily) retained after the earthquake. Thus, if an interevent time much longer than the average interevent time preceded the most recent earthquake, then residual strain may have remained on the fault after the earthquake. This simple model can recreate a wide range of earthquake recurrence patterns including clusters and gaps.

To calculate earthquake probabilities, we formulate the LTFM model as a hidden Markov model that yields analytical expressions. Markov models are stochastic models in which the probabilities of future outcomes only depend on the system's most recent state (Çinlar, 1975; Feller, 1968; Girardin and Limnios, 2018). For LTFM, the current state corresponds to the accumulated earthquake probability (commensurate with strain). This Markov process is hidden because we do not observe what state the system is in (the level of strain on the fault) (Rabiner, 1989). Instead, we only observe whether an earthquake occurs. We discretize the probability range 0–1 into N states such that the probability of an earthquake in state 1 is 0 and the probability of an earthquake in state N is 1 (Fig. 2b). In this formulation, $N = 1 + 1/A$, in which A is the probability increment in the original Salditch *et al.* (2020) formulation. The earthquake drop D corresponds to R/A in the original formulation. Conceptually, $1/N$ is the annual increase in earthquake probability (or strain), and D can be thought of as the number of years of accumulated probability (or strain) released by an earthquake.

At each epoch (timestep), an earthquake either occurs or does not. If an earthquake occurs, the system moves from state

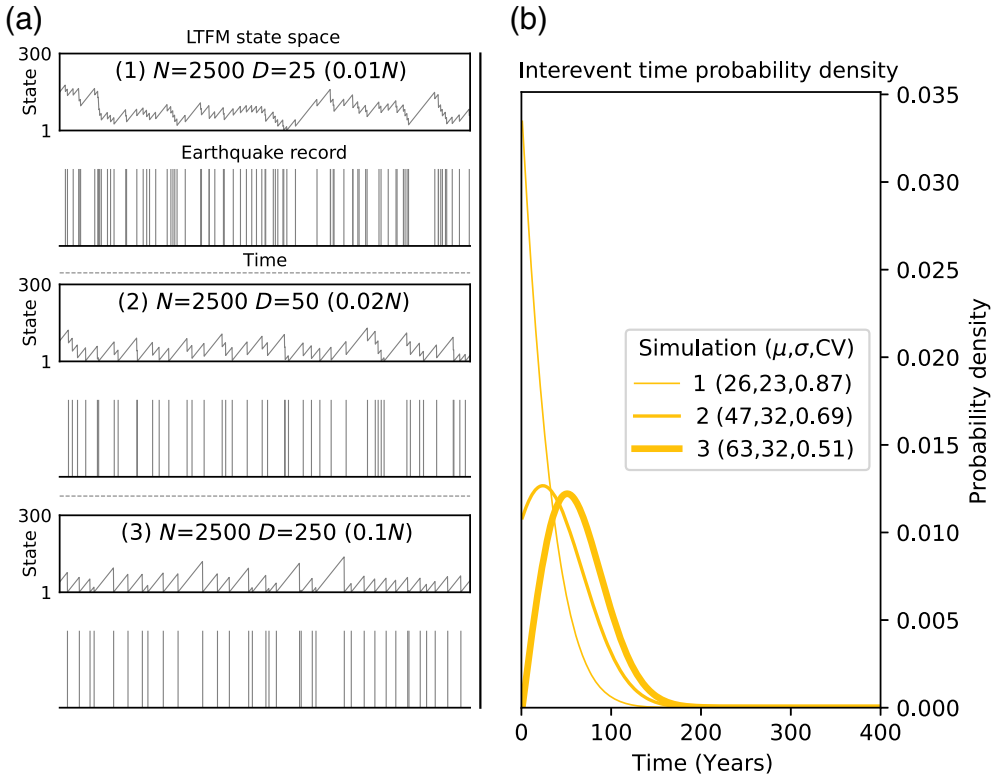


Figure 3. The LTFM simulations with different input parameters. (a) Earthquake record simulations with fixed number of states N and increasing earthquake drop D (1–3). (b) Long-run interevent time probability distributions for the simulations in panel (a). Thicker lines indicate larger values of D. Mean (μ), standard deviation (σ), and CV indicated. The color version of this figure is available only in the electronic edition.

n to a lower state $n - D$ or 1 (whichever is larger—LTFM's formulation does not allow for strain overshoot). If no earthquake occurs, the system moves to state $n + 1$. The conditional probability that an earthquake occurs in the current epoch given that the state is n equals $(n - 1) / (N - 1)$. Figure 2c shows the transition matrix giving the probability of transitioning from state j to state k. This simple example contains 11 states with a fixed earthquake drop of three states. For example, in state 5, the probability of an earthquake is 0.4. Therefore, the probability of having an earthquake and moving down to state 2 is 0.4, and the probability of no earthquake occurring and moving up to state 6 is 0.6. If we assume a fixed earthquake drop D (which we do for the rest of this article), the transition probability matrix is very sparse. Each row contains at most two entries—one corresponding to an earthquake occurring and one for an earthquake not occurring. All values in this $N \times N$ matrix are determined by and calculated from the parameters D and N. The number of states N should be large enough that the probability of ever entering state N (in which the probability of an earthquake equals 1) should be essentially zero.

Using this matrix, we can calculate earthquake probabilities following the procedures in the supplemental material available to this article. Similar to the existing probability models, we do not explicitly account for the different earthquake magnitudes in the paleoseismic record. Instead, we forecast the recurrence of large earthquakes that can have a range of magnitudes, but we treat each earthquake as having the same magnitude. Although the LTFM can be modified to allow for variable magnitude earthquakes, this would introduce additional model parameters, which would be difficult to constrain because most paleoseismic records contain little or no magnitude information.

Markov models have been used in seismic hazard analysis (Anagnos and Kiremidjian, 1988; Ebel *et al.*, 2007; Votsi *et al.*, 2013) but generally did not tie the states to strain increments as LTFM does. Anagnos and Kiremidjian (1985) proposed a Markov formulation with linear strain increments but specified a fixed strain threshold for earthquake occurrence, whereas LTFM allows earthquakes to randomly occur. Although calculating probabilities with LTFM is more complicated than current approaches, LTFM is designed to model the earthquake process and the history of earthquakes on the fault.

With just two estimated parameters—number of states N and earthquake drop size D—LTFM can replicate a wide range of observed earthquake recurrence patterns (Salditch *et al.*, 2020). Figure 3 shows a set of simulations with a fixed N and increasing D and the resulting long-run distribution of interevent times. This long-run PDF shows the expected distribution of interevent times if we observed the fault for a sufficiently long time, analogous to the exponential, lognormal, and BPT PDFs in Figure 1b. We use equation (S.12) from the supplemental material to calculate the LTFM's long-run distributions.

Simulations (1–3) in Figure 3a show how fixing the number of states N and increasing the size of the earthquake drop D changes the behavior of the LTFM model. For each simulation, we show the state space history showing the probability (strain) history and the corresponding earthquake record. In simulation (1), $D \ll N$, so after an earthquake, the system rarely drops to the lowest state (in which the probability of an earthquake equals 0), leaving residual probability for another earthquake.

In this case, residual strain can build up after earthquakes, so the earthquake record contains clusters and long gaps. However, as D increases relative to N (simulations 2 and 3), the system increasingly drops to the lowest state after an earthquake, making residual probability less likely. This reduces the number of short recurrence intervals and so increases the mean expected long-run interevent time and decreases the coefficient of variation (CV, standard deviation divided by the mean) (Fig. 3b). With a lower CV, the earthquake record looks more regular (and periodic) with fewer outlier (short or long) interevent times. If the system ever reached state 2500 (N), the probability of an earthquake would equal 1; however, the probability of reaching this state is essentially zero. In Figure 3a simulations, the system's state never exceeds 250 (0.1 N).

The shapes of the long-run interevent time distributions (Fig. 3b) illustrate how likely memory is to be retained after an earthquake based on the size of D relative to N. Higher y-intercepts indicate more residual strain (probability) after an earthquake in the long run, making very short interevent times possible, as in simulation (1) for $D \ll N$. Conversely, simulation (3) has a y-intercept of essentially zero, so very short recurrence times are very unlikely. For D large relative to N, LTFM behaves in the long run like a renewal model with no residual strain after an earthquake. Beyond a certain point, increasing D relative to N does not impact the expected interevent time distribution. A simulation with D = 500 and N = 2500 would be indistinguishable from simulation (3) with D = 250 and N = 2500, and the CV does not continue to decrease below ~0.5 without the addition of a third parameter (the threshold parameter) in the model.

INTEREVENT TIME ORDER MATTERS FOR LTFM

The advantages of LTFM's partial strain (probability) drop are illustrated by two synthetic paleoseismic records that contain the same interevent times but in a different order (Fig. 4a,b). In Figure 4a, recent interevent times have been short, whereas in Figure 4b, they have been long. We use a maximum-likelihood estimation (MLE) approach to find the distribution parameters that best fit each record. The MLE approach allows us to find the combination of parameters that most likely produced that specific sequence of earthquakes. Because the exponential, lognormal, and BPT models assume interevent time independence, the order does not matter, so the best-fitting parameters (and therefore probability distributions) for these models are the same for both records (Fig. 4c.d). However, LTFM behaves differently. Using an MLE grid search, we estimate the best-fitting number of states N and earthquake drop size D for both records. (Details of the MLE procedure and calculation of the probability density and hazard curves are in the Supplemental Methods section in supplemental material). Because LTFM does not assume interevent time independence, N = 11,300 and D = 100 for the original record, and N = 7100 and D = 180 for the reordered record.

These parameters yield LTFM's long-run distribution of interevent times (solid LTFM-LR lines in Fig. 4c,d). Although both records contain the same interevent times, LTFM fit to the original record indicates that the fault has more memory (retained strain) in the long run compared with the LTFM fit to the reordered record. This is shown by the LTFM-LR distribution fit to the original record having a larger positive y-intercept (Fig. 4c) than when fitted to the reordered record (Fig. 4d). The coefficients of variation for the LTFM-LR models differ from the time-dependent renewal models. For the lognormal and BPT models, the MLE produces parameters that match the paleoseismic record's CV. LTFM-LR, however, assumes that a paleoseismic record is described by both the CV and the order of events, so the MLE parameter estimates reflect both pieces of information. As Salditch *et al.* (2020) demonstrated, records with the same CV can look very different, so the CV only tells part of the story about the earthquake record.

Furthermore, we can use LTFM to calculate short-term earthquake probabilities conditional on the most recent sequence of interevent times (equation S.27 in the supplemental material). These short-term probabilities incorporate extra information—the timing of the most recent prior P interevent times (LTFM-P1, LTFM-P2, and LTFM-P3 for P = 1, 2, and 3). The corresponding short-term forecasts of the distribution of the time until the next earthquake can vary greatly from the long-run probability distribution. In contrast, because the exponential, lognormal, and BPT assume interevent time independence, their short- and long-term forecasts are the same. For clustered paleoseismic records, seismologists often must decide whether to assume the fault is currently in a cluster (shorter average interevent time) or a gap (longer average interevent time) and produce forecasts accordingly (Sieh *et al.*, 1989). By allowing residual earthquake probability, LTFM takes this difficult decision out of the hands of the analyst by effectively making continuous adjustments in a data-driven manner.

LTFM long-run and short-term forecasts estimate the probability that the fault is in a given state to calculate earthquake probabilities. If the fault is in a low state (low earthquake probability), then it will likely be a while until an earthquake occurs. Conversely, a high state (high earthquake probability) means an earthquake is more likely to occur soon. With the additional information about the specific timing of the most recent P interevent times, the short-term forecast updates these state probability estimates to produce an updated earthquake forecast. In our analysis, after conditioning on the first few recent interevent times, the short-term forecast stabilizes and does not change noticeably with the inclusion of additional interevent times further back in time.

For the original record, conditioning on the most recent 1, 2, or 3 interevent times decreases the probability of an earthquake in the near future by shifting the probability distributions (LTFM-P1, LTFM-P2, and LTFM-P3 in Fig. 4c) to the right.

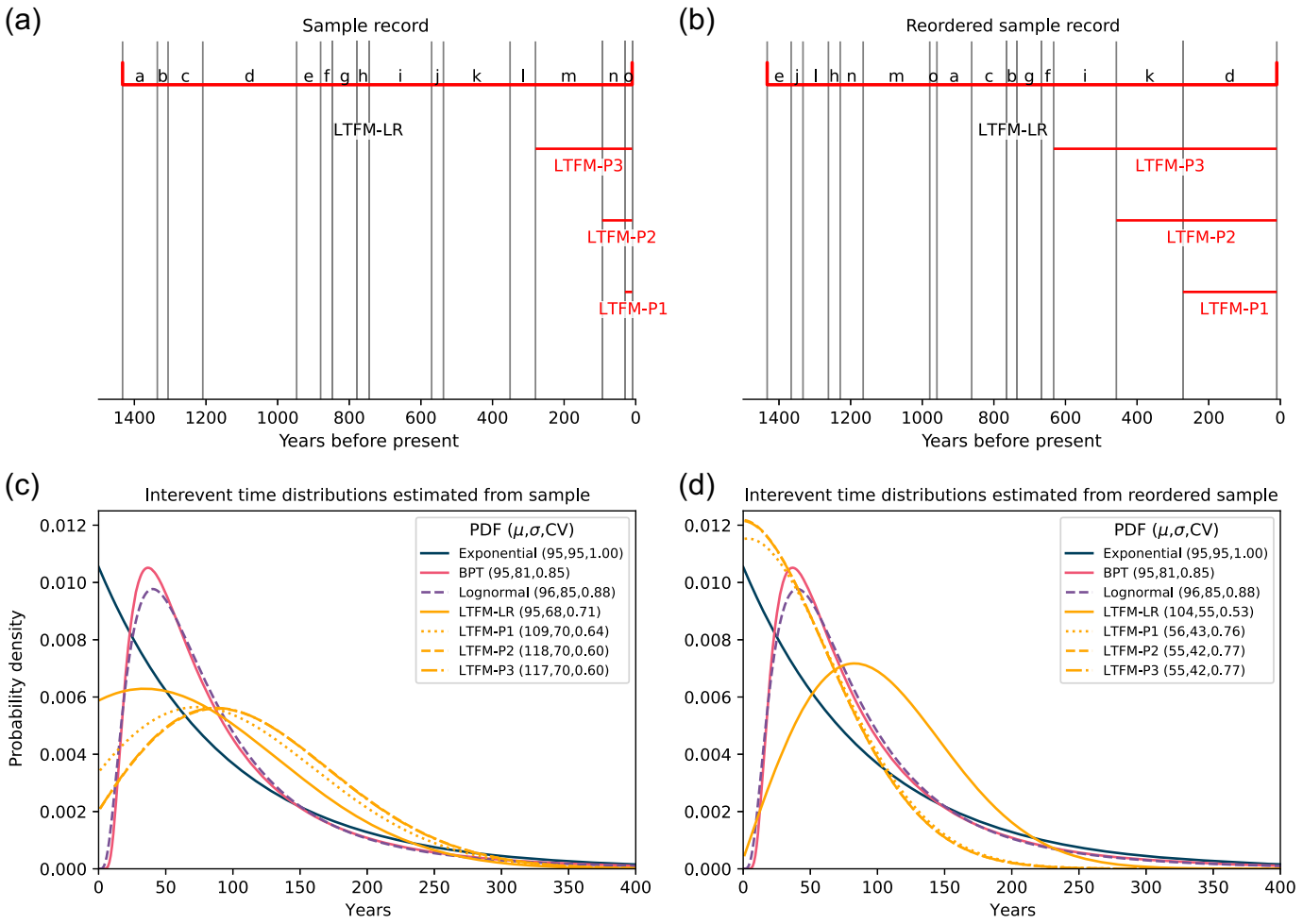


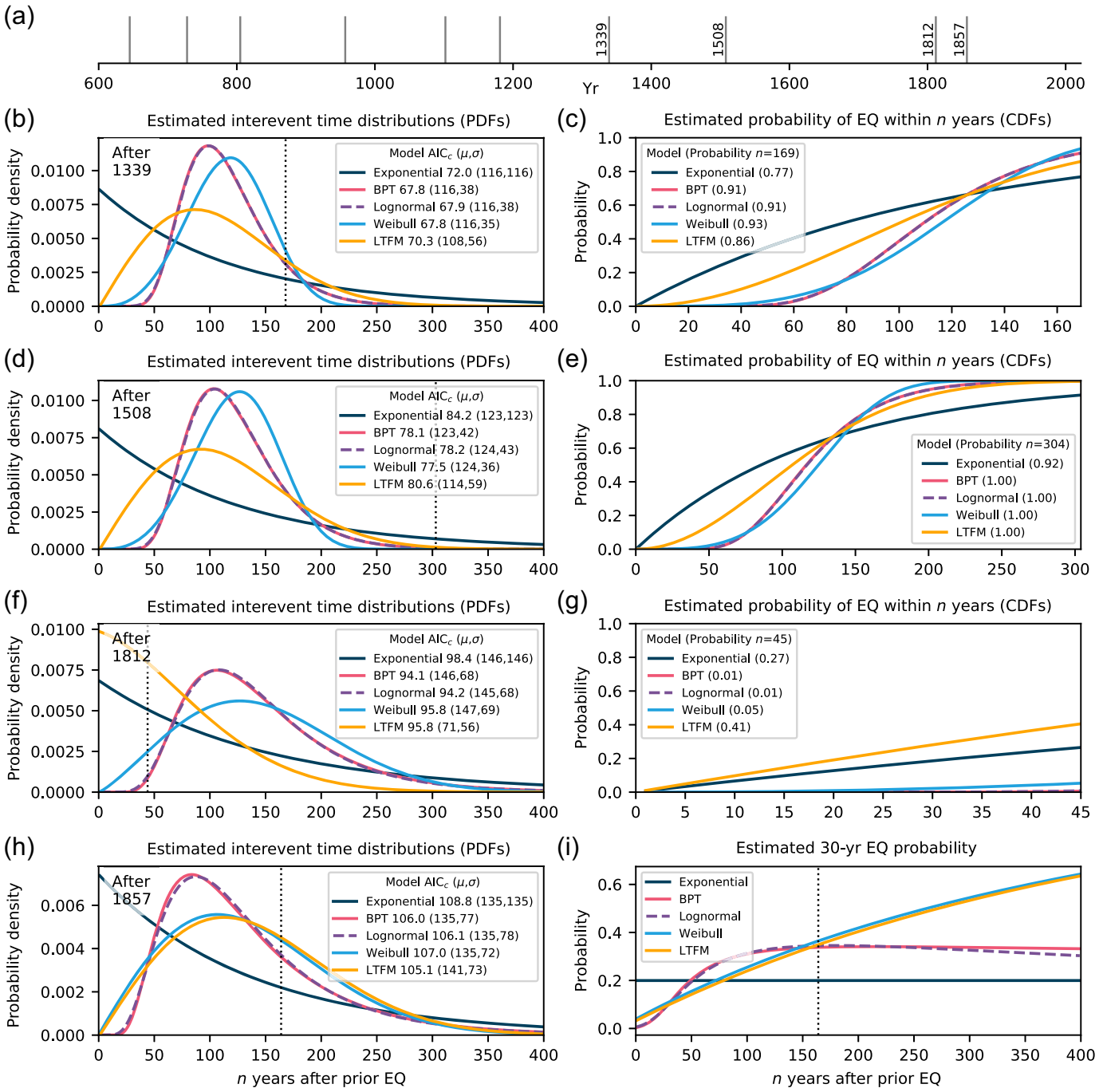
Figure 4. The LTFM applied to sample records. (a) Earthquake record with interevent times labeled. The best-fitting LTFM model parameters for this sequence of interevent times yield long-run probability distributions for the time until the next earthquake (LTFM-LR) and conditional probability distributions for the next earthquake given the most recent P interevent times (LTFM-P1, LTFM-P2, and LTFM-P3). (b) Reordered version of the record in panel (a), such that recent interevent times have been long. (c) Estimated interevent time probability distributions for the record in panel (a) with mean (μ), standard deviation (σ), and CV indicated. (d) Estimated interevent time probability distributions for the record in panel (b). The exponential, lognormal, and BPT forecasts are the same as in panel (c), but the LTFM forecasts differ because they depend on the order of interevent times. The color version of this figure is available only in the electronic edition.

This increases the expected mean interevent time for the next earthquake relative to the long-run distribution (LTFM-LR). The distribution shifts to the right because the most recent interevent times are short compared with the others in the record. LTFM in effect infers that the fault must have released a lot of strain recently, so there is likely little residual strain (probability) left over after the most recent earthquake. LTFM-P2 and LTFM-P3 are similar because when information about the two most recent interevent times is used, the third provides little additional information. However, LTFM behaves differently for the reordered record. Here, conditioning on the 1, 2, or 3 most recent interevent times shifts the distribution to the left (LTFM-P1, LTFM-P2, and LTFM-P3 in Fig. 4d) with a lower expected mean interevent time until the next earthquake. The distribution shifts to the left because the recent interevent times have been exceptionally long, leaving significant residual strain (probability) after the most recent earthquake.

APPLYING LTFM TO THE MOJAVE SECTION OF THE SAN ANDREAS FAULT

To explore how LTFM compares with other models when applied to real paleoseismic records, we use the Pallett Creek record to create forecasts after the four most recent earthquakes in 1339, 1508, 1812, and 1857 (Fig. 5a). For each earthquake, we produce a forecast immediately after the earthquake occurred. These forecasts include the paleoseismic record up to and including the most recent earthquake but no future earthquakes. This allows us to assess how well the various models forecast the next earthquake. We include an additional time-dependent renewal model—the Weibull distribution, discussed shortly—for comparison. The LTFM model shown is the short-term forecast incorporating all the preceding earthquakes.

Figure 5b shows the estimated PDFs for the next interevent time after the 1339 earthquake. The models do not "know" that the next earthquake will occur in 1508. We assess the different



models in two ways. First, we look at how well the models fit the available preceding data using the corrected Akaike information criterion (AICc) (Hurvich and Tsai, 1989), in which lower scores are better. Second, we assess how well the models forecast when the next earthquake occurs. In Figure 5b, the vertical dashed line shows when the next earthquake (1508) occurs. Figure 5c shows the cumulative distribution functions (CDFs), the PDFs integrated from 0 to n, which indicate the probability of an earthquake occurring within n years.

None of the models stand out for the post-1339 forecast. The AICc values are similar (Fig. 5b), and the estimated probabilities of an earthquake within 169 yr (when the next

Figure 5. Earthquake forecasts after recent Pallett Creek sequence earthquakes. (a) Pallett Creek paleoseismic record with the four most recent earthquakes labeled. (b) Estimated probability density functions (PDFs) for interevent time until the next earthquake after the 1339 earthquake. AICc, empirical mean (μ), and standard deviation (σ) indicated for each model. Vertical dashed line indicates when the next earthquake actually occurred. (c) Cumulative distribution functions (CDFs) for models in panel (b). Probability of earthquake occurring within 169 yr (the actual observed interevent time) is indicated. (d) PDFs after the 1508 earthquake. (e) CDFs after the 1508 earthquake. (f) PDFs after the 1812 earthquake. (g) CDFs after the 1812 earthquake. (h) PDFs after the 1857 earthquake for the current quiescent period. Vertical dashed line indicates the year 2022. (i) Estimated 30 yr earthquake probabilities with the year 2022 indicated. The color version of this figure is available only in the electronic edition.

earthquake actually occurs) are nearly indistinguishable (Fig. 5c). The exponential model has the highest AICc and the lowest probability of the next earthquake occurring when it did. For the post-1508 earthquake forecasts (Fig. 5d,e), all the models perform equally poorly. No model stands out based on the AICc, and none forecasted the long 304 yr interevent time. In the PDF plot (Fig. 5d), the dashed line (indicating the next earthquake in 1812) intersects the expected PDFs at relatively low values, indicating that such an exceptionally long quiescent period was unexpected. The occurrence of such a low-probability event should make one carefully consider whether the model fails to hold. Correspondingly, the CDF curves with values of essentially 1 (Fig. 5e) indicate that all models expected the next earthquake to occur long before it actually did.

LTFM, however, distinguishes itself for the forecast after the 1812 earthquake (Fig. 5f,g). The next earthquake occurred in 1857—a remarkably short interevent period—but how well would any of the models have forecast such a short interval? Although the AICc again does not distinguish how well the models fit the past data, LTFM better forecasts the short time until the next earthquake in 1857 (Fig. 5f). LTFM forecasts that on average, the next earthquake will occur in about 71 yr, whereas the other models forecast an interevent time double that. LTFM's large positive y-intercept suggests that a lot of residual strain remains after the 1812 earthquake, which makes intuitive sense considering the exceptionally long, preceding 304-yr interevent time. The CDFs (Fig. 5g) show that the probability of such a short interevent time is extremely low in the lognormal (0.01), BPT (0.01), and Weibull (0.05) models. The exponential model (0.27)—which assumes constant earthquake probability—forecasts a higher probability of a short interevent time compared with the renewal models. LTFM, which conditions on the specific sequence of preceding earthquakes and so explicitly incorporates the recent long prior interevent time, forecasts the highest probability (0.41) of an earthquake by 1857, as actually occurred. This example suggests that allowing residual earthquake strain and conditioning on the prior interevent times provides a more accurate assessment of future earthquake hazards, especially when the most recent earthquake is preceded by a relatively long interevent time.

Figure 5h shows the estimated earthquake interevent time distribution after the most recent earthquake in 1857. This is the forecast for the earthquake that has not happened yet. Despite LTFM's good performance forecasting the 1857 earthquake, the AICc scores are still nearly indistinguishable. The current quiescent period up through 2022 (indicated by the dashed vertical line) is slightly longer than the expected mean interevent times of the models. We have plotted the estimated 30 yr earthquake probabilities for the Weibull and LTFM models (Fig. 5i) and compare them with the 30 yr probabilities previously shown in Figure 1c. The Weibull forecast is nearly identical to the LTFM forecast. In 2022, the LTFM is indistinguishable from the time-dependent models. However, moving forward, the LTFM diverges. If no earthquake occurs by 2050, the lognormal and BPT models give probabilities of 35% and 34%, whereas LTFM predicts 39%. This difference continues to grow as the quiescent period continues. However, if an extremely long (and statistically unlikely) quiescent period occurs, it could indicate that the estimated model parameters are incorrect.

ASSUMPTIONS ABOUT EARTHQUAKE PROBABILITY AND ACCUMULATED STRAIN

The differences in the estimated 30 yr earthquake forecasts (Fig. 5i) reflect assumptions about accumulated strain and earthquake probability. To better understand the accumulated strain (probability) relationship, we examine each PDF's corresponding hazard-rate function. The hazard rate is conceptually similar (although not identical) to the conditional probability of an earthquake occurring in the next time increment (in this case, 1 year) given that one has not occurred since the most recent earthquake. It can be calculated using $f(t) / (1 - F(t))$, in which t is the time since the most recent earthquake; $f(t)$ is the PDF; and $F(t)$ is the CDF, the integral of the PDF. Figure 6a shows four PDFs with the same mean and standard deviation (except for the exponential PDF whose mean is also the standard deviation), and Figure 6b shows the corresponding hazard-rate functions.

Although the hazard rate is a function of time, we can also interpret it in terms of accumulated strain. The agreement between long-term plate motions that load faults and the short-term loading seen geodetically (Gordon and Stein, 1992) indicates that strain steadily accumulates on faults in the interseismic period between large earthquakes, so we can substitute strain for time. The hazard rate thus indicates how earthquake probability changes with accumulated strain, and the derivative of the hazard-rate function (Fig. 6c) shows whether each additional unit of accumulated strain increases or decreases earthquake probability. The exponential distribution's hazard-rate derivative is zero, so additional accumulated strain does not affect the probability of an earthquake. For the lognormal and BPT PDFs, each additional unit of accumulated strain has a different probability increment. The corresponding probability change for each strain increment increases quickly after an earthquake, then begins to decrease, and then ultimately turns negative, indicating that future strain accumulation decreases earthquake probability. How much the BPT hazard-rate curve decreases depends on its parameters, but it eventually levels off, mimicking the time-independent behavior of the exponential model. Unlike the other models, in LTFM each additional strain unit corresponds to a constant increase in earthquake probability.

Interestingly, LTFM's assumptions about accumulated strain (earthquake probability) are shared by the Weibull distribution. This distribution, used in reliability engineering, has

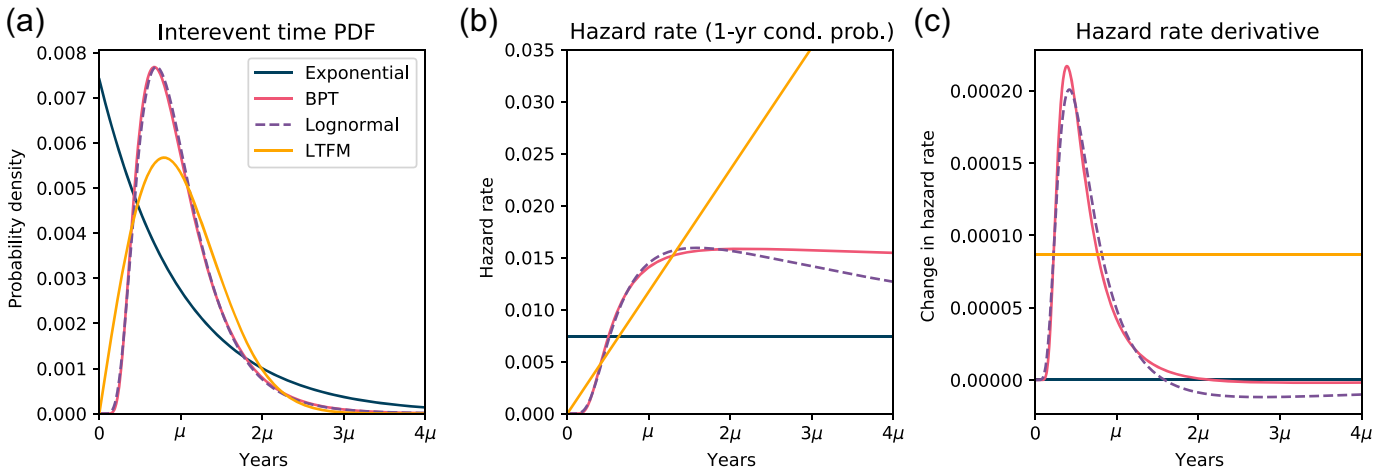


Figure 6. Comparison of probability models. (a) PDFs for the interevent time with same mean and standard deviation—except exponential where mean equals standard deviation. (b) Corresponding hazard-rate functions. The hazard rate at time t is the probability density that an earthquake will occur at t given that one has not occurred since the past earthquake. (c) The derivative of each hazard-rate curve shows whether the hazard rate is increasing, decreasing, or staying constant with time as each new strain increment accumulates. The color version of this figure is available only in the electronic edition.

been used to describe the distribution of earthquake interevent times (Hagiwara, 1974; Chou and Fischer, 1975; Brillinger, 1982). The Weibull distribution has two parameters: a scale parameter η and shape parameter β. The shape parameter β controls the PDF's CV (Fig. 7a) and the shape of the hazard-rate function (Fig. 7b). For $\beta<1$, the hazard rate decreases with time, making failure less likely as time passes. For $\beta=1$, the hazard rate is constant because the Weibull distribution becomes the exponential distribution. For $\beta>1$, the hazard rate increases with time, making failure more likely. For $\beta=2$, the hazard rate increases linearly, implying a linear increase in the earthquake probability, as in LTFM. Brillinger (1982) noted that when $\beta=2$, the linear increase in earthquake probability appropriately mimics the steady strain increases that drive earthquakes.

Plotting observed earthquake interevent times against cumulative hazard probabilities shows that the Weibull distribution is often appropriate (Chou and Fischer, 1975). The cumulative hazard is the integral of the hazard rate and can be calculated by sorting the interevent times and open quiescent period in ascending order (1–K), calculating the reverse rank (for the kth ordered interval, the reverse rank is $K-k+1$), taking the inverse of the reverse rank to get the hazard value, and finally cumulatively summing the hazard value of only the closed interevent times to get the cumulative hazard at each data point. If the resulting data can be fit by a straight line on a log–log plot, then the slope of the line is the inverse of the shape parameter β. Brillinger (1982) and Sieh *et al.* (1989) applied this method to an earlier Pallett Creek dataset, finding shape parameters of ~2 and 1.5 ± .8. A similar analysis (Fig. 8a) shows that the Pallett Creek dataset we used is well-fitted by a line whose corresponding β value is 1.84. This suggests that a model with linearly increasing earthquake probability ($\beta=2$) is reasonable.

The long-run LTFM-LR and Weibull models (either with best-fitting β or with a fixed $\beta=2$) have quite similar PDFs (Fig. 8b). The hazard-rate functions (Fig. 8c) further highlight the similarities, with both indicating a constantly increasing earthquake probability reflecting the linear accumulation of strain. They differ slightly because LTFM assumes that on average, there will be some residual strain after an earthquake, so there may be some very short recurrence times (Fig. 8b). The Weibull model, which is a renewal model, lacks long-term memory and so does not predict these. Hence, Weibull, like the common renewal models, does not describe clusters and gaps and so does much worse than LTFM in forecasting the 1857 Fort Tejon earthquake (Fig. 5g).

UNCERTAINTY IN LTFM FORECASTS

Understanding the uncertainty in the LTFM model is critical for assessing the usefulness of its estimated probabilities. Figure 9a shows the outcome of the MLE grid search to find the best N and D LTFM parameters for the complete Pallett Creek record. The best-fitting solution is marked by the orange x, in which the negative log likelihood is minimized. There is a very broad region of similar log-likelihood values. We use the negative log-likelihood values to approximate the LTFM's 68% and 95% confidence intervals (CIs) using $\mathrm{NLL}_{\min}+\chi^2_2(1-\alpha)/2$, in which $\mathrm{NLL}_{\min}$ is the negative log-likelihood value of the best-fitting N and D pair and χ^2_2 is the chi-squared distribution with two degrees of freedom, and α is the critical value of interest (Bolker, 2008).

The 68% CI has an upper N bound but not an upper D bound. However, the lack of an upper D bound is inconsequential because an ever-increasing D ultimately does not change the predicted behavior, as noted in Figure 3. The 95% CI is

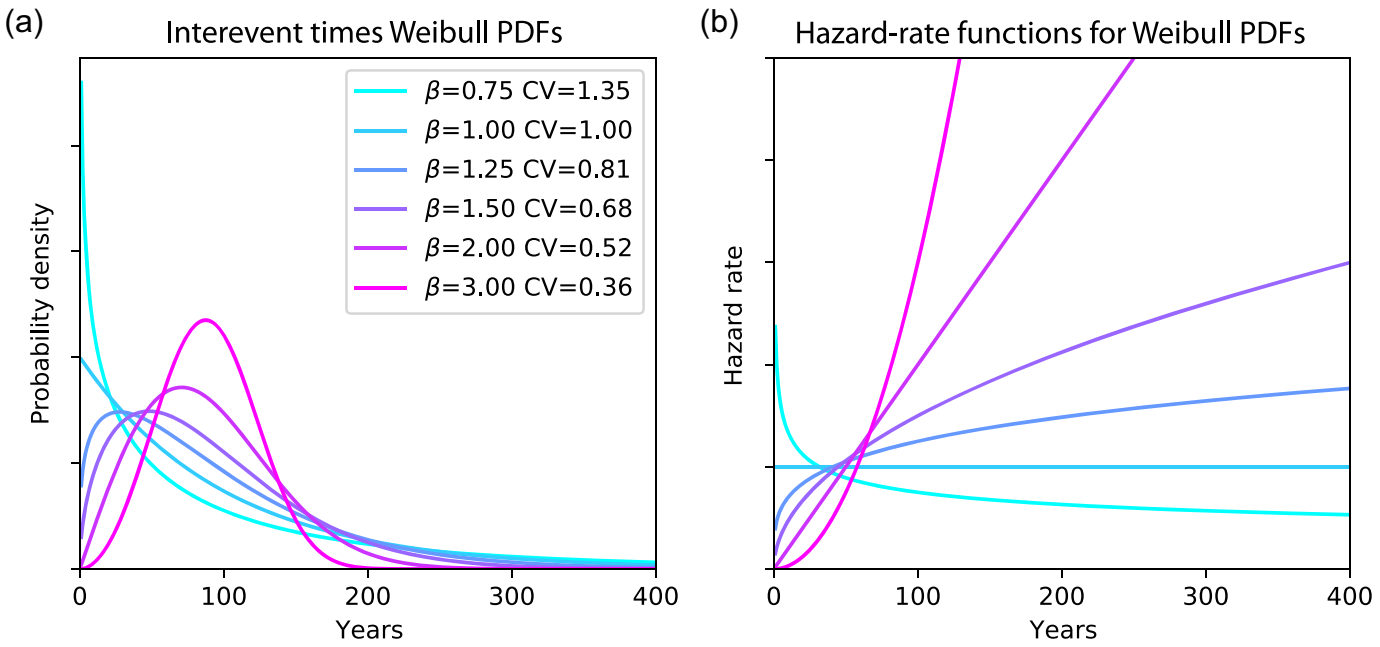


Figure 7. Weibull distribution with different shape parameters (β). (a) Weibull PDFs plotted with same scale parameter η but different β shape parameters. CV indicated. (b) Hazard-rate functions for the PDFs shown in panel (a). The color version of this figure is available only in the electronic edition.

not as well bounded with open contours in both the N and D directions. We selected some parameter combinations within the 95% CI (gray stars in Fig. 9a) and calculated the corresponding 30-yr earthquake forecasts (Fig. 9b). The forecasts for 2022 (Fig. 9b) range from approximately a 15%–45% probability of an earthquake in the next 30 yr. Although this range may seem large, a similar uncertainty analysis of the time-dependent BPT yields a forecast spread just as large. Using the Hessian matrix evaluated at the BPT's MLE solution, we calculate the covariance matrix for the BPT's two parameters and then randomly sample from a 2D Gaussian distribution constructed from the BPT's parameter uncertainties to get new BPT parameters. The corresponding 30 yr earthquake forecasts for 200 of these samples are shown in light red (Fig. 9b). These BPT forecasts show a spread even larger than the LTFM forecasts, ranging from anywhere between 20% and more than 80%.

The LTFM forecasts are constructed based on the estimated current state of the system after the most recent earthquakes. We do not know the initial state at the start of the sequence, but we can calculate the probability of being in a given state using the long-run state probability vector $\boldsymbol{\pi}$ (equation S.3 in the supplemental material). Figure 10 shows histories for the Pallett Creek sequence for the best-fitting MLE parameter combination assuming the system was in state 50, 200, or 300 before the first earthquake. Given the initial starting state, it is straightforward to calculate the history given the number of states (N) and the size of the earthquake drop (D). Because the average earthquake drop D (175) is larger than any of the initial interevent times, all three curves converge after the first few earthquakes, indicating that any initial residual strain at the beginning of the sequence is soon released and so the uncertainty of the starting residual strain is unimportant to the probability estimates today. The history stays in the lowest portion of the state space nowhere near the maximum state of 12,500. To reach such a high state would require an exceptionally long (~10,000 yr) quiescent period, which is very unlikely along the San Andreas fault.

SELECTING THE MOST APPROPRIATE MODEL

Determining which model is most appropriate for earthquake forecasts is challenging. As shown in Figure 5 with the AICc analysis, statistical tests struggle to differentiate between models when paleoseismic records contain so few earthquakes. Other studies have found that time-dependent models tend to fit the observed interevent times better than a time-independent model (Scharer *et al.*, 2010), but differentiating the time-dependent models is challenging because of the relatively few earthquakes in a paleoseismic record (Matthews *et al.*, 2002). As a result, one must look beyond statistical measures when justifying model choice.

Bookstein (2021) argued (using maximum entropy) that in the absence of any additional information besides the earthquake dates, the interevent times will follow a truncated Gaussian distribution. However, we have more information about the earthquake process than just the dates of the earthquakes: we know that the earthquake process is fundamentally about strain accumulation and release. We believe that these strain considerations should not be ignored. In particular, because of the limited length of the paleoseismic time series,

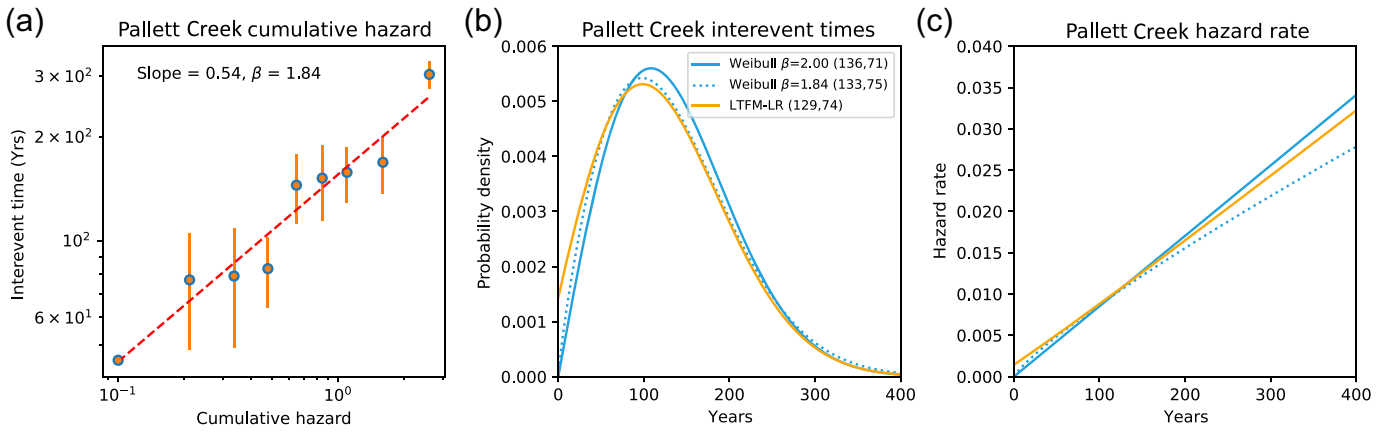


Figure 8. Comparisons of Weibull and LTFM. (a) Pallett Creek interevent time cumulative hazard plot for Weibull distribution with 1σ uncertainties of the interevent times indicated. (b) Comparison of best-fitting Weibull distribution and β=2 Weibull distribution to the long-run LTFM-LR interevent time distribution. Interevent time mean (μ) and standard deviation (σ) are indicated. (c) Corresponding hazard-rate functions for PDFs in panel (b). The color version of this figure is available only in the electronic edition.

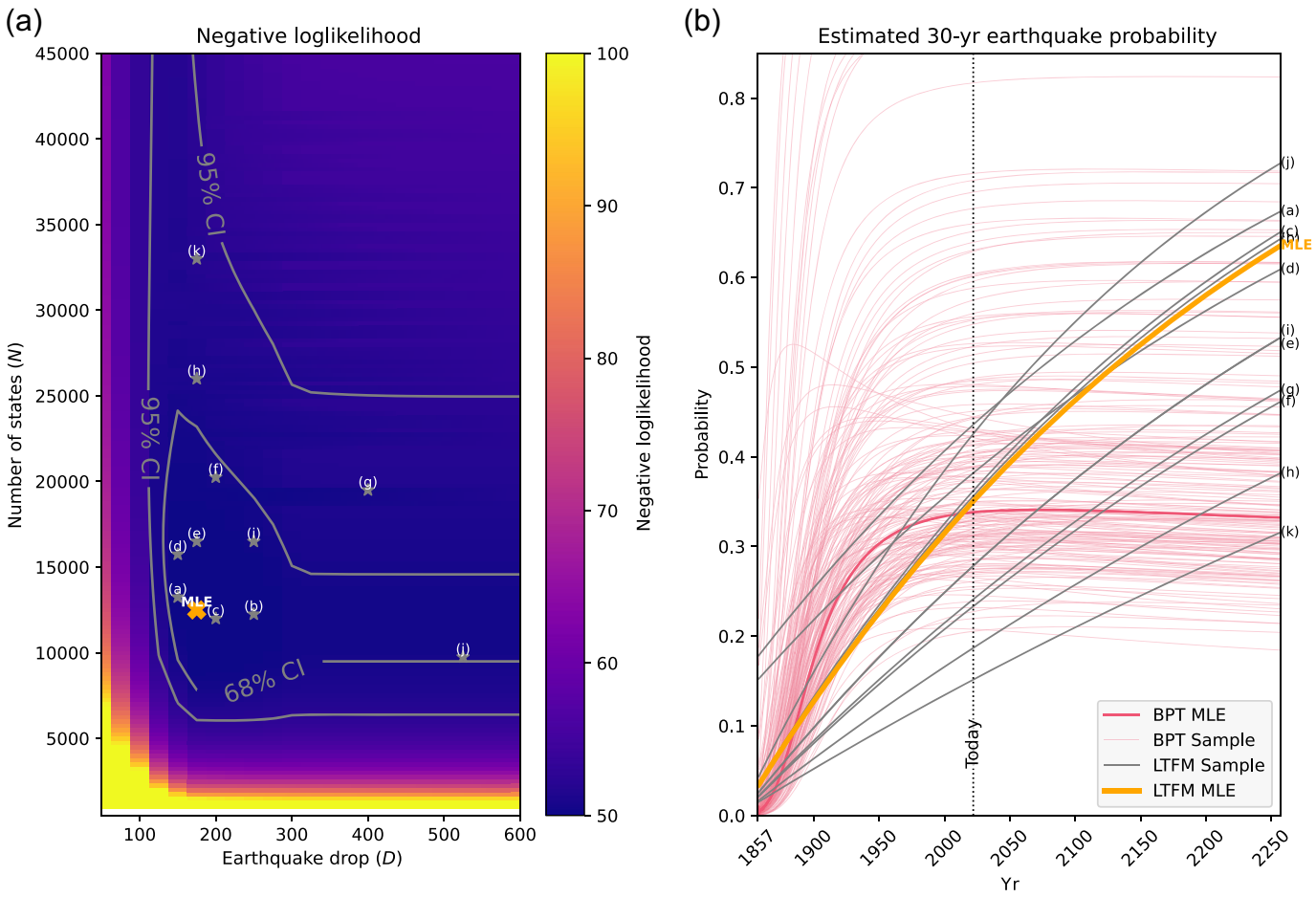


Figure 9. LTFM uncertainty analysis. (a) Plot of LTFM Pallett Creek negative log-likelihood values for different N and D combinations. Best-fitting parameters are indicated by orange X. Gray contours indicate 68% and 95% confidence intervals. Stars indicate parameter combinations used in panel (b). (b) Corresponding 30 yr earthquake forecasts for the parameter combinations. Thick orange line corresponds to the LTFM best-fitting parameters (orange X in panel a). Thin gray lines correspond to the randomly sampled LTFM parameter combinations. Labeling indicates LTFM parameter combination shown in panel (a). The solid pink line is the best-fitting BPT forecast and light pink lines are 200 randomly sampled BPT forecasts based on its estimated parameter uncertainty. The color version of this figure is available only in the electronic edition.

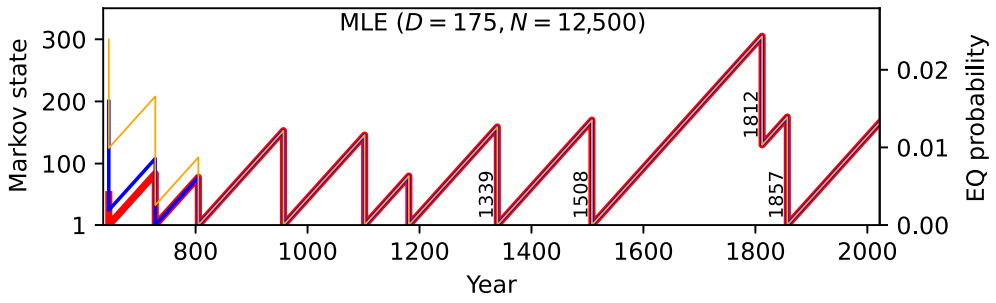


Figure 10. LTFM realizations of Pallett Creek record. Possible Markov state and corresponding earthquake probability histories for the Pallett Creek record for the best-fitting maximum-likelihood estimation (MLE) parameters. Three different initial starting states (300, 200, and 50) are shown (indicated by the orange, blue, and red lines, respectively). The three curves converge after the first three earthquakes indicating how any residual strain at the start of the record is shortly eliminated. The color version of this figure is available only in the electronic edition.

it is important to incorporate as much information as possible about the earthquake process to produce more robust forecasts.

Currently, attention is primarily paid to how the shape of the PDF fits the distribution of interevent times but not to the strain implications of the hazard function. As shown in Figure 6, the constant probability of the time-independent exponential model means it is strain agnostic—no matter how much strain accumulates on a fault, the earthquake probability stays the same. The time-dependent lognormal and BPT models, however, reflect implicit assumptions about accumulated strain and earthquake probability. Traditionally, the lognormal distribution was preferred because its PDF's fatter right tail better fits interevent time data that are mildly bimodal, but as Matthews *et al.* (2002) noted, the hazard-rate function eventually decreases to zero, indicating the cessation of the process driving earthquakes. Matthews *et al.* (2002) instead proposed the BPT, whose PDF has a similar shape to the lognormal's PDF but whose hazard-rate function levels off for long quiescent periods. The BPT mimics a constant stress loading rate with random stress perturbations reflecting outside stress interactions.

LTFM's explicit modeling of both linearly increasing earthquake probability with time (mimicking linearly increasing strain) and partial probability drops (allowing residual strain) attempts to more accurately reflect the underlying process that drives earthquakes. LTFM's accurate forecast for the 1857 Fort Tejon earthquake demonstrates the usefulness of incorporating our understanding of the strain process into the model's design. However, it is just one earthquake, and because of the stochastic nature of the forecasts, there are likely instances where LTFM performs worse. Although all earthquake probability models make simplifying assumptions to make calculations tractable, it is important to understand and acknowledge such trade-offs when selecting a model.

CONCLUSIONS

LTFM advances beyond the current time-dependent renewal models by assuming earthquake probability increases with accumulated strain and allowing residual strain after an earthquake. It incorporates the specific earthquake history, allowing estimates of earthquake probabilities conditioned both on the time of the last earthquake and the specific sequence of preceding earthquakes. Incorporating the specific history impacts earthquake probability calculations, especially when the most recent earthquake is preceded by a relatively long interevent time. Although LTFM is computationally more complex than the existing earthquake probability models, it is far more powerful. It allows forecasts based on more information than the current models can include and should lead to more accurate earthquake probability estimates, especially when an earthquake record contains long gaps and clusters of earthquakes.

DATA AND RESOURCES

All data used in this article came from published sources listed in the references. All methods and equations are provided in the supplemental material to replicate results.

DECLARATION OF COMPETING INTERESTS

The authors declare that there are no conflict of interests recorded.

ACKNOWLEDGMENTS

The authors thank the Institute for Policy Research at Northwestern University for helping to fund this work. The authors thank Sue Hough and Norman Abrahamson for helpful discussions. The authors also thank Glenn Biasi and an anonymous reviewer for their thoughtful reviews of this article.

REFERENCES

Anagnos, T., and A. S. Kiremidjian (1985). A stochastic earthquake recurrence model with temporal and spatial dependence, *Report No. 76, John A. Blume Earthquake Engineering Center*, Stanford University, Stanford, California.

Anagnos, T., and A. S. Kiremidjian (1988). A review of earthquake occurrence models for seismic hazard analysis, *Probab. Eng. Mech.* **3,** 3–11, doi: 10.1016/0266-8920(88)90002-1.

Biasi, G. P., R. J. Weldon, T. E. Fumal, and G. G. Seitz (2002). Paleoseismic event dating and the conditional probability of large earthquakes on the southern San Andreas Fault, California, *Bull. Seismol. Soc. Am.* **92,** 2761–2781, doi: 10.1785/0120000605.

Bolker, B. (2008). *Ecological Models and Data in R*, Princeton University Press, Princeton, New Jersey.

Bookstein, F. L. (2021). Estimating earthquake probabilities by Jaynes's method of maximum entropy, *Bull. Seismol. Soc. Am.* **111,** 2846–2861, doi: 10.1785/0120200298.

Brillinger, D. R. (1982). Seismic risk assessment: Some statistical aspects, *Earthq. Predict. Res.* **1,** 183–195.

Chou, I. H., and J. A. Fischer (1975). Earthquake hazard and confidence, *Proc. of the U.S. National Conf. on Earthquake Engineering 1975*, 34–42.

Çinlar, E. (1975). *Introduction to Stochastic Processes*, Prentice Hall, Englewood, Colorado.

Cornell, C. A., and S. R. Winterstein (1988). Temporal and magnitude dependence in earthquake recurrence models, *Bull. Seismol. Soc. Am.* **78,** 1522–1537, doi: 10.1785/BSSA0780041522.

Ebel, J. E., D. W. Chambers, A. L. Kafka, and J. A. Baglivo (2007). Non-Poissonian earthquake clustering and the hidden Markov model as bases for earthquake forecasting in California, *Seismol. Res. Lett.* **78,** 57–65, doi: 10.1785/gssrl.78.1.57

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. I, Third Ed., Wiley, Somerset, New Jersey.

Field, E. H., G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer, D. A. Jackson, K. M. Johnson, T. H. Jordan, C. Madden, A. J. Michael, *et al.* (2015). Long-term time-dependent probabilities for the third Uniform California Earthquake Rupture Forecast (UCERF3), *Bull. Seismol. Soc. Am.* **105,** 511–543, doi: 10.1785/0120140093.

Friedrich, A. M., B. P. Wernicke, N. A. Niemi, R. A. Bennett, and J. L. Davis (2003). Comparison of geodetic and geologic data from the Wasatch region, Utah, and implication for the spectral character of Earth deformation at periods of 10 to 10 million years, *J. Geophys. Res.* **108,** doi: 10.1029/2001JB000682.

Girardin, V., and N. Limnios (2018). *Applied Probability: From Random Sequences to Stochastic Processes*, Springer.

Goldfinger, C., Y. Ikeda, R. S. Yeats, and J. Ren (2013). Superquakes and supercycles, *Seismol. Res. Lett.* **84,** doi: 10.1785/0220110135.

Gordon, R. G., and S. Stein (1992). Global tectonics and space geodesy, *Science* **256,** 333–342, doi: 10.1126/science.256.5055.333.

Hagiwara, Y. (1974). Probability of earthquake recurrence as obtained from a Weibull distribution analysis of crustal strain, *Tectonophysics* **23,** 313–318, doi: 10.1016/0040-1951(74)90030-4.

Hecker, S., S. B. DeLong, and D. P. Schwartz (2021). Rapid strain release on the Bear River fault zone, Utah–Wyoming—The impact of preexisting structure on the rupture behavior of a new normal fault, *Tectonophysics* **808,** doi: 10.1016/j.tecto.2021.228819.

Hough, S. E. (2016). *Predicting the Unpredictable: The Tumultuous Science of Earthquake Prediction*, Princeton University Press, Princeton, New Jersey.

Hurvich, C. M., and C.-L. Tsai (1989). Regression and time series model selection in small samples, *Biometrika* **76,** 297–307, doi: 10.1093/biomet/76.2.297.

Lomnitz-Adler, J. (1983). A statistical model of the earthquake process, *Bull. Seismol. Soc. Am.* **73,** 853–862, doi: 10.1785/BSSA0730030853.

Matthews, M. V., W. L. Ellsworth, and P. A. Reasenberg (2002). A Brownian model for recurrent earthquakes, *Bull. Seismol. Soc. Am.* **92,** 2233–2250, doi: 10.1785/0120010267.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE* 77**,** 257–286, doi: 10.1109/5.18626.

Reid, H. F. (1910). The mechanics of the earthquake: The California earthquake of April 18, 1906, *Report of the State Investigation Commission Carnegie Institution of Washington*, Washington, D.C., 2 pp.

Rockwell, T. K., S. Lindvall, M. Herzberg, D. Murbach, T. Dawson, and G. Berger (2000). Paleoseismology of the Johnson Valley, Kickapoo, and Homestead Valley faults: Clustering of earthquakes in the Eastern California shear zone, *Bull. Seismol. Soc. Am.* **90,** 1200–1236, doi: 10.1785/0119990023.

Salditch, L., S. Stein, J. S. Neely, B. D. Spencer, E. Brooks, M. A. Agnon, and M. Liu (2020). Earthquake supercycles and long-term fault memory, *Tectonophysics* **774,** doi: 10.1016/j.tecto.2019.228289.

Scharer, K. M., G. P. Biasi, and R. J. Weldon (2011). A reevaluation of the Pallett Creek earthquake chronology based on AMS radiocarbon dates, San Andreas fault, California, *J. Geophys. Res.* **116,** doi: 10.1029/2010JB008099.

Scharer, K. M., G. P. Biasi, R. J. Weldon, and T. E. Fumal (2010). Quasi-periodic recurrence of large earthquakes on the southern San Andreas fault, *Geology* **38,** 555–558, doi: 10.1130/G30746.1.

Schorlemmer, D., M. J. Werner, W. Marzocchi, T. H. Jordan, Y. Ogata, D. D. Jackson, S. Mak, D. A. Rhoades, M. C. Gerstenberger, N. Hirata, *et al.* (2018). The collaboratory for the study of earthquake predictability: Achievements and priorities, *Seismol. Res. Lett.* **89,** 1305–1313, doi: 10.1785/0220180053.

Sieh, K., D. H. Natawidjaja, A. J. Meltzner, C.-C. Shen, H. Cheng, K.-S. Li, B. W. Suwargadi, J. Galetzka, B. Philibosian, and R. L. Edwards (2008). Earthquake supercycles inferred from sea-level changes recorded in the corals of West Sumatra, *Science* **322,** 1674–1678, doi: 10.1126/science.1163589.

Sieh, K., M. Stuiver, and D. Brillinger (1989). A more precise chronology for earthquakes produced by the San Andreas fault in Southern California, *J. Geophys. Res.* **94,** 603–623, doi: 10.1029/JB094iB0P100603.

Votsi, I., N. Limnios, G. Tsaklidis, and E. Papadimitriou (2013). Hidden Markov models revealing the stress field underlying the earthquake generation, *Phys. A* **392,** 2868–2885, doi: 10.1016/j.physa.2012.12.043.

Wallace, R. E. (1970). Earthquake recurrence intervals on the San Andreas fault, *GSA Bull.* **81,** 2875–2890.

Weldon, R. J., T. E. Fumal, G. P. Biasi, and K. M. Scharer (2005). Past and future earthquakes on the San Andreas fault, *Science* **308,** 966–967, doi: 10.1126/science.1111707.

Weldon, R. J., T. E. Fumal, and G. P. Biasi (2004). Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work, *GSA Today* **14,** 4–10.

Manuscript received 16 May 2022
Published online 27 December 2022